Proceedings of the Eastern Asia Society for Transportation Studies, Vol.12,2019

A Study on the Characteristics of MFD determined by Probe Data

Stephanie Wanjiku GITURU[a], Toshio YOSHII[b], Takahiro TSUBOTA[c], Hirotoshi SHIRAYANAGI[d]

[a,b,c,d] *Graduate School of Science and Engineering, Ehime University, Ehime, 790-8577, Japan*
[a] *E-mail: Stephanie.gituru.18@cee.ehime-u.ac.jp*
[b] *E-mail: yoshii@cee.ehime-u.ac.jp*
[c] *E-mail: t.tsubota@cee.ehime-u.ac.jp*
[d] *E-mail: shirayanagi@cee.ehime-u.ac.jp*

Abstract: This paper investigates the characteristics of the macroscopic traffic states represented by the Macroscopic Fundamental Diagram (MFD). Factors that have influence on the MFD shapes have been intensively reported. However, comprehensive study on the MFD shapes from various types of cities has not been investigated. This study investigates the characteristics of the MFD shapes from 47 cities in Japan. The MFDs are defined by the nationwide probe vehicle system in Japan. The analysis is performed in the 47 prefectural capitals. The MFDs are obtained by performing a piecewise linear regression of the area states consisting of standardized traffic flow and density. The results of the k-means method identified 4 typical shapes of MFDs, indicating that the MFD shapes could be affected by types of cities. Future research attempt includes understanding how the traffic characteristics and road network structures in the 47 prefectural capitals are affecting the shape of their MFDs.

Keywords: Macroscopic Fundamental Diagram, Probe Data, Piecewise Linear Regression

1. INTRODUCTION

Traffic congestion and other challenges such as traffic accidents are issues experienced by cities all around the world. In order for the management of these challenges, an understanding of the traffic conditions is paramount. The Macroscopic Fundamental Diagram (MFD) has been proposed as a method to grasp the network traffic conditions. The MFD, proposed by Daganzo(2007), represents network-wide traffic states by relating the traffic throughput at given density levels. Thus, the MFD brings about a comprehensive understanding of network traffic states at an aggregate level. Since the existence of the MFD in a real urban network was empirically revealed by Daganzo and Geroliminis (2008), many research attempts have been made to understand the property of the MFD shapes (Geroliminis and Sun, 2011; Zhang *et al. (*2013); Gayah *et al.* (2014)) and to develop control strategies based on the MFD concept Aboudolas and Geroliminis (2013).

The greatest challenge with empirical MFD is obtaining sufficient data. Among researchers, one of the most common data sources for the MFD is the use of detector data; however, these may not be fully representative as they depend on where the detectors are placed on road sections. Courbon and Leclercq (2011) analyzed the impact of loop detector positioning and found out that, for the MFD to be properly determined, the detectors should be spread and well positioned to capture the different traffic situations; otherwise the resulting MFD will be significantly biased. Another promising data source is probe vehicle data. Advancement in GPS technology has made the vehicle movement data easily available in urban networks. Research has shown that the MFDs for a network can be estimated by the use of trajectory data from probe vehicles (Nagle and Gayah(2013)).

Understanding the characteristics of the MFD is crucial for traffic management strategies. Various studies examining the factors that influence the characteristics of MFDs have been undertaken. The spatial variability of vehicle density has been found to affect the shape, scatter and the existence of the MFD by Geroliminis and Sun(2011). The shape of the MFD has also been found to depend on the particular signal system used and the level of heterogeneity in the system (Zhang *et al* (2013), Gayah *et al.* (2014)). However, there is not much information on how the shape of the MFD is affected by the type of the city such as mega or rural cities and their networks.

Clustering techniques have been applied in various studies to group different kind of traffic data. Clustering has been used to group regions with very high traffic congestion based on speed(Liu *et al.* (2010)). Traffic flow has been previously classified into five different states using K-means clustering; and studied to identify the reasons as to why some states are more susceptible to crashes than others (Xu *et al.* (2012)). Non-recurrent congestion occurrence was identified using clustering analysis based on space and time(Anbaroglu *et al.* (2014)). In this research clustering analysis is applied to group MFDs with similar shapes together.

The main purpose of this study is to understand the characteristics of the MFD shapes in different cities. The study is conducted using the probe vehicle data from 47 prefectural capitals in Japan. The analysis of the MFD is conducted using a piecewise linear regression model. The results of the analysis are then employed to cluster the cities based on the characteristics of the MFD shapes.

The remaining part of this paper is organized as follows; section 2 describes the data and the definition of the MFD. Section 3 details the analysis based on a piecewise regression model, followed by the clustering analysis based on the MFD shapes. Section 4, outlines and discuss the analysis results, and section 5 gives the concluding remarks.

2. DATA DESCRIPTION

2.1 Probe Vehicle Data

This study utilizes the nationwide probe vehicle data from ETC 2.0 system to derive the MFD. The system is a form of Intelligent Transport System (ITS), that collect data from on board units in vehicles using antennas via dedicated short-range communication (DRSC). The antennas are approximately 1,600 units all over Japan and are mostly situated on expressways and major roads. The location of these antennas tends to create a bias in the sampling as short distance trips remaining on arterials are acquired in small amounts compared to long distance trips (Yasuda *et al.* (2011)). The system provides with the trajectory data of vehicles equipped with the ETC 2.0 onboard unit. The unit stores the data approximately every 200-meter drive, and it includes the timestamp and vehicle location such as latitude and longitude. The distance covered by the equipped vehicles in a study area is calculated as the linear distance between two consecutive points from the location data, and the time spent by the vehicle calculated as the time difference of two timestamps. This study utilizes the data aggregated in an hour and collected over a year from April 2015 to March 2016.

