

Event-preserving feature engineering for intermittent demand forecasting using SHOS

Received: 10 March 2026

Accepted: 5 June 2026

Published online: 13 June 2026

Cite this article as: Nathan B.S., B V.S.R., Sastry C.C. *et al.* Event-preserving feature engineering for intermittent demand forecasting using SHOS. *Sci Rep* (2026). <https://doi.org/10.1038/s41598-026-57129-6>

B. Sendhil Nathan, Veera Siva Reddy B, C. Chandrasekhara Sastry, Sachin Salunkhe & Robert Cep

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

Event-Preserving Feature Engineering for Intermittent Demand Forecasting using SHOS

B. Sendhil Nathan^{1,2}, B. Veera Siva Reddy^{1*}, C. Chandrasekhara Sastry^{1*},
Sachin Salunkhe³, Robert Cep⁴

¹Department of Mechanical Engineering, Indian Institute of Information
Technology Design and Manufacturing Kurnool (IIITDMK), Kurnool 518008,
Andhra Pradesh, India

²Ford Motor Private Limited, Ford Global Technology & Business Center,
Chennai 600119, Tamil Nadu, India

³Department of Mechanical Engineering, Gazi University, Ankara, Turkey

⁴Department of Machining, Assembly and Engineering Metrology, Faculty
of Mechanical Engineering, VSB-Technical University of Ostrava, 70800
Ostrava, Czech Republic

*Corresponding authors: sivareddy.bobbili@gmail.com,
chandrasekhar@iiitk.ac.in

Corresponding Authors

Mr. Bobbili Veera Siva Reddy
Department of Mechanical Engineering
Indian Institute of Information Technology Design and Manufacturing
Kurnool (IIITDM Kurnool) Kurnool, Andhra Pradesh - 518008.
Mobile No.: +91-7799683467
(Centrally Funded Technical Institute (CFTI)
Institute of National Importance under Ministry of Education, Govt. of
India)

Dr. C. Chandrasekhara Sastry
Assistant Professor (Grade I), Department of Mechanical Engineering
Indian Institute of Information Technology Design and Manufacturing
Kurnool (IIITDM Kurnool), Kurnool, Andhra Pradesh - 518008,
INSA Visiting Scientist 2024, IIT Hyderabad - IIITDM Kurnool
Office No. : +918518-289100 (Extension: 210)
Email Address: chandrasekhar@iiitk.ac.in;
chandrasekhar041991@gmail.com
URL : <https://iiitk.ac.in/Dr.-C.-Chandrasekhara-Sastry/page>
(Centrally Funded Technical Institute (CFTI))

Abstract

Intermittent demand forecasting remains a persistent challenge in large-scale supply chains due to extreme sparsity, irregular demand occurrence, and highly variable demand magnitudes. While recent studies have increasingly emphasized architectural complexity to address these challenges, comparatively less attention has been directed toward the role of statistically grounded feature representation under sparse-demand conditions. This study introduces the Smoothed Hybrid Occurrence-Size (SHOS) framework as an event-preserving feature-engineering approach for intermittent-demand forecasting. SHOS decomposes demand into latent occurrence probability and conditional demand size components and generates adaptive, series-specific statistical representations using sparsity-aware exponential smoothing. Rather than functioning as a standalone forecasting model, SHOS operates as a feature-generation mechanism that embeds occurrence-size structure directly into supervised learning pipelines. The proposed framework was evaluated using a large-scale automotive aftermarket dataset consisting of approximately 56,000 dealer-part time series and approximately 1.4 million monthly observations generated after preprocessing and zero-padding. Models were evaluated using rolling-window cross-validation, signal-preservation analysis, and robustness studies across heterogeneous demand segments. Under the evaluated sparse-demand setting and model configurations, SHOS-augmented tree-based models substantially improved forecasting accuracy and stability relative to

raw-feature baselines. In particular, the SHOS-augmented LightGBM configuration reduced the mean absolute error by approximately 46.3% relative to the baseline model, while the weighted mean absolute percentage error improved by more than 40%. Beyond conventional forecasting accuracy, the study additionally demonstrates that SHOS preserves event timing, peak structure, and frequency-domain characteristics that are often distorted by conventional smoothing approaches. The results further suggest that statistically grounded feature representations can substantially improve sparse-demand learning behaviour under the evaluated automotive aftermarket forecasting environment. Rather than establishing a universal hierarchy between representation and architectural complexity, the findings demonstrate the effectiveness of event-preserving representation-oriented forecasting under highly intermittent demand conditions.

Keywords: Event preservation; Feature engineering; Intermittent demand forecasting; Machine learning; SHOS; Sparse time series

1. Introduction

Intermittent demand forecasting remains a major challenge in supply chain and inventory management, due to long periods of zero demand interspersed with irregular, low-frequency demand events. This problem is especially critical in automotive aftermarket and spare-parts systems, where organizations manage large populations of slow-moving stock-keeping units (SKUs) across distributed dealer networks [1-3]. Under such conditions,

inaccurate forecasts can simultaneously increase inventory holding costs, obsolescence risk, and service-level failures.

Classical intermittent-demand forecasting methods such as Croston's method, the Syntetos-Boylan Approximation (SBA), and the Teunter-Syntetos-Babai (TSB) framework address this challenge by decomposing demand into occurrence and conditional size components [4-6]. These approaches remain widely used because they provide stable and interpretable forecasting behaviour under sparse-demand conditions. However, they operate primarily as univariate forecasting procedures and cannot directly exploit auxiliary variables or cross-sectional information increasingly available in modern enterprise systems.

Recent forecasting research has therefore increasingly explored machine-learning and deep-learning approaches capable of learning nonlinear relationships from large multivariate datasets [7-10]. Tree-based learners such as LightGBM and XGBoost have demonstrated strong predictive performance in tabular forecasting problems, while deep learning architectures aim to capture complex temporal structure [7,8]. However, intermittent-demand environments remain challenging for such models due to extreme sparsity, irregular demand occurrences, and weak temporal continuity, which limit the availability of stable learning signals.

At the same time, forecasting competitions and hybrid forecasting studies increasingly suggest that statistically meaningful feature

representations may substantially improve forecasting performance even without highly specialized architectures [11,12]. This raises an important question in intermittent-demand forecasting: can statistically grounded feature representations improve sparse-demand learning behaviour more effectively than increasing architectural complexity alone.

To address this problem, this study introduces the Smoothed Hybrid Occurrence-Size (SHOS) framework as an event-preserving feature-engineering approach for intermittent-demand forecasting. SHOS decomposes intermittent demand into a latent-occurrence probability and a conditional-demand-size component via sparsity-aware adaptive smoothing. Rather than functioning as a standalone forecasting model, SHOS serves as a representation-generation framework whose outputs are directly incorporated into supervised learning pipelines.

The proposed framework is evaluated using a large-scale automotive aftermarket dataset containing approximately 56,000 dealer-part time series and approximately 1.4 million monthly observations after preprocessing and zero-padding. Forecasting models are evaluated using rolling-window cross-validation together with signal-preservation and robustness analyses under highly intermittent demand conditions. In the present study, the term “event-preserving” refers to the ability of the SHOS representation framework to retain meaningful demand-event characteristics, including occurrence timing, peak magnitude behaviour, and frequency-domain structure, while reducing instability arising from extreme sparsity.

The results demonstrate that SHOS-derived representations substantially improve forecasting accuracy and stability under the evaluated sparse-demand setting while preserving event timing and demand-structure characteristics that are often distorted by conventional smoothing approaches. Rather than proposing a universal replacement for complex architectures, the study investigates the role of statistically grounded feature representation within the evaluated automotive aftermarket forecasting environment.

2. Literature review

2.1 Classical intermittent-demand forecasting

Intermittent-demand forecasting methods are fundamentally based on the principle that demand occurrence and demand magnitude should be modelled separately. Croston's method introduced this decomposition by independently smoothing inter-demand intervals and positive demand sizes [4]. Subsequent studies demonstrated that Croston's estimator exhibits positive statistical bias, motivating several refinements [13].

The Syntetos-Boylan Approximation (SBA) introduced a debiasing correction that improved forecast accuracy and inventory performance under intermittent demand [5,14]. Later, the Teunter-Syntetos-Babai (TSB) method replaced interval smoothing with direct smoothing of demand occurrence probability, enabling improved responsiveness to changes in demand frequency [6,15]. These approaches remain foundational because they

explicitly preserve the occurrence-size structure characteristic of intermittent demand.

Despite their interpretability and robustness, classical methods remain limited by their primarily univariate formulation. They cannot naturally incorporate rich auxiliary information, such as product attributes, dealer behaviour, or cross-sectional demand patterns, which are increasingly available in modern supply-chain systems.

2.2 Machine learning and deep learning approaches

Machine-learning and deep-learning methods have increasingly been adopted for forecasting due to their ability to learn nonlinear relationships from large-scale multivariate data [7-10]. Tree-based ensemble learners such as LightGBM and XGBoost have demonstrated strong predictive performance in structured forecasting environments [7,8], while neural network architectures aim to capture complex temporal dynamics [16].

However, highly intermittent demand presents substantial challenges for purely data-driven learning approaches. Long sequences of zero demand reduce signal availability, increase noise sensitivity, and may destabilize sequential learning behaviour [9,10]. As a result, machine-learning performance under extreme sparsity often depends strongly on the quality of the underlying feature representation rather than solely on architectural complexity.

2.3 Hybrid forecasting and representation-learning gap

Recent forecasting research increasingly emphasizes hybridization strategies that combine statistical forecasting structure with machine-learning flexibility [11,12,17,18]. Existing hybrid intermittent-demand approaches generally integrate statistical estimators with machine-learning or deep-learning models through multi-stage, ensemble, or hurdle-based architectures.

Examples include Croston-LSTM hybrids, SBA-based feature integration within tree-based learners, and two-stage occurrence-size forecasting frameworks. While these methods often improve forecasting performance, they typically rely on architectural decomposition or model stacking to capture the structure of intermittent demand.

More recently, feature-based forecasting approaches have explored embedding statistical information directly into machine-learning pipelines as informative representations rather than treating it as a separate forecasting stage [18]. Nevertheless, there remains limited research on adaptive, statistically grounded feature representations specifically designed for highly sparse intermittent-demand environments.

The SHOS framework proposed in this study addresses this gap by generating adaptive occurrence-size representations that are incorporated directly into supervised learning models as input features. Unlike multi-stage hybrid architectures, SHOS performs hybridization at the representation

level while preserving computational simplicity and forecasting scalability under sparse-demand conditions.

3. SHOS Framework and Theoretical Formulation

3.1 Problem formulation

The foundational work on intermittent demand forecasting [4,6], a time series of demand $\{y_t\}_{t=1}^T$ is modelled as the product of two latent processes: an occurrence process and a size process. Let y_t denote the observed demand quantity in period t , with $O_t = I(y_t > 0)$ indicating whether a demand event occurred, and $z_t = y_t | y_t > 0$ representing the conditional size. The forecasting objective is to estimate the conditional expectation:

$$E[y_{t+1} | F_t] = P(y_{t+1} > 0 | F_t) \cdot E[y_{t+1} | y_{t+1} > 0, F_t], \quad (5)$$

where F_t is the information set available up to time t . This work focuses on generating robust, series-specific estimates of these two components and incorporating them as features into machine learning models, rather than directly optimizing a standalone forecasting equation. In this formulation, the forecasting objective is to estimate the conditional expectation of demand given the information set F_{t-1} , which includes all historical demand observations and available covariates up to time $t - 1$.

3.2 Theoretical Basis of SHOS

Intermittent demand processes are fundamentally characterized by two distinct sources of uncertainty: whether demand occurs in a given period, and

the magnitude of demand conditional on occurrence. This structure is naturally represented as a point process, where observed demand is generated by the interaction of a binary occurrence process and a conditional size process. Classical intermittent-demand models such as Croston-type estimators are built on this decomposition, reflecting the renewal nature of demand events separated by irregular intervals of inactivity. The SHOS framework adopts this same conceptual foundation, explicitly treating demand occurrence and demand size as separate latent components that evolve over time.

SHOS suppresses stochastic sparsity-induced noise without introducing substantial phase lag, peak attenuation, or distortion of intermittent event structure. Although SHOS is conceptually motivated by the occurrence-size decomposition principle introduced in Croston-type intermittent-demand methods, the SHOS framework does not directly reuse Croston forecast outputs as replacement predictions or embedded forecast estimators. Instead, SHOS independently generates adaptive latent representations of occurrence probability and conditional demand size through sparsity-aware smoothing and recency-adaptive updating mechanisms operating on the zero-padded demand series.

A central challenge in intermittent demand forecasting arises from the amplification of variance under extreme sparsity. When long sequences of zero demand dominate the series, direct smoothing or regression on raw demand values conflates non-occurrence with small magnitudes, leading to

unstable estimates and high noise sensitivity. SHOS addresses this issue by smoothing the occurrence process independently of the size process. By updating the probability of demand occurrence at every time step, including zero-demand periods, SHOS stabilizes frequency estimation and prevents isolated demand events from exerting disproportionate influence. Conditional demand size, in contrast, is updated only when demand occurs, preserving meaningful magnitude information without contamination from inactivity periods. This separation acts as a variance-stabilizing mechanism, particularly effective in highly sparse regimes. The SHOS Decomposition and representation are illustrated in Figure 1.

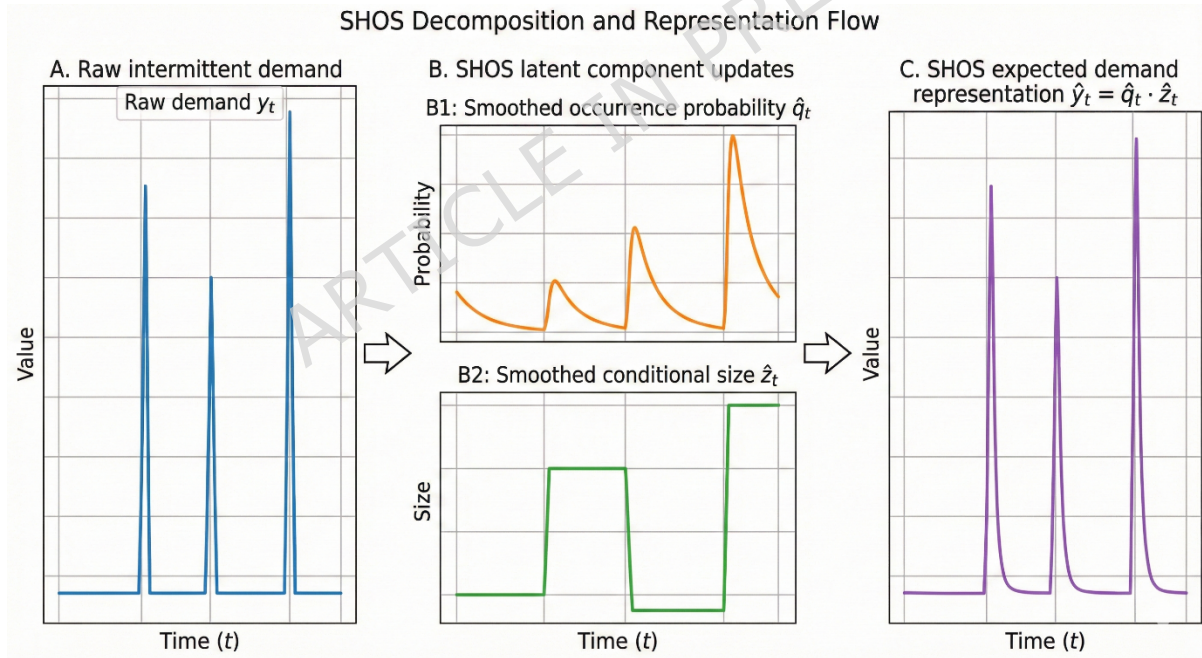


Figure 1. Illustration of the SHOS framework decomposing raw demand into latent components that preserve event timing and peak magnitude without phase lag

Importantly, SHOS does not perform smoothing via convolution or moving-average operations on the raw demand series. Instead, it applies event-preserving smoothing at the level of latent demand components. The timing of demand events is retained through the occurrence process, while magnitude information is preserved through conditional size updates. As a result, SHOS suppresses stochastic noise arising from sparsity without introducing phase lag or attenuating genuine demand bursts. This distinction explains why SHOS fundamentally differs from low-pass filtering approaches and why it can retain cyclical and burst-driven structure while improving statistical stability. These theoretical properties motivate the use of SHOS-derived occurrence and size estimates as statistically grounded input features for downstream machine learning models.

