IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 7, July 2022

Exploring ARIMA Modeling for Accurate Forecasting of Telecom Data in Tamilnadu

K.C.Bhanu[1], Dr.P.Uma Maheswari Devi[2]

[1]Research Scholar, Department of Commerce and Management Studies
Adikavi Nannayya University ,Rajahmundry
[2]Associate Professor , Department of Commerce and Management Studies
Adikavi Nannayya University ,Rajahmundry
bhanu1605@gmail.com,umadevi_4@yahoo.com

ABSTRACT

This study analyzes Tamil Nadu telephone data per 100 people from 2004 to 2021 and forecasts trends for the following five years using the ARIMA model. The dataset includes Tamil Nadu's 2004–2021 telephone connections per 100 people. The Augmented Dickey-Fuller (ADF) test checks data stationarity before analysis. Non-stationary data can cause erroneous forecasts in time series analysis. The ADF test will detect if data needs differencing for stationarity. To determine ARIMA model orders, the ACF and PAF will be investigated. These functions provide data point correlations at different lags, helping pick ARIMA parameters. The ARIMA (0,2,1) model fits telephone data best based on ACF and PACF graphs. To achieve stationarity, the data must be differenced twice, and the model will have one moving average component lag. The ARIMA (0,2,1) model has AIC and BIC values of 126.19 and 127.73, respectively. Lower numbers imply greater model fit. The ARIMA (0,2,1) model predicts 2022–2025 telephone statistics. This projection will reveal Tamil Nadu's five-year telephone connection trend per 100 people. Finally, the Box test will evaluate ARIMA model goodness-of-fit. The Box test validates model predictions by detecting considerable residual autocorrelation. This detailed research of Tamil Nadu's telephone data patterns will assist stakeholders make educated decisions about communications infrastructure and services for the future.

Keywords: Telephone, ARIMA, Forecasting.

INTRODUCTION

The field of telecommunication occupies a central position in the contemporary global landscape, facilitating uninterrupted communication and facilitating the retrieval of information. The telecommunication sector in the state of Tamil Nadu, situated in the southern region of India, has experienced notable progressions throughout its history. The monitoring and prediction of telephone connections per 100 population in Tamil Nadu hold significant importance for policymakers, service providers, and other relevant stakeholders in order to efficiently strategize and allocate resources.

Time series forecasting techniques, such as the Autoregressive Integrated Moving Average (ARIMA) model, have demonstrated their efficacy in projecting future patterns by leveraging historical data. The ARIMA model is highly advantageous in the analysis of data that exhibits temporal dependencies, rendering it a suitable selection for forecasting telephone connections over a given time period.

The aim of this study is to examine the historical telephone data per 100 population in Tamil Nadu spanning from 2004 to 2021. Additionally, the study intends to utilize the ARIMA model to predict the future trend for the subsequent five-year period, specifically from 2022 to 2025. Through the utilization of this model, our objective is to offer significant insights pertaining to the prospective expansion and demand for telecommunication services within the state.

The analysis will be performed through a series of sequential steps. To begin with, we will assess the stationarity of the telephone data by employing the Augmented Dickey-Fuller (ADF) test. The concept of stationarity holds significant importance in the field of time series analysis, as it guarantees the constancy of statistical characteristics within the data throughout the course of time. In the event that the data is determined to lack stationarity, suitable differencing techniques will be employed in order to attain stationarity.

Subsequently, an examination of the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) will be conducted in order to ascertain the suitable order for the ARIMA model. These functions offer valuable insights into the relationship between the present observation and its previous lags, facilitating the selection of suitable autoregressive and moving average components for the model.

After determining the most suitable ARIMA model, we will proceed to estimate its parameters and assess its goodness-of-fit by employing the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). These criteria will aid in assessing the efficacy of the model in capturing the inherent patterns present in the telephone data.

After validating the ARIMA model, it will be employed to predict the telephone connections per 100 population in Tamil Nadu for the upcoming five years. The projected values will offer significant insights to stakeholders, empowering them to anticipate potential expansion, strategize infrastructure development, and make well-informed decisions pertaining to telecommunication services.

In summary, this research endeavor aims to enhance the collective comprehension of telephone usage patterns in Tamil Nadu, thereby offering significant insights for policymakers, telecommunication entities, and other relevant stakeholders. The utilization of the ARIMA

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 7, July 2022

model will enhance the precision of forecasting telephone connections per 100 population, thereby aiding in the development of strategies to address the changing telecommunication demands of the state in the future.

