OD VOLUME VARIATION AND ITS PREDICTION BASED ON URBAN EXPRESSWAY ETC-OD DATA

Hiroaki NISHIUCHI, Assistant Professor, Dr. Eng., Department of Science and Technology, Nihon University, JAPAN

Masao KUWAHARA, Professor, Ph.D, Institute of Industrial Science, Tohoku University, JAPAN

Marc Phillip MISKA, Lecturer, Ph.D, Institute of Industrial Science, The University of Tokyo, JAPAN

Hiroshi WARITA, Dr. Eng., Metropolitan Expressway Co.Ltd, JAPAN

ABSTRACT

This paper presents characteristics of OD volume variation using ETC-OD Data and propose stochastic short-term OD volume prediction model applying Bayesian Network. Since Electronic Toll Collection (ETC) system is installed at tollgates of expressways in Japan, Drivers with an ETC on-board unit in their car do not have to stop at tollgates and the payment is done electronically when the vehicle passes the sensors. In addition penetration of ETC users smoothly reached to 80% in case of Metropolitan Expressway in Japan. This means the system is collecting 80% of dynamic OD volume data in expressway network every time. Therefore this research focusing on OD volume variation which was difficult to collect the data until now, and stochastic short-term OD volume prediction model is proposed by applying Bayesian Network. Results of OD variation analysis shows that several patterns of OD pairs' location can identify characteristics of OD volume variation, and it also shows proposed OD volume prediction model can make it predict better results than historical average OD volume for target OD pairs.

Keywords: OD Volume Variation, Short-term OD Volume Prediction, ETC-OD Data, Bayesian Network

1 INTORODUCTION

Usage of dynamic traffic information is recommended to mitigate traffic congestion by Dynamic Traffic Management (DTM). For example, Tokyo Metropolitan Expressway (MEX) is developing a real time traffic simulation system using dynamic origin-destination (OD) volumes (Shiraishi et al. 2004). However, most of current OD volume as input data is estimated based on census survey result that carried out at one day every several years. Therefore, it is difficult to reflect the real time traffic condition to DTM system. Therefore many researchers have already proposed OD prediction model to bring accurate OD volume as input data of DTM systems. Okutani(1987), Ashok and Ben-Akiba(1993), Kachroo et al.(1997), Yong et al(1997) and Zhou and Mahmassani(2007) are developed dynamic OD prediction model applying Kalman filtering technique using traffic simulation and inflow/outflow data. These approaches estimate OD volume at current state, and predict future OD volume at one time step further using the state of measured data at current time and estimated state of data at last one time step as input data of Kalman filtering model. Moreover, another traffic data resource has also contributed to develop OD prediction model by promoting Intelligent Transportation Systems (ITS); e.g. Auto Vehicle Identification (AVI) Data and ETC Data. For example, Tanabe et al. (1995) and Asakura et al. (2000) developed dynamic OD prediction model using AVI data. However, quality of OD data from AVI is depending on number of video camera and its location. That means it is required to make some assumptions like off-ramp choice behaviour similar with the case using inflow/outflow data. On the other hand, ETC service is started in several countries by progress of ITS technique to make it possible to do money transaction by road to vehicle communication. In Japan, ETC system is applied on expressways from March of 2001 at 11 tollgates in MEX. Drivers with an ETC on-board unit in their car do not have to stop at toll gates and the payment is done electronically when the vehicle passes the sensors. The penetration of ETC users is smoothly increasing, being now over 80% of drivers on MEX. The system records the usage of on-ramps and off-ramps. With that, we can acquire the origin and destination of drivers. By aggregating ETC users' information, we can obtain time-dependent OD volumes which have been a difficult to collect the data. Therefore it can be said that OD volume data in expressway network is collecting automatically from ETC system, and OD estimation/prediction model using ETC data is also proposed not only in Japan, but also in Italy, US and so on. (By Camus et al. (1997), Kwon and Varaiya(2005) and Kurauchi et al. (2008)) However, although they success calculate accurate OD volume, their study was focused on estimation of definite value of OD volume. On the other hand, several study on traffic volume variation e.g; Iida and Takayama(1981), Murakami et al.(1999), Tanaka et al.(2001), Iryo et al(2007), Rakha and Aerde(1995), Stathopulos and Karlaftis(2001) and Weijermars and Berkum(2005) presented that traffic volume is varying by many factors such as weather condition, difference of day of week, and of course time of day by occurring of accident and other reason. Therefore, predicted OD volume which aim to used for DTM systems should react to above factors. However, there is no study on analysis of characteristics of OD volume variation using real OD Data, existing studies were analyzed traffic volume variation through Inflow data and detector data on road section. Although Tanaka et al. used AVI data as similar characteristics with OD volume, it is limited number of OD pair and sampling.

Since OD Data from ETC records are time-dependent and huge in amount, we can analyze characteristics of OD volume variation and identify how OD volume is varying by day-to-day and hour-by-hour. Therefore this study analyzes how much OD volume is varying using ETC-OD data measured in MEX. In addition, this paper also identifies relationship between characteristics of OD volume variation and location pattern of OD pairs. And focusing on the knowledge of OD volume variation analysis, this research propose stochastic short-term OD prediction model to include range of OD variation by learning Bayesian Network model, and accuracy of prediction model is validated focusing on characteristics of OD volume variation based on results of its analysis.

2 OUTLINE OF ETC-OD DATA

ETC users' OD information is collected when driver pass the ETC tollgate (Entrance) and exit. ETC-OD Data used here is collected from June 2006 to March 2007 (169 weekdays in this period), and time-period of data collection is 12 hours (from 7am to 7am). In this research, we used aggregated ETC-OD Data as 5 minutes, 30 minutes and 1hour aggregation. In addition toll system of MEX is used fixed price system. Therefore, price of toll is decided by toll area e.g. Tokyo price area, Kanagawa toll area and Saitama toll area. That means when driver travel from Tokyo to Kanagawa, they need to pay twice at mainline toll gate which is located on border of each toll area. However, we focused on trips in Tokyo price area, and choose On and Off ramps on 5 routes which is describes by bold line on figure1. Using above definition of data collection, number of target OD pair for this study is 870 OD pairs.

