City Research Online

City St George's, University of London

Citation: Kaparias, I., Rossetti, C. & Trozzi, V. (2015). Analysing passenger arrivals rates and waiting time at bus stops. Paper presented at the 94th Annual Meeting of the Transportation Research Board, 11-01-2015 - 15-01-2015, Washington, DC, USA.

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/6257/

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

ANALYSING PASSENGER ARRIVALS RATES AND WAITING TIMES AT BUS STOPS

Ioannis Kaparias
Collaborative Transport Hub,
School of Mathematics, Computer Science and Engineering,
City University London,
Northampton Square,
London EC1V 0HB,
United Kingdom
Email: kaparias@city.ac.uk

Cesar L. Rossetti
Collaborative Transport Hub,
School of Mathematics, Computer Science and Engineering,
City University London,
Northampton Square
London EC1V 0HB
United Kingdom
Email: cesar.rossetti.1@city.ac.uk

Valentina Trozzi
Transport for London,
5Y7 Palestra,
197 Blackfriars Road,
London SE1H 8NJ,
United Kingdom
Email: valentinatrozzi@tfl.gov.uk

(5818 + 7 · Figures (250) = 7,568 words)

This paper was submitted to the Transportation Research Board.
1 November 2014

ANALYSING PASSENGER ARRIVALS RATES AND WAITING TIMES AT BUS STOPS

I. Kaparias [1], C.L. Rossetti [1], V. Trozzi [2]

1 City University London, UK
2 Transport for London, UK

ABSTRACT

The present study investigates the rather under-explored topic of passenger waiting times at public transport facilities. Using data collected from part of London's bus network by means of physical counts, measurements and observations, and complemented by on-site passenger interviews, the waiting behaviour is analysed for a number of bus stops served by different numbers of lines. The analysis employs a wide range of statistical methods and tools, and concentrates on three aspects: passenger interarrival time, passenger actual waiting time, and passenger perceived waiting time. The results suggest that there is a clear difference in terms of the passenger arrivals rate between stops served by up to two lines and stops served by three lines or more, as it appears that passengers in the former time their arrival at the stop to coincide with bus arrivals as much as possible. Also, it is found that waiting time at such stops is best approximated by the exponential distribution, with the gamma distribution also offering an adequate fit. Finally, as concerns the passengers' perception of waiting time, it is found that this follows a lognormal or a gamma distribution, and generally overestimates the actual waiting time; however, this effect fades as the actual waiting time increases.

1 INTRODUCTION

The fact that travel time in road networks is not constant, but entails an element of variability, resulting in uncertainty when attempting to predict it, has been recognised in the literature for a long time. Already in a very early study by Wardrop (1) it was noted that travel times follow a skewed distribution with a long 'tail' representing the few very slow vehicles, such that it is very likely for the mean travel time to be exceeded. In a later study by Thomson (2), travel time variability was identified as an important characteristic of road networks and it was pointed out that the unpredictability of travel time is one of the most important sources of time losses.

The importance of travel time variability has been the objective of much research in the past and has therefore been extensively analysed from the traveller's perspective. Many studies have concluded that although travel time is an important factor affecting the traveller's route choice behaviour, travel time variability can be even more important. Travellers are interested in how long it will take them to reach their destination, but are even more concerned with how reliable their prediction of total travel time is. A wrong travel time prediction results in either an early arrival at the destination or in a delay. None of these situations are appreciated by the traveller, with delays usually having more severe consequences for him/her (e.g. late arrival at the workplace) and therefore not being tolerated (3-5).

Hence, much research has focussed on quantifying and modelling travel time variability. This has almost exclusively concentrated on motorised vehicular traffic. As opposed to road networks, however, where traffic congestion can be fairly easily identified as the sole source of uncertainty, passengers in public transport networks of large cities may be exposed to delays arising from a number of sources, such as service reliability and overcrowding. For instance, an important component of travel time in public transport, currently only superficially considered by journey planners, is waiting time. In fact, travellers are more sensitive to the waiting time at public transport

stops than the actual time spent on board (6), and from an operational perspective, this may lead to them changing their route and mode choice, and sometimes even their final destination (7). And while this is a major problem in large cities' public transport networks, there does not seem to be any broad agreement in the literature on how waiting time should be measured and quantified.

The aim of this study is, hence, to shed light into the under-explored topic of passenger waiting time at public transport facilities, with a view of providing results for input into transport models. Focussing on part of London's bus network and using data obtained from counts and measurements, the study proceeds in three stages. First, the rate of arrivals of passengers at bus stops of different characteristics (mainly differing by the number of lines served) is looked at. Then, coupled with observations of bus and passenger departures, waiting time measurements are carried out, and conclusions with respect to the statistical distribution of waiting time are drawn. Finally, the study is complemented by a passenger survey exercise investigating the perceived waiting time along with a number of other characteristics, and comparisons are made between the perceived level of service and the actual one.

The present paper is structured as follows: Section 2 presents the background of the study, focussing primarily on previous research on the topic of the public transport waiting time modelling. Section 3 then goes on to present the study area and the data collection methods employed, and describes the analysis methodology. Section 4 presents the results of the analysis, and discusses the patterns of passenger interarrival times, actual waiting times and perceived waiting times at bus stops. Section 5, finally, concludes the paper and identifies areas of future research.

2 BACKGROUND

A number of studies in the past have attempted to understand and determine passenger arrival rates and waiting times at public transport facilities (7-12). All of them agree that passenger behaviour is schedule-dependent when the service has long headways (low frequency). In this case, most passengers time their arrival at the stop to coincide with the service arrival as much as possible, minimising thus their waiting time. In the case of shorter headways (high frequency) on the other hand, passenger arrivals become random. The main difference between the various studies in the literature is in the waiting time threshold that schedule-dependent arrivals become random.

The first study that looked into the topic of waiting time was that of Weber (6), which concluded from an analysis of Stuttgart's bus network that passengers start behaving schedule-dependently when the headway exceeds 7-8 minutes. The same threshold value was found in a later study by Seddon and Day (9), where a linear equation relating passenger waiting times and headways was derived, based on an analysis of bus headways in Manchester. A different threshold value, however, was found in the study by Joliffe and Hutchinson (10), which used bus data from London and concluded that passenger arrivals are random for waiting times up to 12 minutes. This threshold was additionally confirmed in a later study by O'Flaherty and Mangan (12) in Leeds, but it was pointed out that it applied only to off-peak periods; for peak periods, the threshold reduced to 5 minutes.

