

# Modeling of public transport waiting time indicator for the transport network of a large city

*Olga Lebedeva*<sup>1</sup>, and *Marina Kripak*<sup>2,\*</sup>

<sup>1</sup>Federal State Autonomous Educational Institution of Higher Education «Sevastopol State University», Russian Federation, 299053, Sevastopol

**Abstract.** The need to develop and improve public passenger transport in major cities was noted. It was reflected that waiting time at bus stops is one of the factors that have a big impact on the passenger quality assessment of transport services. The results of an empirical study of the actual and anticipated waiting time at bus stops were given. It was noted that the reliability functions were used in the field of ride duration modeling, traffic restoration time after an accident, and length of making the decision to travel. The waiting time distribution functions using the lognormal function and the Weibull function were chosen. The results of modeling were objective, the dependent variables in it were the expected waiting time of passengers and the difference between the anticipated and the actual waiting time. The explanatory variables were sex, age, time period, purpose of the trip and the actual waiting time. The results of the research showed that the age, purpose of the trip and the time period influence the waiting time perception, prolong it and lead to its reassessment.

## 1 Introduction

Public municipal passenger transport is an important element of the transport system, which allows passengers to receive transport services at an affordable price. Despite its advantages, the share of transport services in the market is significantly lower than that of personal transport, which is explained by the low quality of the services provided.

Bus transportation is the most popular in the world's transport systems, but it has a much lesser attractiveness for passengers and has a negative impact on passenger traffic, which is mainly related to performance characteristics such as low speed, low frequency and non-compliance with the traffic schedule [1, 2, 3].

The waiting time at bus stops is the factor that has a great influence on the passenger quality assessment of transport services. The passengers waiting for a bus can be exposed to unfavorable weather conditions, crowding and emissions of harmful substances, which will increase the feeling of anxiety while waiting. As a result, the perception will be biased, the waiting time will seem longer compared to the actual waiting time, especially when

---

\* Corresponding author: [marikol@yandex.ru](mailto:marikol@yandex.ru)

passengers are not given any information about the arrival time of the means of transport [1, 2, 3].

Let's present the results of an empirical study of the actual and the anticipated waiting time at bus stops (the survey was conducted on a large bus network using a duration model). The analysis is based on a survey conducted at more than 30 bus stops of the transport network in Sevastopol, which consists of more than 300 carriers and handles 1.3 million passengers on a daily basis.

The issue of waiting time has been considered in the research works for several decades, with both actual and anticipated waiting times taken into account.

In this article, we will trace the relationship between the actual and the anticipated waiting time based on data, observations and surveys and use the appropriate statistical methods to analyze the influence of various factors on the passenger's perception of the waiting time. The data are most often based on a field survey, which combines observations of the actual waiting time by passengers at bus stops and a survey of their perception of this time. The measurements are made with random passengers arriving at bus stops. The time of their arrival is recorded by the surveyors. Sometime after their arrival, the surveyors conduct a survey of passengers about their perception of their waiting time at the bus stop, and also specify other information. The interview beginning time is indicated in the questionnaire. Then, when the expected bus arrives, and the passengers get on it, the actual waiting time is recorded. Thus, we get the actual waiting time, the anticipated waiting time, and the total actual waiting time. As a survey location, bus stops with various intervals of bus arrival times are chosen.

The received data include: the actual arrival time of passengers at the bus stop, the time of survey, the total actual waiting time, the anticipated waiting time, their gender, age and purpose of the trip, time period (morning, afternoon, evening).

The survey was conducted at 30 bus stops in the area of Sevastopol, from 8.00 am to 20.00 pm. As a result of the experiment, a total of more than 1000 passengers were interviewed.

A preliminary statistical analysis of the results showed a relative balance between men and women (47% vs. 53%), while the majority of passengers were adults between 18 and 65 years of age. Moreover, more than 35 percent of the respondents were travelling to/from their place of work. Tables 1 and 2 summarize the actual and the anticipated average waiting times for different time periods, gender and age groups, and trip purposes.

