Australian Journal of Basic and Applied Sciences, 1(2): 126-136, 2007
ISSN 1991-8178

Comparison Between Regression and Arima Models in Forecasting Traffic Volume

[1]M. Sabry, [1]H. Abd-El-Latif and [2]N. Badra

[1]Department of Public Works, Faculty of Engineering, Ain Shams University, Cairo, Egypt.
[2]Department of Physics and Engineering mathematics, Faculty of Engineering,
Ain Shams University, Cairo, Egypt.

Abstract: This paper investigates the application of two time series forecasting techniques, namely logistic regression and univariate auto regressive integrated moving average (ARIMA), to predict daily traffic volume for the Egyptian intercity roads. The data of three permanent traffic volume stations was selected in this paper. A total of 14-year traffic data for the selected stations (from 1990 to 2003) were used in the development of the two models to predict traffic volume. All of the data sets were best represented by both logistic regression model and an autoregressive integrated moving – average (ARIMA) model. The data set containing average annual, average monthly and average weekly daily traffic volume for years 1990 through 2001 was used to fit a time series model. The resulting models were used to forecast traffic volumes for year 2002 and 2003. The forecasted traffic volumes for the two models were then compared with the actual traffic volumes. According to the analysis, the ARIMA model seems to be the best forecasting method especially for the average monthly and average weekly daily traffic volume.

Key words: Traffic volume, ARIMA model, Traffic volume, Regression model, Logistic function, Forecasting method.

INTRODUCTION

Studies on time-series and forecasting procedures have been developed since the recent 50 years. Forecasting plays a major role in traffic engineering. Considering the amount of time, effort and cost involved in model building, it would be reasonable to expect a high degree of accuracy in the forecasts made. Time series is a set of statistical observations, arranged in chronological order. An observed series theoretically consists of two parts: the series generated by real process and the noise, which is the result of outside disturbances. Elimination of this noise is the main aim of time series analysis. One of the problems created in forecasting is the selection of a suitable model for a specific purpose. Selection of model depends on several factors including the context of forecasts, the availability of historic data, the degree of desirable accuracy, the cost of evaluation and the time period to be forecasted. The main purpose of this research is to analyze the daily traffic volume data collected automatically and to evaluate predictive models and then produce a set of forecasts for the site at which the data was collected. It has become apparent that the degree of accuracy has rarely been achieved with conventional models and investigations into their performance in transport studies have resulted in considerable criticism. Atkins (1977), Blanchard (1983) and Horowitz and Enslie (1978) stated that the amount of data available to the transport planner is limited due to time and financial restrictions. This lack of relevant data could be one of the main contributions to the poor performance of accurate forecasting techniques. Time series analysis has been used in transportation studies to reduce the dependence on conventional models. These simplified models are relatively easy to build and are used when several years of observed information is available. They are particularly suitable for short - term forecasting and accuracy of the results is impressive. Comparative studies of forecasting techniques have been developed by Layton, Defris and Zehnwirth (1986). It is clear from these studies that the Box - Jenkins model (1994) is the method providing best forecasts for the majority of the series tested. Box - Jenkins produced the best forecasts for 74% of the series used in a study by Reid (1971). The cost associated with the Box - Jenkins approach in a given situation is generally greater than many other quantitative methods. Williams and Lester (2003) have studied the theoretical basis for modeling univariate traffic condition data streams as seasonal autoregressive integrated moving average processes.

Corresponding Author: N. Badra, Department of Physics and Engineering mathematics, Faculty of Engineering, Ain Shams University, Cairo, Egypt.

This paper investigates the application of two time-series forecasting techniques: namely logistic regression and univariate ARIMA (Auto regressive integrated moving average) to predict daily traffic volume for three stations, Zazeeq – Ismeilia agricultural road (Station 5), Tanta – Mansoura road (Station 9) and Meet – Ghamr – Aga road (Station 11). Regression analysis and ARIMA are two commonly used statistical time series forecasting techniques. A total of 14 years traffic volume data for the three stations from 1990 to 2003 were used in the development of the two models to predict traffic volume. All of the data sets were calibrated by both logistic regression model and ARIMA model. The data set containing average annual, average monthly and average weekly daily traffic volume on three stations for years 1990 through 2001 is used to fit a time series model. The resulting model is used to forecast traffic volumes for year 2002 through 2003. The forecast traffic volumes for the two models are then compared with the actual traffic volumes.

