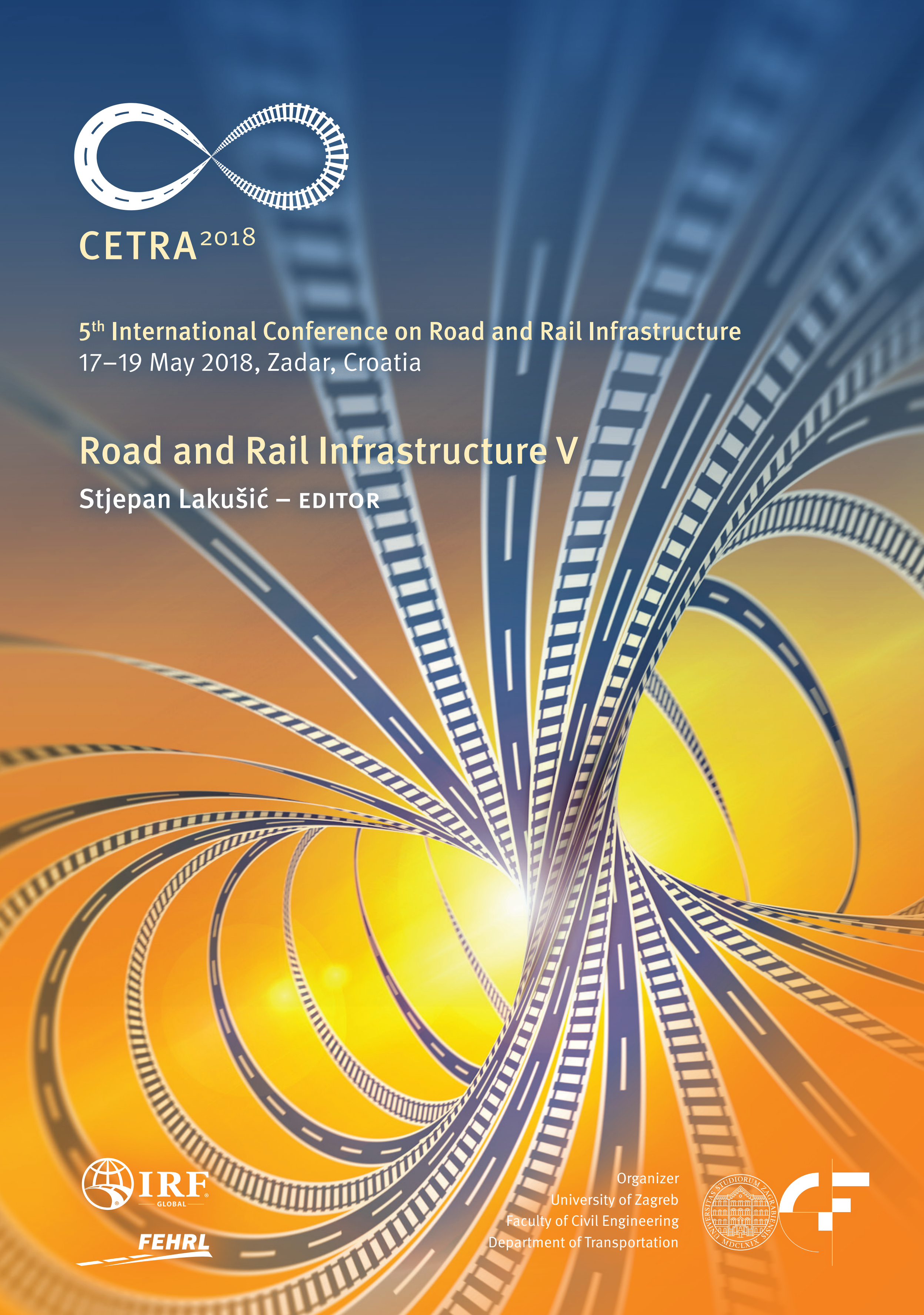

CETRA[2018]

5th International Conference on Road and Rail Infrastructure
17–19 May 2018, Zadar, Croatia

# Road and Rail Infrastructure V

Stjepan Lakušić – EDITOR

IRF GLOBAL

FEHRL


Organizer
University of Zagreb
Faculty of Civil Engineering
Department of Transportation

