

## RESEARCH ARTICLE

# Challenges faced in heterogeneous traffic data collection: a comparison of traffic data collection technologies

DND Jayaratne\*, CJ Vidanapathirana and HR Pasindu

*Department of Civil Engineering, Faculty of Engineering, University of Moratuwa, Moratuwa.*

Submitted: 18 February 2019; Revised: 12 January 2020; Accepted: 08 April 2020

**Abstract:** Traffic data are the fundamental inputs to traffic flow analysis and simulation studies, which facilitate decision making in the field of traffic engineering. Hence, the accuracy of traffic data is of paramount importance. This study compares new technologies available for traffic data collection considering their accuracy and applicability in the Sri Lankan context. Traffic in Sri Lanka is of heterogeneous nature, as opposed to the homogeneous nature observed in most developed countries. Hence, collection of traffic data poses several challenges that affects its accuracy. Three techniques, the infrared driven TIRTL instrument, the video image processing-based TRAZER application and the traffic data collection method using the Google distance matrix application programming interface (API), with respect to their data collection accuracy are reviewed in this study. The fundamental macroscopic traffic data variables (flow and speed) were evaluated against control surveys. It was found that each technology has its strengths and weaknesses and needs to be used appropriately. The TIRTL instrument fared better on road sections on level terrain when the crossfall did not obstruct the infrared beams. Such occasions provided a rich set of microscopic traffic data. The TRAZER software delivered data up to a 100 % accuracy. However, this required the user to go through a lengthy post-processing routine to extract the final set of traffic data. Google traffic data collection provides highly accurate results when estimating link speeds. This method is ideal for collection of bulk data with spatio-temporal variations and the process can be fully automated to reduce the human resource requirement.

**Keywords:** Google distance matrix, heterogeneous traffic, TIRTL, traffic data collection, TRAZER.

## INTRODUCTION

Traffic engineering theories are mainly developed for homogeneous traffic conditions where operating speeds are constant, driver behaviour is uniform, and the sizes of the vehicles do not vary. However, the actual nature of traffic on roads differ from this. Vehicles with different dimensions operate at non-uniform speeds while depicting diverse driver behaviour resulting in heterogeneous traffic flows. There are variations in headways, lateral spacings, and acceleration/deceleration rates. Furthermore, vehicles possess many other diverse operating characteristics. Therefore, when collecting field data for such heterogeneous traffic flows, the use of accurate and appropriate data collection methods is exceedingly important.

Traffic surveys generally provide the necessary empirical data for traffic analyses used in transport planning, traffic management, safety studies, etc. Depending on their application, several types of traffic data are collected in the field. The fundamental types of macroscopic traffic data are speed (km/h), flow (veh/h), density (veh/km), flow direction, number of turning movements, queue length, vehicle class, occupancy, headway and presence (Versavel, 2007). These data along with the geometric characteristics of a given section provide the necessary information to address

\* Corresponding author (nalinj@uom.lk;  <https://orcid.org/0000-0001-9493-5808>)



This article is published under the Creative Commons CC-BY-ND License (<http://creativecommons.org/licenses/by-nd/4.0/>). This license permits use, distribution and reproduction, commercial and non-commercial, provided that the original work is properly cited and is not changed in anyway.

typical traffic engineering problems. Microscopic traffic data include information such as the dynamics of individual vehicles and the manner in which they interact with adjacent vehicles in the traffic stream. Some of these microscopic parameters are individual vehicle speed, headway, position and characteristics of vehicles such as class, type, length, height, weight, length of the wheelbase and axle count (Versavel, 2007).

Both manual and automated traffic data collection methods are available at present for the purpose of data collection. Manual traffic data collection is the oldest as well as the most basic method currently in practice. This is usually carried out by employing enumerators to collect the relevant traffic data. This method is still useful at present since automated methods cannot accurately gather some data types such as vehicle occupancy, vehicle classification and pedestrian details (Leduc, 2008). However, the manual data collection method is neither cost effective nor reliable.

With the increase in transport infrastructure development projects in Sri Lanka, the need for accurate traffic data for feasibility studies, traffic forecasting studies, etc., is pertinent. At present, the authorities collect data using conventional manual methods which are time-consuming and prone to producing errors unless stringent data quality control measures are in place. The use of pneumatic tubes can be observed on some roads maintained by the Road Development Authority of Sri Lanka. Even so, the authorities are still in the process of finding and introducing alternative traffic data collection methods to increase the efficiency of project planning and designing. Although there are several methods available to cater to the above need, the responsible authorities have not yet studied the viability of such methods in the local context (Jayaratne *et al.*, 2016). This study is focused on conducting a review on the traffic data collection technologies and challenges faced in heterogeneous traffic data collection.