MATERIALS AND METHODS

This research uses both regression and ARIMA approaches in forecasting the road traffic volume. The following describes the theoretical background of these two approaches.

Regression:

Regression analysis is used to utilize a relation between a variable of interest, called the dependent or response variable and one or more independent variables. Regression analysis is used to predict the response variable from knowledge of the independent variables. At other times, regression analysis is utilized primarily for examining the nature of the relationship between the independent variables and the response variable (Neter, Wasserman and Whitmore (1988). Regression models containing one independent variable are called simple regression models. While, regression models containing two or more independent variables are called multiple regression models.

The general form of the multiple regression model can be set in the following form:

$$y = f(x_1, x_2, \ldots, x_n) \quad (1)$$

where,

y is the response variable and

$x_1, x_2, \ldots, x_n$ are the independent variables.

This form can be linear or nonlinear functions of the independent variables.
The simple nonlinear regression model can be set in the following form:

$$y = A_o + A_1 x + A_2 x^2 + A_3 x^3 + \ldots\ldots \quad (2)$$

where,

$A_o, A_1, A_2, A_3 \ldots$ are the regression model parameters.

These model parameters are estimated using the least square error approach. Goodness of fit measures available for the least squares regression analysis such as coefficient of determination (R^2) and t - test should be also estimated.

ARIMA:

ARIMA models are flexible and widely used in time series analysis. ARIMA combines three processes: autoregressive (AR), differencing to strip off the integration (I) of the series and moving averages (MA). Each of the three process types has its own characteristic way of responding to a random disturbance. The resulting general linear model is in the form (Pankratz 1983):

$$Z_t = \varphi 1\, Z_{t-1} + \varphi_2 Z_{t-2} + \ldots + \varphi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q} \quad (3\text{-a})$$

where,

φ_i, i=1,2,…,p are the autoregressive parameters,
θ_j, j=1,2,…,q are the moving average parameters,
Z_t is obtained by differencing the original time series d times, (equal $V_t - V_{t-1}$ if d=1),

a_t is the white noise component at t,
p is the order of autoregressive parameters,
q is the order of moving average parameters and
d is the order of differencing.

Identification is a critical step in building an ARIMA (p, d, q) model.

To estimate the parameters φ_i and θ_j for fixed p and q, perform the linear multiple regression indicated by the model:

$$Z^{\backslash} = Z\,\Phi - T\,\theta + \mu \qquad (3\text{-b})$$

where,
$Z^{\backslash}$ is the original time – series realization (Z_t, Z_{t+1}, …),
Z is the vector of lagged regressors, (Z_{t-1}, Z_{t-2}, …, Z_{t-p}),
Φ is the vector of unknown lagged parameters, (φ_1, φ_2, …, φ_p) and
θ is the vector of unknown lagged parameters, (θ_1, θ_2, …, θ_q).

The variables φf_i and θ_j are called lagged variables.

Equation (3-a) can be written as:

$$(1 - \varphi_1 B - \varphi_2 B^2 - \ldots - \varphi_p B^p)\, Z_t = (1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q)\, a_t \qquad (4)$$

$$\varphi(B)\,(1 - B)^d\, V_t = \theta(B)\, a_t \qquad (5)$$

$$V_t = \varphi^{-1}(B)\,(1 - B)^{-d}\,\theta(B)\, a_t \qquad (6)$$

where,
B is the backward shift operator, i. e., $Z_{t-1} = B\,Z_t$.

Identification is a critical step in building an ARIMA (p, d, q) (sp, sd, sq)$_L$ model.

where,
p is the AR order which indicates the number of parameters of φ,
d is the number of times the data series must be differenced to induce a stationary series,
q is the MA order which indicates the number of parameters of θ,
sp is the seasonal AR order which indicates the number of parameters of φ,
sq is the seasonal MA order which indicates the number of parameters of θ and
sd is the number of times the data series needs to be seasonally differenced to induce a seasonally stationary series.

