International Journal on Advanced Computer Theory and Engineering

ISSN: 2319 - 2526

Volume 15 Issue 2s (2026) | Pages: 117-124

# Effective Procurement and Demand Forecasting for Ready-Mix Concrete Plants

Ansh Mahajan[1], Yuganti Kumkar[2], Swathi Pepeti[3], Anjali Bijwe[4]

[1,2,3] Student, Department of AI&DS, GSMCOE Balewadi, Maharashtra, India

[4]Professor, Department of AI&DS, GSMCOE Balewadi, Maharashtra, India

*Email: [1]anshmahajan345@gmail.com, [2]yugantikumkar@gmail.com, [3]swathipepeti@gmail.com, [4]anjaliabijwe@gmail.com*

**Peer Review Information**

***Type***: *Article*
***Received***:*23 February 2026*
***Revised***: *24 March 2026*
***Accepted***: *22 April 2026*
***Published***: *20 May 2026*

**Abstract**

Accurate forecasting of Ready-Mix Concrete (RMC) demand helps optimize procurement, reduce waste, and lower environmental impact. This study presents an AI-based system that predicts daily demand using 18 operational features and recent 5,000 data points (80:20 split). Multiple models were tested, where LSTM performed poorly (MAPE 145.10%), while a hybrid model combining XGBoost (0.35) and Random Forest (0.c5) achieved the best results (RMSE 3.c8 $m^3$, MAPE 1.55%). Feature analysis shows cement density (32.81%) as the most influential factor. The system also estimates $CO_2$ emissions, enabling cost-efficient and sustainable procurement decisions, demonstrating the effectiveness of hybrid ensemble learning in construction forecasting.

**Keywords:** Ready-Mix Concrete Forecasting; Ensemble Machine Learning; XGBoost; Sustainable Procurement; Carbon Emission Estimation; Predictive Analytics

**How to Cite This Article**

Mahajan, A., Kumkar, Y., Pepeti, S., & Bijwe, A. (2026). *Effective Procurement and Demand Forecasting for Ready-Mix Concrete Plants*. *International Journal on Advanced Computer Theory and Engineering,* 15(2s), 117–124.

## Introduction

The rapid integration of artificial intelligence (AI) is transforming supply chain management from reactive to predictive systems [7]. In construction, the Ready-Mix Concrete (RMC) sector presents complex logistical challenges due to perishability, strict delivery timelines, and variable localized demand, requiring predictive models that capture multidimensional, non-linear patterns rather than relying on historical estimates [4].

Existing research remains fragmented, focusing on material properties like compressive strength using supervised learning [2] and dispatch optimization by replicating human decision-making for post-order logistics [3]. These approaches are reactive and fail to address proactive volumetric demand forecasting.

In related fields such as power engineering, LSTM models are used for time-series forecasting [5, 6], but they often perform poorly in non-stationary environments like RMC demand influenced by unpredictable operational factors. Additionally, limited research integrates demand prediction with environmental impact analysis despite high emissions from cement production, highlighting a gap in developing unified machine learning frameworks for both demand forecasting and carbon footprint estimation [1].

## Literature Review

The intersection of artificial intelligence, supply chain logistics, and construction materials has gained significant research attention, yet most studies treat material behavior and demand

forecasting separately. This section reviews prior work in supply chain forecasting, concrete modeling, and dispatch operations to identify key gaps.

*A in Supply Chain and Logistics Forecasting*

AI has improved logistics by enabling adaptive and optimized systems. Pattnaik et al. [7] highlight its shift from rigid scheduling, while Wang and Chen [4] show deep learning outperforms traditional methods by incorporating multiple influencing factors. However, these approaches are not tailored for the high variability and perishability of the Ready-Mix Concrete (RMC) sector.

*Machine Learning in Concrete Applications*

In concrete engineering, machine learning mainly focuses on material properties rather than demand. Wang et al. [2] demonstrate accurate compressive strength prediction using mix parameters, but this does not address volumetric demand. Nandal et al. [1] forecast cement demand using macroeconomic indicators, yet such macro-level models fail to capture localized, short-term demand required for real-time operations.