The study of Braendli and Mueller (11) was the first to acknowledge that regular passengers on a route are more likely to know the timetable, as opposed to occasional users. It proposed a passenger arrival model, in which passengers are divided into schedule-dependent with knowledge of the timetable, and schedule-independent without knowledge of the timetable. The study then derived distribution curves of passenger arrivals with respect to the public transport headways at stops according to these two types of passenger, revealing that passengers arrived near scheduled departure times within a short headway of 6 minutes.

Other studies have developed passenger waiting time models based on distribution curves of passenger arrival rates and the integral calculus method (7, 13). Specifically, Luethi et al (7) also divided passengers into schedule-dependent and schedule-independent, and proposed a model for

passenger arrival rates using a logarithmic function. The study proposed that passenger arrivals distribution can be modelled as a superposition of a uniform distribution (for the timetable-independent passengers) and a shifted Johnson SB distribution (for the timetable-independent passengers). The study also concluded that passengers begin to arrive at stations near the scheduled departure times, even for very short headways. Guo et al (6), on the other hand, fitted normal, exponential, lognormal and gamma distributions to arrival rates of passengers transferring from rail to buses and concluded that the lognormal and gamma distributions had the most appropriate fit for passengers transferring directly and non-directly.

A limitation of all studies performed to date is that they have only analysed stops served by a single line, and have estimated passenger waiting time only by statistical analysis of the bus headways and the so-called Platform Waiting Time (PlatWT), which has been defined by Luethi et al (7) as the actual time the passenger has waited on the platform (or the bus stop). Furthermore, while analysis of the data collected by these studies can provide useful information on waiting times from a purely operational perspective, it does not offer any insight into passengers' perceived waiting time, which is additionally influenced by psychological factors and has been found to be more significant from the point of view of customer satisfaction (14).

The present study extends existing knowledge in two ways: on one hand it conducts actual passenger counts and observations at bus stops served by multiple lines and analyses passenger arrival rates and waiting times; and on the other hand it complements this analysis with additional data on waiting time perception, collected through surveys. The methodology is presented in the next section.

3 METHODOLOGY

The data collection and analysis methodology are outlined in this section. This includes first a short description of the study area, and is followed by an account of the data collection methods and tools employed. Then, the analysis methods used are explained, in preparation for the reporting of the results in Section 4.

3.1 Study area

The area selected for this study is Harrow in North-West London. It is one of the 32 boroughs of Greater London, with an area of 50 km^2 and a population of roughly 250,000 (15). The rationale behind the selection of Harrow is that the area benefits from good rail connections to Central London, but these are part of London's largely radial rail and underground networks, and as such do not accommodate trips within Harrow itself, which rely almost entirely on the local bus network. The latter consists of 16 bus lines, serving the entire borough and connecting it to the adjacent areas. The location and the structure of the Harrow bus network are shown in Figure 1.

A set of bus stops from the study area have been chosen to perform counts, observations and interviews. In order to ensure that the passenger behaviour monitored is as representative as possible, it has been deemed appropriate to survey stops located within approximately 15-20 minutes' walk from the nearest railway station, so as to avoid the case when walking and other modes may act as an alternative to buses. In total 18 bus stops have been surveyed, of which six are served by one line, another six by two lines, two by three lines, two by five lines, and another two by six lines.

The data collection has been carried out between 5 November 2013 and 17 March 2014. Appropriate one-hour survey slots have been identified on school day mornings from 08:00 to 09:00, so as to ensure that the network is surveyed during the peak period, thus maximising the amount of data collected. Each bus stop has been surveyed only once, as surveying on additional days would introduce a bias in the results due to double-counting, given the fact that the vast majority of the travellers at that time are commuters and are likely to perform the same journey every day.

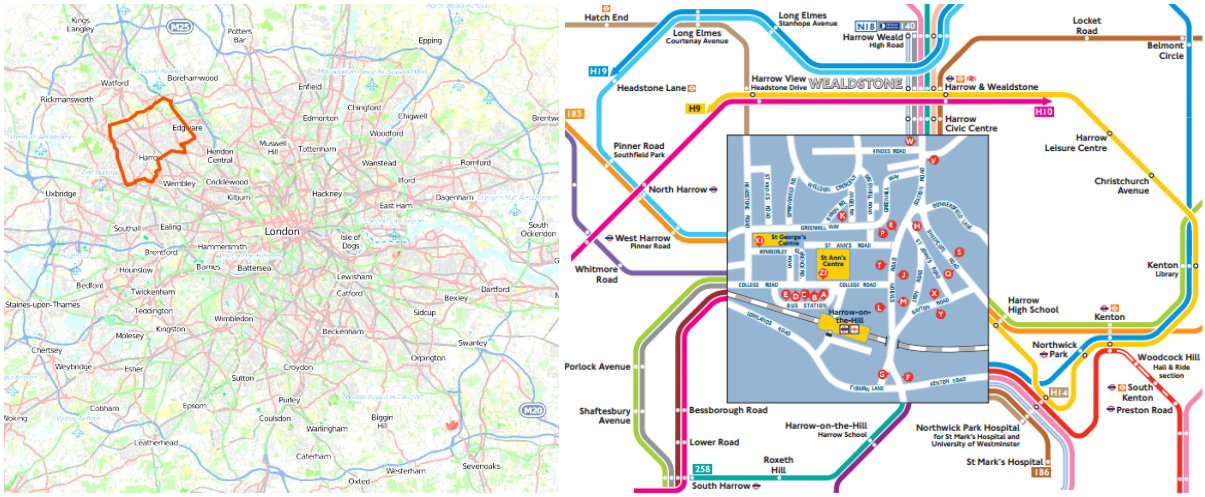


FIGURE 1: Left: Location of Harrow (16); Right: Harrow bus network (17)

3.2 Data collection methodology

Two data collection methods have been used in this study, namely physical counts and on-site passenger interviews at bus stops.

