DATA-DRIVEN DEMAND FORECASTING FOR BLOOD COMPONENTS FOR ACHIEVING NET ZERO

Kriti Karmakar
Calcutta Business School
Kolkata, India
kritik@calcuttabuinessschool.org
DOI: https://doi.org/10.64558/bllp/870070xjaqrm

ABSTRACT

Introduction: Achieving net zero in healthcare requires minimizing waste, optimizing energy use, and improving resource efficiency. Blood banks, as an integral part of the healthcare system, contribute to carbon emissions through energy-intensive storage, transportation, and disposal of expired blood units. Accurate demand forecasting plays a vital role in reducing both over-collection and wastage, thereby contributing to the net zero goal. This study applies data- driven forecasting methods to analyze and predict the demand for Packed Red Blood Cells (PRBC) across all eight major blood types in a stand-alone blood bank in Kolkata, India.

Methodology: Historical PRBC demand data were analyzed individually for each blood group to identify underlying trends, seasonality, and variability. Various forecasting models such as Exponential Smoothing, Winter's method, STL+ARIMA and SARIMA were applied to determine the most efficient and sustainable model. The study evaluated model performance based on forecasting accuracy and its potential to reduce overcollection and inventory-related wastage, which indirectly lowers the energy and carbon footprint of storage and disposal processes.

Results: The Exponential Smoothing model demonstrated superior forecasting performance for most blood groups due to its adaptive error correction and high smoothing coefficient, effectively capturing recent demand variations. Exceptions were observed for AB+, B-, and AB- blood types, where alternative models performed better. Accurate forecasts achieved through this approach can reduce PRBC wastage by aligning collection with true demand, leading to a measurable reduction in refrigeration energy use and disposal-related emissions.

Conclusion: Data driven demand forecasting provides a technological pathway toward net zero blood banking by minimizing waste and optimizing collection quantity, thereby reducing energy consumption. Integrating such predictive models into blood bank management systems supports sustainable operations, ensuring that environmental goals are met without compromising patient care.

Keywords: Blood Bank, Healthcare, Forecasting, Data analysis

INTRODUCTION

Efficient operation of a blood bank is crucial for any healthcare facility. In a typical blood bank, blood units are collected, tested for infectious diseases, processed into different components, stored under specified conditions, and issued to patients after cross-matching when requested by a physician [1]. However, the workflow of a blood bank is not as linear as explained. Characteristics unique to the blood banking system such as supply and demand uncertainty, the perishability of blood products, varying proportions of people with different blood types (A+, B+, AB+, O+, A-, B-, AB-, O-) in the global population, and the heterogeneity of patients contribute to the complexity of the system.

In the pursuit of net zero healthcare, blood banks play a vital role as their operations are energy- intensive, particularly due to refrigeration, testing, and disposal processes. Achieving net zero requires not only reducing direct energy consumption but also optimizing the entire supply chain to minimize wastage and unnecessary resource use. Here, technology and data analytics emerge as transformative tools. By using historical and real-time data, analytics can predict blood demand with greater accuracy, optimize inventory levels, and reduce overcollection, which in turn decreases energy usage and emissions associated with storage and discard of expired units.
In blood banking, uncertainty in the demand for blood components affects the decisions taken at each stage of the system. Determination of the number of blood collection drives, blood units to be collected, and blood components to be produced depends upon the forecasted demand for blood components. For synchronization of blood collection and component preparation, the requirement for various blood components must be known. This necessitates a well-specified forecasting model so that the demand for various blood components is known accurately and a sufficient amount of blood units in the inventory is maintained to avoid any shortage. A number of researchers have analysed several methods for forecasting the demand for blood. Time series models like ARMA, ARIMA, SARIMA, and the Exponential smoothing model, along with various machine learning algorithms, have been applied to estimate the demand for blood components of the blood banks of different countries [3-9]. However, Indian blood banking system is heterogeneous with various types of blood banks such as hospital blood banks, stand-alone blood banks and central blood banks. The objective of this study is to identify the suitability of forecasting models in the context of Indian blood banking system. Data from a representative Indian blood bank is collected and various types of forecasting methods are applied to determine the most accurate model.