*Sequence Models and Dispatch Optimization*

Research in logistics optimization often models dispatch decisions. Maghrebi et al. [3] use rule-based systems to replicate expert strategies, improving scheduling but remaining reactive. Sequence models like LSTM, effective in power demand forecasting [5], [6], assume stable temporal patterns and perform poorly in non-stationary environments like RMC demand influenced by unpredictable operational factors.

*Identification of Research Gap*

Existing research remains fragmented, focusing on macro demand [1], material properties [2], or reactive dispatch [3] without integration. There is a lack of proactive systems that use material composition to predict localized daily RMC demand, along with minimal integration of environmental impact analysis, particularly carbon emissions. To address this, the study proposes a hybrid ensemble-based framework combining accurate demand prediction with carbon emission estimation, enabling efficient procurement and sustainable decision-making.

## Methodology

This study adopts a multi-stage, data-driven methodology to develop a hybrid machine learning framework for forecasting Ready-Mix Concrete (RMC) demand. The proposed approach integrates predictive modeling with procurement planning and carbon emission estimation, transforming traditional heuristic-based estimation into a structured and data-driven process.

*Research Design, Data Collection, and Sample Size*

The predictive framework is developed using a multidimensional dataset representing historical Ready-Mix Concrete (RMC) dispatch operations. The dataset consists of 146,970 observations collected over a five-year period (2020–2024), ensuring coverage of seasonal trends and operational variability. The system processes 24 input features spanning material, operational, meteorological, and temporal dimensions, enabling the model to capture physical mix properties and external factors influencing demand patterns.

To ensure robust model evaluation and generalization, the complete dataset of 146,970 observations is utilized to maximize learning capability and improve predictive accuracy, and the dataset is divided using an 80:20 train–test split, where approximately 117,576 samples are used for training and 29,394 samples are reserved for testing, with model performance evaluated using RMSE and MAPE.

*Table 1. Input Feature Classification*

| Feature Category | Attributes | Purpose |
|---|---|---|
| Material & Operational | cement_density, carbon_coeff, fleet_availability, procurement_lead_time | Captures mix composition and logistical constraints |
| Meteorological | temperature, humidity, precipitation, wind_speed | Represents environmental factors affecting construction activity |
| Temporal & Lag Data | year, month, day, day_of_week, is_weekend, lag_1d, lag_7d, rolling_mean | Captures seasonality, trends, and historical dependencies |

To address differences in feature scales, all numerical variables are standardized using z-score normalization via the StandardScaler function from scikit-learn, ensuring that variables with larger numerical ranges do not dominate the learning process.

*Data Preprocessing and Feature Engineering*

Prior to model training, the dataset undergoes a structured preprocessing pipeline to ensure data consistency and model robustness. Physical Concrete Properties include parameters such as cement_kg_m3, water_binder_ratio, slump_mm, and compressive_strength_mpa, which define material composition and structural characteristics. Aggregate Composition features such as aggregate_10mm_pct, aggregate_20mm_pct, and agg_sieve_pass_4_75mm_pct represent the distribution of coarse and fine aggregates, enabling the model to capture density-related effects on volume demand. Logistical Attributes including transport_time_min, batching_time_min, and truck_capacity_m3 account for delivery constraints and batching efficiency.

*Automated Model Selection*

To avoid reliance on a single predictive model, an automated evaluation process compares multiple regression algorithms: LSTM (Long Short-Term Memory), evaluated for sequence modeling but excluded due to poor performance under highly non-stationary demand conditions; XGBoost, selected for its capability to model complex non-linear relationships and handle structured data effectively; and Random Forest, selected for its robustness and ability to reduce overfitting through ensemble averaging.

*Hybrid Ensemble Model*

Experimental results indicated that no single model consistently achieved optimal performance. Therefore, a hybrid ensemble approach is proposed, combining the strengths of XGBoost and Random Forest.

$Df = 0.35 \,.\, \hat{y}XGB + 0.65 \,.\, \hat{y}RF$

where ŷ_XGB and ŷ_RF represent the predicted outputs of the XGBoost and Random Forest models, respectively, and D_f denotes the final forecasted demand. The Random Forest component provides stability and reduces sensitivity to outliers, while XGBoost enhances predictive precision by capturing fine-grained variations in the data, with weights determined empirically based on performance evaluation.