CETRA[2018]
5th International Conference on Road and Rail Infrastructure
17–19 May 2018, Zadar, Croatia

TITLE
Road and Rail Infrastructure V, Proceedings of the Conference CETRA 2018

EDITED BY
Stjepan Lakušić

ISSN
1848-9850

ISBN
978-953-8168-25-3

DOI
10.5592/CO/CETRA.2018

PUBLISHED BY
Department of Transportation
Faculty of Civil Engineering
University of Zagreb
Kačićeva 26, 10000 Zagreb, Croatia

DESIGN, LAYOUT & COVER PAGE
minimum d.o.o.
Marko Uremović · Matej Korlaet

PRINTED IN ZAGREB, CROATIA BY
"Tiskara Zelina", May 2018

COPIES
500

Zagreb, May 2018.

Proceedings of the
5th International Conference on Road and Rail Infrastructures – CETRA 2018
17–19 May 2018, Zadar, Croatia

# Road and Rail Infrastructure V

**EDITOR**
Stjepan Lakušić
Department of Transportation
Faculty of Civil Engineering
University of Zagreb
Zagreb, Croatia

CETRA$^{2018}$
**5th International Conference on Road and Rail Infrastructure**
17–19 May 2018, Zadar, Croatia

# ORGANISATION

## CHAIRMEN

Prof. **Stjepan Lakušić**, University of Zagreb, Faculty of Civil Engineering
Prof. emer. **Željko Korlaet**, University of Zagreb, Faculty of Civil Engineering

## ORGANIZING COMMITTEE

Prof. Stjepan Lakušić
Prof. emer. Željko Korlaet
Prof. Vesna Dragčević
Prof. Tatjana Rukavina
Assist. Prof. Ivica Stančerić
Assist. Prof. Maja Ahac
Assist. Prof. Saša Ahac
Assist. Prof. Ivo Haladin
Assist. Prof. Josipa Domitrović
Tamara Džambas
Viktorija Grgić
Šime Bezina
Katarina Vranešić
Željko Stepan
Prof. Rudolf Eger
Prof. Kenneth Gavin
Prof. Janusz Madejski
Prof. Nencho Nenov
Prof. Andrei Petriaev
Prof. Otto Plašek
Assist. Prof. Andreas Schoebel
Prof. Adam Szeląg
Brendan Halleman

## INTERNATIONAL ACADEMIC SCIENTIFIC COMMITTEE

**Stjepan Lakušić**, University of Zagreb, president
**Borna Abramović**, University of Zagreb
**Maja Ahac**, University of Zagreb
**Saša Ahac**, University of Zagreb
**Darko Babić**, University of Zagreb
**Danijela Barić**, University of Zagreb
**Davor Brčić**, University of Zagreb
**Domagoj Damjanović**, University of Zagreb
**Sanja Dimter**, J. J. Strossmayer University of Osijek
**Aleksandra Deluka Tibljaš**, University of Rijeka
**Josipa Domitrović**, University of Zagreb
**Vesna Dragčević**, University of Zagreb
**Rudolf Eger**, RheinMain Univ. of App. Sciences, Wiesbaden
**Adelino Ferreira**, University of Coimbra
**Makoto Fujiu**, Kanazawa University
**Laszlo Gaspar**, Széchenyi István University in Győr
**Kenneth Gavin**, Delft University of Technology
**Nenad Gucunski**, Rutgers University
**Ivo Haladin**, University of Zagreb
**Staša Jovanović**, University of Novi Sad
**Lajos Kisgyörgy**, Budapest Univ. of Tech. and Economics
**Anastasia Konon**, St. Petersburg State Transport Univ.
**Željko Korlaet**, University of Zagreb
**Meho Saša Kovačević**, University of Zagreb
**Zoran Krakutovski**, Ss. Cyril and Methodius Univ. in Skopje
**Dirk Lauwers**, Ghent University
**Janusz Madejski**, Silesian University of Technology
**Goran Mladenović**, University of Belgrade
**Tomislav Josip Mlinarić**, University of Zagreb
**Nencho Nenov**, University of Transport in Sofia
**Mladen Nikšić**, University of Zagreb
**Andrei Petriaev**, St. Petersburg State Transport University
**Otto Plašek**, Brno University of Technology
**Mauricio Pradena**, University of Concepcion
**Carmen Racanel**, Tech. Univ. of Civil Eng. Bucharest
**Tatjana Rukavina**, University of Zagreb
**Andreas Schoebel**, Vienna University of Technology
**Ivica Stančerić**, University of Zagreb
**Adam Szeląg**, Warsaw University of Technology
**Marjan Tušar**, National Institute of Chemistry, Ljubljana
**Audrius Vaitkus**, Vilnius Gediminas Technical University
**Andrei Zaitsev**, Russian University of transport, Moscow