MATERIAL AND METHODOLOGY

Data Collection

Data for this study is collected from a stand-alone community blood bank situated in Kolkata, India. The studied blood bank provids blood components to the nearby hospitals along with a number of thalassemia patients on a regular basis. The demand for Packed Red Blood Cells (PRBC) is considered for forecasting in this work as it is the most transfused blood component around the world as well as in the blood bank under study. The data regarding the daily demand of PRBC for various blood groups for the period of three years is obtained from the blood bank. A total of nine time series, including total demand for PRBC and individual demand of PRBC of eight blood groups are considered for forecasting. As the shelf-life of PRBC is 35-42 days, the demand is aggregated for a month for ease of planning. A summary description of the collected data is presented in Table 1.

Table 1: Mean and Standard Deviation of Monthly Demand

Blood Type	Percentage	Mean	Standard Deviation
A+	21.68	78	46.71
B+	35.52	128	80.87
O+	30.47	109	67.70
AB+	9.61	35	23.42

A-	0.63	2	2.04
B-	1.11	4	2.98
O-	0.74	3	1.99
AB-	0.23	1	1.07
PRBC		359	218.53

Time Series Decomposition

In order to identify the presence of the general trend and seasonality of the demand of PRBC for different blood groups, the data is decomposed into different components i.e. trend, seasonality, and residual. Fig. 1 displays the decomposed data for all nine time-series.