*Carbon Emission Estimation*

To incorporate sustainability considerations, the predicted RMC demand is integrated with a carbon emission estimation module. The system utilizes a predefined carbon coefficient associated with cement usage to estimate the total carbon dioxide ($CO_2$) emissions corresponding to the forecasted volume. This integration enables real-time assessment of environmental impact, supporting informed and sustainable procurement decisions.

**System Architecture**

The overall system follows a structured pipeline:

1. Data ingestion from historical dispatch records
2. Feature preprocessing and transformation
3. Model training and evaluation
4. Hybrid ensemble prediction
5. Carbon emission estimation
6. Final output generation for decision support

This pipeline ensures a seamless transition from raw data processing to actionable forecasting and sustainability analysis.

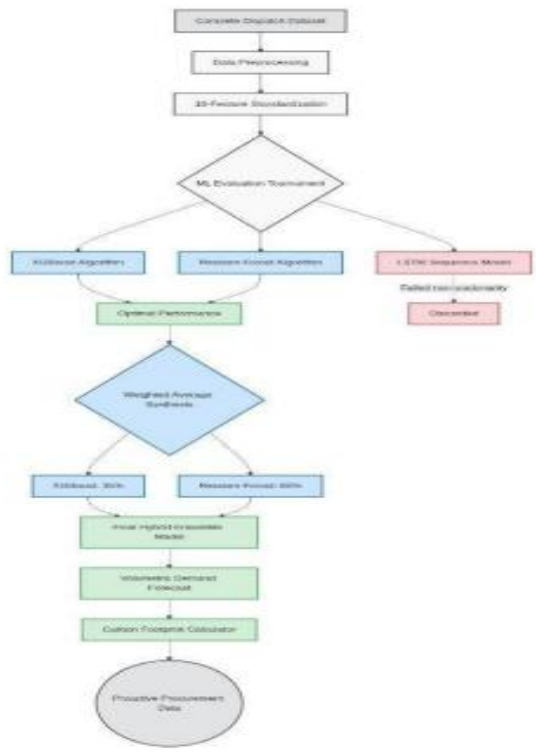

***Fig. 1.*** *Proposed Hybrid Machine Learning Pipeline and Tournament Architecture*

**Results/Findings**

To evaluate predictive performance, standard regression error metrics are employed. Let $y_i$denote actual demand, $\hat{y}_i$predicted demand, and $n$the total observations.

*Performance Evaluation Metrics*

The framework was evaluated using 146,970 RMC dispatch records, demonstrating accurate forecasting and operational reliability on a 20% test set.

*Root Mean Squared Error (RMSE):*

RMSE penalizes larger errors, important in procurement scenarios.

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

*Algorithmic Evaluation and Model Comparison*

Multiple algorithms were evaluated under identical conditions (Table 2).

*Table 2: Out-of-Sample Regression Validation Metrics*

| Machine Learning Algorithm | RMSE ($m^3$) | MAE ($m^3$) | MAPE (%) |
|---|---|---|---|
| Long Short-Term Memory (LSTM) | 259.43 | 221.01 | 145.10 |
| Linear Regression | 7.99 | 6.96 | 3.44 |
| XGBoost | 3.71 | 1.60 | 1.57 |
| Random Forest | 3.69 | 1.63 | 1.57 |
| Proposed Hybrid Ensemble | 3.68 | 1.59 | 1.55 |

LSTM exhibited extremely high error (MAPE: 145.10%), indicating poor suitability due to non-stationary demand. Linear Regression achieved moderate performance (MAPE: 3.44%) but was limited in capturing non-linear relationships. XGBoost and Random Forest demonstrated strong predictive capability (MAPE: 1.57%). The hybrid ensemble achieved the best performance (RMSE: 3.68 $m^3$, MAPE: 1.55%), confirming improved stability and accuracy.

