AN INVESTIGATION OF POSITIONING ACCURACY TRANSMITTED BY CONNECTED HEAVY VEHICLES USING DSRC

Adriana Simona Mihăiţă, Corresponding author
Research Scientist, DATA61|CSIRO
13 Garden St, Level 5, Eveleigh, 2015, NSW
Tel: (+61) 02 9490 5615; Email: simona.mihaita@data61.csiro.au

Paul Tyler
Senior Research Scientist, DATA61| CSIRO
13 Garden St, Level 5, Eveleigh, 2015, NSW
Email: paul.tyler@data61.csiro.au

Aditya Menon
Senior Research Scientist, DATA61| CSIRO
13 Garden St, Level 5, Eveleigh, 2015, NSW
Email: aditya.menon@data61.csiro.au

Tao Wen
Research Assistant, DATA61| CSIRO
13 Garden St, Level 5, Eveleigh, 2015, NSW
Email: tao.wen@data61.csiro.au

Yuming Ou
Senior Research Scientist, DATA61| CSIRO
13 Garden St, Level 5, Eveleigh, 2015, NSW
Email: yuming.ou@data61.csiro.au

Chen Cai
Senior Researcher, Leader of Advanced Data Analytics in Transport group, DATA61| CSIRO
13 Garden St, Level 5, Eveleigh, 2015, NSW
Email: chen.cai@data61.csiro.au

Fang Chen
Research Group Manager, DATA61|CSIRO
13 Garden St, Level 5, Eveleigh, 2015, NSW
Email: fang.chen@data61.csiro.au

Word count:
Abstract (249) + Text (4476) = 4 725 words,
(2tables+6 Figures) x 250 words (each) = 2 000 words,
References = 770 words
Total = Abstract + Text + Tables/Figures + References = **7 495** words

Submission Date
01/08/2016

ABSTRACT

Recent developments in advanced transport technologies such as vehicle-to-vehicle communications and Dedicated Short Range Communications (DSRC) led to an increased interest in building safety vehicular applications that would prevent traffic collisions. Such applications need a high level of performance and positioning accuracy in order to meet critical levels of road safety. However, there is still a lack of practical performance measurements of DSRC equipped systems, especially on a high number of heavy vehicles operating in large and diverse areas.

This paper presents the results obtained from a research investigation undertaken into the capabilities of DSRC technology for meeting the positioning accuracy of road safety applications. The available data sets contain almost 400 million Basic Safety Messages (BSM) transmitted by 58 heavy vehicles equipped with DSRC, operating on a daily basis on a 42 km test bed area in Illawarra, Australia. Firstly, as ground truth is not available, we conduct a comparative analysis of positioning in the transmitted BSMs by using both Open Street Map and Google Street Map as reference, and show that the latter provides better accuracy in positioning error computation. Secondly, we present the results obtained when analyzing the five most active trucks of the fleet, as well as the noise-prone areas in which false collision alerts can be generated. Thirdly, we apply gradient boosted decision trees on the data sets and identify the three most important factors that influence DSRC transmitted positioning error in heavy vehicles.

Keywords: DSRC, connected vehicles, gradient boosted decision trees, positioning accuracy.

1. INTRODUCTION

Traffic congestion and road vehicles collisions are one of the most important problems in concentrated urban areas around the globe, leading to almost 1.24 million road traffic deaths per annum (*1*). Current trends suggest that by 2030 road traffic accidents will become the leading cause of deaths unless urgent action is taken (*2*). In order to address this issue, intelligent transportation systems (ITS) have become essential in investigating problems of vehicular transportation and improve road safety (*3*). Advanced transport technologies such as Vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications are already being tested and recent studies show the benefits of adopting these technologies in terms of life-savings and economic impact (*4*).

Recent advancements in wireless communication technologies have led to the emergence of dedicated short-range communication (DSRC), which has been designed to support V2V communications, enhance mobility and improve road safety (*5*). As vehicular communications need fast interoperability, in the U.S., a dedicated bandwidth of 75 MHz in the 5.850-5.925 GHz band has been assigned for DSRC, together with the IEEE 802.11p standard (*6*). Similarly, Europe and Japan have also established dedicated DSRC bandwidths (*7*). In order to assess the performance and safety benefits of DSRC, various projects and test bed initiatives have concentrated on: testing the effective communication range between two vehicles and security protocols (*8*), analysing the probability of successful message reception (*9*), detecting collision situations and send drivers early alerts (*10*), analysing collision timing (*11*), or investigating signal priority for connected vehicles (CV) at signalized intersections (*12*). Despite a high DSRC reliability indicated by these studies, in 2014, the National Highway Transportation Safety Administration (NHTSA) published the need to further investigate open research problems before establishing rule-making for a deployment-level V2V communication system mandate (*13*).

One of the biggest problems to address when using DSRC for building safety applications such as proximity collision alerts, automated braking, intersection signal alerts, etc., is to have an accurate vehicle positioning capability. Currently this is provided by a Global Navigation Satellite System (GNSS) (*14*). Although in ideal operating conditions (clear sky, no obstructions), GNSS can usually meet the positioning accuracy for most DSRC applications, in dense urban areas, high multi-paths or tunnels the GNSS signal can be limited or contains inaccurate positioning (*15*). Some CV applications need sub-meter accuracy at the lane level, especially for real-time situational awareness (*16*). Bridging the gap between positioning accuracy and the necessary availability for CV applications represents an important challenge still to be tackled. In (*17*) the authors proposed a Bayesian approach for using received signal strength data from roadside equipment (RSE) to update and improve GPS positioning. While this approach can work well when RSE is available and ready to use, many test beds have insufficient RSE or they are located at sparse locations throughout the study network. Other studies propose integrating GNSS and navigation information such as map data (*14; 18*), which contains “metadata” for travellers. However, until such maps are developed and shared across a large fleet, the cost to maintain a huge map database can become prohibitive especially for rapidly growing cities.

Recent developments have investigated the use of cooperative positioning (CP), which aims to enhance location accuracy of GNSS or to provide position data when GNSS is not available in vehicular ad hoc networks (VANETs). CP systems use data fusion methods to combine position-related data transmitted among a group of participating vehicles that can communicate with each other, and thus improve positioning accuracy. While conventional CP systems (differential GPS, real-time kinematic GPS, assisted GPS, etc.) may suffer from limitations such as low signal coverage, weak signals or accuracy (*19*), modern CP methods are defined based on vehicle-to-vehicle and vehicle-to-infrastructure communications and are applicable even when GNSS position data is unavailable (*20*) or when the number of visible satellites is small (*21*). However, ongoing studies have confirmed that even modern CP systems present constraints in radio ranging/ range rating and are not yet capable of bridging the gap and addressing the positioning accuracy required for safety applications, which is under a meter (*15*). In addition, these models normally require particular sensors with high computational complexity.

