

The missing puzzle piece: Contextual insights for enhanced pharmaceutical supply chain forecasting

Arebu Issa Bilal^{a,*}, Bahman Rostami-Tabar^b, Harsha Halgamuwe Hewage^b, Umit Sezer Bititci^c, Teferi Gedif Fenta^a

^aAddis Ababa University, Department of Pharmaceutics and Social Pharmacy, School of Pharmacy, Ethiopia.,

^bCardiff University, Data Lab for Social Good Reserach Group, Cardiff Business School, UK.,

^cHeriot-Watt University, Edinburgh Business School, Heriot-Watt University, UK.,

Abstract

Accurate forecasting in pharmaceutical supply chains is critical for ensuring the continuous availability of essential medicines, particularly in developing countries where resource constraints and logistical challenges are more pronounced. Effective demand forecasting supports procurement and inventory management, which in turn could help to prevent stockouts, reduce wastage due to overstocking, and enhance overall healthcare delivery. In such settings, reliable forecasts, that acknowledge uncertainties, are essential to ensure that limited resources are used efficiently to meet health needs. Despite advancements in forecasting, significant gaps remain in effective forecasting in poor resource countries. Most existing literature on pharmaceutical supply chain forecasting focuses primarily on point estimation of consumption while neglecting the inherent uncertainty of these forecasts. Additionally, these studies often rely solely on historical consumption data without considering the broader context that influences consumption. Furthermore, many works lack rigorous methodological design, transparency, and reproducibility. This study addresses these gaps by integrating domain-specific knowledge, gathered through expert interviews and engagement with key members of the Ethiopian Pharmaceutical Supply Service (EPSS), into forecasting models. Using a dataset spanning five years (December 2017 to July 2022) from EPSS, we developed forecasting models that integrate expert-identified variables such as stock replenishment schedules, fiscal inventory counts, and disease outbreaks. Evaluation metrics including Mean Absolute Scaled Error (MASE), Root Mean Squared Scaled Error (RMSSE), and Continuous Ranked Probability Score (CRPS) are used to report forecast accuracy. The findings underscore the significance of contextual data in developing robust forecasting models that are adaptable to complex, real-world conditions. Our results also highlight the effectiveness of foundational time series models for forecasting. These models are particularly appealing for resource-constrained countries that may lack advanced analytical expertise. To promote usability, generalizability, and reproducibility, we share the complete dataset and code in R and Python, and the entire paper is written in Quarto via a GitHub repository to facilitate these practices.

Keywords: Forecasting, Pharmaceutical supply chain, Domain knowlege, Forecast accuracy, Developing countries

*Corresponding author

Email addresses: arebu.issa@aau.edu.et (Arebu Issa Bilal), rostami-tabarb@cardiff.ac.uk (Bahman Rostami-Tabar), HalgamuweHewageHR@cardiff.ac.uk (Harsha Halgamuwe Hewage), U.S.Bititci@hw.ac.uk (Umit Sezer Bititci), tgedif@gmail.com (Teferi Gedif Fenta)

1. Introduction

Universal access to essential medicines is a fundamental goal of effective healthcare systems, yet ensuring equitable availability continues to be a critical global challenge (Quick, 2003; Organization et al., 2004a). Many countries, particularly in Africa and Asia, face significant barriers to accessing these medications, often due to high costs and inefficient supply chain management (Organization et al., 2004b). A survey conducted in eight sub-Saharan African countries revealed alarmingly low availability of essential medicines. The average availability of women’s priority medicines ranged from just 22% to 40%, while children’s medicines were only slightly more accessible, with availability ranging from 28% to 57% (Droti et al., 2019). However, drug shortages are not limited to low-income regions; they also occur in developed areas like the United States and Europe, where they disrupt healthcare delivery and compromise the quality of patient care (Fox and Tyler, 2003; Kaakeh et al., 2011; Johnson, 2011; Huys and Simoens, 2013; Le et al., 2011).

The consequences of drug shortages extend beyond immediate healthcare delivery, impacting healthcare costs, treatment quality, and patient safety (Alspach, 2012; Kaakeh et al., 2011; Baumer et al., 2004). Such shortages can result in treatment delays, suboptimal care due to the unavailability of preferred therapies, increased reliance on costly secondary markets, and serious patient outcomes, including complications, longer hospital stays, and even deaths (Alspach, 2012). Addressing these challenges requires a resilient pharmaceutical supply chain that can ensure the timely availability of medicines. In this context, accurate demand forecasting becomes critical for effective procurement and inventory management (Subramanian, 2021).

Demand forecasting in pharmaceutical supply chains is a highly complex process, shaped by several factors. These include the quality of available data, the specific position within the supply chain, and external influences such as market trends, regulatory changes, and disruptive events such as pandemics, conflicts, flooding (Schneider et al., 2010; Hyndman and Athanasopoulos, 2018). Recently, major donors such as Bill & Melinda Gates Foundation have invested in data collection technologies in sub-Saharan Africa, recognizing the potential of data to provide valuable insights. This investment is essential and offers significant benefits for improving supply chain delivery. However, much of the data being collected focuses on transactional data, such as consumption and sale. Almost all platforms collecting data tend to overlook the critical contextual factors and events that influence these transactions—factors that are just as crucial for accurate forecasting. This presents a major barrier, as understanding the broader context, including administration procedures, local policies, disruptions, conflicts, and so on is vital for building reliable forecasting models. This gap—known as business context and domain knowledge—requires urgent attention in Pharmaceutical supply chain forecasting. Without incorporating this knowledge, any forecasting model is unlikely to produce reliable results. We must prioritize collecting and integrating this contextual information to improve the accuracy of demand forecasts in pharmaceutical supply chains. Combining statistical models with domain insights offers a more refined forecasting process, as evidenced by research that highlights the value of judgmental adjustments to enhance forecast accuracy (Fildes and Goodwin, 2007; Taylor-Phillips and Freeman, 2022; Soyiri and Reidpath, 2013). Domain knowledge enables analysts to incorporate relevant events, such as supply chain disruptions, into the forecasting model. This integration is particularly beneficial in fields like healthcare, where expert insights can align algorithms with real-world applications (Dash et al., 2022).

This study aims to bridge that gap by employing advanced machine learning models combined with domain expertise to address the distinctive characteristics of Ethiopia’s pharmaceutical sector. The challenges in this sector, such as diverse product categories, limited data from service delivery points, communication issues, long lead times, forecast inaccuracies, and complex policies, necessitate tailored forecasting approaches (Bilal et al., 2024; Boche et al., 2022). Accurate demand forecasting in Ethiopia is essential for optimizing resource allocation, streamlining supply chain operations, and shaping effective healthcare policies (Rostami-Tabar et al., 2022).

This research uses consumption data from 33 health commodities, covering the period from December 2017 to July 2022, provided by the Ethiopian Pharmaceutical Supply Service (EPSS), to develop models for six-month-ahead forecasting. Initially, we implement univariate models that rely solely on consumption data. Following this, we conduct exploratory data analysis and engage in interviews with EPSS experts

to gather domain knowledge on events influencing consumption patterns. Through these interviews, the experts identified three key factors collectively impacting consumption: stock replenishment period, fiscal calendar, and the seasonal malaria. We then develop forecasting models that incorporate the data gathered through the interview process and compare their performance with univariate models that rely solely on consumption data.

Our primary aim is to assess whether incorporating domain knowledge improves forecast accuracy. Moreover, we also fill a gap in the pharmaceutical demand forecasting literature by not only generating point forecasts but also producing probabilistic forecasting techniques that capture the range of potential outcomes with associated probabilities, offering a more comprehensive view of demand uncertainty than traditional methods. Model performance is evaluated using point accuracy measure including MASE, RMSSE, and probabilistic measure using CRPS.

The remainder of this paper is structured as follows: Section 2 reviews the literature and identifies gaps to position our work; Section 3 outlines the experimental design, including data sources, forecasting methods, and evaluation criteria; Section 4 presents the results and analysis; and Section 5 concludes with a summary of findings and future research directions.

