CEDR Transnational Road Research Programme
Call 2013: Aging Infrastructure Management

funded by Denmark, Germany, Ireland, Netherlands, UK and Slovenia

CEDR
Conférence Européenne des Directeurs des Routes
Conference of European Directors of Roads

HiSPEQ: Hi-speed survey Specifications, Explanation and Quality

Identifying the key requirements for Accreditation and Quality Assurance – report for consultation

Report No: WP1-M1
January, 2015

HiSPEQ

CEDR Call 2013: Aging Infrastructure Management
HiSPEQ
Hi-speed survey Specification, Explanation and Quality

Identifying the key requirements for Accreditation and Quality Assurance – report for consultation

Due date of deliverable: Not applicable – interim report for consultation

Start date of project: 14/04/2014 End date of project: 14/04/2016

Author(s) this deliverable:
E Benbow, A Wright (TRL)

PEB Project Manager: Rolf Rabe

Version: Final

Table of contents

1 Introduction and purpose of this document

Road administrations rely on high quality condition data to understand the condition of the asset and plan and undertake maintenance programmes on their networks. High speed surveys have become a key source of this information, providing data on the shape and condition of the road surface and, in recent years, the structural robustness and the structure of the pavement itself. These high speed systems bring the advantage of network wide data collection without interfering with the traffic flow. They can provide coverage of the network which would be impractical for traditional surveys to achieve. They have lower survey costs per km than slow speed surveys and bring data that does not suffer from the subjectivity or inaccuracy of manual surveys. The data can be provided in a very structured manner (for example condition parameters reported every 10m accurately located relative to section, distance and geographical position) and can be easily fed into pavement or asset management systems.

High speed surveys therefore bring significant practical advantages to condition assessment, to support robust asset management. However, previous research (the HeRoad project) found a wide range of policies across countries to define the requirements for the survey equipment, the survey frequencies and the data delivered. Each country appears to adopt its own requirements, each subtly different from one another. This is perhaps unexpected, given that the equipment used to collect this data within different countries is likely to be quite similar. A factor that contributes to this situation is the lack of standardisation for many of the measurements, and where standardisation does exist (e.g. for profile) it is limited in its practicality and may be too complex for road administrations to understand. Hence there is a need for information to be made available to road Authorities to assist them in confidently defining the requirements for their surveys, to help them have confidence in selecting equipment and to help them ensure that the data that is provided is accurate and fit for purpose.

The objective of the HiSPEQ project is to develop guidance, advice and templates that can be used by road Authorities to help them understand high speed road survey equipment, to help them specify survey requirements and quality regimes, and to help obtain good value from the data delivered.

To make sure that the project scope is practical, HiSPEQ is concentrating on the aspects of high-speed survey data collection that help in the assessment of pavement structural robustness. This does not mean that the survey methods considered are only those that measure structural robustness directly (e.g. pavement deflection), but the measurements should be related to structural condition. For example cracking is measured under surface condition surveys, but the defect arises from deterioration in the structural condition of the pavement. As a result, measurements relating to friction, e.g. skid resistance, surface texture, are excluded from HiSPEQ, and we have not discussed devices whose only purpose is to measure texture.

The approach taken has been to combine the technical expertise of the project consortium with a review of previous research and a review of many existing survey specifications to propose the key requirements that should be considered by a road administration when developing a specification for high-speed condition surveys of their network. In particular **this document concentrates on the Accreditation and Quality Assurance regimes that could be applied to ensure the quality of the data delivered (carried out under Work**

Package 3 of HiSPEQ). The summary findings and recommendations for the implementation of Accreditation and Quality Assurance regimes are presented within the next section (2).

It is the intention of HiSPEQ, within the next phase, to formalise the Accreditation and Quality Assurance requirements proposed in this document into a set of templates that could be used by road administrations to support the development of survey specifications. The advantage of this would be that more commonality will be achieved across Europe in the definition of surveys, and in the accuracy requirements for the data, enabling improved consistency between the measurements collected and also ensuring data quality throughout the duration of a survey contract, and between different contracts. This will also allow survey providers to develop equipment that could be more easily adapted to carrying out surveys in different countries.

