MODELLING AND FORECASTING THE INTERNET AND MOBILE SUBSCRIBERS IN BANGLADESH: A TIME SERIES SEASONAL ARIMA APPROACH

ABU SAYEED MD. RIPON ROUF[1], MD. RASHIDUL HASAN[2*],

MD. SAJADUL ISLAM SARKER[3]

[1]*Department of Statistics, Jagannath University, Dhaka- 1100, Bangladesh*
[2]*Department of Public Health, Leading University, Sylhet- 3112, Bangladesh*
[3]*Department of Tourism and Hospitality Management, Leading University, Sylhet- 3112, Bangladesh*

**Corresponding Author: rashidulhasan067@gmail.com*

ABSTRACT: The internet and mobile devices are considered the backbone of today's world. A country can only be called a digital country when it becomes an e-state and all of its activities are carried out over the internet. The main objective of this study is to build the best seasonal ARIMA model to forecast the internet and mobile users in Bangladesh. This study takes into account monthly data from the Bangladesh Telecommunication Regulatory Commission (BTRC) from January 2012 to December 2021. This study also used data from January 2022 to December 2022 to validate the model. The Box-Jenkins approach was applied to build the model as well as to forecast the internet and mobile subscribers. The findings show that SARIMA (1, 1, 1) (1, 0, 0)12 and SARIMA (1, 1, 0) (1, 0, 0)12 are the best models to forecast the internet and mobile subscribers, respectively, in terms of the lowest AIC, AIC_C, and BIC value. The value was forecasted with a 95% confidence interval using the best-fitted models for both internet and mobile subscribers. The original and forecasted series were compared, and the findings reveal that the selected models provide satisfactory results. The proposed models may help policymakers forecast accurately and know the future trends of internet and mobile users in Bangladesh.

KEY WORDS: *Seasonal ARIMA Model, Internet, Mobile, Forecast, Bangladesh*

1. INTRODUCTION

The internet and mobile technology played a vital role in forming digital Bangladesh. In Bangladesh, four mobile phone operators are in operation, such as Grameen Phone Limited, Robi Axiata Limited, Banglalink Digital Communications Limited, and Teletalk Bangladesh Limited (BTRC, n. d.). These operators are currently selling not only voice calls but also many sorts of mobile entertainment or content such as SMS, MMS, welcome tune, miscall alert service, cell Bazar, job search, and so on. Bangladesh Telecommunications Company Limited (BTCL) has been a Public Switched Telephone Network (PSTN) operator in Bangladesh since its inception. The ultimate goal of PSTN services is to improve voice and data communication services so that people can benefit from them. There are numerous Internet Service Providers (ISPs) companies that provide internet connections and services to individuals and companies. ISPs may provide software packages, such as browsers, e-mail accounts, and a personal website or home page, in addition to internet connectivity. Effective forecasting of internet and mobile users indicates a country's progress toward digitization.

Time series forecasting is a crucial forecasting area in which historical data of the same variable is gathered and analyzed to develop a model that describes the underlying relationship. The model is then utilized to forecast the future time series (Zhang, 2003). There have been several methods for forecasting time series data. One of the most widely used models in time series forecasting over the past three decades is the ARIMA, which was provided by Box and Jenkins. Due to its forecasting capacity and deeper information on time-related changes, the Box-Jenkins seasonal ARIMA (SARIMA) model offers various advantages over other models, especially over exponential smoothing and neural networks (Mishra and Desai, 2005). The ARIMA model takes into consideration the serial correlation, which is the most significant feature of time series data. This model also includes a systematic search for an appropriate model at each stage (Chatfield, 1996; Zhang, 2003). The SARIMA family of models also has the benefit of having a small number of model parameters needed to explain time series that exhibit non-stationarity within and between seasons. McKerchar and Dellur (1974), Cline (1981), Govindaswamy (1991), and Yurekli et al. (2005) all discussed the applications of these models. The Box-Jenkins methodology, according to Dizon (2007), is particularly well suited for the development of models with strong seasonal behavior.

This is the first study in Bangladesh to forecast the number of internet and mobile subscribers using the seasonal ARIMA model. Consequently, there is not enough literature available in this context to the best of our knowledge. This study is intended to fill this literature gap. However, several studies have been conducted to analyze trends and forecast the behavior of time series data in the telecommunications sector. For instance, Karunarathna et al. (2018) identified the modeling trend in the Sri Lankan telecommunication sector and also found an accurate mechanism for predicting demand for internet and cellular phone connections. Zamil and Hossen (2012) discussed the present condition of the telecommunications sector of Bangladesh, identified the major challenges faced by the operators, and attempted to determine the prospects of the sector. The Box-Jenkins approach has been used in this study to develop a seasonal ARIMA model of monthly internet and mobile subscriber data. The development of an appropriate forecasting model will aid in forecasting the number of internet and mobile users. The forecasted data series also expresses the probable future trend. Thus, to track Bangladesh's digital advancement, it is essential to forecast internet and mobile users. Hence the study's main objective is to identify the best SARIMA model to forecast the internet and mobile users in Bangladesh. The main contribution of this research is to accurately forecast the internet and mobile users in Bangladesh and to know its future trends.

The general structure of this paper is as follows: The first section provides an introduction. Materials and methods are presented in Section 2 and include the data source, time series decomposition, Box-Jenkins methodology, ARIMA, and SARIMA models. The statistical software and packages utilized, the results, and the discussions are represented in sections 3 and 4, respectively. Finally, the study's conclusion is presented in Section 5.

