A novel data-driven approach for integrated inventory replenishment, safety stock and lead time optimisation

Published Version:
Gabellini, M., Regattieri, A., Bortolini, M., Ronchi, M. (2026). A novel data-driven approach for integrated inventory replenishment, safety stock and lead time optimisation. INTERNATIONAL JOURNAL OF PRODUCTION RESEARCH, [in press], 1-21 [10.1080/00207543.2026.2696440].

Availability:
This version is available at: https://hdl.handle.net/11585/1075860 since: 2026-08-04

Published:

DOI: http://doi.org/10.1080/00207543.2026.2696440

A novel data-driven approach for integrated inventory replenishment, safety stock and lead time optimization

Matteo Gabellini[a*], Alberto Regattieri[b], Marco Bortolini[b],
Michele Ronchi[b]

[a]*Department of Engineering, University of Ferrara, Via Saragat 1, 44122 Ferrara, Italy*

[b]*Department of Industrial Engineering, University of Bologna, Viale del Risorgimento 4, 40136 – Bologna – Italy*

---

[*]Corresponding author.

Email addresses: gblmtt@unife.it (Matteo Gabellini) *Tel.: 3478201725*

Abstract

Effective inventory management requires deciding when to order, how much to order, and how much protection to keep in the form of safety stock and safety lead time. Many recent data-driven approaches still treat lead times, purchasing costs, and defect rates as fixed inputs, and usually optimize replenishment separately from safety buffers. This study proposes a data-driven framework that jointly determines replenishment decisions, safety stock, and safety lead time while accounting for time-varying lead times, purchasing costs, and defect quantities. The approach also uses a custom loss function to reflect the asymmetric consequences of forecast errors in the inventory setting. The method is tested on a real automotive dataset. Results show that it produces inventory decisions that are close to those of an ideal decision-maker and improves cost performance compared with the benchmark approaches considered. Overall, the study shows that integrating forecasting, safety-parameter tuning, and replenishment optimization can support more effective inventory decisions in dynamic operating environments.

Keywords: Inventory control, safety stock, safety lead time, data-driven methods, machine learning, custom loss

1. Introduction

Inventory management, which requires determining replenishment decisions together with safety stock and safety lead time levels, plays a crucial role in ensuring operational continuity and cost efficiency, as firms face substantial holding costs, shortage risks, and variability across supply, cost, and quality processes (Gonçalves et al., 2020; Inderfurth, 1995; Jaggi & Aggarwal, 1994; Van Kampen et al., 2010). However, these decisions are increasingly challenging due to the stochastic and interdependent nature of modern supply chains, where fluctuations in lead times, procurement costs, and defect rates combine with strong interactions between replenishment policies and safety parameters, making isolated or static optimization inadequate.

Fortunately, recent advances in Artificial Intelligence (AI) and Machine Learning (ML) have driven significant progress in operations (Calabrese et al., 2023; Cannas et al., 2023; Gabellini, Calabrese, Regattieri, et al., 2024; Grover et al., 2022; Helo & Hao, 2021; Venkatesh et al., 2024) and supply chain management (Ahmed et al., 2023; Badakhshan & Ball, 2023; Baryannis et al., 2019; Culot et al., 2024; Gabellini, Calabrese, Civolani, et al., 2024; Gabellini, Calabrese, Civolani, Regattieri, & Galizia, 2023; Gabellini, Calabrese, Civolani, Regattieri, & Mora, 2023; Gabellini, Civolani, et al., 2023, 2024; Gabellini et al., 2022, 2025; Ganesh & Kalpana, 2022; Kassa et al., 2023; Min, 2010; Pournader et al., 2021; Regattieri et al., 2024; Riahi et al., 2021; Smyth et al., 2024). Specifically, within inventory control, data-driven methods have emerged mainly through two paradigms: Predict-Then-Optimize (PTO), in which forecasts feed optimization models, and End-to-End (E2E) architectures, which jointly learn predictions and decisions. Although E2E approaches are conceptually appealing, their application in inventory control may be limited by the difficulty of learning cost-optimal decisions from historical data. In industrial settings, observed decisions typically reflect heuristic or suboptimal policies implemented by planners, rather than optimal decisions derived from a well-defined cost function. As a result, E2E models trained directly on historical data may replicate past decision patterns without guaranteeing cost-optimality.

Consequently, PTO approaches—often implemented through Deterministic Lookahead Models (DLM) (Powell, 2022)—remain the most widely adopted in practice due to their interpretability and the fact that they do not require such labelled data.

However, despite their growing adoption, current DLM-based PTO approaches present notable limitations. They rely on deterministic point forecasts and therefore cannot account for forecast uncertainty. They typically consider only demand variability, overlooking fluctuations in purchasing costs, lead times, and yield rates. Moreover, they generally optimize replenishment decisions alone while treating safety stock and safety lead time separately, even though these quantities are intrinsically intertwined. Addressing these issues requires extending DLMs to incorporate richer sources of uncertainty and explicit mechanisms to manage forecast misspecification. Parametrized Deterministic Lookahead Models (PDLMs), which integrate tunable parameters directly into the optimization process, provide a suitable foundation for such advancements.

Based on this evidence, the main novel contributions this study aims to bring are thus threefold:

- First, this study proposes a novel ML-based data-driven approach in the form of a PDLM that jointly determines replenishment decisions, safety stock, and safety lead time within a unified PTO framework.
- Second, the approach expands the scope of typical DLMs by incorporating a broader set of forecasted values beyond demand, including purchasing costs, lead times, and defect rates.
- Lastly, a tailored custom loss is adopted to train the adopted ML predictive model to consider the different impacts that overestimating or underestimating the predicted values can have.

The remainder of this paper is structured as follows. Section 2 reviews prior work on PTO and E2E inventory approaches and on safety stock and safety lead time optimisation. Section 3 presents the proposed methodology and describes its validation through a real-world automotive case study.

Section 4 reports the results, while Section 5 discusses their managerial implications. Section 6 concludes with key insights and directions for future research.

2. Literature review

Modern inventory management increasingly relies on advanced methodologies to address the uncertainties inherent in supply chain systems. This literature review delves into two critical dimensions of this evolution, focusing on the growing role of data-driven approaches. Section 2.1 explores data-driven methods for inventory replenishment, highlighting advancements in determining when and how much to order. Sections 2.2 and 2.3 shift focus to safety stock and safety lead time optimization, where progress has primarily involved traditional mathematical models, linear programming, and simulation-based approaches. Lastly, Section 2.4 synthesizes these findings to provide a schematic literature overview, highlighting the current gap.

2.1. Data-driven approaches for inventory replenishment

Over time, data-driven approaches for inventory replenishment have evolved into two main categories: PTO and E2E models. E2E methods integrate forecasting and optimization into a single framework and have shown strong potential. Huber et al., (2019) demonstrated how combining demand estimation with optimization—interpreted through quantile regression—highlights the importance of accurate forecasts and the flexibility of ML methods such as gradient-boosted trees. Oroojlooyjadid et al., (2020) applied deep learning to the newsvendor problem, removing distributional assumptions and proving robust under volatile and sparse data. Punia et al., (2020) extended E2E modelling to multi-item settings using ML-based quantile regression, while Pirayesh Neghab et al., (2022) used deep learning to capture cross-period dependencies with observable and latent features. Qi et al., (2023) advanced E2E replenishment models by jointly learning decisions across multiple periods without relying on parametric assumptions, and Clausen & Li, (2022) showed that risk-minimization techniques combined with ML can outperform traditional approaches.

Together, these works underscore the transformative potential of E2E methods in complex, volatile, data-rich environments.

Despite these advantages, E2E approaches require historical optimal decisions to serve as labels—data that are often costly or impractical to obtain. This limitation has sustained the relevance of PTO approaches, which preserve a sequential structure between forecasting and optimization. Representative contributions include Ban et al., (2019), who developed a data-driven procurement model integrating forecasting with integer programming under realistic business constraints; Lin et al., (2022), who proposed a risk-averse newsvendor formulation using nonparametric estimation and VaR regularization; and Demizu et al., (2023), who leveraged probabilistic forecasting and meta-learning for multi-product, multi-store inventory settings. Other PTO studies address domain-specific challenges, such as Mete Ayhan & Kır, (2024) in automotive casting, integrating ML prediction with classical inventory models; Schmidt & Pibernik, (2024), who introduced global learning across heterogeneous products in a rolling-horizon framework; and van der Haar et al., (2024), who enhanced PTO with supervised learning and custom loss functions for lost-sales and perishable-goods environments. Collectively, these studies confirm the practicality and versatility of PTO approaches while highlighting remaining gaps in handling uncertainty and jointly optimizing replenishment and safety parameters.

2.2. Approaches for safety stock optimization

Safety stock optimization has evolved from traditional analytical formulas to advanced linear programming and simulation-based methods, reflecting the growing complexity of demand and supply uncertainties. Early analytical contributions—such as Alicke, (2005); Gudehus, (2006) and Herrmann, (2011)—extended classic service-level-based formulas by incorporating forecasting errors, undershooting effects, and cycle-disruption adjustments. More flexible optimization-based approaches subsequently emerged. For demand uncertainty, Janssens & Ramaekers, (2011) proposed a linear program for worst-case lead-time demand, while Beutel & Minner, (2012) integrated causal forecasting with LP techniques, and Chen et al., (2013) applied Monte Carlo simulation to clinical trial supply chains. For lead time variability, M. A. O. Louly & Dolgui, (2009) developed an efficient branch-and-bound method for assembly systems. Yield-related uncertainty was addressed by de Armas & Laguna, (2020), who modelled stochastic production rates in capacitated lot-sizing settings. Multi-uncertainty studies, including Chung et al., (2005); Hung & Chang, (1999) and Keskin et al., (2015), further demonstrated how nonlinear programming, throughput-driven models, and integrated capacity–inventory policies can jointly capture fluctuations in flow times, yields, and production constraints. Collectively, these works show a progression from simple formulas to sophisticated optimization and simulation frameworks capable of handling demand, lead time, yield, or combined uncertainties, thereby strengthening the robustness and adaptability of modern safety stock strategies.

2.3. Approaches for safety lead time optimization

Safety lead time, like safety stock, serves as a key buffering mechanism against demand and supply uncertainty, and research has progressively explored when and how it should complement or substitute inventory-based protection. Early studies such as Buzacott & Shanthikumar, (1994) examined trade-offs in MRP systems, showing that safety lead time is preferable when demand forecasts are reliable, whereas safety stock becomes more effective under high demand or lead time variability. Building on this, Molinder, (1997) used simulation and simulated annealing to assess how variability and cost structures shape the optimal balance between the two mechanisms. Ould-Louly & Dolgui, (2004) extended the analysis to assembly systems, proposing a cost-minimizing model that incorporates lead time uncertainty within a generalized Newsboy framework, while Van Kampen et al., (2010) demonstrated through simulation that the relative effectiveness of safety stock vs. safety lead time depends on the dominant source of uncertainty in multi-product supply chains. Further refinements include M. A. Louly & Dolgui, (2013), who optimized MRP parameters under stochastic lead times without distributional restrictions, highlighting the importance of adjusting planning parameters to lead time variability. More recently, Silva et al., (2022) introduced a hybrid bi-objective simulation–optimisation approach that jointly tunes safety stock and safety time, identifying Pareto-efficient trade-offs between holding costs and service levels under dynamic demand and stochastic lead times. Collectively, these studies show that safety lead time is a valuable complement to safety stock, particularly when supply variability dominates, and that the optimal buffering strategy depends on the interplay of demand characteristics, lead time behaviour, and system cost structures.

2.4. Literature gap analysis

Table 1 synthesizes prior research by mapping the inventory problems addressed, the uncertainties considered, and the solution approaches adopted.