Whilst developing these specifications and templates HiSPEQ wishes to ensure that the proposals for the key survey requirements are aligned with the experience and expectations of stakeholders. Therefore we are issuing this report to stakeholders to invite views on the recommendations that have been made. The project team welcomes comment and views from stakeholders, which will be taken into consideration when confirming the requirements summarised in section 2, and in the subsequent development of survey specification templates*.*

As a guide to this document, it contains the following key sections:
1 Introduction and purpose of this document: This introduction section
2 Summary recommendations for Accreditation and Quality Assurance Regimes for high-speed surveys: Here we present our summary recommendations for the key requirements for Accreditation and Quality Assurance regimes, for review and comment.
3 Technical background: This section presents a technical background for Accreditation and QA regimes and reviews existing specifications in this area, which we have drawn upon in developing our recommendations.
4 Definitions: A summary of the definitions of technical terms used in this document.

2 Summary recommendations for Accreditation and Quality Assurance Regimes for high-speed surveys

2.1 The need for Accreditation and Quality Assurance

Because of the complexity of collecting and delivering the survey data there can be problems obtaining accurate, high quality and consistent measurements across different survey devices and different networks. Indeed, there are many examples, from established high speed condition survey regimes, of delivered data being inconsistent between devices (either owned by the same survey contractor, or a different one), and delivered data not being accurate, despite a high equipment specification. There are also examples of the data quality deteriorating, or changing through the duration of a survey contract, due to wear of the equipment. These will be presented as case studies in a later output of the project.

The experience gained from these well-established survey regimes suggest that there is a great need for both Accreditation of survey equipment and continuing Quality Assurance, in order to obtain confidence in, and good value for money from, the data that will be delivered from network surveys.

2.2 Requirements for Accreditation

The following subsections present summary recommendations for a potential approach to the testing of raw data and derived parameter accuracy as part of a formal accreditation regime.

2.2.1 Which parameters and measurements should be tested within Accreditation?

When commissioning a network survey the quality of delivered data cannot be guaranteed and therefore needs to be tested, before surveys commence. Furthermore, as high quality measurement data has little value if it cannot be referenced back to the exact position on the road network on which it was measured there is a need for any accreditation test regime to include tests of:

- All parameters delivered by the survey, or calculated from delivered measurements, e.g. rutting, IRI, deflection slope.
- The location referencing of the data e.g. distance measurements, GPS coordinates.

The parameters calculated from data measured on the test sites may match the reference data well but may exhibit different behaviour on the road network. One way to ensure that this will not be the case is to test the raw measurement data e.g. transverse profile. Therefore, it would be highly recommended that Accreditation would also contain tests for:

- The measurements used to calculate parameters e.g. transverse profile.

2.2.2 What aspects of data quality should be tested?

The Accreditation tests should demonstrate that the survey device is capable of delivering the accurate and consistent data it needs to help the road operator obtain confidence that the data is accurate (i.e. matches a reference), and repeatable. Furthermore, where multiple devices are to be employed to measure the network there is a need to be confident that the measurements across the fleet are consistent.

Therefore, it is recommended that Accreditation assesses:

- Core/fundamental Accuracy;
- System repeatability;
- Fleet consistency, if multiple devices are employed to deliver data.

2.2.3 How should data for testing be collected?

In order to test the fundamental accuracy of data, there is a need to obtain reference data. The majority of “Golden Devices”, used for collecting such reference data, are either slow-speed, or stationary devices. Thus, if they were to be used on the road network, traffic management and potential road closures would be required. Since this has a high cost associated with it, both in terms of inconvenience to road users and safety risks for the operators, the majority of current Accreditation regimes include tests carried out on private road networks, or test tracks. This enables the data, collected by the survey device to be compared with that collected by the Golden device.

Whilst test tracks and private roads are important for the collection of data for in-depth assessment and comparison with Golden Device data, they do not often contain road construction or conditions that are representative of the network to be surveyed. Therefore, there is also a need to assess data collected on lengths of the road network, particularly to assess system repeatability and fleet consistency.

Some current specifications include static tests for the survey devices. Since the survey device will make most of its measurements when travelling at traffic speed, and not when it is stationary, it is felt that it would be more beneficial to demonstrate its capability to measure accurately when measuring at speed, rather than in conditions that would rarely be found in the survey. However, it may be useful to obtain an initial understanding of performance using static tests, if these are complemented by driven tests.

Therefore, the tests carried out for Accreditation should include:

- Measurements on a private road or test track (we recommend a minimum 1km length);
- Measurements on selected sites on the road network, which represent the range of conditions found on the network (we suggest a minimum length of 10km, but up to 100km is recommended).