2. MATERIALS AND METHODS

2.1. Data Source

The study took into account monthly data from the website www.btrc.gov.bd of the Bangladesh Telecommunication Regulatory Commission. The data consists of the number of internet subscribers (Mobile Internet, ISP+PSTN) and mobile subscribers (GP, Robi, Banglalink, Teletalk) covering 120 months from January 2012 to December 2021.

2.2. Decomposition of Time Series

The decomposition of time series is applied to describe the trend and seasonal pattern in time series data (Ray et al., 2021). When measurements are taken daily, weekly, or monthly, the frequency of measurements will be high, while quarterly or yearly observations will be low. According to Dagum (2010), the four variations have traditionally been assumed to be mutually independent of one another and specified using an additive decomposition model: $Y_t = T_t + S_t + C_t + R_t$, where Y_t denotes the observed data series, T_t the long-term trend, S_t the seasonal components, C_t the cyclic components, and R_t the random or irregular components.

2.3. Box-Jenkins Approach

The ARIMA model, also known as the Box–Jenkins approach, is applied to univariate time series modeling (Box and Jenkins, 1976). For estimating a good ARIMA model, at least 50 observations are required, and a seasonal time series requires a reasonably large sample size (Pankratz, 1983). The ARIMA model is particularly well suited to short-term forecasting as well as forecasting seasonally augmented data series (Ray and Bhattacharyya, 2020). The steps for the Box–Jenkins approach are identification, estimation, diagnostic checking, and forecasting. Before constructing the Box-Jenkins technique, it is essential to examine the stationarity of the time series and any major seasonality that needs to be modeled. In this work, several well-known statistical tests, including the Augmented Dickey-Fuller (1979) test and the Phillips-Perron (1988) test, were employed to examine the stationarity of time series data. In the identification step, the model's order was estimated using the ACF and PACF functions. The parameter estimation procedure was performed by the maximum likelihood method after the model's order had been determined. The best model has been chosen using model selection criteria such as Akaike's information criteria (Akaike, 1974), Akaike's information corrected criteria (Akaike, 1974), and Bayesian information criteria (Schwarz, 1978). The standardized residuals plot, ACF plot of residuals, and normal curve over the histogram of residuals were used to perform diagnostic checking. Additionally, the Ljung-Box test (Ljung and Box, 1978) was employed to test the autocorrelation among residuals. In this case, the null hypothesis is $H_0 : \rho_1 = \rho_2 = \cdots\cdots = \rho_k = 0$ that is the autocorrelation among the residuals is zero is tested with the Ljung-Box statistic $Q^* = n(n+2)\sum_{k=1}^{m}(n-k)^{-1}\rho_k^2 \sim \chi^2_{(m)}$, where n is the number of observations used to estimate the model, ρ_k is the sample autocorrelation at lag k, and m is the number of lags being tested. If Q^* is significantly large from zero, the residuals of the fitted model are said to be autocorrelated. The accuracy of the selected model was measured using the mean absolute percentage error (MAPE). The MAPE is perhaps the most commonly used indicator of forecasting accuracy (Armstrong and Collopy, 1992; Goodwin and

Lawton, 1999; Ren and Glasure, 2009). MAPE is defined by the following formula: $MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{\hat{y}_t - y_t}{y_t}\right| \times 100$, where *n* is the number of observations, y_t is the observed value at time *t*, and $\hat{y}_t$ is the predicted value by the model for time *t*. Finally, the forecasted and actual values were compared to check the validity of the chosen model (Ray et al., 2016).

2.4. Autoregressive Integrated Moving Average (ARIMA) Model

Most of the real-time series are typically non-stationary because of their increasing and declining trends; follow integrated (Ray et al., 2021). Considering a time series of data Y_t, the ARMA model is an instrument for understanding and potentially predicting future values of data. The model is divided into two sections: an autoregressive (AR) component and a moving average (MA) component (Mishra et al., 2021). The model is usually known as ARIMA (*p*, *d*, *q*) model, where *p* is the AR term, *d* is the difference, and *q* is the MA term. The model can be generated as follows:

Let, a time series Y_t is said to follow an autoregressive moving average of order *p* and *q* denoted by ARMA (*p*, *q*) if $Y_t = \mu + \phi_1 Y_{t-1} + \cdots\cdots + \phi_p Y_{t-p} - \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots\cdots - \theta_q \varepsilon_{t-q}$ (1)

Where, $\varepsilon_t \sim WN(0,\ \sigma^2)$, $\phi's$ and $\theta's$ are constants such that (1) is both stationary and invertible. Suppose (1) is put as $\phi_p(B)Y_t = \theta_q(B)\varepsilon_t$, where *B* is the backshift operator. $\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots\cdots - \phi_p B^p$ (The order *p* of AR term) $\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots\cdots - \theta_q B^q$ (The order *q* of MA term) and $B^k Y_t = Y_{t-k}$

Let the *d*th difference of Y_t be denoted by $\nabla^d Y_t$ where $\nabla = 1 - B$. A replacement of Y_t in (1) by $\nabla^d Y_t$ yields an ARIMA model of orders *p*, *d*, and *q* denoted by ARIMA (*p*, *d*, *q*) in Y_t . If *d* = 0 then ARIMA (*p*, *d*, *q*) = ARIMA (*p*, *q*).

2.5. Seasonal Autoregressive Integrated Moving Average (SARIMA) Model

The SARIMA model is commonly referred to as the seasonal ARIMA model where the time series follows the seasonal consequence. The model is built with the seasonal behavior of the series in consciousness. The overall multiplicative SARIMA model can be written as SARIMA (*p*, *d*, *q*) (*P*, *D*, *Q*)$_S$ (Afrifa-Yamoah et al., 2016).