2. Research background

Accurate sales forecasting is crucial in the pharmaceutical industry, where it directly impacts profit maximization, cost minimization, and the ability to respond to market changes. Forecasting in healthcare not only influences clinical decisions but also plays a pivotal role in supply chain management, ensuring that the right drugs are available when needed. The challenge lies in balancing customer demand with inventory costs, which is particularly complex in pharmaceuticals due to factors like short shelf life and quality constraints (Gupta et al., 2000; Makridakis et al., 2020).

In this review, a range of forecasting methods has been applied, from traditional statistical models to advanced machine learning techniques, with many studies using a combination of simple and advanced methods to enhance accuracy (Nikolopoulos et al., 2016; Zhu et al., 2021; Anusha et al., 2014). Some studies, such as those by Newberne and NV (2006), and Restyana et al. (2021), applied traditional forecasting methods like the Holt-Winters method, Single Moving Average (SMA), and Single Exponential Smoothing (SES), which are simpler but effective for certain applications. In contrast, other researchers explored more advanced approaches, including hybrid models and machine learning techniques, to address complex forecasting challenges, as seen in the studies by Siddiqui et al. (2022), de Oliveira et al. (2021), Kim et al. (2015), Candan et al. (2014), and Ribeiro et al. (2017).

Early research in pharmaceutical forecasting primarily focused on traditional statistical methods. These methods, including ARIMA, exponential smoothing, and moving averages, were widely used due to their simplicity and effectiveness in relatively stable environments (Zahra and Putra, 2019). These models typically assess forecast accuracy using metrics like Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). For example, Newberne and NV (2006) applied the Holt-Winters method to forecast healthcare data, focusing on prescription trends. Their findings highlighted the method's utility in short-term planning and resource management, validated through metrics like MAPE, MAD, and MSD. Similarly, Restyana et al. (2021) compared Single Moving Average (SMA) and Single Exponential Smoothing (SES) in drug demand forecasting, concluding that SES provided more accurate results, as indicated by lower MAD and MSE values.

Recognizing the limitations of traditional models, particularly in capturing the complexities of pharmaceutical demand (e.g., seasonality, external influences), researchers began developing hybrid models. These models combine linear methods like ARIMA with nonlinear techniques such as neural networks, aiming to improve forecasting accuracy by addressing diverse data patterns. Khalil Zadeh et al. (2014) introduced a hybrid model that integrated ARIMA with neural networks. Their study demonstrated that this approach

significantly outperformed traditional models, especially in scenarios with limited historical data. The hybrid model’s effectiveness was confirmed by improvements in key performance indicators like MAPE and RMSE.

The pharmaceutical industry’s increasing complexity and the availability of large datasets have led to the adoption of advanced machine learning techniques. These methods, such as Support Vector Regression (SVR), Random Forest (RF), and Long Short-Term Memory (LSTM) networks, offer greater flexibility and predictive power. ? showed that incorporating downstream information into machine learning models, like LASSO and SVR, greatly enhances forecast accuracy. Their study on multi-echelon supply chains revealed that advanced models, particularly those integrating external variables, consistently produced lower forecast errors, as measured by AvgRelRMSE. Similarly, [Rathipriya et al. \(2023\)](#) compared shallow neural networks and deep learning models for drug demand forecasting. The study found that shallow models outperformed deep learning models for most drug categories, while ARIMA was more effective for the remaining categories. This suggests that no single model is universally optimal, and model selection should be context-specific.

More sophisticated models, such as Neuro-Fuzzy Systems, have emerged as powerful tools for pharmaceutical forecasting. These systems combine the learning capabilities of neural networks with the reasoning abilities of fuzzy logic, offering a balanced approach that incorporates both empirical data and expert knowledge. Deep learning models, while powerful, present challenges such as the need for large datasets and substantial computational resources. Despite these challenges, models like LSTM have shown potential in capturing long-term dependencies in time series data, as demonstrated by [Sousa et al. \(2019\)](#) in their study on drug distribution in the Brazilian Public Health System. [Candan et al. \(2014\)](#) employed an Adaptive Neuro-Fuzzy Inference System (ANFIS) to forecast pharmaceutical demand. Their study highlighted the system’s ability to capture complex patterns in the data, resulting in highly accurate forecasts. The effectiveness of this approach was further validated through statistical tests, such as paired T-tests and mean difference analysis, which confirmed its superiority over traditional methods.

[Nguyen et al. \(2023\)](#) explore how sentiment analysis of news media can be utilized to improve demand forecasting for pharmaceutical products during disruptive events, such as pandemics and scandals. The authors employed a VARX (Vector Autoregressive with Exogenous Variables) model to forecast demand volatility. [Papanagnou and Matthews-Amune \(2018\)](#) investigate the impact of big data analytics, specifically through text mining, on improving forecasting accuracy for pharmaceuticals in retail pharmacies, using HMX pharmacy stores in Nigeria as a case study. They use multivariate time series analysis technique known as VARX, which incorporates customer-generated content from sources like Google, YouTube, and online newspapers to predict demand for medicines. [Fourkiotis and Tsadiras \(2024\)](#) focus on enhancing pharmaceutical sales forecasting through the application of both traditional statistical methods and advanced machine learning techniques. Various forecasting approaches, including traditional models like Naïve and ARIMA, Facebook Prophet, and machine learning methods such as XGBoost and LSTM neural networks are compared. The results demonstrated that machine learning models, particularly XGBoost, significantly outperformed traditional methods. [Belghith et al. \(2024\)](#) present a rolling forecasting framework employing moving average and three exponential smoothing methods (Brown’s, Holt’s, and Holt-Winters) implemented using Microsoft Power BI for enhanced data visualization.

Table 1 provides a summary of key studies in the literature on forecasting for pharmaceutical products. We identify several limitations both in the existing research and in practice, which highlight important gaps that motivate our current study. These gaps are summarized as follows:

1. Despite significant investment in data collection technologies and logistics management systems, these systems primarily collect transactional data, such as consumption. However, no data is collected on events that influence consumption variability—information that would be crucial for building reliable forecasting models. No studies to date have explored how domain expertise could enhance pharmaceutical demand forecasting.
2. Current research on forecasting for pharmaceutical supply chain predominantly focuses on generating point forecasts. There is a lack of studies that consider the entire forecast distribution of monthly

consumption, which would better capture the uncertainty of future demand and provide a valuable risk management tool for decision-makers.

3. Reproducibility remains a major challenge in this field. It is often difficult for readers to reproduce previous studies without direct assistance from the authors, limiting the practical application and validation of existing research.

