

Beyond One-Size-Fits-All: A Demand Classification and Decomposition Approach for Intermittent Demand Forecasting

Fitrotun Hasanah^{1*}, Melia Eka Lestiani²
Universitas Logistik dan Bisnis Internasional, Bandung, Indonesia
*Corresponding author email address: hasanahfnfn@gmail.com

Abstract. This study evaluated the performance of Croston, Syntetos–Boylan Approximation (SBA), and Teunter–Syntetos–Babai (TSB) methods in forecasting intermittent demand using a demand classification and decomposition approach in the document solutions industry. The analysis was conducted using monthly SKU-level data for 134 items from August 2023 to November 2025. Demand was classified into smooth, intermittent, erratic, and lumpy types based on the Average Demand Interval (ADI) and squared Coefficient of Variation (CV^2), and further analyzed by separating demand into total, rental, and sales categories. Forecast accuracy was evaluated using RMSE MAE. The results showed that forecasting performance varied across demand types, demand categories, and evaluation metrics. TSB was more effective in reducing large forecasting errors, while SBA provided more stable performance. These findings indicated that RMSE was more relevant for controlling extreme deviations, such as stockout risks, whereas MAE supported more stable inventory planning. In addition, demand aggregation tended to mask underlying variability, while demand decomposition provided more detailed insights. Overall, the study demonstrated that forecasting should not be applied uniformly across heterogeneous SKUs. A demand-based approach that integrates classification and decomposition enabled more accurate forecasting, reduced inventory risk, and supported more effective operational decision-making.

Keywords: Intermittent Demand, Demand Classification, Croston, Syntetos-Boylan-Approximation, Teunter-Syntetos-Babai

I. Introduction

Demand forecasting plays a critical role in inventory planning, production scheduling, and service level management. Forecast accuracy directly affects how organizations balance product availability with inventory-related costs, particularly at the Stock Keeping Unit (SKU) level, where replenishment and inventory control decisions are made (Babai et al., 2022). In environments characterized by demand uncertainty, inaccurate forecasts often lead to stockouts, excess inventory, and inefficient resource utilization, directly affecting operational performance.

At the SKU level, demand patterns are rarely smooth or stable. Many products exhibit intermittent, erratic, or lumpy demand, characterized by infrequent demand occurrences and high variability in demand size. Such demand behavior is commonly observed in spare parts, consumables, and service-related items. Although these items often contribute a limited share of total sales volume, empirical evidence shows that they may account for a substantial proportion of total inventory value, making accurate forecasting operationally critical (Babai et al., 2014).

Empirical studies have consistently shown that intermittent demand poses fundamental challenges to forecasting accuracy and inventory control. Forecast performance varies substantially across different demand patterns and no single forecasting approach performs best across all intermittent demand situations (Babai et al., 2020; Heinecke et al., 2013). As a result, applying forecasting methods uniformly across heterogeneous SKUs may reduce the operational relevance of forecasting outcomes, particularly in service-oriented distribution environments.

One sector particularly exposed to such demand uncertainty is the document solutions industry, which provides document management services, including printing, and scanning devices, consumables, spare parts, software solutions, and technical support. Companies operating in this industry typically manage large portfolios of SKUs with heterogeneous demand characteristics, especially for service-related items such as toner, drums, and spare parts. These characteristics are consistent with intermittent demand environments discussed in the spare parts and service supply chain literatures (Babai et al., 2014, 2020).

This research is motivated by operational challenges observed in Company X, a tier-1 and tier-2 distributor supplying printer and scanner products and related services for several global brands. Company X's business model includes device rental and sales with service support, requiring continuous fulfillment of consumables, spare parts, software, and technical services.

To illustrate the demand characteristics faced by Company X, Table 1 summarizes monthly demand occurrences across the observed SKU-level dataset for consumables from Brand Z covering the period from August 2023 to November 2025. The distribution shows that a large proportion of the observations correspond to zero-demand periods, indicating an intermittent demand environment at the operational level.

Table 1. Demand Distribution

Category	Count	Percentage (%)
Demand = 0	2201	58.66
Demand > 0	1551	41.34
Total_Demand	3752	100.00

Source: Authors' computation, 2026

Such demand irregularity implies that forecasting performance is likely to vary across SKUs with different demand characteristics, making uniform application of forecasting methods less effective. To address demand heterogeneity, demand classification has been proposed as a practical approach to support forecasting evaluation. Classification based on the Average Demand Interval (ADI) and the squared Coefficient of Variation (CV^2) of demand sizes enables demand to be grouped into smooth, intermittent, erratic, and lumpy types, providing a transparent basis for linking demand characteristics to forecasting analysis (Syntetos et al., 2005). Empirical studies have shown that classification-based approaches improve the interpretability and relevance of forecasting performance assessment at the SKU level (Heinecke et al., 2013).

Existing studies often focus on methodological comparisons, as evidenced by empirical results that primarily report improvements in aggregate forecasting accuracy across competing approaches. For example, (Tulli, 2020) showed that machine learning models such as Random Forest and LSTM consistently achieve lower RMSE and MAE than traditional statistical methods, including ARIMA and exponential smoothing, across pooled datasets. Similarly, (Ren, 2023) reported that machine learning and ensemble-based models outperform classical time-series methods in inventory demand forecasting, with overall error reductions observed at the dataset level. (Tang, 2024) further demonstrated that a hybrid LightGBM–PSO–LSTM model provided superior overall forecast fit compared to standalone LSTM, linear regression, and random forest models, thereby improving inventory turnover and OTIF performance. Likewise, (Seyedan et al., 2023) found that ensemble deep learning approaches reduced aggregate forecast errors and improved safety stock efficiency relative to ARIMA and exponential smoothing benchmarks.

