PROJECT REPORT ACA109

The application of crowdsourced data in road condition assessment

R Lee, M Courtier and A Wright

Report details

Report prepared for:	TRL Academy
Project/customer reference:	SIP23003
Copyright:	© TRL Limited
Report date:	November 2024
Report status/version:	Final
Quality approval:	

James Peeling (Project Manager)	**James Peeling**	Alex Wright (Technical Reviewer)	**Alex Wright**

Disclaimer

Table of Contents

Executive Summary

On the Strategic Road Network (SRN) monitoring of pavement asset condition is currently achieved using dedicated surveys that measure the functional (e.g. user experience such as roughness, and visual appearance), safety (e.g. skid resistance) and structural (e.g. bearing capacity) condition. This data is provided by annual network level surveys to support long term planning, and more frequent data collection to detect defects that are rapidly developing (e.g. potholes).

In recent years a new source of data has become available that is referred to as "crowdsourced data". This can be provided by vehicles or mobile phones, on a network wide basis, from multiple sources. Crowdsourcing has the potential to change the way in which data is used to manage the condition of the SRN. The study presented in this report investigates whether and how crowdsourced data has the potential to be applied in the management of the condition of SRN.

The study obtains example data from two crowdsource data types – data provided by sensors installed on standard vehicles (vehicle telemetry data) and data from mobile phones (mobile phone data) – over two geographical regions and over two specific time periods. A focussed, case study-based investigation is carried out to determine if, and to what extent, the crowdsourced data types could be applied to understand road condition and how it complements the assessment of the functional and safety level of service measured by current data collection methods.

The vehicle telemetry data provided to the study consists of point "events", each accompanied by its geolocation and timestamp. The study focusses on the examination of telemetry "slip" and "bump/pothole" events. The mobile phone data is also limited to the reporting of events, in this case "deceleration" events, accompanied by their geolocation and timestamp. No detail was provided on the methodology deployed by the providers to determine/identify these events.

It is noted that the crowdsourced data available to the study covered only a short period of time, and a limited number of data types ("events"). The dataset is not representative of the coverage and content of crowdsourced data that could be achieved if routinely applying the data for road condition assessment. The conclusions of the study must be considered in the light of the data available to it. However, the study provides an insight into the future potential of this data.

The study finds that the crowdsourced data is relevant to road condition management. However, to enable the data to be applied alongside current data sources the data must be aligned, summarised and clustered. Significant effort is required to achieve this in the study. It is concluded that there would be benefit in developing tools within Road Asset Management Systems to facilitate smoother access, fitting, alignment and analysis of crowdsourced data. Refinements in the way crowdsourced data is delivered by providers might help to make this more straightforward.

When considering functional condition (roughness/bumpiness) specific case studies show potential for the identification and tracking of potholes. However, as the crowdsourced data provided is limited to a range of thresholded values (or because of the short collection time-

period) there is limited agreement between the crowdsourced data and conventional roughness measurements. The provision of greater detail in the bump event data (e.g. the vertical acceleration or an appropriate derived parameter) might provide a more useful measure to assist in understanding condition. Data relating to the vehicle type/suspension characteristics may also useful.

The study directly compares telemetry slip events with conventional measurements of skid resistance (SCRIM). This direct comparison does not show a strong correlation. This is not surprising as the time period covered by the telemetry data is short, and it may be expected that measurements would have to “build up” on individual sites to improve the relationship with the SCRIM. A coverage assessment suggests that it would take a significant length of time to achieve network-wide measurements and that the crowdsourced data would not, as it stands, be ready to replace fundamental network measurements such as SCRIM.

Because the mobile phone deceleration data can be related only to the demand for skid resistance requested by the vehicle, and not the supply of skid resistance, the mobile phone data is not compared directly with SCRIM measurements. Instead, as SCRIM site category has a closer relationship with the expected demand for skid resistance on a particular site, the study compares the deceleration event data with site category. There is a reasonable relationship, confirming that the site category established for the sites used as case studies in the work were reflected by the level of braking demand recorded on the sites in the mobile phone data. The observations suggest that deceleration data could provide further insight to support site categorisation.

Although there is little direct correlation with conventional measurements of skid resistance, there is strong evidence to support the use of the crowdsourced data in the assessment of risk. This includes deciding how to define site categories, and to support decisions on whether interventions should be made. Where sites have been found to have low skid resistance, the crowdsourced data could be used to determine whether a site has a disproportionate number of vehicles “demanding” this skid resistance, and it could also be used to bring sites to the road operator’s attention which may not have been flagged by the SCRIM data. As for the functionality data, the provision of greater levels of detail in the data may help provide a more useful measure to assist in understanding the condition.

It is noted that, although there is strong potential for the application of crowdsourced data to manage safety condition, the data may not provide a comprehensive understanding of risk. There may be sites on which collisions are likely (or have occurred), for which low numbers of events have been recorded. The risks presented by these sites will be misrepresented in the crowdsourced dataset. Future studies into the application of crowdsourced data should also consider these types of site.

Finally, given the potential for this data, it is recommended that data providers and end users (e.g. asset management system providers or road administrations) collaborate to refine the approach taken to reporting the data, to ensure that the content aligns with the requirements of the user.

Acknowledgements

The research presented in this report would not have been possible without the pavement condition data sourced from National Highways' databases. We would like to express our gratitude for the permission to use the data in this work.

We would also like to express our gratitude to the providers of the telemetry and mobile phone datasets that were used in this work.

Map data is from OpenStreetMap (openstreetmap.org/copyright).

1 Introduction

To manage the Strategic Road Network (SRN) asset there is a need to understand its condition. This is currently achieved using dedicated surveys that measure the functional (e.g. user experience such as roughness, and visual appearance), safety (e.g. skid resistance) and structural (e.g. bearing capacity) condition. For medium and long-term planning, this data is provided by annual network level surveys. For short/immediate term maintenance planning further data collection is undertaken on a more frequent basis to detect defects that are rapidly developing, typically to resolve failures in the functional or safety level of service (e.g. potholes).

In recent years a new source of data has become available that is referred to as "crowdsourced data". This can be provided by vehicles or mobile phones, on a network wide basis, from multiple sources. Crowdsourcing has the potential to change the way in which data is used to manage the condition of the SRN. However, although these sources of data have potential in this application, they were not necessarily developed for the purpose of measuring road condition.

The study presented in this report investigates whether and how crowdsourced data has the potential to be applied in the management of the condition of SRN. The approach taken in the study is to:

- Obtain example data from two crowdsource data types – data provided by sensors installed on standard vehicles and data from mobile phones – over two geographical regions of the network, collected over two specific time periods.
- Obtain the data provided by conventional condition data sources over the same regions.
- Undertake a focussed, case study-based investigation to determine if, and to what extent, the crowdsourced data types could be applied to understand the condition of the network with respect to the assessment of the functional and safety level of service. The approach has been to formulate a question regarding the application of the crowdsourced data in a particular application, and then to explore the data to determine whether there is strong, weak or inconclusive evidence to support the use of the data in this application.

A summary of the overall observations and conclusions that can be drawn from the study is presented, with recommendations for steps that could be taken to begin the introduction of this data into road condition management on the SRN.

It should be noted that the crowdsourced data available to this study covers only a short period of time, and a limited number of data types. The dataset is not representative of the coverage and content of crowdsourced data that could be achieved if routinely applying this data for road condition assessment. The conclusions of the study must be considered in the light of the data available to it. However, this study does provide an insight into the future potential of this data.

2 Background – current approach to the collection and use of road condition data on strategic roads

Currently, the management of the condition of pavements on the SRN draws on a number of data sources, primarily obtained using dedicated network surveys. Data is collected, fitted to the network and loaded into network databases. Analysis is then carried out in which the data is assessed, both to report condition at the network level and to identify locations in need of further investigation or maintenance. The following provides a summary. Note that this is not a comprehensive summary; it aims to provide a context to the application considered for crowdsourced data in later sections of this report.

2.1 Measurement of functional condition (shape and appearance)

The functional condition of road pavements is primarily associated with the provision of a road surface that enables vehicles to travel with an acceptable level of comfort, at speeds appropriate to the pavement design. The assessment of functional condition is achieved through measurement of the surface profile of the pavement. The transverse profile is typically used to determine the depth of rutting, which can affect vehicle handling and cause aquaplaning. The longitudinal profile can be used to quantify the ride quality and identify bumps and potholes.

On the SRN functional condition is measured by the TRACS survey (Figure 2-1). The TRACS survey vehicle measures 3D road shape (profile) and collects images of the road surface at traffic-speed. Algorithms convert the profile measurements into condition parameters that include pavement rutting and longitudinal profile (reported as enhanced Longitudinal Profile Variance (eLPV) in three different wavelengths - 3m, 10m, 30m). Whilst the eLPV data provides an indication of the general roughness of the pavement, a further parameter is calculated, called the Bump Index, which reports the location of significant bumps which may occur at localised defects such as potholes, failed joints and failed patches. The Image data is also deployed to identify visual deterioration (cracking). TRACS surveys are carried out annually in all lanes.

Figure 2-1: TRACS survey vehicle

In addition to TRACS annual surveys of visual condition and functional condition, further visual assessments of the condition of the SRN are carried out on a more frequent basis. These include daily safety inspections, in which engineers identify defects which have developed in the pavement, primarily to support the rapid response/resolution of safety hazards, rather than for the purpose of data collection for longer term asset management.

2.2 Measurement of safety (skid resistance)

The measurement of skid resistance is used as a primary source data to quantify safety information, because insufficient skid resistance (friction) may reduce the ability of vehicles to safely navigate bends or stop within an acceptable distance. On the SRN skid resistance is measured by surveys carried out using the Sideway-force Coefficient Routine Investigation Machine (SCRIM). The SCRIM uses a test wheel toed-in towards the centre of the machine to create a sideways force as the machine travels (Figure 2-2). The apparatus measures the ratio of the sideways force to the reaction force on the test wheel, which is reported as the SCRIM coefficient. SCRIM surveys are carried out annually on lane 1 of main carriageways, slip roads and roundabouts. Due to seasonal variations in skid resistance, pavements are measured at a different time of year and the SCRIM coefficient averaged over a three-year period to provide the Corrected SCRIM Co-efficient (CSC).

Figure 2-2: Test wheel within the SCRIM survey vehicle

2.3 Measurement of structural condition (deflection)

Pavement structural condition assesses its ability to support vehicle loads. On the SRN the overall structural condition is determined by measuring the deflection of the pavement under load, using both continuous measurements carried out at traffic speed, and devices that work statically or at low speed within road closures.

Network level assessment of structural condition is undertaken on the SRN under the TRASS (Traffic-speed Structural Survey) survey using the SATTS device (which is also referred to as the TSD, Traffic-speed Deflectometer). The device measures pavement deflection velocity using laser sensors mounted under an HGV. The data is processed after the survey to convert the measurements into a set of Network Structural Condition (NSC) categories for each 100m length of the network. Further data is collected at the scheme, or project, level using the Deflectograph (slow speed device) and Falling Weight Deflectometer (static device) to provide further insights into the structural condition and to design treatments.

For some types of pavement construction, the visual condition data (cracking) provided by the TRACS survey can be in indicator of deterioration in the structural condition.

2.4 Further sources of data

Further sources of information are also used to support condition assessment, for example:

- Traffic: Traffic data is collected on the SRN by the MIDAS sensor network and manually by the Department for Transport using traffic counts.
- Collision data: Records of collisions that have occurred on the network provide an important source of data when assessing the risks of particular sites. The STATS19 Road Collision data is published bi-annually, providing a database of all road traffic casualty collisions reported to UK police. Data is available for the past five years. This includes the location and time of the incident, number of vehicles and casualties, road class, function and environment and contributing factors such as weather (Figure 2-3).

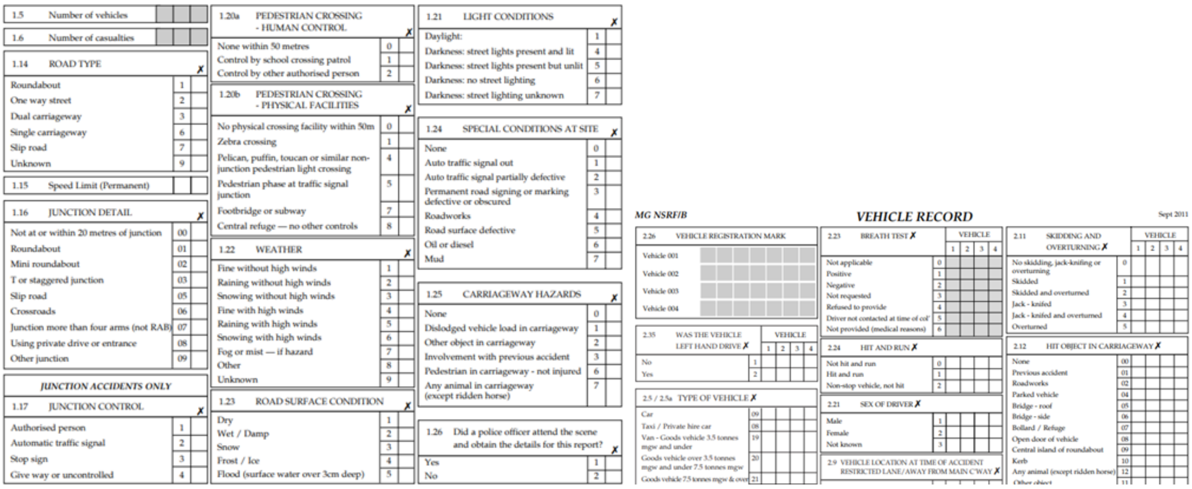


Figure 2-3: STATS19 record sheet

- Construction: Construction records include the pavement material, the date the material was laid and the thickness for each layer of the pavement.
- Flood events: the Highways Agency Drainage Data Management System (HADDMS) contains records on all reported flooding events occurring on the SRN. The frequency of these events may accelerate pavement degradation.

- Treatments: records of previous minor treatments such as patching and major treatments such as resurfacing can support decisions on future maintenance needs.

2.5 Use of current data in asset condition assessment

There are many applications of condition data in the management of asset condition, of which the following are of relevance to this work:

Reporting network level condition. KPI3.1 quantifies the percentage of the pavement asset considered to be in “good condition”. This network-level pavement condition indicator is calculated using data from TRACS and SCRIM surveys. Standards CS 230 (National Highways, 2022) and CS 228 (Highways England, 2021) define four condition categories for the SRN based on the ride quality, rutting and skid resistance present on each 100m length of the network. The indicator reports the percentage of lengths falling into categories 3 or 4.

Determining interventions to restore skid resistance (safety). Standard CS 228, Skid Resistance describes the process for collecting, analysing and acting on SCRIM skid resistance survey data. The skid resistance requirements for each length of the network are established by setting a site category (related to the likelihood of frequent or hard braking on this length) and investigatory level (the extent of skid resistance required on this length). Under guidance from CS 228, the site category is established by assessing the pavement usage, geometry and local knowledge available to the assessor. Where a SCRIM survey returns a measurement value below this investigatory level for any length an investigation will begin which will determine whether any action is required (e.g. to improve its skid resistance). Note: a summary of site category definitions is provided in Table 5 in section 7.

Determining interventions to restore structural and functional condition. CS 230 Pavement Assessment Procedure describes three types of treatment that may be applied on any length to restore pavement condition where the survey identifies deterioration in the structural or functional condition. A Technically Simple Scheme may be carried out with limited approvals where there are no structural defects present, and all defects are confined to the surface layer (i.e. functional deterioration). For this treatment only the surface layer needs to be renewed. A scheme is considered Technically Complex where structural damage is present within the pavement, requiring reconstruction and greater design oversight. Structural defects are inferred from TRASS structural condition data, and rut depth measurements from TRACS surveys.

In addition to the above planned/programmed maintenance interventions, the data from routine visual condition surveys is used to support decisions on maintenance need, in particular the reactive maintenance required to repair defects which develop rapidly, such as potholes.

3 Crowdsourced data

3.1 What is crowdsourced data?

Data is crowdsourced if it is collected from a wide range of participants interacting with or assessing a system. There is growing interest in crowdsourcing data for measuring road condition because of the potential benefits to the Road Authority, for example:

- Crowdsourced data may have lower cost than data collected using dedicated survey activities (although this will depend on the frequency and coverage of the crowdsourced data).
- For some applications (e.g. assessing risk) the data might contain information that cannot be provided using traditional methods (as it is collected by users interacting with the pavement rather than using models derived from survey measurements).
- Crowdsourced data may provide more frequent insights into condition and provide a more robust method to track changes in condition than traditional annual surveys.

Crowdsourced pavement data can be collected through a variety of means, which include:

- User reporting
- Connected devices, such as dash cams
- Smart phones
- Vehicle telemetry, form connected vehicles or fleet monitoring systems

However, crowdsourced data has some inherent disadvantages compared to traditional sources of road condition data. Because crowdsourced data is collected from a large number of independent sources, it is not feasible to calibrate and validate each individual member of the crowd – creating uncertainty around the accuracy and repeatability of metrics derived from crowdsourced data. Although it is believed that, with the collection of a sufficient quantity of data, the inaccuracies present within individual members of the crowd will average out, this cannot be assumed, as there may not be sufficient data or the inaccuracies may be biased in some way amongst the whole crowd. There are also concerns around the coverage of data required to create condition metrics on all roads. For example, the density of crowdsourced data on low trafficked roads will be lower.

3.2 Types of crowdsourced data within this study

This study considers two sources of data – vehicle telemetry data and mobile phone data.