Identification methods are rough procedures applied to a set of data to indicate the kind of representational model that is worthy of further investigation. The aim of this method is to obtain some idea of the values of p, d and q needed in the general linear ARIMA model and to obtain best estimates for the parameters. Parameter estimation algorithm tries all combinations of parameters, which are limited to an integer lying between zero and two. The combination with the least fitting error will be searched. The range limitations of the parameters are set to restrict the search to a reasonable scope. Parameters greater than two are rarely used in practice. This principle reflects not only a view of nature but also a view of the relationship between a time series model and nature. In almost all cases, a social science time series process can be modeled as a probabilistic function of a few past inputs and time series observations.

Data Collection:

Daily traffic volume statistics of three stations (Station 5, Station 9 and Station 11) from a period of January 1990 to December 2003 were considered for models' calibration and validation in this paper. The average annual, monthly and weekly daily traffic volumes were calculated for the data. The collected data were divided into two sets, calibration data and testing data, in order to testify the performance of the two suggested forecasting methods. To achieve a more reliable and accurate result, a long period (from January 1990 to

December 2001) was selected as the calibrating period, while, the period from January 2002 to December 2003 was considered as the testing period.

RESULTS AND DISCUSSIONS

Regression Approach:

Several forms of regression models have been tried for calibration in order to get the best one. For instance, logistic, simple linear and quadratic models were tried. The logistic form was found the best results in terms of R^2 and t – test values. In addition, Benjamin recommended using this model form, due to permitting accurate models when there are temporary fluctuations in traffic volume (Benjamin 1986).

The logistic model can be set in the following form (Benjamin 1986):

$$\ln \left(\frac{V - V_{min}}{V_{max} + V_{min} - V}\right) = A_0 + A_1 t + A_2 t^2 + A_3 t^3 + A_4 t^4 + \ldots\ldots \quad (7)$$

where,

V is the current average daily traffic volume,
V_{min} is the minimum traffic volume,
V_{max} is the maximum traffic volume,
t is the time (year or month or week number) and

A_o, A_1, A_2 and A_3… are the coefficients to be determined by the logistic regression model.

The parameters A_o, A_1, A_2, A_3 and A_4 … are calculated by using the least squares regression analysis with ln $(V - V_{min} / V_{min} + V_{max} - V)$ as the dependent variable and t, t^2, t^3 and t^4 … as the independent variables.

The statistical package SPSS was used for the analysis work of this paper. The model parameters were estimated using the least square error approach. Different logistic models were also calibrated taking into consideration the power of the time value (e. g. first order, second order, etc) and the best model in terms of the coefficient of determination R^2 was also chosen.

The calibrating period was used to estimate the parameters using the logistic regression technique. The values of the parameters are shown in Tables 1, 2 and 3 for the three cases of average annual, monthly and weekly daily traffic volume, respectively. (Appendix A)

Table 1: Results of logistic regression analysis for average annual daily traffic.

Parameters		Station 5	Station 9	Station 11
Asymptotes:	V_{min}	5785	9373	9530
	V_{max}	9198	18861	21925
Coefficients:	A_0	-4.125	-4.382	-4.923
	A_1	0.26	0.448	0.375
	A_2	0.003692	-0.00583	0.004061
R^2		0.889	0.975	0.961

Table 2: Results of logistic regression analysis for average monthly daily traffic.

Parameters		Station 5	Station 9	Station 11
Asymptotes:	V_{min}	4204	8476	7282
	V_{max}	11049	22235	27264
Coefficients:	A_0	-2.055	-3.462	-2.885
	A_1	0.01306	0.02505	0.0192
R^2		0.733	0.783	0.777

Table 3: Results of logistic regression analysis for average weekly daily traffic.

Parameters		Station 5	Station 9	Station 11
Asymptotes:	V_{min}	4006	7667	6804
	V_{max}	11146	22447	27439
Coefficients:	A_0	-1.90	-2.810	-2.823
	A_1	0.002728	0.004752	0.00472
R^2		0.599	0.813	0.809

Results of observed and estimated values of average annual daily traffic volume in the period from 1990 to 2003 at stations 5, 9 and 11 are shown in Figures 1, 2 and 3, respectively.

