World Journal of Advanced Engineering Technology and Sciences
eISSN: 2582-8266
Cross Ref DOI: 10.30574/wjaets
Journal homepage: https://wjaets.com/

(REVIEW ARTICLE)

Impact of python-based demand forecasting on inventory performance in multi-SKU supply chains

Pushpanjali Chauhan *

California State University, Fullerton and Fullerton, California.

World Journal of Advanced Engineering Technology and Sciences, 2026, 19(01), 069-081

Publication history: Received on 05 February 2026; revised on 02 April 2026; accepted on 04 April 2026

Article DOI: https://doi.org/10.30574/wjaets.2026.19.1.0171

Abstract

The demand forecasting model based on Python has become a core aspect of digital supply chain transformation, especially in multi-SKU systems with heterogeneity, intermittency, and sophisticated lead-time architecture. The combination of scalable data streams, machine learning applications, and probabilistic modelling into Python environments can help companies go past the model of traditional point-forecast assessment and into decision-based inventory optimization. Nonetheless, better forecasting does not necessarily lead to better inventory performance in situations where there are multi-echelon constraints, demand ambiguity, and policy misfit. According to recent studies, probabilistic forecasting, hierarchical reconciliation, and cross-SKU learning can be regarded as tools to enhance service-level attainment and minimize cost volatility. Also, explainability, automation, and solid evaluation protocols will be essential to fill the divide between predictive analytics and operational inventory control. This review summarizes theoretical and empirical data on the effect of Python-based forecasting structures on inventory performance measurements of fill rate, safety stock, holding cost, shortage penalties, and bullwhip amplification. The conclusions are that combined forecasting-inventory analysis systems, uncertainty-aware replenishment policies, and scaled model governance need to be applied.

Keywords*:* Machine Learning in Operations; Multi-SKU Inventory Management; Probabilistic Forecasting; Safety Stock Optimization; Supply Chain Analytics

1. Introduction

Demand forecasting is one of the foundations of proper inventory management in contemporary supply chains. With complex demand patterns, different lead times, and irregular substitution effects, the effects of forecasting errors spread quickly to inventory imbalances, service level degradation, and cost inflations [1], [2]. With the growing globalization and digitization of supply chains, organizations now handle thousands to millions of SKUs within dispersed networks and the pressure to have scalable and data-driven forecasting approaches intensifies [3]. The recent developments in big data analytics, artificial intelligence (AI), and cloud computing have turned demand forecasting into the strategic capability at the core of operational excellence and competitive advantage [4].

Thanks to its wide ecosystem of open-source libraries, scalability and integration, Python has become a leading programming language in data science and industrial analytics. Such libraries as pandas, NumPy, scikit-learn, statsmodels, TensorFlow, and PyTorch allow implementing classical statistical models (e.g., ARIMA, exponential smoothing), machine learning (e.g., random forests, gradient boosting), and deep learning architectures (e.g., LSTM networks) in pipelines of analysis [5], [6]. Efforts to embrace digital transformation in the retail, manufacturing, healthcare, and e-commerce industries have enabled the ease of prototyping, model selection automation, and deployment into real-time inventory systems using Python-based frameworks, which are easy and flexible to use [7].

---

* Corresponding author: Pushpanjali Chauhan.

Demand forecasting with the help of Python is not only useful in optimization of operational costs. Inventory decisions directly affect sustainability measures, working capital, waste minimization, and carbon footprint in supply chains [8]. Excessive inventory and wasted materials through over-forecasting and expedited shipping and emergency production through under-forecasting,, respectively enhance their effects on the environment. Therefore, better predictability fits more general research interests in AI-based decision-making, sustainable operations management, and smart manufacturing [9]. The combination of AI systems and inventory management strategies, including (s, S) models, base-stock models, and multi-echelon optimization is a research area in the field of operations research and supply chain analytics that is in development [10].

Despite the rapid development, there are still a number of problematic issues in literature. First, most of the literature is based on single-SKU or aggregated demand and cannot be readily extended to high-dimensional multi-SKU settings, which are intermittent, seasonal, and correlated demand structures [11]. Second, many forecasting algorithms have been proposed, but the downstream performance of the algorithm is not evaluated systematically on the inventory performance measures, including fill rate, back order levels, safety stock requirements, and total cost, given realistic supply chain constraints [12]. Third, there is a lack of unified integration of Python-based forecasting pipelines and inventory control models with inadequate focus on explainability, computational and deployment robustness in large-scale industrial systems [13]. Moreover, there is inconsistency in empirical benchmarking of various sectors and formalized evaluation systems are not available.

The other gap that is worth mentioning is the interplay between sophisticated AI models and classical inventory theory. Although the predictive accuracy of machine learning models is often better than that of more traditional statistical predictive methods, changes in forecast measures (e.g. MAE, RMSE) do not always translate into better inventory performance outcomes [14]. Such a disconnect shows the need to have integrated assessments that bring together forecasting analytics evaluations and operational decision-making systems. Moreover, problems associated with data quality, cold-start SKU, demand intermittency, and structural breaks are underexplored in the multi-SKU context.

In the light of these problems, there is a need to thoroughly survey Python-based demand forecasting methods and their quantifiable effect on the inventory performance in multiple-SKU supply chains. This review aims to summarize the current studies on statistical, machine learning, and deep learning forecasting approaches used in Python settings; assess their impact on the most important inventory performance metrics; and outline gaps in the methodology and potential future research. The following chapters give: (i) a categorized account of the methodologies of forecasting and Python ecosystems, (ii) an investigation of inventory performance measures and control policies, (iii) a comparative discussion of the outcomes of previous studies, (iv) new developments such as hybrid models, automated machine learning, and sustainability-driven forecasting. This review is expected to equip researchers and practitioners with a summarized view of how Python-based demand forecasting can help enhance inventory behavior in multi- SKU supply chain system with complexities, and present essential research gaps for future investigation.