Table 1:Summary of the literature review

STUDY	INVENTORY PROBLEM	SOLVING APPROACH	ADDRESSED UNCERTAINTY	CUSTOM LOSS
Mete Ayhan & Kır, (2024)	IR	PTO with DLM	D	No
Schmidt & Pibernik, (2024)	IR	PTO with DLM	D	No
Demizu et al., (2023)	IR	PTO with DLM	D	No
Qi et al., (2023)	IR	E2E	D; LT	No
Oroojlooyjadid et al., (2020)	IR	E2E	D	No
Ban et al., (2019)	IR	PTO with DLM	D	No
Lin et al., (2022)	IR	PTO with DLM	D	No
Pirayesh Neghab et al., (2022)	IR	E2E	D	No
Punia et al., (2020)	IR	E2E	D	No
van der Haar et al., (2024)	IR	PTO with DLM	D	Yes
Huber et al., (2019)	IR	E2E	D	No
Clausen & Li, (2022)	IR	E2E	D	No
Mete Ayhan & Kır, (2024)	IR	PTO with DLM	D	No
Alicke, (2005)	SS	AF	D	No
Herrmann, (2011)	SS	AF	D	No
Gudehus, (2006)	SS	AF	D	No
Janssens & Ramaekers, (2011)	SS	LP	D	No
Beutel & Minner, (2012)	SS	LP	D	No
Chen et al., (2013)	IR & SS	DES + MILP	D	No
Prak et al., (2017)	SS	-	D	No
M. A. O. Louly & Dolgui, (2009)	SS	MILP	LT	No
de Armas & Laguna, (2020)	SS	MIIP + S	Y	No
Hung & Chang, (1999)	SS	LP	Y; LT	No
Chung et al., (2005)	SS	NLP	D; LT	No
Keskin et al., (2015)	SS	MILP	D; Y	No
Alecchang, (1985)	SS & SLT	AF	D; LT	No
Buzacott & Shanthikumar, (1994)	SLT	AF	D; LT	No
Molinder, (1997)	SLT	HP	D; LT	No
Ould-Louly & Dolgui, (2004)	SLT	NLP	LT	No

Van Kampen et al., (2010)	SS & SLT	S	D; LT	
M. A. Louly & Dolgui, (2013)	SLT	NLP	LT	No
Silva et al., (2022)	SS & SLT	S + HP	D; LT	No
Proposed study	**IR & SS & SLT**	**PTO with PDLM**	**LT; C; Y**	**Yes**

IR: Inventory Replenishment, SSO: Safety Stock Optimization; PTO: Predict Then Optimize, E2E: End To End, DLM: Deterministic Lookahead Model, PDLM: Parametrized Deterministic Lookahead Model; AF: Analytic Formula; LP: Linear Programming; DES: Discrete Event Simulation; MILP: Mixed Integer Linear Programming; NLP: Non-Linear Programming; S: Simulation; HP: Heuristic Procedure; D: Demand, LT: Lead Time; Y: Yield; C: Cost

The literature shows a clear separation: most studies optimize either replenishment or safety buffers, rarely both. Replenishment is typically approached through data-driven PTO or E2E methods, yet these models predominantly address demand uncertainty and generally rely on standard forecasting losses that cannot reflect the asymmetric cost impact of over- and under-estimation—an important limitation in inventory contexts. Conversely, safety stock and safety lead time are still largely determined using traditional formulas, linear programming, or simulation models, which again focus mainly on demand variability and seldom incorporate additional uncertainties.

A key gap emerging from the literature is the lack of integrated models that jointly optimize replenishment, safety stock, and safety lead time despite their strong interdependence in practice. Only a few works, such as Chen et al., (2013) and Silva et al., (2022), address combined buffer optimisation, but these do not incorporate ML-based forecasting with custom loss functions, nor do they embed joint buffer design within a data-driven replenishment framework. Moreover, existing studies seldom consider multiple concurrent sources of uncertainty—such as lead time, cost variability, and yield—within a unified PTO structure.

This study departs from prior work by introducing a PDLM-based PTO framework that integrates replenishment decisions with safety stock and safety lead time optimisation, guided by ML models trained with a tailored custom loss. To our knowledge, it is the first approach to jointly optimize these decisions while simultaneously accounting for multiple evolving uncertainties (lead time, cost, and yield). By closing these methodological gaps, the proposed framework advances the state of the art in data-driven inventory control and offers a comprehensive, adaptive solution for real-world environments.

3. Material and methods

The following section presents a comprehensive overview of the proposed approach, the experimental design developed to evaluate its effectiveness, and the experimental setup used to implement and test the proposed methodology.

3.1. Problem definition

This study addresses a single-item, multi-period inventory control problem under uncertainty in which replenishment decisions, safety stock, and safety lead time are determined jointly over a finite planning horizon. The objective is to minimize total inventory-related cost, namely purchasing and holding costs, while enforcing a no-stockout service policy through feasibility constraints. The main decision variables are the order quantities to be placed over time, the associated receipt quantities, and the resulting inventory levels, safety stock and safety lead time.

3.2. Proposed approach

The proposed approach is grounded in the class of PDLM policies, as formalized by Powell, (2022). In the terminology of sequential decision problems, a DLM computes decisions by optimizing over a projected future horizon using point forecasts of exogenous information. Let S_t denote the system state at time t, $x_t$the decision, and $\widehat{W}_{t+1:t+H}$the point forecasts of the exogenous information process $W_{t+1:t+H}$. Following the deterministic lookahead formulation in Powell (2022), a classical DLM determines $x_t$through:

$$x_t^{DLM} = \arg min_{x_{t:t+H}} \sum_{\tau=t}^{t+H} C(S_\tau, x_\tau) \tag{1}$$

where future costs and constraints are evaluated deterministically using the forecasted values of the exogenous variables. As discussed by Powell, (2022), such deterministic lookahead policies are widely used in practice for complex resource allocation problems; however, they do not explicitly address uncertainty or incorporate mechanisms to adjust suboptimal behavior arising from forecast errors.

In contrast, a PDLM enriches the deterministic lookahead by introducing a vector of tunable parameters θ that modifies the optimization problem to improve robustness under uncertainty. Formally, a PDLM is defined as:

$$x_t^{PDLM}(\theta) = \arg min_{x_{t:t+H}} C(S_{t:t+H}, x_{t:t+H}; \theta) \tag{2}$$

where the parameter vector $\theta \in \Theta$ alters either the objective function or the constraints of the lookahead optimization model, thereby shaping the decision policy to better accommodate uncertainty in the exogenous information process. Powell, (2022) emphasizes that such parametrization acts as a structured mechanism for incorporating domain-informed adjustments—such as inventory buffers or lead-time slack—directly into the decision model.

In the present work, the PDLM is instantiated for the inventory control problem by defining $\theta = (\theta_{SS}, \theta_{SLT})$, where θ_{SS} represents a tunable safety stock level and θ_{SLT} a tunable safety lead time. These parameters directly modify the feasible region of the MILP model: θ_{SS} enters the minimum inventory constraint, while θ_{SLT} shifts the receipt-timing constraints to anticipate uncertainty in realized lead times. Given their operational role, θ functions as a policy parameter in the sense of Powell, (2022) cost-function approximation and PDLA framework.

The optimal parameter vector is obtained through validation-based policy search, again in accordance with following formulation:

$$\theta^* = \arg min_{\theta \in \Theta} \mathbb{E}[J(\theta)] \tag{3}$$

Where $J(\theta)$ denotes the realized cost arising from applying the lookahead policy $x_t(\theta)$ over historical data. This tuning embeds the effect of uncertainty without constructing an explicit stochastic lookahead.

Figure 1 presents how this PDLM structure is operationalized for joint determination of inventory replenishment, safety stock optimization, and safety lead time calibration.

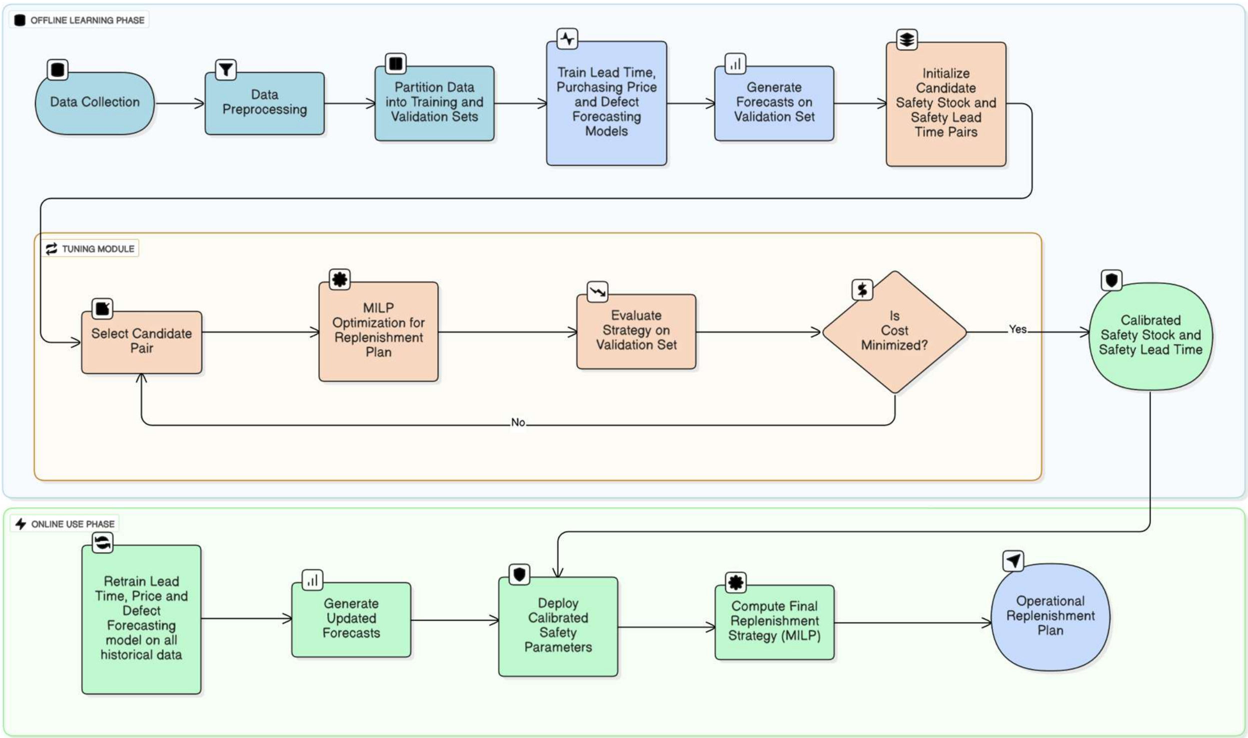

Figure 1: Proposed approach workflow

The proposed approach operates through two sequential phases: an offline learning phase, in which the model components are built and calibrated, and an online use phase, in which these calibrated elements are employed to generate operational replenishment decisions.

In the offline phase, historical data on lead time, purchasing cost, and defect rates are collected, preprocessed, and partitioned into training and validation sets. Forecasting models are trained on the training set and used to generate predictions over the validation horizon. These forecasts serve as inputs to the MILP inventory control model, which computes replenishment plans under different combinations of safety stock and safety lead time. A tuning module then iteratively explores candidate safety parameter pairs, evaluates the resulting replenishment strategies on the validation set by simulating their realized performance, and identifies the pair that minimizes total cost. This process yields calibrated safety stock and safety lead time values together with the trained forecasting models and the MILP structure that constitute the decision engine.