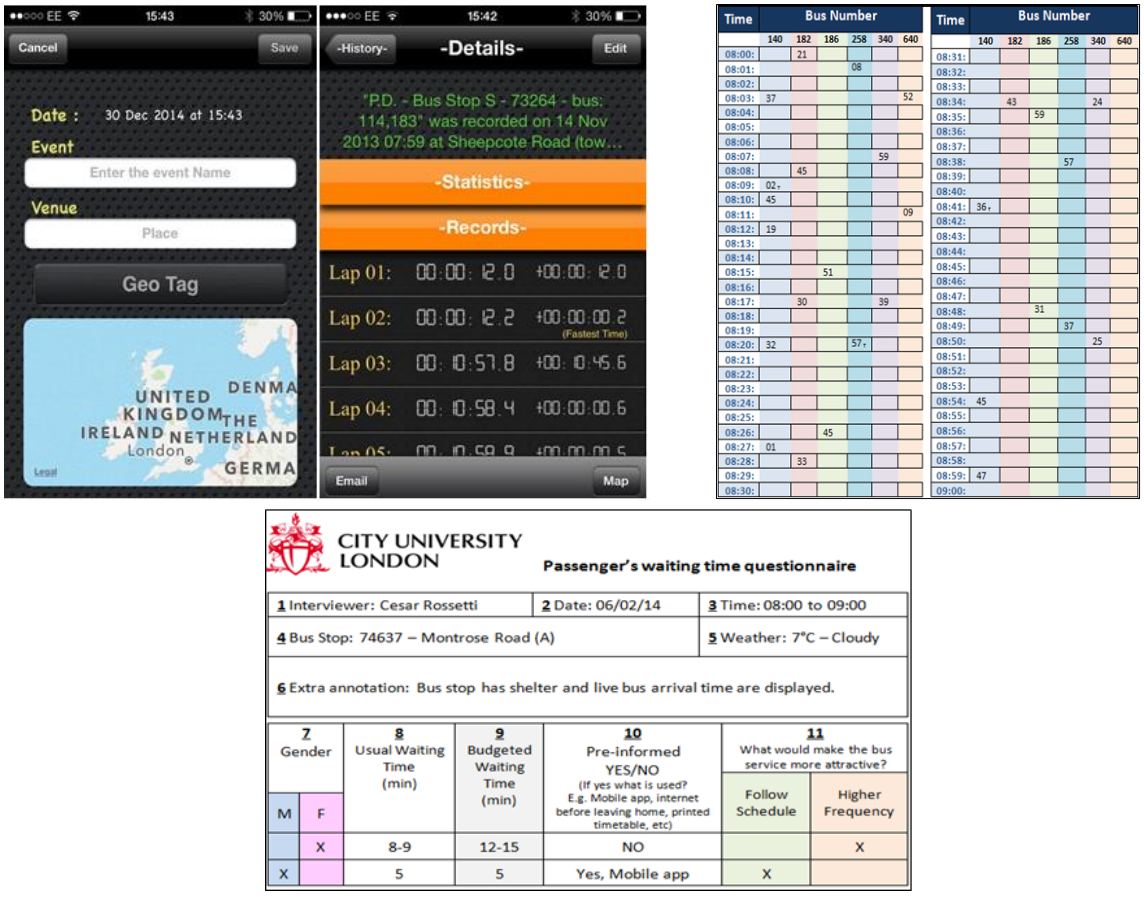


FIGURE 2: Top left: The StopWatch app (18) used in the recording of passenger arrivals and departures; Top right: Counting sheet used in the recording of bus arrivals; Bottom: Interview questionnaire

Physical counts have been conducted with two observers at each bus stop surveyed, who have been allocated separate tasks. The first observer has been responsible for recording the time of passenger arrival occurrences at the bus stop, while the second observer has recorded the times of bus arrival and passenger bus boarding occurrences. The recording of passenger arrival and departure times has been facilitated by the use of two mobile devices running the StopWatch app (18). The recording of bus departure times, on the other hand, has been carried out with the help of counting sheets (Figure 2).

Passenger interviews have been conducted by one interviewer at the same time as the physical counts. They have been facilitated by a questionnaire, whose aim has been to collect passenger perception-specific information as possible, while also being non-intrusive and allowing for the interview to be fast. As such, following a pilot experiment, demographic questions and other general questions on the respondent's journey have not been included, and interviews have instead focused entirely on waiting-time-specific questions, such as the traveller's usual waiting time (i.e. the waiting time usually experienced by a regular traveller) and budgeted waiting time (i.e. the total time allowed for waiting at the bus stop in order to minimise the chances of arriving late at the final destination in case the service does not run on schedule (14), as well as the use of travel information. The questionnaire used is shown in Figure 2.

It should be noted that due to practical limitations relating to the conduct of the counts and interviews, passenger departure times have not been recorded at stops served by more than two lines, and interviews have been conducted only at stops served by one line.

3.3 Analysis methodology

The analysis of the collected data consists of three main tasks, in each one of which different characteristics of the waiting behaviour of bus passengers are observed.

1. Analysis of passenger arrival rates

Considering all 18 stops surveyed, these are categorised by type, which in this case is based on the number of lines serving the stop. Hence, three stop categories are defined, namely "1 line", "2 lines" and "3+ lines". Frequency distributions of the passenger interarrival time (i.e. the time between individual passenger arrival occurrences) for each of the three stop categories are extracted, and key statistics, such as mean, standard deviation, skewness and kurtosis are computed. The analysis then proceeds by investigating similarities and differences between the distributions of the three stop types, verified by means of statistical significance tests, and draws conclusions with respect to the arrival patterns of passengers at bus stops.

2. Analysis of passenger actual waiting times

Focussing on the "1 line" and "2 line" stops, the actual dwelling time for each passenger observed/counted at each stop is established by subtracting the arrival time from the departure time; this, naturally, corresponds to the passenger's PlatWT. At the same time, the so-called "Average Waiting Time" (AvWT) is calculated for each stop and each line from the measured bus headways (time between bus departure occurrences). This is defined by Furth et al (14) as the sum of the halves of the squares of the measured headways, divided by the total time between the first and last observed bus departure. The PlatWT values are then normalised by the AvWT, and distributions of the normalised PlatWT for each of the two stop types are derived, and key statistics, such as mean, standard deviation, skewness and kurtosis are computed. The analysis then goes on to identify similarities and differences between the distributions of the two stop types, verified through appropriate statistical significance tests. It is additionally investigated whether the observed data can fit a standard probability distribution with the same characteristic values.

3. Analysis of passenger perceived waiting times

Focussing on stops served by one line only, the responses to the interviews are looked at, and in particular the survey responses relating to the perceived waiting time (PerWT) and the budgeted waiting time (BudgWT) are analysed. Following the same analysis method as for the actual waiting times, the values of PerWT and BudgWT are normalised by the AvWT for each stop, and probability distributions are deduced from them. These are compared with the normalised PlatWT distribution, and differences and similarities between the three are identified and verified by means of statistical significance tests. It is additionally investigated whether the observed PerWT and BudgWT data can fit a standard probability distribution with the same characteristic values.

4 RESULTS

The data collected are analysed using the methods described in Section 3.3, and the results are presented in this section according to each of the three tasks outlined.

4.1 Passenger arrival rates

A total of 870 measurements of passenger interarrival times have been collected across the 18 stops surveyed. This corresponds to 164 measurements of "1 line" stops, 189 measurements of "2 lines" stops, and 517 measurements of "3+ lines" stops. The interarrival time distributions for each of the three stop categories are presented in the form of histograms in Figure 3, and their characteristic values are also given.

