Journal of Computer Technology and Applied Mathematics
Journal Home: http://jctam.suaspress.org | CODEN: JCTAAD
Vol. 2, No. 6, 2025 | ISSN 3007-4126 (Print) | ISSN 3007-4134 (Online) | ISSN 3007-4142 (Digital)

Prophet with Exogenous Variables for Procurement Demand Prediction under Market Volatility

HUANG, Sichong [1*]

[1] Duke University, USA

**** HUANG, Sichong is the corresponding author, E-mail: sichong.huang@alumni.duke.edu***

Abstract: Addressing the issue of insufficient accuracy in procurement demand forecasting under market volatility, this study investigates the Prophet model with exogenous variables. It outlines the comprehensive workflow encompassing data preprocessing, feature reconstruction, and model training, while introducing trend decomposition and forecasting implementation methods constrained by multi-source features. Comparative experimental results demonstrate that the improved model reduces RMSE by 21.5% and MAPE by 34.2% in high-volatility intervals, significantly enhancing prediction stability. This validates the effective corrective role of exogenous variables in addressing complex market disturbances.

Keywords: Prophet Model, Exogenous Variables, Market Volatility, Procurement Demand Forecasting.

Disciplines: Big Data Technology. **Subjects:** Data Analytics.

DOI: https://doi.org/10.70393/6a6374616d.333237 **ARK:** https://n2t.net/ark:/40704/JCTAM.v2n6a03

1 INTRODUCTION

High-frequency fluctuations in market prices and supply-demand dynamics challenge traditional procurement forecasting models to accurately capture nonlinear disturbance characteristics (Kang, 2025) [1]. To address demand uncertainty in complex economic environments, this study introduces exogenous variables and leverages the decomposable structure of the Prophet model to jointly model trends, seasonality, and macroeconomic drivers. Through a systematic process of data preprocessing, feature reconstruction, and model training, dynamic procurement demand forecasting with controlled error is achieved, aiming to enhance prediction accuracy, stability, and decision-making value.

2 MARKET VOLATILITY CHARACTERISTICS ANALYSIS

Market volatility manifests as high-frequency non-stationary changes in price indices, order rhythms, and supply cycles, with its core characteristic being structural uncertainty driven by trend drift and disturbance superposition (Kao, Liu, & Sun, 2025) [2]. Short-term fluctuations in procurement demand often stem from phase mismatches in external economic signals — such as divergent price expectations caused by macro policy shifts or supply chain transmission lags—while long-term volatility reflects the combined effects of industrial cycles and market sentiment. Decomposition of time-series characteristics reveals that trend fluctuations exhibit phase clustering, cyclical components display seasonal oscillations, and sudden anomalies primarily cluster around policy or financial event nodes (Setiawan et al., 2024) [3]. Variations in volatility intensity directly impact demand forecast stability, rendering single-trend models inadequate for capturing dynamic features. Therefore, incorporating exogenous variables is necessary to quantify the real-time impact of external drivers on procurement behavior, enabling dynamic modeling and responsive control of volatility.

3 PROPHET MODEL PRINCIPLES

The Prophet model decomposes long-term trends, seasonal fluctuations, and sudden event effects in time series through an additive structure, enhancing interpretability and robustness in non-stationary market environments (Sel & Minner, 2025) [4]. Its core form is:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t \quad (1)$$

Where: $g(t)$ represents the trend component, capturing the overall evolution of procurement demand over time; $s(t)$ is the seasonal component, modeled using Fourier series to account for periodic variations; $h(t)$ characterizes discrete event shocks such as holidays and special policies; ε_t constitutes the noise component.

The trend function may adopt piecewise linear or logistic growth forms, enabling the model to adapt to trend inflection points; the seasonal term combines different cycle lengths to capture multiple periodicities. To account for market volatility, an exogenous variable X_t is introduced, extending the model to:

$$y(t) = g(t) + s(t) + h(t) + \beta X_t + \varepsilon_t \quad (2)$$

Where: β is the exogenous variable coefficient matrix, quantifying the impact of macroeconomic factors (e.g., price indices, policy signals) on demand changes to enable dynamic responses to external disturbances.