In the online phase, the forecasting module is retrained on the available historical data to update predictions for lead time, cost, and defects. These forecasts, combined with the calibrated safety parameters, are provided to the MILP model to compute replenishment decisions for the operational horizon. Given the limited complexity of the forecasting models, retraining remains computationally negligible the majority of industrial setting. However, for longer horizons, scalability can be maintained by adopting rolling training windows or scheduled updates, and by recalibrating safety parameters only when structural changes in demand, lead times, or costs are observed. It is important to clarify that the proposed framework does not constitute a multi-objective optimization problem. Instead, it is structured as a sequential decision pipeline in which different cost functions are employed at different stages for distinct purposes. In the forecasting module, loss functions are used to train predictive models, shaping the statistical properties of the forecasts. In the optimization module, the MILP objective function determines cost-optimal replenishment decisions based on these forecasts. Finally, in the tuning module, a validation-based cost function evaluates the realized performance of these decisions under uncertainty to calibrate safety stock and safety lead time parameters. These cost functions are therefore not competing objectives but coordinated components of a unified framework aimed at minimizing overall inventory cost.

Moreover, the proposed approach does not follow an end-to-end differentiable architecture. In line with the taxonomy discussed in Section 2 and reflected in Table 1, the method is explicitly a PTO approach. The forecasting module and the MILP optimization module remain sequential and are not connected through gradients. The term “joint” refers to determining replenishment, safety stock, and safety lead time within a single PDLM decision structure, with the tunable parameters θ calibrated through validation-based policy search, providing an indirect form of feedback without implying end-to-end learning.

3.2.1Forecasting module

The Catboost model (Dorogush et al., 2018), a gradient boosting tree algorithm model, has been selected for the forecasting models in charge of predicting the future value of lead time, cost, and defect rates for future periods. Catboost has emerged as a leading ML technique for various forecasting applications, consistently demonstrating superior performance in diverse settings due to its ability to capture complex patterns in data effectively. Its robust performance in time series prediction tasks has been validated in numerous forecasting competition (Bentéjac et al., 2021; Pan & Zhang, 2020). In particular, three Catboost models, one for lead time, one for cost, and one for defect predictions, have been trained to adopt a univariate forecasting approach, relying solely on the historical values of the target variable to generate predictions. Despite the availability of more complex multivariate models, univariate approaches have proven highly effective in many practical scenarios, offering a balance between simplicity and predictive accuracy. This choice enables the forecasting module to focus on extracting temporal patterns directly related to each target variable while avoiding the potential noise introduced by additional predictors.

However, while for the CatBoost model adopted to generate purchasing cost predictions, the Root Mean Squared Error (RMSE) is adopted as a loss function, a custom loss has been instead adopted for the CatBoost models providing predictions for lead time and quality. Indeed, while errors in purchasing cost estimation do not have a different inventory cost impact if they overestimate or underestimate the real value, different cost impacts are instead generated by lead time and quality predictions if they overestimate or underestimate real values. Underestimating lead time or defect quantities increases the risk of inventory insufficiency under realized operating conditions, whereas overestimating them primarily leads to higher holding requirements. The custom loss is therefore designed to reflect the asymmetric operational consequences of forecasting errors and to bias predictions toward safer replenishment decisions. For this reason, traditional RMSE loss, which, according to Equation 4, does not allow weight forecasting error based on the fact that they are over-estimating or underestimating the predicted metric, has been preferred for purchasing price predictions:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{N}(y_t - \hat{y}_t)^2}{N}} \tag{4}$$

On the contrary, a problem-specific custom loss introduced in Equation 5 is adopted to better capture the imbalance nature of forecasting error in the problem of predicting lead time and quality:

$$CUSTOM\ LOSS = \begin{cases} C^{holding}\sqrt{\frac{\sum_{t=1}^{N}(y_t - \hat{y}_t)^2}{N}}\ if\ y_t \geq \hat{y}_t \\ C^{stockout}\sqrt{\frac{\sum_{t=1}^{N}(y_t - \hat{y}_t)^2}{N}}\ if\ y_t < \hat{y}_t \end{cases} \tag{5}$$

In both Equation 4 and 5, y_t and $\hat{y}_t$ respectively represent the true value and the forecasted value at time t and N is the length of the training set provided to the model. However, in Equation 5, the introduction of the parameters $C^{holding}$ and $C^{stockout}$can be adopted to represent the relative impact of overestimation (holding-related effects) and underestimation (stockout-related effects). The asymmetry parameters are not arbitrary hyper-parameters, but are defined to reflect the same economic asymmetry underlying the inventory problem: underestimation leads to shortage-related risks, whereas overestimation mainly increases holding-related effects. While these weights are aligned with the asymmetric operational consequences of underestimation and overestimation, they are not imposed to be numerically identical to the MILP cost coefficients, as the loss is defined in physical units whereas the inventory objective is monetary and multi-period. Therefore, the loss is cost-aligned rather than cost-identical. It is important to emphasize that this custom loss is not intended as a methodological innovation per se, but rather as a practical mechanism to ensure that the forecasting module is aligned with the asymmetric operational consequences considered in the overall decision framework. The contribution of this work lies in the integration of this cost-aligned forecasting component within the proposed PDLM framework—connecting forecast bias shaping, multi-variable prediction, and the Bayesian tuning of safety stock and safety lead time for the first time in a PTO setting—rather than in the loss function alone.

3.2.2 Inventory control module

This subsection presents the inventory control module adopted in the proposed framework. For clarity, the presentation is organized as follows: Section 3.2.2.1 introduces the model assumptions, Section 3.2.2.2 reports the notation, and Section 3.2.2.3 provides the mathematical formulation of the model.

3.2.2.1 Model assumption

The inventory control module has been formulated as a MILP model. The following assumptions define the operational scope of formulation:

- Demand is predetermined based on a rolling horizon planning approach and an MRP explosion of component requirements derived from the production schedule.
- Purchasing costs increase linearly with the ordered quantity, with no quantity discounts.
- Variable holding cost increases linearly with the quantity of inventory held and the duration of storage across periods.
- Fixed per-order costs are assumed negligible in the considered setting, where ordering is largely automated and such costs are effectively embedded in unit purchase prices.

This assumption defines the boundary of applicability of the model. In settings where setup costs, minimum order quantities, or batch production are relevant, the problem becomes a more general lot-sizing formulation, which can be incorporated by extending the MILP with standard fixed-cost and capacity constraints..

3.2.2.2 Notation

The sets, parameters, and decision variables used in the MILP formulation are summarized in Table 2.

Table 2: Inventory model sets, parameters and variables

TYPE	TERM	DESCRIPTION	DOMAIN
Sets	T	Time periods	$t \in T$
	$\mathcal{T}^{ord}$	Feasible order periods with arrivals inside horizon	$\{(\tau \in T \mid \tau + f_\tau^{leadtime} + \theta_{SLT} \leq T)\}$
	$\mathcal{T}^{ric}$	Receipt periods	$\{(t \in T \mid \exists \tau \in \mathcal{T}^{ord}: t = \tau + f_\tau^{leadtime} + \theta_{SLT})\}$
Parameters	$f_t^{c\_ord}$	Forecasted unitary ordering cost in period t [euro/items]	$\mathbb{R}_{\geq 0}$
	$f_t^{leadtime}$	Forecasted lead time in period t [periods]	$\mathbb{Z}_{\geq 0}$
	D_t	Demand in period t [items]	$\mathbb{Z}_{\geq 0}$
	f_t^{defect}	Forecasted defect in period t [items]	$\mathbb{Z}_{\geq 0}$
	$c^{holding}$	Unitary holding cost [euro/(items*periods)]	$\mathbb{R}_{\geq 0}$
	I_0	Initial inventory level [items]	$\mathbb{Z}_{\geq 0}$
	θ_{SS}	Tuned value of safety stock [items]	$\mathbb{Z}_{\geq 0}$
	θ_{SLT}	Tuned value of safety lead time [periods]	$\mathbb{Z}_{\geq 0}$
	OLS	Order Lot Size [items]	$\mathbb{Z}_{\geq 1}$
Variables	y_t^{ord}	Number of lots ordered at time t [items]	$\mathbb{Z}_{\geq 0}$
	y_t^{ric}	Received lots at time t [items]	$\mathbb{Z}_{\geq 0}$
	I_t	Inventory level at time t [items]	$\mathbb{Z}_{\geq 0}$

3.2.2.3 Mathematical formulation:

Using the notation introduced in Table 2, the inventory control problem is formulated as follows.

The objective function to minimize is reported in Equation 6:

$$Z = \sum_{t=1}^{T} f_t^{c\_ord} * y_t^{ord} * OLS + c^{holding} * I_t \tag{6}$$

According to Equation (6), the objective function aims to minimize the overall cost related to the forecasted ordering and holding costs by adjusting the ordered quantity and inventory level in each planning period. Because service protection is enforced through feasibility constraints, shortage costs do not appear in the nominal objective. This formulation reflects settings where stockouts are prevented by design and should be distinguished from the policy-calibration stage, where realized costs—including stockout penalties—are used to evaluate the impact of forecast errors and assess the robustness of the resulting policy under uncertainty.

The main constraints of the model are represented by Equations (7) to (11):

$$I_t = I_{t-1} + y_t^{ric} * OLS - D_t - f_t^{defect} \qquad \forall\, t \in T \tag{7}$$

$$I_t \geq \theta_{SS} \qquad \forall\, t \in T \setminus t_0 \tag{8}$$

$$y_t^{ric} * OLS = y_\tau^{ord} * OLS \qquad \forall\, (t,\tau) \in \{(\tau, \tau + f_\tau^{leadtime} + \theta_{SLT} | \tau \in \mathcal{T}^{ord})\} \tag{9}$$

$$y_t^{ric} = 0 \qquad \forall\, t \in T \setminus \mathcal{T}^{ric} \tag{10}$$

$$y_t^{ric}, y_t^{ord}, I_t \in \mathbb{Z}_{\geq 0} \qquad \forall\, t \in T \tag{11}$$

Equation (7) ensures that the ending inventory in each period is determined by the previous period's inventory plus the quantity received (adjusted by the order lot size) minus the demand and the defective units in the current period. For the first period, the initial inventory I_0 is used. In contrast, for subsequent periods, the previous period's inventory I_{t-1} carries forward. Equation (8) constraint enforces a minimum inventory level equal to the tuned safety stock values at the end of each period, serving as a buffer against uncertainties in lead time and yield. The formulation adopts a service-constrained perspective in which shortages are not explicitly modeled through backorders or lost-sales variables. Instead, inventory is constrained to remain non-negative and above the safety stock threshold in each period. This modeling choice reflects operational contexts in which stockouts are considered unacceptable and must be prevented by design. Consequently, the optimization problem minimizes procurement and holding costs subject to a no-shortage service requirement, rather than explicitly pricing shortage costs in the objective function. Equations (9) and (10) regulate the timing consistency between orders and receipts. To formalize this, we introduce two index sets $\mathcal{T}^{ord}$ and $\mathcal{T}^{ric}$. $\mathcal{T}^{ord}$is the set of order periods whose arrivals fall within the planning horizon, while $\mathcal{T}^{ric}$ is the corresponding set of receipt periods. Equation (9) ensures that for every feasible order period $t \in \mathcal{T}^{ord}$, the receipt in its associated arrival period $t = \tau + f_{\tau}^{leadtime} + \theta_{SLT}$ matches the quantity ordered. Equation (10) complements this by enforcing that receipts in all other periods—those not in $\mathcal{T}^{ric}$—must be zero. Together, these two constraints guarantee a consistent and unambiguous mapping between order placement and receipt, removing any ambiguity about time shifts or lead-time indexing. Finally, Equation (11) specifies the domain of all decision variables.

3.2.3 Safety stock and safety lead time tuning module

A Bayesian algorithm is proposed to tune the safety stock and lead time values. Bayesian optimization offers several advantages when dealing with complex, expensive-to-evaluate functions (Imani et al., 2022). Using a probabilistic surrogate model, often a Gaussian Process, selects new candidate points based on exploitation and exploration, thereby reducing the number of costly evaluations. It does not rely on gradient information, which facilitates its application to non-differentiable or discontinuous functions. The method systematically and intelligently navigates the design space, typically reaching a global optimum more efficiently than random or grid search methods.