While most of the research studies that focus on positioning accuracy problems are undertaken on a small number of vehicles equipped with DSRC and on a limited test area bed, there is a real need for analysing the GPS positioning accuracy transmitted by a large number of vehicles, over a longer period of time and under various traffic conditions. The Cooperative Intelligent Transport Initiative (CITI) is a project currently undertaken by Transport for New South Wales (TfNSW), with the aim of building Australia's first semi-permanent test-bed for testing the DRSC technology over an area of 917 km^2 in the Illawarra Region of NSW south of Sydney (*22*). Currently, sixty vehicles (mostly heavy vehicles), three signalised intersections and one roadside location have been equipped with DSRC units. In order to ensure road safety, one of the main problems of the project is to address the generation of false collision alerts that would hinder driving and might results in drivers ignoring or not truting the DSRC on-board-unit warning device. The first step to identify the possible cause of false alerts is to investigate the accuracy of the transmitted positioning between the trucks, as reported from Basic Safety Messages (BSMs).

In this paper, we present the procedure, results and analysis we have undertaken in order to investigate the current GPS positioning accuracy of selected DSRC equipped vehicles involved in CITI. The main objectives of this study are:

a) investigating and characterising the error (noise) in the DSRC GPS positioning,
b) identifying "noise – prone sections" of the road network that would cause high levels of noise to be registered,
c) identifying potential factors that would impact noise in the GPS positioning.

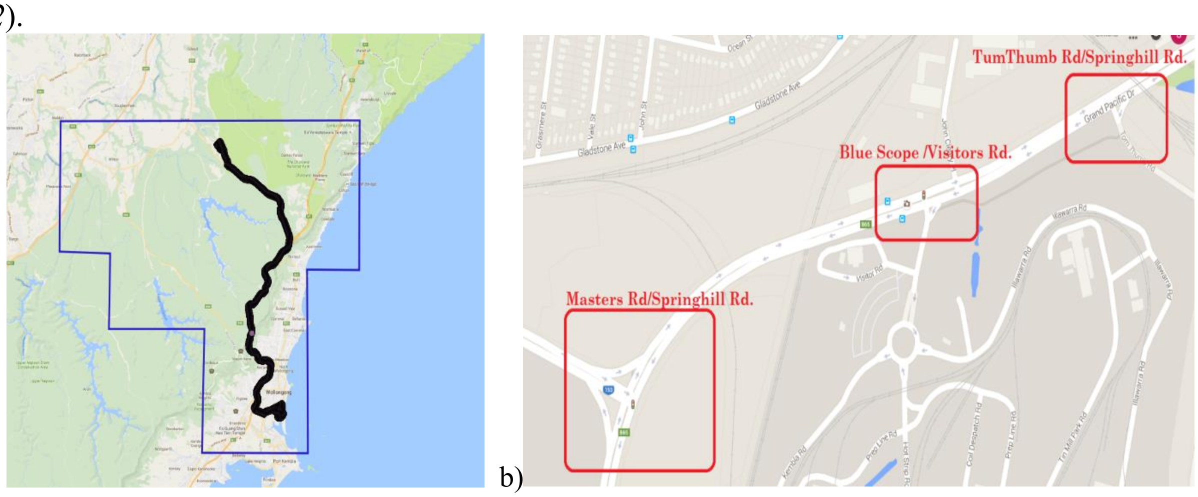
In Section 2, we present the CITI project background and main challenges. Section 3 presents the data sources and processing, as well as the map-matching procedure for computing the noise in transmitted GPS. In Section 4, we conduct a positioning analysis and comparison for the five most active trucks which have been selected for this study. We also analyse the most important features that influence noise, as obtained from applying gradient boosted decision trees over the collected data sets. Conclusions and further perspectives of this work are addressed in Section 5.

2. CITI PROJECT BACKGROUND

The Cooperative Intelligent Transport Initiative (CITI) is a project deployed by Transport for NSW (TfNSW) in partnership with Data61 and the Australian Federal Government's Heavy Vehicle Safety Productivity Program. The main goal of the project is to assess V2V/V2I communication technology that could reduce the number of road accidents, with a focus on the Illawarra region. According to TfNSW, 18% of the traffic on Picton Rd (a road in the CITI area) consists of heavy vehicles, which are involved in 63% of fatal crashes (*23*). Recent studies in Australia have shown that the total cost of crash accidents with fatal injuries for the years 2006-2010 is estimated at almost $6.9billion in economic loss (*24*). In order to address this problem and the high cost generated by truck accidents, CITI project aims at building a semi-permanent test bed for evaluating and further testing of the Cooperative Intelligent Transport Systems (CITS) technology, especially DSRC equipped vehicles.

2.1 Current deployment and location

The focus area for the CITI project represents a 42 km route between Port Kembla and the Hume Highway/Picton Road intersection (as represented in Figure 1a). During the first stage, the project has installed DSRC devices in 58 heavy vehicles, 2 light vehicles, 3 signalised intersections (Figure 1b) and 1 roadside unit at the top of Mt. Ousley near Wollongong, NSW. CITI currently utilises Cohda Wireless MK4 and MK5 DSRC (*25*) units running Cohda's alert software in vehicles and roadside software for infrastructure deployment. Cohda DSRC systems are using the US standards of IEEE 1609 family, SAE J2735 and IEEE 802.11p standards. The heavy vehicles are usually equipped with 2 MobileMark ECO6-5500 DSRC antennas placed near the mirrors of the trucks, and one MobileMark SM-1575 GPS Antenna often placed in the vehicle, under the dashboard. Software on the units include a dead-reckoning feature. Inside the vehicles, the DSRC unit is connected to a Nexus 7 tablet for audio and visual display of generated alerts, such as Forward Collision Warning (FCW), Intersection Collision Warning (ICW), Electronic Brake Light Warning (EBLW), as well as two custom alerts. The custom alerts are a red light ahead warning based on Signal Phase and Timing broadcasts and a heavy vehicle speed restriction monitoring application that alerts drivers if they exceed a 40km/h restriction on a steep descent in the trial area (*22*).



a) b)

FIGURE 1 a) CITI area with with an example of daily truck trip. b) DSRC equipped intersections (Google maps).

