The modeling of effective parameters on public bus passengers' boarding time prediction

Metin Mutlu Aydın

Civil Engineering Department, Ondokuz Mays University, Samsun, Turkey
Corresponding Author: metinmutluaydin@gmail.com

ABSTRACT

Passengers' boarding times at bus stops have a great importance to calculate dwell and travel time for scheduling process in public transport operations. However, there are not so much observed boarding times data in the actual bus transport systems, and it may cause some prediction problems in scheduling process of public transport operations. For this reason, accurate estimation of the boarding times will ensure correct calculation of dwell and total travel time for bus transport systems. Based on this idea, this study aims to model boarding times of each passengers by evaluating different parameters using two different (statistical and optimization analysis) methods. For this purpose, a comprehensive data collection process was conducted in total seven different cities of Turkey based upon their population. Two new models were developed for boarding time estimation by evaluating various parameters using a multiple Ordinary Least Square (OLS) regression and Artificial Bee Colony (ABC) algorithm as statistical and optimization methods, respectively. Study results showed that modeling of boarding times by considering various parameters is an effective strategy to improve the performance of bus transport systems by using developed two models.

Keywords: Boarding time prediction; Statistical analysis; Passenger characteristics; Metaheuristic optimization; Public transport.

1. INTRODUCTION

Bus transport is one of the famous and primary public transport systems in many countries because of its many advantages such as easy accessibility and sustainability. Efficiency of public bus transport systems has a great importance for planners and researchers to make it more attractive. These systems are still the largest and most important parts of the public transport systems in many countries (especially in developing countries). For all these reasons, transportation planners, operators, and decision makers conduct numerous studies to improve their service quality, efficiency, attractiveness, and sustainability. Especially, planners and operators give a great importance to increase public demand and utilization of these systems (Ghasemlou *et al.,* 2012; Tirachini, 2013; Fletcher and El-Geneidy, 2013; Aydin *et al.,* 2016; El-Geneidy *et al*., 2017; Stevanovic, 2018). For the planning process, many modeling and optimization studies were conducted to improve the performance and efficiency of these systems by considering passengers' perceptions and comments on service performance. Additionally, many other studies were conducted to understand passengers' travel decisions to develop new passenger friendly strategies.

According to conducted studies, bus dwell well time was found to be one of the important and effective parameters on the total travel time of a public bus (El-Geneidy *et al*., 2017). Several studies in different countries have been conducted to identify the effect of dwell time on the operational performances of public bus transport systems and traffic flows since 1970s (e.g., Kraft and Bergen, 1974; Levinson, 1983; Guenthner and Hamat, 1988; Levine and Torng, 1994; Genivar, 2011; Chen *et al*., 2013; El-Geneidy *et al*., 2017). All these studies have found that boarding and alighting times of passengers to a bus have a great effect on dwell times of bus transport systems (Dueker *et al*., 2004; El-Geneidy and Vijayakumar, 2011). According to Levinson (1983)'s study, dwell time usually becomes 9–26% of the total travel time for a public bus. He clearly emphasized that dwell time has great importance on travel times of bus transport systems.

It was obtained from the previous study results that there are many effective parameters on dwell time such as passenger activities (boarding-alighting time) and characteristics (normal, elder, with child, handicapped passengers, etc.), bus fare collection types, bus door and floor properties, and bus bay existence. In literature, many studies were carried out to determine all these effective factors on dwell time calculation. For example, Fernandez et al. (2010) investigated the effects of bus geometric properties (bus door width and floor height) on alighting time, and they found that wider bus doors can reduce the average alighting time by almost 40%. It was also found that lower floor height only reduces the average alighting time from 1 to 9%. In a different study, Fletcher and El-Geneidy (2013) focused on the effect of alternative fare payment methods on dwell time. They found that on-board sale and ticket validation can cause an increasing on boarding times. In another study, Tirachini (2013) studied on ages of passengers and crowding and/or friction effects among passengers' boarding time. He determined that when passengers' boarding forms two queues through a single door, this can result in long boarding times. In a similar conducted study, Currie et al. (2013) calculated that crowding had a more significant effect on boarding time than the presence of entrance steps. Larwin (2012) also found that when crowding is reduced at the front doors of the bus, boarding time of a passenger is significantly reduced, by up to 0.5 seconds. Dueker *et al*. (2004) calculated that low-floor buses move 0.13 seconds faster per stop than buses with steps. It was also found from the study that the bicycle and wheelchair loadings have a great positive effect on boarding time. In a different study, Fernandez *et al*. (2010) calculated that boarding and alighting times become slower for trunk bus services when there are more than 40 passengers boarding and less than 15 passengers alighting. In the study, Dueker *et al*. (2004) used squared terms on boarding and alighting times. They calculated that each additional boarding passenger takes 0.04 seconds less, with the same figure being 0.03 seconds for each additional passenger alighting. According to the experimental study of Fernandez (2011), there are negative and linear relationships between crowding and boarding time.

It can be concluded from the literature that the numbers and mean boarding and/or alighting times can be explained as the most important and effective factors on bus dwell time and total travel time. Boarding and/or alighting times include many different and important pieces of information in planning and operating process for public bus transport systems. For this reason, more passenger based studies need to be conducted in this field to estimate these times precisely. Previous studies generally focused on the determination of total boarding and alighting times at a bus stop for each boarding or alighting to estimate dwell time and total travel time. Thus, literature review clearly shows that there is not any study that investigates and models the boarding time of each passenger by considering various effective factors, whereas there are many different factors on boarding behaviors of each passenger. Based on this idea, this study aims to model the boarding times of each passengers by evaluating different parameters as mentioned in next section for boarding behaviors of each passengers by using statistical analysis and optimization methods.