Unlike classical Croston-type estimators that explicitly model inter-demand intervals as renewal durations, the SHOS framework operates by adaptively smoothing latent occurrence probability and conditional demand size representations over a fixed temporal grid. Consequently, SHOS is less dependent on exact interval estimation and instead focuses on generating statistically stable occurrence-size representations suitable for supervised sparse-demand learning.

3.3 The SHOS algorithm: adaptive statistical priors

The SHOS (Smoothed Hybrid Occurrence-Size) algorithm is an adaptive statistical feature generation framework designed to operate robustly under extreme demand intermittency, where time series are

dominated by long sequences of zero demand punctuated by sporadic, irregular non-zero observations. The algorithm decomposes intermittent demand into two latent components estimated at each time step: the probability of demand occurrence and the conditional demand size given that demand occurs. This formulation allows SHOS to update continuously over time, including during prolonged zero-demand periods, which is essential for maintaining stable estimates in highly sparse demand environments. The overall logical flow of the SHOS algorithm is illustrated in [Figure 2](#). The SHOS algorithm explicitly leverages zero-filled observations as informative indicators of demand non-occurrence, enabling continuous updating of occurrence probability during inactivity periods and ensuring statistically consistent estimation under extreme sparsity [19].

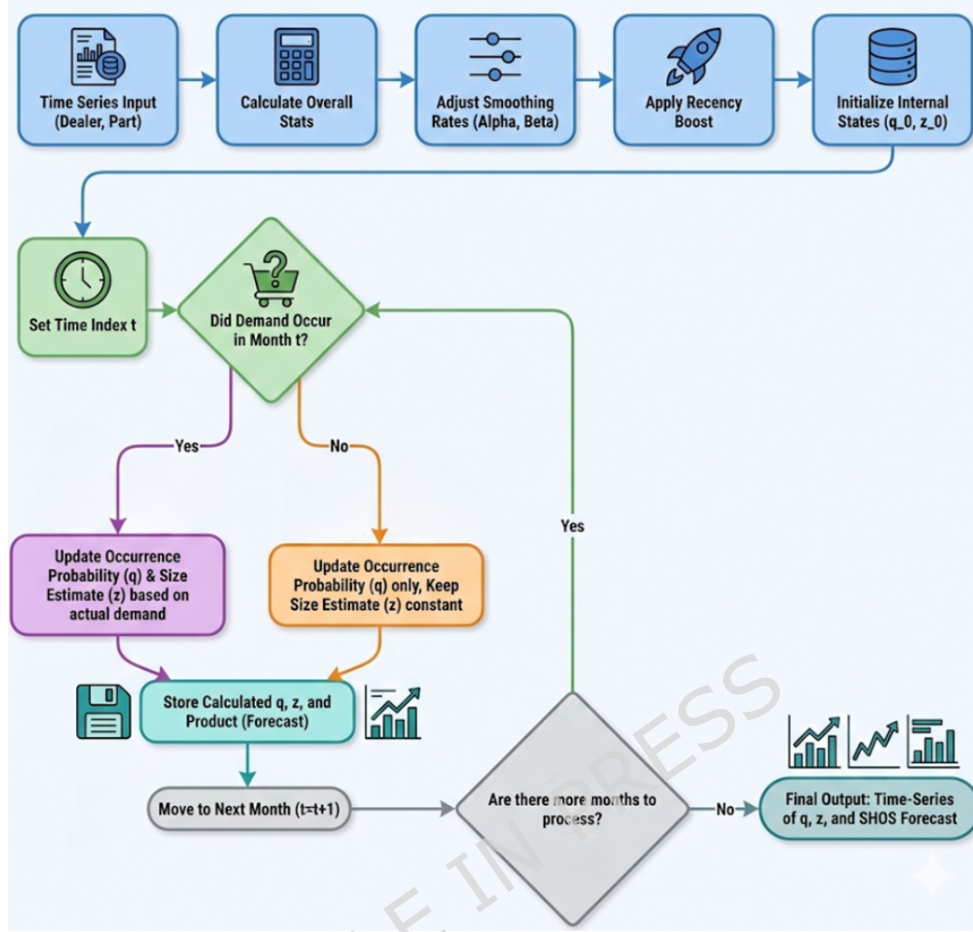


Figure 2. Algorithmic flow of the SHOS framework, illustrating adaptive smoothing, recency adjustment, and occurrence size updates

Algorithm. SHOS feature generation procedure

Input: Demand series $\{y_t\}_{t=1}^T$; base smoothing parameters $\alpha_{base}, \beta_{base} \in (0,1)$; sparsity scaling factor $k \in [0,1]$; recency window $r \in \mathbb{N}$; cap $\alpha_{max}, \beta_{max} \in (0,1]$.

Output: SHOS features $\{\hat{q}_t, \hat{z}_t, \hat{y}_t\}_{t=1}^T$.

1. **Compute demand indicator:** for all t , set $l_t \leftarrow 1(y_t > 0)$.

2. **Compute series sparsity:**

$$\text{sparsity} \leftarrow 1 - \frac{1}{T} \sum_{t=1}^T l_t. \text{Set sparsity adaptive smoothing:}$$

$\alpha \leftarrow \alpha_{\text{base}}(1 - k \cdot \text{sparsity}), \beta \leftarrow \beta_{\text{base}}(1 - k \cdot \text{sparsity})$. **Initialize** $\hat{q}_0$ **and** $\hat{z}_0$: compute from the first three periods if sufficient positive observations exist; otherwise, set $\hat{q}_0$ to the empirical occurrence rate in the initialization window and $\hat{z}_0$ to the global mean of positive demands (computed on the training data of the fold).

4. **For** $t = 1, 2, \dots, T$ **do:**

1. **Recency-aware boost (optional):** let τ_t be the number of periods since the most recent positive demand. If $\tau_t \leq r$, set

$$\alpha_t \leftarrow \min(\alpha_{\text{max}}, \alpha), \beta_t \leftarrow \min(\beta_{\text{max}}, \beta); \text{otherwise set } \alpha_t \leftarrow \alpha \text{ and } \beta_t \leftarrow \beta.$$

2. **Update occurrence probability:**

$$\hat{q}_t \leftarrow \beta_t l_{t-1} + (1 - \beta_t) \hat{q}_{t-1}, \text{with } l_0 \text{ defined from the initialization window.}$$

3. **Update conditional size:**

$$\hat{z}_t \leftarrow \begin{cases} \alpha_t y_{t-1} + (1 - \alpha_t) \hat{z}_{t-1}, & \text{if } y_{t-1} > 0, \\ \hat{z}_{t-1}, & \text{if } y_{t-1} = 0, \end{cases} \text{with } y_0 \text{ defined from the initialization}$$

window.

4. Compute SHOS point forecast:

$\hat{y}_t \leftarrow \hat{q}_t \cdot \hat{z}_t$. **Return** $\{\hat{q}_t, \hat{z}_t, \hat{y}_t\}_{t=1}^T$.

Note: In the rolling window cross-validation protocol, all statistics required for initialization (e.g., global mean of positive demands) are computed using only the training window of the corresponding fold to avoid leakage.

Let y_t denote the observed demand quantity in the time period t , and let $I_t = 1(y_t > 0)$ be a binary indicator that equals 1 when demand occurs and 0 otherwise. SHOS maintains a smoothed estimate of the probability of demand occurrence, denoted by $\hat{q}_t$, and a smoothed estimate of the conditional demand size, denoted by $\hat{z}_t$. The SHOS point forecast of expected demand is given by $\hat{y}_t = \hat{q}_t \cdot \hat{z}_t$.

3.3.1 Smoothed occurrence probability and conditional size

In highly intermittent demand series, naive estimators of occurrence probability and demand magnitude are often unstable or unresponsive due to the dominance of zero observations. SHOS mitigates this issue by explicitly updating the demand occurrence probability at every time step, including periods with zero demand. This ensures that the estimated occurrence probability decays gradually during inactivity rather than remaining fixed, thereby preventing persistent overestimation following isolated demand events. The occurrence probability is updated using exponential smoothing as

$$\hat{q}_t = \beta \cdot I_{t-1} + (1 - \beta) \cdot \hat{q}_{t-1}, \quad (6)$$

where $\beta \in (0,1)$ is the smoothing parameter controls the responsiveness of the occurrence estimate. The conditional demand size $\hat{z}_t$ is updated only when positive demand is observed, thereby avoiding contamination of magnitude estimates by zero demand periods. Specifically, the update rule is defined as

$$\hat{z}_t = \begin{cases} \alpha \cdot y_{t-1} + (1 - \alpha) \cdot \hat{z}_{t-1}, & \text{if } y_{t-1} > 0, \\ \hat{z}_{t-1}, & \text{if } y_{t-1} = 0, \end{cases} \quad (7)$$

where $\alpha \in (0,1)$ is the smoothing parameter for the conditional size process. By separating the update mechanisms for occurrence and size, SHOS ensures that magnitude estimates remain statistically meaningful even when non-zero observations are rare.

3.3.2 Adaptive smoothing, sparsity handling, and initialization

To further stabilize estimation under anomalous intermittent demand patterns, SHOS employs sparsity adaptive smoothing. Series-specific sparsity is quantified using the proportion of zero demand periods over the observed horizon T , defined as

$$\text{sparsity} = 1 - \frac{1}{T} \sum_{t=1}^T I(y_t > 0). \quad (8)$$

The effective smoothing parameters are dynamically adjusted as a function of this sparsity measure according to

$$\alpha = \alpha_{\text{base}} \cdot (1 - k \cdot \text{sparsity}), \quad \beta = \beta_{\text{base}} \cdot (1 - k \cdot \text{sparsity}), \quad (9)$$

where α_{base} and β_{base} are base smoothing values and k is a sparsity-scaling factor that controls the strength of adaptation. In this study, $\alpha_{\text{base}} = \beta_{\text{base}} = 0.2$ and $k = 0.1$, which provides stronger smoothing for highly sparse series while preserving responsiveness for less intermittent ones.

To improve robustness in short or extremely sparse series, SHOS employs a conservative initialization strategy. Initial estimates $\hat{q}_0$ and $\hat{z}_0$ are computed from the first three observed periods when sufficient positive demand exists; otherwise, they default to the global mean of positive demand across the dataset. In addition, a recency-aware adjustment mechanism temporarily increases the effective smoothing weights (subject to an upper bound of 0.9) when demand occurs after an extended period of zero demand. This allows SHOS to respond rapidly to genuine demand reactivation without destabilizing long-term estimates [20].

3.3.3 Sensitivity of SHOS to the sparsity scaling parameter k

The sparsity-scaling parameter (k) controls the extent to which SHOS adapts its smoothing behaviour as intermittency increases. To evaluate whether this parameter requires retuning across different demand regimes, a sensitivity analysis was conducted across representative SKU categories defined by demand movement intensity, intermittency, and lifecycle stage. In addition to adaptive SHOS configurations, the analysis also included the case ($k = 0$), corresponding to static SHOS without sparsity-adaptive smoothing.

Table 1. Sensitivity of SHOS forecasting performance to the sparsity-scaling parameter k across SKU categories defined by demand value, intermittency, and lifecycle stage. The case $k = 0$ corresponds to static SHOS without sparsity-adaptive smoothing

Category	k	MAE	RMSE	WMAPE (%)	MAPE (%)	R^2
Fast-Movers	0.00 Static SHOS	0.3815	2.5540	27.82	33.79	0.7565
	0.01	0.3828	2.5634	27.99	33.98	0.7553
	0.05	0.3860	2.5971	28.22	34.51	0.7488
	0.10	0.3894	2.6060	28.47	34.90	0.7471
	0.20	0.3969	2.5762	29.02	35.30	0.7528
	0.50	0.4192	2.6360	30.65	38.03	0.7412
Slow-Movers	0.00 Static SHOS	0.00064	0.0082	1.52	1.49	0.9983
	0.01	0.00059	0.0077	1.40	1.37	0.9985
	0.05	0.00061	0.0080	1.46	1.44	0.9984
	0.10	0.00052	0.0069	1.25	1.22	0.9988
	0.20	0.00042	0.0054	1.00	0.97	0.9993
	0.50	0.00008	0.0031	0.19	0.09	0.9998
Lifecycle: New	0.00 Static SHOS	0.0376	0.6910	19.18	14.92	0.7215
	0.01	0.0381	0.6777	19.54	15.11	0.7248
	0.05	0.0396	0.6474	20.30	15.45	0.7489
	0.10	0.0399	0.6526	20.46	15.68	0.7448
	0.20	0.0412	0.6898	21.14	16.58	0.7149
	0.50	0.0455	0.6657	23.36	18.06	0.7344

Lifecycle: Mature	0.00 Static SHOS	1.0248	4.4870	26.86	35.02	0.7808
	0.01	1.0268	4.5024	26.94	35.17	0.7798
	0.05	1.0294	4.5268	27.01	35.43	0.7774
	0.10	1.0328	4.5002	27.10	35.48	0.7800
	0.20	1.0461	4.5738	27.44	36.13	0.7727
	0.50	1.1090	4.5272	29.10	38.70	0.7773
Lifecycle: Declined	-	No SKU found	-	-	-	-

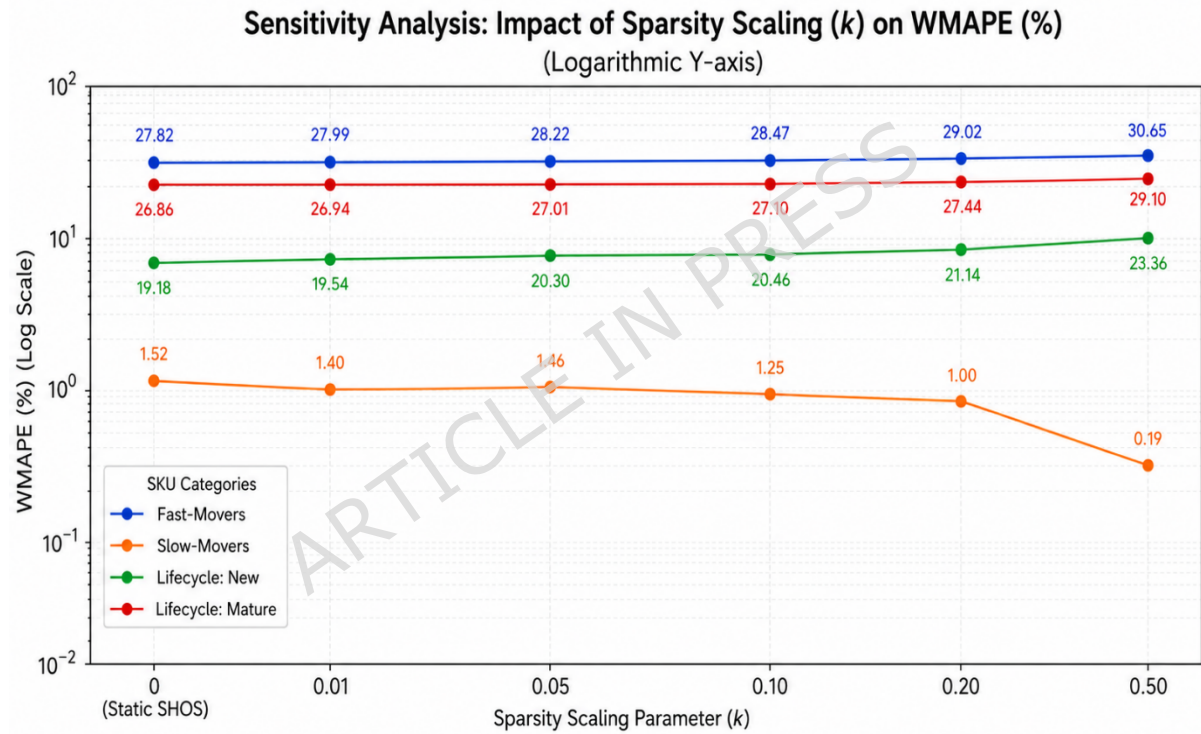


Figure 3. Sensitivity of SHOS forecasting error (WMAPE) to the sparsity scaling parameter k across SKU categories, the point $k = 0$ corresponds to static SHOS (no sparsity adaptation)

For fast-moving items, characterized by frequent demand occurrences and relatively low sparsity, forecasting performance varies gradually with

respect to the sparsity-scaling parameter. As shown in Table 1 and Figure 3, MAE and WMAPE remain relatively stable for $k \in [0, 0.2]$, with only mild degradation observed at larger values. Static SHOS ($k = 0$) already produces competitive performance for dense demand regimes, indicating that the primary forecasting gains arise from the underlying occurrence-size representation rather than from aggressive sparsity adaptation. Increasing k beyond moderate levels introduces additional smoothing, which slightly reduces responsiveness to frequent demand fluctuations [21].