The search domain of the tuning procedure is restricted to non-negative values for both safety stock and safety lead time, as these parameters represent protective inventory and time buffers. Negative values would not correspond to operationally meaningful buffering policies and are therefore excluded from the feasible policy space.

Iteration by iteration, the Bayesian algorithm proposes candidate values for safety stock and lead time, refining these based on observed outcomes. This process aims to minimize discrepancies between the optimization model's ordering decisions and the realizations affected by forecast inaccuracies. Formally, at each iteration, the Bayesian approach evaluates an objective function measuring the error associated with the ordering decision, guiding the selection of new candidate points that systematically improve upon previous solutions and converge toward optimal safety parameters. Specifically, at each iteration, the Bayesian approach evaluated the objective function described by Equation (12):

$$\sum_{t=1}^{T} c_t^{order} * y_t^{ord} * OLS + c^{holding} * I_t^{+,val} + c^{stockout} * I_t^{-,val} \quad (12)$$

Where c_t^{order} is the real cost of ordering a specific component in period t, which can be different from $f_t^{c\_ord}$, and $I_t^{+,val} = \max(I_t, 0)$ and $I_t^{-,val} = \max(-I_t, 0)$ are respectively the positive and negative parts of the realized inventory position generated in period t over the validation set. Therefore, $I_t^{+,val}$ represents on-hand inventory subject to holding cost, whereas $I_t^{-,val}$ represents inventory shortfall subject to stockout cost. This evaluation formulation differs from the nominal MILP objective, as it incorporates stockout costs to quantify the economic impact of potential service violations under realized conditions. This distinction enables a robustness-oriented calibration of safety parameters. The realized inventory position I_t^{val} is computed according to the Equation (13):

$$I_t^{val} = I_{t-1}^{val} + y_t^{val\_ric} - D_t^{val} - Q_t^{val} \quad (13)$$

Where I_{t-1}^{val} is the true inventory at time t-1 in the validation dataset, $y_t^{val\_ric}$ is the true quantity arrived at time t based on the quantity ordered from the optimization module and considering the real lead time required for its arrival in the validation dataset, D_t^{val} is the true demand occurring at time t in the validation dataset and Q_t^{val} is the true number of defective component with quality problems observed in period t in the validation dataset. To ensure a well-defined initialization of the evaluation procedure, the system state at period t=1 is set equal to the actual observed inventory position at the beginning of the validation horizon. This initialization is consistently applied across all candidate parameter configurations and benchmark methods.

3.3. Experimental design

Several benchmarks have been tested against the proposed approach to evaluate its capability to generate cost-effective decisions. In particular, we refer to the following benchmarks: Oracle Approach, Naïve Moving Average (NMA) Approach, Autoregressive Integrated Moving Average (ARIMA) Approach, Catboost Approach, and Formula Based (FB) Approach.

The Oracle Approach has been designed to resemble the decision-making capability of a decision-maker with perfect knowledge of the future. In this approach, safety stock and safety lead time are thus not required as the inventory replenishment decision is made assuming that the decision maker has perfect knowledge of the future lead time cost and defective quantity. The results generated by the Oracle Approach thus represent an ideal lower cost bound.

Instead, the NMA and ARIMA approaches follow the same approach reported in Figure 1. However, the NMA Approach adopt as a forecasting algorithm to estimate future lead time, cost and defect a Naïve Moving Average model, while the ARIMA Approach generates predictions relying on the ARIMA algorithm. These benchmarks have been selected to understand the impact of different forecasting technologies within our proposed approach. Indeed, the comparison allows us to understand if and how much the adoption of more advanced ML forecasting technology can result in better accuracy and inventory costs. This information can be fundamental for industrial practitioners to evaluate the effectiveness of relying on new technology, which, however, requires skilled personnel and investment in both computational infrastructure and training plans.

The CatBoost approach, which uses the same forecasting algorithm as outlined in Section 3 but is consistently trained with the traditional RMSE loss function, was instead selected to assess the effectiveness of the proposed custom loss function.

Lastly, to evaluate the effectiveness of the proposed data-driven strategy for safety stock and lead time optimization, a final benchmark model has been designed to investigate the cost efficiency of the proposed safety stock and lead time determination method. Specifically, while the proposed approach relies on the Bayesian algorithm to iteratively test, simulate and find the best safety stock and safety lead time value, the FB follow the same step reported in the proposed approach with the difference, however that the safety stock dimensioning step is done recurring the traditional analytic formulation (Gonçalves et al., 2020) reported in Equation (14):

$$SS = z\sqrt{LT * \theta_D^2 + D^2\theta_{LT}^2} \tag{14}$$

Where z is the service level factor obtained from the standard normal distribution corresponding to the desired probability of not stocking out, D is the average demand, $\theta_D$is the standard deviation of demand, LT is the average lead time, and θ_{LT} is the standard deviation of lead time. Although this formulation is derived under stylized assumptions, it is widely used in practice as a reference rule for safety stock dimensioning. For this reason, it is adopted in this study as a benchmark to represent a commonly implemented approach in industrial settings. This allows us to evaluate the proposed method against an established baseline and to highlight the benefits of integrating multiple sources of uncertainty within a unified data-driven framework.

Considering these benchmarks, several metrics have been computed to provide an effective comparison. First, the difference ΔTC between the proposed approach cost $TC^{proposed}$ and a benchmark approach cost $TC^{benchmark}$ has been computed according to Equations (15):

$$\Delta TC = \frac{TC^{proposed} - TC^{benchmark}}{TC^{benchmark}} * 100 \tag{15}$$

Where $TC^{proposed}$ and $TC^{benchmark}$ has been computed according to Equations (16) and (17):

$$TC^{proposed} = \sum_{t=1}^{T} c_t^{order} * y_t^{proposed} * OLS + c^{holding} * I_t^{+,proposed} + c^{stockout} * I_t^{-,proposed} \tag{16}$$

$$TC^{benchmark} = \sum_{t=1}^{T} c_t^{order} * y_t^{benchmark} * OLS + c^{holding} * I_t^{+,benchmark} + c^{stockout} * I_t^{-,benchmark} \tag{17}$$

Here, $I_t^{+,proposed}$ and $I_t^{+,be}$ denote the positive part of the inventory position generated at time t by the ordering decisions derived, respectively, from the proposed approach ($y_t^{proposed}$) and the benchmark approach ($y_t^{benchmark}$), when considering the real lead time and defective quantity manifested over the test set. Conversely, $I_t^{-,proposed}$ and $I_t^{-,benchmark}$ denote the corresponding negative part of the inventory position, representing inventory shortfall. Formally, for any inventory level I_t, the positive and negative parts are defined as $I_t^+ = \max(I_t, 0)$ and $I_t^- = \max(-I_t, 0)$. Therefore I_t^+ captures on-hand inventory subject to holding cost, whereas I_t^- captures shortage magnitude subject to stockout cost.

Secondly, the predictive accuracy of the different forecasting algorithms (CatBoost with custom loss, CatBoost, Naïve Moving Average, ARIMA) has been compared considering traditional regression accuracy metrics like the Mean Error (ME), the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Symmetric Mean Absolute Percentage Error (SMAPE) to understand the impact that the forecasting accuracy could has on the final cost. Specifically, the observed predictive accuracy metrics have been computed as follows:

$$ME = \frac{1}{N}\sum_{t=1}^{N} Y_t - \hat{Y}_t \tag{18}$$

$$MAE = \frac{1}{N}\sum_{t=1}^{N} |Y_t - \hat{Y}_t| \tag{19}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N} (Y_t - \hat{Y}_t)^2} \tag{20}$$

$$SMAPE = \frac{1}{N}\sum_{t=1}^{N} \frac{|Y_t - \hat{Y}_t|}{\frac{Y_t + \hat{Y}_t}{2}} \tag{21}$$

Where Y_t is the true historical value of the predicted metric at time t and $\hat{Y}_t$ is the value predicted, and N is the length of the test set.

All performance comparisons in this study—inventory cost differences across forecasting models, and cost differences across safety-parameters methods—were evaluated using the Wilcoxon signed-rank test. This choice reflects the structure of the data: each component provides paired observations (e.g., the cost produced by two methods under identical conditions), and the distributions of these paired differences are non-normal. The Wilcoxon test therefore provides an appropriate, distribution-free procedure for assessing whether the median difference between two methods differs significantly from zero. To complement hypothesis testing, Cliff's delta was computed as a non-parametric effect size. It indicates the probability that the proposed approach outperforms (or underperforms) a benchmark and quantifies the practical magnitude of the observed differences. Values near zero indicate negligible effects, whereas values closer to ± 1 indicate strong directional differences. Specifically, a value of $+1$ means the proposed approach outperforms the benchmark for all components, whereas -1 means the benchmark outperforms the proposed method for all components; values near zero indicate negligible differences. Because multiple comparisons were conducted across metrics and benchmark methods, p-values were adjusted using the Holm–Bonferroni procedure, which controls the family-wise error rate while retaining more statistical power than traditional Bonferroni correction. In this context, a statistically significant adjusted p-value ($\alpha = 0.05$) indicates that the observed performance differences are unlikely to be attributable to random variation. Together, this framework ensures a rigorous, consistent, and interpretable assessment of all statistical comparisons conducted in the study.

Lastly, the computational time required to train the forecasting models, optimize the safety stock and safety lead time, and solve the inventory replenishment problem has been computed to understand the balance between inventory performance and computational time required by the approaches.

3.4. Experimental set up

The experimental setup for evaluating the new proposed data-driven inventory control approach involved a systematic collection and analysis of real-world data, a structured data partitioning strategy, and the reliance on widely adopted computational packages for the forecasting, optimization, and tuning to ensure the validity and reproducibility of the results.

The dataset was sourced from a reputable automotive company, ensuring its relevance and reliability. It included daily detailed records for 94 carefully selected components over four years, chosen to be highly representative of the company's broader inventory. With key variables such as defect rates, lead times, and costs, the dataset provides a strong foundation for testing the proposed approach across a diverse range of inventory management scenarios.

To prepare the experiments, the dataset was split into three distinct subsets based on chronological order: training (from 2021-01-01 to 2023-01-01), validation (from 2023-01-01 to 2023-10-31), and test (from 2023-10-31 to 2024-08-31) datasets. The training set was used to train the forecasting models, establishing their ability to predict lead times, costs, and defect rates. The validation set was employed to fine-tune the safety stock and safety lead time, ensuring that the model's parameters were optimized to balance costs and service levels effectively. Finally, the test set served as an independent evaluation dataset, enabling an objective assessment of the proposed model's performance in unseen scenarios against the benchmarks described in Section 3.3. A summary statistic of the data adopted for the experiments is reported in Table 3, where the min, mean, max, standard deviation, and quantiles of components' order lot size is reported and the lead time, purchasing cost, and defect are computed separately for the training (Tr), validation (Va) and test set (Te).

Table 3: Dataset description

STATISTICS	ORDER LOT SIZE	LEAD TIME			COST			DEFECT		
		Tr	Val	Te	Tr	Val	Te	Tr	Val	Te
Average Mean	5,8	46,52	43,68	43,15	68,38	70,08	69,98	0,02	0,00	0,02
Average Dev std	3.0	9,00	1,77	6,72	4,00	3,45	3,46	0,06	0,06	0,40
Average Min	1	17,86	24,64	24,93	36,42	39,93	39,84	0,00	0,00	0,00
Average Q1	3	42,51	42,52	42,55	68,31	69,98	69,94	0,02	0,00	0,00
Average Q2	6	43,23	43,13	42,62	68,76	69,98	69,98	0,02	0,00	0,00
Average Q3	9	45,98	44,10	43,24	68,77	69,99	69,99	0,02	0,00	0,00
Average Max	10	79,89	50,22	47,58	73,33	72,67	70,39	1,46	1,09	6,97

The computational experiments were conducted on a machine with the following specifications: Intel(R) Core(TM) i9-10900 CPU @ 2.80GHz (2.81 GHz), 64.0 GB of RAM (63.8 GB usable), running a 64-bit operating system based on x64 architecture. The system was powered by Windows 10 Pro, version 22H2. The code was developed using Python and Darts; Pyomo and Optuna libraries were adopted, respectively, to develop the forecasting, optimization, and tuning models. Lastly, the hyperparameter adopted respectively for the CatBoost, ARIMA, and NMA forecasting models and for the Bayesian algorithm adopted in the safety stock and safety lead time tuning phase are reported in Table 4. On the other hand, the hyperparameter related to the Bayesian algorithm has been selected recurring to the elbow method (Racolte et al., 2022) to guarantee optimal results balancing at the same time computational cost.

