FRAMEWORK FOR BLUETOOTH JOURNEY TIME ANALYSIS FOR MODEL VALIDATION

Author: Thanura Rabel, BE (Hons) Civil, M.Phil. (Cambridge)
GIPENZ
Transportation Engineer

thanura.rabel@opus.co.nz
04 471 7246

Opus International Consultants
L10 Majestic Centre, 100 Willis St
PO Box 12 003, Wellington, 6144

ABSTRACT

Smart technologies such as Bluetooth systems are rapidly enabling cost effective data collection and analysis of journey times on our transport networks. These new sources can be particularly useful for use in the validation traffic models. However, the intricacies and the vast sources of data that is created requires a robust and consistent analysis procedure for use in the model validation context.

Using a case study from the recent update of the Wellington Strategic Saturn model, this paper proposes a framework to assist in the development of best practice guidelines for analysis of Bluetooth data (or other smart data sources) for use in validation of traffic models.

A significant portion of the study was related to the analysis stage to identify potentially erroneous data. One of the key outcomes was that an appreciation and understanding of the ‘transport context’ of the study area can point towards where and how sources of error can occur. The study also recommends the use of percentile functions as a robust and effective statistical technique in which to compare observed and modelled outputs. However, further investigation/research is required to determine “what is an acceptable tolerance?” between modelled and observed data with the percentile function.

The final framework incorporated the lessons learnt from the study and considered feedback from the model peer reviewer.

1. INTRODUCTION

The Petone to Grenada Link Road is a proposed East West connection linking SH 1 and SH 2 between Tawa and Petone as shown by Figure 1. Due to its transformative nature on the Wellington Transport network, the scheme will have a major impact on how the regions transport network functions once built.

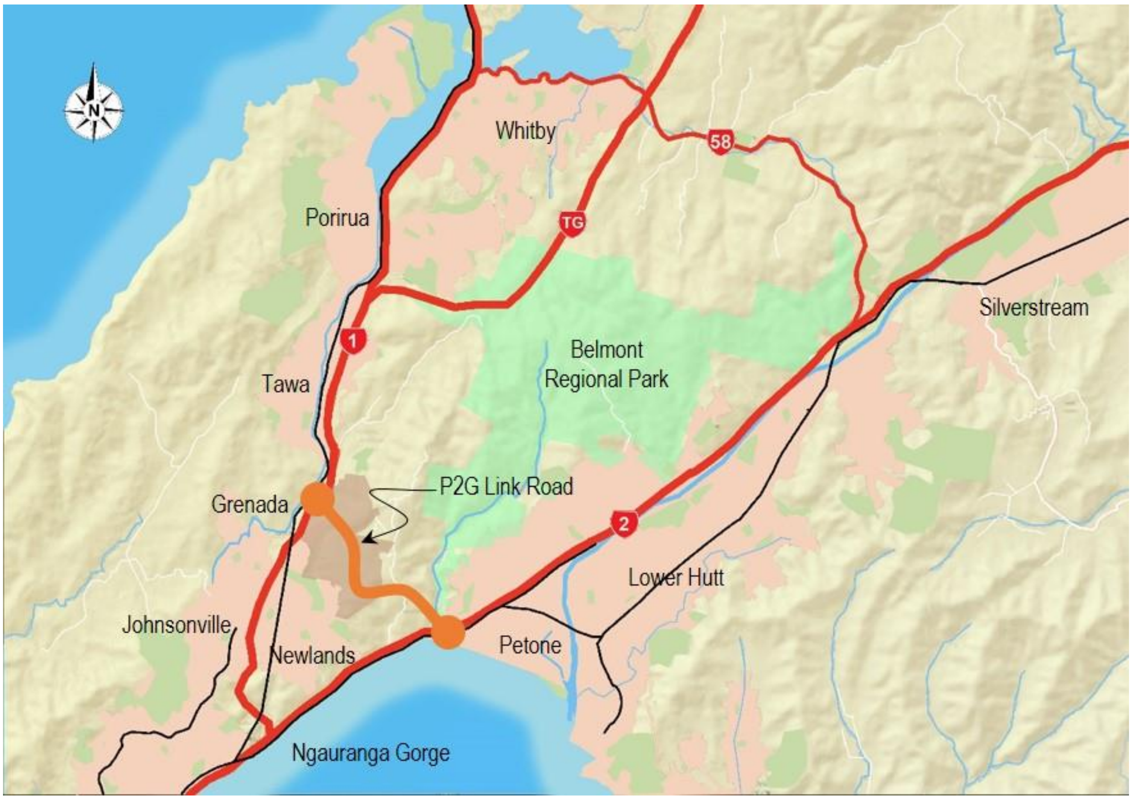


Figure 1: Petone to Grenada Indicative Alignment

During the Scheme Assessment, the North Wellington Saturn Model (NWSM), a traffic assignment and simulation model) was developed and recently updated to test various options and examine the transportation impacts in some key areas. During this update process, an array of observed journey time data sources was collected using traditional methods such as floating car surveys and loop detectors as well as using new smart technologies such as Bluetooth sensor systems. The observed data sets were then analysed and prepared for comparison between the base model outputs and formed the evidence base for the model update.

Bluetooth sensors can obtain travel times with two sensors set-up some known distance apart. These send/receive signals when Bluetooth compatible devices (such hands-free kit, GPS navigation systems and smart phones of users inside cars) drive passed the sensor. Each sensor logs and timestamps a Media Access Control (MAC) address (a unique device identifier) and when a match is made, the time taken to travel between one sensor and the other is reported. The Bluetooth technology enables cost-effective data collection with studies indicating it is 500 to 2,000 times more cost effective compared to floating car surveys in terms of the number of data points produced (Young, 2010).

