VECTOR ARMA TIME SERIES MODELS FOR SHORT-TERM PREDICTION OF TRAFFIC DEMAND

Takashi UCHIDA
Associate Professor
Department of Civil Engineering
Tohoku University
Aoba, Aoba-ku, Sendai,
980-77 Japan
Fax: +81-22-217-7477
E-mail: uchida@plan.civil.tohoku.ac.jp

Satosi YAMASITA
Associate Professor
Institute of Statistical Mathematics
4-6-7 Minami-Azabu, Minato-ku,
Tokyo,
106 Japan
Fax: +81-3-5421-8796
E-mail: yamasita@ism.ac.jp

abstract: When ATIS (Advanced Traveler Information System) is applied as a traffic control measure, short-term prediction of traffic demand is a critical issue. This study aims to a simple prediction model of traffic demand based on on-line observation through traffic detectors. Multivariate time-series (VARMA; Vector Auto-Regressive Moving Average) models which predict 5-minutes traffic volume considering correlation between sites are proposed to apply. Applicability of VARMA model to actual fluctuation of traffic demand depends on adequacy of the hypothesis of stationary stochastic process. This study confirms empirically the validity of stationarization and VARMA model through estimating models using field data.

1. INTRODUCTION

Short-term prediction of traffic conditions is a basis for traffic control. Its importance gets large when the ATIS (Advanced Traveler Information System) is applied as a traffic control measure, because the effects of providing information for dispersing traffic depend on the precision of forecast information. Short-term prediction of traffic flow transition revealed may be attained using some sort of dynamic traffic flow simulation, however, predicting its input—traffic demand—is a critical issue. Minute to minute predictions of traffic demand of which horizon is about an hour is necessary.

Many efforts have been devoted for traffic demand prediction based on on-line observation through traffic detectors. They include “flexible” or “real-time” modeling approach such as Kalman filter techniques and Neural Network, as well as traditional statistical models and time-series models. However the “flexible” models might have potentials to improve prediction, they have not achieved sufficient prediction comparing their computational loads required [van der Voort et. al., 1996]. On the other hand, time-series models, which are easier to compute, have not yet been examined thoroughly. In many cases, single-variate time-series models have been considered. Multivariate (vector) models have been scarcely applied.

This study aims to a simple prediction model of traffic demand, considering the applicability to a large-scale road network. Therefore, multi-variate time-series (VARMA; Vector Auto-Regressive Moving Average) models are considered. Some VARMA/VAR models have already achieved moderate prediction [Okutani, 1990]. But their time periods where traffic volumes aggregated are such lather long as 15 minutes or 30 minutes. On the other hand, traffic control and providing information are conducted usually based on 1 to 5 minutes period. It gets harder to predict traffic volume when its aggregation time period shorten, because fluctuation of volume grows. This study tries to overcome the difficulty of predicting 5-minutes traffic volume by considering correlation between sites. We examine the capability of VARMA model for short-term detail prediction.

Time-series models assume stationary stochastic process. It means that the variables are generated through a stochastic process of which mean and variance are stable. Therefore, some data transformation, that is called as "stationarization", is necessary to model an actual phenomenon. This study examines empirically the validity of stationarization and VARMA through estimating models using field data. Firstly, preliminary analysis is conducted, where single-variate ARIMA for an on-ramp traffic volume is examined. Then multivariate model—VARMA for section traffic volumes—is estimated.

2. THE MODEL

Applicability of ARMA model to actual fluctuation of traffic demand depends on adequacy of the hypothesis of (weakly-)stationary stochastic process, which means the process generating fluctuation is a stochastic process which has constant mean and variance. It can hardly be expected that raw traffic volume data achieve stationarity. So, data transformation, that is called as "stationarization" is necessary. Here, an observed value of traffic volume is decomposed into two elements:

$$x_t = m_t + z_t, \tag{1}$$

where, x_t : traffic volume at time period t,

m_t : mean for time period t,

z_t : stochastic fluctuation of traffic volume at time period t.

The mean m_t, which may be a function of exogenous variables, is not stochastic. The z_t is assumed to be a stationary stochastic time-series.

The general ARMA(p,q) model is formulated as follows:

$$z_t = \sum_{i=1}^{p} \phi_i z_{t-i} - \sum_{j=1}^{q} \theta_j a_{t-j} + a_t, \tag{2}$$

where, a_t : white noise $\sim$ N$[0, \sigma^2]$,

ϕ : parameters for AR (Auto Regressive) terms,

θ : parameters for MA (Moving Average) terms.

Equation (2) includes two structural parameters p, q (≥ 0), which denote orders of AR and MA, respectively. The optimum orders of ARMA(p, q) should be identified for a concrete dataset based on a suitable criterion. In this study, the AIC (Akaike's Information Criterion) is used. It is one of the criteria which are based on likelihood, and can be used for selection between different types of statistical model [Hiromatsu *et al.* (1990), Naidu (1996)]. The AIC is defined basically as follows:

$$\text{AIC} = -2 \times [\text{maximum logarithmic likelihood}] + 2 \times (\text{number of parameters}). \quad (3)$$

Based on AIC, the model which minimizes the value of AIC is selected as the optimum model.

3. THE DATA

The data used are observed during October, 1994 at Hanshin Expressway Route 11 in Osaka, Japan. It contains 5-minutes traffic volumes of road sections and on-ramps, which are collected by supersonic vehicle detectors. Figure 1 shows the configuration of observation sites. Each road segment has a single section where traffic volumes of entire carriageways are observed. Notations of variables are given below:

x_{ijk} : traffic volume at road segment i,
$y_{i'jk}$: traffic volume at on-ramp i',
i, i': no. of site,
j: no. of day,
k: no. of time period (5-minutes).

