論文 / 著書情報
Article / Book Information

題目(和文)	角度を用いた交通システム分析：モデリングと実証分析
Title(English)	Transportation Systems Analysis with Angle: Modeling and Empirical Studies
著者(和文)	長崎滉大
Author(English)	Kota Nagasaki
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第12756号, 授与年月日:2024年3月26日, 学位の種別:課程博士, 審査員:瀬尾 亨,室町 泰徳,髙山 雄貴,花岡 伸也,沖 拓弥,朝倉 康夫
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第12756号, Conferred date:2024/3/26, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Transportation Systems Analysis with Angle: Modeling and Empirical Studies

by

Kota Nagasaki

A Dissertation

Submitted to the

Department of Civil and Environmental Engineering

School of Environment and Society

in partial fulfillment of the requirements for the degree of

Doctor of Engineering

at the

Tokyo Institute of Technology

March, 2024

Abstract

This dissertation is entitled “Transportation Systems Analysis with Angle: Modeling and Empirical Studies” by Kota Nagasaki, submitted to the Department of Civil and Environmental Engineering, School of Environment and Society in partial fulfillment of the requirements for the degree of Doctor of Engineering at the Tokyo Institute of Technology, March, 2024. The dissertation supervisor is Associate Professor Toru Seo, Department of Civil and Environmental Engineering, Tokyo Institute of Technology. The examiners are Associate Professor Toru Seo, Professor Yasunori Muromachi, Professor Shinya Hanaoka, Professor Yuki Takayama, Associate Professor Takuya Oki (Tokyo Institute of Technology) and Honorary Professor Yasuo Asakura (the University of Tokyo).

This dissertation summarizes modeling and empirical studies of angular analysis for transportation analysis. It consists of six chapters.

In Chapter 1, the background and objectives are summarized. Angle is a periodic value and appears in various aspects of the world. In transportation, the direction in which a traveler moves, the direction of the road, and the time of day are examples of angular indicators. However, angular indicators are difficult to analyze mathematically due to their periodicity. For this reason, angular indicators have not received much attention in transportation. Nevertheless, modeling with angle can contribute to the transportation systems analysis. In this dissertation, the author builds a framework for transportation systems analysis with angle by applying directional statistics, which is statistics for angular data. Specifically, behavior, temporal, and spatial transportation are modeled from an angular perspective, and the features of the proposed models are empirically analyzed using actual data.

In Chapter 2, related literature is reviewed and summarized. First, directional statistics, which is statistics to analyze angles, is summarized. Specifically, the procedure of angular data, drawing, probability distribution, and regression analysis in directional statistics are explained, and examples of applications other than transportation are

summarized. Next, the transportation analysis is classified into demand, supply, behavior, and transportation states. The elementary analytical methods and the recent research on each concept are summarized. In addition, studies that employed angular analysis for each concept are reviewed.

In Chapter 3, a route choice model with an angular indicator was built and validated by empirical studies. In discrete choice models, the turn angle is often described by setting dummy variables as right/left turns or U-turns. However, the results vary depending on the predetermined threshold value. In addition, although both behaviors are essentially changes of travel direction, the two are rarely analyzed as continuous variables without separating them. Therefore, the model describes route choice behavior with the angular indicator, named the turning angle at an intersection. The route choice model is applied to the actual route choice behavior of drivers observed in Tokyo. The results showed that the angular indicator significantly affected drivers' route choice behavior.

In Chapter 4, a model of the daily variation of traffic volume by converting time into angle in order to represent periodicity was built. Since the time in a day is looped in 24 hours, conversion of the time into angles is necessary to describe the time variation. Traffic volume varies with time and has two peaks: morning peak on the way to work and evening peak on the way home. The shape of the two peaks is significantly different, with the morning peak rising quickly and sharply while the evening peak is gentle. In describing these variations in traffic volume over time using a probability distribution, a circular distribution that can represent multimodality and flexible shapes must be employed. Therefore, the mixture of Kato–Jones distribution that satisfies this property was applied to the time variation of traffic volume. Furthermore, a quick and stable estimation method for the mixed Kato–Jones distribution was built. The stability of the estimation method was assessed and interpreted when the time variation was divided into multiple parts.

In Chapter 5, a model that describes the states of OD demand for a certain direction by the directional distribution of overall demand and the road network was built. Describing the entire demand pattern within an area is difficult by conventional approaches. By

aggregating the angles between the origin and destination of each demand, the overall demand pattern can be quantitatively described. In addition, the relationship between transportation states and the morphology of road network, such as whether the road network is grid-like or not, is also difficult to describe using conventional methods. The distribution of the direction of each road in an area can intuitively and quantitatively describe the morphology of the road network. In this chapter, a novel descriptive model that captures the transportation state by overall demand pattern and network morphology is modeled. In addition, the proposed model is validated by transportation data for Tokyo.

In Chapter 6, achievements and future directions are summarized. The achievements of this research are the establishment of a framework for transportation systems analysis with angles by building models for each concept of transportation based on angle. In addition, the proposed models are validated through empirical studies, and the potential of transportation systems analysis with angles is revealed. The future directions include more precise analysis using multidimensional angular variables.

Acknowledgements

The traffic counter data was provided by the Hanshin Expressway Co. Ltd., Japan. The probe data was freely provided by Tokyo Metropolitan Government (”offering of several datasets on mobility and transportation in Tokyo 2020 Games”). Mr. Zhong Hengyi, a doctoral student at Tokyo Institute of Technology, helped prepare network data and route choice data.

Parts of this research are supported by the Research Fellow (DC2) program of the Japan Society for the Promotion of Science entitled “Propose a method for evaluating road networks using directional statistics.” (KAKENHI Grant-in-Aid for JSPS Fellows 22J11294).

This research could not have been achieved without the support and advice of many people. First of all, I would like to thank my supervisor, Prof. Toru Seo. He gave me valuable opinions about my research and told me how to be a researcher and what I should do in the future with his words and attitude. In addition, I would like to express my gratitude to Prof. Yasuo Asakura (The University of Tokyo), who was my supervisor until the first year of the doctoral program. He has given me a great deal of support not only in my research but also in my daily life. If he had not been my first supervisor, I would not have decided to become a researcher. I would also like to express my gratitude to Prof. Wataru Nakanishi (Kanazawa University), who was an assistant professor in the Asakura laboratory. He gave me a weapon of angle and helped me a lot in analytical methods and their applications. Furthermore, I would like to thank Prof. Shogo Kato (The Institute of Statistical Mathematics) for teaching me much about directional statistics.

I would like to thank Prof. Tetsuo Yai, Prof. Yasunori Muromachi, Prof. Shinya Hanaoka, Prof. Yuki Takayama, Prof. Takuya Oki, Dr. Choi Sunkyung, Dr. Kashin Sugishita, Dr. Riki Kawase, and Dr. Takara Sakai for advice in seminar presentations and in daily conversation.

Furthermore, the members of Asakura laboratory and Seo laboratory, who accompanied me on daily talks, meals, trips, and other leisure time during my six years in the

laboratory, were my emotional support during my hard research life. In addition, I would like to thank my old friends, members of my baseball team, and Ramen Jiro, who have colored my life outside of my work. I would further like to thank numerous people who kindly supported and warmly encouraged me during this research life.

Last but not least, I would like to express my great appreciation to my mother, who raised me all by herself and allowed me the unfilial piety of entering a doctoral program.

Contents

List of Figures

List of Tables

List of Abbreviations

Following abbreviations are used throughout in this dissertation.

OD	Origin–Destination
FD	Fundamental Diagram
UE	User Equilibrium
SO	System Optimum
NFD	Network Fundamental Diagram
MFD	Macroscopic Fundamental Diagram
RL	Recursive Logit
MoKJ	Mixture of Kato–Jones distribution
ETM	Empirical and Theoretical trigonometric Moments
EM	Expectation and Maximization

Chapter 1 Introduction

1.1 Background

Angle is a periodic value and appears in various aspects of the world. Angles are often used to describe the direction of objects. For example, wind direction is a typical *angular indicator*, which is the indicator described on the angular axis. In addition, time in a day and date in a year can also be described in terms of angles because time and date are looped in a day and a year, respectively. Therefore, they can correspond to angles one-to-one.

Angles also appear in various aspects in transportation, which widely covers the movement of people and geographical objects. Transportation consists of a complex interplay of various factors. They can be classified roughly into four concepts: *demand*, *supply*, *behavior*, and *state*. They are used with the following meanings throughout in this dissertation.

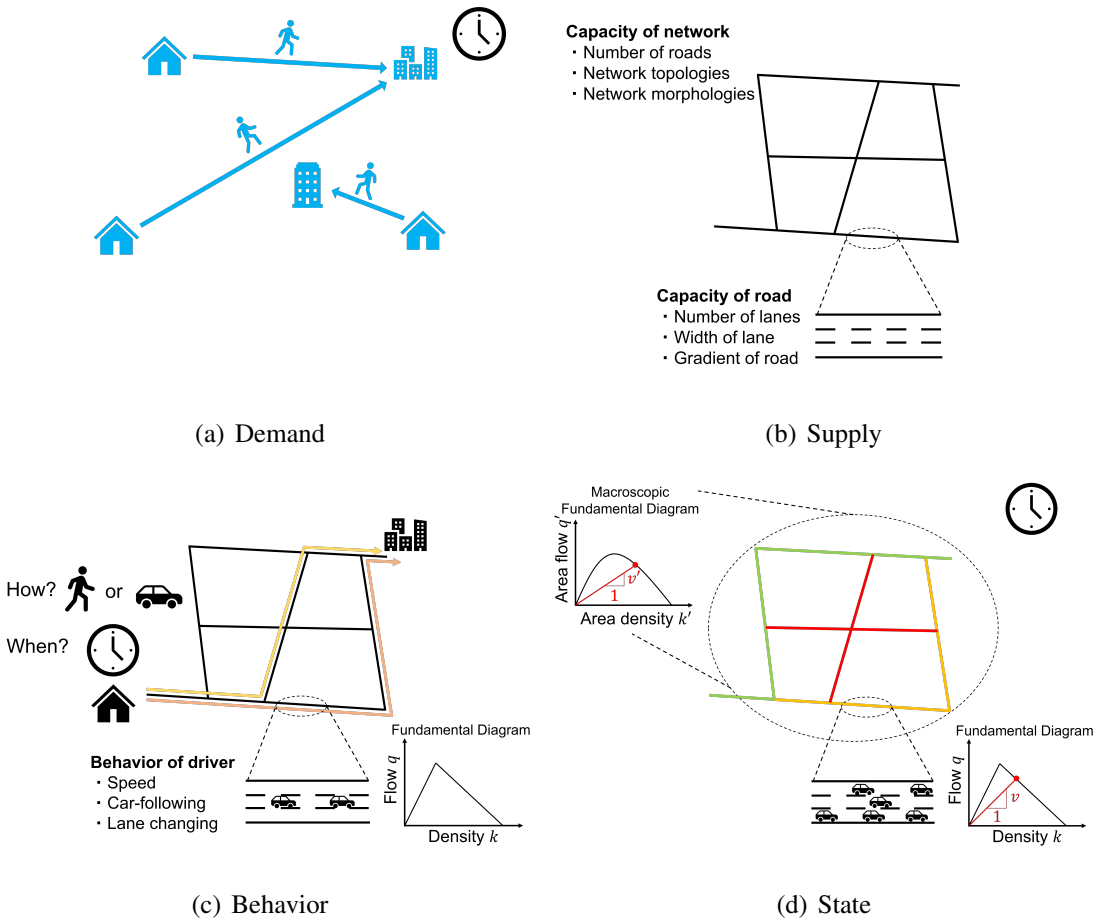
Demand More specifically, collective demand. Collection of intention of movement to achieve an activity.

Supply Links or networks and their performances used by transportation.

Behavior More specifically, individual choice behavior. The way an individual chooses an action, such as route choice.

State The results of transportation from the demand, supply, and behavior of an individual or a group, such as speed and travel time.

A diagram of these four concepts in transportation is shown in Fig. 1.1. In addition, angles appeared in each concept are shown in Fig. 1.2. Angles appear in the demand, supply, and behavior as they can describe the direction of objects, the direction in which people move, and its change. Furthermore, since demand and state constantly changing, they can be analyzed as angular indicators by converting the time variable into angular



(a) Demand (b) Supply

(c) Behavior (d) State

Figure 1.1: Four concepts in transportation.

variables.

Limited number of studies have analyzed transportation with angular indicators. A few examples are studies explained route choice behavior with angular indicators (Dalton, 2003; Turner, 2007) and analyzed the road network and demand using angles (Mohajeri and Gudmundsson, 2012; Boeing, 2019; Zhou, 2015). Moreover, intra-day traffic volume variations and travelers' seasonal variations have specific patterns of concentrating users at certain periods (Transportation Research Board, 2000).

One reason for the limited number of studies with angular indicators is that despite the versatility of angles, they are difficult to analyze mathematically due to their periodicity. The traditional statistical methodology cannot be applied since the angles loop in 360

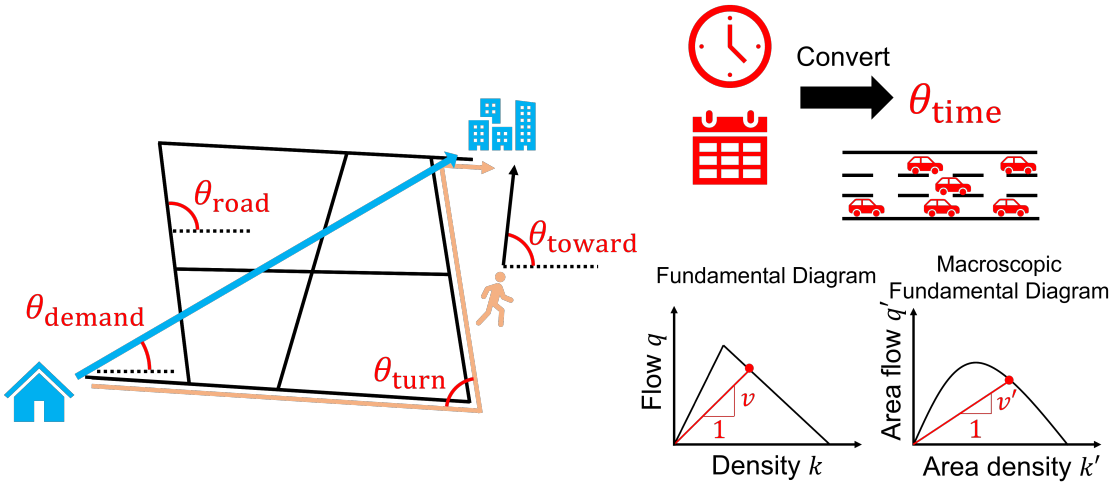


Figure 1.2: Angles in transportation.

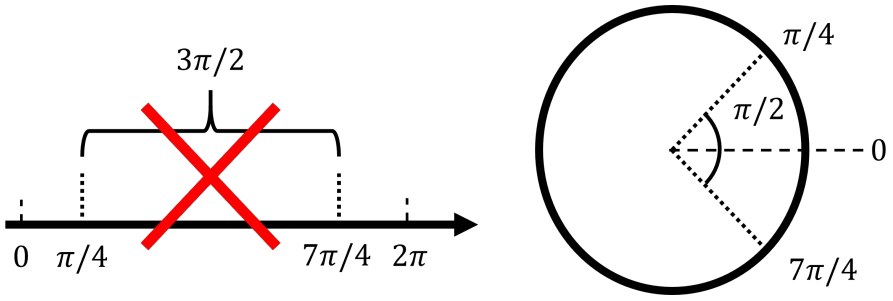


Figure 1.3: The discrepancies in the calculation of angle due to periodicity.

degrees or 2π radians. For example, the difference between $\pi/4$ and $7\pi/4$ radians is not $3\pi/2$ radians, but $\pi/2$ radians is appropriate due to its periodicity as shown in Fig. 1.3. In addition, the usual method of calculating means for $\pi/4$ and $7\pi/4$ radians results in π radians, which is also counterintuitive. Because of the difficulty of mathematical analysis, angular indicators have not been used frequently or correctly in transportation, which requires mathematical analysis in general.

The *directional statistics* (Mardia and Jupp, 2000) is a branch of statistics that is specifically developed to analyze *angular data*, which is the data described on the angular axis. In directional statistics, angular data is treated as a unit vector with its angle. In addition, probability distributions and regression methods that adjust for periodicity have been developed. Some studies in other fields, such as meteorology (Bowers et al., 2000)

or biology (Morales et al., 2004), have analyzed mathematically angular data by applying directional statistics.

Modeling with angle can contribute to the transportation systems analysis, as it has been successfully applied in other fields. Specifically, the following two contributions are possible:

- Simplified description of phenomena
- Universal modeling without preset condition

First, angular indicators enable a simple description of some phenomena that are difficult to represent with existing methods. Angular indicators can describe aggregate patterns with only one-dimensional information. The description of the positional relationship between two points, such as OD, requires at least two-dimensional information of vertical and horizontal displacement. In some cases, only one-dimensional information of the angle between two points is sufficient to understand a phenomenon. For example, only the angular information is necessary to determine whether the demand pattern is concentrated or not, or whether the road network is grid-like or not. Several studies have shown that these representations can visually classify demand patterns and road networks (Mohajeri and Gudmundsson, 2012; Boeing, 2019; Zhou, 2015). On the other hand, to grasp these patterns without using angles, higher dimensional information is required. In addition, quantifying the pattern of angular indicators facilitates mathematical comparisons with other indicators.

Second, direct analysis with angles by using the techniques of directional statistics enables modeling that is independent of preset conditions. Conventional studies with angles have solved the periodicity of angular variables by discretizing them with dummy variables or by cutting the angular axis and converting it to a linear axis. These approaches produce different results depending on the preset thresholds. For example, Mohajeri and Gudmundsson (2012) tried to estimate the angular distribution by extracting a specific section of the angular axis instead of the entire axis to avoid problems caused by periodicity. In the route choice model, Fosgerau et al. (2013) analyzed angle changes

as dummy variables of right/left turns and U-turns. In addition, when analyzing time series data over a day or a year, in some cases, converting the data to angles is better to consider periodicity. For example, by estimating the time variation of traffic volume at a given point as a probability distribution, drivers' behaviors could be detected from easily observable roadside traffic volume data. Conventional statistical methods cannot perform this statistical analysis because they cannot represent the periodicity of time. In order to model such time-varying data, it is indispensable to apply angular analysis, which can represent periodicity.

1.2 Objectives

The primary objective of this dissertation is to build a framework for transportation systems analysis with angle. Modeling with angle can contribute to the transportation systems analysis, as it has been successfully applied in other fields. In order to maximize the potential of angular analysis, the knowledge of directional statistics is used to leverage it to the fullest extent.

Specifically, angle-based models are built for transportation phenomena at three scales: traveler route choice behavior, within-day transportation dynamics, and area-wide transportation state. Then, the features of the proposed models are empirically analyzed using actual data.

This dissertation is organized as follows. Chapter 2 reviews the literature on related topics to this dissertation: directional statistics, demand, supply, behavior, and state in transportation. Chapter 3 builds a behavior model with the angle that deals with individual route choice behavior with angular indicator, named turn angle. Chapter 4 builds a temporal transportation model with the angle that applies a mixture probability density function of angles to daily variations in transportation. Chapter 5 builds a spatial transportation model with angle that describes traffic state by two angular distributions that describe overall demand pattern and network morphology. Finally, Chapter 6 concludes the achievements and future direction of the proposed methodology and angular

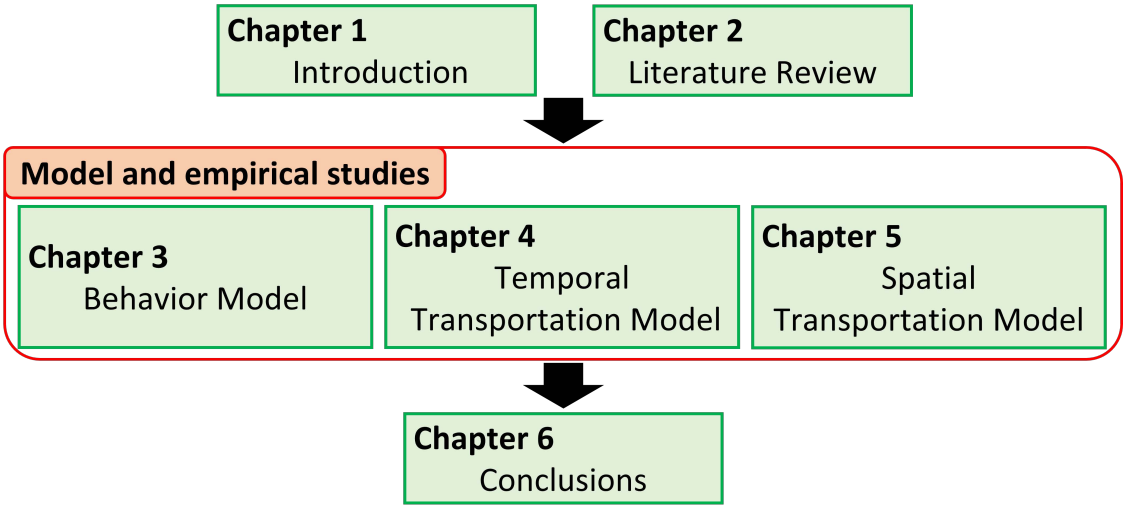


Figure 1.4: The structure of this dissertation.

analysis in transportation. Fig. 1.4 illustrates the structure of this dissertation.

Chapter 2 Literature Review

This chapter reviews the literature on related topics to this dissertation. Section 2.1 summarizes the concept, probability distribution, regression model, and application of directional statistics. Section 2.2, Section 2.3, Section 2.4, and Section 2.5 review the existing studies for demand, supply, behavior, and state analysis in transportation, respectively, which are the fundamental concepts in transportation. Each of them is reviewed separately with conventional studies and studies using angles. Finally, Section 2.6 summarizes the reviews. Note that angles are written in radians unless otherwise noted.

2.1 Directional statistics

2.1.1 Basis of directional statistics

(1) Preliminaries

Directional statistics (Mardia and Jupp, 2000) is the statistics for processing angular data. Angular data has a periodicity that repeat once in 360 degrees or 2π radians.

Due to the periodicity, angular values 0 and 2π, or $-\pi/4$ and $7\pi/4$ represent essentially the same. This periodicity causes trouble in conventional statistical processing. For example, the difference between $\pi/4$ and $7\pi/4$ on a number line is $3\pi/2$. However, since $-\pi/4$ (rad) equals $7\pi/4$ (rad), the angular difference between $\pi/4$ and $7\pi/4$ derives two different values: $3\pi/2$ and $\pi/2$.

Directional statistics solves the problem of periodicity by regarding angular data as a unit vector with its angle. In other words, the angular data of θ is processed as $(\cos\theta, \sin\theta)^\top$ or $z = e^{i\theta} = \cos\theta + i\sin\theta$, where $^\top$ means the transpose of vector. For example, an angular data with $\pi/3$ and π are converted into $(1/2, \sqrt{3}/2)^\top$ and $(-1, 0)^\top$,

respectively. This conversion is independent of the domain of the angles and always results in the same vector. The difference in angles is obtained by taking the angle between the two vectors, i.e., the inner product. In the case of $\pi/4$ and $7\pi/4$, each vector is $(1/\sqrt{2}, 1/\sqrt{2})^\top$ and $(1/\sqrt{2}, -1/\sqrt{2})^\top$. Since the inner product of the two vectors is zero, the difference is θ with $\arccos\theta = 0$, i.e. $\theta = \pi/2$.

(2) Mean and variance of angular data

Mean and variance are defined by employing composite vectors. Suppose that sample set consists n angular samples θ_j, where $j = 1, 2, \ldots, n$. In this case, the composite vector of all samples $(\bar{C}, \bar{S})^\top$ is calculated as

$$\bar{C} = \frac{1}{n}\sum_{j=1}^{n}\cos\theta_j, \quad \bar{S} = \frac{1}{n}\sum_{j=1}^{n}\sin\theta_j. \tag{2.1}$$

Therefore, the sample mean direction $\bar{\theta}$ is the solution of the equations

$$\bar{C} = \bar{R}\cos\bar{\theta}, \quad \bar{S} = \bar{R}\sin\bar{\theta}, \tag{2.2}$$

where $\bar{R}$ is called the sample mean resultant length that is given by

$$\bar{R} = (\bar{C}^2 + \bar{S}^2)^{1/2}. \tag{2.3}$$

$\bar{R}$ takes the value between 0 and 1. If the samples are concentrated around the sample mean direction, the composite vector is larger and $\bar{R}$ takes larger values. For consistency with the usual statistics, the sample angular variance V is calculated by subtracting $\bar{R}$ from 1, i.e.,

$$V = 1 - \bar{R}. \tag{2.4}$$

V takes the value between 0 and 1.

The mean and variance of $\pi/4$ and $7\pi/4$ are taken as an example as shown in Fig. 2.1. The composite vector of the two vectors is $(2/\sqrt{2}, 0)^\top$. The sample mean direction of them is 0 and the sample angular variance is $1 - 1/\sqrt{2} \approx 0.293$.

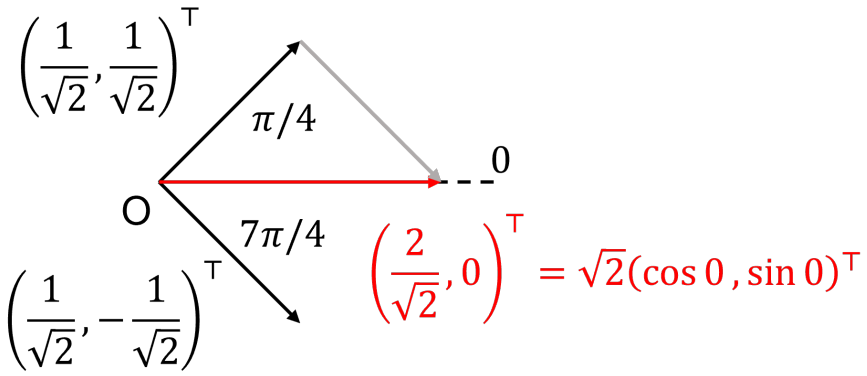


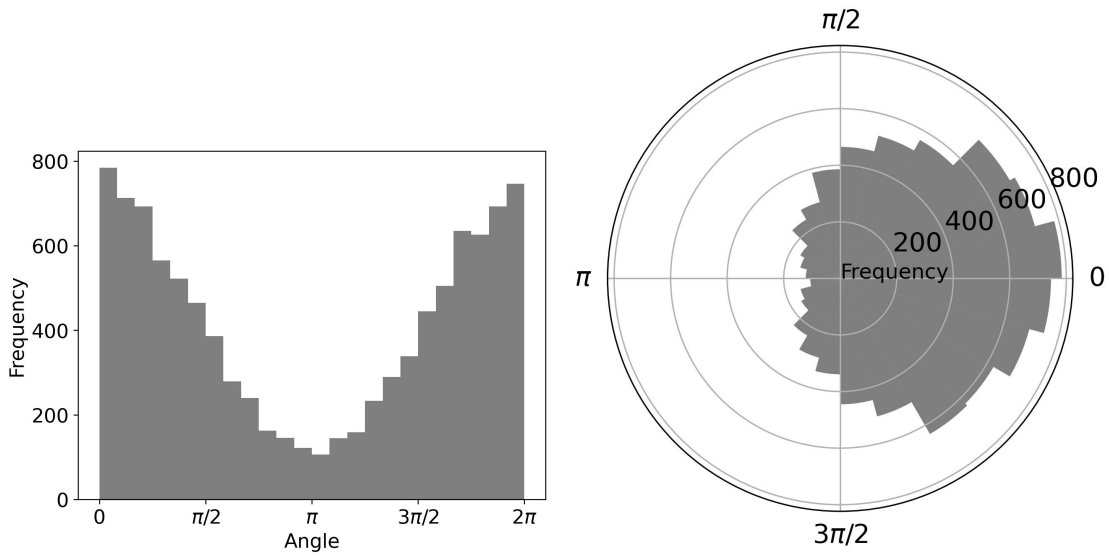
Figure 2.1: Example of calculation of mean and variance for angular data.

(3) Drawing angular data

When depicting angular data, the periodicity should also be taken into account. The usual histogram, in which both ends of the histogram are separated from each other, may not be able to capture the trend of angular data correctly. For this reason, a rose diagram, a circular graph with both ends of the histogram connected, is often used to depict angular data. Fig. 2.2 shows examples of the usual histogram and rose diagram. In the usual histogram without both edges connected, two peaks are seen, while in the rose diagram, with both edges connected, only one peak can be clearly seen.

(4) Multidimensional angular descriptions

The procedure for angular data on the circle in $\mathbb{R}^2$ is discussed above. On the other hand, there are also descriptions to process angular data of higher dimensions. First, angular data with a p-dimensional direction is represented by a unit vector in p-dimensional Euclidean space $\mathbb{R}^p$. A typical example is the direction of a star seen from the earth, which is 3-dimensional angular data. In addition, when p equals 2, the data is on the circle described above. Next, the data with two independent angular data, such as the wind direction at a given time, are represented as data on a torus. A torus is the surface of a donut, which can have two independent angular axes as shown in Fig. 2.3. Finally, the data with linear and angular axes, such as wind speed and direction, are represented



(a) Usual histogram (b) Rose diagram

Figure 2.2: Examples of usual histogram and rose diagram.

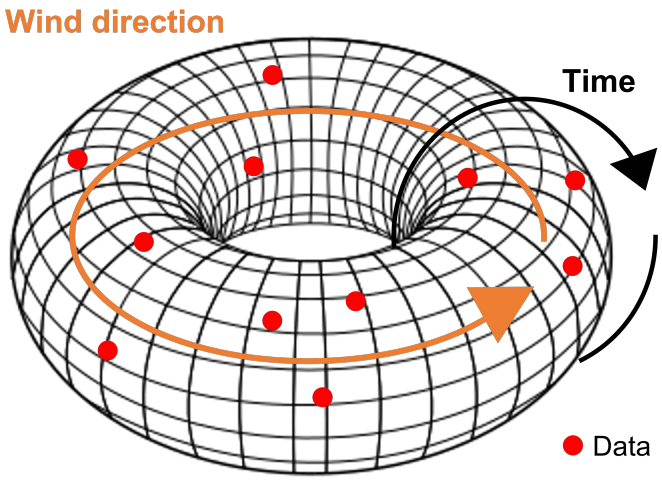


Figure 2.3: Example of torus.

as data on a cylinder. The cylinder is the surface of the cylinder and can have a linear axis and a angular axis independently as shown in Fig. 2.4.

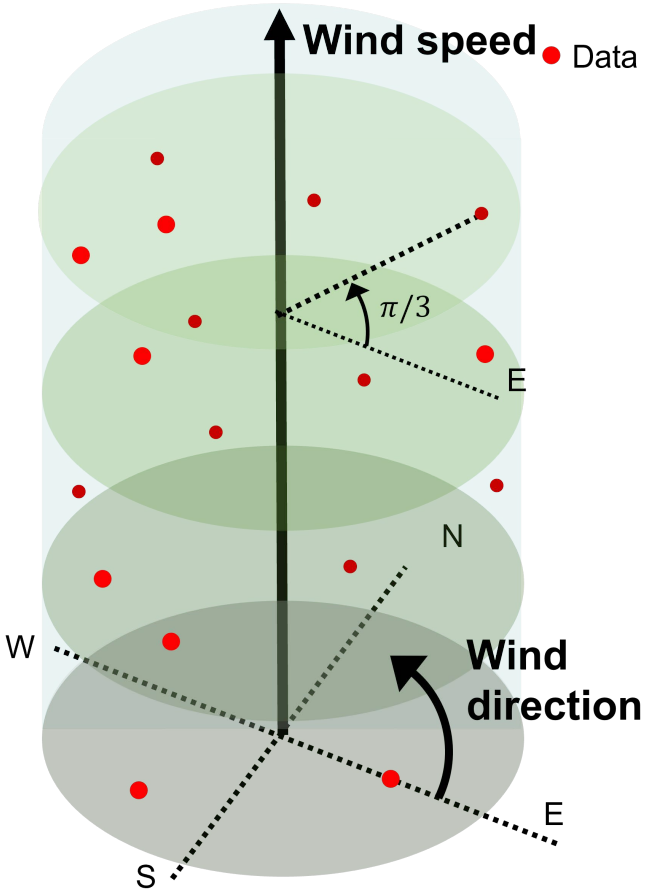


Figure 2.4: Example of cylinder.