### Modern traffic data collection methods

Automated data collection methods are developed based on various scientific principles and phenomena. Time-lapse photography was used by Chari and Badrinath (1983) to take aerial photographs of traffic streams to estimate space-mean speeds. Several studies have used the simple videography technique where the videos are manually analysed later to collect traffic data (Nagaraj *et al.*, 1990; Kumar, 1994; Singh, 1999; Chandra, 2004). The Inductive Loop Detector (ILD) is another

type of non-intrusive traffic sensor, which is installed underneath the road pavement for traffic data collection. Piezoelectric sensor-based instruments (e.g. pneumatic tubes) convert mechanical energy into electrical energy, measuring the speeds and weights of vehicles (Swann, 2010). The Doppler principle is utilised in microwave radar-based instruments while infrared technology is used in other instruments for vehicle detection. Further, image processing which is a rapidly developing technology, is used to extract traffic data from videos captured at roadside locations (Leduc, 2008).

All aforementioned automated data collection methods have their advantages and disadvantages in the context of the present level of technological development. These instruments are vulnerable to the changes in the environment such as lighting conditions, weather, traffic conditions, obstructions, etc. Therefore, the reliability of these instruments is not 100 %. This can be observed by comparing control manual counts with the automated counts. Videography as a method for data collection is popular as the data can be visually observed later, which is useful for quality control purposes. Further, the cost incurred in this method is low compared to the other methods and has low human resource requirement. However, extraction and analysis of traffic data from a video is a time-consuming process. Hence, software programs such as TRAIS, COUNTcam, TrafficVision, TRAZER, MediaTD, Picomixer STA, etc., have been developed to automate this process. These software programmes primarily use image processing techniques for data analysis with the facility to verify outputs manually if required (Kalaanidhia *et al.*, 2015). Infrared based data collection is another commonly used method in automated traffic counters. The principle behind this system is the intervention of infrared beams. When a vehicle passes by and obstructs the infrared rays, it detects and counts the vehicle. This method has a vast number of capabilities based on how the technology is used including the ability to measure the speed, length and lateral placement of a vehicle.

In this research, the TRAZER software, TIRTL (the infrared traffic logger) instrument and the Google distance matrix application programming interface (API) based technique are tested for their suitability for data collection on Sri Lankan roads. TRAZER is a video-based software that provides traffic flow and speed data. TIRTL is an infrared (IR) based instrument that provides a wide array of traffic data including flow and speed. The Google Distance Matrix API is used in this study to collect traffic stream speed data using a method developed by Kumarage *et al.* (2017).

**METHODOLOGY**

**Field study locations**

To evaluate the three selected automated data collection methods, traffic surveys were conducted at three locations using each of the respective methods, and in addition a verified manual traffic count was carried out as the control study. The outputs of TRAZER and Google Distance Matrix API do not depend on the road cross-

section geometry as TIRTL does. Therefore, those two methods were only used at a single location. The analysis was conducted by comparing the manually collected control dataset with the data collected from the alternative techniques. A summary of the survey locations is shown in Table 1.

Since the most reliable method to obtain an accurate control sample is by analysing visual evidence, videos were recorded at all survey locations (Figure 1) along with the automated methods.

**Table 1:** Details of the survey locations

Location ID	Road name	Location coordinates	No. of lanes	Lane width (m)	Survey methods used
P1	A4 - High-level Road	6.844972, 79.954278	4	3.35	TIRTL, TRAZER, Videography
P2	AB11 - New Galle Road	6.728306, 79.898389	4	3.4	TIRTL, Google Distance Matrix Method, Videography
P3	A8 - Horana Road	6.706694, 79.941222	2	3.5	TIRTL, Videography



**Figure 1:** Data collection locations - P1, P2, P3

**TIRTL instrument**

The TIRTL instrument consists of two units, the transmitter and the receiver. Each has to be placed on either side of the road, next to the edge of the road carriageway. IR beams which are transmitted between these two units at tyre level are used for vehicle detection. TIRTL has the ability to classify vehicles into fifteen categories (Kalaanidhia *et al.*, 2015). They are as follows; bicycles, cycle rickshaws, two-wheelers, three-wheelers, tractors, tractors with trailers, SCV (2 axle small commercial vehicles), LMV (2 axle light motor vehicles), LCV (2 axle light commercial vehicles), MCV (medium commercial vehicle; includes 2 axle rigid truck and bus), HCV (heavy commercial vehicle; includes 3 axle rigid truck, articulated truck and bus), MAV (multi axle vehicle, includes rigid truck and articulated truck) and OSV (oversized vehicle).

In a study in 2010, Shou *et al.* compared the classification capability of vehicles of the TIRTL instrument under different weather conditions. It was observed that in clear weather conditions, fog, snow and rain, the TIRTL vehicle counts agreed very well with the actual counts, although during thunderstorms, the TIRTL instrument underestimated the number of vehicles.

Further, it was detected that the accuracy of counts was not equally distributed among different vehicle classes.

In this study, the TIRTL instrument was tested during sunny conditions at all three test locations P1, P2, and P3. The TIRTL instrument has to be set-up in such a manner that the IR beams are located approximately 60 mm above the road surface with a tolerance of -25 mm to +35 mm. This can be observed in Figure 2. Since various types of roads are available in Sri Lanka, the three test locations were selected in such a way that it test the instrument’s accuracy over varied road geometries. The road geometries of the locations P1, P2, and P3 are as listed below and illustrated in Figure 3 (not to scale).

