IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 12 December 2023, accepted 4 January 2024, date of publication 12 January 2024, date of current version 22 January 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3353462

RESEARCH ARTICLE

Development of a Method for Estimating Traffic Volume Fluctuations That Considers Calendar Information and Road Network Geometry

TAKASHI KOBAYASHI [1] **AND MASAYOSHI TANISHITA** [2]

[1]National Traffic Safety and Environment Laboratory, Chofu, Tokyo 182-0012, Japan

[2]Faculty of Science and Engineering Department of Civil and Environmental Engineering, Chuo University, Bunkyo-ku, Tokyo 112-8551, Japan

Corresponding author: Takashi Kobayashi (takoba@ntsel.go.jp)

This work was supported by the Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Scientific Research (KAKENHI) under Grant 21K14272.

ABSTRACT This study aims to develop a method for estimating the traffic volume at locations and time periods where past data are unavailable. To achieve this, we constructed a model based on seven years of traffic volume observations on 22 main roads in Tokyo. The estimation method comprises two components: a traffic volume fluctuation pattern model that explains pre-clustered patterns and an annual average daily traffic model. Through our analysis, we made three key observations. First, the distance to the city center is a significant explanatory factor for monthly fluctuations, the month number for weekly fluctuations, the day-to-night population ratio for daily fluctuations, and the city center direction angle for hourly fluctuations. These calendar and road geometry variables account for 70–90% of the traffic volume fluctuation. Second, when estimating traffic volume for specific time periods using the traffic volume fluctuation pattern model, we explained approximately 60% of routes where no traffic volume observations were conducted and approximately 80% of routes where observations were made infrequently. Finally, the optimal number of clusters for the traffic volume fluctuation pattern, which maximized the coefficient of determination of the traffic volume model, was 8 for monthly fluctuations, 4 for weekly fluctuations, 6 for daily fluctuations, and 11 for hourly fluctuations. These findings made it possible to estimate traffic volumes on routes where traffic volume observations are not conducted and to interpolate traffic volumes for any given month or date using the results of periodic surveys conducted every few years.

INDEX TERMS Road traffic, traffic volume fluctuation, estimating, cyclical fluctuation patterns, calendar information, road network geometry.

I. INTRODUCTION

Different forms of mobility, including low-speed, compact, and personal mobility, are emerging to cater to diverse mobility needs. However, the coexistence of various types of vehicles with different speeds and sizes on roads leads to congestion and accidents [1], [2]. Therefore, considering road traffic conditions, travel demand, and mobility performance, effectively organizing each form of mobility within the hierarchical road structure is crucial.

Understanding traffic volume is essential for assessing road traffic conditions. However, the spatial and temporal densities of traffic volume observations have variations. While traffic volume is mainly observed on expressways and heavily trafficked main roads, observations are not conducted on some routes with significant traffic volume. Furthermore, routes with low traffic volumes may have available capacity that can be utilized for new mobility solutions. Thus, it is necessary to comprehend the traffic volume at unobserved locations and time periods when considering the utilization of different forms of mobility. However, owing to practical limitations, it is challenging to observe all routes at all times, necessitating the development of an estimation method.

The associate editor coordinating the review of this manuscript and approving it for publication was Amin Zehtabian.

When estimating traffic volume, the representative measures, annual average daily traffic (AADT) and traffic volume fluctuation, vary significantly based on the observation point. Takashi [3] demonstrated that there are differences in the way traffic volume fluctuates depending on the geometry of the road, such as belt lines and radial roads. Moreover, it considerably fluctuates even at the same location, with values doubling or halving depending on the time and day. Liu and Sharma [4] compared weekday and holiday traffic volume fluctuations and observed significant differences in the time of day when traffic peaks occur. Hence, simultaneously estimating the fluctuations in traffic volume along with the AADT is crucial.

Previous studies on traffic volume estimation can be broadly classified into two categories: 1) Development of a future forecasting method for routes with observed traffic volumes and a spatio-temporal interpolation method for observation points. 2) Development of estimation methods for routes where traffic volumes have not yet been observed. The former is has been established; however, the latter is has not been sufficiently studied.

Therefore, time-series modeling methods such as autoregressive (AR) and moving average (MA) using historical data are effective on routes where traffic volumes are observed, and many modeling methods have been developed [5]. However, estimates of traffic volume fluctuations have not been targeted for routes where traffic volumes are unobserved and historical data are not available (Table 1).

Traffic volume fluctuates periodically, primarily on monthly, weekly, daily, and hourly bases. This periodic variation can be patterned to categorize continuously changing traffic volume fluctuations. By categorizing traffic volume fluctuations, their relationship with the characteristics of time periods and points of interest can be analyzed without relying on historical data, and traffic volume fluctuations can be estimated even for unobserved routes for which historical data are not available.

Based on the above background, this study aims to develop a method for estimating traffic volume in the absence of past data. To achieve this, we employed clustering techniques to identify pre-defined patterns of traffic volume fluctuations and created a decision tree-based model to explain these patterns. Additionally, we constructed a separate model to estimate the AADT. By utilizing the estimated values of the fluctuation coefficient and the AADT, we successfully estimated the traffic volume for specific time periods.

In the next Chapter, we summarize related previous studies and present the position of this study. Chapter 3 describes the model construction and validation methods. In Chapter 4, clustering of traffic volume fluctuation patterns, clarification of the relationship between clusters, calendar variables, and road geometry, and modeling to explain the fluctuation patterns with calendar variables and road geometry are presented. Chapter 5 summarizes this study and presents the future work.

II. LITERATURE REVIEW

In this paper, we categorize previous studies into two groups: those that propose methods for predicting future traffic volume on routes where traffic volume is observed and those that propose methods for estimating traffic volume on routes where traffic volume is not observed.

A. TRAFFIC VOLUME PREDICTION METHODS FOR ROAD SECTIONS WHERE TRAFFIC VOLUME IS OBSERVED

Traffic volume prediction methods for road sections where traffic volume is observed include models that mainly focus on temporal similarities and those that focus on both temporal and spatial similarities.

1) MODELS FOR TEMPORAL SIMILARITIES IN TRAFFIC FLUCTUATIONS

In previous studies, various approaches have been proposed to analyze temporal similarities in traffic volume. These include time-series modeling using ARIMA, neural network and deep learning models, estimation methods based on random forest or SVM (support vector machine) for defining traffic volume categories, as well as Bayesian and non-parametric approaches.

Approaches based on the ARIMA model [5] apply ARIMA and a nonlinear AR model to predict short-term traffic volumes. Hamed et al. [6] proposed an ARIMA model to predict traffic volume on main roads in urban areas and examined its order. Lee and Fambro [7] suggested the application of an ARIMA model that is divided into subsets for each of several observed road sections for the short-term prediction of traffic volume. Shuvo et al. [8] constructed the ARIMA, ETS, SNAIVE, and PROPHET time-series models to predict traffic volumes in congested areas of Bangladesh. To predict traffic volumes for data that lack a sufficient number of locations to apply the ARIMA model, which is frequently used for traffic volume predictions, Omkar and Kumar [9] proposed a method for decomposing and estimating fluctuating components using multiplicative decomposition. These studies emphasize the importance of considering temporal similarities (periodicity) while estimating the traffic volume.

In deep learning approaches, Nohekhan et al. [10] proposed using a deep learning model to estimate the hourly traffic volume of highway off-ramps. The downstream traffic volumes, probe speeds, and infrastructure characteristics of the road segments were used as explanatory variables. Based on a robust recurrent neural network, Khan et al. [11] developed a multi-step-ahead forecasting model for time-series data. Lingras et al. [12] compare statistical methods with neural network methods for time-series analysis to predict short-term highway traffic volumes.

Approaches focusing on defining boundaries between traffic volume categories using random forest or SVM [13] aimed to estimate traffic volumes on low-traffic volume routes in the US. These studies constructed random forest models using variables such as time, date, weather conditions, and traffic

TABLE 1. Summary of the advantages and disadvantages of existing methods.

Method	Subject of estimation			
	AADT	Future Traffic Volume Prediction (Traffic Volume Variation)	Spatial interpolation between observation points	Unobserved road section
Road sections where traffic volume is observed				
Models for temporal similarities in traffic fluctuations	○	○	×	×
Models for temporal and spatial similarities in traffic volume fluctuates	○	○	○	×
Road sections where traffic volume is not observed				
Regression models, Spatial interpolation through kriging	○	×	○	○
This method	○	○	○	○

volume from the previous day. Das [13] compared the estimation of boundaries using random forests, support vectors, and model trees, identifying model trees as the optimal choice among the three. These studies highlight the importance of incorporating route categorization into traffic volume estimation models, as traffic volumes significantly vary across different route categories.

Another approach presented by Tsapakis et al. [14] involves estimation methods utilizing regression and Bayesian analyses. These methods employ explanatory variables such as road function class, population density, spatial position, regional classification, and average daily traffic volume to estimate the average annual daily traffic (AADT). Smith and Demetsky [15] developed four prediction models for traffic volumes on highways in northern Virginia. These models included historical averages, time-series analysis, neural networks, and non-parametric regression. Their comparative analysis revealed the effectiveness of the non-parametric regression model in predicting traffic volumes 15 minutes into the future.

2) MODELS FOR TEMPORAL AND SPATIAL SIMILARITIES IN TRAFFIC VOLUME FLUCTUATIONS

The proposed approaches that focus on temporal and spatial similarities in traffic volume include models that consider spatiotemporal autocorrelation, models that consider spatiotemporal autocorrelation via deep learning, approaches that use data fusion to merge data from static and mobile observations, and models trained via deep learning on data from static and mobile observations.

