scientific reports

OPEN

A hybrid grey wolf optimized eXtreme gradient boosting-based machine learning model for hospital pharmaceutical demand forecasting

Wilasinee Samniang[1,2], Syed Muhammad Tariq Shah[2,4], Yin May Tun[1], Adeel Munawar[2,3] & Sarin K C[1]

Accurate forecasting of pharmaceutical demand is essential for maintaining the availability of medicines and minimizing waste in hospital supply systems. This study presents a hybrid Grey Wolf Optimized eXtreme Gradient Boosting (GWO–XGBoost) model designed to predict hospital-level medicine demand using real-world dispensing records and meteorological variables. The Grey Wolf Optimizer is applied to select the most informative predictors and fine-tune model parameters, improving the learning efficiency of the eXtreme Gradient Boosting algorithm. Weekly data from two provincial hospitals in Lamphun Province, Thailand were used to evaluate the model's predictive capability. The proposed hybrid model was benchmarked against five machine-learning baseline models and evaluated using three standard performance metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2). By capturing the influence of temporal and environmental factors on medicine utilization, this model supports data-driven hospital planning and more reliable pharmaceutical supply management. The findings highlight the potential of optimization-based machine-learning methods to enhance forecasting performance in healthcare operations.

Keywords Grey wolf optimizer, eXtreme gradient boosting, Hybrid model, Machine learning, Pharmaceutical demand forecasting, Hospital supply chain

Ensuring the continuous availability of essential medicines remains a persistent challenge for healthcare systems worldwide. Stockouts, overstocking, and wastage not only strain hospital budgets but also compromise patient outcomes[1,2]. Accurate forecasting of pharmaceutical demand is therefore vital for strengthening hospital supply chains, supporting efficient procurement planning, and ensuring equitable access to essential medicines.

Traditional forecasting methods in healthcare have often relied on manual estimation or classical statistical models such as moving averages and exponential smoothing[3–5]. While these techniques are easy to implement, they are limited in capturing the nonlinear and dynamic interactions among temporal, clinical, and environmental factors that influence medicine utilization. For instance, seasonal flu can trigger sudden spikes in antiviral demand, while outbreaks such as dengue may rapidly increase the need for antibiotics or supportive medications. As a result, the predictive accuracy of conventional approaches is often inadequate, particularly in resource-constrained settings.

Recent advances in machine learning (ML) have opened new possibilities for improving demand forecasting by modeling nonlinear relationships and exploiting high-dimensional healthcare datasets. ML algorithms have been successfully applied in domains such as patient admissions, disease incidence prediction, and pharmaceutical supply management[6,7]. However, these models often struggle to identify the most relevant

[1]Environmental Economics Unit, Health Intervention and Technology Assessment Program Foundation (HITAP), Ministry of Public Health, Nonthaburi 11000, Thailand. [2]Sirindhorn International Institute of Technology (SIIT), Thammasat University, Pathum Thani 12120, Thailand. [3]Institute of Economics, Karlsruhe Institute of Technology (KIT), Kaiserstraße 12, Karlsruhe 76131, Germany. [4]Department of Civil Engineering, HITEC University, Taxila 47040, Pakistan. email: adeel.munawar@kit.edu

features within complex datasets and may suffer from redundancy or overfitting. To address this challenge, nature-inspired metaheuristic algorithms such as the Grey Wolf Optimizer (GWO)[8] have gained attention for feature selection and hyperparameter tuning, allowing ML models to focus on the most informative predictors and achieve better generalization.

In this study, we propose a Hybrid Grey Wolf Optimized eXtreme Gradient Boosting (GWO–XGBoost) model for forecasting hospital-level pharmaceutical demand. The model integrates hospital dispensing records with high-resolution meteorological and temporal data from the Open-Meteo Archive API[9], covering Pasang and Banthi Hospitals in Lamphun Province, Thailand (October 2023–May 2025). The GWO algorithm is employed to optimize feature subsets and XGBoost parameters simultaneously, enhancing the model's predictive stability and efficiency. By incorporating variables such as temperature, precipitation, and wind speed, the hybrid model captures environmental influences that affect medicine consumption patterns. The proposed GWO–XGBoost model is benchmarked against five machine-learning baseline models and evaluated using three standard performance metrics (MAE, RMSE, and R^2) to provide an early and transparent overview of the comparative evaluation framework.

The key contributions of this study are summarized as follows:

- Proposes a hybrid GWO–XGBoost model for hospital pharmaceutical demand forecasting, integrating meta-heuristic optimization with gradient boosting to enhance predictive accuracy.
- Demonstrates how environmental and temporal features can be leveraged to improve the understanding of medicine utilization dynamics within hospital supply systems.
- Validates the proposed model using real-world weekly data from two provincial hospitals in Thailand, achieving superior performance over traditional statistical and machine learning baselines.

The remainder of this paper is organized as follows. "Literature review" section reviews related work on pharmaceutical demand forecasting. "Methodology" section describes the study setting, data sources, and model development process. "Results and discussion" section presents the experimental results and comparative analysis. "Discussion" section discusses key findings and implications, while "Conclusion" section concludes with major insights and directions for future research.

Literature review

Effective forecasting of pharmaceutical demand has long been a key challenge in health supply chain management. Early approaches to forecasting primarily relied on classical statistical techniques, including time series models such as exponential smoothing and the autoregressive integrated moving average (ARIMA) model. For instance, Sundariyah et al. [10] applied single exponential smoothing to forecast sales and benefits of medicines in a pharmacy case study. Similarly, Rodríguez González et al. [11] utilized ARIMA to estimate medicine demand in a pharmaceutical organization. Similar techniques have been adopted in public health applications, such as Zarghami et al. [12], who used time series modeling to analyze drug-related deaths in Iran, and Bindel and Seifert [13], who employed ARIMA to project long-term antibacterial drug consumption in Germany. Although these approaches are straightforward to implement, they often struggle to capture nonlinearities, external factors, and dynamics specific to hospitals

To overcome these limitations, hybrid models combining statistical techniques with adaptive or probabilistic approaches have been introduced. Huang et al. [14] proposed a hybrid model integrating ARIMA with self-adaptive filtering to improve medical service demand forecasting. In Thailand, Punyapornwithaya et al. [15] demonstrated the use of seasonal time series modeling to predict canine rabies cases, highlighting the applicability of epidemiological forecasting at the national level. Similarly, Lawrence et al. [16] applied a Bayesian network model to analyze vulnerabilities in the U.S. pharmaceutical supply chain following Hurricane Maria, illustrating the role of probabilistic methods in capturing supply risks influenced by external shocks. These hybrid approaches represent an evolution toward more context-aware forecasting but still rely heavily on predefined assumptions.

With the growth of large-scale healthcare datasets, ML has become increasingly important for demand forecasting. Zhu et al. [17] integrated supply-chain features with ML algorithms to enhance pharmaceutical demand forecasting, while Rathipriya et al. [18] developed deep neural networks tailored to pharmaceutical time-series data. Burinskienė [19] applied ML to retail pharmaceutical demand, achieving higher accuracy compared with classical models. Vollmer et al. [20] proposed a unified ML approach for hospital demand forecasting, using ensemble methods to predict emergency department visits. At a broader level, Subramanian [21] systematically reviewed health supply chain forecasting, identifying emerging trends, enablers, and barriers to adoption. More recently, Porto et al. [22] demonstrated how ML models can improve the forecasting of emergency department volumes, highlighting their practical utility for hospital resource allocation. Beyond predictive modeling, optimization techniques have been increasingly integrated into machine learning pipelines to improve feature selection and model efficiency in healthcare applications, contributing to more robust forecasting performance[23,24]. Recent studies have demonstrated an increasing adoption of hybrid machine learning frameworks that integrate optimization techniques to enhance forecasting performance in complex healthcare demand settings. In particular, hybrid optimization techniques have been used to tune deep learning model parameters for time-series prediction tasks, leading to improved predictive accuracy[25]. Recent advances have also explored graph neural network-based temporal models[26] and hypergraph neural networks for capturing higher-order relationships[27], highlighting the potential of these approaches to model complex temporal dependencies in forecasting tasks.