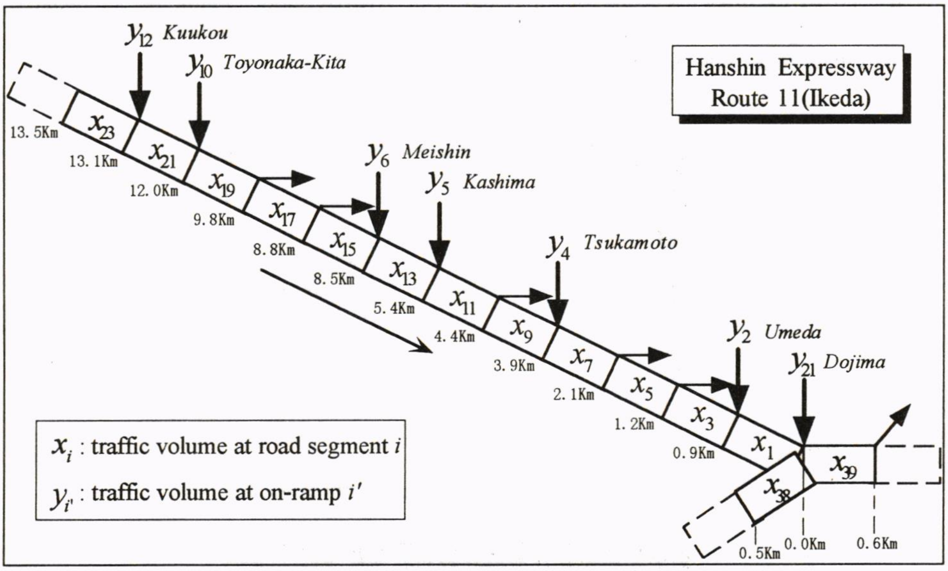

Figure 1 Configuration of observation sites: road segments and on-ramps.

4. PRELIMINARY ANALYSIS – Single Variate ARIMA for On-Ramp Traffic Volume

Applicability of ARMA models to traffic volumes depends on the pre-process of stationarization. In this section, single-variate model is examined in order to assess the validity of stationarization. We assume stationary mean function m_t as shown in equation (1). In addition such transformation, difference, which is a standard technique for stationarization, is also tested in this chapter. The ARMA(p, q) model based on differenced data-set is denoted as ARIMA(p, d, q), where d is the order of difference.

Two cases of d = 0 and 1 are estimated for some combinations of p, q. The dependent variable is an on-ramp traffic volume y_4. (Single-variate model is suitable for on-ramp traffic volume because it would have little relation with other site traffic volumes.) The models are estimated following the procedure shown below. The data used is 7 days data from 1 Oct. to 7 Oct.

1) In order to attain stationary variance, transformation $y_{r4} = \sqrt{y_4}$ is conducted.
2) Deviation $z_4 = y_{r4} - m_4$ is calculated, where m_4 is given by 30 minutes moving average of y_{r4}.
3) SACF (Sample Auto Correlation Function) and SPACF (Sample Partial Auto Correlation Function) are calculated to assess stationarity of z_4 and to limit the range of p, d, *and* q in following procedure.
   Figure 2 plots SACFs and SPACFs of z_4 (d=0) and differenced data z'_4 (d=1). These figures show that a) MA term is dominant and MA(4) or MA(5) models are suggested as shown by drastic declination of SACF, but ARMA model might be appreciated considering the contraction of parameters, b) there seems no periodic factor to reside, stationary time-series are achieved for both d=0 and d=1.

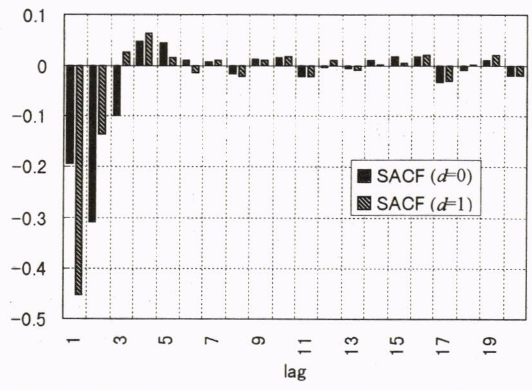

Figure 2 (a) Sample auto correlation function of z_4.

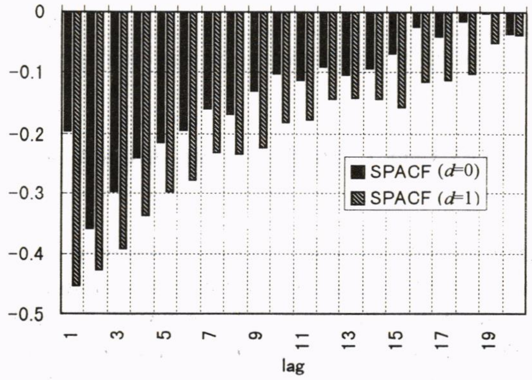

Figure 2 (b) Sample partial auto correlation function of z_4.

4) ARIMA models of z_4 are estimated for combinations of p=0 to 4, d=0 to 1, q=0 to 4. The optimal models for each d are selected based on AIC.
Table 1 shows AICs. As for d=0, ARIMA(2,0,3) (p=2, q=3) results in minimum AIC, and ARIMA(4,1,4) (p=4, q=4) for d=1. Comparing these two models, ARIMA(2,0,3) is superior in any diagnostics (except Adjusted R-squared) as shown in Table 2. The model achieves good reproducibility of fluctuation as shown by Figure 3 which plots actual time series y_4, squared mean function m_t^2 and one-term predictions by ARIMA(2,0,3).