2.2 Definition of the MFD

This study defines the MFD using the probe vehicle data described in 2.1. The MFD is the relationship of area-wide traffic flow and density, and they are equivalent to the total vehicle kilometers and the total vehicle hours within an area, respectively, as defined by Edie L.C (1963). Equation (1) and (2) defines the variables from the probe vehicle data. The time interval t is one hour.

$$Q^{m,d,t} = \sum_{i \in I^{m,d,t}} q_i^{m,d,t} \tag{1}$$

$$K^{m,d,t} = \sum_{i \in I^{m,d,t}} k_i^{m,d,t} \tag{2}$$

where,

$Q^{m,d,t}$:total vehicle kilometers in an area in month m on day d at time interval t,
$K^{m,d,t}$:total vehicle hours in an area in month m on day d at time interval t,
$q_i^{m,d,t}$:distance traveled in an area by vehicle i in month m on day d at time interval t,
$k_i^{m,d,t}$:time spent in an area by vehicle i in month m on day d at time interval t, and
$I^{m,d,t}$:set of vehicles in an area in month m on day d at time interval t.

2.3 Standardized MFD

A challenge faces when analyzing the ETC2.0 probe vehicle data over a year, because the number of equipped vehicles is rapidly growing; the number increases from 607,601 [veh] in April 2015 to 1,019,809 [veh] in March 2016, over 60% increase during the study period. Therefore the number of probe samples is also increasing every month; hence it is necessary to standardize the variables of the MFD, $Q^{m,d,t}$ and $K^{m,d,t}$ considering the growth in sample size. The standardized variables, $stQ^{m,d,t}$ and $stK^{m,d,t}$ are calculated as Equation (3) - (6).

$$stQ^{m,d,t} = Q^{m,d,t} / Q^m \tag{3}$$

$$stK^{m,d,t} = K^{m,d,t} / K^m \tag{4}$$

$$Q^m = \sum_{d=1}^{D^m} \sum_{t=1}^{T} Q^{m,d,t} / (D^m \cdot T) \tag{5}$$

$$K^m = \sum_{d=1}^{D^m} \sum_{t=1}^{T} K^{m,d,t} / (D^m \cdot T) \tag{6}$$

where,

Q^m :average vehicle kilometers during a unit time interval in month m, and
K^m :average vehicle hours during a unit time interval in month m.
D^m :number of days in month m.
T :number of time intervals within a day (i.e., $T = 24$).

Hereafter, the standardized total vehicle kilometers, $stQ^{m,d,t}$, and the standardized total vehicles hours, $stK^{m,d,t}$, are denoted as flow and density, respectively for simplicity. As well, the term “MFD” represents the relationship between the standardized variables, $stQ^{m,d,t}$ and $stK^{m,d,t}$, unless otherwise noted.

3. CHARACTERISTICS OF THE MFD

3.1 Piecewise Linear Regression Modeling for the MFD

The MFD takes non-linear form with different ranges of density in line with the change in flow, hence different linear relationships could occur. This study applies a piecewise linear regression (PLR) model to characterize the MFD shape. The PLR portions the independent variable into intervals, and a separate line segment is assigned to each interval. The boundaries between the segments are known as thresholds or breakpoints (Al-sobky and Ramadan, 2015).

The PLR is used to fit regression lines based on the shape of the MFD. The existence of breakpoints on the MFD plots is assumed to represent the change in traffic states. Figure 1 illustrates the piecewise linear approximation of the MFD; the shape is fitted by three-line segments with two breakpoints. The slopes of each line, β_1, β_2 and β_3, represent the shockwave speed in each phase. When density is smaller than the first breakpoint P_1, the network is in phase1, where the traffic is assumed to be free flow because the network density is smaller. Once density exceeds the P_1, the slope of the MFD changes to β_2, and the traffic state moves to Phase2, in which the network is slightly congested. In the traffic state with density over P_2, the network is in the most congested state (Phase3). This study anticipates that the different urban network will take unique form of the MFD with different number of breakpoints and slopes. Therefore, the MFD shapes are simplified based on the PLR model, and the set of the parameters will be utilized for clustering analysis later in this study.

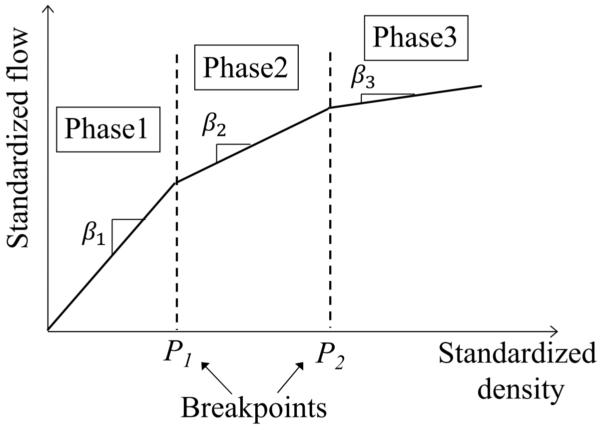

Figure 1. Piecewise linear approximation of the MFD and parameters

The slope parameters in the PLR are estimated using Equation 7. Multiple linear regression analysis is performed for a given set of breakpoints, P_1 and P_2, to estimate the slopes of line segments, β_1, β_2 and β_3. The most suitable set of breakpoints are searched iteratively as described in 3.2,