2. DATA COLLECTION AND DESCRIPTIVE STATISTICAL ANALYSIS

In the scope of the study, a comprehensive data collection process was conducted to relate different parameters on boarding times (T_B) of the passengers. In the study, total seven different cities of Turkey (İstanbul, Ankara, İzmir,

Antalya, Samsun, Trabzon, and Gümüşhane) were chosen to analyze and model boarding times (Figure 1). To reflect the general bus transport structure of these cities, they were selected as the pilot cities based upon their population (Doxiadis, 1968). İstanbul (the most crowded among the examined cities), Ankara, İzmir, and Antalya were chosen as high density cities (more than one million), Samsun and Trabzon as mid density cities (less than one million), and Gümüşhane (the least crowded among the examined cities) as low density city (less than one hundred thousand).
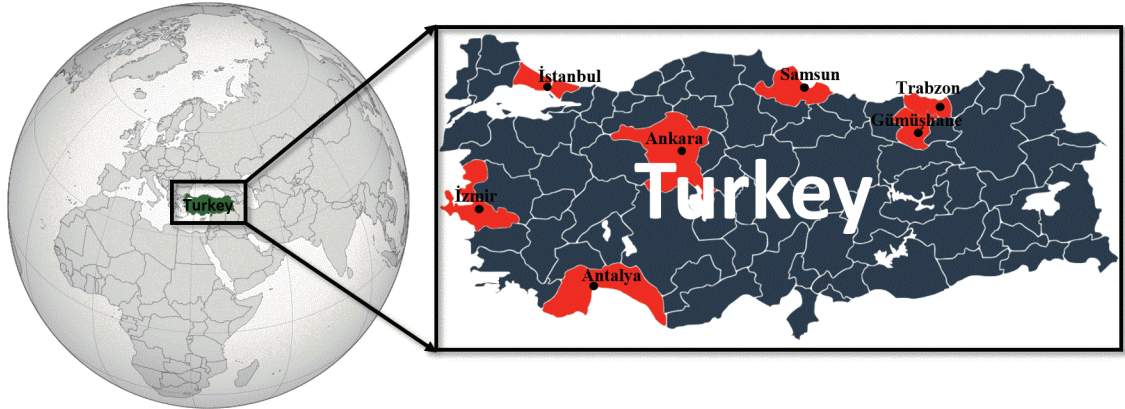
Figure 1. Locations of examined seven different cities in Turkey.

All data were collected by using site observations and video recordings from the different bus stops. To reflect the general bus transport structure and boarding behaviors, the bus stops were determined randomly from the different locations of the cities. The video recordings and site observations were made between 07:00–08:30 hours and 17:00–18:30 (peak) hours during the weekdays. In addition, the data collection process was repeated in different times in winter and summer seasons. Site observations and video camera recordings were not made from so close distance to the bus stops. Hence, passengers could not notice the observations and recordings and show normal boarding behaviors (Figure 2).
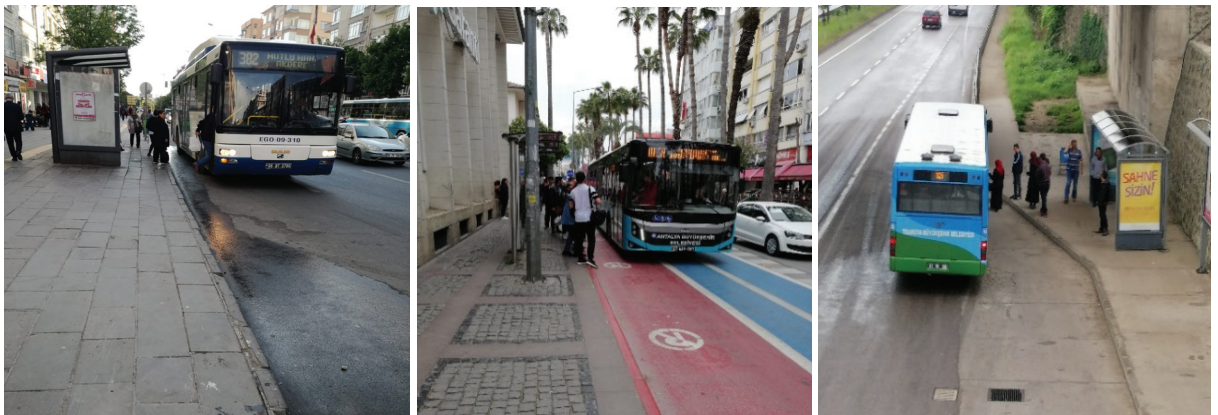
Figure 2. Some example visuals from data collection process at bus stops.

Analysis data were obtained by the application of extraction process from the video recording using digital counters. After counting process, the following data were obtained from the detailed analysis of site observations and recordings:

- Bus type (midibus, single bus, or articulated bus),
- Existence of steps to board and alight and number of steps (if available),
- Fee payment type (all examined cities use smart card as fare collection method),
- Passenger density of the buses (empty, low, mid, and high levels),
- Passenger property (normal, elder, with child, and handicapped),
- Existence of a bus bay (yes/no),
- Bus bay utilization of the bus drivers,
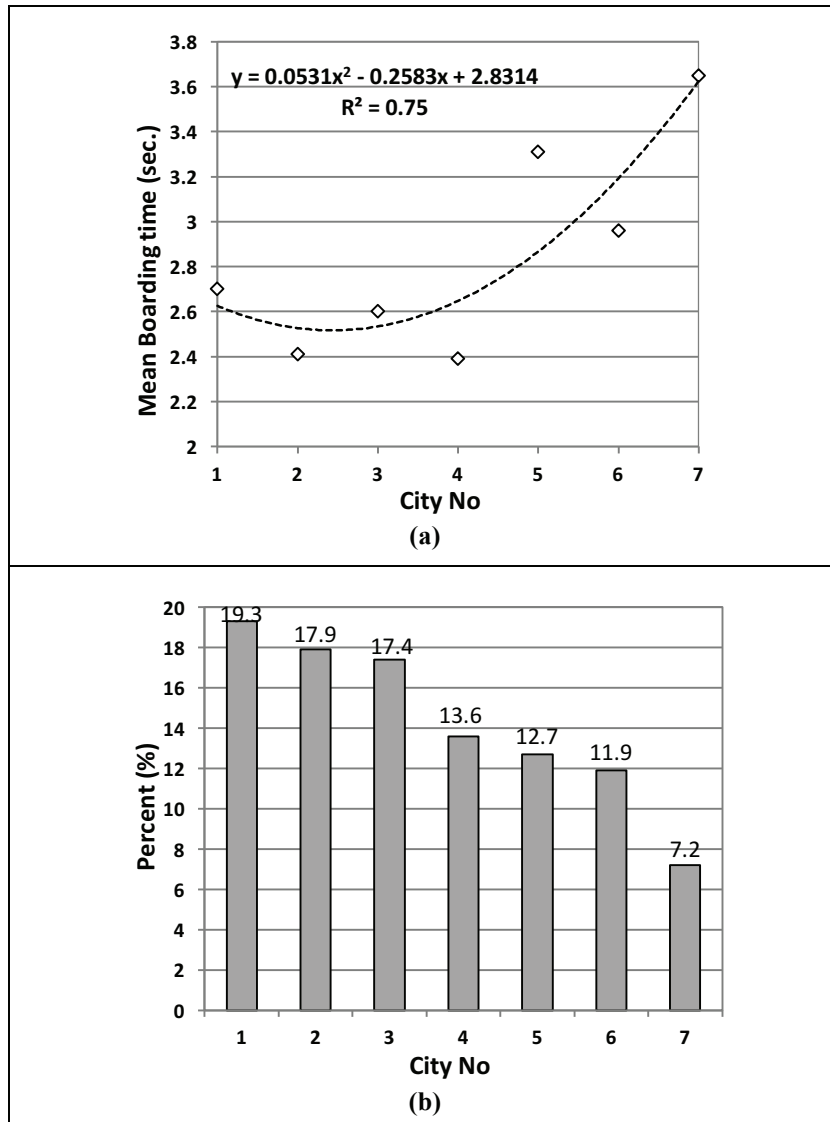- Boarding times (sec.).

