Modelling and Forecasting the Usage of Cellular and Landline Phones in Sri Lanka: Univariate Time Series Approach

A.W.S.P Karunarathner[1], M.S.H Perera[2] and U.P Liyanage[3]

Abstract

Phones have become a mandatory commodity in human life. Nowadays, there is a very strong increase in the cellular phone market, so we tend to forget landline phone services. According to statistics, cellular phones and landline phones usage up to December 2018 is 32,528,104 and 2,484, 616 respectively. That is, the teledensity (per 100 inhabitants) is 150 for cellular phones and 11.5 for landline phones. Due to the increment of the cellular phones and decrement of the landline phones, it is vitally important to study their behaviour. Therefore, the objective of this paper is to model and forecast the usage of cellular and landline phones in Sri Lanka. The model was developed using 80% of the data and validated with 20%. The usage was modelled with Autoregressive Integrated Moving Average (ARIMA) technique. Several models were fitted and based on the lowest Akaike's Information Criteria (AIC), ARIMA (1,2,1) and ARIMA (2,2,1) were identified as the best-fitted models with forecasting accuracy measured by Mean Absolute Percentage Error (MAPE) values 1.403 and 0.976 for cellular and landline phones usage respectively, concluding that two ARIMA models have a strong potential for forecasting the usage of cellular and landline phones. This model would be important to those who are with the telecoms market to achieve their business goals.

Keywords: *AIC, ARIMA, MAPE, ACF, PACF*

---

[1] **Corresponding Author**
Department of Statistics & Computer Science, University of Kelaniya, Sri Lanka
E-mail: sachinikarunarathne94@gmail.com

[2] Department of Statistics & Computer Science, University of Kelaniya, Sri Lanka

[3] Department of Statistics & Computer Science, University of Kelaniya, Sri Lanka

1. Introduction

Among all the essential modern telecommunication equipment, telephones are at the forefront. Phones have become an absolutely necessary commodity in human life. Today the world has become a village which has no boundaries to communicate between countries and people. There are two major categories of phones namely cellular phones and landline phones. Landline phones can be classified into another two subcategories as fixed wire landline phones and fixed wireless landline phones. In this study, the total usage of those two categories has considered as the landline phone usage.

Considering the evolution of the telephones, the modern telephone is the result of work of many people. Alexander Graham Bell was awarded the first U.S. patent for the invention of the telephone in 1876. Many people have contributed to the development of the telephone. So gradually the telephone has become incredibly advanced. The innovation of cellular phones in 1973 was a turning point in the history of telephone invention. Currently, the cellular phones have become tinier and lighter including many more features such as touch screen, camera, fingerprint sensor, internet connectivity, long-lasting battery, GPS, Bluetooth etc. So, with these features, cellular phones have converted into a wonderful innovation which is called a smartphone. Therefore, the usage of mobile phones has rapidly increased as a result of the invention of these smartphones. This makes our lives much easier and more convenient because they provide us with a lot of services such as online shopping, e-banking, e-mailing, taking photographs, accessing the internet, entertaining with games, music, movies etc. That is, comparing to landline phones, mobile phones are multitasking and super easy. Therefore, there can be seen a significant decline in landline phones market. In our opinion, the main reason for this decline is the immobility of these landlines and also lacks of technology, more expensive, no visual communication tools etc. But compared over cell phones, the biggest advantage of landline phones is the consistency of the telephone signals.

The objective of this study is to model and forecast the usage of cellular and landline phones in Sri Lanka using Autoregressive Integrated Moving Average (ARIMA) models and to use the best-fitted models to predict the usage of the cellular and landline phones in future years more accurately. Our study will be helpful to the people those who are with the telecoms market to achieve their business goals by considering the usage of consumers. Also, it will be much easier in decision making by analysing the variations of usage and to get an idea about the gross income of a county from telephones market.

This manuscript is collated as follows: In section II the literature reviews are presented. Section III is used to present the methodology of this study. In section IV and V results, conclusion and future work of the study are pointed out.

2. Literature Review

Yuan et al. (2004) have presented a paper based on the results they have gained through a survey on cell phones. They have conducted their survey for US households because in the last five years the usage of the cellular phone has strongly increased. The survey results showed that, in general, more efforts were required to get a complete interview from people with both telephone services than from people with only a cell phone. McBurney, Parsons, & Green (2002) presented an introductory paper on forecasting market demand for new telecommunications services. There they have given a brief introduction about marketing theory and then talks about how to identify the key stakeholders in the forecasting process in a new telecommunication company. Throughout the study, they have portrayed the main forecasting techniques and highlighted some of the conceptual and practical challenges involved in forecasting the demand for new telecommunication services.