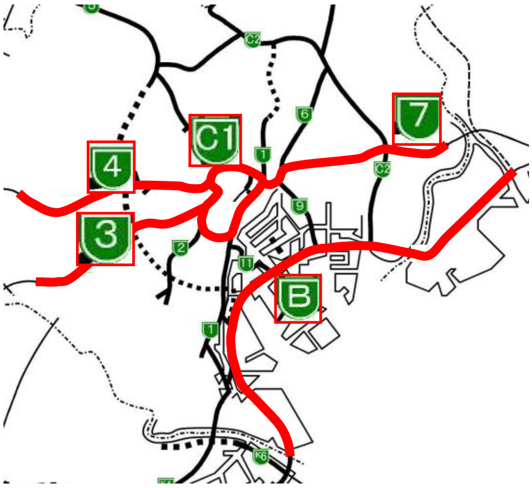


Figure 1 – Metropolitan Expressway Network and Study Routes (Bold Line)

3 ANLAYSIS OF OD VOLUME VARIATION

3.1 Overview of Analysis of OD Volume Variation

This chapter, variation of OD volume is analyzed using ETC-OD Data. Analysis here is focused on followings to understand its characteristics.

1. Fundamental Analysis of OD Volume Variation

   To understanding fundamental characteristics of OD volume variation, distribution of average of OD volume is checked. In addition relationship between average OD volume and coefficient of variance also analyzed to understand fundamental characteristics of OD volume variation.

2. Daily Variation of OD volume

   If MEX users are not frequently changing their entrance according to traffic condition e.g.; whether, incidents, traffic accident and so on, we can expect that we can measure the almost same amount of traffic volume on everyday and every time. However it is actually expected OD volume is varying by above factors as some researchers are discussed. Therefore, variance of OD volume is analyzed here to know how much OD volume is varying every day and hourly. In addition to that the characteristics of the variance are analyzed focusing on location of OD pairs on the network.

3.2 Fundamental Analysis of OD Volume Variation

3.2.1 Distribution of Average OD Volume

This part is described the difference of level of average OD volume by different OD pairs as fundamental characteristics of OD volume variation. The average of OD volume is calculated by equation (1).

$$\mu_{rs}^{\min} = \frac{1}{hmk} \sum_{h} \sum_{m} \sum_{k} q_{rs}^{\min}(h,m,k) \qquad (1)$$

where,

$\mu_{rs}^{\min}$: Average OD volume from ramp *r* to *s* by aggregated min minutes

$q_{rs}^{\min}(h,m,k)$: OD volume from ramp *r* to *s* by aggregated min minutes measured at *k*th during an hour *h* in *m*th measured day

The characteristics of the distribution are analyzed using 5 minutes aggregated OD volume. Figure 2 shows the distribution of 5minutes average OD volume. From the figures only 17% of OD pairs is averaged 5 minutes OD volume is greater than 1.0 [veh/5min]. Therefore, another 83% of OD pairs in study site network are varying by low level of OD volume.

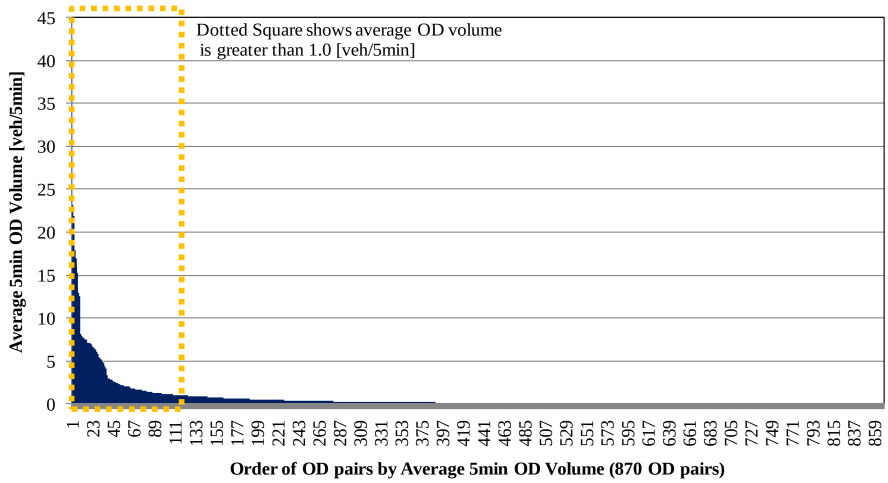


Figure 2 – Distribution of 5min Average OD Volume

3.2.2 Relationship between Average OD Volume and its Coefficient of Variance

More than 80% of OD pairs are varying by less than 1.0[veh/5min]. However, important thing to get knowledge of OD volume fluctuation is to know how much varying OD volume against their average OD volume by day to day. Therefore, coefficient of variance (CV) which is the one of index to describe the variation is determined here. CV is calculated according to equation (2).

$$CV_{rs}^{\min} = \sigma_{rs}^{\min} \Big/ \mu_{rs}^{\min} \qquad (2)$$

$$\sigma_{rs}^{\min} = \sqrt{\frac{\sum_h \sum_m \sum_k (q_{rs}^{\min}(h,m,k) - \mu_{rs}^{\min})^2}{(hmk-1)}} \qquad (3)$$

where,

$CV_{rs}^{\min}$: Coefficient of Variance of OD pair *rs* by aggregated min minute

$\sigma_{rs}^{\min}$: Standard Deviation of OD pair *rs* by aggregated min minute

The relationship between Average OD volume and CV is describes as Figure 3 (5minutes aggregation data). Figure shows that if amount of average OD volume become large, CV become getting smaller on both time aggregation units. Same tendency has been confirmed by Tanaka et al.(2001) which is used Number Plate Matching Data on MEX. Figure 6 shows that same relationship with Figure 4, but it is focused on OD pairs which 5 minutes average OD volume is greater than 1.0[veh/5min]. From figure, there is 20 to 150% of variance against to their average value in case of time unit is 5 minutes. That means if 5minutes average OD volume is around 10 [veh/5min], range of plus minus 5 vehicle can be increase/decrease day by day. Moreover if unit of time aggregation is changed to longer like 1 hour, the range of OD volume variation also become wider than shorter time unit. Therefore considering fluctuation of OD volume should be one of the key factors for applying dynamic traffic control.