From the counts, the random sampling group of the study was calculated as 5030 passengers. Descriptive statistics (means, std. dev., and min. and max. values) of numerical variables are summarized in Table 1.

Table 1. Descriptive statistics of observed data.

| City No | Examined Cities | Passenger Numbers (n) | Boarding Time (sec.) | | | | Rank Number (RN) |
			Minimum (T_{Bmin})	Maximum (T_{Bmax})	Mean ($T_{B\mu}$)	Std. Deviation ($T_{B\sigma}$)	
1	İstanbul	970	1.05	18.66	2.70	1.92	4
2	Ankara	901	1.06	16.88	2.41	2.21	2
3	İzmir	873	1.04	17.10	2.60	1.87	3
4	Antalya	685	0.98	16.40	2.39	1.75	1
5	Samsun	639	1.19	17.30	3.31	2.05	6
6	Trabzon	601	1.08	18.01	2.96	1.91	5
7	Gümüşhane	361	1.41	17.46	3.65	2.16	7

RN 1 shows the best; *RN* 7 shows the worst boarding times among examined cities.

As can be seen in the table, max. and min. mean boarding times ($T_{B\mu}$) are observed in Gümüşhane (the least crowded) city ($T_{B\mu} = 3.65$) and Antalya city ($T_{B\mu} = 2.39$), respectively. It can also be seen from the table that boarding time differences have the highest value in Ankara ($T_{B\sigma} = 2.21$) and the lowest value in Antalya ($T_{B\sigma} = 1.75$). Antalya city has lowest mean boarding time ($T_{B\mu}$) and std. deviation ($T_{B\sigma}$) for boarding time values. It means that Antalya city has better boarding times than the other six cities, which are located in different regions of Turkey. On the other hand, it is clear from the descriptive statistics that there are not big differences between min. and max. boarding times of the examined cities. In Table 1, seven different cities were also ranked according to their mean boarding times ($T_{B\mu}$) from lowest to highest level, and it was found that Antalya city has the best and Gümüşhane city has the worst boarding times. The relation among the mean boarding times ($T_{B\mu}$) is shown in Figure 3(a), and also, the percentages of collected data for 7 cities are given in Figure 3(b).

Figure 3. The relation among (a) mean boarding times ($T_{B\mu}$) and
(b) data percentages (%) of collected data for seven cities.

As seen in Figure 3(a), there is good relation among the mean boarding times ($T_{B\mu}$) of cities, and there are not high and important differences among the mean boarding times. It can be summarized as mean boarding time values are closer to each other, and this result makes a common (standard time) mean boarding time utilization idea possible in dwell time and travel time analysis for public bus transport system in all over the country. To gain a better understanding of mean boarding times ($T_{B\mu}$) differences, the obtained data are categorized and summarized, given in Table 2.

Table 2. Percentage description of obtained data for boarding times of the passengers from examined seven cities.

Parameters	Properties	Observation Numbers from Research Cities							Total Sample Size (%)
		1-İstanbul	2-Ankara	3-İzmir	4-Antalya	5-Samsun	6-Trabzon	7-Gümüşhane	
Bus Type (B_T)	Midibus	0	0	0	0	252	0	260	512 (10.2)
	Single Bus	695	643	421	685	387	405	101	3337 (66.3)
	Articulated Bus	275	258	452	0	0	196	0	1181 (23.5)
Existence of Bus Step (E_{BS})	Yes	375	189	301	0	252	258	341	1716 (34.1)
	No	595	712	572	685	387	343	20	3314 (65.9)
Numbers of Steps at Boarding Door (N_S)	None	595	712	572	685	387	343	20	3314 (65.9)
	Two steps	375	189	301	0	0	258	0	1123 (22.3)
	Three steps	0	0	0	0	252	0	341	593 (11.8)
Passenger Density inside the Bus (P_D)	Empty	102	159	77	240	18	36	30	662 (13.2)
	Low	553	459	633	207	111	231	75	2269 (45.1)
	Mid	179	158	113	128	187	156	116	1037 (20.6)
	High	136	125	50	110	323	178	140	1062 (21.1)
Passenger Property at Boarding Door (P_P)	Normal	802	732	703	498	501	462	270	3968 (78.9)
	with a child	45	34	30	36	35	22	19	221 (4.4)
	Elder	114	123	129	143	93	110	63	775 (15.4)
	Handicapped	9	12	11	8	10	7	9	66 (1.3)
Existence of a Bus Bay (E_{BB})	Yes	271	351	315	277	173	161	159	1707 (33.9)
	No	699	550	558	408	466	440	202	3323 (66.1)
Utilization of Bus Bay by the Drivers (U_{BB})	Yes	518	468	422	329	323	290	191	2541 (50.5)
	No	452	433	451	356	316	311	170	2489 (49.5)
Total (Σ)		970	901	873	685	639	601	361	5030 (100)

