Available online www.ejaet.com

European Journal of Advances in Engineering and Technology, 2018, 5(9): 690-696

Research Article ISSN: 2394 - 658X

Robust Forecasting for Total Blood Demand

Harish Kumar Taletia and Gitesh Vijay

Mechanical Engineering, Rajasthan Technical University, Kota, India, 324005
h.k.taletia@gmail.com

ABSTRACT

The aim of this paper was to apply different forecasting tools for comparative statistical study for the optimal value of total blood requirements of year 2017 on the basis of data from 2011 to 2016. With this objective, the study was undertaken with four appropriate forecasting models having statistical projections like Least-Square, Decomposition multiplicative, winter's, and time series models SARIMA were used to forecast of Total Blood Demand. The final results pertained to the both statistical and time series analysis revealed that the SARIMA (2, 1, 1) (1, 1, 1)$_{12}$ model had most efficient method of forecasting of TBD as given the data. The accuracy of the forecast is also applied with the least Akaike Information Criterion (AIC) of 854.89, Mean square error (MSE) have 98053.15, Mean Absolute Percent Error (MAPE) of 10.81% and Root Mean Square Error (RMSE) of 313.13 to evaluated forecast value for the year of 2017. The model was further validated by the Portmanteau test (L-Jung box test) with no significant correlation between residuals at different lags.

Key words: Decomposition, Least-square, Total Blood Demand, Holt-Winters', SARIMA, performance measures criteria

---

INTRODUCTION

In contemporary world, blood supply system has become more complex than ever before, since consumers' demand in terms of services and better products are changing. The field of Operations Management deals with designing, operating and continuously improving productive systems [1]. In terms of blood operation management, despite many achievements in modern world, people and organizations are still unable to address all practical reasons related to blood operation management [2]. operation management of blood is seen through operation research and distribution through the individual hospital blood banks (HBB) as well as regional blood banks (RBB) systems [3]. According to Times of India reviled a news that "Maharashtra tops in blood collection, but also wastes most" by right to information (RTI) query. In other state Uttar Pradesh (UP) and Karnataka wasted the maximum units of blood. In 2016-17, over 3 lakh units of fresh frozen plasma was discarded. These figures shows that the blood shortage is a chronic problem in our country mostly metro cities like Delhi and Mumbai said Kothari. India has an annual shortfall of three million units of blood. The shortage of blood, plasma or platelets has becomes the main cause for maternal mortality and deaths in mostly accident cases [4].Operation Management community through research work sees significant and rapid progress in the field of blood bank inventory management in the past few years accordingly. Operation management of blood through such research has been studied from both individual hospital blood banks HBB and RBB to show the operation pattern [3]. All types' blood are really needed.

This work has been done at "Krishna Rotary Blood Bank and Diagnostic Centre" which is situated at Kota, Rajasthan. It is associated with the government of Rajasthan. It established at Kota in October 2004 and successfully running by S. K. Sadhak Memorial Trust. It is certified with ISO 9001:2000. The vision and mission of Krishna rotary blood bank is unity, serviceability, creativity perfectness, accountability and availability 24 hours for blood and its components. Blood is a perishable product because of having limited sources and cannot be produce when increasing in demand. Blood transfusion has become one of the most essential part in dealing with modern health since it helps to save millions of lives every year. Blood is a unique product, having limited resources because of artificial substitutes are not yet to be found, making blood donations to be in great need [9].

LITERATURE REVIEW

This paper focused on appropriate forecasting methods applied by past researcher. A paper was to develop "Reducing uncertainty in demand of blood" by using ARMA and VARMA methodology. This Data management system is the repository for huge data, and overtime huge data deposit has come to be required for blood data management [4]. Simamora & Silitonga, (2014) in a simulated study highlighted the effectiveness of ARIMA in assessing the blood demand forecast in Indonesia. In case of blood transfusion series in addition to this, Raman, (2011) also has listed ARIMA and Exponential Smoothening to be effective in forecasting blood transfusion demand, however, only up to the case of time series data of 1 year, forecasting dengue incidence in Dhaka, Bangladesh[10]. A related study was presented a "Forecasting incidence of dengue in Rajasthan" using time series analysis (SARIMA) [11]. The **Least-Square** method forecasts on the other hand, is used for trend estimation, known for being the most cost-effective model in providing valuable information of the desired data, as compared to other models. It is known for determining regression in time series data [21], **Decomposition multiplicative** method is another forecasting method, whose method is based on trend, cyclic, seasonality and irregular variations forecasting system in time series analysis. The equation of follows the trend, de-trend, seasonal, de-seasonal factors *[22]. A comparative study was developed on SARIMA model with other relative techniques for Natural Rubber production in India* Since. The present study has been ultimate use of SARIMA and other techniques for time series data. The above underlined studies have a great supported for prepared this paper.