Table 1 summarizes selected studies on forecasting pharmaceutical and health demand, highlighting the diverse methods, contexts, and findings. The literature demonstrates progression from classical statistical approaches to hybrid models, and more recently, to ML methods. While classical models are easy to interpret, they are limited by linear assumptions. Hybrid approaches incorporate adaptive mechanisms but still depend

Study	Method	Context/data	Limitations/strengths
Sundariyah et al.[10]	Exponential smoothing	Pharmacy sales data (Indonesia)	Could not capture seasonality or nonlinear demand; limited applicability in dynamic hospital settings.
Rodríguez González et al.[11]	ARIMA	Pharmaceutical organization (Cuba)	Focused only on one medicine; lacked scalability and external factor integration.
Zarghami et al.[12]	Time series models	Drug-related deaths (Iran)	Epidemiological focus; not directly applicable to pharmaceutical demand forecasting.
Bindel & Seifert[13]	ARIMA	Antibacterial drug consumption (Germany)	Long-term trends captured, but highly sensitive to shocks and external disturbances.
Huang et al.[14]	Hybrid ARIMA + adaptive filter	Medical service demand (China)	Improved accuracy over ARIMA but dependent on parameter tuning; limited generalizability.
Punyapornwithaya et al.[15]	SARIMA	Canine rabies surveillance (Thailand)	Suitable for seasonality but constrained to epidemiological use case.
Lawrence et al.[16]	Bayesian networks	U.S. pharmaceutical supply chain	Captured risk propagation, but scalability issues for large-scale, high-dimensional data.
Zhu et al.[17]	ML with supply chain features	Pharma industry datasets (China)	Required extensive manual feature engineering; lacked integration of clinical/environmental factors.
Rathipriya et al.[18]	Deep/shallow neural networks	Pharmaceutical product sales (small-scale)	High accuracy but limited dataset size; high computational complexity.
Burinskienė[19]	Holt–Winters, Moving Average	Retail pharmaceutical demand (Lithuania)	Retail-focused, limited applicability for hospital-level forecasting.
Vollmer et al.[20]	Ensemble ML	Emergency department visits (UK)	Improved accuracy, but scope restricted to emergency visits, not pharmaceutical supply chains.
Porto et al.[22]	ML ensemble forecasting	Emergency department volumes (USA, Australia, Netherlands)	Effective for hospital visits, but not focused on pharmaceutical demand.
This study	Hybrid GWO–XGBoost model integrating hospital, weather, and temporal features	Weekly data from two Thai provincial hospitals (3.4M+ records)	Combines metaheuristic optimization and gradient boosting to enhance forecasting accuracy and stability for hospital medicine demand.

Table 1. Summary of selected studies on pharmaceutical and health demand forecasting.

on predefined parameters. In contrast, ML methods can capture nonlinear and high-dimensional relationships, consistently delivering superior performance.

Despite significant progress, most existing studies have focused on retail markets, epidemic surveillance, or large-scale supply chains, with limited attention to hospital-level forecasting, especially in low- and middle-income countries. Furthermore, few studies have considered environmental and temporal factors, such as weather variability, which can greatly impact medicine utilization. This research contributes to the field by introducing a GWO–XGBoost model that integrates hospital, meteorological, and temporal features. This approach leverages metaheuristic feature optimization and gradient boosting to improve predictive performance and capture nonlinear relationships in medicine utilization patterns.

Methodology

This study proposes a Hybrid GWO–XGBoost model for hospital-level pharmaceutical demand forecasting. The model integrates real-world hospital dispensing data, standardized drug catalogues, meteorological variables, and temporal features to improve predictive accuracy and model generalization. The overall workflow of the proposed approach is illustrated in Fig. 1, consisting of three main stages: (i) Data Preparation, (ii) Model Optimization and Training, and (iii) Forecast Evaluation. While the conceptual framework highlights possible extensions for hospital decision support, this study focuses primarily on the forecasting and evaluation stages. To assess forecasting performance, the proposed Hybrid GWO–XGBoost model is benchmarked against five baseline machine-learning models and evaluated using three standard performance metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2).

Study setting and data sources

This study utilized real-world operational data from two provincial hospitals:Pasang Hospital and Banthi Hospital in Lamphun Province, Thailand. The dataset spans from October 2023 to May 2025 and includes over 3.4 million records of outpatient visits, pharmacy dispensing events, and associated service items. Of these, approximately 1.4 million entries were successfully linked to valid Drug Identifier (DID) codes, representing 588 distinct medications, yielding a catalogue match rate of 92.7%. The unmatched records primarily consisted of non-medication items such as laboratory tests, consumables, and service charges, which were excluded from the forecasting analysis to ensure focus on pharmaceutical utilization patterns. Table 2 presents a summary of the integrated dispensing dataset used in this study.

Figure 2 presents the distribution of the top 15 dispensed medicines during 2024 across Pasang and Banthi Hospitals, showing substantial variation in consumption levels among different therapeutic categories. This diversity highlights the heterogeneity of medicine demand across time and between hospitals, emphasizing the importance of robust forecasting techniques.

In addition, localized meteorological data were obtained from the Open-Meteo Archive API[9], covering parameters such as temperature, precipitation, humidity, wind speed, and solar radiation. These variables were synchronized with the hospital datasets according to date and geographic location. Temporal predictors,

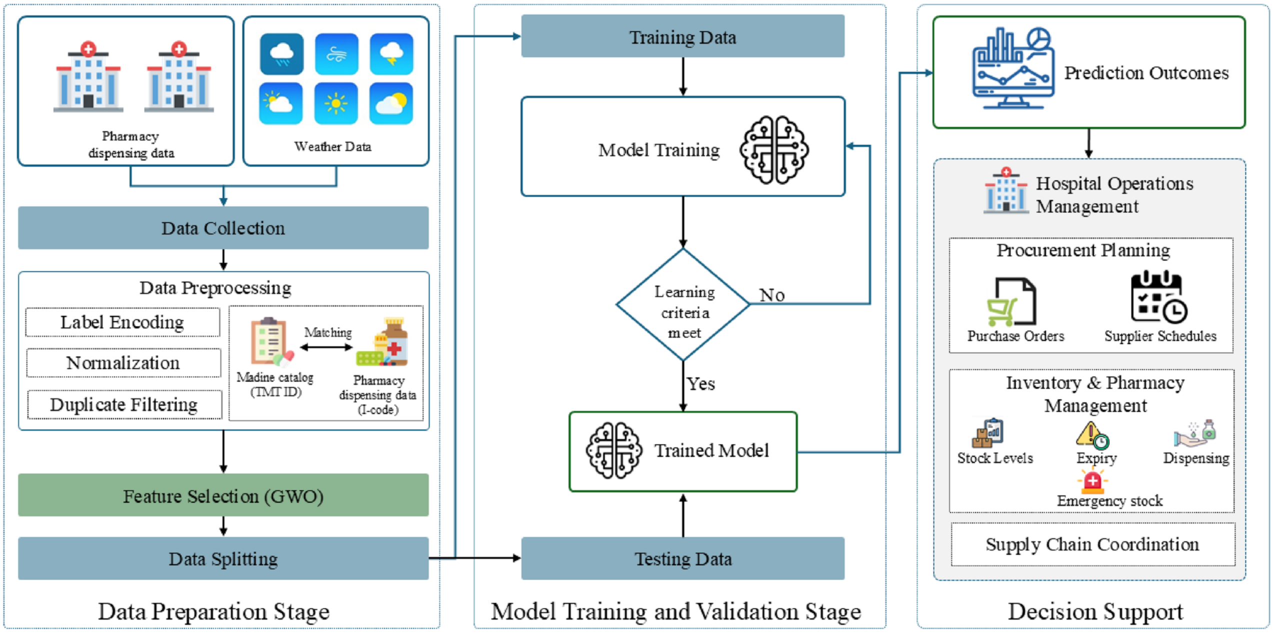


Fig. 1. Overall workflow of the Hybrid GWO–XGBoost model for hospital pharmaceutical demand forecasting.

Dataset information	Count/value
Total dispensing records	3,456,539
Data attributes (fields)	13
Records with valid Drug Identifier (DID)	1,426,720
Records without DID	2,029,819
Unique DID codes	588
Matched records (linked to national drug catalogue)	1,322,093
Unmatched records (with DID but not linked)	104,627
Catalogue match rate (%)	92.67
Observation period	Oct, 2023 – May, 2025

Table 2. Summary of the integrated dispensing dataset used in this study.

including the year, month, and week number, were incorporated to represent seasonality and healthcare utilization patterns. These multi-domain datasets were subsequently used to train and evaluate the Hybrid GWO–XGBoost model, enabling assessment of how environmental and temporal factors contribute to variations in hospital medicine demand. Table 3 summarizes the main predictor domains integrated into the forecasting framework.

Data preparation and preprocessing

The data preparation process standardized, cleaned, and integrated the multi-source datasets prior to model development. Each hospital's internal drug catalogue (I-code) was mapped to Thailand's National Drug Code and the Thai Medicines Terminology (TMT), which is the official standardized terminology for medicines in Thailand, enabling interoperability across hospital information systems. The preprocessing pipeline consisted of the following steps:

- **Cleaning** Removal of duplicate, incomplete, or irrelevant entries (e.g., non-drug service items).
- **Transformation** Log-transformation of weekly demand values, $\log(1 + y)$, to reduce skewness and stabilize variance.
- **Encoding** Categorical variables were one-hot encoded, and continuous variables were standardized to ensure uniform feature scaling.
- **Integration** Weekly aggregation and temporal alignment of dispensing, weather, and time-series variables across hospitals.

Finally, the processed dataset was divided chronologically into training (80%) and testing (20%) subsets to preserve temporal dependencies and reflect real-world forecasting conditions.