Mathematically, $\phi_p(B)\Phi_P(B^S)(1-B)^d(1-B^S)^D(Y_t - \mu) = \theta_q(B)\Theta_Q(B^S)\varepsilon_t$

$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots\cdots - \phi_p B^p$ (The order *p* of AR term)

$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots\cdots - \theta_q B^q$ (The order *q* of MA term)

$\Phi_P(B^S) = 1 - \Phi_1 B^S - \Phi_2 B^{2S} - \cdots\cdots - \Phi_P B^{PS}$ (The order *P* of seasonal AR term)

$\Theta_Q(B^S) = 1 - \Theta_1 B^S - \Theta_2 B^{2S} - \cdots\cdots - \Theta_Q B^{QS}$ (The order Q of seasonal MA term)

and $\varepsilon_t \sim WN(0,\ \sigma^2)$, the difference *d* is a positive integer, *s* is an integer greater than one. It is to be noted that $\mu = 0$, if *d* > 0 or *D* > 0

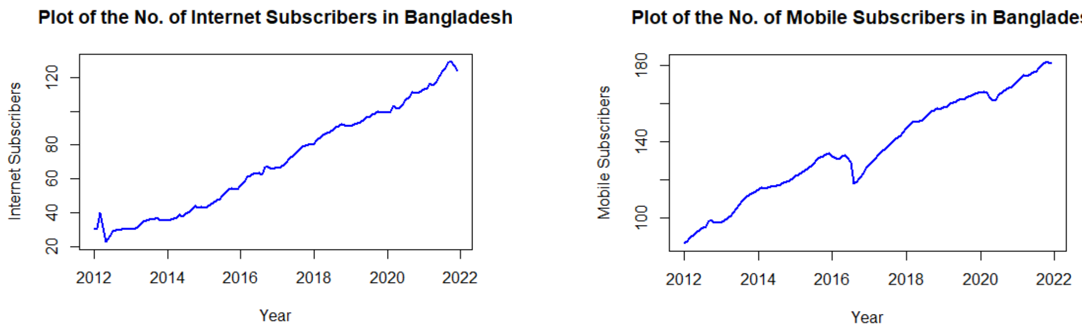
3. STATISTICAL SOFTWARE AND PACKAGES USED

The analysis of this study was carried out entirely using the open-source statistical programming software R Studio (version 3. 5. 3) for Windows. TSA, forecast, MASS, and tseries are some of the library packages used for analysis.

4. RESULTS AND DISCUSSIONS

4.1. Plot of the Original Time Series

In this section, the plot of the original internet and mobile subscriber's data has been depicted.



(a) **(b)**

Fig. 1. Time series plot of the number of internet and mobile subscribers.

Fig. 1(a) and 1(b) show the graphical representation of the number of internet and mobile subscribers (in millions) from January 2012 to December 2021 in Bangladesh. It is obvious from Fig. 1 that the number of internet and mobile subscribers fluctuated over the study period. That is there is an indication of a trend and seasonality in both time series.

4.2. Decomposition of the Original Time Series

In this section, time series of internet and mobile subscribers' data have been decomposed to find the trend and seasonality. Time series decomposition revealed that the series has been broken down into its trend, seasonal, and random components.

Fig. 2(a) and 2(b) depict the observed time series (top), the trend component (second from top), the seasonal component (third from top), and the random component (bottom) of the internet and mobile subscriber series. The trend component for both the internet and mobile subscribers' data shows an upward trend over the period, although the irregular character of the mobile subscribers' series for the years 2016 to 2017 may be of interest. Again, the seasonal component of the data series for both internet and mobile subscribers shows an up-and-down pattern over time. The upward and downward trend of the observed time series provides evidence of seasonality in the data series. Thus from Fig. 2(a) and 2(b) both the internet and mobile subscriber, data have a seasonal effect, with the usual upward and downward pattern occurring yearly throughout the study time. This means that seasonal factors influenced the internet and mobile subscriber's data series each year. Since there is both trend and seasonality in the internet and mobile subscribers data series, apply a seasonal difference (month = 12) of order D = 1 for both data series and re-evaluate the trend (PennState, n. d.).

