Stream travel time reliability using GPS-equipped probe vehicles

Sharmili Banik¹, B. Anil Kumar²,* and Lelitha Vanajakshi¹

¹Department of Civil Engineering, Indian Institute of Technology Madras, Chennai 600 036, India
²Department of Civil and Environmental Engineering, Indian Institute of Technology Patna, Patna 801 103, India

Travel time reliability (TTR) is an important measure to quantify the variation in travel times. Currently, there is no single reliability metric appropriate across all locations, that is easily understandable and can be used to compare across facilities. Moreover, reliability analysis of facilities from developing countries is limited due to the non-availability of extensive data required for such an analysis. The present study addresses these gaps. It identifies a reliable data source for such analysis of heterogeneous, lane-less traffic, compares existing reliability measures for the data, highlights the advantages and disadvantages, proposes a measure that may be more suitable for such traffic with high variability, and finally illustrates how reliability analysis under such conditions can be done with limited data sources such as GPS-fitted transit vehicles. Using such commonly available data for traffic stream reliability analysis is the ultimate aim of this study. For validation, stream travel time from Wi-Fi scanners is used. The study analyses the performance of various reliability measures and identifies the most suitable ones. Following this, a reliability measure, i.e. capacity buffer index (CBI), is developed to identify the unreliable congested regimes or periods, keeping time taken to travel at capacity conditions as the benchmark. From the results, it has been observed that CBI is in agreement with the real-field conditions in 94% of the cases, whereas it is 75% buffer time index. Finally, the feasibility of using bus probes to measure stream TTR is checked. Results show that bus probes can be an indicator of stream reliability and the developed measure can effectively capture the relationship between stream and bus TTR.

Keywords: Bus probes, contingency tables, mixed traffic, travel time reliability, Wi-Fi sensors.

Travel time, a fundamental measure in transportation, is the time taken to traverse a route between any two points of interest. High traffic demand and limited road capacity make travellers spend much more time on their daily journeys. Urban travel times are prone to higher degrees of variability either due to recurring (fluctuations in traffic demand, presence of control features) or non-recurring (traffic incidents or occurrence of any unexpected events) traffic events. Travel time reliability (TTR), defined as the consistency or dependency in travel conditions over time, has been increasingly considered an important measure of the performance of the transportation system, as it is related to the travel experience and is important to both users and operators. For example, user groups may want to know which paths they should use to minimize the probability of unforeseen delays. Transit agencies/operators will want to know which network segments make the travel times vary.

Initially reported studies on reliability were mainly qualitative in nature, mostly questionnaires ascertaining travellers’ experiences with respect to the quality of their commute. Nowadays, with the latest advancements in traffic sensing technologies, huge amounts of data are being generated in both real-time and historical contexts. As is widely known, the majority of the existing automated traffic-sensing technologies may not perform well under heterogeneous and less lane disciplined traffic, such as that existing in India, making city-wide data collection difficult. For example: (i) use of inductive loop detectors under such conditions has encountered a number of operational issues related to error propagation; (ii) use of any video image processing software needs to undergo sufficient parameter calibration process, which may be cumbersome and (iii) use of advanced sensors such as Bluetooth/Wi-Fi sensors has penetration rate constraints in different areas of a city. Therefore, there is a need to identify a robust method to obtain detailed traffic information at a low cost that can work at all locations. In this regard, public transit vehicles equipped with tracking devices that travel all over the network and record the travel time and location information of vehicles at a certain interval can be the best source of traffic data under such conditions. Hence, reliability analysis using these data is a good solution for the heterogeneous and lane-less traffic conditions.

Different types of measures have been reported by researchers and applied to assess traffic reliability across urban arterials and freeways. Chase Jr et al. examined a host of TTR measures and reported that no single measure was ideal for representing reliability, as each reported measure is derived from a part of the entire travel-time distribution. It has to be noted that though a single best measure can simplify decision-making, one has to carefully consider the variability in travel-time data and identify the best indicator. None of the studies that were reported under heterogeneous and less-lane disciplined traffic conditions focused on...
identifying the best reliability measure that can depict reality. Therefore, the present study aims to evaluate TTR measures for the selected study stretch under mixed traffic conditions and to check the feasibility of using probe vehicle data to efficiently measure the reliability of the traffic stream.

