eISSN: 2976-1190

Journal of Engineering and Sciences

Vol. 2 Issue 1 November 30, 2023

Journal homepage: ***Journal of Engineering and Sciences (nepjol.info)***

Publisher: Research Management Cell, Pashchimanchal Campus, Lamachour, Pokhara

Analysis of Forecasting Techniques for the Growth of Vehicular Population in Nepal

Bibek Mishra[1], Ramesh Bastola[1], Narayan Prasad Dawadi[1]

[1]*Department of Civil Engineering, IOE, Pashchimanchal Campus, Tribhuvan University, Nepal*

(Manuscript Received:06/11/2023; Revised:28/11/2023; Accepted:30/11/2023)

Abstract

Forecasting of vehicular population is a critical process that entails predicting future values based on analyzing past trends and considering explanatory variables, such as economic and demographic factors. This becomes especially vital in the context of Nepal, where the accurate prediction of the future growth of the vehicular population is paramount for achieving sustainable transportation systems. This study employs three distinct forecasting methods: trend line analysis, econometric analysis, and time series analysis. These methods have been rigorously evaluated to assess their respective levels of accuracy in predicting Nepal's vehicular population. Notably, the results from time series analysis, particularly the ARIMA model, have demonstrated a remarkable level of precision compared to the traditional trend line and econometric analysis approaches. The superiority of the ARIMA model underscores its efficacy as the preferred method for accurate vehicular population forecasting, providing a reliable foundation for future planning and policy implementation. The forecasted figures for Nepal's vehicular population indicate anticipated counts of 8,914,793 for 2030 AD, 1,482,842,6 for 2040 AD, and 2,203,801,2 for 2050 AD. These predictions offer transportation planners invaluable insights for the effective implementation of new projects, ensuring that resources are optimally allocated and transportation sustainability is realized.

Keywords: ARIMA; Econometric; Forecasting; Time Series; Trend Line; Vehicular population

1. Introduction

In recent decades, the growth of the vehicular population has become increasingly problematic due to the inability of the existing road network to handle the demand, resulting in social and economic inconvenience. Knowledge of future traffic flow is essential in planning, implementing, and developing a transportation system. It also helps in its operation, management and control[1]. Traffic forecasting is the process of predicting how vehicles will move on a particular road or network of roads in the future. To do this, historical traffic data is analyzed and factors such as population growth, land use patterns, and economic conditions are considered. Traffic forecasting plans transportation infrastructure, improves traffic flow, and reduces congestion. It also plays a crucial role in transportation policy and decision-making by informing stakeholders of the potential impacts of different transportation projects. Furthermore, forecasting is necessary for conducting economic analyses related to transportation projects [2].

The growth of the vehicular population has become quite challenging in the metropolitan cities of Nepal, and the urban streets of Nepal have faced congestion problems. Forecasting vehicular numbers is the prime concern for infrastructure planners and Engineers as it is essential for planning and developing infrastructure projects related to transportation and its networking. Accurate estimation of vehicular growth in a region is essential for transportation planning, implementation of traffic rules and regulations, pavement design, environmental concerns and other road infrastructure elements. Predicting vehicular growth is an essential factor for efficient planning of future road networks. Developments in automobile technology and the rise in vehicles on the streets have made traffic management quite challenging [3].

This paper includes forecasting of the vehicular population of Nepal based on the past trend of vehicular growth. The three methods adopted are Trend Line Analysis, Econometric Analysis and Time Series Analysis. Among these methods,

*Corresponding author. Tel.: +977- 0000-00-000,
E-mail address: bibekmishra792@gmail.com

the level of accuracy is questionable to the estimators or researchers. Based on the accuracy level, suitable methods are selected for the future prediction of Nepal's vehicular population.

1.1 Problem Statement and Objective of the Study

Transportation infrastructure projects like highways, bridges, and tunnels are designed to accommodate a certain traffic volume level. Accurately forecasting the expected growth in vehicular population is crucial for determining these infrastructure projects' appropriate size and capacity. However, inaccurate forecasts can lead to overestimation or underestimation of traffic volume, resulting in over or under-designing of the infrastructure and finally to wastage of resources.
This study concludes the following two objectives:

- To evaluate the effective forecasting technique for the growth of Nepal's vehicular population.
- To predict the anticipated figure of vehicular population for the target year.