Table 1: Summary of some studies on forecasting in pharmaceutical supply chain

| References | data granularity | Lenth of data | Forecast variable | Horizon | Probabilistic | Method | Metric | Items | Data | Code |
|----------------------------|------------------|---------------|---|------------|---------------|---|-------------------------------|---------|------|------|
| Current study | Monthly | 60 months | Consumption of pharmaceutical products | 6 months | Yes | TimeGpt (with and without predictors), LSTM (with and without predictors), Exponential Smoothing, ARIMA, Dynamic Regression (ARIMAX), Multiple Regression (with and without predictors) | RMSSE,CRIPS,MASE | 33 | Yes | Yes |
| Siddiqui et al. (2022) | Monthly | 58 months | Sales of pharmaceutical products | Un known | No | ARIMA, HOLT WINTER, ETS, Theta, ARHOW | MEAN, MAD, MSE, RMSE, MAPE | 31 | No | No |
| de Oliveira et al.(2021) | Daily | 365 days | Lead time | NA | No | k-NN, SVM, RF, LR, MLP | MSE | 11 | Yes | No |
| Rathipriya et al. (2023) | Weekly | 260 weeks | Sales of pharmaceuticals product | Un known | No | SNN, ARIMA P_NN GR_NN RBF_NN ARIMA P_NN GR_NN RBF_NN | RMSE, Normalized RMSE | 57 | Yes | No |
| Kim et al.(2015) | Monthly | 44 month | Medicines sales data | 1 month | No | VARX | Prediction error rate | 4 | No | No |
| Merkuryeva et al. (2019) | Weekly | 20 weeks | Sales of pharmaceuticals | 1-week | No | MA, Symbolic Regression, LR | R ² , MAD | 1 | No | No |
| Nikolopoulos et al. (2016) | Yearly | 21 year | Prescription records of pharmaceuticals | 1-5 | No | Diffusion models, ARIMA, ES, LR | R ² , ME, MAE, MSE | 14 | No | No |
| Van Belle et al.(2021) | Weekly | 253 weeks | Medicine sales | 1-5 weeks | No | ARIMA, ETS, MLP, SVR, RF, LR | RMSE | 50 | No | No |
| Khalil Zadeh et al. (2014) | Monthly | 36 months | Sales pharmaceutical products | 1 month | No | ARIMA, Graph-based analysis and ANN, | R2,MSE, MAE | 217 | No | No |
| Zhu X et al.(2021) | Weekly | 520 weeks | Medicine sales | 1-2 months | No | MA, LR, Clustering and RNN | NME, NMAE, NMSE | 245 | No | No |
| Anusha et al.(2014) | Monthly | 36 months | Sales of pharmaceuticals | Un known | No | 6 Month MA, SES, WES | MAD, MSE, MAPE | 2 | Yes | No |
| Burinskiene (2022) | Weekly | 13 weeks | Medicine sales | 36 days | No | SES, MA, Naive, Holt | MAPE,MSE, U ² | 8 | No | No |
| Candan et al. (2014) | Yearly | 6 years | Medicine sales | 3 months | No | NF | SE Mean | 1 | No | No |
| Ribeiro et al.(2017) | Monthly | 24 months | Medicine sales | 3 months | No | DPM | SMAPE | 357 | No | No |
| Bon and NG (2017) | Monthly | 68 months | Medicine sales | 1 month | No | SMA, SES, DMA, DES, RA, HW-A, SA, HW-M, ARIMA-SM | RMSE, MSE, MAD, MAPE | 1 | No | No |
| Newberne and Mike (2006) | Monthly | 36 months | Pseudoephedrine prescriptions | 3 months | No | HW-M | MAPE | 1 | No | No |
| Sousa et al.(2019) | Monthly | 79 months | drug distribution | quarterly | No | Naive, Seasonal Naive, ARIMA, DLM, LST, Recurrent Neural Network, SES, Holts linear and damped trends, HW, ES-SSM, Theta | MAE | 30 | Yes | Yes |
| Zahra and Putra (2019) | Monthly | 36 months | Consumption | 12 months | No | ARIMA, ES | RMSE, MAPE, MAD, MSE | 1 | No | No |
| Nguyen et al. (2023) | Un known | Un known | Pharmaceutical products | Un known | No | VARX | ME, MAE,MAPE,RMSE | Unknown | No | No |
| Christos et al.(2017) | Weekly | 129 weeks | Web-scraping of analgesic medicines | Un known | Yes | VARX | AIC, BIC, MSE, MAE | Unknown | No | No |
| Belghith et al.(2024) | Daily | 1095 days | Sales of pharmaceuticals | 1 year | No | Brown's method, Holt's method, and Holt-Winters method, moving average | MAD, MAPE | 125 | No | No |
| Konstantinos et al.(2024) | Weekly | 260 weeks | Medicine sales | 1 year | No | Naive, ARIMA, XGBoost , LSTM, Prophet | MAPE,MSE | 57 | No | No |

3. Experiment setup

3.1. Data

For this study, we use a dataset spanning five years (Dec. 2017- July 2022) of consumption data obtained from EPSS. Rigorous checks were conducted to ensure the consistency and completeness of the collected data. From the extensive pool of pharmaceutical products within EPSS, we selected a set of 33 key pharmaceuticals, including various programs and representing different classes of drugs.

3.1.1. Exploratory Data Analysis with consumption data

Given the number of products in this study, we created several data plot and computed features of the time series, including the strength of trend and seasonality to better understand data. Figure 1 shows the strength of trend versus seasonality. Each point represents one time series with the strength of trend in x-axis and the strength of seasonality in y-axis. Both measures are on a scale of $[0,1]$. The strength of trend and seasonality were calculated using the “STL” (Seasonal and Trend decomposition using Loess) decomposition method, as described by [Bandara et al. \(in press\)](#).

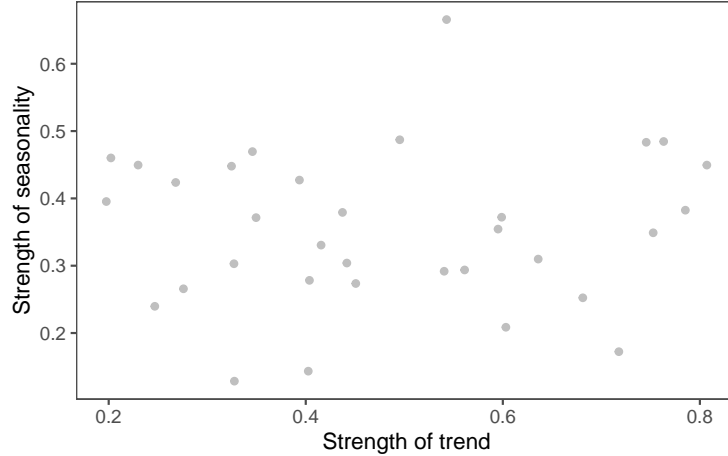


Figure 1: The strength of the trend and seasonality in the time series of pharmaceutical product consumption. The scatter plot shows 33 data points, with each point corresponding to a product.

It is evident that some time series display strong trends and/or seasonality, while the majority exhibit low trends and minimal seasonality. A number of products show pronounced trends, and only a few demonstrate clear seasonal patterns. Beyond assessing the strength of trends and seasonality, we also visualized all time series to understand data and various patterns, including trends and instances of erratic consumption behavior during certain months. For example, some series show low consumption volumes over consecutive months, followed by peak consumption in specific months, making them more volatile and difficult to forecast. This underscores the diversity of monthly pharmaceutical time series patterns within the dataset and highlights the importance of understanding the factors driving these consumption behaviors. Figure 2 illustrates the time plots for a few selected products.

Visualizing the data revealed that various events significantly impacted the consumption of different products, but these were not reflected in the available data. Expert insights were crucial for understanding the nature and timing of these events, filling gaps in the system, and incorporating qualitative factors that EPSS logistics system miss. Therefore, we conducted interviews with domain experts to collect information on external factors that influence consumption. These interviews allowed us to account for unusual patterns and customize our forecasting models to more accurately reflect the unique consumption behaviors of each product.

3.1.2. Collaborative expert review to identify factors affecting consumption

To understand the external factors influencing product consumption, we began by visualizing the time series data for consumption as highlighted in the previous section. This initial step helped identify specific instances within the data that showed unusual patterns, such as peaks, unusually high or low observations, and consecutive periods of low consumption. These observations indicated potential events or factors affecting consumption levels.

To gather domain knowledge about factors affected the consumption of products, we contacted experts from EPSS who work in the distribution unit. Next, we collaborated with six experts, each with over 10 years