- P1 – Four-lane road (normal cross-fall)
- P2 – Four-lane road (super-elevated section)
- P3 – Two-lane road (normal cross-fall)

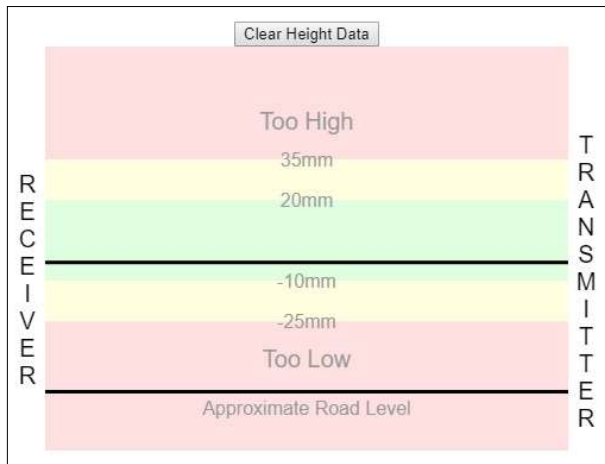


Figure 2: The data collection range of the TIRTL instrument

**TRAZER**

The TRAZER software uses image processing techniques on videos of traffic flows to collect speed and flow data. The videos to be processed through the software should be recorded parallel to the road and aligned to the centre of the lane/lanes with the vehicles moving towards the camera. The version of TRAZER software used for this research provides the user with the facility to detect 4 vehicle categories: namely, light moving vehicles (LMV), heavy moving vehicles (HMOV), three-wheelers (3W) and two-wheelers (2W). HMOVs can be classified further as buses (BUS) and trucks (TRUCK) manually through the software interface. Extracting flow and speed data using the TRAZER software is a four-step process, which includes a manual component where the user has to review the automatically identified vehicles. Through this process, the final accuracy of flow data can be elevated to 100 %. Mallikarjuna *et al.* (2009) used the TRAZER software to collect classified traffic volume, average occupancy, and average speeds. They observed that the detection accuracy was dependent upon the placement of the video camera with respect to the road. If the camera position deviates from the central lane, the detection accuracy decreases. Hence, the TRAZER software was tested at location P1 (Figure 1) where setting up the camera at the centre of the road was achievable.

**Google distance matrix API**

The Google Distance Matrix Application Programming Interfaces (APIs) facilitate traffic data collection such as travel distance and travel time for a matrix of origins and destinations. The API calls return the requested information based on the inputs given such as start and end

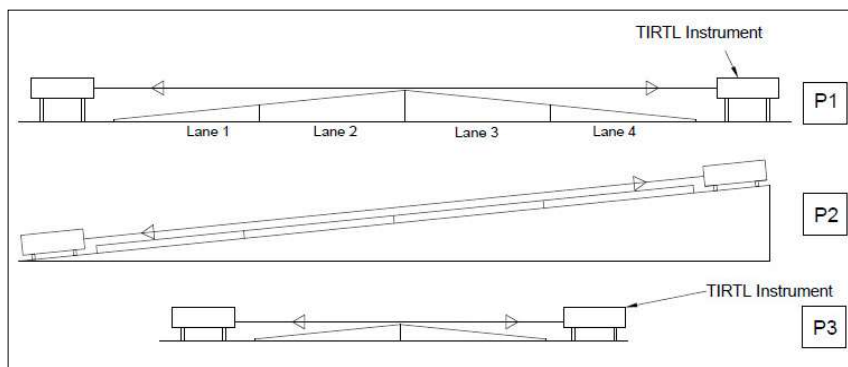


Figure 3: Loci of infrared beams-TIRTL instrument

points. The estimations such as travel time are calculated by the algorithms of Google Maps application. This feature can be used for traffic stream speed estimations of road segments of varying length. A study carried out by Kumarage *et al.* in 2017 estimated that the travel time can be predicted using Google Distance matrix API data to an accuracy of up to 99 %. The same methodology can be extended to predict traffic stream speeds of road links. For this study a 4.2 km straight section at location P2 was selected. The methodology developed by Kumarage *et al.* (2017) was used to collect travel time data through Google Maps. The calibration parameters are the ‘Travel mode’ and ‘Traffic model’. The use of ‘Driving’ travel mode ensures that the distance calculation is executed along the road network for a typical journey. Other travel modes such as ‘Walking’, ‘Cycling’ and ‘Transit’ provide data of either pedestrian sidewalks (where available), cycle lanes (where available) or public transit modes. Similarly, the traffic model - ‘Best Guess’ indicates that the returned ‘duration\_in\_traffic’ is the best estimate of travel time given what is learnt through both historical traffic data and live traffic data. More emphasis is placed on live traffic by the program when the time of request for travel time data is close to the actual ‘departure\_time’. Travel time data were collected at 1-min intervals on both sides of the road at the location.