2.2 Problems and challenges

Currently, there are over 150 drivers from 3 transport companies involved in daily trips from Port Kembla near Wollongong NSW to a colliery near Appin Rd, NSW. Many of the truck drivers make up to 7 trips per day with the trucks operating in two shifts 24 hours a day, 7 days a week. Vehicles in the trial are broadcasting their position 10 times a second in a message known as the Basic Safety Message (BSM). The positioning information in these messages is based on GPS measurements and in the case of brief loss of the GPS signal the position may be extrapolated from last known data in a process known as "dead reckoning". The fusion of GNSS and Inertial Navigation System data for dead-reckoning is common in intelligent vehicles field. For example, during GNSS outages, the dead reckoning estimates the location of the vehicle (*26*). However in CITI, no additional sensors are connected to the DSRC unit and dead reckoning is restricted to interpolations from GPS locations. For the remainder of this article, reference to "GPS" is actually a reference to "GPS-based positioning information" as broadcasted by a vehicle in a BSM. Since the beginning of the project, the vehicles have generated more than 400 million BSMs to be analyzed and tested for positioning accuracy.

To date there has been little data analysis for DSRC equipped vehicles operating in CITI. An initial aim is to examine the DSRC positioning accuracy in the Australian setting, which includes a range of urban and mountain environments, with isolated rural areas and coalmines. Large variations in the transmitted location to other connected vehicles can trigger false collision alerts, or hinder driver response to alerts. As road safety is the main focus of the CITI project, a major concern is to identify the risks that divers face when exposed to false alarms or when false and correct alarms cannot be distinguished. Therefore, the main objective is to understand how the positioning accuracy of DSRC equipped vehicles changes over time and how much the GPS positioning error varies relatively to previously reported locations. This is an important topic to be explored and to understand if the BSM based GPS data is suitable for conducting further analysis or for detecting changes in the driving behavior when collision alerts are received.

This investigation looks at the noise evolution in locations reported in the BSMs over time, in various places and from various vehicles. Due to limited space, we limit our analysis to the 5 most active trucks in the data sets. As ground truth is not available for identifying the accuracy of the reported GPS location, this investigation uses the closest mapped road position from both Google Street Maps (GSM) and Open Street Map (OSM) to determine the "error" or noise in the DSRC transmitted GPS position. A detailed description of data processing and map matching method is provided in Section 3. As well, another important challenge is to identify the factors that lead to significant errors in the transmitted GPS positioning, based on the available data sets. For this purpose, we apply gradient boosted decision trees and identify the most important factors that can influence noise in GPS positioning, as discussed in Section 4.

3. DATA PROCESSING METHOD

3.1 Data sources and processing

For the purpose of this study, we have received from TfNSW, almost 400 million DSRC messages transmitted by the trucks operating in the CITI project, collected between July 2015 and November 2015. The data was collected by two equipped trailers in Port Kembla and contains all transmitted and received DSRC messages, including BSMs. After initial data format reading and verification, we batch process and extract only the necessary messages and fields for data analysis. In our case, we process the Basic Safety Messages, including positioning, speed, heading, acceleration, brakes, elevation, timing, etc.

3.2 Map-Matching

After the positioning points are extracted from the raw positioning data, an important step is to establish the basis for comparison of these positions to "ground truth". Such a comparison would establish the error/noise in the positioning information broadcasted in the BSM from the true position. Unfortunately a proper "ground truth" – the true position of the vehicle – is not available and very expensive to measure. However, by using Map-Matching (MM) algorithms we can integrate positioning data with spatial road network data (roadway centerlines) to identify the correct link on which a vehicle is travelling and to determine the location of a vehicle on a link (*27*). Due to the nature of the data sets, we apply a classical post processing map-matching algorithm and emphasize more on accuracy than computation efficiency (*28*). As our main purpose is to be able to identify noise-prone road sections, we focus on the data analysis and noise comparison by using either GSM or OSM for noise calculation, and the regression models for identifying factors that influence noisy GPS observations. While OSM shapefiles for identifying road centers are free to access and use in the MM procedure, in order to map positioning the GPS observations to road centers reported by GSM, we use the Google Snap to Roads API. The mapped positioning points have then been used to compute the Vicenty distance (*29*) between transmitted GPS locations and GSM.

3.3 Notations and noise calculation procedure

In the following, we denote D as the total number of trucks under the study. For each truck, $d \in \{1,..D\}$, we have a total number of GPS observations N_d^{GPS} extracted from BSMs. Each GPS observation is described by its location: $x_i = (L_i, l_i),\ i \in \{0,..N_d^{GPS}\}$ registered at time t_i, where L_i and l_i denote the longitude, and the latitude respectively. The total time travelled by a single truck is denoted by T_d, which can contain various trips conducted by the truck over multiple days since the beginning of the trial. Let $\Delta t_i = (t_{i-1}, t_i), i \in \{0,..N_d^{GPS}\}$ be the time interval between two consecutive GPS observations, which in our case is set to 0.1 seconds, according to the DSRC specifications. Each GPS observation $(x_{i\,,} t_i)$ can be mapped to a specific road segment $g_j, j \in \{1, G\}$, which can contain sequential GPS observations with the same spatial-temporal characteristics. A symbolic graphical representation of three consecutive GPS observations over a selected road section is provided in Figure 2.

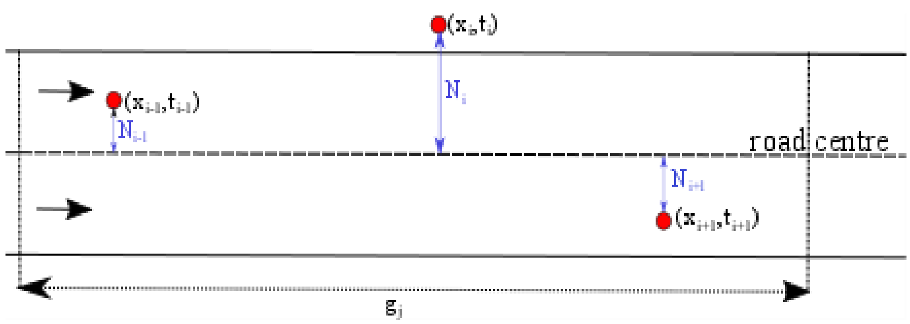


FIGURE 2. Examples of GPS observations

Let N_i be the distance (deviation/noise) between a registered GPS location and the centre of the road section at time t_i, and $\overline{N}$ the mean noise observed on a selected road section. We also note A_i as the anomaly detected at time t_i :

$$A_i = \begin{cases} N_i, if\ N_i > 8\ m, \forall i \in \{1, .., N_d^{GPS}\} \\ 0, otherwise. \end{cases}$$

As the road sections we are investigating have in general 2 lanes, each of 3.5 meters, we consider that any computed distance which is bigger than 8 meters to be recorded as an anomaly in the noise computation. Therefore, the steps we have applied for detecting noise anomalies for each vehicle, are the following:

1) Consider a road section $[\mathrm{A}, \mathrm{B}]$ defined by a starting point A and ending point B.
2) Apply a MM procedure for identifying the trajectory of the DSRC GPS positioning.
3) Compute N_i deviations from the road center for each intermediary points between $[\mathrm{A}, \mathrm{B}]$.
4) Compute mean deviations (noise) on the selected road section ($\overline{N}$), for all available trips undertaken during the total travel time of a truck T_d.

