Australasian Transport Research Forum 2018 Proceedings

30 October – 1 November, Darwin, Australia

Publication website: http://www.atrf.info

Implications of running time variability for passengers and operators – a case study using Bondi's 333 buses

Mathew Hounsell[1]

[1]Institute for Sustainable Futures | Transport + Cities | University of Technology Sydney

Level 10, Building 10, 235 Jones Street Ultimo NSW 2007 (PO Box 123)

Email for correspondence: mathew.hounsell@uts.edu.au

Abstract

This paper examines the impact of runtime variability on the ability of an operator to maintain on-time running and deliver the service promised to customers.

Transport for New South Wales (NSW) published collected observations on operations of buses in NSW. From these observation the inbound running times from the 333-route travelling from North Bondi to Sydney CBD were computed and analysed to determine if the route was reliably delivering the promised running times to passengers.

There is a deficit of andragogic materials on public transport planning, especially on tasks that both cover the operators and customers perspective. This paper also aims to provides a method for undergraduates to follow in studying other bus routes with the same methods.

1. Introduction

Transport planning and operations management are a constant balancing act between costs and benefits. It is rational decisions about how to deliver services and allocate the city's limited land and capital require quality data. With the rollout of the Opal smart-card electronic ticketing monopoly, NSW decision makers have access to a rich source of data on vehicles and customer behaviour. This rich dataset has allowed researchers to re-examine the current NSW metropolitan network and identify key corridors where improvements would have significant benefits.

The Opal system is constant recording information on all bus services in the NSW Greater Metropolitan Area (GMA). This research was a case study to determine if there were a set of simple metrics the NSW transport operators could measure with the existing system to facilitate constant service improvements and thus improve farebox recovery ratios. The research also aimed to determine if there were simple metrics computable from public data that allowed passengers and the community to argue for reasonable service improvements.

Opal analysis shows that the most popular bus routes in NSW carry thousands of passengers per day. The 333 is a limited stops version of the longer all-stop 380 route that runs along Bondi Rd and Oxford St. These routes, together with the 381, were measured with a combined patronage of 25,000 embarkations on an average workday in November 2016. As such this is one of the most important surface transport corridors in the nation.

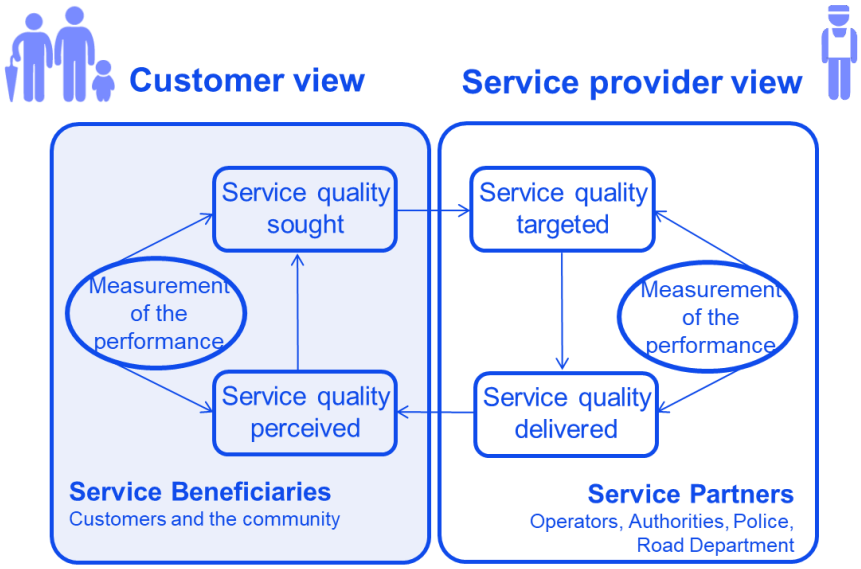
Public transport is vital to the operation of the urban agglomerations that power our national economy. Inefficient public transport is not just unnecessarily expensive, but it is a drag on the nations productivity and prevents Australia from achieving our national goals.

Table 1: Number of embarkations per day for Top 10 bus route (Tue 2016-11-08 to Thu 2016-11-10)

Rank (Mean)	Route	Minimum	Maximum	Mean	95% CI	Standard Deviation	Samples
1	400	15,050	16,180	15,430	750.0	649.5	3
2	343	13,750	14,330	13,987	351.4	304.4	3
3	389	13,140	13,470	13,307	190.6	165.0	3
4	440	11,860	12,830	12,427	583.4	505.2	3
5	380	11,060	12,600	11,997	949.5	822.3	3
6	M52	11,230	11,370	11,280	90.2	78.1	3
7	333	10,270	11,740	11,057	855.0	740.4	3
8	M20	10,220	10,500	10,340	166.5	144.2	3
9	T80	9,460	10,180	9,720	461.3	399.5	3
10	M30	8,170	8,680	8,457	301.2	260.8	3

'We place the customer at the centre of everything we do' has been adopted by Transport for NSW (TfNSW) as their key value. To deliver that value requires researchers and operators to perceive the world from the customers view. From the customers view running time variability increases travel times, reduces reliability, and can increase prices. Significant running time variability leads to customer dissatisfaction as the perceived service quality does not meet the service quality customers sought.

Figure 1: Service Quality Loop (EN13816)



Running times variability will also increase in waiting time for customers, then increase dwell times, and then cause bus bunching; the specifics are outside the scope of this paper.

Running time variability was chosen as the metric examined for this study as it impacts service quality and can be computed using Excel from the public dataset and it is not widely measured nor reported in NSW.

Van Oort (2014) is a survey of international measurements of service reliability is focused on schedule adherence not assessing runtime variability; as is much of the literature for example *Strathman (1993)* which focuses on headway variations. However, the van Oort survey does identify Co-efficient of Variation and Headway Adherence which are used by *Mazloumi (2008)* and *Mazloumi (2009). Currie (2012)* considers a range of measures not directly measuring runtime and again Headway Adherence and Coefficient of Variation. Currie also identified the interesting Reliability Buffer Index. *Bates (2001)* and many others are econometric papers focused on Value of Travel Time. *Kieu (2013)* was also interesting but their method of Vehicle to Vehicle variability was too complex for the community.

Headway adherence does not directly measure running time variability is not applicable. Further research will be required into the Reliability Buffer Index measure.

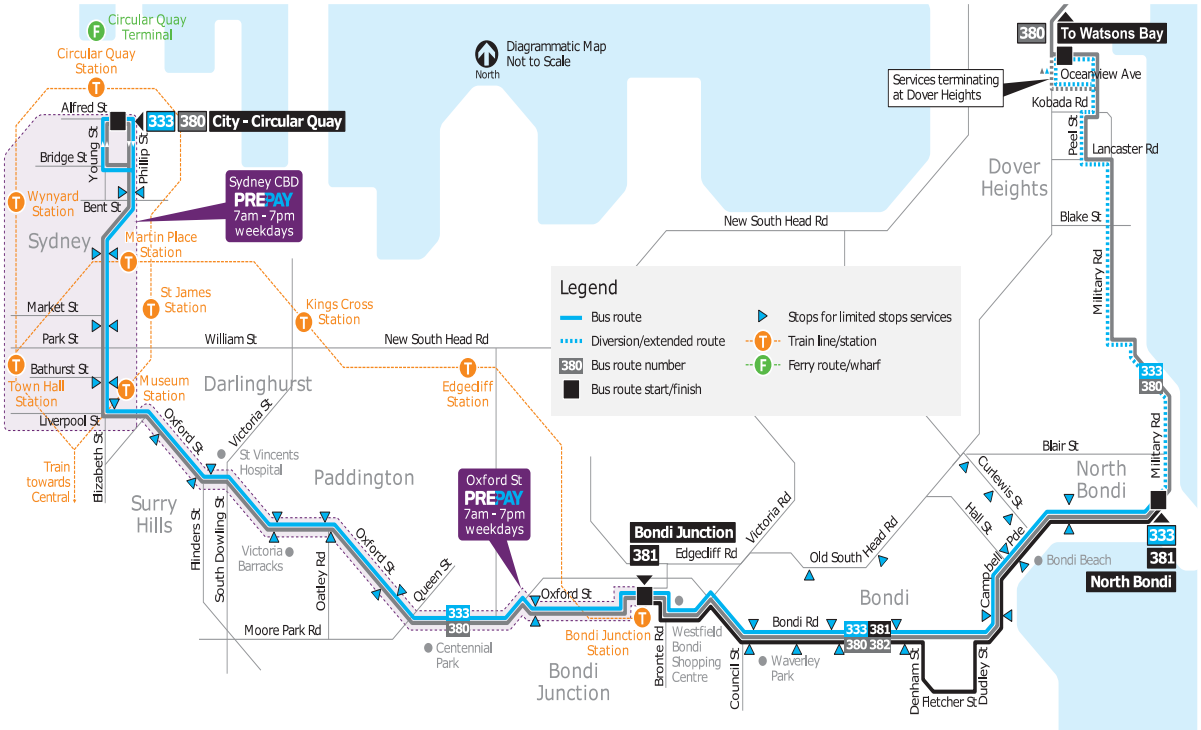
Working with students and the community has demonstrated they need for simple measures calculated relatively easily in tools like Excel or LibreOffice and in simple units like minutes. Unfortunately, complex calculations quickly overwhelm the general public. Meanwhile Coefficient of Variation is useful for determining relative variability but is not useful in determining how late a service can be. As such this case study determined mean and standard deviation are sufficient as a starting point.

The case study examined the 380-route and 333-route from North Bondi to Sydney using a **public dataset** of three weeks in August, November & December 2016. The **public dataset** included timetabled and actual stop arrival times and an approximation of vehicle loading.

This paper discusses the 333 because the running time variability of the limited stop service is less impacted by stop skipping, dwell-times, and lane re-entry delays. The paper considers the schedule within the dataset, then the observed variability of running times.