3. RESEARCH METHOD AND ANALYSIS

3.1. Modeling of Boarding Times Using OLS Regression

Ordinary Least Squares (OLS) regression is one of the most common prediction and modeling techniques in different research areas. The technique is most commonly used to examine and evaluate the relationship between a response variable and more than one explanatory variable (Mert, 2016; Aydin et al., 2019). It can be used in single and multiple explanatory and categorical variables (Aydin, 2020). In summary, OLS regression method defines the relationship between dependent and independent variables as given in

$$\hat{Y} = \beta_0 + \sum_{i=1}^{n} \beta_i X_i + \varepsilon \tag{1}$$

where

$\hat{Y}$: Regression model
β_0 : A constant term
β_i : Regression coefficients ($i = 1 \dots, n$)
X_i : Independent variables' column vectors.
ε : A disturbance term.

Multiple OLS regression technique was used to examine and model the relation between boarding times (T_B) of passengers and effective parameters. For this purpose, a total of 5030 passenger boarding times were investigated from seven different cities. The coefficients of the regression model are defined as the change in the expected value of $\hat{Y}$ associated with a one-unit increase in an independent variable, with the other independent variables being held constant (Table 3). The disturbances are accepted as being distributed normally with an expected value of zero and a common variance (Mert, 2016). Coefficients of the model are estimated by using the least squares that minimizes the weights of the sample error sum of squares (Ardahan and Mert, 2013). According to the OLS regression model, disturbances of the model should be distributed normally. Previous studies found that if a model has big sample size ($n>30$), it can be accepted to be normally distributed based on the Central Limit Theorem (CLT) (Baltagi, 2008). In the study, disturbances of the model were accepted as being distributed normally based on the CLT. In the model analysis multicollinearity, heteroscedasticity and model specification error problems were investigated by using diagnostic tests to verify that the examined model is meaningful. The dependent and independent variables of the OLS regression are given in Table 3.

Table 3. Dependent and independent variables of the regression model.

Dependent Variable	
T_B	Boarding times of the passengers (sec.) as dependent variable
Covariates	
N_S	Number of steps at boarding door
Dummy Variables	
B_{T_2}	(if bus type is articulated:1, otherwise: 0)
B_{T_3}	(if bus type is midibus:1, otherwise: 0)

E_{C_2}	(if examined city is Gümüşhane: 1, otherwise: 0)
E_{C_3}	(if examined city is İzmir: 1, otherwise: 0)
E_{C_4}	(if examined city is İstanbul: 1, otherwise: 0)
E_{C_5}	(if examined city is Ankara: 1, otherwise: 0)
E_{C_6}	(if examined city is Trabzon: 1, otherwise: 0)
E_{C_7}	(if examined city is Samsun: 1, otherwise: 0)
E_{BS}	(if bus has a step: 1, otherwise: 0)
P_{D_2}	(if passenger density at bus is low: 1, otherwise: 0)
P_{D_3}	(if passenger density at bus is mid: 1, otherwise: 0)
P_{D_4}	(if passenger density at bus is high: 1, otherwise: 0)
P_{P_2}	(if there is a child with passenger: 1, otherwise: 0)
P_{P_3}	(if passenger has an old age: 1, otherwise: 0)
P_{P_4}	(if passenger is handicapped (with wheelchair): 1, otherwise: 0)
E_{BB}	(if there is a bus bay at bus stop: 1, otherwise: 0)
U_{BB}	(if bus driver use bus bay: 1, otherwise: 0)

In the OLS regression, developed boarding time estimation model 1 (BTEM-1) includes qualitative and quantitative variables as given in Table 4, and the model is named as Analysis of Covariance (ANCOVA) model. The suggested model equation can be written as

$$T_B = a_0 + a_1 \times B_{T\_2} + a_2 \times B_{T\_3} + a_3 \times E_{C\_2} + a_4 \times E_{C\_3} + a_5 \times E_{C\_4} + a_6 \times E_{C\_5} + a_7 \times E_{C\_6}$$

$$+ a_8 \times E_{C\_7} + a_9 \times E_{BS} + a_{10} \times P_{D\_2} + a_{11} \times P_{D\_3} + a_{12} \times P_{D\_4} + a_{13} \times P_{P\_2} + a_{14} \times P_{P\_3}$$

$$+ a_{15} \times P_{P\_4} + a_{16} \times E_{BB} + a_{17} \times U_{BB} + b_1 \times N_S + \varepsilon \qquad (2)$$

where

a_0 : A constant term of the model,
a_i : Used dummy variables ($i \neq 0$),
b_j : Coefficients of the determined variables in the model ($j = 1, \dots, 2$),
ε : Disturbance term of the model.

To calculate the coefficient of Eq. (2), an OLS estimator was used. Obtained analysis results are summarized in Table 4. After controlling multicollinearity, heteroscedasticity, and model specification error problems, disturbances are accepted as being distributed normally because of the huge sample size (5030 passengers) (Baltagi, 2008). It was seen in Table 4 that the suggested model is statistically significant ($F = 10467.9$; $P = 0.000 < 0.01$ and $R^2 = 0.919$).

Table 4. Results of predicted model for boarding times (T_B) of the passengers.