While these studies provided clear empirical evidence of methodological performance differences, forecasting results are typically reported in aggregate and improvements in accuracy are not explicitly analyzed conditional on demand characteristics such as intermittency. In contrast, this study draws on real-world operational data from the document solutions industry, where a combination of equipment rental and product sales shapes demand.

This setting is rarely examined in prior forecasting studies, which are predominantly based on retail or distribution data. For instance, (Abolghasemi et al., 2020) analyzed demand forecasting performance using retail demand series affected by promotional activities, while (Mitra et al., 2022) evaluated forecasting models in a multi-channel retail company where sales transactions drive demand. Similarly, (Jahin et al., 2025) developed forecasting models using multi-channel sales data from retail and distribution networks, focusing on transaction-based demand rather than service- or rental-driven consumption.

The contribution of this research is threefold. First, it evaluates forecasting performance across demand categories, enabling transparent comparison under heterogeneous demand conditions. Second, it proposes a practical and interpretable decision-support framework by integrating demand classification and demand decomposition for forecasting evaluation. Third, it provides managerial insights into mitigating stockouts and overstocking by identifying the most suitable forecasting methods for specific demand behaviors.

Accordingly, this study aims to empirically evaluate the performance of intermittent demand forecasting using a classification-driven approach. By combining demand classification with empirical evaluation, this study seeks to provide practical insights into forecasting performance under intermittent demand while maintaining applicability beyond a single industry context.

II. Literature Review

2.1 Intermittent Demand and Forecast Challenges

Intermittent demand is characterized by infrequent non-zero demand occurrences combined with irregular demand magnitudes (Syntetos & Boylan, 2005). Early studies emphasized that conventional

forecasting methods designed for smooth demand perform poorly in these contexts due to the dominance of zero-demand periods and unstable parameter estimation (Croston, 1972; M. Williams, 1984). As a result, forecasting errors under intermittent demand tend to be biased and operationally misleading.

Subsequent empirical studies confirm that intermittent demand poses fundamental challenges distinct from those encountered under regular demand. Forecast accuracy deteriorates significantly as demand intermittency increases, particularly when demand series contain long sequences of zeros (Babai et al., 2014). As a result, applying forecasting methods uniformly across heterogeneous SKUs may lead to misleading accuracy assessments and reduced operational relevance.

At the SKU level, forecasting accuracy directly affects replenishment decisions, safety stock determination, and service level performance (Babai et al., 2014). However, empirical evidence suggests that SKU-level demand patterns are rarely smooth or stable. Many products, particularly spare parts and service-related items, exhibit intermittent, erratic, or lumpy demand behavior (Babai et al., 2014, 2020).

To address demand heterogeneity, demand classification has been proposed as a practical framework for structuring forecasting analysis. (Syntetos et al., 2005) introduced a widely adopted classification approach based on the Average Demand Interval (ADI) and the squared Coefficient of Variation (CV^2). CV^2 is computed from non-zero demand observations, while ADI represents the frequency of demand occurrence over time.

2.2 Croston-Based Forecasting Methods

Croston's method represents one of the earliest and most influential approaches developed specifically for intermittent demand forecasting. By decomposing demand into two independent components, demand size and inter-demand interval, Croston's method avoids updating forecasts during zero-demand periods. Empirical evidence suggests that this decomposition improves forecast stability compared to traditional exponential smoothing when demand is sporadic (Croston, 1972).

Despite its advantages, analytical investigations have demonstrated that the original Croston method produces biased estimates of demand per period under certain demand conditions. In particular, the method tends to exhibit bias, which may lead to systematic inaccuracies in forecasting (Syntetos & Boylan, 2005; R. Teunter & Sani, 2009).

To address this issue, Syntetos and Boylan proposed a bias-adjusted modification of Croston's method, commonly referred to as the Syntetos–Boylan Approximation (SBA) (Syntetos & Boylan, 2005). SBA introduces a correction factor designed to reduce the bias of the original estimator while retaining its core structure. Comparative studies have shown that SBA generally improves forecast accuracy relative to the original Croston method, although its performance remains sensitive to underlying demand characteristics (Babai et al., 2014; Syntetos et al., 2005).

The Teunter–Syntetos–Babai (TSB) method further extends this framework by explicitly modelling the probability of demand occurrence. Rather than relying solely on demand intervals, TSB updates the probability of demand occurrence at each period, enabling forecasts to adapt more quickly as demand patterns change. Empirical evidence suggests that this property makes TSB more suitable in environments where demand intermittency evolves (Babai et al., 2014; R. H. Teunter et al., 2011).