1.1 About the 333

Opal charges fares based on the distance travelled and most (93%) of passenger Tap-On upon boarding a bus and to Tap-Off when alighting that bus. This new ticketing system allows transport operators to empirically measure the number of passengers on each bus route.



The limited stop 333-route can be broken up into four nearly equal sections.

1. Up the Hill: from Campbell Parade [tram] Terminus (Stop #200634) at North Bondi to the first definite stop after the beach is Bondi Rd near Denham St (#202659); ~2km with 5 stops inclusive, through 7 signalised intersections
2. To Bondi Junction Interchange: the next section along Bondi Rd to Grafton St outside of the Bondi Junction Station & Interchange (#202268); ~ 2km with 5 stops inclusive; through 10 signalised intersections.
3. To Taylor Square: the third section to Oxford St opposite Palmer St (#201051) in Darlinghurst; ~3km with 5 stops inclusive; through 18 signalised intersections.
4. To CBD: the fourth section to Elizabeth St near Martin Place Station (#2000421) in the Sydney CBD; ~2km with 5 stops inclusive; through 12 signalised intersections.

All signalised intersections are controlled by SCATS.

The buses used are either a standard STA three-door articulated bus with an approximate capacity of 115 persons (71% of buses) or a standard STA two-door bus with approximate capacity of 60 persons (29% of buses). Fare collection pre-paid (Opal) collected on the vehicle. NSW procedures prohibit all-door boarding requiring a single point of entry to ensure revenue protection and this regulation is usually enforced by drivers.

Opal uses highly accurate clocks and augmented GPS to determine the buses geospatial coordinates for recording the vehicles arrival time at each bus stop. The system is also used to provide real-time bus position data to Transport for NSW. Transport for NSW used recording of bus locations and passenger load estimates to create the "Bus Occupancy - Aug 2016 to Jan 2017" **public dataset** for their Open Data portal.

The paper considers the operator to be Transport for NSW (TfNSW), which is the overseeing agency of the Transport cluster for the NSW government. TfNSW has oversight over bus timetables, networks and service levels; and oversight over bus stop locations and the bus stop removal program. The agency also has oversight over Roads and Maritime Services (RMS), the operator that controls bus lanes. TfNSW also has oversight over the Transport Management Centre and RMS who manage the signalised intersections. State Transit Authority provides the drivers and maintains the buses under contract to TfNSW.

2 Methodology

This section outlines a general methodology used to process the data used for the analysis discussed below. This outline omits specific details and unnecessary processing for brevity and clarity, additional detail can be found in "Data Processing' at the end of the paper.

This study demonstrated that processing the public "Bus Occupancy" dataset with a common spreadsheet program could assist customers (and undergraduates) in assessing the performance of important bus corridors. The 333-route was chosen because it was limited stops and the reduced number of stops would reduce the impact of dwell times on the service times. The paper does not break the analysis into a consideration of individual components such as dwell-time as that is extraneous to the customers perception.

When calculating the number of vehicles needed for a service the results will be a rational (fractional) number. Since it is impossible to deploy ¾ of a bus some rational numbers will need to be rounded to integer (whole) numbers. To clarify the mathematical notation used, the ceiling operation ⌈ ⌉ rounds-up and finds the smallest integer greater than or equal to a given rational number and the floor operation ⌊ ⌋ rounds-down and finds the largest integer

less than or equal to a given rational number. Ceiling is written in *Vuchic (2005)* as $[\quad]^{+}$. For example, $\lfloor \Pi \rfloor = \lfloor 3.1459 \rfloor = 3$ and $\lceil \Pi \rceil = 4$ and $\lfloor 42 \rfloor = \lceil 42 \rceil = 42$.

2.2 Loading the dataset

2.2.1 Route variations

NSW operates many varying routes badged under a single route number. The 333 has several long route variations that extend beyond Bondi Beach. Some variations terminate at Martin Place while others continue to Circular Quay. Analysing the variations was required before developing the runtime extraction strategy. It was determined that most common route variations were the beach bound 333-30 (43.0%) and city bound 333-42 (42.7%).

2.2.2 Extracting the actual end-to-end running times

From the primary table, an end-to-end running times table was derived by keeping only records relating to the first and last stop of the 333-30. The table was transformed with each new record containing the actual arrival time of the bus at the first and last stops; with a record for every trip on every date. Then end-to-end running times were calculated by subtracting the actual arrival time at the first stop from that at the last stop. Any records where the actual time was missing, were excluded from the table of running times.

Table 2 below contains the statistic derived from these running times calculated for the 333-30 (outbound) during the weeks recorded in August and November 2016. The December-January week was excluded because it covered the lower demand period between xmas and new year. Then the procedure was repeated to derive the same set of end-to-end running time statistics for the 333-42 (inbound) and those are listed in Table 3 below.

These actual running time statistics show a significant range of running times for both outbound and inbound services. The inbound services have a large range (75 minutes), a larger standard deviation (7.6 minutes) and a larger variance in the running times.

Table 2: Running time statistics for 333-30 outbound from Sydney to Bondi in Aug & Nov 2016

333-30	Samples	Min (minutes)	Max	Mean	95% C.I.	Std Deviation	Variance
Monday	129	30	54	40.4	±0.7	3.9	15.3
Tuesday	136	32	60	42.1	±0.9	5.1	26.5
Wednesday	135	31	60	41.9	±1.0	6.0	36.0
Thursday	145	32	52	41.2	±0.8	4.6	21.5
Friday	132	33	54	41.5	±0.8	4.6	21.5
Saturday	115	32	56	43.9	±1.1	5.7	32.3
Sunday	70	30	47	38.4	±1.0	4.3	18.3
All Days	**862**	**30**	**60**	**41.5**	**±0.4**	**5.1**	**26.5**

Table 3: Running time statistics for 333-42 inbound from Bondi to Sydney in Aug & Nov 2016

333-42	Samples	Min (minutes)	Max	Mean	95% C.I.	Std Deviation	Variance
Monday	132	29	72	45.0	±1.2	6.7	45.4
Tuesday	134	31	101	48.3	±1.8	10.2	103.7
Wednesday	135	30	69	47.5	±1.4	7.9	62.4
Thursday	144	32	64	47.4	±1.1	6.9	47.3

Friday	136	31	59	47.0	±1.0	6.0	36.1
Saturday	118	31	61	47.7	±1.1	6.2	39.0
Sunday	65	26	61	42.4	±1.9	7.7	59.8
All Days	**864**	**26**	**101**	**46.8**	**±0.5**	**7.6**	**58.3**

2.2.3 Timing Locations

Due to the greater variation, it was decided that further analysis would focus on inbound runtimes. To increase the number of samples a new runtime table was generated for all inbound services. However, Campbell Parade Tram Terminus (Stop #202634) at North Bondi was treated as the first stop and Elizabeth St near Martin Place Station (#2000421) in the Sydney was treated as the last stop. This common inbound 333 route from North Bondi to Sydney has 20 stops and is approximately 10km in length between #200634 and #2000421. Running times for the analysis were calculated by subtracting the actual arrival time at the Martin Place stop from the actual arrival time at the Campbell Parade stop.

2.2.4 Sample Size

The runtime tables use the actual time of arrival reported for each stop. All journeys without a recorded actual time were excluded. There were 1,716 applicable 333 services scheduled in the dataset; however only 1,267 contained actual arrival times.

Four (4) running times over 180 minutes were excluded because they were deemed outliers and possibly represented invalid data; those runtimes ranged from 219 to 221 minutes. There were only 21 samples with a runtime exceeding 60 minutes. During the AM peak on Tuesday 22nd of November 2016 there are a continuous range of long running times up to 97 minutes (101 to circular quay) as can be seen in Figure 2 below; so those values were included. However, there is over 200 minutes between the fifth longest runtime and the fourth longest; as such they were excluded for being unusual.

In Table 4, Table 5, and Table 6 the number of inbound 333 running time observations between the timing locations is broken down per month, day of the week, and hour of the day. The hour of a service was derived from its actual departure time at North Bondi. The dataset provided three separate weeks of observations running from Monday to Sunday.

Table 4: Number of 333 observations used in analysis per month

Month	*August 2016*	*November 2016*	*December 2016 + January 2017*
Coverage	*Mon Aug 8th to Sun 14th*	*Mon Nov 21st to Sun 27th*	*Mon Dec 26th to Sun Jan 1st*
Count	487	489	291

Table 5: Number of 333 observations used in analysis per day

Day	*Monday*	*Tuesday*	*Wednesday*	*Thursday*	*Friday*	*Saturday*	*Sunday*
Count	167	210	158	246	235	145	106

Table 6: Number of 333 observations used in analysis per hour of the day

Hour	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*
Count	20	38	127	102	81	75	79	79	78	87	76	91	75	49	54	50	46	43	17

Of the samples used there are many outliers that can be seen in Figure 2 and Figure 3 below. As can be seen in the plots the outliers increase the mean slightly above the median.

It would be useful to know what caused the outliers. However, there is no publicly available dataset recording traffic accidents on NSW roads, nor is there a publicly available dataset recording the delays on taxpayer funded public transport services. Breakdowns occur regularly and unfortunately so do traffic accidents. As the responsible agency, TfNSW should already mitigate the risk of blockages on one of the nation's most important bus corridors.

3 Timetabled services

To understand the operators promised service delivery the schedule for the 333 was extracted from the *TIMETABLE_TIME* column in the dataset. The promised services between North Bondi (#200634) and Sydney (#2000421) are summarised in Table 7 and Table 8 below. The 333 operates one of three timetables: either workday, Saturday, or Sunday/Public Holiday.

The timetable summary in Table 7 below represents the average promised inbound running time for the 333 in August and November 2016 for each hour of the day. Table 7 shows that the operator expects that the 333-route could achieve a 31-minute inbound running time but scheduled a slower 37- to 39-minute runtime during the daytime on workdays and on the weekends an additional 13 minutes was added to schedule for daytime runtimes.

