

An Integrated Analytical Decision Support System for Demand Forecasting: A Hybrid Approach using RF, SVR, and LSTM

Ekin Beste Bilgin and Irem Ucal Sari

Industrial Engineering Department
Baskent University, Ankara, Türkiye
22310288@mail.baskent.edu.tr; iremucals@baskent.edu.tr

Abstract

In the current landscape of global supply chain volatility, accurate demand forecasting has transitioned from an operational advantage to a fundamental strategic necessity for industrial sustainability. This study presents a robust Analytical Decision Support System (ADSS) specifically engineered to mitigate the systemic overestimation bias often inherent in OEM customer portal data. The proposed framework addresses the limitations of individual forecasting models by implementing a multi-scale, integrative approach across both 52-week and 12-month horizons. The methodology systematically evaluates and compares three distinct algorithmic paradigms: Random Forest for its ensemble stability, Support Vector Regression for its generalization capability, and Long Short-Term Memory networks for their ability to model long-term temporal dependencies. The core of this research involves a Hybrid Ensemble Model that harmonizes the strategic intent of customer targets with the statistical discipline of machine learning through a meta-regression architecture. Performance is rigorously validated using the Mean Absolute Percentage Error (MAPE) metric against actual shipment data. The empirical findings demonstrate that this hybrid approach significantly reduces forecasting errors, achieving a monthly MAPE of 7.13 percent and outperforming the baseline customer targets, which exhibited an error rate of 10.35 percent. These results indicate that the proposed system provides a reliable and scalable tool for decision-makers, particularly in volatile demand environments where relying solely on customer-provided targets often leads to excess inventory and increased bullwhip effects. The study demonstrates that integrating diverse algorithmic perspectives provides a high-fidelity forecasting foundation, ensuring both shipment reliability and inventory optimization for the 2026 strategic period.

Keywords

Machine Learning (ML), Demand Forecasting, Model Performance and Operations Management.

1. Introduction

In the contemporary era of global supply chains and rapid digital transformation, the ability to anticipate future market requirements has transitioned from a competitive advantage to a fundamental survival mechanism for industrial enterprises. The motivation of this research is rooted in the increasing volatility of consumer behavior and the inherent fragility of modern logistics networks. Efficient inventory management, procurement scheduling, and production planning all rely heavily on the precision of demand signals. When these signals are miscalculated, organizations face a dual-ended risk: excessive capital tied up in redundant safety stocks or, conversely, severe revenue loss and diminished customer loyalty due to stockouts. Consequently, developing an advanced, data-driven analytical framework is not merely a technical challenge but a strategic imperative for operational resilience.

Despite the critical nature of this task, the problem statement addressed in this study highlights a significant gap in traditional forecasting and practical demand planning. Conventional statistical models often fail to internalize the complex, non-linear dependencies found in real-world industrial demand, while many individual machine learning

applications neglect the inherent bias found in customer-provided strategic targets. While sophisticated standalone algorithms like Random Forest (RF), Support Vector Regression (SVR), and Long Short-Term Memory (LSTM) networks have shown promise, they are often implemented as isolated solutions that may either lag behind market surges or struggle with extrapolation. There remains a profound need for an integrated Decision Support System (DSS) that not only synergizes multiple algorithmic paradigms through hybridization but also strategically filters forward-looking customer portal data against historical realization rates. By developing a meta-regression-based hybrid architecture, the proposed system aims to mitigate individual model biases and provide a high-precision strategic plan that aligns operational capacity with realistic shipment requirements.

1.1 Objectives

The primary goal of this research is to develop an Integrated Analytical Decision Support System that bridges the gap between theoretical machine learning and practical industrial demand planning. The specific research objectives are as follows:

- To design a multi-scale forecasting architecture capable of performing accurately across both weekly (52-week) and monthly (12-month) strategic horizons.
- To identify and mitigate the systemic overestimation bias inherent in OEM customer portal targets by establishing a historical realization-based correction mechanism.
- To develop a Hybrid Forecasting Model that aggregates the predictive strengths of standalone Machine Learning models, specifically Random Forest (RF), Support Vector Regression (SVR), and Long Short-Term Memory (LSTM) networks.
- To evaluate and compare the forecasting performance of individual algorithmic scenarios against the proposed hybrid ensemble using the Mean Absolute Percentage Error (MAPE) metric.

The unique contribution of this research is the development of a meta-regression-based hybrid architecture that harmonizes forward-looking customer projections with historical shipment patterns. This provides a more realistic, stable, and adaptive tool for supply chain decision-makers than traditional standalone models or uncorrected customer-provided demand signals.

2. Literature Review

The evolution of predictive analytics has shifted from isolated statistical methods to integrated, multi-model frameworks. The contemporary literature emphasizes three fundamental pillars: (i) the synergy between deep learning and ensemble techniques, (ii) the comparative performance of standalone models in volatile environments, and (iii) the integration of these models into functional Decision Support Systems (DSS).