## RESULTS AND DISCUSSION

The accuracy of classified vehicle flow counts by the TIRTL and TRAZER software is discussed in this section. The errors in the automated methods are calculated using equation 1.

$$Error = \frac{True\ count - x}{True\ count} \times 100\ %; \quad \dots(1)$$

where  $x = \text{count of automated method}$

### Flow analysis – TIRTL instrument

As can be observed from Table 2, the flow count values have varying error percentages for the three test locations. Locations P2 and P3 have low error values (less than 5 %) but location P1 has a significant error in the total count. Location P1 is a 4-lane road with a carriageway width of approximately 14 m and a normal crossfall of - 2.5 % in each direction. Hence the IR beams of the TIRTL instrument are out of its specified range for a significant portion of the road. Since the instrument was set up in the same manner at all locations it can be concluded that the errors were due to the geometry of the road.

**Table 2:** TIRTL flow summary - locations P1, P2, P3

Location	P1		P2		P3	
	Count	% Error	Count	% Error	Count	% Error
TIRTL	2994	- 15%	4146	-2%	5121	-4%
True Count	3529		4210		5325	

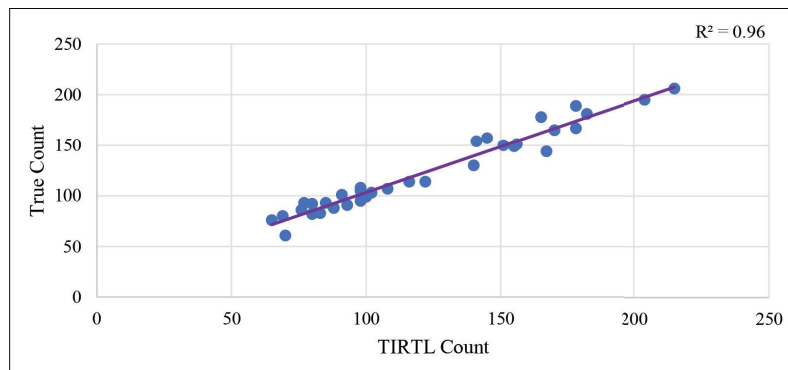
**Table 3:** Error percentage per lane in TIRTL location P1

Location P1	lane1	lane 2	lane 3	lane 4	Total
Actual count	855	829	945	900	3529
TIRTL	704	720	845	725	2994
Error	-18 %	-13 %	-11 %	-19 %	-15 %

As can be observed from Table 3, the errors in lanes 1 and 4 (outer lanes) are higher compared to those of the inner lanes. This is because the gap between the TIRTL instrument’s IR beam and the road surface is higher than the recommended range.

A similar issue was encountered at location P3 (two-lane road with normal -2.5 % crossfall), but due to the shorter carriageway width (7 m) the vertical rise of the road is lesser. Hence, an error of only -4 % is observed in the results at this location. On the other hand, the error in the vehicle count estimate was minimum (-2 %) at location P2 since there was no cross-fall at that section.

The TIRTL counts were plotted against the actual counts for locations P2 (Figure 4) and P3. The R<sup>2</sup> values of 0.96 and 0.98 obtained for the two respective locations indicate that the TIRTL instrument accurately estimates the flow values of Sri Lankan traffic on two-lane roads and multi-lane roads at super-elevated sections.



**Figure 4:** TIRTL counts against the true counts for location P2

At location P1, the  $R^2$  value was 0.85, indicating a lower accuracy in the predicted count. It was observed that motorcycles were the least captured vehicle category by the instrument. Only 67 % of the motorcycles were recognised, whereas 92 % of the other vehicle categories were identified. Consequently, it was attempted to build a model taking into account the percentage of motorcycles, along with the total TIRTL count and lane flow to predict the actual flow. However, no statistically significant relationship was found in the data sample considered.

Accordingly, it was established that the TIRTL instrument is not suitable to be used to estimate flow data on multi-lane roads with normal crossfalls. Alternatively, one unit of the TIRTL set up may be placed on the centre of the road to capture flow data on one direction of a multi-lane road. However, this is bound to cause disruptions to the traffic flow.

#### **TRAZER software**

As shown in Table 4, a total of 3,529 vehicles were analysed at location P1 using the TRAZER software. The analysis procedure of TRAZER has 4 main steps.

**Step 1:** Inputting geometric and vehicle class dimensions to the software and processing the video.

**Step 2:** Reviewing the automatically identified vehicles and deleting false vehicle recognitions.

**Step 3:** Reviewing the identified vehicles and confirming/classifying vehicles. In this step, vehicles that are identified but are in the wrong category are moved to the correct one.

**Step 4:** Adding the unidentified vehicles by reviewing the video manually using TRAZER software.

As seen in Table 4, the estimate provided by the TRAZER software after step 1 is incorrect by a margin of 843 vehicles. The error was calculated using equation 1 substituting counts from steps 1 to 4 to 'x'. An error of + 24 % was observed after step 1. Through further analysis, it was observed that LMV and 2W categories were overestimated by the software, whereas 3W and HMV categories were underestimated. Out of the 2W count of 1303, only 579 were accurate identifications.