METHODOLOGIES

A) Trend (Least square) Method: It is a quantitative procedure for estimating mathematically the average relationship between the independent variables and dependent variables. It involves one depended variable in demand function. One of the best method Least-square method having straight line upward or downward trend is obtained according to time series. Least square method can be formulated as

$$\sum Y' = na + b\sum X$$

$$b = \frac{n\sum XY - (\sum X)(\sum Y)}{n\sum X^2 - (\sum X)^2}$$

$$a = \left(\frac{\sum Y}{n}\right) - b\left(\frac{\sum X}{n}\right)$$

Where Y= Total blood demand, Y'= Forecast demand, a= Intercept, b= slope, n=sample size (72), x= Time period,
The trend equation found that A relatively low R^2 =0.028 indicates that there is a lot of room for improvement in our estimated equation (Y'= -4.8689*X + 2172). The results of accuracy of forecast shows that Mean absolute error (MSE) was found 345119.9, Mean absolute percentage error (MAPE) was to be 23% and Root mean square error (RMSE) was to be 587.46.

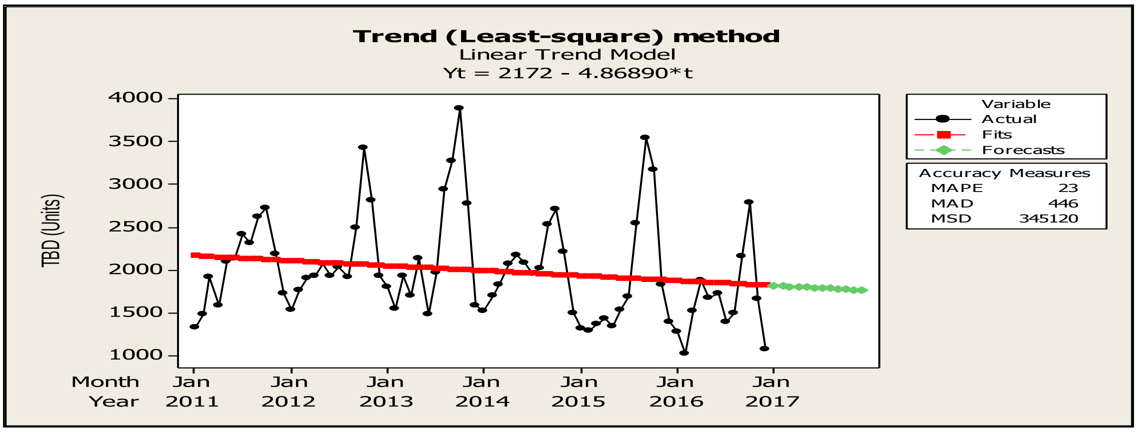

Fig. 1 Forecasting for year of 2017 by Least-square method

B) Multiplicative Decomposition Model

The multiplicative decomposition method is used for modelling that time series data containing seasonal (s), Trend (T), Cyclical (C) and Random (R) components. The error terms may be a random and seasonal components for any one season are the same in each year. This model used when the magnitude of the seasonal pattern vary with time, deepened upon the overall level of series. The magnitude of the seasonal pattern increases as the data values increases, and decreases as the data values decrease. Multiplicative Decomposition method follows equation that is

$Y_t = T_t * C_t * S_t * R_t$ Here, Y_t=denote the forecast series, T_t = denotes the linear trend, C_t =denotes cyclical, S_t =denotes seasonality. R_t=denotes random error, t=denotes the time period.

