International Journal of Statistics and Systems
ISSN 0973-2675 Volume 18, Number 1 (2023), pp. 1-10

http://www.ripublication.com

Air Traffic Forecast for Air India Express

Mounika Panjala and Bhatracharyulu N.Ch.

Department of Statistics,
University College of Science, Osmania University.
E-mail: mounikapanjala123@gmail.com and dwarakbhat@osmania.ac.in

Abstract

In this paper, an attempt is made to forecast the Air India express traffic flow in Domestic and International flights using Auto regressive integrated moving average, Seasonal auto regressive integrated moving average and Neural networks based on the data during the period 2009-2019. The accuracy of the models is evaluated using root mean square error, mean absolute error and mean absolute percentage error.

Key words or Phrases: ARIMA, Seasonal ARIMA, Neural Networks, RMSE, MAE, MAPE.

1. INTRODUCTION

Civil aviation industry is one of the key service sectors of the economy that is continually rising and significantly boosts the wealth of both established and developing countries. It is crucial in moving people or things from one place to another through domestic and international airways. In general, about 65%-70% of total passengers travel in domestic flights and remaining 30%-35% in overseas flights.

Many studies have employed time series models to forecast the number of air passengers. Zhang et al. [1998] used Neural Network algorithm to forecast the air traffic. Lim and Mc Aleer (2002) and Kulendran and Witt (2003) used time series models to predict the demand for international business tourist travel to Australia using auto regressive integrated moving average (ARIMA) model. Chen et al. (2009) used Seasonal Autoregressive Integrated Moving Average (SARIMA) model to forecast arrivals of incoming air travellers to Taiwan.

In this paper, an attempt is made to forecast the Air India express traffic flow in Domestic and International flights using ARIMA, Seasonal ARIMA and Neural networks based on the data during the period 2009-2019.

2. MODELS FOR INTERNATIONAL TRAFFIC

In this section, various time series models like ARlMA, Seasonal ARIMA and Neural Networks are used to analyze the data considered and evaluated. The analysis is done for number of passengers travelled through Air India Express through both Domestic and International flights.

1. Let X_t be the observed value of the time series at time 't', t = 1, 2,..., k ; various time series models used to fit the data are presented below.
   (i) ARIMA (p,d, q) model where p is lagged dependent variables, q is lagged error terms and d is an order of differencing. The model can be expressed as $\phi(B)(1-B)^d X_t = \theta(B)\, a_t$ where $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B_p$; and $\theta(B) = 1 + \theta_1 B + \cdots + \theta_q Bq$. The resulting ARIMA model is with parameters p =1, q= 1 and d=0 and the AR1 estimated parameter value is 0.0024 with Standard error 0.0962 and the ARIMA model is,
   $$(1 - 0.0024\, B)\, X_t = a_t$$
   (ii) SARIMA model with parameters $(p, d, q)(P, D, Q)_s$ is $\phi(B)\, \Phi(B_S)(1-B)^d (1-B_S)^D Xt = \theta(B)\, \Theta(B_S)\, \varepsilon_t$ where: ε_t be the white noise with mean zero and variance σ^2; 'd' be the number of regular differences; $(B) = 1 - \phi_1 B - \cdots - \phi_p B_p$ and $\Phi(B) = 1 - \Phi_1 B - \cdots - \Phi_P B_P$; $(B) = 1 + \theta_1 B + \cdots + \theta_q Bq$ and $\Theta(B) = 1 + \Theta_1 B + \cdots + \Theta_Q B_Q$. SARIMA Model (1,0,0) (2,1,0) [12] is fitted with AR1, SAR1 and SAR2 parameters 0.927,-0.4233 and-0.2724 with Standard error in the parameter estimation is 0.0359, 0.0934 and 0.0857. And the model is,
   $$(1 - 0.927B)(1+0.4233B+0.2724\, B_2)(1 - B_S)\, X_t = 0$$
   (iii) Neural Network Auto Regressive Model (12,1,2) [12] is fitted with variance 5.35E+08.
2. A comparison among the three models with respect to RMSE, MAPE, MASE, AIC, BIC values is tabulated in table 2.1.