Table 1(a) AICs for ARIMA(p,d,q) (d=0, dependent variable: z_4).

Dependent Variable = z_4 d = 0

p \ q	0	1	2	3	4
0		2,072.8	1,842.0	1,827.6	1,778.7
1	2,761.2	N/A	1,836.7	1,778.9	N/A
2	2,485.3	1,775.5	1,750.5	1,748.5	1,750.3
3	2,299.8	1,759.1	N/A	N/A	N/A
4	2,183.7	1,755.5	1,763.1	N/A	N/A

Number of observations = 2,010

Table 1(b) AICs for ARIMA(p,d,q) (d=1, dependent variable: z'_4).

Dependent Variable = z'_4 d = 1

p \ q	0	1	2	3	4
0		2,839.8	3,120.3	3,055.3	2,627.7
1	4,127.3	N/A	3,518.5	N/A	N/A
2	3,724.3	2,604.4	2,557.6	2,395.1	N/A
3	3,392.8	2,325.7	2,488.9	N/A	N/A
4	3,154.0	2,189.9	2,191.7	N/A	1,852.0

Number of observations = 2,010

Table 2 Diagnostics of optimum models for each d.

	ARIMA(2,0,3)	ARIMA(4,1,4)
Dependent variable	z_4	z'_4
Number of observations	2010	2010
Mean of dependent variable	-0.197943E-03	0.205327E-03
Std. dev. of dependent var.	0.490162	0.757778
Sum of squared residuals	279.487	293.355
Variance of residuals	0.139395	0.146604
Std. error of regression	0.373356	0.382890
Adjusted R-squared	0.420991	0.744888
Durbin-Watson statistic	1.99939	2.06712
AIC	1748.5	1852.0

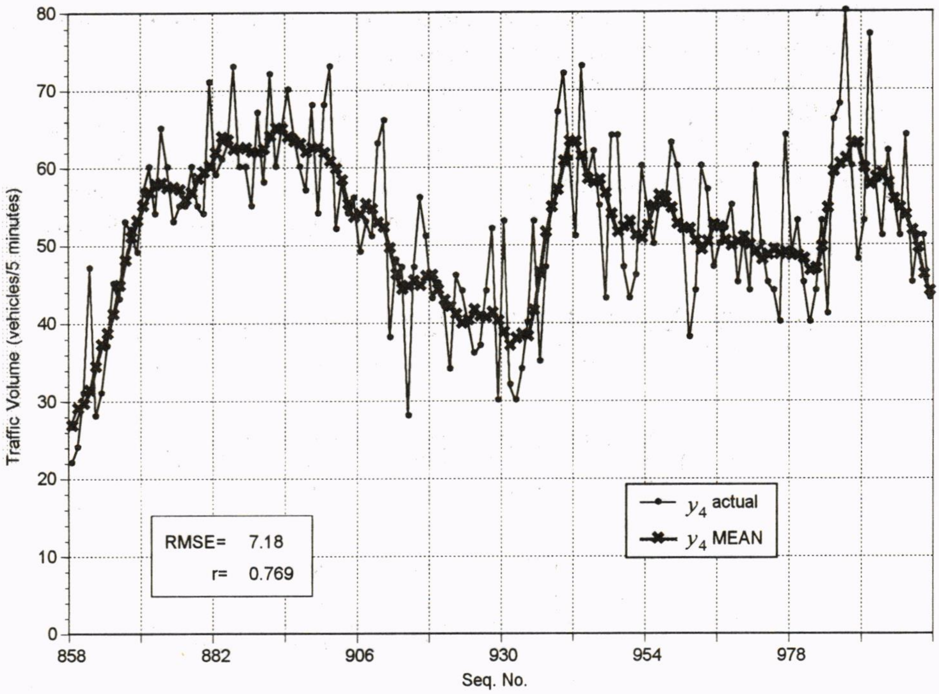

Figure 3(a) Time series of y_4 and squared mean function m_t^2.