2.2.4 Frequency of testing

The frequency with which a survey device is subjected to Accreditation will be associated with the length of time over which poor quality data may be delivered, as the next accreditation test is likely to identify the problem (if not already spotted in the QA regime). The survey device should be accredited before starting any survey contract, and this should be sufficient, if the survey duration is less than 12 months. However, for survey contracts with a longer duration, it would be recommended that yearly tests are implemented. More frequent tests would be beneficial if the survey is covering very long lengths (several thousand km each year). Thus, the frequency of accreditation testing should be:

- Accreditation at the start of survey contract for all survey contracts
- Re-accreditation every year of the survey contract, or more frequently if the survey length presents a risk of significant volumes of poor data being delivered.

2.2.5 Who should be responsible for checking the data?

Ideally, the contractor should not be solely responsible for checking the quality of the data, without there also being independent checks. If the survey commissioner does not have an in depth understanding of the data being delivered for the survey, then the use of an Independent Auditor is beneficial, particularly since the Auditor can help the survey commissioner gain a deeper understanding of the data, without having to have technical knowledge. The Auditor can also act as an intermediary between the survey contractor and the survey commissioner. However, an appropriate body, to provide this service, may not be available. Thus there are three levels recommended for checking the data quality:

- Highly recommended: Checks carried out by Independent Auditor;
- Recommended: Checks carried out by survey commissioner;
- Acceptable: Checks carried out by contractor, but only if overseen (e.g. by an auditor or by the survey commissioner).

2.2.6 Testing for the effects of external influences on the data

Speed is known to have an effect on some of the data delivered by high-speed surveys e.g. TSD, longitudinal profile from inertial profilers. Therefore, it would be beneficial to include a test in the Accreditation tests, where the equipment is tested over a range of speeds, and any lower or upper speed limits set for the survey, based on those results. In addition a requirement for a minimum effect of speed on the data could also be implemented.

The **Season and Temperature** could affect the data delivered, for example, temperature has a large effect on the TSD deflection measurements, and for this reason, there is a suggested temperature range for these surveys. No tests for this were identified during the review of Accreditation tests of surface condition measurement, presumably because no significant influence has been identified. Therefore, testing the influence of temperature and the season is not an essential requirement for Accreditation of all devices, but should be considered on a case by case basis.

The effect of the presence of **water** on the road surface can have very large effects on the data, particularly when measuring with lasers. Therefore, the Accreditation test should include an assessment of the conditions in which the surveys can be undertaken.

There can also be an influence on the data by the **humans** involved in the survey i.e. the driver and the operator. Poor driving line can result in biased data being delivered, as can late/erroneous button pressing by the operator. Thus, accreditation tests should check that the drivers and operators have been trained to an acceptable level and that they are able to demonstrate that they can drive/operate the device well.

To summarise, the Accreditation tests should include:

- An assessment of the effect of speed on the data;
- An assessment of the conditions in which surveying can take place;
- An assessment of the ability of the survey crew.

Additional tests could include:

- The effect of road/ambient temperature on the data – according to need;
- The effect of surveying in different seasons – according to need;
- Requirements setting the tolerance for the effect of survey speed on the data.

2.2.7 Requirements for Accreditation testing of Location Referencing

Location referencing is commonly achieved by measurement of vehicle position, using GPS data, distance measurement, or recording of the location of node points (e.g. 100m markers).

If distance is measured, we recommend that this should be tested to an accuracy of ±0.1% of the actual distance. As this could become a demanding absolute error over short lengths (e.g. 0.1% is 0.1m in 0.1km), we suggest that a minimum tolerance of 1m is defined for measured distances of <1km. Test sites should be chosen so that it can be demonstrated that Road Geometry and vehicle speed do not affect the measurement. Repeatability can be assessed in a similar manner.

If GPS data is used, the test sites chosen should include a range of lengths where there is good and/or poor GPS signal. We recommend a requirement that 95% of the measured locations lie within a horizontal distance of 2m from the actual location, with a maximum error of 10m. The measurement of altitude is often not essential for condition assessment and therefore there is scope to relax these requirements for altitude data.

Where the location of node points is measured, the requirement will vary, depending on how the node points are located. If GPS or distance are used the above requirements would be appropriate.

2.2.8 Requirements for Accreditation testing of transverse profile

Testing the quality of the raw transverse profile data during Accreditation is desirable, but it is demanding and requires access to a dedicated reference device and therefore road authorities would need to decide whether to include such tests within their accreditation programme and hence could be considered optional. Where raw tests are to be carried out, the following recommendations are made.

Before testing for accuracy, the measured data will need to be aligned transversally with the reference data: Comparison of reference data measured on the full width of the lane with measurements from only some of the lane may lead to different results if alignment is not carried out.

A suggested process for accrediting the transverse profile is: The difference between the profile heights of the measured profile and the reference profile should be calculated. 85% of the measured profile heights should lie within 0.5mm of the reference, 95% within 1mm.