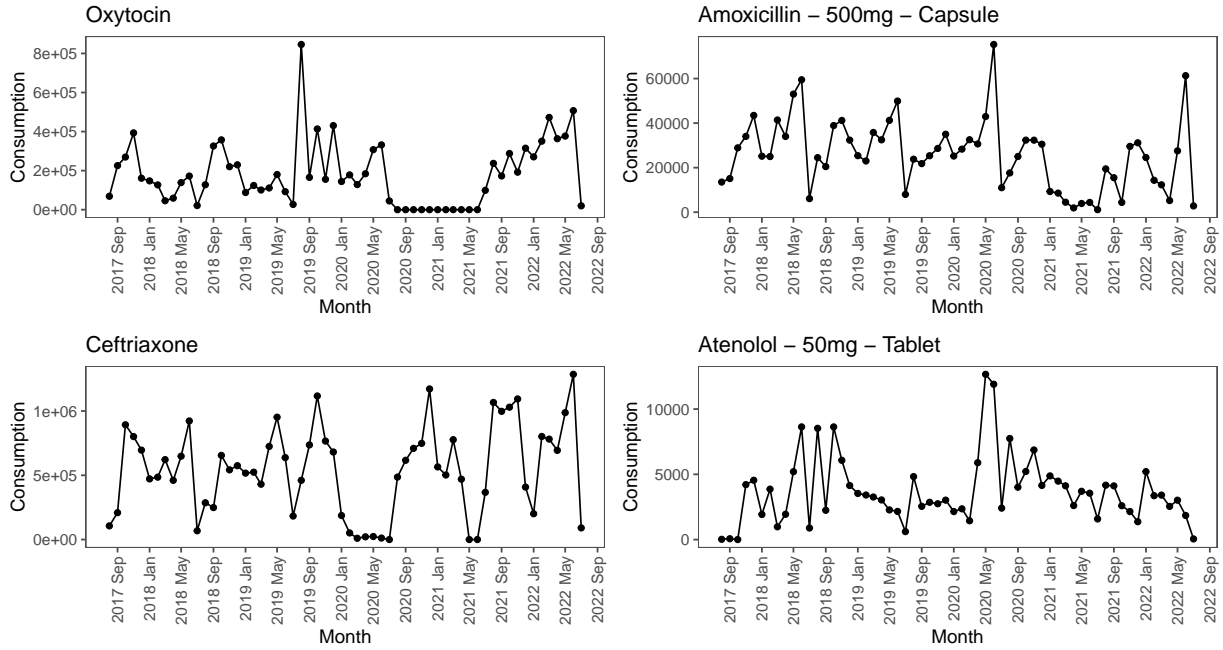


Figure 2: Monthly time plot of consumption. X-axis shows the month, consisting of 60 data points (months) and y-axis shows the consumption. The panels display data from four products to give a glimpse of the consumption patterns.

of experience managing pharmaceutical distribution. In these sessions, we reviewed the time series of all products together to understand the potential drivers behind these anomalies. The experts provided insights based on their extensive experience and knowledge of distribution practices. When there was uncertainty or gaps in the data for certain products, additional experts were consulted to gather further context. If all experts reached a consensus, the influencing factors were confirmed and included in the analysis.

As part of this process, we examined the electronic bin cards for 33 pharmaceutical products and conducted detailed data exploration. A bin card is a physical or electronic record used to track the inventory levels of individual items in a storage location, such as a warehouse or stockroom. The term ‘bin’ refers to the specific storage location for an item. Information typically recorded on a bin card includes the Item ID, unit of measure, product description, account details, date, references, beginning balance, manufacturer, batch number, expiry date, quantity received, quantity issued, and balance. In summary, a bin card is a fundamental tool in inventory management, especially in environments where physical stock tracking is essential. It provides a straightforward and immediate way to monitor item quantities, helping to ensure effective inventory management. Over a period of two weeks, we meticulously traced consumption data through these bin cards. Moreover, We ensured accuracy by converting dates from the Ethiopian calendar to the Gregorian calendar and excluded transactions related to internal storage or warehouse relocation within EPSS. Through this collaborative approach, we were able to identify several predictors affecting distribution and consumption that are summarised as followings:

- **Stock replenishment:** refers to the process of restocking or refilling inventory to ensure that there are sufficient quantities of products or materials available to meet demand. Whenever there was stock replenishment at the central EPSS, consumption and distribution to hubs and health facilities increased. This increase was attributed to the need to restock depleted inventories and the push from central EPSS to manage space constraints.
- **Physical fiscal year inventory counting:** refers to the process of manually counting and verifying the actual quantities of pharmaceutical products available in stock at a specific location. The process is

critical for maintaining the accuracy of inventory records, ensuring that medicines are available when needed, and preventing stockouts or overstocking. Physical inventory counting periods also influenced consumption. Stores closed during these periods, halting transactions. We observed increased consumption before inventory counting periods, as hubs and facilities stocked up. July and August were identified as physical counting periods each year.

- **Malaria seasonality:** Refers to the predictable patterns and fluctuations in malaria incidence throughout the year, typically influenced by climate and environmental conditions. In many regions, malaria transmission peaks during and shortly after the rainy season, when conditions such as stagnant water pools create ideal breeding sites for the *Anopheles* mosquitoes that transmit the disease. Conversely, malaria cases often decline during the dry season when mosquito breeding sites are reduced. During peak malaria seasons, there is a significant surge in the demand for antimalarial drugs and other related treatments. Malaria seasonality was another significant predictor. Certain pharmaceuticals, like Artemether + Lumefantrine and Rapid Diagnostic Test kits, were affected by malaria outbreaks. We identified epidemic periods affecting consumption: September to December 2017, March to May 2018, September to December 2018, March to May 2019, September to December 2019, March to May 2020, September to December 2020, March to May 2021, September to December 2021, and March to May 2022.

The key predictors identified are deterministic values, since both past and future values of these predictors are known in advance. We then integrated them into the consumption data, creating the complete dataset required for model building and running the experiment.

3.2. Forecasting models

We evaluate a range of univariate models and their counterparts that include predictors, spanning from simpler methods like regression, exponential smoothing, and ARIMA to more complex models such as long short-term memory (LSTM) networks and advanced foundational time series forecasting models. Below, we provide a brief overview of these approaches. Detailed implementation codes in R and Python are available in a GitHub repository and accessible for public.

Exponential Smoothing State Space model (ETS): ETS models, as described by [Hyndman and Athanasopoulos \(2021\)](#), combine trend, seasonality, and error components in time series using different configurations that can be additive, multiplicative, or mixed. The trend component can be specified as none (“N”), additive (“A”), or damped additive (“Ad”); the seasonality can be none (“N”), additive (“A”), or multiplicative (“M”); and the error term can be additive (“A”) or multiplicative (“M”). To forecast consumption, we use the ETS() function from the fable package in R, which automatically selects the optimal model for each time series based on the corrected Akaike’s Information Criterion (AICc). In our study, an automated algorithm determines the best configuration for trend, seasonality, and error components across each time series, leveraging the ets() function’s use of AIC to identify the optimal model. Given the high volume of time series (1530), manual selection of components is impractical, so the algorithm customizes model forms for each series based on its unique characteristics. This results in a tailored combination of additive or multiplicative components, adapting to the specific patterns of each time series.

Multiple Linear Regression (MLR): We use Multiple linear regression to model the relationship between a consumption and potential variables influencing its variation. In our first model, we use multiple linear regression with a trend component to capture the underlying progression over time, i.e., regression. We also incorporate dummy variables for each month to account for seasonal fluctuations, without including any additional predictors, i.e., regression_reg. This approach helps us establish a baseline model focused solely on temporal trends and seasonal patterns. We then extend this model by introducing additional predictors that include variables such as replenishment cycles, fiscal year indicators, and periods with malaria prevalence. These additional predictors allow the model to see if capturing external factors can provide a better understanding of the factors influencing the consumption and result at enhanced accuracy. We produce forecasts using TLSM() function from the fable package in R.

ARIMA and ARIMA with regressors: ARIMA (AutoRegressive Integrated Moving Average) is a powerful statistical model designed to forecast time series by capturing temporal dynamics and are widely used in time series forecasting due to their ability to model complex trends and patterns over time without relying on external predictors. ARIMAX (AutoRegressive Integrated Moving Average with eXogenous variables) extends ARIMA by incorporating external variables, or exogenous predictors, into the model. This modification allows to include relevant information from external factors such as malaria season, and fiscal year period, and stock replenishment period that may explain variations in the series beyond its internal time dynamics, we refer to this in the result as `arima_reg`. By adding these predictors, ARIMAX combines the strengths of ARIMA’s time-series structure with the flexibility of regression models. In our study, we use an automated algorithm to determine the optimal configuration for ARIMA components, following the approach outlined by [Hyndman and Athanasopoulos \(2021\)](#) and We use the `ARIMA()` function from the `fable` package in R.