Table 4: Hyperparameter values adopted for the forecasting and the tuning models

MODEL	**HYPERPARAMETER**	**HYPERPARAMETER VALUE**
NMA	Window	5
ARIMA	P	1-5
	D	0-2
	Q	0-2
CatBoost	Input chunk length	5
	N estimators	100
	Learning rate	0.1
	Max depth	No limit
	Num leaves	31
Bayesian Optimization	N trials	20
Safety stock analytical formulation	z	95%

4. Results

This Section provides the results of the experimental design built relying on the real automotive data described in Section 3.43.44. First, the results of the comparison between the cost generated by the decisions reported by the proposed approach and the decisions reported by a perfect decision maker (Oracle approach) are reported. Then, a comparison between the cost generated by the proposed approach and the approach guided by different forecasting technologies (NMA Approach, ARIMA Approach, and CatBoost Approach) is presented. Afterward, the difference between the cost generated from the proposed approach and the cost generated when safety stocks are computed relying on a formula-based Approach is analyzed. Lastly, the average computational costs required by the proposed and the benchmarks to solve the inventory problem are summarized.

4.1. Benchmarking against a perfect decision-maker

Figure 2 illustrates the distribution of cost differences generated over the 94 examined components between the total ordering and inventory costs obtained following the proposed approach and those generated by the Oracle Approach. As previously described in Section 3.33, the Oracle Approach assume perfect knowledge of the future lead time, cost, and defect rates and thus the resulting cost represent the total inventory lower bound cost. The histogram and boxplot thus provide a detailed view of how closely the cost realized by the proposed approach are to the minimum obtainable ones.

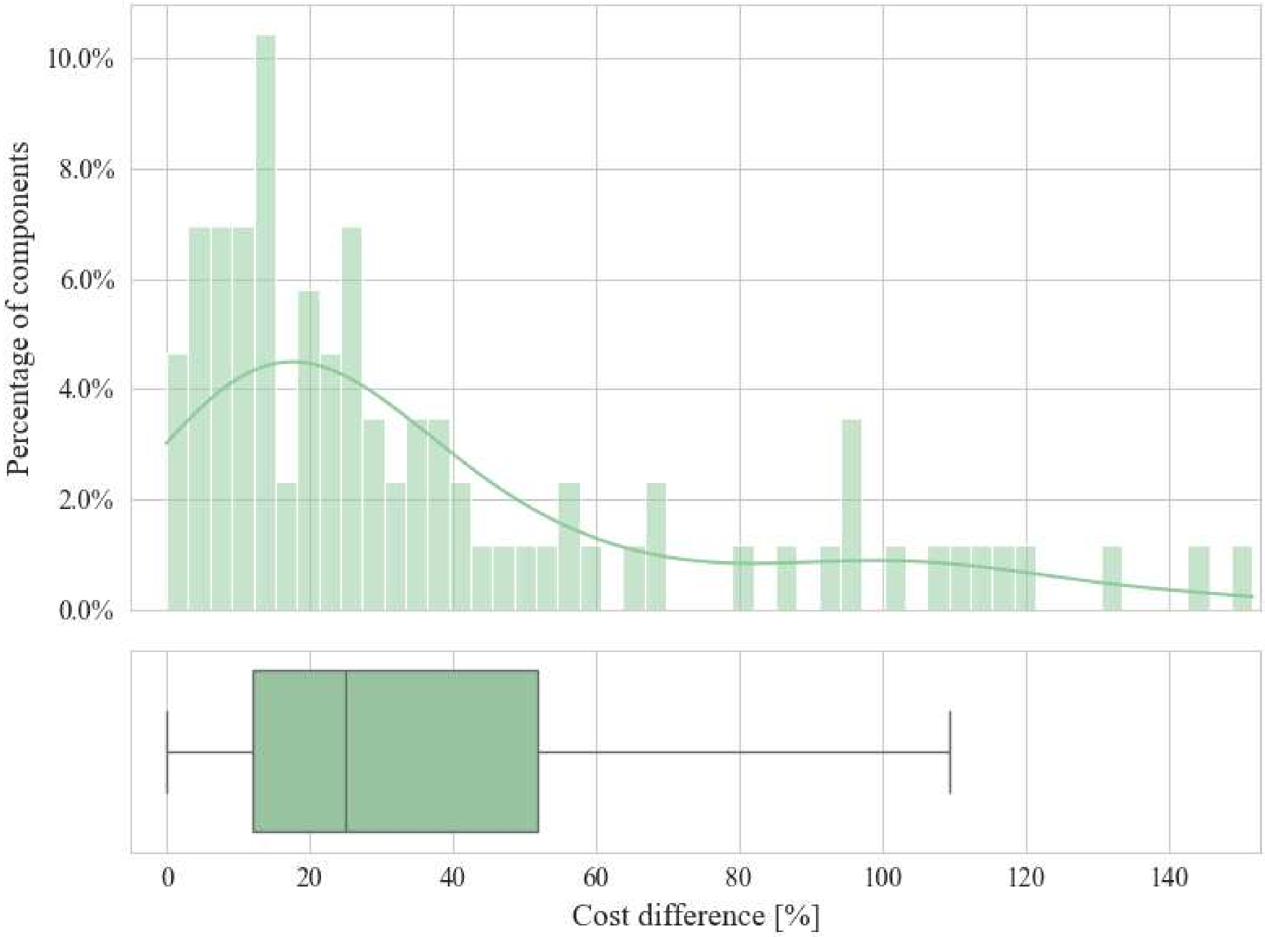

Figure 2: Distribution of the cost difference between the Proposed approach and the Oracle approach

The histogram indicates that a substantial share of components exhibits relatively contained cost differences between the proposed approach and the Oracle benchmark. In particular, around one quarter of the components show differences below 13%, and the median value of approximately 26% confirms that, for most items, the proposed strategy achieves outcomes that remain reasonably close to the ideal decisions of a perfect-information planner. The boxplot further highlights that the central portion of the distribution is compact, with limited dispersion around the median. Although a small number of components display higher deviations—reflected by an extended upper tail—these cases represent isolated situations rather than the general trend and are likely associated with items whose behaviors is inherently more difficult to predict or manage.

4.2. Comparing the effectiveness of different forecasting technologies for the proposed approach

Figure 3 presents a histogram and a boxplot summarizing the effect of replacing the proposed forecasting component, namely CatBoost trained with the custom loss function, with alternative forecasting algorithms: NMA, ARIMA, and CatBoost trained with a traditional loss function. The comparison is performed within the same inventory-control framework, thus isolating the effect of the forecasting model on total ordering and inventory costs. The x-axis reports the percentage cost difference between the proposed approach and the corresponding benchmark-based variant. Under this convention, negative values indicate cases in which the proposed approach achieves lower total cost, whereas positive values indicate cases in which the benchmark forecasting model achieves lower cost. The y-axis reports the percentage of components in each cost-difference interval. The overlaid kernel-density curves summarize the empirical distributions..

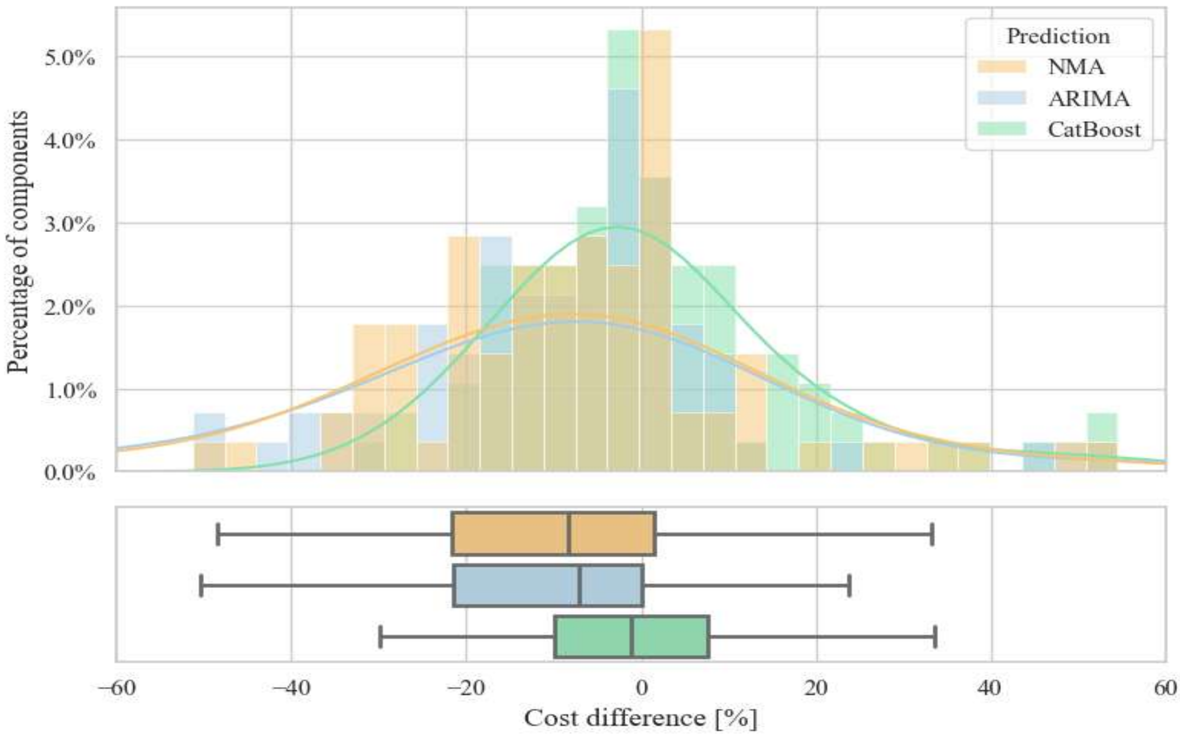

Figure 3: Distribution of the cost difference between the proposed approach and the same approach when guided by different forecasting algorithms

According to Figure 3, the distributions associated with NMA and ARIMA are shifted toward negative values, indicating that replacing the proposed forecasting component with these classical methods generally leads to higher total inventory costs. These descriptive patterns are confirmed by the statistical analysis in Table 5..

Table 5: Statistical significance of forecasting model choiche on generated inventory cost

Examined Comparison	N Samples	Cliff delta	P value	Significant Difference
Proposed VS NMA	94	0,4	0,000043	True
Proposed VS ARIMA	94	0,5	0,000002	True
Proposed VS CatBoost	94	0,1	0,567799	False

The comparisons with NMA and ARIMA both show statistically significant differences, with Cliff's delta values between 0.40 and 0.50, demonstrating that the proposed forecasting model produces lower inventory costs for most components. This supports the value of the proposed forecasting design relative to classical time-series benchmarks. The comparison with standard CatBoost is more balanced. Its distribution contains observations on both sides of zero and is centered close to the no-difference region, although the median lies slightly on the negative side. This indicates that the proposed custom-loss CatBoost provides lower or equal cost for a substantial share of components, while standard CatBoost performs better for others. Consistently, the difference between the proposed approach and standard CatBoost is not statistically significant in Table 5. Therefore, Figure 3 should not be interpreted as showing a general advantage of the traditional CatBoost-based strategy. Rather, it shows that the two CatBoost-based variants perform similarly overall, while the custom-loss formulation yields cost reductions for specific components located on the negative side of the distribution. These components correspond to cases in which realized inventory cost is more sensitive to the forecasting-loss specification. In this sense, Figure 3 provides an ablation analysis of the proposed framework, highlighting both the contribution of ML-based forecasting relative to classical benchmarks and the component-level effect of aligning forecasting loss with inventory-cost asymmetry.