2. SCOPE AND OUTCOMES FROM STUDY

The aim of this conference paper is not to consider the model validation itself, but rather report as a practice paper on the intricacies involved in the preparation and analysis of Bluetooth journey time data to compare with NWSM. Currently the NZ Transport Modelling Guidelines (NZTA, 2014) has not considered how Bluetooth or related “Big Data” can be used in the validation of transport models. More importantly this should entail the processes required to ensure like for like comparisons can take place between the different tiers of models. An outcome of this study would be to form a basis for the next NZTA modelling guidelines update/revision to include a section on "Big Data" and its uses in Transport Models.

It is important that a robust framework is developed as these sources of data are becoming more common and is likely to be used in a similar sense to assist in validation and update of existing and new models in the future. The processes and lessons learnt from this practical example will form the basis of this paper to recommend a favoured methodology and framework for Bluetooth Journey Time Analysis.

3. CASE STUDY BACKGROUND

A critical section of the existing road network, which will be affected by the proposed P2G link is the highway (SH 2) between Ngauranga and Petone and The Esplanade (local road arterial). As indicatively shown by Figure 2, an existing Bluetooth system[1] (commissioned by NZ Transport Agency) was used to retrieve travel times between Petone and Ngauranga and along the Esplanade.

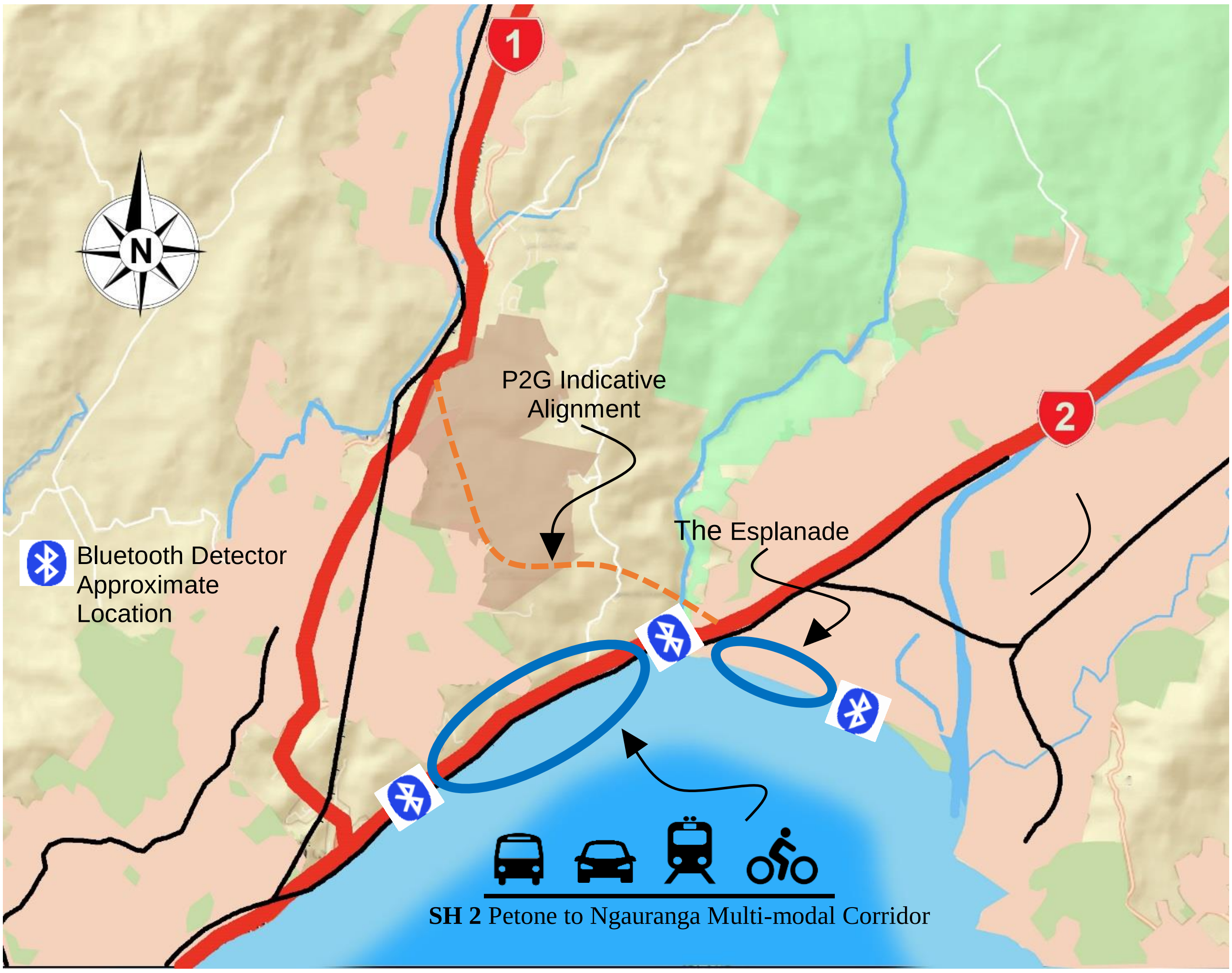


Figure 2: Study Area Context

This section of SH 2 is a significant corridor in the Wellington transport network, serving private vehicles, buses, trains and cyclists. It also battles severe congestion in both the AM and PM peaks with the Petone interchange serving as a local bottleneck. Similarly, traffic along the Petone Esplanade will also be affected by the P2G scheme, particularly with a new proposed interchange at Petone. Therefore, it is imperative that travel times are well represented in the base model.