$$stQ^{m,d,t} = \beta_1\left(stK^{m,d,t}d_1 + P_1 d_2 + P_1 d_3\right) + \beta_2\left((stK^{m,d,t} - P_1)d_2 + (P_2 - P_1)d_3\right) + \beta_3(stK^{m,d,t} - P_2)d_3 \quad (7)$$

where,

P_1, P_2 : breakpoints ($P_1 \le P_2$)

d_1 : Dummy variable (1 if $stK^{m,d,t} \le P_1$, otherwise 0)

d_2 : Dummy variable (1 if $P_1 < stK^{m,d,t} \le P_2$, otherwise 0)

d_3 : Dummy variable (1 if $stK^{m,d,t} > P_2$, otherwise 0)

$\beta_1, \beta_2, \beta_3$: Slopes of line segments

3.2 Break-point Determination

The breakpoints on the MFD are defined as the points where the traffic state changes. The breakpoints are searched iteratively as illustrated in figure 2. In each iteration, a set of breakpoints are given, and the multiple linear regression (Equation 7) is performed and the coefficient of determination R^2 is calculated. The search is conducted for every combination of breakpoints among the range of density from its minimum, min(*stK*), to the maximum, max(*stK*); The set of breakpoints that results in the highest R^2 is determined as the most suitable set. In case that the highest R^2 is obtained when the two breakpoints are the same (i.e., P_1=P_2), the MFD is approximated as the PLR model with one breakpoint; otherwise with two breakpoints.

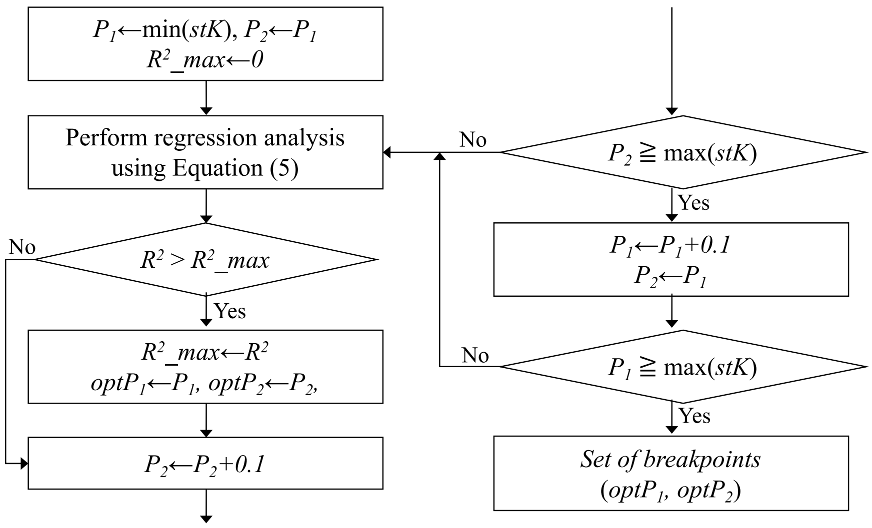

Figure 2. Iterative search for breakpoints

3.3 Clustering based on the K-Means Method

The MFD shapes are grouped based on the parameters estimated from the PLR model detailed in 3.2. The k-means method is applied for the clustering (MacQueen, 1967). The method partitions the group of data $\{\boldsymbol{x_1}, \ldots, \boldsymbol{x_N}\}$ into K clusters. Each data $\boldsymbol{x}$ is a vector to characterize the sample, which corresponds to the set of parameters of the PLR model, $\beta_1, \beta_2, \beta_3, P_1$ and P_2.

K-Means clustering is a distance-based clustering algorithm that divides data into several clusters in numerical attributes. First, the number of clusters K and the number of maximum iterations is decided. Initially, the algorithm begins with estimates of the cluster centroids which can be randomly generated or chosen from the data set. The algorithm then iterates between two steps which are data assignment step and centroid update step. In the data assignment step, each centroid defines each of the clusters. All data points are assigned to their nearest centroid in order to minimize the intra-cluster distance. The objective function J is formulated as Equation 8, and gives an algorithm that minimizes the squared distance between each center and its assigned data points (Hamerly and Elkan, 2002).

$$J = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{n,k} \|\boldsymbol{x_n} - \boldsymbol{c_k}\|^2 \quad (8)$$

where,

$r_{n,k}$:if data $\boldsymbol{x_n}$ belongs to cluster k, $r_{n,k} = 1$; otherwise $r_{n,k} = 0$, and

$\boldsymbol{c_k}$:centroid of cluster k.

The centroids are recalculated in the centroid update step. This is done by taking the mean of all data points assigned to that centroid's cluster equation 9.

$$c_k = \frac{1}{|s_k|}\sum_{x_n \in s_k} x_n \tag{9}$$

where,

s_k :set of data point assignments for each cluster k

The algorithm iterates between these steps until a stopping criterion is met which could be, no data points change clusters, the sum of the distances is minimized, or the maximum number of iterations is reached. The number of clusters K should be predefined, thus a method is required to determine the suitable number of clusters. The elbow method is often utilized to discuss the appropriate number of clusters. The method applies k-means clustering on a dataset for a range of K (e.g., K from 1 to 8), and the sum of squared error (SSE) is calculated for each K. The SSE for K cluster $SSE(K)$ is calculated as Equation 10 for each value of K.

$$SSE(K) = \sum_{n=1}^{N}\sum_{k=1}^{K} r_{n,k}\|\boldsymbol{x_n} - \boldsymbol{c_k^*}\|^2 \tag{10}$$

where,

$\boldsymbol{c_k^*}$:optimal centroid of cluster k.