Literature review

From the measurement perspective, reliability on any given section/network is quantified from the distribution of travel times over a significant amount of time. Asakura and Kashiwadani\textsuperscript{17} defined reliability as the probability that a trip between a given origin and destination pair can be made successfully within a given time interval. Based on the existing literature, TTR measures can be broadly categorized as: (i) statistical range measures, (ii) buffer time measures, (iii) tardy trip measures, (iv) probabilistic measures and (v) congestion/volume-based measures\textsuperscript{9,11,18}.

The statistical range measures analyse travel-time variability based on historical data or real-time information. These measures generally use standard deviation or spread of the distribution to represent traffic conditions experienced by travellers\textsuperscript{17}. Measures that fall into this category include travel-time window, per cent variation and variability index. Day et al.\textsuperscript{25} studied travel times on arterial routes in Indiana, USA; using measures of central tendency and variability. They showed that routes with a greater density of traffic signals tended to have high travel times, leading to less reliability. However, standard deviation indicates the spread of travel time around some expected value, i.e. average travel time, implicitly assuming travel times to be distributed symmetrically and can occur only under free-flow conditions\textsuperscript{9}.

Buffer measures indicate how much worse 95th percentile travel times are compared to typical travel conditions. Measures in this category include buffer time, buffer time index (BTI) and planning time index (PTI). Early research focused on an average-based BTI. However, it has been shown that this index can decrease as variability increases under highly skewed distributions\textsuperscript{20,21}. Therefore, a buffer index based on the median has been used more recently, such that BTI continues to increase as variation increases. Gong and Fan\textsuperscript{22} reported that using 95th percentile travel times while calculating PTI is more vulnerable to severe traffic crashes or inclement weather. They suggested using 80th percentile travel times and proved the case valid while ranking recurrent bottlenecks.

The tardy trip measures indicate unreliable conditions by quantifying how much delay the worst trips experience. These measures provide an idea of how bad the worst trips are and to what extent of extreme travel times the travellers are exposed. Measures in this category include on-time arrival and misery index. On-time arrival captures the number of trips with travel times beyond an acceptable threshold range. The misery index generally captures the variation from an expected travel time to the average travel times of some percentage of the worst trips. Previous studies have reported this percentage of the worst trips to be between 5% and 20% (ref. 11). These measures, however, may not capture reliability effectively because of the thresholds used, which do not vary linearly for all trips\textsuperscript{11}.

Probabilistic travel measures often use a threshold travel time or a predefined window to differentiate between reliable and unreliable travel times\textsuperscript{10,23}. The Dutch Ministry of Transport, Public Works and Water Management proposed a reliability measure stating that all trips should be made within 20% of median travel time\textsuperscript{10}.

Congestion-based measures help view the variability in traffic conditions and determine the unreliable periods or days. Measures in this category include frequency of congestion (FOC). This is defined as the percentage of days or time or trips during which the travel times exceed some threshold value. The threshold value can be chosen by the analyst. Higher values of FOC indicate unreliable conditions. FOC is relatively easy to compute if continuous traffic data are available\textsuperscript{18}.