For slow-moving items exhibiting extreme intermittency and prolonged zero-demand intervals, adaptive sparsity scaling provides a stabilizing effect. Increasing values of k progressively improve MAE, WMAPE, and R^2 , with the strongest performance observed at higher smoothing levels. In this regime, adaptive smoothing reduces estimation instability caused by sparse demand occurrence while preserving event-level responsiveness when demand reappears. The comparison between static SHOS and adaptive SHOS shows that sparsity-aware scaling provides incremental robustness under severe intermittency conditions, rather than serving as the sole source of forecasting improvement.

A similar trend is observed across lifecycle-based SKU segmentation. Newly introduced parts show relatively modest variation in forecasting accuracy across different values of k , remaining within a narrow error band despite changes in smoothing intensity. Mature lifecycle parts exhibit near-

flat forecasting behaviour for $k \in [0, 0.2]$, followed by gradual degradation at larger values. These observations indicate that SHOS generalizes consistently across different lifecycle stages without requiring aggressive lifecycle-specific parameter tuning [22]. No SKUs satisfied the minimum activity criteria for the declined lifecycle category during the evaluation window, so this segment was excluded from the sensitivity analysis.

Overall, the sensitivity analysis demonstrates that SHOS performance varies smoothly rather than abruptly with respect to the sparsity-scaling parameter across heterogeneous SKU groups. The results additionally indicate that adaptive sparsity scaling functions primarily as an incremental stabilization mechanism, while the dominant forecasting gains arise from the combined occurrence-size decomposition and event-preserving feature representation framework. Based on the balance between stability under sparse demand conditions and responsiveness under dense demand regimes, a fixed value of ($k = 0.1$) was adopted throughout this study as a stable and operationally practical default configuration.

3.3.4 SHOS versus TSB: bias mitigation and adaptivity

Although SHOS and the Teunter-Syntetos-Babai (TSB) method share a common conceptual foundation by modelling demand as the product of an occurrence probability and a conditional size, their estimation mechanisms differ fundamentally. In TSB, both components are updated using fixed exponential smoothing parameters, implicitly assuming homogeneous

intermittency and stationarity across all series. While effective under moderate intermittency, this assumption can lead to delayed responsiveness and variance-induced bias in highly sparse or non-stationary demand environments.

SHOS addresses these limitations by introducing sparsity-adaptive smoothing, in which the degree of smoothing is dynamically scaled to series-specific intermittency. This reduces overreaction to isolated demand events in highly sparse series while maintaining responsiveness in less intermittent ones. Furthermore, the recency-sensitive adjustment mechanism enables SHOS to adapt more rapidly than TSB when demand reappears after prolonged inactivity, mitigating the delayed-response bias commonly observed in fixed-parameter smoothing approaches [23].

Finally, the robust initialization strategy employed by SHOS stabilizes early estimates in short or highly intermittent series, further reducing estimation bias relative to classical methods. Collectively, these mechanisms enable SHOS to generate stable, informative, and adaptive estimates of demand occurrence probability and conditional size, making it particularly well suited for feature generation in anomalous intermittent demand scenarios.

4. Experimental study design

This section describes the methodological framework for studying event-preserving feature representation for intermittent demand forecasting.

Rather than emphasizing model architecture, the experimental design is structured to isolate the effect of feature representation on learning behaviour under sparsity.

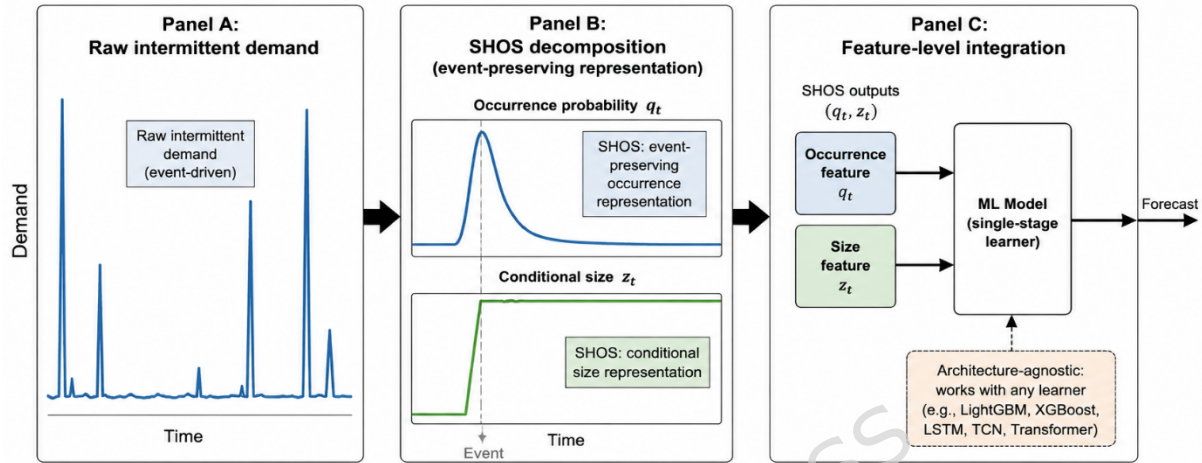


Figure 4. SHOS feature representation logic (event-preserving decomposition and integration)

Figure 4 provides a conceptual overview of the proposed SHOS feature representation. Starting from raw intermittent demand, the figure illustrates how conventional smoothing distorts demand events through averaging and phase lag, whereas SHOS instead decomposes demand into a smoothed occurrence probability and a conditional demand size, preserving event structure. These SHOS-derived signals are subsequently used as input features within a single-stage, architecture-agnostic supervised learning model.

The experimental workflow described in the following subsections operationalizes this representation. Raw transactional data are aggregated

to a monthly frequency and zero-padded to establish a consistent temporal grid. A comprehensive feature engineering process then generates a rich set of predictors, including SHOS-derived statistical features. Forecasting models are trained under a rolling-window cross-validation protocol that respects temporal ordering, enabling controlled evaluation of representation effects across heterogeneous demand regimes.

4.1 Dataset description

The experiments use a large-scale automotive aftermarket intermittent-demand dataset derived from dealer-level spare-parts sales transactions collected from a single national automotive aftermarket network. The original transactional records span the period from August 2023 to August 2025 and include approximately 59 dealers and 7,702 unique spare parts. Following preprocessing and dealer-part filtering, the dataset yielded approximately 56,000 active dealer-part demand series for forecasting analysis. Because raw transactional demand is extremely sparse and irregular, the original daily sales records were aggregated to a monthly frequency prior to forecasting analysis. Monthly aggregation was selected to reduce excessive day-level fragmentation while preserving the underlying intermittency structure relevant to spare-parts demand forecasting.

To establish a consistent temporal representation across all dealer-part combinations, the monthly demand series were zero-padded over the complete aggregation horizon. This preprocessing step produced a final

forecasting dataset containing approximately 1.4 million monthly observations. All forecasting experiments, SHOS feature-generation procedures, rolling-window validation analyses, significance testing, and model comparisons reported in this study were conducted exclusively on this final processed monthly-aggregated dataset.

The zero-padding operation is statistically meaningful in the context of intermittent-demand forecasting because zero-valued observations explicitly represent periods of demand non-occurrence rather than missing information. These inactive periods contain essential occurrence-level information required for modelling renewal behaviour, demand sparsity, and event reactivation dynamics. Treating such periods as missing values would distort occurrence probabilities, alter sparsity distributions, and weaken the statistical representation of intermittent demand processes.

The present study interprets zero-valued periods as meaningful non-occurrence events rather than missing observations. This distinction is important because zero-demand intervals carry essential information regarding sparsity structure, renewal behaviour, and demand reactivation dynamics. Consequently, zero-padding was retained intentionally to preserve occurrence-level information required for SHOS-based feature generation and comparative sparse-demand learning [24].

The processed dataset preserves the extreme sparsity characteristics of the original aftermarket demand environment. Across the active dealer-part series, the median Average Demand Interval (ADI) was approximately

25, while the median squared coefficient of variation (CV^2) remained approximately 24, indicating highly intermittent and burst-driven demand behaviour. Figure 5 presents the empirical distribution of ADI and CV^2 across all evaluated dealer-part series. The forecasting target variable is the total monthly quantity sold for each dealer-part combination (PartQty). Additional covariates include static part-level attributes such as part price, weight, and volumetric cube dimensions, which were incorporated as auxiliary explanatory variables during feature generation and supervised learning.

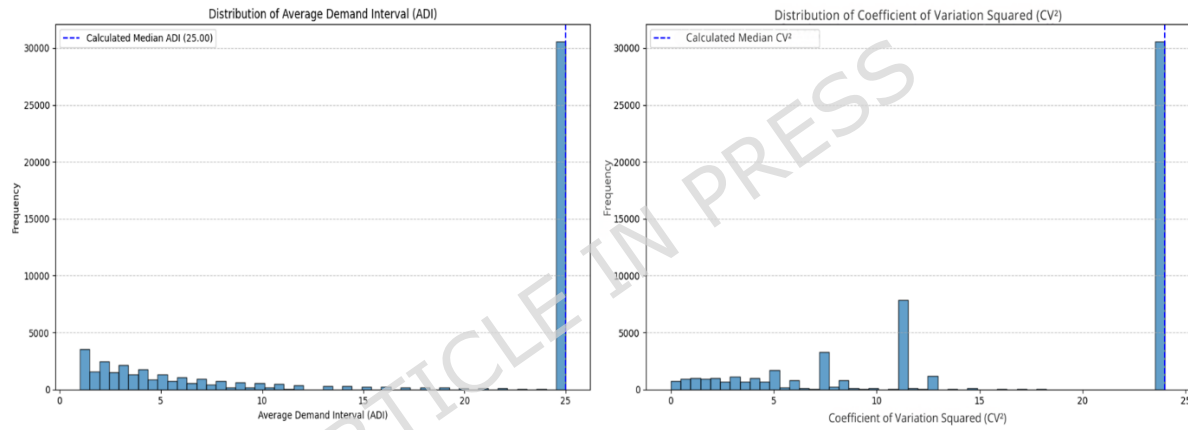


Figure 5. Empirical distribution of intermittency characteristics (ADI and CV^2) across 56,000 active dealer-part time series

This work additionally acknowledges that temporal aggregation may influence the observed renewal structure of intermittent demand. Finer-grained aggregation frequencies, such as weekly or daily representations, may preserve additional short-term event dynamics and inter-arrival variability. However, under the present large-scale dealer-part forecasting environment, such finer temporal resolutions substantially increased zero

inflation, sparsity fragmentation, and noise sensitivity during preliminary exploratory analysis. The monthly aggregation adopted in this study, therefore, reflects an operationally motivated representation scale appropriate for the evaluated automotive aftermarket setting rather than a universally optimal temporal discretization strategy.

Because the dataset originates from a single national automotive aftermarket ecosystem, the reported findings should be interpreted within the context of the evaluated market structure, dealer behaviour, and sparsity characteristics. Extension of the SHOS framework to alternative industrial sectors, geographical markets, or multi-echelon supply-chain environments remains an important direction for future validation.

Monthly aggregation was selected as a practical compromise between temporal fidelity and statistical stability under highly sparse dealer-part demand conditions. At the raw daily transactional level, a substantial proportion of dealer-part series contained isolated low-volume events separated by extended inactive periods, resulting in highly fragmented temporal sequences with weak statistical learning signals. Aggregation to monthly frequency reduced excessive day-level fragmentation while preserving the broader intermittency structure required for representation-oriented forecasting and comparative model evaluation.

The present study interprets zero-valued periods as meaningful non-occurrence events rather than missing observations. This distinction is important because zero-demand intervals carry essential information

regarding sparsity structure, renewal behaviour, and demand reactivation dynamics. No artificial imputation procedure was used to replace missing demand observations. Zero-padding was applied exclusively to represent valid inactive dealer-part periods within the aggregation horizon.

4.2 Preprocessing and feature engineering

A structured feature engineering pipeline transforms the padded monthly data into a high-dimensional feature set. The pipeline includes:

- Temporal features: Lagged values of sales quantity up to 12 months, rolling sums and means (3- and 6-month windows), count of non-zero periods in the last 6 months, and time elapsed since the last sale.
- Algorithmic features: Forecasts from Croston's method [4] and the three core outputs of the SHOS algorithm: smoothed occurrence probability (q_{shos}) smoothed conditional size (z_{shos}), and the SHOS point forecast. Crucially, SHOS features are computed using an adaptive, sparsity-aware smoothing mechanism designed to preserve demand events while stabilizing occurrence and size estimates with sparsity scaling factor $k = 0.1$, and an improved initialization that uses the first three months of data or falls back to the global mean of positive demands.
- Global and latent features: Dealer- and part-level historical averages together with low-rank latent embeddings derived from the dealer-part activity matrix were incorporated to capture cross-sectional interaction structure within the sparse-demand environment. The latent

embeddings were intended to summarize broad dealer purchasing behaviour, part-level co-occurrence structure, and recurring interaction patterns across dealer-part combinations rather than to function as direct probabilistic demand estimators. To generate these representations, Truncated SVD with 12 latent dimensions was used due to its computational efficiency and numerical stability with large sparse matrices.

To strictly preserve temporal integrity and prevent data leakage, the Truncated SVD factorization was recomputed independently within each rolling-window cross-validation fold. Specifically, the dealer-part activity matrix was constructed exclusively from the training subset corresponding to the current fold, and the learned low-rank projection was subsequently applied to the associated test observations. No future demand information from the test horizon was used during latent embedding construction.

The present study additionally acknowledges that alternative sparse-factorization approaches such as zero-inflated probabilistic matrix factorization, Poisson factorization, or nonnegative matrix factorization (NMF) may provide stronger probabilistic alignment with highly sparse count-based intermittent-demand data. However, the objective of the current work was to evaluate the contribution of SHOS-derived event-preserving statistical representations under a computationally scalable and controlled forecasting framework rather than to optimize latent factorization methodology itself. To isolate the effect of feature representation, baseline

ML models are trained on base features only (temporal, global, and latent), while SHOS features are reserved for ablation studies.

4.3 Forecasting models and comparative architectures

To evaluate the effectiveness of SHOS-derived feature representations under intermittent demand, a diverse set of statistical, machine-learning, hybrid, and deep-learning forecasting configurations was examined under a unified experimental framework. All models were trained and evaluated using identical preprocessing procedures, rolling-window cross-validation folds, feature-generation protocols, and evaluation metrics to ensure fair and controlled comparison across architectures. The forecasting target for all supervised learning configurations was the monthly dealer-part demand quantity (PartQty). Unless otherwise stated, all models operated on the same zero-padded monthly-aggregated dataset described in Section 4.1.

Classical intermittent-demand methods were incorporated into the study under three distinct roles. First, Croston and TSB methods were evaluated independently as standalone statistical forecasting baselines. Second, Croston- and SBA-derived signals were incorporated within selected machine-learning hybrids as feature-level integration benchmarks. Third, SHOS-derived occurrence and conditional-size representations were generated independently through adaptive smoothing procedures and evaluated separately within the supervised learning framework. This separation allows the contribution of SHOS-derived representations to be

distinguished from the influence of externally generated Croston-type statistical features.

The machine-learning (ML) models evaluated in this study include: LightGBM, XGBoost, Random Forest, Ridge Regression, and ElasticNet. In addition, deep-learning (DL) architectures were evaluated to examine the impact of sequential learning complexity under highly sparse demand conditions. These configurations include: Long Short-Term Memory (LSTM) networks, Temporal Convolutional Networks (TCN), and Transformer-based sequence models. For each architecture, two forecasting configurations were evaluated:

- Base model: trained using temporal lag variables, intermittency indicators, latent embeddings, and static dealer-part attributes.
- Model + SHOS: trained using the identical feature set augmented with SHOS-derived occurrence probability q_{shos} and conditional demand size z_{shos} representations.