The test should also include an assessment of the shapes of the profile, to ensure that the shape of the measured profile is similar to the reference. This could either be via a visual assessment of the profile, or by calculating the correlation coefficient between the two profiles. It is suggested that the requirement for this is that $r^2 \geq 0.9$.

System repeatability could also be tested in the same way. Note: If testing the accuracy of repeat survey data, then system repeatability can be implied from this. However, more stringent requirements can be applied if repeatability is directly tested. Therefore, it would be recommended that this is included in Accreditation testing.

2.2.9 Requirements for Accreditation testing of rutting parameters

Accreditation tests should assess the rut parameters. A proposed accreditation requirement is: 95% of measured rut depth values lie within ±3mm of the reference. However, before implementing such an approach and requirement, the survey commissioner should consider whether it is appropriate in relation to the rutting parameters defined for their survey and the intended use for the data.

Repeatability and fleet consistency could be tested in the same way, with tighter requirements placed on repeat data collected by the same device E.g. 95% of repeat rut depth values lie within ±2.5mm of the original data. In addition, the standard deviation of the repeat runs could be calculated, with the requirement that the standard deviation between the repeat runs be within ±3mm.

If the implementation of the rutting calculation is left to the contractor, then the initial accreditation should also include tests of this calculation e.g. transverse profiles should be chosen at random from the survey data and the contractor's rutting parameters for this transverse profile compared with rutting parameters calculated by the survey commissioner or Auditor.

We have identified that some specifications carry out specific tests to check for the effect of road markings on the measurement of rutting. It is recommended that the adverse effect of these markings be taken into account and checks made during accreditation.

2.2.10 Requirements for Accreditation testing of longitudinal profile

Testing the quality of the raw longitudinal profile data during Accreditation is desirable, but it is demanding and requires access to a dedicated reference device and therefore road authorities would need to decide whether to include such tests within their accreditation programme and hence could be considered optional. Where raw tests are to be carried out, the following recommendations are made:

- It may be necessary to align and stretch/compress the measured raw data with the reference data, depending on the accuracy of the location referencing, as local differences in alignment can cause significant problems.
- To ensure that the survey device is responding correctly to the range of wavelengths of interest for ride quality, it is recommended that the profile is filtered before assessment. For example, to attenuate longer wavelengths. Potential filter lengths are 3m, 10m, 30m.
- The differences in height between the filtered measured profile and the filtered reference profile should then be calculated and a tolerance defined for the required performance. We suggest that 95% of differences between the measured Longitudinal Profile and the Reference Profile should fall within a specified limit (e.g. ±2.0mm)
- We have observed that some specifications also define a requirement for checking the phase between the reference and measured profile.

The data from tests for the effects of speed and deceleration can be used to determine the range of speeds and accelerations for which the equipment delivers acceptable data, using these criteria.

2.2.11 Requirements for Accreditation testing of ride quality parameters

Accreditation tests should assess the ride quality parameters. Many different parameters are used across Europe, and these will require different tolerances for the accuracy. However, the same approach should be used to assess them:

- Calculate the parameter values from the measured and reference longitudinal profiles, for the reporting length required by the survey specification.
- Calculate the differences between these parameter values.
- The requirement will then be that *x*% of the differences lie within a range of ±*y*. A commonly used ride quality parameter is IRI and we suggest for IRI that 95% of the measurements should lie within 0.5% of the reference.

If the implementation of the ride quality parameter(s) calculation is left to the contractor, then the initial Accreditation should also include tests of this calculation e.g. longitudinal profiles should be chosen at random from the survey data and the contractor's ride quality parameters for this profile compared with ride quality parameters calculated by the survey commissioner or Auditor.

2.2.12 Requirements for Accreditation testing downward facing images

Images: Where the visual condition of the surface is to be derived from the downward images it is essential that these images have high quality and assessment should be included in the accreditation regime. However, quantitative testing the quality of downward facing images is quite a difficult task, and survey commissioners may have to rely on manual (subjective) checks.

It is recommended that accreditation tests include manual visual assessment of the quality of the downward facing images on selected test sites. This would provide a subjective assessment of whether the images are capturing the correct level of detail for surface deterioration assessment. The subjective assessment needs to consider the evenness of the illumination, the focus and the resolution. The use of a controlled test carried out on a test mat, which includes specific patterns, is recommended to assist in this test. Further "network" tests would be carried out via examination of the Downward Facing Images collected on the accreditation network sites. Network tests should be carried out under varying ambient conditions to check that the system is able to perform consistently under the range of conditions to be encountered during the survey.