From a first glance, it becomes apparent that the distributions of the interarrival times for the three stop types are heavily left-skewed, with a significantly higher frequency of observations of smaller interarrival times than higher ones. This is in agreement with what is found in the literature, where passenger arrivals are often modelled as a Poisson process (and hence interarrival time follows an exponential distribution).

Comparing the three histograms and the corresponding statistical measures, it can be seen that the "1 line" and "2 line" distributions are fairly similar. Indeed, a two-sample two-tailed homoscedastic t-test confirms that there are no significant differences between them (p-value = 0.718, which means that the null hypothesis that the two samples are the same cannot be rejected at the 0.05 level). However, the histogram of "3+ lines" is different to the other two, and has a lower mean and standard deviation and a much higher kurtosis, which points to a narrower and more "peaked" histogram. T-tests confirm this finding (p-values of 0.000 in the comparison with both the "1 line" and the "2 lines" distributions mean that the null hypothesis that they come from the same population is rejected at the 0.05 level). This effect becomes more apparent in the cumulative distribution plots of Figure 4a (the "3+ lines" plot is steeper and narrower than the other two), and suggests that the interarrival times at stops served by three or more lines are generally much shorter; in other words, passengers arrive more frequently and more regularly at stops served by three or more lines.

It can, of course, be argued that this is an intuitive finding, as when three lines serve a stop, it will attract more passengers than if it was served by one or two lines, and so it will result in more passenger arrival occurrences during the same time, and consequently in shorter interarrival times. To shed more light on this effect, a normalisation process of the interarrival times is carried out, whereby the inverse of the interarrival time (i.e. the rate of arrivals) is divided by the average rate of arrivals at each stop (i.e. the total number of arrivals in the observation period divided by the duration of the period).

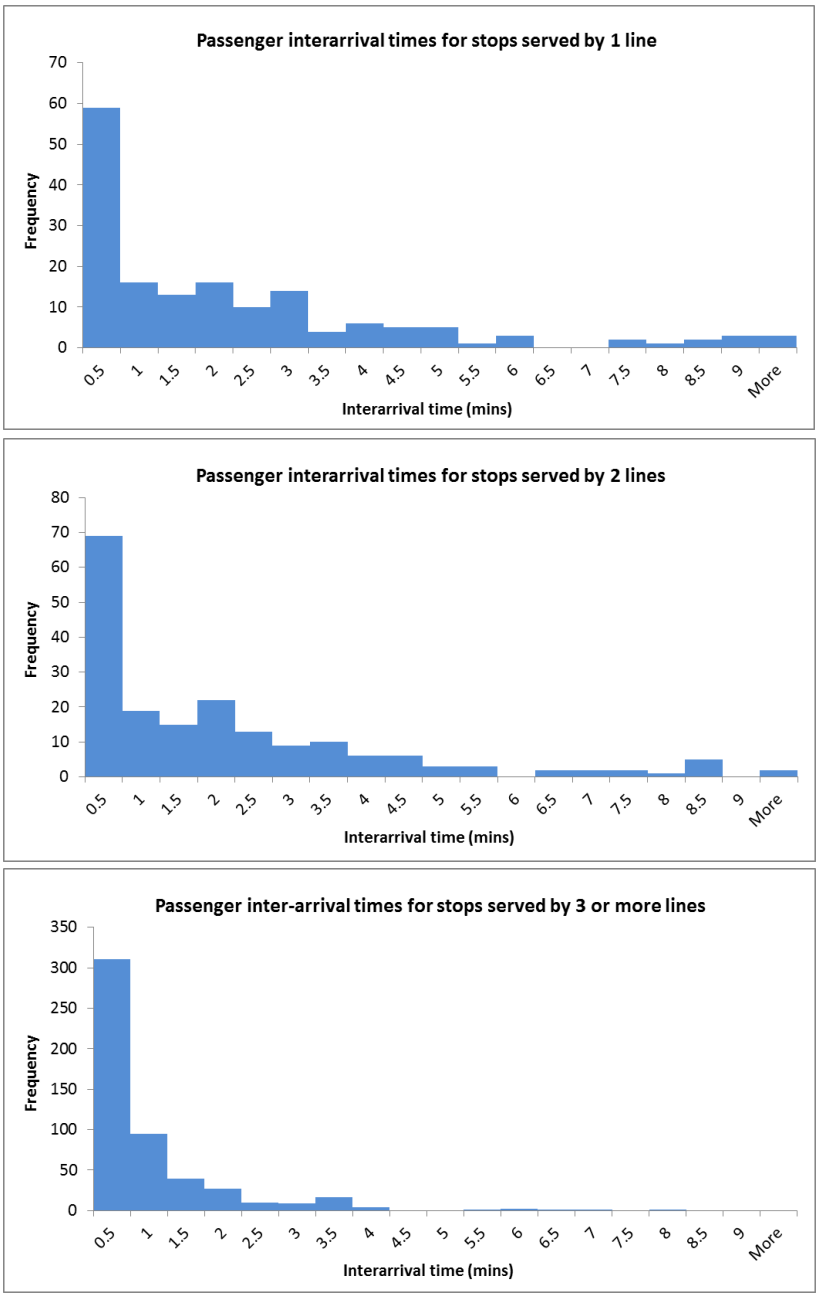


FIGURE 3: Passenger interarrival times at bus stops served by (a) one, (b) two, and (c) three or more lines

The normalised cumulative distribution plots for each of the three stop categories are shown in Figure 4b. T-tests between the three normalised samples (in pairs) again show that "1 line" and "2 lines" distributions are similar, and that "3+ lines" is significantly different. But the most important effect that can now be recognised is that, even if the bias introduced by the different passenger volume is offset, there is still a difference in the way passengers arrive at different stop types. Specifically, the normalised "1 line" and "2 lines" distributions have a high concentration of very high and very low arrivals rates, such that their cumulative plots are steep for small values, "bend" at a cumulative probability of roughly 0.55, and then become almost horizontal. This indicates a high concentration of very short and very long interarrival times, which is consistent with schedule-dependent behaviour, where passengers arrive at the stop shortly before arrival of the bus. On the other hand, the "3+ lines"

cumulative normalised arrivals rate plot "bends" at a lower probability value (0.40), and then makes a much more gradual ascent. This suggests that interarrival times are much more evenly distributed, such that passengers arrive at stops serving three or more lines independently of bus arrivals.