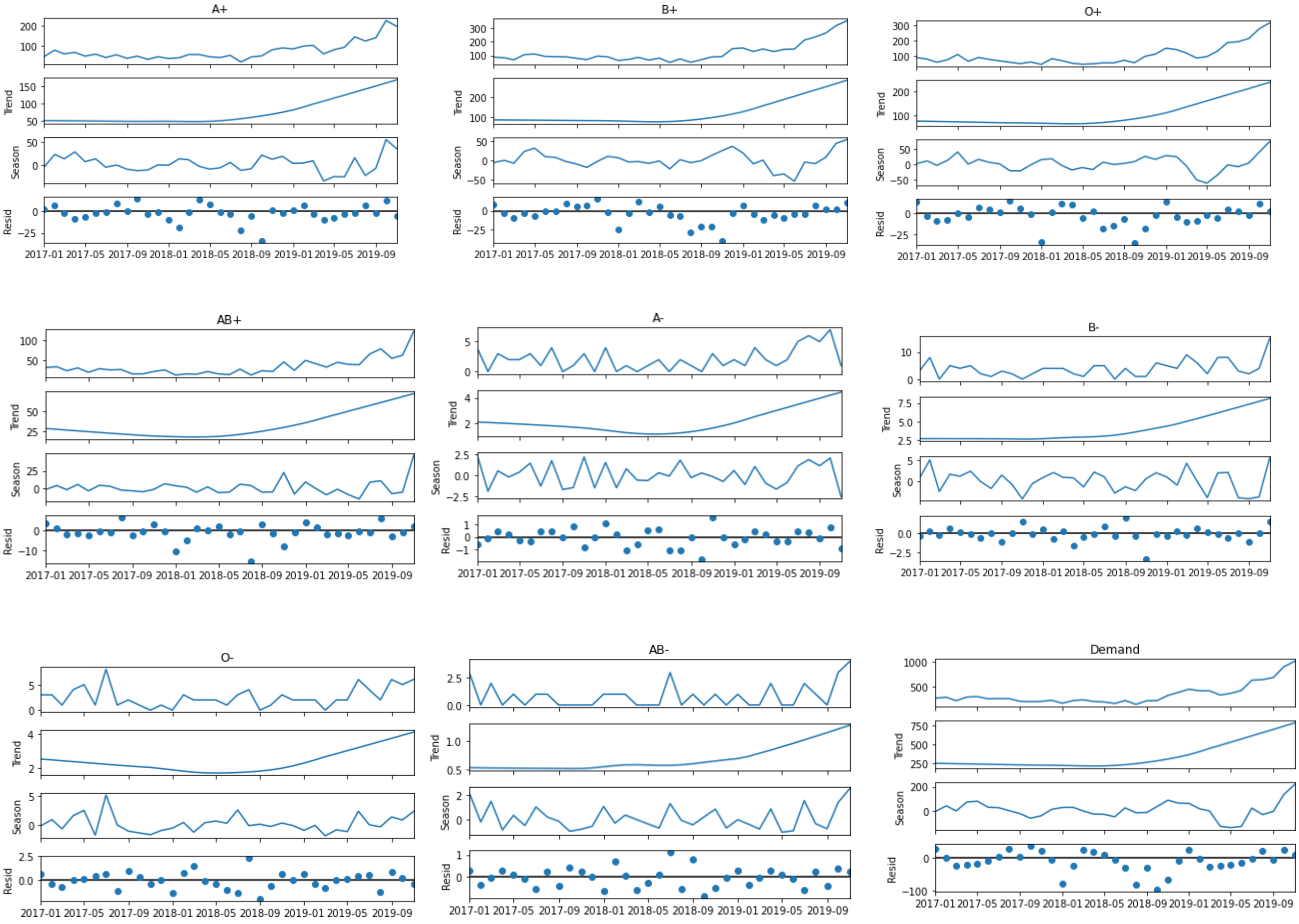

Fig. 1. Time Series Decomposition of Demand of Different Blood Types

Stationarity Analysis

The stationarity of the time-series data is checked by performing Augmented Dicky Fuller (ADF) test. Table II represents the ADF test results for the studied time series. From the presented result in Table 2, it is observed, the total demand for PRBC, demand for blood group A+, B+, O+, Ab+, A-, and B- is non-stationary, and the demand for PRBC of the blood group O-, and AB- is stationary in nature.

Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF)

As autocorrelation is the key statistic for time series analysis, ACF and PACF for the data are plotted to determine the correlation between different time periods [10]. Figure 2a-2j displays ACF and PACF of select time series. It is observed that for all the time series, the demand of one time period depends upon the immediately preceding time period.

Table 2: ADF Test Results

	Test Statistic	Critical Value at 1%	Decision
PRBC	2.005	-3.67	Non-stationary
A+	1.036	-3.71	Non-stationary
B+	1.646	-3.63	Non-stationary
O+	0.093	-3.63	Non-stationary
AB+	2.97	-3.65	Non-stationary
A-	-1.656	-3.63	Non-stationary
B-	-0.849	-3.66	Non-stationary
O-	-4.5	-3.63	Stationary
AB-	-6.323	-3.53	Stationary