Long Short-Term Memory neural network (LSTM): The LSTM model is a specialised form of recurrent neural network (RNN) used to model sequential data by capturing long-term dependencies ([Graves and Graves, 2012](#)). Unlike traditional RNNs, LSTMs can learn to retain information for longer time periods due to their unique architecture, which consists of several gates that control the flow of information. This makes LSTMs particularly effective for time series forecasting. In our implementation, we used a sequential model architecture, comprising one LSTM layer with 50 units, followed by dense layers, with the final output being a single linear unit. The Adam optimizer was employed to minimize the mean squared error, and the model was trained for 100 epochs using the `keras_model_sequential()` function from the Keras package in R. We use LSTM models both with and without predictors, referring to the LSTM model with predictors as `lstm_reg`.

TimeGPT: TimeGPT is the first pre-trained foundational model designed specifically for time series forecasting, developed by Nixtla ([Garza et al., 2023](#)). It uses a transformer-based architecture with an encoder-decoder configuration but differs from other models in that it is not based on large language models. Instead, it is built from the ground up to handle time series data. TimeGPT was trained on more than 100 billion data points, drawing from publicly available time series across various sectors, such as retail, healthcare, transportation, demographics, energy, banking, and web traffic. This wide range of data sources, each with unique temporal patterns, enables the model to manage diverse time series characteristics effectively. Furthermore, TimeGPT supports the inclusion of external regressors in its forecasts and can generate quantile forecasts, providing reliable uncertainty estimation. We use TimeGpt models both with and without predictors, referring to the model with predictors as `timegpt_reg`.

In all models, probabilistic forecasts were generated using bootstrapping to create 1,000 possible future scenarios for each period within the forecast horizon ([Hyndman and Athanasopoulos, 2021](#)).

3.3. Performance evaluation

To assess the performance of our forecasting methods, we split the data into a series of training and test sets and apply time series cross-validation with a forecast horizon of 6 months. Each training set is expanded in monthly increments, allowing model development and hyperparameter tuning to be performed on the training data, with errors evaluated on the corresponding test sets. We assess forecasting performance using both point and probabilistic error measures.

The error metrics presented here consider a forecasting horizon denoted by j , which represents the number of time periods ahead we are predicting, with j ranging from 1 to 6 months in our study. Point forecast accuracy is measured using the Mean Squared Scaled Error (MSSE) and the Mean Absolute Scaled Error (MASE). The Mean Absolute Scaled Error (MASE) ([Hyndman and Koehler, 2006](#); [Hyndman and Athanasopoulos, 2021](#)) is calculated as follows:

$$\text{MASE} = \text{mean}(|q_j|),$$

where

$$q_j = \frac{e_j}{\frac{1}{T-m} \sum_{t=m+1}^T |y_t - y_{t-m}|},$$

and e_j is the point forecast error for forecast horizon j , $m = 12$ (as we have monthly seasonal series), y_t is the observation for period t , and T is the sample size (the number of observations used for training the forecasting model). The denominator is the mean absolute error of the seasonal naive method in the fitting sample of T observations and is used to scale the error. Smaller MASE values suggest more accurate forecasts. Note that the measure is scale-independent, thus allowing us to average the results across series.

Here, e_j represents the point forecast error for forecast horizon j , with $m = 12$ (since we are dealing with monthly seasonal data). The term y_t denotes the observation at time t , and T is the sample size, or the number of observations used for training the forecasting model. The denominator in the MASE formula is the mean absolute error of the seasonal naive method over the training sample of T observations, providing a basis for scaling the forecast error. Lower MASE values indicate more accurate forecasts. Notably, this measure is scale-independent, allowing us to average results across different series for broader performance comparison.

A related measure is MSSE (Hyndman and Athanasopoulos, 2021; Makridakis et al., 2022), which uses squared errors rather than absolute errors:

$$\text{MSSE} = \text{mean}(q_j^2),$$

where,

$$q_j^2 = \frac{e_j^2}{\frac{1}{T-m} \sum_{t=m+1}^T (y_t - y_{t-m})^2},$$

Again, this is scale-independent, and smaller MSSE values suggest more accurate forecasts.

To measure the forecast distribution accuracy, we calculate the Continuous Rank Probability Score (Hyndman and Athanasopoulos, 2021). It rewards sharpness and penalizes miscalibration, so it measures overall performance of the forecast distribution.

$$\text{CRPS} = \text{mean}(p_j),$$

where

$$p_j = \int_{-\infty}^{\infty} (G_j(x) - F_j(x))^2 dx,$$

where $G_j(x)$ is the forecasted probability distribution function for forecast horizon j , and $F_j(x)$ is the true probability distribution function for the same period.

Calibration refers to the statistical consistency between the distributional forecasts and the observations. It measures how well the predicted probabilities match the observations. On the other hand, sharpness refers to the concentration of the forecast distributions — a sharp forecast distribution results in narrow prediction intervals, indicating high confidence in the forecast. A model is well-calibrated if the predicted probabilities match the distribution of the observations, and it is sharp if it is confident in its predictions. The CRPS rewards sharpness and calibration by assigning lower scores to forecasts with sharper distributions, and to forecasts that are well-calibrated. Thus, it is a metric that combines both sharpness and miscalibration into a single score, making it a useful tool for evaluating the performance of probabilistic forecasts.

CRPS can be considered an average of all possible quantiles (Hyndman and Athanasopoulos, 2021, Section 5.9), and thus provides an evaluation of all possible prediction intervals or quantiles. A specific prediction interval could be evaluated using a Winkler score, if required.

4. Results and discussion

In this section, we compare the forecasting performance of various approaches, examining models that incorporate expert-identified business context predictors versus those that rely solely on historical consumption data. Point forecast performance is reported using MASE and RMSSE, while probabilistic forecast accuracy is reported using CRPS.

The forecasting performance is reported in Figure 3, in which the average forecast accuracy over forecast horizon and across all products is calculated for each origin. We report the distribution of accuracy metrics across all rolling origins. This shows how models varies in providing accuracy across different origins. The y-axis displays models sorted by their MASE and RMSSE values, with the model exhibiting the lowest error positioned at the bottom. This model is the Ariam model, which incorporates exogenous variables. Additionally, Figure 3 indicates that predictors obtained through interactions with domain experts enhance point forecast accuracy across most models. This underscores the critical importance of systematically collecting such expert-informed data alongside transactional consumption data.