Aker & Mbiti (2010) presented a study on mobile phones and economic development in Africa. They have shown that the mobile phone subscriptions have increased by 49 precent annually between the year 2002 and the year 2007 as compared with the past decade. They have shown that the growth of the mobile phone coverage across Africa has shown strong positive growth with population density as well as other factors. Thanh et al. (2005) have described that mobile phones will continue to increase in numbers and many more people will have mobile phones in future. They are predicting that in future the mobile phones will have more and more functions such as personal data storage, Personal Information Manager (PIM), MP3 player, camera portable storage etc. They have written and published this paper in 2005. It has completely proven that they were successful in doing predictions about mobile phones since all the smartphones in the market today have almost all the features they have mentioned in their paper.

Rashid & Elder (2009) have presented a paper to emphasize the connection between mobile phones and the development of a country. They have stated that in most of the developing countries have skipped fixed-line telephones and leapfrogged directly into the mobile phones. One of their key findings in this paper is that mobile phones are increasingly accessible to lower-income groups in developing countries. Nawaz (2012) has carried out statistical analysis to find the impact of mobile phones on students' life in India. This paper reviews the impact of the mobile phone on youth peer relationships, on family relationships and the institution in both positive and negative ways. For this purpose, a survey has been conducted by taking five well-known colleges of Gujarat city.

Zajdel, Śmigielski, & Nowak (2013) have published a paper to evaluate the influence of the sound of a ringing mobile phone on the complex reaction time score in a healthy person and to check if there are any differences in reaction time when a landline phone and mobile phone ring. This study concludes that the relationship between a person and their private phone can significantly obstruct their attention and thus affect the attention-demanding activities. Another survey has been conducted by Priya & Jeevitha (2017) in India to analyse mobile phone usage and the academic performance of college students. 200 college students were selected for this survey who are using mobile phones. The collected data has been analysed using a chi-square test. The study concluded that there is a significant relationship between mobile phone users and their academic performance.

Fowdur, Hurbungs, & Beeharry (2016) have objected to do a statistical analysis of the energy consumption of mobile phones for web-based applications in Mauritius. They have concerned the two most common types of mobile phones namely, Android OS-based phones and Nokia phones with

Symbian OS. They have concluded that the average energy consumption over a month for running the web-based applications on mobile phones was estimated to be 16.76 MWh for Mauritius.

3. Research Methodology

The data required for this study were collected from the Central Bank of Sri Lanka. The data set contains quarterly data from the year 2000 quarter 1 to the year 2018 quarter 2.

Stationarity

A stationary time series can be identified as a time series whose statistical properties, mainly mean and variance are constant over time. Graphical methods and statistical tests are commonly used to identify whether a time series is stationary or not. The most accurate way is to use statistical tests and in this study, Kwiatkowski-Phillips-Schmidt-Shin (KPSS), Augmented Dickey-Fuller (ADF) and Phillips Perron (PP) tests were used to check the stationarity of the time series.

i. Kwiatkowski-Phillips-Schmidt-Shin (KPSS) Test

   The corresponding hypothesis is,

   $H_0: the\ series\ is\ stationary$

   $H_1: the\ series\ is\ not\ stationary$

ii. Augmented Dickey-Fuller (ADF) Test

   The corresponding hypothesis is,

   H_0: the series is not stationary

   H_1: the series is stationary

iii. Phillips Perron (PP) Test

   The corresponding hypothesis is,

   H_0: the series is not stationary

   H_1: the series is stationary

Time Series Forecasting Methods

A time series is a collection of observations where they are indexed in time order. There can be numerous time series where the sequence of observations is varied depending on the frequency like weekly, monthly, quarterly or annually etc. The ultimate goal of time series analysis is to derive the future behaviour of the time series considering the past behaviour. Time series forecasting is tremendously used in many sectors such as finance, agriculture, health, education etc. In this study, a univariate time series approach has been used in forecasting the corresponding data set.

Univariate Time Series Approach: ARIMA Model

If the previous values of a time series are used to predict its future values, it is called a 'Univariate Time Series Forecasting'. ARIMA modelling is a particular type in univariate time series forecasting.