The $SSE(K)$ is the optimized value of the objective function J. The elbow method plots the relationship between the number of cluster K and the $SSE(K)$. The SSE(K) tends to decrease towards 0 as the number of clusters K increase. A small value of K that has a low SSE(K) is chosen and the elbow represents where diminishing returns by increasing K begin; thus, one can examine the appropriate number of cluster K.

4. ANALYSIS OF THE MFD CHARACTERISTICS

4.1 Study Area

The analysis was carried out in the 47 prefectural capitals in Japan. The study focuses on the central business districts (CBD) of the capital cities; an area of approximately 9km square is selected considering the sizes of the CBD.

4.2 Characteristics of MFD

The flow and density are calculated for each of the 47 study areas according to Equation 3 and 4. In order to examine the influence of the day of the week, the MFD is calculated for weekdays and holidays. Thus, 94 diagrams are available in total.

The PLR model, as detailed in 3.1 and 3.2, is applied to the MFDs. Estimation results from 4 cities are presented in Table 1 as an example, which summarizes the breakpoints and slopes of each segment in the PLR line as well as the adjusted R^2 and sample size. The full results from 47 cities are summarized in the Appendix. The corresponding MFD plots as well as the PLR lines, are also presented. The results show that the MFDs from 3 cities have two breakpoints (Figure 3, 5 and 6), whereas the diagram from Ehime has one break point (Figure 4). Further, the values of the breakpoints are different; for instance, the first breakpoint P_1 is larger in Tokyo and Ehime than in Saitama and Fukui. The difference in the number and the values of breakpoints suggest that the traffic states could be more stable against the increase of density in some cities.

When looking at the values of slopes, the MFD from Tokyo has a negative β_3, whereas the slopes are all positive in other cities. This could suggest that the majority of the road network in Tokyo is heavily congested when density is higher; on the other hand, other cities do not experience serious breakdown in traffic states.

Another noticeable observation is found in the MFD from Fukui weekdays (Figure 6). The slope decreases from β_1 to β_2, but increases from β_2 to β_3, while the slope is anticipated to decreases with an increase in density as the other 3 cities demonstrate (Figure 3, 4, 5). This counterintuitive result of Fukui weekday could be due to the factors that are not considered in the present study. This study utilizes one-year data without considering extreme weather, such as heavy rain and snow. The influence of such factors should be further investigated.

Table 1. Results of the PLR model from 4 cities

	Tokyo weekday		Ehime holiday		Saitama holiday		Fukui weekday	
Variable	Coef.	t value	Coef.	t value	Coef.	t value	Coef.	t value
β_1	1.26	313.0*	1.17	242.4*	1.36	250.9*	1.68	179.0*
β_2	0.57	77.4*	0.68	118.6*	0.57	111.1*	0.33	29.2*
β_3	-0.11	-8.2*	-	-	0.16	5.0*	0.59	67.2*
Adj. R^2		0.98		0.99		0.99		0.98
Sample		5,905		2,880		2,880		5,905
Breakpoints	P_1=0.81	P_2=1.52	P_1=0.9	P_2= N/A	P_1=0.68	P_2=1.96	P_1=0.52	P_2=1.13