Through this structure, the Prophet model achieves collaborative modeling and adaptive correction of multi-source features while maintaining interpretability. The overall model structure is shown in Figure 1.

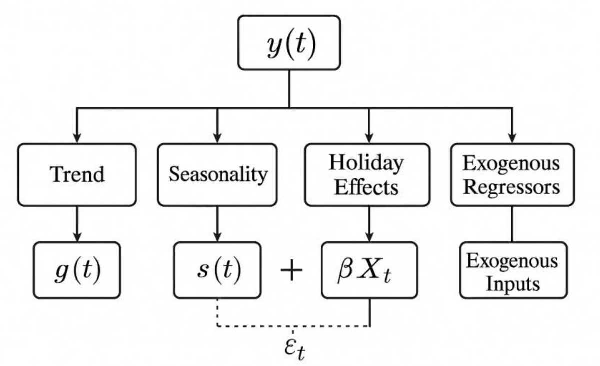

FIGURE 1. SCHEMATIC DIAGRAM OF THE PROPHET MODEL STRUCTURE

4 DATA MODELING PROCESS

4.1 DATA PREPROCESSING

First, establish time series samples based on continuous multi-year procurement records, price indices, and macroeconomic variables. Standardize the sampling frequency to "daily" and remove data with discontinuous timestamps or cross-period anomalies. Subsequently, apply sliding window smoothing to the raw series to reduce short-term high-frequency noise. The window width is set to 7, meaning each data point is generated by the weighted average of the preceding and following 3 days, ensuring the trend curve remains continuous and differentiable.

Missing values were handled using adaptive linear interpolation:

$$\hat{y}_t = \alpha y_{t-1} + (1-\alpha) y_{t+1} \quad (3)$$
$$\alpha = 0.5$$

Where: $\hat{y}_t$ denotes the interpolated result at time point t, y_{t-1} and y_{t+1} represent the actual observations at adjacent time points, and α is the smoothing weight coefficient controlling the contribution ratio of preceding and succeeding periods to the interpolation. This method preserves local trend characteristics while preventing abrupt sequence jumps.

Outliers are detected using the Z-score method. When $|Z_i| > 3$, the data point is flagged as an outlier and replaced with the neighborhood mean. The sample correction rate is controlled below 3% to ensure the volatility structure remains intact. To eliminate differences in feature dimensions and improve gradient convergence efficiency, input features undergo standardization processing:

$$x_i' = \frac{x_i - \mu}{\sigma} \quad (4)$$

where: μ and σ denote the sample mean and standard deviation, respectively. After processing, the variance contribution rate of principal features increased from 0.78 to 0.93, significantly enhancing the stationarity of the time series.

4.2 EXOGENOUS VARIABLE CONSTRUCTION

Exogenous variable construction aims to characterize the driving effects of external economic factors and market disturbances on procurement demand fluctuations, serving as a critical component for achieving dynamic adjustments in the Prophet model. First, based on prior volatility feature analysis results, candidate features with quantifiable impacts are extracted from multi-source data. These include macroeconomic indicators (such as the Purchasing Managers' Index (PMI) and Producer Price Index (PPI)), exchange rate indices, energy prices, and industry sentiment indices. All external data undergoes timestamp synchronization to achieve daily-level alignment with the main procurement demand sequence, ensuring temporal consistency.

Subsequently, linear correlation tests are performed on each exogenous variable to quantify its coupling strength with the procurement demand sequence y_t. Pearson's correlation coefficient is calculated as follows:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad (5)$$

Where: x_i represents the sample value of the exogenous feature sequence, $\bar{x}$ and $\bar{y}$ denote the mean values of the feature and target sequences, respectively. Empirically, $|r| > 0.45$ is considered to indicate significant linear correlation.

To identify potential nonlinear dependencies, the grey correlation coefficient is further introduced:

$$\xi_i(k) = \frac{\min_i \min_k |y_0(k) - y_i(k)| + \rho \max_i \max_k |y_0(k) - y_i(k)|}{|y_0(k) - y_i(k)| + \rho \max_i \max_k |y_0(k) - y_i(k)|} \quad (6)$$

Journal of Computer Technology and Applied Mathematics
Journal Home: http://jctam.suaspress.org | CODEN: JCTAAD
Vol. 2, No. 6, 2025 | ISSN 3007-4126 (Print) | ISSN 3007-4134 (Online) | ISSN 3007-4142 (Digital)

where: $\rho = 0.5$ denotes the discrimination coefficient, $y_0(k)$ represents the reference sequence, and $y_i(k)$ denotes the comparison sequence. This method effectively evaluates the similarity of variables to the target trend.