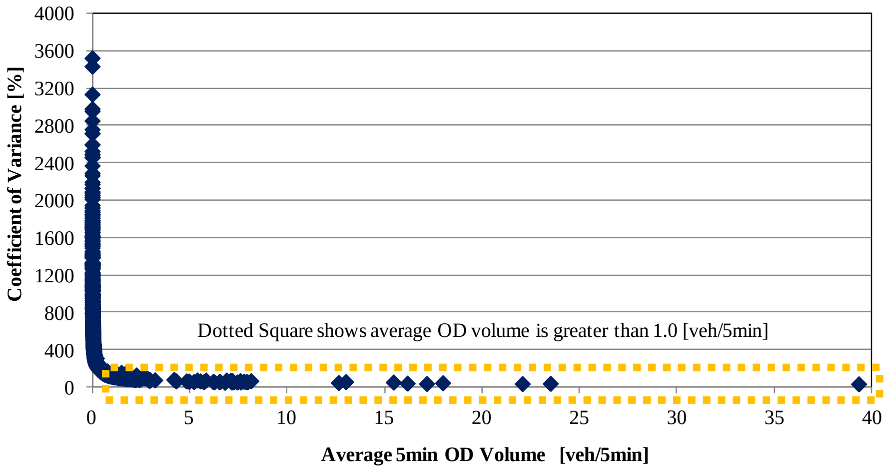


Figure 3 – Relationship between Average 5min OD Volume and Coefficient of Variance (870 OD Pairs)

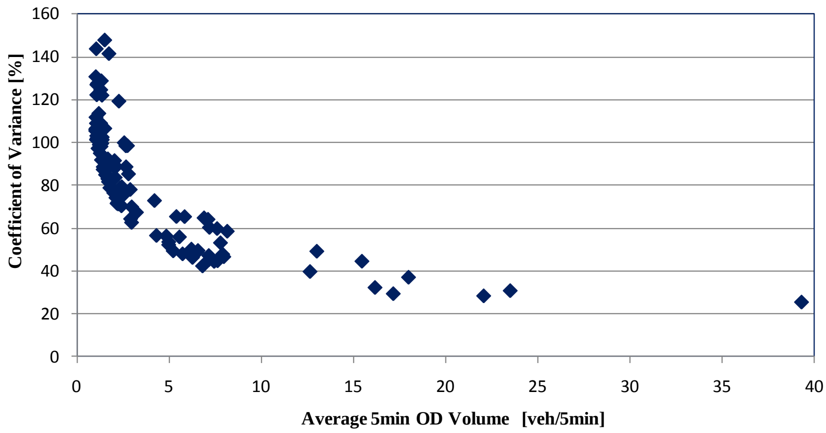


Figure 4 – Relationship between Average 5min OD Volume and Coefficient of Variance (117 OD pairs, Inside of dotted square in Figure 3)

3.3 Daily Variation of OD Volume

In above section, relationship between average and CV of OD volume has analyzed. However the analysis is used averaged weekday's OD volume aggregating 5 minutes, and it showed the range of variation of OD volume. On the other hand, it is well known that traffic condition on the network is changing by factors which are traffic accident, other incident, weather condition, time of day, day of week and so on. Then OD volume is also varying by such kind of factors. In this section, therefore, impact of daily and hourly variation of OD volume is investigated using 5 minutes OD volume data. Two types of variance were used. One is Between-Group (BG) Variance to represent daily OD variation, and another one is Within-Group (WG) Variance to describe variation in the hour. The relationship between BG variance and WG variance analyzed that characteristics of OD variation can be identified by pattern of OD location on MEX. Therefore, first, BG and WG variance is defied, and analyzed to see relationship both variance. And then, characteristics of relationship between BG and WG variance focusing on difference of location pattern of OD pairs are confirmed.

3.3.1 Definition of Between-Group Variance and Within-Group Variance

BG and WG variance are well known used for Analysis of Variance (ANOVA). So this paper is also use the same calculation steps as shown on equation (4) to (11). From the definition of both variances, BG variance is describing how much OD volume is fluctuating day by day in specific time of day. WG variance is meaning fluctuation of OD volume within specific hours on all weekdays. Then if ETC users are and travelling same OD at same time, low level of BG variance at the OD pair should be calculated. In addition if ETC users are travelling same time on every weekday but change a lot entering time in the hour, then high level of WG variance should be calculated.

BG and WG variance is calculated by following steps.

$$V_{A,rs}^{5}(h) = S_{A,rs}^{5}(h) \Big/ \phi_A \quad (4)$$

$$V_{e,rs}^{5}(h) = S_{e,rs}^{5}(h) \Big/ \phi_e \quad (5)$$

$$S_{A,rs}^{5}(h) = \frac{1}{m}\left\{ \sum_m q_{rs}^{60}(h,m)^2 - \frac{\left(\sum_m q_{rs}^{60}(h,m)\right)^2}{m} \right\} \quad (6)$$

$$S_{e,rs}^{5}(h) = S_{T,rs}^{5}(h) - S_{A,rs}^{5}(h) \quad (7)$$

$$S_{T,rs}^{5}(h) = \sum_m \sum_{k=1}^{12} (q_{rs}^{5}(h,m,k))^2 - \frac{\left(\sum_m \sum_{k=1}^{12} q_{rs}^{5}(h,m,k)\right)^2}{12m} \quad (8)$$

$$\phi_A = m-1 \quad (9)$$

$$\phi_e = \phi_T - \phi_A \quad (10)$$

$$\phi_T = 12m-1 \quad (11)$$

where,

$V_{A,rs}^{5}(h)$: Between-Group Variance of 5 minutes OD Volume at ramp *r* to *s* during an hour *h*

$V_{e,rs}^{5}(h)$: Within-Group Variance of 5 minutes OD Volume at ramp *r* to *s* during an hour *h*

$S_{e,rs}^{5}(h)$: Within-Group Square Sum of 5 minutes OD Volume at ramp *r* to *s* during an hour *h*

$S_{A,rs}^{5}(h)$: Between-Group Square Sum of 5 minutes OD Volume at ramp *r* to *s* during an hour *h*

$S_{T,rs}^{5}(h)$: Square Sum of All Data of 5 minutes OD Volume at ramp *r* to *s* during an hour *h*

ϕ_A: Degree of freedom of $S_{A,rs}^{5}(h)$

ϕ_e : Degree of freedom of $S_{e,rs}^{5}(h)$

ϕ_T : Degree of freedom of $S_{T,rs}^{5}(h)$

3.3.2 Characteristics of BG and WG variance of OD Volume

Figure 5 shows the relationship between BG and WG variance by different hours. From the figure, BG variance is drastically increasing according to increasing of VG variance, and it can see the same tendency at all hours. Therefore, different level of OD volume is measured everyday at OD pairs which has high level of WG variance, even if OD volume is measuring on same time every day. And it can be said that the tendency of the relationship is looks same on all time period.