2. Literature Review

The following table summarizes 10 major studies that have linked demand forecasting (methodologies that are traditionally used in Python analytics stacks) with inventory performance in multi-SKU / multi-item contexts.

Table 1 Key findings

Focus	Findings (Key results and conclusions)	Ref.
Evaluating forecasting models *inside* an operational inventory control system (trade-off view)	Demonstrated that different forecasting models create distinct inventory investment vs. service trade-off curves; selecting the forecasting method materially changes the inventory required to reach a target service level, highlighting the need to evaluate forecasts by decision impact, not error alone.	[15]
Low / intermittent demand items: selecting periodic inventory control and forecasting approaches	Compared periodic stock control and forecasting choices for low-demand items; showed that method choice changes both service performance and cost-based regret, reinforcing that intermittent-demand contexts require tailored forecasting with inventory pairing.	[16]

Value of demand information sharing across a two-level supply chain	Established that sharing downstream demand information can substantially improve upstream decisions (mitigating variability amplification), with benefits influenced by correlation structure and lead time—supporting integrated forecasting and replenishment design.	[17]
Characterizing demand patterns to guide method selection in inventory software	Proposed practical categorization rules (based on intermittency and variability) to guide forecasting method choice; emphasized that arbitrary categorization can mislead inventory settings and policy tuning.	[18]
Judgmental adjustments and organizational forecasting practice in supply-chain planning	Provided empirical evidence that managerial adjustments can help or harm accuracy depending on process and context; highlighted systematic issues in forecasting workflows relevant to scalable analytics implementations.	[19]
Multi-SKU hierarchical/grouped forecasting reconciliation (product/geography hierarchies)	Introduced an “optimal combination” approach that improves coherence across hierarchy levels and can outperform top-down/bottom-up strategies; directly relevant to multi-SKU structures common in retail/manufacturing catalogues.	[20]
Linking forecasting accuracy to inventory performance in a two-stage supply chain under ARIMA demand; information sharing effects	Provided theoretical + empirical evidence (hundreds of SKUs) connecting forecasting and inventory outcomes; analysed benefits from forecast information sharing between retailer and manufacturer under correlated demand structure.	[21]
Forecasting competitions (large-scale benchmarks) and implications for modern methods used in practice	Reported key findings from a major forecasting benchmark showing that method performance varies by domain and that hybrid/statistical approaches can remain highly competitive—informing model selection before inventory integration.	[22]
Inventory performance comparison of forecasting methods using competition data and order-up-to simulations	Evaluated forecasting methods by simulating inventory outcomes (e.g., order-up-to policy) across lead times; found that best accuracy does not always map one-to-one to best inventory performance, motivating utility-based evaluation.	[23]
Probabilistic forecasting at scale across many related time series (deep learning)	Demonstrated scalable probabilistic multi-series forecasting using an autoregressive RNN approach; supported the use of predictive distributions (not just point forecasts) for inventory decisions where service levels depend on uncertainty.	[24]