2.1.2 Circular distribution

(1) Properties of circular distribution

In order to represent the periodicity of angles, a distribution function on angular axis $f(\theta)$ must satisfy the following three properties.

1. $f(\theta) \geq 0$ everywhere on $(-\infty, \infty)$
2. $f(\theta) = f(\theta + 2\pi)$ everywhere on $(-\infty, \infty)$
3. $\int_0^{2\pi} f(\theta)d\theta = 1$

A distribution that satisfy these three properties is called circular distribution.

(2) von Mises distribution

The von Mises distribution (von Mises, 1918) is the most representative circular distribution. The probability density function of the von Mises distribution $M(\mu, \kappa)$ is

$$g(\theta; \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu)}, \tag{2.5}$$

where I_0 denotes the modified Bessel function of the first kind and order 0, which can be defined as

$$I_0(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} e^{\kappa \cos\theta} d\theta, \tag{2.6}$$

where μ is the mean direction and κ is a parameter that manipulates concentration; the larger the value, the sharper the peaks of the distribution. The ranges of parameters are $0 \leq \mu < 2\pi$ and $\kappa \geq 0$, respectively. The shape of the distribution is unimodal and symmetrical. The von Mises distribution has played a central role in directional statistics because of its similarity to the normal distribution on the linear axis.

(3) Wrapped distribution

The wrapping is one of the ways to convert from a probability distribution on the linear axis to a circular distribution. The wrapping corresponds a random variable x on the linear axis to a random variable x_w on the circular axis by

$$x_w = x \pmod{2\pi}. \tag{2.7}$$

Therefore, a wrapped distribution $f_w(\theta)$ can be obtained from a probability density function $f(\theta)$ by

$$f_w(\theta) = \sum_{k=-\infty}^{\infty} f(\theta + 2\pi k). \tag{2.8}$$

Many wrapped distributions have been proposed, such as the wrapped Poisson distribution (Lévy, 1939) and the wrapped Normal distribution (de Haas-Lorentz, 2013). However, most of them are difficult to handle because they include infinite series in their probability

density functions. The wrapped Cauchy distribution (Lévy, 1939) is one of the wrapped distributions whose probability density function can be described in a form that does not include an infinite series. The probability density function of the wrapped Cauchy distribution $WC(\mu, \rho)$ is

$$c(\theta; \mu, \rho) = \frac{1}{2\pi} \frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos(\theta - \mu)}, \tag{2.9}$$

where μ is the mean direction and ρ is the mean resultant length. The ranges of parameters are $0 \leq \mu < 2\pi$ and $0 \leq \rho < 1$, respectively. The shape of the distribution is unimodal and symmetrical.

(4) Kato–Jones distribution

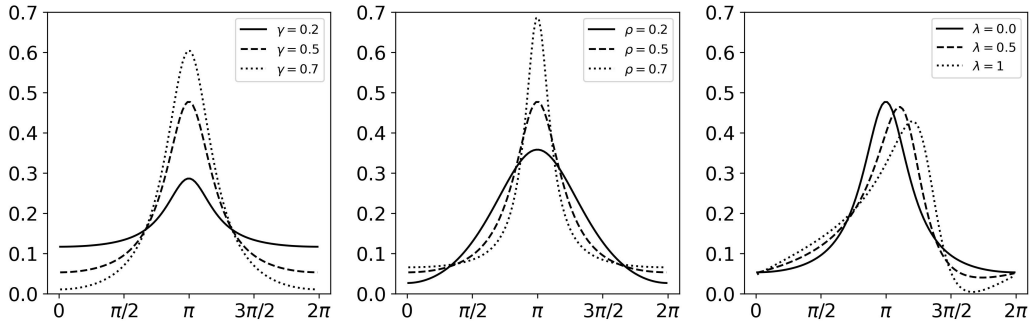
Kato–Jones distribution (Kato and Jones, 2015) is an extended distribution based on the wrapped Cauchy distribution, and its probability density function is

$$g_{\mathrm{KJ}}(\theta; \mu, \gamma, \lambda, \rho) = \frac{1}{2\pi} \left\{ 1 + 2\gamma \frac{\cos(\theta - \mu) - \rho \cos \lambda}{1 + \rho^2 - 2\rho \cos(\theta - \mu - \lambda)} \right\}. \tag{2.10}$$

The parameter μ and γ controls the location or mean direction, the concentration or the mean resultant length, respectively. The parameters λ and ρ influence the skewness and kurtosis of the distribution. The ranges of the parameters are $0 \leq \mu < 2\pi$, $0 \leq \gamma < 1$, $0 \leq \rho < 1$, and $0 \leq \lambda < 2\pi$ satisfying $(\rho \cos \lambda - \gamma)^2 + (\rho \sin \lambda)^2 \leq (1 - \gamma)^2$. Four parameters allow for flexible representation of changes in distribution shape, including kurtosis and skewness as shown in Fig. 2.5. In addition, the ease of parameter estimation and interpretation are advantages of the Kato–Jones distribution. Note that the use of the Kato–Jones distribution is recommended in recent review papers of Ley et al. (2021) and Pewsey and García-Portugués (2021) as a flexible circular distribution.

2.1.3 Regression

In regression, the periodicity of angular indicators must also be taken into account. There are three types of regressions involving angular variables: (1) regression with angular



(a) Difference by γ (b) Difference by ρ (c) Difference by λ

Figure 2.5: Difference in distribution shape by parameter values. The value of parameters is $\gamma = 0.5, \rho = 0.5, \lambda = 0$ if not noted.

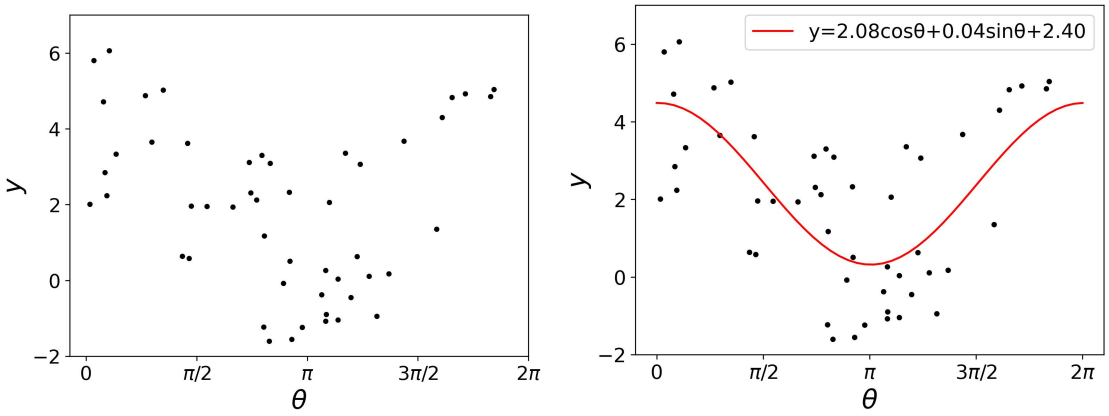
inputs and linear outputs, (2) linear inputs and angular outputs, and (3) both inputs and outputs angular, respectively. (2) was formulated in Fisher and Lee (1992) and Gill and Hangartner (2010), and (3) was formulated in Downs and Mardia (2002), Kato et al. (2008) and Kato and Jones (2010). Fig. 2.6(a) plots the data with the angular variable θ and the linear variable y on the horizontal and vertical axes, respectively. Johnson and Wehrly (1978) proposed a regression model using trigonometric functions to take into account the periodicity of angles as

$$\hat{y} = \alpha + \beta_1 \cos\theta + \beta_2 \sin\theta. \tag{2.11}$$

Eq. (2.11) allows for regression with angular variables as explanatory variables as shown in Fig. 2.6(b). In addition, higher-order regression is possible in the angular regression model as well as polynomial regression in usual regression models as

$$\hat{y} = \alpha + \sum_{k=1}^{K} \beta_{ck} \cos k\theta + \sum_{k=1}^{K} \beta_{sk} \sin k\theta, \tag{2.12}$$

where K is an external parameter, an integer that determines up to which degrees the distributions are combined. The regression model enables the reflection of minor trends in the sample by adding short-period trigonometric functions as Fourier series expansion.



(a) Plot for regression (b) Result of regression

Figure 2.6: The plot and result of regression with angular inputs and linear outputs.

2.1.4 Application

Analysis with angles is effective in various fields and contributes to the understanding of phenomena. A typical application of angular analysis is in meteorology, which focuses on wind direction and other factors. Wind direction has historically been analyzed using the findings of directional statistics because it provides rich angular data (Stephens, 1969). Recently, more advanced directional statistical techniques have been applied to the study of wind direction. Carta et al. (2008) proposed a probability density function on the cylinder that simultaneously describes wind direction and wind speed and showed better fits than existing models by case studies using actual weather data. Modlin et al. (2012) and Jona-Lasinio et al. (2012) analyzed spatial-scale wind and wave direction data by a hierarchical Bayesian model that incorporates temporal trends and wind autocorrelation, and showed better fit than existing models. Lang et al. (2020) proposed a regression model of angular input and output using regression tree and random forest, and applied it to wind direction forecasting. The techniques of directional statistics have provided significant advances in the analysis of meteorology.

In addition, directional statistics is often employed to analyze the direction of animal movement. Many studies have applied the findings of directional statistics to random

walks in animals and plants (Codling et al., 2008). Rivest et al. (2016) constructed a model using the concept of random walk and regressed the direction of movement of animals and plants on variables such as the travel distance traveled. Morales et al. (2004) and McClintock et al. (2012) developed a model in which multiple random walk patterns can switch and classify different behavior patterns. The model was applied to the movements of elk and grey seals, and daily movements were classified into two categories: those with large angle changes and short travel distances and those with long, straightforward travel distances. These models focus on the movement of animals and are closely related to transportation, which deals with the movement of people.

Analyses applying directional statistics are frequently performed in fields other than those mentioned above. For example, bioinformatics research for probabilistic modeling of protein structure (Boomsma et al., 2008), astrophysics research to find hints of dark matter (Morgan et al., 2005), acoustics research on the human ability to detect sound sources (McMillan et al., 2013), and image analysis for modeling three-dimensional rotation with spherical functions (Esteves et al., 2018).

2.2 Demand in transportation

2.2.1 Conventional approach

(1) OD matrix

Transportation demand estimation and measurement are essential to the proper design of transportation facilities. In particular, it is important to identify demand between arbitrary locations.

OD matrix is The traditional method for describing demand between locations. This method describes the demand between locations using a matrix whose (i, j) elements are the demand between locations i and j.

The four step model (McNally, 2007) is a representative method for obtaining an OD matrix. This model estimates how many people will flow to each transportation

mode and route within an area. Each step is performed in the following order: trip generation/concentration, mode assignment, distribution, and network assignment, or trip generation/concentration, distribution, mode assignment, and network assignment. Mode assignment and network assignment are explained in detail in Section 2.4. Trip generation/concentration estimates the number of trips originating from a certain zone and the number of trips destined for a certain zone. These are calculated based on socioeconomic indicators such as the daytime and nighttime population of the zone. In estimating the trips between zones, the gravity model (Anderson, 2011) or entropy model (Wilson, 2011) is often used based on the generated and concentrated traffic volumes of the two zones. The number of trips between each zone is summarized in an OD matrix. These methods have shown some good results in the analysis of transportation demand estimation.

Rather than estimating demand, there are also methods to measure demand. A national travel survey is conducted every few years to gather information on trip objectives, start time, age, gender, license status, and other characteristics (e.g., National Household Travel Survey in U.S. (Santos et al., 2011)).

Many technologies can help estimate the demand by observing a part of the movement. GPS data from mobile phones is rich demand data because many people carry it all day (Huang et al., 2019). In addition, taxi, Uber and bike-sharing services can also provide rich data with information on the location and time of the user (Yao et al., 2018; Dong et al., 2018; Sathishkumar et al., 2020). In public transportation, IC cards provide convenience to users and automatically obtain a massive amount of user information, such as the origin and destination (Pelletier et al., 2011). Although these data are not potential traffic demand because they observe the achieved movement, they are used in many analyses in transportation planning.

(2) Demand analysis

Most studies on demand are focused on its forecasting and the reasons for its variation. Toole et al. (2015) forecast travel demand using big data based on noisy mobile phones,

geospatial data, and census data. Yao et al. (2018) proposed a method to forecast demand from taxi demand data using the Deep Multi-View Spatial-Temporal Network (DMVST-Net) framework to model both spatial and temporal relations, and showed that it is more effective than existing methods. Eren and Uz (2020) reviewed the reasons for variations in demand for bike sharing and categorized factors, such as environmental ones including weather, location of the station, and age group.

(3) Variation of traffic volume

The traffic volume is defined as the number of vehicles traveling in certain time period, such as 15 minutes, an hour or a day. The traffic volume varies depending on the time of day and the season (Transportation Research Board, 2000). In particular, traffic volume is large during the morning commute and evening return trip, and small during the midnight period in a city. This trend results from commuters, who account for a large share of urban transportation. Especially for commuting to work, the departure time choice is closely related to the variation in a day, and a large amount of transportation is concentrated in the wide time period around the start of the working (Arnott et al., 1990). In addition, in tourist areas, the daily traffic volume varies depending on the season (Butler, 1998).

These variations have been confirmed by demand perspectives and traffic volume observations. Furthermore, due to the importance of understanding trends in traffic volume, many studies have been conducted to forecast traffic volume by time of day or date (Smith and Demetsky, 1997; Vlahogianni et al., 2014; Lana et al., 2018; Hong, 2011).

(4) Demand visualization

Visualization of demand is discussed in several studies. Since there is no location information in the OD matrix, another visualization method is preferred to capture demand trends. Wood et al. (2010) drew the OD demand over a large area while preserving the positional relationship between origin and destination by describing the

OD as cells rather than vectors. Andrienko et al. (2017) reviewed methods for visualizing demand and flow on the network, summarized various visualization methods and their properties, and explained the importance of visual analytic in transportation.

2.2.2 Angular approach

For the analysis of variation of traffic volume, a study written in Korean has only conducted the angular analysis, which used a mixed circular distribution to estimate the daily variation of transportation (Na and Jang, 2011).

For visualization, Zhou (2015) clustered massive bike-sharing data in Chicago spatiotemporally. In this study, each trip is described as the angle of a vector from the origin to the user's destination. The angles between the origin and destination of the flows belonging to each cluster were tabulated, and differences in trends were found between clusters and time. Hamedmoghadam et al. (2019) clustered the travel direction data of 16 million taxis by the characteristics of their origin and destination. The differences in the characteristics of the travel direction and travel distance for each cluster were discussed. Andrienko et al. (2016) proposed a visualization method using the direction of movement as a way to abstract origin-destination movement data, which is high dimensional data. They proposed several circular plotting methods and evaluated their visibility through a questionnaire survey. In addition, they performed visualization of actual data and discussed their advantages and disadvantages.

The angular analysis of demand is often used to visualize and classify trends in the overall direction of movement of demand, which is challenging to identify in OD matrices or plots on a map.

2.3 Supply in transportation

2.3.1 Conventional approach

(1) Capacity of supply

Supply in transportation means the links and networks that the transportation uses, and their performance (Button, 2010). Performance includes the capacity, which is the amount of transportation that can flow per unit of time, and the frequency of operations in public transportation. In vehicular transportation, supply refers the road or road network and its performance, such as its capacity. These have been widely analyzed for a long time.

The capacity of a road depends on the number of lanes, lane width, design speed, horizontal-vertical alignments, and other factors (Transportation Research Board, 2000). It is difficult to determine the capacity analytically using these factors. Therefore, traffic capacity is measured by observing actual transportation (Greenshields, 1935). In addition, Fundamental Diagrams (FDs), which describe the relationship between transportation density and flow, can also be used to calculate transportation capacity analytically. The detail of FD is mentioned in Section 2.5.

In addition, the capacity of road networks has also been studied extensively (Yang and H. Bell, 1998). The Macroscopic Fundamental Diagram (MFD), which describes relationships between flow and density on the road network scale, is a particularly successful theory with respect to the capacity of the road network (Geroliminis and Daganzo, 2008). The capacity of the road network can be obtained from the density and flow rate relationships at the road network scale in this model, as in the usual FD. The detail of MFD is also mentioned in Section 2.5.

(2) Network analysis

In addition to capacity, there are many indicators that can evaluate the performance of a road network. In particular, analyses based on network science (Barabási, 2013) have been conducted with great success. In these analyses, road networks are represented

as networks with intersections as nodes and roads between intersections as edges. This representation has long been performed by Garrison (1960), Kansky (1963), and others, however, the limited computational resources made complex analysis impossible. In recent years, advances in computational and observational techniques have enabled topology-based analysis of complex real-world road networks. For example, Latora and Marchiori (2001) compared the shortest route lengths and linear distances between all node pairs in a real road network to evaluate network efficiency. Boeing (2020) calculated tabulation measures such as road density and intersection density within a region for a network of 27,000 U.S. cities. Moreover, the vulnerability of the network is also an important indicator. Jenelius et al. (2006) discussed the vulnerability of the entire network by calculating an indicator of link importance based on the amount of increase in travel costs when links are unavailable. Furthermore, centrality is one of the indicators to measure the importance of a node or link, and has been applied in many transportation network analyses. Lämmer et al. (2006) showed a clear hierarchical order of roads as the frequency distribution of centrality based on travel time follows a power law. Crucitti et al. (2006) evaluated the road networks of 18 cities by degree, closeness, straightness and information centrality.

2.3.2 Angular approach

Angular indicators on the supply often focus on the direction of the road. This road direction is referred to as orientation hereafter.

Swan and Sandilands (1995) is the first study to apply the rose diagram, a histogram in directional statistics, to the visualization of road orientation. Mohajeri and Gudmundsson (2012) focused on the orientation of the road network in Dundee, East Scotland. They compared the road network with and without weighting by the length of the roads, and also compared a rose diagram drawn from all the roads in the area and that of only the named major roads. In addition, the city was divided into several parts and rose diagrams were compared by district. In this study, the rose diagram is segmented into

multiple angular zones when estimating the angular distribution. However, this method is not a correct procedure for angular data. Boeing (2019) drew rose diagrams of the orientation of each road in 100 cities around the world using the Open Street Map road data. He calculated entropy from the rose diagram and proposed an indicator of how close to the grid the city's road network is. Then, he analyzed the trends of the indicators by region, such as Asia and Europe, and perform regressions of these indicators. Nagasaki et al. (2019) quantified the shape of the road network around the station using the orientation of each road. They estimated the probability density function of the distribution of orientation of all roads to quantify the road network, and clustered the networks around stations using the parameters. Fig. 2.7 shows the examples of road networks and corresponding rose diagrams. The left road network has a grid shape as a whole, and each peak in the rose diagram is very sharp. On the other hand, the right road network is partially grid-like, yet its inclination varies from place to place, and each peak in the rose diagram are gentler than those in the left one.

2.4 Behavior in transportation

2.4.1 Conventional approach

(1) Discrete choice model

In the discrete choice model, it is assumed that an individual makes a choice with the maximum utility from a set of alternatives (Train, 2009). The random utility U_{ij} of individual i for alternative j is expressed as

$$U_{ij} = V_{ij} + \varepsilon_{ij}, \tag{2.13}$$

where V_{ij} is the deterministic utility that can be systematically explained out of U_{ij}, and ε_{ij} is an error term that includes the effects of unobserved factors and individual preferences. Let $j \in \{1, 2\}$ be the choice set, the probability $p_i(1)$ that individual i chooses alternative $j = 1$ is expressed by the assumption that individual i chooses the alternative with the

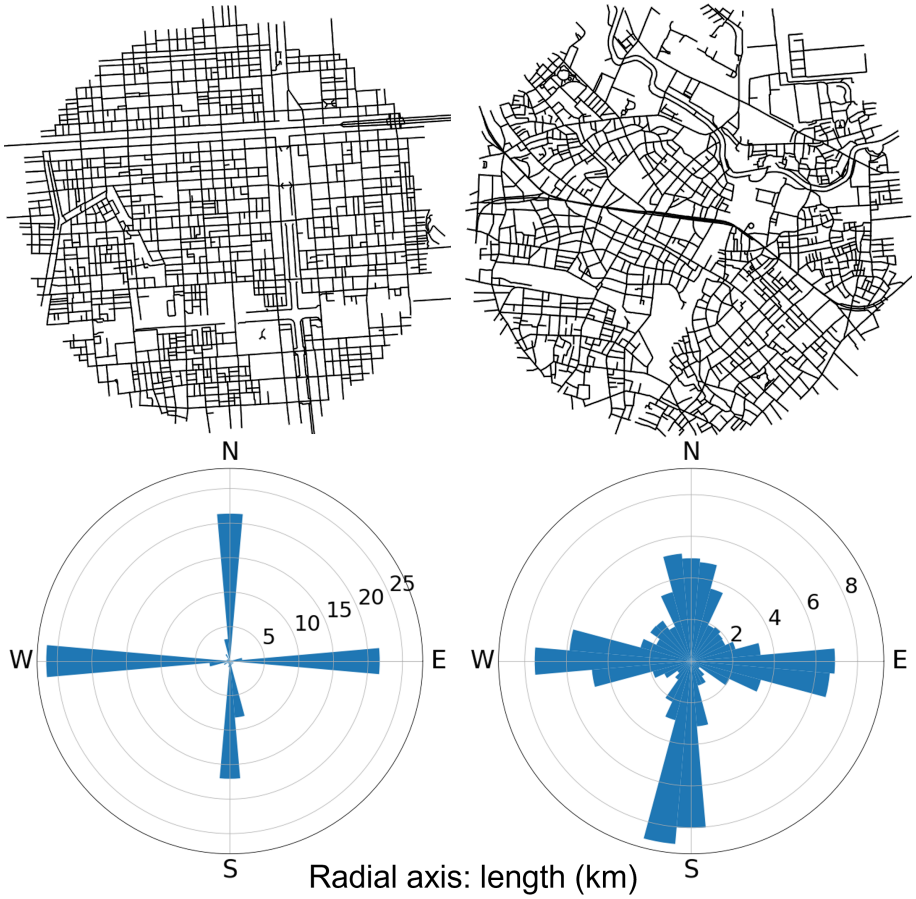


Figure 2.7: Examples of road networks and corresponding rose diagrams.

maximum utility as

$$\begin{aligned} p_i(1) &= \Pr[U_{i1} > U_{i2}] = \Pr[V_{i1} + \varepsilon_{i1} > V_{i2} + \varepsilon_{i2}] \\ &= \Pr[\varepsilon_{i2} - \varepsilon_{i1} > V_{i1} - V_{i2}] = F[V_{i1} - V_{i2}], \end{aligned} \tag{2.14}$$

where $\Pr[A]$ is the probability that event A occurs and F is the cumulative distribution function of $\varepsilon_{i2} - \varepsilon_{i1}$. A model that assumes a standard normal distribution for F is called the *probit model*, and a model that assumes a Gumbel distribution is called the *logit model*. In particular, the logit model with three or more alternatives is called the *multinomial logit model*, and many extended models have been proposed, e.g., the nested logit model, which solves the ratio problem of alternatives, and the mixed logit model, which can express error terms more flexibly.

(2) Route choice model

The discrete choice model is also used for route choice. However, the number of routes on the network is uncountable. Therefore, a model that requires enumeration of a set of alternatives is not directly applicable. Instead of a route-based model, a link-based route choice model is proposed in which the traveler chooses a link at each intersection (Fosgerau et al., 2013; Mai et al., 2015).

Note that these choice behaviors correspond to the mode assignment and network assignment of the four step model mentioned in Section 2.2.

In vehicular route choice models, travel distance, tolls, and travel time have traditionally been employed as utility functions (Yamamoto et al., 2002; Ben-Akiva et al., 2004; Prato, 2009). In addition, the number of right/left turns and U-turns are also regarded as essential factors. Duckham and Kulik (2003) proposed *simplest path* that minimizes the number of turns from origin to destination. Fosgerau et al. (2013) incorporates right/left turns and U-turns as dummy variables in the model validation. Furthermore, many different factors have been incorporated into the function, such as socio-economic characteristics (Ramming, 2001), information of traffic (Mahmassani and Liu, 1999), and latent variables such as attitudes toward risk (Sun et al., 2012), memory, spatial ability, and so on (Prato et al., 2012). These variables have been confirmed to influence route choice behavior in many models and case studies.

(3) Departure time choice

The departure time choice is also a part of behavior in transportation. Travelers choose a departure time based on congestion and a cost function for early or late arrivals to arrive at their destination at the desired arrival time, such as the commute to work time. These problems originated as single-bottleneck problems (Vickrey, 1969) and have been extended to the analysis of multi-bottleneck networks (Kuwahara and Newell, 1987) and multiple bottlenecks on a corridor (Akamatsu et al., 2015), and so on.

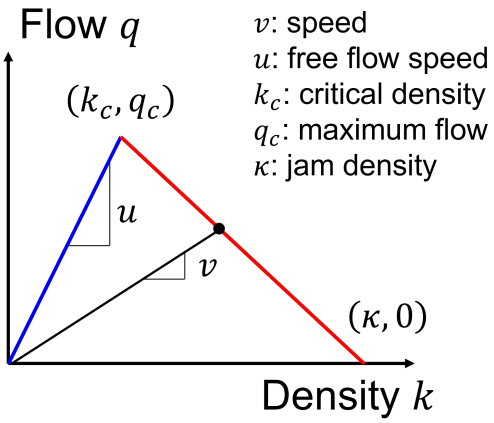


Figure 2.8: Example of triangular FD.

(4) Fundamental Diagram

The car-following behavior on the road is a microscopic transportation behavior. Drivers follow vehicles in front of them, and disturbances in the movement of one vehicle propagate to the following vehicles. Pipes (1953) described this phenomenon using differential equations, and Kometani and Sasaki (1958) developed a car-following model that includes a delay time for the reaction. This model is the basis for the macroscopic representation of transportation flow.

Based on this theory, a Fundamental Diagram (FD) describing the relationship between flow, velocity, and density was developed. Greenshields (1935) plotted the relationship between flow and density by observing actual transportation flows. Thereafter, FDs were described theoretically using car-following model, and various shapes of FDs have been proposed. One of the most widely used is the triangular FD proposed by Newell (1993) as shown in Fig. 2.8. In the density region below the critical density k_c, flow q is linearly increasing with density k, and above k_c, q is decreasing with increasing k. The flow q_c at critical density k_c is called the maximum flow, and the blue line below k_c is called the free flow state, while the red line above k_c is called the congested state.

2.4.2 Angular approach

Dalton (2003) experimentally observed subjects' route choice behavior in a virtual reality space and analyzed the human ability in terms of spatial comprehension. Each subject travels through a complex road network without the information of a route for his/her destination but the direction to the destination from the initial point. The experiment results show that subjects do not make large turns at each intersection and choose the straightest path possible in the direction of their destination. Turner (2007) proposed *shortest angular path model* that minimizes the sum of turning angle from origin to destination. He found some correlation between the observed link traffic on the London road network and the betweenness centrality calculated by the model. Manley et al. (2015) compared the routes between actually chosen routes and minimize some simple indicators between an origin and destination by 677,411 route data obtained from 2,970 minicabs in London. The comparison results showed that the route that minimizes the product of the number of left-right turns and the travel distance was the most consistent with the route actually chosen. In addition, the time-based model deviated more from the actual route choice than the other models.

Akamatsu (1996) introduced a rotation angle, which is the angle formed between links, in the traffic assignment model under the premise of the importance of considering the geometric characteristics of routes. Subsequently, a method was developed to classify paths based on the rotation angle, distinguishing whether they include cycles. This approach ensured that only routes without cycles were reflected in the model. In this model, the angle was used only to determine whether a cycle is included or not, and the effect of the turning angle itself on route choice was not discussed.

Raveau et al. (2011) introduced an angular indicator called angular cost in the analysis of metro users' route choices. Angular cost is the difference between the direction of the destination and the direction of travel after changing metro lines. The estimation of the model using the data of metro users showed that angular cost significantly affects the choice behavior. Antonini et al. (2006) analyzed the short-term movement choice

behavior of pedestrians using a discrete choice model. They divided the pedestrian's movement into several segments based on distance and angle. The utility of choice was modeled using distance, change in direction of travel, and angle to the direction of the objective. They showed that the angle has a significant effect on the pedestrian's movement. Nakanishi and Fuse (2015) analyzed pedestrian traffic line with applying directional statistics. They regarded traffic line as data consisting of the pedestrian's position and angle of moving direction at each time. A correlation between pedestrian position and direction of travel was discovered by basic analysis of pedestrian trajectories in front of ticket gates, and a regression model was constructed. Bongiorno et al. (2021) developed a vector based model for route choice behavior of pedestrians on the network, which includes both distance and direction to the destination of a pedestrian. The proposed model showed better prediction of pedestrian routes in two cities with different road networks than a probabilistic model only based on distance.

2.5 State in transportation

2.5.1 Conventional approach

(1) Variables and their relationships on a link

State in transportation is the result of demand, supply, and behavior. The following three quantities are often used to describe transportation states.

flow q Number of vehicles traveling per unit of time.
density k Number of vehicles per unit distance.
speed v Average speed of vehicles.

In addition, their inverses, spacing s, headway h, and pace p, are sometimes used. Note that the following relation must be satisfied.

$$q = kv. \tag{2.15}$$

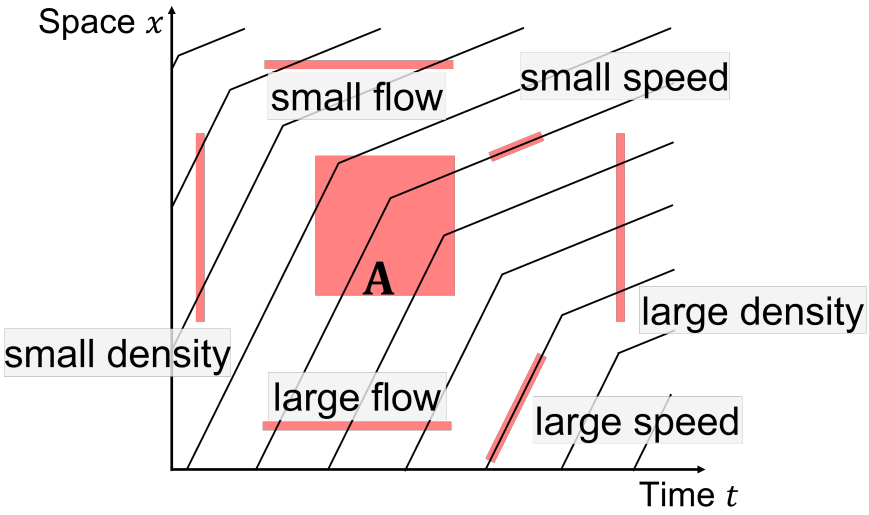


Figure 2.9: The time space diagram.

Therefore, the degrees of freedom for these three variables are two. By using the FD shown in Fig. 2.8, one variable gives the remaining two variables.