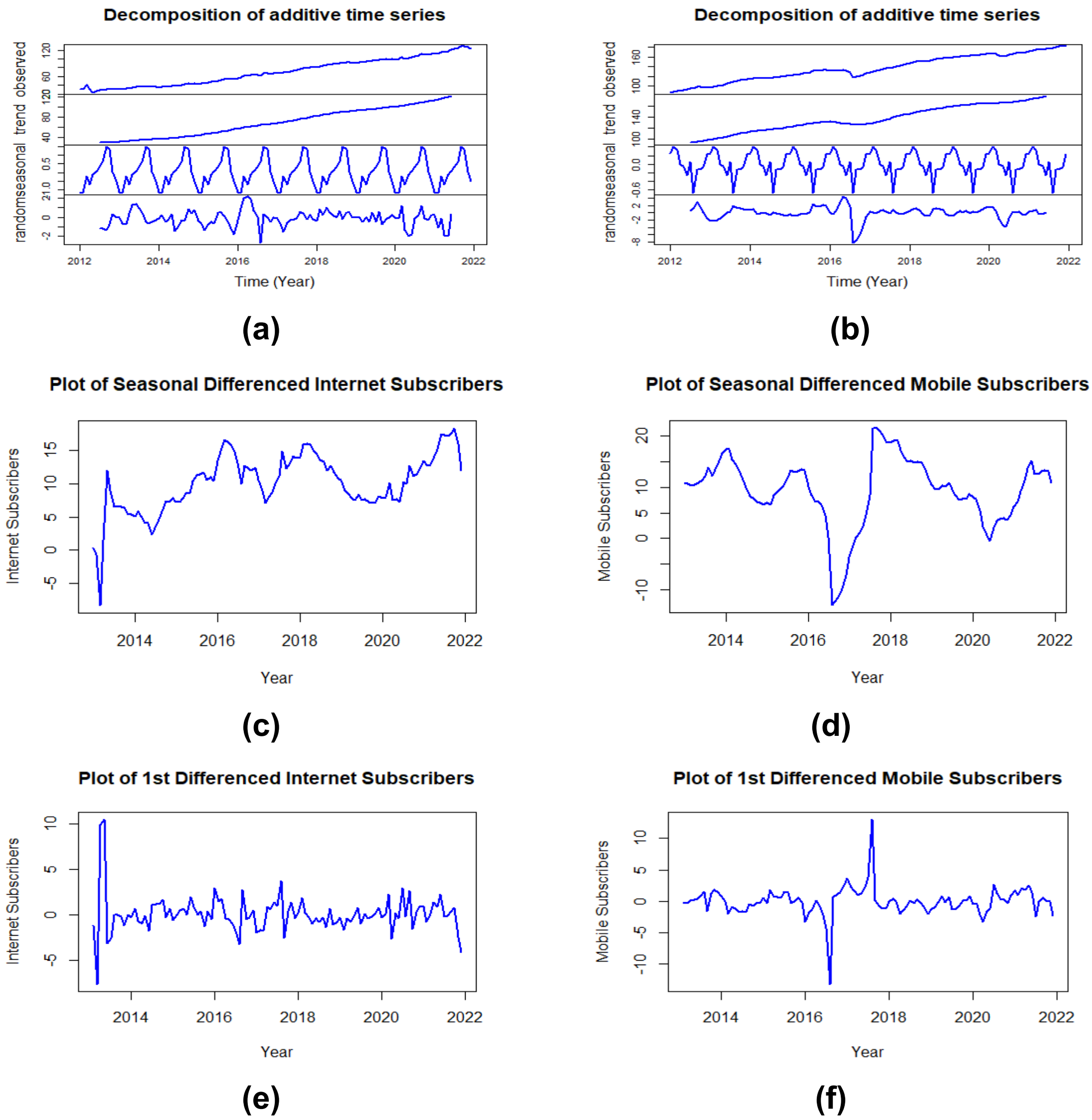


Fig. 2. Decomposition of additive internet and mobile subscriber series.

Fig. 2(c) and 2(d) show that the seasonal differenced internet and mobile subscribers' data series have an upward trend. To remove the trend component that is to make the series stationary we take the non-seasonal difference of order $d = 1$ for both internet and mobile subscribers' data series. Fig. 2(e) and 2(f) show that the first differenced internet and mobile subscriber time series have no significant trend pattern. This is an indication of stationary behavior. Moreover, we apply the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests to ensure stationarity.

4.3. Stationary Test

Data must be checked for stationarity before being used in time series modelling. Under this test, the statistical hypothesis is:

H_0: The series has a unit root.

H_1: The series does not have a unit root; that is, the data series is stationary.

Table 1: ADF and PP tests for first-differenced internet and mobile subscriber's time series

Time series	Tests	Calculated value	Truncation lag parameter	p-value	Comment
Internet	ADF	-5.1725	4	0.01	Stationary
Subscriber	PP	-84.673	4	0.01	Stationary
Mobile	ADF	-3.7927	4	0.02	Stationary
Subscriber	PP	-63.537	4	0.01	Stationary

The p-value of the first differenced internet and mobile subscriber series is lower than the significance level of 0.05 indicating that both the data series are stationary.

4.4. Identification of Model

To detect non-seasonal and seasonal AR and MA terms, inspect the ACF and PACF of first-differenced internet and mobile subscriber data.

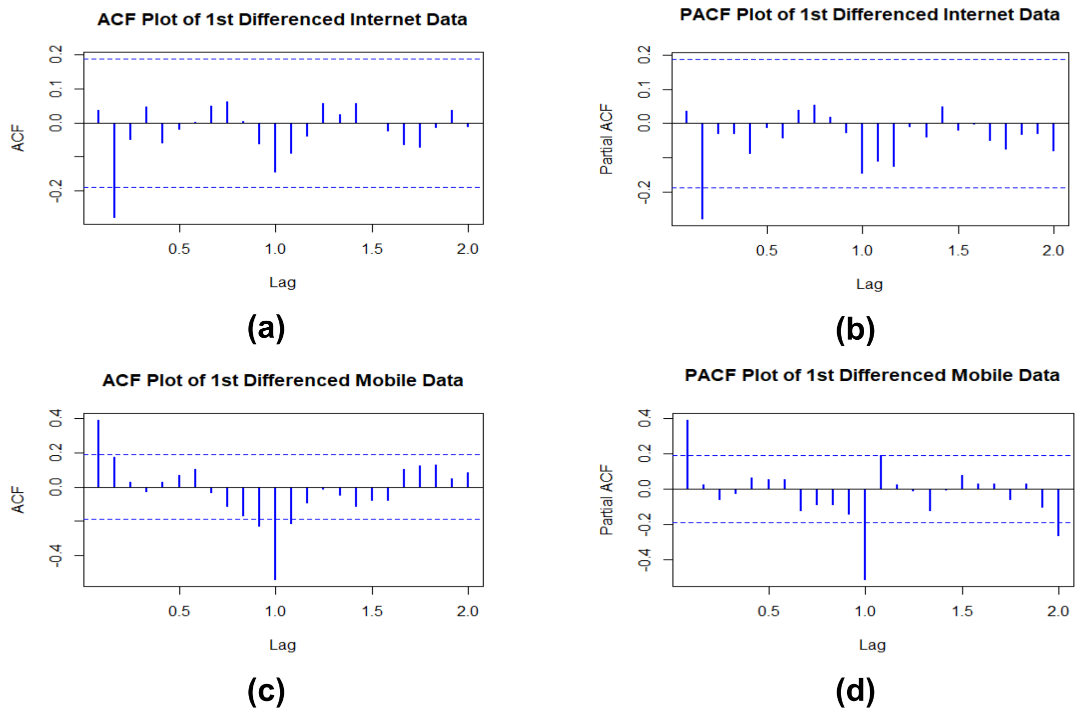


Fig. 3. ACF and PACF plot of first-differenced internet and mobile subscribers.