This procedure has been applied for all heavy vehicles and some selected results will be presented in the following section.

4. POSITIONING ANALYSIS FOR TRUCKS

4.1 Single transmission file analysis

Before presenting the noise results of the trucks, we show the analysis conducted over a single transmission file, belonging to the most active truck, which contains a typical daily trip of a truck from Port Kembla to a nearby colliery, as represented in Figure 1a).

This transmission file contains 87,165 BSMs, recorded between 20:46:13 and 23:12:58 on the 20th of July 2015. For an accurate analysis, we filter GPS positions that indicate stopping in parking areas or inside the mine area. By using GSM as ground truth, we obtain an average noise of 2.9762 meters, with certain GPS points exceeding 8 meters threshold and reaching a maximum of 12.1415 meters from the road center, as represented in Figure 3 a) right.

A special area of the selected road section is the Mt. Ousley area (Figure 3b) left) which has a speed restriction of 40 km/h for trucks descending the mountain. Therefore, on this road section the GPS accuracy will generally be obtained at lower travelling speeds. The average noise obtained in this area is lower (2.3836 meters) and the overall noise under 7 meters (Figure 3b) right). The good accuracy in the GPS positioning can be influenced by truck speed, as we will discuss in Section 4. Regarding the continuity of the GPS signal, and the consistency between consecutive GPS points we have observed noise variations that can go up to a maximum of 15 cm between consecutive BSMs with $\Delta \mathrm{t}_i = 0.1\, sec$ (Figure 4). While a continuous variation in between consecutive GPS points can indicate that the truck is changing lanes and heading in another direction, some other variations between registered BSM seem not to be consistent from previous ones, and might indicate some deviations in the GPS location transmitted by the DSRC system.

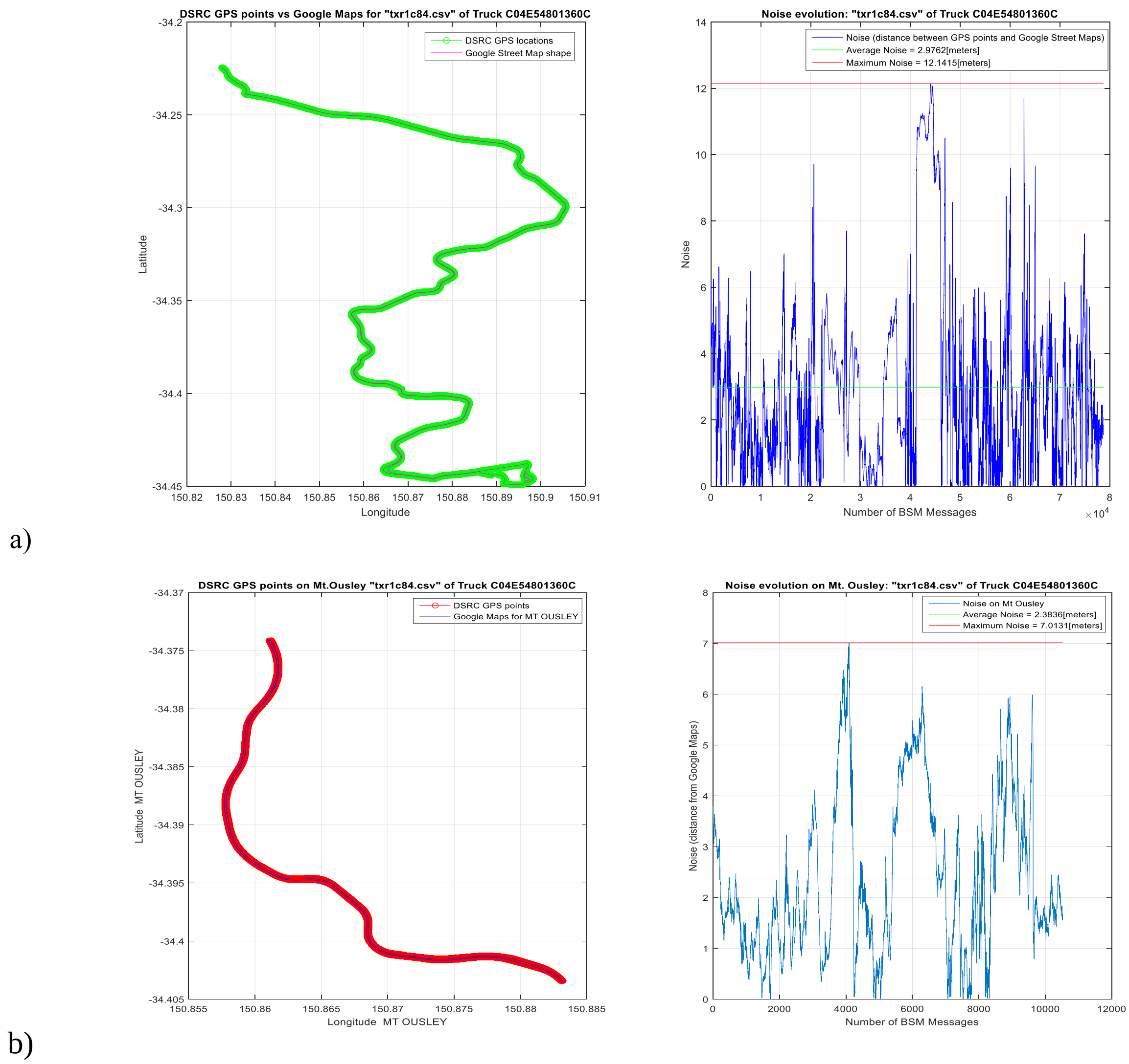


FIGURE 3 Noise mapping and evolution on: a) selected road section b) Mt. Ousley.