To construct a real-time road traffic prediction method that considers spatiotemporal autocorrelation, Min and Wynter [16] developed a method that learns spatial and temporal correlations from time-series data on traffic conditions and predicts various traffic parameters, such as traffic volume, speed, and congestion level. Su et al. [17] studied a single highway in Guangzhou, China, and conducted an analysis using a spatiotemporal autocorrelation function (ST-ACF) and a cross-correlation function (CCF) to clarify whether seasonal patterns exist within spatiotemporal correlations. Du et al. [18] used the connectivity between roads to propose a new method for calculating spatiotemporal correlations in traffic flows.

As models that consider spatiotemporal autocorrelation using deep learning, Zhan et al. [19] combined deep neural networks (DNN) with spatiotemporal auto-regression (ST-AR) using GPS location data to estimate traffic volume across regions in Shanghai. Zheng et al. [20] proposed a fusion deep learning model that incorporated spatiotemporal correlations for predicting traffic volumes on urban roads. Zhu et al. [21] proposed a method that incorporated temporal and spatial autocorrelations in traffic flow, utilizing traffic volume data from the previous day and a radial basis function neural network. Their model aimed to predict traffic volume on main roads in the city of Shenyang, China.

Among the approaches that utilize data fusion by combining static and mobile observation data, Cui et al. [22] presented a method for analyzing spatiotemporal correlations. They merged loop detection sensor data and mobile traffic signal data to predict traffic volume on highways in Fujian Province, China.

Among models trained via deep learning on both static and mobile observation data, Tang et al. [23] focused on two large areas encompassing the provincial capital in China. They proposed an estimation method that utilized deep reinforcement learning and traffic simulations. The method aimed to estimate traffic volumes across the entire city by leveraging high-density GPS locations and incomplete locations obtained from surveillance camera systems. Jun-Fang and Yan [24] introduced a traffic flow status prediction method that considered spatiotemporal correlations. Their approach involved a GPS-based vehicle location fusion optimization deep model (BN-LSTM-CNN) to predict short-term traffic volume.

Thus, previous studies have demonstrated the development of accurate methods for predicting short-term traffic volumes based on historical traffic volume data.

B. TRAFFIC VOLUME ESTIMATION METHODS FOR UNOBSERVED ROAD SECTIONS

Approaches for estimating traffic volume in road sections where direct observations are unavailable encompass regression models, spatial interpolation through kriging, and estimation using Origin-Destination (OD) traffic volumes.

Using regression models, Seaver et al. [25] explored the use of statistical models to estimate traffic volume on local roads in Georgia, USA.

Sfyridis and Agnolucci [26] proposed a combined clustering and regression modeling method to estimate road traffic volume in the UK and Wales. Their approach involves clustering sections of the road network based on similar characteristics and using regression modeling to estimate traffic volumes within each cluster. By integrating clustering and regression, they achieved higher accuracy estimates compared to using regression modeling alone. Apronti et al. [27] aimed to estimate traffic volumes on low-traffic roads in Wyoming, USA. They proposed a linear regression model that considered road pavement type, highway access, land use type, population, and a logistic regression model to estimate traffic using five levels. Caceres et al. [28] suggested a clustering and regression modeling method that considered road and socioeconomic attributes of nearby cities to estimate traffic volumes at locations with scarce traffic volume data between cities. Takashi et al. [29] employed multiple regression models to explain road-section traffic volumes and density using road functions (width, link length, etc.) and information about local land use (commercial or residential areas). They demonstrated that the models' explanatory power varied when analyzed based on different traffic volume categories. Takashi et al. [30] proposed an index of dependence on main roads to determine the traffic volume category for a given route. This index measures the similarity of traffic volume on an estimated route to that of main roads in the region. These studies emphasize the significance of dividing routes into distinct groups before estimating traffic volume.

Baffoe-Twum et al. [31] conducted a comprehensive review of commonly used methods to estimate AADT data. They performed a meta-analysis of prevalent approaches, including regression analysis, artificial neural network technology, existing traffic demand models, traditional factorial approaches, and spatial interpolation (kriging) methods, for estimating AADT on roads with low traffic volume. Their findings indicated the effectiveness of spatial interpolation (kriging). Zhang et al. [32] proposed a method to estimate the traffic volume of an entire network by incorporating observed traffic volume and probe car data into a geometric matrix interpolation model. They also developed a unique spatial smoothing index to enhance the estimation process. Wang et al. [33] developed a data-driven traffic volume estimation method for large-scale road networks by combining license plate recognition (LPR) data and GPS location data from taxis. This approach allowed for accurate estimation of traffic volume across extensive road networks.

Zhong and Hanson [34] used geographic information system-based transportation data management (TDM) to predict traffic volume in two regions in New Brunswick, Canada. By utilizing national census data and the Institute of Transportation Engineers (ITE) Quick Response Method, they demonstrated the feasibility of predicting traffic volume for a significant portion of the road network. Toi et al. [35] estimated traffic volume in non-observed road segments by leveraging the observed traffic volume of other segments and the OD volume composition ratio within the road network.

These previous studies highlight the limited research on traffic volume estimation methods for unobserved routes compared to the observed routes. Estimation methods for non-observed routes are still underdeveloped. Existing methods, such as regression models and co-kriging approaches, explain representative traffic volumes, such as AADT, using variables related to land use and road function. However, these methods have limitations in terms of temporal resolution and the ability to predict traffic volume fluctuations. Moreover, these studies emphasize the importance of considering road functions and categorizing roads based on their functions in traffic volume estimation.

C. POSITIONING OF THIS STUDY

Many models utilizing neural networks or deep learning have been developed for continuously observed routes, incorporating spatiotemporal correlations and time-series modeling. However, these models require past traffic volume information for the observed road section.

In the case of unobserved routes, methods such as regression models and co-kriging have been developed to estimate AADT. However, these methods lack the ability to predict traffic volume fluctuations. To address this, we aim to identify patterns of traffic volume fluctuations by utilizing calendar-based information (e.g., month and day) and map-derived information (e.g., circular and radial road forms). By analyzing seven years of traffic volume observations on 22 main roads in the Tokyo metropolitan area, we uncovered occurrence conditions for traffic volume fluctuation patterns not previously identified in existing studies. Incorporating these findings into our model, we devised a method to predict traffic volume at unobserved locations, enabling the estimation of more temporally decomposable information.

This study proposes an approach that combines regression modeling and cluster analysis, similar to the studies conducted by Sfyridis and Agnolucci [26], Apronti et al. [27], and Caceres et al. [28]. However, our approach and the previous research have distinct differences. First, while previous studies focused on estimating AADT, our study aims to estimate fluctuations in traffic volumes. Second, in previous research, the routes were grouped into clusters based on road functions and other factors, and a regression model was developed for each cluster. In contrast, our study employs cluster analysis to categorize fluctuation patterns and constructs a model to estimate the type of cluster. Therefore, our study differs from

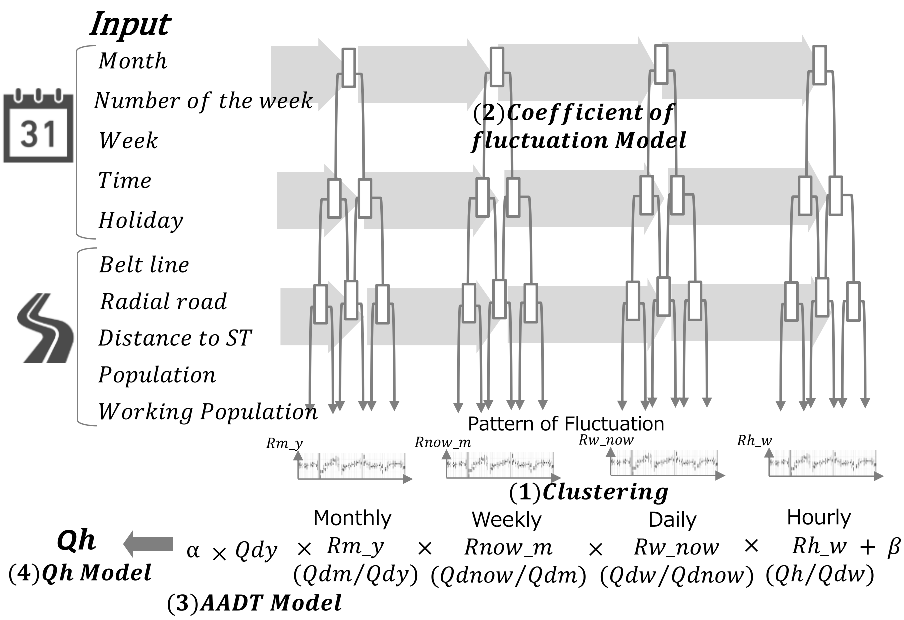


FIGURE 1. **Model structure. The hourly traffic volume (Qh) is calculated from AADT, coefficient of variation, and parameters α and β. The coefficient of variation is determined using the traffic volume fluctuation pattern estimation model. The model is explained for several pre-clustered variation patterns with calendar information and road network Geometry. The AADT is determined using the AADT model.**

previous research as we specifically focus on estimating these fluctuations and constructing a model to predict their patterns.

III. MODEL CONSTRUCTION AND VERIFICATION METHODS

A. MODEL CONSTRUCTION METHOD

In this study, we constructed a structural model, as shown in Fig. 1, to estimate the traffic volume at locations and time periods of unobserved traffic volume. A fluctuation-coefficient estimation model was constructed to estimate the fluctuation pattern on each time axis clustered in advance using explanatory variables obtained from calendars and maps. Next, a model for estimating the AADT was constructed, and the time-period-specific traffic volume was estimated from the estimated traffic volume fluctuation coefficient and AADT.