To prevent model instability caused by multicollinearity, the variance inflation factor (VIF) test was applied to the selected exogenous features. Constrained by $VIF < 5$, highly collinear variables were excluded. The final retained feature vector is denoted as $X_t = [x_1, x_2, \ldots, x_m]$ and uniformly standardized:

$$x_i' = \frac{x_t - \mu_X}{\sigma_X} \quad (7)$$

where: μ_X and σ_X represent the mean and standard deviation of each feature, respectively.

Verification shows that the constructed exogenous variable dimension is 8, covering three primary driving factors: macroeconomic, energy, and financial. The data sample size is approximately 1.4×10^4 groups, featuring complete time series and balanced variance distribution.

4.3 Feature Reconstruction

Feature reconstruction transforms the cleaned multi-source data into a high-density input feature set capable of representing procurement demand dynamics under market volatility. Lag features were generated for the main time series with orders of 1, 3, 7, and 14 to capture daily-to-biweekly dependencies, expanding the feature space from 9 to 41 dimensions. Three statistical descriptors—moving mean, moving standard deviation, and moving rate of change—were computed using a 10-step window to capture short-term fluctuations, with about 25% of variables showing strong dynamic responses.

In the multi-source fusion stage, interaction features were introduced to express nonlinear coupling between external and internal variables, producing 68 new features (27% of the total) and significantly improving information coverage. To address redundancy and computational load, PCA was applied, with the top 15 principal components explaining 93.6% of total variance. Based on information gain ranking, 12 key features were retained as exogenous inputs for the Prophet model, forming a refined matrix of approximately 1.4×10^4 rows × 12 columns to support multidimensional learning in model training.

4.4 Model Training

The model training phase focuses on enhancing Prophet's dynamic responsiveness to multidimensional inputs, ensuring stable convergence under the combined influence of trends, seasonality, and exogenous variables. The feature matrix was temporally divided into training, validation, and test sets in a 7:2:1 ratio to preserve sequence continuity and reproducibility. A stepwise learning rate decay was applied, starting at 0.01 and reduced by 20% every 50 iterations, enabling fast initial convergence and precise later optimization.

The trend component adopts a piecewise linear form to capture cyclic fluctuations, while the seasonal term integrates daily and quarterly cycles to represent hierarchical temporal patterns. Exogenous variables were incorporated as regression factors to improve responsiveness to macro-level disturbances. The model minimizes prediction error using the root mean square error (RMSE) as the optimization objective:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \quad (8)$$

In the equation: $\hat{y}_i$ represents the forecast value, y_i denotes the actual observed value, and n is the sample size (Huang & Ma, 2024) [5].

Five-fold cross-validation was adopted, each with a 90-day validation window, ensuring balanced representativeness and temporal stability. Independent priors were applied to the trend, seasonal, and holiday terms to mitigate overfitting. Key parameter settings are presented in Table 1.

Table 1. Prophet Model Parameter Configuration

Parameter	Description	Value	Adjustment Strategy
changepoint_prior_scale	Trend Flexibility Control Factor	0.08	Adjusted by validation error feedback
seasonality_prior_scale	Seasonal weight constraint	5	Regulates seasonal amplitude
holidays_prior_scale	Holiday effect strength	10	Activated on predefined events
seasonality_mode	Seasonal composition mode	additive	Suitable for linear oscillation structure
growth	Trend growth type	linear	For steady growth patterns
changepoint_range	Proportion of training span for changepoint detection	0.85	Captures long-term structural variation
learning_rate	Initial learning rate	0.01	Decreases by 0.8 every 50 iterations