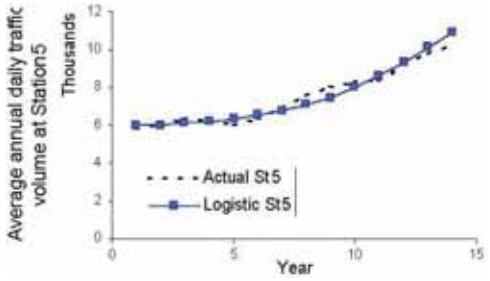

Fig. 1: Observed and forecasted traffic volume for 2002 and 2003 at Station5.

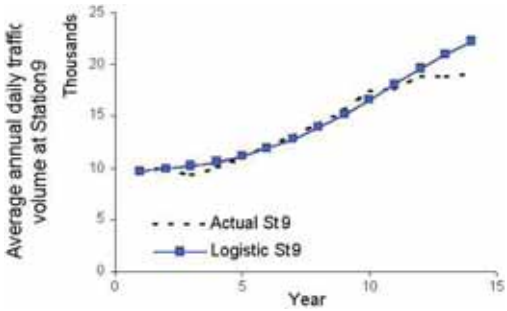

Fig. 2: Observed and forecasted traffic volume for 2002 and 2003 at Station9.

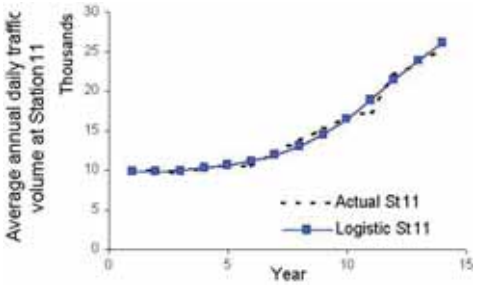

Fig. 3: Observed and forecasting traffic volume for 2002 and 2003 at Station11.

ARIMA:

The calibrating period was used to estimate the parameters using the described searching algorithm. The values of the calibrated parameters are shown in Tables 4, 5 and 6 for the average annual, monthly and weekly daily traffic, respectively. (Appendix B)

Table 4: Results of parameter estimates for average annual daily traffic volume using ARIMA

	ARIMA (p,d,q)	Parameters	μ	Lag 1	Lag 2	Lag 3	R^2
Station 5	(2, 0, 0)	AR	-814.559	1.037	-----	0.133	0.916
Station 9	(3, 0, 0)	AR	1027.586	0.918	0.221	-0.14	0.975
Station 11	(1, 1, 0)	AR	912.823	0.344	-----	-----	0.942

Table 5: Results of parameter estimates for average monthly daily traffic volume using ARIMA.

	ARIMA (p,d,q) $(sp,sd,sq)_s$	Parameters	μ	Lag 1	Lag 2	Lag 2	R^2
Station 5	$(2, 0, 0)(2, 0, 0)_{12}$	Non-Seasonal AR	-109.97	0.374	-----	0.143	0.768
		Seasonal AR		0.438	0.07915	-----	
Station 9	$(3, 0, 0)(2, 0, 0)_{12}$	Non-Seasonal AR	-55.302	0.638	0.08375	-0.026	0.959
		Seasonal AR		0.793	-0.462	-----	
Station 11	$(2, 1, 0)(1, 1, 0)_{12}$	Non-Seasonal AR	54.387	0.113	0.101	-----	0.938
		Seasonal AR		0.306	-----	-----	

Table 6: Results of parameter estimates for average weekly daily traffic volume using ARIMA.