Based on the above review of the literature, it can be observed that reported studies suggest the use of different reliability metrics or indices and no unique measure can be used for reliability analysis\textsuperscript{9–13}. Most of the studies were reported from homogeneous traffic environments, and the applicability of the measures to capture the wide variations in a mixed traffic condition has not been evaluated. It is a well-known fact that wide variations in travel time exist under mixed traffic conditions. For example, it is not likely that the travel-time distributions are symmetric in mixed traffic. Therefore, we consider that measures focusing on the spread around average travel time may fail to capture reliability. Also, threshold-based measures such as on-time arrival may not be appropriate, as thresholds used in such measures should be strictly different for different vehicle types in mixed traffic conditions. Very few studies focused on stream reliability were reported under heterogeneous and less lane-disciplined traffic conditions such as those in India\textsuperscript{14–16,24}. Possibly, this can be due to the unavailability of a robust, continuous data source. Collecting the same under Indian traffic conditions may be difficult due to site-specific sensor calibration, lack of sophisticated technologies that can capture heterogeneity in traffic, huge cost of deploying sensors like Wi-Fi over the entire network, etc. Also, privacy issues deter the monitoring of private vehicles and collect the travel-time information. It is a known fact that the sample size of travel times plays a vital role in the accuracy of TTR estimates. Therefore, there is a need to identify an efficient way to get detailed traffic information that can work under such conditions and explore the applicability of the reliability measures. In this regard, public transit vehicles that travel all over the network can be the best source of traffic data under such conditions. In countries like India, with less lane discipline and heterogeneous traffic, buses share the same road space as other
vehicles (no separate bus lane available), and face similar traffic interruptions and interactions. Also, many buses operate, especially during peak hours with other vehicles on the arterials. This motivated us to explore using buses as probes to gather stream reliability information. For mapping bus-to-stream reliability, effective measures that are representative of the traffic conditions are vital. To address these issues, the present study aims to (i) evaluate and identify a unique TTR measure that can effectively represent travel-time variations under mixed traffic conditions, and (ii) check the feasibility of using probe vehicle data to efficiently measure the reliability metrics of the stream.

Study area and data

The study was carried out by collecting data from a 1.6 km urban arterial connecting the Bharathi Nagar and Vijaya Nagar intersection, located in the southern part of Chennai city, Tamil Nadu, India (Figure 1). The data were collected using two sources, namely GPS and Wi-Fi.

GPS data from devices fitted in the Metropolitan Transport Corporation (MTC) bus number M1 running through the identified study corridor were collected for one month. Data were available from 5:00 am to 10:00 pm daily. The raw GPS data provide information about the position (latitude and longitude) of the bus at every 5-sec interval, which is processed to obtain the distance and travel time between each consecutive position of the bus. The time needed to travel the corridor for every bus trip was obtained and the generated bus travel times were used for analysis.

Wi-Fi sensors fairly capturing all modes in the stream were used to collect stream travel information. Wi-Fi sensors were placed at predetermined locations, i.e. origin and destination of the corridor. Each Wi-Fi device has a unique ID. When a Wi-Fi has enabled device in the stream communicates with the sensors placed along the roadside, it responds to the inquiry scan with its unique MAC address and timestamp. The sensors capture this information and process the data in real-time to a server, and the server stores the data in a database. The data comprising the detected unique MAC addresses and corresponding timestamps were collected from the server for the same one-month period. The stored data were further utilized to estimate travel times using MAC address matching technique\textsuperscript{25}. Once travel times were computed from both sources, outliers were identified by applying thresholds and inter-quartile range (IQR) based on speed. Thresholds were set considering the maximum speed on the road as 80 km/h and minimum speed as 5 km/h. The lower bound speed is expected to exclude data points from pedestrians or vehicles that stopped (for personal reasons) between two Wi-Fi monitoring stations. The upper limit speed was taken to exclude any unusual speeds that can cause due to sensor reading errors. Following this, IQR-based outlier removal was done, and the resulting smoothened and clean data were used for modelling. One of the main characteristics of bus travel time that makes it different from the stream travel time is the frequent stopping of buses at stops and related acceleration and deceleration near these bus stops. This has to be taken into account before calculating the reliability indices. To address this, the non-stop trajectory extraction methodology proposed by Kumar et al.\textsuperscript{26} to remove bus stop dwell time has been adopted in this study. The processed data were used for reliability analysis.
TTR analysis

Reliability has been defined in multiple ways in the existing studies. Some measures focus on relating reliability to variability and some to traffic congestion, etc. In this study, unreliability is considered to be essentially associated with the variability in travel conditions experienced by the users. This notion especially becomes vital in case of mixed traffic conditions, where the variability in traffic conditions is very high. Various measures were analysed under varying traffic conditions to identify suitable measures of reliability under such traffic conditions and compared.