Table 7: Timetabled running time 333 inbound from North Bondi to Sydney CBD in Aug. & Nov. 2016

Hour	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*
Workday	31	32	39	39	38	37	37	37	37	36	37	39	38	37	33	31	31	31	31
Saturday		29	29	31	37	44	44	44	44	44	44	41	36	33	33	30	31	31	31
Sunday			29	32	38	44	44	44	44	44	44	44	41	36	30	31	31	31	31

The average line capacity, as discussed in *Vuchic (2005)*, was calculated by totalling the *VEHICLE_CAPACITY* scheduled in the primary table for each day type and hour of the day. The 333 provides 2,365 passenger spaces inbound during the workday morning peak.

Table 8: Timetabled spaces per hour from North Bondi to CBD in August & November 2016

Hour	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*
Workday	230	345	1,385	980	690	690	690	690	690	805
Saturday		175	230	345	690	690	690	805	690	690
Sunday			230	345	460	405	460	405	460	350

Hour	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*
Workday	750	690	575	460	345	345	345	345	115
Saturday	690	575	460	345	290	345	405	345	115
Sunday	460	575	405	230	345	416	345	230	115

4 Running time variability of the 333

4.1 Initial plots

Basic statistical plots were used to determine if the 333 running times showed significant variability and if there were particular periods of variability that were worth analysing.

The graph Figure 2 below shows the distribution of observed 333 inbound running times. There is significant variability with normal ranges from 26 to 62 minutes but with many outliers exceeding 70 minutes. Remember during the study period that approximately 450 services (26.2%) did not report actual runtimes.

Graphing the reported running times per month of observation Figure 2 below shows that during the xmas to new year holiday period there was less variability in running times. The summer holidays are a popular period for passengers returning from Bondi Beach.

Graphing the distribution of running times per weekday in Figure 3 below shows workday running times have greater variability with higher maximum running times and higher average running times. Figure 3 also shows that over 75% of services exceed the expected 37-minute workday scheduled running time.

Figure 3 shows the weekend has lower maximum running times, and a smaller range of running times. Sunday running times were usually faster than other days of the week.

The histogram plot in Figure 4 below shows that most inbound running times were longer than 40 minutes, with over 784 (61.9%) exceeding the achievable minimum runtime by over 10 minutes.

Figure 2: Whisker plot of runtimes for 333 inbound from North Bondi to Sydney CBD per month

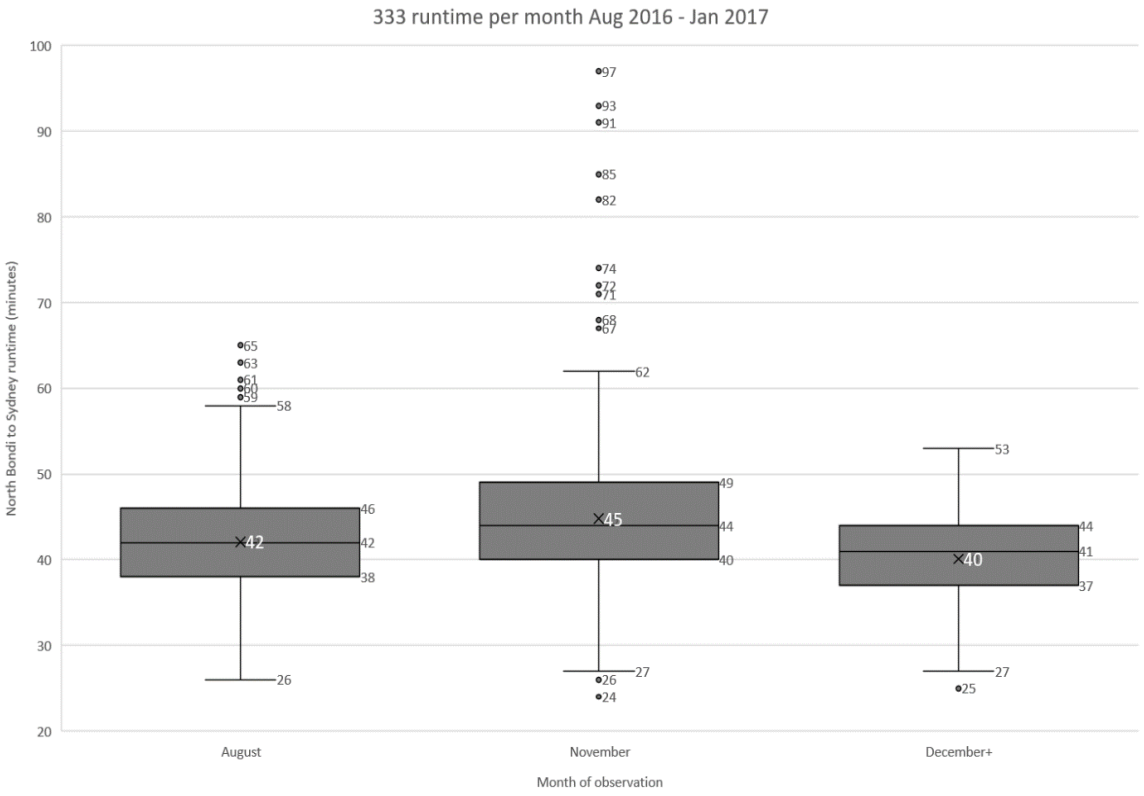


Figure 3: Whisker plot of runtimes for 333 inbound from North Bondi to Sydney CBD per weekday

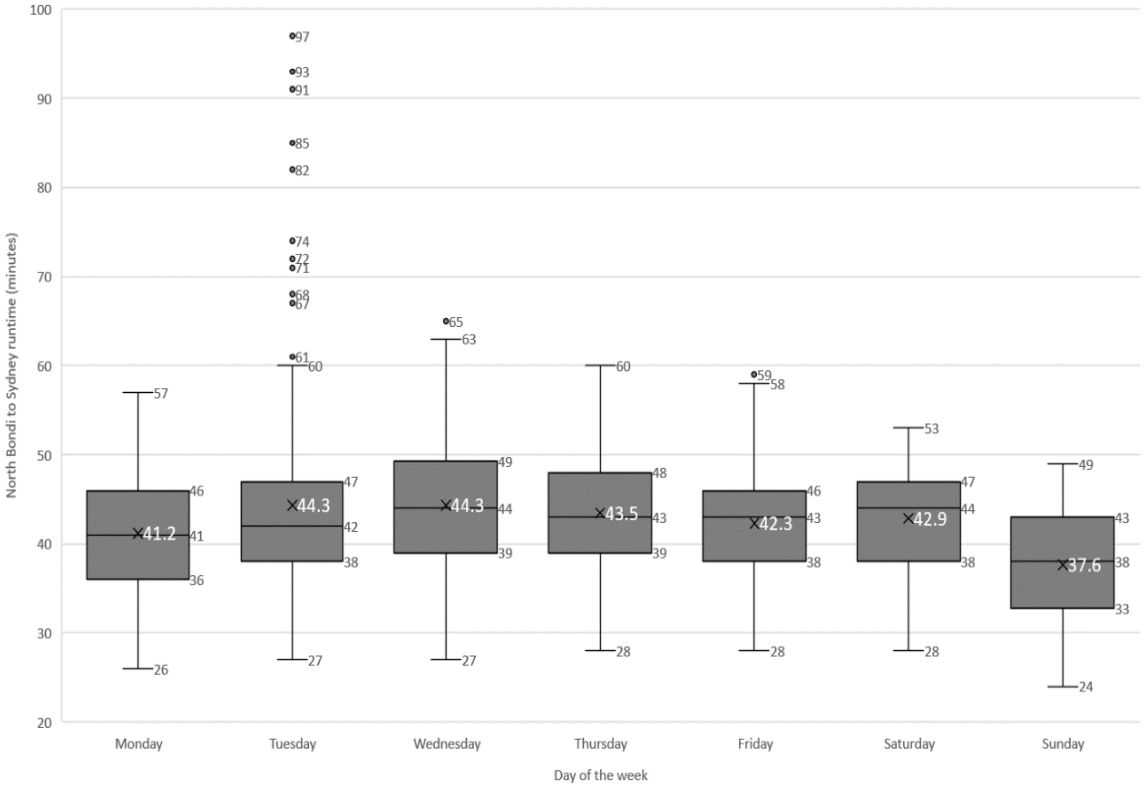
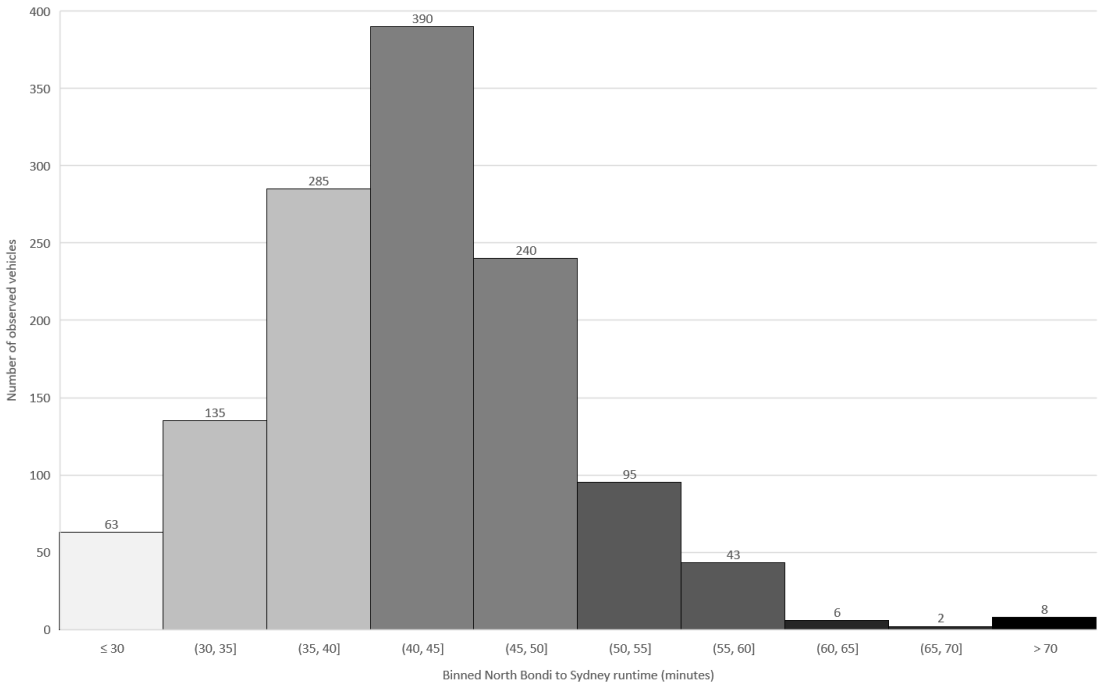