ARIMA short for 'Auto-Regressive Integrated Moving Average'. This can be used to forecast future values using its own lags and the lagged forecasting errors. ARIMA models are applied when the data shows evidence of non-stationarity such as trend, random errors etc. In ARIMA models, differencing steps (corresponding to the "integrated" part of the model) are applied to eliminate the non-stationarity. The differencing steps can be further applied for one or more times.

$$ARIMA(p, d, q) \tag{1}$$

where, p: autoregressive order

d: degree of differencing

q: moving average order

$$\hat{y}_t = c + \emptyset_1 \hat{y}_{t-1} + \emptyset_2 \hat{y}_{t-2} + \cdots + \emptyset_p \hat{y}_{t-p} + z_t + \theta_1 \hat{z}_{t-1} + \theta_2 \hat{z}_{t-2} + \cdots + \theta_q \hat{z}_{t-q} \tag{2}$$

where, $\hat{y}_t$ is the differed series.

c is a constant

$\emptyset_1, \emptyset_2, \ldots, \emptyset_p$ are autoregressive parameters

$\theta_{1,} \theta_2, \ldots, \theta_p$ are moving average parameters

z_t is the White Noise term

The autoregressive order can be identified using the PACF (Partial Autocorrelation Function) plot and the moving average order can be identified using ACF (Autocorrelation Function) plot. ARIMA modelling has four major steps as model building, identification, estimation, diagnostics and forecasting. In this study, the model is developed using 80% of the data and validated with 20%.

Trend Elimination Methods

There are mainly four trend elimination methods namely Transformations, smoothing with Moving Average Filter, Exponential Smoothing and Differencing. Transformation methods estimate and eliminate the trend component. There are several transformation methods such as logarithmic transformations, Box-Cox transformations etc.

Smoothing with Moving Average filter is a nonparametric method that is used to estimate the trend component.

Let q be a nonnegative integer and consider the two-sided moving average.

$$W_t = \frac{1}{2q+1}\sum_{j=-q}^{q} X_{t-j} \quad (3)$$

Trend estimates $\hat{m}_t = W_t$. But in this method choosing the value of q is a hard task.

Exponential Smoothing is based on a moving average of past values only. This is often used for forecasting, the smoothed value at present being used as the forecast of the next value. And exponential smoothing uses weighted averages of the past data.

Differencing the data apply the difference operator to the original time series to obtain a new time series.

Backward shift operator : $BX_t = X_{t-1}$ (4)

Lag-1 difference operator : $X_t = X_t - X_{t-1}$ (5)

Correlogram

Correlograms are plots of autocorrelations and partial autocorrelations associated with time-series data. Correlograms show the relationship between autocorrelation coefficients against increasing time lags and they are useful in visualizing the correlation structure in time-series data.

Model Adequacy Checking

Model adequacy checking is used to check the validity of a model. In this study heteroscedasticity and serial correlation, tests were drawn to check whether the model is adequate.

Heteroscedasticity Test

When the variances of errors are constant, we call it homoscedasticity. Concerning real-world scenarios, a non-constant error variance can be observed frequently. The violation of constant error variance is called as heteroscedasticity.

There are several reasons for heteroscedasticity such as, usual variation structures that occur due to natural factors, omitting important variables from the model, incorrect data transformations etc.

There are several tests to check heteroscedasticity, such as ARCH test, Park test, White's general heteroscedasticity test, Goldfeld-Quaindt test etc. Among them, we have used the ARCH test in this study.

The corresponding hypothesis for the heteroscedasticity test is,

$$H_0: No\ presence\ of\ ARCH\ effect$$

$$H_1: Presence\ of\ ARCH\ effect$$

Serial Correlation Test

Serial correlation occurs when the error terms in a time series transfer from one period to another period. That means the error for one time period is correlated with the error for a subsequent time period. Serial correlation can happen in the excluded variable case, incorrect functional form, inertia (sluggishness) etc. To check the serial correlation, we can use several tests such as Runs test, Durbin-Watson test, Ljung-Box Q test etc.

Here we have used Durbin-Watson test and the test statistic is given below,

$$d = \frac{\sum_{t=2}^{n}(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^{n}\hat{\varepsilon}_t{}^2} \tag{6}$$

where, d is the Durbin Watson statistic

$\hat{\varepsilon}_t, \hat{\varepsilon}_{t-1}$ are the error terms

The corresponding hypothesis for the serial correlation test is,

H_0: *No presence Serial Correlation*

H_1: *Presence of Serial Correlation*

Forecasting Accuracy

To know whether how well the model has been performed it is needed to consider the difference between actual and the forecasted values. It is essential to minimize the difference between actual and forecasted values because the model performance relies on that. That is the smaller the difference, the better the model is. Mean Absolute Percentage Error (MAPE) value has been used to assure the forecasting accuracy.