The SHOS framework, therefore, operates as an independently generated representation-learning mechanism rather than as a direct extension or embedded reuse of Croston forecast outputs. This framework operates exclusively as a feature-generation mechanism and does not alter the prediction target, loss function, or forecasting objective. This design ensures that any observed performance improvement can be attributed specifically to representation-level augmentation rather than architectural modification or hybrid forecast blending [25]. To further evaluate whether

explicit architectural decomposition provides additional benefit beyond feature-level decomposition, a two-stage hurdle forecasting framework was also implemented. The hurdle configuration separates demand occurrence and conditional demand magnitude into two sequential learning stages:

- Stage 1 (Occurrence model): A LightGBM classifier estimates the probability of observing non-zero demand, $P(y_t > 0)$, using the base feature set augmented with the SHOS occurrence representation (q_{shos}).
- Stage 2 (Conditional size model): A LightGBM regressor with a Tweedie objective predicts expected demand magnitude $E(y_t | y_t > 0)$ using the same feature set augmented with the SHOS conditional size representation (z_{shos}).

The final forecast is obtained by multiplying the outputs of the occurrence and conditional-size stages, thereby generating an estimate of expected demand accounting for both event likelihood and demand magnitude. Gradient-boosted architectures (LightGBM and XGBoost) employed a Tweedie objective function to better model the zero-inflation and high skew of intermittent demand distributions.

Classical intermittent-demand methods were incorporated into the study under three complementary roles. First, the TSB framework served as the primary conceptual basis for SHOS because both approaches decompose intermittent demand into components of occurrence probability and conditional demand size. Second, standalone Croston and TSB

implementations were evaluated directly as statistical forecasting baselines within the same rolling-window evaluation framework used for all supervised learning models. Third, hybrid configurations such as SBA-LightGBM, Croston-LightGBM, and related variants were evaluated as feature-level integration strategies within common machine-learning pipelines. These comparisons were intended to assess the relative effectiveness of statistical feature representations rather than to compare fundamentally different forecasting paradigms.

All preprocessing steps, latent embedding construction procedures, feature-generation pipelines, train-test splits, and evaluation folds were kept identical across all forecasting architectures unless explicitly stated otherwise. This controlled design ensures that observed performance differences arise from representation quality and model structure rather than inconsistencies in preprocessing, optimization, or evaluation methodology.

4.4 Hyperparameter configuration

All forecasting models use fixed hyperparameters determined through a preliminary hyperparameter optimization stage conducted prior to final evaluation. Hyperparameter selection was performed using rolling time-series cross-validation on a dedicated validation subset to preserve temporal ordering and avoid look-ahead bias. For each forecasting architecture, a grid-based search was conducted over the principal learning parameters, including tree depth, learning rate, regularization strength, number of

estimators, and sampling ratios where applicable. Mean Absolute Error (MAE) on the validation folds was used as the primary optimization criterion because of its robustness under highly intermittent and skewed demand distributions.

After identifying the best-performing configuration for each architecture, the selected hyperparameters were fixed and consistently applied across all rolling-window evaluation folds. This design ensures fair comparison between forecasting models and isolates the effect of feature representation from optimization variability. For the SHOS framework, the sparsity-scaling parameter was fixed at ($k = 0.1$) based on preliminary sensitivity analysis across representative demand categories. The recency-aware update mechanism used a recency window of three consecutive inactive periods prior to adaptive boosting, while smoothing coefficients were capped at 0.9 to maintain numerical stability under extreme sparsity [26].

Deep-learning architectures were additionally stabilized using fixed training configurations including: Adam optimizer, early stopping, dropout regularization, and fixed sequence-window construction across all folds. The final hyperparameter settings used throughout the experimental study are summarized in Table 2.

Table 2. Hyperparameter settings for evaluated forecasting models

Model	Hyperparameter	Value
Ridge Regression	alpha	0.01
XGBoost	N estimators	500

	Max depth	9
	Learning rate	0.05
	subsample	0.8
LightGBM	N estimators	500
	Max depth	9
	Learning rate	0.05
	Num leaves	31
Hurdle Model	base regressor	LightGBM
SHOS	Alpha base	0.2
	Beta base	0.2
	sparsity scaling (k)	0.1
	recency window	3 periods
	smoothing cap	0.9
	initialization window	first 3 periods

Deep-learning architectures were trained using a fixed sequence-window construction, with the same rolling-window validation framework applied to all machine-learning models. Model optimization employed the Adam optimizer with learning-rate stabilization, mini-batch training, dropout regularization, and early stopping to reduce overfitting under highly sparse demand conditions. Hyperparameters including hidden dimensions, dropout rates, learning rates, batch size, and epoch limits were selected during preliminary rolling-validation tuning and subsequently fixed across all evaluation folds to ensure fair comparison between architectures.

All forecasting models were trained on identical preprocessing pipelines, rolling-window folds, and evaluation windows unless explicitly stated otherwise. This controlled experimental design ensures that reported performance differences arise primarily from feature representation and model structure rather than inconsistencies in preprocessing or optimization procedures.

4.4.1 Deep-learning configuration summary

The evaluated deep-learning architectures were implemented under a standardized training and evaluation framework to ensure controlled comparison across all forecasting configurations. All models were trained using identical preprocessing pipelines, rolling-window validation folds, feature-generation procedures, and train-test splits in order to isolate the effect of representation quality from differences in optimization or data preparation.

The LSTM architecture employed a single recurrent hidden layer with dropout regularization and fixed sequence-window input construction to model temporal demand dependencies under sparse-demand conditions. The TCN configuration utilized causal temporal convolution blocks with residual connections and fixed dilation scheduling to capture localized temporal structure while preserving sequential ordering. The Transformer architecture employed a multi-head self-attention encoder with positional encoding and feed-forward projection layers to evaluate attention-based sequence learning behaviour in highly intermittent-demand environments.

All deep-learning models were optimized using the Adam optimizer together with mini-batch gradient updates, learning-rate stabilization, dropout regularization, and early stopping to improve optimization stability and reduce overfitting. Hyperparameters including hidden dimensions, dropout rates, learning rates, batch size, sequence-window length, and epoch

limits were selected during preliminary rolling-validation tuning and subsequently fixed across all evaluation folds to ensure reproducibility and fair comparison between architectures [27].

Despite these stabilization strategies, the evaluated deep-learning configurations remained sensitive to the extreme sparsity and prolonged inactive demand periods characteristic of the present automotive aftermarket forecasting environment. Consequently, the deep-learning results reported in this study should be interpreted within the context of the evaluated dataset characteristics, sparsity regime, feature representations, and rolling-window forecasting framework.

4.5 Evaluation protocol and metrics

Model performance is evaluated using a rolling-window cross-validation protocol that preserves temporal ordering and prevents look-ahead bias. In each fold, models are trained on a contiguous historical window and evaluated on the immediately following test window, thereby closely simulating real-world deployment conditions. All data-dependent transformations, including feature scaling and the construction of low-rank latent embeddings using truncated singular value decomposition (SVD), are fitted exclusively on the training data of each fold and then applied to the corresponding test data. This ensures strict temporal integrity and avoids information leakage. Model performance is reported as the mean and

standard deviation of evaluation metrics computed across all eleven rolling validation folds.

Let y_i denote the observed demand and $\hat{y}_i$ the corresponding forecast for observation i , with N representing the total number of observations in the evaluation set and $\bar{y}$ denoting the mean of the observed demand. Forecast accuracy is assessed using multiple complementary metrics that capture both absolute and relative error characteristics relevant to intermittent demand forecasting.

The Root Mean Squared Error (RMSE) is used to quantify sensitivity to large forecast errors and is particularly important for avoiding severe underprediction that can lead to stockouts [28]. It is defined as

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}. \quad (1)$$

The Mean Absolute Error (MAE) measures the average magnitude of forecast errors and provides an intuitive measure of typical deviation between predicted and observed demand. It is computed as

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|. \quad (2)$$

The coefficient of determination, R^2 , evaluates the proportion of variance in the observed demand explained by the model and serves as a normalized measure of explanatory power. It is given by

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}. \quad (3)$$

Although R^2 provides a normalized measure of variance explanation, its interpretation under highly intermittent and zero-inflated demand distributions requires caution. In sparse-demand environments dominated by extended inactive periods, variance-based metrics may become sensitive to isolated nonzero observations and may not fully reflect event-detection behaviour or demand-reactivation responsiveness. Accordingly, R^2 is interpreted in this study only as a complementary diagnostic metric alongside absolute and weighted forecasting error measures.

To assess relative forecast accuracy in a volume-weighted manner, the Weighted Mean Absolute Percentage Error (WMAPE) is employed. Unlike conventional percentage errors, WMAPE remains well-defined for intermittent demand series with frequent zero values. It is defined as

$$\text{WMAPE} = \frac{\sum_{i=1}^N |y_i - \hat{y}_i|}{\sum_{i=1}^N |y_i|} \times 100\%. \quad (4)$$

In addition to aggregate accuracy metrics, overfitting is diagnosed by comparing training and test errors across validation folds to assess generalization behavior. Model interpretability and feature-level contributions are further examined using SHAP (SHapley Additive exPlanations) analysis, providing insight into the drivers of predictive performance.

In addition to reporting aggregate error metrics, statistical significance testing is performed to assess whether observed performance differences between competing models are meaningful. Given the paired structure of rolling-window cross-validation and the non-normal error distributions typical of intermittent demand, the Wilcoxon signed-rank test is employed to compare per-fold MAE and WMAPE values between models. All tests are conducted at a 5% significance level.

5. Results and discussion

The results presented in this section follow directly from the experimental protocol described in Section 3. All models are trained and evaluated on the same zero-padded, monthly-aggregated automotive aftermarket dataset using an identical rolling-window cross-validation framework. The dataset comprises approximately 1.4 million monthly observations derived from 56,000 dealer-part-time series. By holding data selection, preprocessing, feature construction, and evaluation protocols constant across all experiments, observed performance differences can be attributed solely to differences in feature representation and learning models.

5.1 Performance of machine learning models

This subsection presents a global comparison of all evaluated forecasting models under a unified experimental protocol. All models were trained and evaluated on the same zero-padded, monthly-aggregated

automotive aftermarket dataset using identical rolling-window cross-validation splits of 11 folds, preprocessing steps, and evaluation metrics. As a result, differences in forecasting performance can be attributed solely to differences in modeling approach rather than experimental artifacts.

Table 3. Comparative Performance of Machine Learning Models
(Mean $\pm$ Std)

Model Group	Model Architecture	MAE	RMSE	MAPE (%)	WMAPE (%)	R ²
Integrated SHOS (ML)	SHOS + LightGBM	0.0712 $\pm$ 0.0028	1.0310 $\pm$ 0.1416	17.73 $\pm$ 0.47	20.39 $\pm$ 0.57	0.8430 $\pm$ 0.0460
	SHOS + Random Forest	0.0702 $\pm$ 0.0032	0.8748 $\pm$ 0.0903	17.63 $\pm$ 0.804	20.11 $\pm$ 0.905	0.8879 $\pm$ 0.0245
	SHOS + XGBoost	0.0793 $\pm$ 0.0240	1.2222 $\pm$ 0.1378	18.54 $\pm$ 0.660	22.74 $\pm$ 0.992	0.7799 $\pm$ 0.0545
Standard ML	LightGBM	0.1326 $\pm$ 0.0046	1.2900 $\pm$ 0.1079	38.25 $\pm$ 1.20	38.00 $\pm$ 0.68	0.7578 $\pm$ 0.0418
	Random Forest	0.1364 $\pm$ 0.0046	1.3411 $\pm$ 0.1057	40.28 $\pm$ 1.54	39.08 $\pm$ 1.03	0.7377 $\pm$ 0.0484
	XGBoost	0.1413 $\pm$ 0.0045	1.3960 $\pm$ 0.1074	40.41 $\pm$ 1.37	40.49 $\pm$ 1.26	0.7156 $\pm$ 0.0494
Regression	Ridge	0.3439 $\pm$ 0.0134	1.5101 $\pm$ 0.1161	69.99 $\pm$ 0.61	98.57 $\pm$ 4.28	0.6685 $\pm$ 0.0535
	ElasticNet	0.3439 $\pm$ 0.0134	1.5101 $\pm$ 0.1160	70.00 $\pm$ 0.61	98.59 $\pm$ 4.26	0.6685 $\pm$ 0.0535
	Hurdle_Baseline	0.1341 $\pm$ 0.0050	1.2911 $\pm$ 0.0934	37.91 $\pm$ 1.29	38.41 $\pm$ 0.71	0.7577 $\pm$ 0.0379

Hybrid (Two-stage)	Hurdle_SHOS	0.0701 $\pm$ 0.0023	1.1136 $\pm$ 0.1102	16.84 $\pm$ 0.44	20.10 $\pm$ 0.47	0.8186 $\pm$ 0.0413
Statistical	TSB	0.2831	1.2905	59.94	82.2	0.7619
	Croston	1.1558	3.968	82.1	335.63	-1.2511
Deep Learning	LSTM	0.6539 $\pm$ 0.055	3.3033 $\pm$ 0.089	103.06 $\pm$ 3.86	187.14 $\pm$ 11.77	-0.5775 $\pm$ 0.100
	TCN	0.6534 $\pm$ 0.025	3.4656 $\pm$ 0.207	101.86 $\pm$ 2.92	187.24 $\pm$ 6.49	-0.7405 $\pm$ 0.211
	Transformer	0.6801 $\pm$ 0.050	3.3057 $\pm$ 0.064	101.20 $\pm$ 7.17	195.03 $\pm$ 15.81	-0.5806 $\pm$ 0.105
Integrated SHOS (DL)	SHOS + LSTM	0.6370 $\pm$ 0.025	3.4152 $\pm$ 0.151	102.79 $\pm$ 1.47	182.53 $\pm$ 6.15	-0.6851 $\pm$ 0.121
	SHOS + TCN	0.6681 $\pm$ 0.044	3.5537 $\pm$ 0.142	100.54 $\pm$ 5.12	191.46 $\pm$ 12.61	-0.8265 $\pm$ 0.146
	SHOS + Transformer	0.6660 $\pm$ 0.050	3.5936 $\pm$ 0.133	106.34 $\pm$ 4.19	190.93 $\pm$ 15.07	-0.8679 $\pm$ 0.153
Naive	Naive Zero	0.3492	2.6557	-	100	-0.0176
Note: MAPE is undefined for the Naive Zero model due to zero actual demand values, rendering percentage error metrics meaningless.						

Table 3 reports the mean performance across all validation folds for statistical baselines, standard machine learning models, SHOS augmented machine learning models, deep learning architectures, and hybrid hurdle formulations. Performance is evaluated using absolute error metrics (MAE, RMSE), relative error metrics (WMAPE, MAPE), and goodness-of-fit (R^2) [29]. Accordingly, aggregate forecasting metrics such as MAE, RMSE, WMAPE, and R^2 should be interpreted jointly under intermittent-demand conditions, because sparse-demand forecasting effectiveness depends on both

magnitude accuracy and stability under zero inflation, as well as responsiveness to irregular demand-occurrence behaviour.

Across all model groups, SHOS-augmented machine learning models consistently achieve the strongest performance. In particular, the SHOS + LightGBM model attains a mean MAE of 0.0712 ± 0.0028 , an RMSE of 1.0310 ± 0.1416 , and a WMAPE of $20.39\% \pm 0.57$. The proposed model achieves a substantial reduction in MAE compared to the baseline (LightGBM without SHOS), computed as $(0.1417 - 0.0712)/0.1417 \approx 49.75\%$, which corresponds to an approximately 50% improvement. This improvement is achieved without increasing model complexity, indicating that the gains arise from the statistical structure embedded in the SHOS feature representation rather than from architectural sophistication.

The two-stage Hurdle SHOS formulation achieves comparable absolute accuracy ($\text{MAE} = 0.0701 \pm 0.0023$), but does not consistently outperform the single-stage SHOS augmented models. Given its increased modelling complexity and training overhead, this result suggests that explicit architectural decomposition into occurrence and size is unnecessary when high-quality occurrence-size priors are already encoded at the feature level.