CETRA2018 17–19 May 2018, Zadar, Croatia
5th International Conference on Road and Rail Infrastructure

# MIXTURE DISTRIBUTION MODEL OF DEPARTURE TIME DECISION FOR RAILWAY COMMUTERS

**Kota Miyauchi, Kazuyuki Takada**
*Tokyo Denki University, Japan*

## Abstract

In recent years, many studies of travel time reliability energetically are performed. However, these studies didn't focus on considering heterogeneity between individuals. In this study, the research of considering heterogeneity between individuals which focused on railway commuters in Tokyo metropolitan was performed. First, it was performed online survey and collected data about the commuter situation in the railway. As a result of the online survey, the railway commuters recognize train travel time as uncertain and tend to decide the departure time considering buffer time. As a result, the arrival situation of railway commuters became variety, it is difficult to arrive on their ideal arrival time. To define arrival situation, the arrival distribution was constructed to their arrival situation. This distribution consists of three types of considering the delay recognition. In this study, the arrival distribution for railway commuters assumed that mixture distribution. The divided ratio of mixture distribution is different every railway commuter. Next, using this result, the factors of buffer time of railway commuters estimated by MCMC method. MCMC method could estimate the heterogeneity between individuals. As a result of estimation, the factor of increase buffer time was travel distance(km), number of transfers(times), and the experience time encountering train delay (min/week). The application of mixture distribution model was verified by this study.

*Keywords: travel time reliability, mixture distribution, optimal departure time*

## 1 Introduction

In recent years, the number of delayed trains tend to increase in Tokyo metropolitan area. As a result, their recognitions against to railway changes by increasing of delay trains. It means that the railway commuters recognize train travel time as uncertain. Thus, without the guarantee of the punctuality, railway commuter must set the buffer time to arrive at their ideal arrival time. The influence of lower travel time reliability of railway service has become an issue in Tokyo metropolitan area.
It was cleared that the railway commuters hope for the accurate operation or shortening stop time of operating than shortening travel time or construct the new route [1]. Therefore, it means that the railway commuters would hope to improve the travel time reliability. There are many studies investigating travel time reliability from various aspects now. In this study, it focused on the buffer time which considering the delay for railway commuters in Tokyo metropolitan. This research purpose is to construct the optimal departure time decision model. This model can estimate the buffer time for railway commuters. In addition to, it performed to assume that the buffer time obeys mixture distribution model.

## 2 Previous studies

### 2.1 Analysis applying mixture distribution model

This section explains about previous studies applying the mixture distribution model. It is energetically performed by fields which the pattern recognition using big-data. For example, it was used to analyse the traffic accidents data. This method is meant to explain the difference the ratio points where to occur the traffic accidents. This estimation method explains the difference region where occur the traffic accident by applying the mixture Poisson model. On the other hand, there are also the studies which analyse the damage diagnosis of construction by applying the mixture Gaussian distribution model.

### 2.2 Estimation method for travel time reliability

There are many previous studies which faced on the travel time reliability, it was performed various aspects. There are two types of approach to develop the behave model under travel time uncertainty. One is "Mean-variance approach" and the other is "Scheduling approach." The former is to formulate the optimal departure time under uncertainty of travel time by applying probability distribution of the deviation of travel time. On the other hand, the latter is to consider the decision of departure times to minimize inconvenience due to early/late arrival. Additionally, Fosgerau and Karlstrom [2, 3] developed a new approach method that combines these two approaches. In their integrated model, utility function of the scheduling approach type was defined to explain the optimal behavior. In this study, the estimation was used to this model.