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i-\hat{y}_i|}{y_i} \times 100\% \qquad (7)$$

where, y_i = the actual value

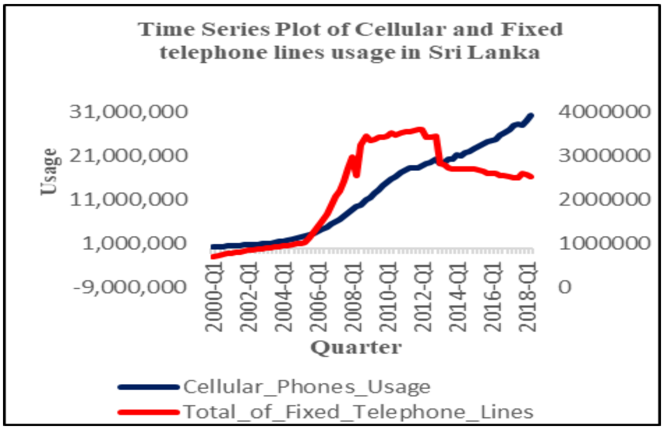
$\hat{y}_i$ = the fitted value

n = number of observations

4. Results & Interpretation

In our study, we have analysed the increment and the decrement of cellular and landline phones usage in Sri Lanka and it is illustrated in the below Figure 1.

Figure 1: Time Series plot of cellular and landline telephone lines usage in Sri Lanka



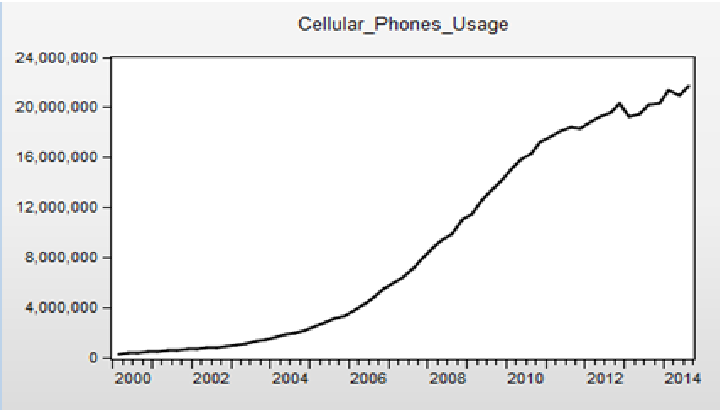
Source: Central bank of Sri Lanka (2018)

According to Figure 1, in the year 2000, the consumption of fixed telephone lines is higher than the consumption of cellular phones by 403,100. And it is clearly visualized that the fixed telephones consumption has phased out by the year 2018.

Time Series Plots

For model building, we have used 80% of the total dataset and it contains the quarterly data starting from the year 2000 quarter 1 to the year 2014 quarter 3. The rest of the 20% of data starting from the year 2014 quarter 4 to the year 2018 quarter 2 were used to validate our results.

Figure 2: Time Series plot of quarterly cellular phones usage

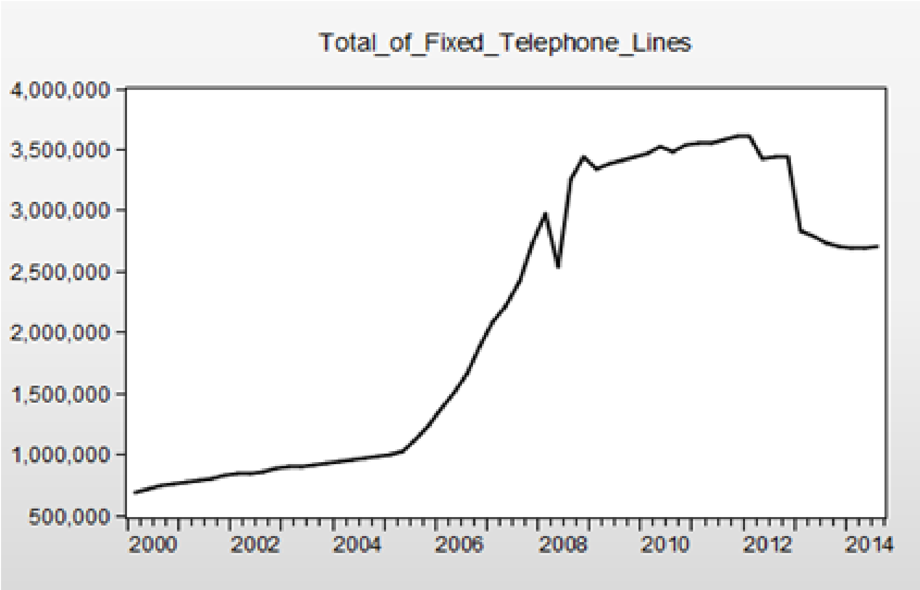


Source: Eviews Output (2020)

Figure 2 implies the variation of the quarterly cellular phones' usage in Sri Lanka. There can be observed consequent upward nonlinear trend implying the series is non-stationary. In the year 2000 quarter 1 the cell phones usage is 293,120 and in the year 2014 quarter 3, it has become 21,727,589. The difference between the usage in 2000 quarter 1 to 2014 quarter 3 is more than 21.4 million. This huge difference proves that within the past 14 years the cellular phones have become an absolutely necessary commodity in day to day life.