Fig. 2.9 shows the trajectory of several cars traveling in a certain spatio-temporal region. This diagram is called a time-space diagram, with time on the horizontal axis and position on the vertical axis, and is very important for visualizing transportation states. The number of trajectories across horizontal and vertical lines of unit length corresponds to flow and density, respectively. In addition, the slope of the trajectory corresponds to the speed. To obtain the three variables satisfying Eq. (2.15), Edie et al. (1963) proposed the following definition.

$$q(\mathbf{A}) = \frac{d(\mathbf{A})}{|\mathbf{A}|}, \tag{2.16}$$

$$k(\mathbf{A}) = \frac{t(\mathbf{A})}{|\mathbf{A}|}, \tag{2.17}$$

$$v(\mathbf{A}) = \frac{d(\mathbf{A})}{t(\mathbf{A})}. \tag{2.18}$$

where $\mathbf{A}$ is the region in the time-space diagram, $d(\mathbf{A})$ is the total distance traveled by all vehicles in $\mathbf{A}$, $t(\mathbf{A})$ is the total time spent by all vehicles in $\mathbf{A}$, and $|\mathbf{A}|$ is the time–space area of $\mathbf{A}$.

(2) Transportation network analysis

Once the transportation states of a road can be estimated, the cost of the traffic can be calculated. Based on the cost per link, it is possible to assign the OD demand to each mode and route. Wardrop (1952) defined the User Equilibrium (UE) in the assignment when all travelers choose the route with the minimum cost based on complete information as,

> *The journey times in all routes actually used are equal and less than those that would be experienced by a single vehicle on any unused route.*

This is called Wardrop's first principle and is widely used in assignments. On the other hand, the System Optimum (SO) is the state in which the traveler's total travel time is minimized. This is called Wardrop's second principle and is also widely used in assignments. In general, the UE and SO assignments are not matched, and individual travelers can reduce their own costs in the SO condition. In addition, while these assignments are static, dynamic assignments have also been extensively analyzed in order to more closely describe realistic phenomena. For dynamic assignment, Dynamic System Optimum (DSO), Dynamic User Optimum (DUO), and Dynamic User Equilibrium (DUE) assignment principles have been proposed (Kuwahara and Akamatsu, 1993; Merchant and Nemhauser, 1978; Kuwahara and Akamatsu, 1997).

(3) Area-based approach

An area-based transportation states model has also been proposed. The continuous modeling approach (Beckmann, 1952) regards the areal traffic flow as a two-dimensional flow and analytically obtains the traffic volume at an arbitrary point. It has the advantage of handling fewer variables because it does not need to capture traffic flow for each road, but it has problems in capturing actual transportation phenomena because it is difficult to handle multiple OD pairs.

The Network Fundamental Diagram (NFD) (Mahmassani et al., 1984) is an extension of the fundamental diagram defined on a link to the network scale. NFD can derive

the relationships between flow, density, and velocity in a network scale using the same procedure as FD, but it is not a universal model because it requires many constraints on network geometry and transportation control and so on.

Geroliminis and Daganzo (2008) revealed the existence of a Macroscopic Fundamental Diagram (MFD) linking space-mean flow, density and speed based on transportation data observed in Yokohama in Japan. Some conditions exist for the existence of an MFD, and Geroliminis and Sun (2011) summarizes those conditions.

(4) Transportation state models

The transportation state is a dynamic phenomenon that changes temporally. Therefore, dynamic models are often constructed in modeling transportation states. However, modeling transportation phenomena dynamically sometimes involves significant obstacles in formulation and implementation, and many static models have been proposed.

Transportation state analysis at a link scale is often modeled using the knowledge of hydrodynamics, in which each individual vehicle or segment is regarded as a single object. In this context, such analysis is usually referred to as "traffic flow analysis". This modeling allows the transportation flow on a link to be described from a macroscopic perspective rather than focusing on individual vehicles. The kinematic wave theory (Lighthill and Whitham, 1955) is a representative method for modeling transportation flows on a link. This theory can precisely describe transportation phenomena such as traffic congestion and queue generation and extension. As analytical methods for this theory, the minimum envelop principle (Newell, 1993) and the variational theory (Daganzo, 2005) have been proposed, and the cell transmission model (Daganzo, 1994) has been proposed as a numerical method for the theory.

Transportation states analysis at a network scale is also important and widely analyzed (Sheffi, 1985). In transportation network analysis, static traffic flow on a link or the overall network performance is often focused on. For example, studies to identify critical links based on the amount of change in travel time across the network (Scott et al., 2006; Sullivan et al., 2010), studies analyzing the relationship between the impact of

transportation accidents on the transportation network and the network features explained by graph theory (Sun et al., 2018). In addition, some studies have aimed at achieving SO states by imposing congestion pricing on transportation states in UE states (Yan and Lam, 1996; De Palma and Lindsey, 2011). Furthermore, many studies focus on public transportation as well as vehicular transportation (Daganzo, 2010; Farahani et al., 2013).

In addition, area scale transportation states have also been extensively analyzed using the continuous modeling approach (Sasaki et al., 1990; Wong, 1998), the network fundamental diagram (Williams et al., 1987), and MFD (Ji et al., 2014; Saeedmanesh and Geroliminis, 2017; Dantsuji et al., 2022). These analysis methods allow to reduce variables, improve computational efficiency, and understand dynamical macroscopic transportation states by capturing transportation flows on a complex network from an areal perspective.

2.5.2 Angular approach

In area-based studies, several studies have independently analyzed demand or supply using angles, as described above. However, to the author's knowledge, no studies have used angles for transportation state analysis, which corresponds to the results of demand and supply.

2.6 Summary

Directional statistics has developed as statistics for analyzing angles. It has been applied to many fields, and contributed to the elucidation of phenomena, as described in Section 2.1. On the other hand, transportation has been analyzed from various aspects for a long time, as described in Section 2.2 to Section 2.5.

For each aspect, the analysis by angle is conducted sparsely. Naturally, the analysis with angle still leaves room for further development.

A method focusing on the direction of each demand has been developed, drawing rose

diagrams to grasp and classify demand trends. However, no attempts have been made to evaluate these trends quantitatively. Quantitative evaluation of demand trends can model the relationship with traffic conditions. Nevertheless, no attempts have been made to model traffic state using this idea as well quantitatively.

The effect of the turning angle on route choice behavior has been shown. However, most analyses that evaluate the economic worth of angular change compared with monetary or metric indices depend on a preset threshold value set by the analyst. Direct use of angles is necessary to evaluate the impact of angle change with a universal method that does not depend on a preset threshold value.

In addition, the analysis using knowledge of directional statistics, such as multidimensional angular variables and regression analysis, has been conducted in many fields. On the other hand, very few studies of this kind have been conducted in the field of transportation. In particular, time series data used in the field of transportation seem to be compatible with directional statistics, which deals with periodic indices.

In order to fulfill the three gaps mentioned above, this dissertation builds an angular route choice model, a temporal traffic model, and a spatial traffic model, and investigates the potential of transportation systems analysis with angle.

Chapter 3 Behavior Model with Angle: Individual Route Choice Behavior

3.1 Introduction

The choice of explanatory variables is crucial for the route choice model. Therefore, various explanatory variables are discussed in the literature (Prato, 2009). Among geographic variables, metric indicators such as the length of links and routes are widely used.

On the other hand, angular indicators are rarely used in the route choice model. For example, indicators such as the direction in which the driver is currently heading and the direction of the destination can be easily recognized by drivers. Dalton (2003) experimentally observed subjects' route choice behavior in a virtual reality space and analyzed the human ability in terms of spatial comprehension. Each subject travels through a complex road network without the information of a route for his/her destination but the direction to the destination from the initial point. The experiment results show that subjects do not make large turns at each intersection and choose the straightest path possible in the direction of their destination.

In particular, the indicator of the change in direction between neighboring links, i.e., the turn angle, influences route choice behavior. Several studies have analyzed the effect of turn angle on route choice. Duckham and Kulik (2003) proposed *simplest path* that minimizes the number of turns from origin to destination. They showed that the simplest path is 16% more traveled than the path that minimizes the travel distance. Turner (2007) proposed *shortest angular path model* that minimizes the sum of turn angle from origin to destination. He found some correlation between the observed link traffic on the London

This chapter is mainly based on the research with Prof. Toru Seo, published in *EWGT 2023* (Nagasaki and Seo, *in press*) and *Journal of JSCE* (Nagasaki and Seo, *in press*).

road network and the betweenness centrality calculated by the model. Manley et al. (2015) compared the routes between actually chosen routes and minimize some simple indicators between an origin and destination by 677,411 route data obtained from 2,970 minicabs in London. The comparison results showed that the route that minimized the product of the number of left-right turns and the travel distance was the most consistent with the route actually chosen. In addition, the time-based model deviated more from the actual route choice than the other models. Raveau et al. (2011) introduced an angular indicator called angular cost in the analysis of metro users' route choices. Angular cost is the difference between the direction of the destination and the direction of travel after changing metro lines. The estimation of the model using the data of metro users showed that angular cost significantly affects the choice behavior.

As described above, the behavior of the drivers' angle change is often processed as a dummy of a right/left turn or a U-turn. In this case, the analyst sets a threshold value and regards any angular change greater than this as a right/left turn or a U-turn. However, the results of the estimation are varied by the threshold of the dummy. In contrast to discrete processing that sets a threshold value, continuous processing that uses the turn angle directly enables analysis of angle-changing behaviors, such as a right/left turns and a U-turn, independent of the threshold value setting. In addition, these models reproduced the actual route choice better than models that minimize the travel distance simply. However, there is no discussion of the ratio of the parameter of the travel distance variables to that of the turn angle variable. Furthermore, there is no discrete route choice model that employs turn angle on road network in a utility function. Raveau et al. (2011) analyzed discrete choice model with introducing angular cost, however, this is route choice model in metro. By using a distance variable in addition to the turn angle in the utility function, the resistance to angle change relative to distance change can be quantified from the ratio of the parameters.

The objective of this chapter is to build a route choice model on a network based on angle. Specifically, the proposed route choice model incorporates the turn angle directly in addition to the distance variable in the utility function. This model allows to evaluate

the effect of angle change on route choice and the resistance of angle change compared to distance change, independent of a preset threshold value. The chapter is organized as follows. Section 3.2 describes a route choice model on a road network that uses angular and distance indicators. Section 3.3 explains the data used to validate the model. In Section 3.4, the results of estimating the parameters of the model in Section 3.3 using the data in Section 3.2 are shown. In Section 3.5, the proposed model is compared with other models. Section 3.6 concludes this chapter.

3.2 Route Choice Model

3.2.1 The definition of turn angle

It is assumed that the angle at which a driver turns at a point is θ and define it as *the difference of angles between the direction of travel before the turn and after the turn*. Fig. 3.1 shows the definition of θ. The domain of θ is set to $(-\pi, \pi]$. Because angular variables are periodic and difficult to handle, θ is converted to a non-periodic variable x_θ by the Eq. (3.1) (Hillier and Iida, 2005).

$$x_\theta = 1 - \cos\theta, \tag{3.1}$$

This value is 0 when there is no turn, and 1 when the turn is square. A larger turn angle results in a larger value of x_θ. Taking the cosine of θ solves the discontinuity and makes it a positive variable for all θ by subtracting it from 1.

3.2.2 Recursive Logit model

Recursive Logit model Fosgerau et al. (2013), which is one of the link-based route choice models, is employed as the route choice model (hereafter referred to as the RL model). In the RL model, drivers determine the next link a on the link k where drivers are, from among the set of multiple links $A(k)$ connected to link k. In other words, an individual n maximizes the utility described by the sum of the instantaneous utility $u_n(a|k) = v_n(a|k) +$

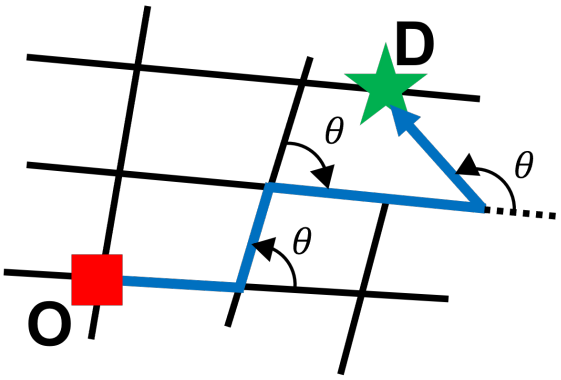


Figure 3.1: The definition of the turn angle θ.

$\mu\varepsilon_n(a)$ and expected downstream utility to his/her destination d denoted as

$$V_n^d(k) = E[\max_{a \in A(k)} v_n(a|k) + V_n^d(a) + \mu\varepsilon_n(a)], \tag{3.2}$$

where $v_n(a|k)$ is a deterministic term, μ is a scale parameter, and $\varepsilon_n(a)$ is a error term.

Link-based models have the advantage of low computational cost because they do not require restrictions on the choice set. In addition, turn angle, which are calculated from the angle of links, are compatible with link-based route choice models. For these reasons, the RL model is used, which is a representative link-based route choice model.

The utility function $v(a|k)$ of choosing the next link a on link k is defined as Eq. (3.3).

$$v(a|k) = \beta_l x_{l(a)} + \beta_\theta x_{\theta(a|k)}, \tag{3.3}$$

where $x_{l(a)}$ is the length of link a (in meters), $x_{\theta(a|k)}$ is the variable of the turn angle from link k to the next link a, and β_l and β_θ are their parameters. $x_{l(a)}$ and $x_{\theta(a|k)}$ can be easily obtained by the information of road network. Because $V(a|k)$ is less than or equal to 0 due to the property of RL model, β_l and β_θ are negative values.

3.2.3 Consideration

This model is very simple, using only link lengths and angles between neighboring links. The reason for using link lengths instead of travel times is that it is difficult to calculate accurate travel times in this analysis, which uses only probe data.

In the link-based route choice model, travelers maximize the utility of the entire route between the origin and destination. In other words, travelers minimize the sum of turn angles and the sum of travel distance at each intersection based on each parameter. Since drivers would be expected to take the shorter travel distance throughout the route choice and not make many direction changes, both the parameter β_l for the link length and the parameter β_θ for the turn angle variable are expected to be negative and significant.

The ratio of the two parameters β_θ/β_l can be regarded as *the cost of converting one right-angle turn into β_θ/β_l meter travel* since $x_{\theta(a|k)} = 1$ for a right-angle turn. Furthermore, this model can evaluate the effect of turn angles independent of preset thresholds since it uses continuous variables rather than dummy variables. Therefore, this model enables to evaluate the resistance of arbitrary turn angles, which is a unique feature of this model.

Dalton (2003) suggested that the behavior of moving toward a direction close to the direction of the destination is important in route choice. However, the purpose of this model is to evaluate the effect of the turn angle on the route choice behavior, and to obtain the ratio of the turn angle to the distance traveled. Furthermore, since the travel distance between the origin and destination shortens when the moving direction is closer to the direction of the destination at each intersection, this behavior is expected to be substitutable with the travel distance variable. Therefore, a model with only turn angle and travel distance is used.

3.3 Data

Vehicle trajectory data obtained within a 10 km square area in the southern part of Tokyo on August 18, 2021 (Wed.) is used as route choice data. The vehicle's position, speed, and time are recorded when the vehicle travels 200 meters or the vehicle's travel direction changes by 45 degrees or more. Note that the threshold values for vehicles using the old type of onboard device are 100 m and 22.5 degrees, respectively. The network data is generated from only the trajectory data (Zhong et al., 2022; Zhong, 2023; Zhong et al.,

2023). The following procedure was conducted to deal with the data.

1) Extract trips that pass through the region or are completed within the region from the vehicle trajectory data.
2) The vehicle trajectory data extracted in the process 1) is used to form a road network using the (Zhong et al., 2022; Zhong, 2023; Zhong et al., 2023) method, and map-matching is performed. The data after map-matching is hereafter referred to as trip data.
3) Delete the data before the first node and after the last node from each trip data, making the first node the origin of the trip and the last node the destination of the trip.
4) Delete trip data that passes through the same node more than once.
5) For each link, remove trip data that passes through links whose traffic volumes of probe car in the day below 100.
6) Delete and replace links that obviously do not match reality in comparison with the actual map, and reflect them in the trip data.

1) and 2) are used to obtain road network data and trip data (i.e., which link each vehicle is traveling on) from the vehicle trajectory data alone. 3) clarifies the origin and destination of the trip. 4) excludes data that obviously do not represent a normal route choice, such as detours. 5) removes trip data that obviously passes through places where no roads exist. 6) is used to correct data that skipped several intersections. The road network generated by the above procedure is shown in Fig. 3.2. The thickness of the links corresponds to the daily traffic volume of the probe car. The number of nodes is 382, the number of links is 836, and the number of trips is 107,280. As described above, the method of Zhong et al. can generate a network consisting of only those links that are important for route choice.

In addition, the actual road network is shown Fig. 3.3. In the area, arterial roads extend radially from the city center in the northeast, and circular arterial roads are installed to connect the radial roads.

In this method, links below a certain traffic volume are not reflected in the network

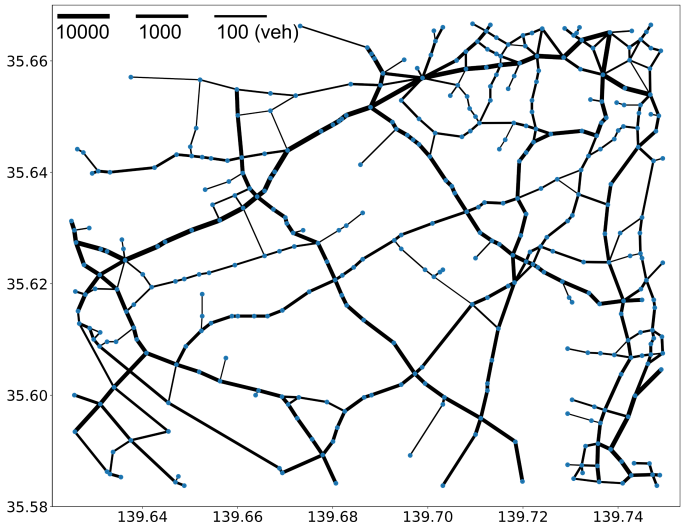


Figure 3.2: Road network (horizontal axis: longitude, vertical axis: latitude)

Figure 3.3: Map of the subject area (GSI: Geographical Survey Institute map with author's additions)

data, which causes some links to be unnaturally disconnected. In addition, some intersections are represented by a single node. Therefore, the estimated parameters and the simulation described below may not reflect the actual phenomenon.

Note that the turn angle θ is calculated from the difference between the angles of the two endpoints of the link before and after the turn. For this reason, this chapter ignores any curvature of road. In addition, the data around the origin and destination of

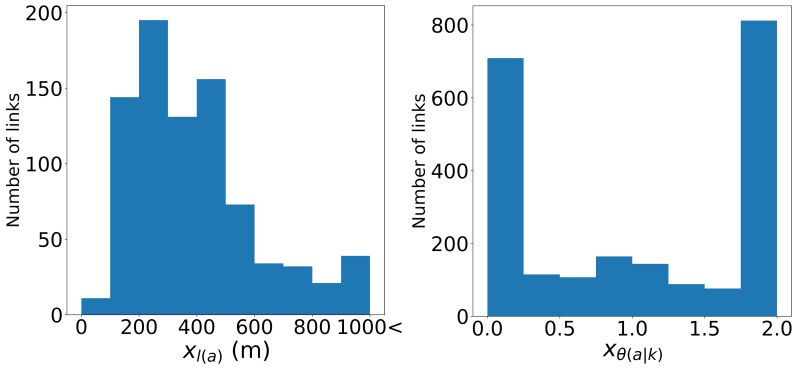


Figure 3.4: Histogram of $x_{l(a)}$ (left) and $x_{\theta(a|k)}$ (right)

the trajectory data has been deleted for the anonymization. Therefore, the right/left turn behavior on the local streets near the origin and destination cannot be reflected, and β_θ may be overestimated.

3.4 Results

3.4.1 Basic Analysis

Histograms of all $x_{l(a)}$ and $x_{\theta(a|k)}$ for the network data in Section 3.3 are shown in Fig. 3.4. The average of $x_{l(a)}$ is 405.4 m and the maximum is 2483.3 m. About 80% of the links are distributed between 100 m and 500 m in length. $x_{\theta(a|k)}$ is distributed around 0 and 2. This means that the links go straight and make U-turns. Since a directed graph is used, U-turns are possible when links in both directions are connected between nodes, and the corresponding frequencies around 2 are also high. In addition, the distribution around 1, which means a right-angle turn, is also distributed to some extent, with about 20% distributed between $0.7 < x_{\theta(a|k)} < 1.3$. Note that $0.7 < x_{\theta(a|k)} < 1.3$ means a left-right turn of approximately 70° to 110°. Moreover, since $x_{\theta(a|k)}$ is widely distributed from 0 to 2, drivers can make with some freedom of angular change. Therefore, the estimation of β_θ is not expected to be affected by the shape of the road network.

Table 3.1: The estimation result.

	Coefficient	Std. err.	t value
β_l	-1.545×10^{-3}	1.225×10^{-5}	-126.1
β_θ	-1.843	5.980×10^{-3}	-308.2
The number of samples			$107,280$
$LL(\hat{\beta})$			$-113,711$

3.4.2 Estimation results

The results of estimating the route choice model in Section 3.2 using the data in Section 3.3 are shown in Table 3.1. According to the t values, both parameters are significant. The ratio of the parameters is $\beta_\theta/\beta_l \approx 1193$, which means that one right-left turn at a right angle corresponds to a travel of about 1.2 km.

3.4.3 The degree of coincidence of route choice

As an indicator of the degree to which the obtained model reflects the actual route choice behavior, the degree of coincidence CO_n is introduced given by Eq. (3.4) used in the previous study (Manley et al., 2015).

$$CO_n = \frac{\sum_{a \in M_n} \delta_a^S x_{l(a)}}{\sum_{a \in S_n} x_{l(a)}}, \tag{3.4}$$

where S_n is the set of all links in the route actually chosen on trip n. M_n is the set of all links in the route obtained by the model on trip n. $x_{l(a)}$ is the length of link a, and δ_a^S is a binary variable that is 1 if link a is included in S_n and 0 otherwise. In other words, CO_n is the sum of the lengths of the common links in the route obtained by the model and actually chosen, divided by the total length of the route actually chosen. CO_n is 1 if the route obtained by the model exactly matches the actual route choice.

Table 3.2 shows the distribution of CO_n for optimal routes along the model in all 107,280 trips. The optimal route is the route with the minimum cost weighted by the travel distance from the origin to the destination and the turn angle, respectively.

Table 3.2: The distribution of CO_n for the proposed model.

CO_n	The number of trips	Ratio
0	4,460	4.16%
(0, 1)	17,141	15.98%
1	85,679	79.86%

In addition, the overall average was 89.95%. The results show that most of the trips reproduced perfectly.

3.4.4 Discussion of the estimated results

According to the value of estimated parameters, one right-left turn at a right angle corresponds to a travel of about 1.2 km. In addition, an angular change of 45 degrees corresponds to a travel of 0.35 km.

In the case study done in Torino by Prato et al. (2012), one turn corresponds to a travel of 0.26 km, and in the combined result of the case study of Borlänge by Fosgerau et al. (2013) and Prato et al. (2012), one turn corresponds to a travel of 0.23 km. Although the ratio of the parameters with the angle change in the right angle is larger than in the previous studies, that in 45 degrees angular change, which is treated as a turn in the previous studies, is close to that in the previous studies.

The results with the model that minimizes only the distance excluding the effect of the turn angle are shown in Table 3.3. Comparing the two log-likelihood values for 3.1 and 3.3, the turn angle term improves the value of the log-likelihood function, indicating that the model performance is improved.

Table 3.3: The estimation result of the model without turn angle.

	Coefficient	Std. err.	t value
β_l	-9.42×10^{-3}	4.145×10^{-5}	-227.3
The number of samples			$107,280$
$LL(\hat{\beta})$			$-222,092$

3.5 Comparison with other models

3.5.1 Compared models

Compare the model proposed in this chapter (called Model A) with a route that minimizes only the sum of the distance $x_{l(a)}$ (Model B), the sum of the turn angle variable $x_{\theta(a|k)}$ (Model C), and the product of the travel distance and the number of left-right turns (Model D). In Model D, a left-right turn is considered to have been made when the value of $x_{\theta(a|k)}$ is greater than 0.5, i.e., when $\pm\pi/3$ or more turns are made. About 60% of all cases where the value of $x_{\theta(a|k)}$ is 0.5 or more (Fig. 3.4). Model B is the commonly used model that minimizes the travel distance, Model C is the shortest angular path proposed by Turner (2007), and Model D is the best model proposed by Manley et al. (2015). In all models, the routes were determined by a shortest path search based on the link costs along each definition.

3.5.2 Comparison result

Table 3.4 shows the distribution of CO_n for each model. The average of CO_n of the proposed Model A is the highest. However, the difference with Model B, which minimizes the travel distance, is negligible.

The reason why CO_n is fairly high overall is that this model limits the region to approximately 10 km square, and trips that use highways to go out of the region only choose routes that use highways within the region, so that all routes are the same in all models. In addition, the links and routes that were selected by only a few travelers were

Table 3.4: The distribution and average of CO_n for each model.

	The distribution of CO_n			
Model	0	(0, 1)	1	The average of CO_n
A	4.16%	15.98%	79.86%	0.8995
B	4.10%	15.33%	80.57%	0.8991
C	4.16%	19.49%	76.35%	0.8749
D	4.50%	18.73%	76.77%	0.8668

removed during network generation. Therefore, CO_n may be overestimated because only representative routes chosen by many travelers are used as route choice data. Furthermore, even if a trip started or ended in the region, the data around the origin and destination of the trip was deleted. CO_n is expected to change due to these causes as well.

On the other hand, CO_n is zero for some routes. Such routes deviated significantly from the shortest route and from the route with the smallest angle change, and CO_n was often zero in all models.

3.5.3 Detail comparison

First, a trip where Model A has the highest CO_n and the other three models have the lower CO_n are shown in Fig. 3.5. The red line is the route actually chosen, the yellow line, the green line, and the blue line are the optimal route for Model A, Model B, and Models C and D, respectively (the same is shown in the following figures). Note that the color distinction is not made between Models C and D, because the optimal routes for the ODs presented in this comparison were all identical for Models C and D.

The optimal route for Model A consists of two left-right turns and shortens the travel distance between origin and destination, while the optimal route for Model B consists of four left-right turns and shortens the travel distance further. On the other hand, the optimal routes in Models C and D connect the origin and destination points with only one right turn, but are slightly more detouring. The travel distance for each route is 11,482

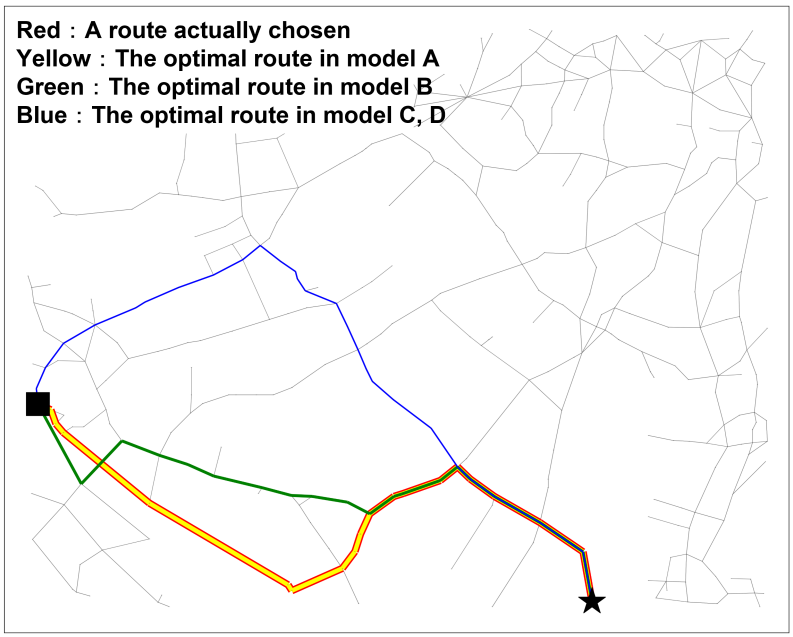


Figure 3.5: Example where only CO_n in Model A equals 1.

m for Model A, 11,046 m for Model B, and 12,434 m for Models C and D, with Models C and D taking a large detour. In addition, the sum of the variables for the turn angles is 3.045 for the route of Model A, 4.039 for Model B, and 2.308 for Models C and D. Model B sacrifices the angular variable to achieve the shortest route, while Models C and D take large detours to reduce the sum of angular variables. Furthermore, for the values of the utility functions substituting the parameters estimated in Section 3.4, the route for Model A is -23.35, Model B is -24.51, and Models C and D are -23.46. By reducing the sum of both the travel distance and the angular variables for the turn angle, the route of Model A is obtained. In fact, only the route of Model A is chosen between this origin and destination.

Next, a trip where Model B has the smallest CO_n and the other three models CO_n equal to 1 is shown in Fig. 3.6. The optimal routes for Models A, C, and D are those that make two left-right turns near the origin and halfway between the origin and destination, while the optimal route for Model B is one that makes four left-right turns and shortens the travel distance between the origin and destination. The travel distance for routes A, C, and D is 13,584 m, while that for Model B is 12,738 m. In addition, the sum of the

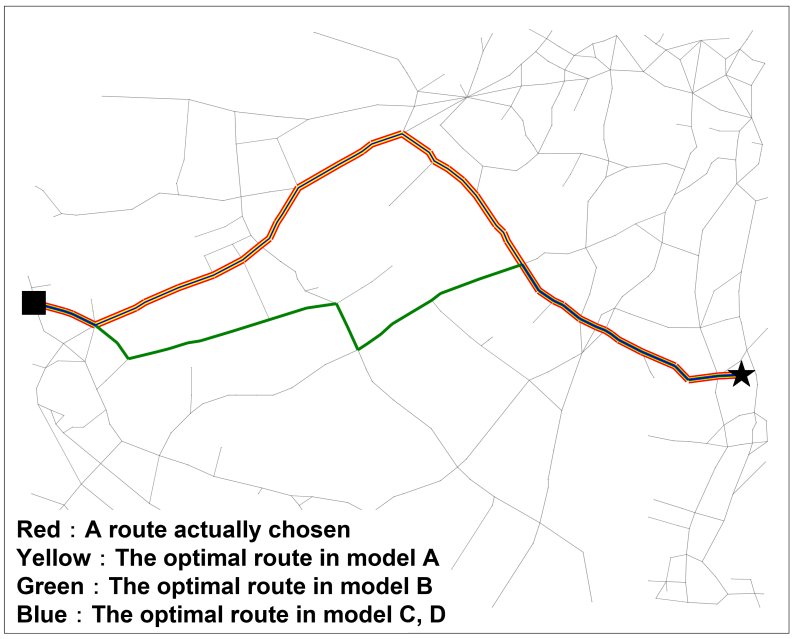


Figure 3.6: Example where only CO_n in Model B is small.

variables for the turn angles is 2.261 for the routes of Models A, C, and D and 4.279 for the route of Model B. The route in Model B is not a good choice in terms of Models A, C, and D because it makes a total of four left-right turns to achieve the shortest route. Only routes A, C, and D are route choices between this origin and destination.