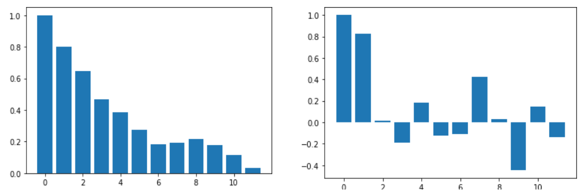

Fig 2a. ACF and PACF of Demand for PRBC for A+ Blood Type

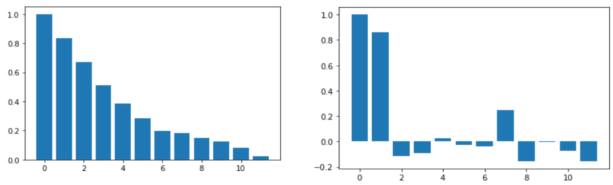

Fig 2b. ACF and PACF of Demand for PRBC for B+ Blood Type

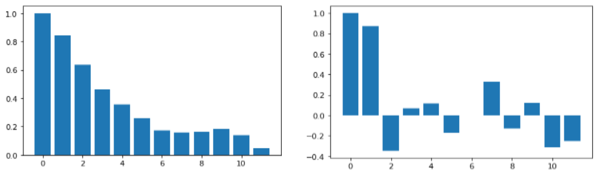

Fig 2c. ACF and PACF of Demand for PRBC for O+ Blood Type

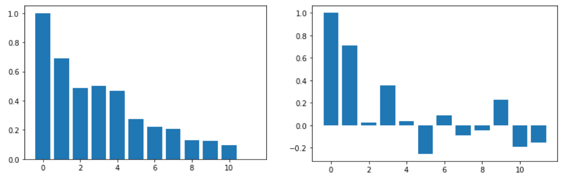

Fig 2d. ACF and PACF of Demand for PRBC for AB+ Blood Type

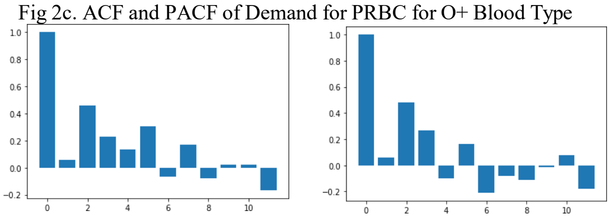

Fig 2e. ACF and PACF of Demand for PRBC for A- Blood Type
Type

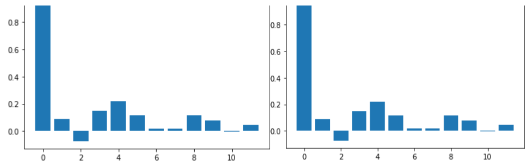

Fig 2f. ACF and PACF of Demand for PRBC for B- Blood

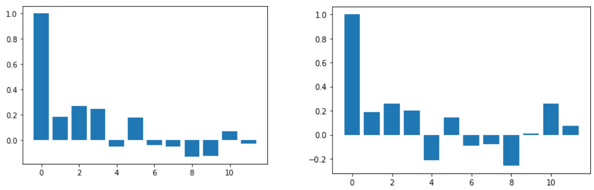

Fig 2g. ACF and PACF of Demand for PRBC for O- Blood Type

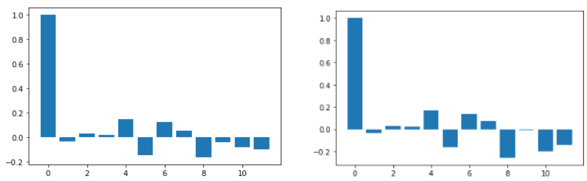

Fig 2h. ACF and PACF of Demand for PRBC for AB- Blood Type

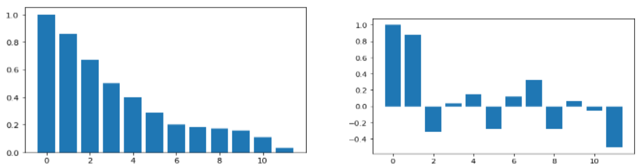

Fig 2j. ACF and PACF of Total Demand for PRBC

Evaluating Forecasting Models

After analyzing the time-series data for the demand of total PRBC and all the blood groups individually, four forecasting techniques i.e. Exponential Smoothing (ES), Winter's method (WM), Seasonal Auto-Regressive Integrated Moving Average (SARIMA), and Seasonal- Trend decomposition using LOESS and Auto-Regressive Integrated Moving Average (STL+ARIMA) are applied to the data. While applying SARIMA for forecasting, the grid search algorithm is used to determine the best model with minimum Akaike's Information Criterion (AIC). In order to forecast the demand of PRBC over a continuous period of time, the rolling horizon method is applied to all the techniques. Data for 24 months is selected as training data, and the models are validated by using the data from 12 months. Errors of the forecasting models are compared by calculating Root Mean Squared Error (RMSE). Python 3 is used to incorporate all the models. Table 3 displays the RMSE of all the applied forecasting models on the time series data for the demand of PRBC.

Table 3: Accuracy Measure (RMSE) for Forecasting Models

	ES	**WM**	**SARIMA**	**STL+ARIMA**
PRBC	107.67	151.63	113.4	178.63
A+	35.07	52.63	39.96	43.68
B+	32.6	70.17	38.4	67.01
O+	37.11	56.37	41.01	58.39
AB+	23.42	22.25	21.22	22.46
A-	2.97	3.25	3.52	3.39
B-	5.12	4.89	4.69	4.65
O-	1.96	2.66	2.99	2.94
B-	1.66	1.6	1.66	1.53

RESULTS

After analyzing the time series data for the demand for PRBC for different blood groups, several forecasting models were applied. It is found that despite using complex forecasting methods, the ES model is able to incorporate the uncertainty and forecast the demand most accurately for all the time series except for the demand of blood group AB-, B-, and AB-. In ES models, it is observed in Fig 3 that the value of the smoothing coefficient is closer to 1 which indicates that substantial adjustment of the forecasting error of the previous period is necessary. Table 4 indicates the best models with relevant parameters for forecasting the demand of PRBC of different blood groups.