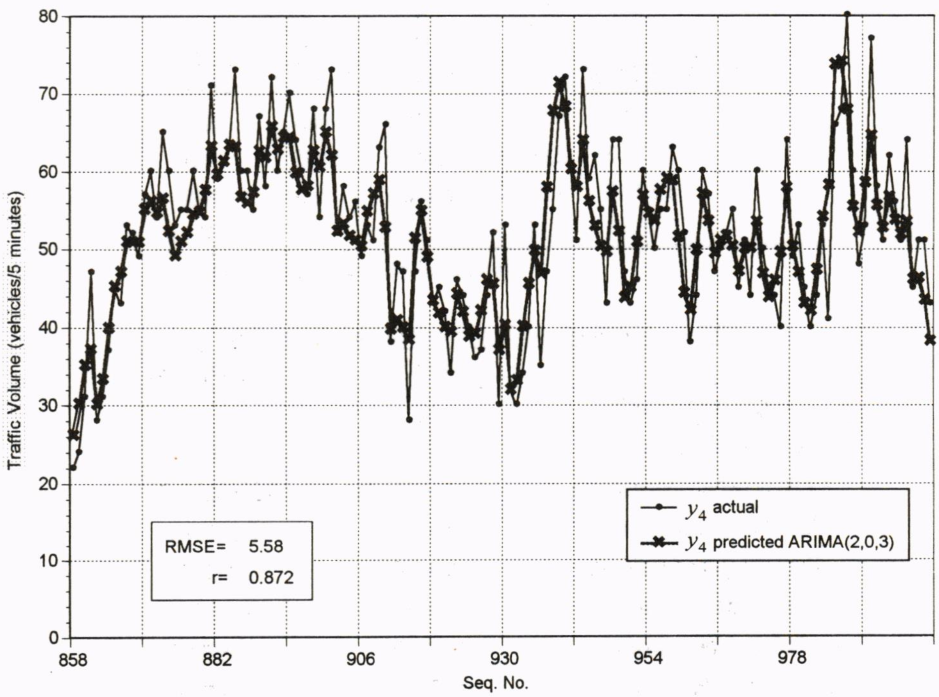

Figure 3(b) Time series of y_4 and predictions by ARIMA(2,0,3).

The preliminary analysis described above concludes as a) using mean function m_t (d=0) is sufficient for stationarization and the difference (d=1) is not necessary to consider, b) good reproducibility of traffic volume fluctuation can be achieved by using ARMA model.

The capability of ARMA model, mentioned above as conclusion b), is confirmed by estimating another model for a different site, of which dependent variable is z_5 (y_5). Table 3 shows that the optimum model is ARIMA(2,0,2). Figure 4 shows that the optimum model for y_5 also achieves good reproducibility. The result certifies the applicability of the ARMA model. In addition, comparing the parameter estimates of the models for z_4 and z_5, these models have similar values of parameters as shown in Table 4, in spite of the difference of the order q. It suggests that the ARMA model for on-ramp traffic volume could be spatially transferable.

Table 3 AICs for ARIMA(p,d,q) (d=0, dependent variable: z_5).

Dependent Variable = z_5			d = 0		
			q		
p	0	1	2	3	4
0		2,088.2	1,830.8	1,829.5	1,801.0
1	2,684.3	2,113.9	N/A	1,823.9	N/A
2	2,437.9	1,796.9	1,790.75	1,792.4	1,793.0
3	2,254.2	1,790.76	1,794.0	1,796.6	N/A
4	2,126.1	1,792.3	1,794.1	N/A	N/A

Number of observations = 2,010

Table 4 Parameter estimates of the optimum ARIMA for z_4 & z_5.

Dependent variable	z_4		z_5	
	ARIMA(2,0,3)		ARIMA(2,0,2)	
	Parameter Estimate	t-statistic	Parameter Estimate	t-statistic
ϕ_1	0.706563	12.7762	0.57741	7.58152
ϕ_2	-0.31026	-5.79272	-0.32972	-12.0267
θ_1	1.39184	24.1766	1.21207	15.1327
θ_2	-0.31806	-3.51996	-0.24739	-3.16017
θ_3	-0.1072	-1.95393	-	-

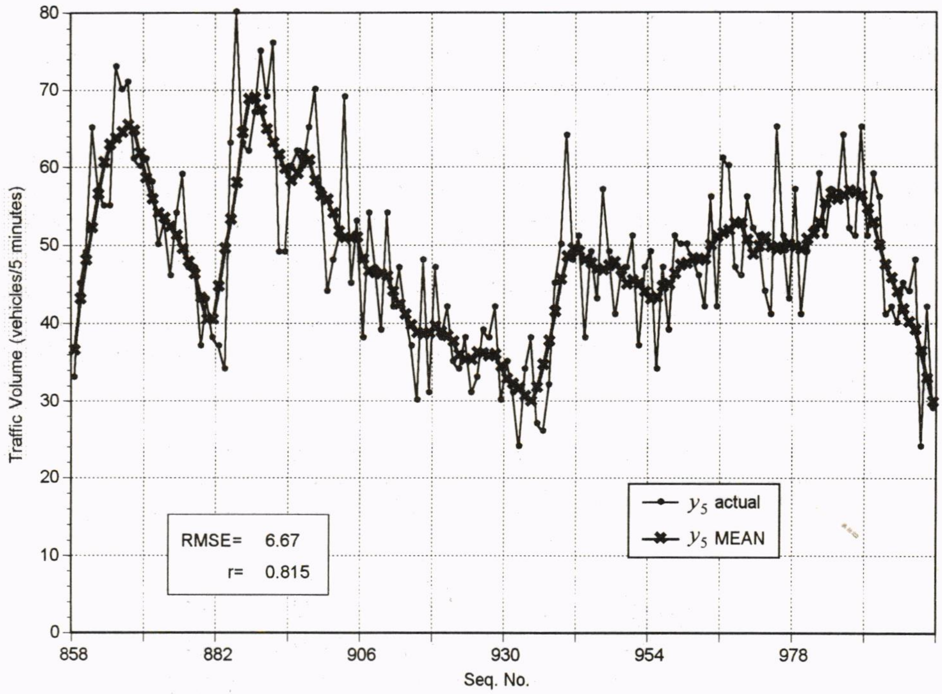

Figure 4(a) Time series of y_5 and squared mean function m_t^2.