Finally, a trip with CO_n equal to 1 for Model B and lower CO_n for the other models is shown in Fig. 3.7. Model A and Model B are taking different links in the eastern part of the region. In Model A, a slightly detouring route is optimal to make only two left-right turns, one near the origin and the other near the destination. In addition, for Models C and D, the best route is the one that makes much detour to make only one left-right turn near the destination. The travel distance for each route is 9,338 m for Model B, 9,656 m for Model A, and 11,983 m for Models C and D. On the other hand, the sum of the variables for the turn angles is 5.233 for the route of Model B, 4.529 for Model A, and 3.818 for Models C and D. Furthermore, for the values of the utility functions substituting the parameters estimated in Section 3.4, are -24.07 for the route of Model B, -23.26 for the route of Model A, and -25.55 for the routes of Models C and D. The actual route

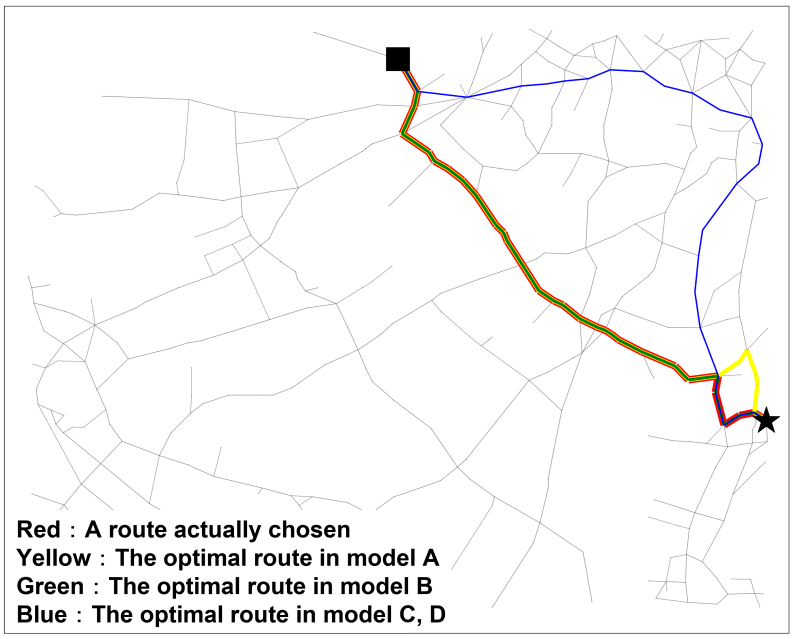


Figure 3.7: Example where only CO_n in Model B is large.

chosen between these origin and destination points was consistent with that of Model B. However, Model A can reproduce a similar route to the actually chosen route, although models with only angular variables or models that minimize the number of turns, the optimal route and the route actually used are very different.

3.5.4 Consideration of comparison

Model A, which considers both turn angle and travel distance, represents actual route choice behavior quite well. However, Model B, which minimizes only the travel distance, performs almost as well. Based on the current analysis, it cannot be concluded that Model A is overall superior to Model B.

Fig. 3.8 shows the plot of the sum of $x_{\theta(a|k)}$ on the horizontal axis and the sum of $x_{l(a)}$ on the vertical axis for the OD pairs with different CO_n in Model A and Model B for the actual route choice. The color of the points is determined by the model with larger CO_n. The number of points is 135. When the sum of the $x_{\theta(a|k)}$ variables for the actual route choice is greater than 3, i.e., when the number of right/left turns near the right angle

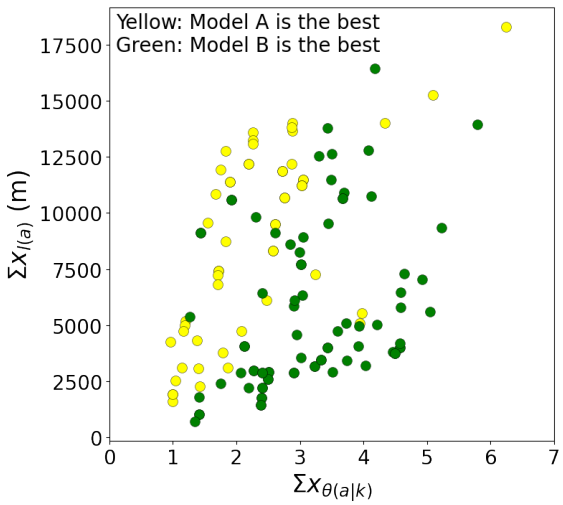


Figure 3.8: The example of optimal route in each model

exceeds two to three, CO_n of Model A among almost all OD pairs is inferior to that of Model B. This suggests that the proposed model may not capture route choice behavior, such as making multiple right and left turns to get to the main road near the origin and destination. In other cases, Model A is more accurate than Model B. Furthermore, by introducing a turn angle variable, unnatural routes that make multiple left-right turns in order to minimize the travel distance are unlikely to be selected. If the formulation of the utility function is modified to take these points into account, it is expected that a more accurate model can be built that incorporates angles and distances more appropriately.

3.6 Conclusions

In this chapter, a variable for turn angle was introduced to route choice behavior on the network to examine how directional changes at intersections within the road network affect driver route choice behavior. The author built a model in which the utility function is a linear combination of a variable for travel distance. The model parameters were estimated from the route choice behavior data obtained from the probe data, and it was confirmed that the turn angle variable strongly influenced route choice behavior. In the case study, it was observed that directional changes significantly affect route choice

behavior, with a single left or right turn equivalent to an additional travel distance of approximately 1.2 km. In addition, the proposed model was compared with a model based only on travel distance, a model based only on turn angle, and a model based only on the number of turns. The proposed model reproduced the actual route choice behavior with more accuracy than the other models.

In this analysis, it is impossible to distinguish whether the behavior of reducing the angle change is caused by the driver or by other factors such as the navigation system. Although Dalton (2003) suggests that the angle change is reduced as an inherent human behavior, the behavior of drivers on the road network has not yet been analyzed. Comparing behavior differences by vehicle type or trip purpose using this analysis method may help to understand drivers' inherent resistance to angle change.

In this model, the estimated parameters were applied to the data of the target area to measure the reproducibility of the route. As a result, the degree of coincidence of route CO_n may have been overestimated. The validity of the model can be evaluated more accurately by estimating parameters in other areas and time periods using the model, and by validating more robust reproducibility such as cross-validation. This is one of the future tasks of this chapter.

Since the angle change is described only by the cosine term in this case, the distinction between a right turn and a left turn is not taken into account. In addition, since only a first-order cosine term is used, the angle changes of going straight and turning left/right, and turning right/left and making a U-turn are equally 1 different on the variables. However, these two angle changes are considered to be very different in nature. To express these differences in the variables, a possible solution is to add a sine or cosine term of first or higher order (Johnson and Wehrly, 1978). This would allow a more detailed analysis of the sensitivity of the angular change and the traveler's choice behavior.

Although only travel distance and turn angle were used as variables, other variables, such as a dummy of arterial roads and the number of lanes, further allow us to examine turn angle's influence on actual route choice behavior. In addition, it is possible to construct a route choice model using only angular indicators, where the turn angle is

considered as a micro angle indicator in route choice and the angle of heading toward the destination as much as possible is considered as a macro angle indicator. This may enable a deeper consideration of the influence of angular indicators on drivers' route choice behavior.

In addition, since this model used a discrete choice model, it is possible to calculate the route choice probability for all users. Using the route choice probabilities, the model can be developed into an assignment problem.

Chapter 4 Temporal Transportation Model with Angle: Daily Variation of Traffic Volume

4.1 Introduction

Traffic volume is one of the most essential variables in the transportation engineering field. In addition, it is easy to observe by installing a traffic counter and counting the number of vehicles passing by at the observation site. Many traffic counters exist along the arterial roads in Japan and other countries and the traffic flow data are accumulated day by day (Leduc et al., 2008; Anacleto et al., 2013). These data are mainly utilized so far in the following two manners. The first is the evaluation of road performance, or throughput, i.e., the maximum number of vehicles the road can accommodate within a unit time (Edie et al., 1963; Transportation Research Board, 2000). This value is represented in units of vehicles per hour, for example. If more vehicles than this value come to the road, traffic congestion occurs. The second is so-called "traffic state estimation", which comprises the three essential variables: the traffic flow rate, the average vehicle density, and the average vehicle speed. Understanding the relationship among these three variables is quite important to predict traffic congestion that is spread along the road network (Muñoz et al., 2003; Wang and Papageorgiou, 2005).

The traffic volume varies depending on the time of day (Transportation Research Board, 2000). In particular, traffic volume is large during the morning commute and evening return trip, and small during the midnight period in a city. Many studies have been conducted to forecast traffic volume with the characteristics of daily variation of

This chapter is mainly based on the research with Prof. Wataru Nakanishi, and Prof. Yasuo Asakura, published in *Journal of JSCE* (Nagasaki et al., 2023), and the research with Prof. Shogo Kato, Prof. Wataru Nakanishi, and Prof. M. C. Jones, in *arXiv* preprint (Nagasaki et al., 2022).

traffic volume (Smith and Demetsky, 1997; Vlahogianni et al., 2014; Lana et al., 2018; Hong, 2011). However, when analyzing daily variation, time should be converted into angles because the time axis has periodicity.

On the other hand, the utilization of probability distribution of the daily variation of traffic volume also contains potential for understanding traffic phenomena. For example, traffic will be classified according to characteristics based on traffic count data, which is easy to obtain. In general, in order to determine the purpose of traffic on a road, not only stationary data such as traffic count data but also other data such as vehicle trajectory, questionnaire surveys, and person trip data are necessary. On the other hand, by fitting a probabilistic model with multiple components, it is possible to classify the traffic by characteristics based on only traffic count data. Additionally, it is useful for the verification and validation of microscopic traffic simulations that deal with each vehicle as an agent. To the best of the authors' knowledge, no study except one written in Korean (Na and Jang, 2011) estimates the probability distribution of traffic volume.

However, the complexity and periodicity of daily variation of traffic makes the estimation of stochastic models quite difficult. Traffic volume is generally determined by the dynamic system of the interaction between the supply side (how many vehicles the road can accommodate) and the demand side (how many drivers wish to use the road). Due to this complexity, the distribution of traffic volume in a day is not expected to be unimodal nor symmetric. Generally, the timestamp of each vehicle's passing does not follow a Poisson process because of some bottlenecks such as traffic congestion and traffic lights (Daganzo, 1997). Actually, it is empirically bimodal and asymmetric in weekdays (explained in detail in Section 4.2). In addition, if we consider an average weekday, time is periodic and should be represented as points on a circle.

Employing a mixture distribution on the circle, where individual components can represent the complex distribution shapes, is essential for estimating a probabilistic model of daily variation of traffic volume. Some mixtures of circular distributions have been proposed in the literature. The most attention has been paid to mixtures of von Mises distributions (e.g., (Wallace and Dowe, 2000; Mooney et al., 2003; Banerjee et al., 2005;

Mulder et al., 2020)). The components of these mixtures, the von Mises distributions, are symmetric distributions with two parameters controlling location and mean resultant length. Recently, Miyata et al. (2020) discussed mixtures of sine-skewed distributions whose components can adopt mildly asymmetric shapes. However, none of these existing models are appropriate for the traffic data because one of the clusters in the data is strongly skewed. Therefore, for the modeling of the data, it seems reasonable to consider the distributions of Kato and Jones (2015) as possible components of the mixtures because these distributions can control a wide range of skewness and kurtosis.

Based on the above, the objective of this chapter is to build a model of daily variation of traffic volume by converting time into angles in order to represent periodicity. Specifically, the daily variation of traffic volume is modeled as a probability distribution on a circle. The mixture of the Kato–Jones distributions, which is a circular distribution capable of representing complex distribution shapes, is employed to apply a probabilistic model that solves the problems of complexity and periodicity of the variation. In addition, the author contributes a method of parameter estimation for mixtures of Kato–Jones distributions as there is no established method thus far.

The chapter is organized as follows. The traffic count data of interest are introduced in Section 4.2. The mixtures of the distributions of Kato and Jones (2015) are defined, which are applied to the traffic data, and investigate basic properties of the mixtures and their components in Section 4.3. Two methods for parameter estimation for the proposed mixtures are presented in Section 4.4. A simulation study is conducted to compare the two proposed methods for parameter estimation in Section 4.5. The proposed mixtures are applied to the traffic data using the proposed inferential methods in Section 4.6. The interpretation of the estimated model is also discussed. The fit of the proposed mixtures to the data is compared with the fits of some other mixture models in Section 4.7. Finally, Section 4.8 concludes the chapter.

4.2 Data

First, the author briefly explain about traffic counters in general. Traffic counters record the timestamps of all vehicles passing. Many counters are installed along the arterial roads in Japan, often at intervals of one kilometer. Usually, these data are aggregated to 1- or 5-minute vehicle counts. Then, the aggregated data are utilized for traffic control such as detecting congestion and providing information to drivers. On the other hand, the data without aggregation is used in this chapter, the data being raw timestamps, to model the probability of a vehicle passing at each particular time. In particular, this chapter focuses on the averaged distribution for weekdays. Then, the time axis is circular; 0:00 (0 a.m.) and 24:00 (12 p.m.) represent the same time, and 1:00 (1 a.m.) and 23:00 (11 p.m.) are not 22 hours apart, but only 2 hours apart. To consider these characteristics accurately, the concept of directional statistics is employed.

Concretely, the data of a traffic counter provided by the Hanshin Expressway Co. Ltd., Japan via personal communication is used. Specifically, the data are obtained at the 20.4 kilopost of the Kobe route, Hanshin Expressway. Fig. 4.1 shows a map of the target area including the location of the traffic counter and the Kobe route of the Hanshin Expressway. The background map is based on a Digital Map by the Geospatial Information Authority of Japan. The Kobe route exists in the Osaka metropolitan area and connects two large cities, Osaka and Kobe. The Osaka metropolitan area is the second largest megacity in Japan following the Tokyo metropolis and its population is around 20 million. Osaka city is the main city of the Osaka metropolitan area with a population of around 2.8 million. Kobe city is the second largest city in the area with a population of 1.5 million and is located about 30 km west of Osaka city.

The counter records the timestamps of vehicles passing in the direction from Kobe to Osaka. The data period is 46 weekdays; from June 6th to July 7th and from August 22nd to October 10th in 2016. It is not unnatural to aggregate data for all weekdays because the distribution of weekday traffic is generally similar (Transportation Research Board, 2000). A histogram of the data of the summation of all 46 days is shown in Fig. 4.2. The

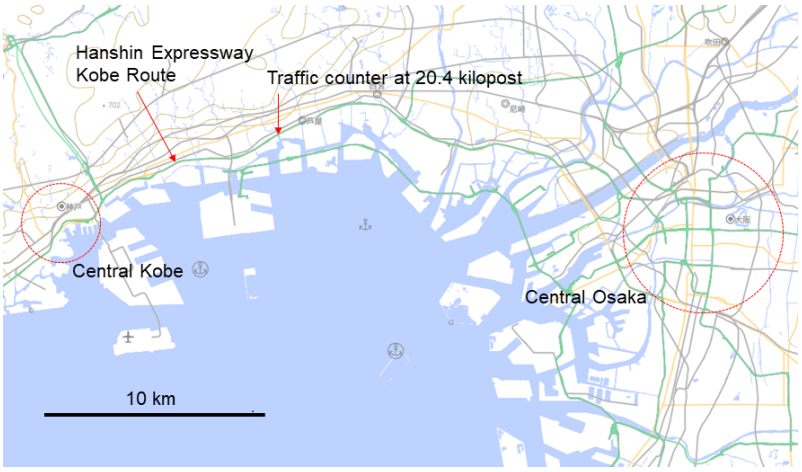


Figure 4.1: Map of target area including the location of the traffic counter and the Kobe route of the Hanshin Expressway. The background map is based on a Digital Map by the Geospatial Information Authority of Japan.

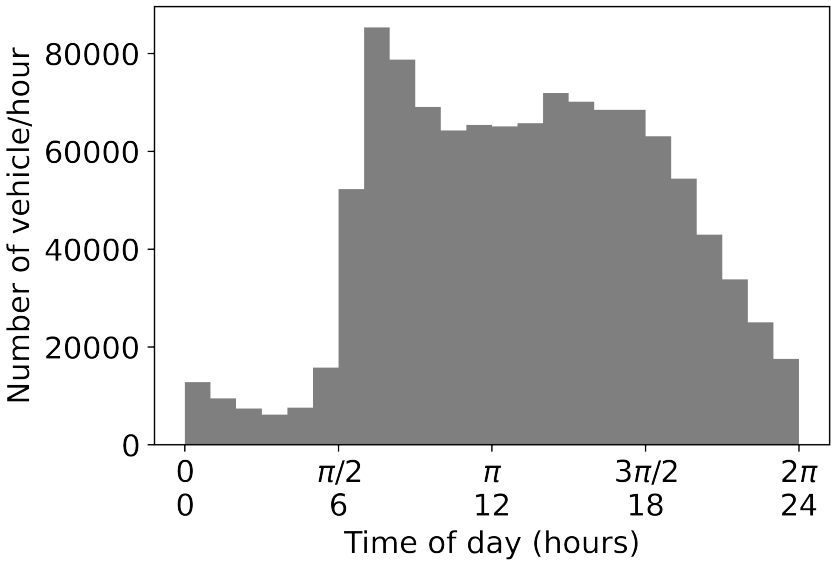


Figure 4.2: A histogram of the data for the summation of all 46 weekdays.

bin width is set as 1 hour for this figure.

The total number of observed vehicles for 46 days is 1,121,262, which equals 24,375 per day on average. Here, the histogram has two peaks at around 7:00 and 15:00. It is assumed that the former peak represents the morning rush hour and the latter does the evening rush hour, from Kobe to Osaka, as the data are for weekdays. Note that

all recorded traffics are in the same direction (i.e., from Kobe to Osaka). It means that vehicles commuting to their office and returning home are observed only once in the morning or evening.

In addition, the peak in the morning is higher and sharper than that in the evening. Also, each peak has an asymmetric shape. For example, as many commuters should arrive at their offices in Osaka at around the same time, typically, at 8:30 or 9:00, the histogram of the morning peak suggests a negatively-skewed shape. A model that cannot represent flexible shapes cannot describe the shape and characteristics of traffic data adequately. To fully represent the characteristics of this distribution, a mixture of probability density functions that allows a flexible shape is necessary. Hereafter, it is assumed that the data follow IID property. Although this might be a strong assumption, it seems to be acceptable from the transportation engineering viewpoint for following reasons. First, if there is no congestion and all vehicles run by their desirable speed, the traffic flow is usually regarded as independent (Transportation Research Board, 2000). Majority of the applied data belong to this situation. Next, even though there is a speed decline, almost all of "congestions" in this data is not serious ones; all vehicles run at least 20 or 30 [km/h] and no complete stop is observed. In addition, they have enough gap between the following and successive vehicles such as 5 to 10 [m]. In the transportation engineering field, this is called "stagnation" and distinguished from congestion. It should be noted that each driver can somehow choose their speed in the stagnation (Brackstone et al., 2009). This directly means that the arrival timing is mainly decided by some randomness caused by each vehicle and driver.

4.3 Model

4.3.1 Kato–Jones distribution

This section builds models that can represent flexible shapes adaptable to the data. As shown in Fig. 4.2, the variation of data has two features: bimodality and different degrees

of skewness and kurtosis in the peaks of the morning and evening rush hours. Therefore, a mixed distribution with the component as the distribution whose skewness and kurtosis can be flexibly adjusted should be introduced. The Kato–Jones distribution (Kato and Jones, 2015) is one of the circular distributions with the desirable properties. Its density function is given by

$$g_{\mathrm{KJ}}(\theta;\mu,\gamma,\lambda,\rho) = \frac{1}{2\pi}\left\{1 + 2\gamma\,\frac{\cos(\theta-\mu)-\rho\cos\lambda}{1+\rho^2-2\rho\cos(\theta-\mu-\lambda)}\right\}, \quad 0 \le \theta < 2\pi, \tag{4.1}$$

where the ranges of the parameters are $0 \le \mu < 2\pi$, $0 \le \gamma < 1$, $0 \le \rho < 1$, and $0 \le \lambda < 2\pi$ satisfying $(\rho\cos\lambda-\gamma)^2+(\rho\sin\lambda)^2 \le (1-\gamma)^2$. In addition, ρ and λ can be converted to the circular skewness and kurtosis of Batschelet (1981) by $\bar{\beta}_2 = \rho\gamma\sin\lambda$ and $\bar{\alpha}_2 = \rho\gamma\cos\lambda$, respectively. The Kato–Jones distribution can be reparametrized to have the following density with the parameters $\mu, \gamma, \bar{\alpha}_2$ and $\bar{\beta}_2$:

$$\begin{aligned} &g^{*}_{\mathrm{KJ}}(\theta;\mu,\gamma,\bar{\alpha}_2,\bar{\beta}_2) \\ &\quad = \frac{1}{2\pi}\left[1 + 2\gamma^2\,\frac{\gamma\cos(\theta-\mu)-\bar{\alpha}_2}{\gamma^2+\bar{\alpha}_2^2+\bar{\beta}_2^2-2\gamma\{\bar{\alpha}_2\cos(\theta-\mu)+\bar{\beta}_2\sin(\theta-\mu)\}}\right]. \end{aligned} \tag{4.2}$$

where $0 \le \mu < 2\pi$, $0 \le \gamma < 1$, and $(\bar{\alpha}_2,\bar{\beta}_2) \ne (\gamma,0)$ satisfy $(\bar{\alpha}_2-\gamma^2)^2+\bar{\beta}_2^2 \le \gamma^2(1-\gamma)^2$. Thanks to this formulation of the density function, all the parameters can be clearly interpreted: the parameters μ, γ, $\bar{\alpha}_2$ and $\bar{\beta}_2$ control the mean direction, mean resultant length, circular kurtosis and circular skewness, respectively. Note that the formulation of Eq. (4.1) is used in parallel because it is convenient for estimating the distribution.

This distribution has three benefits to introduce as the distribution for the traffic data. First, the Kato–Jones distribution is easily interpreted as mentioned above. The mean direction, mean resultant length, circular kurtosis and circular skewness are fundamental statistics in directional statistics, and the correspondence between them and the parameters is very useful for interpretation.

Second, the Kato–Jones distribution is a flexible unimodal distribution that can provide a wide range of skewness and kurtosis. As can be seen in Fig. 2.5, the degrees of skewness and kurtosis of the Kato–Jones distribution can be widely controlled by parameters. Unimodality is also important to interpret each component of the mixture.

Finally, parameter estimation for the Kato–Jones distribution is straightforward by method of moments. This is due to the property that trigonometric moments and probability density function of the Kato–Jones distribution Eq. (4.1) can be expressed in simple and closed form. This property allows for analytical estimation by the method of moments. In addition, since the estimators by the method of moments and the maximum likelihood estimation are consistent estimator, the maximum likelihood estimator can be obtained by numerically maximizing the likelihood function using the parameters estimated by the method of moments as the initial values.

Note also that the use of the Kato–Jones distribution is recommended in recent review papers of Ley et al. (2021) and Pewsey and García-Portugués (2021) as a flexible distribution on the circle.

4.3.2 Mixture of Kato–Jones distribution

The mixture distribution of Kato–Jones distribution of Eq. (4.1) is formulated as

$$f(\theta) = \frac{1}{2\pi} \sum_{k=1}^{m} \pi_k \, g_{\mathrm{KJ}}(\theta; \mu_k, \gamma_k, \lambda_k, \rho_k), \quad 0 \le \theta < 2\pi, \tag{4.3}$$

where $m \in \mathbb{N}$ is the number of the components of the mixture and $0 < \pi_1, \ldots, \pi_m < 1$ are the weights of the components satisfying $\sum_{k=1}^{m} \pi_k = 1$. The parameters of each component, namely, μ_k, γ_k, ρ_k and λ_k, are defined as in Eq. (4.1).

On the other hand, the mixture distribution based on Eq. (4.2) is

$$f(\theta) = \sum_{k=1}^{m} \pi_k \, g^*_{\mathrm{KJ}}(\theta; \mu_k, \gamma_k, \bar{\alpha}_{2k}, \bar{\beta}_{2k}). \tag{4.4}$$

$\bar{\alpha}_{2k}$ and $\bar{\beta}_{2k}$ are defined as in Eq. (4.2). The number of parameters for the mixture is $5m-1$. Note that the mixture of Kato–Jones distribution is called MoKJ hereafter.

4.4 Estimation

Since the traffic volume data has two peaks, fitting is performed using MoKJ, which can represent a multimodality. In the estimation of the distribution, as in the usual Kato–Jones

distribution, the method of moments is employed first, and the obtained parameters are used as initial values to achieve maximum likelihood estimation. However, the formulation of Eq. (4.3) contains redundant parameters, which can cause trouble during optimization calculations. Therefore, the reparameterization is described first, followed by a discussion of the method of moments and maximum likelihood estimation.

4.4.1 Reparameterization

First alternative expressions for the densities of Kato–Jones distributions and their mixture is introduced. The formulation of Eq. (4.1) can be converted by reparameterization as

$$\begin{aligned} &g_{\mathrm{KJ}}(\theta;\mu,\bar{\gamma},\rho,\lambda) \\ &\quad = \pi' \cdot \frac{1}{2\pi}\left\{1 + 2\bar{\gamma}\,\frac{\cos(\theta-\mu)-\rho\cos\lambda}{1+\rho^2-2\rho\cos(\theta-\mu-\lambda)}\right\} + (1-\pi')\,\frac{1}{2\pi}, \end{aligned} \tag{4.5}$$

where $\pi' = \gamma/\bar{\gamma}\,(\in (0,1])$ and $\bar{\gamma} = (1-\rho^2)/\{2(1-\rho\cos\lambda)\}$ is the upper bound of the range of γ for given ρ and λ. The mixture of Eq. (4.5) distributions is described as

$$\begin{aligned} f(\theta) &= \frac{1}{2\pi}\sum_{k=1}^{m}\pi'_k \cdot g_{\mathrm{KJ}}(\theta;\mu_k,\bar{\gamma}_k,\rho_k,\lambda_k) + \frac{\pi'_{m+1}}{2\pi} \\ &= \frac{1}{2\pi}\sum_{k=1}^{m}\pi'_k\left\{1 + 2\bar{\gamma}_k\,\frac{\cos(\theta-\mu_k)-\rho_k\cos\lambda_k}{1+\rho_k^2-2\rho_k\cos(\theta-\mu_k-\lambda_k)}\right\} + \frac{\pi'_{m+1}}{2\pi}, \end{aligned} \tag{4.6}$$

where $\pi'_k = \pi_k\gamma_k/\bar{\gamma}_k$ $(k = 1,\ldots,m)$, $\pi'_{m+1} = 1-\sum_{k=1}^{m}\pi'_k$, and $\bar{\gamma}_k = (1-\rho_k^2)/\{2(1-\rho_k\cos\lambda_k)\}$. This implies the mixture Eq. (4.3) can be divided into a mixture of Kato–Jones submodels $g_{\mathrm{KJ}}(\mu_k,\bar{\gamma}_k,\lambda_k,\rho_k)$ and a uniform distribution. The formulation Eq. (4.6) suggests that π_k and γ_k cannot be uniquely determined in parameter estimation because the information on both parameters is contained in the single parameter π'_k. The number of free parameters in Eq. (4.6) is actually $4m$ since $\bar{\gamma}_k$ is determined by ρ_k and λ_k, which reduces $m-1$ free parameters compared to Eq. (4.3). With this parameterization, it is possible to prove the identifiability of the mixtures.

4.4.2 Modified method of moments

Let random variables $\Theta_1, \ldots, \Theta_n$ be independent and identically distributed from the reparametrized mixture Eq. (4.6). Denote the parameters of the mixture Eq. (4.6) by

$$\boldsymbol{\Psi} = (\mu_1, \ldots, \mu_m, \rho_1, \ldots, \rho_m, \lambda_1, \ldots, \lambda_m, \pi'_1, \ldots, \pi'_{m+1}).$$

In this case, the pth trigonometric moment of Θ_j is given by

$$E(e^{ip\Theta_j}) = \sum_{k=1}^{m} \pi'_k \bar{\gamma}_k \left(\rho_k e^{i\lambda_k}\right)^{-1} \left\{\rho_k e^{i(\mu_k+\lambda_k)}\right\}^p, \quad p \in \mathbb{N}, \tag{4.7}$$

where i is the imaginary unit (Kato and Jones, 2015).

In the usual method of moments estimation, the parameters satisfying Eq. (4.8) are regarded as estimators.

$$\frac{1}{n}\sum_{j=1}^{n} e^{ip\Theta_j} = \sum_{k=1}^{m} \pi'_k \bar{\gamma}_k \left(\rho_k e^{i\lambda_k}\right)^{-1} \left\{\rho_k e^{i(\mu_k+\lambda_k)}\right\}^p, \tag{4.8}$$

for some selected values of p. Eq. (4.8) means equating the theoretical trigonometric moments Eq. (4.7) to empirical ones. However, estimated ρ_k and λ_k by Eq. (4.8) are not guaranteed to always be in the domain.

Therefore, consider the following function that takes the difference between theoretical trigonometric moments and empirical ones.

$$\begin{aligned}
\mathrm{ETM}(\boldsymbol{\Psi}) &\equiv \sum_{p=1}^{q} w(p) \left| \frac{1}{n}\sum_{j=1}^{n} e^{ip\Theta_j} - E(e^{ip\Theta}) \right|^2 \\
&= \sum_{p=1}^{q} w(p) \left| \frac{1}{n}\sum_{j=1}^{n} e^{ip\Theta_j} - \sum_{k=1}^{m} \pi'_k \bar{\gamma}_k \left(\rho_k e^{i\lambda_k}\right)^{-1} \left\{\rho_k e^{i(\mu_k+\lambda_k)}\right\}^p \right|^2,
\end{aligned}$$

where $q \in \mathbb{N}$ and $w(p)$ is a function used to weight the lower-order moments to stabilize the estimation. This function can also be expressed as

$$\begin{aligned}
\mathrm{ETM}(\boldsymbol{\Psi}) = \sum_{p=1}^{q} w(p) \Bigg[&\left\{ \frac{1}{n}\sum_{j=1}^{n} \cos p\Theta_j - \sum_{k=1}^{m} \pi'_k \bar{\gamma}_k \cdot \rho_k^{p-1} \cos\left(p\mu_k + (p-1)\lambda_k\right) \right\}^2 \\
&+ \left\{ \frac{1}{n}\sum_{j=1}^{n} \sin p\Theta_j - \sum_{k=1}^{m} \pi'_k \bar{\gamma}_k \cdot \rho_k^{p-1} \sin\left(p\mu_k + (p-1)\lambda_k\right) \right\}^2 \Bigg].
\end{aligned}$$

When ETM equals 0 within the domain, the estimator coincides with the moment estimator. On the other hand, the estimator that minimizes ETM within the domain is the parameter that is closest to the moment estimator within the domain, although it does not match the moment estimator.

Therefore, a modified method of moments estimator $\hat{\boldsymbol{\Psi}}$ is defined as the minimizer of ETM as

$$\hat{\boldsymbol{\Psi}} = \underset{\boldsymbol{\Psi}\in\Omega}{\operatorname{argmin}} \operatorname{ETM}(\boldsymbol{\Psi}), \tag{4.9}$$

where Ω is the parameter space of $\boldsymbol{\Psi}$.

The result of the estimation depends on the choice of q and $w(p)$. Since the mixture Eq. (4.6) has essentially $4m$ free parameters, there are potentially multiple solutions to the equation $\operatorname{ETM}(\hat{\boldsymbol{\Psi}}) = 0$ for $q < 2m$. To avoid this potential problem, it is recommended to assume $q \geq 2m$ and, in this analysis, $q = 2m$ is assumed. As for the choice of the weight function $w(p)$, $w(p) = c^p (0 < c < 1)$ was adopted in this data analysis to place much more importance on the low-order trigonometric moments than on the high-order ones.

The modified method of moments aggregates the data into multiple moment values, which makes the optimization computation faster. However, since ETM is not a convex function, it is necessary to minimize from many initial values to obtain the parameter.