Relative to the baseline LightGBM configuration ($\text{MAE} = 0.1326$), the SHOS-augmented LightGBM model ($\text{MAE} = 0.0712$) achieved an approximate 46.3% reduction in mean absolute error under the evaluated rolling-window validation framework. Similarly, the weighted mean absolute percentage error decreased from 38.00% to 20.39%, corresponding to an

improvement of approximately 46.3% in weighted forecasting error. These results confirm that SHOS-derived occurrence-size representations substantially improve sparse-demand forecasting performance under the evaluated experimental conditions.

Deep learning architectures, including LSTM, TCN, and Transformer models, perform markedly worse than both SHOS-augmented and standard machine learning models, yielding negative R^2 values and extremely high WMAPE. Even when augmented with SHOS features, deep learning architectures remained less effective under the evaluated configurations, highlighting the difficulty of training high-capacity sequence models under extreme sparsity and low signal-to-noise conditions. Finally, to verify that performance gains are not artefacts of zero inflation, a Naive Zero baseline is included. This model fails to capture variance ($R^2 = -0.0176$) and exhibits the highest error across all metrics, confirming that SHOS-based models capture genuine demand signals rather than trivially exploiting sparsity.

To assess the statistical significance of the performance difference between the single-stage SHOS model and the two-stage hurdle structure, a Wilcoxon signed rank test was applied to per-fold MAE and WMAPE values across the 11 rolling validation folds. The test results indicate statistically significant differences for both metrics (MAE: $W = 0.0$, $p = 0.00098$; WMAPE: $W = 0.0$, $p = 0.00098$) are presented in [Table 4](#). This finding

confirms that the single-stage SHOS model consistently outperforms the hurdle-based formulation under the evaluated experimental setting.

Table 4. Statistical significance testing between SHOS and Hurdle models

Metric	Test	Test statistic (W)	p-value	Significance ($\alpha = 0.05$)
MAE	Wilcoxon signed-rank	0.0	0.00098	Significant
WMAPE	Wilcoxon signed-rank	0.0	0.00098	Significant

5.2 Effect of SHOS feature representation

This subsection isolates the impact of SHOS-derived feature representations on forecasting performance by comparing models trained on identical architectures with and without SHOS features. By holding the model class, hyperparameters, training protocol, and evaluation windows constant, this analysis serves as a controlled ablation study that attributes performance differences solely to the feature representation.

Across all evaluated machine learning architectures, the introduction of SHOS features leads to substantial and consistent reductions in forecasting error. Tree-based ensemble models exhibit the largest absolute gains. For example, augmenting LightGBM with SHOS features reduces MAE by approximately 46% relative to its raw-feature counterpart, alongside

comparable reductions in RMSE and WMAPE. Similar improvements are observed for XGBoost and Random Forest models, indicating that the benefits of SHOS are not model specific but arise from the statistical structure embedded in the feature representation.

Importantly, these improvements are achieved without altering the prediction target or introducing architectural decomposition. SHOS outputs are used strictly as additional input features, enabling standard supervised learners to internalize demand occurrence and conditional size dynamics without modifying the learning objective [30]. This design ensures that observed performance gains reflect enhanced representation of the underlying demand process rather than changes in model formulation. The observed forecasting improvements associated with SHOS-augmented models cannot be attributed solely to indirect interaction effects with Croston-type statistical signals because SHOS-based configurations consistently outperform both:

- (i) raw-feature forecasting baselines, and
- (ii) Croston/SBA-integrated hybrid feature configurations evaluated under the same rolling-window validation framework.

The relative magnitude of performance improvement is particularly pronounced for metrics sensitive to sparsity and scale, such as WMAPE. This suggests that SHOS features improve not only point accuracy but also the stability of forecasts across heterogeneous demand volumes, a critical consideration for large-scale supply chain applications.

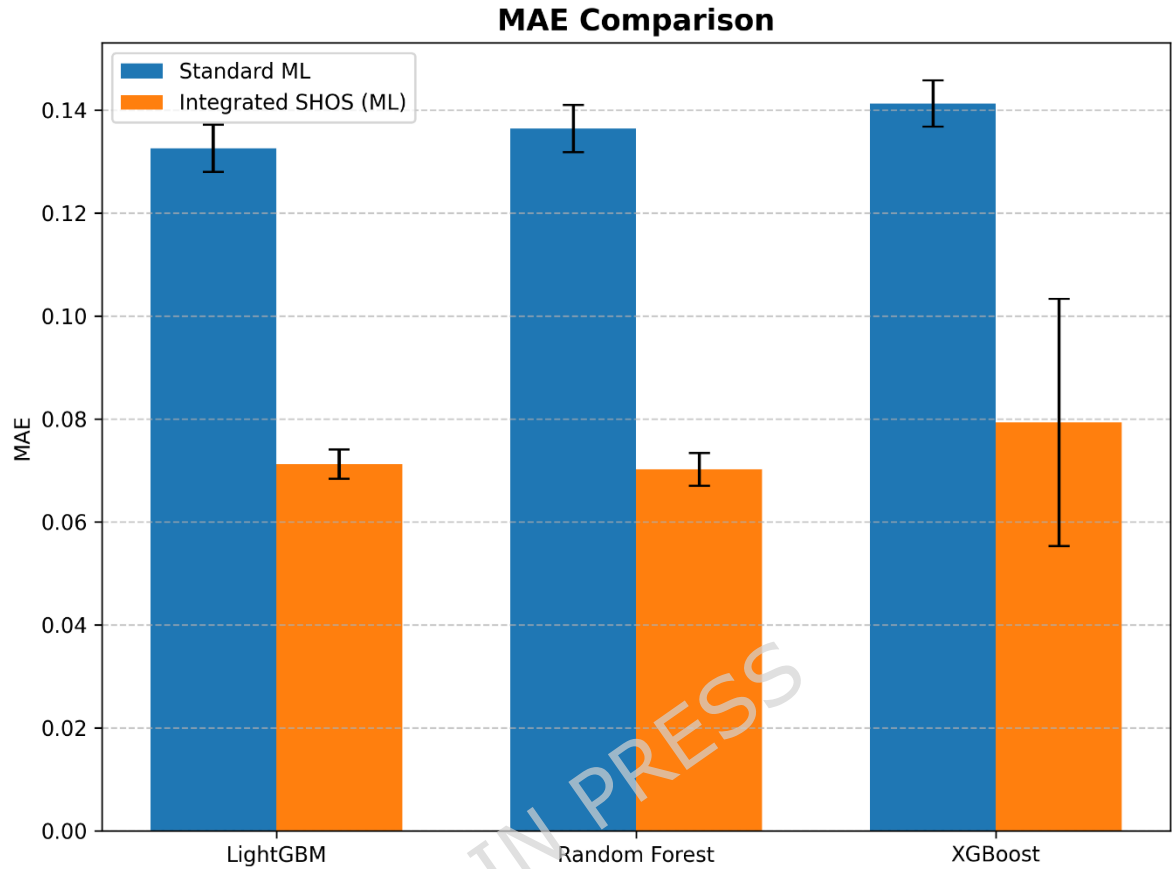


Figure 6. Percentage reduction in MAE obtained by augmenting standard machine learning models with SHOS-derived features

Figure 6 illustrates the percentage reduction in MAE obtained by introducing SHOS features across representative machine learning models. The consistent downward shift across all architectures confirms that SHOS acts as a general-purpose representation enhancement rather than a model-specific tuning mechanism.

Taken together, these results demonstrate that statistically grounded, event-preserving feature engineering can yield larger and more reliable gains than architectural modification alone. By embedding adaptive estimates of

demand occurrence probability and conditional size directly into the input space, SHOS enables standard machine learning models to learn from highly intermittent data more effectively, without requiring complex multi-stage structures or specialized loss functions.

5.3 Feature versus architecture decomposition

While Section 4.2 established that SHOS-derived features substantially improve forecasting accuracy, this subsection examines whether these gains are attributable to feature representation or to model architectural complexity. To this end, we explicitly compare models across different architectural classes, linear models, tree-based ensembles, and deep learning architectures under both raw feature and SHOS augmented settings.

The results reveal a clear and consistent pattern; improvements in forecasting accuracy align primarily with the quality of feature representation rather than with increasing architectural sophistication. Models trained on raw features show limited or negative returns from added complexity. [31] In particular, deep learning architectures such as LSTM, TCN, and Transformer exhibit significantly worse performance than simpler tree-based models when trained directly on sparse intermittent demand data, despite their higher expressive capacity.

In contrast, introducing SHOS features leads to substantial performance gains across all architectural classes. Notably, SHOS-augmented tree-based models outperform both raw-feature deep learning

models and SHOS-augmented deep learning models. For example, SHOS + LightGBM and SHOS + Random Forest achieve the lowest error across all evaluated metrics, while deep learning models, even when provided with SHOS features, remain substantially less accurate. This indicates that although SHOS improves learning across architectures, complex sequence models struggle to fully exploit sparse, event-driven signals relative to simpler learners.

Table 5. Decomposition of forecasting performance by feature representation and model architecture class (raw features vs SHOS features)

Model Group	Model Architecture	MAE	RMSE	MAPE (%)	WMAPE (%)	R ²
Integrated SHOS (ML)	SHOS + LightGBM	0.0712 ± 0.0028	1.0310 ± 0.1416	17.73 ± 0.47	20.39 ± 0.57	0.8430 ± 0.0460
	SHOS + Random Forest	0.0702 ± <i>0.0032</i>	0.8748 ± <i>0.0903</i>	17.63 ± <i>0.804</i>	20.11 ± <i>0.905</i>	0.8879 ± <i>0.0245</i>
	SHOS + XGBoost	0.0793 ± <i>0.0240</i>	1.2222 ± <i>0.1378</i>	18.54 ± <i>0.660</i>	22.74 ± <i>0.992</i>	0.7799 ± <i>0.0545</i>
Standard ML	LightGBM	0.1326 ± 0.0046	1.2900 ± 0.1079	38.25 ± 1.20	38.00 ± 0.68	0.7578 ± 0.0418
	Random Forest	0.1364 ± 0.0046	1.3411 ± 0.1057	40.28 ± 1.54	39.08 ± 1.03	0.7377 ± 0.0484
	XGBoost	0.1413 ± 0.0045	1.3960 ± 0.1074	40.41 ± 1.37	40.49 ± 1.26	0.7156 ± 0.0494

Table 5 summarizes this decomposition by reporting representative performance metrics grouped by architectural class and feature representation. The table highlights that switching from raw features to SHOS features yields a larger reduction in error than switching from a simple architecture to a more complex one. In several cases, a simple SHOS-augmented model outperforms a significantly more complex raw-feature model by a wide margin [32].

These findings demonstrate that, under extreme demand intermittency, architectural complexity alone is insufficient to compensate for weak or noisy representations. Instead, performance gains arise from embedding statistically grounded structure into the input space. By explicitly encoding demand occurrence and conditional size dynamics, SHOS enables even relatively simple models to outperform more complex architectures trained on raw demand data. This decomposition provides strong empirical support for the central thesis of this study in intermittent demand forecasting: feature representation exerts a substantial influence in the evaluated sparse-demand setting.

5.4 Hybrid dominance analysis

To avoid ambiguity between standalone forecasting baselines and feature-level hybridization strategies, the hybrid comparisons presented in this subsection were evaluated within a unified supervised-learning framework. Specifically, classical intermittent-demand estimators such as

Croston and SBA were not treated as complete forecasting systems in these experiments, but rather as statistical signal generators embedded within identical downstream machine-learning pipelines. This design enabled controlled comparison of representation quality while holding the learning architecture constant.

Hybrid forecasting methods are frequently proposed for intermittent demand under the assumption that combining statistical estimators with machine learning or deep learning architectures improves performance relative to purely data-driven models. To examine whether the observed gains of SHOS arise from hybridization in general or from the specific design of SHOS features, we benchmark SHOS-based models against established statistical machine learning hybrid formulations reported in the intermittent demand literature. These include Croston-based and Syntetos-Boylan Approximation (SBA) based hybrids combined with tree-based learners and deep learning models.

Table 6. Performance comparison of classical statistical feature-signal integrations within common machine-learning forecasting pipelines for intermittent-demand forecasting

Model Architecture	MAE	RMSE	MAPE (%)	WMAPE (%)	R²
SBA-LightGBM	0.0848 ± 0.006	1.3685 ± 0.091	29.09 ± 2.11	25.92 ± 1.15	0.7686 ± 0.038
SBA-XGBoost	0.0891 ± 0.009	1.3902 ± 0.102	32.24 ± 2.33	27.16 ± 1.25	0.7582 ± 0.048

Croston-LightGBM	0.1504 ± 0.012	1.4725 ± 0.115	44.46 ± 2.44	43.66 ± 2.15	0.6900 ± 0.042
Croston-XGBoost	0.1581 ± 0.014	1.4974 ± 0.122	47.60 ± 2.61	45.90 ± 2.31	0.6794 ± 0.048
Croston-LSTM	0.6319 ± 0.045	3.1514 ± 0.285	105.46 ± 5.50	183.47 ± 15.4	-0.4200 ± 0.105
SBA-LSTM	0.6979 ± 0.051	3.9382 ± 0.312	106.96 ± 6.12	202.65 ± 18.2	-1.2175 ± 0.146

Table 6 demonstrates that incorporating classical intermittent-demand estimators into machine learning models yields heterogeneous and often limited performance gains. SBA-based hybrids consistently outperform Croston-based hybrids, reflecting SBA's debiasing advantage under intermittency. However, even the best SBA-based hybrids remain substantially inferior to SHOS-based models reported in Section 4.1 across all error metrics. Croston-based hybrids exhibit particularly poor performance, with elevated absolute and percentage errors and reduced explanatory power. This degradation reflects the sensitivity of fixed-parameter smoothing approaches to extreme sparsity, where isolated demand events exert disproportionate influence on the estimator.

Notably, extending these classical hybrids to deep learning architectures does not recover performance. Both Croston-LSTM and SBA-LSTM exhibit severe degradation, including negative R^2 values and inflated error measures. These results indicate that simply injecting classical statistical estimates into complex architectures does not resolve the fundamental challenges of intermittent demand and may amplify instability when demand signals are sparse and irregular [33].

In contrast, SHOS-based hybrids consistently dominate all alternative statistical-machine learning combinations across MAE, RMSE, WMAPE, and R^2 . This dominance confirms that hybrid performance is not driven by the inclusion of *any* statistical signal, but by the quality, adaptivity, and event-preserving nature of the statistical representation. By embedding series-specific occurrence and size priors at the feature-representation level rather than stacking or averaging model outputs, SHOS provides a statistically coherent foundation for learning under extreme intermittency.

5.5 Deep learning performance analysis

Table 7 summarizes the forecasting performance of recurrent (LSTM), convolutional (TCN), and attention-based (Transformer) architectures evaluated with both raw input features and SHOS-derived feature augmentation. Across all evaluated configurations, deep learning models exhibited substantially higher forecasting error than SHOS-augmented tree-based machine learning models. All deep learning variants produced negative R^2 values and elevated WMAPE, indicating limited ability to model highly sparse intermittent demand under the present experimental setting.

Table 7. Forecasting performance of deep learning models with and without SHOS feature augmentation

Model	MAE	RMSE	MAPE (%)	WMAPE (%)	R^2
LSTM	0.6539 ± 0.055	3.3033 ± 0.089	103.06 ± 3.86	187.14 $\pm$ 11.77	- 0.5775

					$\pm$ 0.100
TCN	0.6534 $\pm$ 0.025	3.4656 $\pm$ 0.207	101.86 $\pm$ 2.92	187.24 $\pm$ 6.49	- 0.7405 $\pm$ 0.211
Transformer	0.6801 $\pm$ 0.050	3.3057 $\pm$ 0.064	101.20 $\pm$ 7.17	195.03 $\pm$ 15.81	- 0.5806 $\pm$ 0.105
SHOS + LSTM	0.6370 $\pm$ 0.025	3.4152 $\pm$ 0.151	102.79 $\pm$ 1.47	182.53 $\pm$ 6.15	- 0.6851 $\pm$ 0.121
SHOS + TCN	0.6681 $\pm$ 0.044	3.5537 $\pm$ 0.142	100.54 $\pm$ 5.12	191.46 $\pm$ 12.61	- 0.8265 $\pm$ 0.146
SHOS + Transformer	0.6660 $\pm$ 0.050	3.5936 $\pm$ 0.133	106.34 $\pm$ 4.19	190.93 $\pm$ 15.07	- 0.8679 $\pm$ 0.153

Negative R^2 values observed for several deep-learning configurations primarily reflect the difficulty of stable variance explanation under highly sparse and zero-inflated demand environments rather than implying complete predictive failure. Under such conditions, variance-based metrics become highly sensitive to isolated nonzero demand events and should therefore be interpreted jointly with MAE, RMSE, and WMAPE.