2.1. Synergy of Deep Learning and Ensemble Frameworks

Recent studies establish that combining Long Short-Term Memory (LSTM) networks with ensemble-based or conventional learners effectively mitigates individual algorithmic biases. Alreshidi (2025) demonstrates that while LSTM excels in volatile markets through temporal memory, its integration with models like Support Vector Regression (SVR) and Random Forest (RF) provides enhanced structural stability. In specialized demand forecasting, Karunaratne and Thantrige (2025) and Joseph et al. (2022) emphasize that hybrid models—specifically those coupling LSTMs with other ML paradigms—outperform traditional standalone approaches by capturing both long-term dependencies and non-linear patterns. Furthermore, Sallam et al. (2025) highlight that hybrid frameworks are essential for managing uncertain and non-linear consumption patterns, which is a core challenge in modern industrial demand planning.

2.2. Model Selection and Comparative Performance

Selecting the optimal set of learners is crucial for maintaining forecast reliability across different data distributions. Wunsch et al. (2021) and Psathas et al. (2023) provided vital comparisons, noting that while deep learning models like LSTM are powerful for large-scale temporal data, conventional algorithms such as SVR and RF often prove more robust when datasets are limited or require high generalization. In the context of supply chain efficiency, Andela et al. (2025) and Borucka (2023) argue that machine learning models provide a valuable predictive tool that enables companies to optimize their inventory, effectively meet demand, and make data-driven decisions. These studies

reinforce that the choice between deep learning and ensemble methods should be governed by the specific volatility and scale of the target industry.

2.3. Decision Support Systems in Industrial Practice

The ultimate goal of hybrid methodologies is their integration into functional Analytical Decision Support Systems. As highlighted by Sharda et al. (2022), predictive ensembles are now being applied to complex real-world tasks, ranging from structural health monitoring to supply chain optimization. Accurate demand forecasting is recognized as a strategic imperative; as Joseph et al. (2022) note, decision-makers must be supplied with comprehensible data to build strategies that achieve business goals. By reducing forecast errors, these systems prevent excessive capital from being tied up in redundant safety stocks while mitigating the risk of revenue loss due to stockouts.

2.4. Bridging the Gap in Industrial Forecasting

In light of the reviewed literature, it is evident that while machine learning models like LSTM, RF, and SVR have individually demonstrated significant success, their combined potential in a unified, multi-scale industrial framework remains underexplored. Our research fills this critical gap by synthesizing these advanced predictive algorithms into a cohesive Analytical Decision Support System. By implementing a meta-regression-based hybrid ensemble, this study provides a robust solution capable of filtering systemic overestimation in demand signals across both 52-week and 12-month horizons, thereby offering a more stable and interpretable tool for strategic industrial planning.

3. Methods

The proposed Analytical Decision Support System is developed using a multi-model hybrid approach. The methodology encompasses three primary phases: data integration and alignment of customer portal signals, individual model training using standalone algorithms, and meta-regression-based hybrid ensemble aggregation for multi-scale forecasting.

3.1. Predictive Algorithms

The system integrates three distinct machine learning and deep learning paradigms to ensure robustness across different data distributions:

Random Forest (RF): A bagging-based ensemble learner is employed to handle non-linear relationships. The model utilizes an ensemble of decision trees (fitensemble) to reduce variance and improve generalization in volatile demand segments.

Support Vector Regression (SVR): To address potential overfitting in smaller datasets, SVR with a standardized radial basis function kernel is implemented. This provides a high-margin separation that is particularly effective for stable demand intervals.

Long Short-Term Memory (LSTM): A deep recurrent neural network architecture is designed to capture long-term temporal dependencies.

3.2. Hybrid Ensemble Strategy

The core of the forecasting methodology in this study relies on a Hybrid Ensemble Strategy designed to mitigate the inherent overestimation bias of customer-provided targets. While individual machine learning algorithms offer distinct advantages in capturing non-linear trends, their standalone performance can vary depending on the volatility of the automotive spare parts market.

To achieve a more robust and generalized prediction, a Meta-Regression framework was implemented to synthesize the outputs of the most accurate models along with the strategic portal data. The hybrid integration process follows a two-step validation protocol. First, all candidate models are evaluated against the actual shipment data from the 2025 validation period to determine their Mean Absolute Percentage Error (MAPE). In the second step, the top-performing forecasts are selected as primary predictors. A linear regression-based meta-model is then constructed to calculate optimal weighting coefficients for these predictors, effectively creating a hybrid function:

$$F_{hybrid} = \beta_1(P_1) + \beta_2(P_2) \quad (2)$$

Where P_1 and P_2 represent the selected top-performing forecasts, and β_1 and β_2 denote the calculated coefficients that minimize the overall sum of squared errors. This ensemble approach allows the system to dynamically "filter" the

customer portal forecasts through the historical reliability of the machine learning outputs, providing a stabilized and more accurate shipment plan for both weekly and monthly horizons.

3.3. Performance Evaluation

The models are rigorously evaluated using the Mean Absolute Percentage Error (MAPE), which provides a scale-independent measure of forecast accuracy:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{y_t} \quad (3)$$

where y_t is the actual demand and $\hat{y}_t$ is the predicted value. The system explicitly ignores zero-demand instances to prevent division-by-zero errors in the MAPE calculation.