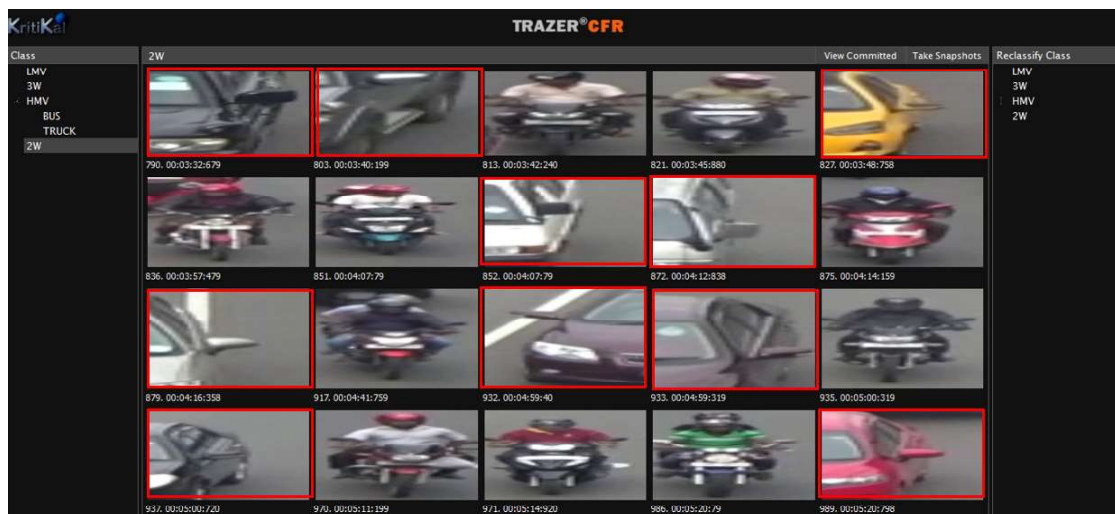
The remaining 724 were either false positives or incorrect classifications. This is a major factor that affects the initial estimate of vehicles. It was observed that vehicle side mirrors are identified by the software as 2W's leading to this error. This observation is shown in Figure 5.

Once steps 2 and 3 (deletion and reclassification) were completed, the total vehicle count estimated by the software was found to be 20 % less than the actual value. The HMV and 2W categories were incorrect by a margin of - 48 % and -27 %, respectively. This shows that the software is less capable of identifying vehicles with irregular dimensions (large and small). This is observed within the HMV category where only 37 % of buses were identified as opposed to the 67 % of trucks. The reason a higher percentage of trucks were identified is because medium sized trucks were misidentified by the software as 'LMV's. The estimate of LMV's were at an acceptable level of 87 %.

The final step is the addition of unidentified vehicles manually. This is a tedious and time-consuming process as the video needs to be analysed frame by frame to detect vehicles that have not been identified by the software. However, at the end of this process 100 % accuracy can be achieved.

**Table 4:** Summary of TRAZER results

TRAZER analysis		Step 1: process	Step 2: deletion	Step 3: reclassification	Step 4: true count
LMV	Count	2260	1540	1500	1720
	Error	31 %	- 10 %	- 13 %	-
3W	Count	573	497	544	667
	Error	- 14 %	- 25 %	- 18 %	-
HMV	Count	236	147	184	352
	Error	- 33 %	- 58 %	- 48 %	-
BUS	Count	-	-	66	177
	Error	-	-	- 63 %	-
TRUCK	Count	-	-	118	175
	Error	-	-	- 33 %	-
2W	Count	1303	623	579	790
	Error	65 %	- 21 %	- 27 %	-
Total	Count	4372	2807	2807	3529
	Error	24%	- 20%	- 20%	-



**Figure 5:** Misdetection of side mirrors as 2W's in TRAZER software

**Estimation capability of TRAZER software with flow**

The precision of the TRAZER software with the change flow values were analysed in the following manner. The flow was divided into 1-min time interval groups. This varied from 16 veh/min to 49 veh/min. Therefore, this was divided into three categories for ease of analysis as shown below,

- Low flow : 16–26 veh/min
- Medium flow : 27–37 veh/min
- High flow : 38–49 veh/min

The precision of the TRAZER software in each category was calculated using equation 2,

$$Precision = \frac{\sum_1^N \frac{Y}{Actual\ no.\ of\ vehicles\ per\ min.}}{N} \dots(02)$$

Where,

$Y$  = no. of vehicles per min after step 1, step 2  
 $N$  = no. of 1-minute intervals per category ( $N \geq 20$ )

From the results in Table 5 it is observed that precision of the software is not affected by the rate of flow.