Figure 3 implies the variation of the quarterly fixed telephone lines usage in Sri Lanka. It can be observed an upward nonlinear trend up to the year 2012 quarter 3 and then from the year 2013 quarter 1 the usage of the fixed telephone starts to decrease gradually. Therefore, this series seems to be non-stationary.

Figure 3: Time Series plot of quarterly fixed telephone lines usage



Source: Eviews Output (2020)

In the year 2000 quarter 1 the usage of the fixed telephone is 696,220 and in the year 2014 quarter 3 that has become 2,711,717. The difference between the year 2000 quarter 1 to 2014 quarter 3 is more than 2 million. Compared to the year 2000, in the year 2014, 19 million people have to get used to cellular phones more than using landline phones in Sri Lanka. From these two graphs, it can be observed that both series are non-stationary. To confirm these result stationarity tests were performed.

Checking Stationarity

For KPSS test the hypothesis is given below,

$H_0: the\ series\ is\ stationary$

$H_1: the\ series\ is\ not\ stationary$

For ADF and PP tests the hypothesis is given below,

$H_0: the\ series\ is\ not\ stationary$

$H_1: the\ series\ is\ \ stationary$

Table 1: Stationarity test for the original data set

Stationary Test	Test statistic	For cellular phones usage	For fixed telephone lines usage	Result
KPSS Test	LM stat	0.9075	0.7929	Non-stationary
ADF Test	Prob	0.7661	0.6948	Non-stationary
PP test	Prob	0.9964	0.6873	Non-stationary

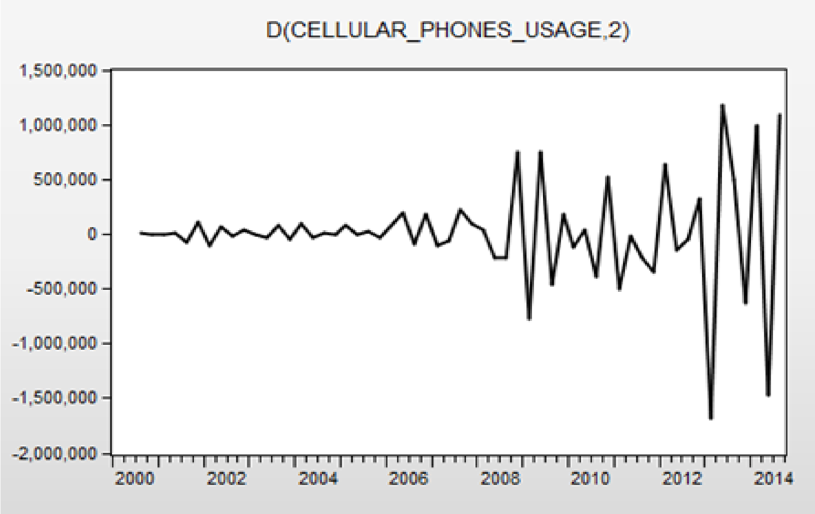
Source: Author Creations (2020)

Graphical and stationarity test results show that both series are not stationary. In order to convert these non-stationary series into stationary series, the differencing technique was applied. The first differencing has been performed and the results verified that the series is not stationary. So, the second differencing was further carried out.

Graphical and Stationarity Test Results After Second Differencing

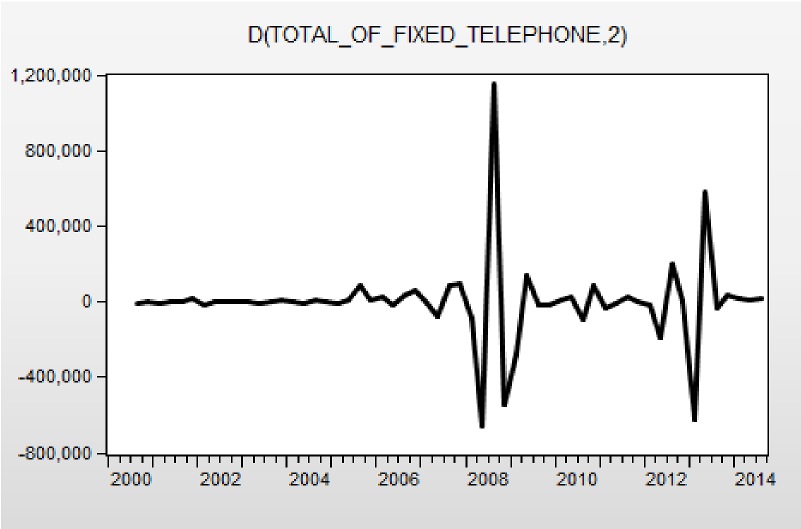
Figure 4 and 5 illustrates the graphical representation of the cellular and landline phones usage after applying the second differencing.