## 3 Analysis data and construct the arrival distribution

### 3.1 Analysis data and sampling attribute

This section explains the summary of the online survey. In this study, it used the research monitors of Macromill Inc. Table 1 shows outline survey. The targeted respondent was employed the person who is living in Tokyo metropolitan area (Tokyo, Kanagawa, Chiba, or Saitama) and who is using the train for commuting at least five days in a week. The survey was conducted on February 23 and 24, 2015, and the number of responses was 1000.

**Table 1** Outline of the internet survey

| The mode of survey | Internet survey |
|---|---|
| Questionnaire | Residents in Tokyo metropolitan area (Saitama, Chiba, Tokyo and Kanagawa) |
| | Jobholder including part time jobs (over 15 years old) |
| | Frequent railway user (at least 5 days a week) |
| Personal attributes | Gender, age, place of residence, etc. |
| Railway usage | Origin and destination stations |
| | Transfer stations |
| | Frequency of railway usage |
| | Estimated travel time |
| Experience of delay | Frequency of encountering train delay |

## 3.2 Construct the arrival distribution for railway commuters

In this survey, it was cleared that the railway commuters set the ideal arrival time to their destination. Table 2 shows the survey asked to the arrival status for railway commuters against to their ideal arrival time. In addition to, the arrival distribution for railway commuters was constructed by results of Table 2. Takada et al. proposed to use the mixture distribution for arrival distribution of railway commuters and cleared that it was better than using the single distribution [4].

However, the number of clusters which optimal applying mixture distribution don't know at this stage. So in this study, it uses Gap statistics to decide the number of clusters. Gap statistics was proposed by Tibshirani et al. It is utilized to determine the appropriate number of clusters [5]. In this study, the number of clusters was set from one to five. The cluster which obtains the maximum Gap statistic become the optimal number of it. The estimation for Gap statistic used the questionnaire results of Table 2. Figure 1 shows the result of Gap statistic for every cluster. Gap statistic became largest when the number of clusters is three. Therefore, the arrival distribution for railway commuters was divided to be three.

**Table 2** Survey contents (some extracts)

| Questions | Answer (example) |
|---|---|
| Q1. How many times can you arrive within 1 minute before and after the ideal arrival time (out of 100) | 10 |
| Q2. The situation of arriving earlier or later than the ideal time (out of 100) | (early/late) arrival times |
| · How many times do you arrive early/late by 1 minute? | 35/5 |
| · How many times do you arrive early/late by 2 minutes? | 15/5 |
| · How many times do you arrive early/late by 3 minutes? | 10/0 |
| · How many times do you arrive early/late by 4 minutes? | 5/0 |
| · How many times do you arrive early/late by 5 minutes? | 5/0 |
| · How many times do you arrive early/late by 10 minutes? | 5/0 |
| · How many times do you arrive early/late by 20 minutes? | 3/0 |
| · How many times do you arrive early/late by 30 minutes? | 2/0 |