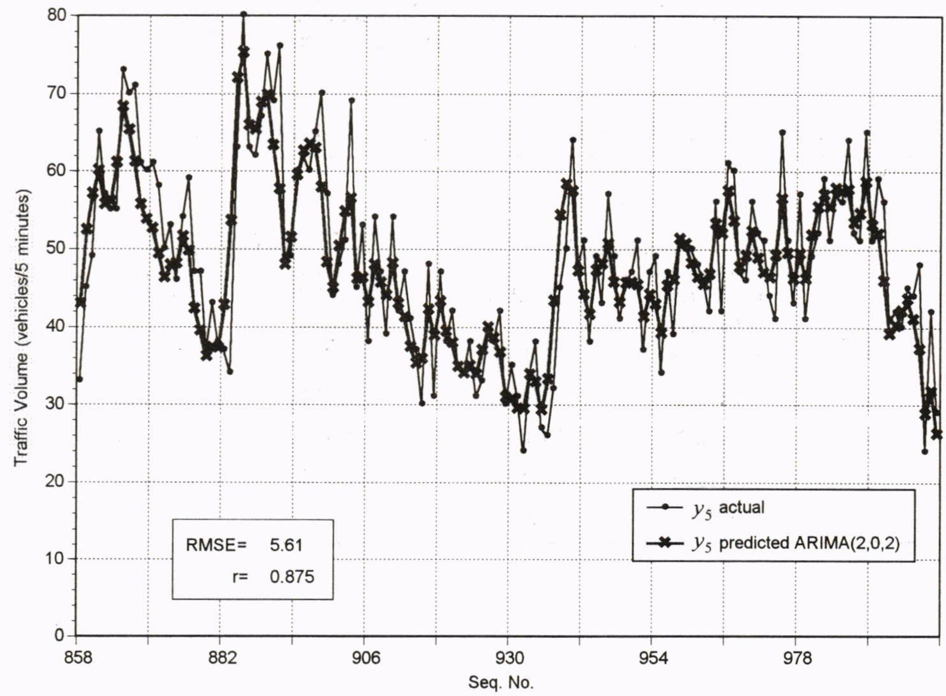

Figure 4(b) Time series of y_5 and predictions by ARIMA(2,0,2).

5. VARMA MODEL FOR TRAFFIC VOLUME PREDICTION

Traffic volumes of road segments are related mutually, especially with successive up/down-stream segments. Considering correlation between sites, the single-variate ARMA model is extended to multivariate model, that is called as VARMA (Vector ARMA). In this chapter, VARMA model is estimated and its ability of forecast is assessed empirically. The data used in this chapter consists of 23 weekdays data in October 1994. The number of observations (5 minutes traffic volumes) for each site is 12×24(hour)×23(days)=6,072.

Here, we consider all sites' traffic volumes depicted in Figure 1 as dependent variables simultaneously. The dependent variables consist of 14 x's (road segment traffic volumes) and 7 y's (on-ramp traffic volumes). The VARMA model is formulated as an extension of eq. (2) by using vector expression as follows:

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \sum_{m=1}^{p} \boldsymbol{\phi}_m \begin{pmatrix} x_{t-m} \\ y_{t-m} \end{pmatrix} - \sum_{n=1}^{q} \boldsymbol{\theta}_n \boldsymbol{a}_{t-n} + \boldsymbol{a}_t \tag{4}$$

The model is estimated through the manner that is similar to the manner mentioned in chapter 4 for a single-variate model. Exceptions are as follows; (There is no substantial difference between x and y. So, x represents y for simplicity.)

1) Stationarization is conducted by only subtracting mean from raw data complying eq.(1).

2) Mean m_t in eq. (1) is defined as:

$$m_t = m_{ik} = \frac{1}{J}\sum_{j=1}^{J} x_{ijk} \tag{5}$$

where, time sequence t is transformed to day j and time k for each day.

It means that only a daily cycle is considered as a systematic component of a traffic volume fluctuation. Note that only weekday data are used.

3) Parameter matrices ϕ, θ are restricted as that at most two segments upstream/downstream are correlated. Eq. (4) is reformulated as eq. (6) by using Δ, which is given as an incidence matrix (Table 5):

$$\begin{aligned} x_{i,t} = \sum_{\Delta}\left(\sum_{m=1}^{p} \boldsymbol{\phi}_m^x(i,\Delta) x_{i'(i,\Delta),t-m} - \sum_{n=1}^{q} \boldsymbol{\theta}_n^x(i,\Delta) a_{i'(i,\Delta),t-n} \right) \\ + \sum_{\Delta}\sum_{m=1}^{p} \boldsymbol{\phi}_m^y(i,\Delta) y_{i'(i,\Delta),t-m} + a_{i,t} \end{aligned} \tag{6}$$

where, $i'(i,\Delta)$: no. of site defined by incidence matrix.

The element of an incidence matrix gives a distance between sites. The null element means that correlation between corresponding sites is <u>not</u> considered. The parameter matrices ϕ, θ have similar structure to the incidence matrix. The parameters corresponding to null elements are eliminated.

Table 5 Incidence matrix.

Δ	x_1	x_3	x_5	x_7	x_9	x_{11}	x_{13}	x_{15}	x_{17}	x_{19}	x_{21}	x_{23}	x_{38}	x_{39}	y_2	y_4	y_5	y_6	y_{10}	y_{12}	y_{21}
x_1	0	+1	+2											-1	+1						
x_3	-1	0	+1	+2										-2							
x_5	-2	-1	0	+1	+2											+2					
x_7		-2	-1	0	+1	+2										+1					
x_9			-2	-1	0	+1	+2										+2				
x_{11}				-2	-1	0	+1	+2									+1	+2			
x_{13}					-2	-1	0	+1	+2									+1			
x_{15}						-2	-1	0	+1	+2											
x_{17}							-2	-1	0	+1	+2								+2		
x_{19}								-2	-1	0	+1	+2							+1	+2	
x_{21}									-2	-1	0	+1								+1	
x_{23}										-2	-1	0									
x_{38}													0	-1							
x_{39}	+1	+2												0							
y_2															0						
y_4																0					
y_5																	0				
y_6																		0			
y_{10}																			0		
y_{12}																				0	
y_{21}																					0

ARMA models are estimated for some combinations of p, q for 3 cases of incidence matrix:

Case I: only considers auto-correlation. $\Delta \neq 0$ are eliminated.