2. Literature Review

In the context of Nepal, past research has mainly concentrated on short-term forecasting [4] and the traffic volume levels of the Kathmandu Ring Road have been predicted using a multiplicative decomposition forecasting model. Frequently used forecasting models such as ARIMA, SARIMA, etc., require extensive traffic data collection, which may not be feasible for predicting the short-term traffic volume. Approximate non-parametric regression method, traffic forecasting is a process predicting a dynamic variable [5]. That is why several approaches may be adopted for traffic forecasting depending upon the situation at hand. Although there can be various traffic volume forecasting methods, three of the most relevant methods were chosen for comparative analysis in this study due to data availability constraints.
Research conducted in India demonstrated the implication of three different analyses of traffic forecasting, i.e., trend line analysis, econometric analysis, and time series analysis, on data from the past 25 years. It forecasts the traffic volume after ten years. They considered vehicular population, per capita income, gross national product, and people. The research paper concluded that econometric and time series analyses are more accurate than trend line analyses. The study suggested that trend line analysis may be suitable for long-term forecasting [6]. The econometric model has vital significance in the transportation sector. The study of the air passenger demand model shows a strong statistical correlation between GDP, tourist arrivals, and passenger air transport demand [7].
The research paper titled "Growth Rate of Motor Vehicles in India - Impact of Demographic and Economic Development" explores the factors influencing the rapid growth of motor vehicles in India. The study finds that economic development, population growth, and urbanization contribute to increasing motor vehicle ownership. The rise in personal vehicles, driven by economic factors such as rising incomes, has led to concerns regarding environmental impacts and strain on urban transport infrastructure [8].
A study addresses the challenge of requiring a large amount of data when using the ARIMA model by introducing a seasonal ARIMA (SARIMA) model, which can be used with limited data. The study focused on traffic flow in a specific roadway section in India, and the researchers developed a SARIMA model to forecast the traffic volume data. The model's predictions were compared with the actual data, and the mean absolute percentage error (MAPE), which fell within an acceptable range [9], was calculated.

3. Methodology

The methodology for the study was designed or developed with a series of steps extracted from the previous literature review of the eminent issues that confront traffic forecasting with minimum errors that might distort the evaluation outcome.
The period selected for the collection of data is 1989-2019. The necessary data for this study are Vehicle Population, Population, Gross National Product, Gross Domestic Product, Per Capita Income, and Labor Force. The prediction for Nepal's vehicular population has been made for the target years 2030, 2040, and 2050 AD. These predictions provide anticipated vehicle numbers for these specific years and can serve as valuable inputs for transportation planners when implementing transportation projects effectively.

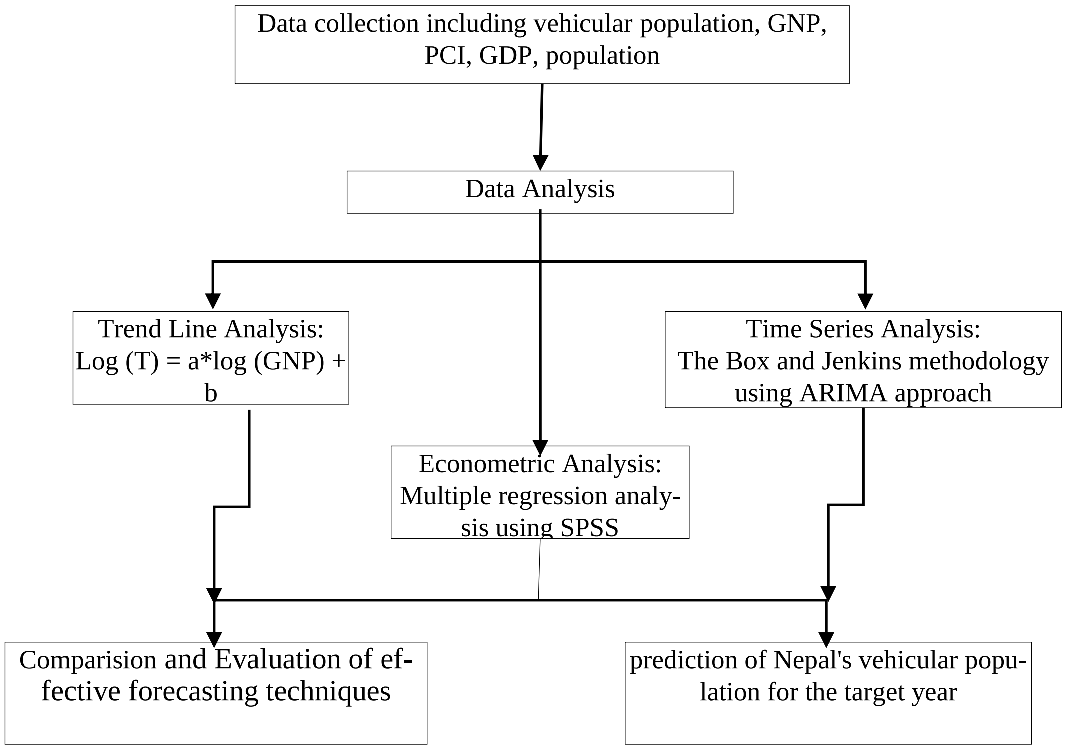

Figure 1. Flowchart of Research Methodology

3.1 Data Collection

For vehicle registration, the Vehicle & Transport Management Act, 2049 (1992) and Vehicle & Transport Management Rule, 2054 (1997) of Nepal classify vehicles into the following 5 main categories based on size and capacity as in Table 1.