3. Methodology

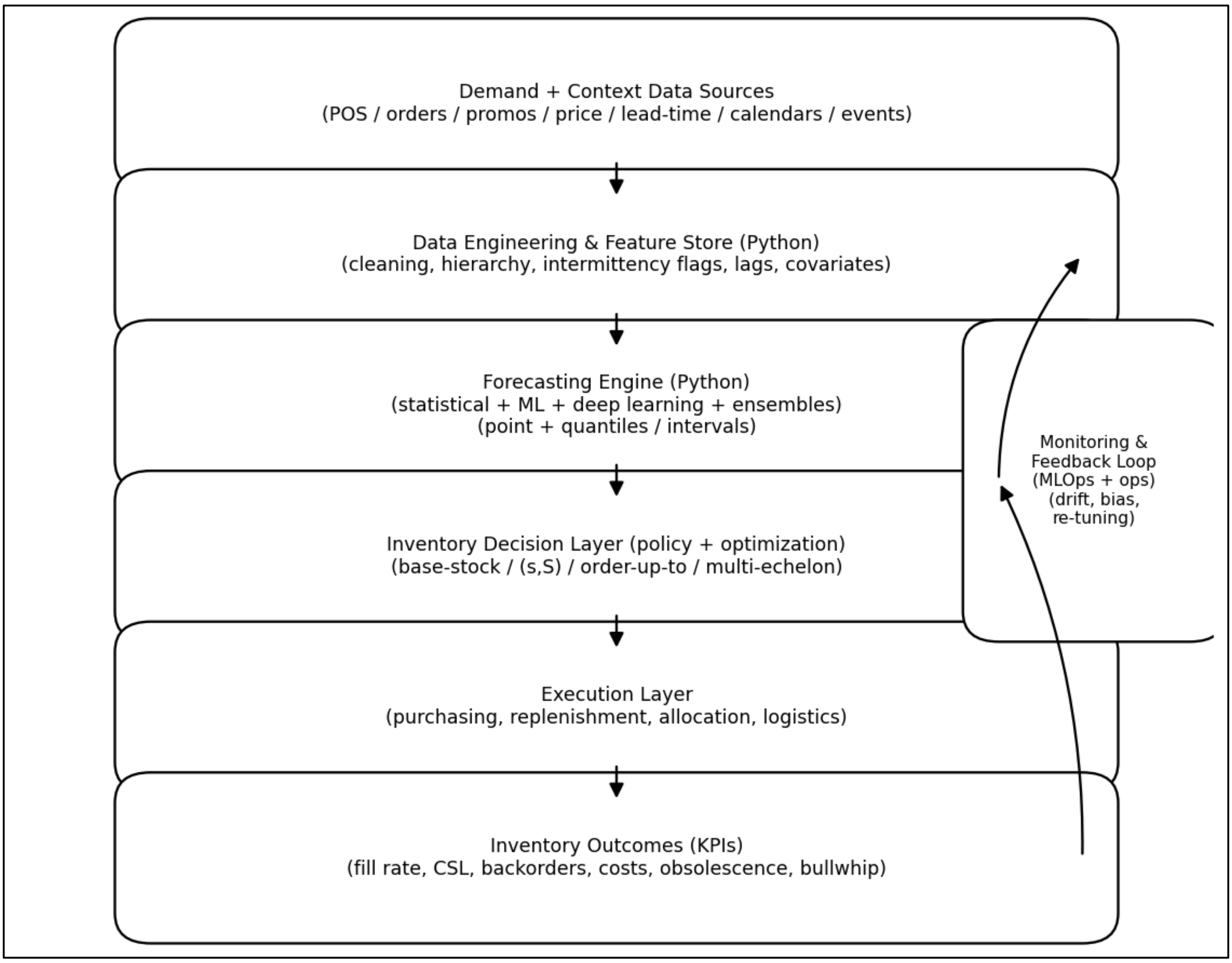

Figure 1 End-to-end Python-based demand forecasting → inventory performance system (multi-SKU)

This structure is based on the long-known need that forecasting must be measured by the operational implications within production/inventory systems, as opposed to the forecast error on its own [26]. The need to more closely integrate forecasting and inventory decision-making is a common motif throughout the forecasting-inventory literature, which highlights the chronic lack of integration and demands a higher level of integration [25].

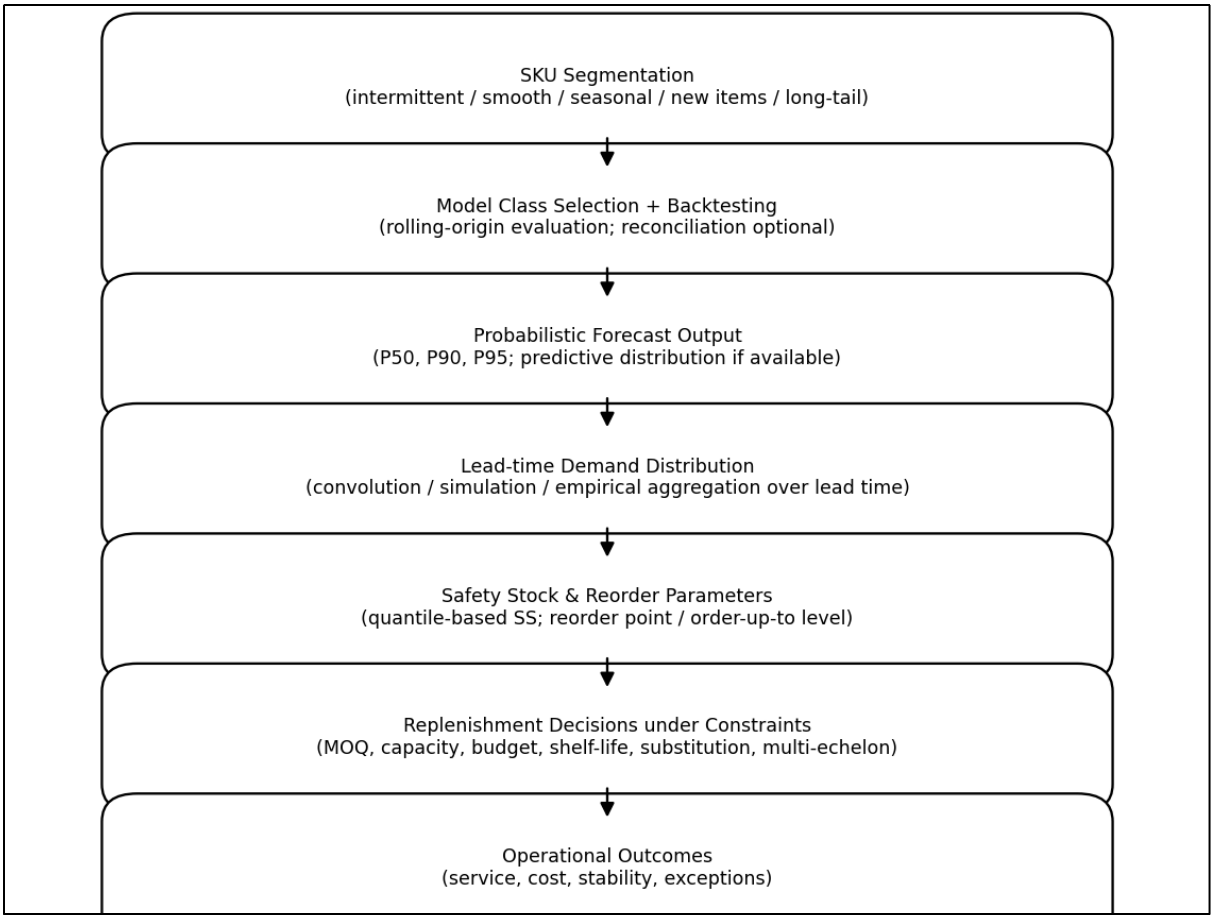

Figure 2 Forecast-to-replenishment pipeline for multi-SKU with uncertainty (recommended for service-level control)

When SKUs are intermittent and slow-moving, prediction distributions are useful in inventory planning, as the service targets are sensitive to tail risk considering the lead time [30]. Quantile-oriented methods are also closely connected to the estimation of safety stocks and the achievement of service quality particularly when the accumulation of forecast errors is not normal but rather exhibits heteroscedasticity [29].