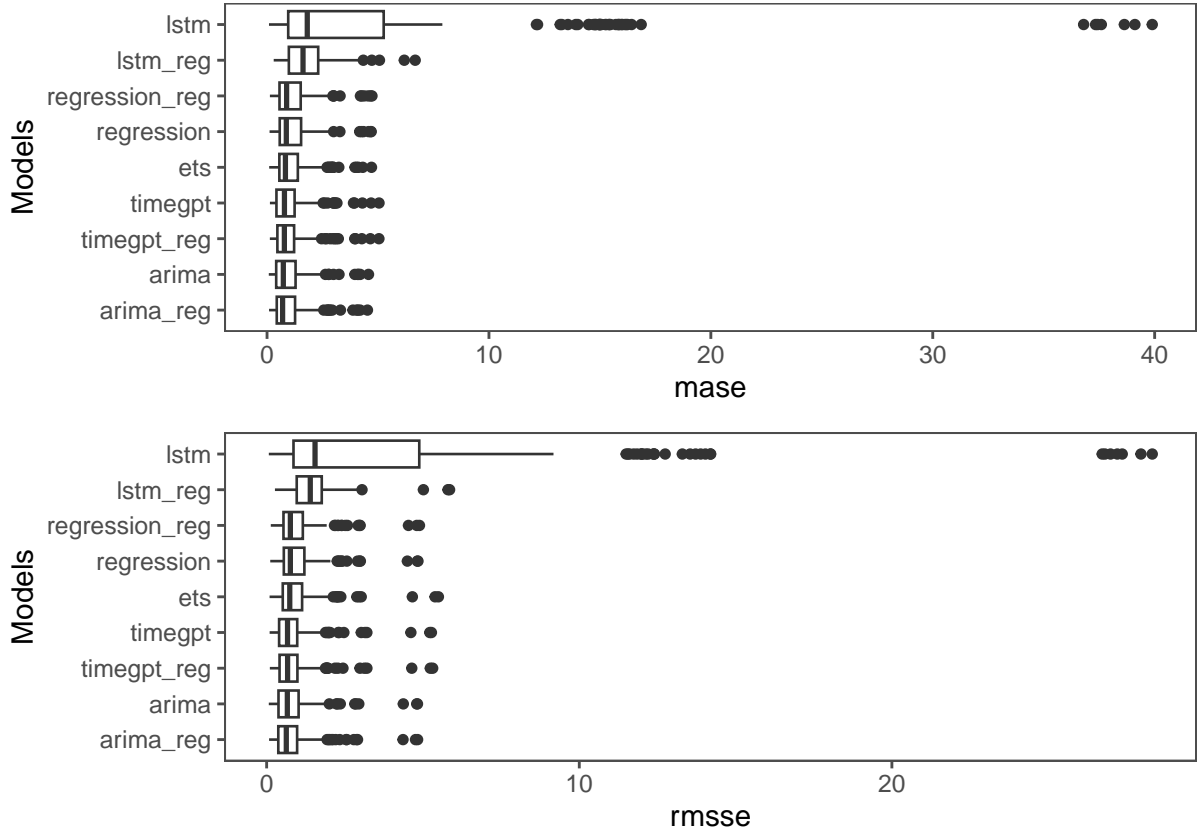


Figure 3: Distribution of point forecast accuracy across different origins, averaged across the forecast horizon and all products. The total number of months used to calculate the accuracy in the test set is 12 months for each product. MASE and MSSE are relative to the corresponding values for the training set.

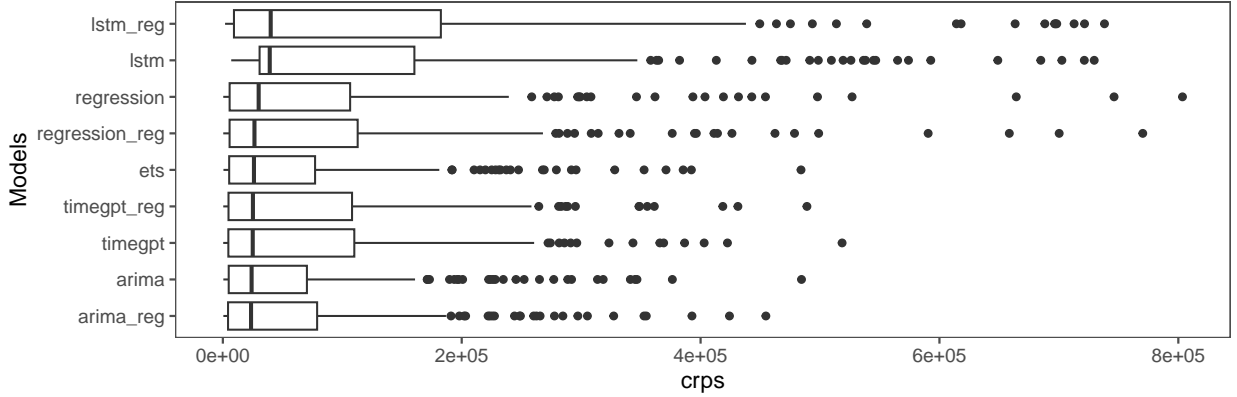


Figure 4: Distribution of probabilistic forecast accuracy across different origins, averaged across the forecast horizon and all products. The total number of months used to calculate the accuracy in the test set is 12 months for each product.

Figure 4 illustrates the forecast distribution accuracy measured by CRPS, which evaluates both forecasting calibration and interval sharpness. A smaller CRPS value indicates better overall performance. We observed that incorporating domain knowledge improved forecast accuracy for most models, enhancing not only point forecasts but also probabilistic forecasts. Notably, ARIMA achieved the highest probabilistic forecast accuracy, while LSTM models had the lowest. This aligns with the findings related to point forecast accuracy, reinforcing the earlier explanations for these results.

We observe that LSTM models exhibit a widespread error distribution compared to other models. This is attributed to their performance variability, which can significantly depend on the characteristics of demand patterns across different products. For example, the product “Amlodipine - 5mg - Tablet” demonstrates periods of extreme variability, with spikes in demand followed by periods of very low or zero demand. Such patterns align well with the strengths of LSTM models, which excel at capturing long-term dependencies and managing complex temporal fluctuations. In contrast, the demand for “Anti-Rho (D)” is erratic and sporadic, characterized by frequent, random fluctuations without any discernible long-term trends or seasonal patterns. This chaotic nature, combined with inconsistent and often low volumes, poses significant challenges for LSTM models. The lack of clear structure or stable patterns limits the models’ ability to generalize effectively, resulting in poorer forecast performance. Although LSTM models can handle irregular data to some extent, they require sufficient and well-distributed training samples to achieve optimal performance. In our study, the majority of products exhibit erratic and sporadic demand patterns similar to “Anti-Rho (D)”, resulting in this widespread error distribution.

The findings emphasize the importance of incorporating relevant domain knowledge into forecasting models. In pharmaceutical supply logistics administration systems, data often consists solely of transactional records on consumption and distribution. Including exogenous variables such as administrative procedures, seasonal patterns, or conflicts, or any other relevant factor that could be identified by those with domain knowledge proved essential for reducing forecast errors. This approach improved both point and probabilistic forecast accuracy, enabling a more confident assessment of uncertainty. Therefore, the systematic collection of information about significant events and their impact on consumption is vital. Recording details of events such as policy changes, administration procedures, conflicts, and incorporating this information into forecasting models creates a comprehensive understanding of consumption. This practice enhances modeling reliability and allows institutions in developing countries and humanitarian organizations to better forecast demand, allocate resources effectively, and respond proactively. Establishing robust data collection systems is thus a critical step for strengthening forecasting capabilities and operational resilience.

Our analysis also highlights the value of using foundational time series forecasting models like TimeGpt, particularly in developing countries and the global health and humanitarian sectors. In these contexts,

the lack of advanced analytical skills and systematic data collection can hinder effective forecasting and decision-making.

4.1. An illustration of probabilistic forecast for Pharmaceutical product consumption

In this section, we present an illustrative example of a probabilistic forecast for future consumption of Sodium Chloride (Normal Saline) product. Due to the complexity of including such visualizations for all products, only one example is shown here. However, it is feasible to generate these plots for all products if needed.

In practice, point forecasts are commonly used despite their limitations, but they do not account for the inherent uncertainty associated with forecasts. The future is inherently uncertain, and effective planning requires considering alternative scenarios. Probabilistic forecasts offer a comprehensive approach by assigning likelihoods to a range of possible outcomes, recognizing that different consumption levels may occur with varying probabilities. The primary purpose of probabilistic forecasting, as illustrated in Figure 5, is to quantify and communicate uncertainty. This figure displays the forecast distribution of consumption over a 6-month horizon using a density plot. For each month within the forecast period, a separate distribution is generated. The plot also includes the point forecast alongside 80% and 90% prediction intervals to illustrate potential variability.

It is important to note that while point forecasts and prediction intervals can be derived from probabilistic forecasts, the reverse is not true. A single-point forecast cannot inherently provide the probabilistic context needed to capture the range of possible outcomes. Although prediction intervals can indicate a range of potential values, they do not convey the detailed probabilities of low or high consumption. This distinction highlights the value of probabilistic forecasting in supporting informed decision-making by offering a clearer view of future uncertainties.