III. Research Method

3.1 Research Design and Data Source

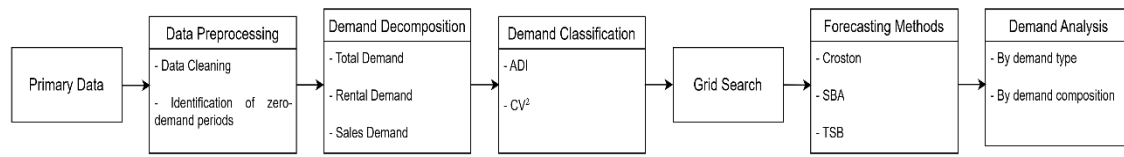
This study employed a quantitative empirical research design using primary data obtained directly from Company X, a document solutions company. The dataset consisted of historical SKU-level monthly demand records extracted from the company's ERP for toner and consumable product categories.

The research was conducted to analyze forecasting performance under heterogeneous demand conditions in a distribution and service-oriented supply chain environment. The observation period spanned from August 2023 to November 2025, enabling the identification of distinct demand patterns across SKUs. The analysis focused on SKU-level demand behavior in a service-oriented supply chain environment where demand arises from both printer rental usage and product sales.

3.2 Analytical Framework

The analytical framework consisted of sequential stages to examine forecasting behavior conditional on the demand structure and the demand composition, rather than relying solely on aggregated accuracy metrics. Figure 1 illustrates the analytical framework employed in this study, showing the flow from primary data to demand-aware comparative analysis.

Figure 1. Analytical Framework



Source: Author's illustration, 2026

3.3 Data Processing and Analytical Tools

Data processing, from data cleaning to analysis, was conducted using Python to ensure accuracy and computational efficiency when handling 134 SKUs, 3,752 observations and a total demand volume of 27,509 units. Data cleaning procedures were applied to remove duplicate records, handle missing values, and ensure temporal consistency across SKUs. Only active SKUs were retained in the dataset, as demand forecasting is required only for products currently in operation. All demand series were aggregated to a monthly frequency prior to analysis.

Data preprocessing and demand classification were performed using pandas and NumPy. To ensure higher forecasting precision, a grid search algorithm was employed for smoothing parameter optimization instead of adopting fixed values. The parameter range for α was defined based on commonly recommended values in Croston-based methods (0.05–0.20), and was extended up to 0.30 to incorporate the higher smoothing values explored in the TSB study. For the TSB method, the smoothing parameter β was tested between 0.01 and 0.30, following the experimental setup used in the TSB study. A grid search with increments of 0.01 was applied to obtain more precise parameter estimates.

Forecast accuracy was evaluated using MAE and RMSE computed with the sklearn.metrics module. In cases where demand occurred at most once across the entire observation period, default parameter values $\alpha = 0.05$ and $\beta = 0.01$ were applied to ensure minimal model updates, as lower smoothing parameters reduce the influence of isolated demand occurrences under extremely sparse demand conditions.

Demand decomposition was divided into rent, sales, and total demand. Demand by rent refers to toner and consumables delivered for rented printers, while demand by sales corresponds to products sold directly. Total demand represents the aggregation of both components. This decomposition was used to compare demand behavior and assess intermittent demand characteristics across different demand sources.

3.4 Demand Classification

Demand classification was conducted using the ADI and CV^2 , widely adopted measures for characterizing intermittent demand patterns. The ADI measured demand intermittency and was calculated as the ratio between the total number of periods and the number of non-zero demand occurrences:

$$ADI = \frac{N}{\sum_{t=1}^N I(y_t > 0)}$$

- N : total number of periods
- t : time period index
- y_t : actual demand in period t
- $I(y_t > 0)$: demand occurrence indicator in period t , 1 when demand occurs and 0 otherwise

Demand variability was measured using the CV^2 , calculated based on non-zero demand observations:

$$CV^2 = \left(\frac{\sigma}{\mu}\right)^2$$

- σ : standard deviation of demand
- μ : mean of demand

Based on established threshold values, SKUs were classified into smooth, intermittent, erratic, or lumpy demand types. The classification criteria are summarized in Table 2.

Table 2. Demand Pattern Classification Based on ADI and CV^2

Demand Type	Demand Intermittency (ADI)	Demand Variability (CV^2)	Key Characteristics
-------------	----------------------------	-------------------------------	---------------------

Smooth	$ADI \leq 1.32$	$CV^2 \leq 0.49$	Demand occurs in almost every period with low variability
Intermittent	$ADI > 1.32$	$CV^2 \leq 0.49$	Demand occurs infrequently but demand sizes are relatively stable when they occur
Erratic	$ADI \leq 1.32$	$CV^2 > 0.49$	Demand occurs frequently but demand sizes are highly variable
Lumpy	$ADI > 1.32$	$CV^2 > 0.49$	Demand occurs infrequently and demand sizes are highly variable

Source: Adapted from Syntetos, Boylan and Croston, 2005

Although demand was classified into four categories, the primary focus of this study remained on intermittent demand conditions. Intermittent, erratic, and lumpy demand types were considered representative of intermittent demand environments due to their irregular demand occurrence and higher variability. Smooth demand was included to provide a baseline comparison and to ensure a comprehensive evaluation across all demand patterns.

3.5 Forecasting Methods

This study evaluated and compared the performance of three widely recognized methods specifically designed for intermittent demand, namely Croston's method, SBA, and the TSB method. These methods are considered the standard benchmarks in the field due to the frequent adoption in industrial inventory management systems.