*:$p < 0.001$

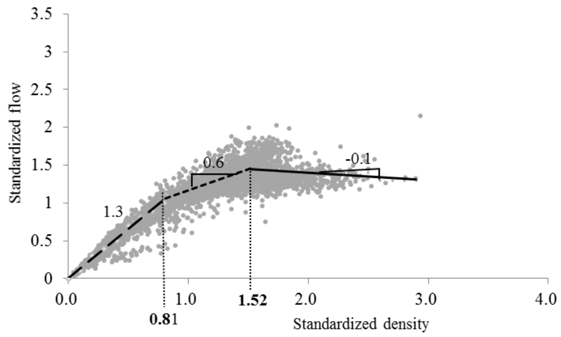

Figure 3. MFD for Tokyo on weekdays

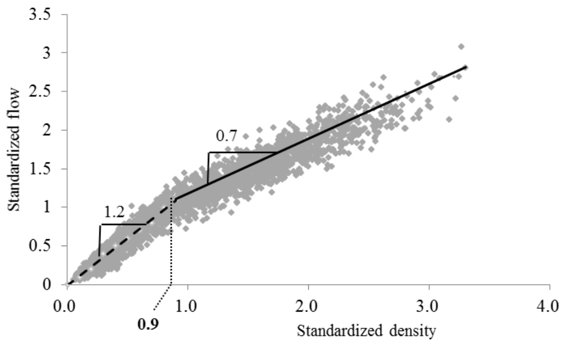

Figure 4. MFD for Ehime on holidays

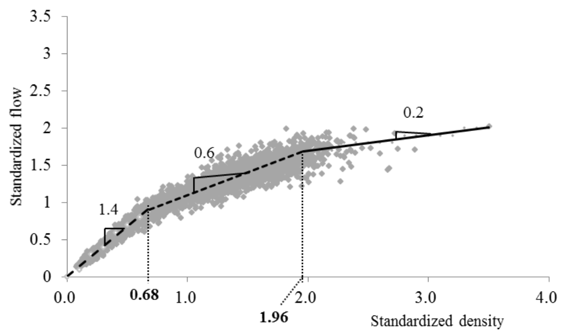

Figure 5. MFD for Saitama on Holidays

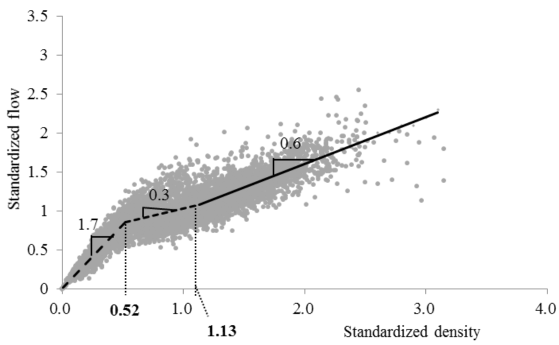

Figure 6. MFD for Fukui on weekdays

4.3 Clustering of the MFDs

The k-means method is applied to the MFDs from 47 cities both in weekdays and holidays. The number of clusters is set to 4 considering the elbow method; the value of $SSE(K)$ does not improve much after K = 4 (Figure 7).

The K-means method identifies four typical shapes of the MFD. The results are summarized in Table 2. For a graphical understanding of the characteristics of each cluster, Figure 8 - 11 present the average MFD shapes from four clusters; the MFDs are defined by using the parameter sets of cluster centroid. Cluster 4 consists of the MFDs with one breakpoint (Figure 11), while Cluster 2, 3 and 4 have two breakpoints.

It is noticeable that the average MFD of Cluster 3 shows negative β_3 (Figure 10) as observed in the MFD from Tokyo weekdays (Figure 3). However, the city of Tokyo does not belong to Cluster 3; instead, it is in Cluster 1 (Table 2). A unique feature of the MFD in Cluster 3 is that the value of second breakpoint P_2 is particularly higher than other clusters. As defined in Equation 4, the density is defined as the ratio of the hourly vehicle hours to the monthly average vehicle hours; Thus, the breakpoint at higher density levels represents the possibility that the network could have experienced severe breakdown under unusually high demand, such as during major events.

Majority of the cities are grouped in Cluster 1 and 2 (Table 2). The average MFDs of both clusters look similar to each other, but the value of second breakpoint P_2 tends to be higher in Cluster 1, implying that the characteristics of the MFD shapes could be further clarified by considering road network structures and traffic demand patterns in each cluster. This study focuses on a nation-wide survey, these groups can be used as a standard for Japanese cities to give an indication of traffic situations. This will, in turn, reduce the need for an intense study to understand the characteristics, as this will be indicated in the group characteristics.

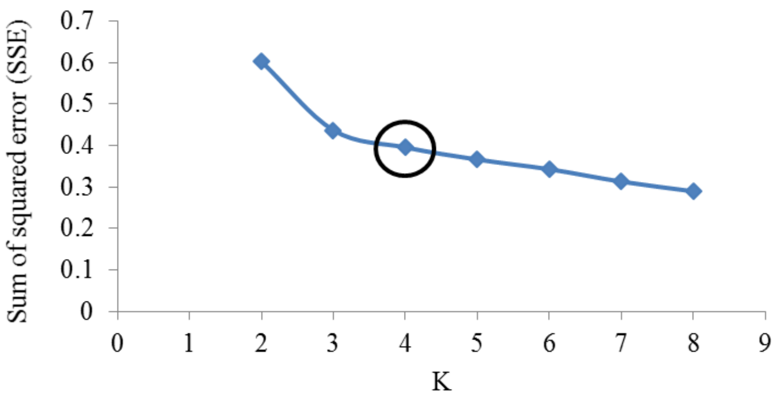

Figure 7. Determination of cluster number K based on the elbow method

Table 2. MFD clustering results

Cluster	Day of week	Prefectures
1	Weekday	Hokkaido, Iwate, Akita, Tokyo, Kanagawa, Niigata, Kyoto, Wakayama, Kagawa, Ehime, Fukuoka, Nagasaki, Miyazaki, Kagoshima, Okinawa.
	Holiday	Aomori, Fukushima, Tochigi, Saitama, Chiba, Tokyo, Shiga, Kyoto, Hyogo, Okinawa.
2	Weekday	Aomori, Miyagi, Yamagata, Tochigi, Ibaraki, Gunma, Saitama, Ishikawa, Fukui, Yamanashi, Nagano, Shizuoka, Mie, Osaka, Nara, Tottori, Shimane, Okayama, Yamaguchi, Saga, Kumamoto.
	Holiday	Hokkaido, Iwate, Miyagi, Akita, Yamagata, Ibaraki, Gunma, Kanagawa, Toyama, Ishikawa, Fukui, Gifu, Shizuoka, Aichi, Mie, Nara, Wakayama, Tottori, Shimane, Hiroshima,

Cluster	Day of week	Prefectures
		Yamaguchi, Tokushima, Kagawa, Kochi, Fukuoka, Saga, Nagasaki, Kumamoto, Oita.
3	Weekday	Fukushima, Chiba, Toyama, Hyogo, Tokushima, Aichi.
	Holiday	Yamanashi.
4	Weekday	Gifu, Shiga, Hiroshima, Kochi, Oita.
	Holiday	Niigata, Nagano, Osaka, Okayama, Ehime, Miyazaki, Kagoshima.