LITERATURE REVIEW:

Zhang et al., (2021) Research and application of traffic Forecasting in customer service center based on ARIMA Model and LSTM Neural Network Model. The correlation between traffic data and the allocation of call center seats is a fundamental principle in the field of artificial intelligence research. Optimal configuration of call center human resources can be achieved by arranging corresponding agents based on varying traffic volumes. This paper employs the ARIMA model and LSTM neural network model, both of which are grounded in time series analysis, to forecast traffic patterns. An empirical investigation is being conducted on Python software, utilizing the power call center traffic data from Hebei Province as a representative case study. The empirical findings indicate that the LSTM neural network model exhibits superior predictive accuracy in comparison to the ARIMA model.

Bianchi et al., (1998) Improving Forecasting for telemarketing centers by ARIMA modeling with intervention. This study examines approaches for anticipating telemarketing center calls for planning and budgeting. We compare additive and multiplicative Holt–Winters (HW) exponentially weighted moving average models to Box–Jenkins (ARIMA) intervention analysis. HW and ARIMA models are tested for telemarketing data forecasting. ARIMA models with intervention analysis outperform "simple models" like Holt–Winters for the time series analyzed, despite recent work suggesting otherwise.

Grambsch et al., (1990) Forecasting demand for special telephone services: A case study. ARIMA models and squared error criterion-based optimum rules are used to predict future demand for products and services. We discovered that a model with independent increments with stable distributions was superior for predicting Special Services in the telephone company when evaluating a large number of time series. It accurately characterized predicted mistakes. This article describes the model, compares it to a state space model utilized for the problem, and uses numerous data analytic processes to determine how well the model matches the data. The report concludes with comments on projected forecast error magnitude.

Zhang et al., (2010) Predicting social ties in mobile phone networks. Since social interactions evolve, a social network does too. Social ties drive social network development. Even within a group, social-tie strengths vary. We use mobile phone call-detail records to assess and forecast social connection strengths. We present a reciprocity index-based affinity model to quantify social-tie strengths. The affinity model maps call-log data to social-tie strengths over time as human social interactions vary. ARIMA model predicts social-tie strengths. over confirmation, we used actual call records of 81 users acquired over 8 months at MIT by the Reality Mining Project group and 20 users collected for 6 months by UNT's Network Security team. These users

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 7, July 2022

have 5000 contacts. Our model worked in experiments. We predict socially close and near members with 95.2% accuracy. This research aids homeland security, spam detection, and marketing.

METHOD AND APPROACH

Time series :- A time series is a sequence of data points or observations collected and recorded over a specific period of time, where the data points are ordered based on their corresponding time index. In other words, it is a collection of data points that are indexed or labeled by time.

Time series data is commonly encountered in various fields, including finance, economics, engineering, environmental science, and many others. It provides valuable insights into the behavior and patterns of phenomena that evolve over time.

The characteristics of time series data include:

1. Time Index: Each data point in a time series is associated with a specific time index or timestamp, indicating when the observation was made.

2. Sequential Order: The data points in a time series are arranged in chronological order, with each subsequent observation occurring after the previous one.

3. Temporal Dependence: Time series data often exhibits a certain level of dependence or correlation between observations. The value at a given time point can be influenced by its previous values or exhibit patterns over time.

4. Seasonality: Many time series exhibit regular patterns or variations that repeat at fixed intervals, known as seasonality. For example, retail sales may have higher values during holiday seasons.

5. Trend: Time series data often shows a long-term trend or systematic changes over time. Trends can be increasing (upward trend), decreasing (downward trend), or exhibit more complex patterns.

6. Randomness: Time series data can also contain random or unpredictable fluctuations, known as noise or random variations. These random components make it challenging to accurately forecast future values.

Analyzing time series data involves various techniques and methods, including trend analysis, seasonality detection, forecasting, and modeling. Time series analysis aims to understand and

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 7, July 2022

extract useful information from the data, uncover underlying patterns, and make predictions about future behavior based on historical observations.

Box-Jenkins Methodology

Box-Jenkins, also known as the Box-Jenkins methodology or Box-Jenkins approach, is a widely used and powerful technique for time series analysis and forecasting. It was developed by George Box and Gwilym Jenkins in the 1970s and has become a standard approach in the field of time series modeling. The Box-Jenkins methodology consists of three main steps: model identification, model estimation, and model diagnostic checking.