4.4.3 Maximum likelihood estimation

Next, the maximum likelihood estimation is discussed. The log-likelihood function for the sample $(\theta_1, \ldots, \theta_n)$ is given by

$$\ell(\boldsymbol{\Psi}) = C + \sum_{j=1}^{n} \log \left[\sum_{k=1}^{m} \pi'_k \left\{ 1 + 2\bar{\gamma}_k \frac{\cos(\theta_j - \mu_k) - \rho_k \cos\lambda_k}{1 + \rho_k^2 - 2\rho_k \cos(\theta - \mu_k - \lambda_k)} \right\} + \pi'_{m+1} \right], \tag{4.10}$$

where $C = -n\log(2\pi)$. As is the case for $p = 1$, there is no closed form expression for the maximum likelihood estimator for MoKJ. Therefore, the EM algorithm (McLachlan and Krishnan, 2007) is employed as a numerical approach to obtain the maximum likelihood estimator. The following algorithm is established:

1. Take an initial value $\boldsymbol{\Psi}^{(0)}$, where

$$\boldsymbol{\Psi}^{(0)} = (\mu_1^{(0)}, \ldots, \mu_m^{(0)}, \rho_1^{(0)}, \ldots, \rho_m^{(0)}, \lambda_1^{(0)}, \ldots, \lambda_m^{(0)}, \pi'^{(0)}_1, \ldots, \pi'^{(0)}_{m+1}).$$

2. For $r = 1, \ldots, N$, compute the following until the value of $\boldsymbol{\Psi}^{(N)}$ is virtually unchanged from $\boldsymbol{\Psi}^{(N-1)}$:

$$\pi'^{(r)}_k = \frac{1}{n}\sum_{j=1}^{n} w_{kj}^{(r-1)}, \quad \pi'^{(r)}_{m+1} = 1 - \sum_{k=1}^{m} \pi'^{(r)}_k,$$

$$\left(\mu_k^{(r)}, \lambda_k^{(r)}, \rho_k^{(r)}\right) = \underset{(\mu_k, \lambda_k, \rho_k)}{\operatorname{argmax}} \left[\sum_{j=1}^{n} w_{kj}^{(r-1)} \log\left\{g_{\mathrm{KJ}}(\theta_j; \mu_k, \bar{\gamma}_k, \lambda_k, \rho_k)\right\}\right], \tag{4.11}$$

where $k = 1, \ldots, m$ and

$$w_{kj}^{(r-1)} = \frac{\pi'^{(r-1)}_k g_{\mathrm{KJ}}(\theta_j; \mu_k^{(r-1)}, \bar{\gamma}_k^{(r-1)}, \lambda_k^{(r-1)}, \rho_k^{(r-1)})}{\sum_{h=1}^{m} \pi'^{(r-1)}_h g_{\mathrm{KJ}}(\theta_j; \mu_h^{(r-1)}, \bar{\gamma}_h^{(r-1)}, \lambda_h^{(r-1)}, \rho_h^{(r-1)}) + \pi'^{(r-1)}_{m+1}/(2\pi)}.$$

3. Record $\hat{\boldsymbol{\Psi}}^{(N)}$ as the maximum likelihood estimate of $\boldsymbol{\Psi}$.

The calculation of the likelihood function requires a long computation time when the data size is large. Therefore, an initial value that is close to the maximum likelihood estimation is desirable. The estimator by the modified method of moments is close to the maximum likelihood estimator due to its consistency. Therefore, it is suitable for the initial value of maximum likelihood estimation.

4.5 Simulation study

The performance of the modified method of moments estimation and the maximum likelihood estimation by EM algorithm is compared via a Monte Carlo simulation study. Random samples of sizes $n = 50, 100, 500, 1000$, and 5000 were generated from the estimated distribution whose parameters are shown in Table 4.4(b). For each sample size, $r = 2000$ samples were generated.

For comparison of the two estimators, the generalized mean squared error is employed that is defined by $\det(\Sigma)$, where $\Sigma = E\{(\hat{\xi} - \xi)(\hat{\xi} - \xi)^\top\}$ and $\hat{\xi}$ is an estimator of $\xi =$

$(\mu_1, \mu_2, \rho_1, \rho_2, \lambda_1, \lambda_2, \pi'_1, \pi'_2)^\top$. An estimate of the generalized mean squared error is given by replacing Σ by its sample analogue $\hat{\Sigma} = r^{-1} \sum_{j=1}^{r} (\hat{\xi}_j - \xi)(\hat{\xi}_j - \xi)^\top$, where the $\hat{\xi}_j$'s, $j = 1, ..., r$, are the estimates from the r simulation samples. It is also considered the estimated relative generalized mean squared error of the estimate via the modified method of moments with respect to the maximum likelihood estimate, which is defined as

$$\widehat{\mathrm{RGMSE}} = \frac{\det(\hat{\Sigma}_{MM})}{\det(\hat{\Sigma}_{ML})}.$$

Here, $\hat{\Sigma}_{MM}$ and $\hat{\Sigma}_{ML}$ are sample estimates of the mean squared error matrices of modified method of moments estimation and maximum likelihood estimation, respectively.

For numerical optimization of ETM, the scipy.optimize.minimize package in Python with SLSQP method is employed Jones et al. (2001); Kraft (1988). This method is a kind of quasi-Newton method and enables us to solve the constrained optimization problem. In this analysis, the optimization of ETM is terminated when the difference in the value of ETM between two successive steps is lower than 1×10^{-10}. The initial values of μ_k, ρ_k and λ_k are random samples from the uniform distribution on the intervals $[0, 2\pi)$, $[0, 1)$ and $[0, 2\pi)$, respectively. Also, those of π'_1, π'_2 and π'_3 are in $[0, 1)$ and $\pi'_1 + \pi'_2 + \pi'_3 = 1$. For this sampling, two random values are drawn from the uniform distribution on $[0, 1)$ and the smaller one is regarded as r_1 and the other as r_2. Then, $\pi'_1 = r_1$, $\pi'_2 = r_2 - r_1$, and $\pi'_3 = 1 - r_2$ are calculated. Furthermore, the moment weight parameter c in $w(p)$ is set to 0.9 because the experiments suggest that estimates with small c are not stable with regard to fitting the data.

100 estimates are obtained via this process and the estimate whose value of ETM is smallest is set as the initial value for the maximum likelihood estimation. For numerical optimization in the M-step, namely, Step 2 in the EM algorithm the same Python package as in the optimization of ETM is employed. In this analysis, each M-step is terminated when the difference of the value of the objective function between two successive iterations is lower than 1×10^{-6}. Also, the whole EM algorithm is terminated when the difference of the log-likelihood function between two successive M-steps is lower than $n \times 10^{-6}$. In both estimates, it is assumed that $\hat{\mu}_1 < \hat{\mu}_2$.

Table 4.1: The values of $\det(\hat{\Sigma}_{MM})$, $\det(\hat{\Sigma}_{ML})$ and $\widehat{\mathrm{RGMSE}}$.

n	$n = 50$	$n = 100$	$n = 500$	$n = 1000$	$n = 5000$
$\det(\hat{\Sigma}_{MM})$	5.35×10^{-9}	7.00×10^{-10}	2.70×10^{-13}	4.25×10^{-15}	4.01×10^{-20}
$\det(\hat{\Sigma}_{ML})$	5.05×10^{-8}	6.41×10^{-9}	3.56×10^{-12}	1.33×10^{-14}	1.23×10^{-22}
$\widehat{\mathrm{RGMSE}}$	0.106	0.109	0.076	0.322	325.3

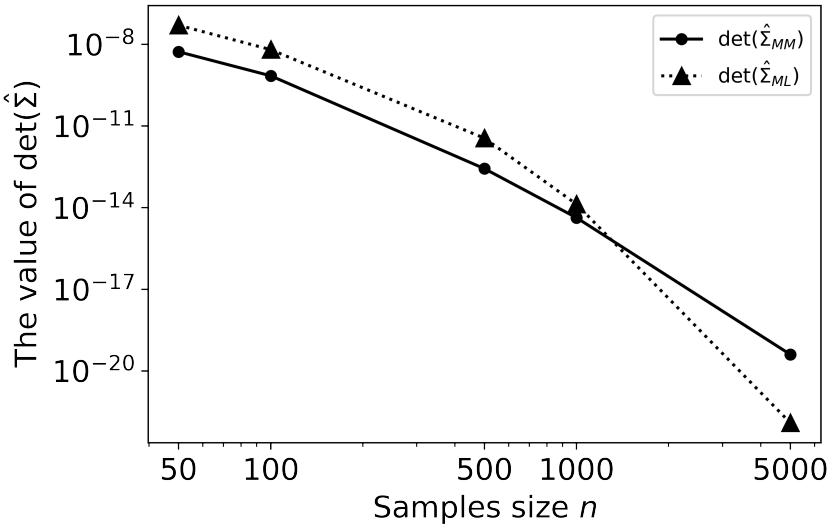


Figure 4.3: Plot of the value of $\det(\hat{\Sigma}_{MM})$ (bold) and $\det(\hat{\Sigma}_{ML})$ (dotted) in simulation study.

Table 4.1 shows the value of $\det(\hat{\Sigma}_{MM})$, $\det(\hat{\Sigma}_{ML})$ and $\widehat{\mathrm{RGMSE}}$ for each sample size. In addition, Fig. 4.3 shows the trend of the change of the value of $\det(\hat{\Sigma}_{MM})$ and $\det(\hat{\Sigma}_{ML})$ in different sample sizes. As the sample size n increases, the value of $\det(\hat{\Sigma}_{MM})$ and $\det(\hat{\Sigma}_{ML})$ decreases. The value of the $\det(\hat{\Sigma}_{MM})$ at $n = 50$ is smaller than $\det(\hat{\Sigma}_{ML})$, however, the value of the $\det(\hat{\Sigma}_{MM})$ at $n = 5000$ is larger than $\det(\hat{\Sigma}_{ML})$ because the degree of decline in value with respect to sample size is greater for $\det(\hat{\Sigma}_{ML})$ than for $\det(\hat{\Sigma}_{MM})$. Therefore, the values of $\widehat{\mathrm{RGMSE}}$ are smaller than 1 for $n \leq 1000$ and larger for $n = 5000$. These results imply modified moments estimation is preferable for estimation from a small number of samples in terms of $\widehat{\mathrm{RGMSE}}$. On the other hand, maximum likelihood estimation is preferable for estimation from a large number of samples such as the data in this chapter. Note that modified moments estimation yields not only an estimate but also a good initial value for maximum likelihood estimation. Although the value of $\widehat{\mathrm{RGMSE}}$ is relatively

large for $n = 5000$, the very small value of $\det(\hat{\Sigma}_{MM})$ for $n = 5000$ indicates that the estimate via the modified moments estimation still has reasonable performance as the initial value for maximum likelihood estimation. Therefore, modified moments estimation is useful to obtain the maximum likelihood estimates efficiently for a large sample size as well as small and medium sample sizes.

It is noted that the degree of decline of the $\det(\hat{\Sigma})$ with respect to sample size depends on the estimation method. Therefore, for $n = 500$, the value of $\widehat{\mathrm{RGMSE}}$ is slightly smaller than for $n = 100$. The tendency for $\widehat{\mathrm{RGMSE}}$ to not increase monotonically is also observed in the unimodal Kato–Jones distribution (Kato and Jones, 2015, Supplementary Material).

4.6 Application

4.6.1 Application of modified method of moments estimation

The proposed inferential methods are applied to the traffic count data of interest. The number of samples is 1,121,262, as mentioned in Section 4.2. In this section, the number of components is fixed to two because two main peaks are recognized in the histogram (see Section 4.7 for results for other numbers of components). The conditions of the modified method of moments estimation are the same as in the previous section.

Table 4.2: Parameter estimates associated with the modified method of moments estimation.

(a) Initial value

k	μ_k	ρ_k	λ_k	π'_k
1	2.2755	0.5322	4.9979	0.0316
2	3.9458	0.8131	2.0455	0.2331
3				0.7353

(b) Estimated parameter

k	$\hat{\mu}_k$	$\hat{\rho}_k$	$\hat{\lambda}_k$	$\hat{\pi}'_k$
1	2.7514	0.7322	5.3162	0.4543
2	4.0106	0.1947	1.1589	0.4820
3				0.0637

The optimization is carried out for 10,000 initial values. Fig. 4.4 and Table 4.2 show the best result in the sense that the value of ETM is the smallest among 10,000 estimations.

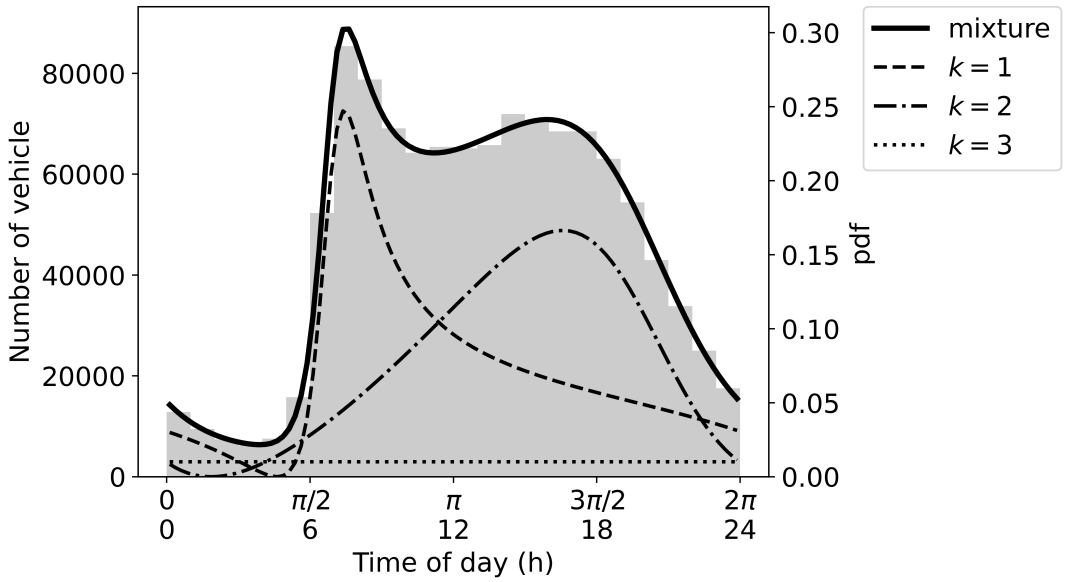


Figure 4.4: Plot of the densities of the reparametrized mixture Eq. (4.6) (bold) and its components (dashed) estimated via the modified method of moments estimation.

Table 4.3: Values of the empirical and theoretical trigonometric moments (t.m.'s) and squares of their differences in the best result.

		empirical t.m.	theoretical t.m.	square of difference
cosine	$p=1$	-3.29×10^{-1}	-3.29×10^{-1}	2.95×10^{-14}
	$p=2$	-7.07×10^{-2}	-7.07×10^{-2}	5.64×10^{-15}
	$p=3$	9.46×10^{-2}	9.46×10^{-2}	3.07×10^{-15}
	$p=4$	-1.61×10^{-2}	-1.61×10^{-2}	8.35×10^{-15}
sine	$p=1$	-1.23×10^{-1}	-1.23×10^{-1}	2.81×10^{-14}
	$p=2$	-1.18×10^{-1}	-1.18×10^{-1}	1.21×10^{-13}
	$p=3$	1.29×10^{-2}	1.29×10^{-2}	4.21×10^{-16}
	$p=4$	6.97×10^{-2}	6.97×10^{-2}	7.43×10^{-16}

Table 4.3 provides the values of the empirical and theoretical trigonometric moments. The solid line in Fig. 4.4 represents the estimated density of the mixture and the dashed lines refer to the estimated densities of the components. The same convention will be used in the following figures of fitted densities. The value of ETM is 7.06×10^{-13} and that of the log-likelihood function is $-1{,}840{,}939$. As shown in the right column of Table 4.3,

the difference between the values of empirical and theoretical trigonometric moments of all the degrees are approximately 0, which means that this result is very close to the result of ordinary method of moments estimation. Nonetheless, it is generally needed the proposed modified method of moments to ensure that the parameters are within their ranges, as explained in Section 4.4. Actually, when the number of components m is set to be greater than two, much larger values of ETM are obtained for some cases and the ordinary method of moments estimates are out of range for some parameters. Also, the estimated densities in Fig. 4.4 fit fairly well to the histogram. The sharp and negatively skewed peak around $\pi/2$ is represented by the component with $k = 1$, while the gentle peak around $3\pi/2$ is modelled by the component with $k = 2$.

The computation time for 10,000 trials is about 14 sec. The calculation was run on a Windows 10 computer with an Intel Core i9-12900 processor running at 3.20 GHz using 64.0 GB of RAM. Furthermore, about 25% of 10,000 trials converge to almost the same value as the best result. In this case, 10,000 initial values were used to verify the performance of the modified method of moments estimation. The estimates that provide the smallest value of ETM can be obtained from 25% of the initial values, which implies that estimation from a much smaller number of initial values (e.g., 100) would be sufficient.

Note that the modified method of moments estimation is carried out to obtain a good initial value for maximum likelihood estimation that is more appropriate for such a large sample size as discussed in Section 4.5.

4.6.2 Maximum likelihood estimation

The EM algorithm is carried out by setting the initial values as the best estimates from the modified method of moments to achieve the maximum likelihood estimation efficiently. This choice of the initial values seems reasonable because of the large sample size of the data and the consistency of the modified method of moments estimator and maximum likelihood estimator. The conditions of the maximum likelihood estimation are the same

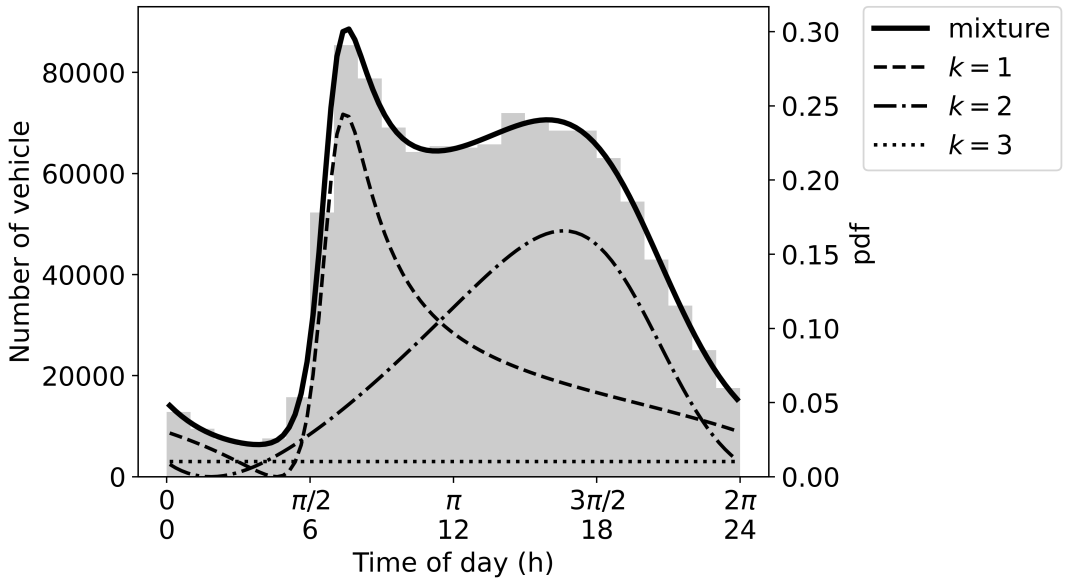


Figure 4.5: Plot of the densities of the reparametrized mixture Eq. (4.6) (bold) and its components (dashed) estimated via the EM algorithm.

as in the previous section except for the threshold for the termination of EM algorithm, which is set to 1. This threshold is set by the following reasoning. Since the number of samples in this chapter is more than 1,000,000, the absolute value of the log-likelihood function is expected to be of that order of magnitude. Thus, the threshold 1 is regarded as small enough to claim convergence.

Table 4.4: Parameter estimates associated with the maximum likelihood estimation for the reparametrized mixture Eq. (4.6) via the EM algorithm.

k	$\hat{\mu}_k$	$\hat{\rho}_k$	$\hat{\lambda}_k$	$\hat{\pi}'_k$
1	2.7572	0.7266	5.3136	0.4536
2	4.0107	0.1970	1.1895	0.4825
3				0.0639

The result of the EM algorithm is shown in Fig. 4.5 and Table 4.4. It takes just two steps to terminate the EM algorithm. The value of ETM increases along with that of the log-likelihood during the EM algorithm. The difference between the parameter estimates by the modified method of moments and those by the maximum likelihood estimation

is not much. This implies that the modified method of moments estimate is close to the maximum likelihood estimate, the estimate produced by another consistent estimator.

In addition, the standard errors for the parameter estimates by the modified method of moments and those by the maximum likelihood estimation are calculated. The standard errors for the parameter estimates by the modified method of moments is calculated by the bootstrap method (Davison and Hinkley, 1997) and those by the maximum likelihood estimation is calculated by the Fisher information matrix for the log-likelihood function due to computation time.

Table 4.5: The value of standard errors for the parameter estimates by the modified method of moments and those by the maximum likelihood estimation.

	k	$\hat{\mu}_k$	$\hat{\rho}_k$	$\hat{\lambda}_k$	$\hat{\pi}'_k$
the modified method of moments	1	5.77×10^{-2}	2.46×10^{-2}	1.87×10^{-1}	8.91×10^{-3}
	2	5.74×10^{-2}	2.47×10^{-2}	1.90×10^{-1}	5.92×10^{-3}
	3				6.59×10^{-2}
the maximum likelihood estimation	1	6.36×10^{-3}	3.50×10^{-3}	2.88×10^{-2}	3.27×10^{-3}
	2	6.82×10^{-3}	1.98×10^{-3}	9.03×10^{-3}	3.76×10^{-3}
	3				3.30×10^{-3}

The results are shown in Table 4.5. For all parameters, the values of standard error for the parameter estimates by the maximum likelihood estimation are smaller than those by the modified method of moments. The values of standard error for the parameter estimates by the maximum likelihood estimation are in between the order of the second and third power of 10, and those by the modified method of moments are in between the order of the one and third power of 10.

As for the computation time, it takes 93 sec. for the EM algorithm although only two M-steps are conducted in the calculation. The calculation environment is the same as that of the previous subsection. Compared to the modified method of moments that takes only about 14 sec. for 10,000 trials, the EM algorithm requires a much longer time. Therefore, setting the initial value close to the maximum likelihood estimator is necessary to achieve the maximum likelihood estimation by the EM algorithm in a feasible computation time.

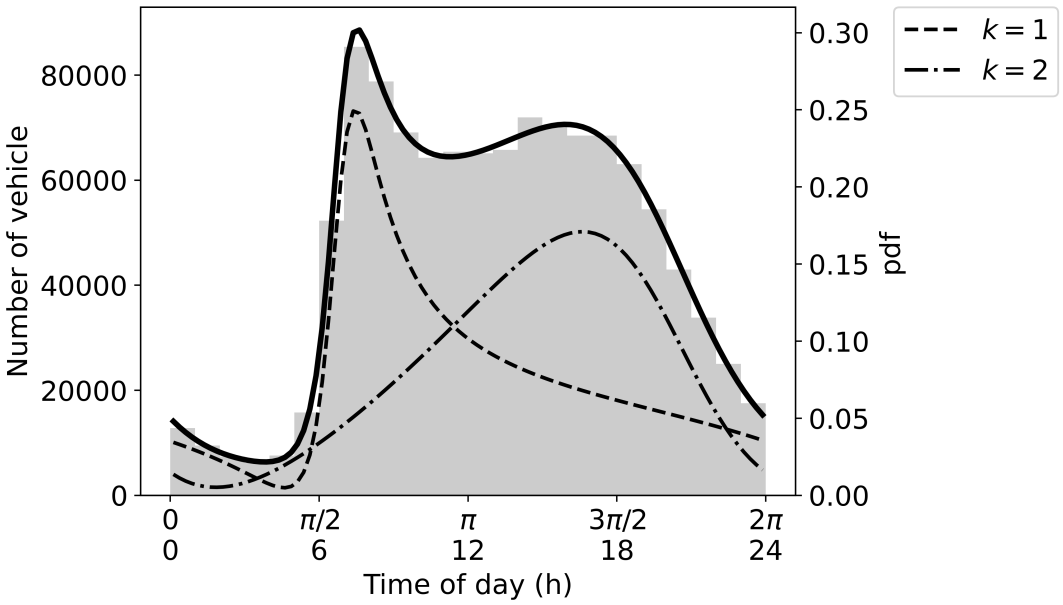


Figure 4.6: Plot of the estimated densities of the recovered mixture Eq. (4.3) (bold) and their components (dashed).

The proposed modified method of moments is one possible solution for this purpose.

Although two parameter estimation methods are introduced in this chapter, the maximum likelihood estimation is preferred in terms of mean squared error of the estimators and usage for model comparison. The modified method of moments is useful to obtain a good initial value for the maximum likelihood estimation with a short computation time for large sample sizes.

Table 4.6: Maximum likelihood estimates for the mixture with the recovered original parameterization Eq. (4.3).

k	$\hat{\mu}_k$	$\hat{\gamma}_k$	$\hat{\rho}_k$	$\hat{\lambda}_k$	$\hat{\pi}_k$
1	2.7572	0.3751	0.7267	5.3136	0.4845
2	4.0107	0.4855	0.1970	1.1895	0.5155

Here, the parameters are estimated in the form of the mixture of two Kato–Jones distributions and the uniform distribution (Table 4.4(b) and Eq. (4.6). Actually, this mixture of three components can be regarded as the estimated model. Nonetheless, it

is also able to reparameterize and recover the estimates of the original parameters Eq. (4.3) as discussed in Section 4.4. As it is initially aimed to estimate the mixture of two components, the latter is employed in the following. The maximum likelihood estimates obtained via the EM algorithm (Fig. 4.5 and Table 4.4) are recovered to $\hat{\mu}_k, \hat{\gamma}_k, \hat{\rho}_k, \hat{\lambda}_k$ and $\hat{\pi}_k$ as shown in Fig. 4.6 and Table 4.6. The shapes of the two estimated components with $k = 1, 2$ in Table 4.4 slightly change by adding the restored uniform component with $k = 3$.

In addition, $\hat{\mu}_k, \hat{\gamma}_k, \hat{\rho}_k, \hat{\lambda}_k$ are reparameterized into $\hat{\mu}_k, \hat{\gamma}_k, \hat{\bar{\alpha}}_{2k}, \hat{\bar{\beta}}_{2k}$ for interpretation because $\hat{\bar{\alpha}}_{2k}$ and $\hat{\bar{\beta}}_{2k}$ can be interpreted as the circular skewness and kurtosis of Batschelet (1981) as mentioned in Section 4.3. The recovered parameters are reparameterized as shown in Table 4.7. The discussion of parameter is carried out only for the parameter in Table 4.7 because the ease of interpretation for $\hat{\bar{\alpha}}_{2k}$ and $\hat{\bar{\beta}}_{2k}$.

Table 4.7: Maximum likelihood estimates for the reparametrized mixture Eq. (4.4).

k	$\hat{\mu}_k$	$\hat{\gamma}_k$	$\hat{\bar{\alpha}}_{2k}$	$\hat{\bar{\beta}}_{2k}$	$\hat{\pi}_k$
1	2.7572	0.3751	0.1542	-0.2248	0.4845
2	4.0107	0.4855	0.0356	0.0888	0.5155

4.6.3 Discussion

Here, the interpretation of the estimated models with the recovered parameterizations Eq. (4.4) given in Table 4.7 is discussed. As shown in Fig. 4.6, the densities of the mixture fit fairly well to the histogram for the entire day, such as the sharp peak in the morning and the gentle peak in the evening. The numerically calculated morning mode and evening mode of the whole distribution of the mixture are 7:32 and 15:56, respectively.

The two components divide the mixture into 48.5% and 51.5% as denoted by each value of π_k. Each peak roughly represents the morning and evening traffic. The peak of the component with $k = 1$ is sharp and negatively skewed. Its mean, $\hat{\mu}_1$, is around 10:32, and its mode is around 7:28 (Kato and Jones, 2015, Supplementary Material). The mode of the component with $k = 1$ is close to the morning mode of the mixture because the

component with $k = 1$ is dominant around the morning. The shape is characterized by large $\hat{\tilde{\alpha}}_{21}$ and negatively large $\hat{\tilde{\beta}}_{21}$. On the other hand, the peak of the component with $k = 2$ is gentle and positively skewed. Its mean, $\hat{\mu}_2$, is around 15:19, and its mode is around 16:37. The mode of the component with $k = 2$ and the evening mode of the mixture is apart because the tail of the component with $k = 1$ is contained even in the evening part of the mixture. The shape is characterized by small $\hat{\tilde{\alpha}}_{22}$ and positively small $\hat{\tilde{\beta}}_{22}$.

The interpretation of each component around its peak can be done straightforwardly. The peak period of the component with $k = 1$ is from 7:00 to 8:00, and its shape is remarkably sharp and quick-rising, which is described by large $\hat{\tilde{\alpha}}_{21}$ and negatively large $\hat{\tilde{\beta}}_{21}$. This is explained by the actual situation that many drivers would like to arrive in central Osaka around the same time, like 8:30 and 9:00, which is the typical start time for working in Japan. On the other hand, the peak period of the component with $k = 2$ is from 15:00 to 19:00, and its shape is gradual and positively skewed, which is described by small $\hat{\tilde{\alpha}}_{22}$ and positively small $\hat{\tilde{\beta}}_{22}$. This may be because drivers do not have to leave their office at the same time, which is different from morning hours. Perhaps, some of them decide to make their departure time earlier or later than peak hour to avoid traffic congestion. Actually, the difference of span of peak periods between commuting to and returning from work has been confirmed (Alexander et al., 2015). In addition, the positively skewed shape may imply that they also avoid returning home too late.

These interpretations are made by focusing on the shape around the mode of each peak. It is more difficult to understand the traffic flow volume around the tails of both components. For example, the first component still has a moderate probability at around 16:00, which cannot be included in the morning rush hour. Moreover, the interpretation of the complex crossing points of each component seen early in the morning is impossible. One possible explanation is that the first component represents the outward trip that includes morning commuting, commercial vehicle, and so on. For example, many commercial vehicles travel from Kobe to Osaka during the daytime. In addition, some people make their first trip in the evening because they work at night or go shopping.

People who work at night are more likely to use their vehicles for commuting than those who work in the daytime as public transport is not operating after midnight. These could explain the first component having a larger probability than the second component after around 22:00. Similarly, the second component represents the day's return trips, including evening commuting. As round and migration trips during the daytime, for example, may be included in the second component, it already has a moderate value even in the morning hours. Note that outward and return trips are observed only once because all recorded traffics are in the same direction.

Finally, the results are compared with some phenomena that are typically mentioned in the traffic engineering field. First, the difference in the value of $\hat{\hat{\alpha}}_2$ between the two components is consistent with the concept of departure time choice in traffic engineering (Noland and Small, 1995). In general, drivers choose their departure time by estimating backward from the time they want to arrive. Alexander et al. (2015) indicates that many drivers depart within short periods in the morning (i.e., traffic is concentrated), while drivers depart within longer periods in the evening and later. Second, 30% of Hanshin Expressway users are for commercial purposes (Hanshin Expressway, 2017, (in Japanese)). If it is assumed that people make outward commercial trips after they start working, the tail of the first components around noon is likely to contain commercial vehicles at a certain level. Therefore, the aforementioned interpretation that the first component is for the outward trips of the day and the second one is for the return trips is reasonable to some extent. Note that the share of heavy vehicles on this road is 13.6% in total and less than 30% even in the midnight period, although one might expect that the freight traffic are the main component during midnight. Therefore, no peak for freight traffic in the midnight period are found in the histograms, nor are such peaks found in the estimated model.

Since these discussions are only derived from traffic volume data, it is necessary to further develop the interpretations by combining data such as trip purpose data and origin and destination data. In addition, combining with other data, such as the spatio-temporal distribution of demand discussed in Chapter 5 provides substantiation for

the interpretation of each component. Nonetheless, it would be the contribution of this chapter that the proposed model that only uses traffic volume information is consistent with various concepts used in traffic engineering.