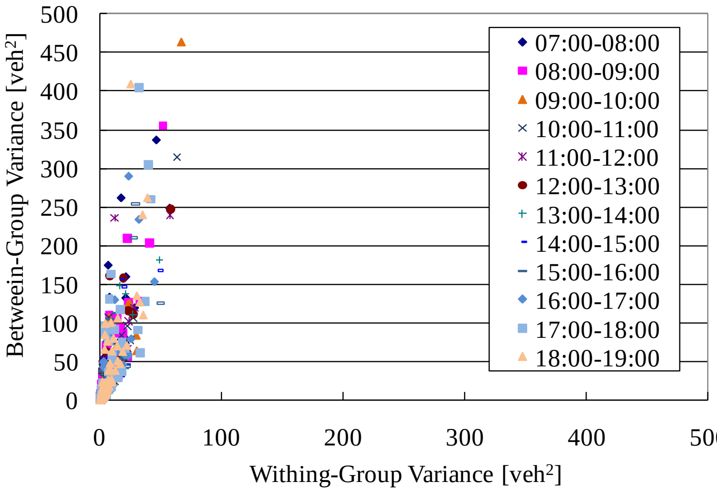


Figure 5 – Relationship between WG Variance and BG Variance by Each Hours

On the other hand, OD pairs which value of both variances is low level are located on bottom of the figures. Therefore, ratio of variance is defined by equation (12) and checked distribution of number of OD pairs by level of ratio of variance. Then the ratio of variance is defined by equation (12), and histogram of ratio of variance is shown on Figure 6.

$$F_{rs}^{5}(h) = V_{A,rs}^{5}(h) \Big/ V_{e,rs}^{5}(h) \qquad (12)$$

where,

$F_{rs}^{5}(h)$: Ratio of Variance of 5 minutes OD Volume at ramp *r* to *s* during an hour *h*

From Figure 6, about 70% of OD pairs are the ratio is 1 to 1.5. That means BG and WG variance of such OD are almost same value, and other OD pairs are BG variance is higher than WG variance since the ratio is calculated greater than 2. About this tendency, it can be considered that average OD traffic volume of most of OD pairs becomes low by definition of calculation of both variances. Therefore, it is clearly understand that value of variance for most of OD pairs becomes also low value.

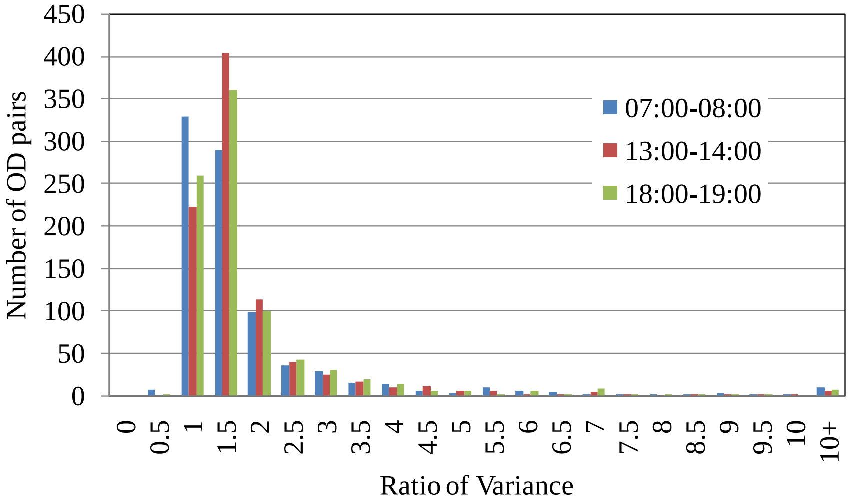


Figure 6 – Number of OD Pairs by Ratio of Variance (In case of 07:00-08:00, 13:00-14:00 and 18:00-19:00)

3.3.3 Relationship between OD volume variation and location of OD pairs

The characteristics of OD volume variation is used for identify location of OD pairs in this part. In particular, it is used normalized BG and VG variance by average OD volume, and checks relationship between these value and location of OD pairs. The analyzed OD pairs are focused on 117 OD pairs which 5 minutes weekday's average OD volume is greater than 1.0[veh/5min]. Moreover the condition of entrance and exit of MEX is classified following 3 patterns in this paper.

Table 2 -Categorized Location of Entrance and Exit on MEX

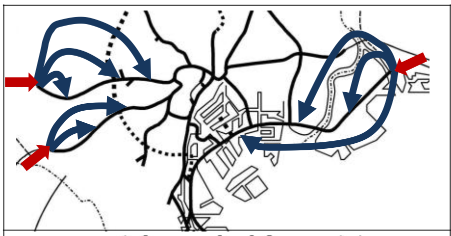
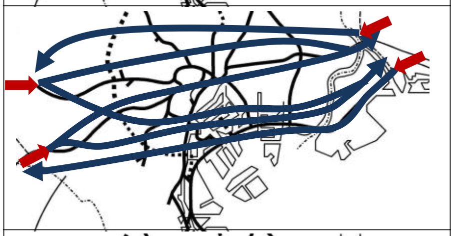
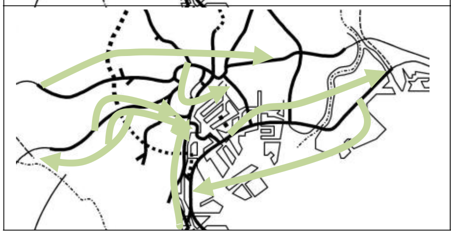
	**OD Group (a) : Short Distance OD pairs** **Entrance :** **Another Expressway** **(Different Toll System)** **Exit :** **MEX Off Ramp** **(Short Distance from Entrance)**
	**OD Group (b) : Through Traffic OD pairs** **Entrance :** **Another Expressway** **(Different Toll System)** **Exit :** **Another Expressway** **(Different Toll System)**
	**Other types of OD pairs in MEX**

Figure 7 shows the relationship between Normalized BG variance and WG variance by different OD location pattern. From the figure, BG and WG variance of Pattern (a) is distributed higher values compare to other patterns e.g. WG variance is plotted from 1.1[veh] to 2[veh], and 15[veh] to 20[veh] in case of BG variance. That means fluctuation of Pattern (a)'s OD has high level of daily and hourly fluctuation every day. In this characteristic, it can consider the phenomena what entrance of these OD pairs are located at mainline tollgate. Since MEX users can check the traffic condition of MEX though the traffic information board at the entrance when they enter from another expressway to MEX as shown on Figure 8, it can guesses MEX users are deciding whether he/she should keep continue to travel on MEX, or get off from MEX. By the results of the decision making by many MEX users, it is influencing to fluctuation of OD volume variation for these OD pairs.