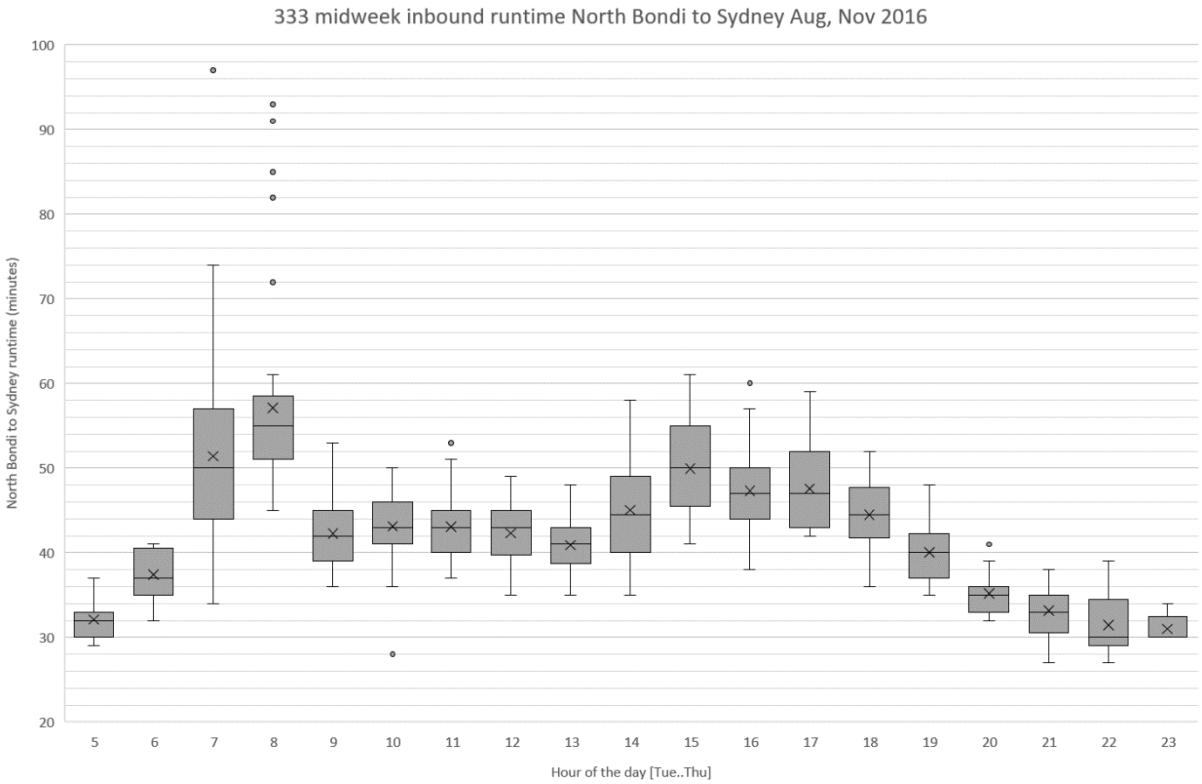
Figure 4: Histogram of runtimes for 333 inbound from North Bondi to Sydney CBD



4.2 Focusing the analysis

To reduce the impact of passenger behaviour on variability, the analysis was further narrowed to August and November 2016. The distribution of inbound runtime for Tuesday, Wednesday, and Thursday (TWT) in August & November 2016 is plotted in Figure 5 below. Figure 5 shows that the inbound runtime varies substantially during each hour of a workday.

Figure 5: Whisker plot of runtimes from North Bondi to Sydney CBD per hour TWT in Aug & Nov 2016



Passenger numbers and environmental conditions will impact the dwell times and thus the running times; see *Vuchic 2005*. Passenger numbers and traffic congestion are known to vary between workdays and weekends, and across the hours of the day.

Over all observed workdays (Monday-Friday) the mean inbound running time was 43.9 minutes (±1.5 minutes) with a standard deviation of 8.3 minutes. In contrast the mean Sunday running time was 37.6 minutes (±0.5 minutes) with a standard deviation of 6,4 minutes.

The inbound running times for services that hour starting 07:00 varied from 31 to 74 minutes, with outliers. The timetabled runtime for that period was 39 minutes, the average observed inbound runtime was 49.5 minutes. In addition, that hour of the day saw a standard deviation of 9.5 minutes, further indicating a substantial degree of variability; see Table 9 below.

Table 9: Observed runtime 333 inbound from North Bondi to Sydney CBD in Aug, Nov 2016

Hour	Workday				Saturday				Sunday			
	Min	Avg	Max	StdDev	Min	Avg	Max	StdDev	Min	Avg	Max	StdDev
All Hours	26	43.9	97	8.3	28	43.4	53	5.8	24	37.6	49	6.4
5	29	32.5	37	2.2								
6	32	36.3	41	2.8	28	30.5	33	3.5				
7	**31**	**49.5**	**97**	**9.5**	32	35.3	37	2.9	24	26.5	29	3.5
8	**43**	**54.3**	**93**	**9.8**	32	34.5	37	2.4	30	31.3	33	1.5
9	36	41.8	53	3.5	35	39.0	43	2.8	34	35.0	37	1.4
10	28	42.8	50	4.1	39	42.0	48	3.0	38	43.0	46	4.4
11	37	43.2	53	4.0	38	45.5	51	4.1	44	45.0	46	1.4

12	35	41.8	49	3.2	39	45.8	52	3.3	44	44.0	44	
13	28	40.9	48	3.5	45	48.4	53	2.7	45	46.5	48	1.3
14	35	44.9	58	5.4	44	48.2	53	3.3	43	43.0	43	0.0
15	41	49.2	61	5.3	41	45.9	53	3.1	37	41.0	43	3.5
16	38	47.0	60	4.6	44	48.3	53	2.6	39	44.6	49	3.8
17	40	46.6	59	4.4	47	48.6	51	1.5	36	39.2	42	2.4
18	36	43.5	52	4.2	37	43.2	46	3.6	37	39.4	43	2.3
19	34	39.7	48	4.0	33	40.0	44	5.0	31	34.3	39	3.1
20	32	35.3	41	2.4	30	36.8	43	5.3	29	32.4	38	3.5
21	27	33.2	40	3.0	31	36.3	43	6.1	26	29.8	32	2.9
22	26	31.8	39	3.6	36	39.0	45	3.9	30	31.8	33	1.3
23	28	30.9	34	2.1	30	30.0	30		27	27.5	28	0.7

Due to page width Table 10 below contains the number of samples per day type per hour.

In August and November, the mean observed inbound 333 running times were longer than the scheduled runtimes on workdays and Saturdays. However, the mean observed Sunday running times were often near to or less than the scheduled running time. In addition, on Sunday the maximum runtimes did not exceed the scheduled runtimes by much, especially compared to Workday and Saturday runtimes. This might indicate that longer 44-minute inbound runtimes for the 333 on Sunday are not required during August and November.

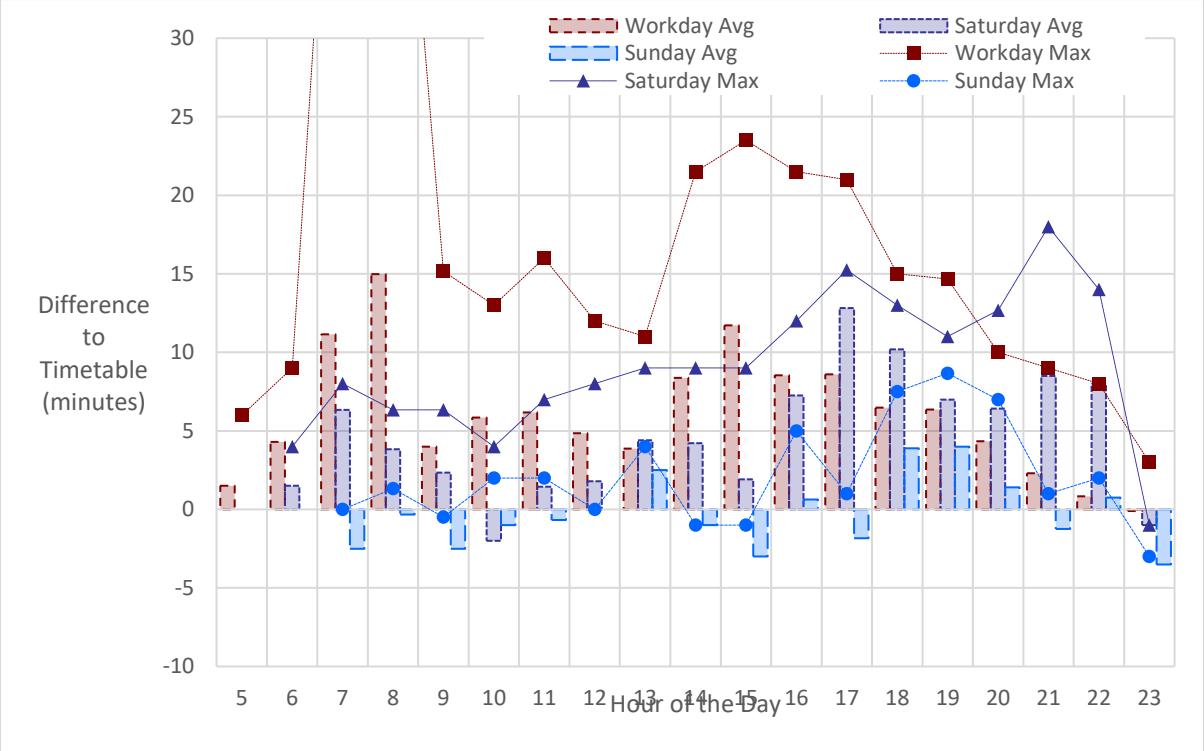
Table 10: Difference to timetable of inbound 333 running time in Aug, Nov 2016 per hour of the day