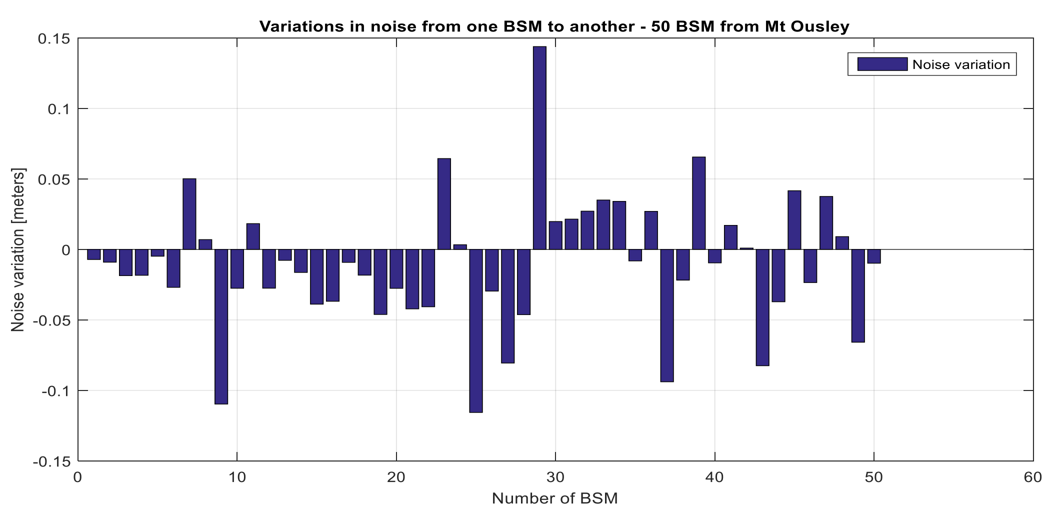


FIGURE 4. Variations between consecutive GPS locations from 50 BSMs on Mt. Ousley.

4.1.2 GSM and OSM noise comparison

We apply the same noise computing method by using as well Open Street Map. Table 1 reports the average and maximum noise level obtained for the single transmission file when using both GSM and OSM.

Table 1 Comparison between GSM and OSM noise for one transmission file example.

	Google Street Maps		Open Street Maps			
	Average noise[m]	Maximum noise[m]	Average noise [m]	Maximum noise[m]	Difference [meters]	Error [%]
Road section	2.9762	12.1415	3.2883	12.6679	0.3121	10.48 %
Mt. Ousley	2.3836	7.0131	2.7480	8.0559	0.3644	15.28 %

These initial results on a single transmission file show a more accurate GPS positioning when using Google Street Map as the "ground-truth". The average noise when using GSM is smaller than the noise obtained when using OSM. We observe that there is a difference that can vary between 31cm and 36cm between the two pseudo ground truth references, which can influence the final noise results. Based on these initial findings, for the rest of the results presented in this paper we will consider GSM as the ground truth for noise calculation.

4.2 Truck positioning analysis

In this section, we present the data analysis and interpretation we have conducted for the five most active trucks over the selected road section including Mt. Ousley, presented in Section 3. For easing the notations, we will denote the trucks as "$Truck\ i, i = 1, ..5$".

A summary of the total number of investigated BSMs, total dates and detected anomalies for each truck is provided in Table 2. All trucks are coal carriers and are doing daily trips on from Port Kembla to a coal mine near Wollongong. In terms of total number of BSMs, we note that Truck 1 appears to be the most active, with almost 4.75 million BSMs transmitted during the testing period in the Illawarra region, followed by Truck 2. From the total number of transmitted messages, after filtering the BSMs sent on the selected road section, we observe that each truck has different and sometimes unique activity. Truck 2 seems to have a higher transmitting activity in this area, gathering 903,209 BSMs. In terms of detected anomalies, Truck 1 and 2 present again a higher number of deviations from the road center, compared to the last 3 trucks. Truck 2 is the one which registered the biggest number of anomalies, representing 4.69% of its total number of BSM positioning points. As well, on Mt. Ousley road section, Truck 2 has registered 24,234 anomalies (8.65%) compared to Truck 1 (1.08%), which is comparably bigger than noise recorded for other trucks.

Table 2 Description of BSM statistics and noise anomalies for each truck.

		Start Date	End Date	Number of BSMs
Truck 1	All road sections registered by DSRC	Jul 3, 2015 15:03:45.189162000	Nov 3, 2015 19:47:24.243018000	4,749,912
	Selected road section	Jul 15, 2015 16:38:07.995023000	Oct 27, 2015 15:56:14.490091000	711,601
	Anomalies on selected road	Jul 15, 2015 16:40:15.195829000	Oct 27, 2015 08:44:00.326029000	42,342 (5.95%)
	Mt. Ousley road section	Jul 15, 2015 16:55:09.595133000	Oct 27, 2015 08:44:06.946928000	186,907
	Anomalies on Mt. Ousley	Oct 27, 2015 08:40:38.025210000	Oct 27, 2015 08:44:00.326029000	2,024 (1.0829%)
Truck 2	All road sections registered by DSRC	Aug 22, 2015 23:12:01.866107000	Oct 30, 2015 05:26:16.002918000	3,732,178
	Selected road section	Aug 24, 2015 01:04:00.246601000	Oct 29, 2015 03:31:14.630633000	903,209
	Anomalies on selected road	Aug 24, 2015 01:47:23.246482000	Oct 28, 2015 12:44:06.390510000	42,363 (4.69%)
	Mt. Ousley road section	Aug 24, 2015 01:06:09.146352000	Oct 29, 2015 03:30:38.730758000	280,057
	Anomalies on Mt. Ousley	Sep 6, 2015 13:48:56.872012000	Oct 19, 2015 09:33:46.141116000	24,234 (8.65%)
Truck 3	All road sections registered by DSRC	Aug 22, 2015 10:50:13.875742000	Nov 2, 2015 23:14:16.176066000	2,766,201
	Selected road section	Aug 22, 2015 17:04:35.079759000	Nov 2, 2015 22:48:51.875910000	362,506
	Anomalies on selected road	Aug 22, 2015 17:22:21.080043000	Nov 2, 2015 22:43:45.277766000	6904 (1.904%)
	Mt. Ousley road section	Aug 22, 2015 17:06:49.080062000	Nov 2, 2015 22:46:35.075914000	121,358
	Anomalies on Mt. Ousley	Oct 5, 2015 02:02:01.981338000	Nov 2, 2015 22:43:45.277766000	4360 (3.59%)
Truck 4	All road sections registered by DSRC	Aug 23, 2015 07:04:38.270970000	Oct 23, 2015 03:31:06.089445000	2,853,832
	Selected road section	Aug 24, 2015 13:29:16.984854000	Oct 15, 2015 08:04:36.729368000	329,612
	Anomalies on selected road	Aug 24, 2015 13:29:23.385231000	Oct 15, 2015 07:52:24.529124999	3345(1.01%)
	Mt. Ousley road section	Aug 24, 2015 13:43:16.085094000	Oct 15, 2015 08:02:32.529147000	90,830
	Anomalies on Mt. Ousley	Oct 7, 2015 06:10:03.925373000	Oct 7, 2015 06:13:05.025705000	450 (0.49%)
Truck 5	All road sections registered by DSRC	Aug 22, 2015 14:11:35.674519000	Oct 26, 2015 15:42:51.925330000	1,670,058
	Selected road section	Aug 22, 2015 14:57:18.611686000	Oct 26, 2015 15:26:14.425329000	345,849
	Anomalies on selected road	Aug 23, 2015 15:35:39.695137000	Oct 26, 2015 15:13:58.525482000	2093 (0.6%)
	Mt. Ousley road section	Aug 22, 2015 14:59:24.211495000	Oct 26, 2015 15:24:08.829763000	103,870
	Anomalies on Mt. Ousley	Sep 27, 2015 01:15:24.979748000	Sep 27, 2015 01:16:36.743136000	720 (0.69%)