4.7 Comparison with other models

4.7.1 Mixtures of other distributions

Other multimodal distributions are compared with the proposed model in terms of practicability and fit. The distributions considered here are the mixture of von Mises distributions (MovM), the mixture of wrapped Cauchy distributions (MowC), the mixture of sine-skewed von Mises distributions (MossvM) and the mixture of sine-skewed wrapped Cauchy distributions (MosswC) whose densities are given by

$$\begin{aligned}
f_{\mathrm{MovM}}(\theta) &= \sum_{k=1}^{m} \frac{\pi_k}{2\pi I_0(\kappa_k)} \exp(\kappa_k \cos(\theta - \mu_k)), \\
f_{\mathrm{MowC}}(\theta) &= \sum_{k=1}^{m} \frac{\pi_k}{2\pi} \frac{1 - \rho_k^2}{1 + \rho_k^2 - 2\rho_k \cos(\theta - \mu_k)}, \\
f_{\mathrm{MossvM}}(\theta) &= \sum_{k=1}^{m} \frac{\pi_k (1 + \lambda_k \sin(\theta - \mu_k))}{2\pi I_0(\kappa_k)} \exp(\kappa_k \cos(\theta - \mu_k)), \\
f_{\mathrm{MosswC}}(\theta) &= \sum_{k=1}^{m} \frac{\pi_k}{2\pi} \frac{1 - \rho_k^2}{1 + \rho_k^2 - 2\rho_k \cos(\theta - \mu_k)} (1 + \lambda_k \sin(\theta - \mu_k)),
\end{aligned}$$

respectively, where $\mu_k \in [0, 2\pi)$ is the mean direction, $\kappa_k \in [0, \infty)$ and $\rho_k \in [0, 1]$ is a concentration parameter, $\lambda_k \in [-1, 1]$ is a skewing parameter, $I_0(\cdot)$ is the modified Bessel function of the first kind and order zero, $m \in \mathbb{N}$ is the number of components of the mixture and $0 < \pi_1, \ldots, \pi_m < 1$ are the weights of the components satisfying $\sum_{k=1}^{m} \pi_k = 1$. The number of free parameters for MovM and MowC are $3m - 1$ and that for MossvM and MosswC are $4m - 1$. These distributions are popular mixtured discussed, (e.g., in Wallace and Dowe, 2000; Mooney et al., 2003; Banerjee et al., 2005; Mulder et al., 2020; Miyata et al., 2020). Note that the mixture of sine-skewed wrapped Cauchy is a submodel

of proposed model (Kato and Jones, 2015, Supplementary Material).

The 50-folds cross-validation is conducted to compare the performance of the proposed model with that of four existing models. The cross-validated log-likelihood is employed as the criterion for comparison (Smyth, 2000). The number of components is set to two as in the proposed model.

Table 4.8: The average value of cross-validated log-likelihood function in 50-folds cross validation for MovM, MowC, MossvM, MosswC and proposed model.

model	log-likelihood function
proposed	-36818.5
MovM	-37059.4
MowC	-37731.7
MossvM	-37026.3
MosswC	-36855.3

Table 4.8 shows the average value of cross-validated log-likelihood function for each model. The proposed model's predictive performance is superior to the other models because the value for proposed model is the largest. In contrast to the proposed model that allows varying values of kurtosis and skewness, MovM and MowC do not allow both, while MossvM and MosswC do not allow varying kurtosis. In terms of the number of parameters, MovM and MowC have m less than the proposed model, and MossvM and MosswC have the same. The proposed model is superior to the other models in terms of predictive performance because the proposed model can vary its shape flexibly with a small number of parameters.

4.7.2 The proposed mixture with more than two components

The mixed Kato–Jones distributions Eq. (4.4) whose number of components m is more than two are estimated. In this subsection, models with $m \leq 6$ are discussed because of stability of estimation; estimation for the model with $m > 6$ was carried out but the results were unstable. Recall that the histogram shows two peaks that can be interpreted as shown

above and the assumption of more than two peaks is not necessarily insightful from the transport engineering viewpoint. As such, the main purpose of this subsection is not to find the most accurate number of components but to find out what shapes of distributions will be obtained.

The 50-folds cross-validations for the proposed model whose number of components m from 2 to 6 are conducted. The conditions for the estimation are the same as those in Section 4.5.

Table 4.9: The average value of cross-validated log-likelihood function in 50-folds cross validation for proposed model whose number of components m from 2 to 6.

m	log-likelihood function
2	-36818.5
3	-36818.7
4	-36808.6
5	-36817.6
6	-36820.2

Table 4.9 shows the average value of cross-validated log-likelihood function for each number of components. The model whose number of components $m = 4$ is the best because the value for proposed model is the largest. The plot and value of the parameter for the estimated model whose number of components $m = 4$ are shown in Fig. 4.7 and Table 4.10. Note that the number of bins in Fig. 4.7 is 96 (i.e., the width of the bin equals 15 minutes) for visualization.

In the results of $m = 4$, two prominent components with $k = 1, 2$ have a large value of $\hat{\pi}_k$. As well as the results of $m = 2$, the components with $k = 1, 2$ represent the peak around $\pi/2$ and $3\pi/2$, respectively. However, the component with $k = 2$ for $m = 4$ is not skewed compared to the component with $k = 2$ for $m = 2$ shown in Fig. 4.6. Not only the component with $k = 2$ but also that with $k = 4$ in the results of $m = 4$ represent the shape between $3\pi/2$ and 2π that is represented by only one component with $k = 2$ in the results of $m = 2$. In addition, the component with $k = 3$ has a small value of $\hat{\pi}_3$ and represents the small peak around 16:00 seen in the histogram.

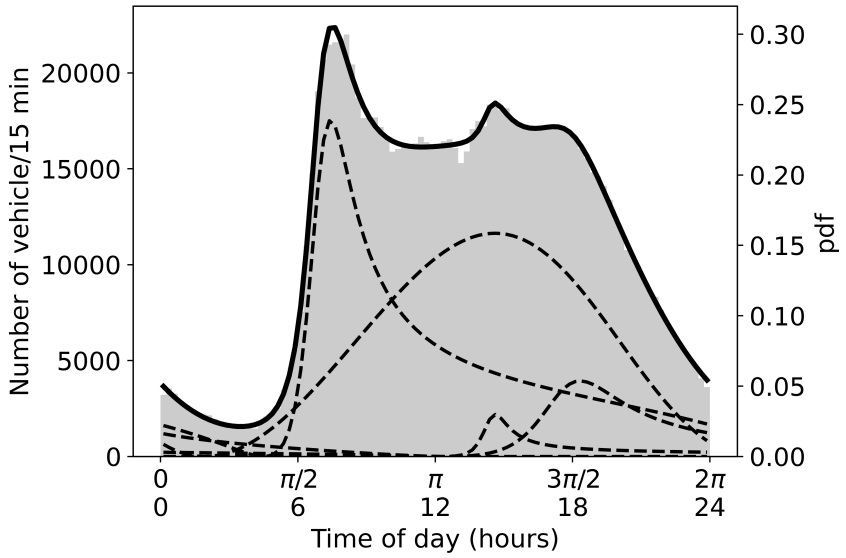


Figure 4.7: Plot of maximum likelihood fits of the densities with the numbers of components $m = 4$.

The components with $k = 3, 4$ may not be meaningful from the transport engineering viewpoint; they only represent the trivial peak along with the improvement of the value of the log-likelihood function. The interpretation of the parameters for these smaller components is difficult with only the traffic volume data. Some additional information such as each vehicle's origin, destination, and trip purpose may help the interpretation for the parameters.

Table 4.10: Maximum likelihood estimates of the parameters of Kato–Jones mixtures Eq. (4.4) whose numbers of components $m = 4$.

k	$\hat{\mu}_k$	$\hat{\gamma}_k$	$\hat{\bar{\alpha}}_{2k}$	$\hat{\bar{\beta}}_{2k}$	$\hat{\pi}_k$
1	2.6665	0.4376	0.2128	-0.2452	0.3855
2	3.7380	0.5065	0.0076	0.0225	0.4907
3	4.3642	0.4887	0.3402	-0.2284	0.0285
4	5.2889	0.6037	0.2628	-0.2166	0.0952

4.8 Conclusions

In this chapter, the probability distribution of the daily variation of traffic volume is estimated. The mixture of Kato–Jones distributions is employed, which can provide a wide range of skewness and kurtosis for each component. To estimate the parameters of these distributions, the modified method of moments developed. This method ensures that the estimated parameters always belong to the parameter space. Then the maximum likelihood method for the proposed mixture was established using the EM algorithm.

The proposed method was applied to some traffic counter data from Japan. As a result, the data were classified into two components, which were interpreted as the outward and return trip. Some reasonable explanations for these components are suggested, that the difference in the shapes of the components might be caused by the different behaviours underlying the commuting trips between morning and evening hours. In addition, the modified method of moments was seen to allow for fast calculation and provided a reasonable initial value for maximum likelihood estimation. With this initial value, the EM algorithm provided reasonable maximum likelihood estimates.

According to the cross-validated log-likelihood function, the model whose number of components $m = 4$ is the best model. However, the interpretation of the trivial components in the model are difficult thus far. As future work, further consideration of missing factors that generate third and later components may be necessary. Some additional information such as each vehicle's origin, destination, and trip purpose would be useful for this.

Chapter 5 Spatial Transportation Model with Angle: Traffic State by Travel Directions

5.1 Introduction

Traffic congestion occurs when the supply of road network is insufficient compared to demand at urban-scale levels. Therefore, the relationship between supply and demand has been extensively investigated in the literature. Specifically, the impact of the topology of network (i.e., connectivity) on transportation state has been widely studied (Sheffi, 1985; Yang and H. Bell, 1998; Yin and Ieda, 2001; Farahani et al., 2013; Rempe et al., 2016; Sun et al., 2018). Alternatively, simpler area-based approaches that focus on aggregated transportation state over a wide area have also been employed, such as the continuous modeling approach (Beckmann, 1952; Sasaki et al., 1990; Wong, 1998) and the network or macroscopic fundamental diagram (Mahmassani et al., 1984; Williams et al., 1987; Daganzo and Geroliminis, 2008; Ji et al., 2014; Saeedmanesh and Geroliminis, 2017).

These topology- or area-based approaches have been highly successful in providing valuable insights into urban-level transportation state analysis. For example, it has been observed that roads in urban centers experience higher congestion as a result of their higher accessibility , which can be evaluated by betweenness centrality. Furthermore, the implementation of detour routes has been shown to reduce overall travel time. Additionally, the capability of throughput of demand depends on traffic volume (i.e., fundamental diagram) at the network level, which does not depend on demand patterns (Daganzo and Geroliminis, 2008).

However, these approaches encounter challenges in explaining some phenomena. For

This chapter is mainly based on the research with Prof. Toru Seo, in *arXiv* preprint (Nagasaki and Seo, 2023).

example, although it is possible to quantify the transportation state of each path, providing a comprehensive explanation of their sensitivity to the overall demand pattern is difficult. Moreover, it is challenging to clarify the relationships between state and the physical characteristics of the road network, such as whether it follows a grid pattern or not.

These can be easily described by adding the variable of "angle". It appears on the demand side as each origin-destination (OD) pair associated with a traveler also possesses an angle. The pattern of demand can be captured by aggregating the individual OD angles. On the supply side, the shape of a road network can be characterized by the distribution of the angles or *orientation* of each road in the network, which can be called *morphology* of road networks. Analyzing transportation state by specifically focusing on these angles could potentially offer insights and contribute to a better understanding of transportation dynamics.

In concrete, an angle-based approach may offer certain advantages over topology- or area-based approaches. For example, by representing travel demand using angles, it becomes easier to describe the relationship of that particular demand within the overall demand pattern. This could potentially enable a more comprehensive explanation of state sensitivity to demand patterns. Furthermore, it could facilitate describing the relationship between the physical shape of the road network, characterized by angular indicators, and transportation dynamics. By considering angles, a deeper understanding of how the road network's geometry influences state could be attained.

However, angular indicators are difficult to analyze mathematically due to its periodicity as mentioned in Chapter 2. To overcome the difficulty, the field of statistics has developed a specialized branch known as *Directional Statistics* (Mardia and Jupp, 2000), which is designed to analyze angular data. However, its application to transportation research, particularly in the context of traffic state analysis, remains limited.

There have been several studies that have utilized angular indicators to describe various aspects of road traffic. For example, Boeing (2019) and Nagasaki et al. (2019) employed a rose diagram, which is a histogram for angular data, to depict the distribution of road orientations within a road network. These studies focused on describing

the physical shape of the road network (i.e., morphology) rather than the traditional approach of analyzing the network based on connectivity relationships between nodes (i.e., topology). On the other hand, Zhou (2015) analyzed massive bike-sharing data in Chicago by employing a rose diagram to visualize the direction of the trip. In this study, each trip is described as the angle of a vector from the origin to the user's destination. However, no studies have investigated transportation state with employing angular analysis in both the demand and supply sides.

In this context, the aim of this chapter is to build a novel descriptive model that captures the transportation state by two angular distributions: the direction of demand and the shape of the road network. The proposed model utilizes directional statistical methodology to predict direction-dependent transportation state based on the angular distributions of supply and demand. To validate the model, actual traffic data from Tokyo is utilized.

The chapter is organized as follows. Section 5.2 describes a state model with angular distribution and details of two distributions, the estimation process, and the expected property of the model. Section 5.3 describes a case study on the actual road network with the actual data and the interpretation of the results. Section 5.4 concludes this chapter.

5.2 Model

5.2.1 Overview

In the proposed model, the author aim to capture the state for travel in a specific direction $\theta \in [0, 2\pi)$ by the angular distribution of demand and the road network.

Let $c(\theta)$ represent the state for travel in the direction θ within a given area, $d(\theta)$ denote the angular distribution of demand, and $n(\theta)$ do the angular distribution of the road network.

Here, $c(\theta)$ is expected to be influenced by demand and road network at angles different from θ itself. For example, high demand in the north-south direction or insufficient road

network development in the north-south direction may result in congestion in the east-west direction. The aim is to build a model that accounts for such complexities. However, as a pioneering study in state analysis using angles, this chapter proposes a simplified model to demonstrate the concept.

5.2.2 State model

To account for the influence of demand and road network characteristics on state levels $c(\theta)$ in different directions, two additional variables are introduced, $\phi \in [-\pi, \pi)$ and $\eta \in [-\pi, \pi)$. These variables allow us to describe $c(\theta)$ in terms of the demand distribution at angles different from θ, denoted by $d(\theta+\phi)$, and the road network characteristics at angles different from θ, denoted by $n(\theta + \eta)$.

Furthermore, the impact of these factors on $c(\theta)$ is expected to depend on the differences in angle, ϕ, and η. When ϕ and η are close to zero, the effect on $c(\theta)$ is anticipated to be significant. Conversely, when ϕ and η are around $\pm\pi$, indicating opposite directions, the effect is expected to be smaller. At $\pm\pi/2$, indicating orthogonal directions, there may be a special effect due to the presence of intersections at right angles.

To incorporate these influences into the model, $\alpha(\phi)$ and $\beta(\eta)$ are introduced that manipulate the degree of impact on $c(\theta)$ based on the values of ϕ and η.

By considering these variables and functions, the proposed model aims to capture the complex relationship between state, demand distribution, road network characteristics, and the angular differences between them.

Based on the above, $c(\theta)$ can be described as

$$c(\theta) = \int_{\phi}\int_{\eta} f\big(d(\theta + \phi)\alpha(\phi), n(\theta + \eta)\beta(\eta)\big)d\eta d\phi. \tag{5.1}$$

Here, it is expected that $d(\theta)$ and $n(\theta)$ intricately influence $c(\theta)$. For simplicity, $d(\theta)$ and $n(\theta)$ are assumed to affect $c(\theta)$ independently as the simplest case. Then, Eq. (5.1) can be converted to a linear combination of $d(\theta)$ terms and $n(\theta)$ terms as

$$c(\theta) = \int_{\phi} d(\theta + \phi)\alpha(\phi)d\phi + \int_{\eta} n(\theta + \eta)\beta(\eta)d\eta + \gamma, \tag{5.2}$$

where γ is a constant term.

Here, it is difficult to handle the $\alpha(\phi)$ and $\beta(\eta)$ because these variables are angles. Therefore, the following transformation is performed to achieve linear regression with angle variables (Johnson and Wehrly, 1978).

$$\begin{aligned} \alpha(\phi) &= \sum_{k=1}^{K}(\alpha_{ck}\cos k\phi + \alpha_{sk}\sin k\phi), \\ \beta(\eta) &= \sum_{k=1}^{K}(\beta_{ck}\cos k\eta + \beta_{sk}\sin k\eta), \end{aligned} \tag{5.3}$$

where K is an external parameter, an integer that determines up to which degrees the distributions are combined. As with the Fourier series expansion, employing the terms with higher degrees K allows us to obtain a finer distribution of $\alpha(\phi)$ and $\beta(\eta)$.

Substituting Eq. (5.3) into Eq. (5.2) gives

$$\begin{aligned} c(\theta) = \int_{\phi}\Big\{d(\theta+\phi)\sum_{k=1}^{K}(\alpha_{ck}\cos k\phi + \alpha_{sk}\sin k\phi)\Big\}d\phi \\ + \int_{\eta}\Big\{n(\theta+\eta)\sum_{k=1}^{K}(\beta_{ck}\cos k\eta + \beta_{sk}\sin k\eta)\Big\}d\eta + \gamma. \end{aligned} \tag{5.4}$$

If each distribution is discrete rather than continuous, then

$$\begin{aligned} c(\theta) = \sum_{\phi\in\Phi}\Big\{d(\theta+\phi)\sum_{k=1}^{K}(\alpha_{ck}\cos k\phi + \alpha_{sk}\sin k\phi)\Big\} \\ + \sum_{\eta\in H}\Big\{n(\theta+\eta)\sum_{k=1}^{K}(\beta_{ck}\cos k\eta + \beta_{sk}\sin k\eta)\Big\} + \gamma. \end{aligned} \tag{5.5}$$

where Φ and H are the set of values of the difference between the representative point of each bin of the discretized angular distribution of demand d and the road network n and θ, respectively. The parameters to be estimated are α_{ck}, α_{sk}, β_{ck}, β_{sk} and γ. Note that this model is a linear regression model. The estimation method is discussed later.

Fig. 5.1 describes the conceptual diagram of the model. In the proposed state model, $c(\theta)$ represents the state, while $d(\theta)$ and $n(\theta)$ correspond to the angular distributions of demand and road network, respectively.

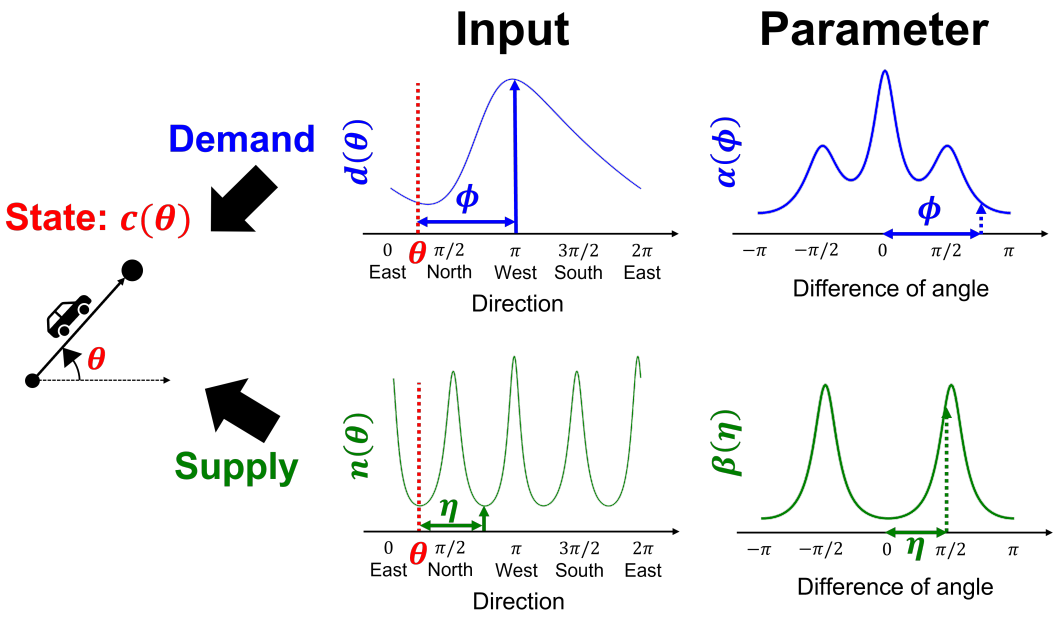


Figure 5.1: Conceptual diagram of the model.

Considering that $c(\theta)$ can be influenced by the entire distribution of $d(\theta)$ and $n(\theta)$, the model incorporates the effects of these distributions based on the values of ϕ and η. To quantify this influence, functions $\alpha(\phi)$ and $\beta(\eta)$ are introduced, representing the degrees of influence depending on the values of ϕ and η.

Overall, this state model describes $c(\theta)$ for each trip using the distributions of $d(\theta)$ and $n(\theta)$ within the target area as explanatory variables, along with $\alpha(\phi)$ and $\beta(\eta)$ as corresponding parameters. By considering these factors, the model aims to capture the complex relationship between state, demand distribution, road network characteristics, and the influence of angular differences.

5.2.3 Specification

The proposed model does not restrict the method of deriving $c(\theta)$, $d(\theta)$ and $n(\theta)$. This section introduces an example of a method for deriving them.

(1) State

Typical examples of indicators for traffic state include flow, speed, and density. Speed and its inverse, known as pace, which can be easily obtained from vehicle trajectory data, are easily applicable for representing the traffic state of a trip in a particular direction. Alternatively, direction dependent traffic states can be described by introducing an direction variable into the spatio-temporal domain, using the method described in Chapter 2.5.

(2) Demand

The calculation of $d(\theta)$ relies on data containing origin and destination information. $d(\theta)$ is obtained by aggregating the directions from the origin to the destination for demand within the target area. The origin and destination points can also be the coordinates of entering or exiting the target area rather than the origin and destination points of the entire trip. Therefore, if trajectory data is used to derive $c(\theta)$, the data can be used as demand data. However, note that demand, in this case, is only the observed traffic volume, which is different from the actual travel demand.

(3) Network

The angular distribution of the road network $n(\theta)$ is the aggregate of the angles that the roads face, called orientation, within the target area (Boeing, 2019; Nagasaki et al., 2019). In this distribution, the peak is sharper when many roads are oriented in a particular direction, like a grid network of roads as shown in the left of Fig. 5.2. When the road network is not a grid, the peaks are gradual, and the skirts are higher, as shown in the right of Fig. 5.2. The large number of roads oriented in a particular direction can be regarded as a high capacity to serve demand in that direction. In addition, by applying corrections based on the number of lanes or road types for each road, it is possible to represent the performance differences among road characteristics. Therefore, this method is a suitable distribution for $n(\theta)$, the supply indicator in this analysis.

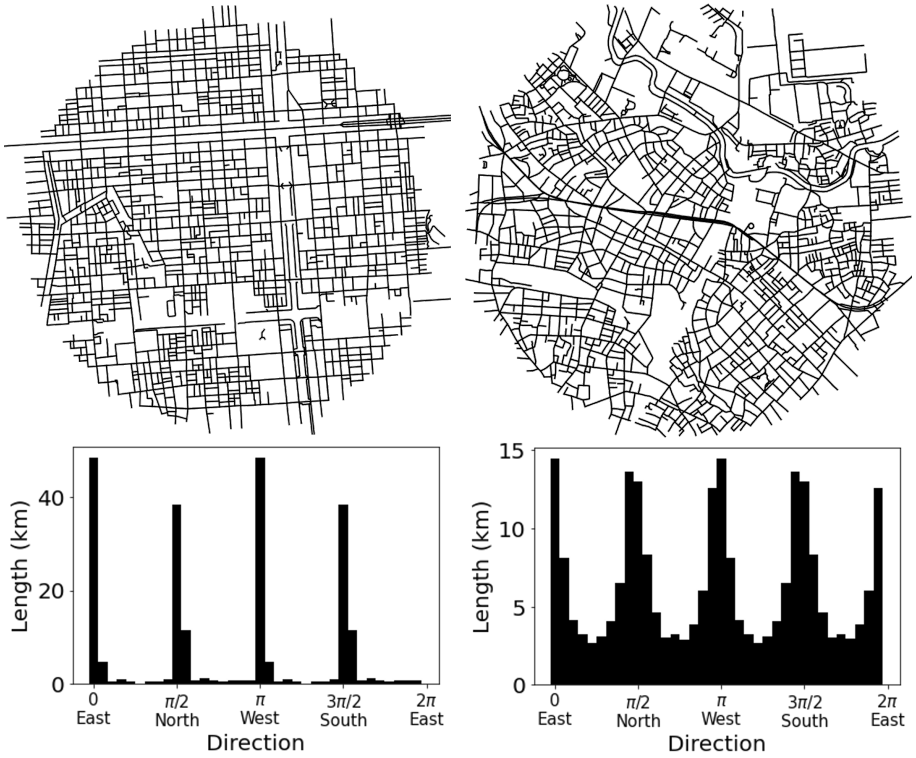


Figure 5.2: Example of $n(\theta)$ and corresponding road networks for a 1 km radius area. Maps exported from digital national land information by Ministry of Land, Infrastructure, Transport and Tourism.

Note that these distributions can be transformed into continuous circular distributions easily by employing the distributions used in directional statistics and their estimation methods (Nagasaki et al., *submitted*).

5.2.4 Estimation

The parameters to be estimated in the model are α_{ck}, α_{sk}, β_{ck}, β_{sk} and γ. In addition, K is an external parameter that determines the fine granularity of the estimated $\alpha(\phi)$ and $\beta(\eta)$ and needs to be determined a priori before the estimation. A larger K may provide a better estimate; however, it is not recommended because of the interpretability of the estimation results.

Eq. (5.4) and Eq. (5.5) are linear regression models and their parameters are obtained

by the least squares method. Eq. (5.5) can be transformed as

$$c(\theta) = \sum_{k=1}^{K} \Big\{ \alpha_{ck} \sum_{\phi \in \Phi} d(\theta+\phi) \cos k\phi + \alpha_{sk} \sum_{\phi \in \Phi} d(\theta+\phi) \sin k\phi + \beta_{ck} \sum_{\eta \in H} n(\theta+\eta) \cos k\eta + \beta_{sk} \sum_{\eta \in H} n(\theta+\eta) \sin k\eta \Big\} + \gamma. \quad (5.6)$$

The parameters can be estimated by performing a linear regression with the given $d(\theta)$ and $n(\theta)$ and the trigonometric values substituted, and the state indicator $c(\theta)$ as the objective function.

In addition, by substituting the estimated α_{ck}, α_{sk}, β_{ck} and β_{sk} into Eq. (5.3), the distributions $\alpha(\phi)$ and $\beta(\eta)$ of the magnitude of the effect of $d(\theta+\phi)$ and $n(\theta+\eta)$ on $c(\theta)$ by the values of ϕ and η can be obtained. The shape of the distributions provides a quantitative indication of the magnitude to which demand in a given direction is affected by other direction demand and road, e.g., orthogonal.

Note that if the distribution of the road network is point symmetric like Boeing (2019), i.e., all roads are aggregated in both directions, then the distribution of $\beta(\eta)$ is also point symmetric. Therefore, the terms β_{ck} and β_{sk} can be deleted in the case k is odd.

5.2.5 Discussion

This model is a simple state model using the angular distribution of demand and road network. The parameters to be estimated are α_{ck}, α_{sk}, β_{ck}, β_{sk} for each degree k. However, the interpretation of the model is enhanced by examining the shape of the $\alpha(\phi)$ and $\beta(\eta)$ distributions formed by these parameters. In addition, the change in state $c(\theta)$ when the angular characteristics of demand or network change can be facilitated by using the estimated parameters $\alpha(\phi)$ and $\beta(\eta)$.

Because various factors cause the transportation phenomenon, the distribution of $\alpha(\phi)$ and $\beta(\eta)$ could be complicated and multimodal. Some expected results are: (1) $\alpha(0)$ is expected to be positive because much demand in the same direction is expected to cause more congestion in that direction. (2) $\beta(0)$ is expected to be negative because the lack of

road construction in a given direction is expected to cause congestion in demand toward that direction.

5.3 Case Study

5.3.1 Data and variables

The case study was conducted in the vicinity of the Tokyo metropolitan area. The data used to derive $c(\theta)$ and $d(\theta)$ were obtained from vehicle trajectory data collected between July 12, 2021 (Monday) and July 16, 2021 (Friday). Hereafter, the penetration rate of probe vehicle is assumed to be uniform within the region.

All trips that enter or commence in the area within a time period are reflected in $d(\theta)$ and $c(\theta)$ is derived from that pace. For each trip, the start and end times are determined when the probe car enters or exits the target area, or when the car commences or concludes a trip within the area. The duration of the trip is calculated as the difference between these times, resulting in the trip time (in seconds). $c(\theta)$ is computed as the ratio of travel time to the distance covered (in kilometers) for each trip. On the other hand, $d(\theta)$ represents the aggregate of the angles formed between the origin and destination of each trip.

To ensure data quality, the top 10% of $c(\theta)$ values were excluded, as some trips were deemed ongoing despite the vehicle staying in one location for an extended period. Similarly, the bottom 10% of $c(\theta)$ values were omitted due to their apparent high speeds.

The network distribution $n(\theta)$ was derived using OpenStreetMap. The analysis focused on major roads classified as ”motorway”, ”trunk”, ”primary”, and ”secondary”. The capacity is assumed to be equal for all roads, i.e., $n(\theta)$ is calculated using only the road lengths.

Both $d(\theta)$ and $n(\theta)$ were normalized frequencies represented as discrete histograms with 32 bins. The parameter K was set to 8 for improved interpretability.

Four case studies were conducted. Cases 1 and 2 are the 8:00 a.m. and 6:00 p.m. trip data for a 10 km square area located north of the center of Tokyo. Cases 3 and 4 are the

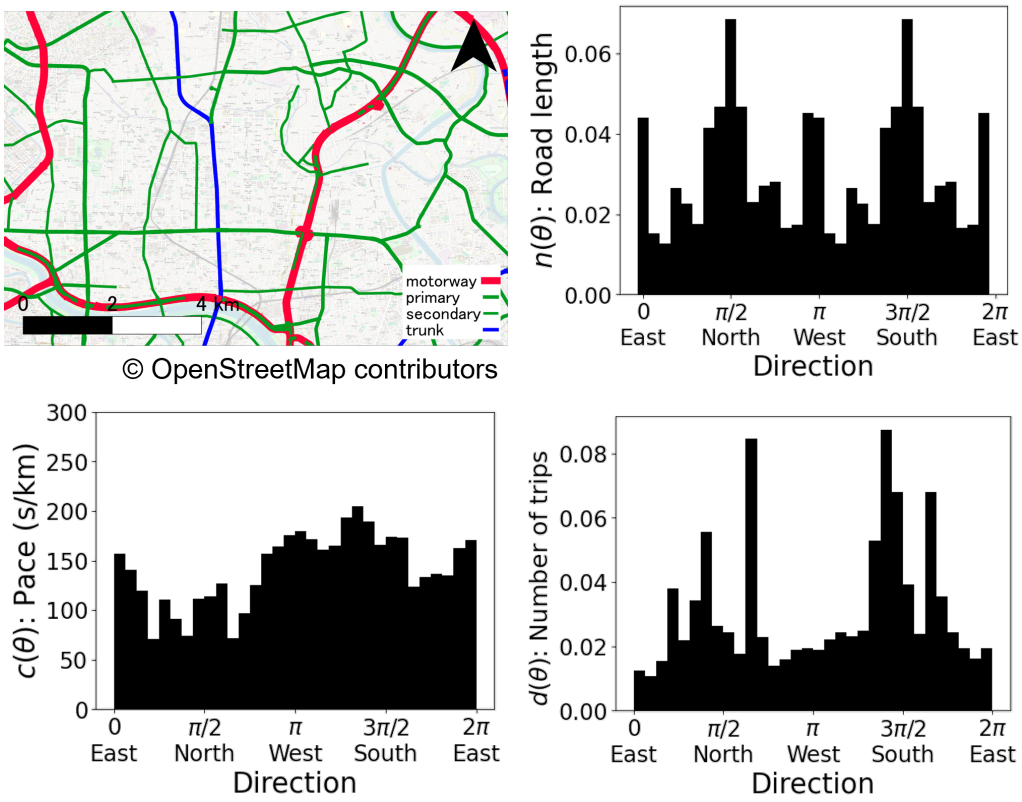


Figure 5.3: The map of the road network and the distributions (upper left) of $c(\theta)$ (bottom left), $d(\theta)$ (bottom right), and $n(\theta)$ (upper right) for Case 1 (North of the center of Tokyo in morning). The map exported from OpenStreetMap.