1) The first step is calculates L-step moving average to measure the combination of trend–cyclical (TC) components of data series, centred at the time period (t) for length (L) '12' (monthly series). Seasonal and random components, S_t and R_t will be eliminated by this step, this step is expressed as

$$M_t = \sum_{1}^{12} Y_t /12$$

2) Centre moving average (CMA) = Trend* Cycle

Ratio of actual to centre moving average

3) This step is calculate the trend and remove the trend means de-trend the data series by $T_t=a+b_t$

4)This step is calculate cycle term by dividing the moving average by the computed trend and seasonal ratio for de-seasonal the data. It can be expressed as $C_t=M_t/T_t$

5) This step is calculate seasonality by dividing the series by moving averages. It can be expressed as

$S_t=Y/(T*C)$

6) After obtaining the seasonal index, remove the seasonal effect from the time series, called de-seasonal the data, dividing the data by the seasonal indices. This is done by accomplished by dividing the K series by S_i where the values of S_1, S_2,… Sg are repeated. It can be expressed as

$R_t= K_t/S_t$

7) Develop the forecast using the trend equation and these factor. Table is showing the estimated Trend, seasonal and residual for calculate the forecast for year 2017

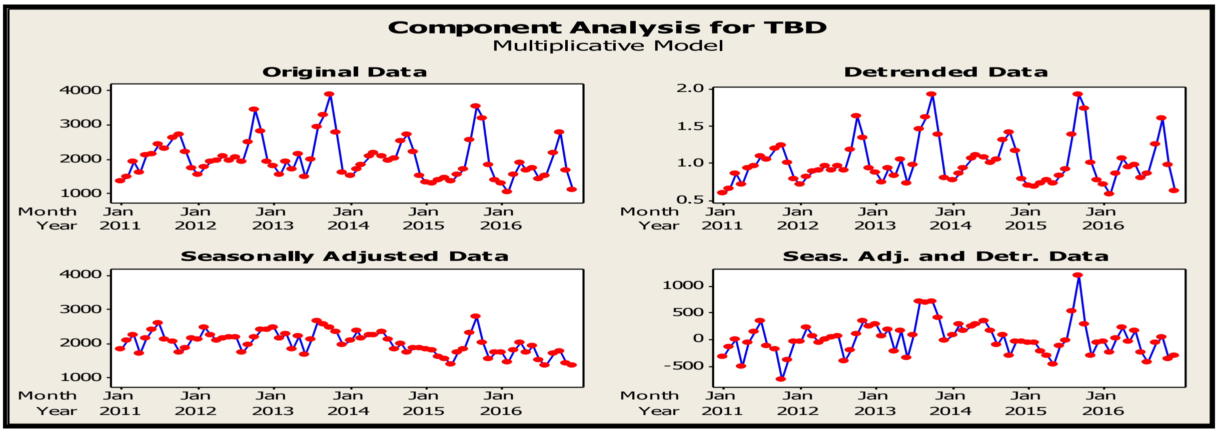

Fig. 2 Decomposition method (Parameters estimation)

Fig.2 shows that the various components de-trended data, seasonal adjusted data, and combination of both. Figure component analysis of total blood demand because of the de-trend data looks same as the original observations and in the seasonal adjusted data look quite different from the original observations that concluded the trend component is not present in data but a seasonal component were presented in data. The residual graph shows that the fitted values are under predicted in the part of 5th annual cycle, having a large positive residuals in this regions.

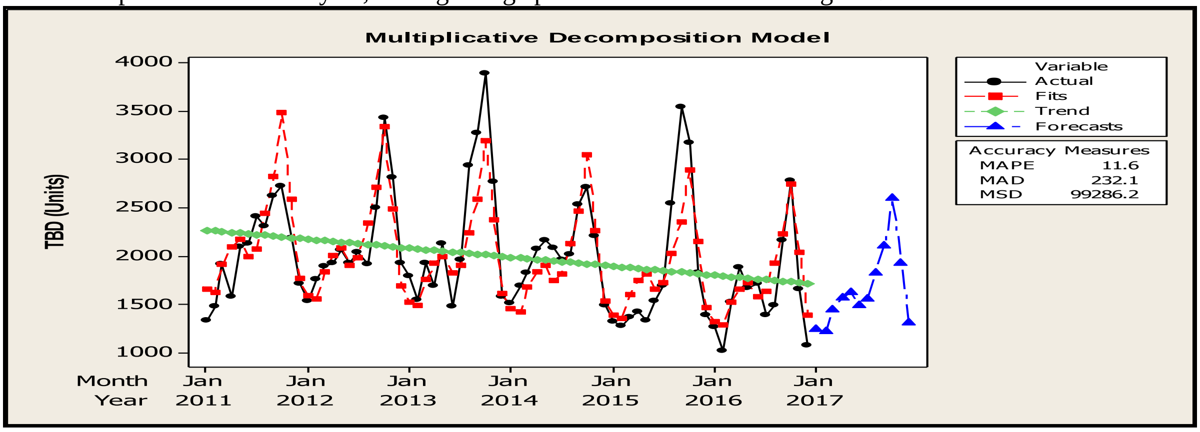

Fig. 3 Forecasting for year of 2017 by decomposition method

c) Winters' (Triple Exponential smoothing method)