- Vehicle telemetry data can be collected from sensors installed within the vehicle by the manufacturer as part of the standard functionality of the vehicle, or sensors that have been retro fitted, for example for fleet monitoring purposes. This data may be routinely collected by manufacturers of connected vehicles[1], or collected by fleet managers as part of their operational activities.

[1] I.e. each vehicle is connected to the collector of the data. This is typically provided through a SIM (3G, 4G, 5G) installed in the vehicle by the manufacturer.

- Mobile phone data is collected by installing an app onto a smartphone which collects data from the smartphone's sensors and uploads it to the data collector.

These sources are described further in Table 1.

Table 1: Summary of crowdsourced data

	Vehicle Telemetry Data	Mobile Phone Data
Sensors	Sensors may include: GPS; accelerometers; driver action (steering wheel position, braking force etc.,); traction force applied at the wheel; wheel speeds and differences (e.g. for ABS activation); suspension position (compression); weather (e.g. windscreen wiper activity); and cameras (e.g. for parking or lane assistance).	Mobile phone sensor data is typically limited to the GPS and accelerometer sensors within the phone. However, may include audio and camera input.
Potential for road condition that may be inferred from the data	The large number of sensors available from vehicle telemetry offers the potential for a wide variety of condition characteristics to be inferred from the data. E.g.: Suspension telemetry provides insights into longitudinal profile (roughness) and potholes; ABS sensors indicate where braking takes place, and where the demand for skid resistance exceeded the supply; recording driver actions provides insight into behaviour; vehicle imaging and sensing systems can assess the legibility of signs or quality of road markings.	Accelerometer data can indicate potentially rough surfaces. High levels of deceleration of the vehicle identifies locations where braking demands occur, but not whether the skid resistance matched driver demand. A smartphone within a vehicle cannot separate driver intention, from vehicle behaviour. As such, smartphone data is limited to analyses of the combined vehicle:driver:pavement system.
Issues / Limitations	There are several issues specific to vehicle telemetry that create variances in the data collected: vehicle characteristics can vary considerably across different types and manufacturers of vehicle; sensor types and locations will differ; the condition of the vehicles will vary; the vehicles may have been modified in some way.	If the phone is not secured within the vehicle, there may be false reports of acceleration as the device moves independently of the vehicle. Without any defined orientation of the smartphone relative to the vehicle, the accelerometer inputs may need to be calibrated each journey, and potentially mid journey if the smartphone is moved.
	Both sources of data have shared limitations listed below: • Low trafficked roads will have limited crowd coverage • Fleets have very limited coverage and potential biases • Driving lines / speed will vary • The sensors are not validated / calibrated • The condition / quality of sensors is not controlled or known	

3.3 The crowdsourced data used in this study

For this study TRL engaged with two providers of crowdsourced data, who provided anonymised data covering two regions of the country, as described in the following sections.

3.3.1 Vehicle telemetry data (Region 1)

Vehicle telemetry data was provided by a major vehicle manufacturer for a region of the country covering a circle extending approximately 200km from Manchester (Figure 3-1). The manufacturer provided access to an API that enabled data to be downloaded from their cloud database. Access was provided to data transmitted from all connected vehicles travelling in this region between 9th and 29th February 2024. The data covered all road types, both strategic and local. However, this study has considered only the data that was collected on strategic roads.

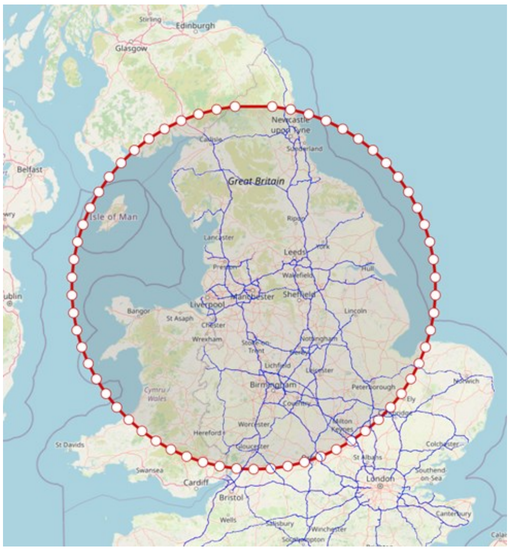


Figure 3-1: Region 1, covered by vehicle telemetry data

The data provided by the vehicle manufacturer was not “raw sensor” data. The vehicle telemetry data consisted of “events” that, it is assumed, were derived from the raw data by the manufacturer. No detail was provided on the methodology deployed to determine/identify an “event”. The event data contained in the vehicle telemetry dataset included:

- ‘Slip’ events
- Bumps & Potholes, including severity & wheel path
- Weather events
- Emergency braking, accidents, breakdowns and hazard lights

Each event was accompanied by its geolocation and timestamp, with a total of 296,000 events provided in the in the sample dataset (see further discussion of the content in Section 4.2).

It can be seen from the above that vehicle telemetry data covers two aspects of road condition. The bumps and potholes events are potentially associated with road functional condition, and the slip events are likely to be associated with safety.

3.3.2 Mobile phone data (Region 2)

Mobile phone data was provided by a major technology/data provider for a region of the country covering a circle extending approximately 60km, centred on the southwest of Greater London (Figure 3-2). For this study the manufacturer provided an example dataset in CSV format, but typically the data would be accessed via a web API. The data provided was that collected from all connected/participating mobile phones travelling in this region between 1st and 7th October 2023. Again, the data covered all road types, both strategic and local, with only the data collected on strategic roads considered in this study.

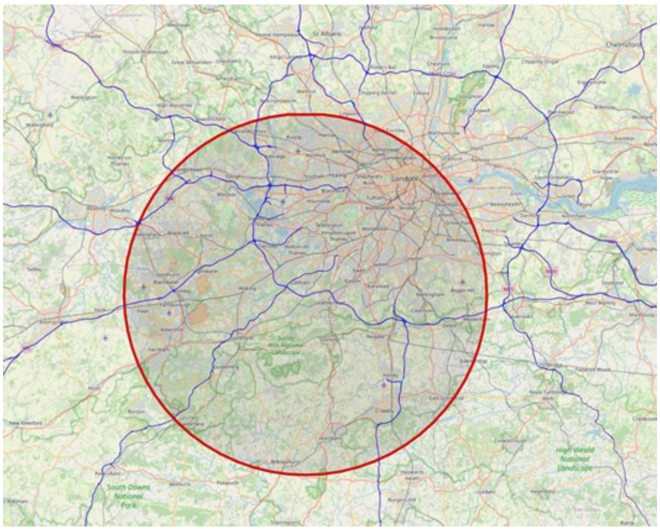


Figure 3-2: Region 2, covered by mobile phone data

The data provided was limited to the reporting of "deceleration events", which have potential as an indicator of road safety. Although detail was not provided, these events appear to be associated with locations where the mobile phone has recorded high levels of deceleration, which may have arisen from heavy braking. The dataset included, for each event:

- Geolocation (start and end locations) and timestamp (including the duration of the event)
- Initial speed
- Average deceleration, for any deceleration event greater than 7mph/s (3.2m/s^2)
- An estimate of the traffic volume in the location

There were ~2,100 deceleration events within the sample dataset provided (see further discussion of the content in Section 4.2). It was deduced that this data would be associated with safety condition.

4 Approach to data collation, alignment and analysis

4.1 Current data types

4.1.1 Functional and safety condition data

Given that the crowdsourced data was expected to provide insight into safety and function condition, current data on pavement functional and safety condition were obtained from National Highways Pavement Asset Management System (P-AMS), including:

- Road surface condition from TRACS data, which for the purposes of this study focussed on the TRACS roughness (eLPV and Bump Index) parameters for the two regions.
- Road safety condition from SCRIM, including Skid resistance from SCRIM data. In addition, the Skid risk site category and SCRIM investigatory levels were obtained for the two regions.

The most recent survey data available for each lane of the road was obtained, for lengths where the P-AMS construction records did not indicate the pavement had been treated since the survey took place. This ensured that the data was representative of the current condition of the pavement.

Pavement data stored on P-AMS is locationally referenced to a network defined in terms of section and chainage. A section is a defined length of road with a start and end point (with defined coordinates), and the chainage states the distance from the start of the section. Data was downloaded as 100m and 10m average values for each lane, with section and chainage.

As discussed below, it was not possible to fit the crowdsourced data to a specific traffic lane. Therefore, to enable comparison between the crowdsourced and road condition data it was necessary to combine the road condition data across all lanes. The data was therefore aggregated to report the average, maximum and minimum value across all lanes. The data was also aggregated to report the maximum, minimum and average value per section. The per section data also included the percentage of each skid risk site category present within the section.

4.1.2 Traffic data

Traffic volumes are reported on the SRN on the WebTRIS web portal. Traffic data used in this study was collated from 2022, because a suitably formatted and aggregated dataset had been prepared from 2022 for use in a previous project. The data included road name and direction, and Average Daily Traffic over 24 hours, showing the total number of vehicles passing each measurement point in the carriageway over a full day. The dataset also included the number of hours each counting point was in service during the year, which was used to clean the dataset. Counting points that had a low number of operational hours were removed (this can occur where the traffic counting loops are broken).

As traffic counting sites are referenced only to their coordinates, it was necessary to allocate these to P-AMS section and chainage. The coordinates of the mid-point of the section were approximated by calculating the mid-point between the section start and end co-ordinates provided in P-AMS. The coordinates of each traffic counting point were matched to the closest section mid-point which shared the same road name and direction of travel. This is an imperfect method of matching the traffic data to sections – as the distance between counting point and section increases, so does the likelihood that a junction is present between the two, making the traffic count inaccurate. However, alternative methods would either involve complex analyses of the SRN network definition or manual interpretation, which were beyond the scope of this study.

4.1.3 Collision data

Collisions from STATS19 are reported in a similar way to crowdsourced event data. They are reported with the coordinates of a single point in space. This data was fitted to the SRN (P-AMS network) so it could be compared with other NH data sources and the crowdsourced data, all of which was also fitted to the SRN (see below). Fitting of the STATS19 data followed broadly the same fitting procedure as used for fitting the crowdsourced data to the SRN, which is described below in Section 4.2.1 (noting that STATS19 events had a single location coordinate, whereas most of the crowdsourced data was provided with a series of coordinates).

4.2 Crowdsourced data

4.2.1 Functional and safety data

Crowdsourced data was obtained from telemetry and mobile phone data providers as discussed in Section 3 above. The vehicle telemetry data was collected during a 20-day period in February 2024 and the deceleration data from the mobile phone source was collected during 7 days in October 2023. Table 2 provides a summary of the event data, by type. It can be seen that the vehicle telemetry data contained a higher number of events than the mobile phone data but, following fitting (see Section 4.2.2), the number occurring on the SRN was lower. Hence the mobile phone deceleration data formed the largest dataset of all the event types, both in terms of total count of events and events per km per hour.

Crowdsourcing relies upon an event being experienced by the vehicle or mobile phone. Therefore, we may expect the rate of reporting to vary with traffic flow. Figure 4-1 shows how the rate of event recorded varied over the data collection period. There are clear peaks that occur during the daytime for all event types, the daytime being the busiest travel period. In the peaks which occur on weekdays, a smaller morning and evening peak can also be regularly seen. This matches the expected traffic flow. However, although there is variation between the peaks for each day in the plots in Figure 4-1, the overall trend in rate of events is constant. The only exception to this is there is a large peak in the number of slip events at the start of the data collection period. The reason for this is unknown.

Table 2: Summary of the coverage of the crowdsourced data types

Data source	Vehicle telemetry	Vehicle telemetry	Vehicle telemetry	Mobile phone
Cause	Bump	Pothole	Slip	Deceleration
Month of data collection	February 2024	February 2024	February 2024	October 2023
Timespan of data collection (days)[2]	20.1	20.1	20.1	6.9
Total number of events	11621	3818	27904	2158
Number of events matched to the SRN	443	276	672	1831
Length of SRN covered (lane km)	8785.6	8785.6	8785.6	424.0
Events per km of SRN per hour[3]	0.00251	0.00156	0.00381	0.02599

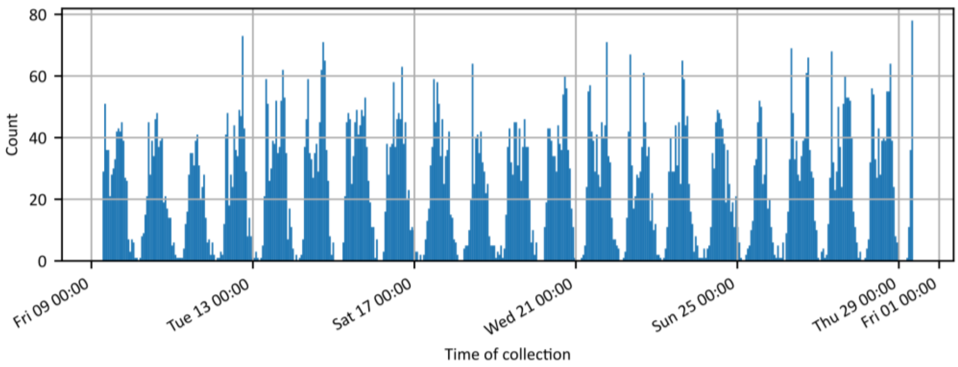
4.2.2 Fitting to the P-AMS network

As noted above, the crowdsourced data was provided as "events". Typically, these included coordinates and information about the event itself. For the mobile phone deceleration events, three location coordinates were provided for each event (at the start of deceleration, the middle and the end of deceleration). For the vehicle telemetry data, all events had one set of coordinates. However, most of the vehicle telemetry data was also provided with a set of coordinates simply showing the vehicle path leading up to the event. The purpose of these is to help locate and contextualise the event.

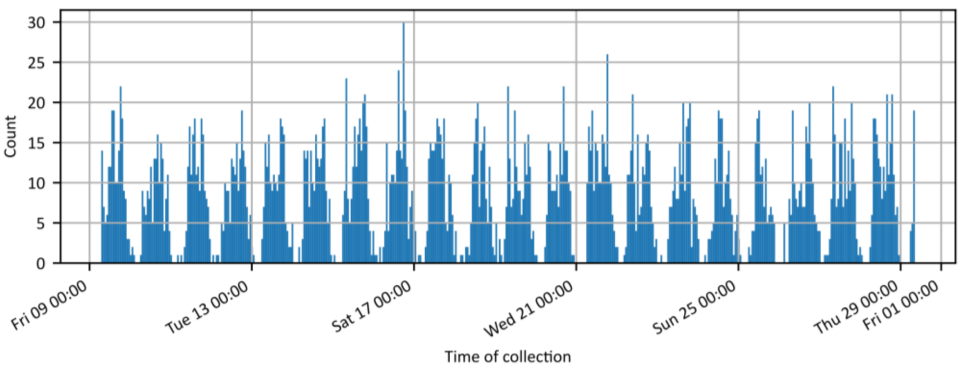
For this study it was necessary to match the events to their location on the network (as defined by P-AMS section and chainage). A fitting procedure was developed in which the location coordinates were compared to a line representation of the carriageway, in order to identify the section, chainage (the distance along that part/section of road) and direction of travel of the event. For most events on straightforward sections the fitting procedure was not complex. However, a more complex approach was required at roundabouts and junctions, particularly where roads were crossing over or under each other. Events that did not occur on the SRN were not fitted.

[2] The timespan of the data collection is not in whole days due to their being a delay between access to the relevant APIs being granted and the data collection commencing.

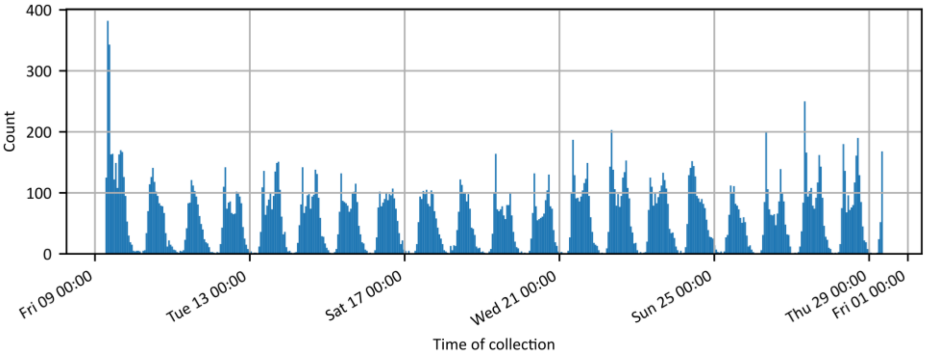
[3] Events per km of SRN per hour = Number of events matched to the SRN / (Length of SRN covered (lane km) * Timespan of data collection (days) * 24)



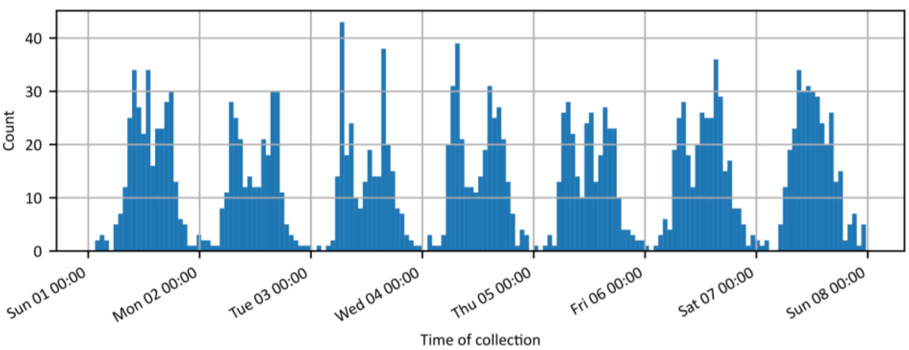
a) Bump



b) Pothole



c) Slip



d) Deceleration

Figure 4-1: The count of events per hour over the data collection periods

Whilst the fitting procedure enabled the crowdsourced data to be compared to current road condition data, it also helped improve the robustness of the analysis by ensuring the direction of travel and the correct road had been allocated to the event. This would not be the case if the analysis had simply clustered similar events using GIS tools (which could erroneously cluster events occurring on different roads).