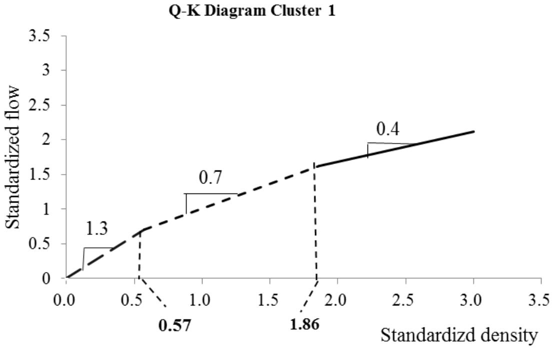

Figure 8.Average MFD for cluster 1

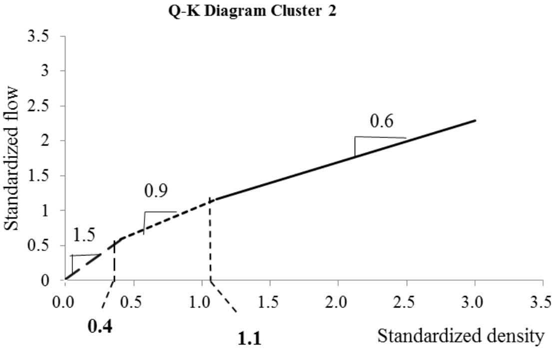

Figure 9.Average MFD for cluster 2

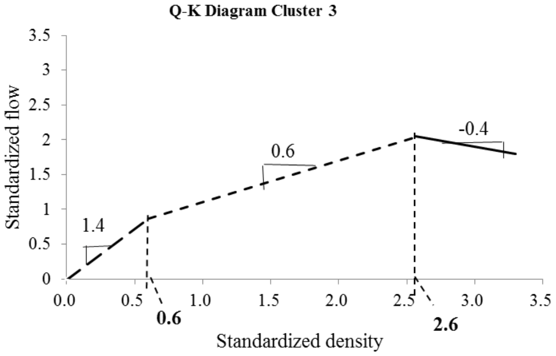

Figure 10.Average MFD for cluster 3

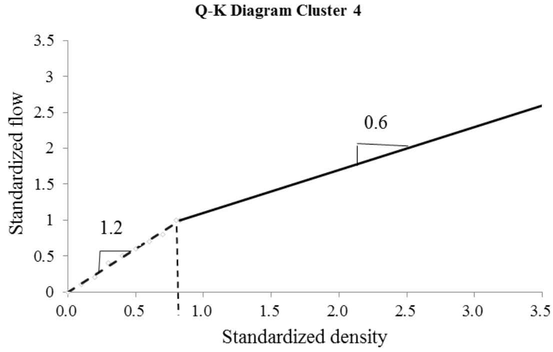

Figure 11.Average MFD for cluster 4

5. CONCLUSION

This research analyzed the MFDs from 47 prefectural capitals in Japan by using the probe data collected by the ETC2.0 system. In order to consider the growth in the sample size of the data, the concept of standardized flow and density was introduced to define the MFD. The piecewise linear regression analysis revealed that the traffic states change as the network density increases. Also, the results suggested that the MFD shapes attributed by the slopes and breakpoints could reflect the characteristics of different cities. The clustering analysis based on the k-means method revealed that the MFDs can be grouped into 4 clusters, each of which takes typical forms of the MFDs in the cluster.

Further research is required to understand how network structure, population density, and the share of taking public transport are affecting the shape of their MFDs. This study used standardized and dimensionless Q-K because it was challenging to obtain detailed sampling rate of the ETC 2.0 vehicles. In future research, flow and density should be estimated from probe data by applying appropriate expansion rate so that the MFD can be obtained in units.

ACKNOWLEDGEMENT

The authors would like to acknowledge the Matsuyama River National Highway Office of Ministry of Land, Infrastructure, Transport and Tourism for their support in providing the ETC2.0 data utilized in this study. We are also grateful to Mr. Shinji Odaka for his support with the analysis and useful comments on the research.

REFERENCES

Aboudolas, K., & Geroliminis, N. (2013). Perimeter and boundary flow control in multi-reservoir heterogeneous networks. *Transportation Research Part B: Methodological*, *55*, 265–281.

Al-sobky, A. A., & Ramadan, I. M. I. (n.d.). *A Generalized Piecewise Regression for Transportation Models*. *129*(17), 16–22.

Anbaroglu, B., Heydecker, B., & Cheng, T. (2014). Spatio-temporal clustering for non-recurrent traffic congestion detection on urban road networks. *Transportation Research Part C: Emerging Technologies*, *48*, 47–65.

Courbon, T., & Leclercq, L. (2011). Cross-comparison of macroscopic fundamental diagram estimation methods. *Procedia Social and Behavioral Sciences*, *20*, 417–426.

Daganzo, C. F. (2007). Urban gridlock: Macroscopic modeling and mitigation approaches. *Transportation Research Part B: Methodological*, *41*(1), 49–62.

Daganzo, C. F., & Geroliminis, N. (2008). An analytical approximation for the macroscopic fundamental diagram of urban traffic. *Transportation Research Part B: Methodological*, *42*(9), 771–781.

Edie, L. C. (1963). *Discussion of traffic stream measurements and definitions*. Port of new york authority.

Gayah, V. V., Gao, X. (Shirley), & Nagle, A. S. (2014). On the impacts of locally adaptive signal control on urban network stability and the Macroscopic Fundamental Diagram. *Transportation Research Part B: Methodological*, *70*, 255–268.

Geroliminis, N., & Sun, J. (2011). Properties of a well-defined macroscopic fundamental diagram for urban traffic. *Transportation Research Part B: Methodological*, *45*(3), 605–617.