3.5.1 Croston's Method

Croston separated intermittent demand into demand size and demand interval. The demand size captured how large the demand was when it occurred, while the interval captured how frequently it occurred. These two components were updated only when non-zero demand was observed, and the forecast was obtained by dividing the estimated demand size by the estimated interval.

$$\begin{aligned}\hat{z}_t &= \hat{z}_{t-1} + \alpha(z_t - \hat{z}_{t-1}) \\ \hat{p}_t &= \hat{p}_{t-1} + \alpha(q_t - \hat{p}_{t-1})\end{aligned}$$

The forecast was computed as:

$$\hat{y}_t = \frac{\hat{z}_t}{\hat{p}_t}$$

\hat{y}_t	: forecasted demand in period t
\hat{z}_t	: estimated demand size in period t
\hat{p}_t	: estimated non-zero demand interval in period t
z_t	: actual demand size in period t
p_t	: actual interval between non-zero demands
α	: smoothing parameter
\hat{p}_{t-1}	: previous period's estimate of the demand interval
\hat{z}_{t-1}	: previous period's estimate of demand size
q_t	: number of periods since the last non-zero demand

3.5.2 Syntetos-Boylan Approximation (SBA)

The SBA method adjusted Croston's estimator to correct for positive bias. The forecast was computed as:

$$\hat{y}_t^{SBA} = \left(1 - \frac{\alpha}{2}\right) \frac{\hat{z}_t}{\hat{p}_t}$$

3.5.3 Teunter–Syntetos–Babai (TSB)

The TSB method explicitly models demand occurrence by estimating the probability of demand occurrence p_t and the expected demand size z_t . The probability of demand occurrence was updated in every period as:

$$\hat{p}_t = \begin{cases} \hat{p}_t = \hat{p}_{t-1} + \beta(0 - \hat{p}_{t-1}), & \text{if } p_t = 0 \\ \hat{p}_t = \hat{p}_{t-1} + \beta(1 - \hat{p}_{t-1}), & \text{if } p_t > 0 \end{cases}$$

The demand size estimate was updated only when demand occurred:

$$\hat{z}_t = \begin{cases} \hat{z}_t = z_{t-1}, & \text{if } p_t = 0 \\ \hat{z}_t = \hat{z}_{t-1} + \alpha(z_t - \hat{z}_{t-1}), & \text{if } p_t > 0 \end{cases}$$

The forecast was given by:

$$\hat{y}_t^{TSB} = \hat{p}_t \times \hat{z}_t$$

p_t	: demand occurrence indicator in period t , where: $p_t = 1$ if demand occurs and $p_t = 0$ otherwise
\hat{p}_t	: estimated probability of demand occurrence in period t
α	: smoothing parameter for demand size
β	: smoothing parameter for demand occurrence probability

3.6 Forecasting Performance Evaluation

Forecast performance was evaluated using MAE and RMSE, computed as:

$$MAE = \frac{1}{T} \sum_{t=1}^T |y_t - \hat{y}_t| \qquad RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2}$$

t	: time period index
T	: total number of periods

MAE and RMSE were selected because they provide stable and reliable accuracy assessments in the presence of data irregularities, such as demand spikes and variability (Khoshvaght et al., 2025). Performance metrics were calculated at the SKU-month level and summarized by demand type, demand composition, and forecasting method.

IV. Results and Discussion

4.1 Results

Forecast performance was analyzed separately for rental, sales, and total demand to capture differences in demand behavior and instability. Results were reported using MAE and RMSE to assess forecast deviations, while comparisons were interpreted conditionally on the demand structure rather than solely on aggregated accuracy.

4.1.1 Demand Category Distribution

Table 3 presented the distribution of zero and non-zero demand periods across the company's operations:

Table 3. Demand Category Distribution

Demand	Qty	Count	Percentage (%)
Rent	0	2565	68.36
Rent	> 0	1187	31.64
Rent TOTAL		3752	100.00
Sales	0	2670	71.16
Sales	> 0	1082	28.84
Sales TOTAL		3752	100.00
Total	0	2201	58.66
Total	> 0	1551	41.34
Total TOTAL		3752	100.00

Source: Author’s computation, 2026

The results indicated that 68.36% of rental periods and 71.16% of sales periods recorded zero demand, confirming that both channels exhibited strong intermittent demand characteristics. Rental activity tended to occur more frequently than sales, suggesting that service contracts generated relatively consistent, though still infrequent, demand occurrences. In contrast, sales were more irregular and were primarily driven by project-based activities.

Overall, 58.66% of the total periods showed no demand, confirming that the toner and consumable categories exhibited irregular and infrequent demand patterns. These results highlighted the inherent variability in demand behavior and emphasized the importance of applying forecasting approaches specifically designed for intermittent demand when supporting operational planning and inventory management decisions.

4.1.2 Demand Type Distribution

The demand classification result, as presented in Table 4, indicated that the intermittent demand dominated the dataset, accounting for 50.75% of total SKUs, followed by lumpy (20.15%), erratic (15.67%), and smooth demand (13.43%). When disaggregated by demand source, rental demand was primarily dominated by intermittent demand (69.40%) and exhibited no erratic demand, while smooth and lumpy demand accounted for 19.41% and 11.19%, respectively. This pattern suggested that rental demand occurred infrequently but with relatively consistent demand sizes.