Hour	Workday				Saturday				Sunday			
	Samples	Min	Mean	Max	Samples	Min	Mean	Max	Samples	Min	Mean	Max
5	16	-2	2	6								
6	27	0	4	9	2	-1	2	4				
7	104	-8	11	58	3	3	6	8	2	-5	-3	0
8	73	4	15	54	4	1	4	6	3	-2	0	1
9	48	-2	4	15	11	-2	2	6	4	-4	-3	-1
10	45	-9	6	13	9	-5	-2	4	3	-6	-1	2
11	51	0	6	16	11	-6	1	7	3	-4	-1	2
12	53	-2	5	12	10	-5	2	8	1	0	0	0
13	53	-9	4	11	10	1	4	9	4	1	3	4
14	55	-1	8	22	9	0	4	9	2	-1	-1	-1
15	40	4	12	24	11	-3	2	9	3	-7	-3	-1
16	50	-1	9	22	8	3	7	12	8	-5	1	5
17	41	2	9	21	7	11	13	15	6	-5	-2	1
18	25	-1	6	15	5	4	10	13	5	2	4	8
19	29	1	6	15	4	0	7	11	6	1	4	9

20	26	1	4	10	4	0	6	13	5	-2	1	7
21	29	-4	2	9	4	0	9	18	4	-5	-1	1
22	23	-5	1	8	5	5	8	14	4	-1	1	2
23	9	-3	0	3	1	-1	-1	-1	2	-4	-4	-3

From a customer's perspective their perception of the 333 is that the service delivered is often slower than the service promised; see Figure 1: Service Quality Loop (EN13816). On an average workday service at 08:00 the passengers can expect that the service delivered will be on average 15 minutes slower than the service promised. The passengers can also expect that on Saturday the average service delivered will be more than five-minutes slower than the service promised even though the promised services are slower on Saturdays.

Figure 6:Plot of difference to timetable of inbound 333 running time in Aug, Nov 2016 per hour of the day



At the time of the study there was no rich public dataset available to analyse the road vehicle travel times in this corridor. Future research could investigate road vehicle travel times. The Bus Occupancy dataset contains a record for every buses arrival at every stop and can be used to calculate the buses travel speed and identify most unreliable route segments; see below.

The nation should not have to accept its productivity being limited and its expenditure being increased because important surface public transport is slow, unreliable, and delivered inefficiently. 1,385 passenger spaces are delivered by just the 333 inbound after 7 on a workday; let alone by the 380, 381, 382, etc. If travel times have value, then the running time and reliability of high capacity public transport on major corridors is important by itself.

4.3 Distribution

In theory the timetable should provide a 'floor' to the travel times, creating a right-skewed distribution with a narrow standard deviation. However as seen in the dataset and Table 10 above, the real-world operations in NSW running times do not have that clear floor. On some NSW routes many services arrive at the end of the line early due to significant schedule

overhead. On others there is a significant leftward skew to the distribution as the services are often very late by the end of the line.

Table 11: Skew of 333 inbound running times per day type

Workday	Saturday	Sunday
1.37	-0.56	-0.04

Table 12: Skew of Workday inbound 333 running times per hour of the day

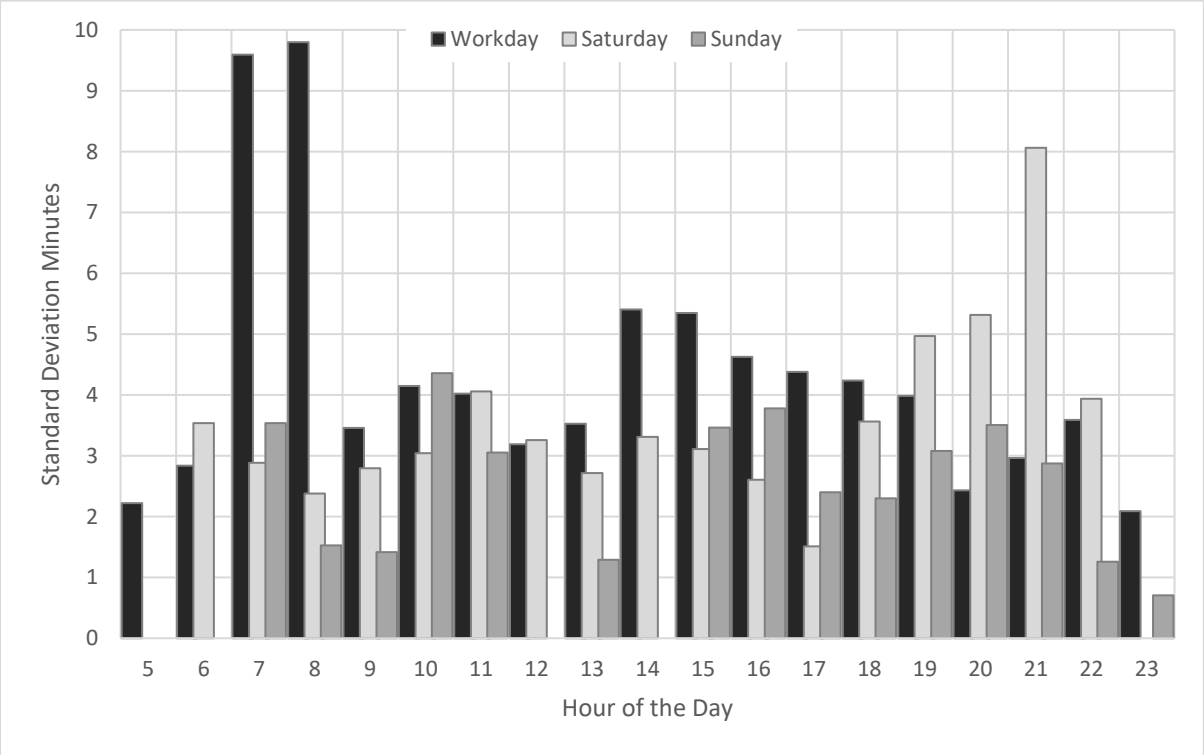
6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0.31	1.43	2.29	1.19	-0.84	0.40	0.12	-0.55	0.38	0.40	0.49	0.87	0.21	0.54	0.92

Table 13: Skew of Saturday inbound 333 running times per hour of the day

7	8	9	10	11	12	13	14	15	16	17
0.00	0.07	0.99	-0.23	-0.28	0.55	0.42	0.98	0.35	0.62	-1.93

As described in the literature including Mazloumi (2009), Standard Deviation is a simple measure of running time variability. There would be no variance of a service operating with perfectly reliable running times and then the standard deviation would be zero. The higher the standard deviation the more variable the runtimes. In a normal distribution 68% of values would be expected to be within ±1 standard deviation from the mean; and 95% of values would be expected within ±2 standard deviations. In theory, if the running times were under control the distribution would be a narrow slightly positively skewed distribution.

Figure 7: Standard deviation of inbound 333 running time in Aug, Nov 2016 per hour of the day



Standard deviation is a well-known measure of variance and as seen here and in section six below can be used to identify times and locations were running times are variable. Since running time standard deviation is measured in minutes it allows the community to compare a variability in units they are familiar with and get a sense of the scale of variability. Although Coefficient of Variation may be a simpler measure for comparing different routes.

Is the standard deviation a useful if simplistic predictor? For instance, if an app developer wanted to provide an estimate of the maximum running time to a passenger using the mean plus twice the standard deviation for a given hour and day type. The difference between the maximum observed value and the mean plus two standard deviations is shown in Table 14.

Table 14: Expected range of inbound 333 running time in Aug, Nov 2016 per hour of the day

	Workday				Saturday			
	Samples	Max	Mean + 2 StdDev	Diff	Samples	Max	Mean + 2 StdDev	Diff
All Hours	797	97	60.4	-36.6	118	53	55	2
5	16	37	36.9	-0.1				
6	27	41	42	1	2	33	37.6	4.6
7	104	97	68.9	-28.1	3	37	41.1	4.1
8	73	93	73.9	-19.1	4	37	39.3	2.3
9	48	53	48.8	-4.2	11	43	44.6	1.6
10	45	50	51.1	1.1	9	48	48.1	0.1
11	51	53	51.2	-1.8	11	51	53.6	2.6
12	53	49	48.2	-0.8	10	52	52.3	0.3
13	53	48	47.9	-0.1	10	53	53.8	0.8
14	55	58	55.7	-2.3	9	53	54.8	1.8
15	40	61	59.9	-1.1	11	53	52.1	-0.9
16	50	60	56.3	-3.7	8	53	53.5	0.5
17	41	59	55.4	-3.6	7	51	51.6	0.6
18	25	52	51.9	-0.1	5	46	50.3	4.3
19	29	48	47.7	-0.3	4	44	49.9	5.9
20	26	41	40.2	-0.8	4	43	47.4	4.4
21	29	40	39.2	-0.8	4	43	55.6	12.6

The table above shows that with sufficient samples the mean plus two standard deviations ($\bar{e}+2\sigma_e$) appears to estimate a reasonable value for maximum travel time; plus-or-minus 4 minutes. However, if the running time in that hour is very large then the standard deviation will still underestimate the maximum observed.