Non-seasonal AR and MA: To identify the non-seasonal terms, look into the early lags (1, 2, 3, and so on). For internet subscriber data, the PACF shows (Fig. 3(b)) a significant negative spike at lag 2 and then nothing else. So, a non-seasonal AR of order p = 1 could be an important component of the model for the internet series. Also, for internet subscriber data, the ACF shows (Fig. 3(a)) a significant negative spike at lag 2 and then nothing else. Hence, a non-seasonal MA of order q = 1 could be an important component of the model for the internet series (PennState, n. d.). Again, the PACF for mobile subscriber data shows (Fig. 3(d)) a significant positive spike at lag 1 and then nothing until around lag 12. As a result, a non-seasonal AR of order p = 1 could be an important element of the model for the mobile data series. In addition, the ACF for mobile subscriber data shows (Fig. 3(c)) a significant positive spike at lag 1 and then nothing until around lag 11. Hence, a non-seasonal MA of order q = 1 could be an important element of the model for mobile data series (PennState, n. d.).

Seasonal AR and MA: Examine the patterns across multiples of S = 12 lags. For seasonal parameters, we are looking at what's going on around lags 12, 24, and so on. There is no significant spike in PACF for internet subscriber data at lags 12 and 24 (Fig. 3(b)). So, a seasonal AR of order P = 0 could be useful in the model for the internet series. There is no significant spike in ACF for internet subscriber data at lags 12 and 24 (Fig. 3(a)). Hence, a seasonal MA of order Q = 0 could be useful in the model for the internet series (PennState, n. d.). At lags 12 and 24, there is a significant negative spike in PACF for mobile subscriber data (Fig. 3(d)). As a result, a seasonal AR of order P = 2 could be useful in the model for the mobile data series. There is a significant negative spike at lag 12 in ACF for mobile subscriber data (Fig. 3(c)), and then nothing else. Hence, a seasonal moving average of order Q = 1 could be useful in the model for mobile data series (PennState, n. d.). Based on several seasonal and non-seasonal AR and MA terms the appropriate model for internet subscriber data is ARIMA (1, 1, 1) (0, 1, 0)12 and mobile subscriber data is ARIMA (1, 1, 1) (2, 1, 1)12. Now all possible tentative seasonal ARIMA models with the appropriate combination of seasonal and non-seasonal AR and MA terms and their summary statistics are listed in the following table.

Table 2: Seasonal ARIMA models for internet and mobile subscriber data and their goodness of fit test

Time series	Models	Goodness of fit		
		AIC	AICC	BIC
Internet subscriber	ARIMA (1, 1, 1) (0, 1, 0)12 with drift	-324.18	-323.95	-316.16
	ARIMA (1, 1, 1) (1, 0, 0)12 with drift	-381.00	-380.64	-369.88
	ARIMA (1, 1, 0) (1, 0, 0)12 with drift	-363.21	-363.00	-354.88
	ARIMA (0, 1, 1) (0, 0, 1)12 with drift	-366.96	-366.75	-358.62
	ARIMA (0, 1, 0) (1, 0, 0)12 with drift	-365.03	-364.93	-359.47
	ARIMA (0, 1, 0) (2, 0, 0)12 with drift	-363.19	-362.99	-354.86
	ARIMA (0, 1, 0) (1, 0, 1)12 with drift	-364.55	-364.34	-356.21
	ARIMA (0, 1, 0) (0, 0, 1)12 with drift	-364.85	-364.75	-359.3
	ARIMA (0, 1, 0) (2, 0, 1)12 with drift	-362.56	-362.21	-351.44
	ARIMA (0, 1, 1) (1, 0, 0)12 with drift	-366.27	-366.06	-357.93
Mobile subscriber	ARIMA (1, 1, 1) (2, 1, 1)12 with drift	-618.27	-617.43	-602.24
	ARIMA (1, 0, 1) (2, 0, 1)12 with drift	-715.34	-714.81	-701.44
	ARIMA (1, 1, 1) (2, 1, 0)12 with drift	-603.5	-602.91	-590.14
	ARIMA (0, 1, 1) (2, 1, 0)12 with drift	-603.17	-602.78	-592.48
	ARIMA (1, 1, 0) (1, 1, 1)12 with drift	-622.27	-621.88	-611.58
	ARIMA (2, 1, 2) (1, 0, 1)12 with drift	-712.68	-711.67	-693.22
	ARIMA (1, 1, 0) (1, 0, 0)12 with drift	-718.65	-718.44	-710.32
	ARIMA (0, 1, 1) (0, 0, 1)12 with drift	-708.55	-708.34	-700.21
	ARIMA (1, 1, 0) (0, 0, 1)12 with drift	-718.65	-718.44	-710.31
	ARIMA (1, 1, 0) (1, 0, 1)12 with drift	-717.27	-716.92	-706.15
	ARIMA (0, 1, 0) (1, 0, 0)12 with drift	-687.75	-687.65	-682.19
	ARIMA (0, 1, 0) (0, 0, 1)12 with drift	-687.45	-687.35	-681.89