Preliminary data analysis of Wi-Fi-based travel-time data was done first to identify high variability and low variability hours within a day. Figure 2a and b shows a sample scatter plot and box plot of the collected travel times using Wi-Fi sensors over a day. From the figure, it can be clearly observed that there are four different groups based on the spread in travel times. They are 5:00–8:00 am with the lowest spread in travel times, 8:00 am–12:00 noon, 12:00 noon–17:00 pm, 17:00–22:00 pm. This classification was used to tag the periods into low variability group-1 (LVG-1), high variability group-1 (HVG-1), low variability group-2 (LVG-2) and high variability group-2 (HVG-2) respectively. The relative spread across the time periods was used as a cut-off or decision criterion to group the periods. This was done based on insights such as the box plots, exploratory data analysis, engineering judgement and past experience of researchers regarding variability on the study route.

The most commonly used reliability measures were calculated for the above four groups. Reliability measures were first calculated for each hour of the day separately and then the percentage of hours that showed unreliability in respective groups was calculated. It is to be noted that the relative reliability across groups has been analysed and an observation period is defined as unreliable if the value of the measure exceeds approximately 60th percentile value of all hourly observations taken together for the corresponding measure. This was done for a week, excluding Sunday. For instance, under HVG-2 (5 pm–10 pm), there are five hours and for six days, there are $5 \times 6 = 30$ hourly observations. The percentage of hours reported to be unreliable was calculated for that group. Table 1 provides a summary of the results. Some of the key observations made from the table discussed below.

Statistical measures

Standard deviation, per cent variation, width, skew and unreliability index are the measures considered under this group. Standard deviation shows the spread of travel time
Table 1. Percentage of unreliable hours

<table>
<thead>
<tr>
<th>Category</th>
<th>Reliability measure</th>
<th>LVG-1</th>
<th>LVG-2</th>
<th>HVG-1</th>
<th>HVG-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistical</td>
<td>Standard deviation</td>
<td>5.56</td>
<td>29.17</td>
<td>40</td>
<td>73.33</td>
</tr>
<tr>
<td></td>
<td>Per cent variation</td>
<td>0</td>
<td>4.17</td>
<td>13.33</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Width</td>
<td>0</td>
<td>4.17</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>72.22</td>
<td>54.17</td>
<td>50</td>
<td>26.67</td>
</tr>
<tr>
<td></td>
<td>Unreliability index</td>
<td>61.11</td>
<td>50</td>
<td>50</td>
<td>30.33</td>
</tr>
<tr>
<td>Buffer</td>
<td>Buffer time index</td>
<td>50</td>
<td>54.17</td>
<td>76.67</td>
<td>46.67</td>
</tr>
<tr>
<td></td>
<td>Planning time index</td>
<td>5.56</td>
<td>25</td>
<td>40</td>
<td>70</td>
</tr>
<tr>
<td>Tardy trip</td>
<td>On-time arrival</td>
<td>0</td>
<td>0</td>
<td>4.17</td>
<td>6.67</td>
</tr>
<tr>
<td></td>
<td>Misery index</td>
<td>50</td>
<td>45.83</td>
<td>73.33</td>
<td>58.33</td>
</tr>
<tr>
<td>Probabilistic</td>
<td>Pr(α)</td>
<td>5.56</td>
<td>8.33</td>
<td>6.67</td>
<td>8.33</td>
</tr>
<tr>
<td></td>
<td>AVV 2004 probabilistic measures</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Congestion</td>
<td>Frequency of congestion</td>
<td>0</td>
<td>6.67</td>
<td>8.33</td>
<td>43.33</td>
</tr>
</tbody>
</table>

around an expected value, mostly around the mean value. This can be expressed as

\[ \text{Standard deviation} = \sqrt{\frac{(x_i - \mu)^2}{N}}, \quad (1) \]

where \( x_i \) is the \( i \)th travel-time observation; \( \mu \) the average travel time and \( N \) is total number of observations. Increased standard deviation from the mean refers to unreliable conditions. From Table 1, it can be observed that HVG shows a higher percentage of unreliable periods (73.33). This indicates that most of the hours in HVG show high deviations from the mean value in the corresponding hour, particularly hours in the HVG-2 (evening hours), agreeing with the field condition. However, from the user perspective, this measure is abstract and does not provide information on much extra time the travellers need to allocate to their trips to reach their destination on time. Also, it treats early and late arrivals equally, while the travellers are concerned more about late arrivals. It has to be noted that the standard deviation type of expression implicitly assumes travel times to be distributed symmetrically. van Lint et al.\(^\text{16}\) showed that a symmetrical distribution could occur only under free-flow conditions. Per cent variation, defined in eq. (2) suffers from similar disadvantages of standard deviation.