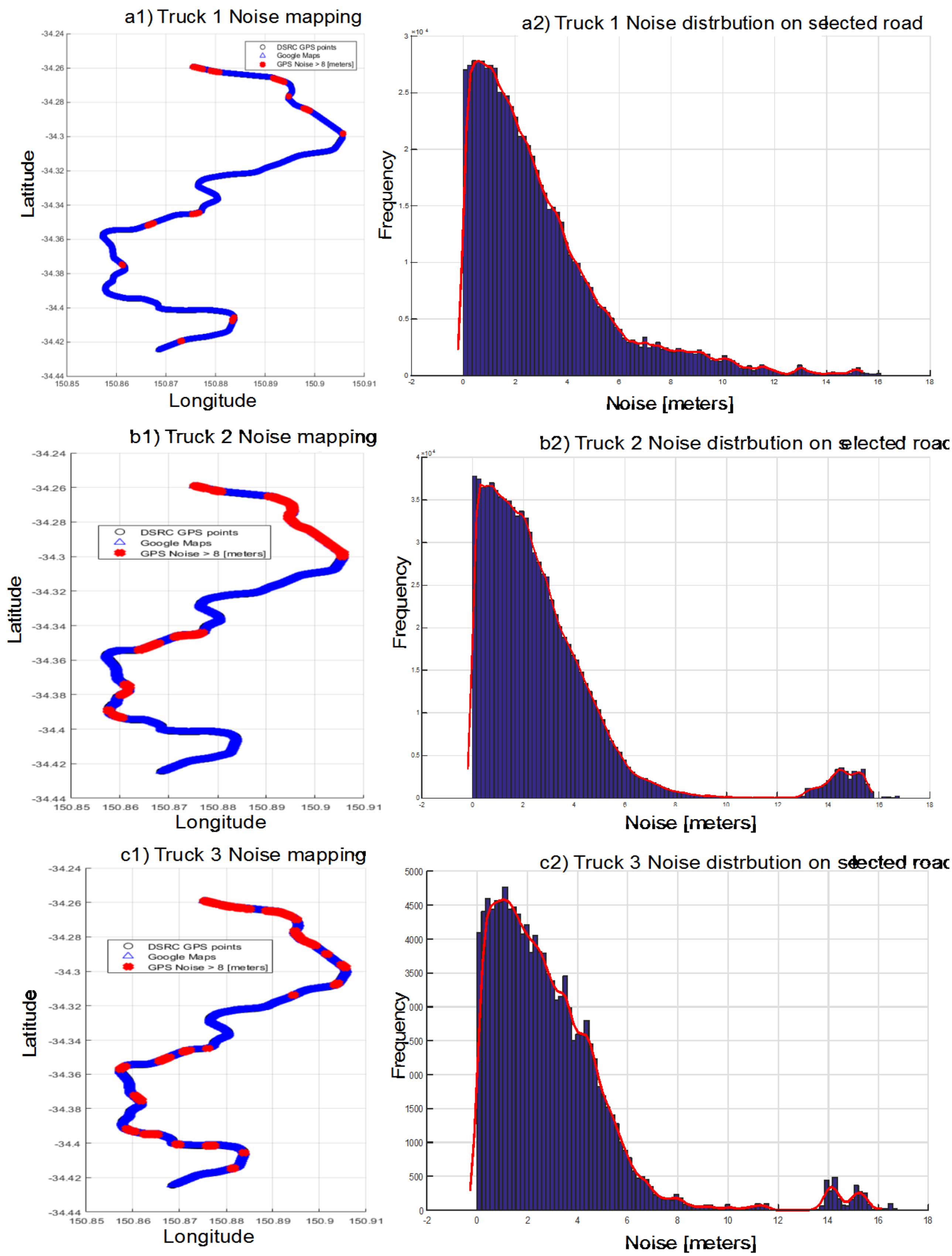


FIGURE 5 Noise mapping and distribution for Truck 1, 2 and 3.

Figure 5 shows the noise mapping and noise distribution on the selected road section for the first 3 trucks. We can therefore identify which road areas are susceptible to register deviations from the road center, which we will define as "noise-prone" areas. Figures 5 a1), b1) and c1) show that the north part of the road is more sensitive to noise, which is near the coal mine where the trucks stop for loading. As well, we can notice that, although Truck 2 registered the biggest number of anomalies, Truck 3 seems to present a large spread in the positioning where the noise is registered. The noise distribution plot (Figure 5 a2), b2), c2)) confirms again a particular behavior for Truck 3 and 2, as the maximal noise can reach 17.0785 meters in certain locations. In terms of average noise on Mt. Ousley, we make the observation that Truck 5 (not represented here) has the lowest noise levels (2.24 average noise from road center), which falls into good levels of positioning on the streets. In furthering the understanding of GPS and BSM accuracy, we suggest these noise-prone locations would be good places to investigate in detail in order to understand the phenomena of common localized issues.

By taking into consideration the global positioning of all the trucks we investigated in the CITI project, we can state that the average noise obtained for almost all trucks fall under 3 meters, which indicate that the GPS location being transmitted has good accuracy in most of the BSMs. Nevertheless, bad positioning accuracy can lead to false alert generation and hinder road safety, especially when trucks are fully loaded with 82 tons of coal. This aspect is not to be neglected and further studies need to be undertaken in order to understand the cause of bad positioning accuracy, possibly including: bad placement of antennas, road geometry, speed, heading etc.