Variables	Coefficient[†]	Std. E.	t	P value[†]
Dependent Variable: T_B				
a_0	1.478*	0.018	82.89	0.000
B_{T_2}	0.239*	0.013	18.22	0.000
B_{T_3}	0.752**	0.027	-27.61	0.011
E_{C_2}	0.724**	0.029	25.38	0.048
E_{C_3}	-0.230**	0.021	-10.96	0.032
E_{C_4}	-0.208**	0.020	-10.26	0.030
E_{C_5}	-0.344**	0.021	-16.60	0.039
E_{C_6}	-0.428**	0.022	-19.70	0.041
E_{C_7}	-0.449**	0.023	-19.55	0.038
E_{BS}	0.239**	0.015	6.47	0.041
P_{D_2}	0.584*	0.018	33.34	0.000
P_{D_3}	1.388*	0.020	68.66	0.001
P_{D_4}	2.553**	0.022	118.53	0.024
P_{P_2}	4.085**	0.025	160.52	0.048
P_{P_3}	0.227*	0.015	15.41	0.008
P_{P_4}	12.651**	0.041	311.43	0.040
E_{BB}	-0.086	0.049	5.639	0.112
U_{BB}	-0.456	0.018	11.390	0.124
N_S	0.241	0.008	30.81	0.000
Max. VIF	3.4 < 10 (there is no multicollinearity problem)			
White Test	$P = 0.546 > 0.10$ (there is no heteroscedasticity problem)			
Shapiro-Wilk W Test	$P = 0.141 > 0.10$ (disturbances of the model are normally distributed)			
Ramsey Reset Test	$P = 0.476 > 0.10$ (model have no model specification error problem)			
n: 5030; $F(16, 4280)$; 10467.9; $P > F$: 0.000; RMSE: 0.3155				

[†]** Significant at 0.05 level, *Significant at 0.01 level.

Model results show that all three (midibus, single bus, or articulated bus) bus types have significant and positive effects on boarding times (coeff.>0; *P*<0.05). On the other hand, only Gümüşhane city's public bus transport has positive effect on boarding times than other cities (coef.=0.724; *P*<0.05). That is, examined bus type in Gümüşhane city (the least crowded) is "Midibus," and it has higher times caused by the passenger property and step existence in boarding doors. The variables, existence of bus step (E_{BS}), and number of steps at boarding door (N_S) have also positive and significant effects on boarding times (coeff.>0; *P*<0.05). Furthermore, the coefficients of passenger density (P_D) are significant and have positive effects on boarding times (coef.=0.584, *P*=0.000<0.01 for P_{D_2}; coef.=1.388; *P*=0.001<0.01 for P_{D_3}; coef.=2.553; *P*=0.024<0.05 for P_{D_4}). Property of the passengers at boarding door has also a positive and significant coefficient in the model. It means that having different passenger characteristics results in positive significant effects on boarding times caused by the characteristic property of the passengers such as a child, the handicapped, or the elderly. It was also seen that examining the other six cities has significant and negative effects on boarding times. It can be concluded from the population of the other six cities that Gümüşhane city is the smallest city among them, and it has low density city property (less than one hundred thousand). Additionally, it was seen from the results that dummy variables, existence of a bus bay (E_{BB}), and utilization of bus bay by the drivers (U_{BB}) have no significant effect on boarding times. This means that they do not have great and significance effect on boarding times in the entrance of public buses.

3.2. Modeling of Boarding Times Using ABC Algorithm

The Artificial Bee Colony (ABC) algorithm is a one of the most used metaheuristic search algorithms for optimizing many different real world problems such as transportation, engineering design, scheduling analysis, cost optimization, and logistic planning, as well as the optimization method (Naghibi and Delavar, 2016). ABC is recently developed by Karaboga in 2005, and it was applied to different problems by many researchers (Karaboga, 2005; Karaboga and Basturk, 2008; Karaboga and Akay, 2011; Karaboga *et al.,* 2014; Szeto and Jiang, 2012; Naghibi and Delavar, 2016; Sonmez *et al.*, 2017; Yao *et al.,* 2017). ABC algorithm is inspired by the intelligent foraging behavior of honey bees, and it is mainly based on foraging behaviors of honey bee swarms in their daily life, and the main steps and behaviors of the bees in the algorithm are given below (Figure 4).
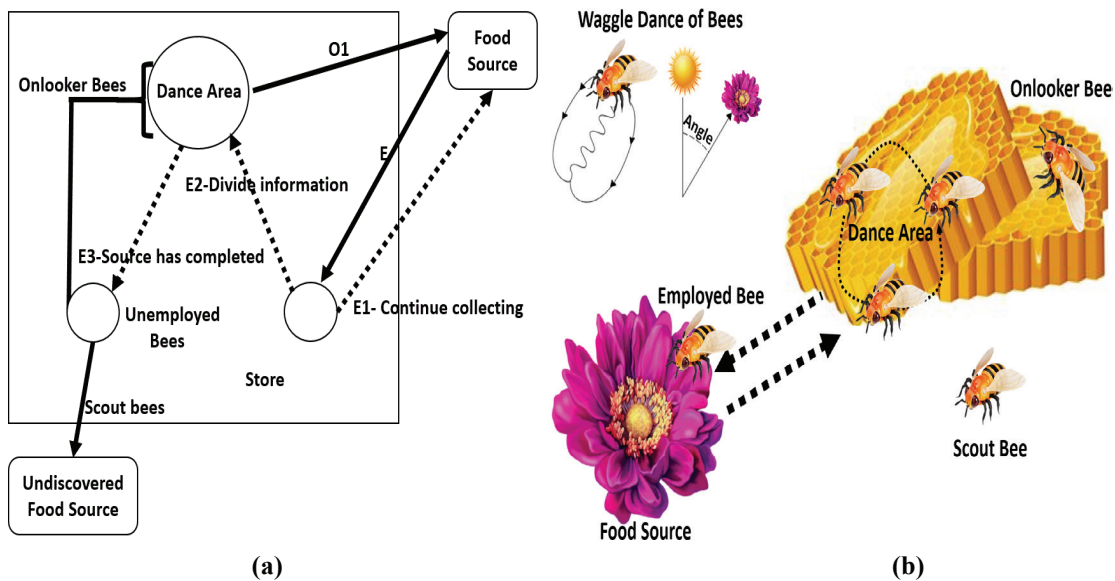

Figure 4. The elements of ABC algorithm. (a) Foraging behavior of bees, (b) schematic representation.