One limitation of the fitting method was that the lane of the crowdsourced event is not identified. This is a result of the expected (but not tested) accuracy limitations of the crowdsourced coordinate data and because the definition of the SRN used does not include the coordinates for each lane. For multiple lane roads (a large proportion of the SRN), there is an uncertainty as to which lane the crowdsourced data is from. This is a particular challenge when considering events such as a potholes, which will only occur in one lane.

4.2.3 Aggregation and cleaning

The fitting process enabled the crowdsourced event data to be reported in relation to section and chainage. Events were reported as both 10m and 100m “values” in the dataset. Aggregate metrics included counts of each type of event in each reporting length, and minimum, mean and maximum values for events that included numeric data.

It was assumed that providers of the crowdsourced data had undertaken reliability checks on the delivered data, and therefore little further cleaning was undertaken. For one data provider, some events were flagged as invalid, so these were removed from the dataset prior to any analysis.

4.2.4 Exploring the clustering of events in the crowdsourced dataset

As discussed in Section 3, crowdsourcing relies on high numbers of participating “sources” (e.g. connected vehicles) to experience a defect or event. The data is anticipated to have low precision/accuracy for individual sources, but collation into clusters should improve confidence that any identified event is a true event that indicates potential deterioration in the condition. However, even with high numbers of vehicles passing a location where there is a potential defect, there is no guarantee that all vehicles will experience the defect. There are many reasons for this, including:

- If there are multiple lanes and the defect is in a single lane, then only vehicles in that single lane will detect it.
- If the feature is a potential hazard (e.g. a pothole), then drivers may change their driving line to avoid the feature and therefore the event will not be detected or reported.
- The detection of a slip is likely to be dependent on the vehicle accelerating or decelerating (braking) in some way. Not all vehicles passing a location may be doing this or doing it in the same way. Therefore, only a subset of the vehicles passing this location may report a slip event.

Unfortunately, the number of connected vehicles or mobile phones passing each location was not available from either data provider. This makes it impossible to determine the proportion of the vehicles travelling past the location of event that reported an event at that location. This would be a useful metric to aid interpreting the data.

To explore the behaviour of the basic clustering of events the geographical locations of the events were analysed using the DBSCAN clustering algorithm[4]. DBSCAN is a density-based algorithm which analyses clusters on the basis of the distance between individual points and the minimum number of points per cluster (Ester *et al.*, 1996). Clusters identified using DBSCAN can therefore contain points which are more than the minimum distance apart, but they will be connected by a chain of other points which do satisfy the maximum distance criteria. For the analysis the distance between points was set to 30m[5] and the minimum number of points was set to four. As deceleration events contained several coordinates the end point was used.

Note that this clustering investigation was based on the coordinate data (not section and chainage). After the DBSCAN algorithm had been applied, each cluster identified was checked to ensure that all events identified in the cluster came from vehicles travelling in the same direction and on roads of the same function. The latter check (roads of the same function) identified some invalid clusters on roundabouts over mainline sections. Where the checks were not satisfied the cluster was either split or removed, depending on whether the new clusters would still contain greater than the minimum number of reported events (4 in this case).

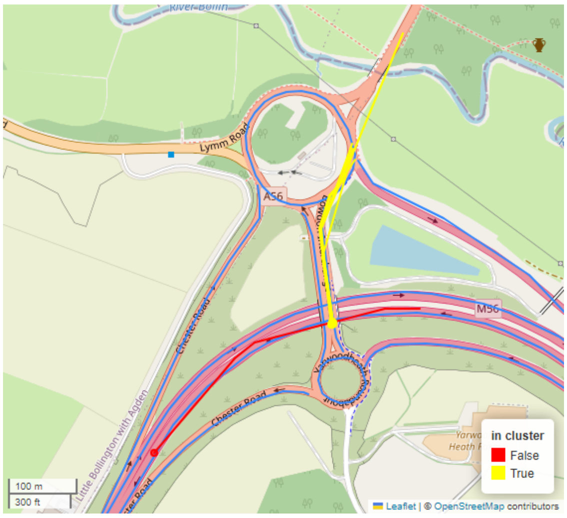
Figure 4-2 shows examples of clusters that were identified in the analysis. Here individual events are shown as points on a map background. Many events can be seen, with the events considered to form part of a cluster presented in yellow. The importance of "road context" becomes apparent in these plots. For example, a simple geographical clustering analysis could incorrectly allocate mainline deceleration events to the same cluster as the roundabout deceleration events in the right of Figure 4-2, which could result in over-estimation of the risks present.

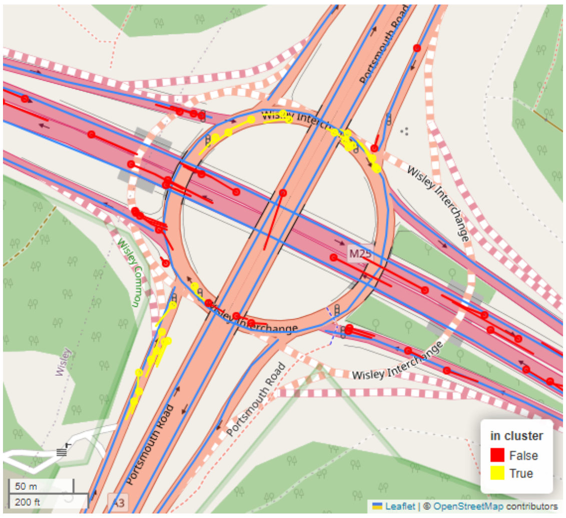
A summary of the clustering results is shown in Table 3. The event for which the largest percentage of events also fell into a cluster is bump, followed by pothole. It may be that bumps and potholes, which are point features, are more likely to be successfully identified as a cluster because they will be reported in similar locations by each passing device, whilst deceleration or slip events are less precise in location. Alternatively, it may be that the data is indicative of high numbers of widely dispersed braking events, with only a small proportion occurring at similar locations.

Bump events also have largest number of events per cluster, followed by slip (Figure 4-3). Note that bump, pothole and slip were all reported by the same fleet of vehicles over the same time period. The cause of the differing sizes of clusters is unclear. It may be that bumps are more likely to be detected than other event causes due to the ability of this to be detected by the vehicle telemetry data.

[4] The DBSCAN implementation used was from the Python library scikit-learn (Pedregosa *et al.*, 2011).

[5] 30m was selected based on the estimated accuracy of the vehicle/phone location measurement system (GNSS).





a) Slip, 1 cluster visible **b) Deceleration, 3 clusters visible**

Figure 4-2: Visualisation of example clusters

The length of clusters was estimated by calculating the minimum bounding rectangle for the points in each cluster and taking the longest side as the cluster length. The distribution of lengths is shown in Figure 4-4. The cluster lengths for bump and pothole are similar and the smallest. This may be expected because they are point features. The length of these clusters is probably indicative of the error in the position measurement. Slip and deceleration occur over lengths, with the cluster length perhaps indicative of both the deceleration length and the position error.

Table 3: The proportion and size of clusters for each type of crowdsourced data

Cause	Bump	Pothole	Slip	Deceleration
Data source	Vehicle telemetry	Vehicle telemetry	Vehicle telemetry	Mobile phone
Number of clusters	34	15	26	32
Count of events not allocated to a cluster	195	188	506	1993
Estimated total number of event sites (sum of above rows)	229	203	532	2025
Total number of events	443	276	672	2158
Percentage of total number of events which are in a cluster	55%	31%	25%	8%

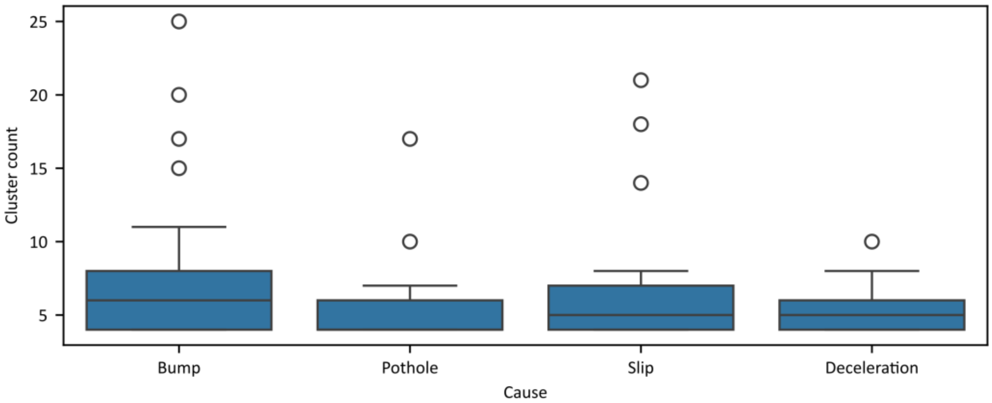


Figure 4-3: Box plots of the count of events per cluster

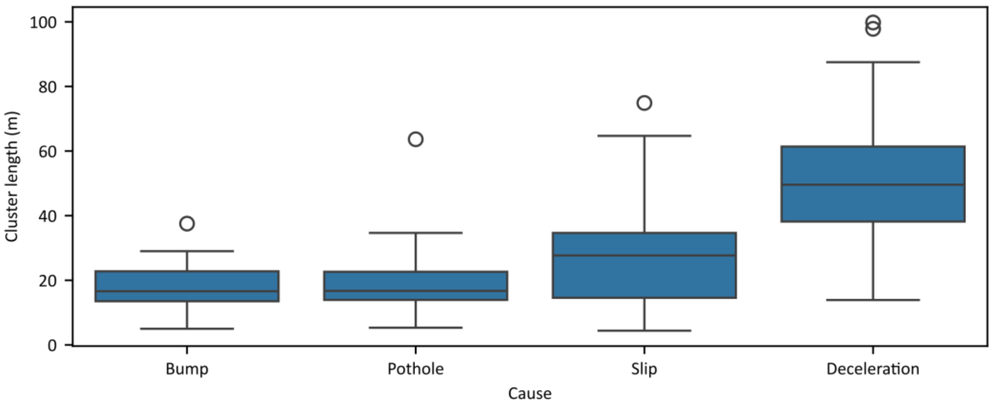


Figure 4-4: Box plots of the estimated length of the clusters

4.3 Approach to analysis

As noted in Section 1, the goal of this work has been to explore how the data provided from crowdsourcing could be applied to better understand or manage the condition of the SRN. The analysis was therefore centred on posing questions regarding the use of the data in specific applications, and then undertaking analysis to determine whether there was evidence that the available dataset could be used to answer that question.

The questions were separated into four areas, as follows:

- As the telemetry dataset provided event data that relates to the experience of road roughness (bump and pothole events, including severity), can telemetry data be used to assess functional condition? This is presented in Section 5. The analysis focusses on the comparison between TRACS functionality (roughness) data and the bump events provided in the telemetry data in Region 1.
- As the telemetry dataset also provided event data that monitored the requirement for the connected vehicles to decelerate or stop on the network (slip and deceleration events), can telemetry data be used to assess safety condition? This is

presented in Section 6. The analysis focusses on the comparison between SCRIM safety (friction and supporting data such as site category) data and the slip events provided in the telemetry data in Region 1. The data is explored from a number of perspectives – the potential to replace SCRIM data, the potential to support the identification / assessment of potential high-risk individual sites, the potential to support more robust / more evidence-based establishment of Site Categories, and the potential to identify sites where there may be higher risks of collisions occurring.

- As the mobile phone dataset provided data the that focussed on vehicle deceleration (deceleration events for any deceleration greater than $3.2m/s^2$), can mobile phone data be used to assess safety condition? This is presented in Section 7. The focus of the analysis is broadly similar to that applied to the telemetry data. Hence comparison is undertaken between SCRIM safety (friction and supporting data such as site category) data and the deceleration events provided in the mobile phone data in Region 2. The data is explored to determine whether the mobile phone data can support the identification / assessment of potential high-risk individual sites, and to identify sites where there may be higher risks of collisions occurring.
- Finally, robust application of crowdsourced condition data will require sources of data that provide adequate coverage of the network. Some implications of the coverage for specific applications are discussed in Section 8.

5 Application of vehicle telemetry crowdsourced data to assess functional condition

5.1 Network level assessment of bumpiness

Question: TRACS data provides a measure called the Bump Index that reports the location of bumps. Can the telemetry data provide network-level assessment of bumpiness?

The experience of road users can be adversely affected by the presence of bumps caused by defects such as potholes, poor drainage covers, failed reinstatements etc. On the SRN bumps are reported using the TRACS Bump Index. The measure was developed via comparison with data from user experience studies, aiming to provide a general assessment of the extent to which a feature identified within the measured profile will affect user experience. However, TRACS provides a snapshot of the condition of the pavement on a single day each year. Bump defects can form rapidly, and hence bump data may be inaccurate within weeks after the survey date. Crowdsourcing data can be collected year-round, to provide an up-to-date representation of the pavement condition. Also, as it reflects the actual vehicle response, it may better reflect real world bump experience. The relationship between vehicle telemetry bump data and the TRACS Bump Index was therefore explored to understand how crowdsourced could complement data provided by the Bump Index.

The 100m lengths where the telemetry crowdsourced data reported at least one bump or pothole was used as the basis to filter the TRACS Bump Index dataset into two datasets – lengths where the crowdsourced data reported a bump, and the whole network. The distributions are shown in the left of Figure 5-1. It can be seen that higher values of Bump Index are reported on a greater proportion of the dataset in which crowdsourced data reported a bump. This analysis was repeated for the TRACS 3m eLPV measurement, which indicates the presence of short longitudinal profile deviations, shown in the right of Figure 5-1. Again, higher values of 3m eLPV are reported on a greater proportion of the dataset in which crowdsourced data reported a bump. These comparisons suggest that the crowdsourced data is indicating the presence of rougher roads and bump-like features that are also reported by TRACS, and indicates that the crowdsourced data may provide the ability to characterise roughness at the network level.

Although there appears to be a broad indication within the distributions that the crowdsourced bump measures are reflecting lengths that TRACS considers to be bumpy, a direct comparison between the TRACS and crowdsourced intensities does not show a strong relationship. Figure 5-2 shows the relationship between the Bump Index and the magnitude of the crowdsourced bump. To obtain this, the dataset was divided into six subsets corresponding the 100m lengths in which the crowdsourced data reported a bump magnitude of 1 to 6. The average TRACS Bump Index was then calculated for each subset. This shows no overall relationship between the TRACS Bump Index and the crowdsourced bump magnitude. This analysis was then repeated, in this case determining the average Bump Index value over lengths reported by the crowdsourced pothole data in the range 1-6. This shows a relationship, but the reverse of that which would be expected. We are not able to explain this behaviour. If the crowdsourced data is reporting rapidly developing and/or the worst bumps (potholes) it may be that TRACS has not identified these because they were either not present when the survey took place, or that local treatments have taken place and they were repaired before the survey took place. Alternatively, the TRACS Bump Index may have failed to identify bumps which do cause bumpiness events in real vehicles – or vice versa.

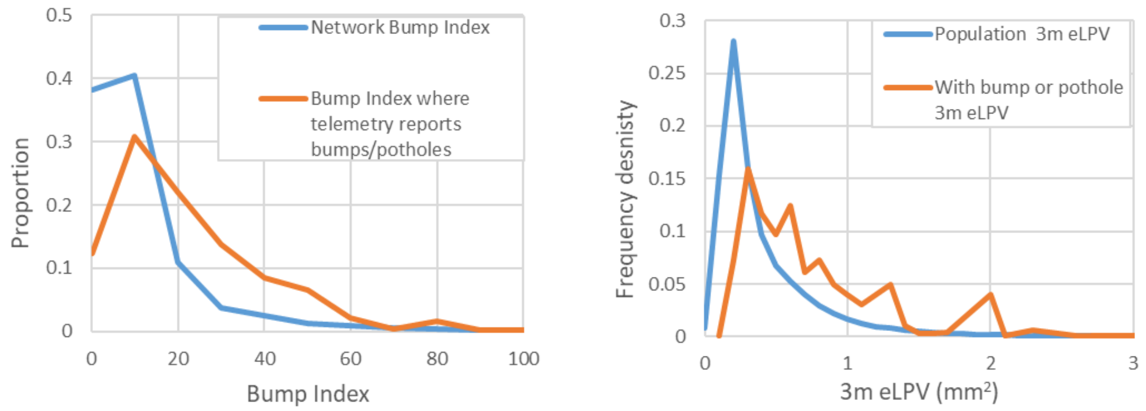


Figure 5-1: Comparison of Bump Index distributions (left) and 3meLPV distributions (right) for the whole SRN and the SRN where crowdsourced data reported a bump

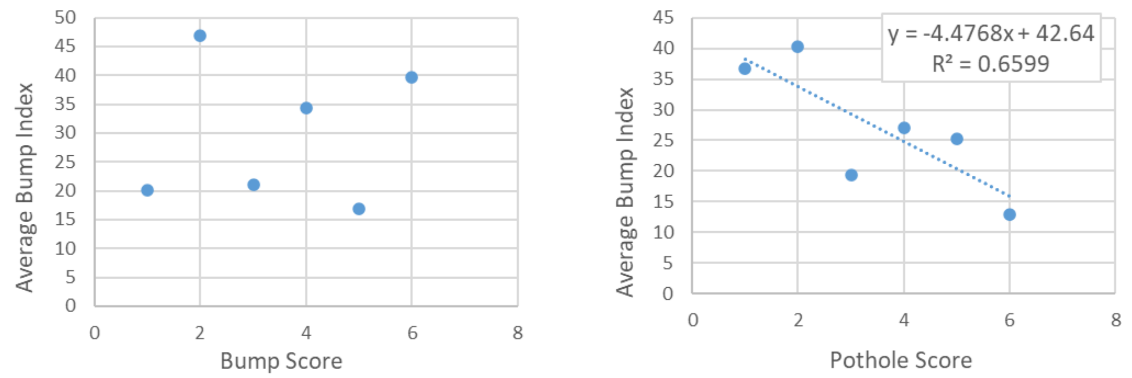


Figure 5-2: Relationship between crowdsourced bump magnitude (left) and pothole magnitude (right) and the average Bump Index

5.2 Identifying and tracking specific bumps / potholes

Question: Can telemetry provide information to identify and track the development of specific bumps / potholes?