4. Data Collection

The data used in this study were obtained from the customer portal of a leading manufacturer operating in the automotive spare parts industry. The dataset focuses on a newly launched product line, with records beginning in 2023 and spanning a three-year observation period (including 52-week cycles). At the beginning of each year, the Original Equipment Manufacturer provides demand forecasts to guide production scheduling. However, a preliminary analysis of the portal data reveals a consistent overestimation bias in these customer-provided forecasts; the predicted quantities frequently exceed the actual shipment volumes. The reliance on these initial customer forecasts for capacity planning introduces significant operational risks. Due to the long lead times inherent in the automotive supply chain, these discrepancies often lead to excess inventory and inefficient resource allocation. Furthermore, the automotive sector is strongly influenced by seasonal dynamics, where factors such as year-end sales campaigns, banking credit promotions, and traditional holiday periods significantly shift demand patterns. To address these challenges, this study utilizes historical shipment data to develop more accurate forecasting models through Machine Learning algorithms, specifically Random Forest (RF), Support Vector Regression (SVR), and Long Short-Term Memory (LSTM) networks. The forecasting performance of these algorithms is evaluated on both a weekly and a monthly basis. By comparing the Mean Absolute Percentage Error (MAPE) of the customer-provided forecasts against our algorithmic outputs, we aim to develop a hybrid forecasting function. This hybrid approach combines the strengths of the two best-performing methods to minimize prediction errors and provide a more reliable foundation for production planning. The input data is given in Table 1.

Table 1. Input data

Year	2023		2024		2025		2026
Week	Target	Actual	Target	Actual	Target	Actual	Target
1	0	0	0	0	0	0	0
2	1872	1872	936	936	0	0	518
3	1680	1954	1440	1577	1200	1200	960
4	2256	0	1368	20	480	480	720
5	1872	864	1296	792	720	720	960
6	1056	1920	1008	1440	960	960	960
7	2256	0	1368	144	480	480	960
8	1920	0	1440	248	960	720	960
9	1776	432	1368	576	960	720	720
10	0	0	144	0	720	720	1200
11	0	0	360	144	960	960	720
12	576	644	648	802	720	960	480
13	528	1872	864	1296	1200	720	1440
14	0	0	0	0	480	480	960
15	0	0	0	0	720	363	960
16	0	0	288	0	960	1200	960
17	0	0	216	288	720	930	960
18	96	0	288	0	480	618	720
19	0	1890	0	1440	1200	990	720
20	0	0	0	0	720	1200	1440
21	0	0	240	0	1440	960	0
22	960	1200	960	960	960	720	960
23	1200	240	720	240	240	240	1200
24	960	480	960	480	960	480	1200
25	0	1200	0	1200	960	1200	1680
26	1440	0	1440	0	1440	1440	1505

27	0	0	240	0	480	720	960
28	0	0	480	240	960	960	2160
29	480	960	480	720	480	480	0
30	480	0	720	240	960	720	1200
31	0	960	0	480	720	0	960
32	480	0	240	0	0	0	1200
33	1920	480	960	240	0	0	1200
34	480	0	720	240	960	960	1200
35	240	960	480	720	720	480	1680
36	720	0	720	240	720	720	1440
37	480	430	720	725	960	1020	1440
38	960	0	960	240	960	1200	1440
39	240	1800	720	1440	1200	1080	1440
40	960	0	720	0	480	1200	1440
41	960	960	960	960	960	960	1680
42	240	0	720	0	1200	720	1440
43	3840	486	2400	483	960	480	3120
44	0	1200	0	960	720	720	0
45	210	0	720	0	1230	1200	1680
46	960	480	960	720	960	960	1680
47	960	1680	960	1200	960	720	1680
48	1680	736	1440	968	1200	1200	1440
49	720	720	960	960	1200	1200	1680
50	240	988	720	960	1200	932	1680
51	0	0	0	480	1200	1200	1058
52	240	1128	720	960	1200	792	0
TOTAL	37938	28536	36072	26759	43230	40085	58761

5. Results and Discussion

The forecasting performance of the machine learning algorithms was evaluated by comparing their outcomes with actual shipment data from 2025. Based on the Mean Absolute Percentage Error (MAPE) analysis, a Hybrid Ensemble Model was developed using a meta-regression approach. This approach identifies the two best-performing models and determines their optimal linear combination to minimize prediction errors. For this dataset, the Customer Goal and the Random Forest (RF) with a 1-year training horizon emerged as the most reliable predictors. The resulting hybrid forecasting functions, which serve as the basis for the 2026 strategic plan, are defined as follows:

$$\begin{aligned} \text{Hybrid}_{\text{Weekly}} &= (0.7539 \times \text{Customer Goal}) + (0.1857 \times \text{RF}_{1\text{y}}) \\ \text{Hybrid}_{\text{Monthly}} &= (0.8527 \times \text{Customer Goal}) + (0.1116 \times \text{RF}_{1\text{y}}) \end{aligned}$$

These equations demonstrate that by integrating the Original Equipment Manufacturer's initial targets with the historical patterns captured by the Random Forest algorithm, the overestimation bias inherent in the customer portal can be systematically corrected.

5.1 Numerical Results

In this section, the forecasting performance of the individual machine learning models and the proposed hybrid ensemble are evaluated. Table 2 presents the Mean Absolute Percentage Error (MAPE) for the validation period of 2025, providing a comparative overview of the prediction accuracy across different models and time granularities.