This model was built using the following five steps (Fig. 2):

Step 1: Cluster the patterns of traffic volume fluctuations for each time axis (month/week/daily/hour).

Step 2: Construct a decision tree model that explains the traffic volume fluctuation pattern, which varies with location and period.

Step 3: Build an estimation model for AADT.

Step 4: Build a model that explains the time-period-specific traffic volume by multiplying the traffic volume fluctuation coefficient for each time axis (month/week/day of week/hour) by the AADT.

Step 5: Determine the optimal number of clusters that provide explanatory power.

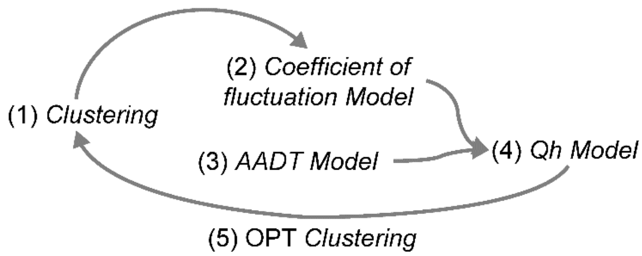


FIGURE 2. **Analysis procedure.**

1) CLUSTERING TRAFFIC VOLUME FLUCTUATION PATTERNS

The fluctuation coefficient for each time axis (month, week, day of the week, and hour) was calculated using (1–4). These calculated coefficients were then clustered across all locations and time periods using the K-means clustering method. To determine the optimal number of clusters in Step 5, we varied the number of clusters from 3 to 20 during the clustering process.

$$Rm_y = Qdm/Qdy \tag{1}$$
$$Rnow_m = Qdnow/Qdm \tag{2}$$
$$Rw_{now} = Qdw/Qdnow \tag{3}$$
$$Rh_w = Qh/Qdw \tag{4}$$

where Rm_y, $Rnow_m$, Rw_{now}, and Rh_w represent the monthly, weekly, daily, and hourly fluctuation coefficients, respectively; and Qdy, Qdm, $Qdnow$, Qdw, and Qh represent the

TABLE 2. Explanatory variables for traffic volume fluctuation pattern decision tree model.

Explanatory variable	Explanation
Month	Month
Week	Week number with the first day of the week every month as 1
Holiday	Holiday
Angle_CTR	Vehicle traveling direction angle with direction angle of the city center at the measurement point as 0° (0°–180°)
Dist_CTR	Straight-line distance to city center [km]
Dist_ST	Straight-line distance to nearest station [km]
Dist_IC	Road distance to nearest expressway IC [km]
Angle_HW_CTR	Vehicle traveling direction angle with direction angle of the city center on the inbound lane of the nearest expressway as 0°
Lane	Number of lanes on one side
Speed	Speed limit
POP_day/night	Ratio of daytime population to nighttime population

annual, monthly, weekly, daily average daily traffic, and hourly traffic, respectively.

2) BUILD A TRAFFIC VOLUME FLUCTUATION PATTERN DECISION TREE MODEL

We developed a decision-tree model to represent the clustered fluctuation patterns. The decision tree model uses explanatory variables listed in Table 2, which include calendar-based variables such as month, week number, day of the week, hour, and holidays, as well as road-map variables such as circular/radial shape, distance to station, roadside residential population, and roadside working population. The selection of variables was automatically performed. The decision tree model was selected owing to its ease of interpretation of the relationship between traffic volume patterns, season, time of day, and road geometry. The development of a more optimal model for practical use, including various state-of-the-art methods, needs to be considered, and we believe that this should be studied in the future.

3) BUILDING AADT ESTIMATION MODEL

In the AADT model, we utilized a multiple regression approach to estimate the relationship between the dependent variable and several explanatory variables. These variables, as presented in Table 3, encompassed the number of lanes, number of vehicles owned by the municipality, resident population, working population, residential land area ratio, and GDP.

TABLE 3. Explanatory variables for AADT model.

Explanatory variable	Explanation
Lane	Number of lanes
Pas_Car	Number of vehicles owned according to municipality
R_PasCar_POP	Number of cars owned per population
POP_day	Daytime population
POP_night	Nighttime population
WPOP_ind1	Working population (primary industry)
WPOP_ind2	Working population (secondary industry)
WPOP_ind3	Working population (tertiary industry)
ARE_residence	Residential area ratio
ARE_industry	Industrial area ratio
ARE_business	Business area ratio
GDP	GDP

4) BUILDING A TIME-PERIOD-SPECIFIC TRAFFIC VOLUME MODEL

To estimate the time-period-specific traffic volume, we used (5) for the regression-model estimation, which incorporates explanatory variables derived from multiplying the fluctuation coefficient of each time axis by the AADT.

$$Qh = \alpha \times Qdy \times Rm_y \times Rnow_m \times Rw_{now} \times Rh_w + \beta, \quad (5)$$

where α and β represent the parameters estimated using the least-squares method.

Because the frequency at which traffic volume can be observed depends on the road cross-section, five models with different traffic volume observation conditions (Models 1–5) were estimated.

Model 1: Estimates time-period specific traffic volume for unobserved routes with no input of traffic volume information.

Model 2: Assumes multiple traffic volume observations throughout the year and performs temporal interpolations to estimate time-period-specific traffic volume.

Model 3: Assumes monthly traffic volume observations and performs temporal interpolations for weekly, daily, and hourly fluctuations to obtain time-period-specific traffic volume.

Model 4: Predicts time-period specific traffic volume by assuming that weekly traffic volume if obtained.

Model 5: Predicts time-period specific traffic volume by assuming that daily specific traffic volume is obtained.

5) DETERMINING THE NUMBER OF CLUSTERS

The optimal number of clusters was determined by confirming the coefficient of determination and the accuracy of the estimated time-period-specific traffic volume model.

B. DATA

The data used for the analysis comprised traffic volumes on 22 main roads in Tokyo (Fig. 3, Table 4), measured using a

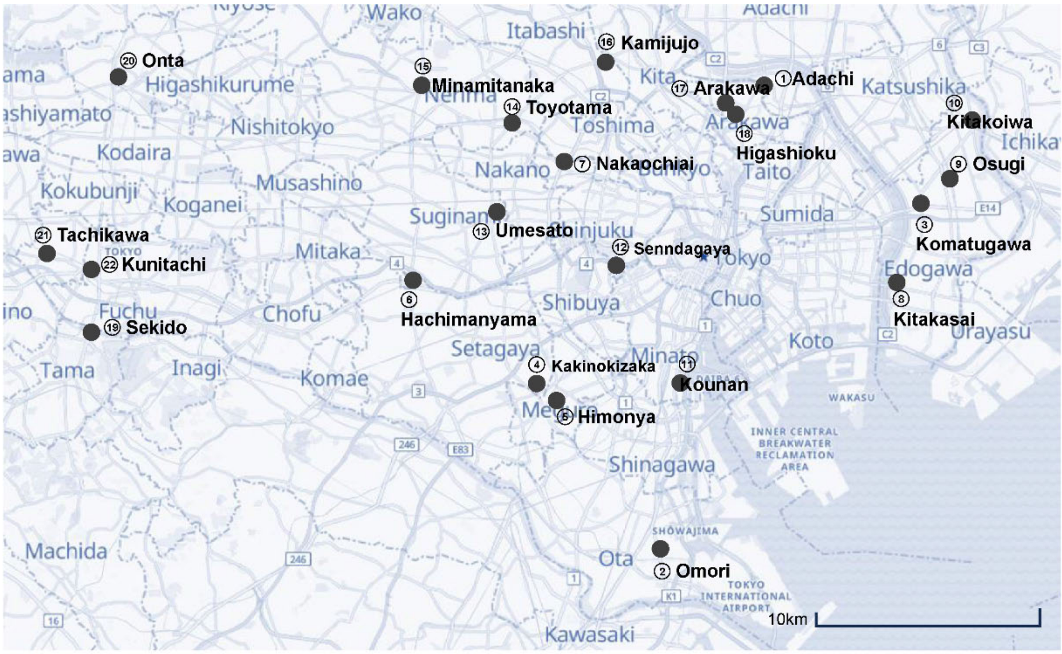


FIGURE 3. Traffic-volume measurement locations.

traffic counter over a seven-year period (1995–2001). Data points with a traffic volume of 0, missing values, or identical values were excluded from the analysis owing to potential challenges with the measurement equipment.

IV. ESTIMATION OF TRAFFIC VOLUME ESTIMATION MODEL

A. CLUSTERING OF TRAFFIC VOLUME FLUCTUATION PATTERNS

Fig. 4 illustrates the relationship between the five clustered fluctuation patterns and calendar and road geometry variables. The vertical axis represents the calendar variable or road shape variable (RO: outbound radial road, RI: inbound radial road, and C: circular road), whereas the horizontal axis represents each time axis (a: monthly fluctuation, b: weekly fluctuation, c: daily fluctuation, and d: hourly fluctuation) along with the clustered traffic volume fluctuation pattern. The density of occurrence probability is depicted through color shading.

1) MONTHLY FLUCTUATION CHARACTERISTICS

Regarding the monthly fluctuations, the following characteristics were observed:

1) Lower values were observed in January, February, August, and September.
2) No significant differences were observed in the fluctuation patterns when comparing clusters.
3) Cluster 1 exhibited a higher amount of RI, whereas Cluster 4 had a higher amount of C.

However, no clear distinctions were observed.