4.3 Regression models for noise analysis

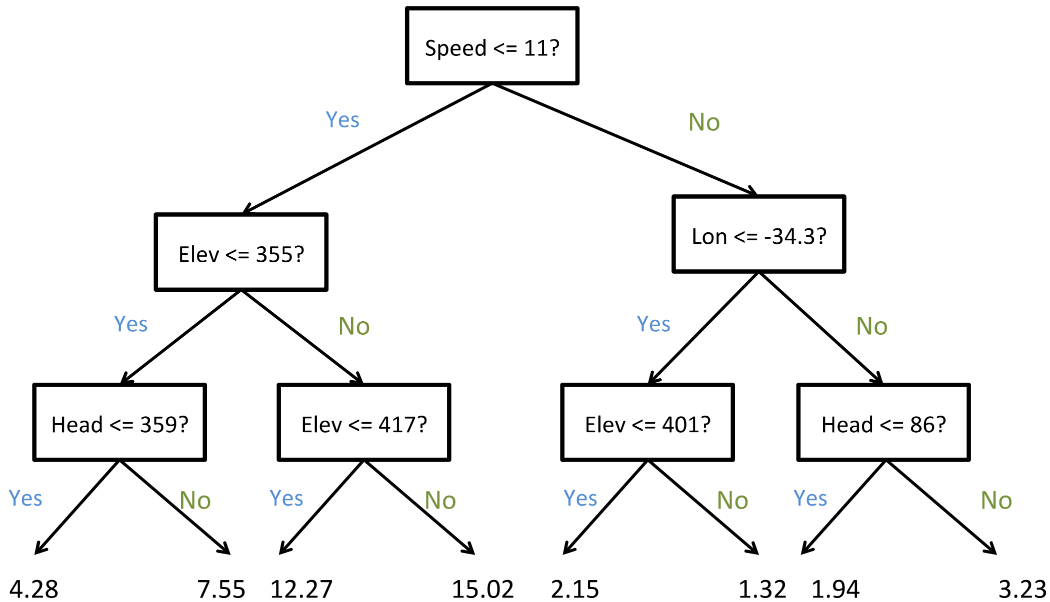
In this section we perform a closer investigation of explanatory factors that can influence DSRC GPS noise. Besides GPS observations with longitude and latitude from transmitted BSMs, we also record the following features (variables): Elevation, Speed, Heading, Brakes, Acceleration Longitude, and Acceleration Latitude, in matrix $\boldsymbol{X_t} = [\boldsymbol{X_{i,j}}]_{i=1,..N_d^{GPS}}^{j=1,..8}$. We also consider the corresponding noise vector $\boldsymbol{N_t} = [N_i]_{i=1,..N_d^{GPS}}$ for this time. We then consider the regression problem of predicting $\boldsymbol{N_t}$ from $\boldsymbol{X_t}$, so as to determine the highly predictive features which influence GPS noise.

To avoid the statistical issue of overfitting (*30*), we separate our data into a training and a testing set. The training set comprised the first 80% of all GPS readings, with the rest falling into the testing set. We then fit a regression model (to be described subsequently) on the training set, and evaluated model performance on the testing set. Performance is evaluated using the mean squared error (MSE). As a baseline, we used the trivial model which predicts the mean of the GPS noise in the training set; any model that performs worse than this is practically useless.

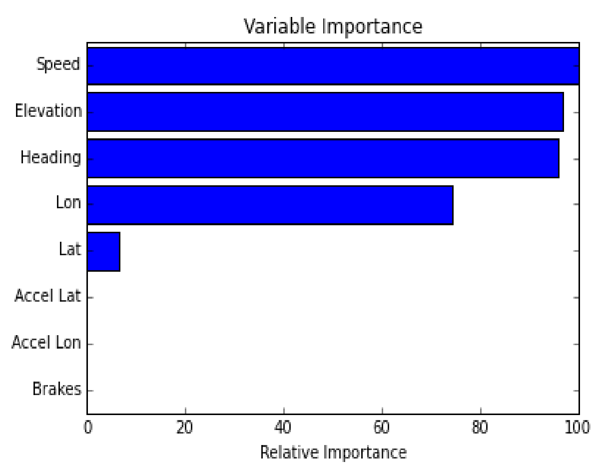
We use two underlying regression models. The first is a decision tree (specifically, one using the CART algorithm (*31*)). This model was chosen because it is intuitive to explain, and can easily fit nonlinear relationships in the data. At a high level, a decision tree involves making a number of splits of the data based on some thresholding of the feature values. Depending on the outcome of this thresholding, one then fits a sub-model, which is recursively another decision tree. Finally, one terminates at a leaf node, where a hard prediction is made for the target value. This is typically done by the average of the points that fall into that leaf node.

We fit a decision tree with a fixed depth of 3 levels. We found this model to give an MSE of **2.4261**, which is a nearly 60% improvement over the baseline MSE of **5.7864**. Further, the output of the tree is shown in Figure 6 a), and seems to be intuitive. We find that the most predictive features are the Speed, Elevation, and Heading. The model is seen to separately treat the cases of very low speed (< 11 km/h). For higher speed, the Longitude is seen to be predictive. While this may seem counterintuitive, in fact the longitude indicates high variations in the movement of the truck along the selected road section, as represented in Figure 5 a1).

To further assess feature importance, we fit a gradient boosted decision tree (GBDT) model (*32*). This is an example of an ensemble method (one that computes a number of individual sub-models, and then considers an appropriately weighted average of them). Such an averaging procedure lends these methods a robustness against overfitting to spurious signals in the data.



a)



b)

FIGURE 6 Outputs of decision tree methods: a) Decision tree with a fixed depth of 3 levels for X_t, b) Features influencing DSRC GPS positioning accuracy.

We fit a GBDT comprising 500 individual sub-models, to a maximum depth of 2 levels. We found this model to give an MSE of **2.2696**, which is a further 6% improvement over the single decision tree model. Compared to a decision tree, it is harder to directly visualize the output of a GBDT, since it comprises hundreds of sub-models. Nonetheless, we can still estimate the overall importance of individual features. We find that the most predictive features that can influence DSRC GPS accuracy are: speed, elevation and heading (Figure 6b)), which is largely consistent with the finding from the single decision tree. These results validate as well our previous finding of noise evolution on Mt. Ousley which has restricted low speed for trucks descending the mountain. The fact that these features are also useful for the GBDT gives confidence that there is a statistically meaningful relationship between these variables and the DSRC GPS noise.

5. CONCLUSIONS

In this paper we conducted a detailed investigation for analysing the GPS positoning accuracy as transmitted in BSMs by DSRC equipped heavy vehicles, operating in the CITI project. After choosing GSM as main ground truth for noise computation, we showcase the DSRC transmitted positioning accuracy of the 5 most active trucks of CITI fleet, and identify the noise-prone areas in which DSRC false generated alerts can be triggered. Lastly, by conducting a regression analysis method based on gradient boost decision trees we found that for the data set we used, speed, elevation and heading were most predictive of GPS positioning error.