In addition, if the scheduled running times have significant overhead for late running then based on NSW operation patterns the observed values for the day type and hour will have a negative skew and ($\bar{e}+2\sigma_e$) will slightly overestimate. If there is a positive skew and minimal schedule overhead, then ($\bar{e}+2\sigma_e$) will underestimate. Additional research is required to prove if these results hold for more bus routes in NSW.

5 What section of the 3333-route causes the variability.

In order to determine what section of the route had the most variable running times, the route was broken up into four nearly equal sections; see above. The distribution of running times on each section are plotted in Figure 8 below. *Note that the y-axis has been trimmed and the several sections on Bondi Rd exceeding 40 minutes in duration are not visible.* It is

interesting to note that half the trips of the 333 took over 14 minutes to travel less than 2km from Taylor Square to Martin Place; which is an average speed of 7.7 km/h.

Figure 8: Whisker plot of the inbound running times as four sections

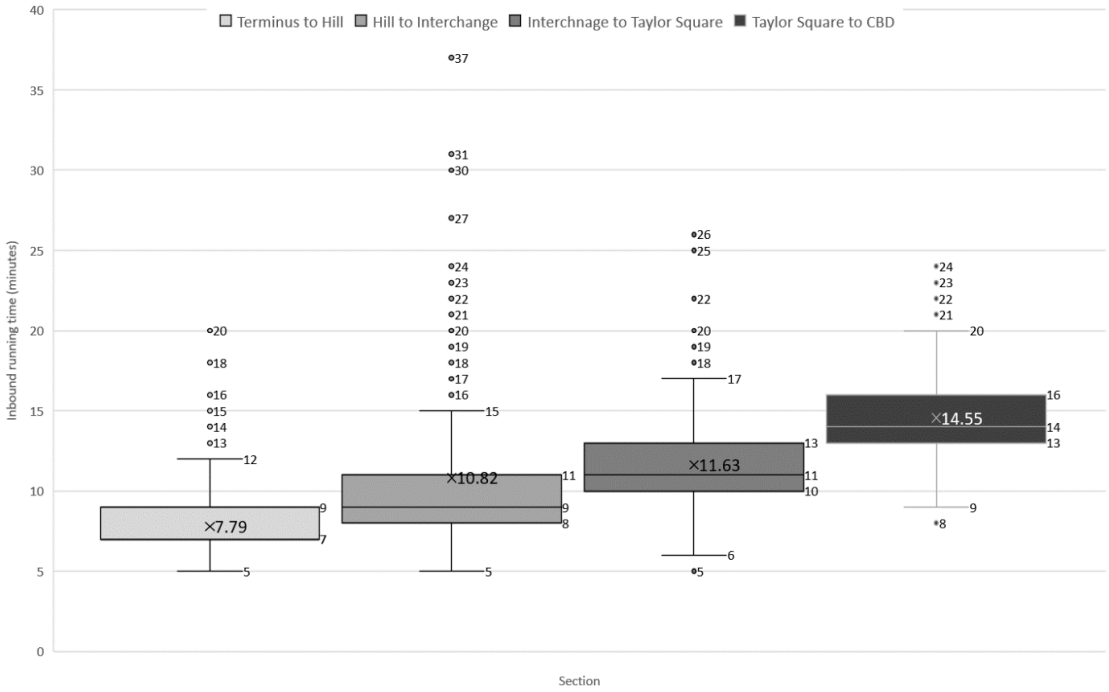


Figure 9, Figure 10, and **Error! Reference source not found.** below show the mean inbound runtimes for the 333 per hour of the day on Tuesday Wednesday Thursday (TWT) in August and November 2016. From 07:00 to 09:00 the 333 suffers a significant increase in the average full route running times. During that period there is a 3.5-minute increase in average runtimes between Bondi Junction and Taylor Square. However, the section with largest increase is Bondi Road with an increase of 10.5 minutes in average running times.

Figure 9: Mean inbound running time per section per hour of the day for Tue-Thu in Aug & Nov 2016

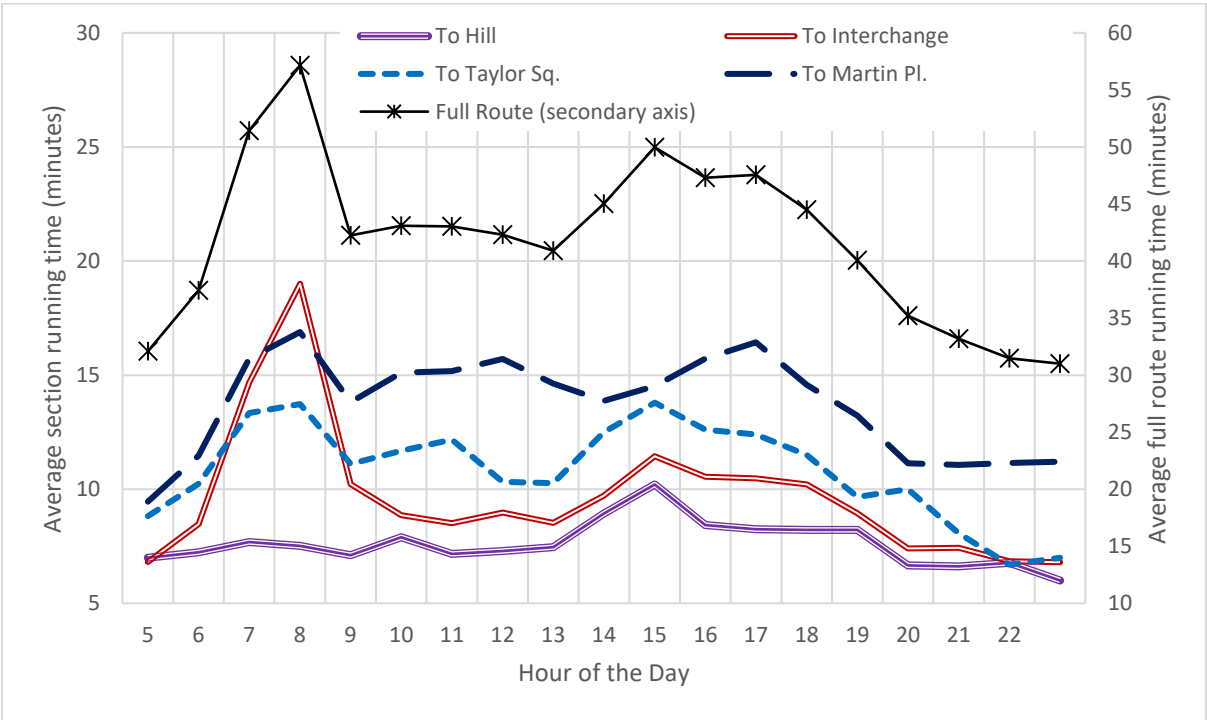
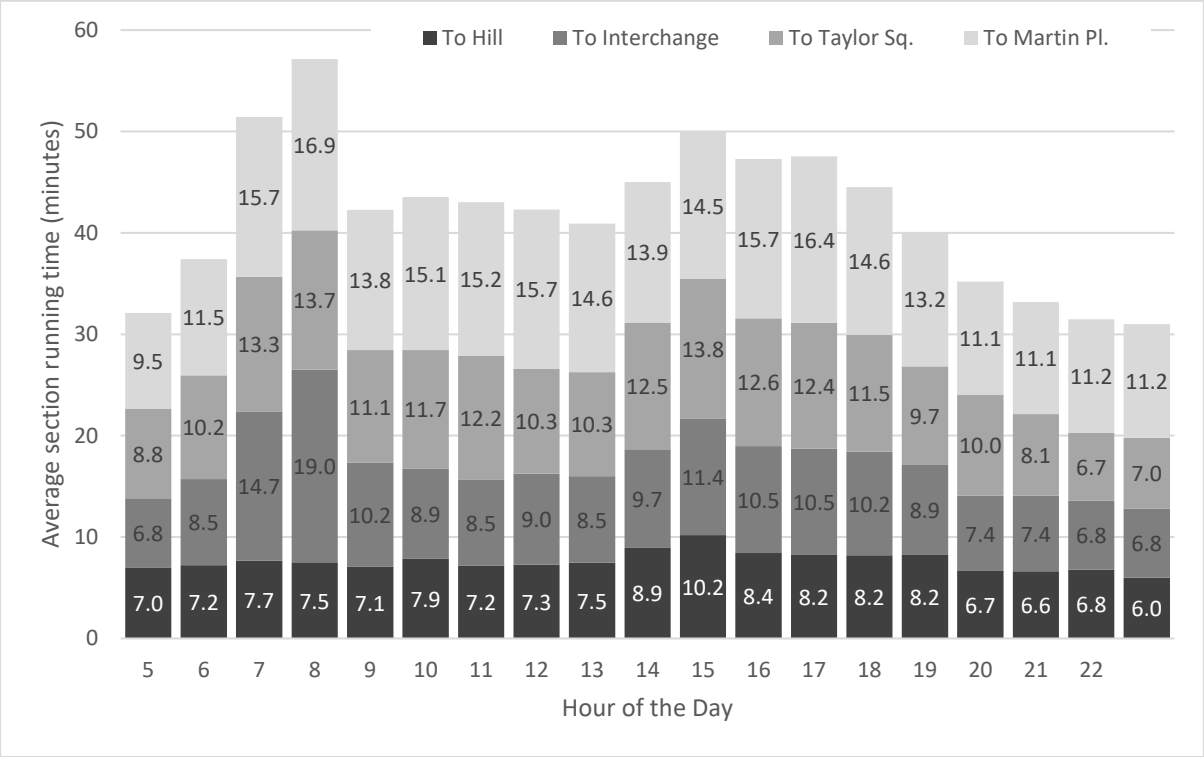


Figure 10: Stacked mean running time per section per hour of the day for Tue-Thu in Aug & Nov 2016



The above graphs show that during the morning peak the increase in Bondi Road running times causes that section to dominate the full route running times. At 8am Bondi Rd takes 8.8 minutes longer than 8pm

The hours with the longest average running times were the hours starting at 07:00, 08:00 and 15:00; that is the morning commute and the afternoon after school finishes.