---

[1] Acyclica Bluetooth System and Analytics

The physical set-up of the detectors is critical for determining accurate travel times. Empirical studies by Cragg (2011) and Vo (2011) indicate that optimisation of the placement location, height and number of detectors in accordance with site context can greatly increase the number match rates and hence the accuracy of travel times. These factors are unknown to the author (and out of scope with this paper) but is noted that in future set-ups it should be considered. Bluetooth sensors also have a "detection zone" as demonstrated by Figure 3. In this zone a vehicle with Bluetooth signal can be picked up three times upon entering and exiting the zone as shown by Figure 3.

In this example, "strength" to "strength" travel times were adopted. However, further research may be required to see which is the best to use depending on the site context and what the travel times will be used for. Strength to Strength travel time is the time which is captured when signal strength from the device and Bluetooth sensor is the strongest both at the origin and destination. It was thought at the time that this would give an indication of the "average" out of the other methods.

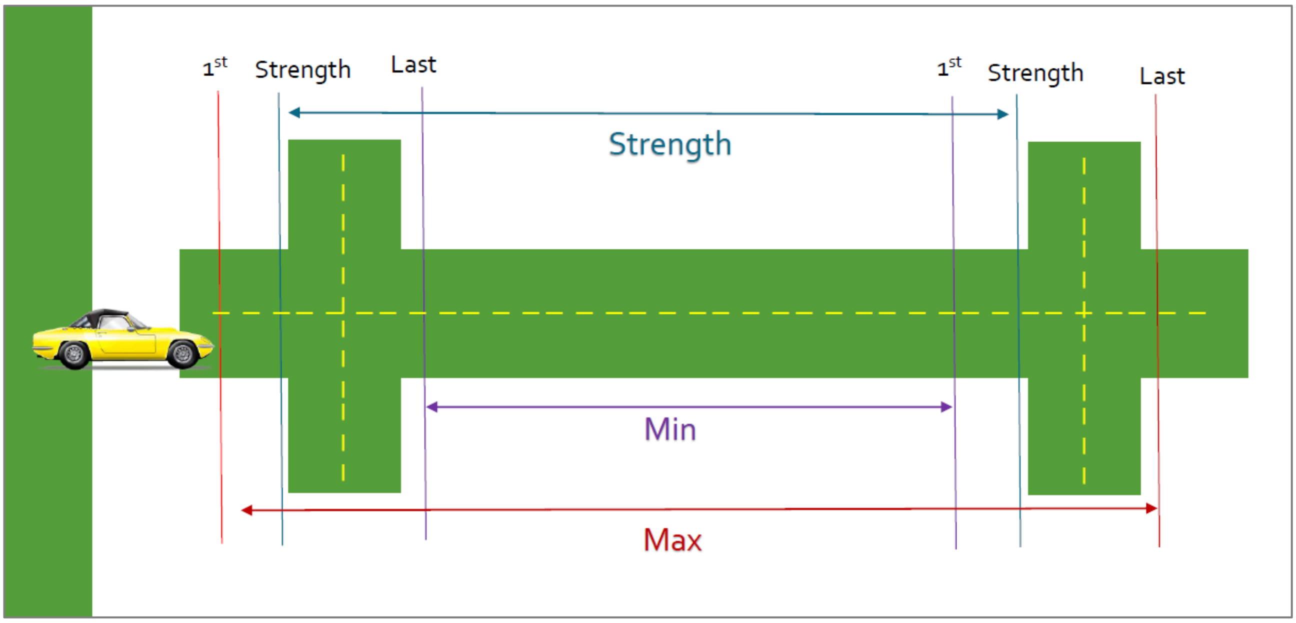


Figure 3: Bluetooth Detection Zone (Source: Acylica Website, 2014)

4. METHODOLOGY

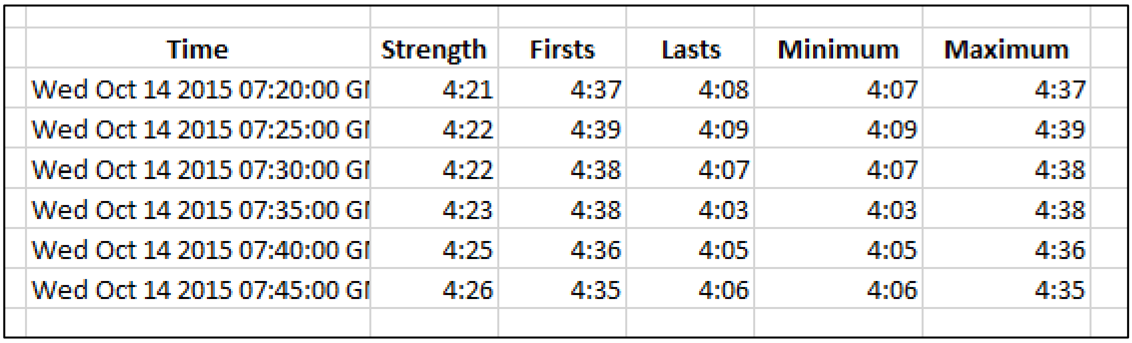
The following key stages of the methodology are summarised below:

4a) Data Extraction

Two months of data was extracted from the Acylica website[2] between 1 August and 27 September in 2015 for the peak periods modelled in NWSM (7:30am – 8:30am & 4:30pm – 5:30pm). No inter-peak period was considered for comparison. Data was extracted in five minute intervals. An example of the raw data is shown below by Figure 4. The column headings with different travel times relate to the variants in the calculation methodology as per Figure 3. As mentioned previously, the "Strength" travel time was adopted for the analysis.