\[ \text{Per cent variation} = \frac{\text{Standard deviation}}{\mu}, \quad (2) \]

Other variables like width, skew and unreliability index do not assume a symmetrical shape of distribution. They are defined as follows

\[ \text{Width } r_{\text{var}} = \frac{T_{90} - T_{10}}{T_{50}} \], \quad (3)

\[ \text{Skew } r_{\text{skew}} = \frac{T_{90} - T_{50}}{T_{50} - T_{10}} \], \quad (4)

Unreliability index \( UI_r = \frac{r_{\text{var}} \ln(r_{\text{skew}})}{L_r} \),

for \( r_{\text{skew}} > 1 = \frac{r_{\text{var}}}{L_r} \), otherwise, \quad (5)

where \( T_{10}, T_{50} \) and \( T_{90} \) are the 10th, 50th and 90th percentile travel times respectively, and \( L_r \) the route length. Table 1 shows that width can identify the high variability groups to some extent, as it shows a relatively higher percentage of unreliable periods under HVG. However, width gives insight into the total variation in the travel times (\( T_{90} - T_{50} \)), but the travellers are concerned mostly about late arrivals.

Skew, which measures the asymmetry of travel-time distributions around the median, is observed to identify the unreliable periods differently and disagree with the field variability conditions. Analysis showed that the left-hand side (\( T_{50} - T_{10} \)) was steeper than the right-hand side of the distribution (\( T_{90} - T_{50} \)), resulting in high value of skew at the low variability hours. This indicates that skew only identifies a period as unreliable when the distribution is highly left-skewed, which occurs mostly during congestion onset or offset, i.e. when the distribution is highly skewed. Unreliability index is a combination of the corresponding skew and width measure. Since the effects of skew are also reflected in this index, the imitations of skew are also valid here, resulting in being inconsistent with variability conditions.

**Buffer measures**

These include BTI and PTI. BTI provides extra travel time to be allocated to a trip so that travellers reach their destination on time. Buffer time provides the difference between the 95th percentile and average travel time. It is dimensional and hence difficult to compare across times. BTI expresses
buffer time as a percentage of average travel time as follows

\[
BTI = \frac{T_{95} - \text{Average travel time}}{\text{Average travel time}},
\]

where \(T_{95}\) is the 95th percentile travel time. Higher values indicate unreliable conditions. The concept of BTI relates to the way travellers make their decisions and hence is a user-friendly reliability measure. It takes into account the variation between average travel times and the 95th percentile travel times. This measure thus is concerned only with late arrivals. However, Florida Department of Transportation, USA\(^2\) reported that BTI might give an unstable value of changes in reliability as it can move in a direction opposed to the mean and percentile-based measures. This is because buffer index uses both the 95th percentile and average travel times, and percentage change in these values can vary over time. If one changes more relative to the other, counter-intuitive results are observed. From Table 1, it can be observed that though BTI captures the unreliability to some extent in HVGs, it overestimates the unreliability in LVGs and underestimates in HVGs. This can be explained by understanding the following sample analysed case. The 95th percentile travel time, average travel time, difference between these values and buffer index were 706.66 sec, 402.1 sec, 304.56 sec and 0.75 respectively at 7 pm and the corresponding values were 349 sec, 182.54 sec, 166.46 sec and 0.91 respectively at 5 am. The average, 95th percentile times and the variation between average and 95th percentile conditions were much higher at 7 pm than 5 am, but the buffer index appeared significantly lower at 7 pm. Thus, this measure cannot compare the variation between average and 95th percentile conditions and hence is not suitable for comparison across time periods and modes. However, it is a meaningful measure for travellers and planners if corrected for the above issues of underestimating variation and difficulty in comparing across entities.