2) WEEKLY FLUCTUATION CHARACTERISTICS

Regarding the weekly fluctuations, the following characteristics were observed:

1) The fluctuations were categorized into the following types: decrease up to Week 3, then increase in Week 4 (Cluster 1); slight increase up to Week 3, then decrease in Week 4 (Cluster 2); constant type (Cluster 3); increase (Cluster 4); and constant up to Week 3, then slight increase in Week 4 (Cluster 5).
2) The majority of the fluctuations belonged to the constant type, observed in Cluster 3 and Cluster 5. This pattern was prevalent in all months except for January, August, and December.
3) The patterns with an increase in Week 4 (Cluster 1 and Cluster 4) frequently occurred in January and August. This can be attributed to the extended holidays at the beginning of January and mid-August, with the traffic volume tending to normalize during the second half of the month.
4) The pattern with a decrease in Week 4 (Cluster 2) was often observed in December. This is likely because the end of December corresponds to a long year-end holiday, leading to a decrease in traffic volume.

3) DAILY FLUCTUATION CHARACTERISTICS

The characteristics of the daily fluctuations are as follows:

1) The fluctuations were categorized into the following types: decrease on Saturdays (Cluster 2, Cluster 5), no change from weekdays (Cluster 1, Cluster 3), and increase from weekdays (Cluster 4).

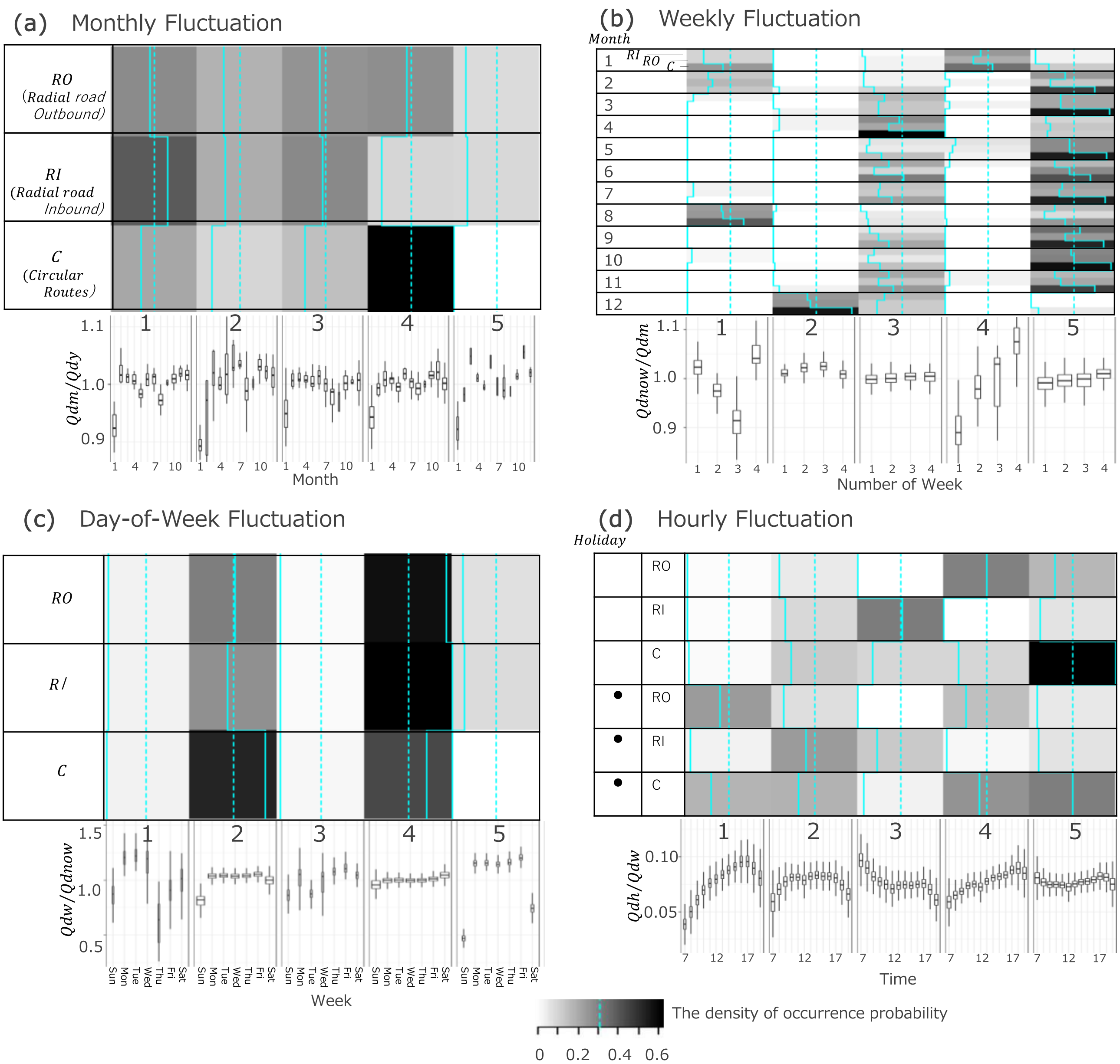


FIGURE 4. Relationship between fluctuation pattern and calendar variable / road shape (heatmap). The probability of occurrence of each variable pattern ((a) monthly, (b) weekly, (c) day-of-week, and (d) hourly) is displayed by representative road shape (RO: Radial road Outbound, RI: Radial road Inbound, C: Circular Routes) and calendar variables (Month, Holiday/Weekday). The darker the color, the higher is the probability of occurrence.

2) Across all clusters, there was a tendency for traffic volume to be lower on Sundays.
3) On weekdays, all clusters showed a relatively constant pattern, with a slight increase observed on Fridays.
4) Cluster 2, which exhibited a decrease on Saturdays, was commonly associated with circular routes, whereas Cluster 4, which exhibited an increase on Saturdays, was relatively more common on radial routes.
5) When the number of clusters was set to five, clusters with low-frequency patterns, such as Clusters 1, 3, and 5, were identified and classified.

4) HOURLY FLUCTUATION CHARACTERISTICS

The fluctuations in hourly traffic volume can be categorized into three types: morning-to-evening increase (Cluster 1, Cluster 4), decrease (Cluster 3), and constant (Cluster 2, Cluster 5).

The decrease type is predominantly observed in the inbound lanes of routes emanating from the city center, and it is more common on weekdays.

The increase type is mostly observed in the outbound lanes of routes emanating from the city center. Cluster 4, characterized by a gradual increase, is more prevalent on weekdays,

TABLE 4. Traffic volume measurement locations and traffic volumes.

	Name	Road	Average Daily Traffic Volume	Standard Deviations
①	Adachi	R4	56,394	0.049
②	Omori	R131	35,664	0.114
③	Komatugawa	R14	61,493	0.069
④	Kakinokizaka	Belt7	71,671	0.072
⑤	Himonya	Meguro-ST	52,791	0.081
⑥	Hachimanyama	Belt8	84,035	0.049
⑦	Nakaochiai	Belt6	38,958	0.067
⑧	Kitakasai	Kasaibashi-ST	36,597	0.106
⑨	Osugi	Kan7	66,825	0.065
⑩	Kitakoiwa	Kuramaebashi-ST	39,699	0.056
⑪	Kounan	Kyuukaigann-ST	33,044	0.253
⑫	Senndagaya	Kan5	33,608	0.062
⑬	Umesato	Oume-ST	49,705	0.061
⑭	Toyotama	Kan7	66,082	0.055
⑮	Minamitanaka	Sasame-ST	65,629	0.051
⑯	Kamijujo	Kan7	74,842	0.066
⑰	Arakawa	Kan5	40,134	0.100
⑱	Higashioku	Okubashi-ST	43,060	0.091
⑲	Sekido	Kamakura-ST	47,248	0.056
⑳	Onta	Shinoume-ST	46,939	0.052
㉑	Tachikawa	Shinokutama-ST	35,957	0.061
㉒	Kunitachi	R20	39,189	0.046

whereas Cluster 1, characterized by a rapid increase, is more common on weekends and holidays.

Although not shown in the figure, separate aggregations revealed similar rapid increases during the year-end and New Year holidays, the Golden Week at the beginning of May, and the Obon festival in mid-August.

Cluster 3, which experiences a decrease from morning to evening, occurs when there is a higher likelihood of morning commuters (weekdays, inbound routes on radial roads). Clusters 1, 2, and 4, which exhibit an increase from morning to evening, occur when there is a lower likelihood of morning commuters (holidays, outbound routes on radial roads).

Cluster 5, characterized by a constant fluctuation from morning to evening, often occurs along circular routes. Cluster 5 generally has a relatively high morning traffic volume and is particularly common on weekdays. Gradually increasing patterns, such as those in Cluster 4, are more likely to occur on holidays.

These findings indicate a clear association between the fluctuation patterns and the variables obtained from calendars and road network shapes. The observed relationships suggest that the temporal and spatial characteristics of traffic volume fluctuations are influenced by calendar-based factors and the configuration of the road network.

TABLE 5. Variable importance and accuracy variable importance means that the higher the value, the greater is the influence of the variable on the discrimination of variation patterns. accuracy represents the overall accuracy of the model.