---

[2] https://cr.acyclica.com/accounts/login/?next=/#



Time	Strength	Firsts	Lasts	Minimum	Maximum
Wed Oct 14 2015 07:20:00 GI	4:21	4:37	4:08	4:07	4:37
Wed Oct 14 2015 07:25:00 GI	4:22	4:39	4:09	4:09	4:39
Wed Oct 14 2015 07:30:00 GI	4:22	4:38	4:07	4:07	4:38
Wed Oct 14 2015 07:35:00 GI	4:23	4:38	4:03	4:03	4:38
Wed Oct 14 2015 07:40:00 GI	4:25	4:36	4:05	4:05	4:36
Wed Oct 14 2015 07:45:00 GI	4:26	4:35	4:06	4:06	4:35

Figure 4: Raw Data Example

An important step was to also recognise the Bluetooth detector locations. Coordinates of the relevant detector were retrieved and noted. This is critical so that “like for like” comparison can take place between the model nodes. NWSM uses nodes at intermediate points on roads or at critical junctions, which had been previously coded.

The closest nodes and the detector location correlated well in the Esplanade region however the Ngauranga detector appeared to be a further 300 north of the closest NWSM node. This needed to be taken into consideration in the analysis stage and results interpreted appropriately.

Scatter plots displayed by Figure 5 and Figure 6 shows the profile of the raw data. Immediately, it can be seen the variation in travel time in The Esplanade corridor. SH 2 shows a reasonably uniform profile with the some variation particularly in the middle / end of the time period. The next section considers data filtering techniques which were used to clean the data and also to investigate why such variable data points were ocurring on The Esplanade.