Table 2.1: Comparison of ARIMA SARIMA & NN Models for International traffic

	ARIMA (1,1,0)	ARIMA (1,0,0) (2,1,0) [12]	NNAR (12,1,2) [12]
RMSE	37131.93	40173.77	22560.04
MAPE	26.63253	71.28765	36.75404
MASE	0.176162	0.264655	0.145112
AIC	2558.41	2331.96	
BIC	2561.08	2342.22	

3. Comparison of Residual Plots for ARIMA, SARIMA and NN is presented in fig1

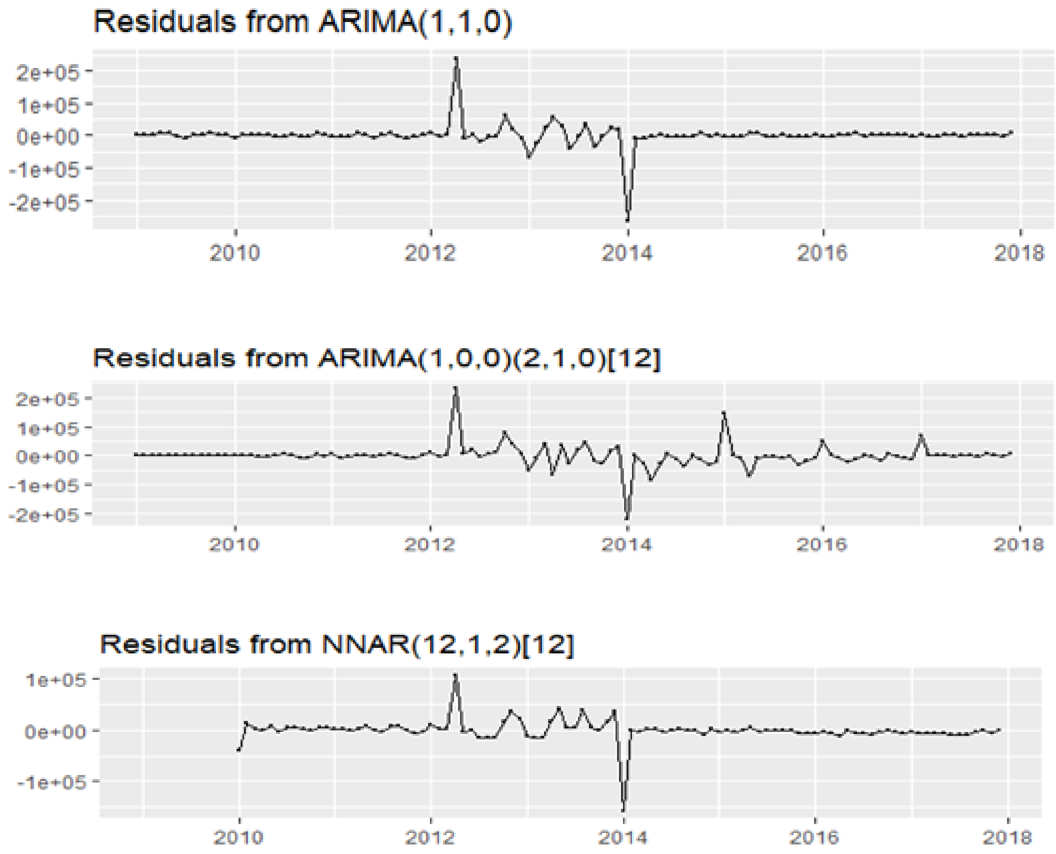
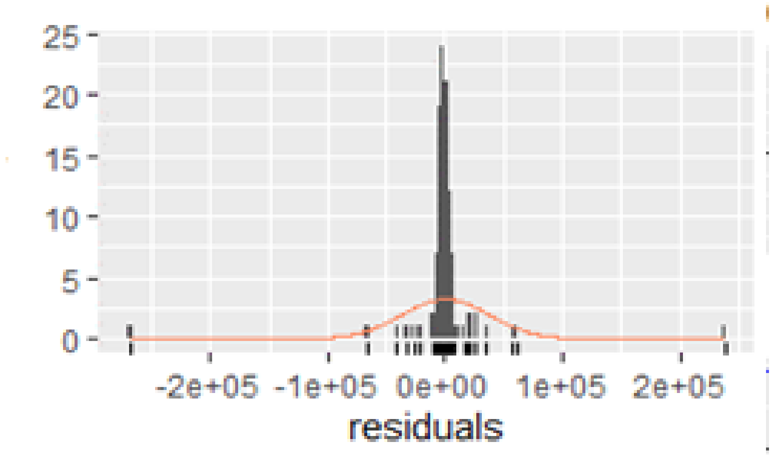