**Table 5:** Precision of TRAZER under different flow conditions

Category	Precision	
	Step 1: Process	Step 2: Deletion
Low	115.2 %	82.6 %
Med	125.4 %	76.6 %
High	123.1 %	79.5 %

### Speed data analysis

#### *TIRTL speed analysis*

Speed data collected through the instrument were compared with speed data computed manually. The manual speed data calculation was carried out by analysing the video and calculating the time taken by vehicles to traverse a known distance. A sample of 177 vehicles were selected for the speed survey ranging between the speeds 81km/h and 12 km/h. The mean absolute error (MAE) of the data was 3.47 and the root mean square error (RMSE) was 4.65 while the mean absolute percentage error (MAPE) was 8.9 % (< 10 %), which are acceptable values denoting that the instrument was able to capture the speeds of individual vehicles with high accuracy. Figure 7 depicts the absolute percentage error of each data point in ascending order with MAPE drawn for reference.

#### *TRAZER speed analysis*

Traffic speed data of a group of 60 vehicles were collected by analysing the captured videos using TRAZER software. The data were compared with the corresponding actual speed data to evaluate the accuracy of the outputs of TRAZER using a similar methodology as used in the TIRTL speed analysis. The speed range of the surveyed vehicles were between 50 km/h and 19km/h. The MAE was 2.57 and the root RMSE was 3.31 while the MAPE was 2.6 %. According to the results

of the study it is observed that the TRAZER software predicts the speeds of vehicles at a higher accuracy than the TIRTL instrument. Figure 8 depicts the absolute percentage error of each data point in ascending order with MAPE drawn for reference.

### Google Maps – distance matrix API

Google Distance Matrix API provides traffic stream speeds as opposed to the individual vehicle speeds obtained through the TRAZER software and TIRTL. One minute stream speed data were collected through the programme and combined to five minute average traffic stream speed counts for a period of two hours and compared with the actual average stream speeds.

Considering the statistical data, it was observed that the speeds predicted are of high accuracy given that the MAE was 0.87, RMSE was 0.97 and the MAPE was 1.7 %.

The ability to evaluate individual vehicle speeds is not available through this method because Google distance matrix provides only the traffic stream speeds and not individual vehicle speeds. Since traffic stream speed is the parameter that is predominantly used in traffic engineering applications, this method can be successfully employed for data collection. A limitation in this study is that the speed values observed had a small spread (47–53 km/h) and the flow volume of vehicles during the study was between low to moderate (maximum directional flow 2470 veh/h). Nevertheless, it is understood that Google Distance Matrix API data can be used to estimate link speed at high accuracy and can be used as a substitute to conventional speed calculation methods.

### Comparison of speed estimation capability of TIRTL and TRAZER methods

In the comparison of the two automated speed detection methods at location P1, it was observed that TRAZER had a comparatively smaller spread in error in speed detection. This can be seen in Figure 9. The TIRTL instrument has a RMSE of 4.65 whereas TRAZER has a RMSE of 3.31. Hence, the TRAZER software has better precision in measuring speeds between the two methods. Since the Google distance matrix API speed data were collected at a different location it was not considered in this comparison.