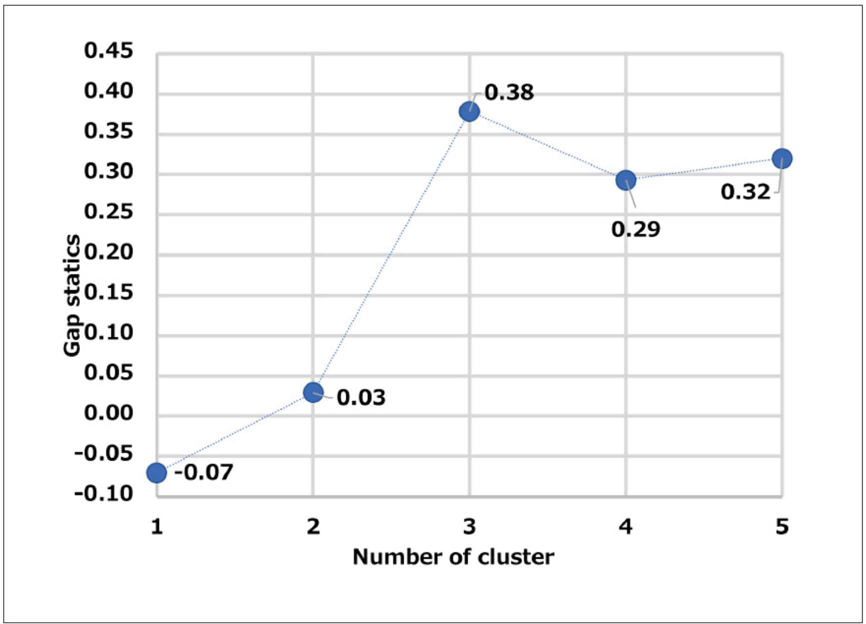


**Figure 1** Gap statics by each cluster

### 3.3 The decision to divided ratio for the mixture distribution model

The feature which was divided into three distribution on the basis of Gap statistic shows below. The railway commuters tend to set the ideal arrival time at the destination. As a result, the early arrival, the ideal arrival, and the late arrival exist against to their ideal arrival time. In this study, it defined these state of arrival as “early arrival recognition”, “ideal arrival recognition”, and “delay arrival recognition”. The arrival distribution consists of combined these arrival recognitions. Moreover, the parameter of each distribution shows below:

- Early arrival recognition · · · $N(\mu_1,\sigma_1)$ s.t.$\mu_1 < 0$
- Ideal arrival recognition · · · $N(0,\sigma_2)$
- Delay arrival recognition · · · $E(\lambda)$

Eqn (1) shows that is the probability distribution of mixture distribution model. This model was constructed from two normal distribution and exponential distribution.

$$p\left(x\middle|\pi,\mu,\sigma^2,\lambda\right)=\sum_{k=1}^{K}\pi_k N\left(x\middle|\mu_k,\sigma^2{}_k\right)+\pi_E f\left(x\middle|\lambda\right) \quad (1)$$

Where $\pi_k$ is the divided ratio of normal distribution and $\pi_E$ is the divided ration of exponential distribution. Figure 2 shows some example applying the mixture distribution model. Figure 3 shows the divided ratio of mixture distribution model by all respondents. It cleared that the divide ratio by railway commuters arises the individual variation.