1. **Model Identification:**

The first step in the Box-Jenkins approach is to identify an appropriate model that best represents the underlying structure of the time series. This involves determining the orders of autoregressive (AR), differencing (I), and moving average (MA) components, denoted as (p, d, q), respectively. Model identification is typically done through the analysis of autocorrelation function (ACF) and partial autocorrelation function (PACF) plots, which help identify the significant lags and the presence of autoregressive and moving average patterns.

2. **Model Estimation:**

Once the model has been identified, the next step is to estimate the model parameters using a suitable estimation method. Maximum likelihood estimation (MLE) is commonly used for estimating the parameters of the ARMA model. The estimation process involves finding the values of the AR and MA coefficients that maximize the likelihood function based on the available data.

3. **Model Diagnostic Checking:**

After estimating the model, it is important to assess the adequacy of the model by conducting diagnostic checks. This involves analyzing the residuals (the differences between the observed values and the values predicted by the model) to ensure they meet certain assumptions, such as being normally distributed and exhibiting no significant autocorrelation. Various statistical tests and plots, such as ACF of residuals and Ljung-Box test, are employed to assess the model's goodness of fit.

If the diagnostic checks reveal that the model assumptions are not adequately met, further iterations of model identification, estimation, and diagnostic checking may be performed to refine the model until a satisfactory fit is achieved.

The Box-Jenkins methodology is known for its flexibility and versatility, allowing for the modeling of complex time series patterns. It has been successfully applied in various fields for forecasting, analyzing, and modeling time series data.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 7, July 2022

ARIMA

The ARIMA (Autoregressive Integrated Moving Average) model is a widely used time series forecasting method. It combines autoregressive (AR), differencing (I), and moving average (MA) components to capture the underlying patterns and dependencies in a time series. Here's the mathematical description of an ARIMA (p, d, q) model:

Let Y_t be the observed value of the time series at time t.

1. **Autoregressive (AR) component:**

The autoregressive part models the relationship between the current observation and a linear combination of past observations.

AR(p) component:

$Y_t = c + φ_1 Y_{t-1} + φ_2 Y_{t-2} + ... + φ_p Y_{t-p} + \varepsilon_t$

Here, c is a constant term, $φ_1$, $φ_2$, ..., φ_p are the autoregressive coefficients, ε_t is the error term assumed to be white noise with zero mean and constant variance.

2. **Differencing (I) component:**

The differencing part helps in removing trend or seasonality from the time series data. It calculates the difference between consecutive observations to achieve stationarity.

I(d) component:

$Y'_t = Y_t - Y_{t-d}$

Here, Y'_t is the differenced series, Y_t is the original series, and d represents the order of differencing.

3. **Moving Average (MA) component:**

The moving average part models the dependency between the current observation and a linear combination of past error terms.

MA(q) component:

$Y_t = μ + φ_1 \varepsilon_{t-1} + φ_2 \varepsilon_{t-2} + ... + φ_q \varepsilon_{t-p}$

Here, μ is the mean of the time series, $θ_1$, $θ_2$, ..., $θ_q$ are the moving average coefficients, and $\varepsilon_{t-1}, \varepsilon_{t-2}, ..., \varepsilon_{t-p}$ are the lagged error terms.

Combining all three components, the ARIMA(p, d, q) model can be written as:

$Y'_t = c + φ_1 Y'_{t-1} + φ_2 Y'_{t-2} + ... + φp Y'_{t-p} + φ_1 \varepsilon_{t-1} + φ_2 \varepsilon_{t-2} + ... + φ_q \varepsilon_{t-p} + \varepsilon_t$

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 7, July 2022

Here, Y'_t represents the differenced series after applying the autoregressive and moving average components, ε_t and is the error term in the model.

The parameters p, d, and q are determined based on the characteristics of the time series data and are typically estimated using methods like autocorrelation function (ACF) and partial autocorrelation function (PACF) analysis.

ACF and PACF

ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function) are statistical tools used to analyze and identify the correlation structure within a time series. Both ACF and PACF provide insights into the relationship between an observation and its lagged values. Here's the mathematical description of ACF and PACF:

1. Autocorrelation Function (ACF):

The ACF measures the correlation between an observation and its lagged values at different time lags.