To note that these cost observations align with the forecasting accuracy distribution reported in Figure 4.

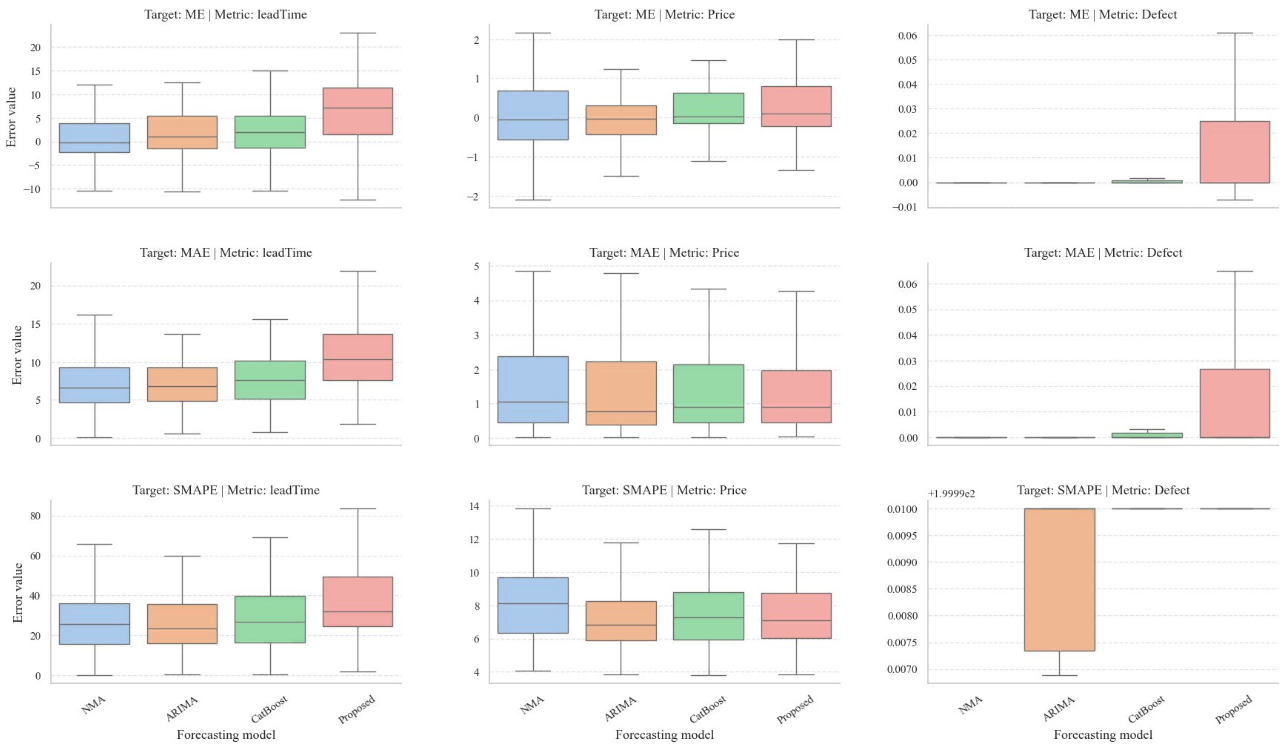

Figure 4: Forecasting accuracy distribution of the different analyzed forecasting algorithms for each considered target variable according to different accuracy metrics

An analysis of the error distributions across the forecasting models highlights how the proposed approach behaves differently than the benchmarks. For lead time and price, median errors across all models remain close, confirming that the proposed method retains solid predictive accuracy despite adopting a loss function that prioritizes operational cost asymmetry rather than pure point precision. This naturally yields a slightly broader error spread for the proposed model, but this behavior is consistent with its design objective: to guide inventory decisions rather than optimise classical statistical error metrics. For defect forecasting, the plots show that the proposed model tends to introduce a controlled positive bias, whereas the other methods collapse near zero. This reflects a purposeful advantage of the custom loss: by slightly overestimating defect quantities, the model systematically reduces the risk of under-protection, which is far more costly in inventory settings than mild over-protection. Thus, what appears as “higher error” under classical metrics actually translates into safer and more cost-effective decisions.

4.3 Benchmarking the effectiveness of the proposed safety stock dimensioning method against a traditional analytical safety stock dimensioning model

Figure 5 reports the distribution of the cost difference generated between the proposed approach when guided with the safety stock and lead time tuning module, which is reported in Section 3.2.3, and the traditional safety stock analytical formula reported in Section 3.33.

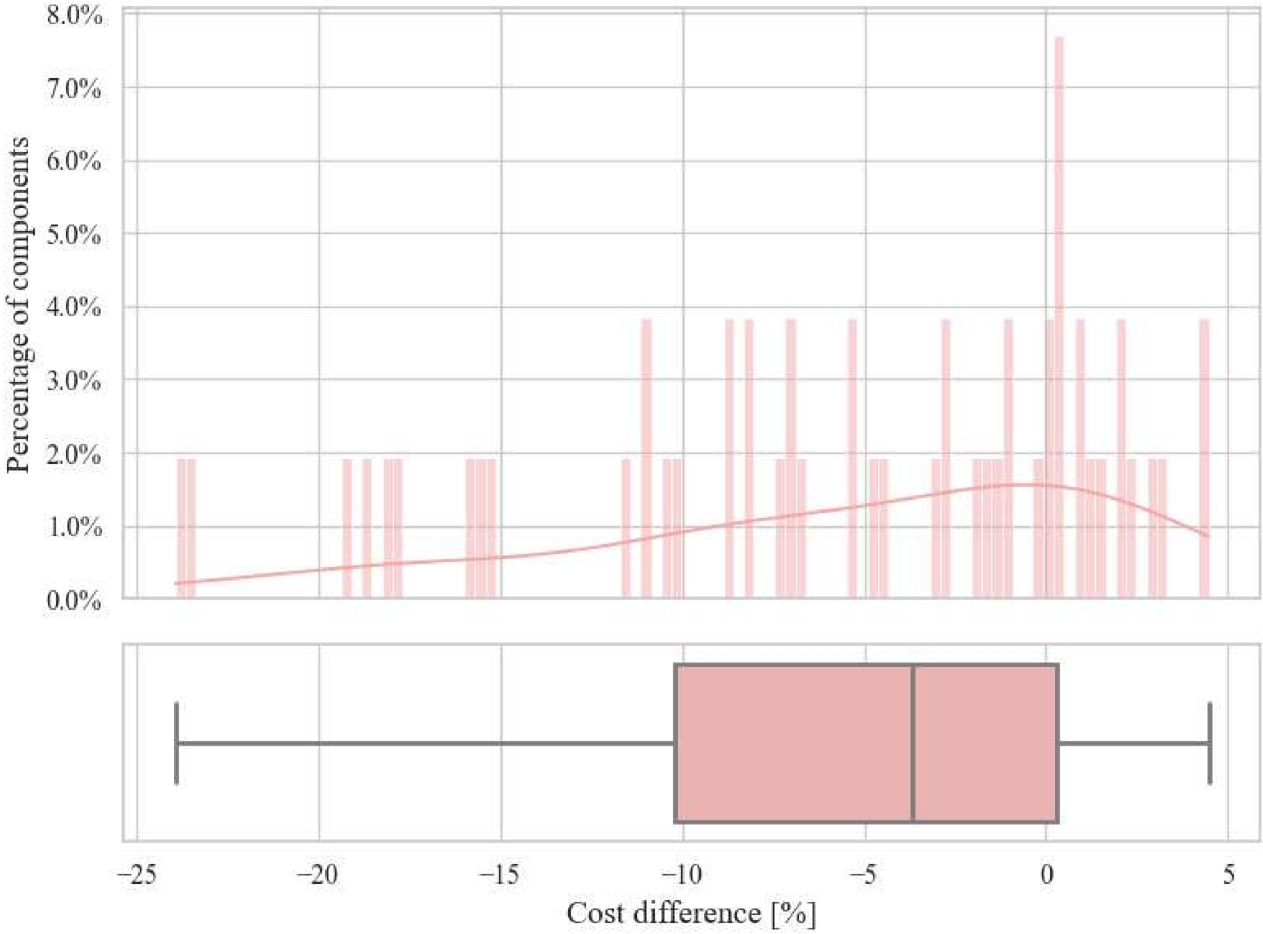

Figure 5: Distribution of the cost difference between the proposed approach and the same approach when guided by different safety stock dimensioning methods

The histogram shows that most components exhibit negative cost differences, confirming that the proposed approach generally achieves lower inventory costs than the traditional formula-based method. The distribution is centered below zero, and the highest concentration of observations lies in this cost-saving region. Although some variability is present, the extended left tail highlights those components for which the proposed tuning mechanism delivers substantial additional reductions relative to the analytic formula. These visual patterns are supported by the statistical evidence in Table 6.

Table 6: Statistical significance of safety stock and safety lead time computation over inventory cost

Examined Comparison	N Samples	Cliff delta	P value	Significant Difference
Proposed VS Formula based	94	0,3	0,00003	True

The comparison yields a statistically significant difference, with a Cliff's delta of 0.3, indicating that the proposed method outperforms the formula-based approach for a clear majority of components. Overall, the results demonstrate that the safety stock and safety lead time tuning module provides consistently superior cost performance compared to traditional analytic dimensioning.

4.4 Benchmarking the computational cost of the different approaches

Table 7 compares the average computational time required by the proposed approach and three benchmark strategies over the 94 observed components. The table provides a detailed breakdown of the computational time associated with each process step, including forecasting model training, safety stock and lead time dimensioning, inventory control optimization, and the overall computational time.

Table 7: Comparison of the average computational time of the proposed and benchmark approach for the observed components

Approach	Average forecasting training time [s]	Average safety stock and lead time dimensioning time[s]	Average inventory control optimization time [s]	Average overall computational time [s]
Proposed approach	6.4	600	20	626.4
CatBoost approach	3.2	600	20	623.2
ARIMA approach	2.5	600	20	622.5
NMA approach	1.4	600	20	621.4
Formula based approach	3.2	1	20	24.2

According to the table, the proposed approach requires 626.4 seconds, with most of the time allocated to safety stock and lead time dimensioning. Forecasting model training contributes only 6.4 seconds, and inventory control optimization accounts for 20 seconds. The benchmark approaches exhibit comparable computational times for safety stock and lead time dimensioning and inventory control optimization. Specifically, the NMA Approach has an overall computational time of 621.4, the ARIMA Approach takes 623.2 seconds and the approach based on traditional CatBoost takes 623.2 s. In contrast, the Formula-based approach shows a substantially lower overall computational time of only 24.2 seconds. This efficiency stems from the minimal time required for safety stock dimensioning while maintaining the same inventory control optimization time of 20 seconds as the other approaches. However, this significant reduction in computational time comes at the cost of reduced effect, as detailed in previous sections. The results in Table 7 highlight the balance between computational efficiency and methodological robustness.

5. Managerial Insight

The empirical findings derived from the real automotive case study provide several actionable implications for industrial practice. The following insights summarize the most relevant managerial lessons emerging from the analysis.