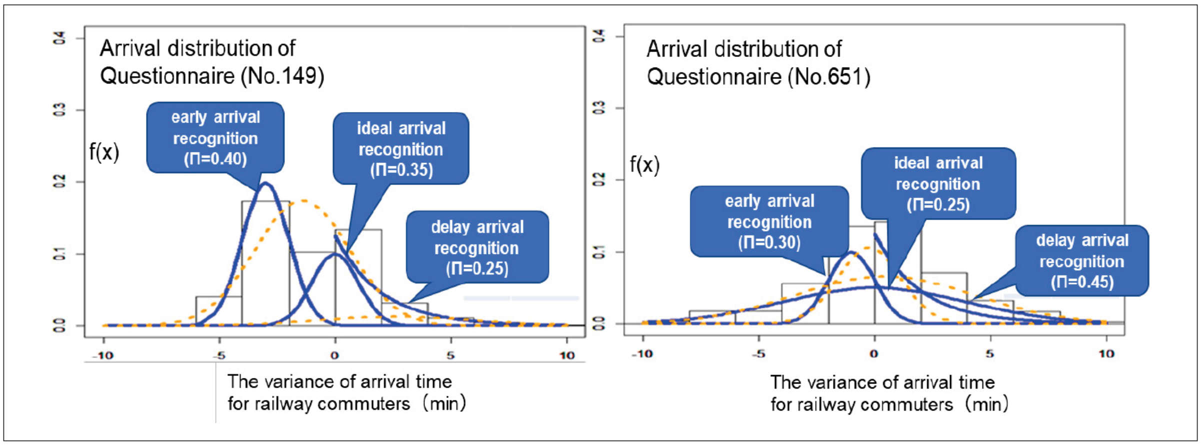


**Figure 2** Construct the arrival distribution applying mix. distribution model for railway commuters (example)

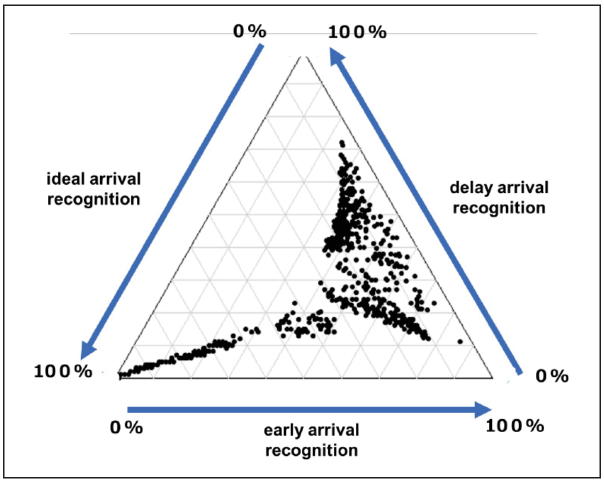


**Figure 3** The divided ratio of mixture distribution model by all respondents (n = 1000)

## 4 Construct the optimal departure time decision model for railway commuters

### 4.1 About the optimal departure time decision model

In this section, it assumes that to decide the buffer time for railway commuters obey considering the arrival distribution. Thus it assumed that also the buffer time obeys mixture distribution model. To estimate the buffer time, it performed that to construct the optimal departure time decision model and estimated by MCMC method. This model is used to estimate the benefit-cost approach for lost-time which the railway commuters was damaged it.
However, it is difficult to estimate the buffer time using only their arrival status. the variable of model don' t enough to explain the factors. In other words, there is the heterogeneity between individual. So it used the Markov chain Monte Carlo (MCMC method) method to consider it.

### 4.2 Construct the optimal departure time decision model applying mix. distribution model

In this section, explains about the optimal departure time decision model. In this study, it proposes the model applying the mixture distribution model. This model improved the proposed it by Fosgerau and Takada et al. Eqn (2) (3) shows original model which estimate the optimal departure time decision model by Fosgerau and Takada et al.

$$D^* - u = \ln\left(\frac{\alpha}{\beta}\right) / \lambda \exp\left(\sum_{i=1}^{n} \theta_i x_{ij}\right) \tag{2}$$

$$D^* - u = \left(\sum_{i=1}^{n} \theta_i x_{ij} + \varepsilon_i\right) \varphi^{-1}\left(1 - \frac{\alpha}{\beta}\right) \tag{3}$$

Where:
- $T$ – travel time;
- $D$ – departure time;
- $D^*$ – optimal departure time;
- $\mu$ – travel time for timedable;
- $\alpha, \beta$ – unknown parameters;
- $\lambda, \varepsilon$ – scale parameters (represent the individual variance);
- $\theta_i$ – parameter of variable;
- $x_i$ – explain variable;
- $\Phi$ – cummrative distribution of N (0 , 1).

Chapter 3 assumed that the buffer time of railway commuters consisted of three distribution which two normal distribution and exponential distribution. So eqn (4) integrated by eqn (2) and eqn (3). The divided ratio uses the results in Figure 2.

$$D^* - u = \left\{\pi_{Nk}\left(\left(\sum_{i=1}^{n} \theta_i x_{ij} + \varepsilon_i\right) \varphi^{-1}\left(1 - \frac{\alpha}{\beta}\right)\right) + \pi_{Ek}\left(-\ln\left(\frac{\alpha}{\beta}\right) / \lambda \exp\left(\sum_{i=1}^{n} \theta_i x_{ij}\right)\right)\right\} + \omega_i \tag{4}$$

Where:
- $\pi_i$ – divided ratio;
- $\theta_{N1k/Ek}$ – parameter of variable.

## 4.3 Estimation result and consideration

Table 4 shows an estimation result. If the values of parameter more significant, the railway commuters tend to set increase the buffer time. The most utility variable became the travel distance (km), the number of transfers (times), and the experience time encountering train delay (min/week). It was cleared that the factor which increases buffer time is three factors. t value of each parameter is high, so it is said that this estimation result is effectiveness. Figure 4 shows scale parameter for all respondent. Figure 5 shows the reproducibility of buffer time. The results of the reproducibility are quite high. So the application of mixture distribution model was verified by this estimation. The optimal delay probability obtained by $\alpha$ and $\beta$ estimation result is 36.2 %.