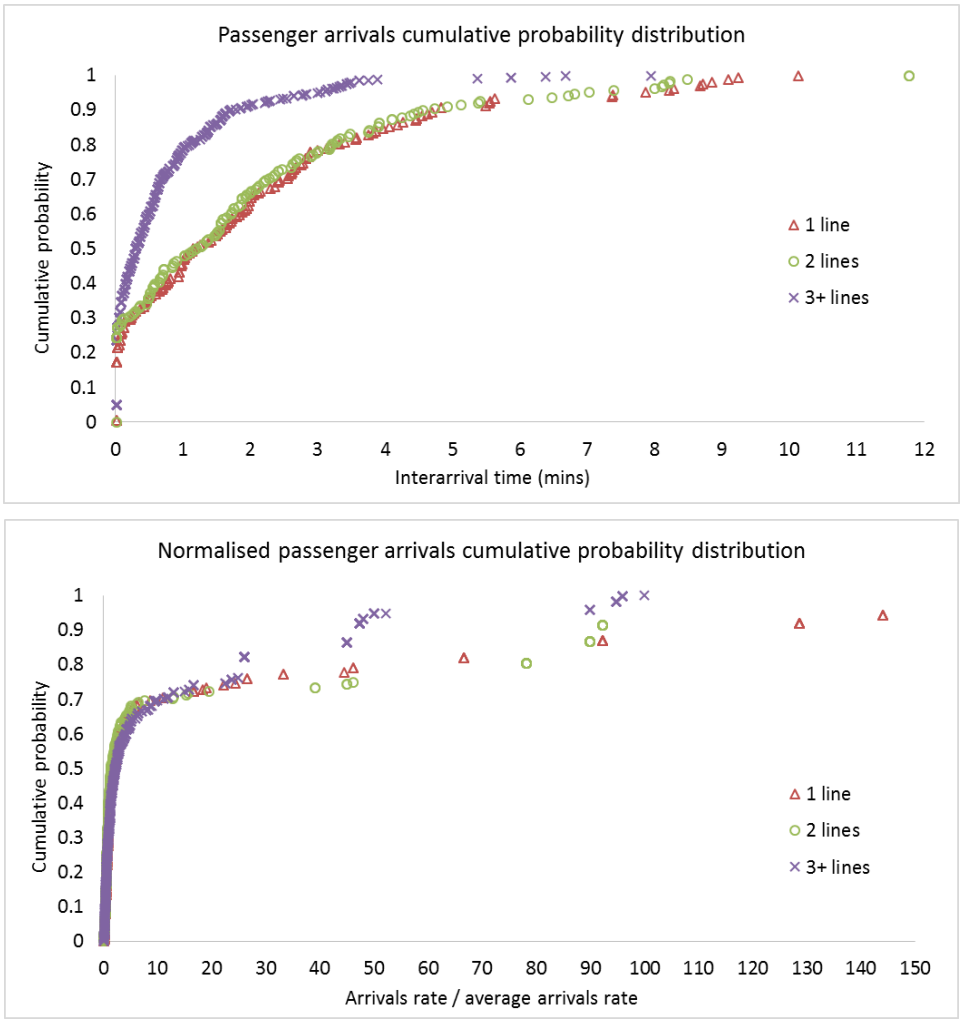
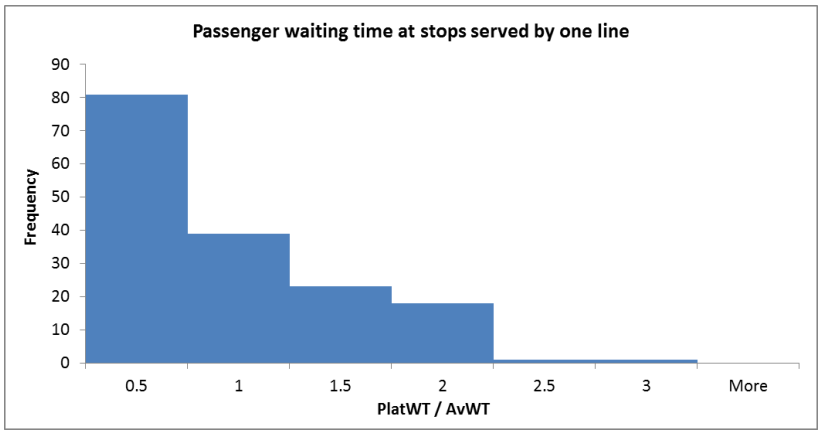


FIGURE 4: Cumulative distribution plots for (a) interarrival time and (b) normalised arrivals rate for each of the three stop types

4.2 Passenger actual waiting times

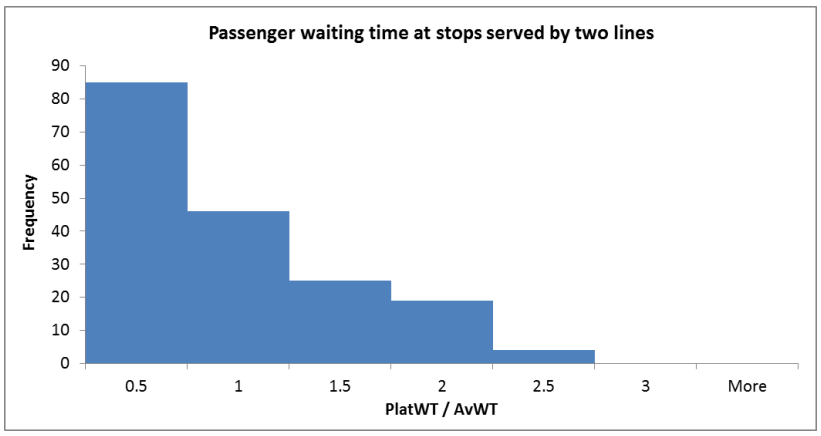
Due to practical limitations relating to the tracking of specific passengers at bus stops, it has not been possible to record the departure time and the bus line boarded of passengers at stops served by three lines or more, and hence waiting time analysis concentrates on stops served by one or two lines. This includes 12 stops and a corresponding 342 measurements of passenger actual waiting time (PlatWT), of which 163 refer to "1 line" stops and 179 to "2 lines" stops. When examining the two resulting distributions in raw form, however, any conclusions could again be biased, as the PlatWT is heavily dependent on the bus line frequency of service, and so measurements relating to different lines may not comparable. To allow for a comparison, the PlatWT measurements are normalised by the AvWT metric as defined in Furth et al (14). The normalised PlatWT distributions for each of the two stop categories considered are presented in the form of histograms in Figure 5, and their characteristic values are given.

From a first glance, it is evident that both distributions are left-skewed, which is consistent with what can be found in the literature. Also, it can be seen that the two distributions have many similarities (the only characteristic value that is different is the kurtosis, which indicates a slightly stronger "peakedness" of the "2 lines" distribution). A two-sample two-tailed homoscedastic t-test confirms this observation, and shows that there are no significant differences between them (p-value = 0.912, which means that the null hypothesis that the two distributions are the same cannot be rejected at the 0.05 level). This is also in line with the finding of the similarity of the "1 line" and "2 lines" interarrival time distributions.