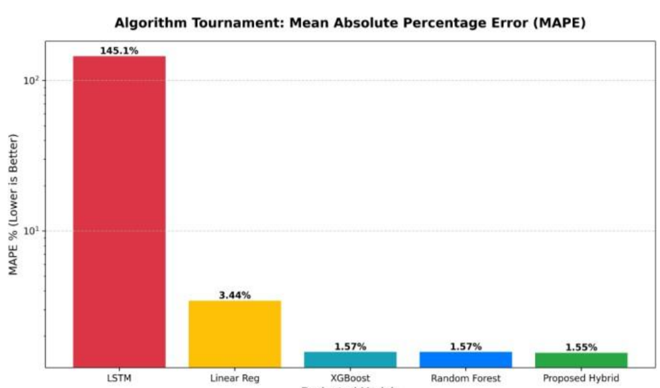


***Fig. 2.*** *Mean Absolute Percentage Error (MAPE) Comparison of Evaluated Algorithms*

*Performance of the Proposed Hybrid Architecture*

The hybrid model achieved the lowest MAPE (1.55%), improving predictive capability. Reducing error to 1.55% improves procurement reliability in RMC logistics. The model achieved RMSE of 3.68 $m^3$, indicating resistance to large prediction errors and stable supply chain operations.

*Feature Importance Analysis*

Feature importance shows physical mix parameters dominate demand prediction. Cement_kg_m3 contributes approximately 32.81%, while coarse aggregate proportions also show significant influence. Temporal variables such as day of the week exhibit minimal contribution, indicating demand is driven by material and operational characteristics.

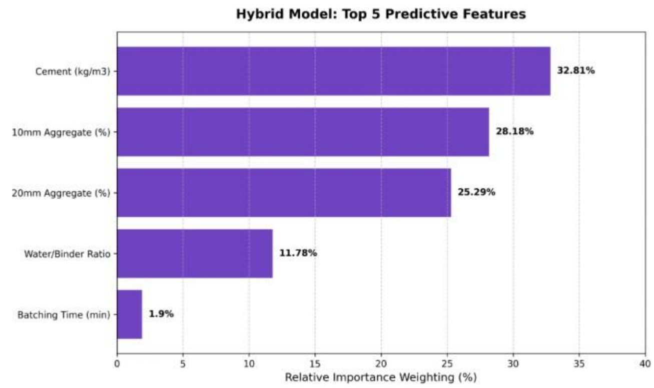


***Fig. 3.*** *Relative Importance Weighting of Dataset Features in Hybrid Demand Prediction*

*Real-World Application Subsystems*

The predictive outputs are integrated into a Node.js-based API and frontend interface with multiple operational modules.

Volumetric Demand and Procurement Dashboard

Provides precise estimation of RMC volume (MAPE: 1.55%), enabling proactive decision-making, better fleet scheduling, and reduced material wastage.


***Fig. 4.*** *Hybrid Model Procurement Interface displaying real-time volumetric demand forecasting.*

Environmental and Carbon Equivalency Tracking View

Translates predicted volume into $CO_2$ emissions using carbon_coeff, enabling environmental assessment, decision support, and sustainability integration.


***Fig. 5.*** *Environmental Tracking Module calculating localized CO2 emissions footprint per RMC dispatch.*

*Algorithmic Input and Operational Parameter Controls*

Allows modification of parameters such as cement density, transport time, and fleet availability, generating real-time forecasts. Supports scenario analysis, adaptability, and flexible decision-making.


***Fig. 6.*** *System operational interface demonstrating parameter input and algorithmic constraints.*

**Discussion**

This study aimed to develop an accurate forecasting framework for Ready-Mix Concrete (RMC) logistics, addressing perishability and environmental concerns. Results from the 146,970-record dataset confirm the effectiveness of the proposed hybrid ensemble model and highlight key insights into model performance and feature relevance.

*Interpretation of Hybrid Model Performance*

The hybrid ensemble model (MAPE: 1.55%) significantly outperforms the LSTM model (MAPE: 145.10%). While LSTM models perform well in stable environments like power systems [6], they are less effective in non-stationary contexts such as RMC logistics. The hybrid approach (XGBoost + Random Forest) captures non-linear relationships without relying on temporal continuity—Random Forest improves robustness for high-variance data, while XGBoost enhances precision—resulting in superior performance, consistent with prior studies [1].