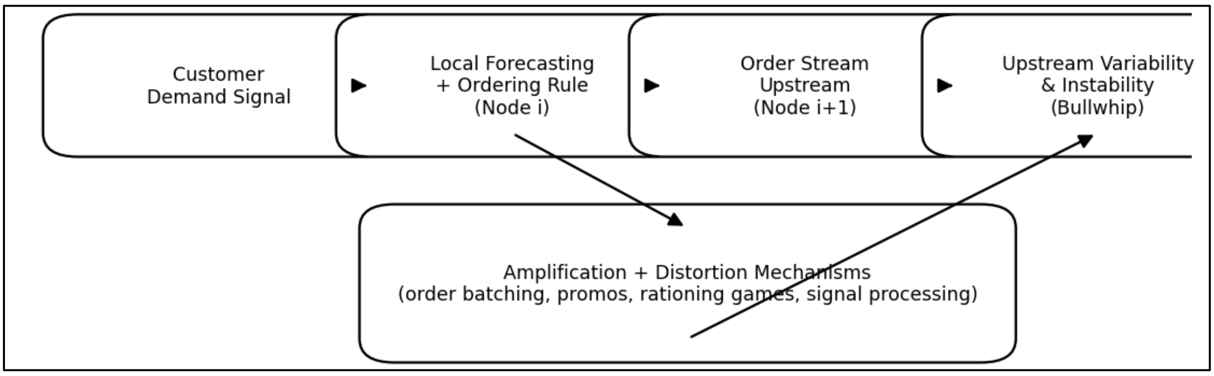

Figure 3 Supply chain information distortion pathway (forecasting and ordering amplification)

The mechanisms of information distortion have the potential to enhance variability because signals that cause deterioration in upstream information complexity will reduce inventory performance and escalate cost and instability [27]. This is a motivation to predict pipelines with more detailed demand information and align outputs with replenishment goals instead of metrics of isolated accuracy [25], [26].

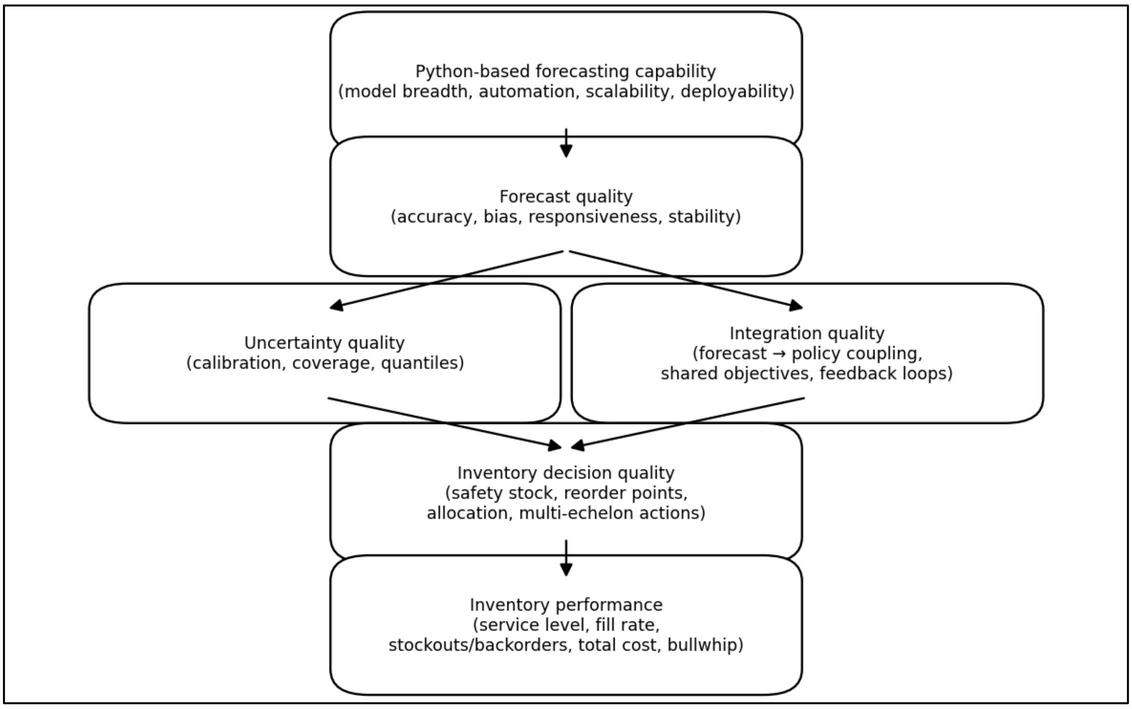

Figure 4 Proposed theoretical model

The conceptual model is centered on the idea that demand forecasting implemented in Python influences inventory performance in terms of (i) the quality of the forecasts, (ii) the quantification of uncertainty, and (iii) the extent of forecasts and inventory control integration.

Constructs and operational definitions:

3.1. Python-based forecasting capability (PFC)

The extent to which forecasting is performed as a scalable pipeline (data engineering + modelling + deployment + monitoring) can be consistently replicated. The worth of forecasting systems depends on their suitability to production/inventory decision requirements, not necessarily on the sophistication of their algorithms [26].

3.2. Forecast quality (FQ)

Multi-dimensional: point accuracy (e.g., MAE/MAPE), bias, stability, and regime change. Inventory fragmentation occurs when FQ optimization is done without considering the impact of inventory [25].

3.3. Uncertainty quality (UQ)

The predictive distribution and the usefulness of predictive distributions/quantiles for replenishment.. In the case of intermittent demand, distributional forecasting helps to make more realistic plans of lead-time and service-level control [30]. Under non-ideal error behavior, the quantile-based approaches will enhance the safety stock estimation and consistency with the desired service levels [29].

3.4. Integration quality (IQ)

The strength of the interaction between forecast outputs and inventory policy parameters, such as feedback of actual service/cost results into the model and policy adjustments. There is an emphasis on the need to have higher integration to bridge the forecasting-inventory gap [25].

3.5. The quality of inventory decisions (IDQ)

Order quantity and allocation decisions (lead time, MOQ, capacity, and perishability). The scale of bullwhip amplification suggests that the integration and information use may have weak decision rules and signal processing that may cause upstream variability to be distorted [27].