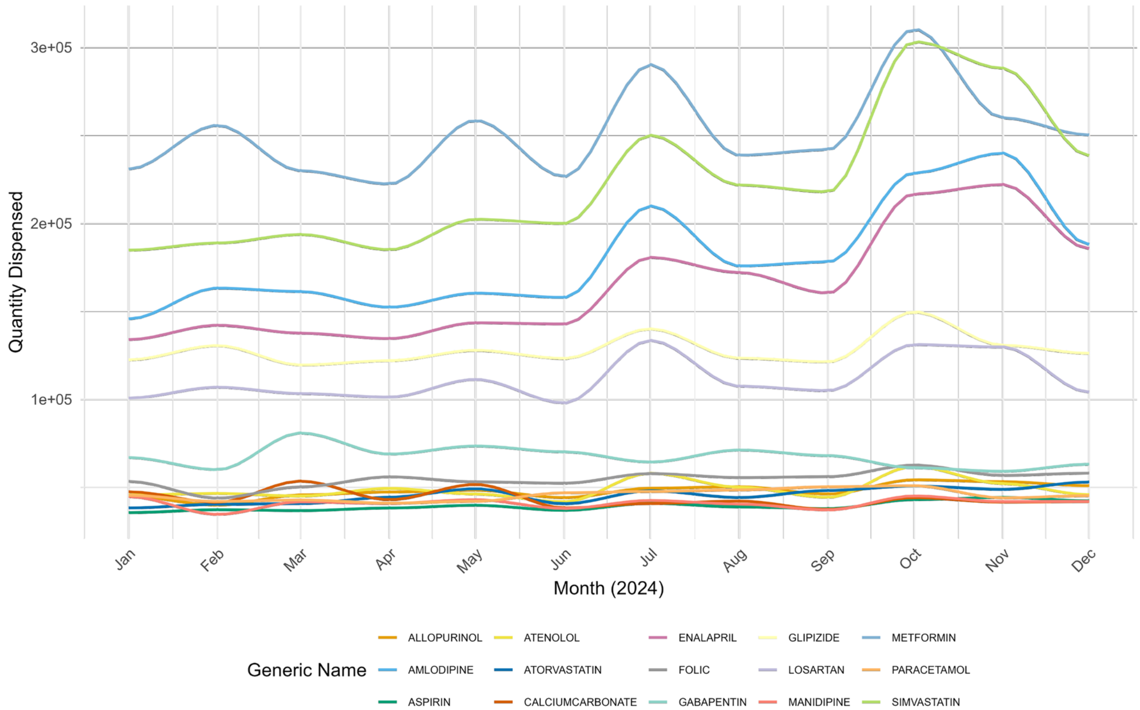


Fig. 2. Monthly demand trends for the top 15 medicines in 2024 across Hospitals.

Domain	Variables
Hospital dispensing	Weekly aggregated demand by medicine and hospital; Drug Identifier (DID); therapeutic class.
Weather (Open-Meteo)	Temperature (min, max, mean), precipitation, rainfall, precipitation hours, humidity, wind speed, wind gusts, evapotranspiration, solar radiation.
Temporal	Year, month, week number.

Table 3. Predictor domains integrated into the forecasting model.

Feature selection using grey wolf optimizer (GWO)

Feature selection plays a crucial role in constructing efficient and interpretable predictive models, particularly for complex healthcare datasets that combine hospital dispensing, meteorological, and temporal variables. Selecting redundant or irrelevant predictors can reduce model accuracy and increase computational cost. To address this challenge, this study employs the GWO[8], a nature-inspired metaheuristic algorithm modeled on the leadership hierarchy and cooperative hunting behavior of grey wolves. GWO has demonstrated strong convergence stability and a good balance between exploration and exploitation, making it suitable for feature selection in nonlinear, high-dimensional healthcare data[28,29].

In the proposed Hybrid GWO–XGBoost model, each wolf search agent represents a candidate binary vector of features, where "1" denotes inclusion and "0" denotes exclusion. The population is initialized randomly and updated iteratively according to the top three wolves–α (best), β (second-best), and δ (third-best)–which guide the others toward the optimal subset.

The encircling mechanism of prey is mathematically expressed as:

$$\vec{D}_\alpha = |\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}|, \quad \vec{D}_\beta = |\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}|, \quad \vec{D}_\delta = |\vec{C}_3 \cdot \vec{X}_\delta - \vec{X}|, \tag{1}$$

where $\vec{X}$ is the position vector of a wolf (feature subset), and $\vec{X}_\alpha$, $\vec{X}_\beta$, and $\vec{X}_\delta$ are the positions of the leading wolves. The coefficient vectors $\vec{C}_1$, $\vec{C}_2$, and $\vec{C}_3$ introduce stochastic variations in movement.

The hunting process is defined as:

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 \cdot \vec{D}_\alpha, \quad \vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot \vec{D}_\beta, \quad \vec{X}_3 = \vec{X}_\delta - \vec{A}_3 \cdot \vec{D}_\delta, \tag{2}$$

and the wolves update their positions as:

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3}. \tag{3}$$

The coefficients $\vec{A}$ and $\vec{C}$ controlling exploration and exploitation are defined as:

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a}, \quad \vec{C} = 2\vec{r}_2, \tag{4}$$

where $\vec{r}_1$ and $\vec{r}_2$ are random vectors in $[0, 1]$, and $\vec{a}$ decreases linearly from 2 to 0 over the iterations to gradually shift focus from exploration to exploitation.

To perform binary feature selection, the continuous values of $\vec{X}$ are converted into binary inclusion decisions using a sigmoid transfer function:

$$F(x) = \begin{cases} 1, & \text{if } \sigma(x) \geq \gamma, \\ 0, & \text{otherwise,} \end{cases} \tag{5}$$

where $\sigma(x)$ is the sigmoid activation and γ is a random threshold within [0, 1].

The fitness function guiding the optimization is defined as:

$$F = w_1(1 - R^2) + w_2 \frac{|S|}{|S_{total}|}, \tag{6}$$

where R^2 is the coefficient of determination from the XGBoost predictions, $|S|$ is the number of selected features, and $|S_{total}|$ is the total number of available predictors. The weights w_1 and w_2 (set to 0.9 and 0.1, respectively) balance forecasting accuracy and model simplicity. Through iterative updates, GWO searches for the subset of features that minimizes F, retaining only the most informative predictors influencing weekly pharmaceutical demand. The optimized subset is then used for XGBoost model training, forming the complete forecasting pipeline.

Model training and validation

Following the feature selection stage, the optimized subset of predictors identified by the GWO was used to train the forecasting model based on the eXtreme Gradient Boosting (XGBoost) algorithm[30]. XGBoost is a scalable and regularized gradient boosting framework designed to enhance predictive accuracy and prevent overfitting. It is particularly effective for structured healthcare datasets because of its ability to model nonlinear dependencies, handle missing values, and integrate heterogeneous variables.

In the training phase, XGBoost constructs an additive ensemble of regression trees, where each new tree iteratively learns to correct the residuals of previous trees. Given a training dataset $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{n}$, where $x_i \in \mathbb{R}^m$ represents the selected input features (hospital dispensing, meteorological, and temporal variables) and $y_i \in \mathbb{R}$ denotes the weekly medicine demand, the model prediction $\hat{y}_i$ is expressed as:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), \quad f_k \in \mathcal{F}, \tag{7}$$

where $\mathcal{F}$ denotes the functional space of regression trees and K represents the total number of boosting rounds.

The optimization objective combines a convex loss function $\ell(y_i, \hat{y}_i)$, which measures prediction error, with a regularization term $\Omega(f_k)$ that penalizes overly complex trees, thereby improving generalization[30]:

$$\mathcal{L}(\phi) = \sum_{i=1}^{n} \ell(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k), \tag{8}$$

where the regularization term is defined as:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \|w\|^2, \tag{9}$$

with T being the number of leaves in the tree, w the leaf weights, γ the complexity penalty, and λ the L2 regularization parameter. The first term of the objective function enforces predictive fidelity, while the second controls model complexity, effectively mitigating overfitting.

The loss function used in this study was the Mean Squared Error (MSE):

$$\ell(y_i, \hat{y}_i) = (y_i - \hat{y}_i)^2, \tag{10}$$

which measures the squared difference between observed and predicted demand values. The final model was trained with tuned hyperparameters, including a learning rate of 0.05, a maximum tree depth of 15, 500 boosting rounds, and regularization parameters $\lambda = 1.5$ and $\gamma = 0.2$.

Decision support integration

The final stage conceptually links the forecasting outcomes to hospital operations intelligence (Fig. 1). Forecasted pharmaceutical demand can be translated into actionable insights across key management domains, supporting proactive and data-driven decision-making:

- **Procurement planning** Supports adjustment of purchase orders and supplier schedules to prevent medicine shortages or overstock situations.
- **Inventory and pharmacy management** Enables dynamic stock-level optimization, expiry monitoring, and preparedness for emergency demand surges.
- **Supply chain coordination** Facilitates improved coordination with regional warehouses and the Government Pharmaceutical Organization (GPO) for timely and efficient distribution.