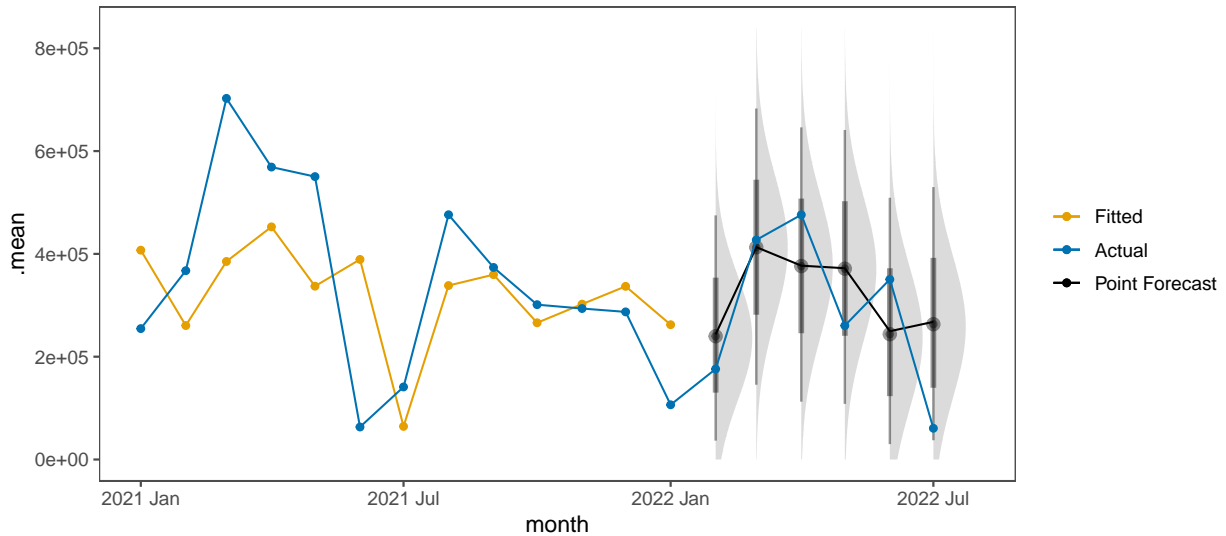


Figure 5: A graphical illustration of the forecast distribution of a pharmaceutical product (i.e. total incidence attended) for the SB health board for a horizon of six month. For each month, we display the point forecast (black point), the histogram, and 80% (thick line) and 90% (thin line) prediction intervals. It also shows a portion of a historical time series as well as its fitted values.

5. Conclusion and future reserach

In this study, we conducted an extensive analysis of pharmaceutical demand forecasting within the EPSS. Using five years of consumption data for 33 key pharmaceutical products, along with additional information

gathered through collaboration with domain experts at EPSS, our goal is to contribute to enhancing the demand forecasting process and emphasize the importance of integrating domain knowledge into model building. This step is critical, as the phase of understanding the data and incorporating valuable contextual information is often overlooked. Too frequently, forecasting efforts rely solely on consumption data, leading to building models with little relevance to reality. Our approach highlights the necessity of including domain insights to construct more accurate and effective forecasting models.

Our findings highlight the importance of collecting and incorporating domain knowledge when building forecasting models. We evaluated both point and probabilistic forecasts using a range of models, from simple univariate approaches like ARIMA to complex models such as LSTM. Notably, we demonstrated that employing complex models does not necessarily lead to more accurate forecasts in the pharmaceutical supply context. Additionally, our research revealed that newly developed foundational time series forecasting models have potential applicability, particularly in environments where advanced analytical skills and knowledge are limited, as often seen in pharmaceutical supply chains in developing countries.

To advance forecasting practices further, exploring additional predictor variables—such as public health campaigns, disease outbreaks, and economic indicators—could provide valuable insights into demand dynamics. Improving the granularity of historical consumption data, through finer temporal intervals and geographic differentiation, also holds potential for more precise predictions. Additionally, hybrid forecasting models that combine judgmental point forecasts with probabilistic machine learning models could enhance predictive performance. Empowering EPSS staff through training in foundational time series models, data interpretation, and collection will enable informed decision-making and strengthen workforce capabilities. Replicating this study across diverse healthcare settings and conducting comparative analyses would help validate the adaptability and applicability of the findings. Additionally, it is recommended that pharmaceutical supply services systematically collect and maintain detailed records of events that may influence consumption, alongside consumption data. This practice is essential for refining demand models and enhancing forecast accuracy.

5.1. Practical challenges and limitations

Despite dedicated efforts to engage with domain experts to better understand the data—and spending significant time over two weeks collaborating with them—we found that the process of interpreting data through domain knowledge is complex and presents unique challenges. These challenges include defining what constitutes domain knowledge and determining how best to incorporate it to enhance model reliability and accuracy. Below, we summarize some of these key challenges. One issue involved capturing accurate information from bin cards and expert review, particularly during periods of disruption like COVID-19 and conflicts. For instance, during the COVID-19 pandemic, experts noted a decrease in demand due to travel restrictions. However, despite the reduced demand, there were still significant distributions of pharmaceuticals from the EPSS central to various hubs and health facilities. This was because, in response to anticipated shortages, a political decision was made to push products to these facilities before the travel restrictions took full effect. The rationale was that if the facilities remained closed, patients would have no access to medications, necessitating the preemptive stockpiling of essential pharmaceuticals. This scenario illustrates how policy decisions can significantly impact supply chain data, making it challenging to model distribution patterns accurately. Similarly, experts indicated that conflicts would disrupt the distribution of pharmaceuticals. While conflict zones did indeed hinder transportation, distribution still occurred whenever roads were temporarily opened, even amidst on-going conflict. This created inconsistencies in the data, as periods of halted distribution were followed by rapid replenishment once access was restored. Such fluctuations add complexity to modelling the supply chain from the central EPSS to hubs and health facilities. Other challenges involved seasonal or periodic activities, such as the fiscal year-end stock counting. During this time, EPSS temporarily halts transactions to conduct a full inventory count, which can take anywhere from one to two months, depending on the circumstances. This inconsistency in the duration of inventory counts adds a layer of unpredictability to supply chain modelling.

5.2. Future reserach

Following the current research, several promising areas could further advance knowledge and practical applications in this area:

- Future research could focus on developing intelligent systems that incorporate domain-specific knowledge into model building. This would involve creating algorithms capable of identifying relevant domain insights and integrating them effectively into model training processes. Such tools would bridge the gap between purely data-driven approaches and expert-knowledge-enhanced modeling, leading to more robust and contextually informed forecasts.
- Large Language Models (LLMs) for time series forecasting, like the one included in this study, offer transformative opportunities for developing countries by making advanced forecasting methods accessible, even in contexts with limited expertise or infrastructure. However, future research should explore their applicability, advantages, limitations, and performance relative to traditional and hybrid forecasting methods.
- The use of distribution-based representation of error metrics across different rolling origins or forecast horizons, such as boxplots, to showcase forecast accuracy provides essential insights into variability in accuracy at different point of time in data. Building on this, future work could aim at designing methodologies that identify diagnostic patterns within the distribution of errors. These approaches could suggest model improvements by analyzing error distributions and help practitioners to refine models based on error behavior.
- While current study focus on monthly aggregate consumption at the organizational level, there is potential in exploring more granular forecasts. Future research could investigate daily or weekly forecasting and examine the impacts at hierarchical levels, such as specific sites or health facilities. Investigating how fine-grained temporal and hierarchical forecasts influence inventory management and operational efficiency could yield valuable insights for reducing stockouts, minimizing costs and reducing waste.

Reproducibility

We use R and Python for model development and accuracy evaluation. The paper is written in Quarto, and all code, data, and files will be shared in a GitHub repository upon acceptance for publication.

References

- Alspach, J.G., 2012. Is the drug shortage affecting patient care in your critical care unit?
- Anusha, S.L., Alok, S., Shaik, A., 2014. Demand forecasting for the indian pharmaceutical retail: A case study. *Journal of Supply Chain Management Systems* 3, 1–8.
- Bandara, K., Hyndman, R.J., Bergmeir, C., in press. MSTL: A seasonal-trend decomposition algorithm for time series with multiple seasonal patterns. *International J Operational Research* .
- Baumer, A.M., Clark, A.M., Witmer, D.R., Geize, S.B., Vermeulen, L.C., Deffenbaugh, J.H., 2004. National survey of the impact of drug shortages in acute care hospitals. *American journal of health-system pharmacy* 61, 2015–2022.
- Belghith, M., Ammar, H.B., Elloumi, A., Hachicha, W., 2024. A new rolling forecasting framework using microsoft power bi for data visualization: A case study in a pharmaceutical industry, in: *Annales Pharmaceutiques Françaises*, Elsevier. pp. 493–506.
- Bilal, A.I., Bititci, U.S., Fenta, T.G., 2024. Challenges and the way forward in demand-forecasting practices within the ethiopian public pharmaceutical supply chain. *Pharmacy* 12, 86.
- Boche, B., Mulugeta, T., Gudeta, T., 2022. Procurement practice of program drugs and its challenges at the ethiopian pharmaceuticals supply agency: A mixed methods study. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing* 59, 00469580221078514.
- Candan, G., fathi Taskin, M., Yazgan, H., 2014. Demand forecasting in pharmaceutical industry using neuro-fuzzy approach. *Journal of Management and Information Science* 2, 41–49.
- Dash, T., Chitlangia, S., Ahuja, A., Srinivasan, A., 2022. A review of some techniques for inclusion of domain-knowledge into deep neural networks. *Scientific Reports* 12, 1040.
- Droti, B., O'Neill, K.P., Mathai, M., Dovlo, D.Y.T., Robertson, J., 2019. Poor availability of essential medicines for women and children threatens progress towards sustainable development goal 3 in africa. *BMJ global health* 4, e001306.