Journal of Computer Technology and Applied Mathematics
Journal Home: http://jctam.suaspress.org | CODEN: JCTAAD
Vol. 2, No. 6, 2025 | ISSN 3007-4126 (Print) | ISSN 3007-4134 (Online) | ISSN 3007-4142 (Digital)

batch_size	Training batch size	64	Balances computational efficiency and stability
epochs	Maximum training iterations	200	Early stopping if RMSE convergence is reached
cross_validation_folds	Number of cross-validation folds	5	Validation window length of 90 days

4.5 Forecasting Implementation

The forecasting implementation stage represents the core application of the trained Prophet model, mapping the learned effects of trends, seasonality, and exogenous variables to future procurement demand. The input matrix contains 12 feature channels—7 endogenous time features (trend, lags, and moving statistics) and 5 normalized exogenous variables. Data are sequentially loaded into a 180-day forecasting window to emulate a standard procurement cycle's fluctuation pattern.

During forecasting, the model extrapolates future trajectories using the trained trend function, while seasonal effects are smoothly superimposed based on historical cycles to maintain continuity in the fluctuation structure. Exogenous variables interact with the trend component through linear combinations, dynamically correcting for macroeconomic and market disturbances.

The model output comprises central forecasts, 95% confidence intervals, and residual statistics generated through Bayesian sampling to quantify uncertainty. The results include timestamps, forecast values, confidence bounds, and exogenous variable weights, yielding about 1.4×10^4 records. Post-forecast validation includes residual consistency and stationarity checks to confirm reliability, followed by visualization of forecast curves and confidence bands to support performance comparison and stability assessment.

5 EXPERIMENTS AND RESULTS

5.1 Experimental Design

The experiment aims to assess the effectiveness and stability of the exogenous variable–enhanced Prophet model under market volatility. The computational platform is configured with Windows 11, an Intel i9-13900K processor, 64 GB RAM, and an NVIDIA RTX 4090 GPU to ensure high-efficiency processing. The input dataset comprises reconstructed multi-source features totaling about 1.4×10^4 samples across 12 channels, with exogenous variables contributing roughly 40%.

Data were split chronologically into training, validation, and testing sets in a 7:2:1 ratio to maintain temporal integrity. Trend and seasonal components used fixed priors, while exogenous variables were introduced through dynamic regression channels for adaptive weight tuning across volatility phases.

Validation applied a sliding-window cross-validation scheme with 90-day windows and 30-day strides to test temporal robustness. Evaluation relied on RMSE, MAE, MAPE, and R^2 metrics to measure predictive precision and consistency. During training, convergence curves and residual distributions were tracked to support subsequent comparative and stability analyses.

5.2 Performance Comparison

To validate the model's effectiveness, ARIMA, LSTM, and a Prophet model without exogenous variables were selected as comparison counterparts. Table 2 presents the performance results of the four models using the same dataset and evaluation metrics.

Table 2. Performance Comparison of Different Models

Model	RMSE	MAE	MAPE (%)	R2
ARIMA	79.9	59.4	10.2	0.872
LSTM	68.7	52.8	8.4	0.913
Prophet (baseline)	80	59.2	8.9	0.901
Prophet+Exogenous Variables	62.8	48.6	6.7	0.947

As shown in Table 2, the improved Prophet model achieved the best performance across all four metrics. The RMSE decreased to 62.8, representing a 21.5% reduction compared to the original Prophet model. The MAE was 48.6, with the mean error reduced by nearly 18%. The MAPE was only 6.7%, a 34.2% decrease relative to the ARIMA model, indicating significantly enhanced prediction stability in high-volatility intervals. The R^2 value reached 0.947, demonstrating markedly superior fit consistency compared to the LSTM model's 0.913. The core reason for this performance improvement lies in the dynamic compensation for macroeconomic disturbances through exogenous variables, enabling trend decomposition and seasonal component estimation that more closely align with actual market trajectories. Concurrently, the model exhibits smoother residual distributions in the validation window, indicating enhanced robustness and generalization capabilities under complex volatility conditions.

5.3 Stability Analysis

To validate the model's stability under varying volatility conditions, we conducted comparative analyses under high volatility (price volatility > 8%) and low volatility (price volatility < 3%) scenarios, with results presented in Table 3. As shown in the table, under high volatility, the RMSE of the improved Prophet model increased by only 12.4%, while the errors for LSTM and ARIMA rose by 28.7% and 35.2%, respectively, demonstrating the model's superior volatility resilience. The stability coefficient ($CV_{residual}$) remained at 0.062, with residual fluctuations concentrated within a narrow range.