Figure 4: Time Series plot of quarterly cellular phones usage after taking the 2nd difference



Source: Eviews Output (2020)

Figure 5: Time Series plot of quarterly fixed telephone lines usage after taking the 2nd difference



Source: Eviews Output (2020)

Since there is no trend seems to be in the above Figure 4 and 5, they seem to be stationary.

The below Table 2 illustrates the stationarity test results after 2nd differencing.

Table 2: Stationarity test results after second differencing

Stationary Test	Test statistic	Cellular phones usage	Fixed telephone lines usage	Result
KPSS Test	LM stat	0.0417	0.0448	Stationary
ADF Test	Prob	0.0000	0.0000	Stationary
PP test	Prob	0.0001	0.0001	Stationary

Source: Author Creations (2020)

Both graphical and stationary test results confirm that both series are stationary after taking the 2nd difference.

Correlogram

Figure 6: Correlogram for cellular phones usage

Correlogram of D(CELLULAR_PHONES_USAGE,2)

Date: 11/04/19 Time: 11:55
Sample: 2000Q1 2014Q3
Included observations: 57

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.656	-0.656	25.879	0.000
		2	0.210	-0.388	28.583	0.000
		3	0.054	-0.018	28.765	0.000
		4	-0.337	-0.443	35.961	0.000
		5	0.417	-0.166	47.226	0.000
		6	-0.117	0.331	48.123	0.000
		7	-0.122	0.131	49.121	0.000
		8	0.214	0.086	52.265	0.000
		9	-0.267	0.131	57.257	0.000
		10	0.151	-0.010	58.887	0.000

Source: Eviews Output (2020)

Figure 7: Correlogram for fixed telephone lines usage

Correlogram of D(TOTAL_OF_FIXED_TELEPHONE,2)

Date: 11/04/19 Time: 14:06
Sample: 2000Q1 2014Q3
Included observations: 57

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.546	-0.546	17.911	0.000
		2	-0.097	-0.564	18.491	0.000
		3	0.224	-0.349	21.619	0.000
		4	-0.023	-0.147	21.652	0.000
		5	-0.062	0.011	21.903	0.001
		6	-0.010	0.006	21.910	0.001
		7	0.061	0.027	22.158	0.002
		8	-0.054	-0.052	22.355	0.004
		9	0.015	-0.060	22.370	0.008
		10	0.021	-0.029	22.402	0.013

Source: Eviews Output (2020)

In Figure 6, the significant cut-off lags of ACF and PACF plots are 1,4,5 and 1,2,4,6 respectively and in Figure 7, the significant cut-off lags of ACF and PACF plots are 1 and 1,2,3.

Since there is no seasonal pattern is visualized in above ACF and PACF plots, an ARIMA model was fitted instead of a Seasonal ARIMA model.

Model Fitting

The corresponding model fitting test results are shown in the below Figure 8 and 9.

Figure 8: ARIMA (1,2,1) model for cellular phones usage

Dependent Variable: D(CELLULAR_PHONES_USAGE,2)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 11/04/19 Time: 11:56
Sample: 2000Q3 2014Q3
Included observations: 57
Convergence achieved after 20 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5286.383	13921.40	0.379731	0.7057
AR(1)	-0.427722	0.117927	-3.627018	0.0006
MA(1)	-0.615109	0.092196	-6.671766	0.0000
SIGMASQ	9.72E+10	1.16E+10	8.388353	0.0000

R-squared	0.569705	Mean dependent var	11992.42
Adjusted R-squared	0.545349	S.D. dependent var	479460.7
S.E. of regression	323289.8	Akaike info criterion	28.29816
Sum squared resid	5.54E+12	Schwarz criterion	28.44153
Log likelihood	-802.4975	Hannan-Quinn criter.	28.35388
F-statistic	23.39048	Durbin-Watson stat	2.041512
Prob(F-statistic)	0.000000		

Source: Eviews Output (2020)

Figure 9: ARIMA (2,2,1) model for fixed telephone lines usage

Dependent Variable: D(TOTAL_OF_FIXED_TELEPHONE,2)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 11/04/19 Time: 14:08
Sample: 2000Q3 2014Q3
Included observations: 57
Convergence achieved after 36 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-781.6311	8641.572	-0.090450	0.9283
AR(1)	-0.525842	0.124134	-4.236098	0.0001
AR(2)	-0.386833	0.170329	-2.271098	0.0273
MA(1)	-0.544235	0.152297	-3.573504	0.0008
SIGMASQ	2.20E+10	3.57E+09	6.153672	0.0000