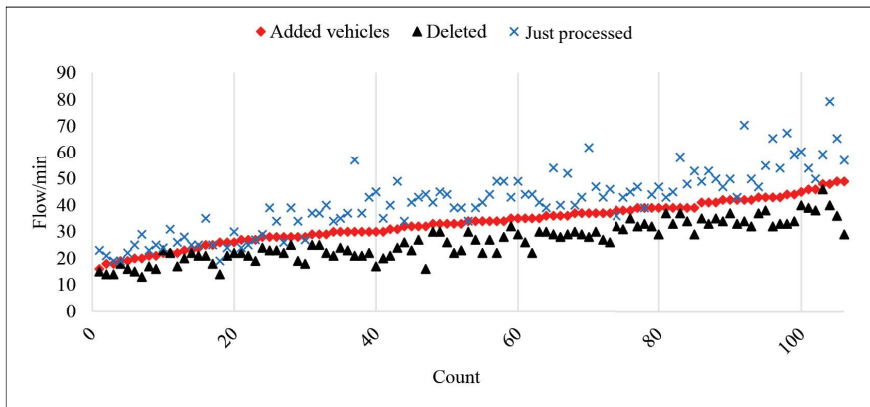


Figure 6: Comparison of TRAZER flow values

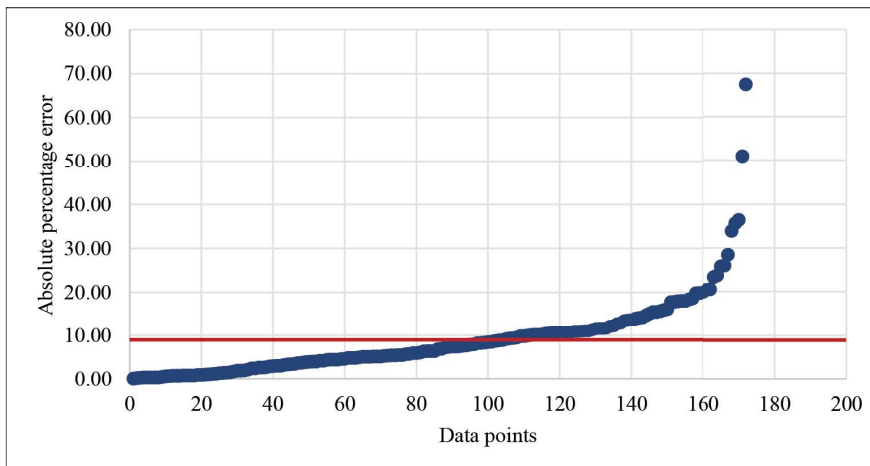


Figure 7: Absolute percentage error of TIRTL instrument

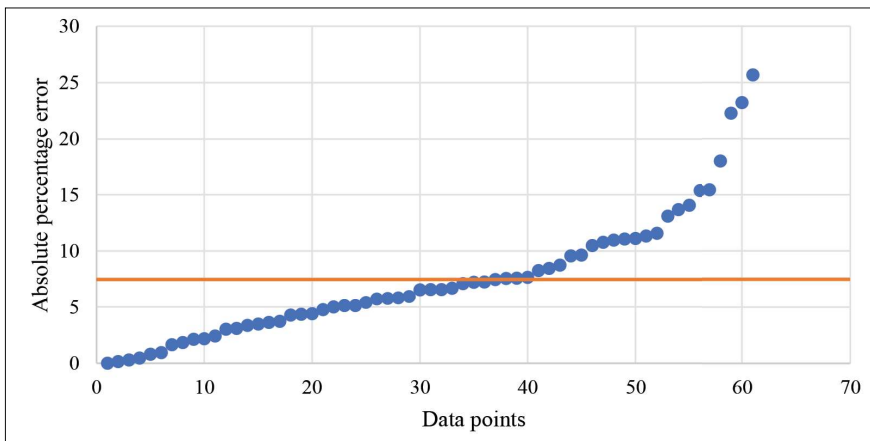
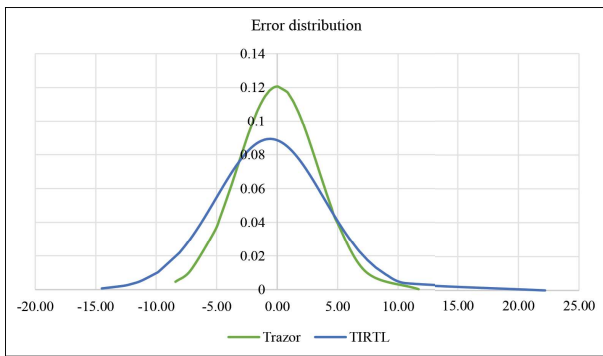


Figure 8: Absolute percentage error of TRAZER software



**Figure 9:** Error distribution in TIRTL and TRAZER

**Comparative analysis of data collection methods**

A comparison of primary traffic data that can be collected and the typical number of man-hours required for the three methods used in the study are shown in Table 6. The TIRTL instrument requires two individuals to position it and the process takes approximately 0.5 hours. A similar time period is required to set up the video camera to collect video data for the TRAZER software. The setting up of the script to collect traffic data using Google Distance Matrix API requires less than 0.5 man-hours. No post-processing is required in the TIRTL instrument and Google Traffic Data to extract the traffic data but the TRAZER software has a time-consuming analysis

**Table 6:** Comparison of traffic data collection technologies

Technology used	Traffic data				Man-hours		
	Speed	Flow	Classification	Headway	12 hour survey	Post processing	Total
TIRTL	Yes	Yes	Yes	Yes	0–1	-	0–1
TRAZER <sup>a</sup>	Yes	Yes	Yes	No	0–1	36–48	36–48
Google Traffic Data	Yes <sup>b</sup>	No	No	No	0	-	0–1

<sup>a</sup> Requires post processing

<sup>b</sup> Traffic stream speed. Individual vehicle speeds cannot be extracted

procedure. This is one of the major drawbacks of the software. When considering the traffic variables presented by the three methods, the TIRTL instrument gives a rich set of data including speed, flow, vehicle classification and headway. The TRAZER software provides speed, flow and vehicle classification, whereas Google Traffic data provides only the traffic stream speed. If Google Traffic data is to be used to collect fundamental traffic data, manual flow counts will have to be carried out along with the speed survey (approximately 60 man-hours: 5 enumerators). When comparing this with the other two methods, it is observed that using Google Traffic data requires a higher investment of labour in cases where all traffic data types are required.