On the other hand, characteristics of OD pattern (b) have opposite tendency against to OD pattern (d). Values of BG and WG variance are 1[veh] to 1.3[veh] for WG variance and BG variance is plotted around 5[veh]. So values of both variance are low level particularly BG variance against to the other OD patterns. This reason can say that OD pairs which are travelled thought the MEX is constantly used every day and every time. For other OD patterns, there is difference of value of BG variance, however WG variance is plotted around 0.8[veh] to around 1.4[veh] and it can say that range of distribution of WG variance is smaller than the one of BG variance.

By above results, it could show the characteristics of Daily and Hourly OD volume variation by calculating the BG variance as daily OD volume variation and WG variance as hourly OD volume variation. In addition it also could be identified location of OD pairs by relationship between BG and WG variation, and it showed that OD pairs which are short distance entering from another expressway are highly varying.

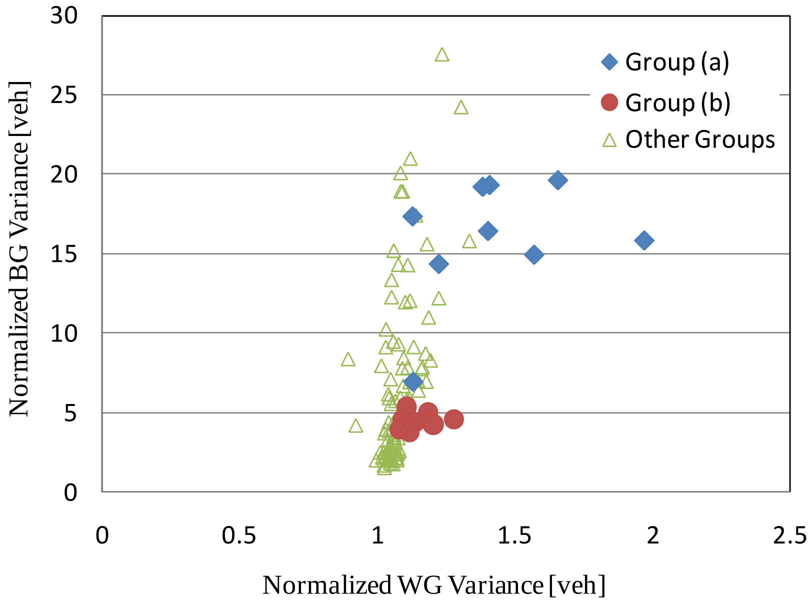


Figure 7 – Relationship between Normalized WG and BG Variance by OD Location Pattern

Figure 8 – Travel Time Information Providing on MEX

4 SHORT-TERM OD VOLUME PREDICTION

4.1 Outline of Proposed OD Volume Prediction Model

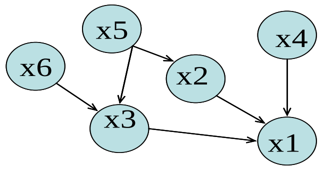
4.1.1 Bayesian Network Model

In Chapter 3, characteristics of OD volume variation are described by BG and WG variance. Through the discussion on chapter 3, characteristics of OD volume variation is different by location of OD pairs and difference of time and day. However, average or one day representative OD volume is mainly used for DTM scheme. Therefore, stochastic short-term OD prediction model is proposed for consider the traffic condition at prediction timing, and it is also considered that output not only definite prediction value but also range of possible OD volume by stochastic model. To develop the model for satisfying above condition, Bayesian Network is applied in this research.

Bayesian Network is the stochastic model by using acyclic network structure as shown at Equation (13) and Figure 11. Equation (13) shows the joint probability distribution of Bayesian Network which $\{X_1, X_2, \ldots, X_n\}$ are probability variables on the Bayesian Network and group of parent nodes $Pa(X_i)$ that are used $\{X_1, X_2, \ldots, X_n\}$ as sun node. In addition, the $P\{X_i \mid Pa(X_i)\}$ describes conditional probability decided by structure of parent nodes (Motomura and Iwasaki. 2006).

$$P(X_1, X_2, \ldots X_3) = \prod_{i=1}^{n} P\{X_i \mid Pa(X_i)\} \tag{13}$$

Figure 9 shows an example of Bayesian Network structure and conditional probability table (CPT) which is output from Bayesian Network. Conditional Probability is calculated based on condition of variables which have causal relationship. BN consist by son node and parent node, and parent node's conditional probability is calculated by condition of their son node. For example, if we want to estimate CPT of x1, it depends on condition of x2, x3 and x4 as CPT in figure 8. This Bayesian inference method have already proposed for OD volume or traffic flow prediction by Sun et al. (2006), Castillo et al. (2008) and Hazelton(2008). However there are no studies for OD volume prediction by Bayesian Network using real OD data.



CPT of $P(x1 \mid x2, x3, x4)$

x2		0				1			
x3		0		1		0		1	
x4		0	1	0	1	0	1	1	1
x1	0	0.2	0.3	0.1	0.15	0.6	0.65	0.7	0.8
	1	0.8	0.7	0.9	0.85	0.4	0.35	0.3	0.2

CPT: Conditional Probability Table

Figure 9 – Example of Bayesian Network Structure and Conditional Probability Table

The software called Unbbayes (http://unbbayes.sourceforge.net/) is used for estimate CPT. Unbbayes is estimating CPT based on frequency of data depending on the network structure as shown at equation (14). Depending on the learning results, CPT is estimated.

$$\hat{\theta}_{ijk} = \frac{N'_{ijk} + N_{ijk}}{N'_{ij} + N_{ij}} \tag{14}$$

$$N'_{ij} = \sum_k N'_{ijk} \tag{15}$$

$$N_{ij} = \sum_k N_{ijk} \tag{16}$$

where,

$\hat{\theta}_{ijk}$: Estimated conditional probability of sun node x_i take *k* when set of parent node $Pa(X_i)$ is pattern *j*

N'_{ijk} : Hyper parameter for decide prior probability before data learning

N_{ijk} : Frequency of data when sun node variable x_i take *k*th value and set of parent node $Pa(X_i)$ against to sun node x_i is pattern *j*

4.1.2 Concept of Proposed OD Volume Prediction Model

Proposed OD volume prediction model basically consider the imminent traffic condition at prediction timing. Therefore, the model consist of time of day, and ratio of 30 minutes OD volume against to their average value at last 5 minutes as explanatory variables, and then OD volume at 30 minutes further is predicted depending on the condition of explanatory variables as shown on Figure 10. That means predicted index is not directly OD volume, but ratio of OD volume against to its average from Bayesian Network under the assumption what average OD volume is updating by record of ETC-OD data and we have already known when OD volume is predicted.