Although no consistent relationship was found above between the bump / pothole score and TRACS derived measurements on a network level, examination of the data on specific sites shows that is possible to use the telemetry data to identify specific features. For example on the A14 we were able to track the development of a pothole, as shown in Figure 5-3. The bump values reported over a 20 day period were tracked manually in the data, as shown in Figure 5-4. The National Highways database marked this feature as repaired approximately two weeks after the final date reported in the telemetry dataset. This suggests that the vehicle telemetry data is identifying significant features, despite a lack of consistent correlation with the TRACS measurements. Further study would be required to validate and make effective use of this data source.

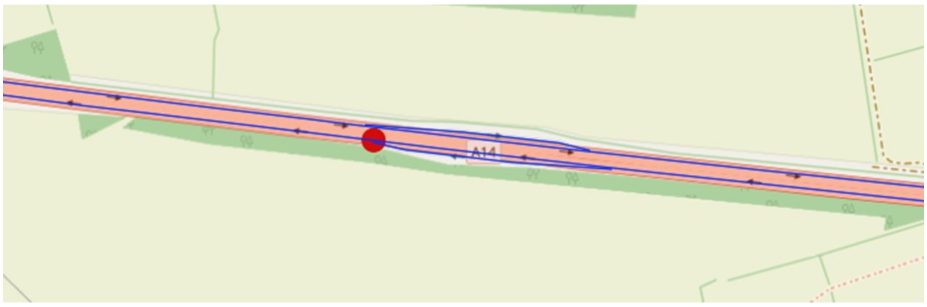


Figure 5-3: Location of pothole on the A14, around the exit of a layby

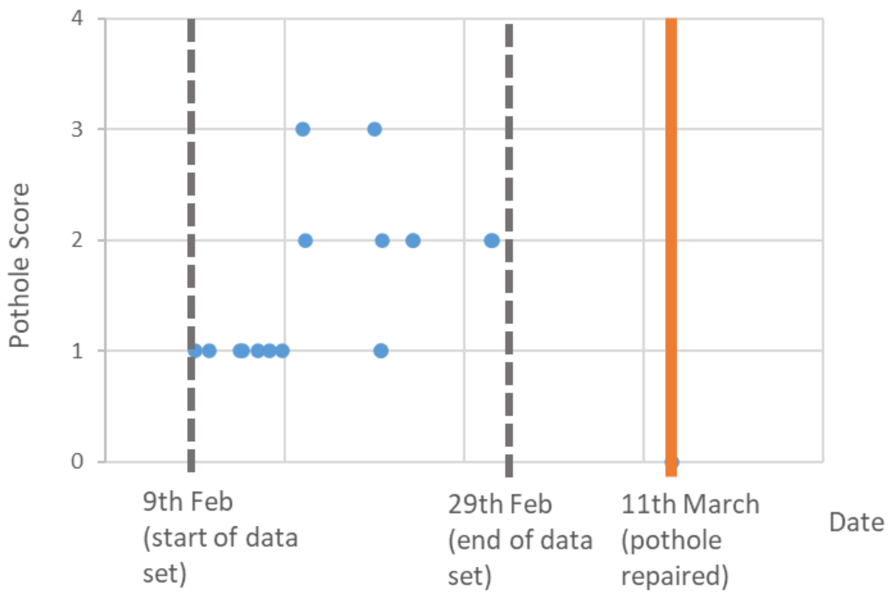


Figure 5-4: Evolution of pothole scores over time from start till end of dataset timespan

6 Application of vehicle telemetry crowdsourced data to assess safety condition

6.1 Telemetry data for the measurement of skid resistance

Question: Can locations where the telemetry reports the road is slippery provide insight into the requirement for skid resistance related maintenance?

The vehicle telemetry data reports a 'slip' event. The method deployed by the telemetry data supplier to determine this was not provided, but it is assumed that the event is intended to indicate a situation where the skid resistance was, to some extent, low, and when the vehicle was braking there was evidence of slip (ABS activation). However, no information is provided regarding the braking force applied and duration (and hence the demand for skid resistance). To explore the behaviour of this data, the skid Site Categories for the locations where slip events occurred were determined. The distribution of slip events by Site Category is shown in Figure 6-1. The figure also shows the distribution of site categories for the whole network covered by the telemetry data (Region 1). This shows that the proportion of slip events reported to occur on sites categorised as Q and R[6] (see also Table 5) far exceeded the proportion to which these site categories are present on the road network, suggesting that drivers were significantly more likely to experience slip on sites categorised as Q and R.

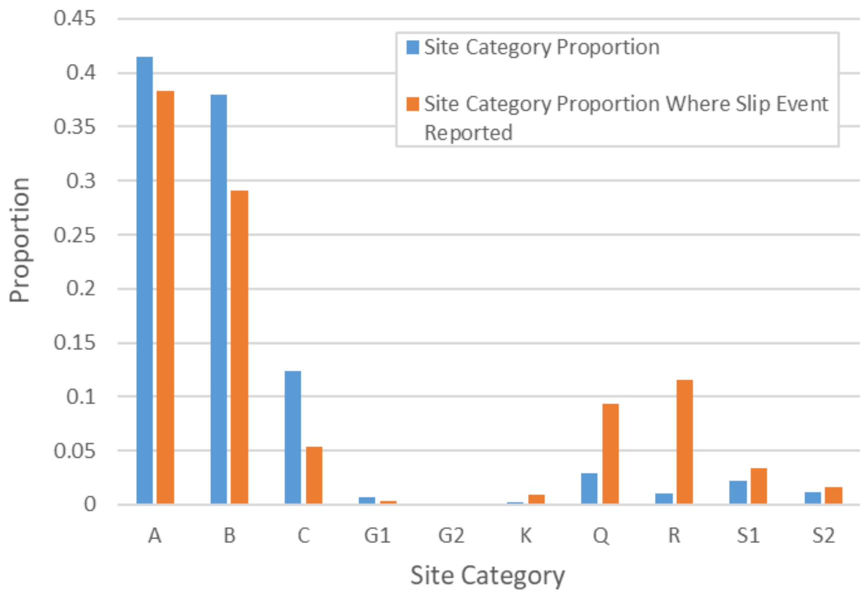


Figure 6-1: Comparison of overall distribution of highway Site Categories compared to the distribution of slip events by Site Category

[6] As defined in standard CS 228 (Highways England, 2021). Q: approaches to and across minor and major junctions and approaches to roundabouts. R: Roundabouts

To explore the data further Figure 6-2 shows the distribution of SCRIM Difference values for the (Region 1) network. Figure 6-2 also shows the distribution SCRIM Difference values for those lengths on which a slip event was reported. Lengths on which a slip event occurred are much more likely to have negative SCRIM Difference compared in comparison to overall network.

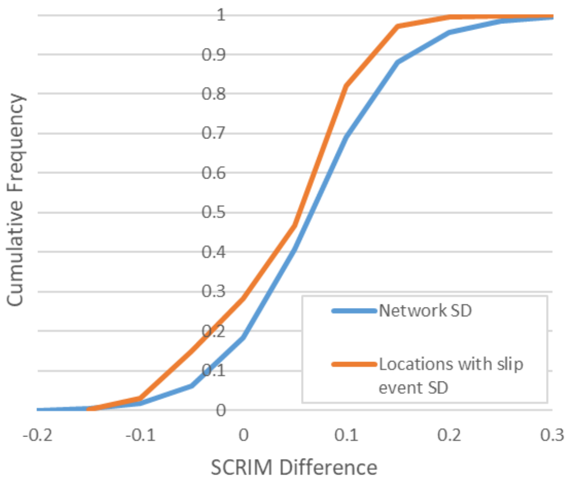


Figure 6-2: Cumulative frequency of SCRIM difference values for whole dataset compared with pavements which have an associated slip event

Although the distributions suggest that slip events may be associated with greater SCRIM difference, direct comparison between these datasets does not show a strong relationship. Figure 6-3 directly compares the average Corrected SCRIM Coefficient values, and SCRIM difference values for each section with the frequency at which slip events occurred in that section. These scatter plots show no appreciable relationship with either CSC or SCRIM Difference.

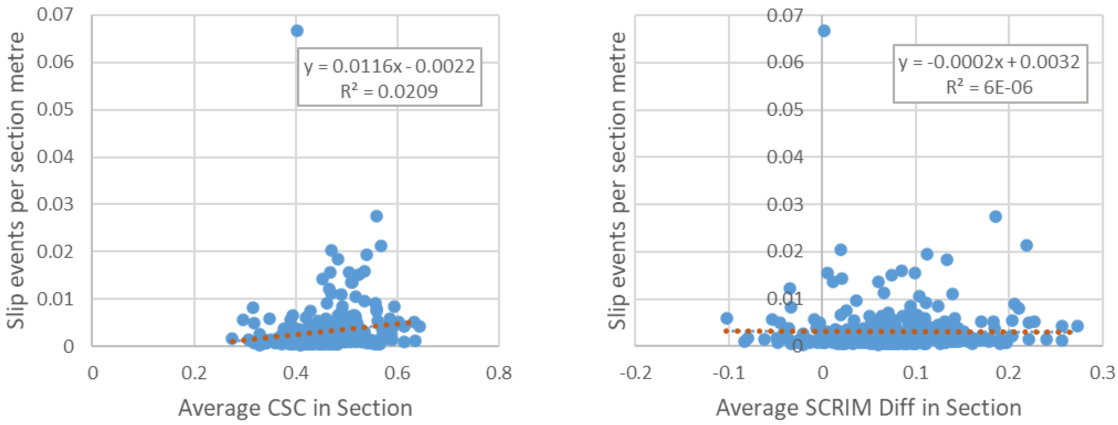


Figure 6-3: Relationship between Slip events and average Corrected SCRIM Coefficient (left) and SCRIM difference (right), normalised by section length

The lack of a direct relationship is, perhaps, not surprising as the interaction of driver behaviour, braking demand and friction supply is complex. In addition, the dataset studied in this work covered a relatively short telemetry measurement period, which may have been insufficient (insufficient vehicle passes) to build up a robust assessment of the friction. The slip events data available to this study therefore does not appear to provide a general/direct indicator of skid resistance. However, this does not mean that the does not contain valid information regarding hard braking events and their significance for providing adequate levels of skid resistance, as discussed further below.

6.2 Identifying higher risk sites – case studies

Question: Can the data be used to identify specific locations that may require higher levels of skid resistance, which are not clearly identified using current methods.

To establish if higher risk sites could be identified using crowdsourced data, the vehicle telemetry data was reviewed to locate clusters of slip events. These are discussed as "case study" locations.

6.2.1 Case study example on M5 mainline northbound, Birmingham

Clustering of slip (i.e. hard braking) events is not typically expected to occur on main carriageways. On these road types heavy braking will often arise as a result of the behaviour of other road users, which will not be confined to any given location, unlike braking for a junction, which is in a fixed position. However, two clusters of slip events were reported over two days at a similar location around a bridge carrying traffic over a small river, the clusters of events are shown in Figure 6-4. The vehicle telemetry data shows, on average, 1 slip event per twenty sections of road, making this cluster very unlikely to occur through chance alone. The pavement at this location has a positive SCRIM difference value, suggesting that it is unlikely that the slip events are only due to low skid resistance, and that they represent hard braking events that lead to slip.

Figure 6-4: Location of slip events on the M5 clustered around an overbridge

Finding a definitive explanation for the braking events is not possible; however, investigation of the site found four sets of stairs allowing access to the main carriageway from each side corner of the bridge. One such set of stairs is shown in Figure 6-5. A potential explanation for the deceleration could be maintenance operatives accessing the carriageway to inspect the bridge, or trespassers using the bridge to cross the river.

Figure 6-5: View from the M5 carriageway showing top of bridge/embankment access stairs

6.2.2 Case study example on the A46, north of Syston

Table 4 shows the SCRIM Difference values for a section of pavement on the A46 which shows a range SCRIM difference values, some less than zero (which suggests a lower level of skid resistance). The site may therefore be subject to a site investigation to determine any requirement for remedial action. Visualisation of the telemetry data (Figure 6-6) shows the occurrence of slip events on this section. Events can clearly be seen clustering around a bus stop. The presence of this bus stop may be causing some drivers to brake hard. The crowdsourced slip data could provide evidence to support action at this location.

Table 4: SCRIM Difference measurements on pavement section on the A46 showing low and deficient values

Start Chainage	End Chainage	Site Definition	SCRIM Difference
0	63	B	-0.04
63	163	B	0.03
163	263	B	0.035
263	314	B	0.027
314	414	B	0.024
414	474	B	0.028
474	526	B	0.029
526	607	B	-0.038
607	677	B	0.049
677	777	B	-0.033
777	790	B	-0.048
790	865	B	0.101
865	916	B	0.058

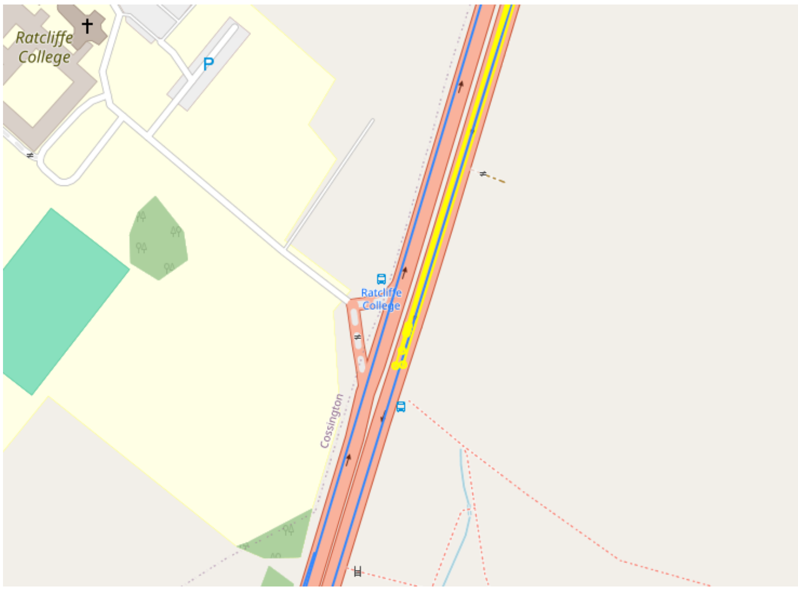


Figure 6-6: Cluster of slip events around a bus stop on the A46 north of Syston

Figure 6-7: Photo of the location of the cluster of slip events near a bus stop and footpath crossing on the A46 north of Syston

6.2.3 Case study example on the A45 eastbound, Wellingborough

On this mainline section (Figure 6-8) the pavement has an average SCRIM difference of -0.0516, which suggests that it may require investigation. The section has two lanes, with a slip road merging into them. This is a point of conflict with high levels of traffic that may suggest that the site has a higher level of risk and hence intervention may be required.

However, there are no slip events on this section, suggesting that hard braking is not occurring and the risk may be low.

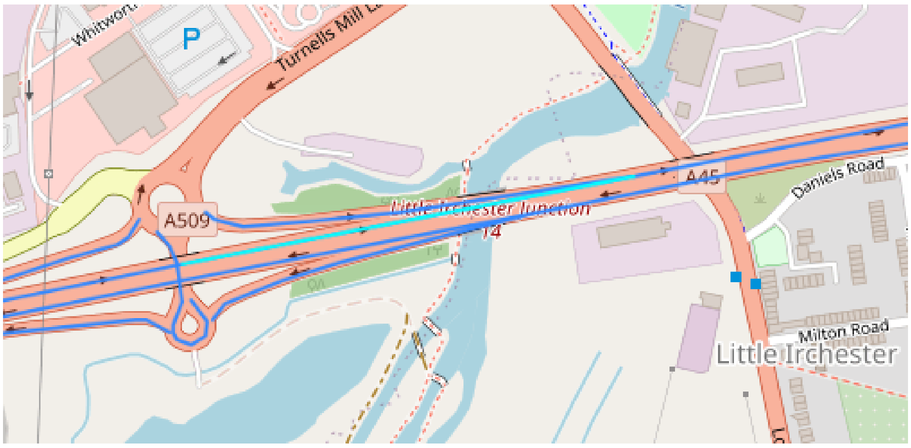


Figure 6-8: Section of east bound main carriageway on the A45

6.3 Identifying higher risk sites – treatment and categorisation

Question: What approach could be used to apply crowdsourced data more objectively to support decisions on treatments?