Quantitative methods should be considered to enable more objective assessment of the image quality. There are two national surveys, to the knowledge of HiSPEQ, that undertake such assessments (UK strategic roads, DE), and the project will be investigating these further.

3D Images: Where 3D images are used, there is a need to also test the height measuring equipment. A standardised approach for this is to survey calibration surfaces with the device (described in Annex D of ISO 13473-1).

2.2.13 Requirements for Accreditation testing of surface deterioration (visual condition)

Accreditation tests should include assessment of the surface deterioration parameters delivered by a survey (for example cracking). Many different approaches are used across Europe to report the level of surface deterioration in any length, and these will require different tolerances for the accuracy. Likewise, there are different approaches used to obtain these parameters (fully automatic assessment, semi-automatic, and manual). The key objective, as for the measures such as ride quality, is to determine that these are reported accurately and that they are repeatable.

It is recommended that survey commissioners obtain a robust and comprehensive reference dataset against which to check the performance of systems, as the measurement of surface deterioration is recognised to be the most challenging (and inconsistent) component of network surveys.

- The reference test sites should be extensive (several km) and include examples of the types of surfaces over which the survey will be carried out;
- It is not recommended that the tests be restricted to a single site, especially not a closed test track;
- The reference data should kept up to date, as the surface is subject to change over relatively short lengths of time;

- The reference should be collected in a way such that the data can be practically and objectively compared with the test data. E.g. by quantifying the differences between the reference and test data, and that performance criteria (e.g. for the size of the differences) are specified.

If the system under test uses manual analysis to obtain the surface deterioration, then all the manual raters used for this should be subject to accreditation tests, to ensure consistency in the data delivered.

2.2.14 Requirements for Accreditation testing of structural condition (TSD) data

Our review identified far fewer examples of the routine application of the TSD, which is recognised to be a new type of device. Primary experience in routine application via a network survey specification has been derived from its use in the UK, which employs regular accreditation tests on the TSD that include quantitative testing of the data. However, an additional complication is that there is not yet available a formal reference that can be used to test the accuracy of TSD survey data. Therefore quantitative accreditation tests are limited to checks on the consistency of the device with itself (or another “identical” device). This is not entirely satisfactory.

In addition the TSD employs many sensors and systems that assist the slope sensors. This includes temperature checking systems and sensors to control the height of the measurement equipment above the surface of the road. These should be checked at accreditation.

Hence we recommend that:

- Accreditation tests of the slope sensors are carried out which check the raw deflection slope. We would recommend the requirement that the measured data lie within ±0.050 of the reference, with the reference data being provided either by a reference TSD, or a previous survey using the same TSD;
- Each TSD slope laser should be checked separately;
- The lasers to control the height of the slope sensors above the pavement are checked during accreditation testing to check that they are working correctly. In the UK this is achieved by treating the output of the device as a longitudinal profile sensor and applying the requirements for longitudinal profile measurements to this data;
- The temperature sensors are tested for accuracy and consistency against a suitable reference device.

Different parameters are calculated from the raw TSD data in different countries. These will require different accuracy tolerances to be applied but the approach of comparing the parameter values, for the reporting length required by the specification, can be used for all of them.

2.2.15 Requirements for Accreditation testing of GPR data

No examples of implemented Accreditation tests applied to GPR data were identified during the review. Therefore, the knowledge of experts within the project team, and their colleagues, has been used to develop the following requirements.

A GPR system consists of sensitive electronic equipment which can deteriorate over time if it is not maintained correctly; this is especially true of the antennae which, if subject to water ingress, can produce very poor quality data. Therefore it is important that the survey commissioner requires the GPR survey contractor to produce documentation to demonstrate that the equipment has been regularly maintained and tested to ensure it is working correctly.

Whilst there is no widely practiced or formally recognised procedures in place, the survey contractor should be able to show evidence that some form of regular in-house testing has been carried out. In the simplest form this could be confirmation that the time base of the system is correct and that the frequency content of the antennae is as should be expected. If the survey contractor holds ISO 9001 quality management system certificate, this should give some surety that a minimum level of equipment checking is being carried out. The testing should occur at a frequency recommended by the manufacturer of the equipment.
A major difficulty in the accreditation of GPR systems is the fact that the data analysis must be carried out using human intervention and therefore objective testing is more complicated. We suggest that the accreditation testing should try to address fundamental equipment issues and human assessment issues.