The variables used in this analysis are vehicular population, PCI (Per Capita Income), GNP (Gross National Product), Population, and GDP per capita. Trend line analysis was performed using two variables i.e., Vehicular population and Gross national product [6], and similarly for econometric analysis population and per capita income have been selected as independent variables. This choice is based on the understanding that an increase in population is likely to contribute to the growth of the vehicular population. As the overall population expands, there tends to be a corresponding rise in private vehicle ownership. Additionally, the inclusion of per capita income as an independent variable is motivated by the expectation that increasing individual income levels will boost people's purchasing power. This, in turn, is anticipated to result in a higher vehicular population, as individuals with higher incomes are more likely to acquire vehicles. Therefore, these variables are considered essential factors in predicting and understanding the trends in the vehicular population in Nepal. These data are collected from a few sites like the CEIC global database, Ministry of Finance, World Bank, Macrotrends, and Economic Survey 79/80 as in Table 2.

Table 1. Number of Vehicles of all Category

Year	Heavy and medium-sized vehicle	Light vehicle	Two-wheeler	Tractor and power-trailer	Three-wheeler	Others	Total	Vehicular Population (T)
1989	10330	23414	34576	5417	2359	102	**76198**	**76198**
1990	1292	2790	5697	965	856	1549	**13149**	**89347**
1991	2055	3092	9336	1342	1207	435	**17467**	**106814**
1992	2097	2451	8513	751	62	381	**14255**	**121069**
1993	2908	3170	10550	1396	213	372	**18609**	**139678**
1994	2479	3126	11401	1814	241	353	**19414**	**159092**
1995	1637	4056	12357	2183	117	58	**20408**	**179500**
1996	1515	4696	15739	1278	185	352	**23765**	**203265**
1997	2190	4269	12306	1265	344	51	**20425**	**223690**
1998	1850	2526	17090	2248	388	37	**24139**	**247829**
1999	1323	3769	19755	2542	789	102	**28280**	**276109**
2000	2474	5402	29291	3519	232	77	**40995**	**317104**
2001	2666	4854	36117	3189	248	86	**47160**	**364264**
2002	2225	3436	29404	2485	17	43	**37610**	**401874**
2003	2687	8200	26547	2191	16	58	**39699**	**441573**
2004	2345	5650	31273	1374	48	21	**40711**	**482284**
2005	3827	5843	44610	635	0	0	**54915**	**537199**
2006	5578	6100	72568	2942	12	1535	**88735**	**625934**
2007	6601	5951	68667	3297	18	206	**84740**	**710674**
2008	6773	7578	83334	4663	20	202	**102570**	**813244**
2009	8387	13193	168707	11460	9	31	**201787**	**1015031**
2010	6666	9995	138907	7937	2	133	**163640**	**1178671**
2011	6399	10036	145135	8413	10	91	**170084**	**1348755**
2012	12017	11081	175381	9795	57	152	**208483**	**1557238**
2013	11233	12962	163945	10070	17	116	**198343**	**1755581**
2014	14030	16762	196383	10524	1541	343	**239583**	**1995164**
2015	17741	34123	267439	9786	14507	169	**343765**	**2338929**
2016	28729	24141	354071	17085	20029	204	**444259**	**2783188**
2017	25468	28245	341623	13396	28534	348	**437614**	**3220802**
2018	26906	27758	282997	12220	19977	380	**370238**	**3591040**
2019	10358	2318	198062	4663	6331	216	**221948**	**3812988**
Total	**232786**	**300987**	**3011781**	**160845**	**98386**	**8203**	**3812988**	

Source: Department of Transport Management (DoTM), 2021; Nepal in Data; CIEC Global Database, 2022 [10-12]

.

Table 2. Data table for Variables used in Analysis

Year	Population	PCI (in US$)	GNP (Billion of US $)	GDP per capita (US$)
1989	18445028	200	3.8	191.12078
1990	18905478	200	3.92	191.8789
1991	19405504	200	4.08	202.08061
1992	19938320	190	4	170.58667
1993	20489975	190	3.94	178.62597
1994	21040904	190	4.11	193.27951
1995	21576071	200	4.42	203.98081
1996	22090352	210	4.69	204.68576
1997	22584775	210	4.9	217.78795
1998	23057883	200	4.85	210.61149
1999	23509964	210	5.05	214.10677
2000	23941110	220	5.41	229.49029
2001	24347106	230	5.78	246.72563
2002	24725627	230	5.86	244.72082
2003	25080872	250	6.4	252.40243
2004	25419344	280	7.21	286.15759
2005	25744500	300	8.02	315.80563
2006	26066693	330	8.85	346.94525
2007	26382581	370	9.83	391.38013
2008	26666576	430	11.58	470.45555
2009	26883535	480	12.95	478.17318
2010	27013212	540	14.61	592.4011
2011	27041220	630	17.13	797.81395
2012	26989136	760	20.69	804.14232
2013	26916793	840	23.05	823.35994
2014	26905978	860	23.74	844.85325
2015	27015031	870	24.13	901.74967
2016	27263433	860	24.07	899.52348
2017	27632681	970	27.23	1048.4538
2018	28095714	1110	31.52	1178.5259
2019	28608710	1220	35.2	1194.9574
2020	29136808	1180	34.59	1147.472

Source: Ministry of Finance,2079/80; World Bank, 2023; MacroTrends, 2023; and CEIC Global Database, 2022. [12-14]

4. Data Analysis

In this process, three techniques were used for to compare results: Trend Line Analysis, Econometric Analysis, and Time Series Analysis.