8:00 a.m. and 6:00 p.m. trip data for a 10 km square area located west of the center of Tokyo. The map of the road network and the distributions of $c(\theta)$, $d(\theta)$, and $n(\theta)$ for each case are shown in Fig. 5.3, Fig. 5.4, Fig. 5.5 and Fig. 5.6.

For Case 1 and Case 2, $d(\theta)$ is higher in the south direction toward the center of Tokyo and in the opposite direction to the north. In addition, $c(\theta)$ in these directions is small because of the motorways. $n(\theta)$ has a significant peak in the north-south direction due to the large number of roads in that direction. For Case 3 and Case 4, $d(\theta)$ is higher in the east direction toward the center of Tokyo and in the opposite direction to the west. However, $c(\theta)$ is generally large because there are no motorways in the area. $n(\theta)$ has a significant peak in the east-west direction due to the large number of roads in that direction.

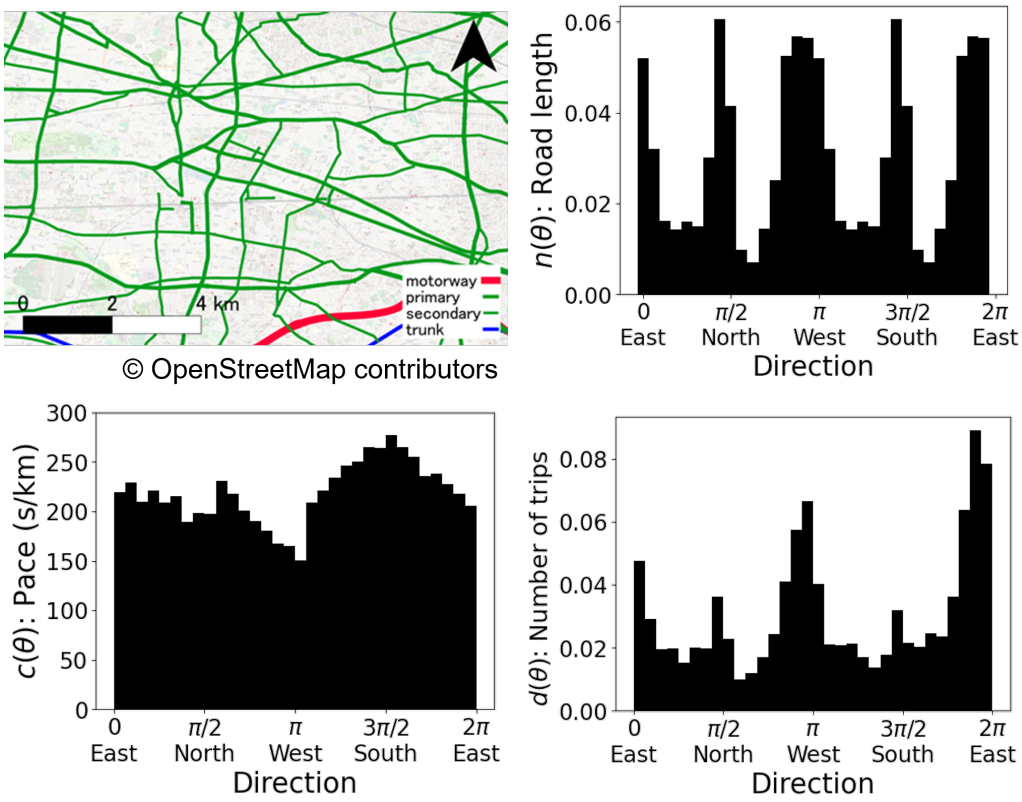


Figure 5.4: The map of the road network and the distributions (upper left) of $c(\theta)$ (bottom left), $d(\theta)$ (bottom right), and $n(\theta)$ (upper right) for Case 2 (North of the center of Tokyo in evening). The map exported from OpenStreetMap.

5.3.2 Estimation results

The distributions of $\alpha(\phi)$ and $\beta(\eta)$ estimated in each case are shown in Fig. 5.7. The value of R^2, F-statistic, Prob(F-statistic) are shown in Table 5.1, Table 5.2, Table 5.3 and Table 5.4. Note that $\alpha(\phi)$ and $\beta(\eta)$ are calculated by only the 5% significant parameters. In each case, the number of parameters that are 5% significant out of a total of 16 α_{ck} and α_{sk} is 15, 13, 8 and 9, respectively. In addition, the number of parameters that are 5% significant out of a total of 8 β_{ck} and β_{sk} is 5, 8, 7 and 5, respectively.

Estimated parameters for each model are shown in Table 5.5, Table 5.6, Table 5.7 and Table 5.8. "*" means 5% significant.

According to the value of Prob(F-statistic), all models are significant, and each parameter, including the constant term γ, is also working significantly.

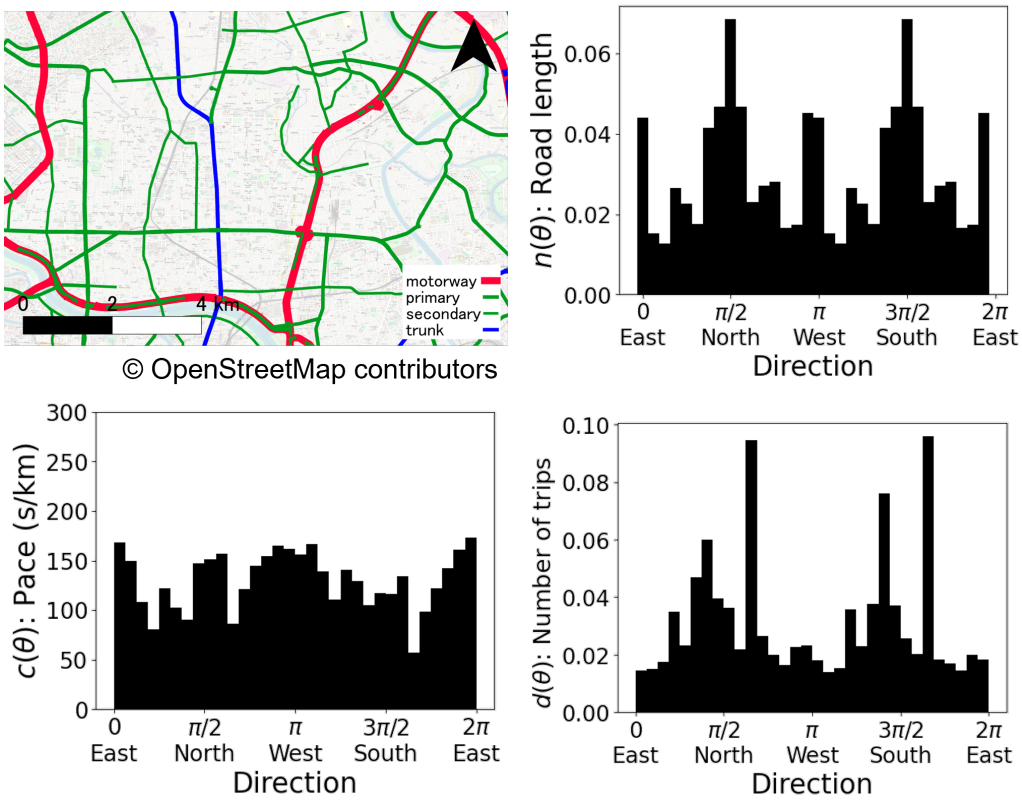


Figure 5.5: The map of the road network and the distributions (upper left) of $c(\theta)$ (bottom left), $d(\theta)$ (bottom right), and $n(\theta)$ (upper right) for Case 3 (West of the center of Tokyo in morning). The map exported from OpenStreetMap.

Table 5.1: Case 1 results.

	Value
Number of samples	10322
R^2	0.278
F-statistic	247.9
Prob(F-statistic)	0.000

In addition, the heatmaps of correlation matrix for all cases are shown in Fig. 5.8. According to the heatmaps, there is a strong correlation between α_k and α_k of the same degree for all case studies. On the other hand, there is no correlation within α_k and β_k, and between α_k and β_k of different orders.

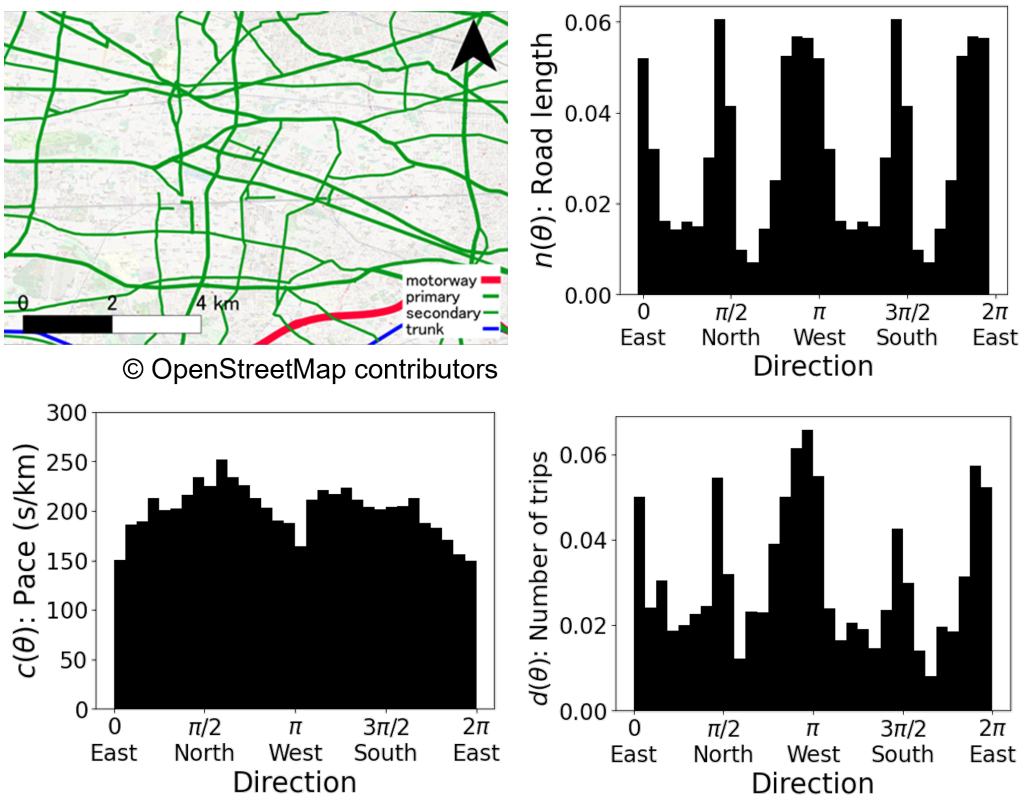


Figure 5.6: The map of the road network and the distributions (upper left) of $c(\theta)$ (bottom left), $d(\theta)$ (bottom right), and $n(\theta)$ (upper right) for Case 4 (West of the center of Tokyo in evening). The map exported from OpenStreetMap.

Table 5.2: Case 2 results.

	Value
Number of samples	8494
R^2	0.168
F-statistic	107.1
Prob(F-statistic)	0.000

5.3.3 Discussion

The distribution depicted in Fig. 5.7 illustrates the impact of $d(\theta + \phi)$ and $n(\theta + \eta)$ values on $c(\theta)$. In Case 1, for instance, if $d(\theta)$ is 0.1, $c(\theta)$ increases by approximately 20, whereas if $d(\theta+\pi)$ is 0.1, $c(\theta)$ decreases by approximately 30. However, it is unlikely that demand

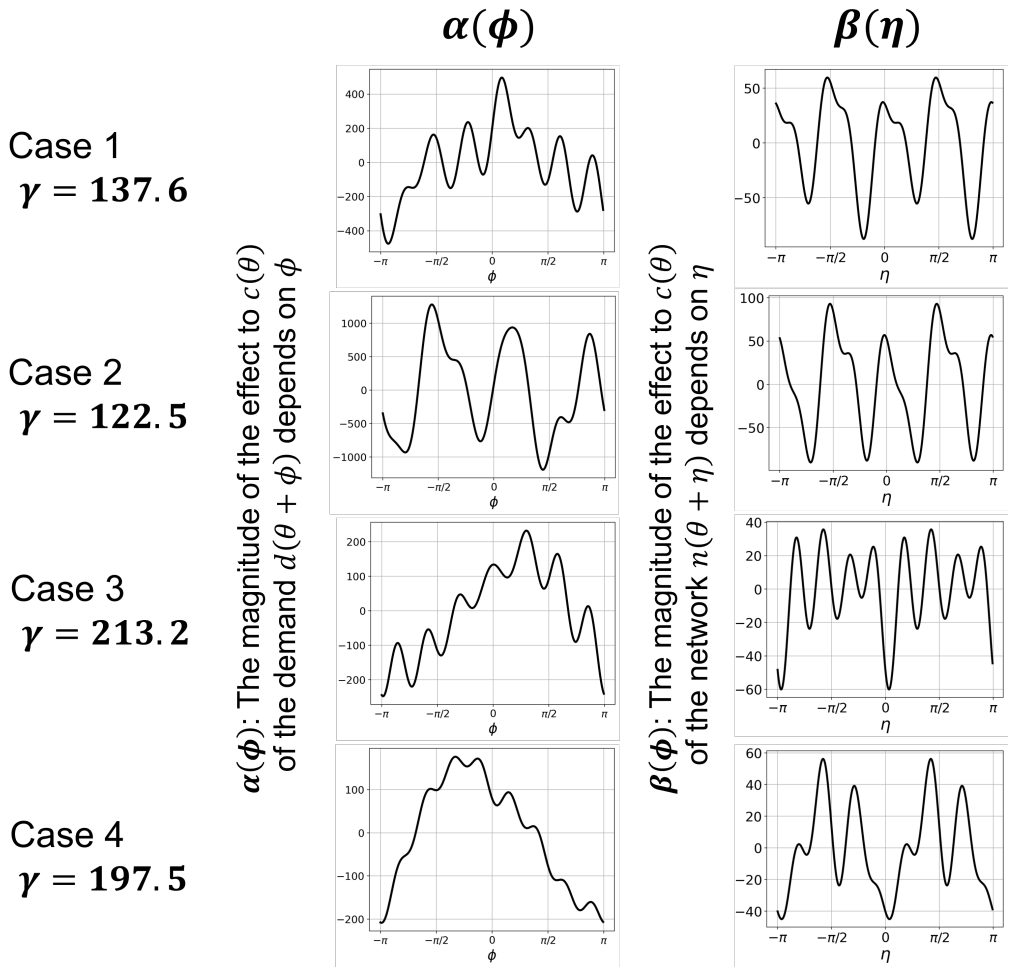


Figure 5.7: The distribution of $\alpha(\phi)$ and $\beta(\eta)$ estimated in each case.

in the opposite direction would decrease congestion. On the contrary, it is reasonable to assume that such demand would have no impact on congestion, rather than reducing it. This discrepancy arises from the possibility of $\alpha(\phi)$ having negative values, inherent to the characteristics of the model. Therefore, it is natural to interpret $\alpha(\phi)$ not based on its value but rather by its incremental change from the smallest value.

Across all cases, $\alpha(\phi)$ is positive and exhibits significant magnitudes near $\phi = 0$. This outcome aligns with expectations, as it indicates that a high volume of demand in the same direction leads to increased congestion in that specific direction. Furthermore, $\alpha(\phi)$ is significantly small around $\phi = \pi$, which is also expected, as demand in the opposite direction is unlikely to affect congestion.

Table 5.3: Case 3 results.

	Value
Number of samples	3127
R^2	0.166
F-statistic	38.59
Prob(F-statistic)	0.000

Table 5.4: Case 4 results.

	Value
Number of samples	2677
R^2	0.165
F-statistic	32.86
Prob(F-statistic)	0.000

However, the overall shape of $\alpha(\phi)$ differs between Case 1 and Case 2, despite the same target area. Although the distributions of the two $d(\theta)$ in Fig. 5.3 and Fig. 5.4 are not significantly different, the shape of $c(\theta)$ in the southern direction differs. This discrepancy can be attributed to the heavy traffic towards central Tokyo during the morning peak period, resulting in congestion even for travelers using motorways. Conversely, congestion is reduced in the evening due to less crowded motorways and faster travel speeds. These factors contribute to the differing shapes of $\alpha(\phi)$ in Case 1 and Case 2.

In the comparison between Case 3 and Case 4, the trend is line symmetric around $\phi = 0$. This may be due to the demand for commuting in the morning and returning home in the evening being in opposite directions, resulting in a linear symmetry of $d(\theta)$.

In Cases 1, 3 and 4, $\alpha(\phi)$ generally exhibits high values near $\phi = 0$, contrasting with the highly oscillatory $\alpha(\phi)$ observed in Case 2. This discrepancy can be attributed to the distribution of demand in directions other than the peak, as seen in $d(\theta)$ for Cases 1, 3 and 4.

Table 5.5: Estimated parameters of Case 1.

	Coefficient	Std. err.	t value		Coefficient	Std. err.	t value
γ	137.58	0.52	*263.58				
α_{c1}	216.36	5.79	*37.37	α_{s1}	43.72	6.93	*6.31
α_{c2}	-16.86	2.04	*-8.26	α_{s2}	22.59	1.92	*11.75
α_{c3}	48.72	12.54	*3.89	α_{s3}	89.02	11.74	*7.58
α_{c4}	2.27	0.46	*4.93	α_{s4}	-7.05	0.44	*-15.99
α_{c5}	-58.45	6.72	*-8.70	α_{s5}	59.25	6.88	*8.62
α_{c6}	-18.11	4.49	*-4.04	α_{s6}	5.05	5.05	1.00
α_{c7}	45.1	14.33	*3.15	α_{s7}	138.76	14.23	*9.75
α_{c8}	-18.96	2.03	*-9.34	α_{s8}	15.20	2.08	*7.30
β_{c2}	-13.70	1.18	*-11.56	β_{s2}	8.93	1.12	*8.00
β_{c4}	45.84	3.18	*14.43	β_{s4}	24.61	3.16	*7.78
β_{c6}	3.84	1.15	*3.35	β_{s6}	2.10	1.08	1.95
β_{c8}	-0.34	1.77	-0.19	β_{s8}	-20.79	1.75	*-11.9

Additionally, $\beta(0)$ is negative in Case 3 and Case 4, whereas it is positive in Case 1 and Case 2. This contradicts initial expectations, as it suggests that congestion occurs when the road network aligns with travel demand. This discrepancy may caused by the differences between the diagonal motorway and the many east-west, north-south, and west-oriented roads. While $c(\theta)$ in the diagonal direction is smaller due to the presence of the motorway, $n(\theta)$ in that direction is also lower compared to other directions. Hence, $\beta(0)$ is positive in Cases 1 and 2.

Furthermore, when comparing $\beta(\eta)$ in Case 1 and Case 2, the impact of the road network on state exhibits similar shapes during both morning and evening periods. This finding aligns with the results of Daganzo and Geroliminis (2008), indicating that the road network's influence on state follows a similar pattern regardless of demand patterns.

However, the shapes of $\beta(\eta)$ in Case 3 and Case 4 are significantly different. In

Table 5.6: Estimated parameters of Case 2.

	Coefficient	Std. err.	t value		Coefficient	Std. err.	t value
γ	122.51	0.61	*202.45				
α_{c1}	176.85	33.97	*5.21	α_{s1}	-264.15	31.55	*-8.37
α_{c2}	-39.93	2.39	*-16.69	α_{s2}	-7.02	2.24	*-3.13
α_{c3}	118.86	84.20	1.41	α_{s3}	857.92	76.55	*11.21
α_{c4}	-6.32	0.34	*-18.67	α_{s4}	-0.10	0.37	-0.26
α_{c5}	-183.21	50.87	*-3.60	α_{s5}	139.63	51.10	*2.73
α_{c6}	-48.26	4.54	*-10.64	α_{s6}	17.65	5.08	*3.47
α_{c7}	225.38	57.00	*3.95	α_{s7}	72.50	60.37	1.20
α_{c8}	-33.20	2.30	*-14.41	α_{s8}	18.87	2.32	*8.12
β_{c2}	-22.17	1.44	*-15.37	β_{s2}	-10.25	1.35	*-7.60
β_{c4}	62.81	3.74	*16.81	β_{s4}	19.86	3.67	*5.41
β_{c6}	6.52	0.98	*6.64	β_{s6}	7.96	0.95	*8.41
β_{c8}	6.16	1.73	*3.57	β_{s8}	-27.54	1.69	*-16.26

Cases 1, 2 and 4, $\beta(\eta)$ exhibits four peaks, which could be attributed to the road network distribution $n(\theta)$ also having four peaks. However, Case 3 displays eight peaks in $\beta(\eta)$, although $n(\theta)$ has four peaks. In addition, the four peaks of $\beta(\eta)$ in Case 1 and Case 2 are equally spaced, whereas those in Case 4 are not. This is due to smaller pace $c(\theta)$ in the east-west direction, where there are many roads, but larger pace $c(\theta)$ in the other directions. In addition, the small $\beta(\eta)$ in the $\pm\pi/2$ direction is a different result from Case 1 and Case 2. The interpretation of these differences in shape remains elusive at present.

Furthermore, the magnitudes of $\alpha(\phi)$ and $\beta(\eta)$ differ by approximately tenfold. This indicates that demand has a stronger influence on state than the shape of the road network.

Table 5.7: Estimated parameters of Case 3.

	Coefficient	Std. err.	t value		Coefficient	Std. err.	t value
γ	213.22	0.70	*304.66				
α_{c1}	126.15	8.22	*15.35	α_{s1}	103.58	10.02	*10.34
α_{c2}	-21.72	2.33	*-9.33	α_{s2}	-4.70	2.70	-1.74
α_{c3}	12.31	16.10	0.76	α_{s3}	-7.70	17.30	-0.45
α_{c4}	-8.38	1.33	*-6.32	α_{s4}	-3.22	1.39	*-2.31
α_{c5}	58.17	36.68	1.59	α_{s5}	47.46	35.92	1.32
α_{c6}	-1.80	4.44	-0.41	α_{s6}	-11.60	4.31	*-2.69
α_{c7}	62.79	30.64	*2.05	α_{s7}	43.56	30.24	1.44
α_{c8}	-25.01	5.36	*-4.67	α_{s8}	0.08	5.29	0.01
β_{c2}	-11.38	1.24	*-9.17	β_{s2}	-3.23	1.44	*-2.25
β_{c4}	-9.99	1.76	*-5.68	β_{s4}	-6.75	1.89	*-3.57
β_{c6}	-13.33	4.93	*-2.71	β_{s6}	1.59	5.08	0.31
β_{c8}	-13.67	5.48	*-2.49	β_{s8}	-22.25	5.64	*-3.95

5.4 Conclusions

This chapter built a novel descriptive model that describes the transportation state by overall demand pattern and network morphology. The proposed model describes direction-dependent transportation state using the two angular distributions using the Directional Statistical methodology. In concrete, the model describes state using the angular distribution of demand, described by the direction of origin and destination of trips in the target area, and the angular distribution of the road network, described by the direction in which all roads are oriented. To incorporate the two angular distributions into the model, the shape of the angular distribution is transformed into a cosine term and a sine term using directional statistics. This made it possible to describe state with a simple linear regression model with angular indicators. Furthermore, a case study of the proposed model was conducted in the Tokyo Metropolitan Area. The model worked

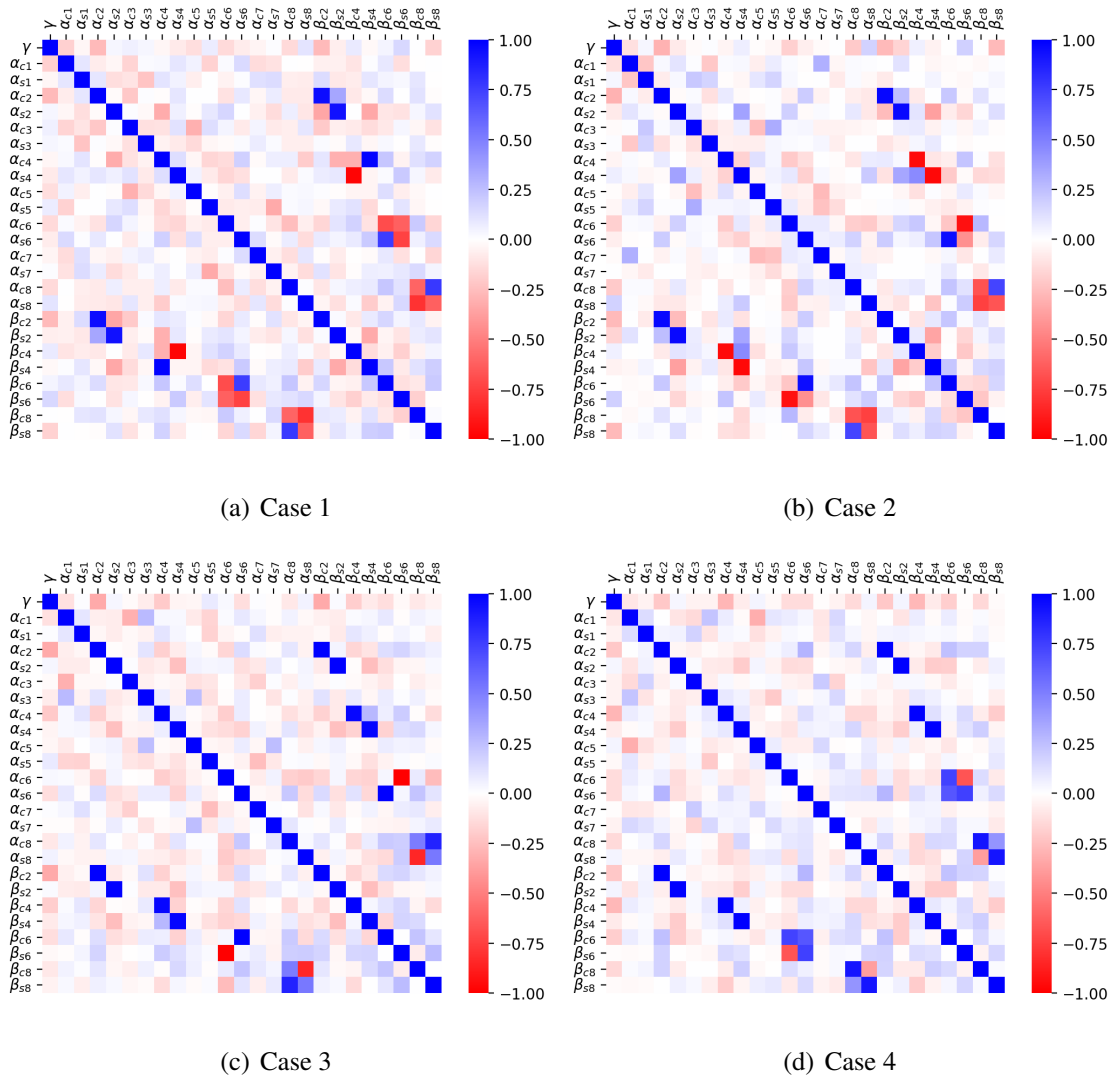
Table 5.8: Estimated parameters of Case 4.

	Coefficient	Std. err.	t value		Coefficient	Std. err.	t value
γ	197.48	0.74	*266.35				
α_{c1}	143.77	12.98	*11.08	α_{s1}	-87.44	12.72	*-6.88
α_{c2}	-31.32	2.85	*-11.00	α_{s2}	2.30	3.47	0.66
α_{c3}	29.94	23.00	1.30	α_{s3}	21.41	25.79	0.83
α_{c4}	-9.87	1.38	*-7.17	α_{s4}	-7.01	1.47	*-4.78
α_{c5}	23.02	54.38	0.42	α_{s5}	-85.00	54.50	-1.56
α_{c6}	-6.35	2.91	*-2.18	α_{s6}	2.18	2.91	0.75
α_{c7}	9.13	41.08	0.22	α_{s7}	-79.11	41.32	-1.91
α_{c8}	-16.39	5.61	*-2.92	α_{s8}	-14.31	5.61	*-2.55
β_{c2}	-26.94	2.45	*-10.97	β_{s2}	3.07	3.00	1.02
β_{c4}	-13.23	1.89	*-7.01	β_{s4}	-10.09	2.02	*-4.99
β_{c6}	-9.23	8.12	-1.14	β_{s6}	16.43	8.23	*2.00
β_{c8}	-7.72	4.61	-1.67	β_{s8}	-16.31	4.70	*-3.47

significantly in the case study.

The current model has several problems, such as multicollinearity and difficulty in interpretation of the results. Improving the proposed model involves integrating it with existing traffic theories and models while concurrently improving the interpretability and statistical robustness of the model.

To the author's knowledge, this is the first attempt to describe transportation state using angular indicators with Directional Statistics. Therefore, there is still room for further development. First, the connection to well-established concepts in transportation researches, such as Macroscopic Fundamental Diagram (Daganzo and Geroliminis, 2008) would be worth investigating. Second, a dynamic model that introduces a time term could be developed by extending this result. Time loops in 24 hours so that it can be described as angles (Nagasaki et al., *submitted*). The data with two angular axes, demand and time,



(a) Case 1 (b) Case 2

(c) Case 3 (d) Case 4

Figure 5.8: "The heatmap of the correlation matrices for the estimated parameters in each case..

can be described by a torus, which is a doughnut-shaped surface. The establishment of more advanced models using the latest findings in directional statistics is one of the development directions of this research.

Chapter 6 Conclusions

Angle is a concept that describes the direction of an object or a looping time axis. On the other hand, transportation covers the arrangement of geographic objects such as road networks and the movement of people and its dynamics. Although both seem compatible, only few attempts have been made to address transportation from an angular perspective. This dissertation modeled the transportation systems with angle from three aspects: behavior, temporal state, and spatial state. This chapter summarizes the main conclusions of this dissertation and shows the possible future directions of transportation systems analysis with angle.

6.1 Achievements

Chapter 2 reviewed concepts of angular analysis and studies related to this dissertation. First, directional statistics is summarized, which is statistics to analyze angles. Specifically, the procedure of angular data, drawing of data, probability distribution, and regression analysis in directional statistics are explained, and examples of applications other than transportation are summarized. Angular analysis has contributed to many fields, and many of these analyses have focused on the movement of animals. Next, the transportation analysis is classified into demand, supply, behavior, and transportation states. The elementary analytical methods and the recent research on each concept are summarized. In addition, studies that employed angular analysis for each concept were reviewed. There were several studies on each concept that used angles. However, some of them did not analyze the angles correctly. Moreover, no studies discussed the state from both supply and demand perspectives by angle.

In Chapter 3, a route choice model with angular variables was built and validated by empirical studies. In discrete choice models, right/left turns and U-turns are often

described by setting dummy variables. However, the results vary depending on the predetermined threshold value. In addition, although both behaviors are essential changes in travel direction, they are rarely analyzed as continuous variables without separating them. Therefore, drivers' route choice behavior on the network was modeled using turn angles at each intersection. The route choice model was applied to drivers' actual route choice behavior in Tokyo. The results showed that the angular indicator significantly affected drivers' route choice behavior. In particular, a 90-degree angular change (i.e., a right-angle left-right turn) corresponds to a travel distance of approximately 1.2 km. In addition, this model enables the evaluation of the resistance of arbitrary turn angles, a unique feature of this model. An angular change of 45 degrees corresponds to approximately 0.35 km, similar to previous studies using right/left turn dummies. The proposed model was compared with models that minimize only the distance, only the turn angle, and the number of turns and distance, showing that the proposed model more accurately reflects actual route choice behavior.