In this method, all bee types in the population are classified into three different main groups, employed bees (E), onlooker bees (O), and scout bees (S), according to their work responsibilities (Karaboga and Basturk, 2008). According to this algorithm, the first group (1st group bees) is employed bees. Each employed bee in the colony determines the food source, considers its nectar amount, and keeps the location of better sources in their memory. Therefore, each food source is determined only by one employed bee. This shows that the number of employed bees in the colony algorithm is equal to number of food sources. Employed bees also share this important information with other bees in dancing area by dancing when they fly back to hive share. The duration of dance shows the amount of nectar in the food area. Onlooker bees (2nd group bees) observe the duration of dance and may decide to fly to the food source if they find it worthwhile to visit the food source. If the amount of the food source is high, this important information attracts more onlooker bees to the food source by the time the source is exhausted. After the food source is exhausted, employed bees of the food source are replaced by scout bees (3rd group bees). The scout bees discover new food areas by making randomized searches. Every new food source is evaluated as a possible solution for the examined optimization problem. Additionally, the amount of nectar also shows the solution quality, which is identified by its fitness value (Karaboga and Basturk, 2008; Aydoğdu *et al.,* 2016). All the algorithm's four steps are summarized with the minimization of the calculated errors. The objective function (minimization) of the optimization model can be expressed as given in Eq. (3) (Sultana *et al.,* 2018).

$$\text{Minimize } f(\vec{x}) = \sum_{i=1}^{V_S}((A_t)_i - (F_t)_i)^2, \ \vec{x} = [x_1, x_2, \cdots, x_n] \tag{3}$$

where

$\vec{x}$	: Weight coefficient vector of the model (design variables),
V_s	: Used data number in error analysis,
A_t	: Real value,
F_t	: Predicted value,
n	: Number of total weight coefficient in model.

In the scope of the study, Artificial Bee Colony (ABC) method is used as the second method to examine and model (BTEM-2) boarding times (T_B) to a bus in the examined cities. For this purpose, modeling of boarding times (T_B) was determined as the aim function of the study. The flow chart of the used ABC algorithm can be summarized as given in Figure 5.
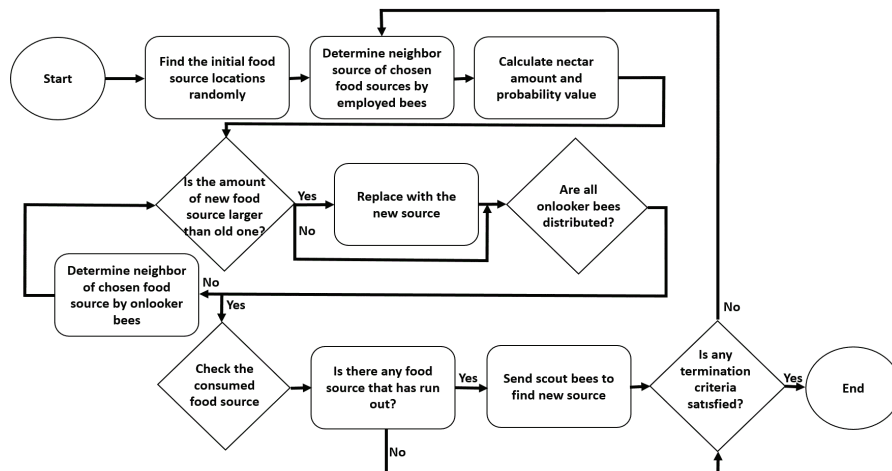

Figure 5. The flow chart of used ABC algorithm.

In the modeling study, statistically significant parameters such as Bus Type (B_T), Examined City (E_C), Existence of Bus Step (E_{BS}), Number of Steps at Boarding Door (N_S), Passenger Density inside the Bus (P_D), and Passenger Property at Boarding Door (P_P) were used as the examined parameters and statistically insignificant parameters. Existence of a Bus Bay (E_{BB}) and Utilization of Bus Bay by the Drivers (U_{BB}) were not used. In the analysis, the population size and maximum iteration numbers are taken as 50×100 and 10000, respectively. In each algorithm, the stopping criteria are to terminate the search process when the maximum number of generations is reached (assumed 1000 generations) or $|f(\vec{x})_{max} - f(\vec{x})_{min}|<10^{-4}$ where $f(\vec{x})$ is objective function value. The equation of the suggested model in the study by using ABC algorithm can be written as given in Eq. (4).

$$T_B = 2.54 \times B_{T\_1} + 2.77 \times B_{T\_2} + 1.90 \times B_{T\_3} + 0.30 \times E_{C\_1} + 1.12 \times E_{C\_2} - 33.24 \times E_{C\_3} - 25 \times E_{C\_4}$$

$$-13.6 \times E_{C_5} + 1.04 \times E_{C_6} - 24.06 \times E_{C_7} + 0.5 \times E_{BS} - 3.73 \times P_{D_1} - 3.10 \times P_{D_2} - 2.34 \times P_{D_3}$$

$$-1.23 \times P_{D\_4} + 2.9 \times P_{P\_1} + 7 \times P_{P\_2} + 3.10 \times P_{P\_3} + 15.55 \times P_{P\_4} + (N_S + 1)^{0.39} \tag{4}$$

where

T_B	: Boarding times of the passengers
B_{T_i}	: Bus Type ($i = 1, ..., n$)
E_{C_i}	: Examined cities ($i = 1, ..., n$)
E_{BS}	: Existence of bus step (yes/no)
P_{D_i}	: Passenger density inside the bus ($i = 1, ..., n$)
P_{P_i}	: Passenger Property at boarding door ($i = 1, ..., n$)
N_S	: Number of steps at boarding door.

To measure the performance of the model, descriptive statistical analysis and an error calculation were done by using Mean Absolute Percentage Error (MAPE). MAPE was obtained the average of absolute errors divided by actual observation values. Calculated descriptive statistics and MAPE (%) results of the suggested model are given in Table 5. As seen in the table, the developed model's estimation error (%) equals 9.9<10, and it shows that BTEM-2 can be used in the analysis. Furthermore, it can be also seen that the observed (μ_{OT_B}) and predicted (μ_{PT_B}) mean boarding times and std. deviations are close to each other. It means that there is not a great difference between them. Thus, it can be concluded that developed model's boarding time estimation has a good performance.

Table 5. Descriptive statistics and MAPE error for observed and predicted values in BTEM-2.