*Interpretation of Feature Importance*

Feature importance analysis shows that material properties dominate demand prediction. Cement content (cement_kg_m3, 32.81%) has the highest influence, followed by coarse aggregates (10 mm and 20 mm). Temporal features like day of the week have minimal impact, indicating demand depends more on material composition and operational factors. These findings extend prior work by Wang et al. [2].

*Implications for Sustainable Construction*

Beyond accuracy, the study contributes to sustainability by integrating demand forecasting with carbon emission estimation. Unlike traditional approaches focused on efficiency [3], this system enables proactive environmental assessment and supports sustainable procurement decisions aligned with green construction practices.

**Conclusion**

This study presents a data-driven framework for forecasting Ready-Mix Concrete (RMC) demand and estimating localized carbon emissions. The proposed system demonstrates the effectiveness of hybrid machine learning models in improving supply chain reliability and supporting sustainable construction practices.

*Summary of Key Findings*

The major findings of this research are summarized as follows: Predictive Performance: Using a longitudinal dataset of 146,970 records spanning five years, the proposed hybrid ensemble model (35% XGBoost and 65% Random Forest) achieved a Mean Absolute Percentage Error (MAPE) of 1.55%, indicating high forecasting accuracy. Model Suitability: The results indicate that sequence-based models such as LSTM (MAPE: 145.10%) are not well-suited for highly non-stationary environments like RMC demand. In contrast, tree-based models effectively capture non-linear relationships and variability in the data. Feature Importance: The analysis reveals that physical mix parameters, particularly cement_kg_m3 and aggregate proportions, play a dominant role in determining demand, while temporal variables contribute comparatively less. Sustainability Integration: The incorporation of a carbon estimation module enables the system to quantify environmental impact prior to procurement, supporting informed and sustainability-oriented decision-making.

*Future Work*

Future research can extend this framework in several directions: Integration with Routing Systems: Demand forecasts can be integrated with real-time routing algorithms to optimize delivery schedules and reduce transportation delays. Real-Time Data Integration: Incorporating IoT-based sensor data (e.g., mixer drum conditions, temperature, and hydration status) can further enhance prediction accuracy and enable real-time monitoring of concrete quality. Expanded Environmental Modeling: Future systems may include comprehensive emission tracking by integrating transportation and operational energy consumption data. In summary, the proposed hybrid machine learning framework provides an effective solution for RMC demand forecasting while incorporating sustainability considerations, offering both practical and environmental benefits for modern construction supply chains.

**Reference**

1. Nandal, S., Sabannawar, S., Pawan, A., Fursule, P., s Depuru, B. K. (2023). Effective Cement Demand Forecasting using Deep Learning Technology: A Data-Driven Approach for Optimal Demand Forecasting. International Journal of Innovative Science and Research Technology, 8(9), 1833-1845.
2. Wang, W., Zhong, Y., Liao, G., Ding, Q., Zhang, T., s Li, X. (2024). Prediction of Compressive Strength of Concrete Specimens Based on Interpretable Machine Learning. Materials, 17(15), 3661. doi: 10.3390/ma17153661.
3. Maghrebi, M., Sammut, C., s Waller, S. T. (2013). Reconstruction of an Expert’s Decision-Making Expertise in Concrete Dispatching by Machine Learning. Journal of Civil Engineering and Architecture, 7(12), 1540-1547.
4. Wang, L., s Chen, W. (2024). Research on AI Forecasting Model of Logistics Demand Driven by Deep Learning. Baidu Online Network Technology/ Hebei Professional College.
5. Tang, T., Li, B., Ma, X., Zhou, Y., Qiu, M., Zhao, W. (n.d.). DBO-LSTM-based Demand-side Resource Load Forecasting Method. North China Electric Power University C China Electric Power Research Institute.
6. Liu, K., Wei, B., Chen, Y., Huang, Z., s Ma, W. (2024). EMD-LSTM Based Demand Forecasting Research on Power Engineering Banded Framework Procurement Materials. Shenzhen Power Supply Co., Ltd, China Southern Power Grid.
7. Pattnaik, M., N.S., V., Sharma, M., Kumar, A., s Sharaschandra, K. S. (2022). A Novel Paradigm to Artificial Intelligence in Transforming Supply Chain Management in the Agile Business World. Princess Nourah Bint Abdulrahman University.