AIC: Akaike Information Criterion; AIC_C: Akaike Information Correction Criterion; BIC: Bayesian Information Criterion

The best seasonal ARIMA model is selected based on the lowest AIC, AIC_C, and BIC values. According to these model selection criteria, seasonal ARIMA (1, 1, 1) $(1, 0, 0)_{12}$ is the best model to forecast the internet subscribers in Bangladesh with the lowest AIC= -381.00, AIC_C = -380.64, and BIC= -369.88. Also, seasonal ARIMA (1, 1, 0) $(1, 0, 0)_{12}$ is the best model to forecast the mobile subscribers in Bangladesh with the lowest AIC= -718.65, AIC_C = -718.44, and BIC= -710.32.

4.5. Parameter Estimation and Diagnostic Test of Models

The estimated parameters, standard error, and Ljung–Box Q statistic for best-fitted models are presented in Table 3.

Table 3: Parameter of models and Ljung-Box test

Models	Parameter estimation			Ljung-Box Q test		
	Parameters	EV	SE	Lag used	Statistics	p-value
ARIMA (1, 1, 1) $(1 ,0, 0)_{12}$	*p*	-0.479	0.117	24	28.500	0.127
	q	0.867	0.063			
	P	0.253	0.158			
ARIMA (1, 1, 0) $(1, 0, 0)_{12}$	*p*	0.505	0.081	24	12.978	0.934
	P	-0.025	0.095			

EV: Estimated Value; SE: Standard Error of Parameter

From Table 3, the p-value ensures that the null hypothesis of all autocorrelation coefficients equal to zero at various lags is accepted for the best-fitted models.

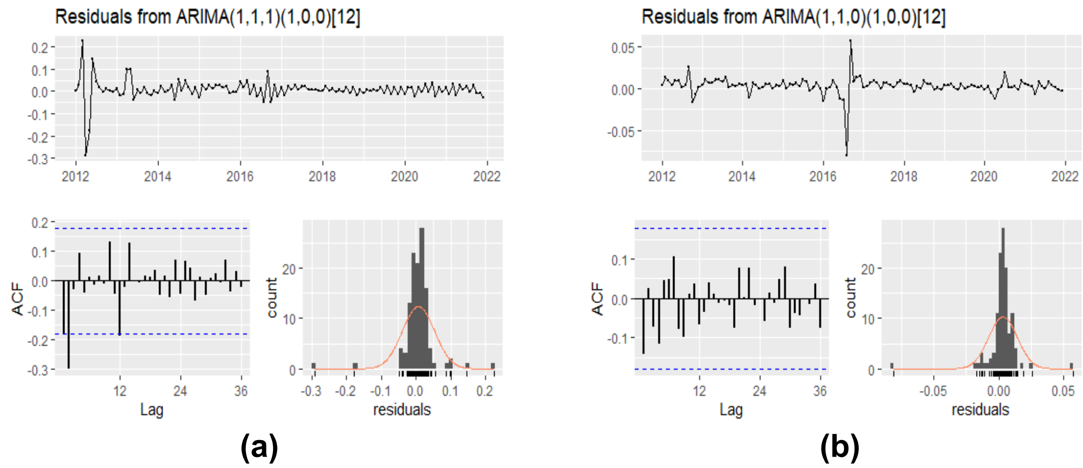


Fig. 4. ACF and residual plot of best-fitted models.

The residual plot of the best-fitted seasonal ARIMA models shown in Fig. 4(a) and 4(b) indicates that the standard errors are roughly constant in their mean and variance over time, though there appears to be some higher variance near the beginning of the internet subscriber series and towards the middle of mobile subscriber series. Furthermore, the ACF of the residual of the best-fitted models shows the different autocorrelation coefficients at different lags within the 95% confidence interval, except for lag 2 of the internet subscriber series, confirming the data “good fit” to the selected models. Finally, the residuals of the best-fitted models are normally distributed, according to the histogram of residuals. However, to investigate the adequacy of the fitted models, forecasting criteria values were presented in the following table.

Table 4: Forecasting criteria value of the fitted models

Models	Forecasting Criteria			
	RMSE	MAE	MAPE	MASE
ARIMA (1, 1, 1) $(1, 0, 0)_{12}$	2.063	1.415	2.501	0.138
ARIMA (1, 1, 0) $(1, 0, 0)_{12}$	1.410	0.839	0.656	0.081

MAPE: Mean Absolute Percentage Error

The MAPE is 2.501 for model ARIMA (1, 1, 1) $(1, 0, 0)_{12}$ and it is 0.656 for model ARIMA (1, 1, 0) $(1, 0, 0)_{12}$ which is less than 10, indicates that the fitted models are highly accurate (Lewis, 1982). Therefore, these models can be applied to accurately forecast the internet and mobile subscribers in Bangladesh.