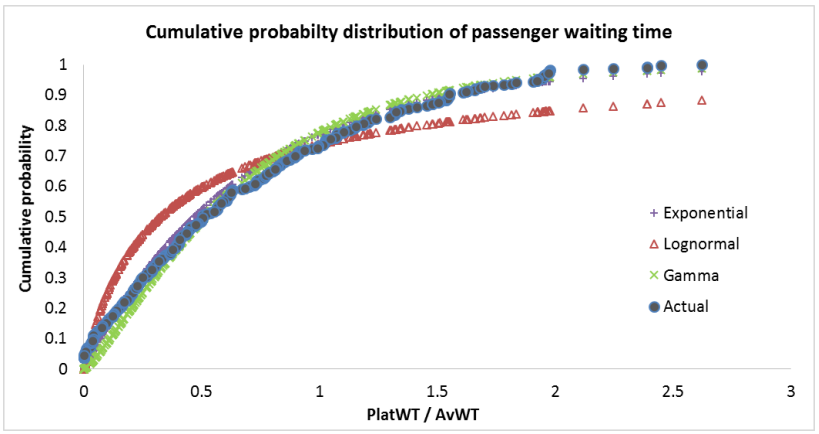
(a)

Mean = 0.673
Median = 0.502
Std. dev. = 0.607
Skewness = 0.926
Kurtosis = 0.116



(b)

Mean = 0.679
Median = 0.506
Std. dev. = 0.571
Skewness = 0.974
Kurtosis = 0.224



(c)

K-S test results:
$D^{0.05}_{crit}$ = 0.074
$D^{0.05}_{crit}$ = 0.088

Exponential:
D = 0.053, accepted at level 0.05

Lognormal:
D = 0.151, rejected

Gamma:
D = 0.084, accepted at level 0.01

FIGURE 5: Passenger actual waiting time (PlatWT) distribution at bus stops served by (a) one and (b) two lines; (c) cumulative distribution plots of PlatWT compared with exponential, lognormal and gamma plots.

Since the "1 line" and "2 lines" PlatWT distributions are essentially the same, it is more appropriate to consider the combined PlatWT distribution of the two stop types when fitting a standard probability distribution. Given the left-skewed shape, appropriate standard distributions would be the exponential, the lognormal and the gamma distributions. The fitting is performed by first deducing the cumulative PlatWT distribution, and then generating exponential, lognormal and gamma cumulative plots on the basis of the same characteristic values (i.e. mean and standard deviation). The resulting cumulative plots are then compared with the PlatWT distribution by means of Kolmogorov-Smirnov (K-S) tests, and their goodness-of-fit is assessed.

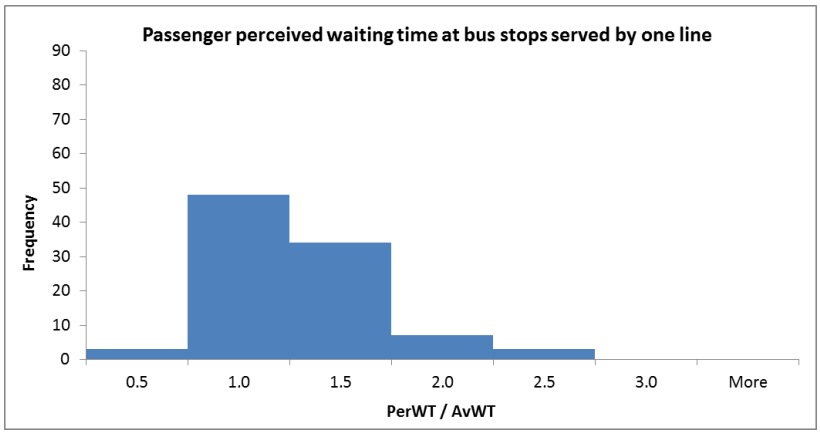
The cumulative PlatWT distribution, and the corresponding exponential, lognormal and gamma cumulative plots, are shown in Figure 5c. In an initial visual assessment it can be seen that that lognormal distribution with the same mean and standard deviation as the actual PlatWT distribution is not a good fit; this is confirmed by the K-S test, which is rejected at both the 0.05 and the 0.01 levels (the K-S statistic of 0.151 is larger than the critical values at both levels). The other two plots, on the other hand, seem to reflect the actual observations much better, and the respective goodness-of-fit tests reinforce this finding. Specifically, the gamma distribution is a fairly good match, as the K-S statistic of 0.084 is rejected at the 0.05 level but accepted at the 0.01 level, and the exponential distribution is an even better match, with the K-S statistic of 0.053 being accepted at the 0.05 level.

4.3 Passenger perceived waiting times

The final stage of the analysis considers the passenger interviews at bus stops served by one line. This includes 6 stops and a corresponding 95 responses, of which 40 (42%) come from male respondents and 55 (58%) from female ones, and of which 38 (40%) have stated that they have been pre-informed on actual bus arrival times using a smartphone app accessing Transport for London's Live Bus Arrivals database (Hardy et al, 2012). Among the data collected, the present study looks at the responses relating to the PerWT and the BudgWT. Again, to allow for comparisons, PerWT and BudgWT values are normalised by the AvWT; the distributions are presented in the form of histograms in Figure 6, and their characteristic values are given.

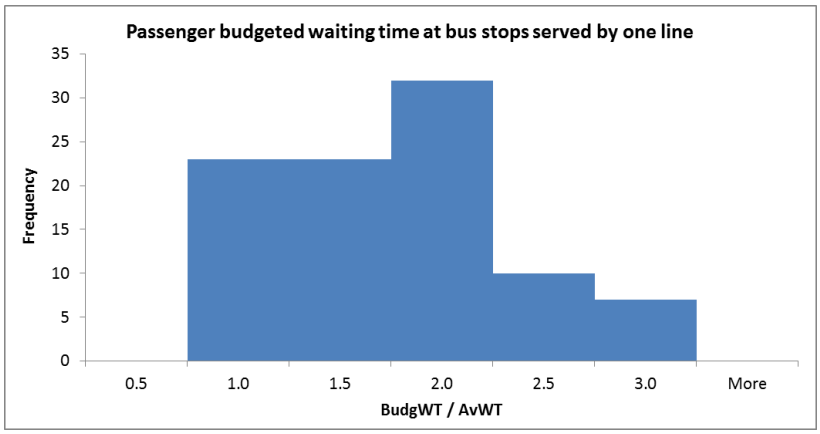
From initial interpretation and comparing the histograms and characteristic values of the PerWT and BudgWT distributions in Figures 6a and 6b with that of PlatWT in Figure 5a, it can be observed that PerWT has a higher mean than PlatWT, and that BudgWT has a higher mean than PerWT. Statistical significance tests confirm this finding (a *p*-value of 0.000 is obtained for both two-sample one-tailed t-tests carried out, thus rejecting the null hypotheses that PlatWT $\geq$ PerWT and that PerWT $\geq$ BudgWT respectively at the 0.05 level). The cumulative plots for the three distributions (Figure 6c) make this finding even more apparent, but also reveal some additional effects. Specifically, it can be clearly seen that the BudgWT is always longer than the PlatWT (which is a sensible finding, as passengers allow for longer waiting times than their usual ones to make sure that they do not miss their bus), but also that it gets closer to the PlatWT, as the latter increases. Additionally, it can be observed that the PerWT lso consistently overestimates the PlatWT, but that this effect is much stronger at lower PlatWT values and disappears at higher values (more than 1.5 times the AvWT).