After examining the results of Tables 1 and 2, we see that the anticipated waiting time in most cases increased by at least 50 percent compared to the actual waiting time.

In addition, older age groups of passengers travelling to work or for personal matters tend to overestimate the waiting times compared to other groups. The same applies to the passengers interviewed in the morning. Preliminary interpretation of this overestimation by specific age groups can be explained by such factors as limited patience of elderly passengers and their fear of being late for work. However, a detailed statistical analysis supported by the model will reveal the actual consequences of the above factors for the actual and anticipated waiting times.

**Table 1.** Actual and anticipated average waiting time for different time periods

Time period	Average waiting time (min)	Average actual waiting time (min)	Ratio of average anticipated waiting time and average actual waiting time
Morning	5.2	4.2	1.2
Afternoon	6.0	4.4	1.4
Evening	5.4	3.6	1.5

**Table 2.** Actual and anticipated average waiting time for different groups according to gender, age and purpose of the trip

	<b>Average waiting time (min)</b>	<b>Average actual waiting time (min)</b>	<b>Ratio of average anticipated waiting time and average actual waiting time</b>
<b>Age group</b>			
<18	8.0	6.3	1.3
18-30	6.1	5.2	1.2
31-45	6.2	3.9	1.3
45-65	7.5	4.3	1.7
>65	7.0	4.1	1.7
<b>Gender</b>			
Male	6.1	3.9	1.2
Female	6.4	4.2	1.5
<b>Purpose of the trip</b>			
Home	7.0	5.0	1.4
Work	7.2	4.8	1.5
Education	8.9	7.7	1.2
Personal matters	7.9	6.0	1.3
Entertainment	7.1	5.1	1.4
Shopping	8.0	5.9	1.4
Other	10.1	4.7	2.1

The data on trip duration is information about the duration before/between the events [4]. Such data, for example, include the transport waiting time, and the duration between transport events. To model such situations, the following indicators are taken into account: the time before the occurrence of an event (or the time period between successive events); a reliability function that expresses the probability of an episode starting at time  $t$  and finishing within the time interval  $(t, t + \Delta t)$ , provided that the event did not occur before the start of the interval. The waiting time is the period (episode) until the passenger gets on the bus. As a result, the probability of getting on a bus after waiting for the duration  $\Delta t$  seems to be a reliability function.

Reliability functions are used in the field of modeling the duration of a trip, the time of restoration of traffic after an accident, and the duration of making a decision on the need for a trip [4].

## 2 Theoretical basis

Let us assume that  $T$  is a non-negative random variable that represents the anticipated waiting time (duration) and the difference between the perceived and the actual waiting time. The distribution of probabilities  $T$  can be represented in several ways, of which the reliability and hazard functions are the most suitable to our task. The reliability function is defined as the probability of the interval  $T$  being at least  $t$  (that is, the anticipated waiting time or the difference between the perceived and the actual waiting time is not less than  $t$  min) [4]:

$$F(t) = P(T \geq t), 0 < t < \infty \tag{1}$$

The function used here assumes that  $F(t)$  is a monotonic continuous function with  $F(0) = 1$  and  $\lim_{t \rightarrow \infty} F(t) = 0$ . The probability density function is represented in the following form:

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t} = \frac{-df(t)}{dt} \quad (2)$$

The reliability function determines the instantaneous failure rate when  $T = t$ , conditional after the occurred event before time  $t$  and can be defined as follows [4]:

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{F(t)} \quad (3)$$

It is important to note that the reliability function is extremely useful in practice. It indicates the speed at which the perceived waiting time increases after the idle time  $t$ . Also, the equation (3) shows that  $l(t)$  determines the distribution  $T$ , since,

$$l(t) = \frac{-d \log F(t)}{dt}$$

There are many functions for the distribution of duration, such as exponential, Weibull, lognormal, and others [5]. The choice of the distribution function is important for the practical significance of the results.