Case II: considers upstream correlation. $\Delta < 0$ are eliminated.

Case III: considers both upstream and downstream correlation.

The optimum model is selected based on AIC. Table 6 which lists number of parameters and AIC shows that ARMA(2,1) in Case III is optimum. Comparing the cases, it is confirmed that considering correlation between sites is effective. Table 7 shows a portion of parameter estimates of the optimum model ARMA(2,1). These tables shows that AR term is dominant, in contrast to the single-variate model for on-ramp described in chapter 4.

The reproducibility of the model is assessed by comparing the fluctuations of one-period predictions and observed. For example, Figure 5 plots actual traffic volumes of road segment 1 x_1 and the values predicted by ARMA(2,1). Generally, the model reproduces fluctuation pattern very well. It reproduces even the fall at time sequence 100-130, in spite of that the model assumes stationary stochastic process. Though too violent fluctuation is not followed properly and results large residual, the model predicts the tendency of increase/decrease correctly.

Table 6 AICs for Vector ARMA(p,q).

	Case I ($\Delta = 0$)	Case II (Δ = 0 to 2)	Case III (Δ = -2 to 2)
AR(1)	21 12887.5	57 9899.7	81 7274.1
AR(2)	42 10969.6	114 8671.3	162 6514.0
ARMA(1,1)	42 12495.7	101 9170.0	149 6685.6
ARMA(2,1)	65 10985.6	158 8755.0	230 6356.5
ARMA(2,2)	84 10933.2	202 8756.6	298 6441.3

upper : no. of parms., no. of variables = 21

lower : AIC -2210000 no. of observations = 127,512 (= 6,072 $\times$ 21)

Table 7 Parameter estimates of VARMA(2,1) (portion).

	x_1	x_3	x_5	x_7	x_9	x_{11}	x_{13}	x_{15}	x_{17}	x_{19}
$\phi_1^x(i,+2)$	0.414	0.521	0.456	0.180	0.199	0.132	0.081	0.060	0.165	0.402
$\phi_2^x(i,+2)$	-0.085	0.080	0.025	0.074	0.019	-0.051	-0.057	0.000	-0.027	-0.069
$\phi_1^x(i,+1)$	0.021	-0.345	0.139	0.367	0.242	0.274	0.214	0.154	0.062	0.080
$\phi_2^x(i,+1)$	0.061	-0.238	-0.102	-0.061	0.060	0.050	0.087	0.086	0.042	0.068
$\phi_1^x(i,0)$	0.161	0.177	-0.074	-0.225	-0.074	-0.127	-0.016	0.040	-0.052	-0.091
$\phi_2^x(i,0)$	0.148	0.241	0.030	-0.230	-0.207	-0.141	-0.088	-0.009	0.043	-0.005
$\phi_1^x(i,-1)$	0.029	0.086	-0.008	0.433	0.291	0.319	0.268	0.280	0.277	0.411
$\phi_2^x(i,-1)$	-0.025	0.013	-0.006	0.251	-0.011	-0.050	0.030	-0.039	0.056	-0.018
$\phi_1^x(i,-2)$		0.045	0.214	0.071	0.158	0.323	0.259	-0.028	0.087	-0.036
$\phi_2^x(i,-2)$		0.024	0.096	0.025	0.180	0.109	0.073	0.128	0.071	0.197
$\theta_1^x(i,+2)$	-0.015	-0.034	0.009	0.107	0.045	0.058	0.017	0.076	0.028	0.016
$\theta_1^x(i,+1)$	0.024	0.102	0.031	-0.055	0.094	0.037	0.014	0.013	0.033	0.012
$\theta_1^x(i,0)$	-0.034	-0.061	-0.030	0.059	-0.136	-0.078	0.003	-0.023	-0.023	-0.031
$\theta_1^x(i,-1)$	0.022	-0.012	-0.015	-0.123	0.066	0.089	0.054	0.009	0.004	-0.004
$\theta_1^x(i,-2)$		0.029	0.029	0.030	-0.044	-0.068	-0.077	0.009	0.003	-0.038
$\phi_1^y(i,+2)$			0.420		0.142	0.218			0.254	0.594
$\phi_2^y(i,+2)$			0.027		-0.065	-0.190			-0.258	-0.201
$\phi_1^y(i,+1)$	0.198			0.289		0.148	0.228			0.264
$\phi_2^y(i,+1)$	0.011			0.042		0.134	-0.030			-0.028