PTI is the ratio of the 95th percentile time to the free flow travel time (eq. (7)). In Table 1, PTI can be observed to capture the wide variation in travel times in HVGs effectively. As this measure captures the total variation, unlike the buffer index, it is good for planners to evaluate infrastructure improvements and policies. However, travellers are more concerned about late arrivals and thus with the variation from an expected value to the extreme travel times they may face.

\[
PTI = \frac{T_{95}}{\text{Free-flow travel time}}.
\]

Tardy trip measures

These include on-time arrival and misery index. On-time arrival calculates the percentage of trips for which the travel times are in a threshold range (between average and average + 110% travel times; eq. (8)). On-time arrival identifies the relative variability in HVGs and LVGs (Table 1). However, the percentage of unreliable periods identified in all groups, particularly HVGs, is less. This may be because on-time arrival considers a threshold range which may be too high to capture the variation within travel times for small routes, as in the present study. The threshold for on-time measures needs to be carefully defined as it does not vary linearly for every trip. Also, this measure is abstract for travellers.

\[
\text{On-time arrival} = \left(100 - \frac{\# \text{ of trips with travel time} > 110\% \text{ of average travel time}}{\# \text{ of trips}}\right). \tag{8}
\]

Misery index evaluates the negative attributes of trip reliability (how much the average travel times of the maximum delayed trips exceed the mean travel time per unit of mean travel time; eq. (9)) and is observed to report a similar situation as BTI. This measure also suffers from the same disadvantages as BTI.

\[
\text{Misery index} = \frac{\text{Average travel time of longest } 20\% \text{ of trips} - \text{Average travel time}}{\text{Average travel time}}. \tag{9}
\]

Probabilistic measures

These include \(\text{Pr}(\alpha)\) and \(\text{Pr}\{\text{travel time} \leq 10 \text{ min} + 750\}\). \(\text{Pr}(\alpha)\) determines the probability of trips that may take longer travel time than the predefined travel-time threshold \(\alpha\) (eq. (10)). Higher value indicates greater unreliable conditions. van Lint et al.\(^3\) suggested 1.2 times median travel time (750) as the threshold \((\alpha = 1.2)\). This measure has similar implications as the on-time arrival and suffers from the same disadvantages as those of on-time arrival.

\[
\text{Pr}(\alpha) = \text{Pr}\{\text{travel time} \geq \alpha \times 750\}. \tag{10}
\]

Probabilistic measure, \(\text{Pr}\{\text{travel time} \leq 10 \text{ min} + 750\}\), can be defined as

\[
\text{Pr}\{\text{travel time} \leq 10 \text{ min} + 750\} > 95\% \text{ for shorter sections (less than 50 km), } \text{Pr}\{\text{travel time} \leq 12 \times 750\} > 95\%, \text{ for longer sections}. \tag{11}
\]

Table 1 shows that in all cases, the probability is greater than 95%, indicating no unreliable periods. Thus, this measure could not capture the reliability condition. This is because it is sensitive to thresholds which need to be modified every time for comparison across modes, space and time. Also, this measure is abstract to travellers.
Congestion measure

Frequency of congestion can be defined as the percentage of time the travel time exceeds the threshold travel time in a given corresponding hour. Here, a threshold travel time corresponding to 15 km/h speed is used. From Table 1, it can be observed that this measure can identify the HVGs as unreliable. However, unreliability during uncongested periods, if it occurs, may not get captured by this measure. Though this measure indicates unreliability, it can be more related to performance assessment than variability or reliability.