Introducing SHOS-derived features produced modest improvements in selected absolute error metrics, particularly for the LSTM configuration. However, these improvements remained substantially smaller than those observed for SHOS-augmented tree-based models and did not alter the overall model ranking. The results, therefore suggest that informative statistical representations alone are insufficient to fully overcome the

optimization and representation challenges associated with extreme demand sparsity in deep sequential architectures.

This behaviour is consistent with the structural characteristics of intermittent demand, in which long inactive periods and irregular event occurrences reduce the availability of stable temporal learning signals. In contrast, tree-based learners operate directly on event-level feature representations and are therefore better suited to exploit sparse occurrence and conditional-size information encoded in SHOS features [34]. Figure 7 further illustrates that forecasting performance clusters more strongly by feature representation quality than by architectural depth. SHOS-augmented tree-based models consistently occupy the lowest-error region, while deep learning models remain comparatively unstable despite feature augmentation.

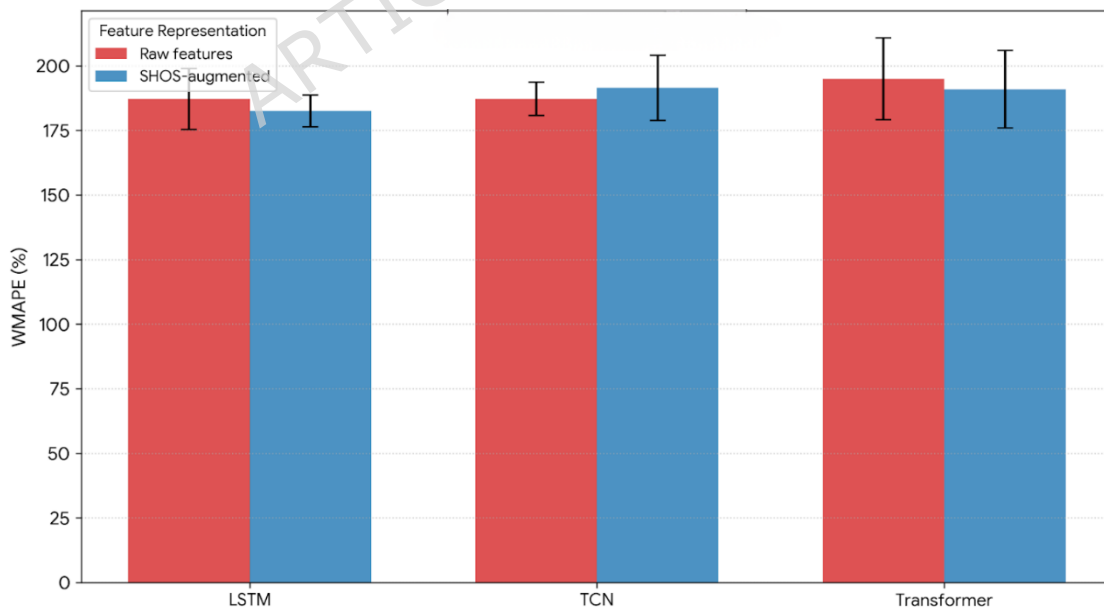


Figure 7. Comparison of forecasting error across deep learning and SHOS-augmented models

Overall, the results indicate that the evaluated deep-learning architectures, under the present dataset characteristics, sparsity regime, feature representations, and rolling-window forecasting framework, remained less effective than representation-oriented machine-learning approaches for highly intermittent demand forecasting. These findings should therefore be interpreted as experimental observations rather than universal conclusions regarding the suitability of deep learning for intermittent-demand forecasting.

5.6 Volatility retention and event-preserving behaviour

To assess whether SHOS provides an event-preserving representation rather than conventional low-pass smoothing, we evaluate its behaviour using time-domain, frequency-domain, and structural metrics that explicitly capture phase distortion, peak attenuation, and temporal dependence. These analyses are designed to determine whether SHOS suppresses informative volatility or preserves the timing and structure of demand events under intermittency.

In intermittent-demand forecasting, preserving the temporal structure of demand events is operationally important because replenishment and inventory decisions depend not only on aggregate forecast accuracy, but also on timely recognition of demand reactivation and event intensity. Excessive smoothing or temporal phase distortion can delay detection of genuine

demand resurgence following prolonged inactivity, potentially increasing service-level risk and stockout exposure. Accordingly, the analyses presented in this subsection were intended to evaluate whether SHOS preserves informative event structure while reducing stochastic sparsity-induced noise.

Table 8. Temporal and structural signal preservation metrics for simple moving average (SMA), SHOS, and SHOS+LGBM across cyclical and intermittent demand

Demand Category	Metric	SMA (Low-Pass)	SHOS (Adaptive)	SHOS+LGBM (Hybrid)
Cyclical	Phase Lag (Months)	0.005	0	0
	Peak Preservation	0.3333	0.1408	0.1395
	FFT Similarity	0.4029	0.1116	0.1093
	Amplitude	0.1859	0.0559	0.0544
Intermittent	Phase Lag (Months)	0.205	0	0
	Peak Preservation	0.3333	0.2408	0.2408
	FFT Similarity	0.4015	0.1304	0.1304
	Amplitude	0.1271	0.0869	0.0869

Phase lag measures temporal displacement between smoothed and raw demand signals, peak preservation quantifies the fraction of original peak magnitude retained, and FFT similarity measures spectral alignment between signals in the frequency domain. Table 8 evaluates temporal and structural preservation properties across smoothing approaches. SMA exhibits a non-zero phase lag and substantial peak attenuation, consistent with low-pass filtering behaviour. In contrast, SHOS introduces zero phase lag for both cyclical and intermittent demand, indicating that the timing of

demand events is fully preserved. Moreover, SHOS achieves higher peak preservation and greater similarity to the raw signal than SMA, demonstrating that noise reduction is achieved without suppressing informative demand bursts. Because SHOS determines the temporal representation, the SHOS+LGBM pipeline inherits these event-preserving properties without introducing additional temporal distortion [35]. While temporal metrics assess event alignment and peak fidelity, frequency domain analysis provides complementary insight into how much of the original signal's spectral content is retained after smoothing.

Table 9. Frequency-domain energy and average amplitude of raw demand, SMA, SHOS, and SHOS+LGBM signals for cyclical and intermittent demand categories.

Demand Category	Metric	Raw Demand	SMA	SHOS	SHOS+LGBM
Cyclical	FFT Energy (Sum)	120.625	40.1528	8.0072	8.1189
	Avg. Amplitude	0.3549	0.1859	0.0559	0.0544
Intermittent	FFT Energy (Sum)	66.25	18.25	21.5375	21.5375
	Avg. Amplitude	0.2606	0.1271	0.0869	0.0869

FFT energy reflects the total spectral content of the signal and provides a coarse measure of volatility retention. Excessive energy loss indicates over smoothing, while uncontrolled energy retention reflects noise amplification.

Frequency-domain analysis provides complementary insight into how smoothing methods affect the structural variability of intermittent demand signals across different temporal scales. In the present context, retaining meaningful spectral structure indicates that the representation preserves recurring event behaviour and burst-driven dynamics rather than excessively flattening the signal through averaging. This property is particularly relevant in intermittent-demand environments where isolated demand events may carry disproportionate operational importance.

Table 9 reports aggregate frequency-domain energy and average amplitude for each smoothing method. SMA removes a substantial portion of spectral energy, particularly for cyclical demand, reflecting aggressive attenuation of higher-frequency components. In contrast, SHOS retains significantly more frequency domain energy while still reducing noise relative to raw demand. For intermittent demand, SHOS maintains a balanced reduction in amplitude without eliminating structurally meaningful variability. The SHOS+LGBM results closely match those of SHOS, indicating that the hybrid pipeline preserves spectral characteristics rather than reintroducing artificial smoothing.

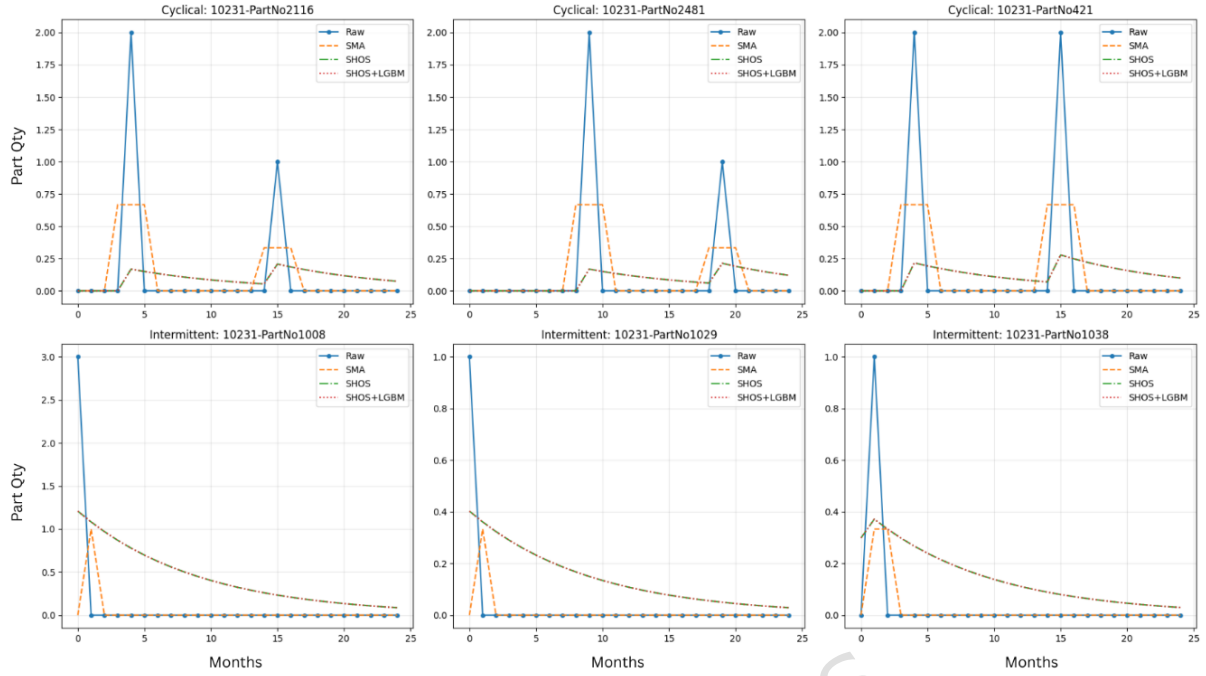


Figure 8. Time-domain comparison for representative cyclical (top row) and intermittent (bottom row) SKUs

Figure 8 illustrates the time domain behaviour for representative SKUs. SMA attenuates and delays demand peaks, introducing phase distortion and flattening event magnitude. In contrast, SHOS preserves both peak timing and magnitude while reducing stochastic noise. These properties remain unchanged when SHOS-derived features are used within the LightGBM model. For cyclical demand, SMA spreads and delays demand peaks, resulting in phase distortion and amplitude loss. SHOS preserves peak timing while reducing high-frequency noise, maintaining the underlying cyclical structure. For intermittent demand, SHOS remains responsive to isolated demand events while avoiding artificial oscillations during extended zero

demand periods. Importantly, these properties are preserved when SHOS-derived features are used within a machine learning model.

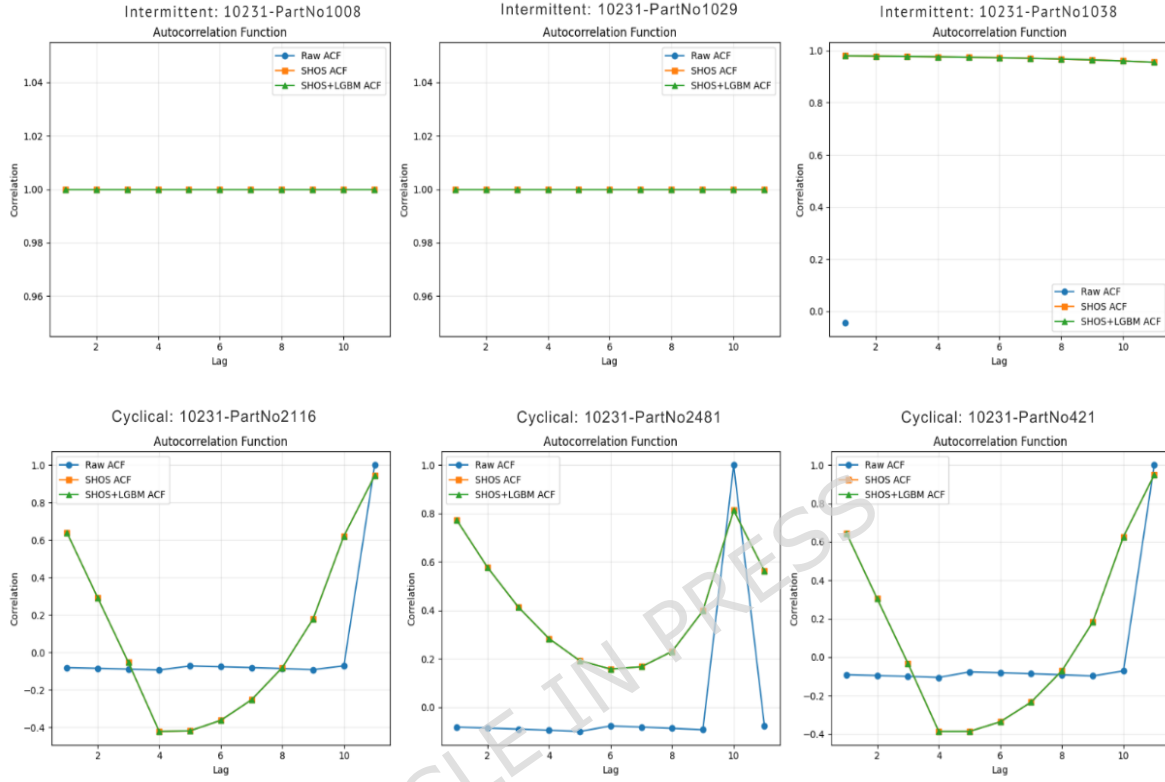


Figure 9. Autocorrelation functions for representative cyclical and intermittent SKUs

Autocorrelation analysis further confirms that SHOS does not behave as a low-pass filter. As shown in Figure 9, SHOS closely preserves the temporal dependence structure of the raw demand signal, particularly for cyclical SKUs where seasonal correlations remain intact. In contrast, conventional smoothing disrupts temporal dependence by averaging across event boundaries. The SHOS+LightGBM pipeline maintains the same

autocorrelation profile, indicating that integrating SHOS features does not introduce artificial temporal structure [36].

The signal-preservation metrics presented in this subsection should therefore be interpreted as diagnostic and mechanistic evidence explaining the behaviour of the SHOS representation under intermittency, rather than as direct proof of downstream inventory or replenishment optimization performance. While preservation of event timing and structural demand characteristics may support improved operational responsiveness, the primary validation of SHOS effectiveness in this study remains the forecasting accuracy, robustness, and sparsity-response analyses presented in the preceding sections.

Taken together, these results demonstrate that SHOS fundamentally differs from conventional low-pass smoothing. Rather than averaging demand values over time, SHOS performs event-preserving smoothing by operating on latent occurrence and conditional size components. This design enables effective noise reduction under sparsity while preserving cyclical structure, peak timing, and temporal dependence. These properties provide complementary evidence supporting why SHOS-derived features may improve learning stability and event responsiveness under intermittent demand without suppressing informative temporal structure.

Although improved responsiveness and event preservation may contribute to better replenishment awareness under intermittent demand, the analyses presented in this study were not intended to directly quantify inventory-policy

performance or operational cost reduction. Rather, these metrics were evaluated as forecasting-oriented indicators of representation quality and sparse-demand learning stability.