Table 4. Demand Type Distribution

Demand_Type	SKU_Total	Percentage_Total (%)	SKU_Rent	Percentage_Rent (%)	SKU_Sales	Percentage_Sales (%)
Erratic	21	15.67	0	0.00	6	4.48
Intermittent	68	50.75	93	69.40	76	56.72
Lumpy	27	20.15	15	11.19	43	32.09
Smooth	18	13.43	26	19.41	9	6.71
Total	134	100.00	134	100.00	134	100.00

Source: Author’s computation, 2026

In contrast, sales demand demonstrated a more heterogeneous distribution. Although intermittent demand remained the largest category (56.72%), lumpy demand represented a substantial share (32.09%), while smooth and erratic demand accounted for 6.71% and 4.48%, respectively. This indicated that sales demand was more variable and prone to fluctuations compared to rental demand. These differences suggested that demand behavior varied significantly across demand sources, thereby supporting the need for separate evaluation of forecasting performance.

4.1.3 Smoothing Parameters Selection

Table 5 summarized the grid search results which demonstrated a clear distinction in parameter behavior across the different demand types. The optimal parameters were selected based on the lowest RMSE and MAE values obtained from the grid search.

Table 5. Grid Search Result

Demand	Method	Total_SKU	Identical_α	Percentage_α (%)	Identical_β	Percentage_β (%)	Most_Common_α	Count_α	Most_Common_β	Count_β	Total_Default
Total	Croston	134	87	64.93	-	-	0.30	49	-	-	8
Total	SBA	134	93	69.40	-	-	0.30	65	-	-	8
Total	TSB	134	87	64.93	97	72.39	0.05	62	0.01	59	8
Rent	Croston	134	103	76.87	-	-	0.05	77	-	-	50
Rent	SBA	134	101	75.37	-	-	0.05	69	-	-	50
Rent	TSB	134	103	76.87	101	75.37	0.05	97	0.01	82	50
Sales	Croston	134	102	76.12	-	-	0.05	65	-	-	18
Sales	SBA	134	98	73.13	-	-	0.30	64	-	-	18
Sales	TSB	134	107	79.85	86	64.18	0.05	78	0.30	57	18

Source: Author’s computation, 2026

For total demand, Croston and SBA were dominated by higher smoothing parameters ($\alpha = 0.30$), indicating the need for more responsive updates to capture overall demand variability. In contrast, TSB was characterized by lower smoothing parameters ($\alpha = 0.05$ and $\beta = 0.01$), reflecting a preference for gradual updates in both demand size and occurrence probability. The relatively low number of default parameters suggested that most SKUs contained sufficient information for parameter optimization.

For rental demand, all methods consistently converged to lower smoothing parameters, with $\alpha = 0.05$ dominating across Croston, SBA, and TSB. This indicated highly sparse demand conditions, where stable updates were required to avoid overreacting to limited observations. TSB also showed consistent β values at low levels, reinforcing the need for gradual updates in demand occurrence. The high number of default parameters further confirmed the limited demand information available in rental demand.

For sales demand, a more heterogeneous pattern was observed. While Croston and TSB remained dominated by lower α values (0.05), SBA shifted toward higher responsiveness with $\alpha = 0.30$. In addition, TSB showed a significant increase in β (0.30), indicating more adaptive updates in demand occurrence probability. This suggested that sales demand contained more observable demand events, allowing for more responsive adjustment compared to rental demand. The lower number of default parameters supported the presence of sufficient demand observations for parameter estimation.

Overall, these results demonstrated that parameter selection was strongly influenced by demand category. Total demand required more responsive updates, rental demand favored stability due to sparsity, and sales demand required a balance between stable demand size estimation and adaptive occurrence modeling. This reflected the trade-off between responsiveness and stability in intermittent demand forecasting.

4.1.4 Forecasting Performance

Table 6 presented the best-performing forecasting methods across demand categories based on RMSE and MAE.

Table 6. Best Method Selection by Demand Category

Demand_Category	Best_Method_RMSE	SKU_Count_RMSE	Best_Method_MAE	SKU_Count_MAE	Total_SKU
Total	TSB	60	TSB	69	129
Rent	TSB	40	TSB	44	84
Sales	TSB	62	TSB	68	130

Source: Author’s computation, 2026

The results showed that the best-performing forecasting method varied across demand categories and evaluation metrics. The best method selection in Table 6 was determined at the SKU level, where each SKU was assigned to the method that produced the lowest RMSE or MAE. In contrast, the performance results in Tables 7–9 were calculated as average error values across demand types, which reflected overall forecasting accuracy rather than the frequency of method selection.

Across all demand categories, TSB consistently emerged as the most frequently selected method under both RMSE and MAE. The dominance of TSB can be attributed to its ability to dynamically update both demand size and the probability of demand occurrence, allowing it to better capture intermittent and irregular demand patterns. As a result, TSB was more adaptable to variations in demand frequency and magnitude, making it more likely to produce the lowest forecasting errors across a wide range of SKUs compared to Croston and SBA, which rely on more static updating mechanisms.

The number of SKUs evaluated under RMSE and MAE did not match the total dataset size. This discrepancy occurred because some SKUs did not provide sufficient data points to generate meaningful error calculations, leading to their exclusion from the RMSE and MAE evaluation. This condition reflected a common characteristic of intermittent demand environments, where certain items lacked adequate demand occurrences for reliable forecasting assessment.