This is an exponential smoothing method that consider both Trend and seasonality pattern in to account. The seasonal "figure" in a seasonal time series should repeat itself for each period (m). This method is applicable when the data series having trend as well as seasonality over time. It consist three factors like Level (l_t) Trend (T_t) seasonality (sn_t) to find the forecast. The Holt-Winters method can be easily applied on a single data series that show a typical seasonal as well as trend pattern, in that it allows the shape of the figure (but not the period) to vary with time.

This model can be represented as:

Forecast Equation $F_{t+m} = (sn_t + mb_t)l_{t-L+m}$

Level Equation $l_t = \alpha\left(\frac{y_t}{sn_{t-l}}\right) + (1-\alpha)(l_{t-1} + b_{t-1})$

Trend Equation $b_t = \beta(l_t - l_{t-1}) + (1-\gamma)b_{t-1}$

Whereas y_t is the observation, b_t is the trend factor,

sn_t is the seasonal index, F_{t+m} is the forecast at m periods ahead, α, β, and γ are smoothing constants between with 0 and 1, consider (0.2) optimal value, L= number of seasons in a year (L= 12 for monthly data)

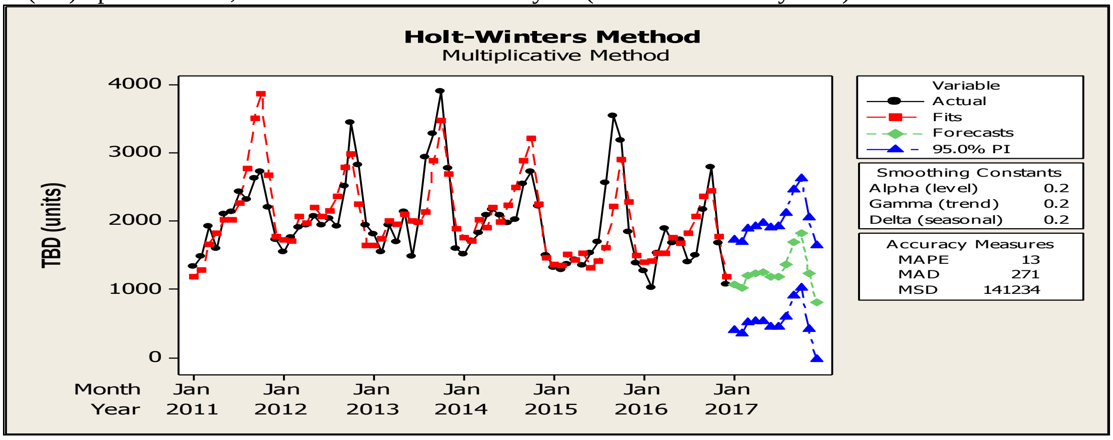

Fig. 4 Forecasting for year of 2017 by Holt-winters method

d) The Box-Jenkins (SARIMA) Methodology

This is a specified statistical, effective method and having iteration approach to solve the complex time series. It is a stochastic process that evolves with time. Box-Jenkins forecasting models are based on statistical concepts based upon multiplicative linear auto regressive integrate moving average method. It is an optimization process that having a wide spectrum of time series behaviour. Seasonal Auto Regressive Integrated Moving Average (SARIMA) is essentially the extension of ARIMA model used to detect the "seasonal" component in the time series data. In this non-seasonal and seasonal component are involved for selecting a better model. A standard notation for the SARIMA model as ARIMA (*p*, *d*, *q*) × (*P*, *D*, *Q*) $_S$.