5.7 Extreme sparsity stress test: responsiveness and recovery under demand reactivation

Intermittent demand forecasting models are particularly vulnerable under extreme sparsity, where long sequences of zero demand are followed by isolated demand reactivation events. In such regimes, learning systems trained on raw demand inputs often exhibit delayed response and unstable adjustment when demand resumes. To evaluate whether SHOS-derived feature representations mitigate these effects, we conducted an extreme sparsity stress test focusing on two operationally relevant criteria: responsiveness, defined as the delay in reacting to demand reappearance, and recovery, defined as the time required for forecasts to stabilize following a demand event [37].

Responsiveness is measured as the number of periods required for the forecast to exhibit a statistically meaningful increase following the first non-zero demand observation, while recovery measures the number of periods required for forecast error to return within a stable tolerance band after the reactivation event.

Table 10. Extreme sparsity stress test

Sparsity Bin	Model	Responsiveness	Recovery
--------------	-------	----------------	----------

0.9+	LightGBM	0.3696	0.4693
	LightGBM + SHOS	0.059	0.0323
0.8 - 0.9	LightGBM	0.5614	1.3321
	LightGBM + SHOS	0.3096	0.7566
0.7 - 0.8	LightGBM	1.4775	2.6577
	LightGBM + SHOS	1.1597	2.5779

Table 10 reports stress test results across three sparsity bins. In the most extreme sparsity regime (sparsity ≥ 0.9), when augmented with SHOS-derived features, LightGBM exhibits a dramatic reduction in both responsiveness and recovery relative to LightGBM trained on raw demand features. Responsiveness is reduced by nearly an order of magnitude, indicating substantially faster reaction to demand reactivation, while recovery time is similarly compressed, reflecting more stable post-event adjustment. These improvements highlight SHOS's ability to preserve actionable demand signals during prolonged periods of inactivity.

In moderately sparse regimes (sparsity 0.8-0.9 and 0.7-0.8), LightGBM + SHOS continues to outperform the raw-feature baseline, though the magnitude of improvement diminishes as demand density increases. This behaviour is expected, as raw demand signals become more informative as sparsity decreases. Importantly, SHOS does not introduce lag or excessive smoothing in these regimes, indicating that its adaptive mechanism remains responsive across sparsity levels.

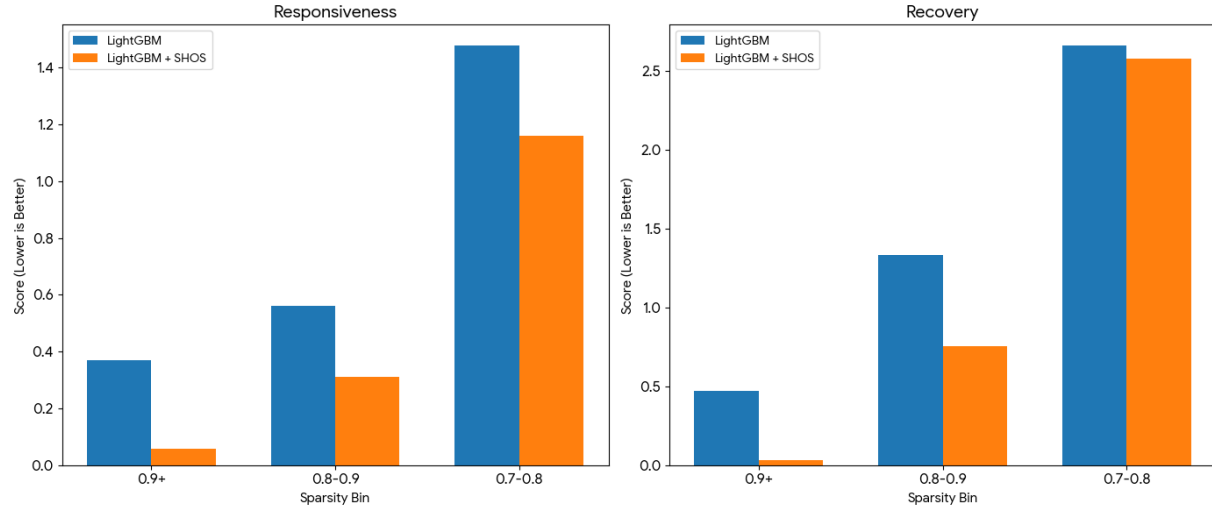


Figure 10. Bar chart comparison of responsiveness and recovery metrics

Figure 10 provides a compact visual summary of the extreme sparsity stress test. Across all sparsity bins, LightGBM + SHOS consistently achieves lower responsiveness and recovery values than the raw-feature baseline, with the largest relative gains observed under extreme sparsity. The monotonic reduction in improvement as sparsity decreases further confirms that SHOS primarily enhances learning behaviour in regimes where raw demand signals are least informative, rather than artificially inflating performance in denser demand settings.

Taken together, the extreme sparsity stress test demonstrates that SHOS-derived features substantially improve model behaviour precisely under conditions where intermittent demand forecasting is most challenging. By explicitly encoding demand occurrence probability and conditional demand size, SHOS stabilizes learning during prolonged zero-demand

periods while enabling rapid adaptation when demand resumes [38]. These findings reinforce the central claim of this study: the primary benefits of SHOS arise at the feature-representation level, improving responsiveness and recovery without increasing architectural complexity. These findings complement the volatility analysis in Section 4.6 by demonstrating that SHOS not only preserves events, but also enables faster recovery when events reoccur after prolonged inactivity. From an operational perspective, faster responsiveness directly reduces the risk of stockouts following demand reactivation, while faster recovery prevents prolonged overreaction after isolated demand events, both of which are critical in spare-parts inventory planning.

5.8 SHAP interpretability analysis

To further understand the mechanisms underlying the observed performance improvements, we analyze model interpretability using SHAP (SHapley Additive exPlanations). SHAP provides a consistent, game-theoretic framework for quantifying the marginal contribution of each input feature to the model's predictions, enabling both global and local interpretability. This analysis focuses on the best-performing SHOS augmented machine learning model and examines which features drive forecast accuracy under extreme demand intermittency.

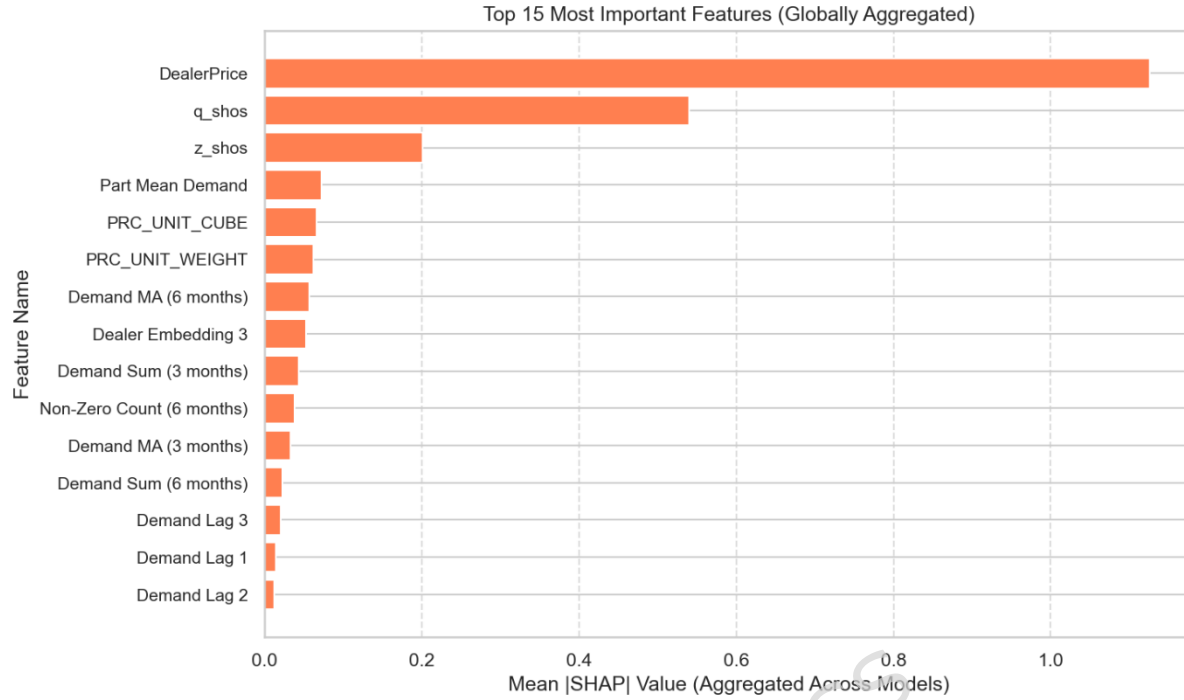


Figure 11. Global SHAP feature importance for the SHOS-augmented model, showing mean absolute contribution of each feature to forecast predictions across all validation folds

Figure 11 presents the global SHAP feature importance summary, ranked by mean absolute SHAP value across all validation folds. The results reveal that SHOS-derived features dominate the importance ranking. In particular, the smoothed occurrence probability q_{shos} and the smoothed conditional demand size z_{shos} consistently emerge as the most influential predictors, surpassing traditional lag-based features, rolling statistics, and static product attributes.

The prominence of q_{shos} indicates that the model places substantial weight on the estimated probability of demand occurrence, reflecting the

critical role of event timing in intermittent demand forecasting. Similarly, the importance of z_{shos} confirms that conditional magnitude estimation provides essential information once demand occurrence is anticipated. Together, these features closely align with the theoretical decomposition of intermittent demand into occurrence and size components, thereby validating the conceptual foundations of the SHOS framework.

In contrast, conventional temporal features such as raw lagged demand values and rolling averages exhibit lower and more dispersed SHAP contributions. This suggests that, in highly sparse settings, such features provide weaker, noisier signals than statistically smoothed occurrence-size representations. Static attributes and latent embeddings contribute additional explanatory power but play a secondary role relative to SHOS-derived signals.

Importantly, the SHAP analysis also highlights the interpretability advantage of the proposed approach. Unlike deep learning architectures, which make it difficult to trace feature contributions, the SHOS-augmented model provides transparent, domain-consistent explanations. Decision-makers can directly relate forecast outputs to intuitive quantities such as recent demand occurrence likelihood and expected conditional demand size, facilitating trust and adoption in operational settings [39].

Overall, the SHAP results confirm that the performance gains achieved by SHOS-enhanced models are not driven by spurious correlations or

overfitting, but by meaningful statistical representations that capture the fundamental structure of intermittent demand. This interpretability evidence complements the quantitative results presented in Sections 4.1-4.7 and reinforces the conclusion that SHOS provides both accuracy and explanatory value for large-scale intermittent demand forecasting.

5.9 Managerial Insights

The findings of this study have several direct implications for practitioners managing large-scale intermittent demand environments, such as automotive aftermarket, spare parts, and service logistics networks. Beyond forecasting accuracy, the proposed SHOS framework offers operational advantages in terms of robustness, interpretability, and deployability that are particularly relevant for real-world decision making.

First, SHOS enables materially improved forecast reliability for highly intermittent items without increasing system complexity. The empirical results show that SHOS-enhanced single-stage machine learning models outperform both standard ML baselines and more complex multi-stage or deep learning architectures. For practitioners, this implies that meaningful accuracy gains can be achieved without introducing additional modelling layers, custom architectures, or specialized training pipelines [40]. This is especially important in enterprise forecasting systems, where model simplicity directly translates to lower maintenance cost, faster retraining cycles, and easier integration with existing planning workflows.

Second, SHOS improves responsiveness to demand reactivation, reducing the risk of stockouts after prolonged inactivity. The extreme sparsity stress tests demonstrate that SHOS-based models recover more rapidly when demand reappears after long periods of zero demand. From an inventory management perspective, this behaviour directly addresses a common operational failure mode in intermittent-demand forecasting: delayed response to genuine demand signals. By stabilizing occurrence estimation while preserving event timing, SHOS supports earlier replenishment decisions when demand resumes, reducing service level failures without inflating inventory during inactive periods.

Third, SHOS preserves demand volatility and event structure, supporting downstream planning decisions. Unlike conventional smoothing or low-pass filtering approaches that distort peaks and introduce phase lag, SHOS retains the timing and magnitude of demand events while suppressing noise induced by sparsity. This property is particularly valuable for planners who rely on forecasts not only for point estimates, but also for understanding demand variability, planning safety stock, and evaluating risk. Preserving event-driven dynamics ensures that forecasts remain operationally meaningful rather than overly smoothed artefacts.

Fourth, the SHOS framework provides transparent and interpretable signals aligned with domain intuition. The SHAP analysis confirms that SHOS-derived features, with occurrence probability and conditional demand size as the dominant drivers, are the dominant drivers of model predictions.

These quantities correspond to intuitive managerial concepts of the likelihood of demand and the size it is expected to be if it occurs. This interpretability facilitates trust, auditability, and stakeholder acceptance, particularly in settings where planners must justify decisions to procurement, finance, or service teams.

Finally, SHOS is operationally scalable and parameter-light, supporting large-scale deployment. The k-sensitivity analysis demonstrates that SHOS performance varies smoothly across a wide range of sparsity conditions and does not require fine-grained parameter tuning across product categories or lifecycle stages. As a result, a single global parameter setting can be applied across tens of thousands of SKUs, enabling consistent and scalable deployment in large assortments without manual intervention or per-item calibration [14].

Taken together, these insights suggest that the primary managerial value of SHOS lies not only in improved forecast accuracy, but also in shifting the focus of intermittent-demand forecasting from architectural complexity to representation quality. By embedding a statistically grounded, event-preserving demand structure directly into the features, SHOS enables simpler, more robust, and more interpretable forecasting systems that closely align with the operational realities of intermittent demand planning.

5.10 Practical Implications for Inventory Management

From an operational perspective, the performance characteristics of the SHOS-augmented model have direct implications for inventory planning and control in intermittent-demand environments. Forecast errors in such settings are inherently asymmetric: underprediction increases the risk of stockouts and service level failures, while overprediction leads to excess inventory, obsolescence, and unnecessary holding costs. The consistently low MAE and RMSE achieved by SHOS enhanced models indicate not only improved point accuracy, but also tighter error distributions, which are critical for managing this trade-off effectively.

For inventory managers, this reliability translates into greater confidence in replenishment decisions, particularly for slow-moving items where even small forecast deviations can have disproportionate cost impacts.

Equally important is the operational simplicity of the proposed framework. The use of a single-stage machine learning model augmented with SHOS-derived features reduces implementation complexity relative to multi-stage or hurdle-based pipelines. Such simplicity lowers barriers to deployment, facilitates routine retraining and monitoring, and improves transparency when forecasts must be communicated to planners and stakeholders. These considerations are especially relevant in large-scale supply chains that manage tens of thousands of SKUs amid frequent demand volatility.

The robustness of the SHOS-enhanced model under extreme sparsity further supports its practical applicability. Overfitting diagnostics and stress test evaluations confirm stable generalization across validation windows, indicating reliable out-of-sample performance. This stability is essential for dynamic inventory systems that must adapt continuously to evolving demand patterns without frequent model recalibration.

Taken together, these findings suggest that effective intermittent demand forecasting for inventory management does not require increasingly complex modelling architectures. Instead, embedding statistically grounded, event-aware representations of demand occurrence and magnitude into the forecasting pipeline enables simpler models to deliver more reliable and operationally useful predictions. This shift from architectural complexity to representation quality offers a practical pathway for improving inventory performance while maintaining system scalability and interpretability.

6. Conclusions and managerial implications

Intermittent demand forecasting remains one of the most operationally critical challenges in large-scale supply chains, particularly in aftermarket and spare-parts environments where thousands of slow-moving SKUs exhibit prolonged inactivity punctuated by irregular demand events. While recent research has largely focused on increasing architectural sophistication, the present study demonstrates that substantial and more robust gains can be achieved at the feature-representation level.

This work introduced the SHOS framework as an event-preserving feature engineering approach that explicitly decomposes demand into an occurrence probability and a conditional size component, and embeds these statistically grounded priors directly into supervised learning pipelines. Extensive large-scale experiments on more than 56,000 intermittent time series show that SHOS-derived features consistently reduce forecasting error by approximately 46.3% relative to conventional feature sets, while also improving generalization stability, responsiveness, and recovery under extreme sparsity.