Figure 1: Residual Plots for ARIMA, SARIMA and NN

4. Comparison of residuals of the three models is presented in fig 2



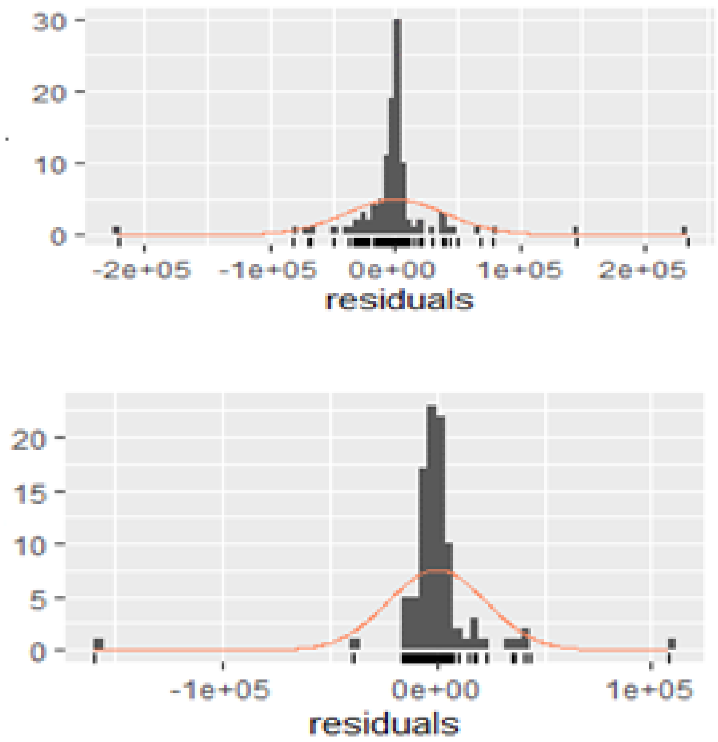
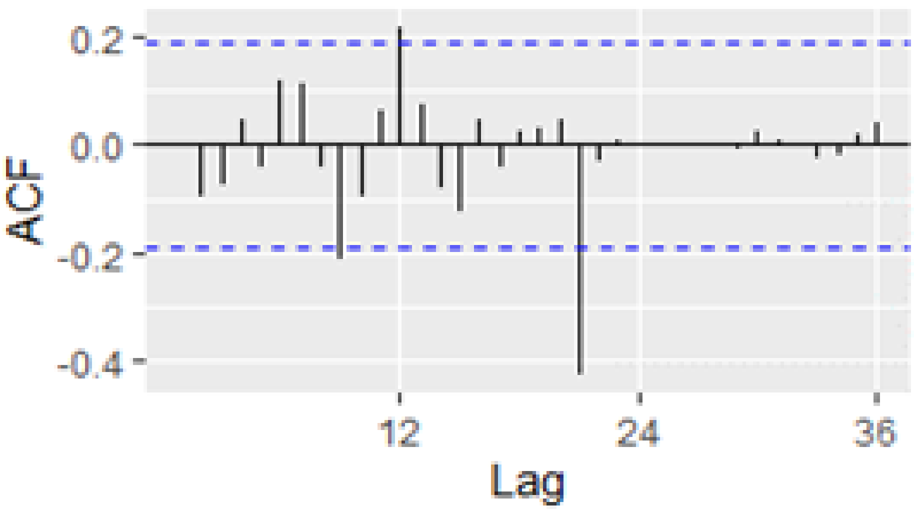