It is recommended that accreditation of the GPR is achieved by undertaking a GPR survey of a route that includes a representative range of the construction found on the network. Repeat surveys would need to be carried out in similar conditions (e.g. temperature, humidity) with the requirement for no rainfall between the runs. Cores from (at a minimum) locations where the construction changes should be provided to the contractor for calibration of velocity measurements, and for use in assessing the accuracy of layer thicknesses obtained from the analysis.

- Each type of material, used for pavement construction, has a range of likely GPR velocities associated with it. The velocities calculated from the GPR data should be assessed, to ensure that they lie within the correct ranges. The repeatability of either the interpreted velocities, or the raw time and amplitude data should be assessed to check system performance. We recommend the requirements would be that:
    - The average time and average amplitude for the repeat survey are within ±5% of the time and amplitude recorded by the original survey;
    - The average velocity values for the repeat survey are within ±5% of the average velocity recorded by the original survey.
- The accuracy of the layer thickness values would be checked. Care must be taken to ensure that the comparison of GPR and core is done at exactly the same location (both transversally and longitudinally), otherwise the measurements are likely to be inconsistent. This may require very accurate position information (e.g. GPS measurements accurate to a few cm) for the location of both the reference and calibration cores. We suggest that accreditation specifies:
    - 95% of layer thickness values (calculated over the relevant reporting length) lie within ±5% of the reference layer thicknesses, i.e. from core measurements.
- If several analysts are be used to interpret the data, then the consistency of the results of their analyses should be tested. This could be done by requiring that all analysts interpret the same data, to obtain layer thickness. A requirement would then be that 95% of values (calculated over the relevant reporting length) lie within ±5% of the average calculated thickness.
    - Note: It may be difficult to ensure that this test is carried out correctly and trust would need to be placed with the contractor, unless an independent test tool is used e.g. an online assessment tool, providing random sections of data. This assessment method is currently being considered by the EuroGPR group, as they would like to provide a means of assessment of (firstly) data collection operatives and, in the future, the data analysts.

Thus, we propose that an Accreditation test for GPR data should include:

- Calibration of the equipment, to the manufacturer's recommendations;
- Assessment of the accuracy of the interpreted data, including layer thickness and velocities;
- Assessment of the consistency of results from different analysts;
- System repeatability tests of the time/amplitude or velocity data.

2.2.16 Summary of information/requirements to be contained in the Accreditation specification

The following should always be included in a specification for Accreditation

- A description of the tests included in the Accreditation i.e. surveys of test sites, surveys of road network, number of repeat surveys required, tests for effect of speed, tests for competence of survey crew etc;
- A description of the reference devices or methods used to provide reference data, including how this data will be collected and how often updated;
- The frequency with which the test will need to be repeated;
- Who will be responsible for checking the data;
- How the data will be assessed and who will do the assessment;
- Requirements for the accuracy of any parameters delivered or calculated from delivered data.

It would be beneficial to also include the following in the specification:

- Requirements for the accuracy of the raw data collected during the survey;
- Requirements for the repeatability of the data (both parameters and measurements);
- Requirements for fleet consistency (if multiple devices are to be used).

2.3 Requirements for Quality Assurance

The following subsections present our recommendations for checking measurement and parameter accuracy throughout the duration of the survey contract. These are derived from the review of current approaches identified in our review.

2.3.1 Which parameters and measurements should be tested within QA?

As stated above, the quality of delivered data can change throughout the duration of the survey contract and therefore needs to be tested at regular intervals. As a minimum, the QA should include tests of:

- All parameters, e.g. rutting, IRI, deflection slope, delivered by the survey, or calculated from delivered measurements.
- Location referencing used e.g. distance measurements, GPS coordinates.

QA tests on raw measurement data e.g. transverse profile can often be quite onerous and therefore QA regimes may choose to omit these tests. However, the quality of images delivered can deteriorate relatively quickly. Therefore, it would be highly recommended that QA would also contain tests for image quality (e.g. focus, illumination, contrast).

2.3.2 What aspects of data quality should be tested?

The QA tests should demonstrate that the survey device is capable of delivering accurate and consistent data throughout the duration of the contract. The QA regime should:

- Include requirements for system consistency/repeatability, for data collected on a select number of sites;

- Include requirements for accuracy;
- Include requirements for fleet consistency, if multiple devices are employed to deliver data (and if this has been tested within the Accreditation test).

2.3.3 Who should be responsible for checking the data?

QA is primarily an internal operation, and it is recommended that the contractor should be made responsible for checking the data collected during the QA regime. If any of the data requirements are not met, the Auditor or survey commissioner should be informed immediately. The contractor should provide the QA data to the Auditor or survey commissioner.