	Monthly	Weekly	Daily	Hourly
Month		78		
Week				16
Angle_CTR			15	31
Dist_CTR	40	9	31	17
Dist_ST	21	3	10	7
Dist_IC	16	3	8	8
Angle_HW_CTR				8
Lane			2	
POP_day/night	23	7	34	12
Accuracy	0.95	0.68	0.82	0.73

B. ESTIMATION OF TRAFFIC VOLUME FLUCTUATION PATTERN MODEL

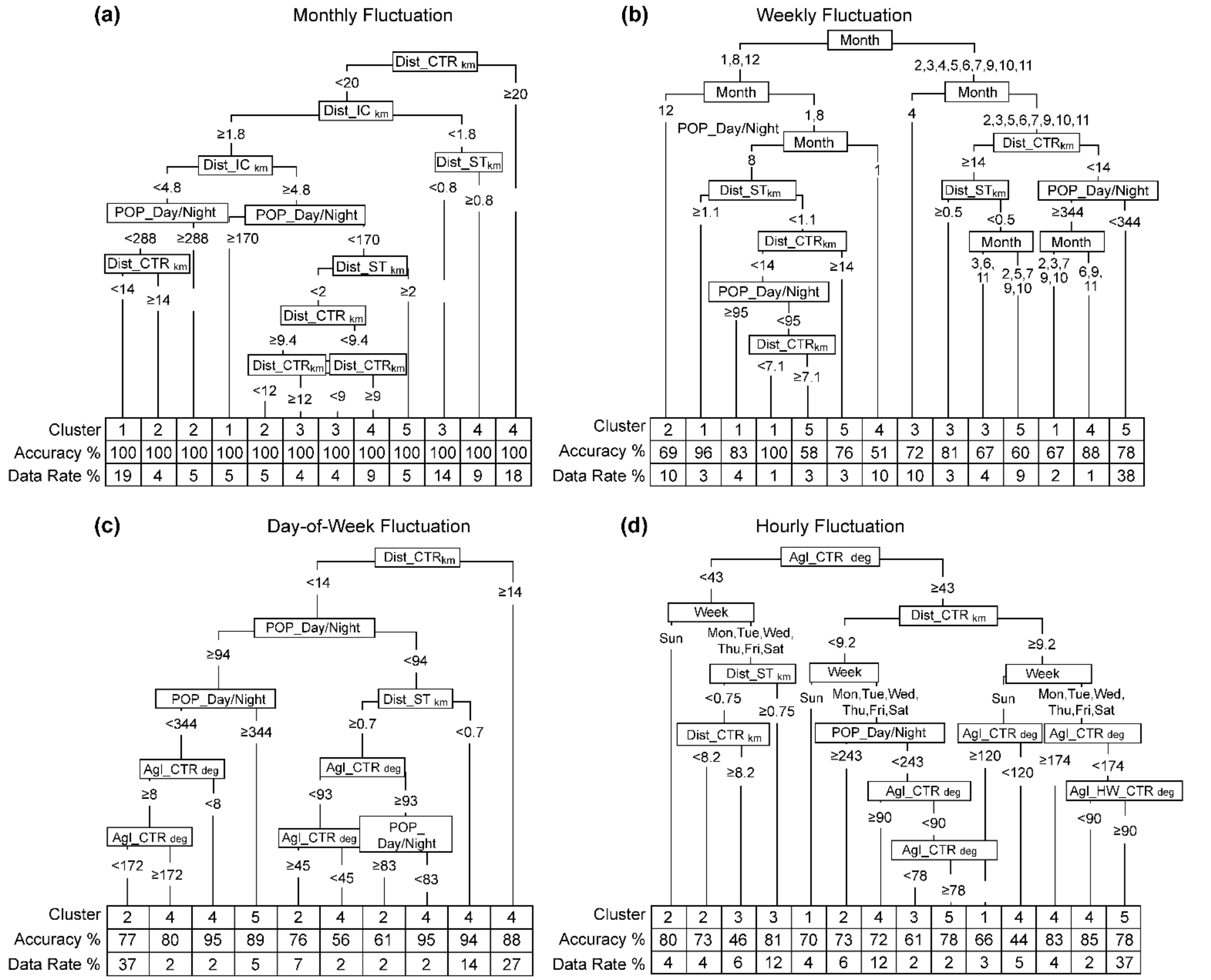
As mentioned earlier, a clear relationship was observed between the variables obtained from calendars and road network shapes. To explain the traffic volume fluctuation pattern, a decision-tree model was developed using these variables. Table 5 presents the importance and accuracy of the explanatory variables in estimating the fluctuation coefficient pattern for each time axis using the decision tree. Additionally, Fig. 5 illustrates the tree structure of the decision tree. Importance is a metric used to assess the level of contribution to the cluster classification.

1) VARIABLE IMPORTANCE

After analyzing the table, the following explanatory variables were chosen for the monthly fluctuation coefficient pattern estimation model: distance from the city center, distance to the station, distance to the expressway interchange, and day-to-night population ratio. Among these variables, distance to the city center showed the highest importance, whereas the other variables had similar levels of importance.

Similarly, for the weekly fluctuation coefficient pattern estimation model, the selected variables were the same as the monthly model, with the addition of the month variable. The month variable exhibited significantly higher importance compared to the other variables, similar to the monthly fluctuation model. As discussed earlier, the weekly fluctuation pattern is influenced by the month. This can be attributed to the varying impact of long holidays in different months (such as year-end/New Year, Obon) on weekly traffic volume fluctuations.

For the daily fluctuation coefficient pattern estimation model, the selected explanatory variables were the day-to-night population ratio, distance to the city center, direction angle to the city center, and distance to the station. Among these variables, the day-to-night population ratio had the highest importance, followed by the distance to the city center and direction angle to the city center. It is noteworthy that the



(a)

Cluster	1	2	2	1	2	3	3	4	5	3	4	4
Accuracy %	100	100	100	100	100	100	100	100	100	100	100	100
Data Rate %	19	4	5	5	5	4	4	9	5	14	9	18

(b)

Cluster	2	1	1	1	5	5	4	3	3	3	5	1	4	5
Accuracy %	69	96	83	100	58	76	51	72	81	67	60	67	88	78
Data Rate %	10	3	4	1	3	3	10	10	3	4	9	2	1	38

(c)

Cluster	2	4	4	5	2	4	2	4	4	4
Accuracy %	77	80	95	89	76	56	61	95	94	88
Data Rate %	37	2	2	5	7	2	2	2	14	27

(d)

Cluster	2	2	3	3	1	2	4	3	5	1	4	4	4	5
Accuracy %	80	73	46	81	70	73	72	61	78	66	44	83	85	78
Data Rate %	4	4	6	12	4	6	12	2	2	3	5	4	2	37

FIGURE 5. Dendrogram of traffic volume fluctuation pattern decision tree model. A legend for the variables in the boxes is provided in Table 2. Cluster numbers correspond to Figure 4. Accuracy represents the positive discrimination rate for each cluster. Data Rate represents the percentage of data for each cluster.

variables with high importance in the daily fluctuation model differ from those in the monthly and weekly fluctuation models. This can be attributed to the fact that the day-to-night population ratio, which strongly influences daily fluctuations, reflects commuting patterns that significantly vary depending on the day of the week (e.g., commuting on weekdays but not on Saturdays and Sundays). These variations in commuting behavior based on the day of the week are believed to be the underlying cause of these differences.

For the estimation model of hourly fluctuation coefficient patterns, the following explanatory variables were chosen: direction angle to the city center, distance to the city center, day of the week, day-to-night population ratio, distance to the IC, and direction angle to the expressway. The direction angle to the city center had the highest variable importance, followed by the distance to the city center and the day of the week. The large importance of the direction angle to the city center can be explained by the following observation: Routes leading to the city center tend to experience high traffic volume in the morning due to morning commuting, which subsequently decreases in the evening. Conversely, routes heading away from the city center have lower morning congestion but higher evening traffic as people return home. These variations in the direction angles from the city center contribute to the temporal fluctuations. Additionally, the day of the week was observed to be significant in the hourly fluctuation model owing to the following reason: Weekdays exhibited higher morning traffic volume due to commuting activities, whereas weekends showed lower morning movement. This variation of hourly traffic volume depending on the day of the week influenced the importance of the variable in the model.

2) DISCRIMINATION ACCURACY

The monthly fluctuations exhibited a high accuracy rate of 95% (Table 5). Notably, clusters other than clusters 3 and 4 achieved a perfect discrimination rate of 100% (Fig. 5). Despite the lack of significant differences in monthly fluctuation patterns between clusters, this consistency likely contributed to the high discrimination accuracy.

Regarding the weekly fluctuations, an accuracy rate of 68% was achieved (Table 5). However, Cluster 4 in January displayed a lower discrimination accuracy of 51% compared to other clusters (Fig. 5). This can be attributed to the unique nature of January, which encompasses the year-end and New Year's holidays, resulting in distinct traffic patterns.

The daily fluctuations demonstrated a discrimination accuracy of 82% (Table 5), with clusters of significant sample sizes achieving at least 70% accuracy (Fig. 5). The daily fluctuation patterns were divided between Clusters 2 and 4. Cluster 2 was frequently associated with routes close to the city center, characterized by a high day-to-night population ratio and circular road geometry. In contrast, Cluster 4 was often observed on routes located further from the city center, exhibiting a relatively lower day-to-night population ratio, and in proximity to stations. The distinguishing factor between Clusters 2 and 4 was the decrease in traffic volume on Saturdays for Cluster 2, whereas Cluster 4 exhibited an increase on Saturdays. The proximity of Cluster 4 to stations, which experience traffic demand even on weekends, and its connection to radial routes from the city center likely influenced this pattern.

In terms of hourly fluctuations, discrimination accuracy reached 73% (Table 5), with clusters of substantial sample sizes achieving at least 70% accuracy (Fig. 5). Cluster 3, which exhibited a decreasing pattern from morning to evening, was primarily associated with routes oriented towards the city center, with angle values typically below 78° at the sixth branch and around 43° for the first branch. However, Cluster 5, characterized by a consistent pattern throughout the day, was predominantly found in circular routes on weekdays (excluding Sundays) and displayed an angle range of at least 78° to a maximum of 174° perpendicular to the city center. The decision tree model appears to capture the influence of the direction angle to the city center on the traffic volume fluctuation patterns from morning to evening.

C. ESTIMATION OF AADT MODEL

In the preceding section, a decision tree model was developed to elucidate the fluctuation patterns utilizing variables derived from calendars and road maps. Now, our focus shifts to estimating a model that explains the AADT, representing the total traffic volume without considering fluctuations.