4.1 Trend Line Analysis

Trendline analysis, also known as trend analysis or trend forecasting, is a statistical technique used in various fields, including finance, economics, and data analysis, to identify and analyze patterns or trends in data over time. It involves plotting data points on a graph and fitting a straight line or curve to the data points to make predictions.

This assumes a linear relationship between the country’s Gross National Product (GNP) and the total vehicular population(T) [6]. The data used in this analysis is for the years 1990-2009, i.e., 20 years and the predictions are done for 2014 to 2019. The equation resulting from the regression analysis can be expressed in

$$Log(T) = 1.816 * Log(GNP) + 4.078 \qquad (1)$$

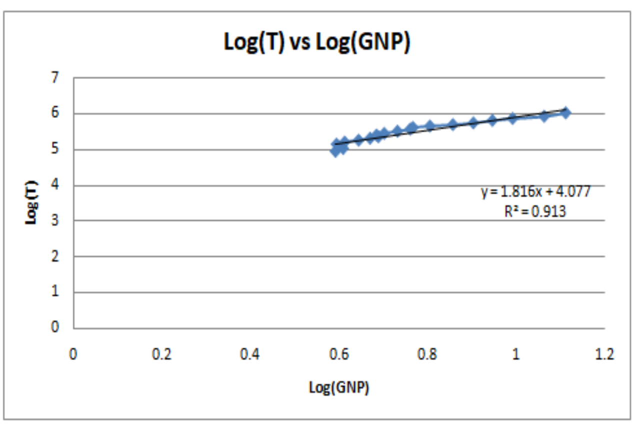

Figure 2. Plot between Log(T) and Log (GNP)

4.2 Econometric Analysis

An econometric model is a statistical framework used to analyze and quantify the relationships among economic variables. It combines economic theory, statistical methods, and real-world data to understand and predict economic phenomena. Traffic growth is often linked to specific financial and demographic factors [15]. These factors include population size, per capita income, and per capita net national product. In this study population and per capita income is predictor variables.

For the analysis, the data was chosen for 21 years (1989-2009) and the estimation is done from 2014 to 2019. The data were subjected to SPSS (Statistical Package for Social Science) for multiple linear regression analysis and the following result came out:

$$Log(T) = 2.861 + 1.0447 * 10^{-7} * P + 4.9074 * 10^{-6} * PCI \qquad (2)$$

Eq. (2) shows that Per Capita Income has more effect on traffic demand than the country's population. This is because higher per capita income often correlates with increased purchasing power, which can lead to higher demand for personal vehicles. As people's incomes rise, they are more likely to afford and desire private transportation options.

Tables 1 and 2 show some important indicators defining the model's fitness—the Durbin-Watson static ranges from 0 to 4. A value towards 0 indicates positive autocorrelation, a value towards 4 indicates a negative correlation and a value near 2 indicates non-autocorrelation.

Table3. Analysis Model Summary

Model Summary[b]										
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.999[a]	0.998	0.998	0.015862929672	0.998	4207.787	2	18	0.000	0.877
a. Predictors: (Constant), PCI, Population										
b. Dependent Variable: LogT										

Table 4. Coefficient and Significance level

Coefficients[a]								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	2.861	0.037		77.983	0.000		
	Population	1.045E-07	0.000	0.865	50.361	0.000	0.402	2.488
	PCI	4.907E-06	0.000	0.166	9.665	0.000	0.402	2.488
a. Dependent Variable: LogT								

Since positive autocorrelation is seen much more in practice than negative autocorrelation, the value in our case is 0.877, which is towards 0, meaning there is approximately positive autocorrelation. Positive autocorrelation indicates that the increase observed in a time interval leads to a proportionate increase in the lagged time interval. The significance value for PCI is 0.000 and for the population, it is 0.000 which is less than 0.05. This means rejecting the null hypothesis and accepting the alternative view that a relationship shows the independent variables do a good job explaining the variation in the dependent variable. Also, we can see the f-change value is 4207.787, ranging from zero to positive infinity depending upon the degree of freedom. A larger value of f-change indicates a more substantial improvement in model fit.

The Histogram(fig.4) is a frequency plot obtained by placing the data in regularly spaced cells and plotting each cell frequency versus the center of the cell. This graph is used to verify that the residuals are normally distributed, as is assumed by the regression model. Due to fewer observations, a perfect average graph has not been obtained here but the plot seems to suggest a normal distribution of residuals. Hence, the error terms can be said to be normally distributed.