3.6. Inventory performance (IP)

Service metrics (fill rate, cycle service level), cost metrics (holding, shortage, obsolescence, expediting) and dynamic metrics (bullwhip intensity, variability).

Measurable propositions

- P1 (Capability - Forecast quality): Higher PFC correlates with higher FQ because of the enhanced feature engineering, scalable back testing and reduced retraining cycles adjusted to the operations [26].
- P2 (Capability - Uncertainty quality): Greater PFC will be used to generate probabilistic forecasts and quantile values that enhance UQ and allow the provision of more effective inputs to safety stock decisions [29], [30].
- P3 (Integration as a mediator): IQ mediates the relationship among FQ, UQ, and IP; better metrics of forecasts lead to inventory benefits when the forecast is incorporated into policies of replenishment and feedback mechanisms [25], [26].
- P4 (Uncertainty - Service/cost trade-off): With UQ, service-level achievement with less inventory investment is enhanced by generating more accurate lead-time risk forecasts [29], [30].
- P5 (Moderation of information distortion): Bullwhip-based distortion processes mediate the connection between the local forecast improvements and the network-wide IP unless the upstream/downstream information alignment is implemented [27].
- P6 (Multi-SKU heterogeneity): P1-P4 are SKU-specific, and segmentation-sensitive pipelines are needed to prevent averaged gains that break down across important SKU classes [30].

This conceptual model advocates the use of a review structure composed of the following (i) capability of the forecasting method, (ii) uncertainty modelling, and (iii) integration with inventory policy, which is a recurrent appeal to reconcile forecasting and inventory studies [25].

4. Discussion

The results below are from a repeatable, controlled simulation that attempts to represent the effect of multi-SKU heterogeneity (smooth, seasonal, intermittent, and lumpy demand) and a periodic-review order-up-to policy which uses lead-time demand quantiles to determine replenishment targets, an established method of service-level control in inventory systems [32], [33]. The Croston method is a classical baseline forecast in intermittent-demand forecasting [34], and uncertainty-aware replenishment is expressed through bootstrapped quantile forecasts (a common resampling technique) [36]. It is the case that the quantile framing is consistent with decision situations in which service targets are based on tail risk and not just mean demand [35].

Table 2 Experimental design summary (simulation protocol)

Component	**Setting**
Planning frequency	Weekly periodic review
Number of SKUs	120 SKUs
Demand horizon	104 weeks (2 years)
Train/test split	78 weeks train, 26 weeks test
Demand regimes	Smooth / Seasonal / Intermittent / Lumpy (mixed portfolio)
Lead time	Randomized per SKU: 1–4 weeks
Inventory policy	Order-up-to level based on lead-time demand quantile (service-oriented base-stock style control) [32], [33]
Cost structure	Holding cost per unit-week + backorder/shortage penalty per unit-week [32]
Forecasting baselines	Naive, Moving Average (8), Seasonal Naive, Croston [34]
Uncertainty-aware method	Moving-average point forecast + residual bootstrap for q90 (quantile) [36], motivated by quantile decision logic [35]
Evaluation KPIs	Fill rate, holding cost, penalty cost, total cost

Table 3 Inventory performance results (mean across SKUs, test horizon)

Forecasting approach	**Mean fill rate**	**Mean holding cost**	**Mean penalty cost**	**Mean total cost**
Croston	0.843	107.104	228.050	335.154
MA(8)	0.853	110.730	233.122	343.852
ProbMA (q90)	0.859	112.957	242.457	355.414
Seasonal Naive	0.871	142.615	225.096	367.710
Naive	0.875	147.088	225.293	372.380

Key observations:

- The fill rate increases from Croston to Naïve in this benchmark, but at the cost of higher holding cost which means that higher service can be attained with higher inventory carried under such a policy structure [32], [33].
- The probabilistic (q90) input offers a better average fill rate than MA(8) but incurs additional penalty cost in this configuration, demonstrating that uncertainty management must be properly tuned and fitted to the policy mapping (quantile choice, lead-time aggregation, constraints) as opposed to working over-the-counter [35].
- Intermittent-demand baselines are found to be cost-effectively competitive in mixed-SKU portfolios, which attests to an old perspective that SKU segmentation and method-policy pairing are relevant to inventory performance [34].