- Fildes, R., Goodwin, P., 2007. Against your better judgment? how organizations can improve their use of management judgment in forecasting. *Interfaces* 37, 570–576.
- Fourkiotis, K.P., Tsadiras, A., 2024. Applying machine learning and statistical forecasting methods for enhancing pharmaceutical sales predictions. *Forecasting* 6, 170–186.
- Fox, E.R., Tyler, L.S., 2003. Managing drug shortages: seven years' experience at one health system. *American journal of health-system pharmacy* 60, 245–253.
- Garza, A., Challu, C., Mergenthaler-Canseco, M., 2023. Timegpt-1. arXiv preprint arXiv:2310.03589 .
- Graves, A., Graves, A., 2012. Long short-term memory. *Supervised sequence labelling with recurrent neural networks* , 37–45.
- Gupta, A., Maranas, C.D., McDonald, C.M., 2000. Mid-term supply chain planning under demand uncertainty: customer demand satisfaction and inventory management. *Computers & Chemical Engineering* 24, 2613–2621.
- Huys, I., Simoens, S., 2013. European drug shortages: a call for action!
- Hyndman, R.J., Athanasopoulos, G., 2018. *Forecasting: principles and practice*. OTexts.
- Hyndman, R.J., Athanasopoulos, G., 2021. *Forecasting: principles and practice*. 3rd ed., OTexts.
- Hyndman, R.J., Koehler, A.B., 2006. Another look at measures of forecast accuracy. *International Journal of Forecasting* 22, 679–688.
- Johnson, T.J., 2011. Drug shortages: an increasing problem for patients and clinicians. *South Dakota Medicine: the Journal of the South Dakota State Medical Association* 64, 14–15.
- Kaakeh, R., Sweet, B.V., Reilly, C., Bush, C., DeLoach, S., Higgins, B., Clark, A.M., Stevenson, J., 2011. Impact of drug shortages on us health systems. *American Journal of Health-System Pharmacy* 68, 1811–1819.
- Khalil Zadeh, N., Sepehri, M.M., Farvaresh, H., 2014. Intelligent sales prediction for pharmaceutical distribution companies: A data mining based approach. *Mathematical Problems in Engineering* 2014, 420310.
- Kim, W., Won, J.H., Park, S., Kang, J., 2015. Demand forecasting models for medicines through wireless sensor networks data and topic trend analysis. *International Journal of Distributed Sensor Networks* 11, 907169.
- Le, P., Seoane-Vazquez, E., Rodriguez-Monguio, R., Fox, E.R., Szeinbach, S.L., Duneheew, A.R., Montagne, M., 2011. The prevalence of pharmaceutical shortages in the united states. *Journal of Generic Medicines* 8, 210–218.
- Makridakis, S., Hyndman, R.J., Petropoulos, F., 2020. Forecasting in social settings: The state of the art. *International Journal of Forecasting* 36, 15–28.
- Makridakis, S., Spiliotis, E., Assimakopoulos, V., 2022. M5 accuracy competition: Results, findings, and conclusions. *International Journal of Forecasting* 38, 1346–1364.
- Newberne, J.H., NV, M.O.F.H.L.V., 2006. *Holt-Winters Forecasting: A Study of Practical Applications for Healthcare Managers*. US Army Medical Department Center and School.
- Nguyen, A., Pellerin, R., Lamouri, S., Lekens, B., 2023. Managing demand volatility of pharmaceutical products in times of disruption through news sentiment analysis. *International Journal of Production Research* 61, 2829–2840.
- Nikolopoulos, K., Buxton, S., Khammash, M., Stern, P., 2016. Forecasting branded and generic pharmaceuticals. *International Journal of Forecasting* 32, 344–357.
- de Oliveira, M.B., Zucchi, G., Lippi, M., Cordeiro, D.F., da Silva, N.R., Iori, M., et al., 2021. Lead time forecasting with machine learning techniques for a pharmaceutical supply chain., in: *ICEIS (1)*, pp. 634–641.
- Organization, W.H., et al., 2004a. *Annual report 2003: essential drugs and medicines policy*. Technical Report. World Health Organization.
- Organization, W.H., et al., 2004b. *WHO medicines strategy 2004-2007: countries at the core*. Technical Report. World Health Organization.
- Papanagnou, C.I., Matthews-Amune, O., 2018. Coping with demand volatility in retail pharmacies with the aid of big data exploration. *Computers & Operations Research* 98, 343–354.
- Quick, J.D., 2003. Essential medicines twenty-five years on: closing the access gap. *Health policy and planning* 18, 1–3.
- Rathipriya, R., Abdul Rahman, A.A., Dhamodharavadhani, S., Meero, A., Yoganandan, G., 2023. Demand forecasting model for time-series pharmaceutical data using shallow and deep neural network model. *Neural Computing and Applications* 35, 1945–1957.
- Restyana, A., Savitri, L., Laili, N.F., Probosiwi, N., 2021. Analysis of drug forecasting with single moving average and single exponential smoothing approach (case study in jombang regency 2017-2019), in: *Journal of Physics: Conference Series*, IOP Publishing. p. 012100.
- Ribeiro, A., Seruca, I., Durão, N., 2017. Improving organizational decision support: Detection of outliers and sales prediction for a pharmaceutical distribution company. *Procedia computer science* 121, 282–290.
- Rostami-Tabar, B., Ali, M.M., Hong, T., Hyndman, R.J., Porter, M.D., Syntetos, A., 2022. Forecasting for social good. *International Journal of Forecasting* 38, 1245–1257.
- Schneider, J.L., Wilson, A., Rosenbeck, J.M., 2010. Pharmaceutical companies and sustainability: an analysis of corporate reporting. *Benchmarking: An international journal* 17, 421–434.
- Siddiqui, R., Azmat, M., Ahmed, S., Kummer, S., 2022. A hybrid demand forecasting model for greater forecasting accuracy: the case of the pharmaceutical industry, in: *Supply Chain Forum: An International Journal*, Taylor & Francis. pp. 124–134.
- Sousa, R.M., Hannachi, S., Ramos, G.N., 2019. Statistical and deep learning models for forecasting drug distribution in the brazilian public health system, in: *2019 8th Brazilian Conference on Intelligent Systems (BRACIS)*, IEEE. pp. 723–728.
- Soyiri, I.N., Reidpath, D.D., 2013. An overview of health forecasting. *Environmental health and preventive medicine* 18, 1–9.
- Subramanian, L., 2021. Effective demand forecasting in health supply chains: emerging trend, enablers, and blockers. *Logistics* 5, 12.
- Taylor-Phillips, S., Freeman, K., 2022. Artificial intelligence to complement rather than replace radiologists in breast screening. *The Lancet Digital Health* 4, e478–e479.

- Zahra, I., Putra, Y., 2019. Forecasting methods comparison based on seasonal patterns for predicting medicine needs with arima method, single exponential smoothing, in: IOP Conference Series: Materials Science and Engineering, IOP Publishing. p. 022030.
- Zhu, X., Ninh, A., Zhao, H., Liu, Z., 2021. Demand forecasting with supply-chain information and machine learning: Evidence in the pharmaceutical industry. *Production and Operations Management* 30, 3231–3252.