	ARIMA (p,d,q) (sp,sd,sq) $_S$	Parameters	μ	Lag 1	Lag 2	Lag 4	R^2
Station 5	(2, 0, 0) (2, 0, 0) $_{50}$	Non-Seasonal AR	81.299	0.85	-----	- 0.02	0.885
		Seasonal AR		0.102	0.062	-----	
Station 9	(2, 0, 0) (1, 0, 0) $_{52}$	Non-Seasonal AR	38.155	0.704	-----	0.122	0.958
		Seasonal AR		0.184	-----	-----	
Station 11	(1, 1, 0) (1, 1, 0) $_{52}$	Non-Seasonal AR	20.056	-----	0.067	0.12	0.971
		Seasonal AR		0.125	-----	-----	

Comparison Between the Two Forecasting Methods:

The predictions of traffic volume in the testing period are done using the two forecasting methods. Figures 4 to 9 illustrate the comparison between the actual and forecasted values for both monthly and weekly daily traffic volume in the three stations, respectively. To compare the prediction performance of the two approaches for the period from January 2002 to December 2003, Standard Error Deviation (STDEV) was calculated.

The standard deviation error (STDEV) is defined as follows (Gaynor and R. Kirkpartrick 1994):

$$STDEV = \sqrt{\frac{\sum e_t^2}{n}} \tag{8}$$

Results with minimal errors, seems to give the best performance of the two methods.

where,

n is the number of testing points and

e_t ($V_{t\text{ - actual}} - V_{t\text{ - forecast}}$) is the error in forecasting traffic volume.

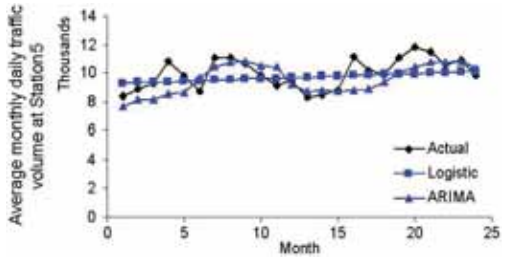

Fig. 4: Observed and forecasted monthly traffic volume for 2002 and 2003 at Station5.

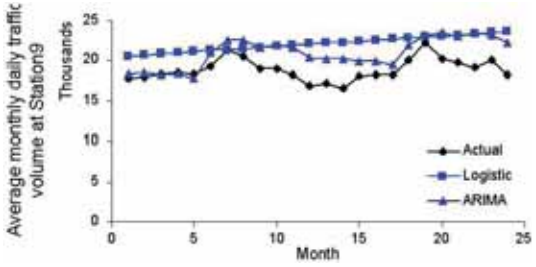

Fig. 5: Observed and forecasted monthly traffic volume for 2002 and 2003 at Station9.

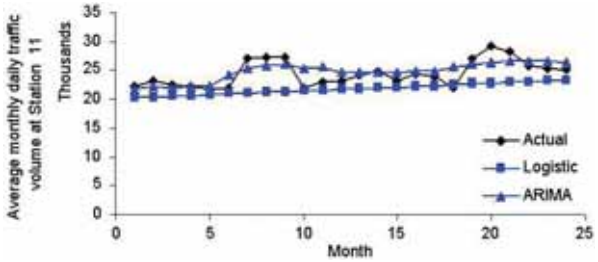

Fig. 6: Observed and forecasted monthly traffic volume for 2002 and 2003 at Station11.

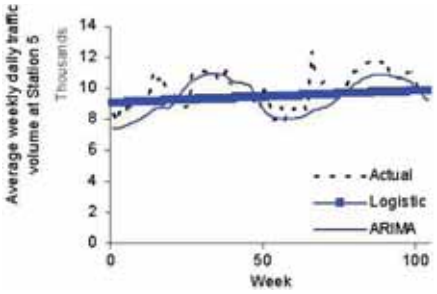

Fig. 7: Observed and forecasted weekly traffic volume for 2002 and 2003 at Station 5.

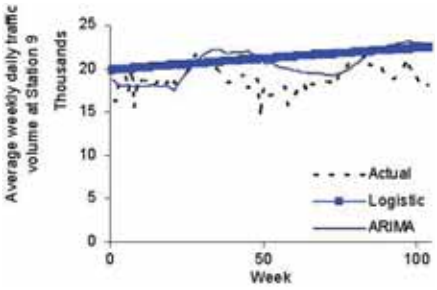

Fig. 8: Observed and forecasted weekly traffic volume for 2002 and 2003 at Station 9.