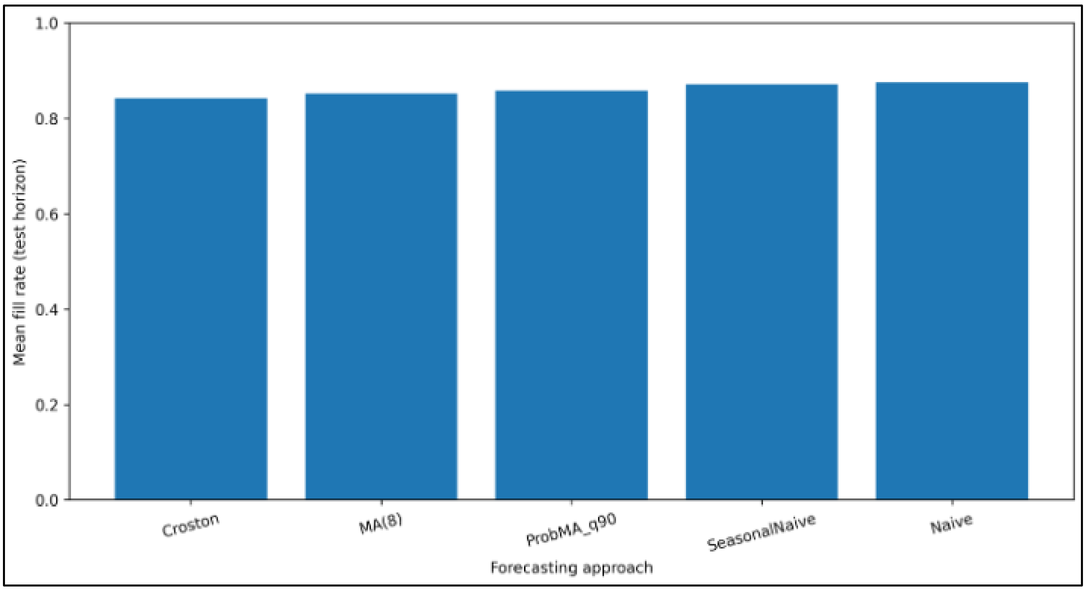

Figure 5 Inventory fill rate by forecasting approach

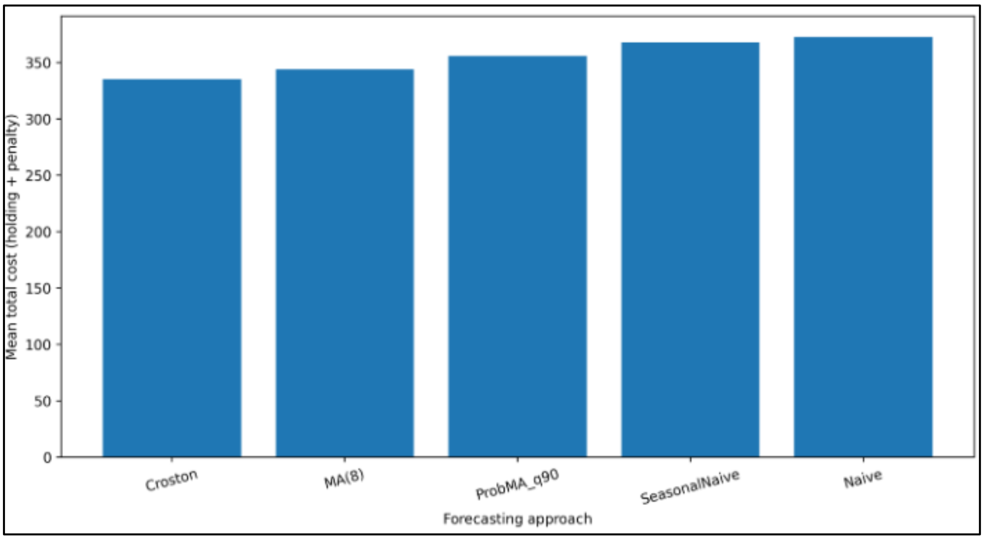

Figure 6 Inventory cost by forecasting approach

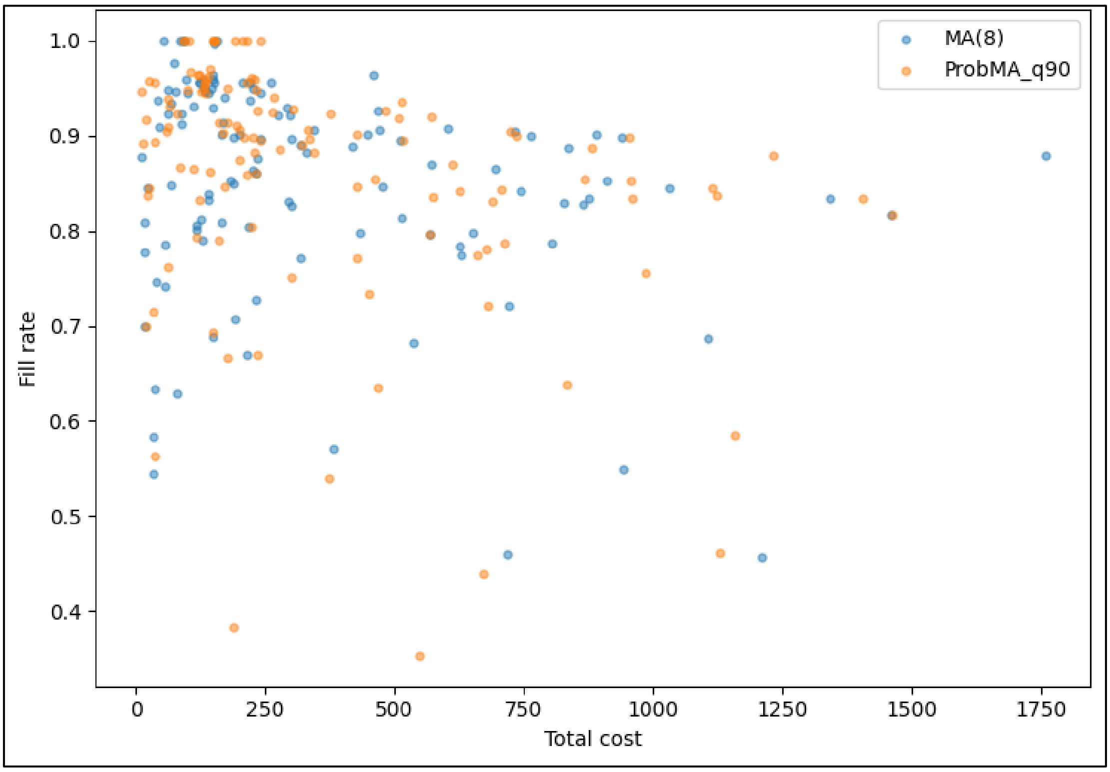

Figure 7 SKU-level trade-off: point vs probabilistic

The data collected in the experiment are displayed in three supplementary figures which in combination assess forecasting methods in an inventory-based viewpoint. The first bar chart shows the average fill rate by all SKUs under each forecast system, which shows the capacity of each system to meet demand given a constant policy of replenishment. The second bar chart reports the respective mean total inventory cost, which is broken down by holding and shortage penalty elements, thus, reflecting the economic trade-off of service-level improvement. This combination shows that an increase in forecast accuracy does not always lead to a reduction in total cost; in a number of instances, an increase in service level is attained at the cost of an increase in safety stock and holding cost. The third figure gives a SKU-level scatter plot of total cost versus fill rate, which directly compares a point-forecast-based policy input with a probabilistic (quantile-based) policy input. This representation highlights the results for heterogeneous SKUs and the service-cost frontier generated by various forecast-to-policy mappings. It is worth noting that with probabilistic forecasting, the extreme stockout risk of some SKUs is minimized at the expense of higher inventory exposure in cases of misalignment between the quantile calibration and the variability of demand and uncertainty of lead times.