Dependent variables are the anticipated waiting time of passengers and the difference between the anticipated and the actual waiting time. Explanatory variables are sex, age, time period, purpose of the trip and the actual waiting time. They are presented in Table 3.

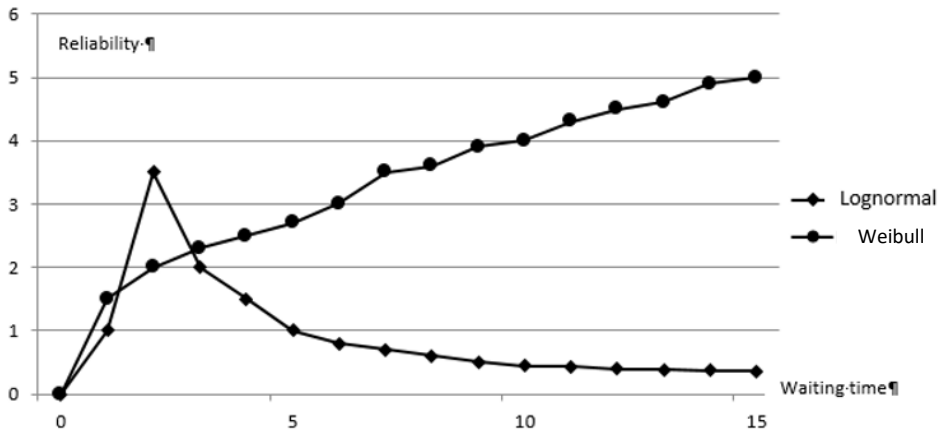
**Table 3.** Dependent and explanatory variables

Name of variable	Description	Value
<b>Dependent variables</b>		
LOGPER	Anticipated waiting time	Any
LOGDIF	Difference between the anticipated and the actual waiting time	Any
<b>Explanatory variables</b>		
Gender	Gender (male, female)	0 for men, 1 for women
AGE18	Gradations by age groups	1 for ages <18, 0 in other cases
AGE1830		1 for ages 18-30, 0 in other cases
AGE3045		1 for ages 30-45, 0 in other cases
AGE4565		1 for ages 45-65, 0 in other cases
AGE65		1 for ages >65, 0 in other cases
<b>Purpose of the trip</b>		
HOME	Home	1 for «returning home», 0 in other cases
WORK	Work	1 for «work», 0 in other cases
EDUC	Education	1 for «education», 0 in other cases
PERS	Personal matters	1 for «personal matters», 0 in other cases
ENTERT	Entertainment	1 for «entertainment», 0 in other cases
SHOP	Shopping	1 for «shopping», 0 in other cases

Name of variable	Description	Value
		cases
TRAVEL	Travelling to the country	1 for «travelling to the country», 0 in other cases
OTHER	Other	1 for «other», 0 in other cases
Time period		
QUE1	Time period 1	from 8:00 to 12:00, 0 in other cases
QUE2	Time period 2	from 12:00 to 17:00, 0 in other cases
QUE3	Time period 3	from 17:00 to 20:00, 0 in other cases
REALTIME	Real time	Actual waiting time before the beginning of survey

In particular, five variables in Table 3 define age groups: under 18 years old, from 18-30 years old, from 31-45 years old, from 46-65 years old and over 65 years old. Eight variables determine the purpose of the trip (returning home, work, education, shopping, entertainment, travelling to the country and others). Three variables define the hour of the day (8:00-12:00, 12:00-17:00, 17:00-20:00).

As a result of the calculations carried out, it should be noted that the reliability functions based on the lognormal function and the Weibull function are given as  $l + \frac{p(lt)^{p-1}}{1}$  and  $(lt)^p, lp(lt)^{p-1}$  respectively; for the same distribution, the reliability function is  $-\frac{l}{t} + (lt)^p, e^{-(lt)^p}$ .