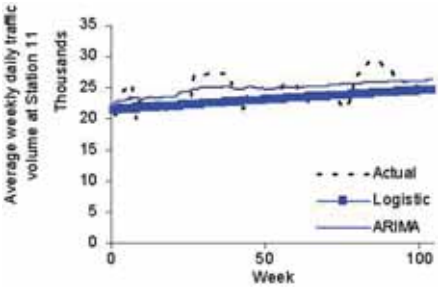

Fig. 9: Observed and forecasted weekly traffic volume for 2002 and 2003 at Station 11.

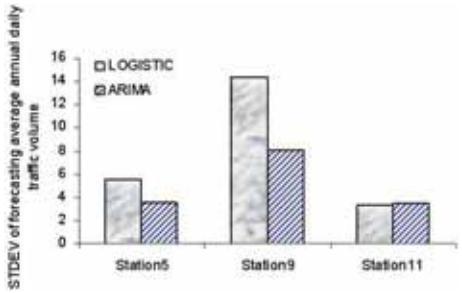

Fig. 10: STDEV of forecasting average annual daily traffic volume.

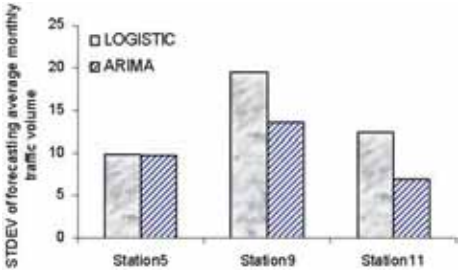

Fig. 11: STDEV of forecasting average monthly daily traffic volume.

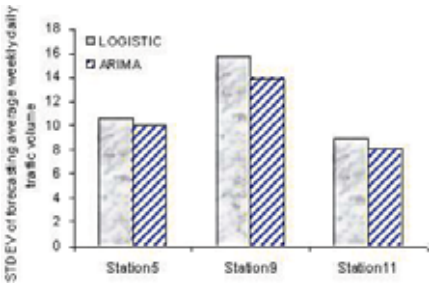

Fig. 12: STDEV of forecasting average weekly daily traffic volume.

Results of the standard deviation error (STDEV) in the testing period using the two forecasting methods for average annual, monthly and weekly daily traffic volumes are shown in Figures 10, 11 and 12, respectively. It can be concluded from the figures that the ARIMA standard deviation is generally less than that of the logistic regression for estimating either annual, monthly or weekly daily traffic volume.

Conclusion:

This paper investigates the application of two time-series forecasting techniques: namely regression and univariate ARIMA to predict traffic volume at three Egyptian intercity roads, Zazeeq – Ismeilia agricultural road (Station 5), Tanta – Mansoura road (Station 9) and Meet – Ghamr – Aga road (Station 11). The period of collected data is from January 1990 to December 2003. The average annual, monthly and weekly daily traffic volumes are calculated for the data. The collected data were divided into two sets, calibrating and testing data, in order to testify the performance of the two suggested forecasting methods. The period from January 1990 to December 2001 was the calibrating period, while; the period from January 2002 to December 2003 was the testing period. Both logistic regression model and an autoregressive integrated moving – average (ARIMA) model were calibrated using the data. The statistical package SPSS was used for the analysis work of this paper. The forecasted traffic volumes for the two models were then compared with the actual traffic volumes. To compare between the predictions performances of the two approaches, the standard error deviation (STDEV) was calculated. According to the analysis, the ARIMA model seems to be the best forecasting method for traffic volume.

APPENDIX A (LOGISTIC MODELS)

The logistic models that were used in forecasting average annual, monthly and weekly daily traffic volume are as follows:

$$V_{forecasting} = (V_{min} + (V_{min} + V_{max})\ {}_eA_0 + A_1\ t + A_2\ t^2) / (1 + {}_eA_0 + A_1\ t + A_2\ t^2)$$

Taking into consideration the calibrated parameters shown in Tables 1, 2 and 3: the above formula can be rewritten as follows:

Station 5:

Average annual daily traffic volume:

$$V(t) = \frac{5785 + 14983\, e^{-4.125 + 0.26\, t\ + 0.003492\, t^2}}{1 + e^{-4.125 + 0.26\, t\ + 0.003492\, t^2}}$$ (t = year - 1989)