Figure 2: Residual Plots for ARIMA, SARIMA and NN

5. Comparison of lag values for the three models



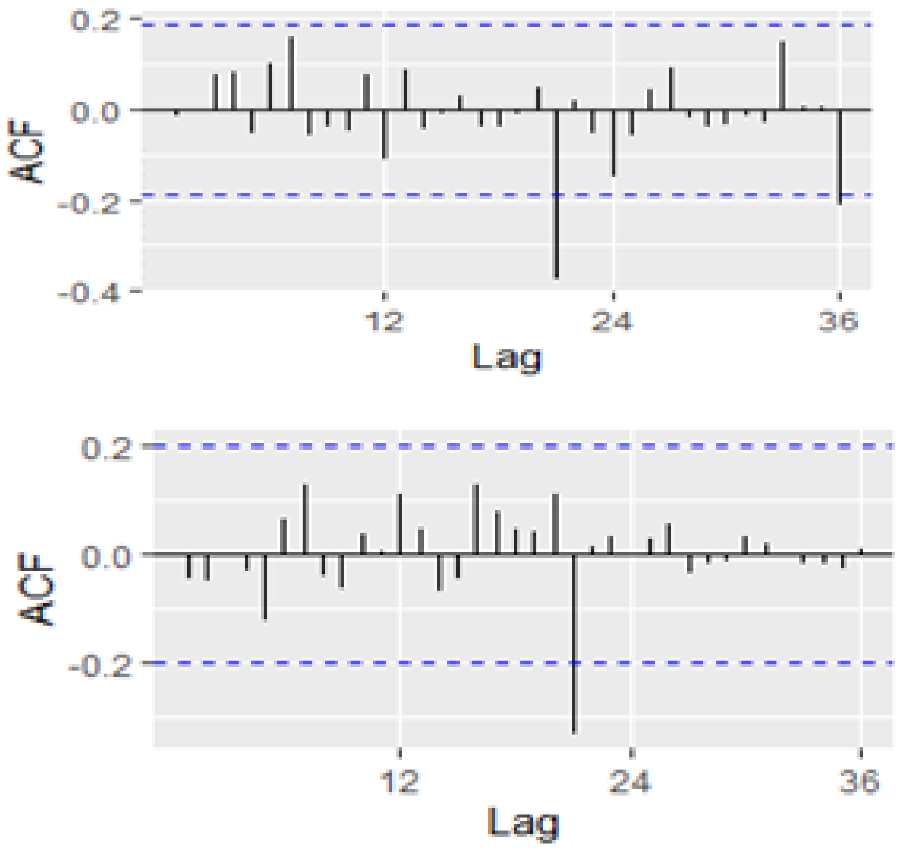


Figure 3: Lag values for ARIMA, SARIMA and NN

6. Comparison of forecasted values with various models is presented below.

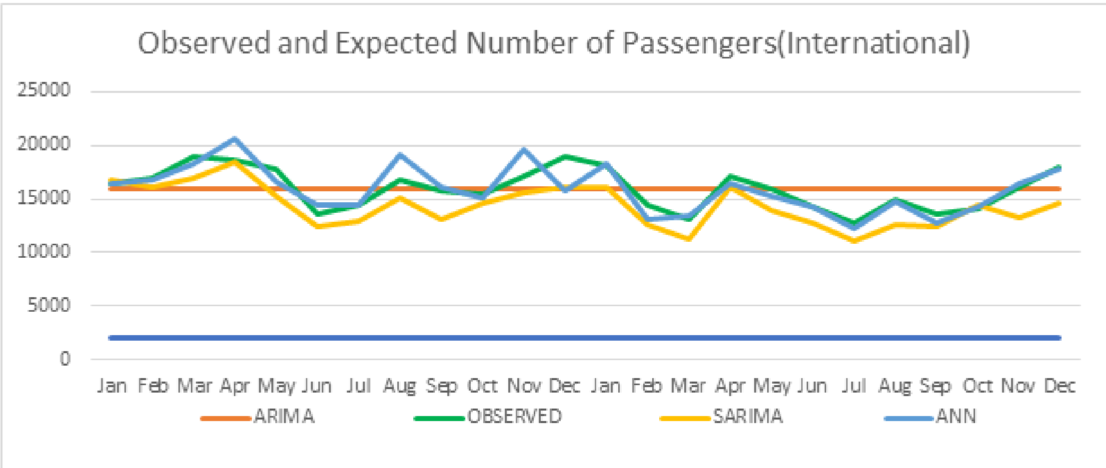


Figure 4: Comparison of forecasted values with various models

3. MODEL BUILDING FOR DOMESTIC TRAFFIC

1. ARIMA (0,1,2) Model is fitted with MA1 and MA2 estimated parameters-0.3409 and-0.2887 with standard errors 0.0953 and 0.0968. And the model is