Hamerly, G., & Elkan, C. (2002). Alternatives to the k-means algorithm that find better clusterings. *Proceedings of the Eleventh International Conference on Information and Knowledge Management - CIKM '02*, 600.

Liu, S., Liu, Y., Ni, L. M., Fan, J., & Li, M. (2010). Towards Mobility-based Clustering Categories and Subject Descriptors. *16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 919–927.

Nagle, A. S., & Gayah, V. V. (2013). A method to estimate the macroscopic fundamental diagram using limited mobile probe data. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, (October 2013), 1987–1992.

Xu, C., Liu, P., Wang, W., & Li, Z. (2012). Evaluation of the impacts of traffic states on crash risks on freeways. *Accident Analysis & Prevention*, *47*, 162–171.

Yasuda, S., Iryo, T., Koyama, Y., & Sakai, K. (2011). *Characteristics of ETC2 . 0 probe data compared to road traffic census data*. 3–6.

Zhang, L., Garoni, T. M., & de Gier, J. (2013). A comparative study of Macroscopic Fundamental Diagrams of arterial road networks governed by adaptive traffic signal systems. *Transportation Research Part B: Methodological*, *49*, 1–23.

APPENDIX

Table 3. Piecewise linear regression results (1/3)

Prefecture	**Period**	P_1	P_2	β_1	β_2	β_3	**R2**	**Sample**
Hokkaido	Weekday	0.5	1.8	1.3**	0.8**	0.6**	0.989	5,901
	Holiday	0.51	1.14	1.2**	1.0**	0.7**	0.993	2,857
Aomori	Weekday	0.18	1.54	1.8**	0.9**	0.7**	0.967	5,870
	Holiday	0.19	1.74	1.6**	0.9**	0.7**	0.971	2,842
Iwate	Weekday	0.3	2.0	1.6**	0.7**	0.6**	0.970	5,897
	Holiday	0.2	1.0	1.6**	1.0**	0.6**	0.978	2,860
Miyagi	Weekday	0.11	0.64	2.1**	1.1**	0.7**	0.985	5,905
	Holiday	0.1	1.0	2.1**	1.0**	0.6**	0.988	2,860
Akita	Weekday	0.69	2.12	1.3**	0.7**	0.5**	0.980	5,878
	Holiday	0.68	1.29	1.2**	0.9**	0.7**	0.984	2,839
Yamagata	Weekday	0.08	0.74	2.2**	1.1**	0.7**	0.976	5,898
	Holiday	0.09	1.09	2.0**	1.0**	0.7**	0.977	2,858
Fukushima	Weekday	0.43	2.78	1.7**	0.5**	-0.2**	0.972	5,907
	Holiday	0.39	1.51	1.6**	0.7**	0.5**	0.975	2,858
Ibaraki	Weekday	0.11	0.61	1.6**	1.2**	0.8**	0.986	5,900
	Holiday	0.19	1.26	1.4**	1.0**	0.6**	0.981	2,860
Tochigi	Weekday	0.52	0.95	1.6**	0.5**	0.6**	0.990	5,907
	Holiday	0.44	1.46	1.5**	0.8**	0.5**	0.989	2,859
Gunma	Weekday	0.4	1.29	1.7**	0.6**	0.8**	0.986	5,906
	Holiday	0.26	1.17	1.6**	0.9**	0.6**	0.986	2,858
Saitama	Weekday	0.35	0.86	1.5**	1.0**	0.5**	0.991	5,908
	Holiday	0.68	1.96	1.4**	0.6**	0.2**	0.992	2,880
Chiba	Weekday	0.58	2.31	1.3**	0.7**	-0.1**	0.990	5,906
	Holiday	0.62	1.72	1.3**	0.7**	0.2**	0.986	2,880
Tokyo	Weekday	0.81	1.52	1.3**	0.6**	-0.1**	0.989	5,907
	Holiday	0.78	1.54	1.2**	0.7**	-0.2**	0.990	2,880
Kanagawa	Weekday	0.64	1.8	1.2**	0.8**	-0.1**	0.991	5,906
	Holiday	0.67	1.27	1.2**	0.8**	0.6**	0.995	2,880
Niigata	Weekday	0.58	1.96	1.3**	0.8**	0.3**	0.988	5,907
	Holiday	1.01	–	1.2**	0.6**	–	0.988	2,858
Toyama	Weekday	0.35	2.83	1.7**	0.6**	-0.4**	0.984	5,906
	Holiday	0.36	1.13	1.5**	0.8**	0.6**	0.984	2,858
Ishikawa	Weekday	0.53	1.3	1.5**	0.4**	0.7**	0.978	5,906
	Holiday	0.07	0.55	1.7**	1.3**	0.7**	0.984	2,857
Fukui	Weekday	0.52	1.13	1.7**	0.3**	0.6**	0.979	5,905
	Holiday	0.33	0.94	1.6**	0.8**	0.6**	0.980	2,857
Yamanashi	Weekday	0.28	1.0	1.6**	0.9**	0.6**	0.980	5,907
	Holiday	0.64	2.64	1.4**	0.6**	-1.0**	0.981	2,880
Nagano	Weekday	0.79	1.42	1.2**	0.6**	0.8**	0.983	5,895
	Holiday	0.84	–	1.2**	0.2**	–	0.988	2,856
Gifu	Weekday	0.78	–	1.2**	0.7**	–	0.989	5,902
	Holiday	0.5	1.0	1.2**	1.0**	0.7**	0.993	2,880
Shizuoka	Weekday	0.6	1.14	1.6**	0.2**	0.7**	0.984	5,906
	Holiday	0.46	1.07	1.5**	0.8**	0.5**	0.988	2,880