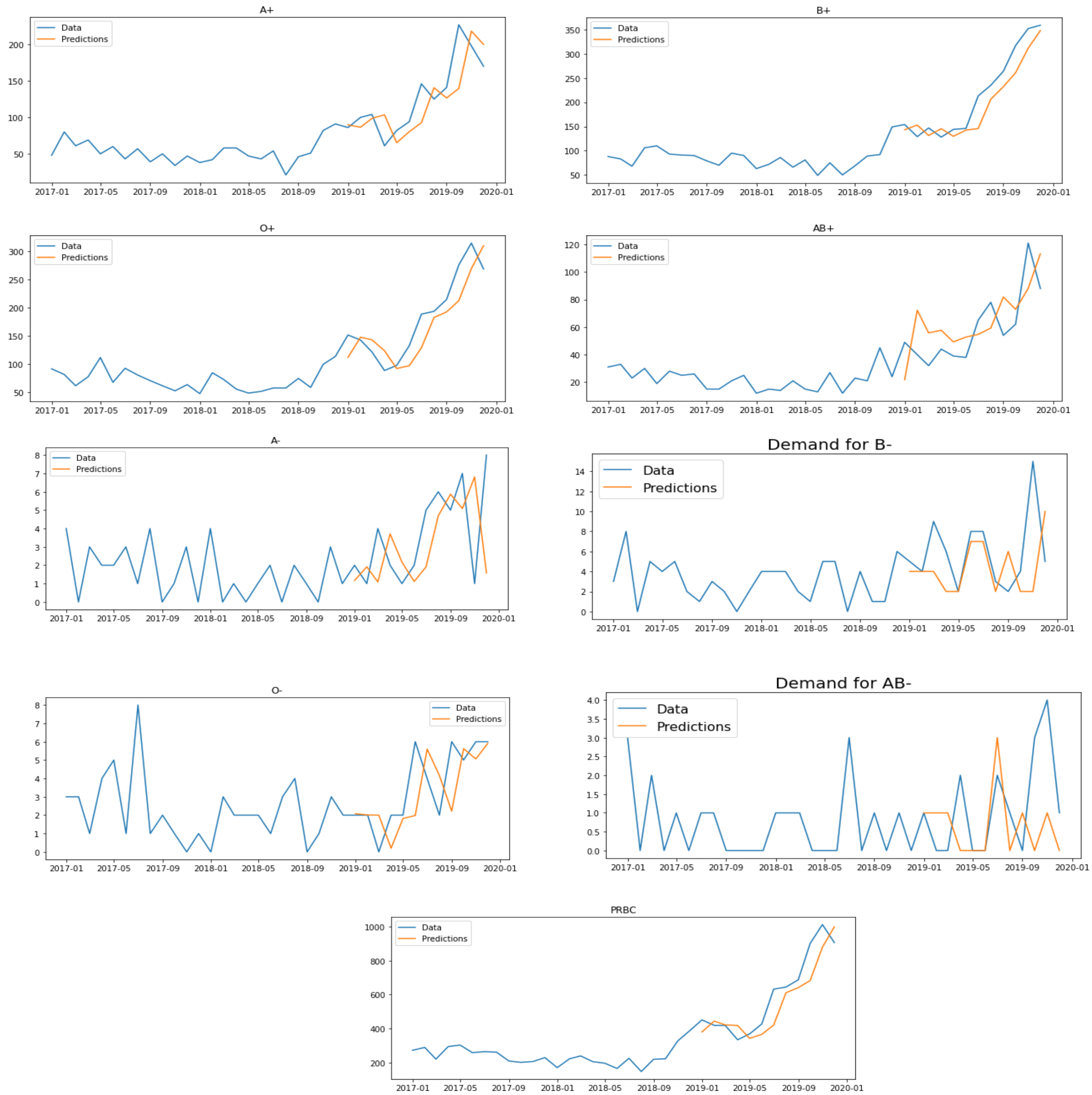

Fig 3. Forecasting Model Fitted to Demand of PRBC of Different Blood Types

Table 4: Best Forecasting Model for Demand of PRBC

Blood Type	Best Model	Parameters
PRBC	Exponential Smoothing	Smoothing constant α = 0.9
A+	Exponential Smoothing	Smoothing constant α = 0.9
B+	Exponential Smoothing	Smoothing constant α = 0.9
O+	Exponential Smoothing	Smoothing constant α = 0.9
AB+	SARIMA	Best Parameter (0, 1, 1)x(0, 1, 1, 12)
A-	Exponential Smoothing	Smoothing constant α = 0.9
B-	STL + ARIMA	Best Parameter (2, 1, 1)
O-	Exponential Smoothing	Smoothing constant α = 0.9
AB-	STL + ARIMA	Best Parameter (1, 1, 1)

CONCLUSION

The uncertain demand pattern of blood poses significant challenge in blood banking system. Blood is a critical resource for any medical system, a continuous supply of safe blood is necessary at any time. In the context of advancing toward net zero healthcare, reducing wastage, optimizing energy use, and improving operational efficiency are key priorities. This paper presents a comparative study of forecasting models to forecast the demand for PRBC. It is identified that despite using complex forecasting techniques, Exponential Smoothing model with a high smoothing coefficient performs the best. An accurate forecast of the demand for PRBC will help the blood bank management official to plan the number of blood collection drives to conduct and the number of donors needed for the next time period so that a proper level of inventory is maintained. By aligning collection and storage more closely with actual demand, unnecessary overcollection and wastage of blood units can be minimized. This reduction in wastage directly decreases the energy consumption and carbon emissions associated with refrigeration, testing, and disposal processes. In case of shortage in blood collection, with the knowledge of future demand, blood bank officials may incorporate replacement donations from the friends and family of a blood recipient and adjust the issuing policy accordingly. In conclusion, integrating data analytics-based forecasting into blood bank operations not only strengthens decision-making and supply reliability but also represents a step toward achieving net zero in healthcare systems through smarter, data-driven, and sustainable resource management.