As has been observed in previous sections, it is likely that the point events provided by the crowdsourced data will accumulate, or cluster, in locations where there are defects present. To make use of this data in condition assessment and asset management it will be necessary to identify and locate these clusters. In this section we consider the use of a Kernel Density Estimator to better understand the locations of point events reported in crowdsourced data. The kernel density estimator (KDE) can be used to estimate the probability density function of a continuous variable using point data. In a KDE, a kernel function, typically a symmetric and smooth function such as the normal distribution, is placed over each data point. The estimator then calculates the density at any point in the distribution by summing the contributions of all kernels. The degree of smoothness is controlled by a bandwidth parameter, which determines the width of each kernel (SciKit - Learn, 2024). The KDE could be used, for example, to locate the onset and development of potholes by identifying increasing probability of bump events occurring at a location, or to identify locations containing higher risk slippery roads as slip events cluster.

In this study we have applied the KDE to convert the point slip event data into a probability density function of slip likelihood. To do this it is necessary to model the distribution of braking distances. However, this was not available in the telemetry data as only point values were provided. Therefore, the distribution of braking distances was estimated using the smartphone deceleration event data (which included information on the start and end positions of braking events). Figure 6-9 shows the distribution of deceleration lengths from this dataset, displaying an approximately normal distribution. This was used as justification for selection of a gaussian Kernel function.

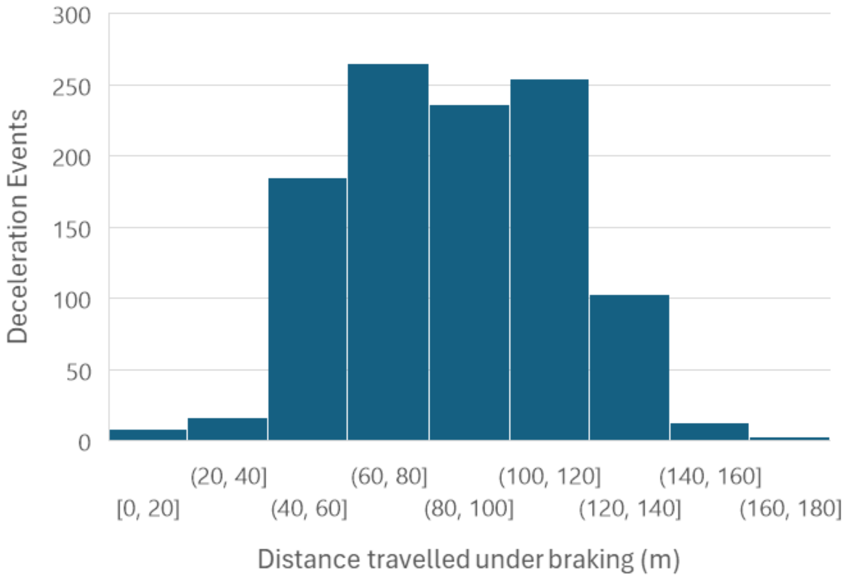


Figure 6-9: Approximately normal distribution of distance travelled under braking

The smartphone deceleration data indicated that the variance of deceleration length was ~32m for A roads. This was used to convert the slip point data into a probability density function for event location. This is shown tor the approach to the bus stop on the A46 north of Syston (used in the case study of Section 6.2.2) in Figure 6-10. The plot shows a high probability density for slip events immediately before the bus stop. There is also a lower peak after the bus stop, which may be a result of drivers braking as vehicles leave the bus stop. Whilst the case study discussed above identified (via manual visual interpretation) a potential higher-risk location at the cluster of slip events, the KDE approach can define, objectively, a length over which there is higher risk, which could assist the road authority when deciding where to apply treatments.

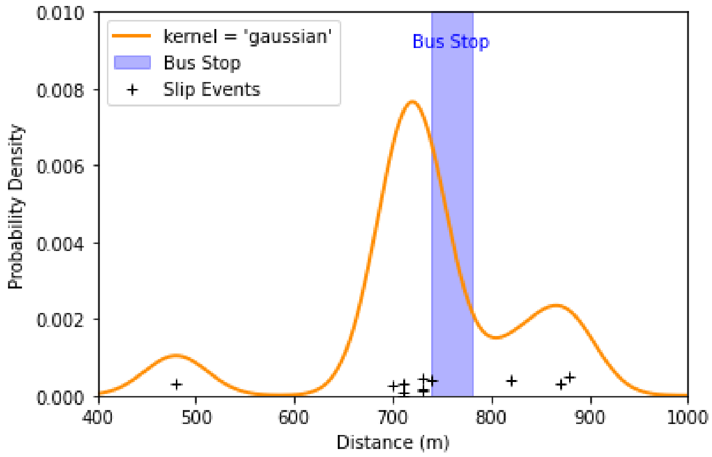


Figure 6-10: Kernel Density Estimator plot of Slip events approaching and beyond a bus stop on the A46 north of Syston

The KDE method may also be applied when reviewing skid site categories. For example, Figure 6-11 shows a slip road leaving the A45, featuring three site categories: the first category is B, representing a featureless carriageway; the second is category G1 representing a road with a steep downhill gradient; and the final category is Q representing a junction approach where vehicles will routinely come to a stop. The complexity of this slip road creates challenges in setting Site Categories. Drivers may need to brake earlier in this slip road due to the steep gradient, and there is a possibility of a queue forming, with and drivers potentially coming to a stop several hundred meters before the junction.

Figure 6-12 shows the KDE plot of the events on this slip road. The KDE plot shows that the events align approximately with the Site Categories, with increasing braking probability as the vehicles approach the junction. This plot implies that the Site Categories have been correctly selected and account for driver behaviour at the junction (noting that the volume of data is low and that with further slip events the distribution could shift).

As the KDE probability densities are normalised, it is difficult to compare sites with different numbers of slip events. However, they can be converted to pseudo frequency by multiplying the PDF by the total event count. Two case studies (the A45 slip road and A46 bus stop at Syston discussed above) are shown in Figure 6-13. It can be seen that there is a higher peak in the frequency of slip events behind the bus stop on the Syston site compared to the Q site category on the junction approach at the end of the A45 slip road. Despite this, the road behind the A46 bus stop is site category B (with a less demanding SCRIM Investigatory Level). The crowdsourced data suggests this may not be appropriate.

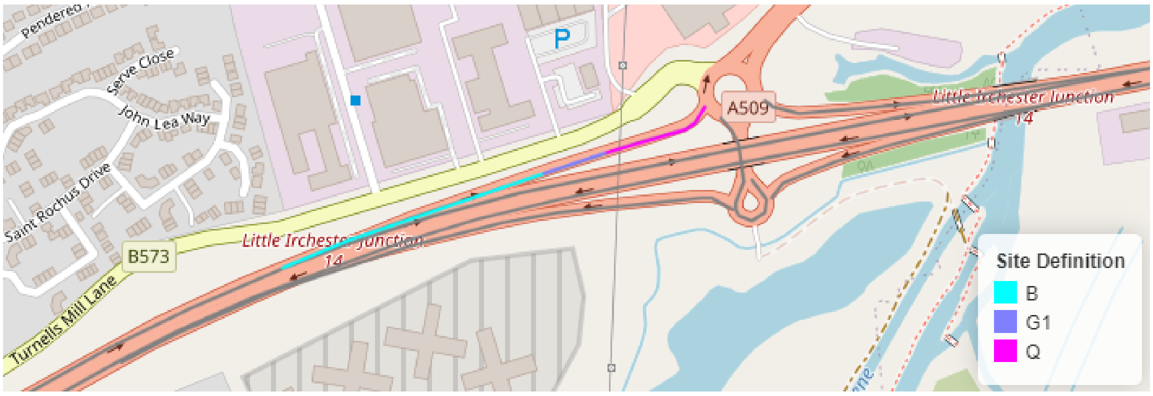


Figure 6-11: The site definitions on a slip road leaving the A45

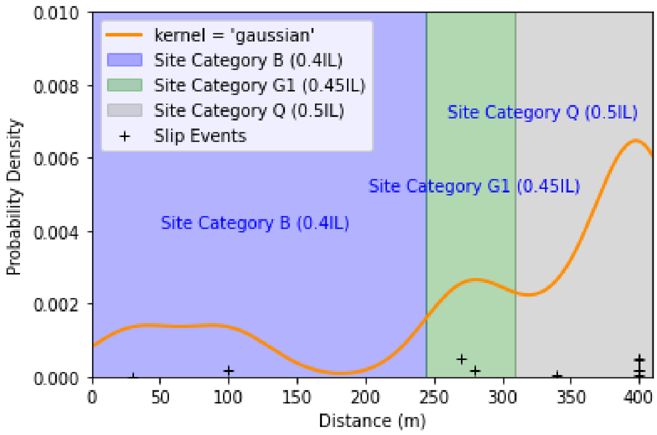
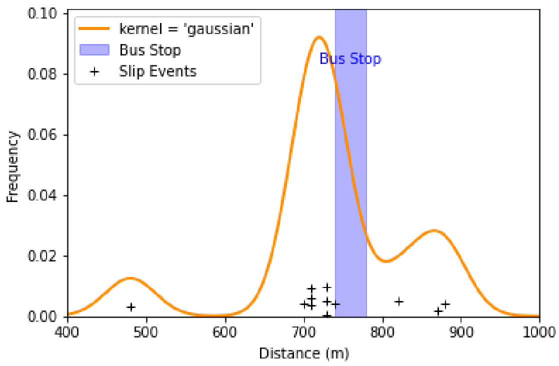


Figure 6-12: Kernel Density Estimator plot of slip events approaching on a slip road leaving the A45



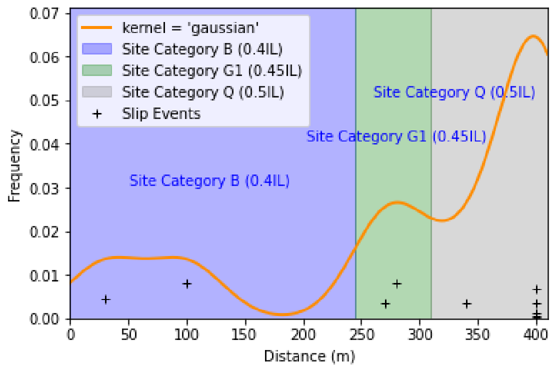


Figure 6-13: Kernel Density Estimator plot of slip events approaching and beyond a bus stop on the A46 (left) and A45 (right) converted to show pseudo frequency

As noted above, the KDE approach could also be applied to understand the development of clusters of other point data provided by crowdsourcing data, such as pothole or bump events relating to functional condition. This would provide an objective tool to bring the developing pothole to the attention of the maintaining engineer. It would also assist in maintaining records of the locations of potholes in the asset database, so that lengths in which higher numbers of potholes have developed can be monitored. This information is of use when making decisions on longer term treatments.

6.4 Is there a relationship between crowdsourced telemetry data and collision risk?

Question: Is there potential to use telemetry data to identify sites with higher risk of collisions occurring?

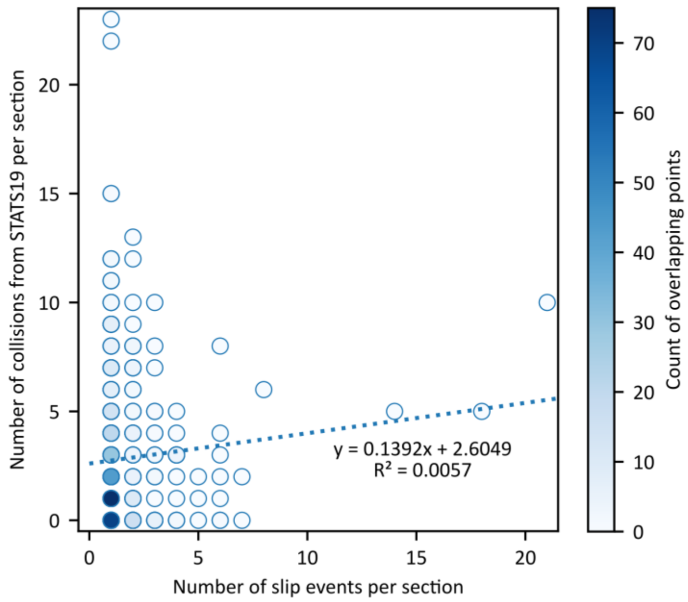
Collision records form part of the risk assessment process when determining the skid resistance requirements for locations on the network. Collisions occur rarely, and therefore for any given location that may be at risk of a collision occurring, this may not have yet occurred. Telemetry data has a higher density of events, which could be considered “near misses”. If there is a relationship between the rate of events and ultimate rate of collisions, telemetry could be used as a proxy for collision records when assessing risk.

To investigate this, the relationship between slip events and STATS19 collisions was evaluated. All available slip events and STATS19 collisions from 2018 to 2023 were used in this analysis. However, there is a large imbalance between the number of STATS19 collisions (19,743) and the total number of slip events (672), which produced an unbalanced dataset.

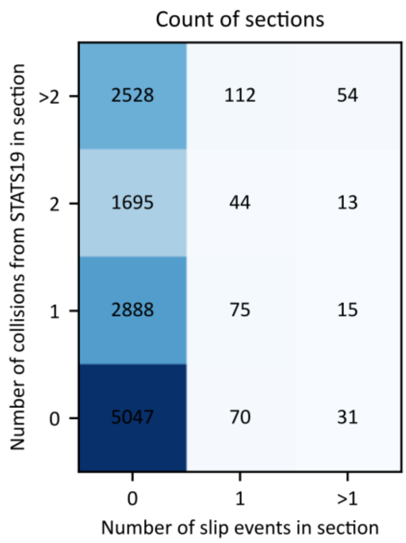
The left of Figure 6-14 plots the count of collisions reported in the STATS19 data against the number of slip events reported per section. Note that the numbers of events (STATS19 or slip) are reported as whole numbers, and therefore there are several overlapping points in the left plot. The colour bar shows the count of these overlapping points. To see if there is any relationship in sections with a non-zero count of slip events, sections with zero slip events (which is most of the dataset) were excluded from the left plot of Figure 6-14. The fit of a linear regression line, as shown in the plot shows poor correlation. There is no clear relationship between the slip events and collisions for this subset of the dataset.

The right plot in Figure 6-14 shows an overview of the dataset. It shows the count of sections which have specific ranges of STATS19 collisions and slip events. These ranges are based approximately on the quartiles of STATS19 collisions and slip events respectively. However, due to the imbalance of this dataset, the “quartiles” are not close to the true quartiles, but instead are chosen to show the range of values in the data. If there was a strong correlation between the number of STATS19 collisions and slip events, then it would be expected that the sections with the lowest quartile of number of STATS19 collisions would also be the sections with lowest quartile number of slip events. This pattern would continue for the other quartiles, so the diagonal from bottom left to top right would contain most of the sections. This is not what is shown in the right plot of Figure 6-14, and instead no strong correlation is shown.

It appears that the quantity of events present in the telemetry dataset is insufficient to conclude that there is a relationship between the slip events and STATS19 collisions across this dataset. It must be noted that the dataset covered a very short reporting period and there may be a stronger relationship obtained for data collected over a longer assessment period.



Scatter plot and liner regression line of best fit where sections with zero slip events have been excluded from the plot and regression calculation. Due to many points in the scatter plot being plotted at exactly the same location, the points have been given a colour scale to show the count of overlapping points.



Histogram with bins based on approximate quartiles, where sections with zero slip events have been included in the plot.

Figure 6-14: Plots of the total number of collisions from STATS19 from 2018 to 2023 and slip events per section

7 Application of mobile phone crowdsourced data to understand safety condition

7.1 Relationship with SCRIM category and investigatory level

Question: Current policy establishes requirements for skid resistance in terms of the demand and risks presented by specific section of road. Does deceleration data from mobile phones provide any insight?

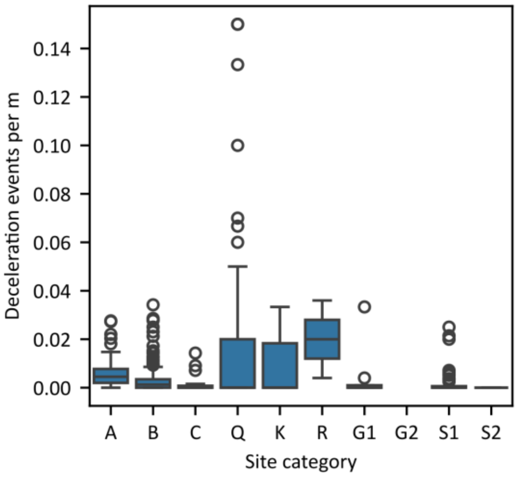
Currently, the required skid resistance (the investigatory level) for a length of road is set based upon its function and geometry. These are summarised in Table 5. The requirements are typically specified in terms of a range of skid resistance, from which engineers select an investigatory level appropriate to their understanding of risk on the site.