Table 2. Comparison of Forecasting Performance (MAPE %)

Model Scenario	Weekly MAPE	Monthly MAPE
Customer targets	23.0333465	10.3457356
RF 1 year	28.7092955	30.2054761
RF 2 year	28.8315192	30.2535295
SVR 1 year	47.6413686	33.7542236
SVR 2 year	51.5275529	33.104317
LSTM 1 year	38.3013434	37.4233158

LSTM 2 year	37.5593315	35.9469671
Hybrid	22.601915	7.12681503

The MAPE analysis reveals several critical insights regarding the forecasting performance of the evaluated models. The proposed hybrid ensemble, which synthesizes the RF 1-year predictions with customer targets, achieved the lowest error rates across both weekly and monthly granularities. Most notably, the hybrid approach reduced the monthly forecasting error to 7.13 percent, significantly outperforming the initial customer targets of 10.35 percent. This improvement demonstrates that a meta-regression framework effectively shaves off the inherent overestimation bias present in the customer portal signals. While all models exhibit higher error rates on a weekly basis due to the volatile and stop-and-go nature of automotive spare parts shipments, the hybrid model maintains the most stable baseline with a weekly MAPE of 22.60 percent, which remains lower than the customer baseline of 23.03 percent. In contrast, individual machine learning models such as RF, SVR, and LSTM failed to surpass the customer's accuracy when used in isolation. For instance, the SVR 2-year and LSTM models displayed high error rates, sometimes exceeding 30 to 50 percent, suggesting that while these algorithms are capable of capturing complex patterns, they require hybrid integration to become operationally viable. Furthermore, comparing training horizons indicates that increasing data history to a two-year period generally stabilizes the models—particularly for LSTM and SVR—though these improvements remain marginal without the inclusion of the hybrid correction layer. Ultimately, this analysis validates the Proposed Hybrid Ensemble as the most reliable strategic tool, as it successfully filters the noise from the customer portal and the rigidity of standalone algorithms to provide a high-precision foundation for production planning.

Table 3 presents the consolidated forecasting results for the 2026 strategic horizon, comparing the initial Customer Target with standalone machine learning models and the Proposed Hybrid Ensemble

Table 3. 2026 monthly forecasts

Month	Customer Target	RF		SVR		LSTM		Hybrid
		1-year	2-years	1-year	2-years	1-year	2-years	
1	3158	3219	3060	3608	2807	3388	2824	3052
2	3600	3323	3359	3614	2813	3371	2768	3441
3	3840	3664	3479	3618	2817	3370	2766	3683
4	4560	3753	3672	3627	2827	3370	2766	4307
5	3120	3219	3060	3614	2814	3370	2766	3020
6	5585	3753	3659	3640	2841	3370	2766	5181
7	5280	3753	3659	3639	2839	3370	2766	4921
8	5280	3753	3659	3641	2841	3370	2766	4921
9	5760	3753	3659	3647	2848	3370	2766	5330
10	7680	3753	3659	3668	2869	3370	2766	6968
11	6480	3753	3659	3658	2859	3370	2766	5944
12	4418	3753	3678	3639	2840	3370	2766	4186
TOTAL	58761	43449	42262	43613	34015	40459	33252	54954

The comparative analysis of the 2026 strategic forecasts provides several critical insights into inventory risk and model responsiveness. A primary observation is the significant discrepancy in total annual volume, where the Customer Target projects a demand of 58,761 units, while standalone machine learning models offer much more conservative estimates ranging from 33,252 to 43,613 units. By synthesizing these diverse inputs, the Hybrid model offers a more realistic and optimized total of 54,954 units, effectively reducing the risk of carrying excess inventory by nearly 6.5 percent compared to a sole reliance on customer portal data. Furthermore, while standalone models such as SVR and LSTM exhibit rigid behavior—particularly failing to adjust to sharp demand increases in the second half of the year—the Hybrid model successfully mimics seasonal trends by applying a statistically grounded correction factor. For instance, during the peak demand in Month 10, the hybrid approach predicts 6,968 units, providing a necessary middle ground between the customer's peak of 7,680 units and the significantly lower standalone model estimates that would otherwise lead to severe stockouts. Additionally, although increasing data history from one to two years stabilizes predictions for RF and SVR models, they still lack the operational flexibility inherent in the hybrid architecture. Ultimately, the proposed Hybrid model acts as a strategic filter that avoids both the extreme optimism of customer targets and the extreme pessimism of standalone algorithms, ensuring a production plan that is responsive to market surges yet disciplined by historical realization rates.