ACF at lag k, denoted as ρ(k), is defined as:

$\rho(k) = \mathrm{Corr}(Y_t, Y_{t-k})$

Here, Y_t and Y_{t-k} are the values of the time series at time t and t-k, respectively. Corr represents the correlation coefficient between the two variables.

The ACF provides a measure of the linear relationship between the current observation and its lagged values. It quantifies the extent to which the values in the time series are correlated with each other over different lags. The ACF plot displays the correlation coefficients for various lags, which helps in understanding the presence of autocorrelation in the data.

2. **Partial Autocorrelation Function (PACF):**

The PACF measures the correlation between an observation and its lagged values, accounting for the intermediate correlations through the removal of the effects of shorter lags.

PACF at lag k, denoted as φ(k,k), is defined as:

$\varphi(k,k) = \mathrm{Corr}(Y_t, Y_{t-k} \mid Y_{t-1}, Y_{t-2}, ..., Y_{t-k+1})$

Here, Y_t, Y_{t-k}, and Y_{t-1}, Y_{t-2}, ..., Y_{t-k+1} represent the values of the time series at the respective time points.

The PACF helps in identifying the direct relationship between the current observation and its lagged values, after accounting for the intermediate correlations. It provides insights into the

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 7, July 2022

unique contribution of each lag to the current observation, which is useful for determining the appropriate lag order in autoregressive models.

Both ACF and PACF are commonly used in time series analysis to identify the order of autoregressive (AR) and moving average (MA) components in models like ARIMA and to understand the correlation structure of the time series data.

ARMA

ARMA (Autoregressive Moving Average) is a popular time series model that combines autoregressive (AR) and moving average (MA) components to capture the underlying patterns and dependencies in a time series. The ARMA model is characterized by two parameters, p and q, representing the orders of the AR and MA components, respectively. Here's the mathematical description of an ARMA(p, q) model:

Let Y_t be the observed value of the time series at time t.

1. Autoregressive (AR) component:

The autoregressive part models the relationship between the current observation and a linear combination of past observations.

AR(p) component:

$Y_t = c + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + ... + \varphi_p Y_{t-p} + \varepsilon_t$

Here, c is a constant term, φ_1 , φ_2 , ..., φ_p are the autoregressive coefficients, ε_t is the error term assumed to be white noise with zero mean and constant variance.

2. **Moving Average (MA) component:**

The moving average part models the dependency between the current observation and a linear combination of past error terms.

MA(q) component:

$Y_t = \mu + \varphi_1 \varepsilon_{t-1} + \varphi_2 \varepsilon_{t-2} + ... + \varphi_q \varepsilon_{t-p} + \varepsilon_t$

Here, μ is the mean of the time series, θ_1 , θ_2 , ..., θq are the moving average coefficients, ε{t-1}, ε{t-2}, ..., ε{t-q} are the lagged error terms, and ε_t is the error term assumed to be white noise with zero mean and constant variance.

Combining the AR and MA components, the ARMA(p, q) model can be written as:

$Y_t = c + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + ... + \varphi p Y_{t-p} + \varphi_1 \varepsilon_{t-1} + \varphi_2 \varepsilon_{t-2} + ... + \varphi_q \varepsilon_{t-p} + \varepsilon_t$

AIC, AICC, BIC

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 7, July 2022

AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion), and AICC (Corrected Akaike Information Criterion) are statistical measures used for model selection and comparison in the context of time series analysis and other statistical modeling. They provide a quantitative assessment of the goodness of fit of different models and help in selecting the most appropriate model based on their respective criteria. Here's the mathematical description of AIC, BIC, and AICC:

1. **Akaike Information Criterion (AIC):**
AIC is a measure of the relative quality of a statistical model. It balances the trade-off between goodness of fit and model complexity, penalizing models with a larger number of parameters.

AIC for a model with parameter p and log-likelihood value L is calculated as:

$$AIC = -2 * L + 2 * p$$

Here, a lower AIC value indicates a better model fit. The term -2 * L represents the maximized log-likelihood of the model, and 2 * p is the penalty term that accounts for the number of parameters in the model.

2. **Bayesian Information Criterion (BIC):**
BIC is a criterion similar to AIC but places a stronger penalty on the number of parameters. It incorporates a prior belief that simpler models are more likely to be true.

BIC for a model with parameter p and log-likelihood value L is calculated as:

$$BIC = -2 * L + p * \log(n)$$

Here, n represents the sample size. BIC penalizes models with a larger number of parameters more strongly than AIC, making it more suitable for model selection when the sample size is relatively small.