Overall, from the above discussion, it can be observed that buffer measures can capture the unreliability better than the other measures considered. Among the buffer measures, PTI captures the wide variation between the lowest travel times (can be free-flow travel time) and the worst travel times (usually 95th percentile travel time) faced by the travellers. BTI, on the other hand, captures the variation from expected conditions to the highest travel times, which is more important. BTI is thus more suitable for practical applications as it is concerned with late arrivals. However, when used for comparison across modes and times, BTI can be an unstable indicator as discussed previously. Also, buffer index calculates the extra time required from average conditions to be on time. While the average is a good benchmark for deciding the buffer, during the peak of congested hours when unreliability is more, optimal travel time or travel time at capacity can be a better benchmark. To address these limitations, a new reliability measure has been developed in the present study.

Development of a reliability measure

Considering the various limitations discussed above, a capacity buffer index (CBI) is proposed here that helps overcome the problem of instability and captures the variability concerned with late arrivals compared to optimal travel condition. CBI is the ratio of the difference between the 95th percentile travel time and travel time at capacity divided by the travel time at capacity (eq. (12)).

\[
CBI = \frac{T_{95} \ - \ Travel\ time\ at\ capacity}{Travel\ time\ at\ capacity}. \tag{12}
\]

Here, the travel time at capacity was determined by assuming a parabolic relationship between speed and flow, i.e. by calculating half of the maximum speed. Advantages of this measure are:

(i) It uses percentile-based value (changes with time and mode) and travel time at capacity (does not change with time and mode), solving the problem of instability due to the use of both mean and percentile-based measures in BTI, making it suitable for temporal and modal comparison. (ii) It uses capacity travel time as a benchmark, making it suitable to identify unreliable congested regimes or periods. Travelling at capacity is more expected or frequent than travelling at free-flow speeds. (iii) It focuses only on late arrivals. (iv) It is dimensionless, making it easy to compare with other measures.

The performance of CBI as a reliability measure was tested across LVGs and HVGs. The average travel time at capacity conditions was obtained from speed–volume analysis as 6 min. Table 2 summarizes the percentage of unreliable periods under each variability group calculated using CBI. It can be observed that CBI is able to capture the high unreliability or variability in HVG groups and relative lower unreliability in LVGs. Figure 3 shows the unreliable periods identified by CBI for each day for two weeks. Red colour indicates an unreliable period and green for a reliable period. From Figure 3, it can be observed that unreliable conditions are dominant during the high variability peak hours (HVG-1 and HVG-2).

CBI also has the advantage of identifying only time periods of unreliability associated with late arrivals, which is of greater concern. Figure 4 shows the reliability values using CBI. Here, negative CBI values highlighted in green show that early morning off-peak periods hardly exceed capacity conditions and unreliability is not high, and vice versa during evening peak hours. This allows us to consider reliability only at critical periods of high congestion. Thus, CBI can be used as a combined performance and reliability measure.

Having identified a suitable measure based on Wi-Fi data, we addressed the data availability issue. Deployment of Wi-Fi sensors is limited and the only data source available across various cities of countries like India is GPS fitted in public transport buses. Hence, we analysed whether GPS data of buses can be used to obtain stream TTR.

Use of buses as probes for measuring stream TTR

Obtaining robust continuous travel-time data for TTR analysis in developing countries like India may be difficult. Wi-Fi that provides stream travel-time information is a slowly developing data source, while GPS installed in public transit vehicles is already a developed and popular data source for travel time across cities. Exploiting probe vehicle data to gather stream reliability information can be a good way to overcome the difficulties in conducting reliability analyses due to lack of data sources. In this study, we checked

<table>
<thead>
<tr>
<th>Variability group</th>
<th>Percentage of unreliable hours in different variability groups using CBI</th>
</tr>
</thead>
<tbody>
<tr>
<td>LVG 1</td>
<td>5.56</td>
</tr>
<tr>
<td>LVG 2</td>
<td>41.67</td>
</tr>
<tr>
<td>HVG 1</td>
<td>53.33</td>
</tr>
<tr>
<td>HVG 2</td>
<td>93.33</td>
</tr>
</tbody>
</table>
the feasibility of using probe vehicle data to efficiently measure the reliability conditions of stream. The developed CBI for different time periods of the day for six days was used to compare against the stream-based capacity index for the same periods. Additionally, the applicability of suitable measures under each group that are realistic was also evaluated. All these were used to check whether the same pattern of TTR condition is indicated by stream and probe vehicle data over a day. The performance was grouped as good, average and poor based on the range of values of the measures. Figure 5 shows the reliability status for the stream (Wi-Fi) and bus (GPS)-based CBI.