2.3.4 Frequency of testing

The frequency with which a survey device is subjected to QA, will determine how long it takes to identify issues with data quality, and hence the length of time over which poor quality data might be delivered.
The survey commissioner and contractor need to consider their approach to risk and their confidence in the stability of the survey device. Appropriate periods between QA checks are daily, weekly and monthly. The regime could include checks over all of these periods, but to increasing levels of complexity.

2.3.5 General approach to QA tests

Our review of current specifications has shown that QA testing is typically carried out using an approach where data is collected on the road network, with reference data being collected either by the same device (to check consistency) or by approved high-speed survey devices (to assess accuracy). We suggest that QA should focus on the ability to ensure consistency, as accuracy is checked under the accreditation regime. Consistency can be tested by assessing data collected during repeat surveys of sites for which previous data exists. We therefore recommend that the QA approach is:

- Carry out Surveys of a number of QA reference sites soon after accreditation, with the sites being representative of the condition found on the network;
- Undertake QA test surveys of these sites, at regular intervals throughout the survey contract.

Where the QA regime is to include tests of fleet consistency, then the survey commissioner could consider a requirement for part (e.g. 5%) of the network to be surveyed by all vehicles in the fleet, and the data would then be compared.

Where the survey commissioner requires an additional accuracy test (in addition to consistency) they could employ an external survey of part (e.g. 5%) of the network surveyed by the contractor, using a reference device, and the data would then be compared.

2.3.6 Suggested technical requirements for specific parameters

To implement the QA tests the survey commissioner will need to define tolerances that must be achieved in the consistency of data provided on the QA test sites. This may depend on the complexity of the test, but in general the tests should consider:

- Location Referencing: Consistency test for measured distance (e.g. within 3m or 0.1%) ;
- Location Referencing: Consistency of measured position (GPS data) (e.g. within 4m of the actual location);
- Transverse Profile: Consistency test of rutting (e.g. 95% within ±3mm, all within ±10mm);

- Longitudinal Profile: Consistency test of ride quality parameters. Will require different tolerances to be applied depending on the parameter (e.g. IRI measurements within 10%);
- Surface Deterioration: Consistency test of surface deterioration parameters. Different approaches to reporting the level of surface deterioration are used across Europe, which will require different tolerances. However, the tests would in general compare the measured and original/reference surface deterioration, reported in the same way and for the same reporting length required by the survey specification and calculate the differences between these parameter values;
- Downward facing images: It is particularly important that image quality remains consistent due to the effect on any parameters obtained from images. However, quantitative tests are difficult to apply. We suggest that tests be considered that:
    - Check the footprint of the image;
    - Check for changes in greyscale values over time;
    - Check for changes in image contrast and focus over time.
- Structural Condition (TSD): Consistency test of deflection slope from each sensor (e.g. 95% within ±0.05, all within ±0.2);
- GPR data: Calibration tests. The equipment should be calibrated in line with the recommendations of the manufacturer i.e. it should be tested for time measurement, and anything else that the manufacturer recommends, at a frequency recommended by the manufacturer. It is possible that the equipment will only need to be calibrated once within the survey contract.
- GPR data: Consistency of GPR data. A requirement for consistency in the time and amplitude values for each reporting length (e.g. 95% of the average time and amplitude values within ±5% of the measurements made at the start of the survey). It may be that a larger tolerance would need to be considered when comparing data collected during the summer months, with data collected during the winter, as the moisture content of the pavement will cause changes in the measurements made.

2.3.7 Summary of information/requirements to be contained in the QA specification

The following should always be included in a specification for QA

- A description of the tests included in the QA regime i.e. calibration, surveys of road network sites, number of repeat surveys required, whether accuracy and fleet consistency will be tested, in addition to system consistency;
- A description of the road network sites (i.e. length, characteristics etc.) to be surveyed and the frequency with which they should be surveyed;
- If appropriate: A description of the reference devices used to provide reference data, including how often this data will be updated;
- Who will be responsible for assessing and checking the data;
- How the data will be assessed;
- Requirements for the system consistency required for all parameters delivered or calculated from delivered data;
- If appropriate: Requirements for the accuracy of the parameters, requirements for the fleet consistency.

3 Technical background

When developing a specification for survey equipment or survey data, a road Authority must consider how they will obtain confidence in the data that will ultimately be delivered. Questions that should be answered by a road Authority commissioning a survey may include:

- What are my technical requirements for the consistency/accuracy and how will I ensure that the data is provided to a suitable level?
- How will I ensure that the data remain consistent during the survey?
- How will I ensure the data, or problems within it, do not introduce changes to the way the condition of the network is reported (for example in comparison with the regime that it has replaced)?
- How will I ensure that the data continues to cover the network at the level of performance required in the specification?