3. **Corrected Akaike Information Criterion (AICC):**
AICC is a modification of AIC that adjusts for small sample sizes. It takes into account both the model complexity and the sample size to provide a more accurate measure of model fit.

AICC for a model with parameter p, log-likelihood value L, and sample size n is calculated as:

$$AICC = AIC + (2 * p * (p + 1)) / (n - p - 1)$$

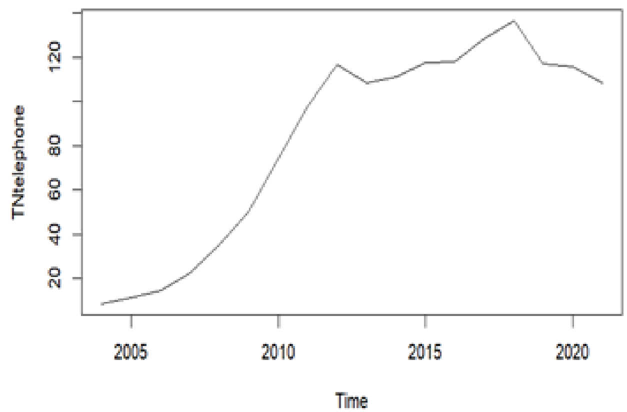
The additional term in AICC accounts for the correction based on the number of parameters and the sample size. AICC is particularly useful when the sample size is small and provides a more reliable measure of model fit compared to AIC.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 7, July 2022

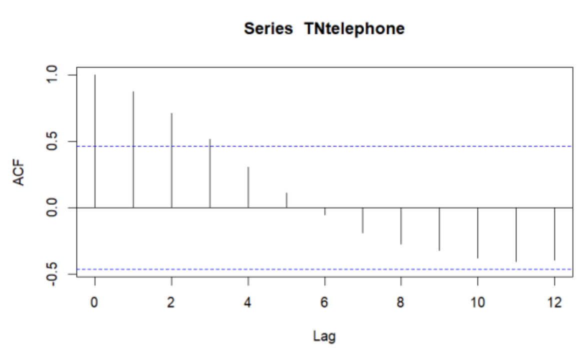
RESULT:

The following graph shows data on the number of telephones per 100 people from 2004 to 2021. 2006 to 2008 marked the beginning of an upward trend in that graph, which will continue until 2021.



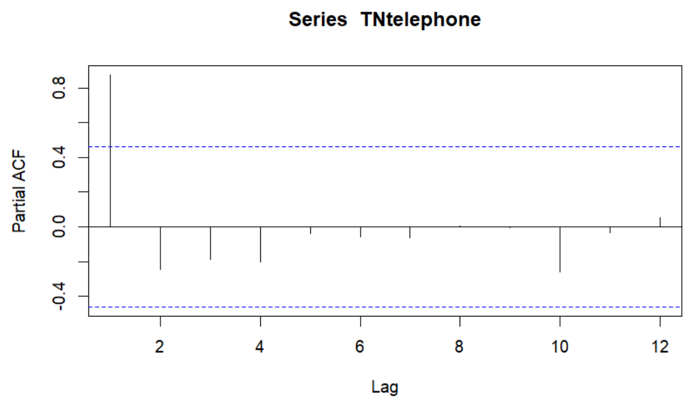
Finding the AR and MA polynomial order for Tamil Nadu Telephone per 100 population by comparing stationary and non-stationary data. Plotting the autocorrelation functions allowed for the determination of the proper order of autoregressive and moving average polynomials, as well as the values of p and q. The fact that virtually all of the autocorrelations lag, or (n/4th), are considerably different from zero demonstrates that the data are not stable. With non-stationary data series pertaining to Tamil Nadu. It was found that the ACF plot in figure was not steady in its behavior. On the other hand, PACF's figure reveals a substantial spike at lag 1, which suggests that the series may have an autoregressive component of order one.

In order to turn non-stationary data series for Tamil Nadu into stationary data series, the initial step was to differentiate the original data series. When we make use of the auto arima function, it displays a number of possible arima models and automatically converts data into the stationary format. This format was sufficient for obtaining a suitable stationary series in Tamil Nadu. The autocorrelations of Telephone series now validates the non-stationary behaviors of telephone usage.



ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 7, July 2022



Modelling Components

Several alternative ARIMA models, each with a unique combination of lags of moving average and autoregressive orders, were utilized in the fitting process for the population projection of Tamil Nadu Telephone 100. During the identification step, the ARIMA(2,2,2), ARIMA(0,2,0), ARIMA(1,2,0), ARIMA(0,2,1), ARIMA(1,2,2), ARIMA(0,2,2), and ARIMA(1,2,2) models were all taken into consideration. The ARIMA estimate was calculated using a method known as non-linear least squares (NLS). For the sake of achieving this goal, the author Marquardt 1913 utilized an approach to parameter estimation under ARIMA that is rather popular. The calculated parameters for the provisional ARIMA models are presented in the table.

ARIMA MODELS	Estimate Value
ARIMA(2,2,2)	131.7182
ARIMA(0,2,0)	126.6666
ARIMA(1,2,0)	127.0533
ARIMA(0,2,1)	126.1863
ARIMA(1,2,1)	127.7301
ARIMA(0,2,2)	127.7919
ARIMA(1,2,2)	129.7755

The table that follows presents many ARIMA models for the Telephone Industry in Tamil Nadu, with ARIMA proving to be the most accurate (0,2,1) model. Mainly due to the fact that the estimate value is lower than that predicted by other models. The Auto.arima() method is responsible for managing the process of automatically include a constant. When d is more than one, the constant is never included; however, when d is less than zero or equal to one, it is always included if the AIC value is improved by its presence. After then, we made our predictions for the post-sample era in Tamil Nadu using these models, which lasted from 2004 to 2021.

The below table shows the coefficients of ARIMA Models(0,2,1) and value of AIC, AICc, BIC:

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 7, July 2022

Coefficients	MA1
Standard error	-0.4749
	0.2873

σ^2=127.4, log likelihood =-61.09, AIC =126.19, AICc=127.11, BIC=127.73.

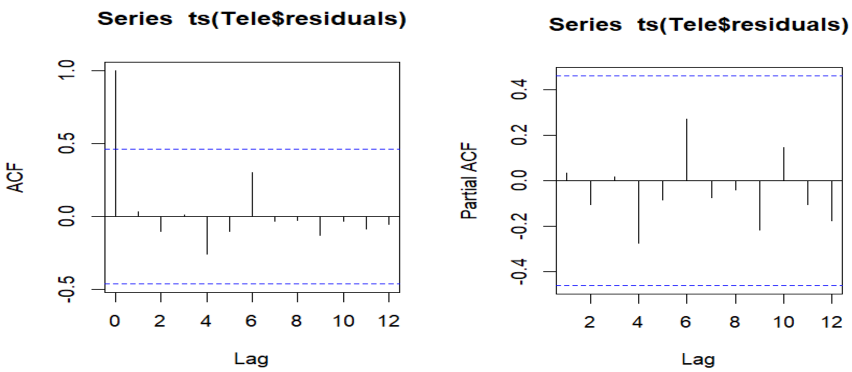
Diagnostic Screening

The model was analyzed by looking at the residuals to discover if there were any systematic patterns that could be removed to improve the selected ARIMA models. If there were any, they were looked for and deleted. After conducting tests using a number of different combinations of delays for the autoregressive and moving average, it was determined that the ARIMA (0,2,1) was the most accurate model for estimating the number of telephones in use by each population in Tamil Nadu.

The fitted model(0,2,1)

LJung-Box Q Statistic			
Model	X^2	df	Sig
ARIMA(0,2,1)	2.2662	5	0.8662

Residual ACFs and PACFs plot of Tamil Nadu based on ARIMA model



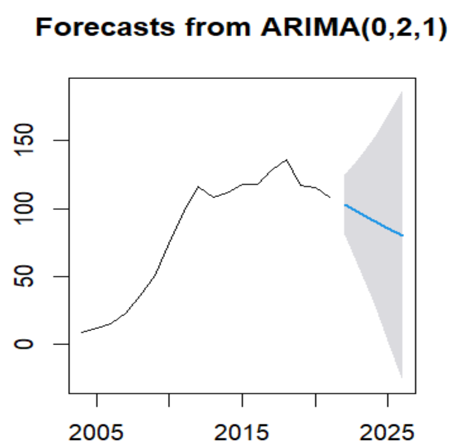
From 2022 to 2026, the forecast year and value for telephone per 100 population in Tamil Nadu

Year	Forecast Value	Lo(95)	Hi(95)
2022	102.54152	80.419087	124.6640
2023	96.86304	56.518176	137.2079
2024	91.18456	30.482027	151.8871

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 7, July 2022

2025	85.50608	2.256155	168.7560
2026	79.82760	-28.037340	187.6925



CONCLUSION:

In conclusion, our research into the patterns and possible future scenarios in Tamil Nadu's telecommunications industry using the ARIMA model (0,2,1) for the number of telephone connections per 100 population from 2004 to 2021 has been quite fruitful.