Table 15: Average headway of inbound 333 timetabled per hour in August and November 2016

Hour	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*
Workday	30	18	6	7	10	10	10	10	10	10	10	10	10	19	21	20	20	20	20
Saturday		29	30	19	10	10	10	10	10	10	10	11	13	21	20	21	20	20	20
Sunday			30	19	14	15	15	15	15	15	15	13	16	20	21	20	20	29	30

Figure 11 below and **Error! Reference source not found. Error! Reference source not found.** show the standard deviation for running times on the 333 per section. The results show that in the morning peak the Bondi Road section contributes the most variability to running times.

Significant levels of runtime variability also indicate that the services are not running optimally and that many services will be running faster than the measured mean runtimes.

As shown in Table 16 below, the 333 inbound running times for services departing in the hour after 08:00 are on average operating 10 minutes slower than the observed minimum running time for a workday morning. Similarly, in the afternoon the services are on average running slower than the minimum running times observed during that hour. Significant differences between the average and the minimum will cause undesirable situations like bus bunching and over-capacity buses.

Figure 11: Plot of standard deviation in runtime per section per hour for TWT in Aug & Nov 2016

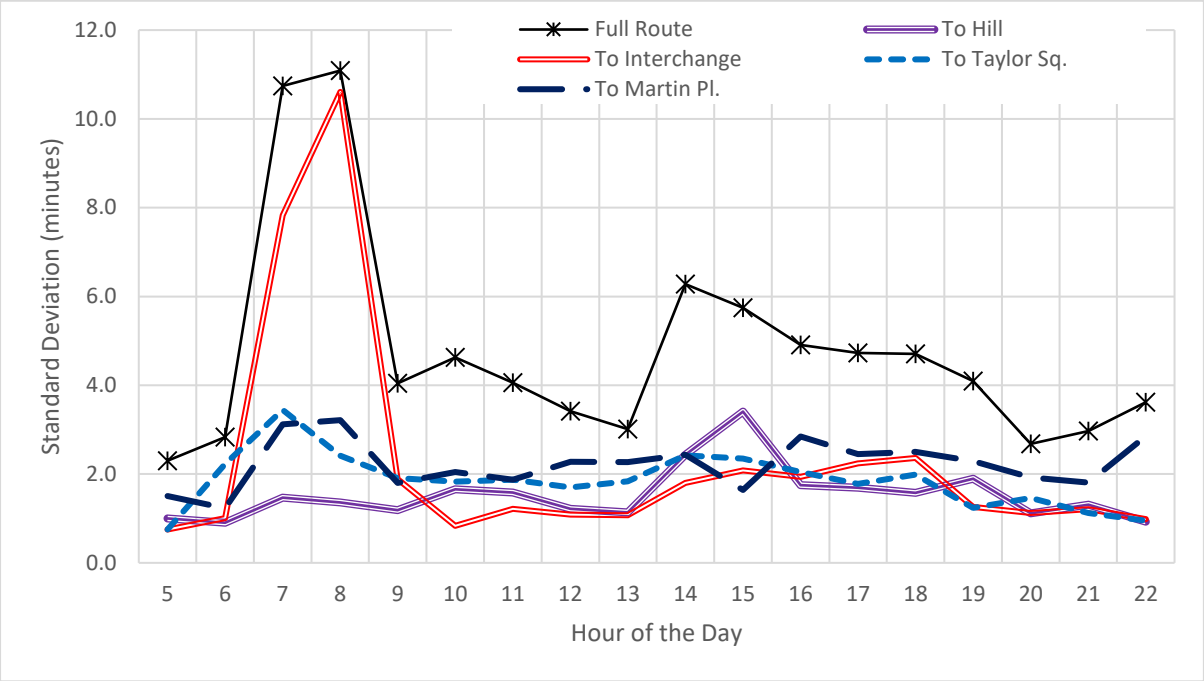


Table 16: Average inbound running times compared to minimum per section per hour

	To Hill			To Interchange			To Taylor Sq.			To Martin Pl.		
Hour	Min	Mean	Diff	Min	Mean	Diff	Min	Mean	Diff	Min	Mean	Diff
5	6	7.0	1.0	6	6.8	0.8	8	8.8	0.8	8	9.5	1.5
6	6	7.2	1.2	7	8.5	1.5	7	10.2	3.2	9	11.5	2.5
7	5	7.7	2.7	7	14.7	7.7	9	13.3	4.3	10	15.7	5.7
8	5	7.5	2.5	9	19.0	10.0	9	13.7	4.7	11	16.9	5.9
9	5	7.1	2.1	7	10.2	3.2	8	11.1	3.1	10	13.8	3.8
10	5	7.9	2.9	7	8.9	1.9	9	11.7	2.7	10	15.1	5.1
11	5	7.2	2.2	7	8.5	1.5	8	12.2	4.2	12	15.2	3.2
12	5	7.3	2.3	7	9.0	2.0	8	10.3	2.3	12	15.7	3.7
13	6	7.5	1.5	7	8.5	1.5	8	10.3	2.3	11	14.6	3.6
14	6	8.9	2.9	6	9.7	3.7	9	12.5	3.5	8	13.9	5.9
15	6	10.2	4.2	7	11.4	4.4	9	13.8	4.8	12	14.5	2.5
16	5	8.4	3.4	7	10.5	3.5	8	12.6	4.6	10	15.7	5.7
17	5	8.2	3.2	7	10.5	3.5	8	12.4	4.4	12	16.4	4.4
18	6	8.2	2.2	8	10.2	2.2	8	11.5	3.5	10	14.6	4.6
19	6	8.2	2.2	6	8.9	2.9	8	9.7	1.7	10	13.2	3.2
20	5	6.7	1.7	6	7.4	1.4	8	10.0	2.0	8	11.1	3.1
21	5	6.6	1.6	6	7.4	1.4	7	8.1	1.1	8	11.1	3.1
22	6	6.8	0.8	5	6.8	1.8	5	6.7	1.7	8	11.2	3.2
Full Day	5	7.8	2.8	5	10.9	5.9	5	11.7	6.7	8	14.6	6.6

Consider that there was only 1.5 minutes difference between the minimum and average running times on Bondi Rd at 13:00 and that increased to 3.7 minutes difference at 14:00. That indicates there might be scope to control running time variability in the early afternoon and improve operating efficiency.

2.4 Bondi Rd vs Oxford St

Table 17 below shows the running times for the route in two halves running on Bondi Rd and Oxford St. These results show operations on Bondi Road are causing the most variability in the morning, and the operations in the CBD are causing the most variability during the day and in the evening.

As such decoupling services operating on Oxford St from services operating on Bondi Road would reduce runtime variability and would allow improvements in on-time running, headway adherence and customer satisfaction on both sections.

Table 17: 333 Inbound runtimes per half per hour for TWT in Aug & Nov 2016

		North Bondi to Bondi Junction				Bondi Junction to Martin Place			
Hour	Samples	Avg	Standarfd Deviation	95% CI	Avg - Min	Avg	Standard Deviation	95% CI	Avg - Min
6	17	15.7	1.3	±0.6	9.3	21.7	2.6	±1.3	8.1
7	63	22.3	7.7	±1.9	11.5	29.1	4.8	±1.2	6.6
8	45	26.5	10.9	±3.3	3.3	30.6	3.3	±1.0	3.9
9	27	17.3	2.4	±0.9	2.8	24.9	2.5	±1.0	12.3
10	29	16.8	1.7	±0.6	2.7	26.3	3.8	±1.4	5.3
11	29	15.7	1.9	±0.7	4.3	27.3	2.9	±1.1	4.0
12	34	16.3	1.6	±0.6	2.0	26.0	2.6	±0.9	4.9
13	30	16.0	1.3	±0.5	5.7	24.9	2.8	±1.0	7.4
14	32	18.7	3.8	±1.3	7.7	26.4	3.8	±1.3	4.3
15	29	21.7	4.5	±1.7	4.0	28.3	2.6	±1.0	8.3
16	33	19.0	2.5	±0.9	4.7	28.3	3.7	±1.3	4.8
17	25	18.7	3.2	±1.3	3.4	28.8	2.7	±1.1	6.1
18	14	18.4	3.3	±1.8	3.2	26.1	3.7	±2.0	3.9
19	18	17.2	2.4	±1.1	3.1	22.9	2.6	±1.2	3.1
20	15	14.1	1.7	±0.9	3.1	21.1	2.1	±1.1	3.1
21	16	14.1	1.9	±1.0	2.6	19.1	1.5	±0.8	2.8
Full Day	485	18.6	6.0	±0.5	1.7	26.2	4.7	±0.4	4.7

Further research is needed to determine if the method in this paper can be replicated by the community and undergraduates.

7 Conclusion

The running times for the 333 are variable and appear dependent on the month, day of the week, and hour of the day. The mean observed running times significantly exceed the scheduled running times. Stabilising then reducing the inbound runtimes would allow the operating authority (TfNSW) to provide the running times promised to the passengers. The current schedule cannot realistically be achieved due to the extended mean running times and the running time variability.

Luckily NSW has Transport for NSW, a single authority that is responsible for the entire transport cluster. The two local councils, Waverley and City of Sydney routinely identify improvements to public transport and that corridor in particular as one of their strategic goals.