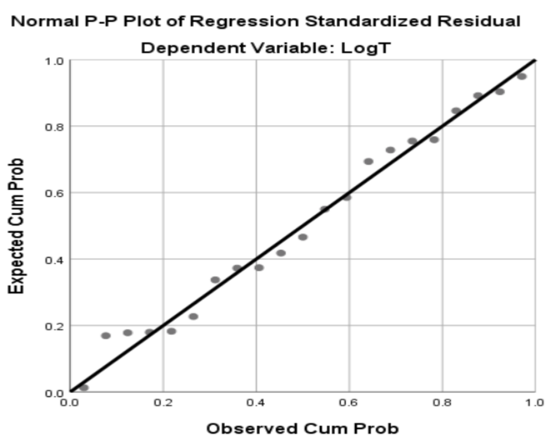

Figure 3. Normal P-P plot regression standardized residual

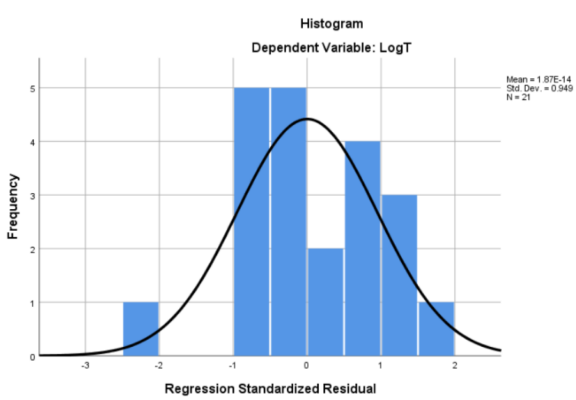

Figure 4. Residual Histogram

The "N" condition of the linear regression model is that the error terms are normally distributed. A normal probability plot of the residuals is a scatter plot with the observed cumulative residual of the normal distribution on the x-axis and the expected cumulative residual of the residuals on the y-axis as shown in fig.3.The relationship between the theoretical percentiles and the sample percentiles is approximately linear. Therefore, the normal probability plot of the residuals suggests that the error terms are indeed normally distributed.

4.3 Time Series Analysis

Time series analysis involves studying and understanding data points collected sequentially over time. This data type often exhibits patterns, trends, seasonality, and other temporal dependencies not present in cross-sectional data.

As Box and Jenkins (1976) suggested, ideally at least 30 observations are required to perform an appropriate Time Series Analysis[16]. The Box and Jenkins methodology has been adopted and analysis has been done using the Auto-Regressive Integrated Moving Average (ARIMA) approach [17]. The analysis has been performed on STATA.ARIMA combines three key components: AutoRegressive (AR), Integrated (I), and Moving Average (MA). Each component addresses a specific aspect of the time series data. Combining these components, the ARIMA model is expressed as ARIMA(p, d, q), where:

- 'p' represents the order of the AutoRegressive component (number of past values to consider).
- 'd' represents the number of differencing operations applied to achieve stationarity.
- 'q' represents the order of the Moving Average component (number of past residuals to consider).

The analysis has been performed on STATA

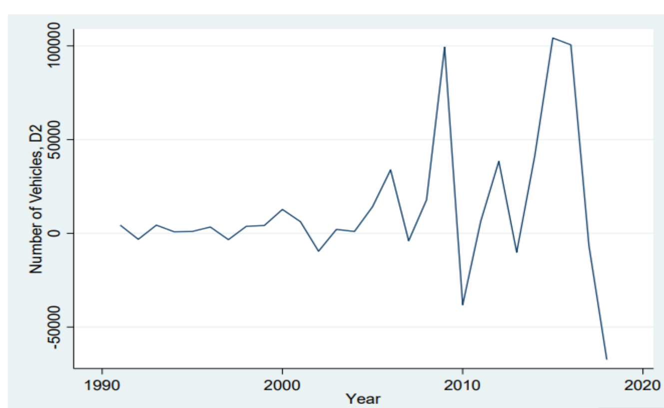

Figure 5. Stationary Vehicular Data (1989-2018)

```
. dfuller d2.VehicularPopulation

Dickey-Fuller test for unit root          Number of obs  = 27
Variable: D2.VehicularPo~n                Number of lags =  0

H0: Random walk without drift, d = 0

                                       Dickey-Fuller
                   Test        ——————— critical value ———————
                 statistic        1%           5%          10%
——————————————————————————————————————————————————————————————
 Z(t)              -3.657      -3.736       -2.994       -2.628
——————————————————————————————————————————————————————————————
MacKinnon approximate p-value for Z(t) = 0.0048.
```

Figure 6. Dickey-Fuller Test of Stationary Data

Time series must be stationary, an assumption confirmed by the Dickey-Fuller test and data trend. Figure 5 shows the mean and covariance of the series do not depend on time and the p-value is 0.0048, which is less than 0.05 as shown in Figure 6. Hence, the series becomes stationary after two differencing, which gives the order of '*d*' as 2.