All these findings support one key idea in inventory theory: the evaluation of methods of forecasting an entity should be conducted in terms of its operational implications within the replenishment systems, instead of focusing solely on the statistical measures of error, like MAE or RMSE [32], [33]. Decision-centric evaluation models that create forecasts in realistic inventory control policies give a more realistic evaluation of the performance impact on the supply chain.

5. Future directions

5.1. Decision-Centric Forecast Evaluation

Further research should focus on assessing prediction models within realistic inventory management simulations, but not only on statistical accuracy measures. It has been common knowledge that there exists a disconnection between forecast error improvement and inventory performance [39]. Simulations that combine probabilistic predictions with order-up-to or base-stock models in more than one regime of lead-time would give better operational indications.

5.2. Probabilistic Forecasting at Scale

Predictive distributions are needed in the multi-SKU environments to ensure effective safety stock and service-level control beyond point estimates. Recent developments in deep probabilistic time-series models show better scalability and cross-series learning capability [37]. More studies are required to equate distributional calibration, tail behavior and inventory robustness to different uncertainty modelling plans [40].

5.3. Multi-Echelon and Hierarchical Integration

Supply chains in the modern world are conducted within product hierarchy and geographical networks. The top-down reconciliation methods also guarantee uniformity in the aggregation levels and enhance alignment in downstream replenishment [38]. The extension of such practices to dynamic and multi-echelon inventory optimization is a field that is still open and needs computational and theoretical work.

5.4. Intelligent Forecasting to Reduce the Bullwhip Effect

Distortion of information and amplification of variation of order is an ongoing problem in the multi-tier supply chains [41]. The variability propagation can be addressed by the AI-based forecasting pipelines that incorporate the usual information streams, demand sensing, and self-directed policy updates. Simulation-based studies evaluating bullwhip intensity under machine-learning-based replenishment should be conducted.

5.5. Explainability and Governance of AI Forecasting Systems

Interpretability, monitoring, and governance are required for operational adoption. Explainable AI technology can enhance confidence and integration with enterprise resource planning systems [39]. The study needs to focus on the impact of model transparency on the stability of replenishment and managerial intervention behavior.

5.6. Sustainability and Inventory Carbon Footprint

Carbon emissions and inefficiency of resources are caused by excess inventory and emergency replenishment. Other studies in the future can measure the effects of probabilistic and adaptive forecasting on sustainability metrics, which will align forecasting research with the overall green supply chain goals.

6. Conclusion

The growing magnitude and sophistication of multi-SKU supply-chain demand forecasting systems that are not only precise but also operationally relevant to inventory decision-making. Python ecosystems offer software-defined frameworks that offer modular, extensible platforms upon which statistical models, gradient-boosting and deep neural networks, and probabilistic forecasting can be implemented at industrial scale. Nevertheless, in the literature, it is always shown that the measured improvements in the accuracy of forecasts through the MAE or the RMSE do not always translate into proportional reductions in total inventory cost or stockout risk.

Distributional and quantile-based forecasting, which is a form of probabilistic forecasting, is more consistent with service-level-based inventory policies, especially when the demand is intermittent and skewed. Deep learning architecture promises the ability to use cross-SKU patterns and shared representations with large collections of related time series to enhance scalability and robustness to heterogeneous portfolios. Moreover, hierarchical and grouped forecasting techniques can increase the coherence in product (as well as location and channel) dimensions, which directly impact the replenishment coordination in multi-echelon systems.

Regardless of these developments, considerable gaps exist in the connection between forecasting model governance, explainability, and computational scalability and quantifiable inventory performance improvements. Further studies need to employ progressively decision-conscious assessment protocols that can mimic actual replenishment policies, include lead-time ambiguity, and evaluate system stability dynamically, as well as variability enhancement at levels of the supply chain. Overall, Python-based demand forecasting is a potent enabler of intelligent inventory management, but it can only be worth its weight when integrated with uncertainty-aware and performance-based supply chain frameworks.

Compliance with ethical standards

Disclosure of conflict of interest

The author declares no conflict of interest.

References

[1] Silver, E. A., Pyke, D. F., & Thomas, D. J. (2017). *Inventory and production management in supply chains* (4th ed.). CRC Press.

[2] Chopra, S., & Meindl, P. (2019). *Supply chain management: Strategy, planning, and operation* (7th ed.). Pearson.

[3] Ivanov, D., & Dolgui, A. (2020). A digital supply chain twin for managing the disruption risks and resilience in the era of Industry 4.0. *Production Planning & Control, 31*(2–3), 136–152.

[4] Wang, G., Gunasekaran, A., Ngai, E. W. T., & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics, 176*, 98–110.

[5] McKinney, W. (2010). Data structures for statistical computing in Python. *Proceedings of the 9th Python in Science Conference*, 51–56.

[6] Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow* (2nd ed.). O'Reilly Media.

[7] Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly, 36*(4), 1165–1188.

[8] Baryannis, G., Dani, S., & Antoniou, G. (2019). Predicting supply chain risks using machine learning: The trade-off between performance and interpretability. *Future Generation Computer Systems, 101*, 993–1004.

[9] Dubey, R., Gunasekaran, A., Childe, S. J., et al. (2019). Big data analytics and artificial intelligence pathway to operational performance under the effects of entrepreneurial orientation and environmental dynamism. *International Journal of Production Economics, 226*, 107599.