** $p<0.01$; * $p<0.05$.

Table 4. Piecewise linear regression results (2/3)

Prefecture	Period	P_1	P_2	β_1	β_2	β_3	R2	Sample
Aichi	Weekday	0.69	2.24	1.3**	0.7**	0.1**	0.990	5,906
	Holiday	0.69	1.3	1.3**	0.8**	0.5**	0.994	2,880
Mie	Weekday	0.17	1.12	1.7**	1.0**	0.6**	0.976	5,894
	Holiday	0.19	1.13	1.6**	1.0**	0.4**	0.981	2,879
Shiga	Weekday	0.48	–	1.4**	0.6**	–	0.983	5,906
	Holiday	0.66	1.73	1.3**	0.6**	0.2**	0.988	2,880
Kyoto	Weekday	0.71	2.56	1.3**	0.7**	1.1**	0.990	5,906
	Holiday	0.79	2.25	1.3**	0.6**	0.3**	0.993	2,880
Osaka	Weekday	0.4	1.1	1.2**	1.0**	0.6**	0.984	5,905
	Holiday	1.1	–	1.1**	0.7**	–	0.989	2,880
Hyogo	Weekday	0.73	2.62	1.3**	0.6**	-1.0**	0.981	5,905
	Holiday	0.61	1.65	1.3**	0.8**	0.3**	0.985	2,880
Nara	Weekday	0.64	1.21	1.4**	0.5**	0.8**	0.987	5,903
	Holiday	0.44	1.01	1.4**	1.0**	0.6**	0.989	2,880
Wakayama	Weekday	0.53	2.15	1.3**	0.8**	0.3**	0.989	5,905
	Holiday	0.42	1.13	1.3**	1.0**	0.6**	0.989	2,880
Tottori	Weekday	0.12	0.95	1.6**	1.0**	0.7**	0.972	5,887
	Holiday	0.31	1.21	1.3**	1.0**	0.7**	0.980	2,869
Shimane	Weekday	0.19	1.39	1.5**	0.9**	0.7**	0.975	5,874
	Holiday	0.1	1.15	1.7**	1.0**	0.7**	0.973	2,868
Okayama	Weekday	0.48	1.0	1.3**	0.9**	0.6**	0.987	5,905
	Holiday	0.86	–	1.2**	0.6**	–	0.989	2,880
Hiroshima	Weekday	0.88	–	1.1**	0.7**	–	0.990	5,906
	Holiday	0.35	0.93	1.1**	1.2**	0.6**	0.993	2,880
Yamaguchi	Weekday	0.2	1.3	1.8**	0.9**	0.7**	0.967	5,617
	Holiday	0.2	1.3	1.6**	1.0**	0.7**	0.971	2,820
Tokushima	Weekday	0.77	2.55	1.2**	0.7**	-0.1**	0.986	5,900
	Holiday	0.67	1.07	1.2**	1.0**	0.6**	0.989	2,879
Kagawa	Weekday	0.53	2.28	1.3**	0.8**	0.6**	0.987	5,901
	Holiday	0.49	1.18	1.3**	0.9**	0.7**	0.987	2,880
Ehime	Weekday	0.65	1.7	1.2**	0.8**	0.7**	0.985	5,904
	Holiday	0.9	–	1.2**	0.7**	–	0.989	2,880
Kochi	Weekday	0.56	–	1.3**	0.8**	–	0.979	5,796
	Holiday	0.51	1.0	1.2**	1.0**	0.7**	0.982	2,858
Fukuoka	Weekday	0.59	1.59	1.2**	0.9	0.4**	0.982	5,906
	Holiday	0.8	1.2	1.1**	1.0**	0.5**	0.987	2,880
Saga	Weekday	0.72	1.21	1.2**	0.7**	0.8**	0.985	1,436
	Holiday	0.5	1.0	1.2**	1.1**	0.7**	0.985	648
Nagasaki	Weekday	0.78	2.29	1.2**	0.7**	0.5**	0.959	5,661
	Holiday	0.71	1.2	1.1**	1.0**	0.7**	0.969	2,832
Kumamoto	Weekday	0.51	1.21	1.4**	0.7**	0.6**	0.977	5,886
	Holiday	0.52	1.0	1.3**	0.9**	0.6**	0.978	2,878
Oita	Weekday	0.68	–	1.2**	0.7**	–	0.988	5,901
	Holiday	0.78	1.39	1.2**	0.9**	0.6**	0.988	2,880

** $p<0.01$; * $p<0.05$.

Table 5. Piecewise linear regression results (3/3)

Prefecture	**Period**	**P_1**	**P_2**	**β_1**	**β_2**	**β_3**	**R2**	**Sample**
Miyazaki	Weekday	0.71	2.1	1.1**	0.8**	0.5**	0.978	5,874
	Holiday	1.0	–	1.2**	0.6**	–	0.983	2,876
Kagoshima	Weekday	0.57	1.87	1.3**	0.8**	0.6**	0.980	5,864
	Holiday	0.92	–	1.2**	0.6**	–	0.986	2,868
Okinawa	Weekday	0.3	1.6	1.6**	0.8**	0.5**	0.942	5,266
	Holiday	0.3	1.68	1.5**	0.8**	0.6**	0.948	2,628

** $p<0.01$; * $p<0.05$.