Table 5: SCRIM site categories and investigatory levels (Highways England, 2021)

Site category	Definition	Investigatory level for skid resistance	
		Low risk site	Standard site
A	Motorway	0.30	0.35
B	Non-event carriageway with one-way traffic	0.30	0.35, 0.40
C	Non-event carriageway with two-way traffic	0.35	0.40, 0.45
Q	Approaches to and across minor and major junctions, approaches to roundabouts and traffic signals	N/A	0.45, 0.50, 0.55
K	Approaches to pedestrian crossings and other high-risk situations	N/A	0.50, 0.55
R	Roundabout	N/A	0.45, 0.50
G1	Gradient 5-10%, longer than 50 m	N/A	0.45, 0.50
G2	Gradient >10%, longer than 50 m	0.45	0.50, 0.55
S1	Bend radius <500 m – carriageway with one-way traffic	N/A	0.45, 0.50
S2	Bend radius <500 m – carriageway with two-way traffic	0.45	0.50, 0.55

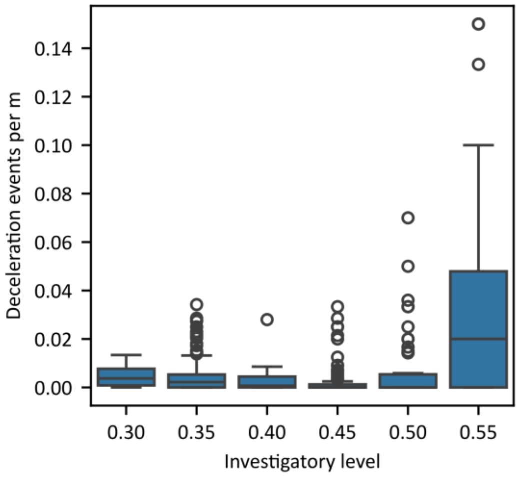
On the SRN P-AMS sections can be divided into subsections which have different site categories (and hence investigatory levels). Therefore, to undertake this analysis the network was divided into lengths as defined by the P-AMS section and the site category (and investigatory level) in the section (to obtain “site category sections”). The count of deceleration events within each “site category section” was then normalised by dividing by the length of that “section”. Note that this did not take the number of lanes in the “section” into account. Box plots of the distributions of deceleration events per m by site category and by investigatory level are shown in Figure 7-1. Tables of the numbers of counts of site category (or investigatory level) “sections” are also included to aid interpretation. It can be seen that:

- For site category, the categories with high upper quartiles of deceleration events per metre reported in the mobile phone data are Q, K and R. As may be inferred from the definitions provided in Table 5, vehicle stopping is likely to take place in these categories of site. The greater number of deceleration events reported by the mobile phone data is therefore consistent with the site categorisation of these sections established by the skid policy.
- When considering the distribution of deceleration events per metre by investigatory level, most of the distributions are similar except for the highest, 0.55. There is a smaller number of “sections” in the distribution of deceleration events per metre for this investigatory level. However, the mobile phone dataset does suggest that the lengths for which higher investigatory levels have been set do experience a greater proportion of harsh barking events. Hence deceleration is occurring where it is expected that there would be more demand for friction.



Site category	Count of “sections”
A	182
B	290
C	15
Q	74
K	11
R	2
G1	8
G2	0
S1	48
S2	3

a) Site category



Investigatory level	Count of “sections”
0.30	18
0.35	434
0.40	30
0.45	75
0.50	50
0.55	26

b) Investigatory level

Figure 7-1: Box plots of the deceleration events and respective counts

7.2 Identifying higher risk sites – case study

Question: Can the data be used to identify specific locations that may require higher levels of skid resistance, which are not clearly identified using current methods?

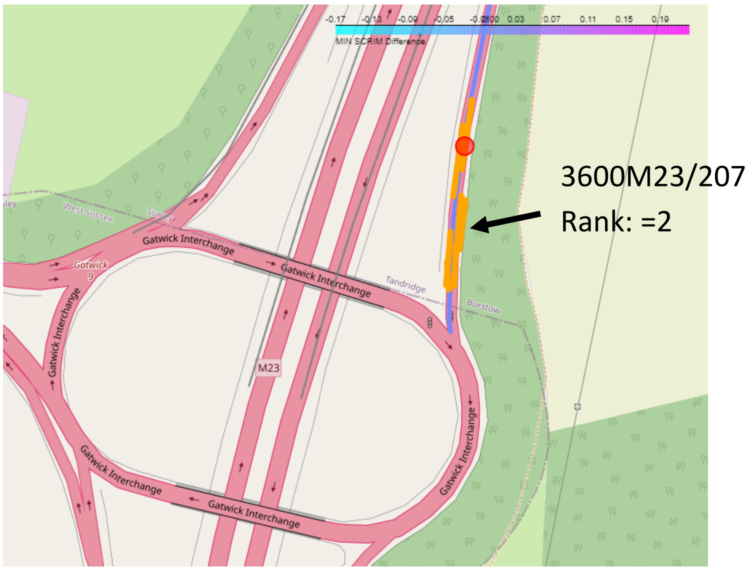
As for the telemetry data, we have investigated whether higher risk sites could be identified using mobile phone crowdsourced data. To conduct this study, sections in Region 2 which contained lengths defined as junction approaches, Q, or roundabouts, R, were identified – there were 75 such sections. Deceleration events and STATS19 collisions (from 2018 to 2023) occurring in these sections were also investigated to identify those containing the highest number of deceleration events or STATS19 collisions. These are listed in rank order in Table 6. It can be seen that several of the sections reported to contain a high number of deceleration events also have a high occurrence of STATS19 collisions. The five sites with the highest number of deceleration events are shown in Figure 7-2. Section 3600M23/207 (Figure 7-2a) has a high number of deceleration events but only one STATS19 collision. However, the next four deceleration sections all feature in the top seven collision sections.

The data suggest that cluster-based analysis of deceleration events could be useful for the identification of higher risk sites. Further insight may be provided through more in-depth investigation of current sites where there is both a strong and a weak relationship. For example, section 3600M23/207 appears to have sustained a higher number of deceleration events with a lower proportion of collisions. Is this due to pavement design or condition, driver behaviour or another factor? Also, we note that the clustering approach itself can affect the analysis when primarily performing analysis based on P-AMS sections. The junction between 3600A3/308 and roundabout 3600A3/195 shown in Figure 7-2b has a high number of collisions, but these are separated across two sections, but this is effectively one location with a high number of collisions. This suggests that breaking the data on such an arbitrary basis could limit the analysis.

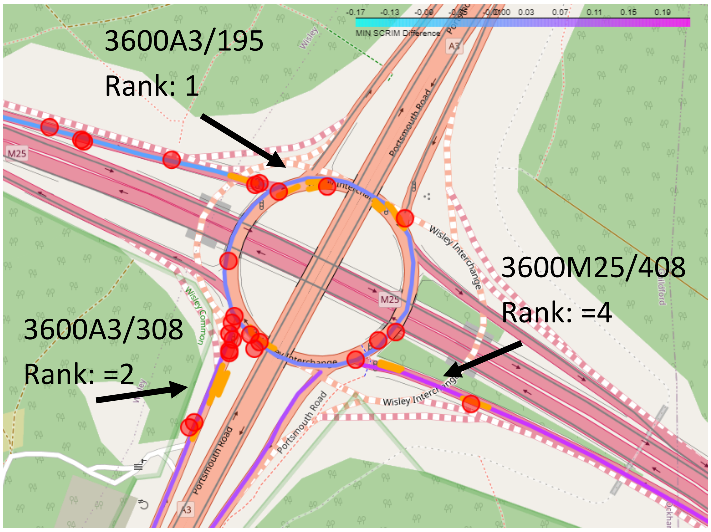
Table 6: The 10 sections with the highest number of deceleration events or STATS19 collisions

Section label	Rank	Number of deceleration events
3600A3/195	1	9
3600A3/308	2	7
3600M23/207	2	7
3600M3/479	3	5
3600A3/143	4	3
3600M25/408	4	3
3600M3/110	4	3
3600A3/148	5	2
3600A3/317	5	2
3600A3/392	5	2

Section label	Rank	Number of STATS19 collisions
3600A3/195	1	12
3600M3/506	2	8
3600A30/181	3	7
3600A3/308	4	6
3600M25/367	4	6
3600M25/337	4	6
3600A30/149	5	5
3600M25/479	6	4
3600A3/355	6	4
3600M3/479	7	3

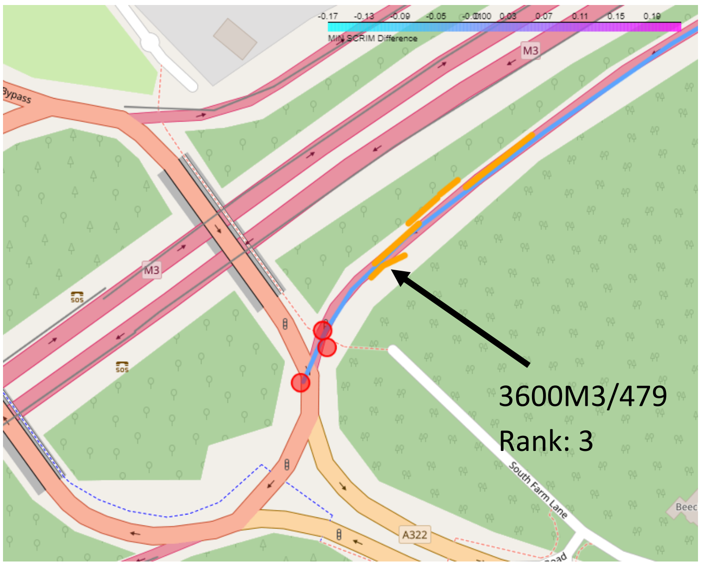


a)



Deceleration event
STATS19 collision

b)



c)

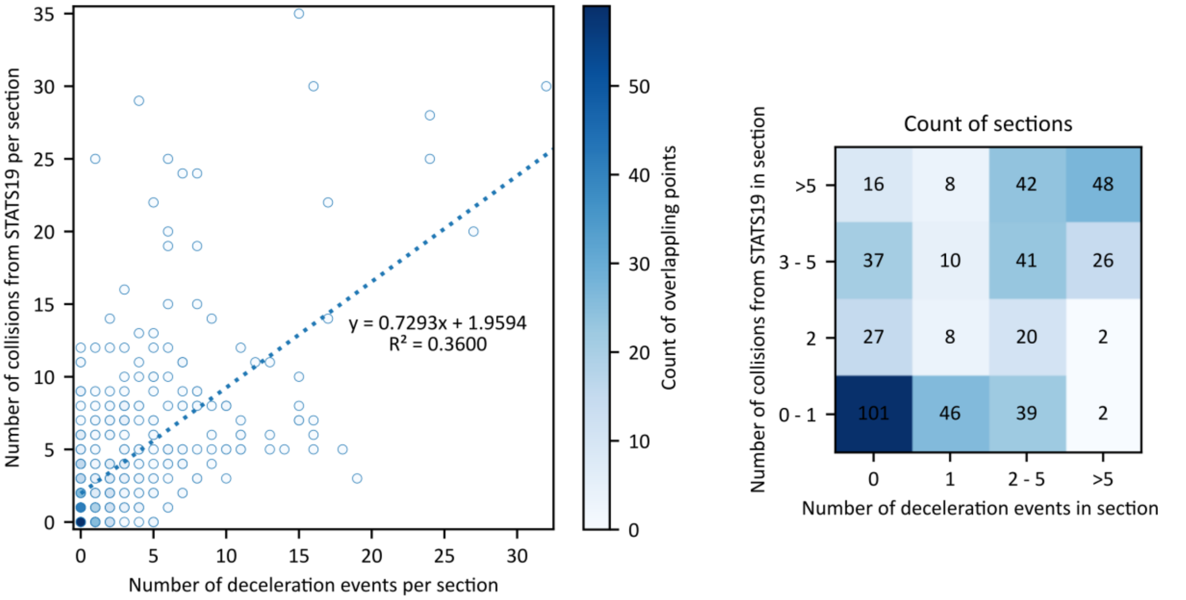
Figure 7-2: The five junction/roundabout sections with the highest number of deceleration events

7.3 Is there a relationship between mobile phone crowdsourced data and collision records?

Question: Is there potential to use mobile phone data to identify sites with higher risk of collisions occurring?

Because collisions occur rarely, using collision data to assess risk may delay the identification of higher risk sites. The higher density of events provided by crowdsourcing (“near misses”) could provide this indication earlier. This question was also asked of the telemetry data in Section 6.4. As for the above, this investigation has related crowdsourced events – in this case the mobile phone deceleration events – to STATS19 collisions.

Figure 7-3 plots the count of collisions reported in the STATS19 data against the number of deceleration events reported in the mobile phone data per section. These plots are the same type as in the analysis of Slip events in Section 6.4 where there is a longer description of the plots. Note that the numbers of events (STATS19 or deceleration) are reports as whole numbers, and therefore there are several overlapping points in the left plot. The colour bar shows the count of these overlapping points.



Scatter plot and liner regression line of best fit

Histogram with bins based on approximate quartiles

Figure 7-3: Plots of the total number of collisions from STATS19 from 2018 to 2023 and deceleration events per section

The right plot in Figure 7-3 shows the data as a histogram with uneven bins, based approximately on the quartiles of each distribution. If there was a strong correlation between the number of STATS19 collisions and number of deceleration events, it would be expected that sections with the lowest quartile of number of STATS19 events would also be sections with the lowest quartile of number of deceleration events, and that this pattern would continue for the other quartiles. We can therefore observe that:

- Comparison of the mobile phone deceleration event data shown in Figure 7-3 with the similar plot presented for telemetry data in Figure 6-14 shows that there were considerably fewer P-AMS sections covered by the mobile phone data in Region 2, but the number of events is much higher. The balance between STATS19 collisions and deceleration events is smaller, with 1,888 STATS19 events and 1,318 deceleration events. For this analysis, the coverage of deceleration events was sufficiently high that sections with zero deceleration events could be included (unlike the analysis of slip events).
- There is a visible relationship between deceleration events and collisions. The linear regression has a coefficient of determination, R^2, of 0.36, showing a moderate relationship. This suggests that the risk of collisions occurring over a period (as covered by the STATS19 data) might be estimated using the number of deceleration events collected over a mobile phone reporting period. It is anticipated that the level of confidence would probably increase if a stronger relationship and coefficient of determination was established using data mobile phone data collected over a longer time period.
- An alternative other approach to identify sections where collisions are likely is to consider, for example, the quartiles (right plot in Figure 7-3). Sections with a number of deceleration events greater than the upper quartile are more likely to represent sections with above the median number of STATS19 collisions.
- Although other information from STATS19 was considered in the analysis (e.g. breaking down collisions by severity and if the vehicle skidded), the strongest relationship with mobile phone braking events was seen for the overall comparison shown in Figure 7-3.

7.4 Influence of traffic

Question: Traffic flow and density affect collision risk. Does the crowdsourced data provide any insight?

This section considers the relationships between traffic density and deceleration events reported by the smartphone data in Region 2. In an initial comparison, a dataset of the count of deceleration events per P-AMS section was compared with the averaged traffic volume for that section. The comparison showed little relationship. The strength of the relationship was not strongly improved following normalisation by section length, or filtering to only include mainline sections on the M25 sections in Region 2.

However, the strength of the relationship does improve when speed is included. The dataset was separated into two sets: the first included vehicles that were travelling at less than 45mph when beginning braking; the second included vehicles travelling at greater than 45mph when beginning braking. These are shown in Figure 7-4. There is evidence of a relationship between deceleration frequency and traffic volume when decelerating from lower initial speeds. As the speed limits are typically higher than 45mph on the M25 when the traffic is free flowing, those travelling at speeds less than 45mph are likely to be experiencing congestion.

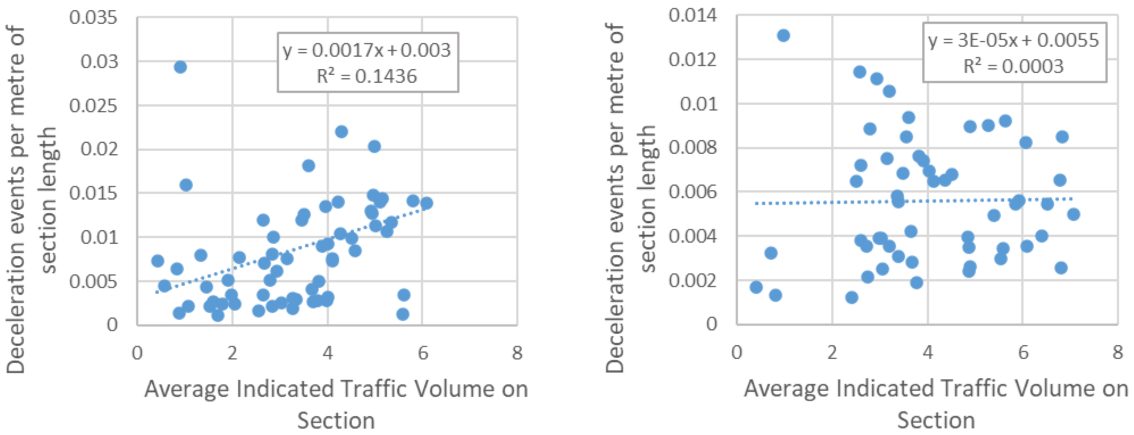


Figure 7-4: Relationship between deceleration event counts and indicated traffic volume on the M25 (normalised for length). Where the vehicle was travelling at less than 45mph (left), and where the vehicle was travelling at greater than 45mph (right)

Given the positive relationship between deceleration events and traffic collisions, there is likely to be an increase in the collision risk at the locations where the crowdsourced data has reported deceleration events. Whilst the discussion in the previous section focussed on the use of the deceleration data as a tool to identify sections at higher risk of collision (primarily to consider long term / planned interventions for skid resistance), the relationship with traffic and speed observed here shows potential application to support ongoing tracking of risk when congestion occurs. For example, by tracking speed, deceleration events and congestion at closures an insight could be provided on any requirement for improved skid resistance on the approach to the closures, or closures could be planned further from the start of the roadworks to avoid regions of poor skid resistance where congestion is likely to occur.