In Chapter 4, a model of the daily variation of traffic volume by converting time into angle in order to represent periodicity was built. Since the time in a day is looped in 24 hours, conversion of the time into angles is necessary to describe the variation depending on time. Traffic volume varies with time and has two peaks: morning peak on the way to work and evening peak on the way home. The shape of the two peaks is significantly different, with the morning peak rising quickly and sharply while the evening peak is gentle. In describing these variations in traffic volume over time using a probability distribution, a circular distribution that can represent multimodality and flexible shapes must be employed. Therefore, the mixture of Kato–Jones distribution that satisfies this property was applied to the time variation of traffic volume. The modified method of moments based on the usual method of moments is proposed for a quick and stable estimation method for the mixed Kato–Jones distribution. This estimation method can obtain estimate within the parameter domain that is almost the same as the estimator of the method of moments, with a small computational cost. In addition, because it is a consistent estimator, it can be helpful as an initial value in maximum likelihood

estimation, which requires numerical optimization. The stability of the estimation method was assessed and interpreted when the time variation was divided into multiple parts. Thanks to the model's flexibility to represent complex distribution shapes, the temporal variations in traffic volume can be divided into two clusters with reasonable accuracy. The two clusters can be interpreted as commuting and returning trips within a day. The possibility of investigating transportation using only pulse data, which is easy to obtain, was indicated.

In Chapter 5, a model that describes the states of OD demand for a certain direction by the demand pattern and the network morphology for the entire area was built. Describing the entire demand pattern within an area is difficult by conventional approaches. By aggregating the angles between the origin and destination of each demand, the overall demand pattern can be quantitatively described. In addition, the relationship between transportation states and the morphology of road networks, such as whether the road network is grid-like or not, is also difficult to describe using conventional methods. The distribution of the direction of each road in an area can intuitively and quantitatively describe the morphology of the road network. In this chapter, the transportation states of demand in a given direction were regressed on the angular distributions of the overall demand pattern and the morphology of the road network. Regressions with angular variables require the use of methods that take into account periodicity. In addition, the proposed model was validated by transportation data for Tokyo. The features of the proposed model, including its consistency with existing transportation theory, are discussed based on the results of the empirical analysis.

Modeling using angles in three scales: behavior, temporal state, and spatial state, is performed, and a framework for transportation analysis with angle is built. Based on the modeling and empirical analysis above, two possibilities for the transportation systems analysis with angles are suggested.

- Investigating phenomena that are difficult to describe with existing methods
- Providing a new perspective on transportation analysis

First, the analysis with angles suggested the possibility of capturing transportation phenomena that are difficult to describe with existing methods. For example, the relationships between the overall demand patterns, the physical shape of the road network, and transportation states targeted in Chapter 5 are difficult to analyze concisely using existing methods. The possibility of quantitatively description this relationship was suggested by describing the overall demand pattern and the physical shape of the road network in terms of angles and performing a regression using the knowledge of directional statistics, as proposed in Chapter 5. In addition, capturing the share of types of transportation from only the traffic volume at a certain point suggested in Chapter 4 is very difficult with the existing analysis.

Second, analysis with angles can provide a new perspective on the existing transportation analysis. The behavior of angle change in route choice targeted in Chapter 3 has been performed using dummy variables in previous studies. By describing changes in direction, such as left-right turns, in terms of angles, behaviors suggested in previous analyses can be captured more precisely. In addition, Chapter 4 shows that the analysis of only traffic volume data at a certain location, which is easily available, can be interpreted commonly with existing traffic analysis methods such as departure time selection. The analysis using data other than traffic volume can potentially analyze transportation phenomena from different perspectives with consistency with existing interpretations.

In conclusion, a new framework for the transportation systems analysis with angles built in this dissertation has the potential to provide results that can be interpreted similarly to existing methods, simultaneously allowing for a more simplified analysis of phenomena that are difficult to describe with existing methods, and this dissertation has shown some of these possibilities.

6.2 Future directions

The usefulness and potential of angle-based transportation analysis are indicated from Chapter 3 to Chapter 5. This section presents technical tasks and future directions of

transportation systems analysis with angle to leverage angular analysis to the fullest extent based on the achievements in this dissertation.

6.2.1 Technical tasks of transportation systems analysis with angle

Technical tasks to be addressed in the future are listed below.

Incorporation of heading angle and other factors in route choice behavior

Chapter 3 used only the turn angle as the angular variable in the utility function. Among angular indicators, the direction in which the driver is heading and the direction to the destination may also likely be easily recognized indicators. In addition, the behavior of desiring to head in the direction of the destination has been shown in previous works (Dalton, 2003). By incorporating these two variables, the impact of angular variables on route choice behavior can be analyzed from a different perspective. In addition, to analyze the economic impact of angular variables, a utility function incorporating a time variable must be developed.

Sensitivity analysis of angular indicators in route choice behavior

Since the angle change is described only by the cosine term in Chapter 3, the distinction between a right turn and a left turn is not taken into account. In addition, since only a first-order cosine term is used, the angle changes of going straight and turning left/right and turning right/left and making a U-turn are equally 1 different on the variables. However, these two angle changes are considered to be very different in nature. To express these differences in the variables, a possible solution is to add a sine or cosine term of first or higher order (Johnson and Wehrly, 1978). This would allow a more detailed analysis of the sensitivity of the angular change and the traveler's choice behavior.

6.2.2 Future directions of transportation systems analysis with angle

For more meaningful transportation systems analysis with angles, analysis from the unique perspective of angular indicators should be conducted. The future directions of

transportation systems analysis with angle based on these objectives listed below.

Integration with other transportation data in temporal state analysis

Chapter 4 showed the possibility of interpreting traveler behavior to some extent, even using only traffic volume data. Combining the traffic volume data with other data such as departure time and origin and destination may explore a more comprehensive analysis. Moreover, the angular demand distribution discussed in Chapter 5 can help the interpretation of each component. For example, it may be possible to interpret each distribution further or the components with $m = 3$ or more. In addition, the potential of combinations with conventional transportation theories, such as departure time choice (Noland and Small, 1995), are also possible.

Derivation of the theoretical model of spatial transportation model

Chapter 5 built a regression model of transportation state using two angular distributions of demand and road network. A theoretical model based on this idea could be developed. For example, an area-based transportation model with angles could be developed by mathematically describing transportation states given a simple network and demand based on existing transportation theory. In addition, simulations using a toy network can be performed to observe the phenomenon more straightforwardly. This study used much noise data, making the fitting result worse and the interpretation difficult. Describing area-based transportation states with angles through simulation or theory is one of the future directions.

Introducing a higher dimension of angular analysis

In this dissertation, angular indicators were used only in one dimension, which means the angle on a circle. On the other hand, there are also descriptions to process angular data of higher dimensions. Among them, the application of a torus (Fig. 2.3) with two angular axes and a cylinder (Fig. 2.4) with a linear axis and an angular axis has a wide range of applications. For example, the simultaneous application of direction and time of demand discussed in Chapter 4 and Chapter 5 is possible using a torus. In other words, dynamic transportation phenomena can be represented using angles. In addition, the direction and travel distance of demand

can be analyzed using cylinders. Higher-order angular data description helps model complex transportation phenomena.

References

Akamatsu, T. (1996). Stochastic traffic assignment with geometric attributes of paths.

Akamatsu, T., Wada, K., and Hayashi, S. (2015). The corridor problem with discrete multiple bottlenecks. *Transportation Research Procedia*, 7:474–498.

Alexander, L., Jiang, S., Murga, M., and González, M. C. (2015). Origin–destination trips by purpose and time of day inferred from mobile phone data. *Transportation research part c: emerging technologies*, 58:240–250.

Anacleto, O., Queen, C., and Albers, C. J. (2013). Multivariate forecasting of road traffic flows in the presence of heteroscedasticity and measurement errors. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 62(2):251–270.

Anderson, J. E. (2011). The gravity model. *Annu. Rev. Econ.*, 3(1):133–160.

Andrienko, G., Andrienko, N., Chen, W., Maciejewski, R., and Zhao, Y. (2017). Visual analytics of mobility and transportation: State of the art and further research directions. *IEEE Transactions on Intelligent Transportation Systems*, 18(8):2232–2249.

Andrienko, G., Andrienko, N., Fuchs, G., and Wood, J. (2016). Revealing patterns and trends of mass mobility through spatial and temporal abstraction of origin-destination movement data. *IEEE transactions on visualization and computer graphics*, 23(9):2120–2136.

Antonini, G., Bierlaire, M., and Weber, M. (2006). Discrete choice models of pedestrian walking behavior. *Transportation Research Part B: Methodological*, 40(8):667–687.

Arnott, R., de Palma, A., and Lindsey, R. (1990). Departure time and route choice for the morning commute. *Transportation Research Part B: Methodological*, 24(3):209–228.

Banerjee, A., Dhillon, I. S., Ghosh, J., Sra, S., and Ridgeway, G. (2005). Clustering on the unit hypersphere using von mises-fisher distributions. *Journal of Machine Learning Research*, 6(9).

Barabási, A.-L. (2013). Network science. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(1987):20120375.

Batschelet, E. (1981). Circular statistics in biology.

Beckmann, M. (1952). A continuous model of transportation. *Econometrica: Journal of the Econometric Society*, pages 643–660.

Ben-Akiva, M. E., Ramming, M. S., and Bekhor, S. (2004). Route choice models. In *Human Behaviour and Traffic Networks*, pages 23–45. Springer.

Boeing, G. (2019). Urban spatial order: Street network orientation, configuration, and entropy. *Applied Network Science*, 4(1):67.

Boeing, G. (2020). A multi-scale analysis of 27,000 urban street networks: Every us city, town, urbanized area, and zillow neighborhood. *Environment and Planning B: Urban Analytics and City Science*, 47(4):590–608.

Bongiorno, C., Zhou, Y., Kryven, M., Theurel, D., Rizzo, A., Santi, P., Tenenbaum, J., and Ratti, C. (2021). Vector-based pedestrian navigation in cities. *Nature Computational Science*, 1(10):678–685.

Boomsma, W., Mardia, K. V., Taylor, C. C., Ferkinghoff-Borg, J., Krogh, A., and Hamelryck, T. (2008). A generative, probabilistic model of local protein structure. *Proceedings of the National Academy of Sciences*, 105(26):8932–8937.

Bowers, J., Morton, I., and Mould, G. (2000). Directional statistics of the wind and waves. *Applied ocean research*, 22(1):13–30.

Brackstone, M., Waterson, B., and McDonald, M. (2009). Determinants of following headway in congested traffic. *Transportation Research Part F: Traffic Psychology and Behaviour*, 12(2):131–142.

Butler, R. (1998). Seasonality in tourism: Issues and implications. *The Tourist Review*, 53(3):18–24.

Button, K. (2010). *Transport economics*. Edward Elgar Publishing.

Carta, J. A., Ramirez, P., and Bueno, C. (2008). A joint probability density function of wind speed and direction for wind energy analysis. *Energy Conversion and Management*, 49(6):1309–1320.

Codling, E. A., Plank, M. J., and Benhamou, S. (2008). Random walk models in biology. *Journal of the Royal society interface*, 5(25):813–834.

Crucitti, P., Latora, V., and Porta, S. (2006). Centrality in networks of urban streets. *Chaos: an interdisciplinary journal of nonlinear science*, 16(1).

Daganzo, C. F. (1994). The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation research part B: methodological*, 28(4):269–287.

Daganzo, C. F. (1997). *Fundamentals of transportation and traffic operations*. Emerald Group Publishing Limited.

Daganzo, C. F. (2005). A variational formulation of kinematic waves: basic theory and complex boundary conditions. *Transportation Research Part B: Methodological*, 39(2):187–196.

Daganzo, C. F. (2010). Structure of competitive transit networks. *Transportation Research Part B: Methodological*, 44(4):434–446.

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

Dalton, R. C. (2003). The secret is to follow your nose: Route path selection and angularity. *Environment and Behavior*, 35(1):107–131.

Dantsuji, T., Hoang, N. H., Zheng, N., and Vu, H. L. (2022). A novel metamodel-based framework for large-scale dynamic origin–destination demand calibration. *Transportation Research Part C: Emerging Technologies*, 136:103545.

Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap methods and their application*. Number 1. Cambridge university press.

de Haas-Lorentz, G. L. (2013). *Die Brownsche Bewegung und einige verwandte Erscheinungen*. Springer-Verlag.

De Palma, A. and Lindsey, R. (2011). Traffic congestion pricing methodologies and technologies. *Transportation Research Part C: Emerging Technologies*, 19(6):1377–1399.

Dong, Y., Wang, S., Li, L., and Zhang, Z. (2018). An empirical study on travel patterns of internet based ride-sharing. *Transportation research part C: emerging technologies*, 86:1–22.

Downs, T. D. and Mardia, K. (2002). Circular regression. *Biometrika*, 89(3):683–698.

Duckham, M. and Kulik, L. (2003). "simplest" paths: automated route selection for navigation. In *Spatial Information Theory. Foundations of Geographic Information Science: International Conference, COSIT 2003, Kartause Ittingen, Switzerland, September 24-28, 2003. Proceedings 6*, pages 169–185. Springer.

Edie, L. C. et al. (1963). *Discussion of traffic stream measurements and definitions*. Port of New York Authority New York.

Eren, E. and Uz, V. E. (2020). A review on bike-sharing: The factors affecting bike-sharing demand. *Sustainable cities and society*, 54:101882.

Esteves, C., Allen-Blanchette, C., Makadia, A., and Daniilidis, K. (2018). Learning so (3) equivariant representations with spherical cnns. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 52–68.

Farahani, R. Z., Miandoabchi, E., Szeto, W. Y., and Rashidi, H. (2013). A review of urban transportation network design problems. *European journal of operational research*, 229(2):281–302.

Fisher, N. I. and Lee, A. J. (1992). Regression models for an angular response. *Biometrics*, pages 665–677.

Fosgerau, M., Frejinger, E., and Karlstrom, A. (2013). A link based network route choice model with unrestricted choice set. *Transportation Research Part B: Methodological*, 56:70–80.

Garrison, W. L. (1960). Connectivity of the interstate highway system. *Papers in Regional Science*, 6(1):121–137.

Geroliminis, N. and Daganzo, C. F. (2008). Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological*, 42(9):759–770.

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

Gill, J. and Hangartner, D. (2010). Circular data in political science and how to handle it. *Political Analysis*, 18(3):316–336.

Greenshields, B. D. (1935). A study in highway capacity. *Highway Research Board Proc.*, pages 448–477.

Hamedmoghadam, H., Ramezani, M., and Saberi, M. (2019). Revealing latent characteristics of mobility networks with coarse-graining. *Scientific reports*, 9(1):7545.

Hanshin Expressway (2017). Csr report on 2017. *viewed 25 May 2023* https://www.hanshin-exp.co.jp/company/csr/files/report/csrreport_201707_2.pdf.

Hillier, B. and Iida, S. (2005). Network and psychological effects in urban movement. In *Spatial Information Theory: International Conference, COSIT 2005, Ellicottville, NY, USA, September 14-18, 2005. Proceedings 7*, pages 475–490. Springer.

Hong, W.-C. (2011). Traffic flow forecasting by seasonal svr with chaotic simulated annealing algorithm. *Neurocomputing*, 74(12-13):2096–2107.

Huang, H., Cheng, Y., and Weibel, R. (2019). Transport mode detection based on mobile phone network data: A systematic review. *Transportation Research Part C: Emerging Technologies*, 101:297–312.

Jenelius, E., Petersen, T., and Mattsson, L.-G. (2006). Importance and exposure in road network vulnerability analysis. *Transportation Research Part A: Policy and Practice*, 40(7):537–560.

Ji, Y., Luo, J., and Geroliminis, N. (2014). Empirical observations of congestion propagation and dynamic partitioning with probe data for large-scale systems. *Transportation Research Record*, 2422(1):1–11.

Johnson, R. A. and Wehrly, T. E. (1978). Some angular-linear distributions and related regression models. *Journal of the American Statistical Association*, 73(363):602–606.

Jona-Lasinio, G., Gelfand, A., and Jona-Lasinio, M. (2012). Spatial analysis of wave direction data using wrapped gaussian processes.

Jones, E., Oliphant, T., Peterson, P., et al. (2001). Scipy: Open source scientific tools for python.

Kansky, K. J. (1963). *Structure of transportation networks: relationships between network geometry and regional characteristics*. The University of Chicago.

Kato, S. and Jones, M. (2010). A family of distributions on the circle with links to, and applications arising from, möbius transformation. *Journal of the American Statistical Association*, 105(489):249–262.

Kato, S. and Jones, M. (2015). A tractable and interpretable four-parameter family of unimodal distributions on the circle. *Biometrika*, 102(1):181–190.

Kato, S., Shimizu, K., and Shieh, G. S. (2008). A circular–circular regression model. *Statistica Sinica*, pages 633–645.

Kometani, E. and Sasaki, T. (1958). On the stability of traffic flow (report-i). *Journal of the Operations Research Society of Japan*, 2(1):11–26.

Kraft, D. (1988). A software package for sequential quadratic programming. *Forschungsbericht- Deutsche Forschungs- und Versuchsanstalt fur Luft- und Raumfahrt*.

Kuwahara, M. and Akamatsu, T. (1993). Dynamic equilibrium assignment with queues for a one-to-many od pattern. *Transportation and Traffic Theory*, 12:185–204.

Kuwahara, M. and Akamatsu, T. (1997). Decomposition of the reactive dynamic assignments with queues for a many-to-many origin-destination pattern. *Transportation Research Part B: Methodological*, 31(1):1–10.

Kuwahara, M. and Newell, G. F. (1987). Queue evolution on freeways leading to a single core city during the morning peak. In *Proceedings of the 10th International Symposium on Transportation and Traffic Theory*, pages 21–40.

Lämmer, S., Gehlsen, B., and Helbing, D. (2006). Scaling laws in the spatial structure of urban road networks. *Physica A: Statistical Mechanics and its Applications*, 363(1):89–95.

Lana, I., Del Ser, J., Velez, M., and Vlahogianni, E. I. (2018). Road traffic forecasting: Recent advances and new challenges. *IEEE Intelligent Transportation Systems Magazine*, 10(2):93–109.

Lang, M. N., Schlosser, L., Hothorn, T., Mayr, G. J., Stauffer, R., and Zeileis, A. (2020). Circular regression trees and forests with an application to probabilistic wind direction forecasting. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 69(5):1357–1374.

Latora, V. and Marchiori, M. (2001). Efficient behavior of small-world networks. *Physical review letters*, 87(19):198701.

Leduc, G. et al. (2008). Road traffic data: Collection methods and applications. *Working Papers on Energy, Transport and Climate Change*, 1(55):1–55.

Lévy, P. (1939). L'addition des variables aléatoires définies sur une circonférence. *Bulletin de la Société mathématique de France*, 67:1–41.

Ley, C., Babić, S., and Craens, D. (2021). Flexible models for complex data with applications. *Annual Review of Statistics and Its Application*, 8:369–391.

Lighthill, M. J. and Whitham, G. B. (1955). On kinematic waves ii. a theory of traffic flow on long crowded roads. *Proceedings of the royal society of london. series a. mathematical and physical sciences*, 229(1178):317–345.

Mahmassani, H. S. and Liu, Y.-H. (1999). Dynamics of commuting decision behaviour under advanced traveller information systems. *Transportation Research Part C: Emerging Technologies*, 7(2-3):91–107.

Mahmassani, H. S., Williams, J. C., and Herman, R. (1984). Investigation of network-level traffic flow relationships: some simulation results. *Transportation Research Record*, 971:121–130.

Mai, T., Fosgerau, M., and Frejinger, E. (2015). A nested recursive logit model for route choice analysis. *Transportation Research Part B: Methodological*, 75:100–112.

Manley, E., Addison, J., and Cheng, T. (2015). Shortest path or anchor-based route choice: a large-scale empirical analysis of minicab routing in london. *Journal of transport geography*, 43:123–139.

Mardia, K. V. and Jupp, P. E. (2000). *Directional statistics*, volume 2. Wiley Online Library.

McClintock, B. T., King, R., Thomas, L., Matthiopoulos, J., McConnell, B. J., and Morales, J. M. (2012). A general discrete-time modeling framework for animal movement using multistate random walks. *Ecological Monographs*, 82(3):335–349.

McLachlan, G. J. and Krishnan, T. (2007). *The EM algorithm and extensions*. John Wiley & Sons.

McMillan, G. P., Hanson, T. E., Saunders, G., and Gallun, F. J. (2013). A two-component circular regression model for repeated measures auditory localization data. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 62(4):515–534.

McNally, M. G. (2007). The four-step model. In *Handbook of transport modelling*, volume 1, pages 35–53. Emerald Group Publishing Limited.

Merchant, D. K. and Nemhauser, G. L. (1978). A model and an algorithm for the dynamic traffic assignment problems. *Transportation science*, 12(3):183–199.

Miyata, Y., Shiohama, T., and Abe, T. (2020). Estimation of finite mixture models of skew-symmetric circular distributions. *Metrika*, 83:895–922.

Modlin, D., Fuentes, M., and Reich, B. (2012). Circular conditional autoregressive modeling of vector fields. *Environmetrics*, 23(1):46–53.

Mohajeri, N. and Gudmundsson, A. (2012). Entropies and scaling exponents of street and fracture networks. *Entropy*, 14(4):800–833.

Mooney, J. A., Helms, P. J., and Jolliffe, I. T. (2003). Fitting mixtures of von mises distributions: a case study involving sudden infant death syndrome. *Computational Statistics & Data Analysis*, 41(3-4):505–513.

Morales, J. M., Haydon, D. T., Frair, J., Holsinger, K. E., and Fryxell, J. M. (2004). Extracting more out of relocation data: building movement models as mixtures of random walks. *Ecology*, 85(9):2436–2445.

Morgan, B., Green, A. M., and Spooner, N. J. (2005). Directional statistics for realistic weakly interacting massive particle direct detection experiments. *Physical Review D*, 71(10):103507.

Mulder, K., Jongsma, P., and Klugkist, I. (2020). Bayesian inference for mixtures of von mises distributions using reversible jump mcmc sampler. *Journal of Statistical Computation and Simulation*, 90(9):1539–1556.

Muñoz, L., Sun, X., Horowitz, R., and Alvarez, L. (2003). Traffic density estimation with the cell transmission model. In *Proceedings of the 2003 American Control Conference, 2003.*, volume 5, pages 3750–3755. IEEE.

Na, J.-H. and Jang, Y.-M. (2011). Modeling on daily traffic volume of local state road using circular mixture distributions. *The Korean Journal of Applied Statistics*, 24(3):547–557.

Nagasaki, K., Kato, S., Nakanishi, W., and Jones, M. C. (2022). Traffic count data analysis using mixtures of kato–jones distributions on the circle. *arXiv preprint arXiv:2206.01355*.

Nagasaki, K., Nakanishi, W., and Asakura, Y. (2019). Application of the rose diagram to road network analysis. In *24th International Conference of Hong Kong Society for Transportation Studies, Hong Kong*.

Nagasaki, K., Nakanishi, W., and Asakura, Y. (2023). Analysis of traffic volume by the estimated parameter of the mixture of circular distribution. *Journal of Japan Society of Civil Engineers, Ser. D3 (Infrastructure Planning and Management)*, 78(5):I_825–I_831.

Nagasaki, K. and Seo, T. (2023). Understanding impact of angle in urban transportation. *arXiv preprint arXiv:2310.16470*.

Nagasaki, K. and Seo, T. (in pressa). Route choice model based on angular indicator and case study. *Journal of Japan Society of Civil Engineers, Ser. D3 (Infrastructure Planning and Management)*.

Nagasaki, K. and Seo, T. (in pressb). Route choice model using angular indicators. In *The 25th EURO Working Group on Transportation Meeting, Santander*.

Nakanishi, W. and Fuse, T. (2015). A preliminary study on analysis of pedestrian traffic line using angular data. *JSTE Journal of Traffic Engineering*, 1(4):A_18–A_23.

Newell, G. F. (1993). A simplified theory of kinematic waves in highway traffic, part I: General theory. *Transportation Research Part B: Methodological*, 27(4):281–287.

Noland, R. B. and Small, K. A. (1995). Travel-time uncertainty, departure time choice, and the cost of morning commutes. *Transportation research record*, 1493:150–158.

Pelletier, M.-P., Trépanier, M., and Morency, C. (2011). Smart card data use in public transit: A literature review. *Transportation Research Part C: Emerging Technologies*, 19(4):557–568.

Pewsey, A. and García-Portugués, E. (2021). Recent advances in directional statistics. *Test*, 30(1):1–58.

Pipes, L. A. (1953). An operational analysis of traffic dynamics. *Journal of applied physics*, 24(3):274–281.

Prato, C. G. (2009). Route choice modeling: past, present and future research directions. *Journal of choice modelling*, 2(1):65–100.

Prato, C. G., Bekhor, S., and Pronello, C. (2012). Latent variables and route choice behavior. *Transportation*, 39:299–319.

Ramming, M. S. (2001). Network knowledge and route choice. *Unpublished Ph. D. Thesis, Massachusetts Institute of Technology*.

Raveau, S., Muñoz, J. C., and De Grange, L. (2011). A topological route choice model for metro. *Transportation Research Part A: Policy and Practice*, 45(2):138–147.

Rempe, F., Huber, G., and Bogenberger, K. (2016). Spatio-temporal congestion patterns in urban traffic networks. *Transportation Research Procedia*, 15:513–524.

Rivest, L.-P., Duchesne, T., Nicosia, A., and Fortin, D. (2016). A general angular regression model for the analysis of data on animal movement in ecology. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 65(3):445–463.

Saeedmanesh, M. and Geroliminis, N. (2017). Dynamic clustering and propagation of congestion in heterogeneously congested urban traffic networks. *Transportation research procedia*, 23:962–979.

Santos, A., McGuckin, N., Nakamoto, H. Y., Gray, D., Liss, S., et al. (2011). Summary of travel trends: 2009 national household travel survey. Technical report, United States. Federal Highway Administration.

Sasaki, T., Iida, Y., and Yang, H. (1990). User-equilibrium traffic assignment by continuum approximation of network flow. In *International Symposium on Transportation and Traffic Theory, 11th, 1990, Yokohama, Japan*.

Sathishkumar, V., Park, J., and Cho, Y. (2020). Using data mining techniques for bike sharing demand prediction in metropolitan city. *Computer Communications*, 153:353–366.

Scott, D. M., Novak, D. C., Aultman-Hall, L., and Guo, F. (2006). Network robustness index: A new method for identifying critical links and evaluating the performance of transportation networks. *Journal of Transport Geography*, 14(3):215–227.

Sheffi, Y. (1985). *Urban transportation networks*, volume 6. Prentice-Hall, Englewood Cliffs, NJ.

Smith, B. L. and Demetsky, M. J. (1997). Traffic flow forecasting: comparison of modeling approaches. *Journal of transportation engineering*, 123(4):261–266.

Smyth, P. (2000). Model selection for probabilistic clustering using cross-validated likelihood. *Statistics and computing*, 10(1):63–72.

Stephens, M. A. (1969). *Techniques for directional data*. Department of Statistics.

Sullivan, J. L., Novak, D. C., Aultman-Hall, L., and Scott, D. M. (2010). Identifying critical road segments and measuring system-wide robustness in transportation networks with isolating links: A link-based capacity-reduction approach. *Transportation Research Part A: Policy and Practice*, 44(5):323–336.

Sun, C., Pei, X., Hao, J., Wang, Y., Zhang, Z., and Wong, S. (2018). Role of road network features in the evaluation of incident impacts on urban traffic mobility. *Transportation research part B: methodological*, 117:101–116.

Sun, Z., Arentze, T., and Timmermans, H. (2012). A heterogeneous latent class model of activity rescheduling, route choice and information acquisition decisions under multiple uncertain events. *Transportation research part C: emerging technologies*, 25:46–60.

Swan, A. R. and Sandilands, M. (1995). Introduction to geological data analysis. In *International Journal of Rock Mechanics and Mining Sciences and Geomechanics Abstracts*, volume 8, page 387A.

Toole, J. L., Colak, S., Sturt, B., Alexander, L. P., Evsukoff, A., and González, M. C. (2015). The path most traveled: Travel demand estimation using big data resources. *Transportation Research Part C: Emerging Technologies*, 58:162–177.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

Transportation Research Board (2000). Highway capacity manual.

Turner, A. (2007). From axial to road-centre lines: a new representation for space syntax and a new model of route choice for transport network analysis. *Environment and Planning B: planning and Design*, 34(3):539–555.

Vickrey, W. S. (1969). Congestion theory and transport investment. *The American Economic Review*, pages 251–260.

Vlahogianni, E. I., Karlaftis, M. G., and Golias, J. C. (2014). Short-term traffic forecasting: Where we are and where we're going. *Transportation Research Part C: Emerging Technologies*, 43:3–19.

von Mises, R. (1918). Über die "ganzzahligkeit" der atomgewichte und verwandete fragen. *Physikalische Zeitschrift*, 19:490.

Wallace, C. S. and Dowe, D. L. (2000). Mml clustering of multi-state, poisson, von mises circular and gaussian distributions. *Statistics and Computing*, 10:73–83.

Wang, Y. and Papageorgiou, M. (2005). Real-time freeway traffic state estimation based on extended kalman filter: a general approach. *Transportation Research Part B: Methodological*, 39(2):141–167.

Wardrop, J. G. (1952). Road paper. some theoretical aspects of road traffic research. *Proceedings of the institution of civil engineers*, 1(3):325–362.

Williams, J. C., Mahmassani, H. S., and Herman, R. (1987). Urban traffic network flow models. *Transportation Research Record*, 1112:78–88.

Wilson, A. (2011). *Entropy in urban and regional modelling*, volume 1. Routledge.

Wong, S. (1998). Multi-commodity traffic assignment by continuum approximation of network flow with variable demand. *Transportation Research Part B: Methodological*, 32(8):567–581.

Wood, J., Dykes, J., and Slingsby, A. (2010). Visualisation of origins, destinations and flows with od maps. *The Cartographic Journal*, 47(2):117–129.

Yamamoto, T., Kitamura, R., and Fujii, J. (2002). Drivers' route choice behavior: analysis by data mining algorithms. *Transportation Research Record*, 1807(1):59–66.

Yan, H. and Lam, W. H. (1996). Optimal road tolls under conditions of queueing and congestion. *Transportation Research Part A: Policy and Practice*, 30(5):319–332.

Yang, H. and H. Bell, M. G. (1998). Models and algorithms for road network design: a review and some new developments. *Transport Reviews*, 18(3):257–278.

Yao, H., Wu, F., Ke, J., Tang, X., Jia, Y., Lu, S., Gong, P., Ye, J., and Li, Z. (2018). Deep multi-view spatial-temporal network for taxi demand prediction. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32.

Yin, Y. and Ieda, H. (2001). Assessing performance reliability of road networks under nonrecurrent congestion. *Transportation Research Record*, 1771(1):148–155.

Zhong, H. (2023). Generation of aggregated road network by only vehicle trajectory data. (Unpablished Master's thesis), Tokyo Institute of Technology, Japan.

Zhong, H., Nakanishi, W., Yasuda, S., and Iryo, T. (2022). Extraction of major intersections by only vehicle trajectory data. *Theory and Applications of GIS*, 30(2):7–18.

Zhong, H., Seo, T., Nakanishi, W., Yasuda, S., Asakura, Y., and Iryo, T. (2023). Generation of aggregated road network by vehicle trajectory data. In *The 10th International Symposium on Transportation Data and Modelling*.

Zhou, X. (2015). Understanding spatiotemporal patterns of biking behavior by analyzing massive bike sharing data in chicago. *PloS one*, 10(10):e0137922.