Although the implementation of a decision-support system is beyond the scope of this paper, the proposed Hybrid GWO–XGBoost forecasting framework provides a foundational analytical layer that can be extended to hospital management dashboards. This integration has the potential to strengthen forecasting precision, optimize resource allocation, and enhance the sustainability and resilience of healthcare supply chains.

Baseline methods

To benchmark the performance of the proposed model, several widely used machine learning (ML) algorithms were implemented as baselines. These models were selected for their strong performance in structured data forecasting and healthcare analytics applications.

Decision tree (DT)

Decision Trees are non-parametric models that recursively partition the input space into homogeneous regions[31]. For regression, the prediction in a terminal node R_m is given by:

$$\hat{y}(x) = \frac{1}{N_m} \sum_{i \in R_m} y_i, \tag{11}$$

where N_m is the number of training samples in R_m. In this study, the DT model was tuned using cross-validation, with a maximum depth of 15, yielding the best validation performance.

Random forest (RF)

Random Forests extend decision trees by constructing an ensemble of trees using bootstrap aggregation (bagging) and random feature selection[32]. The prediction is obtained by averaging across M trees:

$$\hat{y}(x) = \frac{1}{M} \sum_{m=1}^{M} f_m(x). \tag{12}$$

Hyperparameters were optimized using grid search with cross-validation. The optimal configuration employed 500 trees, a maximum depth of 15, and $\sqrt{p}$ feature selection at each split, where p is the number of predictors.

Light gradient boosting machine (LightGBM)

LightGBM is a gradient boosting framework optimized for efficiency and scalability[33]. It uses a histogram-based algorithm and grows trees leaf-wise, selecting the split with the maximum loss reduction. At iteration t, the prediction is expressed as:

$$\hat{y}_t(x) = \hat{y}_{t-1}(x) + \eta f_t(x), \quad f_t \in \mathcal{F}, \tag{13}$$

where η is the learning rate, f_t is the regression tree fitted to the gradient of the loss function at iteration t, and $\mathcal{F}$ denotes the space of regression trees. The final tuned configuration used a learning rate of 0.05, 500 boosting rounds, a maximum depth of 15, and 64 leaves.

CatBoost

CatBoost is a gradient boosting method designed to reduce prediction shift and efficiently handle categorical features[34]. Unlike traditional boosting, it employs *ordered boosting*, which builds each new tree using target statistics computed on a permutation of the training data. Formally, the prediction at iteration t is given by:

$$\hat{y}_t(x_i) = \hat{y}_{t-1}(x_i) + \eta f_t(x_i \mid \sigma_{<i}), \tag{14}$$

where $f_t(x_i \mid \sigma_{<i})$ denotes the tree trained on permutations $\sigma_{<i}$ that exclude the current sample i, ensuring unbiased estimates of categorical feature statistics. This mechanism mitigates overfitting and enhances generalization. The tuned configuration used a learning rate of 0.05, 500 iterations, and a maximum tree depth of 15.

Artificial neural network (ANN)

Artificial Neural Networks (ANNs) approximate complex nonlinear relationships through layers of interconnected neurons[35]. They are commonly used in healthcare forecasting due to their ability to model temporal dynamics and nonlinear dependencies in demand data. Each hidden neuron computes a weighted sum of its inputs followed by a nonlinear activation. The Rectified Linear Unit (ReLU) activation was used in hidden layers, while the output layer used a linear activation suitable for regression tasks.

For an input vector $x \in \mathbb{R}^m$, the transformation at hidden layer l is:

$$h^{(l)} = \max\left(0,\, W^{(l)} h^{(l-1)} + b^{(l)}\right), \quad l = 1, \ldots, L, \tag{15}$$

where $W^{(l)}$ and $b^{(l)}$ denote the weight matrix and bias vector of layer l. The final prediction is obtained as:

$$\hat{y} = W^{(o)} h^{(L)} + b^{(o)}, \tag{16}$$

where $W^{(o)}$ and $b^{(o)}$ are the parameters of the output layer.

Parameter optimization used the Adam optimizer[36], which combines momentum and adaptive learning rates. The best ANN architecture contained three hidden layers with 128, 64, and 32 neurons, respectively, each using ReLU activation. Training employed a batch size of 64, a learning rate of 0.001, and 500 epochs.

All baseline models were tuned through grid search with cross-validation, and the configuration achieving the lowest validation error was selected for performance comparison against the proposed model.

Evaluation metrics

Model performance was evaluated using three standard metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2). These measures jointly assess predictive accuracy, sensitivity to large deviations, and explanatory power.

The Mean Absolute Error (MAE) quantifies the average magnitude of prediction errors, providing an intuitive measure of model accuracy in the same units as the target variable. The Root Mean Squared Error (RMSE) penalizes larger errors more strongly, emphasizing the model's ability to handle extreme variations in demand. The coefficient of determination (R^2) indicates the proportion of variance in observed demand explained by the model, reflecting its overall goodness of fit.

The three evaluation metrics are formally defined in Equations 17 to 19[37,38]:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{17}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{18}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{19}$$

Here, y_i and $\hat{y}_i$ denote the observed and predicted weekly demand, respectively, and $\bar{y}$ represents the mean observed demand. Together, these complementary metrics provide a balanced evaluation of model accuracy, error dispersion, and explanatory strength in pharmaceutical demand forecasting.

Code availability

The code used to develop and evaluate the proposed model is available at DOI: 10.5281/zenodo.19387761.

Results and discussion

This section presents the experimental results of the Hybrid GWO–XGBoost model compared with five baseline GWO–optimized machine learning models: DT, RF, LightGBM, CatBoost, and ANN. All models were trained and evaluated using the same dataset and three performance indicators MAE, RMSE, and the R^2 to ensure fair comparison and consistency.

Model performance comparison

Table 4 summarizes the comparative performance across all models. The proposed model achieved the best overall performance with the lowest MAE (0.81), lowest RMSE (2.65), and the highest R^2 (0.984), demonstrating its superior predictive capability and generalization compared with other GWO–based learners. Among the baselines, GWO–LightGBM and GWO–ANN also performed competitively with R^2 values above 0.96, whereas GWO–RF exhibited the weakest performance due to overfitting and poor residual convergence.

Model	MAE	RMSE	R^2
GWO–DT	1.52	4.38	0.957
GWO–RF	3.76	11.13	0.721
GWO–LightGBM	1.02	3.21	0.977
GWO–CatBoost	1.55	4.23	0.960
GWO–ANN	1.15	3.73	0.969
Proposed model	0.81	2.65	0.984

Table 4. Comparative performance of GWO–optimized models and the proposed model.

To assess evaluation stability, the proposed model was additionally tested under multiple train–test split ratios and random seeds while maintaining identical feature engineering and hyperparameter settings. As shown in Table 5, performance metrics exhibit low variance across configurations, indicating that the model's predictive accuracy and generalization are robust to changes in data partitioning and not dependent on a specific training split.

Regression and fit analysis
Figure 3a–f illustrate the regression performance of each model. Each scatter plot compares the predicted versus actual weekly medicine demand, with color intensity representing the magnitude of prediction error. A dashed diagonal line indicates the perfect fit.

The scatter plots demonstrate a consistent clustering of predictions along the perfect-fit line, indicating strong agreement between observed and predicted values. The proposed model shows the tightest clustering and smallest dispersion, confirming its robustness and higher precision in capturing nonlinear relationships between temporal, meteorological, and dispensing variables.

Error and residual distribution analysis
To further examine model consistency, Fig. 4 presents the absolute error distributions across models. The proposed model achieved the narrowest error spread (MAE = 0.81), indicating both accuracy and stability across different medicines and demand magnitudes.

Figure 5 shows the residual density curves, highlighting that the proposed model yields the sharpest, most symmetric distribution centered around zero reflecting unbiased predictions and minimal variance.

Forecast trend comparison
To further assess temporal forecasting reliability, representative medicines were randomly selected from the testing dataset for trend visualization. Figure 6 illustrates the predicted versus actual weekly demand across all GWO–based models. The proposed model demonstrates superior temporal tracking performance, accurately capturing both peak and trough dynamics with minimal lag and smoother transitions.

The proposed model consistently generalizes across diverse medicine types, exhibiting stable temporal responses and reduced deviation under varying demand fluctuations, confirming its robustness in real-world hospital forecasting scenarios.