8 Coverage of crowdsourced data

8.1 Coverage requirements for network level assessment of skid resistance

Question: What timescales would be required for the data sources used in this work to provide skid resistance data coverage over the SRN?

As noted in Section 2, current methods to assess pavement condition typically cover a high percentage of the SRN (e.g. all of lane 1, or all mainline lanes) to ensure that the network is adequately covered in terms of maintenance needs and to report network level condition indicators. To replace such a dataset, or to provide complementary data that can be used alongside the whole dataset, will require a comparable level of coverage.

When considering safety (the main focus of the data provided to this work), the crowdsourced measure most closely linked to skid resistance measurements would be the slip event. However, this is event-based, which means that data it is more likely to be provided when an event, such as heavy braking, causes the system to record a slip. These events are infrequent and therefore it is likely that it would require considerable time to collect a crowdsourced dataset that could be considered to have “tested” the whole network.

Therefore, a Monte Carlo braking simulation model was developed to explore the time needed to cover the network. The smartphone data was used as the basis for this. It is noted that the smartphone deceleration data does not indicate skid resistance. However, the smartphone data provided higher numbers of events than the telemetry data, hence providing an estimate of the “minimum” level of time that would be required.

8.1.1 Model

Sections of road were selected to be representative of different traffic regimes:

- M25 between junction 8 and 9 as a high traffic road
- A3 westbound at West Clandon, as a lower traffic road

The frequency of deceleration on these roads was calculated and found to be approximately uniformly distributed, allowing the model to use an exponential distribution to model both the time and location of the initiation of a braking event. The distribution of distances covered by braking events was assumed to be normal, as previous sections of this report have found.

The model randomly simulates braking events over each length of pavement, identifying where the pavement has experienced at least one deceleration event, with simulation continuing until a given percentage of coverage has been reached. To prevent boundary issues, braking events passing over the end of the simulated pavement continue at the start of the pavement. For each configuration, the model was simulated 10,000 times to reach convergence of the result, shown in Figure 8-1.

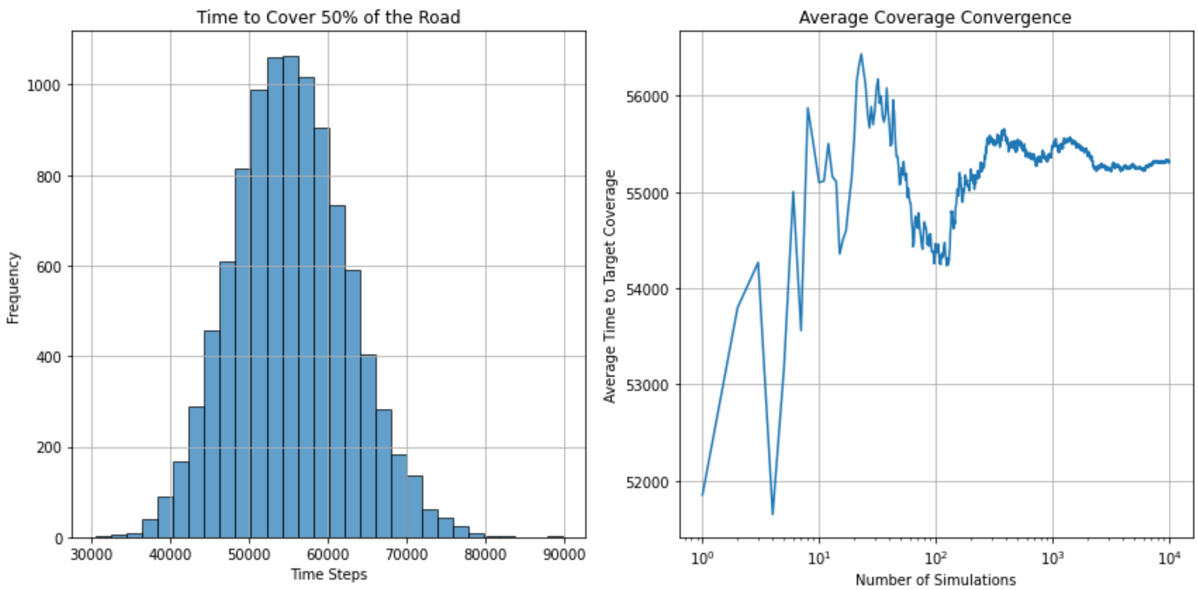


Figure 8-1: Distribution of samples from Monte Carlo model (left), convergence of Monte Carlo model to desired level of accuracy within 10,000 samples (right)

8.1.2 M25 between Junction 8 & 9

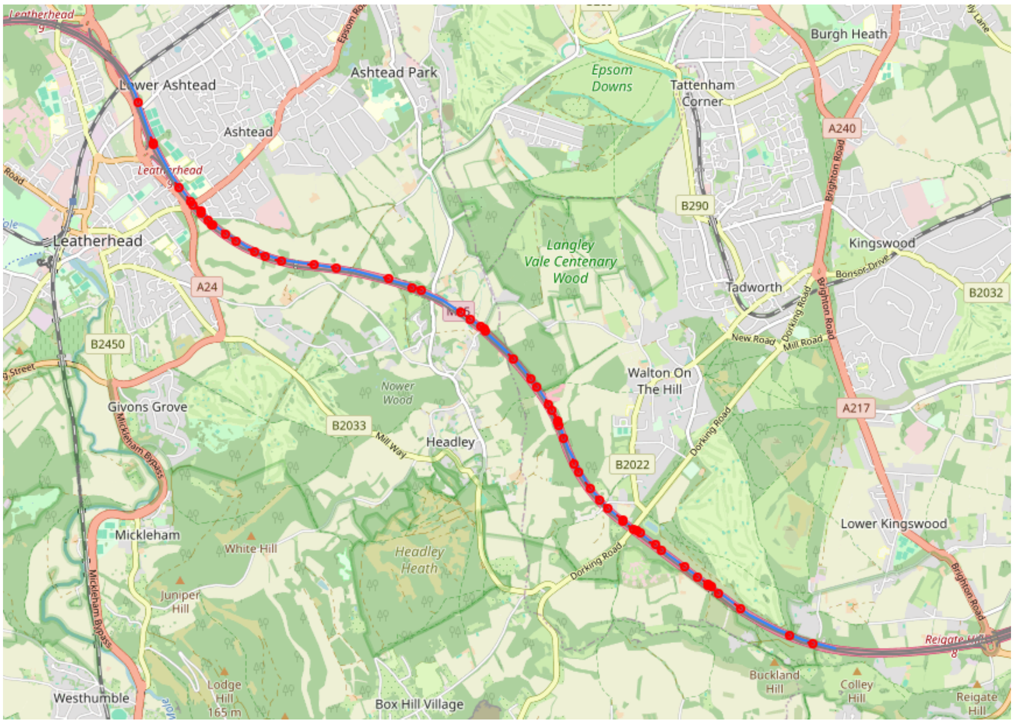


Figure 8-2: Deceleration events on the M25, anticlockwise between junctions 10 and 9

The deceleration data summarised by P-AMS section in Table 7 was used to calibrate the Monte Carlo model to simulate the collection of crowdsourced deceleration data over the

M25. These sections have a total of 57 events, occurring over 7 days, over 4 lanes and 9582m of road. This results in 0.0057 deceleration events per minute, ignoring changes in traffic due to time of day. These 57 deceleration events have a mean travel distance of 88m, with a standard deviation of 31m. To fit the assumptions of the Monte Carlo model, these deceleration events must be uniformly distributed across the network. A Chi squared test was conducted where the count of deceleration events was summed over lengths of around 1.5km and around 1km (6 and 10 bins), as shown in Figure 8-3, and the Chi Squared statistic calculated from the difference between the actual and expected count in each bin.

Table 7: Deceleration event data used to calibrate model to represent the M25

Section	Deceleration Events	Section length
3600M25/439	3	877
3600M25/443	8	970
3600M25/447	8	1999
3600M25/455	11	2016
3600M25/463	17	1994
3600M25/471	10	1996

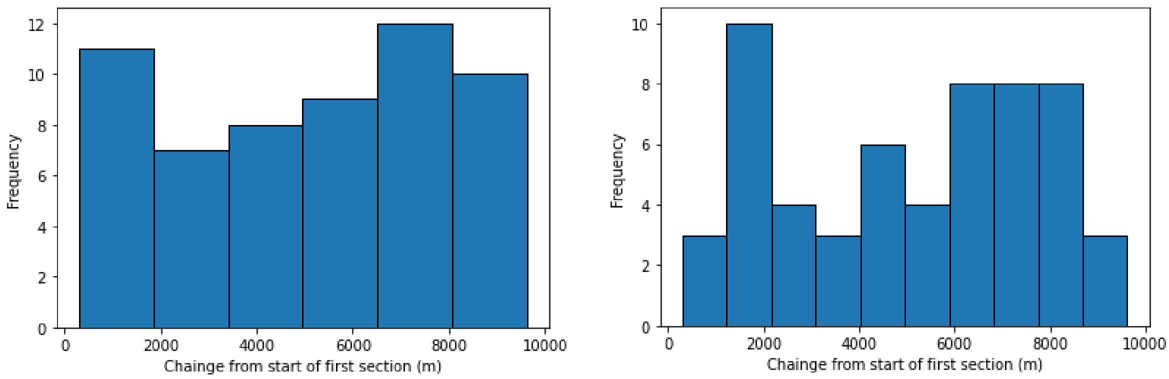


Figure 8-3: Histogram showing braking event frequency along a segment of the M25 using six bins (left) and ten bins (right)

Using 6 bins and performing the Chi Squared test, the probability of the events being uniformly distributed was found to be 0.87. However, this is for approximately one bin per 2000m where ideally the length of each bin would be small enough to capture the events of individual road features such as junctions or signals. Ideally, this would be as small as 200m. However, repeating the Chi Squared test on the ten bins of approximately 1000m revealed a probability of uniform distribution of only 0.28. This poor result may be caused by insufficient data leading to random artifacts in the distribution. Whilst the evidence is inconclusive, the spatial distribution of deceleration events on the M25 may be uniformly distributed, but more data is required for a robust statistical test.

8.1.3 A3 Westbound

The deceleration data summarised by P-AMS section in Table 8 was used to calibrate the Monte Carlo model to simulate the collection of crowdsourced deceleration data on the A3 (Figure 8-4). In this data set there are only 16 deceleration events over 7916m and two lanes of highway, resulting in a deceleration event frequency of 0.00159, around a quarter of the M25 rate. Due to the low number of deceleration events, no statistical testing has been performed to confirm the events are uniformly distributed as this would be inclusive.

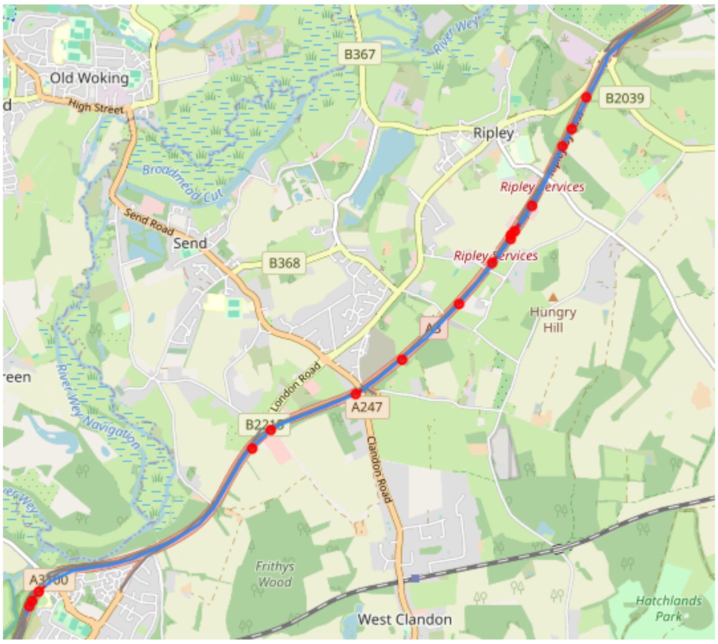


Figure 8-4: Deceleration events on the A3, southbound between Wisley and Guildford

Table 8: Deceleration event data used to calibrate model to represent the A3

Section	Deceleration Events	Section length
3600A3/325	0	325
3600A3/327	3	1218
3600A3/329	5	933
3600A3/335	2	1454
3600A3/345	0	178
3600A3/347	1	368
3600A3/351	2	900
3600A3/357	0	505
3600A3/361	0	484
3600A3/367	0	1168
3600A3/372	3	383

8.1.4 Timescales to achieve high levels of coverage

The Monte Carlo model was used to simulate the time required to reach varying levels of deceleration data coverage on the A3 and M25, as shown in Figure 8-5. The figure shows an exponential increase in the time required as the required level of coverage increases. This occurs because as the percentage of pavement covered by a deceleration event increases, the remaining uncovered pavement length decreases, making the likelihood of a deceleration event over its length reduce.

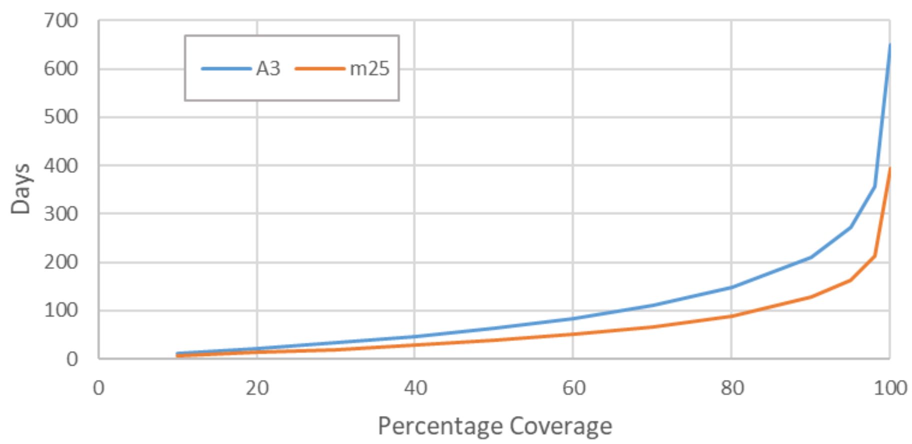


Figure 8-5: Relationship between the percentage of the carriageway covered by at least one deceleration event against the average time needed to collect the crowdsourced data

From this investigation, and Figure 8-5, we can deduce that:

- It would take on average 164 days to achieve 95% coverage on the M25, and 273 days on the A3. Based on our fleet size, we would achieve one deceleration event per length in these timescales.
- The coverage analysis has used smartphone data as the basis of volume. Because of the measurement method, telemetry slip data is more likely to be related to skid resistance. However, we have already found (Section 6.1) that slip events are only weakly (directly) linked to skid resistance. Hence it is likely that multiple slip events would be required to establish a robust assessment. Therefore, the timescales estimated from the smartphone volumes above may be significantly greater (by several times) when applying slip data.
- The direct network level assessment of the skid resistance of each length of the network is therefore unlikely to be practical using this type of crowdsourced data, without very large increases in the “size” of the crowd. Even with this increase, there would still be a need to understand the relationship between the events reported and the level of skid resistance present.
- However, this does not reduce the potential of the data as a valuable source of data for the assessment of localised risk on the network, as discussed in the following section.

8.2 Coverage requirements to identify high risk locations

Question: What timescales would be required for the data sources used in this work to identify developing high-risk clusters on the SRN?

The findings of Section 7.3, in which there was evidence for a relationship between deceleration events and collision risk, suggest that clusters of deceleration events could be used to identify high-risk sites, potentially before collisions occur. As the mobile phone data available for this work covered only one week, this also suggests that such high-risk sites could potentially be identified quite quickly. However, the timescale will vary with crowd density and risk. Therefore we have investigated how statistical analyses might be applied to determine when sufficient mobile phone data had been collected to make recommendations (potentially automatically) on whether further investigations should be made into the risks present on a particular site.

We have selected a section of the M25 between Junction 10 and 11 as a case study to discuss this analysis. During the time the mobile phone deceleration data was collected there were roadworks present in this section, shown in Figure 8-6. Analysis found the section has 14 times the average mobile phone deceleration event frequency for motorway main carriageways. This poses two questions: Is there sufficient confidence in the data to act on it; and how much data was required to achieve this confidence?