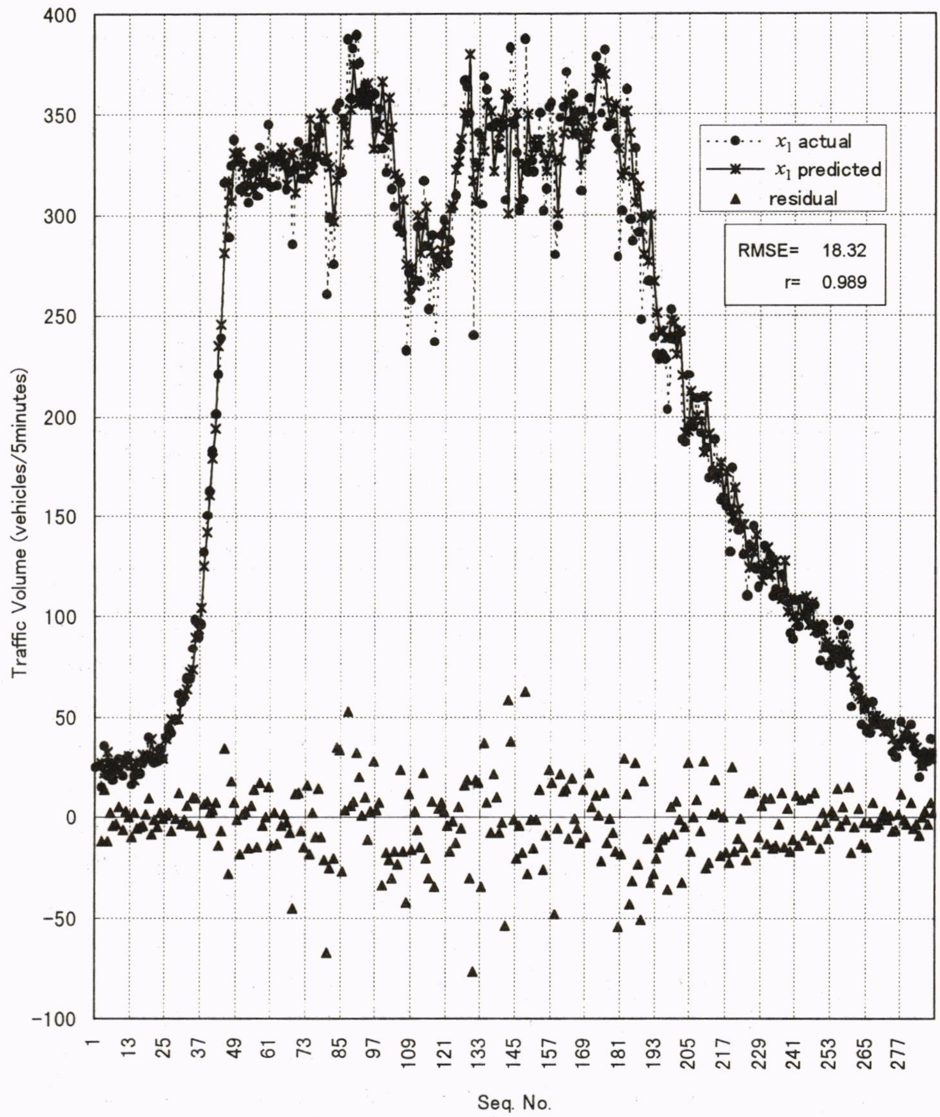

Figure 5 Time series of x_1 and predictions by VARMA(2,1).

Capability of ARMA model to forecast is assessed by applying ARMA(2,1) estimated above to the data of 30 Oct. which is preserved from model estimation. 4-periods predictions are calculated and compared with observed values. Figure 6 shows that the model forecasts generally well, especially even the peak without delay. But accidental fall can not be forecasted such as x_3 and x_{13} during 9:45-11:15. Note again that the model is estimated statically based on the mean function. As shown in Figure 7 which focuses on the discrepancy of x_{13}, the deviations from mean are very large and the change is suddenly. Considering these conditions, it should be noticed that good deals of corrections from mean are attained by the model. Therefore, improvement of forecast would be achieved if more lags, that are larger p and q, were considered when estimating the model.

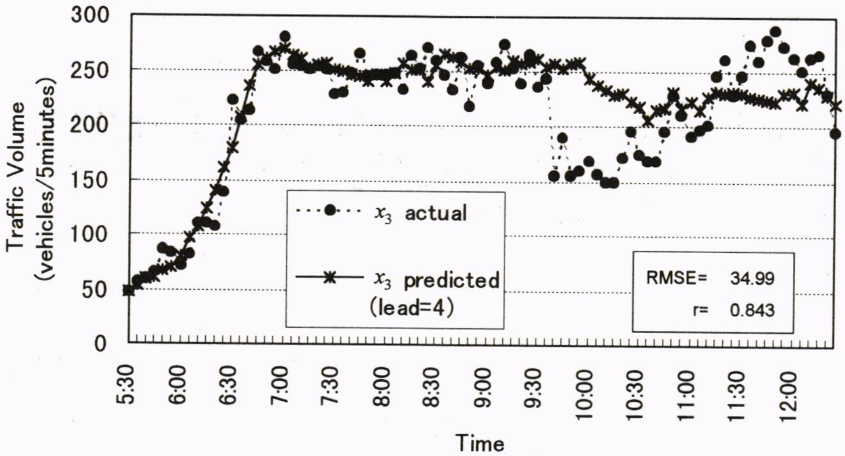

Figure 6 (a) Time series of x_3 and 4-periods predictions by VARMA(2,1).

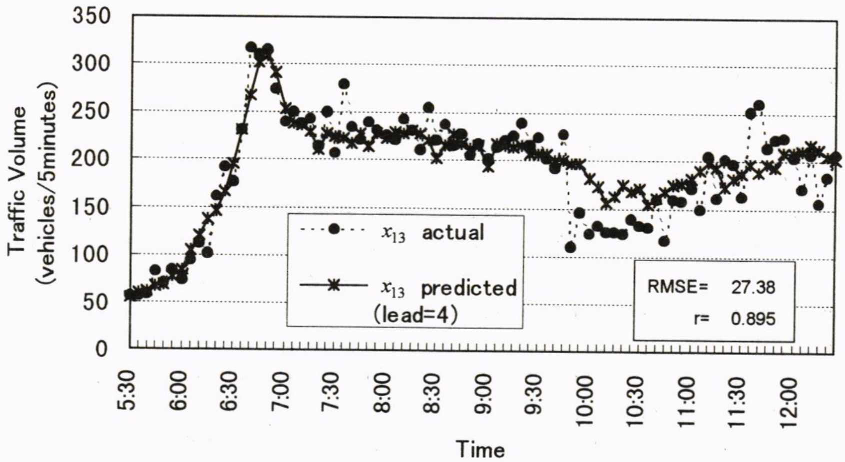

Figure 6 (b) Time series of x_{13} and 4-periods predictions by VARMA(2,1).