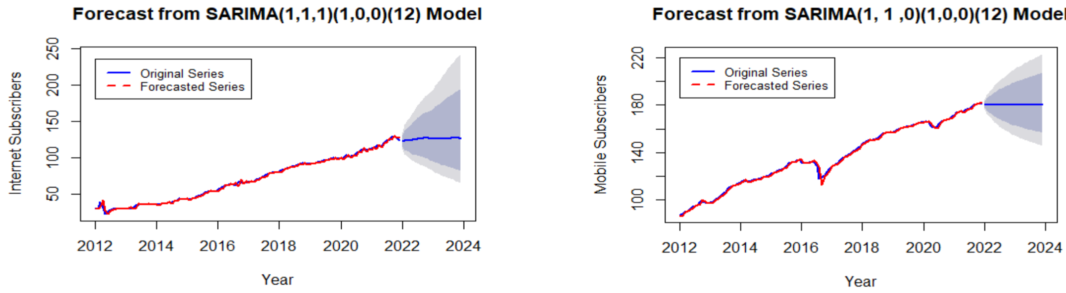
4.6.Forecasting

Forecasting is done for both internet and mobile subscribers after the best-fitted time series models have been developed. Here, the data of the first ten years covering 120 months (January 2012 to December 2021) were used to build models, while the remaining one year covering 12 months (January 2022 to December 2022) was used to ensure the validity of the model.

Table 5: Forecast value of the number of internet and mobile subscribers in Bangladesh

Year	Month	SARIMA (1, 1, 1) $(1, 0, 0)_{12}$			SARIMA (1, 1, 0) $(1, 0, 0)_{12}$		
		Forecast value	Actual value	95% CI	Forecast value	Actual value	95% CI
2022	Jan.	122.79	121.87	(111.87,134.79)	180.74	180.78	(176.68,184.90)
	Feb.	123.40	122.79	(105.21, 144.72)	180.59	181.54	(173.31,188.17)
	Mar.	124.04	124.89	(102.08, 150.73)	180.49	182.92	(170.42,191.16)
	Apr.	123.99	124.20	(98.61, 155.90)	180.48	183.38	(167.99,193.89)
	May	124.43	125.52	(96.24, 160.88)	180.43	184.23	(165.83,196.31)
	June	125.43	126.21	(94.53, 166.43)	180.39	184.45	(163.93,198.51)
	July	126.14	127.55	(92.87, 171.33)	180.38	184.05	(162.23,200.55)
	Aug.	126.59	127.26	(91.19, 175.74)	180.33	183.58	(160.67,202.41)
	Sept.	127.43	126.31	(89.93, 180.56)	180.29	181.43	(159.22,204.14)
	Oct.	127.53	126.18	(88.28, 184.23)	180.26	181.67	(157.89,205.80)
	Nov.	126.88	125.02	(86.23, 186.68)	180.26	180.87	(156.68,207.39)
	Dec.	126.17	124.42	(84.26, 188.92)	180.27	180.20	(155.54,208.93)
2023	Jan.	125.90	---------	(82.09, 193.09)	180.28	---------	(154.49,210.37)
	Feb.	126.06	---------	(80.11, 198.37)	180.28	---------	(153.50,211.74)
	Mar.	126.22	---------	(78.39, 203.24)	180.28	---------	(152.56,213.05)
	Apr.	126.21	---------	(76.64, 207.84)	180.28	---------	(151.66,214.32)
	May	126.32	---------	(75.10, 212.49)	180.28	---------	(150.79,215.55)
	June	126.58	---------	(73.72, 217.34)	180.29	---------	(149.96,216.75)
	July	126.76	---------	(72.39, 221.99)	180.29	---------	(149.15,217.92)
	Aug.	126.88	---------	(71.08, 226.47)	180.29	---------	(148.38,219.06)
	Sept.	127.09	---------	(69.89, 231.07)	180.29	---------	(147.63,220.18)
	Oct.	127.11	---------	(68.67, 235.30)	180.29	---------	(146.90,221.27)
	Nov.	126.95	---------	(67.40, 239.12)	180.29	---------	(146.19,222.35)
	Dec.	126.77	---------	(66.17, 242.86)	180.29	---------	(145.49,223.40)

CI: Confidence Interval



(a) **(b)**

Fig. 5. Forecast from the best-fitted SARIMA model.

From Table 5, the best models predict an increase in the number of internet and mobile subscribers in Bangladesh during the period January 2022 to December 2023. Moreover, the forecast and actual values of the internet and mobile subscriber series from January 2022 to December 2022 lie between the 95% confidence interval. Fig. 5 shows that the forecasted series of internet and mobile subscribers (red color) differed by a small amount from the original series of internet and mobile subscribers (dark blue color), indicating that the fitted model for the number of internet and mobile subscribers is functioning properly. Therefore, the estimated internet and mobile subscribers are a much more accurate representation of the actual internet and mobile subscribers in Bangladesh.

5. CONCLUSION

This study is intended to identify the best SARIMA model to forecast the internet and mobile users in Bangladesh. The model was constructed and the forecast was made using the Box-Jenkins technique in this study. The findings of this study conclude that SARIMA (1, 1, 1) $(1, 0, 0)_{12}$ and SARIMA (1, 1, 0) $(1, 0, 0)_{12}$ are the best modes to forecast the internet and mobile subscribers, respectively, in Bangladesh. The graphical evaluation of the observed and forecasted series shows a slight difference, indicating that the fitted model performs well in forecasting. Both models also predict that there will be an increase in internet and mobile subscribers during the period of prediction. This study may help predict the internet and mobile subscribers by using the estimated models. Hence, policymakers may find this research beneficial.

ACKNOWLEDGMENT

The authors sincerely appreciate the reviewers' insightful comments and suggestions for the manuscript.