$$(1-B)\,X_t = 1\text{-}0.3409\,B\text{-}0.2887B_2$$

2. SARIMA model for domestic traffic ARIMA (0,1,1) (0,1,1) [12] with MA1 and MA2 estimated parameters-0.3664 and-0.5917 with Standard errors 0.1018 and 0.1106.

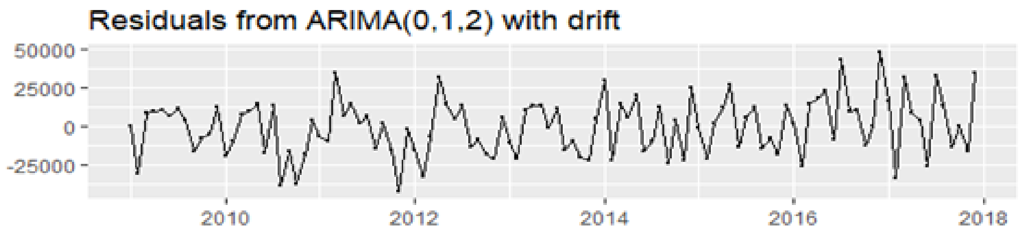
$$(1-B)(1-B_S)X_t = (1\text{-}0.3664B)(1\text{-}0.5917B)\,\varepsilon_t$$

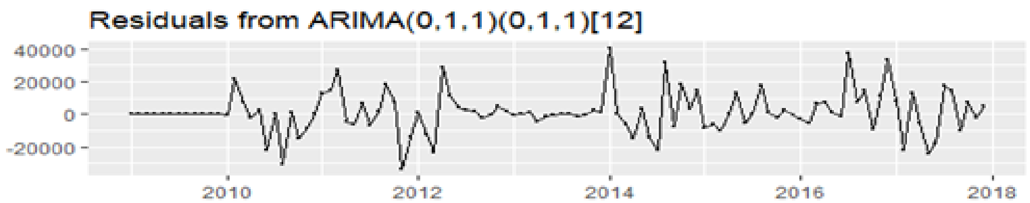
3. NNAR(12,1,2)[12] is fitted with variance 92719331.