Average monthly daily traffic volume:

$$V(t) = \frac{4204 + 15253\, e^{-2.055 + 0.01304\, t}}{1 + e^{-2.055 + 0.01304\, t}}$$ (t = (year – 1990)*12 + month order)

Average weekly daily traffic volume:

$$V(t) = \frac{4006 + 15152\, e^{-1.9 + 0.00272 t}}{1 + e^{-1.9 + 0.00272 t}}$$ (t = 626 + (year – 2002)*52 + week order)

Station 9:

Average annual daily traffic volume:

$$V(t) = \frac{9373 + 28234\, e^{-4.382 + 0.448 t - 0.00583 t^2}}{1 + e^{-4.382 + 0.448 t - 0.00583 t^2}}$$ (t = year - 1989)

Average monthly daily traffic volume:

$$V(t) = \frac{8476 + 30711\, e^{-3.442 + 0.0250 t}}{1 + e^{-3.442 + 0.0250 t}}$$ (t = (year – 1990)*12 + month order)

Average weekly daily traffic volume:

$$V(t) = \frac{7667 + 30114\, e^{-2.81 + 0.00475 t}}{1 + e^{-2.81 + 0.00475 t}}$$ (t = 626 + (year – 2002)*52 + week order)

Station 11:

Average annual daily traffic volume:

$$V(t) = \frac{9530 + 31455\, e^{-4.923 + 0.375 t + 0.00404 t^2}}{1 + e^{-4.923 + 0.375 t + 0.00404 t^2}}$$ (t = year - 1989)

Average monthly daily traffic volume:

$$V(t) = \frac{7282 + 34547\, e^{-2.885 + 0.0192 t}}{1 + e^{-2.885 + 0.0192 t}}$$ (t = (year – 1990)*12 + month order)

Average weekly daily traffic volume:

$$V(t) = \frac{6804 + 34243\, e^{-2.823 + 0.00472 t}}{1 + e^{-2.823 + 0.00472 t}}$$ (t = 626 + (year – 2002)*52 + week order)

It should be noted that the forecasted traffic volumes V (t) that are estimated by these logistic models should never exceed $V_{min} + V_{max}$.

APPENDIX B (ARIMA MODELS)

Taking into consideration the calibrated parameters shown in Tables 4, 5 and 6: the ARIMA models can be rewritten as follows:

Station 5:

Average annual daily traffic volume: ARIMA (2, 0, 0)

$$V_t = -814.559 + 1.037\, V_{t-1} + 0.133\, V_{t-3}$$

Average monthly daily traffic volume: ARIMA (2, 0, 0) (2, 0, 0) $_{12}$

$$V_t = - 109.979 + 0.374\ V_{t-1} + 0.143\ V_{t-3} + 0.438\ V_{t-12} + 0.07915\ V_{t-13}$$

Average weekly daily traffic volume: ARIMA (2, 0, 0) (2, 0, 0) $_{50}$

$$V_t = 81.299 + 0.854\ V_{t-1} - 0.0228\ V_{t-4} + 0.102\ V_{t-50} + 0.062\ V_{t-52}$$

Station 9:

Average annual daily traffic volume: ARIMA (3, 0, 0)

$$V_t = 1027.586 + 0.918\ V_{t-1} + 0.221\ V_{t-2} - 0.14\ V_{t-3}$$

Average monthly daily traffic volume: ARIMA (3, 0, 0) (2, 0, 0) $_{12}$

$$V_t = - 55.302 + 0.638\ V_{t-1} + 0.08375\ V_{t-2} - 0.0268\ V_{t-3} + 0.793\ V_{t-12} - 0.462\ V_{t-13}$$

Average weekly daily traffic volume: ARIMA (2, 0, 0) (1, 0, 0) $_{52}$

$$V_t = 38.155 + 0.704\ V_{t-1} + 0.122\ V_{t-4} + 0.184\ V_{t-52}$$

Station 11:

Average annual daily traffic volume: ARIMA (1, 1, 0)

$$V_t = V_{t-1} + \Delta V_t,\ \Delta V(t) = V(t) - V(t-1)$$

$$\Delta V_t = 912.823 + 0.344\ \Delta V_{t-1}$$

Average monthly daily traffic volume: ARIMA (2, 1, 0) (1, 1, 0) $_{12}$

$$V_t = V_{t-1} + \Delta V_t,\ \Delta V(t) = V(t) - V(t-1)$$

$$\Delta V_t = 54.387 + 0.113\ \Delta V_{t-1} + 0.101\ \Delta V_{t-2} + 0.306\ \Delta V_{t-12}$$

Average weekly daily traffic volume: ARIMA (1, 1, 0) (1, 1, 0) $_{52}$

$$V_t = V_{t-1} + \Delta V_t,\ \Delta V(t) = V(t) - V(t-1)$$

$$\Delta V_t = 20.056 + 0.06751\ \Delta V_{t-3} + 0.125\ \Delta V_{t-52}$$

REFERENCES

Atkins, S., 1977. “Transportation Planning: Is there a Road Ahead ? ”, Traffic Engineering and Control, Vol. 18, pp: 58-62.

Blanchard, R., 1983. “Transportation and the New Planning”, Traffic Engineering and Control, Vol. 24, pp: 306-310.

Benjamin, J., 1986. “A Time – Series Forecast of Average Daily Traffic Volume”, Transportation Research, Vol. 20, 1986, pp: 51-60.

Box, G., G. Jenkins and G. Reinsel, 1994. “Time Series Analysis: Forecasting and Control”, Prentice Hall,

Chen, C. and S. Grant, 2001. “Use of Sequential Learning for Short Term Traffic Flow Forecasting”, Transportation Research, Vol. 9, pp: 319-336.

Davis, C., 2002. “Statistical Methods for the Analysis of Repeated Measurements”, New York.

Gaynor, P. and R. Kirkpatrick, 1994. “Introduction to Time Series Modeling and Forecasting in Business and Economics”, McGraw-Hill, Inc., New York.

Horowitz, J. and R. Enslie, 1978. “Comparison of Measured and Forecast Traffic Volume on Urban Interstate Highways”, Transportation Research, Vol. 12, pp: 29-32.

Kim, J., 2003. "Forecasting Autoregressive Time Series with Bais- corrected Parameter Estimators", International Journal of Forecasting, Vol. 19, pp: 493-502.

Lattin, J., J. Carrol and P. Green, 2003. "Analyzing Multivariate Data", U. K.

Layton, A., L. Defris and B. Zehnwirth, 1986. "An International Comparison of Economic Leading Indicators of Telecommunication Traffic ", International Journal of Forecasting, Vol. 2, pp: 413-425.

Neter, J., W. Wasserman and G. Whitmore, 1988. "Applied Statistics ", New York.

Pankratz, A., 1983. "Forecasting With Uniariate Box-Jenkins Models ", Jon Wiley & Sons, New York.

Reid, D., 1971. "Forecasting in Action – A Comparison of Forecasting Techniques in Economic Time Series ", Joint Conference of OR Society's group on Long Term Research Planning and Forecasting,

Smith, B., B. Williams and R. Oswald, 2002. "Comparison of Parametric and Nonparametric Models for Traffic Flow Forecasting ", Transportation Research, Vol. 10, pp: 257-321.

Williams, B., 2001. "Multivariate Vehicular Traffic Flow Prediction: Evaluation of ARIMAX Modeling", Transportation Research Board, Washington, pp: 194-200.

Williams, M. and A. Lester, 2003. "Modeling and Forecasting Vehicular Traffic Flow as a Seasonal ARIMA Process: Theoretical Basis and Empirical Results", Journal of Transportation Engineering, Vol. 129, pp: 664-672.

Yaffee, R. and M. McGee, 2000. "Introduction to Time Series Analysis and Forecasting", San Diego: Academic Press, Inc.,

Yam, R., R. Whitfield and R. Chung, 2000. "Forecasting Traffic Generation in Public Housing Estates", Journal of Transportation Engineering, Vol. 126, pp: 358-361.