1) PARAMETER ESTIMATION

Table 6 presents the parameter estimation results for the AADT model. The explanatory variables shown in the table were significant, and an explanatory power with a coefficient of 0.66 was obtained.

TABLE 6. Parameter estimation results of AADT model Variable legends are listed in Table 3.

	Estimate	Pr(>|t|)
Int	3.316e+04	2e-16 ***
Lane	2.769e+03	2e-16 ***
Pas_Car	-5.83e+04	2e-16 ***
POP_night	-7.497e-04	2e-16 ***
WPOP_ind2	-1.086e-02	2e-16 ***
ARE_residence	6.544e+03	2e-16 ***
GDP	-7.831e-03	2e-16 ***
MAPE	23.7%	
R2 Adj.	0.66	

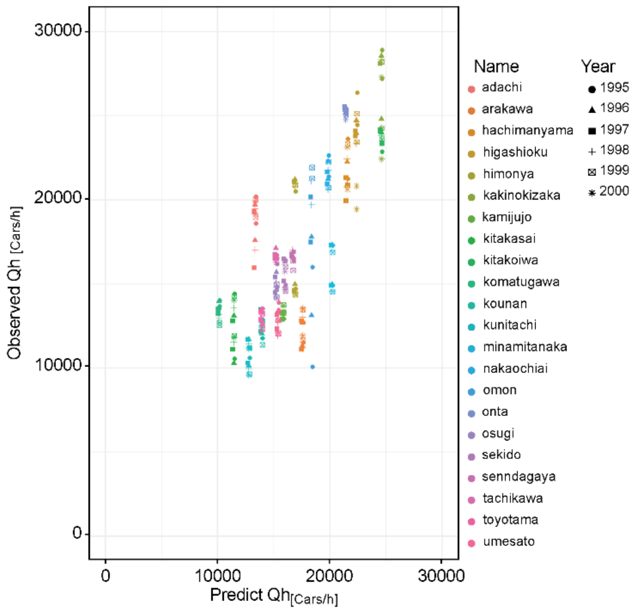


FIGURE 6. Comparison of measured and estimated values of Annual Average Daily Traffic. The vertical axis represents the observed traffic volume, and the horizontal axis represents the estimated traffic volume. Name in the legend represents the observation points and corresponds to Figure 3.

2) COMPARISON BETWEEN MEASURED AND ESTIMATED VALUES

Fig. 6 shows a scatterplot of the measured and estimated values from the AADT model.

By observing the distinctions among the data points (indicated by differences in color), the estimates successfully capture the variations in the actual measured values.

However, when considering the discrepancies between years (indicated by differences in shape), the estimated values exhibit less variability compared to the actual measured values. Consequently, the estimated values do not adequately reflect the year-to-year differences. The AADT estimation model combining clustering and regression models

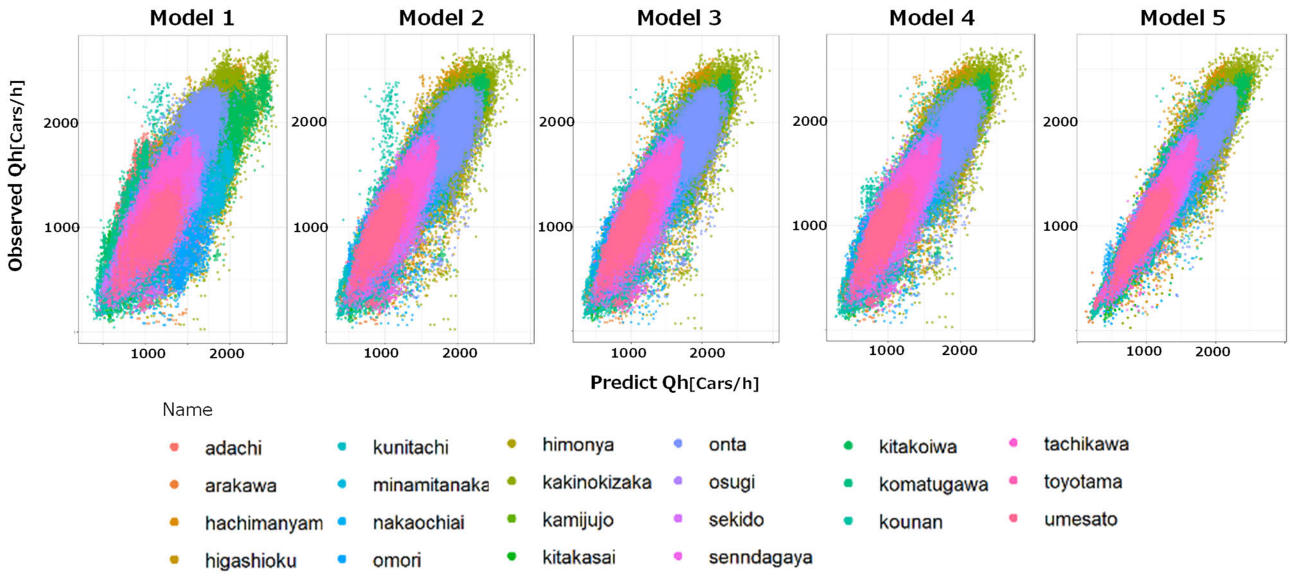


FIGURE 7. Comparison of measured and estimated values of time period-specific traffic volume Models 1–5 correspond to Table 7. Names in the legend represent the observation points and corresponds to Figure 3.

estimated by Sfyridis and Agnolucci [26] yields an estimated MAPE = 14%. The simple regression model estimated in this study yields MAPE = 23%, leaving room for improvement by using existing methods for AADT. As the accuracy of the time-period-specific traffic volume estimation model heavily relies on the AADT, there is a significant need for improvement in the accuracy of the model. This aspect will be addressed in future studies, including an exploration of previous research methods, to enhance the estimation accuracy of the AADT.

D. ESTIMATION OF TIME-PERIOD-SPECIFIC TRAFFIC VOLUME ESTIMATION MODEL

Table 7 presents the parameter estimation results of the time-period-specific traffic volume estimation model, and Fig. 7 shows a scatter plot of the measured and estimated values of the time-period-specific traffic volume estimation model. Model 1 corresponds to the scenario where no traffic volume information is available, whereas Models 2–5 correspond to situations where the yearly average, monthly average, weekly average, and daily traffic volume are obtained, respectively. Model 1 demonstrated an explanatory power with a coefficient of determination of approximately 0.58. Models 2–4 exhibited similar coefficients of determination, ranging from 0.85 to 0.87, and Model 5, which incorporated daily traffic volume, achieved a coefficient of determination of 0.91. Comparing Model 1, without traffic volume information, to Model 2, with the inclusion of AADT, the coefficient of determination increased from 0.58 to 0.85. This highlights the significant impact of the accuracy of the AADT estimation. As mentioned earlier, there is room for improvement in the accuracy of the AADT estimation, thereby suggesting possibilities for enhancing Model 1's estimation accuracy

TABLE 7. Parameter estimation results of time-period-specific traffic volume model Int represents the intercept, and QDY*R represents the explanatory variable obtained by multiplying AADT by the coefficient of variation for the month, week, day of the week, and time.

	Model 1	Model 2	Model 3	Model 4	Model 5
Qdy	-	Given	-	-	-
Qdm	-	-	Given	-	-
Qdnow	-	-	-	Given	-
Qdw	-	-	-	-	Given
Int	39.41	28.66	24.25	18.97	11.59
Pr(>|t|)	2e-16 ***	2e-16 ***	2e-16 ***	2e-16 ***	2e-16 ***
Qdy*R	0.970	0.978	0.98	0.98	0.99
Pr(>|t|)	2e-16 ***	2e-16 ***	2e-16 ***	2e-16 ***	2e-16 ***
R2 Adj.	0.58	0.85	0.86	0.87	0.91

This method effectively addresses a significant portion of traffic volume fluctuations. To determine the hourly traffic volume, it is essential to consider the AADT and the variability in traffic volume. In section IV-B, we have demonstrated how the latter, traffic volume fluctuation, can be quantified using the proposed method. Although estimating the AADT poses challenges, improving its accuracy in the future could enhance the precision of time-specific traffic volume estimation. The AADT can be roughly estimated through periodic surveys conducted every few years or multiple times within a year. This enables interpolation between different time points, even with surveys conducted multiple times annually, by utilizing fluctuation coefficient estimation.