Insight 1: Joint optimization of replenishment, safety stock, and safety lead time can improve cost performance, with benefits varying across components. The results indicate that integrating safety stock and safety lead time tuning within the replenishment optimization process can lead to improved cost performance, although the magnitude of the benefit is not uniform across all components. In particular, the comparison with the traditional analytical formula shows a median cost reduction of approximately 4%, with larger improvements observed for a subset of components (exceeding 23%). These findings suggest that treating safety buffers as decision variables, rather than fixed parameters, allows the system to better adapt to variability in lead time and defect rates. However, the gains are context-dependent and more pronounced for components characterized by higher uncertainty or operational relevance. For managers, this implies that integrated decision frameworks can be particularly valuable in targeted applications, rather than universally replacing simpler formula-based approaches.

Insight 2: Forecast bias aligned with cost asymmetry can enhance economic performance, even when statistical accuracy is comparable.

The results suggest that aligning forecasting behavior with the asymmetric cost structure of inventory systems can enhance economic performance, even when traditional accuracy metrics (e.g., SMAPE, RMSE) are comparable across models. Although the proposed forecasting model is not consistently more accurate in statistical terms, the adoption of a custom loss introduces a controlled positive bias in lead time and defect predictions, which contributes to more cost-effective decisions. Empirically, replacing this model with NMA or ARIMA leads to cost increases of approximately 9%–12% for most components, while differences with standard CatBoost remain limited and not statistically significant. These findings indicate that shaping forecast errors in a cost-consistent direction can provide tangible operational benefits, particularly by reducing exposure to high-impact stockouts, whose costs typically outweigh incremental holding costs. While the magnitude of this effect varies across components, the results consistently show that incorporating cost asymmetry into the forecasting process can improve decision quality. For managers, this highlights the importance of complementing statistical accuracy metrics with an evaluation of the economic consequences of forecast errors as suggested also in Kourentzes et al., (2020), especially in contexts where underestimation carries significant risk.

Insight 3: Data-driven decision-making can achieve near-optimal outcomes without requiring perfect information. Comparison against a hypothetical perfect decision-maker (Oracle Approach) shows that more than 50% of components exhibit a cost deviation below 26%, with 25% of components remaining within 12% of the optimal benchmark. These results demonstrate that organizations can achieve near-optimal performance with realistic data quality and without needing perfect information about future lead times, costs, or defect rates. For managers, this reinforces the value of deploying data-driven inventory systems even when forecasting noise and variability are unavoidable.

Insight 4: Computationally intensive tuning is justified for high-impact components, suggesting a tiered implementation strategy. The proposed method requires approximately 626 seconds per component—primarily driven by the Bayesian tuning process—while simpler benchmarks require 621–623 seconds, and formula-based approaches only 24 seconds. Despite this, the economic benefits observed in Sections 4.1–4.3 justify the computational investment for components where stockout or holding costs are substantial. A selective deployment strategy is therefore recommended: use the full proposed approach for high-value, high-variability, or safety-critical items, and employ simpler formula-based methods for stable or low-cost components. This targeted application ensures cost-efficiency while maintaining high performance where it matters most.

Insight 5: Forecasting, safety parameter tuning, and optimization should be treated as an integrated decision pipeline rather than independent tasks. The study highlights that the performance gains of the proposed method do not arise from any individual module alone but from the interaction between accurate (and intentionally biased) forecasts, data-driven safety buffer calibration, and MILP-based replenishment optimization. Traditional processes often segment these activities—forecasting, then safety stock definition, then ordering decisions—leading to suboptimal outcomes. Managers should therefore transition toward integrated workflows that explicitly couple forecasting characteristics with safety parameter tuning and optimization, enabling more coherent and cost-effective decision-making across the inventory planning horizon.

Overall, these insights emphasize that effective inventory management in uncertain environments requires integrated, data-driven decision processes that jointly align forecasting behaviors, safety parameter calibration, and optimization logic to achieve consistently superior cost performance.

6. Conclusions

Effective inventory management requires jointly deciding when to order, how much to order, and how to set safety stock and safety lead time. This study proposed an ML-based Parametrized Deterministic Lookahead Model that integrates replenishment optimization with safety stock and safety lead time calibration, while accounting for time-varying purchasing costs, lead times, and defect rates through tailored forecasting models.

The experimental results on 94 relevant automotive components show that the proposed approach generates cost outcomes close to those of a perfect decision-maker, with cost differences below 26% for most components. The analysis also shows that forecasting model choice materially affects cost performance: replacing the proposed forecasting component with NMA, ARIMA, or CatBoost trained with traditional loss leads to higher inventory costs. These gains do not stem from uniformly better statistical accuracy, but from a better alignment between forecast behavior and the asymmetric operational consequences of forecast errors. In addition, the proposed safety stock and safety lead time tuning module outperforms the traditional analytical benchmark, with a median cost improvement of about 4%.

These results support the value of integrating forecasting, safety-parameter tuning, and replenishment optimization within a single decision framework. At the same time, the higher computational burden suggests that the full approach is especially suitable for critical components, while simpler methods may remain preferable for less relevant items.

Although the proposed approach demonstrates solid performance and provides valuable managerial insights, several limitations must be acknowledged. First, the computational intensity of the framework—particularly the iterative tuning of safety stock and safety lead time—represents a key limitation. As discussed in Section 4.4, the overall computational burden is largely dominated by this dimensioning phase, while performance differences across forecasting-based approaches remain marginal. Although the method is tractable for medium-scale applications or high-value components, this characteristic may hinder scalability in large-scale settings involving thousands of SKUs. Improving the efficiency of the safety parameter tuning process is therefore a critical area for further development. Promising directions include surrogate-assisted optimization, adaptive or reduced search spaces, and hybrid heuristic–optimization strategies aimed at accelerating convergence while preserving solution quality. Second, the approach relies on historical data to train forecasting models and calibrate safety parameters. In highly volatile or structurally evolving environments, historical patterns may not adequately capture emerging dynamics, potentially reducing the robustness of both predictions and subsequent optimization outcomes. Extending the framework to incorporate real-time data streams, adaptive learning mechanisms, and concept drift detection techniques represents a promising avenue to enhance performance in dynamic contexts. Third, the empirical evaluation is based on a single real-world dataset from an automotive company characterized by medium-to-long procurement lead times (approximately 40 days). Under such conditions, the demand aggregated over lead time is reasonably approximated by a Gaussian distribution, which may partially favor classical assumptions embedded in inventory control logic. Consequently, the findings may not directly generalize to environments with significantly shorter lead times, where demand over lead time may exhibit stronger intermittency, skewness, or heavy-tailed behavior. The performance and robustness of the proposed framework under such conditions remain to be systematically assessed. Finally, while the dataset is rich and representative of a complex industrial setting, it does not fully capture the diversity of supply chain environments across industries. Broader empirical validation—such as multi-industry case studies, cross-company datasets, and controlled simulation experiments—would strengthen the external validity and generalizability of the results.

Future research should therefore focus on (i) improving the computational efficiency of the safety parameter tuning phase, (ii) extending the framework toward adaptive and real-time learning settings, and (iii) systematically evaluating performance under alternative demand–lead time regimes, particularly in short lead time contexts with non-Gaussian and heavy-tailed demand distributions. These developments would further enhance the scalability, robustness, and practical applicability of the proposed methodology in complex and data-rich supply chain environments.

Declaration of generative AI and AI-assisted technologies in the writing process.

During the preparation of this work the author(s) used ChatGPT in order to improve the readability of the manuscript. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Data availability statement

The data that support the findings of this study are available from the corresponding author, [M.G.], upon reasonable request.

Declaration of no competing interests

The authors declare no competing interests.

Funding

Funding: This study was carried out within the MICS (Made in Italy – Circular and Sustainable) Extended Partnership and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.3 – D.D. 1551.11-10-2022, PE00000004). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

Bibliography

Ahmed, T., Karmaker, C. L., Nasir, S. B., Moktadir, M. A., & Paul, S. K. (2023). Modeling the artificial intelligence-based imperatives of industry 5.0 towards resilient supply chains: A post-COVID-19 pandemic perspective. *Computers & Industrial Engineering*, *177*, 109055. https://doi.org/10.1016/J.CIE.2023.109055

Alecchang, C. (1985). The Interchangeability of Safety Stocks and Safety Lead Time. In *JOURNAL OF OPERATIONS MANAGEMENT* (Vol. 6).

Alicke, K. (2005). *Planung und Betrieb von Logistiknetzwerken*. Springer.

Badakhshan, E., & Ball, P. (2023). Deploying hybrid modelling to support the development of a digital twin for supply chain master planning under disruptions. *International Journal of Production Research*. https://doi.org/10.1080/00207543.2023.2244604

Ban, G. Y., Gallien, J., & Mersereau, A. J. (2019). Dynamic procurement of new products with covariate information: The residual tree method. *Manufacturing and Service Operations Management*, *21*(4), 798–815. https://doi.org/10.1287/msom.2018.0725

Baryannis, G., Validi, S., Dani, S., & Antoniou, G. (2019). Supply chain risk management and artificial intelligence: state of the art and future research directions. In *International Journal of Production Research* (Vol. 57, Number 7, pp. 2179–2202). Taylor and Francis Ltd. https://doi.org/10.1080/00207543.2018.1530476

Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, *54*(3), 1937–1967. https://doi.org/10.1007/s10462-020-09896-5

Beutel, A. L., & Minner, S. (2012). Safety stock planning under causal demand forecasting. *International Journal of Production Economics*, *140*(2), 637–645. https://doi.org/10.1016/j.ijpe.2011.04.017

Buzacott, J. A., & Shanthikumar, J. G. (1994a). Safety Stock versus Safety Time in MRP Controlled Production Systems. In *Source: Management Science* (Vol. 40, Number 12).

Buzacott, J. A., & Shanthikumar, J. G. (1994b). *Safety Stock versus Safety Time in MRP Controlled Production Systems* (Vol. 40, Number 12). https://about.jstor.org/terms

Calabrese, F., Regattieri, A., Gabellini, M., Caporale, A., & Epifania, P. (2023). Condition Monitoring of CNC machines: machining process classification through Temporal Convolutional Networks. *28th ISSAT International Conference on Reliability and Quality in Design, RQD 2023*, 51–55. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85174318267&partnerID=40&md5=02cbbbcc96e5afe5e9fe3a90dd585cca

Cannas, V. G., Ciano, M. P., Saltalamacchia, M., & Secchi, R. (2023). Artificial intelligence in supply chain and operations management: a multiple case study research. *International Journal of Production Research*, 1–28. https://doi.org/10.1080/00207543.2023.2232050

Chen, Y., Pekny, J. F., & Reklaitis, G. V. (2013). Integrated planning and optimization of clinical trial supply chain system with risk pooling. *Industrial and Engineering Chemistry Research*, *52*(1), 152–165. https://doi.org/10.1021/ie300823b

Chung, S. H., Kang, H. Y., & Pearn, W. L. (2005). A service level model for the control wafers safety inventory problem. *International Journal of Advanced Manufacturing Technology*, *26*(5–6), 591–597. https://doi.org/10.1007/s00170-003-2028-9

Clausen, J. B. B., & Li, H. (2022). Big data driven order-up-to level model: Application of machine learning. *Computers and Operations Research*, *139*. https://doi.org/10.1016/j.cor.2021.105641

Culot, G., Podrecca, M., & Nassimbeni, G. (2024). Artificial intelligence in supply chain management: A systematic literature review of empirical studies and research directions. In *Computers in Industry* (Vol. 162). Elsevier B.V. https://doi.org/10.1016/j.compind.2024.104132

de Armas, J., & Laguna, M. (2020). Parallel machine, capacitated lot-sizing and scheduling for the pipe-insulation industry. *International Journal of Production Research*, *58*(3), 800–817. https://doi.org/10.1080/00207543.2019.1600763

Demizu, T., Fukazawa, Y., & Morita, H. (2023). Inventory management of new products in retailers using model-based deep reinforcement learning. *Expert Systems with Applications*, *229*. https://doi.org/10.1016/j.eswa.2023.120256

Dorogush, A. V., Ershov, V., & Gulin, A. (2018). *CatBoost: gradient boosting with categorical features support*. http://arxiv.org/abs/1810.11363

Gabellini, M., Calabrese, F., Civolani, L., Regattieri, A., & Galizia, F. G. (2023). A predictive data-driven approach for supply chain quality risks in the automotive sector. *Proceedings of the Summer School Francesco Turco*. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85193746650&partnerID=40&md5=c52a7125ed7e170e4791ac09d49067aa

Gabellini, M., Calabrese, F., Civolani, L., Regattieri, A., & Mora, C. (2023). A data-driven approach to predict supply chain risk due to suppliers' partial shipments. *International Conference on Sustainable Design and Manufacturing*, 227–237.