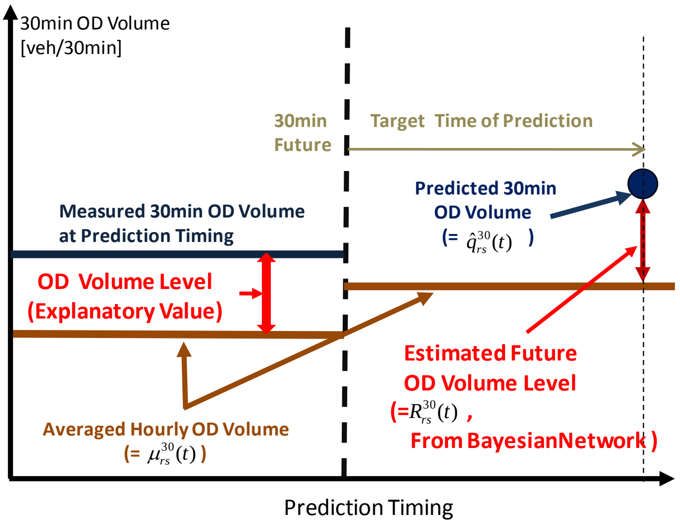


Figure 10 –Concept of OD Volume Prediction Model

According to above concept, the proposed model predicts not value of OD volume itself, but ratio of OD volume against to the average value. Therefore, predicted OD volume is calculated by equation (17)

$$\hat{q}_{rs}^{30}(t) = R_{rs}^{30}(t)\mu_{rs}^{30}(t) \qquad (17)$$

where,

$\hat{q}_{rs}^{30}(t)$: Predicted 30 minute OD Volume of OD pair *rs* at Time *t*

$R_{rs}^{30}(t)$: Ratio of Predicted 30 minute OD Volume of OD pair *rs* against to historical average at Time *t* which is estimated from Bayesian Network

By following concept of OD prediction model, the structure of Bayesian network is described as Figure 11. From the figure, $R_{rs}^{30}(t)$ is chosen from category of son node "Predicted OD Volume Level" which shows higher conditional probability depending on condition of parent nodes.

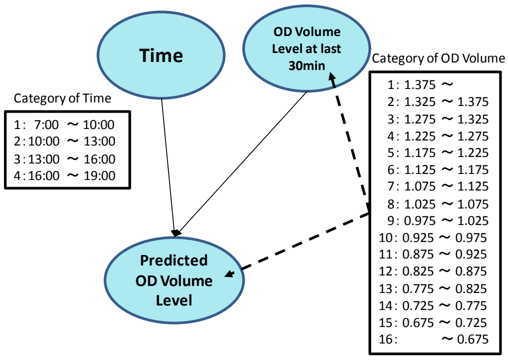


Figure 11 – The Structure of Bayesian Network for OD Volume Prediction

4.1.3 Evaluation of Proposed OD Volume Prediction Model

The Accuracy of proposed OD volume prediction model is evaluated by Rout Mean Square Error (RMSE) as expressed at Equation (18) to (20). Weekday's historical average OD volume and 30min shifted measured OD volume is used as criterion of model evaluation here.

$$RMSE_{rs}^{\mathrm{Pred}} = \sqrt{\frac{\sum_{t=1}^{Num}(\hat{q}_{rs}^{30}(t) - q_{rs}^{30}(t))^2}{T}} \quad (18)$$

$$RMSE_{rs}^{Ave} = \sqrt{\frac{\sum_{t=1}^{Num}(\mu_{rs}^{30}(t) - q_{rs}^{30}(t))^2}{T}} \quad (19)$$

$$RMSE_{rs}^{TimeShift} = \sqrt{\frac{\sum_{t=1}^{Num}(q_{rs}^{30}(t-6) - q_{rs}^{30}(t))^2}{T}} \quad (20)$$

where,

$RMSE_{rs}^{\mathrm{Pred}}$: RMSE between Predicted OD Volume and Measured OD Volume of OD pair *rs*

$RMSE_{rs}^{Ave}$: RMSE between Average OD Volume and Measured OD Volume of OD pair *rs*

$RMSE_{rs}^{TimeShift}$: RMSE between 30 minutes Time Shifted Measured OD Volume and Measured OD Volume of OD pair *rs*

$q_{rs}^{30}(t)$: Measured 30minute OD Volume of OD pair *rs* at Time *t*

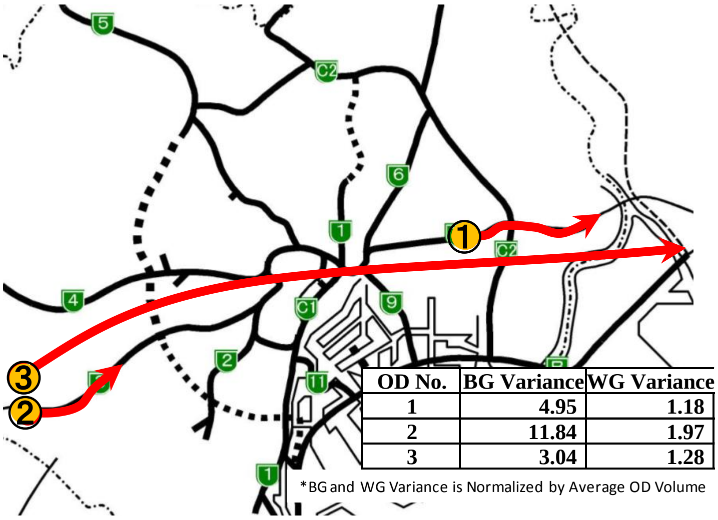
t : Prediction Timing (Each 5 minutes in this study)

T : Number of Prediction Timing

4.2 Accuracy of Proposed OD V4olume Prediction Model

4.2.1 Target OD pairs for Tested the OD Prediction Model and Learned Data

OD prediction model by Bayesian Network is tested by three OD pairs on MEX network. Tested OD pairs' location is shown on Figure 12, and is also described the characteristics of BG and WG variance for target OD pairs.