**Table 3** Estimation result for optimal departure time decision model

| Parameter / explanatory variables | Prior distribution | Construct the model in this study | |
|---|---|---|---|
| $\alpha$ / Early arrival parameter | U (0, $\beta$) | 23.54** | |
| $\beta$ / Late arrival parameter | U (0, 100) | 64.92** | |
| $\theta_{11}\theta_{12}\theta_{13}$ / travel distance (km) | N (0, 1000) | Ideal arrival recognition…<br>Early arrival recognition…<br>Late arrival recognition… | 5.29<br>6.85**<br>3.59* |
| $\theta_{21}\theta_{22}\theta_{23}$ / number of transfers (times) | N (0, 1000) | Ideal arrival recognition…<br>Early arrival recognition…<br>Late arrival recognition… | -2.49<br>2.30**<br>1.85* |
| $\theta_{31}\theta_{32}\theta_{33}$ / Experience encountering train delay (min/week) | N (0, 1000) | Ideal arrival recognition…<br>Early arrival recognition…<br>Late arrival recognition… | 3.07*<br>2.01**<br>2.76* |
| Scale parameter | N (0, 100) | Figure 4 | |
| R2 (Decision parameter) | - | Figure 5 | |
| Number of sample | - | 1000 | |
| $\alpha$ / $\beta$ : Optimal late arrival probability (%) | - | 36.3 | |

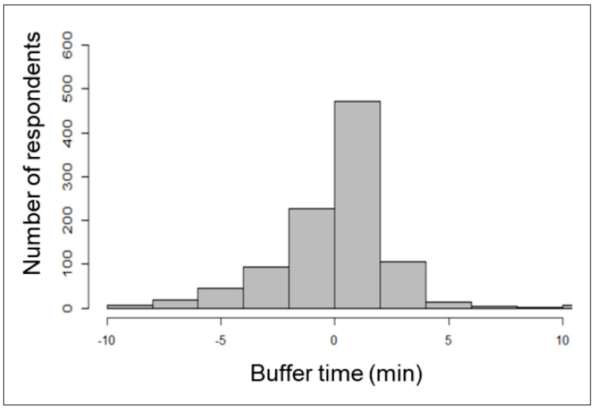


**Figure 4** Scale parameter of all respondent

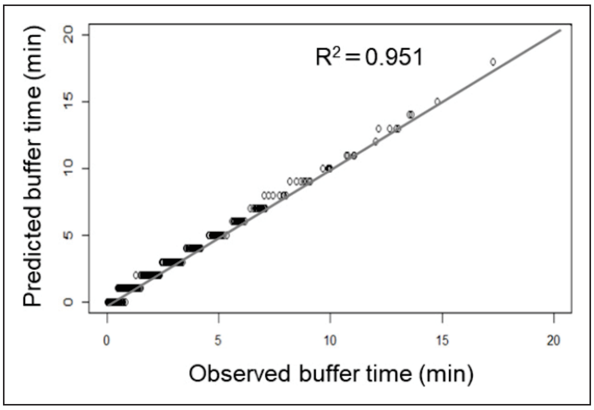


**Figure 5** Reproduction of the buffer time for mixture distribution

## 5 Conclusion

This study has analysed the evaluation of the travel time of railway commuters under the assumption that there exist multiple arrival distributions regarding arrival situation to the optimal arrival time. Assuming that the delay probability varies among individuals and that they obey multiple probability distributions, it formulated the expected utility for departure times and found the optimal departure time by solving the issue of maximizing the expected utility. As a result, the railway commuter decisions for their respective departure times were found to be in agreement with the decision problem of the buffer time. The regression model used to estimate the buffer time was based on the MCMC method. Under the assumption that this model obeys multiple probability distributions, the study demonstrated that it was possible to yield highly convincing estimations by taking into account the variance of railway commuters from estimation results. The study also indicated that the increase in travel distance, number of transfers, and experience encountering train delay correlates to the increase in the expected delay time.

## References


[1] Takada, K., Miyauchi, K., Fujiu, M.: Departure Time Decision of Railway Commuters considering Different Buffer Time Distributions, Journal of Eastern Asia Society for Transportation Studies, 2018

[2] Fosgerau, M., Karlstrom, A.: The value of reliability and the distribution of random durations, European Transport Conference, 2007

[3] Fosgerau, M., Karlstrom, A.: The value of reliability. Transportation Research Part B, Vol.44, No.1, 33-49, 2010

[4] Takada, K., Miyauchi, K., Takanami, Y., Fujiu, M.: Estimation of the arrival distribution of the railway commuters by applying mixture distributions, Journal of Japan Society of Traffic Engineers , Vol.3, No.2, 238-245, 2018

[5] Tibshirani, R., Walther, G., Hastie, T.: Estimating the number of clusters in a dataset via the gap statistic. Journal of Royal Statistical Society, 2001