R-squared	0.584402	Mean dependent var	-176.7719
Adjusted R-squared	0.552433	S.D. dependent var	231941.2
S.E. of regression	155169.6	Akaike info criterion	26.85245
Sum squared resid	1.25E+12	Schwarz criterion	27.03167
Log likelihood	-760.2949	Hannan-Quinn criter.	26.92210
F-statistic	18.28026	Durbin-Watson stat	2.065050
Prob(F-statistic)	0.000000		

Source: Eviews Output (2020)

After considering all possible models, ARIMA (1,2,1) and ARIMA (2,2,1) models were identified as the best-fitted models with minimum AIC equals to 28.298 and 26.852 for forecasting the cellular phones usage and the fixed telephone lines usage respectively.

Model Adequacy Checking

This section outlines the results of model adequacy checking test results.

- *Heteroscedasticity Test*

$$H_0: No\ presence\ of\ ARCH\ effect$$

$$H_1: Presence\ of\ ARCH\ effect$$

According to heteroscedasticity test, the probability values 0.5349 and 0.1042 are greater than 0.05. Thus the null hypothesis is not rejected and therefore, there is no ARCH effect at 5% level of significance.

- *Serial Correlation Test*

$$H_0: No\ presence\ of\ Serial\ Correlation$$

$$H_1: Presence\ of\ Serial\ Correlation$$

Table 3: Durbin-Watson Test Results

	Durbin-Watson Statistic
For Cellular Phones	2.0415
For Landline Phones	2.0650

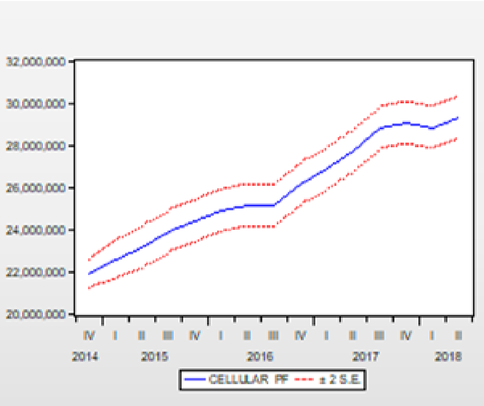
Source: Author Creations (2020)

In Table 3, Durbin-Watson Statistic values are 2.0415 and 2.0650. Since both values are closed to 2, it implies there is no serial correlation at 5% level of significance.

Forecasting

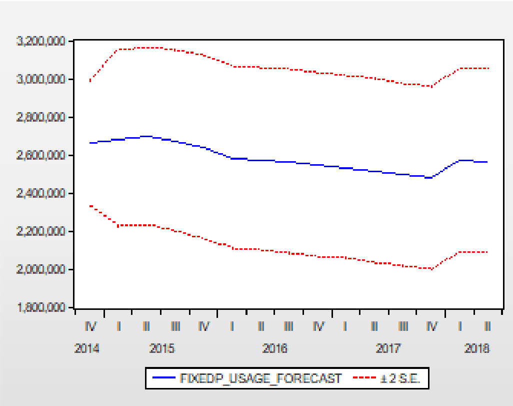
As mentioned in the previous section we have used 80% of data for model fitting and now we are moving to forecast our dataset using the remaining 20% of data.

Figure 10: Forecasting accuracy for ARIMA (1,2,1) model



Source: Eviews Output (2020)

Figure 11: Forecasting accuracy for ARIMA (2,2,1) model

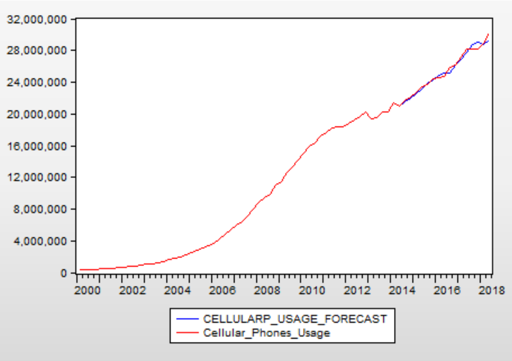


Source: Eviews Output (2020)

The measure the forecasting accuracy Mean Absolute Percentage Errors (MAPE) was used and the corresponding values are 1.403 and 0.976.