**CONCLUSIONS**

This study incorporates a few of the latest technologies available in the field of traffic data collection in order to test their applicability to the heterogeneous traffic conditions observed in Sri Lanka. It was observed that these technologies provided satisfactory results with a few exceptions. The TIRTL instrument provides the

user with a comprehensive set of data with more vehicle classification, as well as macroscopic data such as headway spacing and lateral placement. The instrument can be set up in road sections where the road camber does not interfere with the cross beams of the instrument. The TRAZER software on the other hand provides flow and speed data. A disadvantage of using TRAZER software is that it requires the video to be captured from a high elevation parallel to the roadway for processing purposes. This limits the software’s usability as it is typically inconvenient to set up the camera at higher elevations at the centre of the road. Further, the processing of the software is a time-consuming venture albeit producing a result of 100% accuracy in terms of the vehicle count. In conclusion, the TIRTL instrument was found to be more practical to use and provides the user with a more comprehensive set of data.

Google Distance Matrix API is a convenient way of collecting traffic data. Data collection can be automated, and it is suitable for studies that require collection of data with a high frequency at multiple locations. Google travel times are estimated with both historical and real-time data, and therefore, the accuracy is high as observed

in this study. When compared with manually collected data, it showed an acceptable level of compliance which suggests that this method is suitable to be adopted in transportation engineering related research. Further studies should be carried out by changing traffic, geometric and other variables such as composition, flow, weather conditions, road conditions, network coverage and driver behaviour to evaluate the accuracy and reliability of the Google Distance Matrix API for mining data for transportation engineering related studies.

Overall, from this study, it is understood that technologies such as TIRTL, TRAZER and Google distance matrix APIs, when applied appropriately can be used to reliably collect traffic data in Sri Lankan conditions.

### Acknowledgement

This research was supported by a Senate Research Committee (SRC) Grant of the University of Moratuwa (SRC/LT/2017/04).

### REFERENCES

- Chandra S. (2004). Effect of road roughness on capacity of two-lane roads. *Journal of Transportation Engineering* **130**(3): 360–364.  
DOI: [https://doi.org/10.1061/\(ASCE\)0733-947X\(2004\)130:3\(360\)](https://doi.org/10.1061/(ASCE)0733-947X(2004)130:3(360))
- Chari S.R. & Badarinath K.M. (1983). Study of mixed traffic stream parameters through time lapse photography. *Highway Research Bulletin* **20**: 57–83.
- Jayarathne D.N.D., Jayasinghe P.W.P.R. & Pasindu H.R. (2016). Evaluation of level of service for two-lane roads in Sri Lanka. In: Transactions-2016 PART B Technical Papers. *Proceedings of the 110<sup>th</sup> Institute of Engineers Sri Lanka (IESL) Annual Sessions*, Colombo, Sri Lanka, 14 October. Institute of Engineers Sri Lanka, pp. 367–374.
- Kalaanidhi S., Gunasekaran K., Badhrudeen M.M. & Velmurugan S. (2015). *Review of Data Collection Methods for Establishing the Capacity of Intercity Highway*. Elsevier B.V., Netherlands.
- Kumar V.M. (1994). A study of some traffic characteristics and simulation modeling of traffic operations on two lane highways. *PhD thesis*, Indian Institute of Technology, Kharagpur, India.
- Kumarage S.P., Rajapaksha R.P.G.K.S., De Silva G.L.D.I. & Bandara J.M.S.J. (2017). Traffic flow estimations for urban roads based on crowdsourced data and machine learning principles. *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering* **222**: 263–273.  
DOI: [https://doi.org/10.1007/978-3-319-93710-6\\_27](https://doi.org/10.1007/978-3-319-93710-6_27)
- Leduc G. (2008). Road traffic data: collection methods and applications. In: *Working Papers on Energy, Transport and Climate Change*. European Commission, Joint Research Centre, Institute for Prospective Technological Studies.
- Mallikarjuna C., Phanindra A. & Rao K.R. (2009). Traffic data collection under mixed traffic conditions using video image processing. *Journal of Transportation Engineering* **135**(4): 174–182.  
DOI: [https://doi.org/10.1061/\(ASCE\)0733-947X\(2009\)135:4\(174\)](https://doi.org/10.1061/(ASCE)0733-947X(2009)135:4(174))
- Nagaraj B.N., George K.J. & John P.K. (1990). A study of linear and lateral placement of vehicles in mixed traffic environment through video recording. *Highway Research Bulletin* **42**: 105–136.
- Shou L., Yingzi D. & Yi J. (2010). *Site Verification of Weigh-In-Motion Traffic and TIRTL Classification Data*. Publication FHWA/IN/JTRP-2010/26. Joint Transportation Research Program, Indiana Department of Transportation and Purdue University, West Lafayette, Indiana, United States.  
DOI: <https://doi.org/10.5703/1288284314247>
- Singh B (1999). Simulation and animation of heterogeneous traffic on urban roads. *PhD thesis*, Indian Institute of Technology, Kanpur, India.
- Swann S. (2010). *Albury City Comparison of Traffic Data Collection Methods*. Albury City, New South Wales, Australia.
- Versavel J. (2007). Traffic Data Collection: Quality Aspects of Video Detection. *Proceedings of the 14<sup>th</sup> World Congress on Intelligent Transport Systems*, Beijing, China, Intelligent Transport System (ITS) Japan, Intelligent Transportation Society of America & ERTICO ITS Europe, pp. A–X.