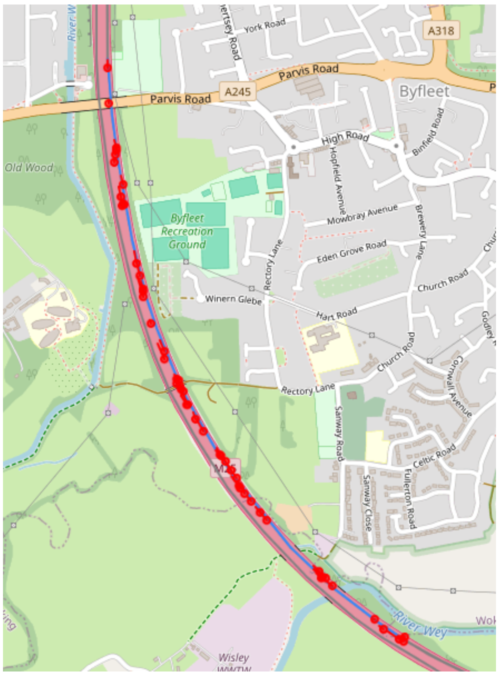


Figure 8-6: Section of the M25 between junction 10 and 11 with frequent deceleration events potentially due to roadworks and traffic management

By assuming that probability of a deceleration event is constant with time and not affected by hourly variations in traffic volume, the observed count of events may be represented by a Poisson Process where the mean is the product of the deceleration rate and time (Ross, 2014).

$$X \sim Poisson(\lambda T) \quad [1]$$

The deceleration event rate, λ, can be estimated using the sample event rate, λs, which is equal to the event count divided by the sampling duration.

$$\lambda s = \frac{n}{T} \quad [2]$$

For large values of n, λs is an estimator for λ and follows an approximately normal distribution, where:

$$E(\lambda s) = \lambda, \quad Var(\lambda s) = \frac{\lambda}{T} \quad [3]$$

$$\lambda \sim Normal\left(\lambda s, \frac{\lambda s}{T}\right) \quad [4]$$

This enables confidence intervals to be applied to the sample event rate as it converges towards the population event rate. This is shown in Figure 8-7. As the sampling time increases the sample rate changes and the 95% confidence interval narrows, giving greater confidence that the deceleration event rate obtained from the smartphone data represents the true behaviour of vehicles on this section, and is not a result of random chance.

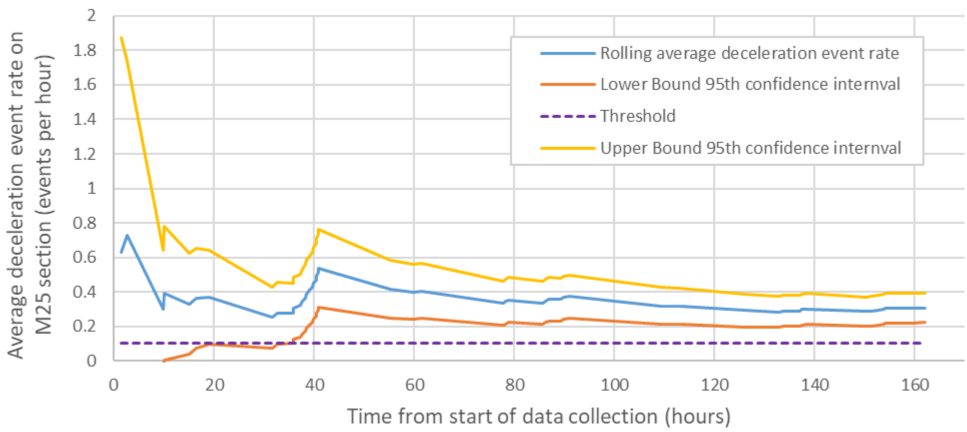


Figure 8-7: Chart showing sample rate converging towards the population rate including 95% confidence interval and arbitrary threshold value

A statistical technique such as this could be applied to identify when a length has exceeded a threshold rate of events, such that there is confidence that the site requires investigation. The threshold would be related to factors such as skid resistance site categories, traffic volumes, and relationships established between rates of events and collision risk.

Figure 8-7 suggests that an arbitrary threshold of five times the average deceleration event rate for motorway main carriageways was exceeded after approximately 36 hours of data collection on this site[7] containing roadworks. The data hence shows potential for tracking increased risks where there have been changes to road layout or management (e.g. at signals) or where there is longer term traffic management in place, to support changes to the design or control of speed limits, in addition to helping to identify sites with higher risks of collision on the overall network.

[7] After 36 hours of collecting event data there is 95% confidence that the rate of occurrence of deceleration events has exceeded the threshold

9 Discussion

The goal of this work has been to explore how the data provided from crowdsourcing could be applied to understand or manage the condition of the SRN. The analysis has centred on asking questions regarding the use of the data in specific applications. To summarise the findings we consider how well these questions have been answered.

9.1 General observations on the collection and collation of the data and its content

The data for this work were provided by two key crowdsource data providers. The telemetry data, which was accessed using an API, was straightforward to obtain and the content was simple to understand. The smartphone data was also accessible. However, the datasets are large and only georeferenced. To use the data alongside current condition datasets available on the SRN required the crowdsourced data to be referenced in the same way as current datasets – i.e. to section and chainage. Without this the analysis is limited to coordinate based (georeferencing) tools, which may not be practical, as the complexity of the network can result in data from different roads being considered to have been collected on the same road (e.g. events in slips adjacent to mainline lengths or bridges over mainline lengths being treated as the same cluster events that occurred on the mainline). This can result in incorrect deductions on road condition.

However, although care was taken in this work to develop a robust approach to fitting the crowdsourced data to the network, inaccuracies remained.

- The condition of the network varies across lanes, and the georeferencing provided in the crowdsourced data may not be sufficiently accurate to enable it to be fitted to the driven lane. As a result, this study had to collate both the crowdsourced data and the current data sources by carriageway, although it was possible to separate by direction.
- Data for which the vehicle path was provided (deceleration events) were more straightforward (and robust) to fit than point events as the vehicle path could be used to better identify the target section and chainage.
- The fitting was more complicated for events that traversed two sections, or sub-section lengths, as a decision has to be made on which section to attribute the event. This problem occurs frequently because many events such as deceleration occur at junctions, which are typically associated with the ends of sections. Fitting across sections is also made more challenging because of the lack of connectivity data in the P-AMS network definition.

Given the challenges achieving robust fitting to support comparison and combination of crowdsourced and traditional data sources for decision making, it may be that a hybrid approach to analysis would be appropriate. For example this could firstly ensure that all data is correctly allocated to the correct road (road name (e.g. A1) and type (mainline, slip, roundabout)), and then geographical analysis could be applied to cluster the data.

During analysis of the data we also noted that having an understanding of the “lack of events” could provide equal value to an understanding of the events themselves. Currently it is difficult to place the events into the context of the proportion of users for which an event occurred[8]. Reporting (in the telemetry data) the number of connected vehicles that have passed each point could provide this context. This data would need to be provided as overall contextual information via the API, as it would not be associated with individual vehicles.

9.2 Use of crowdsourced data to report functional condition?

Information relating to customer experience of road roughness was only provided in the telemetry data (bump and pothole events). There is reasonable evidence, overall, that this data indicates the presence of rougher roads and bump-like features that are also reported by current network data sourced (TRACS), indicating that the crowdsourced data may provide the ability to characterise roughness at the network level. However, we were unable to establish a relationship between the telemetry data and TRACS. Indeed, there appeared to be a negative trend between the datasets, which was the opposite of what would be expected and remains unexplained. It is possible that some of the bumps reported in the telemetry data related to significant defects that had been repaired before the TRACS survey was undertaken, but this would only explain a small proportion of the differences.

Pavement engineers responsible for maintaining the network have suggested that the frequency at which a pavement develops minor and visually apparent defects (which are identified using the frequent visual inspections undertaken by local engineers) can be a strong indicator of pavement condition and may be more informative than the network level information on functional condition provided by annual surveys. It is possible that crowdsourced data could provide a more consistent assessment of these types of defects, to augment the application of existing survey data.

However, there is mixed/uncertain evidence that the crowdsourced data can report specific bumps reported by the current TRACS survey. There appears to be a high level of inconsistency in the data. This may be because there was insufficient coverage in the dataset to enable true bumps to be filtered from false positives. However, there may also be disagreement between what the crowdsourced system data provider defines as a “bump” and what TRACS defines as a “bump”. The telemetry provider had already thresholded the data to classify bumps and therefore our ability to interpret the data was limited. A technical measure such as vertical acceleration or a parameter derived from this would be a more useful measure to assist in understanding the condition. Furthermore, if the measurement of bumpiness is being derived from sensors installed on the sprung mass of the vehicle, the measurement will be influenced by the type of vehicle. In this case additional data relating to the vehicle type or suspension characteristics may be useful.

Collaboration between the data provider and the end user (in this case National Highways) could help to refine the approach taken to reporting bumps so that the data better aligns with the requirements of the end user. For example, if the telemetry provider continues to provide ratings, then the work National Highways has already undertaken to relate TRACS

[8] It is noted that the mobile phone dataset did include a “volume” figure. However, it was not clear how this was calculated and what the number related to. Therefore, we did not include this within our analysis.

roughness measures to user experience could be used to refine the ratings/scales used in the crowdsourced dataset.

9.3 Use of telemetry data to report safety condition?

The telemetry dataset provided event data that monitored the requirement for the connected vehicles to decelerate or stop on the network ('Slip Events'). We have focussed on comparisons between the current SCRIM safety (skid resistance and supporting data such as site category) data and these events. The observation that slip events occur more frequently on sites on which drivers are more likely to have to stop is a positive indicator of the potential of this data. However, individual slip events are not well correlated with individual SCRIM measurements on any specific site. This is not surprising. We may expect that measurements would have to "build up" on individual sites to improve the relationship with the SCRIM measurements. The coverage assessment of Section 8 suggests that building up such a set of data could take a long time and a much larger fleet. However, if there is a need to ensure that every length of the network is "inspected" this suggests that the crowdsourced data would not, as it stands, be ready to replace fundamental network measurements such as SCRIM.

However, there is strong evidence to support the use of crowdsourced data in the assessment of risk. This includes deciding how to define site categories and to support decisions on whether interventions should be made on a site. The clustering of events, even in this small dataset, showed that the data could be used to provide insight in two ways. Where sites have been found to have low skid resistance, the crowdsourced data could be used to determine whether a site has a disproportionate number of vehicles "demanding" this skid resistance through heavy braking (i.e. for site investigation and decisions on treatments). However, the data also suggest that the clustering could be used to bring sites to the road operator's attention which may not have been flagged by the SCRIM data. For example, locations where slip events are occurring for unknown reasons, which would need further investigation. Both applications could improve safety.

It is interesting that the telemetry data was not correlated with collision records, given the case study / localised evidence of its ability to highlight risk. However, there were a very low number of events in the telemetry dataset. This suggests that a significantly longer collection period would be needed for this data source to provide a link to collision risk.

We understand that the telemetry slip event data is only reporting significant events – for example where there is ABS activation or activation of traction control. This approach inevitably reduces the events reported to the most severe and it is possible that not all slip events were the result of braking. It may also introduce a degree of bias to the data associated with the braking technology deployed on each of the connected vehicles. Therefore, as for the functionality data above (bump events), there may be benefit in expanding the content of this data to facilitate its wider use in condition assessment, and in particular if the data is to provide any proxy for SCRIM. The telemetry provider has detailed knowledge of the vehicle that provided the data and (we assume) the context of the event. This may include the vehicle speed, mass and braking demand (force) and perhaps the wheels/tyres in use on the vehicle and the environmental conditions[9]. This data could provide a greater insight into the slip event and the friction demand. It could also be used to better understand the bias between vehicles and to group together data provided by different fleets of telemetry data (e.g. different vehicle manufacturers).

9.4 Use of mobile phone data to report safety condition?

The mobile phone dataset provided event data that monitored the requirement for the vehicles to decelerate (deceleration events). Because the deceleration data can be related only to the demand for skid resistance requested by the vehicle, and not the supply of skid resistance (which we have assumed is more likely to be identified in the telemetry data), we have not compared the data directly with SCRIM measurements in this report. As the SCRIM site category has a closer relationship with the expected demand for skid resistance on a particular site, we therefore compared the deceleration event data with site category. The results show a reasonable relationship, confirming that the site category established for the sites used as case studies in the work were reflected by the level of braking demand recorded on the sites in the mobile phone data. Whilst the process described in CS228 (Highways England, 2021) is mature and technically robust, the initial establishment of site category and investigatory levels are based on rules and limited quantities of data regarding the experience and behaviour of vehicles on each specific site. The observations suggest that deceleration data could provide further insight to support site categorisation.

There again is strong evidence to support the use of this data in the assessment of risk. However, for the mobile phone data the relationship with accidents is stronger. This may be because of the much higher numbers of events in the mobile phone dataset. Whilst (as noted above) the telemetry Slip data is likely to trigger when there is a severe braking event (e.g. ABS activation, specific identification of slip), the mobile phone data requires a much lower "barrier" to triggering the event. It is interesting that there appears to be a link with collisions even though the barrier is lower, suggesting that these braking events are reasonable indicators of a near miss even without on-board vehicle safety systems triggering. The coverage assessments of Section 8 suggest that the larger dataset provided by the mobile phone data could provide a method to identify the development of higher risk clusters in a reasonable timescale.

[9] "Weather" data was included in the telemetry dataset which, we assume, relates to the use of (automated) windscreen wipers/rain sensors.

10 Conclusions and recommendations

It has been proposed that crowdsourcing has potential to change the way in which data is applied to manage the condition of road networks. This work has investigated two types of crowdsourced data (telemetry and mobile phone) and considered how it could be applied alongside, or to replace, conventional pavement condition data sources. In summary:

- Crowdsourced data is becoming accessible and with content that is relevant to road condition management. However, the way the data is delivered means that work is likely to be required to enable it to be applied alongside current data sources. In particular, this will need to consider how it is aligned, summarised and clustered. Road administrations typically deploy Road/Pavement Asset Management Systems to manage condition. The appropriate tools will need to be developed to facilitate access, fitting, alignment and analysis of the data. Refinements in the way crowdsourced data is delivered might help to make this more straightforward.
- Crowdsourced data shows potential to assist in reporting the functional condition of road networks. The functional condition data provided in this work focussed on roughness/bumpiness. However, the data provided was limited to a range of thresholded values. Either for this reason, or because of the short collection time-period, there was limited agreement with conventional roughness measurements. The provision of greater detail (e.g. the vertical acceleration or a derived parameter) might provide a more useful measure to assist in understanding condition. Data relating to the vehicle type/suspension characteristics may also useful.
- For safety, the crowdsourced data is unlikely to provide a direct replacement for SCRIM measurements. The correlation with SCRIM was not strong, and it would take a long time to build up a robust picture of the network. However, there is evidence to support the use of the data in the assessment of risk. This includes deciding how to define site categories, and to support decisions on whether interventions should be made. Where sites have been found to have low skid resistance, the crowdsourced data could be used to determine whether a site has a disproportionate number of vehicles “demanding” this skid resistance, and it could be also used to bring sites to the road operator’s attention which may not have been flagged by the SCRIM data. As for the functionality data, the provision of greater levels of detail in the data may help provide a more useful measure to assist in understanding the condition.
- Although there is strong potential for the application of this data to manage safety condition, crowdsourced data may not provide a comprehensive understanding of risk. There may be sites on which collisions are likely (or have occurred), for which low numbers of events have been recorded. The risks presented by these sites will be misrepresented in the crowdsourced dataset. Future studies into the application of crowdsourced data should also consider these types of site.
- Given the potential for this data, collaboration between data providers and end users (e.g. asset management system providers or road administrations) could help to refine the approach taken to reporting the data, to ensure that the content aligns with the requirements of the user.

11 References

Ester M, Kriegel HP, Sander J and Xu X (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96).*

Highways England (2021). *CS 228 Skidding resistance*. Highways England.

National Highways (2022). *CS 230 Pavement maintenance assessment procedure*. National Highways.

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, et al. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research, 12*, 2825-2830.

Ross SM (2014). *Introduction to Probability Models (11th ed.).*, Academic Press.

SciKit - Learn (2024). *Density Estimation.* Available from World Wide Web: https://scikit-learn.org/stable/modules/density.html.

The application of crowdsourced data in road condition assessment

In recent years a new source of data has become available that is referred to as "crowdsourced data". This can be provided by vehicles or mobile phones, on a network-wide basis, from multiple sources. Crowdsourcing has the potential to change the way in which data is used to manage the condition of the SRN. The study presented in this report investigates whether and how crowdsourced data has the potential to be applied in the management of the condition of SRN.

The study obtains example data from two crowdsource data types – vehicle telemetry data and mobile phone data. A focussed, case study-based investigation is carried out to determine if, and to what extent, the crowdsourced data types could be applied to understand road condition and how it complements the assessment of the functional and safety level of service measured by current data collection methods.

It is concluded that crowdsourced data provides content that is relevant to road condition management. However, tools will need to be developed to facilitate access, fitting, alignment and analysis of the data. Refinements in the way crowdsourced data is delivered might help to make this more straightforward. The data itself shows potential to assist in reporting the functional condition of road networks. However, providing a greater level of detail in the data might provide a more useful measure to assist in understanding condition. The data can also assist in assessing the safety condition of road networks, in particular in the assessment of risk. This includes deciding how to define site categories, and to support decisions on whether interventions should be made.

Other titles from this subject area

ACA108	The application of artificial intelligence to road pavement maintenance assessment. M Courtier, R Lee, P Sanders and R Dhammi. 2024
PPR2042	Road Condition Monitoring Data, Network Study. A Wright and S Brittain. 2024
PPR863	Applications of machine learning in transport. K Nesnas, R Khatry, A Smirnov, D Peeling, S Mistry, M Crabtree and S Reeves. 2018
ACA105	Motorcade – analysis toolkit for monitoring trials of zero emission vehicles. R Dhammi, M Jones, S. Varadarajan, C. Baverstock, S. Chege and J. Zihni. 2023

TRL
Crowthorne House, Nine Mile Ride,
Wokingham, Berkshire, RG40 3GA,
United Kingdom
T: +44 (0) 1344 773131
F: +44 (0) 1344 770356
E: enquiries@trl.co.uk
W: www.trl.co.uk

ISSN 2514-9695

DOI 10.58446/stvr8228

ACA109