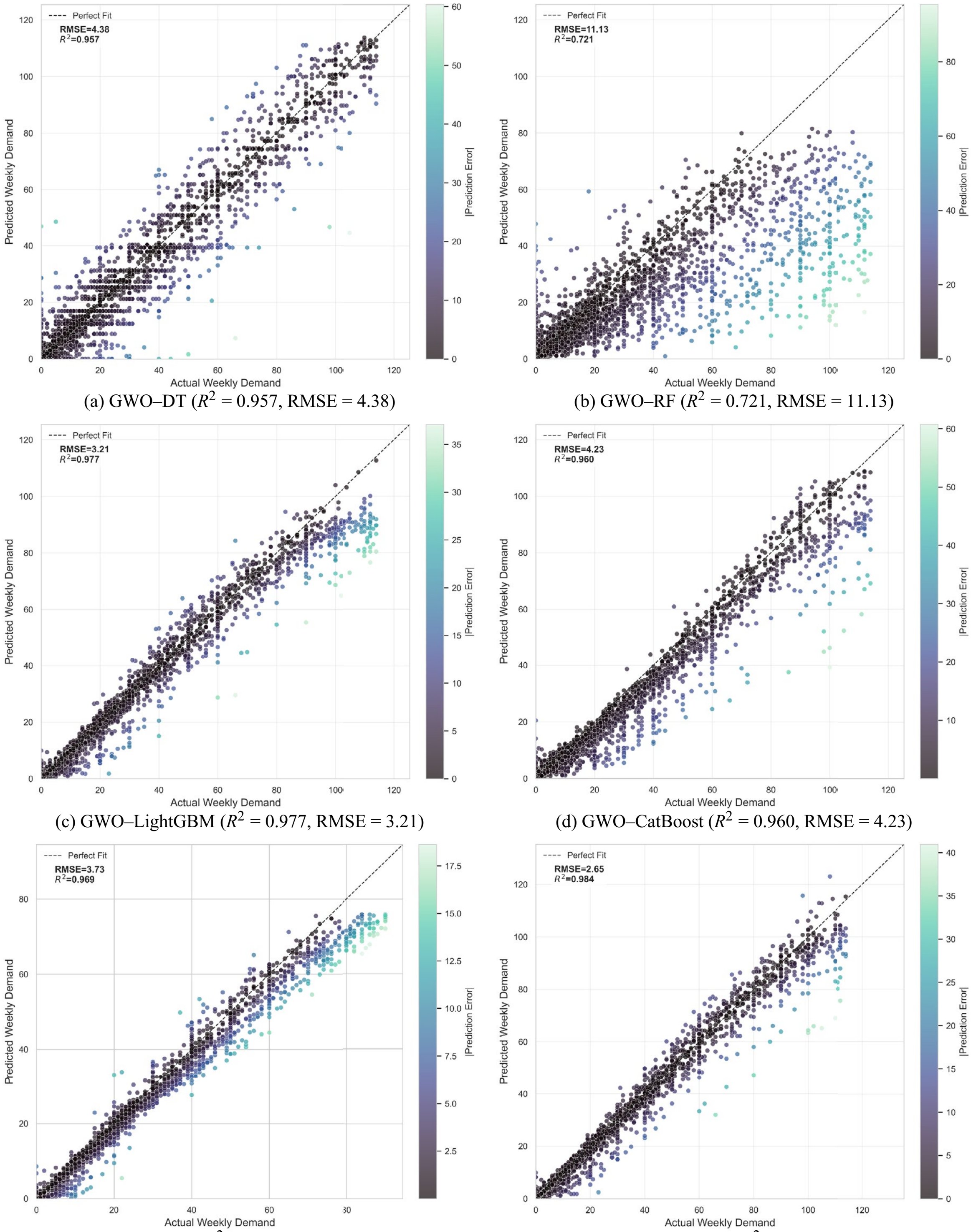
Comparison with existing literature
To contextualize the performance of the proposed model, Fig. 7 compares its coefficient of determination (R^2) with values reported in recent studies on pharmaceutical demand and drug response forecasting. Prior research has explored diverse modeling approaches, including Linear Regression [39], LSTM [40], Random Forest [41], AFT–LSTM [42], and LightGBM [43]. These models reported R^2 values ranging between 0.76 and 0.96, indicating substantial progress in predictive analytics for medicine demand. In comparison, the proposed Hybrid GWO–XGBoost model achieved an R^2 of 0.984, outperforming all previously published benchmarks. This improvement underscores the advantage of integrating metaheuristic feature optimization (GWO) with gradient-boosted ensembles (XGBoost) to enhance generalization, stability, and predictive accuracy in real-world hospital forecasting applications.

In addition to predictive performance, model efficiency is a critical factor for real-world deployment in hospital decision-support systems. Figure 8 compares the average prediction time (in seconds) for 500 test samples across all GWO–based learners and the proposed Hybrid GWO–XGBoost model. Despite its enhanced accuracy, the proposed model maintained competitive computational efficiency, achieving an average prediction time of only 0.009 s faster than most ensemble baselines, including GWO–RF (0.265 s) and GWO–ANN (0.269 s). This demonstrates that the hybrid integration of GWO and XGBoost not only improves predictive reliability but also preserves scalability and responsiveness, making it suitable for operational deployment in hospital information systems. Notably, the Grey Wolf Optimization is applied during model training, whereas the computational times reported here correspond to the prediction stage of the trained models.

Discussion
The comparative analysis demonstrates that integrating the GWO with XGBoost substantially enhances forecasting accuracy and generalization capability. The GWO component effectively identified the most informative predictors from multidomain datasets, improving model interpretability and reducing feature redundancy, while XGBoost captured nonlinear dependencies with robust regularization. Among all GWO–based learners, the Hybrid GWO–XGBoost model consistently achieved the lowest error and highest R^2, confirming its reliability for real-world hospital-level pharmaceutical forecasting.

Test split	MAE (mean ± SD)	RMSE (mean ± SD)	R^2 (mean ± SD)
10%	0.73 ± 0.03	2.36 ± 0.20	0.987 ± 0.002
20%	0.78 ± 0.03	2.54 ± 0.12	0.985 ± 0.001
30%	0.83 ± 0.01	2.70 ± 0.04	0.983 ± 0.000

Table 5. Robustness analysis of the proposed Hybrid GWO–XGBoost model across different train–test splits.



(a) GWO–DT (R^2 = 0.957, RMSE = 4.38)
(b) GWO–RF (R^2 = 0.721, RMSE = 11.13)
(c) GWO–LightGBM (R^2 = 0.977, RMSE = 3.21)
(d) GWO–CatBoost (R^2 = 0.960, RMSE = 4.23)
(e) GWO–ANN (R^2 = 0.969, RMSE = 3.73)
(f) Hybrid GWO–XGBoost (R^2 = 0.984, RMSE = 2.65)

Fig. 3. Regression performance comparison of models.

From an operational standpoint, these improvements translate into stronger anticipation of medicine demand fluctuations and more efficient resource allocation. Even moderate reductions in forecasting error can yield significant managerial benefits including optimized purchase scheduling, reduced emergency procurement, and lower wastage from expired stock. By incorporating temporal and environmental variables, the model provides adaptive insights that support proactive procurement strategies aligned with seasonal patterns and external drivers.

The integration of GWO-driven feature selection was central to these outcomes. By systematically identifying the most relevant predictors, the model achieved better generalization across diverse medicines and hospital

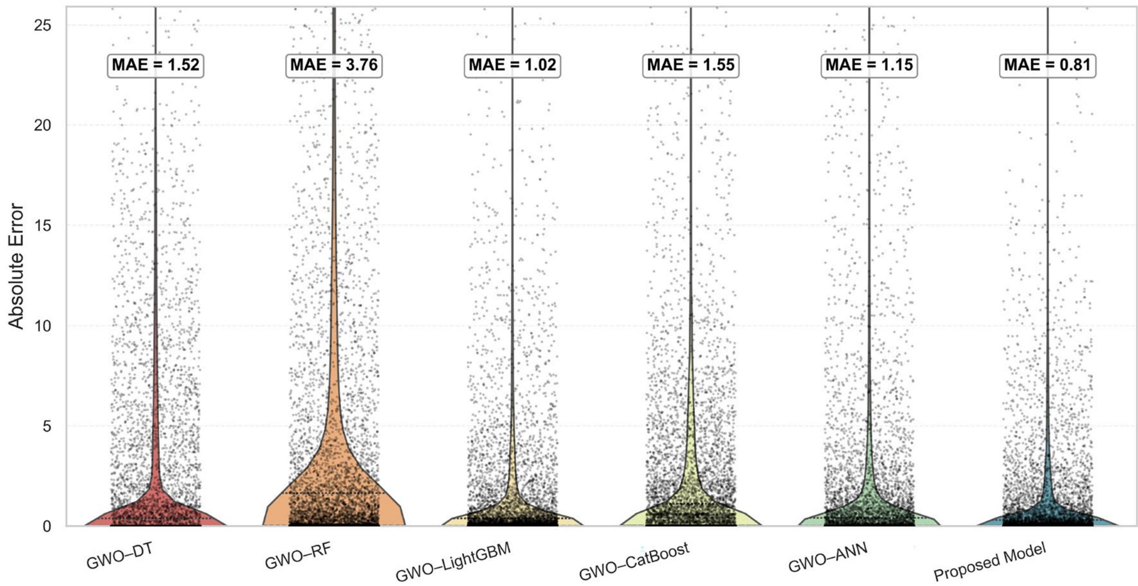


Fig. 4. Comparison of absolute error distributions across all models.