**Fig. 1.** Waiting time distribution using the lognormal and the Weibull functions

As can be seen from Fig. 1, the Weibull distribution increases monotonically, indicating an ever-increasing degree of reliability over time, whereas the lognormal distribution initially increases, and then the reliability level decreases [6].

In order for the model to match the data, the graph for each of the functions must pass through the origin of coordinates. The values obtained show that the Weibull distribution is more suitable for describing both the waiting time and the difference between the anticipated and the actual waiting time.

According to the interpretation of results (for the Weibull distribution), the waiting time in the age group over 18 years old is longer than in other groups. And in the age group from 45 to 65 years old, the anticipated waiting time has increased by 10%.

The waiting time is also influenced by the purpose of the trip. Travelling to work, for example, leads to an increase in the anticipated waiting time by 30%, education and personal matters – by 45% and 20%, respectively, visiting shopping centers – by 30%. Travelling back to the place of residence and entertainment have a lower impact on the anticipated travelling time.

The waiting time in the morning hours decreases by approximately 5%, which is probably explained by the shorter traffic intervals. The results (for Weibull distribution) show that passengers (up to 30 years of age) tend to perceive the waiting time more easily, for these categories the difference between the anticipated and the actual waiting time is 4% -5%.

### 3 Conclusion

Obviously, the waiting time at a stop is one of the key factors affecting the attractiveness of public municipal passenger transport. The research results showed that age, purpose of the trip and time period influence the waiting time perception, prolong it and lead to its reassessment. These factors especially manifest themselves in the population of the older age group, when traveling to work or for education purposes.

The situation can be changed with the implementation of the Intelligent Transport System (ITS). In accordance with the recommendations, ITS includes a non-cash payment of fares in the urban public transport and its dispatching control. Such complex will improve the quality of passenger transportation services.

The non-cash payment will increase the efficiency of payment control, will ensure the introduction of a flexible system of tariffs and personalized accounting for the provision of benefits, it will increase the collection of fares, improve the working conditions of the passenger transport system employees, and provide the possibility of introducing a non-conductor payment scheme. The dispatching control will ensure the quality of traffic management for vehicles carrying out passenger traffic in the city, the quality of scheduling and route analysis, automation of the means of transport register, data analysis on the speed of these means of transport, automation of the public transport arrival forecast and other improvements.

Authors are grateful for financial support Russian Foundation for Basic Research. The scientific research was supported by the Russian Foundation for Basic Research within the framework of the project № 18-410-920016 p\_a "Research of socio-economic and ecological processes of Sevastopol with the growth of industrial, traffic, transit and tourist potentials".

### References

1. Psarros, I., Kepaptsoglou, K., & Karlaftis, M. . *An empirical investigation of passenger waiting time perceptions using hazard-based duration models*. Journal of Public Transportation, **14(3)**, 109–122 (2011)
2. Lebedeva O.A. *Estimation of the degree of unevenness of passenger flows*, Modern technologies and scientific and technical progress. T. 1. № 1. P. 41 (2014)
3. Lebedeva O.A. *Perfection of methods for monitoring passenger traffic on public urban passenger transport routes: diss. cand. of tech. sciences*. Irkutsk, 2014. 171 p.
4. Hensher, D. A., F. L. Mannering, *Hazard-based duration models and their application to transport analysis*. Transport Reviews **14(1)**, 63-82 (1994)

5. Washington, S., M. Karlaftis, F. Mannering, *Statistical and Econometric Methods for Transportation Data Analysis* (Chapman & Hall/CRC Press, Boca Raton, FL, USA, 2003)
6. Lee, E. T., and J. W. Wang, *Statistical Methods for Survival Data Analysis 3rd Edition* (Wiley and Sons, Hoboken, NJ, 2003)