Autocorrelation (ACF) measures the correlation between a time series and its lagged values. In contrast, partial autocorrelation (PACF) measures the correlation between a time series and its lagged values, excluding the influence of intermediate lags. PACF and ACF plots help determine the order of the ARIMA model's AR (AutoRegressive, p) and MA (Moving Average, q) components.

In a multiple regression model, the variable of interest is forecasted by constructing a linear combination of predictors. In an autoregression model, the variable of interest is denoted by constructing a linear combination of past values of the variable. The term autoregression indicates a regression of the variable against itself.

In the partial autocorrelation graph, five lags exceed the confidence band. This means the order of the autoregressive component is five (the order of *p* is 5). Similarly, In the autocorrelation graph, no lags exceed the confidence band. This means the order of moving average component is zero (the order of *q* is 0). Hence, the order of (p,d,q) is (5,2,0), which gives ten possible ARIMA models by substituting the value of p equals 5 to 1 and d equals 2 to 1.

Ten diverse models have been developed, and to select the best-fit model, there are a few model selection criteria.

- Significance of ARIMA components
- Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) should be minimum
- Maximum Likelihood: A bigger value is better
- Sigma Square: Estimate of the error variance (Smaller value is better)
- Least RMSE(Root Mean Square Error)
- White Noise closer to 1

Table 5. Diagnostic Checking and Selection of Model

Model	Criteria				P(White Noise)	RMSE
	SigmaSQ	Log Likelihood	Akaike	Bayeseian		
ARIMA(5,2,0)	**32418.83**	**-330.9089**	675.8177	685.1431	0.9867	**32474**
ARIMA(4,2,0)	32443.15	-330.938	673.8761	681.8693	0.9833	32488
ARIMA(3,2,0)	33800.87	-331.8741	673.7481	680.4091	0.9562	33819
ARIMA(2,2,0)	33946.33	-331.9866	**671.9731**	677.302	0.966	33969
ARIMA(1,2,0)	36029.51	-333.5146	673.0291	**677.0258**	0.8368	36011
ARIMA(5,1,0)	34528.7	-345.807	705.6139	715.185	**0.9999**	41653
ARIMA(4,1,0)	34994.36	-346.1599	704.3198	712.5236	0.9998	42329
ARIMA(3,1,0)	35861.98	-346.8104	703.6207	710.4572	0.9997	43938
ARIMA(2,1,0)	36411.97	-347.1272	702.2544	707.7236	0.9988	42882
ARIMA(1,1,0)	38607.55	-348.882	703.7641	707.866	0.9971	47885

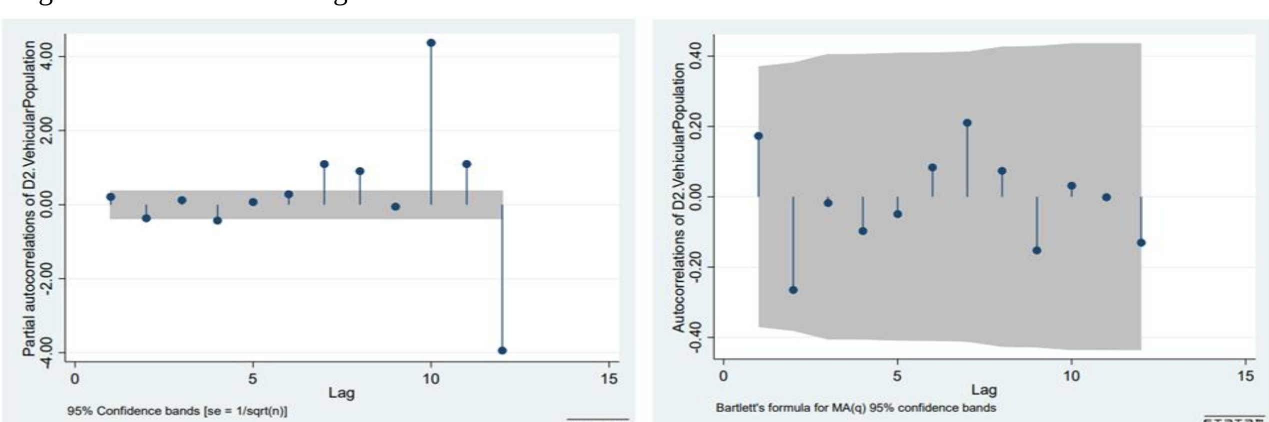

Figure 7. PACF and ACF of Vehicular Data

From Table 3, checking all selection criteria, the model ARIMA(5,2,0) has minimum sigma square, maximum log-likelihood rule, and lowest RMSE, which meets the maximum number of selection crite-

ria(Box, G., & Jenkins, G. (1976). So, this model seems to be the best fit among other forecasting models.
There are two primary approaches to forecasting using an ARIMA model: static and dynamic. In static forecasting, also referred to as one-step ahead forecasting, predictions are made solely for the immediate subsequent period. This entails forecasting one time step ahead using the model trained on historical data. On the other hand, dynamic forecasting involves making predictions for a longer time horizon. Dynamic forecasting is more intricate but better suited to capturing extended trends and patterns, whereas static forecasting offers simpler, immediate predictions. In this case, dynamic forecasting is used for a longer time horizon.