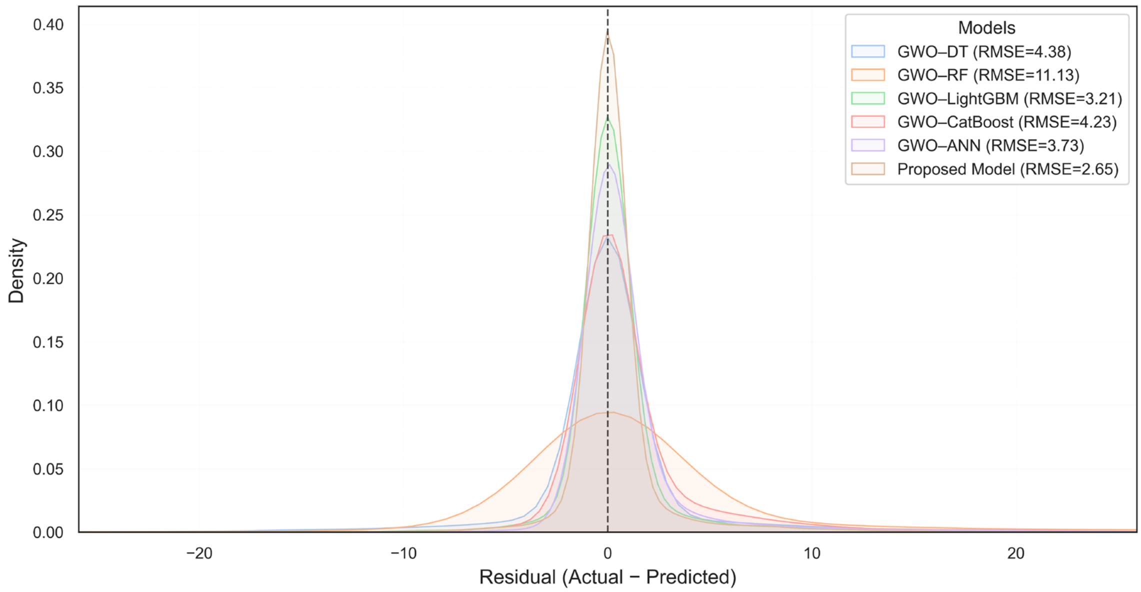


Fig. 5. "Residual distribution comparison among GWO–optimized models and the proposed model".

types, demonstrating robustness under variable demand conditions. This interpretable optimization framework enhances confidence in the adoption of machine learning for healthcare supply chain management. Recent studies indicate that tree-based methods, such as gradient boosting, often remain highly competitive and can outperform deep learning architectures when applied to structured or tabular datasets, particularly in healthcare and operational forecasting contexts. Grinsztajn et al. [44] demonstrated that tree-based models frequently achieve similar or superior predictive performance to deep neural networks on tabular data, while also offering advantages in interpretability and computational efficiency. Consistent with these observations, recent healthcare demand forecasting studies have reported strong performance of gradient boosting approaches, including XGBoost, in hospital and emergency department prediction tasks [22,45]. These results support the suitability of the proposed GWO–XGBoost framework for hospital pharmaceutical demand forecasting using structured dispensing and environmental data. Future research could explore transformer-based and attention-driven architectures to further evaluate their effectiveness in modeling complex temporal dependencies in pharmaceutical demand forecasting.

At the policy level, the proposed model supports Thailand's *Drug System Development Action Plan* [46], fostering data-driven planning and centralized coordination between hospitals and distribution networks

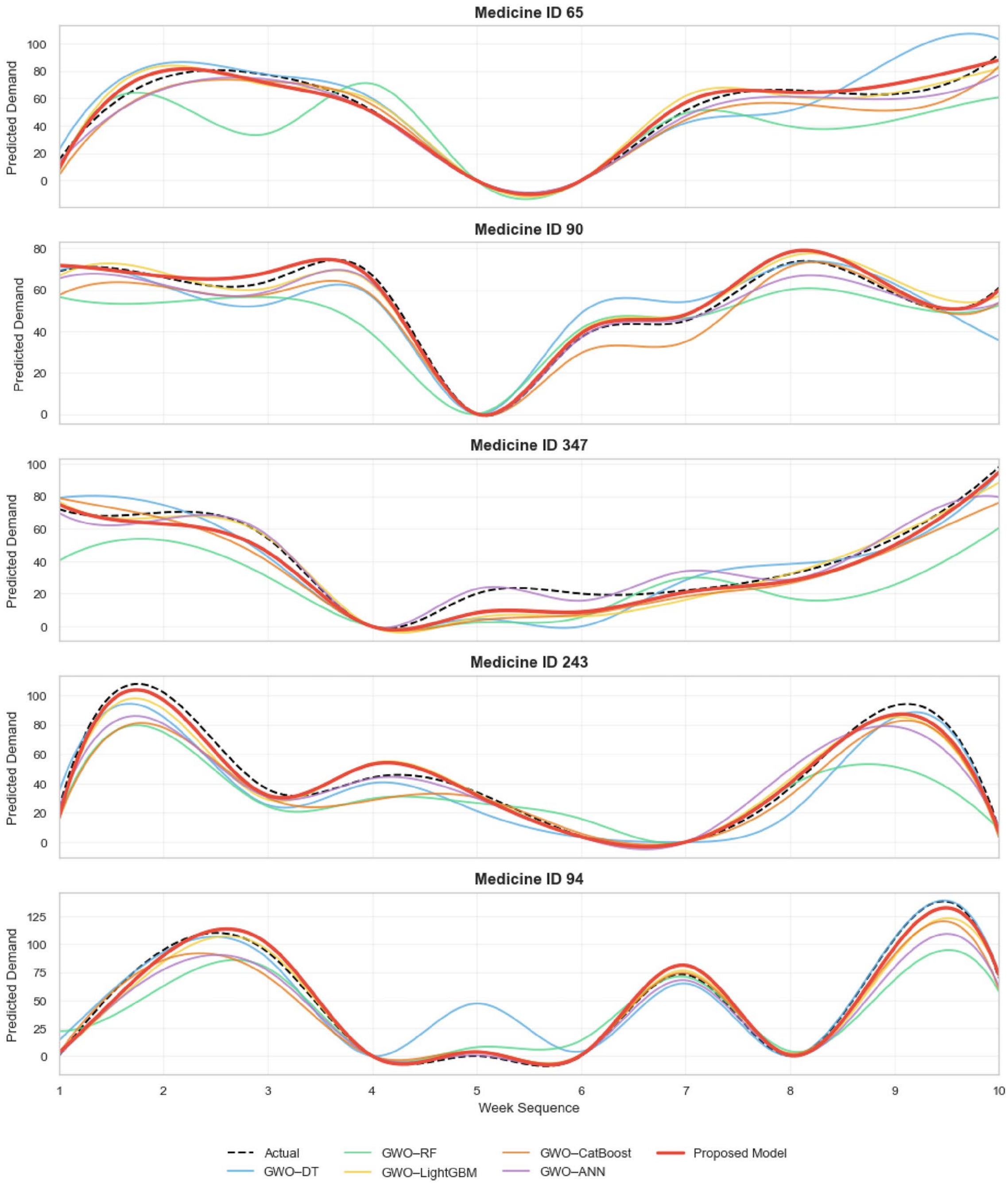


Fig. 6. Comparison of actual and predicted weekly demand trends for medicines across all models.

such as the Government Pharmaceutical Organization (GPO). Embedding predictive analytics within health information systems can strengthen supply security, transparency, and sustainability, directly contributing to national objectives for digital health transformation and efficient resource utilization.