SARIMA model can be formulated as:

$$\varphi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D y_t = \theta_q(B)\Theta_Q(B^s)\alpha_t y_t = (z_t - \mu)$$

$AR_{N\text{-}S}$ (*P*) – Non-Seasonal Autoregressive part of order p

$$\varphi_p(B) = (1 - \Phi_1 B - \cdots - \Phi_p B)$$

$MR_{N\text{-}S}$ (*q*) – Non-Seasonal Moving average part of order q

$$\theta_q(B) = (1 + \theta_1 B + \cdots + \theta_q B^q)$$

AR_S (*P*) – Seasonal Autoregressive part of order p

$$\Phi_P(B^s) = (1 - \Phi_1 B^s - \cdots - \Phi_p B^{sP})$$

MA_S (Q) – Seasonal Moving average part of order Q

$$\Theta(B^s) = (1 + \Theta_1 B^s + \cdots + \Theta_Q B^{sQ})$$

$$(1 - B^s)x_t = x_t - x_{t-s}$$

When = 1, the series is non-stationary,

= Non-seasonal AR component, = Seasonal AR component, θ_q = Non-Seasonal MR component

SARIMA is applicable on stationary time series with having seasonality in data series accomplished in four stages, they are following

Whereas y_t is the observation, b_t is the trend factor,

sn_t is the seasonal index, $F_{t+m}$is the forecast at m periods ahead, α, β, and γ are smoothing constants between with 0 and 1, consider (0.2) optimal value, L= number of seasons in a year (L= 12 for monthly data)

a) Model Identification.
b) Model Estimation.
c) Diagnostic Checking.
d) Forecasting.

Box-Jenkins methodology assumption follows some tests results for data compatibility

Table –1 Assumptions for Box-Jenkins Methodology

Statistical Tests	Score	C.V.	P-Value	Sig. (5.0%)
(a) White noise test	30.30	3.84	0.0%	False
(b) Normality Test				
Jarque-Bera	14.35	5.99	0.1%	False
Doornik Chi-Square	17.65	5.99	0.0%	False
Shapiro –wilk	.927	5.99	0.0%	False
(c) Stationary Test				
Augmented Dickey-fuller	-1.0	-2.0	28.41%	False

This table shows that the white test is not present in data which show the data is applicable for box-methodology but not stationarity (presence of unit root). The first order differencing will make the data stationarity and box-cox transformation will make the data normal which must require for this method. After this formulation various test results is presented below

Table –2 Statistical test result (After Box-Cox transformation-2)

Statistical Tests	Score	C.V.	P-Value	(5.0%)
(a) Normality Test (after Box-cox Transformation 2)				
Jarque-Bera	0.21	5.99	90.2%	True
Doornik Chi-Square	0.02	5.99	98.9%	True
Shapiro–wilk test	0.99	5.99	90.1%	True
(b) Stationary Test (after Differencing of order 1)				
Augmented Dickey-fuller	-6.6	-2.0	0.1%	True

Model identification is based on autocorrelation and partial auto correlation, presented below