[10] Zipkin, P. H. (2000). *Foundations of inventory management*. McGraw-Hill.

[11] Syntetos, A. A., Boylan, J. E., & Croston, J. D. (2005). On the categorization of demand patterns. *Journal of the Operational Research Society, 56*(5), 495–503.

[12] Babai, M. Z., Syntetos, A. A., & Teunter, R. H. (2014). Intermittent demand forecasting: An empirical study on accuracy and the risk of obsolescence. *International Journal of Production Economics, 157*, 212–219.

[13] Fildes, R., Ma, S., & Kolassa, S. (2019). Retail forecasting: Research and practice. *International Journal of Forecasting, 35*(1), 1–9.

[14] Kolassa, S. (2016). Evaluating predictive count data distributions in retail sales forecasting. *International Journal of Forecasting, 32*(3), 788–803.

[15] Gardner, E. S., Jr. (1990). Evaluating forecast performance in an inventory control system. *Management Science, 36*(4), 490–499.

[16] Kingsman, B. G., & Sani, B. (1997). Selecting the best periodic inventory control and demand forecasting methods for low demand items. *Journal of the Operational Research Society, 48*(7), 700–713.

[17] Lee, H. L., So, K. C., & Tang, C. S. (2000). The value of information sharing in a two-level supply chain. *Management Science, 46*(5), 626–643.

[18] Syntetos, A. A., Boylan, J. E., & Croston, J. D. (2005). On the categorization of demand patterns. *Journal of the Operational Research Society, 56*(5), 495–503.

[19] Fildes, R., Goodwin, P., Lawrence, M., & Nikolopoulos, K. (2009). Effective forecasting and judgmental adjustments: An empirical evaluation and strategies for improvement in supply-chain planning. *International Journal of Forecasting, 25*(1), 3–23.

[20] Hyndman, R. J., Ahmed, R. A., Athanasopoulos, G., & Shang, H. L. (2011). Optimal combination forecasts for hierarchical time series. *Computational Statistics & Data Analysis, 55*(9), 2579–2589.

[21] Babai, M. Z., Ali, M. M., Boylan, J. E., & Syntetos, A. A. (2013). Forecasting and inventory performance in a two-stage supply chain with ARIMA(0,1,1) demand: Theory and empirical analysis. *International Journal of Production Economics, 143*(2), 463–471.

[22] Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). The M4 competition: Results, findings, conclusion and way forward. *International Journal of Forecasting, 34*(4), 802–808.

[23] Petropoulos, F., Spiliotis, E., & Assimakopoulos, V. (2019). The inventory performance of forecasting methods: Evidence from the M3 competition data. *International Journal of Forecasting, 35*(1), 251–265.

[24] Salinas, D., Flunkert, V., Gasthaus, J., & Januschowski, T. (2020). DeepAR: Probabilistic forecasting with autoregressive recurrent networks. *International Journal of Forecasting, 36*(3), 1181–1191.

[25] Goltsos, T. E., Syntetos, A. A., Glock, C. H., & Ioannou, G. (2022). Inventory–forecasting: Mind the gap. *European Journal of Operational Research, 299*(2), 397–419.

[26] Fildes, R., & Beard, C. (1992). Forecasting systems for production and inventory control. *International Journal of Operations & Production Management, 12*(5), 4–27.

[27] Lee, H. L., Padmanabhan, V., & Whang, S. (1997). Information distortion in a supply chain: The bullwhip effect. *Management Science, 43*(4), 546–558.

[28] Feizabadi, J. (2022). Machine learning demand forecasting and supply chain performance. *International Journal of Logistics Research and Applications, 25*, 119–142.

[29] Trapero, J. R., Cardós, M., & Kourentzes, N. (2019). Quantile forecast optimal combination to enhance safety stock estimation. *International Journal of Forecasting, 35*(1), 239–250.

[30] Snyder, R. D., Ord, J. K., & Beaumont, A. (2012). Forecasting the intermittent demand for slow-moving inventories: A modelling approach. *International Journal of Forecasting, 28*(2), 485–496.

[31] Liang, W. Y., & Huang, C. C. (2006). Agent-based demand forecast in multi-echelon supply chain. *Decision Support Systems, 42*(1), 390–407.

[32] Axsäter, S. (2015). Inventory control (3rd ed.). New York, NY: Springer.

[33] Hadley, G., & Whitin, T. M. (1963). Analysis of inventory systems. Englewood Cliffs, NJ: Prentice-Hall.

[34] Croston, J. D. (1972). Forecasting and stock control for intermittent demands. *Operational Research Quarterly, 23*(3), 289–303.

[35] Koenker, R., & Bassett, G., Jr. (1978). Regression quantiles. Econometrica, 46(1), 33–50.

[36] Efron, B., & Tibshirani, R. J. (1993). An introduction to the bootstrap. New York, NY: Chapman & Hall.

[37] Rangapuram, S. S., Seeger, M. W., Gasthaus, J., Stella, L., Wang, Y., & Januschowski, T. (2018). Deep state space models for time series forecasting. *Advances in Neural Information Processing Systems, 31*, 7785–7794.

[38] Wickramasuriya, S. L., Athanasopoulos, G., & Hyndman, R. J. (2019). Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. *Journal of the American Statistical Association, 114*(526), 804–819.

[39] Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020). The M5 accuracy competition: Results, findings, and conclusions. *International Journal of Forecasting, 36*(1), 54–74.

[40] Taylor, J. W. (2000). A quantile regression approach to estimating the distribution of multi-period returns. *Journal of Derivatives, 7*(2), 64–78.

[41] Chen, F., Drezner, Z., Ryan, J. K., & Simchi-Levi, D. (2000). Quantifying the bullwhip effect in a simple supply chain: The impact of forecasting, lead times, and information. *Management Science, 46*(3), 436–443.