From an ethical and governance perspective, forecasting errors in pharmaceutical demand may have direct implications for hospital procurement decisions, potentially contributing to medicine shortages or excess stock if not appropriately managed. For this reason, forecasting outputs should be used as decision-support tools rather than automated decision-makers, with human oversight retained at the procurement and inventory planning stages. Transparency and interpretability are therefore important to ensure that hospital managers can understand and trust model behavior. In addition, deployment of predictive models in public health environments requires appropriate risk management strategies, including routine performance monitoring and periodic retraining to prevent model degradation. The use of de-identified data and adherence to institutional data governance and approval processes are essential to ensure responsible use of analytics within hospital information systems. From an operational perspective, the proposed forecasting framework should be regarded as a proof-of-concept. In practical settings, forecast outputs could be incorporated into routine pharmacy and procurement planning as

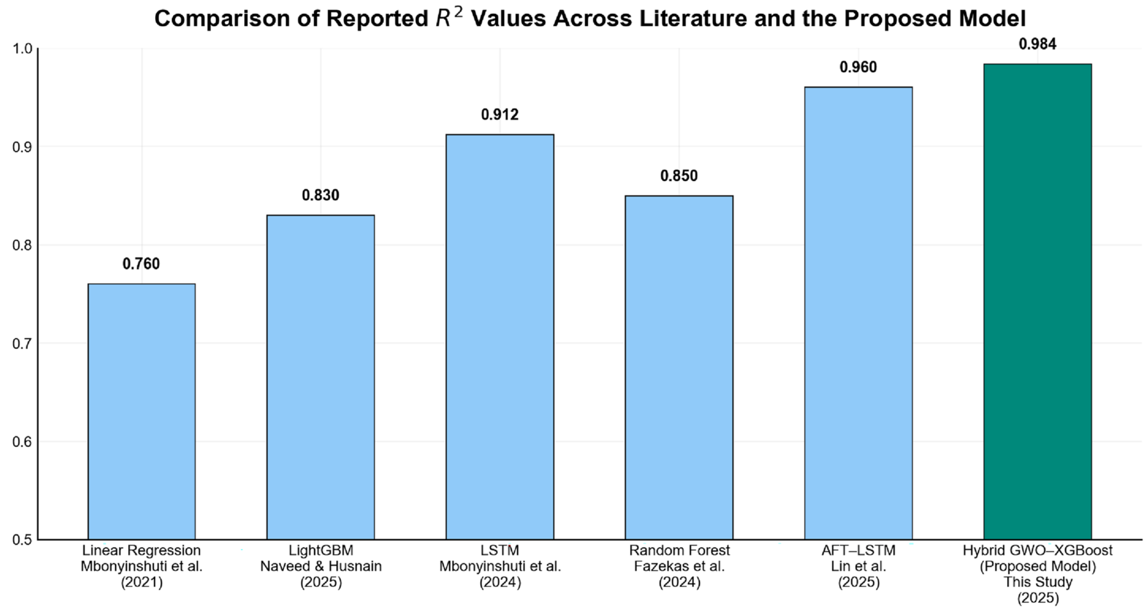


Fig. 7. Comparison of reported R^2 values across literature and the proposed Hybrid GWO–XGBoost model. The proposed model demonstrates superior explanatory performance compared with previous studies [39–43].

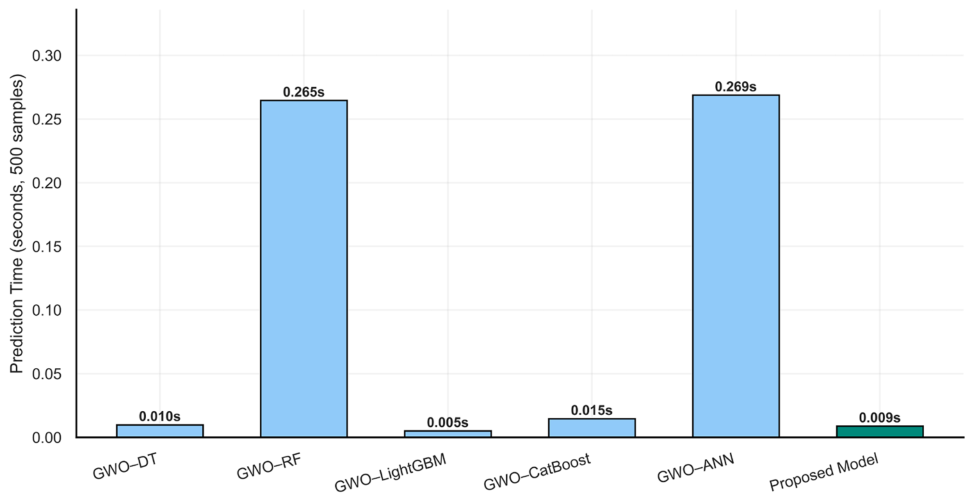


Fig. 8. Prediction time comparison across all models for 500 test samples.

decision-support inputs rather than automated decision rules. Broader deployment across hospital networks would require additional validation and coordination with existing hospital information systems.

Beyond methodological suitability, several practical considerations must be addressed before adoption and scaling. These include investment in adequate computational infrastructure at healthcare facilities, integration with existing hospital information systems to ensure interoperability, continuous input of up-to-date procurement and inventory data, and capacity building for healthcare staff to operate the system and interpret forecasting outputs. Addressing these operational requirements is essential to enable sustainable deployment and effective use of the proposed framework in real-world hospital settings. While the GWO-based framework effectively identifies an optimal subset of predictors and enhances forecasting performance, additional analysis using explainable AI techniques (e.g., SHAP) and complementary ablation studies could further strengthen model interpretability and provide deeper insight into the contributions of individual variables and model components. Future work will explore these directions to advance understanding of the influence of meteorological and temporal drivers on pharmaceutical demand patterns.

However, this study is based on data from two provincial hospitals, which may not fully capture the variability of pharmaceutical demand across different hospital tiers, regions, and procurement systems in Thailand. Accordingly, the findings should be interpreted as a proof-of-concept demonstration. Future research should expand data coverage across multiple regions and hospital tiers, subject to data-sharing and governance approvals, to further assess scalability, transferability, and broader integration potential. Such expansion would support the development of a more comprehensive and adaptive forecasting framework for sustainable and resilient healthcare logistics.

Conclusion

This study proposed and evaluated a Hybrid GWO–XGBoost model for forecasting hospital-level pharmaceutical demand by integrating dispensing records, temporal indicators, and localized meteorological data. Using real-world operational datasets from two provincial hospitals in Lamphun Province, Thailand, the proposed model achieved the strongest predictive performance among all tested approaches, yielding the lowest MAE and RMSE and the highest R^2. By effectively capturing nonlinear dependencies and environmental influences, the model demonstrated the advantages of combining metaheuristic optimization with gradient-boosted ensemble learning for complex healthcare forecasting tasks.

Comparative analyses confirmed that while ensemble models such as Random Forest, LightGBM, and CatBoost achieved competitive accuracy, the Hybrid GWO-XGBoost model consistently outperformed them across all evaluation metrics. Regression and residual analyses further validated its robustness and unbiased predictive behavior, with predictions closely aligned to observed demand patterns. These results highlight the contribution of GWO-driven feature optimization in enhancing generalization, interpretability, and stability within heterogeneous hospital datasets.

Beyond predictive performance, the model serves as a practical decision-support tool for hospital procurement and inventory management. By linking demand forecasts with temporal and meteorological variations, it enables proactive purchasing, minimizes emergency orders, and supports evidence-based strategies to reduce wastage and prevent stockouts. At the policy level, such forecasting systems have the potential to support Thailand's pharmaceutical supply chain management by improving supply security, operational efficiency, and sustainability aligning with the national *Drug System Development Action Plan* [46].

Despite these promising results, validation was limited to two hospitals over a relatively short observation period. Future research should expand the dataset to encompass diverse healthcare facilities and geographic regions, incorporate additional demand drivers such as disease incidence, demographic dynamics, public health interventions, and policy changes, and explore advanced deep learning architectures (e.g., recurrent and graph neural networks) to better capture spatiotemporal dependencies in pharmaceutical demand. Future work will prioritize expanding institutional participation and conducting external validation using hospitals not involved in model development, as data-sharing frameworks and governance mechanisms continue to mature. Integrating the model into hospital information systems and national supply chain platforms would further enhance its scalability, real-time applicability, and long-term sustainability.

Data availability

The datasets used in this study contain confidential hospital information and are not publicly available. However, de-identified or aggregated data can be provided upon reasonable request and with permission from the Ministry of Public Health, Thailand.