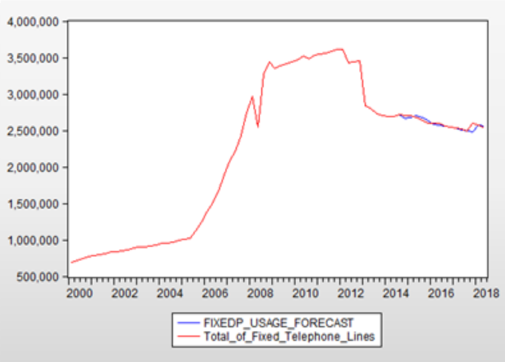
According to Figure 12 and 14, ARIMA (1,2,1) and ARIMA (2,2,1) models have a strong potential for forecasting the usage of cellular phones and landline phones respectively.

Figure 12: Forecasting with ARIMA (1,2,1) model



Source: Eviews Output (2020)

Figure 13: Forecasting with ARIMA (2,2,1) model



Source: Eviews Output (2020)

5. Conclusion and Future Work

This study mainly aims at modelling and forecasting the usage of cellular and landline phones in Sri Lanka. For this study, the quarterly phone's usage data were driven from the Central Bank of Sri Lanka. A univariate time series approach: ARIMA was used to analyse the data set. Here 80% of the data was used for modelling and 20% was used for forecasting. As the original data set of both of these series is not stationary, the differencing technique is used to make the series stationary. Here, the first differencing is also not stationary and so we have moved to second differencing. Second, differencing made both of our series stationary. Then observing the correlogram we have determined the corresponding cut-off lags and according to them AR and MA orders were decided. Then using these AR and MA orders we have tested for several ARIMA models. To finalize the best fitted ARIMA model, the minimum Akaike's Information Criteria (AIC) was used. The best-fitted models were ARIMA (1,2,1) and ARIMA (2,2,1) with minimum AIC 28.298 and 26.852 for modelling cellular and landline phones usage respectively.

Afterwards, we have checked for the assumptions of the constant variance (ARCH effect) and for the serial correlation. Since these assumptions were not violated by our models, they were used for forecasting purposes. The forecasting accuracy was measured by the mean absolute percentage error (MAPE).

The MAPE for cellular and landline phones are 1.403 and 0.976. This result clearly shows that the performance of both ARIMA models selected here is quite impressive and the actual and predicted values seem to be related to each other.

These models could guide the ones who are with the telecoms market to achieve their business goals. A further accurate forecast can be obtained if there were daily or weekly cellular and landline usage data records. And also, we can consider the factors such as the growth of the population, the price of the mobile and landline phones in the market, the tax by the government, monthly charges by corresponding service providers for a more accurate multivariate model. As a further study, machine learning techniques such as artificial neural network models are suggested to be used in achieving more accurate forecasting.

References

Aker, J. C., & Mbiti, I. M. (2010). Mobile Phones and Economic Development in Africa. *Economic Perspective*, 207-232.

Fowdur, T. P., Hurbungs, V., & Beeharry, Y. (2016). *Statistical analysis of energy consumption of mobile phones for web-based applications in Mauritius.* International Conference on Computer Communication and Informatics, ICCCI 2016, 1–8.

McBurney, P., Parsons, S., & Green, J. (2002). Forecasting market demand for new telecommunications services: An introduction. *Telematics and Informatics,* 19(3), 225–249. https://doi.org/10.1016/S0736-5853(01)00004-1

Nawaz, S. (2012). Statistical Study of Impact of Mobile on Student's Life. IOSR *Journal of Humanities and Social Science*, 2(1), 43–49. https://doi.org/10.9790/0837-0214349

Priya, R. K., & Jeevitha, P. (2017). Analysis of mobile phone usage and the academic performance of college students, *International Journal of Multidisciplinary Research and Development*, 4(2), 86–87.

Rashid, A. T., & Elder, L. (2009). Mobile Phones and Development: An Analysis of IDRC-Supported Projects. *The Electronic Journal of Information Systems in Developing Countries,* 36(1), 1–16. https://doi.org/10.1002/j.1681-4835.2009.tb00249.x

Thanh, D. Van, Lehne, P. H., Luke, G., Ling, R., Savage, R., Wilcox, N., … Calvet, L. (2005). Future Mobile Phones. *Telektronikk*, 101(3/4), 22–34. http://www.telenor.com/wp-content/uploads/2012/05/T05_3-4.pdf#page=24

Zajdel, R., Zajdel, J., Śmigielski, J., & Nowak, D. (2013). Cell phone ringtone, but not landline phone ringtone, affects complex reaction time. *International Journal of Occupational Medicine and Environmental Health*, 26(1), 102–112. https://doi.org/10.2478/S13382-013-0080-8

Yuan, A. Y., Allen, B., Brick, J. M., Dipko, S., Presser, S., Tucker, C., … Galesic, M. (2004). Surveying Households on Cell Phones — Results and Lessons Keywords : Cell, 4021–4027.