Table 3.1: ARIMA Model for Domestic

	ARIMA (0,1,2)	ARIMA (0,1,1) (0,1,1) [12]	NNAR(12,1,2)[12]
RMSE	18304.24	12693.39	10144.91
MAPE	7.532806	4.023371	3.787076
MASE	0.676249	0.373292	0.337203
AIC	2413.47	2088.37	
BIC	2424.16	2096.04	

4. Comparison of residuals





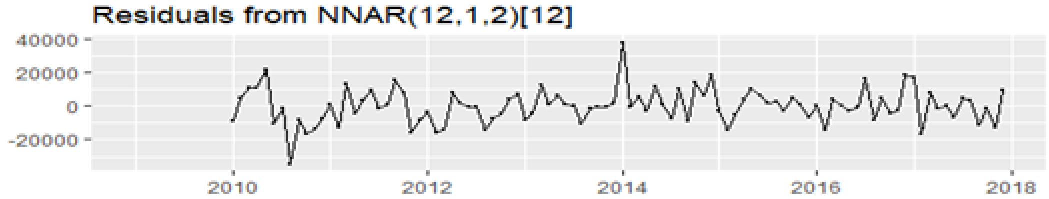


Figure 1: Residual Plots for ARIMA, SARIMA and NN

5. Comparison of Lag values

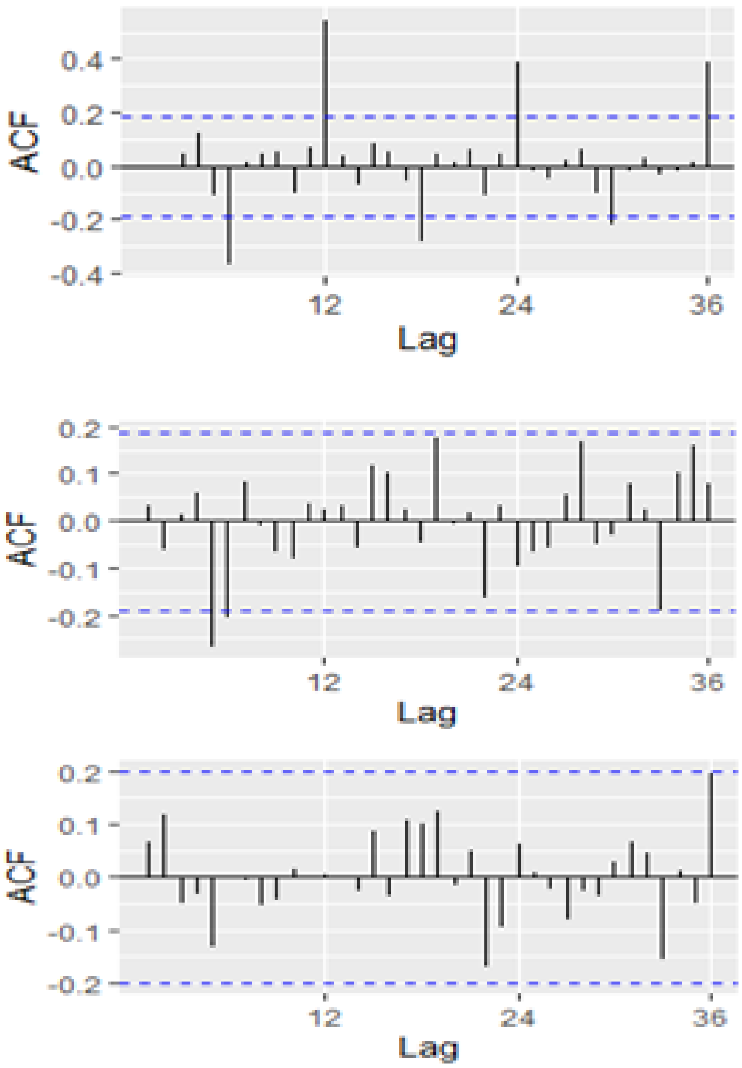


Figure 2: Residual Plots for ARIMA, SARIMA and NN

6. Comparison of Time sequence plot of residuals

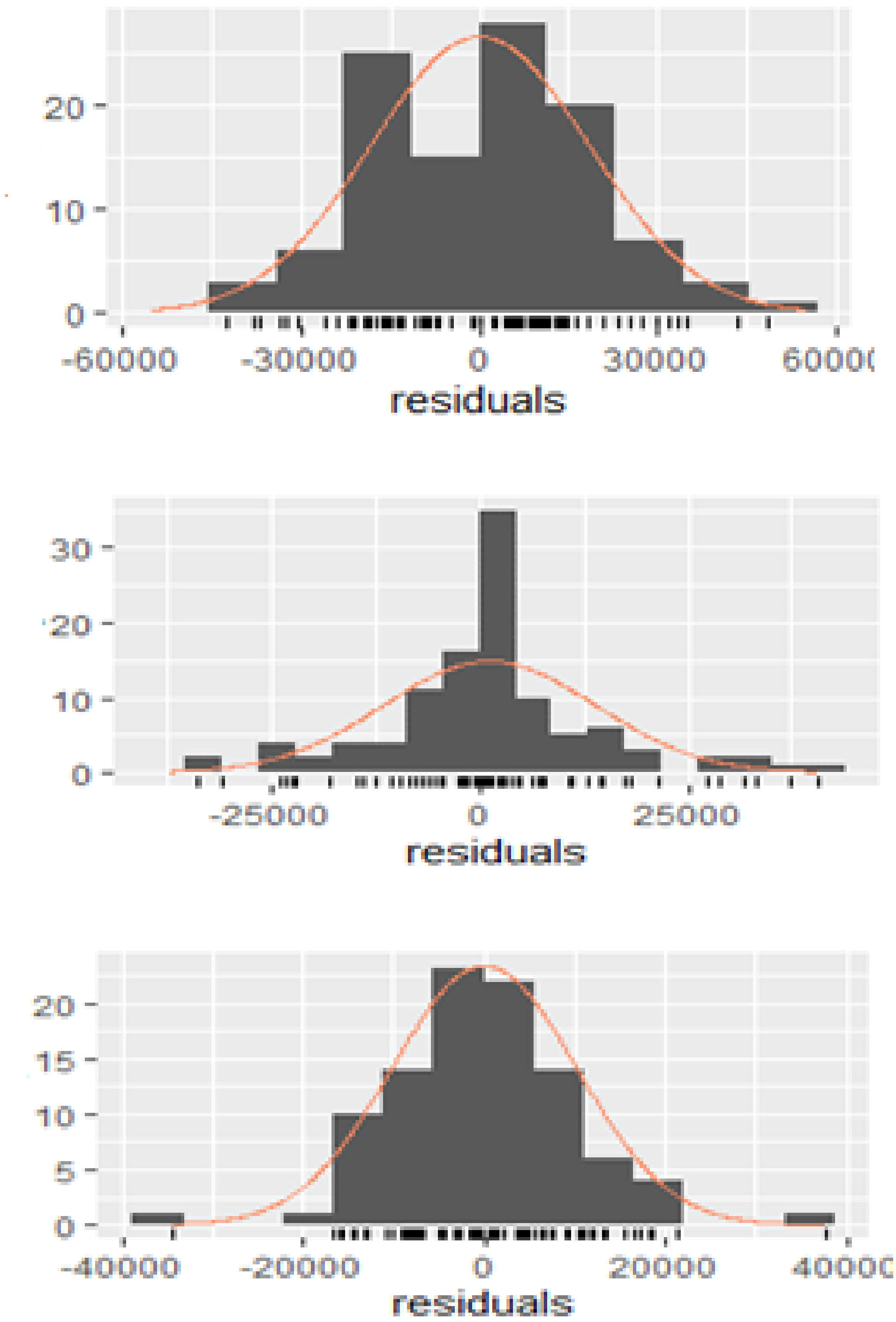


Figure 3: Residual Plots for ARIMA, SARIMA and NN

7. Comparison of forecasted values

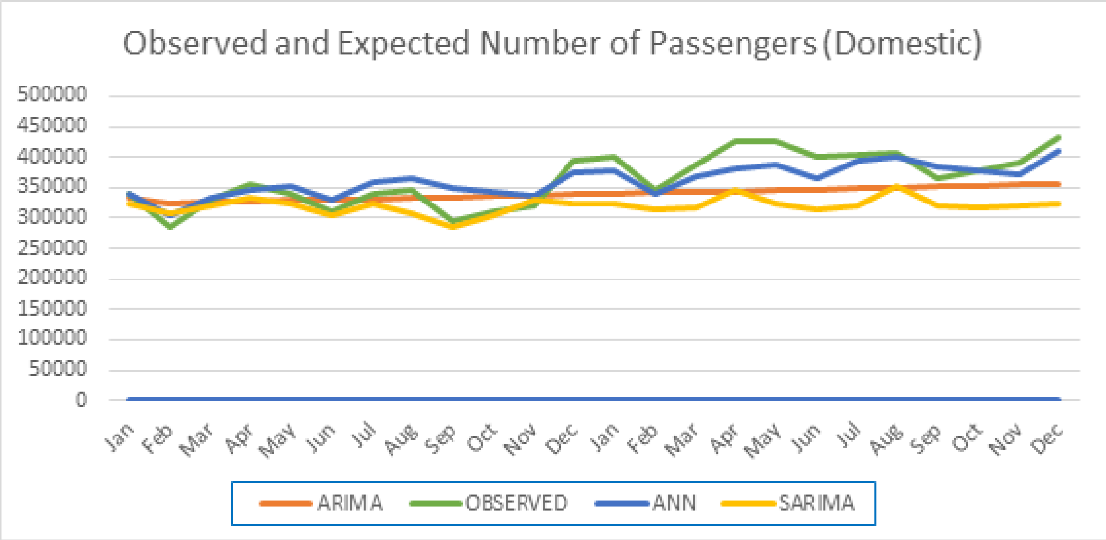


Figure 4: Comparison of forecasted values with various models

Remarks:

1. The mean of the residuals is close to zero in NN residual plot as compared to ARIMA and SARIMA residual plots.
2. There is no significant correlation in the residual series of NN, where as in ARIMA and SARIMA there are few spikes in the graphs indicating a correlation between the lags.
3. The Histogram suggests that the residuals are close to Normal in the NN plots.
4. The White Noise process 'a_t' is assumed to be i.i.d Normal $(0, \sigma^2)$ can be verified from the time sequence plot of residuals.
5. The autocorrelation of lag 1 is insignificant, implies that the 'a_t' are independent.
6. The NN model completely outperformed the ARIMA and SARIMA forecasts all the error terms for both the data sets.
7. From the above it can be noted that the estimated number of Air passengers through the ANN Model is more appropriate and near to the original values when compared with the ARIMA and SARIMA Models for both Domestic and International datasets.

REFERENCES

1. Chen, Ching-Fu., Yu-Hern, C. and Yu-Wei, C. "Seasonal ARIMA forecasting of inbound air travel arrivals to Taiwan." Transport-metrica 5.2 (2009): 125-140.

2. Kulendran, Nada, and Stephen F. Witt. "Forecasting the demand for international business tourism." Journal of Travel Research 41.3 (2003): 265-271.
3. Lim, C. and McAleer, M. "Time series forecasts of international travel demand for Australia." Tourism Management 23.4 (2002): 389-396. vol 23(4), pp 389-396.
4. Zhang, Guoqiang, B. Eddy Patuwo, and Michael Y. Hu. "Forecasting with artificial neural networks: The state of the art" International journal of forecasting 14.1 (1998): 35-62.