Table 7. Forecasting Performance of Total Demand

Method	Demand_Type	SKU_Count	Avg_RMSE	Avg_MAE
Croston	Erratic	21	11.75	8.56
SBA	Erratic	21	11.15	7.96
TSB	Erratic	21	11.75	8.53
Croston	Intermittent	68	2.25	1.94
SBA	Intermittent	68	2.07	1.76
TSB	Intermittent	68	1.77	1.36
Croston	Lumpy	27	7.20	5.03
SBA	Lumpy	27	7.11	4.92
TSB	Lumpy	27	6.92	4.56
Croston	Smooth	18	17.05	13.25
SBA	Smooth	18	16.89	12.73
TSB	Smooth	18	17.05	13.24

Source: Author's computation, 2026

The forecasting performance for total demand, as presented in Table 7, showed clear differences across demand types and forecasting methods. For erratic demand, SBA achieved the lowest RMSE (11.15) and MAE (7.96), indicating that it provided more stable forecasts under highly variable demand conditions.

For both intermittent and lumpy demand, TSB consistently outperformed other methods, achieving the lowest RMSE and MAE values across both demand types. This demonstrated its strong capability in handling irregular demand occurrences and variability through adaptive updating. In contrast, for smooth demand, SBA slightly outperformed other methods with the lowest RMSE (16.89) and MAE (12.73), although the differences between methods were relatively small.

Overall, these results indicated that forecasting performance varied depending on demand characteristics, where TSB performed better under intermittent and variable demand conditions, while SBA provided more stable performance under erratic and relatively smoother demand patterns.

Table 8. Forecasting Performance of Rental Demand

Method	Demand_Type	SKU_Count	Avg_RMSE	Avg_MAE
Croston	Erratic	NaN	NaN	NaN
SBA	Erratic	NaN	NaN	NaN
TSB	Erratic	NaN	NaN	NaN
Croston	Intermittent	93	1.18	1.03
SBA	Intermittent	93	1.11	0.97
TSB	Intermittent	93	0.96	0.78
Croston	Lumpy	15	5.57	4.40
SBA	Lumpy	15	5.58	4.37
TSB	Lumpy	15	5.50	4.33
Croston	Smooth	26	6.38	4.92
SBA	Smooth	26	6.46	4.86
TSB	Smooth	26	6.38	4.91

Source: Author's computation, 2026

The forecasting performance for rental demand, as presented in Table 8, showed a distinct pattern compared to total demand, reflecting its more sparse and stable demand characteristics. No SKUs were classified under erratic demand, resulting in unavailable RMSE and MAE values for all methods.

For both intermittent and lumpy demand, TSB consistently provided the best performance, achieving the lowest RMSE and MAE values across both demand types. This indicated its strong ability to handle irregular demand occurrences and variability through adaptive updating of both demand size and occurrence probability.

In contrast, for smooth demand, the differences between methods were minimal. While TSB and Croston produced identical RMSE values (6.38), SBA achieved the lowest MAE (4.86), suggesting that all methods performed comparably under more stable demand conditions, with slight differences depending on the evaluation metric.

Overall, these findings indicated that TSB demonstrated superior performance under more complex and irregular demand conditions, while differences between methods became less pronounced as demand patterns became more stable. Croston, in contrast, did not outperform other methods in any demand type, reflecting its limited responsiveness and lack of bias correction compared to SBA and TSB.

Table 9. Forecasting Performance of Sales Demand

Method	Demand_Type	SKU_Count	Avg_RMSE	Avg_MAE
Croston	Erratic	6	23.47	18.94
SBA	Erratic	6	22.97	18.39
TSB	Erratic	6	23.29	18.73
Croston	Intermittent	76	2.42	2.13
SBA	Intermittent	76	2.18	1.90
TSB	Intermittent	76	1.88	1.48
Croston	Lumpy	43	8.72	6.47
SBA	Lumpy	43	8.23	6.00
TSB	Lumpy	43	8.41	5.89
Croston	Smooth	9	9.99	8.25
SBA	Smooth	9	9.98	8.19
TSB	Smooth	9	9.82	8.07

Source: Author's computation, 2026

The average forecasting performance for sales demand, as presented in Table 9, showed that forecasting accuracy varied across demand types and evaluation metrics. For erratic demand, SBA achieved the lowest RMSE (22.97) and MAE (18.39), indicating its ability to provide more stable forecasts under highly variable demand conditions.

For both intermittent and smooth demand, TSB consistently outperformed other methods, achieving the lowest RMSE and MAE values in both demand types. This demonstrated its effectiveness in handling irregular demand occurrences as well as maintaining accuracy under more stable demand conditions.

For lumpy demand, the results showed mixed performance, where SBA achieved the lowest RMSE (8.23), while TSB produced the lowest MAE (5.89). This indicated that different methods were preferable depending on the evaluation metric, with SBA better at reducing larger errors and TSB more effective in minimizing absolute deviations.

Overall, these findings suggested that no single method consistently dominated across all demand types in sales demand. While SBA performed well under erratic conditions, TSB demonstrated strong and consistent performance across intermittent and smooth demand, with competitive results in other demand types.