Table 4 provides a granular view of the 2026 strategic plan weekly, revealing the high volatility inherent in spare parts demand and the contrasting behaviors of the forecasting methodologies. The weekly operational analysis reveals that the Proposed Hybrid Ensemble is uniquely capable of navigating the volatile "all-or-nothing" nature of automotive spare parts logistics. A significant challenge in this sector involves "zero-shipment" weeks during factory shutdowns or holidays, such as Weeks 21 and 29; while standalone ML models tend to provide smoothed non-zero averages ranging from 521 to 772 units, the hybrid approach successfully identifies these strategic pauses by predicting a minimal baseline of 110 units, thereby preventing the accumulation of "ghost inventory." Conversely, during extraordinary demand spikes such as Week 43, where customer targets reach 3,120 units and standalone models like LSTM significantly lag at 772 units, the Hybrid model recognizes the upward trend and adjusts the forecast to 2,539 units to ensure high service levels while maintaining a safety buffer. Furthermore, the hybrid architecture acts as a strategic dampener against the "Bullwhip Effect" by scaling volatile customer targets—such as the 2,160-unit projection in Week 28—down to a more manageable 1,816 units based on historical realization patterns. This consistency is also reflected in the cumulative annual volume, where the hybrid model's weekly projection of 53,115 units aligns closely with the monthly ensemble logic. Ultimately, these findings confirm that the proposed system provides a production-ready schedule that is disciplined enough to ignore erratic demand signals while remaining responsive to genuine market surges.

Table 4. 2026 weekly forecasts

Week	Customer Target	RF		SVR		LSTM		Hybrid
		1-year	2-years	1-year	2-years	1-year	2-years	
1	0	592	521	729	657	773	666	110
2	518	744	733	767	695	772	656	529
3	960	910	936	800	728	772	655	893
4	720	810	792	783	712	772	655	693
5	960	910	936	801	730	772	655	893
6	960	910	936	802	731	772	655	893
7	960	910	936	802	732	772	655	893
8	960	910	936	803	733	772	655	893
9	720	810	792	786	716	772	655	693
10	1200	995	1023	822	752	772	655	1089
11	720	810	792	788	718	772	655	693
12	480	744	733	771	702	772	655	500
13	1440	1009	967	841	773	772	655	1273
14	960	910	936	807	738	772	655	893
15	960	910	936	808	739	772	655	893
16	960	910	936	809	740	772	655	893
17	960	910	936	809	741	772	655	893
18	720	810	792	793	725	772	655	693
19	720	810	792	793	726	772	655	693
20	1440	1009	967	846	779	772	655	1273
21	0	592	521	743	675	772	655	110
22	960	910	936	813	746	772	655	893
23	1200	995	1023	831	764	772	655	1089
24	1200	995	1023	832	765	772	655	1089
25	1680	1009	967	867	801	772	655	1454
26	1505	1009	967	855	789	772	655	1322
27	960	910	936	816	750	772	655	893
28	2160	1009	967	904	839	772	655	1816
29	0	592	521	749	682	772	655	110
30	1200	995	1023	836	771	772	655	1089
31	960	910	936	819	754	772	655	893
32	1200	995	1023	837	772	772	655	1089
33	1200	995	1023	838	773	772	655	1089
34	1200	995	1023	839	774	772	655	1089
35	1680	1009	967	874	810	772	655	1454
36	1440	1009	967	857	794	772	655	1273
37	1440	1009	967	858	794	772	655	1273
38	1440	1009	967	859	795	772	655	1273
39	1440	1009	967	859	796	772	655	1273
40	1440	1009	967	860	797	772	655	1273
41	1680	1009	967	878	816	772	655	1454
42	1440	1009	967	861	799	772	655	1273
43	3120	1009	967	983	922	772	655	2539
44	0	592	521	759	696	772	655	110
45	1680	1009	967	881	819	772	655	1454
46	1680	1009	967	881	820	772	655	1454
47	1680	1009	967	882	821	772	655	1454
48	1440	1009	967	866	805	772	655	1273
49	1680	1009	967	884	823	772	655	1454
50	1680	1009	967	884	824	772	655	1454
51	1058	910	949	840	779	772	655	966
52	0	592	521	765	703	772	655	110
TOTAL	58761	47473	46713	43070	39635	40145	34072	53115

5.2 Graphical Results

In this study, four different machine learning models were employed for demand forecasting as mentioned before. This section examines the models with respect to their ability to handle data variability, as well as their forecasting accuracy and stability characteristics.