The operating authority could split the route into a Bondi Rd and Oxford St two sections as discussed in Section 6. That would prevent the impacts of runtime variability on one section reducing customer satisfaction on the other section. In addition, the operating authority could introduce all-door boarding, dedicated lanes, kerb extensions, signal priority and other proven public transport improvement methods to reduce running times and running time variability.

This paper is yet another case study identifying yet another nationally significant public transport corridor that is in dire need of improvement. We have the data, we have the know-how, all we need is the political will.

Appendixes

8 Data Processing

The bus occupancy dataset contains data for all Opal enabled buses operating during three weeks; one each in August, November and December 2016. Each single record detailed the timetabled and actual arrival time of a bus at one bus stop. The record also detailed the trip identifier, the route variation being operated, the passenger capacity and a coarse estimate of the number of passengers onboard.

Note that the week in December started on Monday the 26th and ended on Sunday the 1st of January 2017. The data for these weeks only was made publicly available by TfNSW. The holiday week was included because it 1/3 of available data and from customers use buses on holidays and during the off-peak periods and service quality during those periods still matters to customers placed at the centre of everything we do.

First the complete bus occupancy dataset for August, November, and December was filtered, keeping only the records where *ROUTE* equalled 333; the results built the primary table. New columns were added by extracting the *DAY_NAME*, *MONTH_NAME*, and *YEAR* from the *CALENDAR_DATE;* as well as by extracting the *TIMETABLE_HOUR* from *TIMETABLE_TIME* and *ACTUAL_HOUR* from *ACTUAL_TIME*.

Then three new columns were added: one computing the *VEHICLE_CAPACITY*, and two estimating minimum passengers (*MIN_PAX*) and maximum passengers (*MAX_PAX*). *VEHICLE_CAPACITY* was *SEATED_CAPACITY* plus *STANDING_CAPACITY.* If the *CAPACITY_BUCKET* equalled 'STANDING_ROOM_ONLY' then *MIN_PAX* was *SEATED_CAPACITY* and *MAX_PAX* was *VEHICLE_CAPACITY*. However, if *CAPACITY_BUCKET* equalled 'FEW_SEATS_AVAILABLE' then *MIN_PAX* was half of *SEATED_CAPACITY* and *MAX_PAX* was *SEATED_CAPACITY*. Otherwise *MIN_PAX* was *0* and *MAX_PAX* was half of *SEATED_CAPACITY*.

To identify the most common route variants, a route table was derived containing only year, month, route variation, direction of operation, stop identification, stop coordinates and a count of the number of records for each combination. GIS was used to plot the route table showing the route variations arriving at each stop.

9 Data sources

T4NSW (2017)	Transport for NSW, 2017-09-28. *Bus Occupancy - Aug 2016 to Jan 2017*. Open Data Hub and Developer Portal, Transport for NSW, Sydney, NSW, Australia. Copyright: CC-BY 4.0. UUID: f9393f9c-f8ad-4628-bec2-542dbc8ad9de. https://opendata.transport.nsw.gov.au/dataset/bus-occupancy-aug-2016-jan-2017.

10 Additional Tables

Table 18: Mean running times (minutes) per section per hour for Tue-Thu in Aug & Nov 2016

	Samples	Average per Section (minutes) 95% CI				
Hour	Samples	Full Route	Hill	Interchange	Taylor Sq.	Martin Pl.
5	11	32.1 ±1.4	7.0 ±0.6	6.8 ±0.5	8.8 ±0.5	9.5 ±0.9
6	17	37.4 ±1.4	7.2 ±0.4	8.5 ±0.5	10.2 ±1.1	11.5 ±0.6
7	63	51.4 ±2.7	7.7 ±0.4	14.7 ±2.0	13.3 ±0.9	15.7 ±0.8
8	45	57.1 ±3.3	7.5 ±0.4	19.0 ±3.2	13.7 ±0.7	16.9 ±1.0
9	27	42.3 ±1.6	7.1 ±0.5	10.2 ±0.7	11.1 ±0.7	13.8 ±0.7
10	29	43.1 ±1.7	7.9 ±0.6	8.9 ±0.3	11.7 ±0.7	15.1 ±0.8
11	29	43.0 ±1.5	7.2 ±0.6	8.5 ±0.5	12.2 ±0.7	15.2 ±0.7
12	34	42.3 ±1.2	7.3 ±0.4	9.0 ±0.4	10.3 ±0.6	15.7 ±0.8
13	30	40.9 ±1.1	7.5 ±0.4	8.5 ±0.4	10.3 ±0.7	14.6 ±0.8
14	32	45.0 ±2.2	8.9 ±0.9	9.7 ±0.6	12.5 ±0.9	13.9 ±0.9
15	29	50.0 ±2.1	10.2 ±1.3	11.4 ±0.8	13.8 ±0.9	14.5 ±0.6
16	33	47.3 ±1.7	8.4 ±0.6	10.5 ±0.7	12.6 ±0.7	15.7 ±1.0
17	25	47.6 ±1.9	8.2 ±0.7	10.5 ±0.9	12.4 ±0.7	16.4 ±1.0
18	14	44.5 ±2.5	8.2 ±0.8	10.2 ±1.3	11.5 ±1.1	14.6 ±1.3
19	18	40.1 ±1.9	8.2 ±0.9	8.9 ±0.6	9.7 ±0.6	13.2 ±1.1
20	15	35.2 ±1.4	6.7 ±0.6	7.4 ±0.6	10.0 ±0.8	11.1 ±1.0
21	16	33.2 ±1.5	6.6 ±0.7	7.4 ±0.6	8.1 ±0.6	11.1 ±0.9
22	13	31.5 ±2.0	6.8 ±0.5	6.8 ±0.5	6.7 ±0.5	11.2 ±1.6
Full Day	485	44.8 ±0.8	7.8 ±0.2	10.8 ±0.5	11.6 ±0.3	14.5 ±0.3

Table 19: Standard deviation of runtimes (δ in minutes) per section per hour for TWT in Aug & Nov 2016

Hour	Samples	Full Route δ	Hill δ	Interchange δ	Taylor Sq. δ	Martin Pl. δ
5	11	2.3	1.0	0.8	0.8	1.5
6	17	2.8	0.9	1.0	2.2	1.2
7	63	10.7	1.5	7.8	3.4	3.1
8	45	11.1	1.4	10.6	2.4	3.2
9	27	4.0	1.2	1.9	1.9	1.8
10	29	4.6	1.7	0.8	1.8	2.0
11	29	4.1	1.6	1.2	1.9	1.9
12	34	3.4	1.2	1.1	1.7	2.3
13	30	3.0	1.1	1.1	1.8	2.3
14	32	6.3	2.4	1.8	2.4	2.4
15	29	5.7	3.4	2.1	2.4	1.6
16	33	4.9	1.8	1.9	2.0	2.8
17	25	4.7	1.7	2.2	1.8	2.5
18	14	4.7	1.6	2.4	2.0	2.5
19	18	4.1	1.9	1.3	1.2	2.3
20	15	2.7	1.1	1.1	1.5	1.9
21	16	3.0	1.3	1.2	1.1	1.8
22	13	3.6	0.9	1.0	0.9	2.9
Full Day	485	9.2	1.9	5.6	2.8	3.0

11 References

Bates (2001)	Bates, J., Polak, J., Jones, P. and Cook, A., 2001. The valuation of reliability for personal travel. *Transportation Research Part E: Logistics and Transportation Review*, *37*(2-3), pp.191-229.
Beaud (2016)	Beaud, M., Blayac, T. and Stéphan, M., 2016. The impact of travel time variability and travelers' risk attitudes on the values of time and reliability. *Transportation Research Part B: Methodological*, *93*, pp.207-224.
Chen (2009)	Chen, X., Yu, L., Zhang, Y. and Guo, J., 2009. Analyzing urban bus service reliability at the stop, route, and network levels. *Transportation research part A: policy and practice*, *43*(8), pp.722-734.
Currie (2012)	Currie, G., Douglas, N.J. and Kearns, I., 2012, September. An assessment of alternative bus reliability indicators. In *Australasian Transport Research Forum*.
Hounsell (2017)	Hounsell, M. 2017, 'Resolving transport and land-use questions with data driven enquiry', course material presented at UTS CPD short course in Integrated Transport & Land-use, Sydney, 11-12 September.
Hounsell (2018)	Hounsell, 2018. '*Better service through runtime savings – Inner West Light Rail case study'.* Australian Transport Research Forum 2018 – PAPER 61.
Kieu (2013)	Kieu, L.M., Bhaskar, A. and Chung, E., 2013. Empirical evaluation of public transport travel time variability. In *Australasian Transport Research Forum*.
Mazloumi (2008)	Mazloumi, E., Currie, G. and Rose, G., 2008, September. Causes of travel time unreliability–a Melbourne case study. In *Proc., 31st Australian Transport Research Forum*. Gold Coast, Australia: ATRF.
Mazloumi (2009)	Mazloumi, E., Currie, G. and Rose, G., 2009. Using GPS data to gain insight into public transport travel time variability. Journal of Transportation Engineering, 136(7), pp.623-631.
Monchambert (2014)	Monchambert, G. and De Palma, A., 2014. Public transport reliability and commuter strategy. *Journal of Urban Economics*, *81*, pp.14-29.
van Oort (2014)	van Oort, N., 2014. Incorporating service reliability in public transport design and performance requirements: International survey results and recommendations. Research in Transportation Economics, 48, pp.92-100.
Strathman (1993)	Strathman, J.G. and Hopper, J.R., 1993. Empirical analysis of bus transit on-time performance. *Transportation Research Part A: Policy and Practice*, *27*(2), pp.93-100.
TRB (2003)	Kittelson & Associates, United States. Federal Transit Administration, Transit Cooperative Research Program and Transit Development Corporation, 2003. *Transit capacity and quality of service manual* (No. 100). Transportation Research Board.
Vuchic (2005)	Vuchic, V.R., 2005. *Urban transit: operations, planning, and economics*. John Wiley & Sons.