With respect to the shapes of the histograms, it can be seen that both PerWT and BudgWT are left-skewed, but not as much as PlatWT. As concerns the cumulative plots, it can be observed that the PerWT and BudgWT curves are more S-shaped, as opposed to the PlatWT one, which, as expected, is consistently concave without any points of inflection. This suggests that appropriate standard distributions for PerWT and BudgWT could be the normal, the lognormal and the gamma distributions. Again, the fitting is performed by generating normal, lognormal and gamma cumulative plots on the basis of the same characteristic values, and comparing them with the actual PerWT and BudgWT distributions by means of K-S tests.



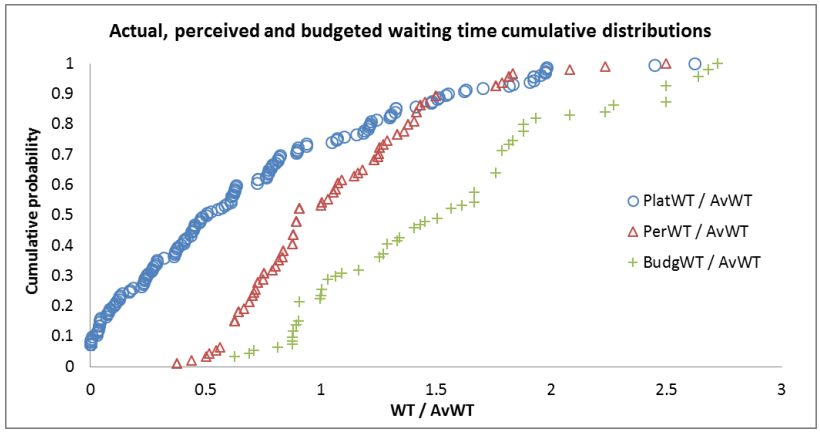
(a)

Mean = 1.047
Median = 0.908
Std. dev. = 0.425
Skewness = 0.956
Kurtosis = 0.833



(b)

Mean = 1.541
Median = 1.567
Std. dev. = 0.589
Skewness = 0.392
Kurtosis = -0.753



(c)

T-test results:

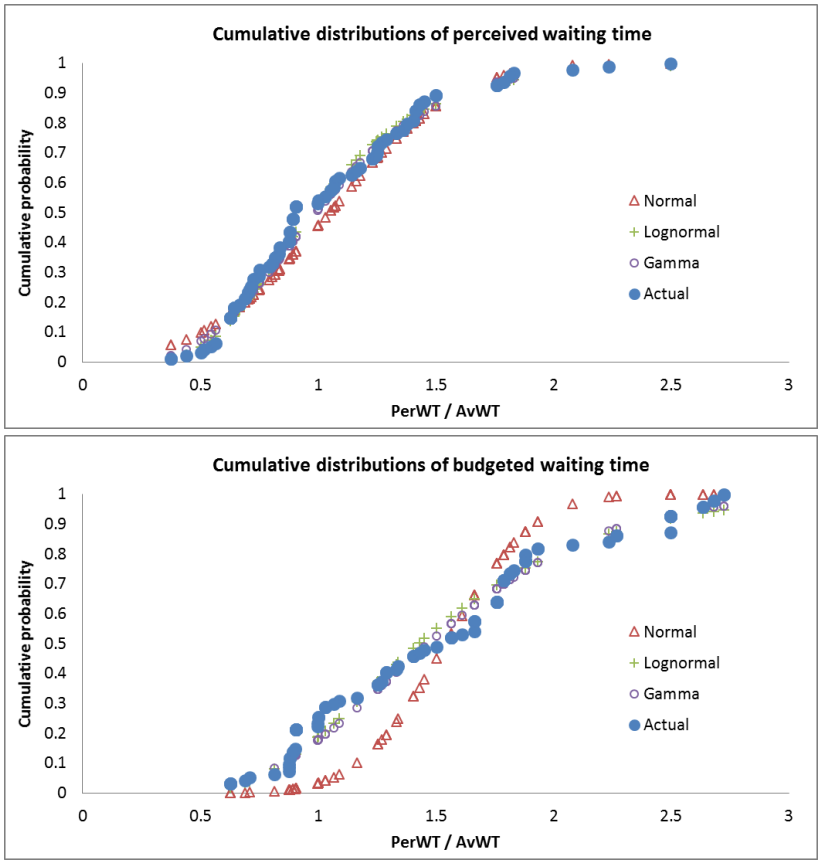
1. H_0: PlatWT $\geq$ PerWT:
   p = 0.000, rejected at 0.05
   ⇨ PlatWT < PerWT
2. H_0: PerWT $\geq$ BudgWT:
   p = 0.000, rejected at 0.05
   ⇨ PerWT < BudgWT

⇨ PlatWT < BudgWT

FIGURE 6: (a) Passenger perceived waiting time (PerWT) distribution at bus stops served by one line; (b) passenger budgeted waiting time (BudgWT) distribution at bus stops served by one line; (c) cumulative distribution plots of PlatWT, PerWT and BudgWT.

The cumulative PerWT and BudgWT distributions, and the corresponding normal, lognormal and gamma cumulative plots are shown in Figure 7. In an initial visual assessment of Figure 7a, it can be seen that that all three standard distributions with the same mean and standard deviation as the actual PerWT distribution are fairly good fits. This is confirmed by the K-S tests, which are accepted for the lognormal and gamma plots at the 0.05 level (the respective K-S statistics of 0.086 and 0.103 are smaller than the critical value of 0.140), as well as for the normal plot, though only at the 0.01 level (the K-S statistic of 0.150 is smaller than the relevant critical value of 0.167). Similarly, from Figure 7b it can be seen that the lognormal and gamma plots are good fits of the BudgWT distribution, and

the acceptance of the relevant K-S tests at the 0.05 level confirm this. On the other hand, the same cannot be said about the normal plot, a finding that is also verified by the K-S test, which is rejected at both the 0.05 and the 0.01 levels.