There is a discernible trend in the number of telephone connections over the years, as shown by the statistics. There was an ongoing rising trend after 2005, suggesting that more people were using and having access to telephones. Increased mobile penetration and the general acceptance of communications services are likely causes of this expansion.

But starting in 2016, we saw a decline in the number of telephone connections per 100 people. This drop might be due to a number of variables, including but not limited to market saturation, changes in customer behavior, or changes in communication preferences.

The ARIMA model's projection for the next five years of Tamil Nadu's telephone connections per 100 people confirms the downward trend shown since 2016. This research demonstrates the need of keeping tabs on and analyzing people's communication habits in order to meet their ever-evolving demands and preferences. This knowledge may be used by policymakers, telecom firms, and other interested parties to develop policies that boost development and sustainability in the industry while keeping up with the evolving needs of the digital sphere.

We used the ARIMA (0,2,1) model based on our study to project the number of telephone connections for the next five years (2022-2026). This model includes a single lag of the moving average component and takes into account twice-differenced data to ensure stationarity.

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, Iss 7, July 2022

The ARIMA model predicts that the number of telephone connections per 100 people will continue to decline during the next few years. As an early warning of future issues, and as a prerequisite for the formulation of strategic plans to manage the diminishing demand for telephone services, this projection is important for stakeholders in the telecommunication sector, government agencies, and service providers.

The declining rate of new phone lines installed might reflect a general trend away from using landlines or the advent of more sophisticated means of communication that have supplanted them. The results show how crucial it is for the telecom sector to develop and change with the times.

REFERENCES

1. Tych, W., Pedregal, D. J., Young, P. C., & Davies, J. (2002). An unobserved component model for multi-rate forecasting of telephone call demand: the design of a forecasting support system. *International Journal of forecasting*, *18*(4), 673-695.
2. Grambsch, P., & Stahel, W. A. (1990). Forecasting demand for special telephone services: a case study. *International Journal of Forecasting*, *6*(1), 53-64.
3. Madden, G., Savage, S. J., & Coble-Neal, G. (2002). Forecasting United States–Asia international message telephone service. *International Journal of Forecasting*, *18*(4), 523-543.
4. Bianchi, L., Jarrett, J., & Hanumara, R. C. (1998). Improving forecasting for telemarketing centers by ARIMA modeling with intervention. *International Journal of Forecasting*, *14*(4), 497-504.
5. Andrews, B. H., & Cunningham, S. M. (1995). LL Bean improves call-center forecasting. *Interfaces*, *25*(6), 1-13.
6. He, W., Qin, X., Jia, Z., Chang, C., & Cao, C. (2014). Forecasting of Busy Telephone Traffic Based on Wavelet Transform and ARIMA-LSSVM. *International Journal of Smart Home*, *8*(4), 113-122.
7. Bianchi, F. M., Scardapane, S., Uncini, A., Rizzi, A., & Sadeghian, A. (2015). Prediction of telephone calls load using echo state network with exogenous variables. *Neural Networks*, *71*, 204-213.
8. Piccolo, D. (1990). A distance measure for classifying ARIMA models. *Journal of time series analysis*, *11*(2), 153-164.
9. de Boer, T. R., Mérelle, S., Bhulai, S., Gilissen, R., & van der Mei, R. (2023). Forecasting call and chat volumes at online helplines for mental health. *BMC Public Health*, *23*(1), 1-13.
10. Antipov, A., & Meade, N. (2002). Forecasting call frequency at a financial services call centre. *Journal of the Operational Research Society*, *53*(9), 953-960
11. Hillmer, S. C. (1982). Forecasting time series with trading day variation. *Journal of Forecasting*, *1*(4), 385-395.
12. Liu, Q., & Wang, H. (2015). Research on the forecast and development of China's public fiscal revenue based on ARIMA model. *Theoretical Economics Letters*, *5*(04), 482.