Beyond predictive accuracy, the analysis demonstrates that SHOS preserves event timing, peak structure, and temporal dependence characteristics that are frequently distorted by conventional smoothing or deep sequence models. These properties explain why SHOS-augmented models outperform both raw-feature machine learning models and more complex two-stage or deep learning architectures. The findings indicate that, in highly intermittent regimes, representation quality dominates architectural complexity in determining forecasting performance.

The findings of this study should be interpreted within the context of the evaluated automotive aftermarket forecasting environment, the selected preprocessing strategy, and the forecasting configurations considered in the present investigation. The dataset originates from a single national automotive aftermarket network, and therefore, the reported behaviour may

reflect market-specific sparsity characteristics, dealer behaviour, and inventory dynamics.

The present work focused primarily on point forecasting and event-preserving representation learning under highly intermittent demand conditions. Although the reported forecasting improvements, robustness behaviour, and signal-preservation characteristics are operationally relevant for intermittent-demand planning, the study did not explicitly evaluate probabilistic forecasting, uncertainty quantification, downstream inventory policies, service-level optimization, replenishment cost behaviour, or multi-echelon inventory performance. Consequently, the forecasting improvements reported here should not be interpreted as direct proof of inventory-level optimization performance.

In addition, the reported forecasting behaviour depends on the evaluated feature-generation framework, temporal aggregation strategy, and preprocessing configuration adopted in the present study. While SHOS demonstrated strong performance under the evaluated sparse-demand conditions, further validation is required across alternative spare-parts ecosystems, retail forecasting environments, geographical markets, and heterogeneous supply-chain structures.

Future work should therefore focus on extending SHOS toward probabilistic forecasting environments, integrating downstream inventory simulation and replenishment optimization, and evaluating representation-

preserving sparse-demand forecasting across broader industrial and retail settings.

Managerial and operational implications

From a practical perspective, the SHOS framework provides several actionable advantages for organizations managing large intermittent-demand portfolios:

- Plug-and-play deployment: SHOS operates as a feature-generation layer rather than a forecasting engine. It can be integrated directly into existing ERP, MRP, or demand-planning systems without restructuring current forecasting models or pipelines.
- Cold-start and slow-moving SKU support: By explicitly modelling demand occurrence probability, SHOS provides meaningful signals even when historical demand is extremely sparse, enabling more informed planning for new, rare, and long-tail spare parts.
- Improved stock-out prevention: The extreme-sparsity stress tests show that SHOS-augmented models react significantly faster when demand reappears after long inactivity periods. This responsiveness directly supports earlier replenishment decisions and reduces the risk of unexpected stock-outs.
- Stabilized safety-stock planning: By reducing noise amplification and preserving genuine volatility structure, SHOS improves forecast consistency across heterogeneous SKUs, supporting more reliable service-level and safety-stock calculations.

- Scalable large-portfolio forecasting: Because SHOS embeds statistical structure into features rather than architectures, it remains computationally efficient and suitable for enterprise-scale deployment across tens of thousands of parts.

Scientific and strategic implications

More broadly, this study contributes to a growing body of evidence that learning performance in sparse and intermittent systems is governed primarily by the quality of representation rather than by model complexity. The success of SHOS suggests that statistically grounded, event-aware feature engineering can reduce reliance on complex hybrid architectures, enabling simpler, more interpretable, and more scalable forecasting systems.

Limitations and future work

Future work will extend SHOS toward adaptive parameter learning, automated lifecycle-aware representations, and integration with prescriptive inventory-optimization frameworks. Further evaluation across multi-echelon supply chains, real-time deployment environments, and additional industry sectors will help establish broader generalizability and operational impact.

The present study evaluates SHOS representations under a fixed monthly aggregation framework. Although monthly discretization improves statistical stability under highly sparse demand conditions, alternative temporal resolutions may preserve additional renewal dynamics and short-term demand structure. Future work should therefore investigate weekly

aggregation, adaptive temporal discretization, and multi-resolution intermittent-demand representation frameworks to evaluate the sensitivity of SHOS-derived features to aggregation bucket size across different industrial forecasting environments.

The present evaluation framework focuses primarily on aggregate forecasting accuracy and representation-oriented comparative analysis under sparse-demand conditions. Future work should extend the evaluation protocol to include event-oriented intermittent-demand metrics, such as the sensitivity and specificity of nonzero-demand detection, conditional evaluation restricted to positive-demand observations, and Average Detection Delay (ADD) to measure responsiveness to demand reactivation after prolonged inactivity periods.

Future work should also investigate alternative representation-learning approaches for sparse dealer-part interaction matrices, including zero-inflated probabilistic matrix factorization, Poisson factorization, and nonnegative matrix factorization methods specifically designed for highly sparse count-based environments. Such approaches may further improve latent behavioural representation under extreme intermittent-demand conditions.

Data Availability

All the data and materials used in this study are available in the manuscript, and further details, if required, will be provided by the corresponding author upon proper requisition.

Code Availability

The custom code and computational scripts used for SHOS-based event-preserving feature engineering and intermittent demand forecasting analysis are publicly available at <https://doi.org/10.5281/zenodo.20525971>.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Author Contributions

S. N. B.: Conceptualization (equal); Data curation (lead); Formal analysis (lead); Investigation (equal); Methodology (equal); Writing-original draft (lead); Writing-review & editing (equal). **B. V. S. R.:** Conceptualization (equal); Data curation (lead); Formal analysis (lead); Investigation (equal); Methodology (equal); Writing-original draft (lead); Writing & editing (equal). **C. C. S.:** Conceptualization (equal); Methodology (equal); Formal analysis (lead); Investigation (equal); Writing & editing (lead). **S. S.:** Formal analysis (lead); Investigation (equal); Methodology (equal); Writing-original draft (lead); Writing & editing (equal). **R. C.:** Formal analysis (lead); Investigation (equal); Methodology (equal); Writing-original draft (lead); Writing & editing (equal).

Funding

This work was co-funded by the European Union under the REFRESH - Research Excellence for Region Sustainability and High-tech Industries

project (Project No. CZ.10.03.01/00/22_003/0000048) via the Operational Programme Just Transition. This article was also supported by the Students Grant Competition SP2024/087, Specific Research of Sustainable Manufacturing Technologies, financed by the Ministry of Education, Youth and Sports (MEYS), Czech Republic, and the Faculty of Mechanical Engineering, VŠB-Technical University of Ostrava.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

The authors gratefully acknowledge the support of the funding agencies and institutions listed above.

References

1. Nikolopoulos, K. (2021). We need to talk about intermittent demand forecasting. *European Journal of Operational Research*, 291(2), 549-559. <https://doi.org/10.1016/j.ejor.2019.12.046>
2. Kourentzes, N., & Athanasopoulos, G. (2021). Elucidate structure in intermittent demand series. *European Journal of Operational Research*, 288(1), 141-152. <https://doi.org/10.1016/j.ejor.2020.05.046>
3. Pinçe, Ç., Turrini, L., & Meissner, J. (2021). Intermittent demand forecasting for spare parts: A Critical review. *Omega*, 105, 102513. <https://doi.org/10.1016/j.omega.2021.102513>
4. Croston, J. D. (1972). Forecasting and stock control for intermittent demands. *Journal of the operational research society*, 23(3), 289-303. <https://doi.org/10.2307/3007885>
5. 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>
6. R. H. Teunter, A. A. Syntetos, M. Z. Babai, Intermittent demand: Linking forecasting to inventory obsolescence, *European Journal of Operational Research* 214 (3) (2011) 606-615 <https://doi.org/10.1016/j.ejor.2011.05.018>

7. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... & Liu, T. Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30.
8. T. Chen, C. Guestrin. XGBoost: A Scalable Tree Boosting System. 785-794 In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016. <https://doi.org/10.1145/2939672.2939785>
9. M. Z. Babai, M. Arampatzis, M. Hasni, F. Lolli, A. Tsadiras. On the use of machine learning in supply chain management: a systematic review. *IMA Journal of Management Mathematics* **36**, 21-49 (2025). <https://doi.org/10.1093/imaman/dpae029>
10. R. S. Gutierrez, A. O. Solis, S. Mukhopadhyay. Lumpy demand forecasting using neural networks. *International Journal of Production Economics* **111**, 409-420 (2008). <https://doi.org/10.1016/j.ijpe.2007.01.007>
11. S. Makridakis, E. Spiliotis, V. Assimakopoulos. The M4 Competition: Results, findings, conclusion and way forward. *International Journal of Forecasting* **34**, 802-808 (2018). <https://doi.org/10.1016/j.ijforecast.2018.06.001>
12. Spyros Makridakis, Evangelos Spiliotis, Vassilios Assimakopoulos. The M5 competition: Background, organization, and implementation. *International Journal of Forecasting* **38**, 1325-1336 (2022). <https://doi.org/10.1016/j.ijforecast.2021.07.007>
13. A. A. Syntetos, J. E. Boylan. On the bias of intermittent demand estimates. *International Journal of Production Economics* **71**, 457-466 (2001). [https://doi.org/10.1016/S0925-5273\(00\)00143-2](https://doi.org/10.1016/S0925-5273(00)00143-2)
14. A. H. C. Eaves, B. G. Kingsman. Forecasting for the ordering and stock-holding of spare parts. *Journal of the Operational Research Society* **55**, 431-437 (2004). <https://doi.org/10.1057/palgrave.jors.2601697>
15. R. H. Teunter, A. A. Syntetos, M. Z. Babai. Intermittent demand: Linking forecasting to inventory obsolescence. *European Journal of Operational Research* **214**, 606-615 (2011). <https://doi.org/10.1016/j.ejor.2011.05.018>
16. G. Zhang, B. E. Patuwo, M. Y. Hu, Forecasting with artificial neural networks: The state of the art, *International Journal of Forecasting* 14 (1) (1998) 35-62. [https://doi.org/10.1016/S0169-2070\(97\)00044-7](https://doi.org/10.1016/S0169-2070(97)00044-7)
17. S. Smyl. A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. *International Journal of*

- Forecasting* **36**, 75-85 (2020).
<https://doi.org/10.1016/j.ijforecast.2019.03.017>
18. L. Li, Y. Kang, F. Petropoulos, F. Li. Feature-based intermittent demand forecast combinations: accuracy and inventory implications. *International Journal of Production Research* **61**, 7557-7572 (2023).
<https://doi.org/10.1080/00207543.2022.2153941>
 19. Nathan, B. S., Siva Reddy, B. V., Sastry, C. C., Krishnaiah, J., & Eswaramoorthy, K. V. (2024). Innovative framework for effective service parts management in the automotive industry. *Frontiers in Mechanical Engineering*, **10**, 1361688.
<https://doi.org/10.3389/fmech.2024.1361688>
 20. Doszyń, M. (2019). Intermittent demand forecasting in the Enterprise: Empirical verification. *Journal of Forecasting*, **38**(5), 459-469. <https://doi.org/10.1002/for.2575>
 21. B, S.N., Reddy, B.V.S., Sujana, V.S.S. *et al.* A machine learning framework for long-term forecasting of spare part demand in end-of-life product scenarios. *Sci Rep* (2025). <https://doi.org/10.1038/s41598-025-31171-2>
 22. Kleen, O. (2024). Scaling and measurement error sensitivity of scoring rules for distribution forecasts. *Journal of Applied Econometrics*, **39**(5), 833-849. <https://doi.org/10.1002/jae.3056>
 23. M. Lichman, P. Smyth. Prediction of sparse user-item consumption rates with zero-inflated poisson regression. 719-728 In *The Web Conference 2018 - Proceedings of the World Wide Web Conference, WWW 2018*. ACM, 2018.
<https://doi.org/10.1145/3178876.3186153>
 24. P. Wallström, A. Segerstedt. Evaluation of forecasting error measurements and techniques for intermittent demand. *International Journal of Production Economics* **128**, 625-636 (2010).
<https://doi.org/10.1016/j.ijpe.2010.07.013>
 25. Bobbili, V. S. R., C, C. S., & A, H. R. (2026). Explainable machine learning for high-strain-rate prediction of LPBF-fabricated A286 lattice structures under SHPB testing. *Rapid Prototyping Journal*, 1-20.
<https://doi.org/10.1108/RPJ-08-2025-0362>
 26. A. C. Cameron, P. K. Trivedi. Regression Analysis of Count Data. Cambridge University Press, 2012.
 27. Nathan, B. S., Aravindh, P. M., Reddy, B. V. S., Sastry, C. C., Salunkhe, S., & Cep, R. (2026). Primacy of feature engineering over architectural complexity for intermittent demand forecasting. *Scientific Reports*. <https://doi.org/10.1038/s41598-026-35197-y>

28. N. Kourentzes. On intermittent demand model optimisation and selection. *International Journal of Production Economics* 156, 180-190 (2014). <https://doi.org/10.1016/j.ijpe.2014.06.007>
29. Bertsimas, D., Digalakis Jr, V., Li, M. L., & Lami, O. S. (2025). Slowly varying regression under sparsity. *Operations Research*, 73(3), 1581-1597. <https://doi.org/10.1287/opre.2022.0330>
30. Hafiz, F., Broekaert, J., La Torre, D., & Swain, A. (2024). A multi-criteria approach to evolve sparse neural architectures for stock market forecasting. *Annals of Operations Research*, 336(1), 1219-1263. <https://doi.org/10.1007/s10479-023-05715-6>
31. Affonso, T. B., Conceição, S. V., Muniz, L. R., de Freitas Almeida, J. F., & de Lima, J. C. (2024). A new hybrid forecasting method for spare part inventory management using heuristics and bootstrapping. *Decision Analytics Journal*, 10, 100415. <https://doi.org/10.1016/j.dajour.2024.100415>
32. Kim, J. D., Hwang, J. H., & Doh, H. H. (2023). A predictive model with data scaling methodologies for forecasting spare parts demand in military logistics. *Defence science journal*, 73(06), 666-674. <https://doi.org/10.14429/dsj.73.19129>
33. Huang, F., Zhang, J., Zhou, C., Wang, Y., Huang, J., & Zhu, L. (2020). A deep learning algorithm using a fully connected sparse autoencoder neural network for landslide susceptibility prediction. *Landslides*, 17(1), 217-229. <https://doi.org/10.1007/s10346-019-01274-9>
34. Kamara, A. F., Chen, E., & Pan, Z. (2022). An ensemble of a boosted hybrid of deep learning models and technical analysis for forecasting stock prices. *Information Sciences*, 594, 1-19. <https://doi.org/10.1016/j.ins.2022.02.015>
35. Liang, L., Liu, B., Su, Z., & Cai, X. (2024). Forecasting corporate financial performance with deep learning and interpretable ALE method: Evidence from China. *Journal of Forecasting*, 43(7), 2540-2571. <https://doi.org/10.1002/for.3138>
36. Tharinda, N., Kosgoda, D., Tibbotuwawa, A., & Nielsen, I. (2025, June). Spare Parts Categorization and Machine Learning Approaches for Demand Forecasting in Automotive Spare Parts Sector of Developing Economies. In *International Conference on Intelligent Systems in Production Engineering and Maintenance* (pp. 262-277). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-99159-2_19
37. Anglou, F. Z., Ponis, S., & Spanos, A. (2021). A machine learning approach to enable bulk orders of critical spare-parts in the shipping

industry. *Journal of Industrial Engineering and Management*, 14(3), 604-621. <https://doi.org/10.3926/jiem.3446>

38. Ifraz, M., Aktepe, A., Ersöz, S., & Çetinyokuş, T. (2023). Demand forecasting of spare parts with regression and machine learning methods: Application in a bus fleet. *Journal of Engineering Research*, 11(2), 100057. <https://doi.org/10.1016/j.jer.2023.100057>
39. Fan, L., Liu, X., Mao, W., Yang, K., & Song, Z. (2023). Spare parts demand forecasting method based on intermittent feature adaptation. *Entropy*, 25(5), 764. <https://doi.org/10.3390/e25050764>
40. Giannopoulos, P. G., Dasaklis, T. K., Tsantilis, I., & Patsakis, C. (2025). Machine learning algorithms in intermittent demand forecasting: a review. *International Journal of Production Research*, 1-43. <https://doi.org/10.1080/00207543.2025.2578701>