4.2 Discussion

The results indicated that forecasting performance was influenced by both the evaluation metric and demand structure. The difference between RMSE and MAE explained variations in method performance, as RMSE placed greater emphasis on large errors, while MAE evaluated errors in absolute terms. From an operational perspective, RMSE was more relevant in situations where large forecasting errors had significant consequences, such as stockouts or emergency replenishment, as it penalized extreme deviations more heavily. In contrast, MAE was more appropriate for maintaining overall forecasting stability, as it provided a balanced measure of average error without disproportionately weighting extreme values.

As a result, methods that adapted more quickly to demand changes, such as TSB, tended to perform better under RMSE, while SBA provided more stable performance under MAE due to its bias-adjustment mechanism. However, this pattern was not consistent across all demand categories, indicating that method effectiveness depended on both the metric and the underlying demand characteristics.

In addition, the difference between SKU-level results (Table 6) and average performance results (Tables 7–9) highlighted an important distinction in forecasting evaluation. SKU-level selection reflected how frequently a method achieved the lowest error and was more relevant for operational decision-making, where forecasting methods were applied at the individual SKU level. In contrast, average performance reflected the magnitude of forecasting errors across all observations and was more appropriate for assessing overall model accuracy and benchmarking forecasting methods. As a result, a method could be frequently selected as the best without necessarily producing the lowest overall error, particularly when large forecasting errors occurred in some SKUs. This implied that relying solely on average performance could overlook methods that were consistently reliable across many SKUs, while relying only on SKU-level selection could ignore the impact of large forecasting errors in some SKUs.

The results also showed clear differences in method performance across demand types. TSB consistently performed better under intermittent and lumpy demand conditions, where demand occurrences were irregular and required adaptive updating. In contrast, SBA performed better under erratic and relatively stable demand conditions, where reducing bias and maintaining consistency were more important. Croston, however, did not outperform other methods across any demand type, confirming its limitations due to the absence of bias correction and adaptive updating mechanisms. These findings were consistent with prior studies on Croston-based methods (Syntetos & Boylan, 2005; R. H. Teunter et al., 2011).

Furthermore, the comparison across demand sources revealed that rental demand was sparser and more stable, while sales demand exhibited higher variability and fluctuation. These differences explained why forecasting performance varied across demand categories and confirmed the importance of demand classification in forecasting evaluation.

The results also demonstrated that demand aggregation influenced the interpretation of forecasting performance. In some cases, total demand produced lower error values compared to disaggregated demand. However, this did not necessarily indicate better forecasting performance, as aggregation masked underlying variability by combining different demand patterns. When demand was separated, higher errors were observed, particularly in sales demand, indicating that the underlying demand was more difficult to predict than suggested by aggregated results.

From operational perspective, these findings implied that forecasting methods should be selected based on demand characteristics rather than applied uniformly across all SKUs. For demand with high intermittency and variability, such as intermittent and lumpy demand, adaptive methods like TSB were more appropriate, particularly to reduce the risk of large forecasting deviations that could lead to stockouts. For demand with relatively stable patterns, SBA provided more consistent performance, supporting more stable inventory control and reducing unnecessary fluctuations in replenishment decisions.

In practice, total demand forecasting remained sufficient for estimating overall demand volume, especially in environments where inventory was managed as a unified pool and demand from rental and sales was interchangeable. In such cases, aggregation enabled risk pooling, allowing variability from individual demand sources to be absorbed at the system level, thereby supporting more efficient inventory utilization. However, demand decomposition provided greater managerial insight by revealing differences in demand behavior across sources, enabling more targeted forecasting strategies, better safety stock allocation, and improved responsiveness to demand variability.

V. Conclusion

The findings of this study demonstrated that no single forecasting method consistently outperformed others across all demand types, confirming that forecasting should not be applied uniformly across heterogeneous SKUs. Instead, forecasting methods should be selected based on demand characteristics identified through classification using ADI and CV^2 .

The empirical results showed that SBA provided more stable and consistent forecasting performance, particularly under MAE evaluation, while TSB was more effective in reducing large forecasting errors under RMSE. These differences had direct operational implications. Minimizing variability in forecasting errors supported more stable inventory levels, enabled smoother replenishment planning, reduced unnecessary inventory fluctuations, and improved cost control (Babai et al., 2022). In contrast, minimizing extreme forecasting deviations was critical for reducing the risk of stockouts, avoiding sudden shortages and preventing lost sales or emergency replenishment costs (Babai et al., 2014; R. H. Teunter et al., 2011)..

The study also showed that method effectiveness varied across demand types. TSB was more suitable for intermittent and lumpy demand due to its adaptive updating mechanism, while SBA performed better under erratic and relatively stable demand conditions by providing more consistent forecasts. For smooth demand, performance differences between methods were relatively small, allowing method selection to be based on operational preference.

In addition, the results consistently showed that the original Croston method underperformed compared to SBA and TSB. This finding supported prior studies that have highlighted the limitations of Croston in intermittent demand environments, particularly due to its lack of bias correction and limited responsiveness to changes in demand occurrence (Syntetos & Boylan, 2005; R. Teunter & Sani, 2009).

In the context of the document solutions industry, where demand arises from both rental and sales activities, these differences were particularly relevant. Rental demand tended to be sparser and more stable, while sales demand exhibited higher variability. Aligning forecasting methods with these characteristics enabled better forecasting accuracy and more effective inventory management.