(a)

K-S test results:
$D^{0.05}_{crit}$ = 0.140
$D^{0.05}_{crit}$ = 0.167

Normal:
D = 0.150, accepted at level 0.01

Lognormal:
D = 0.086, accepted at level 0.05

Gamma:
D = 0.103, accepted at level 0.05

(b)

K-S test results:
$D^{0.05}_{crit}$ = 0.140
$D^{0.05}_{crit}$ = 0.167

Normal:
D = 0.246, rejected

Lognormal:
D = 0.107, accepted at level 0.05

Gamma:
D = 0.091, accepted at level 0.05

FIGURE 7: Cumulative distribution plots of (a) PerWT and (b) BudgWT compared with normal, lognormal and gamma plots.

5 CONCLUSIONS

The aim of the present study has been to investigate the topic of passenger waiting time at stops, using data from actual counts, measurements and interviews on part of London's bus network. Specifically, data from different stop types (differing by the number of lines serving the stop) have been collected and analysed by means of statistical tools and methods. The analysis has looked at three main aspects of passenger waiting behaviour: arrivals rate, actual waiting time, and perceived waiting time.

A number of useful conclusions can be drawn from the results of this study. First and foremost, looking at the findings of the interarrival times analysis, it appears that the arrivals rate at stops served by one or two lines is similar, and tends to intensify in the proximity of the bus departure times, such that passengers behave in a relatively schedule-dependent way and time their arrival at the stop to coincide with that of the bus as much as possible. Stops served by three or more lines, on the other hand, exhibit a much different pattern, where interarrivals are more uniformly spread and appear to be independent of the bus departures. This is an important finding, which identifies the need for further research into the waiting behaviour of passengers at stops served by a higher number of lines, and which has not been investigated here due to practical limitations.

Another conclusion that can be drawn here is from the analysis of the passenger actual

waiting time (PlatWT), and is that of the similarity of the waiting time probability distribution at stops served by one and two lines. From a practical point of view, this means that, provided that a normalisation by the bus service headway is done, the same distribution can be introduced in relevant planning and operations models, such as route finding and transit assignment. It has also been found that the current assumption of exponentially distributed waiting time is a valid one, but also that the gamma distribution can offer an adequately good fit.

Useful conclusions are also drawn from the analysis of the perceived and budgeted waiting times. It is shown that both perceived and budgeted waiting time can be approximated by a lognormal or a gamma distribution, and that for the perceived waiting time the normal distribution also offers an adequate fit. Most importantly, though, it is found that the perceived waiting time consistently overestimates the actual waiting time, but also that this effect is mainly noticeable at lower waiting time values, and tends to fade at higher ones. This challenges the current state-of-practice on the value of waiting time, where it is usually assumed that it is constant, and highlights the need of further research into this topic. For example, it may not be a good business case for a bus service operator to implement a measure or a system that reduces the waiting time from 5 to 3 minutes, as it is likely that it will not be noticed by the passenger; however, it may be worth implementing the same measure if it is to reduce the waiting time from 20 to 18 minutes.

While the present study has shed some light onto the relatively under-explored topic of public transport waiting time, research in this direction continues. Further work will concentrate on the more detailed investigation of the passenger waiting behaviour at bus stops served by three or more lines, particularly as concerns the actual and perceived waiting times. Moreover, a comprehensive investigation of the behaviour of passengers will be carried out, with a view of understanding the factors that lead to specific types of behaviours. A similar analysis is also foreseen for data collected from train and light rail stations, which may result in much different patterns. Overall, it is anticipated that the research will provide a framework, according to which waiting time can be comprehensively modelled and incorporated into transport planning models and algorithms.

ACKNOWLEDGEMENT

This research project is part of the EU COST action TU 1004 “Modelling public transport passenger flows in the era of Intelligent Transport Systems”.

REFERENCES

(1) Wardrop, JG. Some theoretical aspects of road traffic research. Proceedings of the Institution of Civil Engineers, vol. 1, 1952, pp. 325-378.
(2) Thomson, JM. The value of traffic management. Journal of Transport Economics and Policy, vol. 2, 1968, pp. 3-32.
(3) Jackson, WB and Jucker, JV. An empirical study of travel time variability and travel choice behaviour. Transportation Science, vol. 16, 1981, pp. 460-475.
(4) Adbel-Aty, A, Kitamura, R and Jovanis, PP. Investigating the effect of travel time variability on route choice using repeated measurement stated preference data. Transportation Research Record, vol. 1493, 1995, pp. 39-45.
(5) Lam, TC and Small, KA. The value of time and reliability: measurement from a value pricing experiment. Transportation Research E, vol. 37, 2001, pp. 231-251.
(6) Guo, S, Yu, L, Chen, X and Zhang, Y. Modelling waiting time for passengers transferring from rail to buses. Transportation Planning and Technology, vol. 34, 2011, pp. 785-809.

(7) Luethi, M, Weidmann, U and Nash, A. Passenger arrival rates at public transport stations. ETH Zurich report, 2006.
(8) Weber, W. Die Reisezeit der Fahrgäste öffentlicher Verkehrsmittel in Abhängigkeit von Bahnart und Raumlage. Technical University of Stuttgart, 1966.
(9) Seddon, PA and Day, MP. Bus passenger waiting times in Greater Manchester. Traffic Engineering and Control, vol. 15, 1974, pp. 442-445.
(10) Jolliffe, JK and Hutchinson, TP. A behavioural explanation of the association between bus and passenger arrivals at a bus stop. Transportation Science, vol. 9, 1974, pp. 248-282.
(11) Braendli, H and Mueller, H. Fahrplanabhängigkeit des Fahrgastzuflusses zu Haltestellen. ETH Zurich report no. 81/5, 1981.
(12) O'Flaherty, CA and Mancan, DO. Bus passenger waiting times in central areas. Traffic Engineering and Control, vol. 11, 1990, pp. 419-421.
(13) Gong, HB, Chen, XM, Yu L and Wu LJ. Application-oriented model of passenger waiting time based on bus departure time intervals. 93rd Transportation Research Board Annual Meeting, Washington, DC, 2014.
(14) Furth, PG, Hemily, B, Muller, THJ and Strathman, JG. Using archived AVL-APC data to improve transit performance and management. Transportation Research Board Transit Cooperative Research Program (TCRP) report no. 113, 2006.
(15) UK Office for National Statistics (ONS), 2011 Census, 2011.
(16) OpenStreetMap. Available at www.openstreetmap.org, 2014.
(17) Transport for London. Official website, 2014.
(18) Rai, A. Stop-Watch mobile application software (Version 1.1). Retrieved from http://itunes.apple.com, 2013.