Boarding Time (sec.) (T_B)	Mean (μ)	Std. Error	Std. Dev. (σ)	Kurtosis	Skewness	Max.	Min.	MAPE (%)
Observed (OT_B)	2.80	0.03	2.00	22.28	4.15	18.66	0.98	9.92
Predicted (PT_B)	2.72	0.03	1.97	23.45	4.30	17.34	1.07	

The distributions of observed and estimated values, obtained from site observations and developed model estimations, respectively, are also shown in Figure 6. As seen in the figure, distribution characteristics of cumulative frequency (%) and observed passenger numbers according to boarding times are quite similar for observed and predicted values. This result also supports the accuracy of the developed model by using ABC algorithm.
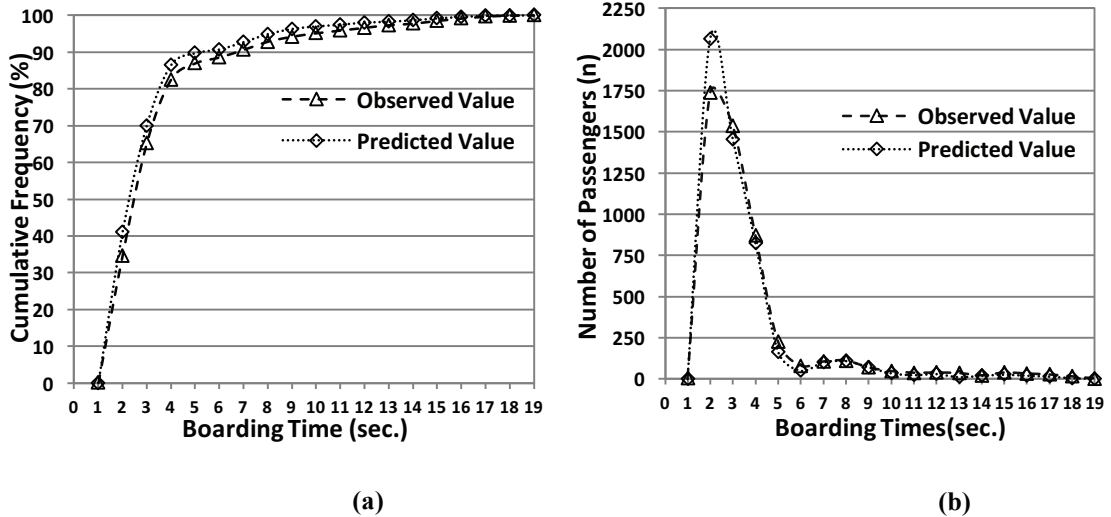
Figure 6. Distributions of boarding times (a) cum. frequency (%), (b) passenger number (n).

4. RESULTS AND DISCUSSIONS

The variations on travel times caused by the boarding and alighting behaviors have also a negative effect on passengers' decisions on mode and route choice. The boarding number and time at each bus stop has a great importance to calculate dwell and travel time for scheduling process. The current literature only gives several different methods that are used to calculate or predict dwell time of buses at bus stops to plan and schedule. However, it is known that boarding times at bus stops are one of the most important parameters on the dwell times. A possible boarding time modeling studies according to different factors can supply a benefit for planners and researchers. For this reason, a comprehensive study was conducted in the study from different cities in order to empirically examine and model the impact of various parameters. All effective factors have been determined on boarding behavior using OLS Regression (statistical analysis method) and ABC algorithm (optimization method) for modeling. Two new boarding time estimation models (BTEM-1 and BTEM-2) were proposed using two different modeling techniques.

The first developed model named as BTEM-1 was found to be significant ($F = 10467.9$; $P = 0.000 < 0.01$ and $R^2=0.919$). It was mainly seen from the model results that that boarding times have been negatively affected from the city population, and it means that if the city population increases, it probably will have a negative effect on boarding times. According to the second proposed model (BTEM-2) results, the developed model's estimation error (%MAPE) was calculated as 9.9<10, and this result shows that BTEM-2 estimation model can be used in the estimation analysis effectively. Distribution characteristics of cumulative frequency (%) and observed passenger numbers also support the accuracy of the developed BTEM-2 model.

All these findings for both models reveal that there is not a great difference between observation and model prediction values. Both developed models' boarding time estimations have a good performance. Overall, the results

of the study have demonstrated that modeling of boarding times by considering various parameters is an effective strategy to improve bus transport systems performance. Study results may become helpful for researcher and public transport planners to have better understanding on determination of different boarding times for public bus transport system planning and scheduling.

5. CONCLUSIONS

In the study, boarding times were modeled, and two new boarding time estimation models (BTEM-1 and BTEM-2) were developed using two different modeling techniques. This research mainly reveals the following results:

- Developed boarding time estimation models (BTEM-1 and BTEM-2) were found as effective analysis (estimation) methods in bus transport planning and operations.
- According to the BTEM-1 model, all bus types (B_T), being passenger in Gümüşhane (the least crowded among the examined cities) city (E_{C_2}), existence of bus step (E_{BS}), number of steps at boarding door (N_S), passenger density (P_D) inside the bus, and property of the passengers at boarding door, have statistically significant and positive effects on boarding times ($P<0.05$).
- BTEM-1 model results show that existence of a bus bay (E_{BB}) and utilization of bus bay by the drivers (U_{BB}) have no significant effect on boarding times of the passengers as dummy variables of the model. This means that they do not have great and significance effect on boarding times in the entrance of public buses.
- BTEM-2 model results reveal that observed (μ_{OT_B}) and predicted (μ_{PT_B}) mean boarding times and std. deviations were found to be so close to each other. This means that BTEM-2 has a good ability to fit the relation between used data and study parameters.
- Developed model results provide useful information for bus transport system planners and operators for improving bus service performance.

6. LIMITATIONS AND FUTURE AIMS

The modeling methodologies in the study are based on the obtained data from seven different cities of Turkey. Modeling methodologies could be applicable to other cities, and the general findings of the study may help for a better understanding on relation between boarding time and effective parameters. On the other hand, the transferability of the model parameters is limited. Both models may not provide desirable prediction accuracy for boarding time (sec.) estimation in other countries' cities that cannot be evaluated within the scope of the study. Especially, different passenger characteristics and bus transport facilities may cause different effects on model performances for other applications. Hence, future studies could investigate dwell and boarding times for bus transport systems for other countries using various methods with more extended data and parameters.

REFERENCES

Aydin, M. M., Yildirim, M. S., Aydin, R., & Arslan, Y. (2016). Effect of different passenger characteristics and bus types on boarding times at bus-stops. Journal of Engineering Research and Applied Science 5(2): 499–509.