From an operational perspective, demand classification played a critical role in improving forecasting relevance. Rather than applying a single method across all SKUs, organizations should adopt a structured, demand-aware forecasting approach, where demand is first classified and decomposed, and then matched with appropriate forecasting methods. This approach enabled more accurate forecasting, better inventory planning, reduced risk of overstock and stockout and improved alignment between forecasting outcomes and operational requirements (Heinecke et al., 2013; Syntetos & Boylan, 2005).

This study was limited to a single company and focused on three classical forecasting methods. In addition, the evaluation relied on RMSE and MAE, which may not fully capture operational impacts such as service level or inventory cost. Future research could extend this study by applying it across different industries, incorporating advanced forecasting methods, and integrating operational performance metrics to better link forecasting results with inventory decision-making.

References

- Abolghasemi, M., Beh, E., Tarr, G., & Gerlach, R. (2020). Demand forecasting in supply chain: The impact of demand volatility in the presence of promotion. *Computers and Industrial Engineering*, 142. <https://doi.org/10.1016/j.cie.2020.106380>
- Babai, M. Z., Boylan, J. E., & Rostami-Tabar, B. (2022). Demand forecasting in supply chains: a review of aggregation and hierarchical approaches. *International Journal of Production Research*, 60(1), 324–348. <https://doi.org/10.1080/00207543.2021.2005268>
- Babai, M. Z., Syntetos, A., & Teunter, R. (2014). Intermittent demand forecasting: An empirical study on accuracy and the risk of obsolescence. *International Journal of Production Economics*, 157, 212–219. <https://doi.org/10.1016/j.ijpe.2014.08.019>
- Babai, M. Z., Tsadiras, A., & Papadopoulos, C. (2020). On the empirical performance of some new neural network methods for forecasting intermittent demand. *IMA Journal of Management Mathematics*, 31(3), 281–305. <https://doi.org/10.1093/imaman/dpaa003>
- Croston, J. D. (1972). *Forecasting and Stock Control for Intermittent Demands*. Palgrave Macmillan Journals and Operational Research Society Are Collaborating with JSTOR to Digitize, Preserve and Extend Access to *Operational Research Quarterly (1970-1977)*, 23(3), 289–303.
- Heinecke, G., Syntetos, A. A., & Wang, W. (2013). Forecasting-based SKU classification. *International Journal of Production Economics*, 143(2), 455–462. <https://doi.org/10.1016/j.ijpe.2011.11.020>
- Jahin, M. A., Shahriar, A., & Amin, M. Al. (2025). MCDFN: supply chain demand forecasting via an explainable multi-channel data fusion network model. *Evolutionary Intelligence*, 18(3), 1–27. <https://doi.org/10.1007/s12065-025-01053-7>
- Khoshvaght, H., Permal, R. R., Razmjou, A., & Khiadani, M. (2025). A critical review on selecting performance evaluation metrics for supervised machine learning models in wastewater quality prediction. *Journal of Environmental Chemical Engineering*, 13(6), 119675. <https://doi.org/10.1016/j.jece.2025.119675>
- M. Williams, T. (1984). Stock Control with Sporadic and Slow-Moving Demand. *Journal of the Operational Research Society*, 35, 939–948.

- Mitra, A., Jain, A., Kishore, A., & Kumar, P. (2022). A Comparative Study of Demand Forecasting Models for a Multi-Channel Retail Company: A Novel Hybrid Machine Learning Approach. *Operations Research Forum*, 3(4), 1–22. <https://doi.org/10.1007/s43069-022-00166-4>
- Ren, H. (2023). Enhanced Inventory Demand Forecasting with Machine Learning (Number Icbdss). Atlantis Press International BV. <https://doi.org/10.2991/978-94-6463-276-7>
- Seyedan, M., Mafakheri, F., & Wang, C. (2023). Order-up-to-level inventory optimization model using time-series demand forecasting with ensemble deep learning. *Supply Chain Analytics*, 3. <https://doi.org/10.1016/j.sca.2023.100024>
- Syntetos, A. A., & Boylan, J. E. (2005). The accuracy of intermittent demand estimates. *International Journal of Forecasting*, 21(2), 303–314. <https://doi.org/10.1016/j.ijforecast.2004.10.001>
- 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. <https://doi.org/10.1057/palgrave.jors.2601841>
- Tang, W. (2024). Improvement of Inventory Management and Demand Forecasting by Big Data Analytics in Supply Chain. *Applied Mathematics and Nonlinear Sciences*, 9(1). <https://doi.org/10.2478/amns-2024-2213>
- Teunter, R. H., Syntetos, A. A., & Babai, M. Z. (2011). Intermittent demand: Linking forecasting to inventory obsolescence. *European Journal of Operational Research*, 214(3), 606–615. <https://doi.org/10.1016/j.ejor.2011.05.018>
- Teunter, R., & Sani, B. (2009). On the bias of Croston's forecasting method. *European Journal of Operational Research*, 194(1), 177–183. <https://doi.org/10.1016/j.ejor.2007.12.001>
- Tulli, S. K. C. (2020). Comparative Analysis of Traditional and AI-based Demand Forecasting Models. *International Journal of Emerging Trends in Science and Technology*, 07(06), 6933–6956. <https://doi.org/10.18535/ijetst/v7i6.02>