OD No.	BG Variance	WG Variance
1	4.95	1.18
2	11.84	1.97
3	3.04	1.28

*BG and WG Variance is Normalized by Average OD Volume

Figure 12 – Target OD Pairs and their WG and BG Variance

ETC-OD Data collected 147 weekdays (June 2006 to February 2007) is learned by Bayesian Network, and 30 minutes of OD volume of 10 days of weekday at March 2007 are predicted. In addition, time of day is categorized to 4 groups by each 3 hours in 12 hours (7:00 to 19:00). And OD volume at 30 minutes further is predicted for target OD pairs.

4.2.2 Accuracy of OD prediction model focusing on difference of characteristics of OD volume variation

RMSE of target OD pairs are showed on Table 2 as the results of prediction for 30 minutes further. When we focus on OD2 which is the highest value of the variances, RMSE by prediction results are lower than the one from historical average OD Volume. On the other hand, OD3 which is the lowest value of the variances but long distance OD pair, average OD volume can be accurate predicted value, but still proposed model can make it calculate better results than historical average OD volume. Figure 13 shows the gap of RMSE between Prediction and Averaged OD volume. From the figure, it can see OD2 could improve on 9 days, and even OD3 which is lowest value of variance improved 5 days. Therefore it could be shown proposed model can improve OD volume prediction accuracy rather than using historical average OD volume as prediction value.

Table 2 - RMSE of Predicted and Averaged OD Volume

Prediction Date		2007/3/5	2007/3/6	2007/3/7	2007/3/8	2007/3/9	2007/3/12	2007/3/13	2007/3/14	2007/3/15	2007/3/16
OD 1	Prediction	***12.191***	10.306	***12.341***	***13.767***	***13.413***	***14.798***	***15.084***	***12.218***	***13.535***	***14.893***
	Average	***13.623***	9.981	***13.018***	***15.333***	***17.010***	***18.202***	***15.844***	***13.725***	***16.593***	***18.566***
	Time Shift	***17.046***	16.455	***16.455***	***21.550***	***21.651***	***21.559***	***24.777***	***20.482***	***19.576***	***22.755***
OD 2	Prediction	***18.494***	***18.012***	***25.868***	***24.105***	***22.351***	***20.502***	***19.410***	***22.750***	***21.280***	22.679
	Average	***28.923***	***23.423***	***28.760***	***35.896***	***21.370***	***24.022***	***21.804***	***30.243***	***25.183***	22.070
	Time Shift	***18.599***	***25.365***	***30.161***	***29.259***	***27.866***	***21.895***	***23.582***	***29.216***	***26.726***	26.726
OD 3	Prediction	10.259	9.102	***8.470***	7.860	6.806	***9.277***	***7.716***	***10.178***	***7.816***	5.982
	Average	13.351	8.463	***9.124***	7.548	6.439	***10.213***	***9.029***	***11.405***	***7.853***	5.107
	Time Shift	10.173	10.606	***10.058***	10.242	9.198	***11.261***	***8.736***	***11.428***	***9.843***	8.791

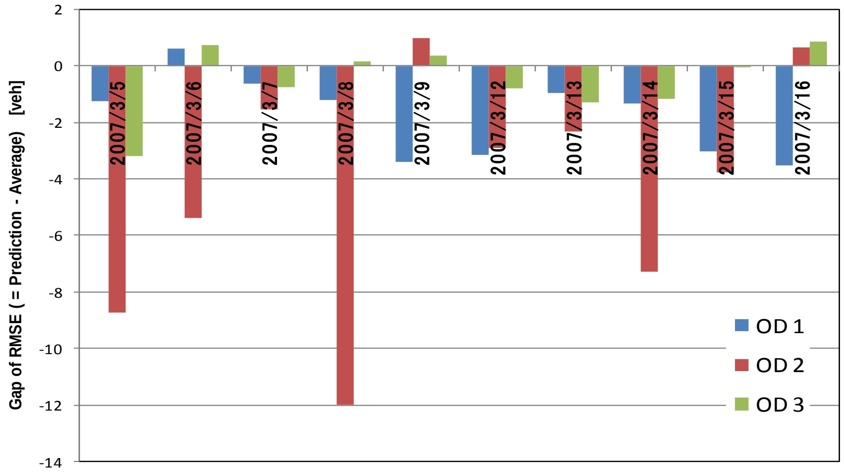


Figure 13 – Comparison of Root Mean Square Error by different OD pairs

(Comparison RMSE between Predicted value and Averaged value)

Figure 14 shows comparison of profile of OD Volumes at OD2 by actual value, average value and predicted value. From the figure, OD profile of predicted value is similarly varying with actual value, rather than the profile of average value.

On the other hand, proposed model is introduced ratio of OD volume at the latest time period, therefore it is concerned about validity to use this index as explanatory variables. However, OD pairs which has higher value of variation and higher improvements of RMSE is shorter distance compared to other OD location pattern. Therefore it is possible to use proposed OD volume prediction model for DTM scheme on MEX.

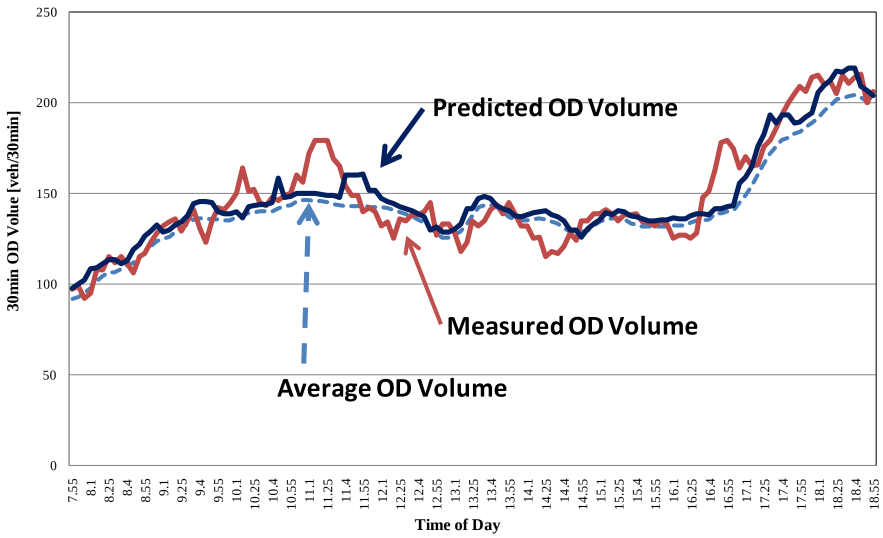


Figure 14 – Comparison of OD Volume Profile in case of 1day of OD1

5 CONCLUSION

This paper analyzed the characteristics of OD volume variation and it is confirmed tendency of prediction accuracy is different by depending on the characteristics of OD volume variation using OD prediction model by Bayesian Network.