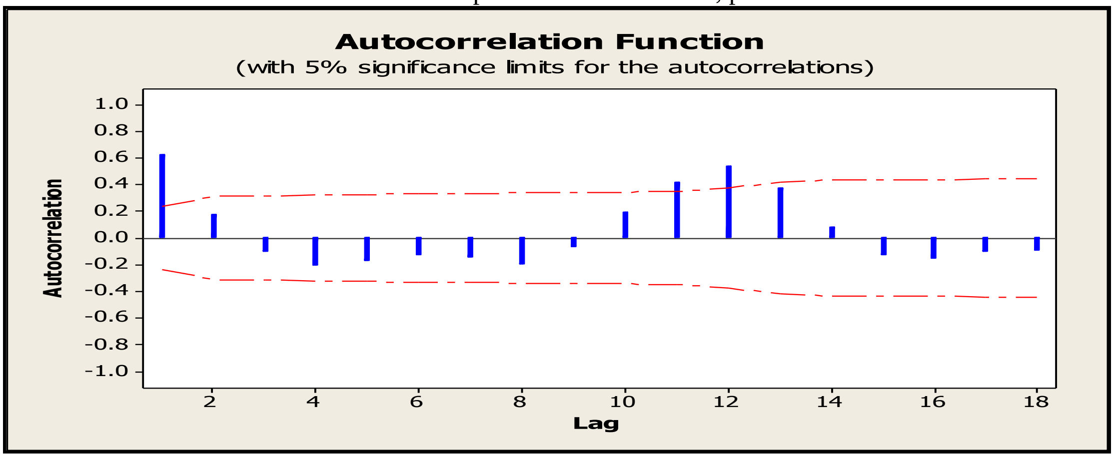

Fig. 5 ACF analysis for model selection of MR term

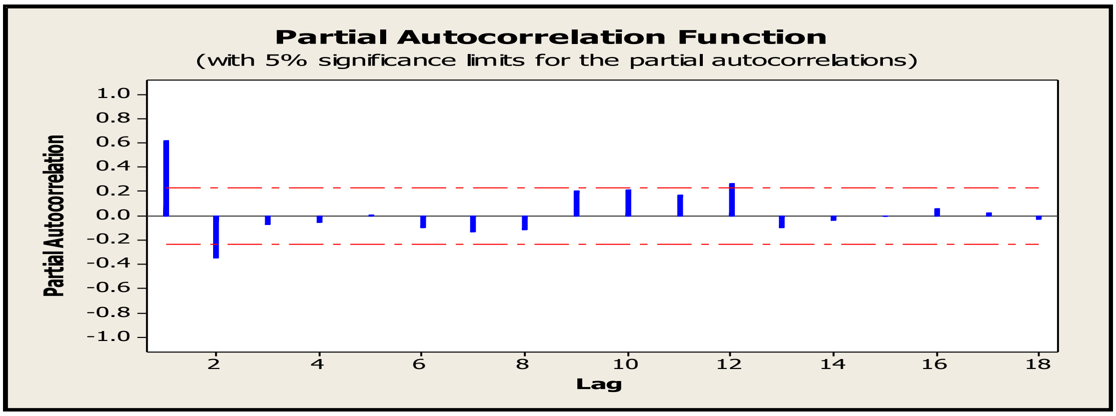

Fig. 6 PACF analysis for model selection of AR term

The ACF and PACF suggested that the five SARIMA models like SARIMA$(2,0,1)(1,1,1)_{12}$, SARIMA$(2,1,1)(1,1,1)_{12}$,SARIMA$(2,0,1)(0,1,1)_{12}$, SARIMA $(2,0,1)(1,1,0)_{12,}$ SARIMA $(1,0,1)(1,1,0)_{12}$ SARIMA $(1,0,1)(1,1,0)_{12}$ estimated for total blood demand for 2017. Therefore, to detect the same, this section presents the SARIMA modeling using the existing blood demand data.

SARIMA model selection shown in table below

Table-3 Model Selection amongst various types of SARIMA models

SARIMA(p,q,d)(P,D,Q)s	LLF	AIC	SBC
SARIMA$(1,1,1)(1,1,1)_{12}$	**-840.92**	**854.92**	**872.34**
***SARIMA$(2,1,1)(1,1,1)_{12}$**	**-838.89**	**854.89**	**871.37**
SARIMA$(2,0,1)(0,1,1)_{12}$	**-853.64**	**865.64**	**878.11**
SARIMA $(2,0,1)(1,1,0)_{12}$	**-853.02**	**865.02**	**877.48**
SARIMA $(1,0,1)(1,1,0)_{12}$	**-854.30**	**864.30**	**874.69**
SARIMA $(1,0,1)(0,1,1)_{12}$	**-853.64**	**863.64**	**874.03**

Table-3 presents the goodness of fit model with appropriate Akaike information criterion (AIC) and Log likely hood function (LLF)and Schwarz's Bayesian Criterion (SBC) that is SARIMA$(2,1,1)(1,1,1)_{12}$ has been found optimal model having least (AIC), Schwarz's Bayesian Criterion higher Log likely hood function (LLF) than any other models.

SARIMA(2,1,1) $(1,1,1)_{12}$ model equation can be written as:

$$Y_t = Y_{t-12} + (Y_{t-1} - Y_{t-13}) * 872e_{t-1} - .444\ e_{t-12} + .872 * .444e_{t-13} + \propto_t + .872 \propto_t$$

CONCLUSION

The forecasts estimated for year 2017 by all the models are significantly different from each other. Amongst these models, SARIMA $(2,1,1)(1,1,1)_{12}$ have least value of MSE (98053.15), MAPE(10.81%) and RMSE (313.13) appeared to be the best model. Considering the results from the analysis, the work has concluded SARIMA $(2,1,1)(1,1,1)_{12}$ model to be the most efficient model with accuracy measure with least MSE, MAPE and RMSE values in forecasting blood demand for reducing wastage, inventory control means preventing excess and shortage of blood. SARIMA (2,1,1)(1,1,1)12 model is regarded as the robust model amongst all other forecasting methods criteria for estimating the Total blood demand**.**

REFERENCES

[1]. R. S. Russell and B. W. Taylor, "Operations Management along the Supply Chain", Sixth Edit. New Delhi, India: John Wiley & Sons, 2008.