E. EXAMINATION OF NUMBER OF CLUSTERS

Fig. 8 shows the cluster discrimination accuracy (upper row), coefficient of determination of the fluctuation coefficient

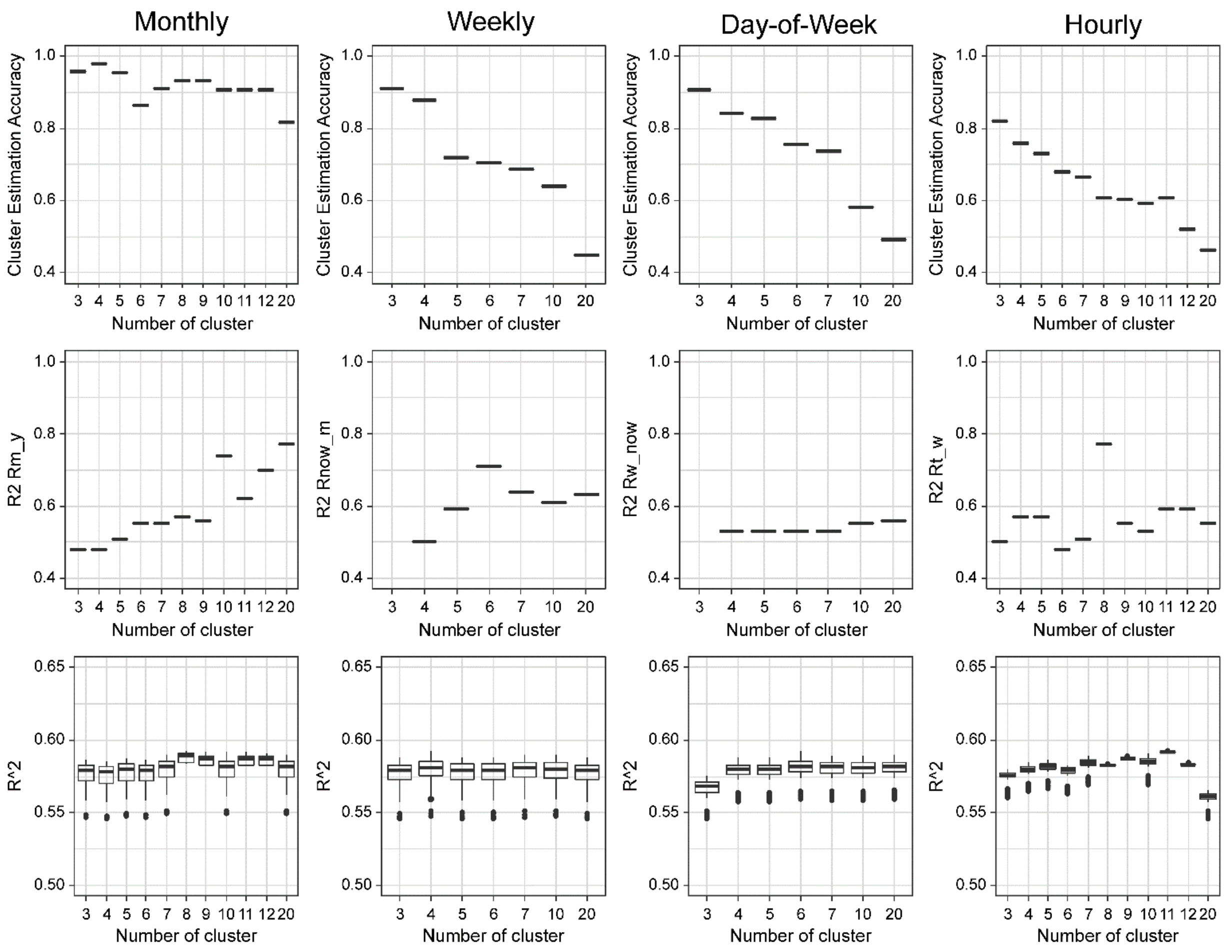


FIGURE 8. **Number of clusters and accuracy of each step. Upper row: the cluster discrimination accuracy. Middle row: coefficient of determination of the fluctuation coefficient. Bottom row: coefficient of determination of the time period-specific traffic volume estimation model across different cluster numbers.**

(middle row), and coefficient of determination of the time period-specific traffic volume estimation model (bottom row) across different cluster numbers.

The cluster discrimination accuracy (upper row) exhibits a decreasing trend as the number of clusters increases, observed consistently across monthly, weekly, daily, and hourly fluctuations. This decline can be attributed to the decision-tree-based estimation method, where the discrimination accuracy decreases with an increasing number of dependent variables.

The coefficient of determination of the fluctuation coefficient (middle row) demonstrates an increasing pattern with the number of clusters. This trend is particularly notable in monthly and daily fluctuations. It arises from the fact that smaller numbers of clusters cannot adequately express the variation in the fluctuation coefficient because it represents the average value of each fluctuation pattern.

The optimal number of clusters for the coefficient of determination of the time-period-specific traffic volume estimation model (bottom row) is as follows: eight clusters for monthly fluctuations, four clusters for weekly fluctuations, six clusters for daily fluctuations, and eleven clusters for hourly fluctuations. These numbers reflect the trade-off between cluster discrimination accuracy, which decreases with more clusters, and the coefficient of determination of the fluctuation coefficient, which tends to increase with a larger number of clusters.

V. DISCUSSION

In this study, we constructed a decision tree model to explain traffic volume fluctuations using variables from calendars and road maps. Additionally, we estimated a model that can explain the AADT. These models have been used to parameterize the time-period-specific volume estimation model.

Our findings revealed previously undisclosed insights into traffic volume fluctuations. Existing research on traffic volume estimation has mainly focused on time-series models or deep machine-learning approaches using historical data to predict future traffic volume. However, determining the traffic volume at specific points in time for roads without observed traffic volumes has remained challenging. In our study, we addressed this by estimating traffic volumes at specific time points on routes with unobserved traffic volumes, leveraging the significance of traffic volume fluctuation patterns through calendar and road geometry variables.

Furthermore, we recognize the potential for improvement in the AADT estimation model, as mentioned earlier. In addition to the simple regression model approach employed in this study, we suggest considering spatial interpolation methods, such as those highlighted by Takashi et al. [29], and examining the impact of neighboring arterial roads, as investigated by Caceres et al. [28], for AADT estimation. Understanding the travel potential upstream and downstream of a route, which is a fundamental factor, requires modeling that incorporates data on population, number of vehicles owned, and other relevant explanatory variables.

VI. CONCLUSION

As new mobility options catering to diverse needs, such as low-speed, compact, and personal mobility, continue to emerge, it becomes crucial to plan and implement suitable regulations for their use on roads. However, in areas with low mobility demand where these new forms of mobility are necessary, accurate observations of traffic volumes are often lacking. Consequently, determining the available traffic capacity in road space poses a significant challenge. Estimating traffic volume fluctuations on different time scales (monthly, weekly, daily, and hourly) has proven particularly difficult due to reliance on past data, making it challenging to estimate traffic volume for unobserved routes.

This study aimed to develop a method for estimating traffic volume in locations and time periods where past data are unavailable. To achieve this, we constructed a traffic volume fluctuation estimation method based on seven years of traffic volume observations on 22 main roads in Tokyo. The estimation method consists of a model for traffic volume fluctuation patterns, which explains pre-clustered traffic volume patterns, and an AADT model. In our analysis, we made the following three key observations:

First, for each time axis, important explanatory variables for traffic volume fluctuations were identified. These included distance to the city center for monthly fluctuations, month number for weekly fluctuations, day-to-night population ratio for daily fluctuations, and city center direction angle for hourly fluctuations. By incorporating these calendar and road geometry variables, we were able to explain 70–90% of the traffic volume fluctuations.

Secondly, when estimating traffic volume by time period using the traffic volume fluctuation pattern model, we were able to explain approximately 60% of routes where no traffic volume observations were conducted and approximately 80% of routes where observations were conducted infrequently (once every few years).

Thirdly, the optimal number of clusters for traffic volume fluctuation patterns, which maximized the coefficient of determination of the time period-specific traffic volume model, were identified as 8 for monthly fluctuations, 4 for weekly fluctuations, 6 for daily fluctuations, and 11 for hourly fluctuations.

These findings enable the estimation of traffic volumes on routes where traffic volume observations are not available. Additionally, they allow for the interpolation of traffic volumes for any given month or date by utilizing the results from periodic surveys conducted every few years.

However, the following three points warrant further research consideration:

First, the AADT volume model needs enhancement. Although our method successfully estimated traffic volume fluctuations, improving the accuracy of AADT estimation can lead to better accuracy in hourly traffic volume estimation.

Second, establishing a logic that determines the direction of human flow toward the city center is important. In the case of Tokyo, with its oblong shape and the city center located in the east, the movement of people has a clear direction. However, for local cities where the flow direction is not as apparent, logic must be developed to determine the direction of human flow.

Third, the traffic volume data obtained in this study were primarily focused on arterial roads with high traffic volumes. Because unobserved routes mostly consist of non-arterial roads, it is essential to investigate the fluctuation patterns of these non-arterial roads. We suggest exploring the applicability of the arterial road dependence concept proposed in [28] to delineate the scope of this method.

By improving the AADT estimation model, refining the traffic volume fluctuation pattern model, understanding urban human flow directions, analyzing non-arterial road patterns, and integrating these findings, we anticipate the development of a comprehensive method capable of estimating traffic volumes at any given point and time in the future.

ACKNOWLEDGMENT

The data used in this study was provided by the Bureau of Environment of the Tokyo Metropolitan Government. The authors express their gratitude to the reviewers. They would like to thank Editage (www.editage.com) for English language editing.

REFERENCES

[1] K. Hunter-Zaworski, "Impacts of low-speed vehicles on transportation infrastructure and safety," *J. Transp. Land Use*, vol. 5, no. 2, pp. 68–76, Oct. 2012, doi: 10.5198/jtlu.v5i2.250.

[2] T. Q. Pham, C. Nakagawa, A. Shintani, and T. Ito, "Evaluation of the effects of a personal mobility vehicle on multiple pedestrians using personal space," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 4, pp. 2028–2037, Aug. 2015, doi: 10.1109/TITS.2014.2388219.

[3] K. Takashi, "Analysis of similarities in traffic fluctuations between non-arterial roads and arterial roads," *JSTE J. Traffic Eng.*, vol. 9, no. 2, pp. 88–93, 2023, doi: 10.14954/jste.9.2_A_88.

[4] Z. Liu and S. Sharma, "Statistical investigations of statutory holiday effects on traffic volumes," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1945, no. 1, pp. 40–48, Jan. 2006, doi: 10.1177/0361198106194500106.

[5] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Statistical methods for detecting nonlinearity and non-stationarity in univariate short-term time-series of traffic volume," *Transp. Res. C, Emerg. Technol.*, vol. 14, no. 5, pp. 351–367, Oct. 2006, doi: 10.1016/j.trc.2006.09.002.