This work is an intial step in the positioning accuracy and accident alerts investigation for improving road safety. CITI project is an ongoing project, with aims to investigate DSRC use at signalised intersections, as well as improving road safety especially in high concern public areas (schools, kindergardens, etc.). As the DSRC systems are applied more widely, there is a real need for testing and investigating the technology on more light vehicles, in order to improve road safety.

REFERENCES

[1] WHO. *Global status report on road safety 2013: supporting a decade of action*. World Health Organization, 2013.

[2] ---. *Global status report on road safety: time for action*. World Health Organization, 2009.

[3] Grace, N., C. Oxley, S. Sloan, A. Tallon, P. Thornton, T. Black, and A. V. Easton. Transforming Transportation through Connectivity: ITS Strategic Research Plan, 2010–2014.In, US DOT's Volpe Center, 2012.

[4] Kockelman, K. M., and T. Li. Valuing the Safety Benefits of Connected and Automated Vehicle Technologies.In *Transportation Research Board 95th Annual Meeting*, 2016.

[5] Zeng, X., K. N. Balke, and P. Songchitruksa. Potential Connected Vehicle Applications to Enhance Mobility, Safety, and Environmental Security.In, Southwest Region University Transportation Center, US, 2012.

[6] 02-302, F. Amendment of the Commission's Rules Regarding Dedicated Short-Range Communication Services in the 5.850–5.925 GHz band.In, Federal Communications Commission, 2002.

[7] Li, Y. An Overview of the DSRC/WAVE Technology.In *7th International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness, QShine 2010*, Springer Berlin Heidelberg, Houston, USA, 2012. pp. 544-558.

[8] Transportation, U. S. D. o. Vehicle Safety Communications – Applications (VSC-A) Final Report.In, Research and Innovative Technology Administration, Washington D.C., 2011.

[9] Jiang, D., V. Taliwal, A. Meier, W. Holfelder, and R. Herrtwich. Design of 5.9 Ghz Dsrc-based Vehicular Safety Communication. *Wireless Communications,* Vol. 13, No. 5, 2006, pp. 36--43.

[10] Pan, J.-S., S. Ma, S.-H. Chen, and C.-S. Yang. Vision-based Vehicle Forward Collision Warning System Using Optical Flow Algorithm. *Journal of Information Hiding and Multimedia Signal Processing,* Vol. 6, No. 5, 2015, pp. 1029 - 20142.

[11] Tang, A., and A. Yip. Collision avoidance timing analysis of DSRC-based vehicles. *Accident Analysis and Prevention,* Vol. 42, 2010, pp. 182-195.

[12] Gende, M., M. Chowdhury, K. Dey, and W. Sarasua. Connected-Vehicle Technology for Allowing Priority Requests at Signalized Intersections: Analysis.In *Transportation Research Board, 95th Annual Meeting*, Washington DC, 2016.

[13] Harding, J., G. Powell, Y. R., R., J. Fikentscher, C. Doyle, D. Sade, M. Lukuc, J. Simons, and J. Wang. Vehicle-to-vehicle communications: Readiness of V2V technology for application.In, National Highway Traffic Safety Administration, Washington, DC, 2014.

[14] Efatmaneshnik, M., A. Kealy, A. T. Balaei, and A. G. Dempster. Information Fusion for Localization Within Vehicular Networks. *The journal of navigation,* Vol. 64, 2011, pp. 401-416.

[15] Alam, N., and A. G. Dempster. Cooperative Positioning for Vehicular Networks: Facts and Future. *IEEE Transactions on Intelligent Transportation Systems,* Vol. 14, No. 4, 2013, pp. 1708-1717.

[16] Alam, N., A. T. Balaei, and A. G. Dempster. An instantaneous Lane-Level Positoning using DSRC Carrier Frequency Offset. *IEEE T. on Intelligent. Tr. Syst.,* Vol. 13, No. 3, 2012.

[17] Jiangchen, L., G. Jie, Z. Hui, and Q. Z. Tony. A RSE-Assisted GPS-RSS Hybrid Lane-Level Positioning System for Connected Vehicles.In *Transportation Research Board 95th Annual Meeting*, Washington D.C., 2016.

[18] Van Hamme, D., P. Veelaert, and W. Philips. Lane Identification based on Robust Visual Odometry.In *16th International IEEE Annual Conference in ITS,* , Hague, 2013.

[19] Hofmann-Wellenhof, B., H. Lichtenegger, and J. Collins. *Global Positioning System: Theory and Practice*. Springer-Verlag Wien, 2001.

[20] Williams, T., P. Alves, G. Lachapelle, and C. Basnayake. Evaluation of GPS-based methods of relative positioning for automotive safety applications. *Transportation Research Part C: emerging technologies,* Vol. 23, 2012, pp. 98-108.

[21] Tan, Y. K. Positioning Techniques with two GNSS satellites over time.In *ION GNSS*, 2010. pp. 208-216.

[22] Tyler, P., J. Wall, and V. Vecovski. CITI- An update on Australia's First Pilot Deployment of CITS.In *23rd ITS World Congress*, Melbourne, 2016.

[23] RMS. *Picton Road crash history*. http://www.rms.nsw.gov.au/projects/illawarra/picton-road-safety-improvements/picton-road-crash-history.html. Accessed 31/07/2016.

[24] NRMA. *Cost of Road Crashes.* NRMA Monitoring and Services. https://www.mynrma.com.au/media/Cost_of_Road_Crashes.pdf. Accessed 29/09/16, 2016.

[25] CohdaWireless. *MK5-OBU*. http://cohdawireless.com/Portals/0/MK5_OBU_10122015.pdf.

[26] Bento, L. C., R. Parafita, and U. Nunes. Inter-vehicle sensor fusion for accurate vehicle localization supported by v2v and v2i communications.In *15th International IEEE Conference on ITS*, IEEE, 2012. pp. 907-914.

[27] Greenfield, J. S. Matching GPS observations to locations on a digital map. Presented at Transportation Research Board, 81st Annual meeting, 2002.

[28] Yang, D., B. Cai, and Y. Yuan. An improved map-matching algorithm used in vehicle navigation system.In *Intelligent Transportation Systems, 2003. Proceedings. 2003 IEEE, No. 2*, IEEE, 2003. pp. 1246-1250.

[29] Vicenty, T. Direct and inverse solutions on the ellipsoid with application of nested equations. *Survey Review XXII,* Vol. 176, 1975, pp. 88--93.

[30] Hastie, T., R. Tibshirani, and J. Friedman. Unsupervised learning.In *The elements of statistical learning*, Springer, 2009. pp. 485-585.

[31] L. Breiman, J. F., R. Olshen, and C. Stone. . *Classification and Regression Trees*. Wadsworth, Belmont, CA, 1984.

[32] Friedman, J. H. Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 2001, pp. 1189-1232.