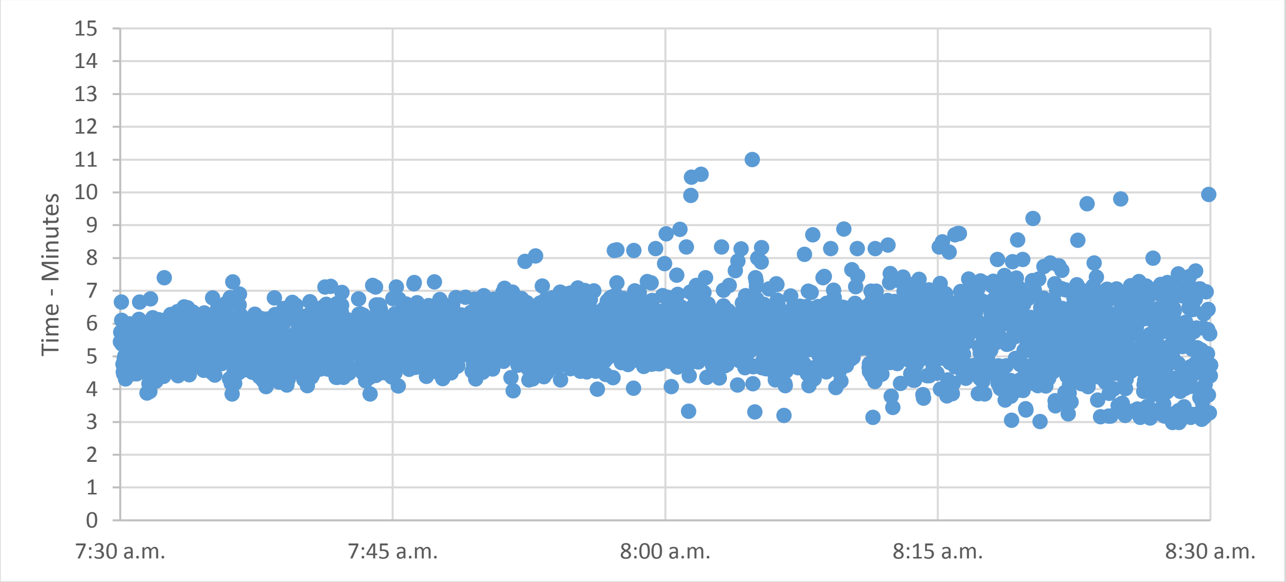


Figure 5: SH 2 Petone to Ngauranga (NBD) Pre-Filtered Travel Times, AM Peak

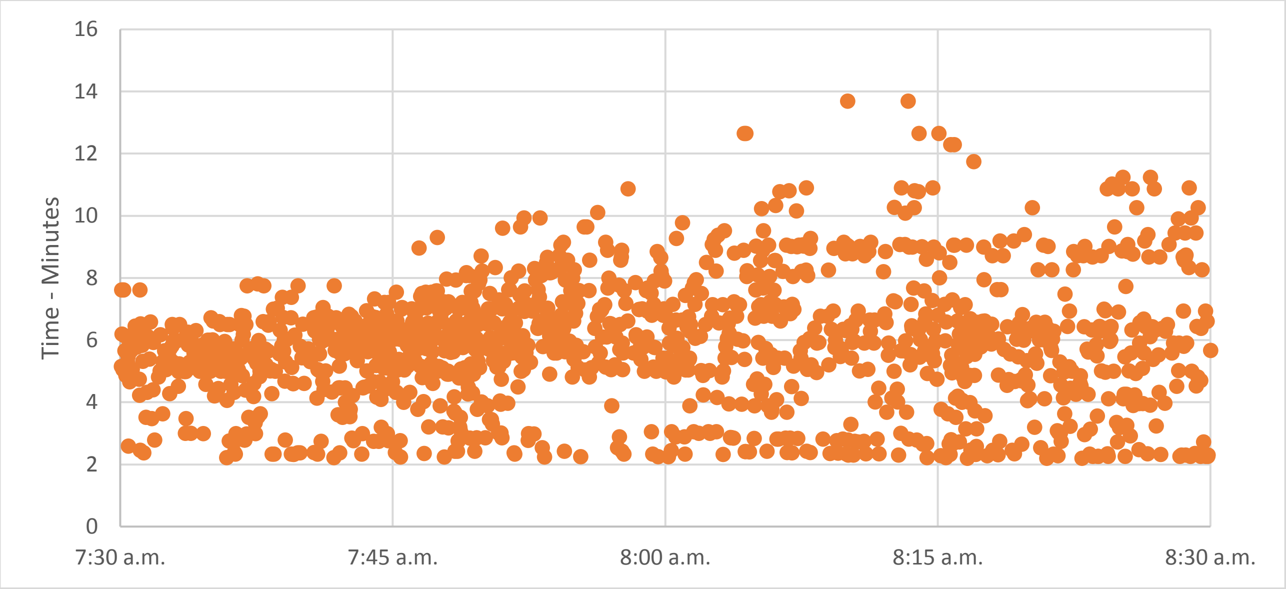


Figure 6: The Esplanade (WBD) Pre-Filtered Travel Times, AM Peak

4b) Data filtering

Data filtering was initially undertaken through removing all dates relating to public/school holidays which would potentially skew travel times. Importantly, it was recognised that SH 2 was prone to incidents, which may also produce irregular travel times through the corridor. Recent incidents include accidents, rockfalls and storms/flooding, which result in partial and full lane closures. Information from the NZ Transport Agency's Traffic Road Event Information System (TREIS) was analysed through liaison with Wellington Transport Operation Centre (WTOC). Using the TREIS database, a query was made for SH 2 within the study area[3] with the dates as noted above. This information was then further reviewed to narrow incidents which occurred in the peak hours and to determine if road closures would have had a significant impact on the corridor. This primarily depended on the duration of closure. As seen most of the incidents in the time periods of interest occur further north on SH 2. There was only one relating to the study area, which is highlighted by red below. After further liaison with WTOC it was concluded that this incident was of little disruption. Nonetheless, that day was excluded from further consideration.

Table 1: TREIS Summary for SH 2

Start date	**Day**	**Duration**	**Location**	**Direction**	**Incident Description**	**End Status**
5/08/2015 5:07	Wednesday	2 h, 24 m	Kelson	NBD	Breakdown	Resolved
9/08/2015 15:48	Sunday	1 d, 14 h, 34 m	Melling Link To Western Hutt Road	NBD	Fallen tree/s across highway	Resolved
10/08/2015 16:54	Monday	13 h, 21 m	Harbour View Road To Melling Link	SBD	Breakdown	Resolved
10/08/2015 17:08	Monday	1 h, 13 m	Harbour View	NBD	Crash	Resolved
14/08/2015	Friday	1 d, 19 h,	Normandale,	SBD	Breakdown	Resolved

---

[3] Note the wider study also included Melling further north on SH 2. This is not included in this paper.

18:48		58 m	Lower Hutt			
17/08/2015 7:40	Monday	10 m	Melling To Normandale Road	SBD	Rock falls	Resolved
20/08/2015 16:48	Thursday	6 d, 17 h, 59 m	Melling	SBD	Rock falls	Resolved
27/08/2015 17:52	Thursday	3 d, 12 h, 14 m	Korokoro	SBD	Crash	Resolved
16/09/2015 8:06	Wednesday	10 m	Ngauranga	SBD	Wandering Stock	Resolved
24/09/2015 16:55	Thursday	1 h, 18 m	Melling	SBD	Breakdown	Resolved

There was no such data for The Esplanade as TREIS is only used for state highways with collaboration with the Network Outcome Contract managers.

Another issue which was discovered during the data filtering stage was repetitive data. These occur in two instances:

1. Duplicate travel times were repeated in successive five minute blocks. Whilst there is no factual confirmation as why this occurs it could be due to the following reasons i) detector fault/maintenance or ii) lack of match rates leading the analytics system duplicating the last travel time of which a suitable number of matches were made. Nevertheless, these data points were excluded from the analysis

2. Clustering of travel times under one time stamp. This is due to either multiple vehicles being detected/matched at the same exact time or multiple devices being picked up inside vehicles. The issue was more noticeable on the SH 2 section. On further investigation, it was noted that this could be due to passengers on trains, explaining the clustering of the data points. Again, these data points were excluded from the analysis.

Variability on The Esplanade was significant and required closer inspection of the area in relation to detector location. The following observations were made:

- The eastern (upstream) detector near Cuba St was very close to a signalised intersection. It therefore may be the case that the “detection zone” was extending out and capturing queued vehicles as well as free flowing vehicles through the intersection both from Cuba Street and Esplanade approaches

- The context of the wider corridor encompasses signalised pedestrian crossings, several side roads, high proportion of heavy vehicles (~10%) and it is interspersed with kerbside parking. This would understandably introduce sources of variation with some vehicles been interrupted in their line of travel through the above and some being able to travel freely. Also it would not be unusual for a vehicles to be detected by the upstream sensor to then reroute through a side road and enter back to The Esplanade again (at the same time or later time) to be captured again by the downstream sensor. This would introduce a variability in travel time as vehicles do not take a direct route as intended.