[6] M. M. Hamed, H. R. Al-Masaeid, and Z. M. B. Said, "Short-term prediction of traffic volume in urban arterials," *J. Transp. Eng.*, vol. 121, no. 3, pp. 249–254, 1995, doi: 10.1061/(ASCE)0733-947X(1995)121:3(249).
[7] S. Lee and D. B. Fambro, "Application of subset autoregressive integrated moving average model for short-term freeway traffic volume forecasting," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1678, no. 1, pp. 179–188, Jan. 1999, doi: 10.3141/1678-22.
[8] M. A. R. Shuvo, M. Zubair, A. T. Purnota, S. Hossain, and M. I. Hossain, "Traffic forecasting using time-series analysis," in *Proc. 6th Int. Conf. Inventive Comput. Technol. (ICICT)*, Jan. 2021, pp. 269–274, doi: 10.1109/ICICT50816.2021.9358682.
[9] G. Omkar and S. V. Kumar, "Time series decomposition model for traffic flow forecasting in urban midblock sections," in *Proc. Int. Conf. Smart Technol. Smart Nation (SmartTechCon)*, Aug. 2017, pp. 720–723, doi: 10.1109/SmartTechCon.2017.8358465.
[10] A. Nohekhan, S. Zahedian, and A. Haghani, "A deep learning model for off-ramp hourly traffic volume estimation," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2675, no. 7, pp. 350–362, Jul. 2021, doi: 10.1177/03611981211027151.
[11] Z. Khan, S. M. Khan, K. Dey, and M. Chowdhury, "Development and evaluation of recurrent neural network-based models for hourly traffic volume and annual average daily traffic prediction," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2673, no. 7, pp. 489–503, Jul. 2019, doi: 10.1177/0361198119849059.
[12] P. Lingras, S. C. Sharma, P. Osborne, and I. Kalyar, "Traffic volume time-series analysis according to the type of road use," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 15, no. 5, pp. 365–373, Sep. 2000, doi: 10.1111/0885-9507.00200.
[13] S. Das, "Traffic volume prediction on low-volume roadways: A cubist approach," *Transp. Planning Technol.*, vol. 44, no. 1, pp. 93–110, Jan. 2021, doi: 10.1080/03081060.2020.1851452.
[14] I. Tsapakis, W. H. Schneider, and A. P. Nichols, "A Bayesian analysis of the effect of estimating annual average daily traffic for heavy-duty trucks using training and validation data-sets," *Transp. Planning Technol.*, vol. 36, no. 2, pp. 201–217, Mar. 2013, doi: 10.1080/03081060.2013.770944.
[15] B. L. Smith and M. J. Demetsky, "Traffic flow forecasting: Comparison of modeling approaches," *J. Transp. Eng.*, vol. 123, no. 4, pp. 261–266, Jul. 1997, doi: 10.1061/(ASCE)0733-947X(1997)123:4(261).
[16] W. Min and L. Wynter, "Real-time road traffic prediction with spatio-temporal correlations," *Transp. Res. C, Emerg. Technol.*, vol. 19, no. 4, pp. 606–616, Aug. 2011, doi: 10.1016/j.trc.2010.10.002.
[17] F. Su, H. Dong, L. Jia, Z. Tian, and X. Sun, "Space–time correlation analysis of traffic flow on road network," *Int. J. Mod. Phys. B*, vol. 31, no. 5, Feb. 2017, Art. no. 1750027, doi: 10.1142/S0217979217500278.
[18] M. Du, L. Yang, and J. Tu, "A novel approach to calculate the spatial–temporal correlation for traffic flow based on the structure of urban road networks and traffic dynamic theory," *Sensors*, vol. 21, no. 14, p. 4725, Jul. 2021, doi: 10.3390/s21144725.
[19] X. Zhan, Y. Zheng, X. Yi, and S. V. Ukkusuri, "Citywide traffic volume estimation using trajectory data," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 2, pp. 272–285, Feb. 2017, doi: 10.1109/TKDE.2016.2621104.
[20] Y. Zheng, C. Dong, D. Dong, and S. Wang, "Traffic volume prediction: A fusion deep learning model considering spatial–temporal correlation," *Sustainability*, vol. 13, no. 19, p. 10595, Sep. 2021, doi: 10.3390/su131910595.
[21] J. Z. Zhu, J. X. Cao, and Y. Zhu, "Traffic volume forecasting based on radial basis function neural network with the consideration of traffic flows at the adjacent intersections," *Transp. Res. C, Emerg. Technol.*, vol. 47, pp. 139–154, Oct. 2014, doi: 10.1016/j.trc.2014.06.011.
[22] Y. Cui, B. Jin, F. Zhang, B. Han, and D. Zhang, "Mining spatial–temporal correlation of sensory data for estimating traffic volumes on highways," in *Proc. 14th EAI Int. Conf. Mobile Ubiquitous Syst., Comput., Netw. Services*, Nov. 2017, pp. 343–352, doi: 10.1145/3144457.3144489.
[23] X. Tang, B. Gong, Y. Yu, H. Yao, Y. Li, H. Xie, and X. Wang, "Joint modeling of dense and incomplete trajectories for citywide traffic volume inference," in *Proc. World Wide Web Conf.*, May 2019, pp. 1806–1817, doi: 10.1145/3308558.3313621.
[24] S. Jun-Fang and C. Yan, "Traffic flow state prediction based on space-time correlation of vehicle trajectory using the deep hybrid model," *J. Intell. Fuzzy Syst.*, vol. 44, no. 4, pp. 6855–6863, Apr. 2023, doi: 10.3233/JIFS-212998.
[25] W. L. Seaver, A. Chatterjee, and M. L. Seaver, "Estimation of traffic volume on rural local roads," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1719, no. 1, pp. 121–128, Jan. 2000, doi: 10.3141/1719-15.
[26] A. Sfyridis and P. Agnolucci, "Annual average daily traffic estimation in England and Wales: An application of clustering and regression modelling," *J. Transp. Geogr.*, vol. 83, Feb. 2020, Art. no. 102658, doi: 10.1016/j.jtrangeo.2020.102658.
[27] D. Apronti, K. Ksaibati, K. Gerow, and J. J. Hepner, "Estimating traffic volume on Wyoming low volume roads using linear and logistic regression methods," *J. Traffic Transp. Eng.*, vol. 3, no. 6, pp. 493–506, Dec. 2016, doi: 10.1016/j.jtte.2016.02.004.
[28] N. Caceres, L. M. Romero, F. J. Morales, A. Reyes, and F. G. Benitez, "Estimating traffic volumes on intercity road locations using roadway attributes, socioeconomic features and other work-related activity characteristics," *Transportation*, vol. 45, no. 5, pp. 1449–1473, Sep. 2018, doi: 10.1007/s11116-017-9771-5.
[29] K. Takashi, S. Yoichi, and K. Shigeru, "Proposal of estimation classification of narrow road traffic volume," *J. Traffic Eng.*, vol. 4, pp. 26–33, Jan. 2018.
[30] K. Takashi, S. Yoichi, and K. Shigeru, "The development of the estimation method of non-arterial road traffic, volume considering the dependence on arterial road traffic volume," *J. Traffic Eng.*, vol. 6, pp. 296–302, Mar. 2020.
[31] E. Baffoe-Twum, E. Asa, and B. Awuku, "Estimation of annual average daily traffic (AADT) data for low-volume roads: A systematic literature review and meta-analysis," *Emerald Open Res.*, vol. 4, p. 13, Mar. 2022, doi: 10.35241/emeraldopenres.14515.1.
[32] Z. Zhang, M. Li, X. Lin, and Y. Wang, "Network-wide traffic flow estimation with insufficient volume detection and crowdsourcing data," *Transp. Res. C, Emerg. Technol.*, vol. 121, Dec. 2020, Art. no. 102870, doi: 10.1016/j.trc.2020.102870.
[33] P. Wang, J. Lai, Z. Huang, Q. Tan, and T. Lin, "Estimating traffic flow in large road networks based on multi-source traffic data," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 9, pp. 5672–5683, Sep. 2021, doi: 10.1109/TITS.2020.2988801.
[34] M. Zhong and B. L. Hanson, "GIS-based travel demand modeling for estimating traffic on low-class roads," *Transp. Planning Technol.*, vol. 32, no. 5, pp. 423–439, Oct. 2009, doi: 10.1080/03081060903257053.
[35] S. Toi, Y. Kajita, T. Yoshitake, and H. Tatsuni, "Traffic counts arrangement for the minimum estimation error of traffic volume at non-observed road sections," *J. Eastern Asia Soc. Transp. Stud.*, vol. 6, pp. 1775–1789, Jan. 2005, doi: 10.11175/easts.6.1775.

TAKASHI KOBAYASHI received the M.S. and Ph.D. degrees in civil engineering from Chuo University, in 2014 and 2016, respectively. From 2016 to 2017, he was an Assistant Professor with the Faculty of Science and Engineering, Chuo University. Since 2018, he has been a Senior Researcher with the National Traffic Safety and Environment Laboratory, Chofu, Tokyo, Japan. His current research interests include cross-sectional traffic volume estimation of unobserved routes, estimation of potential travel demand, and improving public transportation convenience in underpopulated areas.

MASAYOSHI TANISHITA received the Ph.D. degree in civil engineering from The University of Tokyo, in 1992. From 2007 to 2008, he was an Associate Professor with the Faculty of Science and Engineering, Chuo University, Bunkyo-ku, Tokyo, Japan. Since 2008, he has been a Professor with the Faculty of Science and Engineering, Chuo University. His current research interests include social infrastructure (civil engineering, architecture, and disaster prevention), civil planning, and transportation engineering.