[2]. D. R. Tobergte and S. Curtis, "Supply chain network operations management of a blood banking system with cost and risk minimization," J. Chem. Inf. Model., vol. 53, no. 9, pp. 1689–1699, 2013.

[3]. H. Lowalekar and N. Ravichandran, "Management in Blood Banks" in Case Studies in Operations Research: Applications of Optimal Decision Making, K. G. .Murty, Ed. New York: Springer, 2014.

[4]. Times of India" an article is published RTI query" http://timesofindia.indiatimes.com/city/mumbai/maha-topsin-blood-collection-but-also-wastes-most-reveals-rti-query/articleshow/58331352.cms

[5]. M. Singhal, M. Patel, D. Kapoor, and D. Mittal, "Full Length Research Paper A research analysis on blood component usage and wastage in blood bank and blood component center," J. Physiol. Pathophysiol., vol. 4, no. March, pp. 23–28, 2013.

[6]. C. M. Karuppan, M. R. Waldrum, and N. E. Dunlap, Operations Management in Healthcare: Strategy and Practice. New York: Springer Publishing Company, 2016.

[7]. T. C. Lindstrøm and A. Røsvik, "Ambiguity leads to uncertainty: Ambiguous demands to blood donors," Scand. J. Caring Sci., vol. 17, no. 1, pp. 74–77, 2003.

[8]. J. Chapman and J. MacPherson, "Unlocking the essentials of effective blood inventory management," Transfusion, vol. 47, no. SUPPL. 2, pp. 190–196, 2007.

[9]. Delhi Government," Health family welfare: Blood Bank,"http://delhi.gov.in/wps/wcm/connect/DoIT_Health/health/home/blood+bank, 2016.

[10]. Powerful Forecasting with MS Excel sample http://www.xlpert.com/ebook-Forecasting-MS-Excel.html

[11]. An Introductory Study on Time Series Modeling and Forecasting by R. Adhikari. https://gejza.nipax.cz/_media/stochasticke_procesy:1302.6613.pdf

[12]. J. Lee, "Univariate time series modelling and forecasting (Box-Jenkins Method)", Econ 413, lecture 4, by V. Kumar. www.ijesrt.com/issues%20pdf%20file/Archive-2016/December-2016/79.pdf

[13]. "Reducing uncertainty in demand for blood" http://www.sciencedirect.com/science/article/pii/S2211692315300047

[14]. SARIMA Model for Natural Rubber production in India by P. Arumugam*, V Anithakumari, 2013. https://ijcttjournal.org/Volume4/issue-8/IJCTT-V4I8P116.pdf
[15]. Forecasting Incidence of Dengue in Rajasthan, Using Time Series Analyses https://link.springer.com/chapter/10.1007/978-3-319-25939-0_5
[16]. Forecasting dengue incidence in Dhaka, Bangladesh: A time series analysis https://link.springer.com/chapter/10.1007/978-3-319-25939-0_5
[17]. “Red Gold: The Epic Story of Blood”. 2002. Public Broadcasting Service. 15 Oct. 2005. http://www.pbs.org/wnet/redgold/resources.html
[18]. "Econometric Forecasting" homepage.univie.ac.at/robert.kunst/prognos.pdf
[19]. K. Nasserinejad, W. de Kort, M. Baart, A. Komárek, J. van Rosmalen, and E. Lesaffre, “Predicting the hemoglobin levels in whole blood donors using transition models and mixed effects models”. BMC Med. Res. Methodol., vol. 13, p. 62, 2013.
[20]. M. Valipour, “Long-term runoff study using SARIMA and ARIMA models in the United States,” Meteorol. Appl., vol. 22, no. 3, pp. 592–598, Jul. 2015.
[21]. Development of temporal modeling for prediction of dengue infection in North-eastern Thailand. http://www.sciencedirect.com/science/article/pii/S1995764512600340
[22]. S. M. Fortsch and E. A. Khapalova, “Reducing uncertainty in demand for blood,” Oper. Res. Heal. Care, vol. 9, pp. 16–28, 2016