Aydin, M. M., Gunay, B., & Akgol, K. (2019). Performance Comparison of Various Chicane Types: A Driving Simulator Study. International Journal of Civil Engineering 17(11): 1753–1765.

Aydin, M. M. (2020). A new evaluation method to quantify drivers' lane keeping behaviors on urban roads. Transportation letters 12(10): 738–749.

Aydoğdu, İ., Akın, A. & Saka, M. P. (2016). Design optimization of real world steel space frames using artificial bee colony algorithm with Levy flight distribution. Advances in Engineering Software 92: 1–14, DOI: 10.1016/j.advengsoft.2015.10.013.

Baltagi, B. H. 2008. Econometrics. 4th ed. Berlin: Springer.

Chen, S., Zhou, R., Zhou, Y. & Mao, B. (2013). Computation on bus delay at stops in Beijing through statistical analysis. Mathematical Problems in Engineering 2013: 1–9, DOI: 10.1155/2013/745370.

Currie, G., Delbosc, A., Harrison, S. & Sarvi, M. (2013). Impact of crowding on streetcar dwell time. Transp. Res. Rec.: J. Transp. Res. Board 2353: 100–106, DOI: 10.3141/2353-10.

Doxiadis, C. A. (1968). Ekistics; an introduction to the science of human settlements.

Dueker, K. J., Kimpel, T. J. & Strathman, J. G. (2004). Determinants of bus dwell time. Journal of Public Transportation 7(1): 21–40, DOI: 10.5038/2375-0901.7.1.2.

El-Geneidy, A., van Lierop, D., Grisé, E., Boisjoly, G., Swallow, D., Fordham, L. & Herrmann, T. (2017). Get on board: Assessing an all-door boarding pilot project in Montreal, Canada. Transportation Research Part A: Policy and Practice 99: 114–124, DOI: 10.1016/j.tra.2017.03.005.

El-Geneidy, A. & Vijayakumar, N. (2011). The effects of articulated buses on dwell and running times. J. Public Transp. 14: 63–86, DOI: 10.5038/2375-0901.14.3.4.

Fernandez, R., (2011). Experimental study of bus boarding and alighting times. In European Transport Conference, Glasgow, Scotland.

Fernandez, R., Zegers, P., Weber, G. & Tyler, N. (2010). Influence of platform height, door width, and fare collection on bus dwell time. Transp. Res. Rec.: J. Transp. Res. Board 2143: 59–66, DOI: 10.3141/2143-08.

Fletcher, G. & El-Geneidy, A. (2013). Effects of fare payment types and crowding on dwell time: fine-grained analysis. Transportation Research Record 2351(1): 124–132, DOI: 10.3141/2351-14.

Ghasemlou, K., Aydın, M. M., Tanyel, S., Topal, A. & Çalışkanelli, S. P. (2012). Effect of dwell time on performance of signalized intersections. In 10th International Congress on Advances in Civil Engineering, Middle East Technical University, Ankara, Turkey, 1–10.

Guenthner, R. P. & Hamat, K. (1988). Transit dwell time under complex fare structure. J. Transp. Eng. 114: 279–367, DOI: 10.1061/ (ASCE) 0733-947X(1988)114:3(367).

Karaboga, D. & Akay, B. (2011). A modified artificial bee colony (abc) algorithm for constrained optimization problems. Applied soft computing, 11(3): 3021–3031, DOI: 10.1016/j.asoc.2010.12.001.

Karaboga, D. & Basturk, B. 2008. On the performance of artificial bee colony (ABC) algorithm. Applied soft computing 8(1): 687–697, DOI: 10.1016/j.asoc.2007.05.007.

Karaboga, D., Gorkemli, B., Ozturk, C. & Karaboga, N. (2014). A comprehensive survey: artificial bee colony (ABC) algorithm and applications. Artificial Intelligence Review 42(1): 21–57, DOI: 10.1007/s10462-012-9328-0.

Karaboga, D. (2005). An idea based on honey bee swarm for numerical optimization. Technical Report-TR06, Erciyes University, Engineering Faculty, Computer Engineering Department.

Kraft, W. H. & Bergen, T. F. (1974). Evaluation of passenger service times for street transit systems. Transp. Res. Rec.: J. Transp. Res. Board 505: 13–20.

Larwin, T. F. (2012). Off-Board fare payment using proof-of-payment verification - a synthesis of transit practice. Transit Coop. Res. Program 96.

Levine, J. & Torng, G. W. (1994). Dwell time effects of low floor bus design. Journal of Transportation Engineering 120(6): 914–929, DOI: 10.1061/(ASCE)0733-947X(1994)120:6(914).

Levinson, H. S. (1983). Analyzing transit travel time performance. Transp. Res. Rec.: J. Transp. Res. Board 915: No.1–6.

Naghibi, F. & Delavar, M. (2016). Discovery of transition rules for cellular automata using artificial bee colony and particle swarm optimization algorithms in urban growth modeling. ISPRS International Journal of Geo-Information 5(12): 241.

Sonmez, M., Akgüngör, A. P. & Bektaş, S. (2017). Estimating transportation energy demand in Turkey using the artificial bee colony algorithm. Energy 122: 301–310.

Stevanovic, S. (2018). Urban planning indicators in sustainable building assessment methods. Journal of Engineering Research, 6(1): 1–14.

Sultana, U., Khairuddin, A. B., Rasheed, N., Qazi, S. H., & Mokhtar, A. S. (2018). Allocation of distributed generation and battery switching stations for electric vehicle using whale optimiser algorithm. Journal of Engineering Research, 6(3): 70–93.

Szeto, W. Y. & Jiang, Y. (2012). Hybrid artificial bee colony algorithm for transit network design. Transportation Research Record, 2284(1): 47–56: DOI: 10.3141/2284-06.

Tirachini, A. (2013). Bus dwell time: the effect of different fare collection systems, bus floor level and age of passengers. Transportmetrica A 9: 28–49, DOI: 10.1080/18128602.2010.520277

Yao, B., Yan, Q., Zhang, M. & Yang, Y. (2017). Improved artificial bee colony algorithm for vehicle routing problem with time windows. PloS one 12(9): e0181275, DOI: 10.1371/journal.pone.0181275