REFERENCES

Afrifa-Yamoah, E., Saeed, B., & Karim, A. (2016). Sarima modelling and forecasting of monthly rainfall in the Brong Ahafo region of Ghana. World Environment, 6(1), 1–9. http://dx.doi.org/10.5923/j.env.20160601.01

Akaike, H. (1974). A new look at the statistical model identification. IEEE Transaction on Automatic Control, 19(6), 716–723. https://link.springer.com/chapter/10.1007/978-1-4612-1694-0_16

Armstrong, J. S., & Collopy, F. (1992). Error measures for generalizing about forecasting methods: Empirical comparisons. International Journal of Forecasting, 8(1), 69-80. https://doi.org/10.1016/0169-2070(92)90008-W

Bangladesh Telecommunication Regulatory Commission. (n. d.). Mobile Subscribers. https://btrc.gov.bd

Bangladesh Telecommunication Regulatory Commission. (n. d.). PSTN Services. https://btrc.gov.bd

Box, G. E. P., & Jenkins, G. M. (1976). Time series analysis: Forecasting and control. Holden Day, San Francisco. http://dx.doi.org/10.1111/jtsa.12194

Chatfield, C. (1996). The analysis of time series: An introduction with R, 283. Chapman and Hall, New York. https://doi.org/10.4324/9780203491683

Cline, T. B. (1981). Selecting seasonal stream flow models. Water Resources Research, 17(4), 975–984. https://doi.org/10.1029/WR017i004p00975

Dagum, E. B. (2010). Time series modeling and decomposition. Statistica, 70(4), 433–457. http://dx.doi.org/10.6092/issn.1973-2201/3597

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association, 74(366), 427–431. http://dx.doi.org/10.2307/2286348

Dizon, C. Q. (2007). ARMA modeling of a stochastic process appropriate for the Angat reservoir. Philippine Engineering Journal, 28(1), 1-20. https://journals.upd.edu.ph/index.php/pej/article/download/2211/2108

Goodwin, P., & Lawton, R. (1999). On the asymmetry of the symmetric MAPE. International Journal of Forecasting, 15(4), 405-408. https://doi.org/10.1016/S0169-2070(99)00007-2

Govindaswamy, R. (1991). Univariate Box-Jenkins forecasts of water discharge in Missouri river. Water Resource Development, 7(3), 168–177. http://dx.doi.org/10.1080/07900629108722509

Karunarathna, K. A. N. K., Brindha, S., & Paramadevan, P. (2018). Modeling trend in telecommunication in Sri Lanka: A case study on internet and cellular connections. International Journal of Statistics and Probability, 7(4), 32- 42. http://dx.doi.org/10.5539/ijsp.v7n4p32

Lewis, C. D. (1982). Industrial and Business forecasting methods. Journal of Forecasting, 2(2), 194-196. https://doi.org/10.1002/for.3980020210

Ljung, G. M., & Box. G. E. P. (1978). On a measure of lack of fit in time series models. Biometrika, 65(2), 297–303. https://doi.org/10.1093/biomet/65.2.297

McKerchar, A. I., & Dellur, J. W. (1974). Application of seasonal parametric linear stochastic models to monthly flow data. Water Resources Research, 10(2), 246–255. http://dx.doi.org/10.1029/WR010i002p00246

Mishra, A. K., & Desai, V. R. (2005). Drought forecasting using stochastic models. Stochastic Environmental Research and Risk Assessment, 326-339. https://www.researchgate.net/publication/226230747

Mishra, P., Matuka, A., Abotaleb, M. S. A., Weerasinghe, W. P. M. C. N, Karakaya, K., & Das, S. S. (2021). Modeling and forecasting of milk production in the SAARC countries and China. Modeling Earth Systems and Environment, 947-959. https://doi.org/10.1007/s40808-021-01138-z

Pankratz, A. (1983). Forecasting with univariate Box-Jenkins models: concepts and cases. Wiley Series in Probability and Mathematical Statistics, USA. http://dx.doi.org/10.1002/9780470316566.refs

PennState. (n. d.). Applied Time Series Analysis. https://online.stat.psu.edu/stat510/lesson/4/4.1

Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. Biometrika, 75(2), 335–346. https://doi.org/10.1093/biomet/75.2.335

Ray, S., & Bhattacharyya, B. (2020). Statistical modeling and forecasting of ARIMA and ARIMAX models for food grains production and net availability of India. Journal of Experimental Biology and Agricultural Sciences, 8(3), 296–309. http://dx.doi.org/10.18006/2020.8(3).296.309

Ray, S., Bhattacharyya, B., & Pal, S. (2016). Statistical modeling and forecasting of food grain in effects on public distribution system: An application of ARIMA model. Indian Journal of Economics and Development, 12(4), 739–744. http://dx.doi.org/10.5958/2322-0430.2016.00199.2

Ray, S., Das, S. S., Mishra, P., & Al Khatib, A. M. G. (2021). Time series SARIMA modelling and forecasting of monthly rainfall and temperature in the South Asian countries. Earth Systems and Environment, 531-546. https://doi.org/10.1007/s41748-021-00205-w

Ren, L., & Glasure, Y. (2009). Applicability of the revised mean absolute percentage errors (MAPE) approach to some popular normal and nonnormal independent time series. International Advances in Economic Research, 15(4), 409-420. https://doi.org/10.1007/s11294-009-9233-8

Schwarz, G. (1978). Estimating the dimension of a model. The Annals of Statistics, 6(2), 461–464. https://doi.org/10.1214/aos/1176344136

Yurekli, K., Kurunc, A., & Ozturk, F. (2005). Application of linear stochastic models to monthly flow data of Kelkit stream. Ecological Modelling, 183(1), 67–75. https://www.researchgate.net/publication/223244028

Zamil, M. F., & Hossen, M. M. (2012). Problems and prospects of telecommunication sector of Bangladesh. A critical review. New Media and Mass Communication, 16-25. https://api.semanticscholar.org/CorpusID:166558832

Zhang, G. P. (2003).Time series forecasting using a hybrid ARIMA and neural network model.Neurocomputing,159–175. http://dx.doi.org/10.1016/S0925-2312(01)00702-0