Gabellini, M., Calabrese, F., Civolani, L., Regattieri, A., & Mora, C. (2024). A Data-Driven Approach to Predict Supply Chain Risk Due to Suppliers' Partial Shipments. In S. G. Scholz, R. J. Howlett, & R. Setchi (Eds.), *Smart Innovation, Systems and Technologies* (Vol. 377, pp. 227–237). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-981-99-8159-5_20

Gabellini, M., Calabrese, F., Regattieri, A., & Ferrari, E. (2022). Multivariate multi-output LSTM for time series forecasting with intermittent demand patterns. *Proceedings of the Summer School Francesco Turco*. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85176726507&partnerID=40&md5=57a91bccffa761dad0e449d27de71820

Gabellini, M., Calabrese, F., Regattieri, A., Loske, D., & Klumpp, M. (2024). A hybrid approach integrating genetic algorithm and machine learning to solve the order picking batch assignment problem considering learning and fatigue of pickers. *Computers and Industrial Engineering*, *191*. https://doi.org/10.1016/j.cie.2024.110175

Gabellini, M., Civolani, L., Calabrese, F., & Bortolini, M. (2024). A Deep Learning Approach to Predict Supply Chain Delivery Delay Risk Based on Macroeconomic Indicators: A Case Study in the Automotive Sector. *Applied Sciences (Switzerland)*, *14*(11). https://doi.org/10.3390/app14114688

Gabellini, M., Civolani, L., Regattieri, A., & Calabrese, F. (2023). A Data Model for Predictive Supply Chain Risk Management. *Lecture Notes in Mechanical Engineering*, 365 – 372. https://doi.org/10.1007/978-3-031-34821-1_40

Gabellini, M., Mak, S., Schoepf, S., Brintrup, A., & Regattieri, A. (2025). A continuous training approach for risk informed supplier selection and order allocation. *Production & Manufacturing Research*, *13*(1), 2447035.

Ganesh, A. D., & Kalpana, P. (2022). Future of artificial intelligence and its influence on supply chain risk management A systematic review. *Computers &amp Industrial Engineering*, *169*, 108206. https://doi.org/10.1016/j.cie.2022.108206

Gonçalves, J. N. C., Sameiro Carvalho, M., & Cortez, P. (2020). Operations research models and methods for safety stock determination: A review. In *Operations Research Perspectives* (Vol. 7). Elsevier Ltd. https://doi.org/10.1016/j.orp.2020.100164

Grover, P., Kar, A. K., & Dwivedi, Y. K. (2022). Understanding artificial intelligence adoption in operations management: insights from the review of academic literature and social media discussions. *Annals of Operations Research*, *308*(1–2), 177–213. https://doi.org/10.1007/s10479-020-03683-9

Gudehus, T. (2006). *Dynamische Disposition: Strategien zur optimalen Auftrags-und Bestandsdisposition*. Springer.

Helo, P., & Hao, Y. (2021). Artificial intelligence in operations management and supply chain management: an exploratory case study. *Production Planning &amp Control*, *33*(16), 1573–1590. https://doi.org/10.1080/09537287.2021.1882690

Herrmann, F. (2011). *Operative Planung in IT-Systemen für die Produktionsplanung und-steuerung: Wirkung, Auswahl und Einstellhinweise von Verfahren und Parametern*. Springer-Verlag.

Huber, J., Müller, S., Fleischmann, M., & Stuckenschmidt, H. (2019). A data-driven newsvendor problem: From data to decision. *European Journal of Operational Research*, *278*(3), 904–915. https://doi.org/10.1016/j.ejor.2019.04.043

Hung, Y.-F., & Chang, C.-B. (1999). Determining safety stocks for production planning in uncertain manufacturing. In *Int. J. Production Economics* (Vol. 58).

Imani, M., Imani, M., & Ghoreishi, S. F. (2022). Bayesian Optimization for Expensive Smooth-Varying Functions. *IEEE Intelligent Systems*, *37*(4), 44–55. https://doi.org/10.1109/MIS.2022.3163227

Inderfurth, K. (1995). Multistage safety stock planning with item demands correlated across products and through time. *Production and Operations Management*, *4*(2), 127–144.

Jaggi, C. K., & Aggarwal, S. P. (1994). Credit financing in economic ordering policies of deteriorating items. *International Journal of Production Economics*, *34*(2), 151–155.

Janssens, G. K., & Ramaekers, K. M. (2011). A linear programming formulation for an inventory management decision problem with a service constraint. *Expert Systems with Applications*, *38*(7), 7929–7934. https://doi.org/10.1016/j.eswa.2010.12.009

Kassa, A., Kitaw, D., Stache, U., Beshah, B., & Degefu, G. (2023). Artificial intelligence techniques for enhancing supply chain resilience: A systematic literature review, holistic framework, and future research. *Computers and Industrial Engineering*, *186*. https://doi.org/10.1016/j.cie.2023.109714

Keskin, G. A., Omurca, S. I., Aydin, N., & Ekinci, E. (2015). A comparative study of production-inventory model for determining effective production quantity and safety stock level. *Applied Mathematical Modelling*, *39*(20), 6359–6374. https://doi.org/10.1016/j.apm.2015.01.037

Kourentzes, N., Trapero, J. R., & Barrow, D. K. (2020). Optimising forecasting models for inventory planning. *International Journal of Production Economics*, *225*. https://doi.org/10.1016/j.ijpe.2019.107597

Lin, S., Chen, Y., Li, Y., & Shen, Z. J. M. (2022). Data-Driven Newsvendor Problems Regularized by a Profit Risk Constraint. *Production and Operations Management*, *31*(4), 1630–1644. https://doi.org/10.1111/poms.13635

Louly, M. A., & Dolgui, A. (2013). Optimal MRP parameters for a single item inventory with random replenishment lead time, POQ policy and service level constraint. *International Journal of Production Economics*, *143*(1), 35–40. https://doi.org/10.1016/j.ijpe.2011.02.009

Louly, M. A. O., & Dolgui, A. (2009). Calculating safety stocks for assembly systems with random component procurement lead times: A branch and bound algorithm. *European Journal of Operational Research*, *199*(3), 723–731. https://doi.org/10.1016/j.ejor.2007.11.066

Mete Ayhan, H., & Kır, S. (2024). Ml-driven approaches to enhance inventory planning: Inoculant weight application in casting processes. *Computers and Industrial Engineering*, *193*. https://doi.org/10.1016/j.cie.2024.110280

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

Molinder, A. (1997). Joint optimization of lot-sizes, safety stocks and safety lead times in an MRP system. *International Journal of Production Research*, *35*(4), 983–994. https://doi.org/10.1080/002075497195498

Oroojlooyjadid, A., Snyder, L. V., & Takáč, M. (2020). Applying deep learning to the newsvendor problem. *IISE Transactions*, *52*(4), 444–463. https://doi.org/10.1080/24725854.2019.1632502

Ould-Louly, M. A., & Dolgui, A. (2004). The MPS parameterization under lead time uncertainty. *International Journal of Production Economics*, *90*(3), 369–376. https://doi.org/10.1016/j.ijpe.2003.08.008

Pan, Y., & Zhang, L. (2020). Data-driven estimation of building energy consumption with multi-source heterogeneous data. *Applied Energy*, *268*. https://doi.org/10.1016/j.apenergy.2020.114965

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

Pournader, M., Ghaderi, H., Hassanzadegan, A., & Fahimnia, B. (2021). Artificial intelligence applications in supply chain management. *International Journal of Production Economics*, *241*, 108250. https://doi.org/10.1016/j.ijpe.2021.108250

Powell, W. B. (2022). Designing Lookahead Policies for Sequential Decision Problems in Transportation and Logistics. *IEEE Open Journal of Intelligent Transportation Systems*, *3*, 313–327. https://doi.org/10.1109/OJITS.2022.3148574

Prak, D., Teunter, R., & Syntetos, A. (2017). On the calculation of safety stocks when demand is forecasted. *European Journal of Operational Research*, *256*(2), 454–461. https://doi.org/10.1016/j.ejor.2016.06.035

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

Qi, M., Shi, Y., Qi, Y., Ma, C., Yuan, R., Wu, D., & Shen, Z. J. (2023). A Practical End-to-End Inventory Management Model with Deep Learning. *Management Science*, *69*(2), 759–773. https://doi.org/10.1287/mnsc.2022.4564

Racolte, G., Marques, A., Scalco, L., Tonietto, L., Zanotta, D. C., Cazarin, C. L., Gonzaga, L., & Veronez, M. R. (2022). Spherical K-Means and Elbow Method Optimizations with Fisher Statistics for 3D Stochastic DFN from Virtual Outcrop Models. *IEEE Access*, *10*, 63723–63735. https://doi.org/10.1109/ACCESS.2022.3182332

Regattieri, A., Gabellini, M., Calabrese, F., Civolani, L., & Galizia, F. G. (2024). Balancing Data Acquisition Benefits and Ordering Costs for Predictive Supplier Selection and Order Allocation. *Applied Sciences (Switzerland)*, *14*(10). https://doi.org/10.3390/app14104306

Riahi, Y., Saikouk, T., Gunasekaran, A., & Badraoui, I. (2021). Artificial intelligence applications in supply chain: A descriptive bibliometric analysis and future research directions. *Expert Systems with Applications*, *173*, 114702. https://doi.org/10.1016/J.ESWA.2021.114702

Schmidt, F. G., & Pibernik, R. (2024). Data-driven inventory control for large product portfolios: A practical application of prescriptive analytics. *European Journal of Operational Research*. https://doi.org/10.1016/j.ejor.2024.10.012

Silva, P. M., Gonçalves, J. N. C., Martins, T. M., Marques, L. C., Oliveira, M., Reis, M. I., Araújo, L., Correia, D., Telhada, J., Costa, L., & Fernandes, J. M. (2022). A hybrid bi-objective optimization approach for joint determination of safety stock and safety time buffers in multi-item single-stage industrial supply chains. *Computers and Industrial Engineering*, *168*. https://doi.org/10.1016/j.cie.2022.108095

Smyth, C., Dennehy, D., Fosso Wamba, S., Scott, M., & Harfouche, A. (2024). Artificial intelligence and prescriptive analytics for supply chain resilience: a systematic literature review and research agenda. In *International Journal of Production Research*. Taylor and Francis Ltd. https://doi.org/10.1080/00207543.2024.2341415

van der Haar, J. F., Wellens, A. P., Boute, R. N., & Basten, R. J. I. (2024). Supervised learning for integrated forecasting and inventory control. *European Journal of Operational Research*, *319*(2), 573–586. https://doi.org/10.1016/j.ejor.2024.07.004

Van Kampen, T. J., Van Donk, D. P., & Van Der Zee, D. J. (2010). Safety stock or safety lead time: Coping with unreliability in demand and supply. *International Journal of Production Research*, *48*(24), 7463–7481. https://doi.org/10.1080/00207540903348346

Venkatesh, V., Raman, R., & Cruz-Jesus, F. (2024). AI and emerging technology adoption: a research agenda for operations management. *International Journal of Production Research*, *62*(15), 5367–5377. https://doi.org/10.1080/00207543.2023.2192309