In OD volume variation analysis, two types of variance were used to describe characteristics of OD variation. One is Between-Group Variance as daily OD variation, and another one is Within-Group Variance to describe the variation in the hour. Relationship between BG variance and WG variance shows effect of daily OD variation is much higher than effect of hourly OD volume variation in MEX. It is especially prominent in OD pairs which have higher level of WG variance. In addition, it also showed that characteristics of OD variation could identify by pattern of OD location on MEX. Especially short distance OD pair which entering from another Expressway connected to MEX has higher variance. The reason of characteristics has considered MEX users are deciding whether they should continue to use MEX or not according to MEX traffic condition. On the other hand, long distance OD pairs which are just through MEX network are always almost traffic is travelling constantly. Therefore average OD volume of such OD pairs can be a certain level of accurate prediction by historical average OD volume. However, other OD pairs, which daily variations are much bigger, OD volume prediction should be considered the real time traffic condition.

Stochastic short-term OD prediction model was proposed using Bayesian Network considering condition of OD volume at prediction timing, and OD data during half year is learned to proposed Bayesian Network model. And the accuracy of proposed model is validated by RMSE and it was compared with RMSE of historical average OD volume. From the results of accuracy of predicted OD volume for OD pairs which higher OD variation could has improved by proposed model since OD volume is highly varying depending on change of traffic condition momentarily. On the other hand, in case of OD pairs which are low variation could also improve its accuracy, although historical average had already suitable accuracy.

However, target OD pairs are only focused on higher level of OD volume compared to other MEX OD pairs. Therefore, analyzing another OD pair which has low level of OD volume is recommended since number of such kind of OD is huge amount against to number of higher level of OD pairs. And test of the performance of traffic control scheme by applying proposed model is also recommended as future works.

ACKNOWLEDGMENT

We would like to thank to Metropolitan Expressway Co. Ltd to give us opportunity to use ETC-OD data. We also would like to appreciate to Prof. Toshio Yoshii at Ehime University for his valuable comments to this research.

REFERENCES

Asakura Y., Hato E. and Kashiwadani M. (2000). Origin-destination matrices estimation model using automatic vehicle identification data and its application to the Han-Shin expressway network. Transportation, Vol.27, 419-438

Ashok K. and Ben-Akiva M.E. (1993). Dynamic Origin-Destination Matrix Estimation and Prediction for Real-Time Traffic Management System. C.F.Daganzo, Editor, International Symposium on Transportation and Traffic Theory, 465-484

Camus R., Cantarella G.E. and Inaudi D. (1997). Real-time estimation and prediction of origin-destination matrices per time slice. International Journal of Forecasting, Vol.13, 13-19

Castillo E., Menendez J. M. and Santos Sanchez-Cambronero (2008). Predicting traffic flow using Bayesian networks. Transportation Research Part B, 42, 482-509

Fumitaka Kurauchi, Shinei Kim, Nobuhiro Uno and Teruhisa Ishibashi (2008). Time Dependent OD Volume Estimation using ETC Statistical Data. Proceeding of Conference of Japan Society of Traffic Engineer, Vol. 28, 205-208 (In Japanese)

Hazelton M. L. (2008). Statistical inference for time varying origin-destination matrices. Transportation Research Part B, Vol.42, pp452-552

Kachroo P., Ozbay K., and Narayanan A. (1997). Investigating the Use of Kalman Filtering Approaches for Dynamic Origin-Destination Trip Table Estimation. Proceeding of Southeastcon '97 of IEEE, 138-142

Kwon J. and Varaiya P. (2005). Real-Time Estimation of Origin-Destination Matrices with Partial Trajectories from Electronic Toll Collection Tag Data. Transportation Research Record 1923, 119-126

Motomura Y. and Iwasaki H.(2006). Bayesian Network Technology (In Japanese)

Okuani I. (1987). The Kalman Filtering Approaches in Some Transportation and Traffic Problems. In N.H. Gartner and N.H.M. Wilson, editors, International Symposium on Transportation and Traffic Theory, 397-416

Rakha H. and Aerde M. V. (1995). Statistical Analysis of Day-to-Day Variation in Real-Time Traffic Flow Data. Transportation Research Record 1510, 26-34

Shiraishi T., Kuwahara M. and Horiguchi R. (2004) A Development of Traffic Prediction System Based on Real-time Simulation. Proceeding of 30th Conference on Infrastructure Planning and Management, CD-ROM (In Japanese)

Stathopoulos, A. and Karlaftis, M. (2001). Temporal and Spatial Variations of Real-Time Traffic Data in Urban Areas. Transportation Research Record 1768, 135-140

Sun S., Zhang C. and Yu G. (2006). A Bayesian Network Approach to Traffic Flow Forecasting. IEEE Transactions on Intelligent Transportation Systems, Vol.7, No.1, 124-132

Tanabe A., Daito T., Sugiyama M. and Asakura Y. (1995). Real Time Estimation of OD Matrix on Hanshin Expressway from Traffic Counts and AVI data. Proceeding of 18th Conference on Infrastructure Planning and Management, Vol.2, pp193-196 (In Japanese)

Weijermars W. and Berkum E. V. (2005). Analyzing highway flow patterns using cluster analysis. Proceeding of the 8th International IEEE Conference on Intelligent Transportation System, 831-836

Yong L., YuanLi C. and YongXuan H. (2003). Kalman Filtering Based Dynamic OD Matrix Estimation and Prediction for Traffic Systems. Proceeding of Intelligent Transportation Systems of IEEE, Vol. 2, 1515-1520

Zhou X. and Mahmassani H. S. (2007). A Structural state space model for real-time traffic origin-destination demand estimation and prediction in a day-to-day learning framework. Transportation Research Part B, Vol. 41, 823-840