- The corridor also carries a bus lane, which operates in the westbound direction between 7am and 9am on the weekdays, which also allows taxis to use it. The buses and taxis will therefore travel at a faster speed and hence lower travel time. Furthermore, multiple passengers inside the buses may have Bluetooth enabled device which may further skew travel times of the corridor towards the bus travel time.

- Land use around the area contains several cafes, museum, hotels and access to the Petone beach and wharf. It would not be unusual for a vehicle to enter the Bluetooth zone then stop at a café during the journey and continue again sometime later. In this instance, a continuous route was not taken and hence travel time will be longer than what was expected.

It is difficult to account for the above sources of error as the raw data (actual MAC address matches) was not available to be downloaded. Therefore, with knowledge of local context and other data collected by manual surveys, maximum and minimum travel times were defined. Any travel times outside of this range was excluded from further analysis.

5. RESULTS

Following the data filtering stages, travel times for the entire period were then ranked from smallest to largest and a percentile function was formed as per charts shown by Figure 7 and Figure 8.

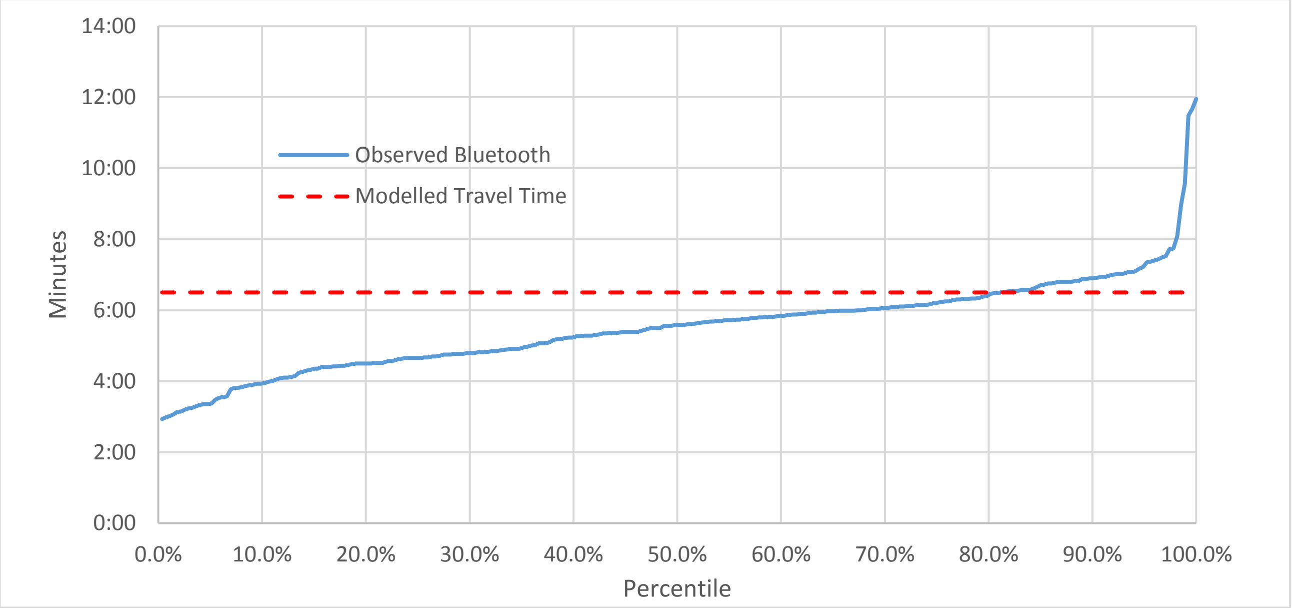


Figure 7: Percentile Observed Travel time vs Modelled – SH 2 Petone to Ngaurana (SBD) AM Peak

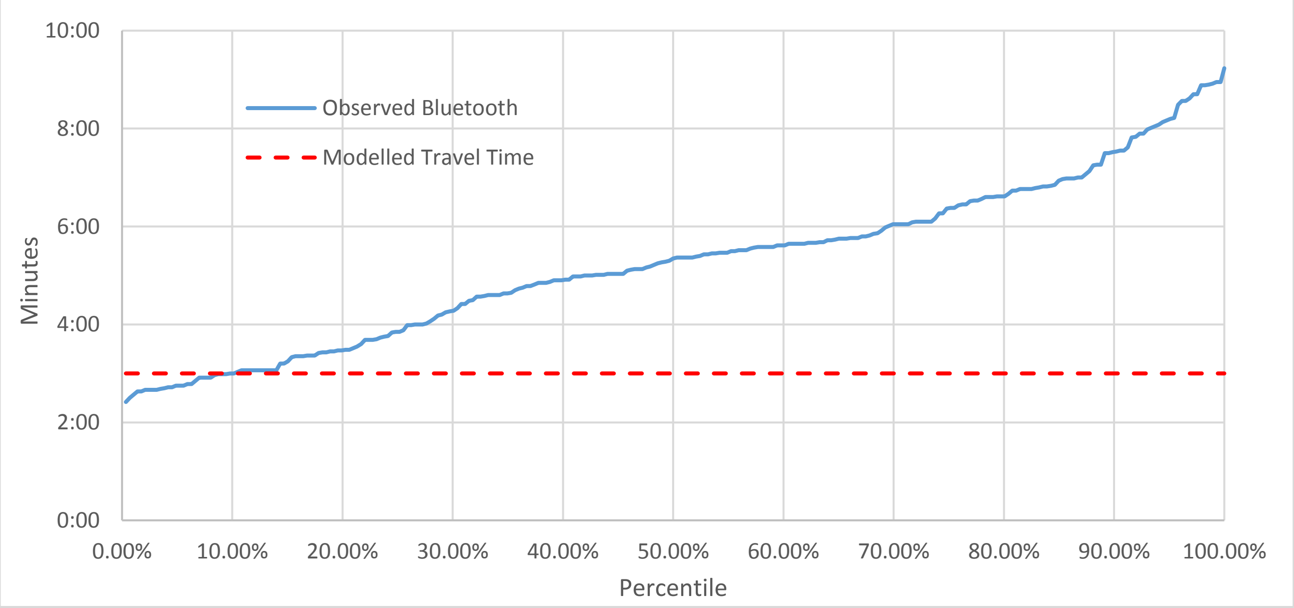


Figure 8: Percentile Observed Travel time vs Modelled – The Esplanade (WBD) AM Peak