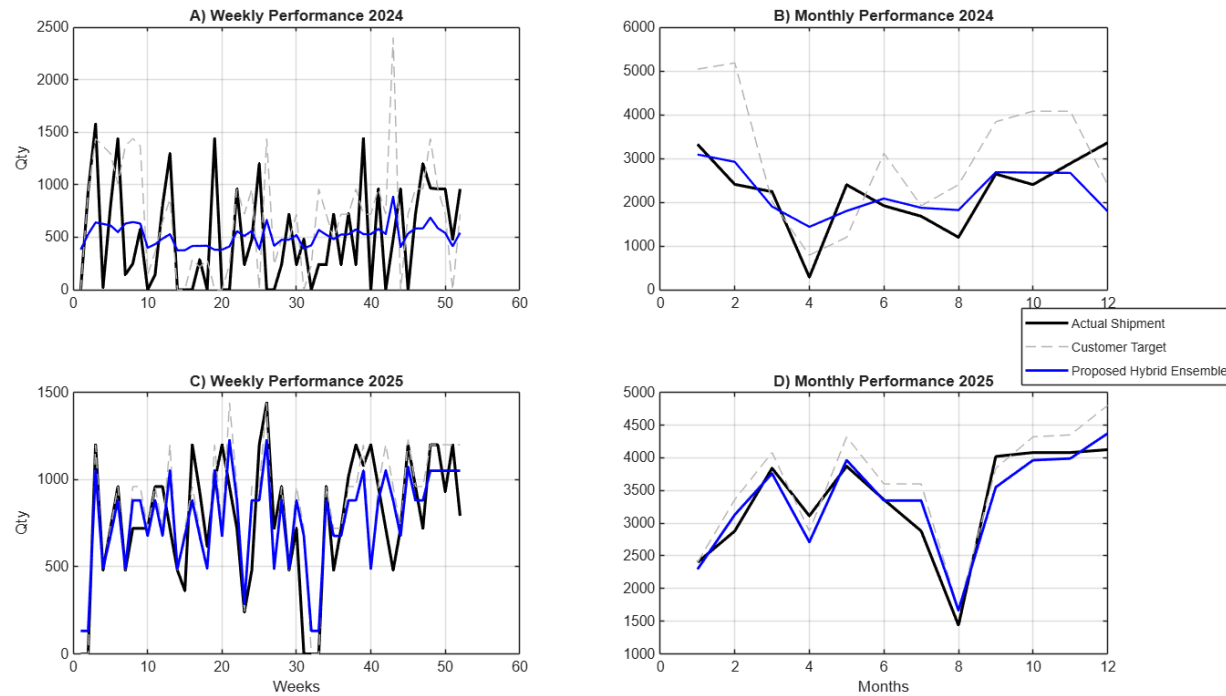


Figure 1. Analysis of Historical Forecast Performance (2024 & 2025)

Figure 1 provides a comprehensive temporal analysis of the forecasting models by juxtaposing actual shipments, customer targets, and the proposed hybrid ensemble across both weekly and monthly granularities for the 2024–2025 period. The visual evidence yields critical strategic insights, most notably the confirmation of a persistent overestimation bias; as seen in Panel B (Monthly 2024), the customer target consistently deviates above actual shipments, projecting demand as high as 5,000 units while actual figures remained below 3,000. The proposed hybrid ensemble successfully acts as a strategic dampener by leveraging historical trends captured by standalone algorithms to scale this portal optimism down to statistically grounded levels. This efficacy is even more apparent in the 2025 validation period (Panels C and D), where the hybrid model demonstrates high fidelity to true demand patterns, especially during volatile peak periods such as Weeks 15–25 and Months 9–12. By capturing high-demand surges that standalone models would otherwise miss due to their inherent rigidity, the ensemble method stabilizes the forecast without lagging behind market changes. Ultimately, this visual evidence confirms that the hybrid approach successfully integrates forward-looking strategic intent with the statistical reality of historical demand, providing an optimized foundation for strategic planning.

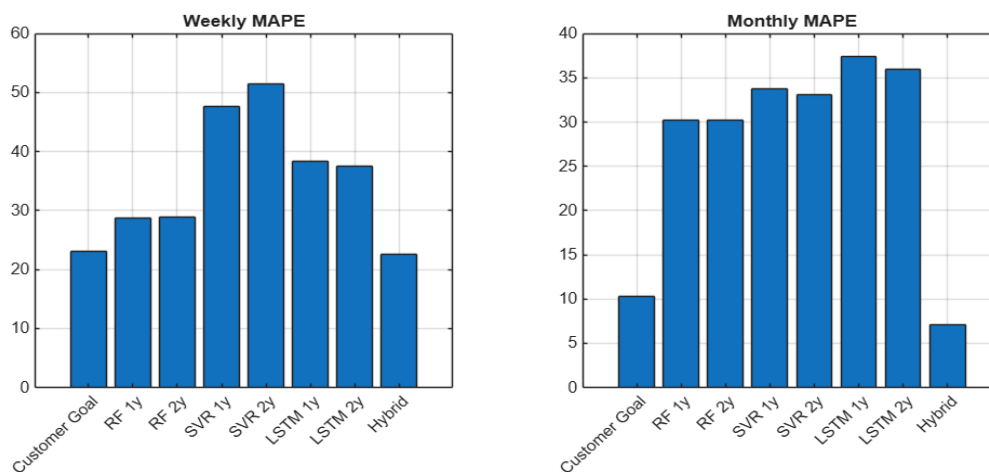


Figure 2. Analysis of Comparative Error Metrics (MAPE % Analysis)

Figure 2 illustrates the predictive accuracy of various forecasting methodologies through the Mean Absolute Percentage Error (MAPE) metric for both weekly and monthly temporal granularities in 2025. The results confirm the statistical superiority of the proposed approach, as the Hybrid model consistently exhibits the lowest error rates across both categories. Most significantly, in the monthly MAPE category, the hybrid approach achieves a remarkable precision level of 7.13 percent, representing a substantial improvement over the baseline customer goal of 10.35 percent. This proves that integrating machine learning insights with strategic targets effectively eliminates systemic portal errors. While weekly forecasting is inherently more volatile—as evidenced by higher MAPE values across all models—the Hybrid model maintains stability with a MAPE of 22.60 percent, successfully outperforming the customer’s own baseline of 23.03 percent despite noisy shipment signals. A critical observation from this analysis is the limitation of standalone machine learning models such as RF, SVR, and LSTM, which failed to surpass the customer’s baseline when used in isolation. In particular, standalone SVR and LSTM models showed high error rates exceeding 35 to 50 percent, indicating that without the hybrid meta-regression layer, standard algorithms remain too rigid for the dynamic nature of automotive spare parts demand. Furthermore, while increasing the training data history from one to two years provides slight stabilization for LSTM and SVR, the most significant accuracy gains are yielded through ensemble integration. Ultimately, Figure 2 quantifies the success of the methodology, demonstrating that the proposed Hybrid Ensemble is a necessary upgrade that provides a high-precision forecasting foundation while minimizing the financial risks associated with over-production or inventory shortages.