Table 6. ARIMA (5,2,0)

ARIMA regression

Sample: 1991 thru 2018 — Number of obs = 28; Wald chi2(5) = 9.39; Prob > chi2 = 0.0946

Log likelihood = -330.9089

D2. VehicularPopulation	Coefficient	OPG std. err.	z	P>|z|	[95% conf.	interval]
VehicularPopulation						
_cons	12966.08	9608.532	1.35	0.177	-5866.298	31798.46
ARMA						
ar						
L1.	.3516046	.2364234	1.49	0.137	-.1117769	.814986
L2.	-.4238265	.2414627	-1.76	0.079	-.8970848	.0494318
L3.	.1378884	.3614413	0.38	0.703	-.5705236	.8463003
L4.	-.3289175	.510948	-0.64	0.520	-1.330357	.6725221
L5.	.0606593	.4005781	0.15	0.880	-.7244593	.8457779
/sigma	32418.83	4575.276	7.09	0.000	23451.45	41386.2

5. Results and Discussion

The analysis results compare three methods adopted in this study to predict the growth of Nepal's vehicular population and forecast the vehicular population for the target years 2030, 2040, and 2050 AD using effective forecasting techniques. The input data used for analysis is from 1990 to 2009 (20 years), and a prediction was made for the years 2014 to 2019 for the comparative study, as shown in Table 5.
These error percentages provide a measure of the deviation between the predicted values and the actual values. A negative error indicates an underestimation, while a positive error indicates an overestimation. The error's magnitude suggests the deviation's extent, with larger values representing larger discrepancies. In trend line analysis, The graphical plot of Log(T) versus (Log(GNP) reveals a linear relationship between vehicular population and GNP, which is represented as a straight line.

Table 7. Results Obtained from Different Methods for the year 2014 to 2019

Methods	Year	Actual	Predicted	Error(%)
Trend line Analysis	2014	1995164	3765936	88.75
	2015	2338929	3879039	65.85
	2016	2783188	3861540	38.75
	2017	3220802	4831105	50.00
	2018	3591040	6301335	75.47
	2019	3812988	7700545	101.96
Econometric Analysis	2014	1995164	1696471	-14.97
	2015	2338929	1767762	-24.42
	2016	2783188	1848799	-33.57
	2017	3220802	2381201	-26.07
	2018	3591040	3280606	-8.64
	2019	3812988	4373990	14.71
Time Series Analysis	2014	1995164	1969131	-1.30
	2015	2338929	2269999	-2.95
	2016	2783188	2703802	-2.85
	2017	3220802	3245570	0.77
	2018	3591040	3629266	1.06
	2019	3812988	3938089	3.28

This suggests that the vehicular population has a corresponding linear growth as GNP increases. However, a significant discrepancy emerges between the actual and expected values when this model is applied to predict vehicular population from 2014 to 2019. To provide context, the Department of Transport Management data indicates that half of Nepal's total vehicles were registered after 2014. This rapid and exponential increase in vehicular population after 2014 is not captured by the initially assumed linear model, leading to inaccurate predictions for 2014 to 2019. However, the econometric model shows a satisfactory level of accuracy where the error is below 33%. The econometric model chooses population and per capita income as independent variables. It shows that the growth of Nepal's vehicular population depends on the country's people and their income level.
The results of the ARIMA (5,2,0) model applied to forecast Nepal's vehicular population are promising, exhibiting a tight error range of -1 % to +3%. This indicates high accuracy and precision in predicting future vehicular population trends. Such reliable predictions within this narrow percentage range can aid policy formulation, urban planning, and informed decision-making in Nepal's transportation sector.
Until 2018, the predicted values generated by the model exhibited a satisfactory level of consistency in the prediction of vehicular population from all three methods. However, a significant prediction accuracy shift became evident in 2019. This

sudden increase in prediction errors in 2019 can be attributed, at least in part, to the outbreak of the COVID-19 pandemic. The pandemic introduced unprecedented disruptions to global economies, supply chains, and socio-economic aspects. The resultant lockdowns, travel restrictions, and shifts in consumer behavior led to a decrement in vehicle registration in this particular year. Overall, econometric analysis shows a satisfactory level of prediction as compared to trend line analysis. This clarifies that the economic/demographic indicator population and per capita income depend more on the prediction of vehicular growth than using a single indicator GNP. Encompassing trend line analysis, econometric analysis, and time series analysis to forecast Nepal's vehicular population, a clear pattern emerges, indicating the superior accuracy of the time series analysis, specifically the ARIMA(5,2,0) model. The error margin of up to 3% achieved by the ARIMA model distinctly outshines the alternatives. This outcome underscores the robustness of the ARIMA model's predictive capability, making it the preferred method for accurate vehicular population forecasting.