Received: 8 November 2025; Accepted: 8 April 2026
Published online: 24 April 2026

References

1. World Health Organization. *Medicine Supply in Health Systems* (World Health Organization, Geneva, 2010).
2. Yadav, Pankaj. Health product supply chains in developing countries: Diagnosis of the root causes of underperformance and an agenda for reform. *Health Systems & Reform* **1**(2), 142–154 (2015).
3. Ireneous, N., Soyiri, & Daniel, D. Reidpath. An overview of health forecasting. *Environmental health and preventive medicine*, 18(1):1–9, 2013.
4. Houfani, Djihane, Slatnia, Sihem, Kazar, Okba, Saouli, Hamza & Merizig, Abdelhak. Artificial intelligence in healthcare: a review on predicting clinical needs. *International Journal of Healthcare Management* **15**(3), 267–275 (2022).
5. Scott, J., Armstrong, Evaluating forecasting methods. In *Principles of forecasting: A handbook for researchers and practitioners*, pages 443–472. Springer, 2001.
6. Rajkomar, Alvin, Dean, Jeff & Kohane, Isaac. Machine learning in medicine. *New England Journal of Medicine* **380**(14), 1347–1358 (2019).
7. Rong, G., Mendez, A., Assi, E.B., Zhao, B., & Sawan, M. Artificial intelligence in healthcare: review and prediction case studies. *Engineering*, 6(3):291–301, 2020.
8. Mirjalili, Seyedali, Mirjalili, Seyed Mohammad & Lewis, Andrew. Grey wolf optimizer. *Advances in Engineering Software* **69**, 46–61 (2014).
9. Open-Meteo. Open-meteo: Free weather api for non-commercial use. https://open-meteo.com/, 2025. Accessed via Open-Meteo Archive API.
10. Sundariyah, R. D. & Adityo, and A. Arizal,. Forecasting the total sales and benefits of drug using the single exponential smoothing method (case study: Bentar pharmacy). *Journal of Electrical Engineering and Computer Sciences (JEECS)* **4**(2), 687–694 (2019).
11. González, Ramón Rodríguez., González, Jorge Luis León. & Sánchez, Yanelys Álvarez. Forecast of the demand for medications by a pharmaceutical organization using the arima model. *Universidad y Sociedad* **13**(1), 119–130 (2021).
12. Zarghami, M. et al. Time series modeling and forecasting of drug-related deaths in iran (2014–2016). *Addiction and Health* **15**(3), 149–155 (2023).
13. Bindel, L. J. & Seifert, R. Long-term forecast for antibacterial drug consumption in germany using arima models. *Naunyn-Schmiedeberg's Archives of Pharmacology* **398**(6), 7409–7428 (2025).
14. Huang, Y., Xu, C., Ji, M., Xiang, W. & He, D. Medical service demand forecasting using a hybrid model based on arima and self-adaptive filtering method. *BMC Medical Informatics and Decision Making* **20**(1), 237 (2020).
15. Punyapornwithaya, V. et al. Time series analysis and forecasting of the number of canine rabies confirmed cases in thailand based on national-level surveillance data. *Frontiers in Veterinary Science* **10**, 1294049 (2023).

16. Lawrence, J.-M., Ibne Hossain, N. U., Jaradat, R., & Hamilton, M. Leveraging a bayesian network approach to model and analyze supplier vulnerability to severe weather risk: A case study of the u.s. pharmaceutical supply chain following hurricane maria. *International Journal of Disaster Risk Reduction*, 49:101607, 2020.
17. Zhu, X., Ninh, A., Zhao, H. & Liu, Z. Demand forecasting with supply-chain information and machine learning: Evidence in the pharmaceutical industry. *Production and Operations Management* **30**(9), 3231–3252 (2021).
18. Rathipriya, R., Abdul Rahman, A. A., Dhamodharavadhani, S., Meero, A. & Yoganandan, G. Demand forecasting model for time-series pharmaceutical data using shallow and deep neural network model. *Neural Computing and Applications* **35**(2), 1945–1957 (2023).
19. Burinskiene, A. Forecasting model: The case of the pharmaceutical retail. *Frontiers in Medicine* **9**, 582186 (2022).
20. Michaela, A. C. et al. A unified machine learning approach to time series forecasting applied to demand at emergency departments. *BMC Emerg. Med.* **21**(1), 1–14 (2021).
21. Subramanian, L. Effective demand forecasting in health supply chains: Emerging trend, enablers, and blockers. *Logistics* **5**(1), 12 (2021).
22. Bruno Matos Porto and Flavio Sanson Fogliatto. Enhanced forecasting of emergency department patient arrivals using feature engineering approach and machine learning. *BMC Medical Informatics and Decision Making* **24**(1), 377 (2024).
23. Sami Khafaga, D. et al. Meta-heuristics for feature selection and classification in diagnostic breast cancer. *Computers, Materials & Continua* **73**(1), 749–765 (2022).
24. Ahmed, M. et al. An ai-based system for predicting renewable energy power output using advanced optimization algorithms. *J. Artif. Intell. Metaheuristics.* **8**(1), 1–8 (2024).
25. Sami Khafaga, D. et al. Improved prediction of metamaterial antenna bandwidth using adaptive optimization of lstm. *Comput. Mater. Continua* **73**(1), 865–881 (2022).
26. Uygun, Yasin & Sefer, Emre. Financial asset price prediction with graph neural network-based temporal deep learning models. *Neural Computing and Applications* **37**(30), 25445–25471 (2025).
27. Alaygut, T., & Sefer, E. Hypergraph neural networks to predict stock movements by exploring higher-order relationships. In *Proceedings of the 6th ACM International Conference on AI in Finance*, pp. 700–708, (2025).
28. Momanyi, Enock & Segera, Davies. A master–slave binary grey wolf optimizer for feature selection in high?dimensional biomedical datasets. *PLoS ONE* **16**(12), e0256762 (2021).
29. Wang, D., Ji, Y., Wang, H. & Huang, M. Binary grey wolf optimizer with a novel population adaptation strategy for feature selection. IET Cyber? *Phys. Syst. Theory Appl.* **8**(4), 392–404 (2023).
30. Chen, T., & Guestrin, C Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794. ACM, (2016).
31. Breiman, L.c., Friedman, J.H., Olshen, R.A. & Stone, C.J. *Classification and Regression Trees*. (Wadsworth International Group, 1984).
32. Breiman, L. Random forests. *Machine Learning* **45**(1), 5–32 (2001).
33. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.Y. Lightgbm: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 3149–3157, (2017).
34. Dorogush, A.V. , Ershov, V., & Gulin A. Catboost: Gradient boosting with categorical features support. arXiv:1810.11363, 2018.
35. Goodfellow, I., Bengio, Y. & Courville, A. *Deep Learning* (MIT Press, 2016).
36. Diederik, P., Kingma. Adam: A method for stochastic optimization. arXiv:1412.6980, 2014.
37. When to use them or not. Timothy O Hodson. Root mean square error (rmse) or mean absolute error (mae). *Geoscientific Model Development Discussions* **2022**, 1–10 (2022).
38. Chicco, D., Matthijs, J., Warrens & Jurman, G. The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation. *Peerj Comput. Sci.* **7**, e623 (2021).
39. Mbonyinshuti, François, Nkurunziza, Joseph, Niyobuhungiro, Japhet & Kayitare, Egide. The prediction of essential medicines demand: a machine learning approach using consumption data in rwanda. *Processes* **10**(1), 26 (2021).
40. Mbonyinshuti, François, Nkurunziza, Joseph, Niyobuhungiro, Japhet & Kayitare, Egide. Health supply chain forecasting: a comparison of arima and lstm time series models for demand prediction of medicines. *Acta Logistica* **11**(2), 269–280 (2024).
41. Fazekas, M., Veljanov, Z., & Borges, A., Oliveira, d. Predicting pharmaceutical prices: advances based on purchase-level data and machine learning. *BMC Public Health*, 24(1):1888, 2024.
42. Lin, Yuhao et al. Emergency drug demand forecasting in earthquakes with xgboost and aft-lstm. *Sustainability* **17**(5), 1910 (2025).
43. Naveed, Sajid & Husnain, Mujtaba. A drug recommendation system based on response prediction: integrating gene expression and k-mer fragmentation of drug smiles using lightgbm. *Intelligence-Based Medicine* **11**, 100206 (2025).
44. Grinsztajn, Léo., Oyallon, Edouard & Varoquaux, Gaël. Why do tree-based models still outperform deep learning on typical tabular data?. *Advances in neural information processing systems* **35**, 507–520 (2022).
45. Zhou, Lingling, Zhu, Qin, Chen, Qian, Wang, Ping & Huang, Hao. Predicting hospital outpatient volume using xgboost: a machine learning approach. *Scientific Reports* **15**(1), 17028 (2025).
46. Thai Food and Drug Administration. Announcement of drug system development action plan. https://en.fda.moph.go.th/news/announcement-of-drug-system-development-action-plan, 2025. Accessed: 2025-10-04.

Acknowledgements
The authors acknowledge the Health Intervention and Technology Assessment Program Foundation (HITAP), Ministry of Public Health; the Government Pharmaceutical Organization (GPO); the Lamphun Provincial Health Office; and Pasang and Banthi Hospitals for their valuable support and collaboration in this study.

Author contributions
A.M. and W.S. conceived the study. A.M. and S.M.T.S. developed the methodology, software, and performed the formal analysis. A.M., S.M.T.S., and S.K.C. validated the results. W.S. and S.K.C. provided resources and supervision. W.S. and Y.M.T. prepared the original draft. A.M. and S.K.C. reviewed and edited the manuscript. All authors reviewed the manuscript.

Funding
This study was co-funded by the Thailand Science Research and Innovation (TSRI) and the National Science, Research and Innovation Fund (NSRF) administered through the Program Management Unit for Human Resources & Institutional Development, Research and Innovation (B41G680024).

Declarations

Competing interests
The authors declare no competing interests.

Ethics statement
This study did not involve human participants or animal experiments. The use of de-identified hospital data complied with the Ministry of Public Health's data governance policy; therefore, ethical approval was not required.

Additional information
Correspondence and requests for materials should be addressed to A.M.