REFERENCES

- Osorio, A. F., Brailsford S.C., Smith, H.K. 2015. A structured review of quantitative models in the blood supply chain: a taxonomic framework for decision-making, International Journal of Production Research, 53:24, pp 7191-7212 doi: 10.1080/00207543.2015.1005766
- Arcot, P. J., Kumar, K., Mukhopadhyay, T., Subramanian, A. 2020. Potential challenges faced by blood bank services during COVID-19 pandemic and their mitigative measures: The Indian scenario. Transfusion and Apheresis Science, 59(5), 10287 Goel, I., Sharma, S., Kashiramka, S. 2021. Effects of the COVID-19 pandemic in India: An

analysis of policy and technological interventions. Health Policy and Technology. 10, pp151–164. doi: 10.1016/j.hlpt.2020.12.001.

- Frankfurter, G. M., Kendall, K. E., Pegels, C.C. 1974. Management Control of Blood Through a Short-Term Supply-Demand Forecast System. Management Science. 21:4, pp 444- 452
- Filho, O. S. S., Cezarino, W., Salviano, G. R. 2012. A Decision-making Tool for Demand Forecasting of Blood Components. Proceedings of the 14th IFAC Symposium on Information Control Problems in Manufacturing Bucharest, Romania, 23-25 May, pp 1499-1504.
- Fortsch, S. M., Khapalova, E. A. 2016. Reducing uncertainty in demand for blood. Operations Research for Health Care. pp 16-28
- Li, N., Chiang, F., Down, D. G., and Heddle, N. M. 2021. A decision integration strategy for short-term demand forecasting and ordering for red blood cell components. Operations Research for Health Care. 29, 100290.
- Shokouhifar, M., Ranjbarimesan, M. 2022. Multivariate time-series blood donation/demand forecasting for resilient supply chain management during COVID-19 pandemic. Cleaner Logistics and Supply Chain. 5, 100078.
- Shih, H., Rajendran, S. 2019. Comparison of Time Series Methods and Machine Learning Algorithms for Forecasting Taiwan Blood Services Foundation's Blood Supply. Journal of Healthcare Engineering. 6123745. doi: 10.1155/2019/6123745
- Makridakis, S., Wheelwright, S. C., McGee, V. E. 1983. Forecasting: Methods and Applications, , 2nd Edition, John Wiley and Sons