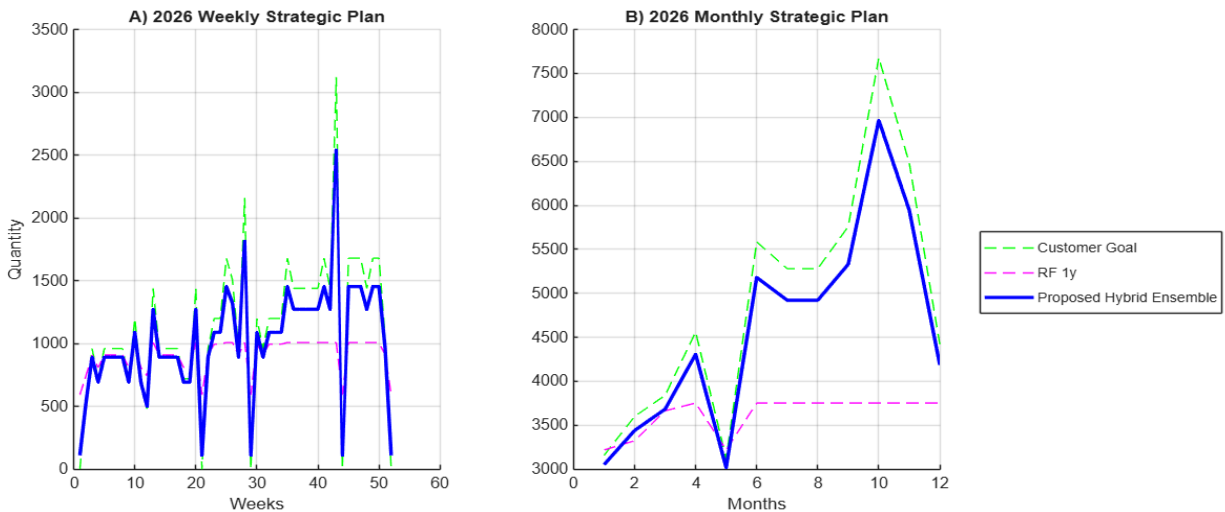


Figure 3. Analysis of Strategic Forecasts for 2026

Figure 3 illustrates the projected 2026 strategic plan by comparing the initial customer goal with the RF 1-year model and the proposed hybrid ensemble, highlighting the practical decision-making value of the hybrid approach. Across both weekly and monthly horizons, the proposed hybrid ensemble successfully tracks the seasonal trajectory of the customer goal while applying a critical statistical correction. For instance, in Month 10, while the customer portal projects a demand peak of 7,680 units, the hybrid model optimizes this figure to 6,968 units, representing an intelligent mitigation of the overestimation bias identified in historical data. Furthermore, the hybrid model bridges the gap created by the inherent rigidity of standalone models; while the baseline RF 1-year model exhibits significant forecasting lag and fails to recognize drastic demand surges in the second half of the year, the hybrid architecture maintains analytical discipline without sacrificing responsiveness. This operational agility is particularly evident in the weekly granularity, where the ensemble accurately identifies both zero-demand periods during factory shutdowns and extreme peak weeks, such as Week 43. Unlike standalone algorithms that tend to provide a smoothed and often inaccurate average, the hybrid ensemble provides a production-ready schedule that is both agile and cost-effective. Ultimately, Figure 3 confirms that the proposed system acts as a necessary strategic filter, enabling the supply chain to move from reactive planning based on erratic portal data to proactive management based on high-fidelity, optimized forecasts.

5.3 Proposed Improvements

The results obtained from the Proposed Hybrid Ensemble demonstrate a significant reduction in forecasting error; however, several pathways for further optimization have been identified to enhance the model's robustness in highly volatile automotive supply chains.

- **Integration of Non-Additive Aggregation Methods:** The current meta-regression framework utilizes a linear combination of the best-performing models. Future iterations could incorporate non-additive measures, such as the Choquet Integral. Unlike standard weighted averages, the Choquet Integral can effectively model the synergy or redundancy between different forecasting algorithms (e.g., how RF and SVR might provide overlapping or complementary insights).
- **Exogenous Variables and Global Supply Chain Disruption Index.** The current model relies on historical shipment data and customer portal inputs. Incorporating exogenous variables, such as global logistics disruption indices, raw material lead times, or macroeconomic indicators (e.g., inflation rates affecting the automotive aftermarket), could provide a more proactive forecasting layer.
- **Real-Time Learning and Adaptive Weighting:** In the current framework, the ensemble weights are determined based on a static historical validation phase. Transitioning to a Sliding Window approach, whereby the hybrid coefficients are dynamically re-optimized at the conclusion of each operational week, would enable the system to respond more agilely to abrupt market fluctuations, such as unforeseen surges in spare parts demand driven by regional vehicle aging cycles.