Similarly, the travel time as determined by the base year NWSM model was also plotted (dotted red line). This is the average travel time experienced during the modelled AM Peak (7:30am to 8:30am).

As can be seen the Petone to Ngauranga section on SH 2 is well represented in the model, showing modelled travel times are around the 80th percentile of the observed Bluetooth data. However, for the Esplanade, the modelled travel time is (more or less) representing the free flow travel time of the observed data (near 10th percentile), which needed further consideration in the model update

This analysis was then used by the NWSM modelling team for the model update along with other data sources that were collected.

6. LESSONS LEARNT

A summary of the lessons learnt are outlined below. This summary also includes comments by the NWSM peer reviewer who also reviewed the Bluetooth analysis.

- 80% of the analysis process was devoted to data cleaning and sorting data into format you require – a known predicament in the wider Big Data community. The Esplanade required a considerable amount of effort in data cleaning due to high variability. The placement of Esplanade detectors near other intersection (on Cuba St intersection and Petone interchange) may not have been the most ideal location in the initial set-up
    - ➢ In future studies on arterials, Bluetooth detectors should be deployed at mid-block locations if possible to avoid biasing the travel time data with oversampling of Bluetooth devices in vehicles that queue at intersections
- Whilst it did not significantly influence the data filtering in this case study, TREIS is a useful database and should be used in conjunction with similar studies as an additional input to assist in filtering of data.
- Due to the nature of Bluetooth detection zones, devices traveling through these zones may dwell longer than a single detection cycle iteration and may be reported more than once. To ensure accurate travel times, a consistent convention must be established.
    - ➢ This study utilised ‘strength-to-strength’ detections to calculate travel times but further research is required if this is the most appropriate
- The use of the statistical travel time metrics such as the percentile function is an effective manner in which to show observed travel times. It is also useful in areas where high travel time variability is experienced as you can effectively class the travel times in accordance with certain percentiles (i.e. anything over 90th percentile could be classed as unwarranted due to influence of external factors – weather, incidents etc.)
    - ➢ However, further research is required to determine “What is an acceptable tolerance level for modelled data?”. This investigation used the 50th percentile as an acceptable benchmark
- Defining maximum and minimum travel times (as outlier detection measure) may be problematic for corridors, which already experience high variability. This could result in exclusion of legitimate travel times indicating incident based congestion or other non-typical behaviour that is occurring
    - ➢ Further statistical tests (time series moving average, Z-test etc.) could be employed if budget/time warrants as more robust outlier detection measure

- In contrast to highway travel time studies, arterial roads typically experience a significant number of vehicles departing and entering the roadway from cross-streets over relatively short distances. This case study showed The Esplanade (arterial local road) with the potential for error as discussed in the Section 4b may not be suitable for Bluetooth detection of travel times. This could hold true for other local roads with similar context. The State Highway (SH 2) on the other hand, was effective as it does not contain any side roads/access points, no intersections and/or bus lanes and therefore, it was not prone to sources of error experienced on The Esplanade.

    - ➢ Further research is required in this area as to what type of road environment is suitable for the effective analysis of Bluetooth travel times.

- Extend data period to one year and conduct further analysis into how travel times vary between months. It is also recommended that analysis is extended into the inter-peak period

- When model nodes and Bluetooth detector locations do not exactly match up, an alternative would be to compare average travel speed rather than travel time. Attention should be given upfront to ensure the correct and most appropriate model nodes (start/end points) are selected for comparison. Consideration also need to be given to the model function used for determining travel times and if this includes delays at intersections or not.

7. RECOMMENDED FRAMEWORK

6a) Establishing Transport Context

One of the most critical things to establish early in the analysis procedure is a desktop review of the ‘transport context’ of the corridor. By transport context, this means answering the following things:

- What modes uses the corridor and what frequency/proportions?
- How many exit and entry points lie within the corridor and are there any major intersections?
- Construction/Road works and or ITS with variable speed limits in place?
- What land uses are present in the corridor?
- Is there high pedestrian activity near the detector locations?
- Is the area prone to high variability in travel times?
- Is the area prone to incidents?

By answering these questions, a lot can be gained about whether the corridor is appropriate in the first place for such analysis. It can also provide guidance about what to look for in the data filtering stage including what dates to remove and what travel times are typically expected. It is highly recommended this desktop review is conducted upfront before beginning the analysis to save a lot of time during the process.