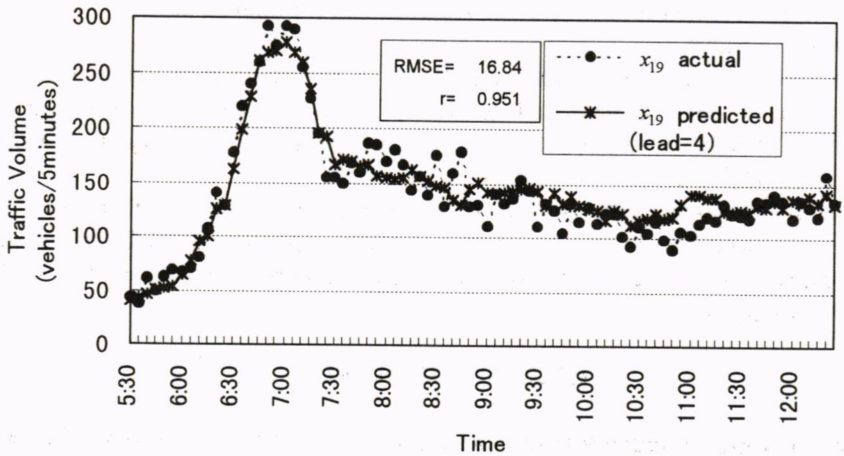

Figure 6 (c) Time series of x_{19} and 4-periods predictions by VARMA(2,1).

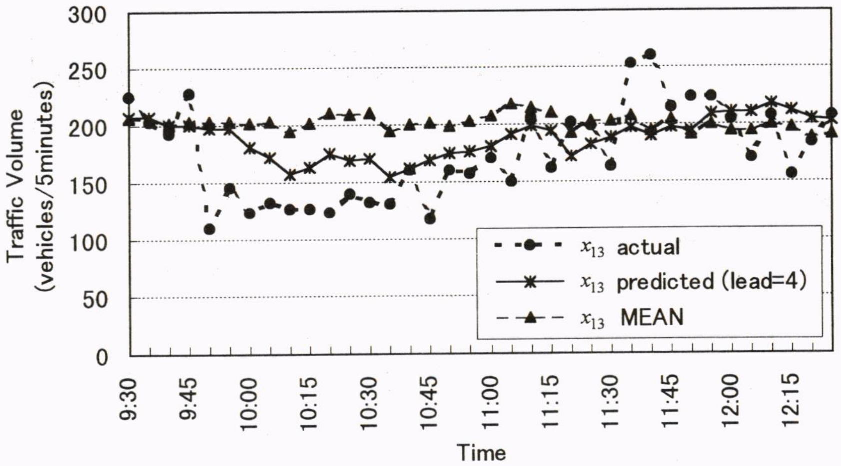

Figure 7 Time series of x_{13}, mean and 4-periods predictions by VARMA(2,1).

6. CONCLUSION

This study aims to a simple short-term prediction model of traffic demand, considering the applicability to a large-scale road network. We assumed stationarity of traffic volume fluctuation and estimated time-series model for short-term prediction empirically. It can hardly be expected that raw traffic volume data achieve stationarity. Therefore, mean function was introduced for stationarization and the deviation was treated as a stochastic time-series. A single-variate ARMA model for on-ramp traffic volume and a Vector ARMA model for road segment traffic volumes were proposed to apply. The validity of stationarization and the capability of ARMA models were examined empirically. The results confirmed basically that traffic volume fluctuation could be modeled by VARMA with stationarization. In spite of simple treatment of considering only daily cycle, fairly good forecast was achieved.

However, this study is on the first step and examined only the simplest model. There exist some issues to improve:

1) stationarity of variance is not always attained. Especially when the traffic volume is large and congested, volatility tends to grow. Therefore, some more complex model such as GARCH (Generalized Auto-Regressive Conditionally Heteroscedastic) model which considers varying variance should be examined.
2) spatial and temporal transferability should be examined on the larger road network. To attain transferability, parameters which are specific to sites should be contracted. Contraction of parameters also contributes computability of more complex model mentioned above.

ACKNOWLEDGMENTS

We wish to express our gratitude to *Hanshin Expressway Public Corporation* for their help in providing us with data.

REFERENCES

Hiromatsu, H. & S. Naniwa (1990) **Economic Time Series Analysis**, Asakura-Shoten, Tokyo (in Japanese).

Naidu, P.S. (1996) **Modern Spectrum Analysis of Time Series**, CRC Press, Boca Raton, Florida.

Okutani, I. (1990) Examinations of the Applicability of Traffic Demand Predictors, **T.IEE Japan**, Vol. 110-D, No.6, pp.673-682 (in Japanese).

van der Voort, M. *et. al.* (1996) Combining Kohonen Maps with ARIMA Time Series Models to Forecast Traffic Flow, **Transpn. Res.-C**, Vol.4, No.5, pp.307-318.