6. Conclusion

This study presents a robust Hybrid Ensemble forecasting framework designed to mitigate the systemic overestimation bias inherent in OEM customer portal data within the automotive spare parts supply chain. By synthesizing the strategic foresight of customer targets with the analytical precision of Random Forest, Support Vector Regression, and Long Short-Term Memory networks, the proposed model achieved a significant improvement in forecasting accuracy. The experimental results for the 2025 validation period confirm that the Hybrid Ensemble reached a monthly MAPE of 7.13 percent, substantially outperforming the baseline customer targets, which exhibited an error rate of 10.35 percent.

The operational projections for 2026 further demonstrate the strategic value of the hybrid approach. While standalone machine learning models often suffer from rigidity and fail to capture seasonal surges, the Hybrid Ensemble successfully identifies critical demand peaks while simultaneously filtering out erratic signals during low-demand weeks. This dual capability allows production facilities to maintain high service levels while reducing the financial risks associated with excess inventory and the bullwhip effect.

Ultimately, the findings suggest that a meta-regression-based integration of diverse algorithmic perspectives provides a more reliable foundation for strategic planning than relying on any single data source or algorithm in isolation. Future research will focus on the integration of non-linear aggregation methods and the incorporation of real-time exogenous variables to further refine the system's responsiveness to global market fluctuations.

References

- Alreshidi, E., Forecasting of Saudi Stock Prices Using Statistical and Machine Learning Models: A Multi-Model Comparative Approach, *International Journal of Advanced Computer Science and Applications*, vol. 16, no. 11, pp. 125-136, 2025.
- Andela, M., Ivan, M., Mora, D. and Arioli, S., Uncovering the Probiotic Supplement Landscape: Market Offerings, Sales Patterns, and Future Forecasts Using Machine Learning Approach — A Case Study of Montenegro, *Probiotics and Antimicrobial Proteins*, vol. 17, no. 6, pp. 4635-4658, 2025.
- Borucka, A., Seasonal Methods of Demand Forecasting in the Supply Chain as Support for the Company's Sustainable Growth, *Sustainability (Switzerland)*, vol. 15, no. 9, pp. 7399, 2023.
- Joseph, R. V., Mohanty, A., Tyagi, S., Mishra, S., Satapathy, S. K. and Mohanty, S. N., A hybrid deep learning framework with CNN and Bi-directional LSTM for store item demand forecasting, *Computers and Electrical Engineering*, vol. 103, pp. 108358, 2022.
- Karunaratne G., Thanthrige U.S.K.P.M., Machine Learning & IoT based Water Distribution System for Reservoirs in Sri Lanka. *Moratuwa Engineering Research Conference, MERCon 2025*, pp. 551 – 556, 2025.

- Psathas A.P., Iliadis L., Papaleonidas A., Strain Prediction of a Bridge Deploying Autoregressive Models with ARIMA and Machine Learning Algorithms. *Communications in Computer and Information Science*, 1826 CCIS, pp. 403 – 419, 2023.
- Sallam M., Deepika, Kaur K., Jindal N., Kaddoum G., Singh M., Singh Y.N., Energy Consumption Forecasting in Net-Zero Energy Buildings: Firefly-Driven LSTM for Smart Consumer Electronics and Edge-Based Energy Optimization. *IEEE Transactions on Consumer Electronics*, DOI 10.1109/TCE.2025.3646842, 2025
- Sharda, S., Singh, M. and Sharma, K., A complete consumer behaviour learning model for real-time demand response implementation in smart grid, *Applied Intelligence*, vol. 52, no. 1, pp. 835-845, 2022.
- Wunsch, A., Liesch, T. and Broda, S., Groundwater level forecasting with artificial neural networks: A comparison of long short-term memory (LSTM), convolutional neural networks (CNNs), and non-linear autoregressive networks with exogenous input (NARX), *Hydrology and Earth System Sciences*, vol. 25, no. 3, pp. 1671-1687, 2021.

Biographies

Irem Ucal Sari is a Professor in the Department of Industrial Engineering at Baskent University. She received her B.Sc. in Environmental Engineering (2004), M.Sc. in Engineering Management (2008), and Ph.D. in Industrial Engineering (2012) from Istanbul Technical University. Her research interests include fuzzy decision-making, fuzzy engineering economics, and fuzzy production planning. She has published more than 30 articles in reputable international journals. Since 2019, she has been serving as Co-Chair of the INFUS – Intelligent and Fuzzy Systems Conference.

Ekin Beste Bilgin is a Planning Engineer and a graduate student in Industrial Engineering at Başkent University. She received her B.Sc. degree in Industrial Engineering from Başkent University. Her professional experience focuses on production planning, demand forecasting, and supply chain coordination in the automotive electronics industry. She is responsible for order management, capacity planning, stock control, and coordination between production, purchasing, and logistics teams. Her research interests include demand forecasting, machine learning-based prediction models, and data-driven decision support systems. She has been working on forecasting models using methods such as Random Forest, Support Vector Regression (SVR), and Long Short-Term Memory (LSTM).