Further literature review was consistent with the findings from the study. The following points outline how errors from the transport context could arise:

- Picking up signals from other sources (mobile phones from pedestrians, cyclists, multiple devices inside a vehicle), which has potential to skew data (Beasley et al., 2015)
- Clustering of data from a bus or train full of passengers with Bluetooth enabled devices, which creates a bias in journey times towards bus/train rather than give overall (Beasley et al., 2015)
- Vehicles not travelling from origin sensor to destination sensor through usual or intended path (alternative path taken to destination which may lead to longer journey times (Morahan, 2015)

- Vehicles not travelling from origin sensor to destination sensor as continuous journey (e.g. stopping at a service station) (Morahan, 2015)
- Vehicles detected at the origin sensor, misses destination sensor but picked up by destination sensor travelling other direction later (Vo, 2011)

If working with Bluetooth supplier at installation/commissioning phase it is also recommended that detector location is appropriate and justified. Other studies have recommended that optimisation of detector offset, height and number of detectors can increase match rates and sample sizes (Vu, 2011). In fact, Malinovskiy et al. (2010) found that placing two omni-directional detectors at a single point differing in width and height has been known to increase detection and match rates. There are also studies indicating that a minimum study length of 1.5 to 3.5 km is recommended on arterial roads. Shorter sections may be influenced by the size of the detection zone (Schneider IV, et al., 2010).

Cragg (2013) attempted to investigate the demography of the type of vehicles that are more likely to contain Bluetooth enabled devices. His work concluded that Bluetooth related travel times may be over-represented by commercial heavy/light goods vehicles as they are more likely to have a Bluetooth transmitting device in their vehicle. Furthermore, he concluded that “business” journey purpose may be slightly over-represented in the sample along with newer vehicles. In the NZ context, this is important as NZ has a relatively old vehicle fleet. Therefore, the disparity in commercial vehicles over-representing the sample size could be greater as these vehicles also tend to be newer. As an implication, care must be taken along corridors with high heavy vehicle use. The Esplanade for example contains roughly 10% heavy vehicles – this could have been another manner which resulted in the high variability in data.

6b) Data Extraction and Filtering

As per the lessons learnt section, a dataset of one year is recommended for a study. This could be dependent on the model type and use of the Bluetooth data. For NWSM, a strategic year around the average model (which is used for scheme assessment/economics), one year would be the recommended size to capture variations. Micro-simulation models may warrant lower timescale.

Data filtering should comprise of the following steps but again this may be dependent on model type and purpose.

- Removal of all public and school holidays
- If a State Highway, review of the relevant TREIS database to establish if any significant incidents have taken place during time of review. Some liaison with the local TOC will be required to establish the consequences was noted incidents
- Evaluate what was determined in step 6a above and investigate if there will be any outliers emerging from the transport context. These can be often seen in the data set through a simple visual inspection
- Establish maximum and minimum expected travel times as a simple method to clean outliers. If time and cost constraints are not limited, then recommend adopting a more robust statistical methodology for outlier detection[4]. This could include looking at spread of data point in comparison with rest of data or filtering data that is more than four times the 25th percentile (Beasley et al., 2015). Thogulava et al., (2015) also used several measures including quartile approach and winsorising[5] of data. Furthermore, UK Department of Transport (DFT, 2014) suggest using 95% confidence level for the mean travel time within ± 10% accuracy threshold over the distance of 3 km or longer. This 10% threshold could be increased for shorter sections i.e. (20%)

---

[4] It is noted the Bluetooth supplier and analytics platform will most likely have its own outlier detection algorithms in place. This needs to be established with early on with the supplier and if any additional data cleaning is required.

[5] Winsorising or winsorisation is the transformation of statistics by limiting extreme values in the statistical data to reduce the effect of possibly spurious outliers.

6c) Analysis

The percentile travel time function is recommended as a tool for comparison of observed vs modelled travel time outputs. The percentile measure is not sensitive to outliers as it ranks travel times from smallest to largest. Theoretically for a simple analysis, the data filtering stage could be bypassed by using the percentile function, with anything over the 90th and below the 10th percentile ignored. It is a robust technique which can see through data abnormality.

The skewness of the data could also be examined by comparing 50th percentile to the 90th (and 50th to the 10th). This will give an indication as to where most of the travel time data points are situated and importantly, assist in seeing if the modelled output is in a range where this majority is.

Using average (mean) travel time is not recommended due to the variability experienced in this case study. If a simple check is required, the median would be a better indicator.

8. REFERENCES

S. Beasley., et al. (2015). “Reading Bluetooth Journey Time Monitoring System”.

S. Cragg. (2013). “Bluetooth Detection – Cheap but Challenging”. Transport for Scotland

Y. Malinovskiy., et al. (2010). "Investigation of Bluetooth-Based Travel Time Estimation Error on a Short Corridor," presented at the Transportation Research Board 2011 Annual Meeting, Washington, D.C

C. Morahan. (2015). “Congestion: What Are Christchurch’s Worst Offending Roads?”. IPENZ Transportation Group Conference 2015, Christchurch.

NZ Transport Agency. (2014). “Transport model development guidelines”, viewed October 2015, < https://www.nzta.govt.nz/assets/resources/transport-model-development-guidelines/docs/tmd.pdf>

W. H. Schneider IV., et al. (2010). "Statistical Validation of Speeds and Travel Times Provided by a Data Service Vendor," The University of Akron

S. Young. (2010). "Bluetooth Traffic Monitoring Technology - Concept of Operation & Deployment Guidelines," University of Maryland - Center for Advanced Transportation Technology.
UK Department for Transport. (2014). “Data Sources and Surveys, Transport Analysis Guidance”, UNIT M1.2

T. Vo. (2011). “An Investigation of Bluetooth Technology for Measuring Travel Times on Arterial Roads: A Case Study on Spring Street”. School of Civil & Environmental Engineering, Georgia Institute of Technology

ACKNOWLEDGMENTS

Acknowledgements to Wellington Modelling Control Group, QTP and Jacobs for their excellent feedback during the process.