From Figure 5, it can be observed that the reliability conditions for stream and probe vehicles follow an almost similar pattern, i.e. when stream reliability decreases, probe-based reliability also decreases, and vice versa. Such pattern is stronger in the morning off-peak hours and evening peak hours with free flow and congested regimes, and all vehicles travel at free-flow and stop–go conditions respectively. Thus, there is a strong correlation between stream and probe vehicle reliability. Table 3 (contingency table) indicates the applicability of these measures. Contingency tables consist of an \( n \times n \) array, where \( n \) is the number of classes. Here the classes are the reliability conditions, viz. good, average and poor. The values in the array are the number of times stream and bus TTR agree. Table 3 shows percentage acceptance of the number of times the stream and probes show comparable reliability status. While calculating this, extreme opposite status (good to poor and poor to good) was considered unacceptable.
From Table 3, it can be observed that CBI shows the maximum number of positive agreements and the least number of negative agreements. It shows a maximum percentage acceptance of 94 and thus can be used for mapping the bus to stream TTR. Frequency of congestion also shows a good percentage acceptance indicating that probe vehicles face similar traffic conditions as those of other vehicles in less lane disciplined, mixed traffic conditions. However, the frequency of congestion is more related to performance assessment than variability or reliability.

Next, statistical analysis was carried out by conducting Pearson’s chi-square test at 5% level of significance to check whether bus and stream TTR are significantly related or not. Such analysis also gives the phi coefficient, which represents the strength of such association, independent of the sample size. In the present study, the null hypothesis is assumed as bus and stream TTR are not associated with each other, and alternative hypothesis is assumed to be that the bus and stream TTR are associated with each other. Table 4 shows the results obtained from the hypothesis testing.

From Table 4, it can be observed that there is an association between stream and bus TTR with a $P$-value of $1.4715\text{e}-10$. Also, the phi value is 0.713, indicating that bus probes are good indicators for stream TTR.

A similar analysis was carried out on another arterial link of Chennai city to check the repeatability of the results. Table 5 presents the percentage acceptance calculated based on contingency values under each category of relatability measure.

From Table 5, it can be observed that CBI and frequency of congestion show a high percentage of acceptance in this corridor also. Thus, a strong relationship between stream and bus TTR can be observed and CBI best captures this reliability correlation compared to other measures. Thus, it can be concluded that bus probes are a good representation of stream reliability.

**Summary**

TTR is an important measure that has gained wide attention among transportation engineers, as it can give useful insights into urban traffic conditions. The applicability of the existing measures to capture stream reliability under mixed traffic conditions has been less explored in past studies. The present study has addressed this gap and examined the applicability of various reliability measures under mixed traffic conditions. Various measured were evaluated and the effective measures were identified using stream travel-time data collected from Wi-Fi sensors. It has been observed that different measures give different notions of reliability, and each has advantages and disadvantages. In this
study, the notion of reliability associated with the variability in travel times was considered and the reliability measures that could capture the variation were identified. It has been observed that a set of measures can capture the variability, but not all these measures are suitable for analysis. This is due to their disadvantages and/or being abstract indicators to travellers. Based on this, buffer measures were found to be the most suitable. However, BTI possesses the problem of instability, making it difficult for cross-comparisons across entities. PTI, on the other hand, gives equal importance to both early and late arrivals, while in reality, only late arrivals are of serious concern. To address these issues, a CBI proposed in this study can capture the variability associated with late arrivals and also be used for cross-entity comparisons. CBI was observed to capture TTR in majority of the cases. The feasibility of using bus probes to measure stream TTR was evaluated using the developed measure and other selected measures. A strong correlation between stream and bus TTR was observed, and CBI best captured this correlation. Thus, in conclusion, bus probes can be used to represent stream reliability.

18. FHWA, Travel time reliability: Making it there on time, all the time. Federal Highway Administration, US Department of Transportation, 2006.