The prediction of Nepal's vehicular population reveals anticipated figures of 8,914,793 for 2030AD,1,482,842,6 for 2040AD, and 2,203,801,2 for 2050AD. These projections offer critical insights for planners and policymakers as invaluable tools for strategic decision-making and urban development. By accurately estimating the future vehicular population, authorities can anticipate the associated demands on transportation infrastructure, energy resources, and environmental impacts. These figures enable planners to allocate resources efficiently, design sustainable transportation systems, and implement policies that cater to the evolving needs of a growing vehicular population.

6. Conclusion

The study's results have shown that the conventional Trend Line Analysis often leads to a significant overestimation of future traffic volumes. It is crucial to improve the accuracy of these estimations to ensure the efficient allocation of limited resources such as land, labor, and funds, especially in developing nations.

This research suggests that adopting more rational, reliable, and advanced analytical methods, such as Econometric Analysis and Time Series Analysis, can yield more credible forecasts. Econometric modeling proves to be superior in capturing the impact of two critical factors: the growth in the overall number of users and their purchasing power to access new transportation developments. This stands in contrast to the traditional Trend Line Analysis, which relies solely on a country's total productivity level (e.g., Gross National Product) for estimating prospective users.

Notably, Time Series Analysis has a well-established track record in short-term forecasting within finance and economics and warrants exploration in transportation engineering. Further investigation can determine its accuracy within different time frames. Research findings indicate that Time Series Analysis is particularly practical for short-term forecasts. When comparing our analyses, Trend Line, Econometric, and Time Series approaches, Time Series modeling displayed significantly lower errors.

The potential of Time Series Analysis for long-term forecasting, given advancements in data availability and technology, is promising. It aligns with the findings of other researchers and could substantially contribute to more precise traffic forecasting in the future.

References

[1] Dhingra, S. L., Mujumdar, P. P., and Gajjar, R. H. Application of time series techniques for forecasting truck traffic attracted by the Bombay metropolitan region, *Journal of Advanced Transportation*, 27(3) (1993) 227-249.

[2] Matas, A., Raymond, J. L., Gonzalez-Savignat, M., and Ruiz, A. Demand Forecasting in the Evaluation of Projects, *Economic Evaluation of Transportation Projects*, (2009) 1-31.

[3] Sinha, A., Puri, R., Balyan, U., Gupta, R., and Verma, A. Sustainable Time Series Model for Vehicular Traffic Trends Prediction in Metropolitan Network Jaypee Institute of Information and Technology. Department of Computer Science Engineering (2020)

[4] Devkota, R., Sitoula, S., Karki, S., and Shahi, P. B. Urban Arterial Mid-Block Traffic Forecasting: Case Study of Kathmandu Ring Road, *IOSR Journal of Computer Engineering,* 23(6) (2021) 01-09.

[5] Oswald, K., Scherer, W. T., and Smith, B. L. Traffic Flow Forecasting Using Approximate Nearest Neighbor Nonparametric Regression, *Thornton Hall: Center for Transportation Studies at the University of Virginia*, (2000).

[6] Jha, K., Sinha, N., Arkatkar, S. S., and Sarkar, A. K. Modeling Growth Trend and Forecasting

Techniques for Vehicular Population in India, *International Journal for Traffic and Transport Engineering (IJTTE)*, 3(2) (2013) 139-158.

[7] Bastoal, D. P. Air Passenger Demand Model (APDM): Econometric Model For Forecasting Demand In Passenger Air Transports In Nepal, *International Journal of Advanced Research and Publications (IJARP)*, 1(4) (2017).

[8] Sharma, R. D., Jain, S., and Singh, K. Growth rate of Motor Vehicles in India - Impact of Demographic and Economic Development, *Journal of Economic and Social Studies*, 1(2) (2011).

[9] Kumar, S. V., and Vanajakshi, L. Short-term traffic flow prediction using seasonal ARIMA model with limited input data, *Eur. Transp. Res. Rev.*, 7(21) (2015)

[10] Transport, M. o. (n.d.). Department of Transport Management. see https://www.dotm.gov.np/

[11] *NEPAL IN DATA*. see https://nepalindata.com/: https://nepalindata.com/data/?request&secid=infrastructure-communication-technology,subsecid=land-transportation,indid=number-of-registered-vehicles-by-type

[12] Innocenti, A. CEIC, Retrieved from CEIC Database (2020), see https://www.ceicdata.com/

[13] Economic Survey 79/80. In G. o. Nepal. Kathmandu: Ministry of Finance (2023)

[14] *Macrotrends*. (2010). see https://www.macrotrends.net/: https://www.macrotrends.net/countries/NPL/nepal/gdp-gross-domestic-product

[15] Gujarati, D. N. *Basic Econometrics*. New-York: The McGraw-Hill companies

[16] Box, G., and Jenkins, G. Time Series Analysis: Forecasting and Control. *San Francisco: Holden-Day*. (1976) 570-576.

[17] Pranratz, A. Forecasting with Univariate Box-Jenkins Models: Concepts and Cases. New York. (1983) 234-370: John Wiley & Sons.