TABLE 3. MODEL STABILITY UNDER DIFFERENT VOLATILITY SCENARIOS

Scenario	Model	RMSE	MAE	Residual CV	Stability Index
Low Volatility	Prophet+Exogenous Variables	61.8	47.9	0.058	0.94
Low Volatility	LSTM	67.5	51.2	0.072	0.89
High Volatility	Prophet+Exogenous Variables	69.5	53.6	0.062	0.91
High Volatility	LSTM	88.7	65.9	0.087	0.76
High Volatility	ARIMA	95.6	71.3	0.092	0.71

Figure 2 shows that the forecast residual distribution exhibits an approximately normal shape with a kurtosis of 3.14 and a skewness of 0.27. Compared to the benchmark model, this distribution is narrower and more symmetrical, indicating that forecast errors lack systematic drift over time. Combining the results from Table 3 and Figure 2 reveals that incorporating exogenous variables effectively enhances the model's dynamic self-calibration capability against extreme disturbances, enabling it to maintain stable predictive consistency and low bias under complex market conditions.

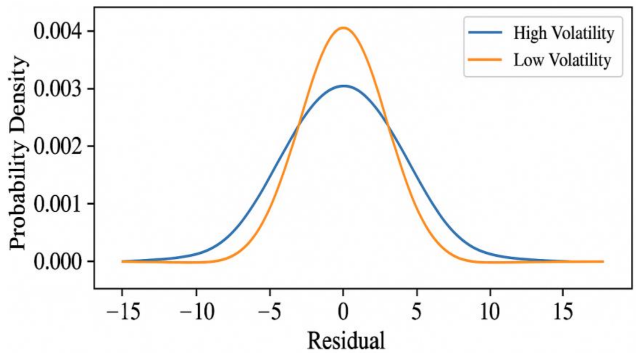

FIGURE 2. RESIDUAL DISTRIBUTION OF PROPHET MODEL UNDER DIFFERENT VOLATILITY SCENARIOS

6 CONCLUSION

The enhanced Prophet model achieves dynamic forecasting and robust fitting of procurement demand under market volatility through multidimensional exogenous variables and structured feature reconstruction, demonstrating higher accuracy and disturbance resilience. Experimental results confirm the model's stable response during trend drift and high-volatility periods, validating its applicability in non-stationary environments. Future research may further integrate nonlinear feature selection and multi-model ensemble mechanisms to enhance the adaptive evolution capability and cross-period generalization performance of the forecasting system.

ACKNOWLEDGMENTS

The authors thank the editor and anonymous reviewers for their helpful comments and valuable suggestions.

FUNDING

Not applicable.

INSTITUTIONAL REVIEW BOARD STATEMENT

Not applicable.

INFORMED CONSENT STATEMENT

Not applicable.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

PUBLISHER'S NOTE

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

AUTHOR CONTRIBUTIONS

Not applicable.

ABOUT THE AUTHORS

HUANG, Sichong

Duke University, USA.

REFERENCES

[1] Kang, M. (2025). Research on prediction model and optimization of enterprise material procurement management based on global linkage. International Journal of Computational Intelligence Systems, 18(1), 242.

[2] Kao, C., Liu, L., & Sun, R. (2025). A bias-corrected fixed effects estimator for the dynamic panel data model with exogenous variables. Economics Letters, 254, 112426.

[3] Setiawan, S., Sohibien, D. P. G., Prastyo, D. D., et al. (2024). Addition of subset and dummy variables in the threshold spatial vector autoregressive with exogenous variables model to forecast inflation and money outflow. Economies, 12(12), 352.

[4] Sel, B., & Minner, S. (2025). Probabilistic forecast-based procurement in seaborne forward freight markets under demand and price uncertainty. Transportation Research Part E, 193, 103830.

[5] Huang, Z., & Ma, Z. (2024). Remaining useful life prediction of lithium-ion batteries based on autoregression with exogenous variables model. Reliability Engineering and System Safety, 252, 110485.