By implementing an approach to data quality assurance that provides answers to these questions the road Authority can gain much greater confidence in their data and use it to greater effect in managing the asset.

3.1 *Information sources on survey data requirements*

A number of specifications containing descriptions of Quality Assurance and/or Accreditation regimes have been reviewed, as listed in Table 1. These include thirteen specifications for surface condition but just one for structural condition. The review has shown that many road Authorities are now beginning to understand the importance of ensuring the quality of their data. For example through the application of challenging accreditation regimes (e.g. those implemented in the UK, Australia, Sweden, Austria, Germany, for surface condition measurements, and in the UK for TSD surveys of structural condition). These regimes check that the equipment is fully compliant with the expectations of the survey commissioner, usually in terms of the data that is to be delivered. These accreditation regimes are typically followed up by on-going Quality Assurance (QA) checks to ensure that the equipment does not deteriorate during the survey and between accreditations.

Table 1: Overview of specifications reviewed which included QA and accreditation requirements

Specification reference	Country/Region	Road network	Specification name
Surf_AU	Australia	National and state roads	AG:AM/S001, S004, T001, T002, T003, T004, T005, T009, T010,T011, T012
Surf_AT	Austria	Motorways	RVS11.06.67, RVS11.06.68, RVS11.06.69, RVS13.01.16
Surf_CA	Canada, British Columbia	Provincial paved road network	Pavement Surface Condition Rating Manual
Surf_DE	Germany	Motorways, Primary roads (all roads maintained by state)	Technical testing for unevenness measurements on

Specification reference	Country/Region	Road network	Specification name
			road surfaces in longitudinal and transverse direction Part: Non-contact measurements
Surf_IE	Ireland	National road network, including motorways	National Roads Network Pavement Condition Survey and Associated Consultancy Services
Surf_NL1	Netherlands	National road network	Pavement surface condition specification
Surf_NL2	Netherlands	Regional road network	Schedule of Requirements for surveying the road surface properties of provincial highways
Surf_NZ	New Zealand	State highway network	High Speed Pavement Condition Surveys 2013-2020
Surf_SI	Slovenia	National and regional road networks	TSC 06.610:2003 "Road pavement surface characteristics Unevenness" TSC 06.620:2003 "Road pavement surface characteristics Skid resistance" TSC 06.630:2002 "Road pavement surface characteristics Bearing capacity"
Surf_SE	Sweden	All road network	TRVMB 150 Road Surface Variables
Surf_UK1	England	Strategic road network	TRACS
Surf_UK2	UK	Regional road network	SCANNER
Surf_US	USA, Louisiana	All paved roads	Pavement Distress Data Collection specification
Surf_US2	USA, California	All paved roads	CALTRANS Automated Pavement Survey
Struct, UK	UK	Strategic road network	TRASS

3.2 Accreditation testing of surface condition equipment

The findings of the reviews have been summarised in the following subsections, for Accreditation regimes applied to the surface condition measurements and parameters identified in Work Package 1.

3.2.1 What parameters and measurements are tested during accreditation?

Thirteen of the fourteen specifications listed in Table 1 describe Accreditation tests. The remaining specification, Surf_US2 states that any survey device will have to go through an accreditation and quality assurance process but does not describe the process. This specification has not been considered for the following analysis.

Location referencing

All thirteen specifications require the delivery of location referencing data and most (10) test the accuracy of this during the Accreditation tests.

Transverse and longitudinal profile measurements/parameters

The tests of the transverse and longitudinal profile measurements/parameters are summarised in Table 2. Just under half of the specifications test what is delivered. However, two specifications (IE and AT) do not test all measurements or parameters, whilst another (DE) tests the calculated parameters, not the raw profiles. A further specification (US) did not specify how rutting would be tested but did specify testing of longitudinal profile, which was not actually a deliverable. Additional parameters or measurements were tested for the remaining three specifications (NZ, UK1, UK2).

Images and surface deterioration

The approach to these parameters is also summarised in Table 2. For the five specifications that required delivery of images (in order to perform surface deterioration analysis), only one tested the quality of the images (UK1) during Accreditation. It is thought that this is because such tests are difficult to devise, carry out and also assess. The majority (5 out of 7) of specifications requiring delivery of a surface deterioration parameter included tests for this in the Accreditation.

Thus, the common practice appears to be that the Accreditation tests only assess the parameters or measurements that are deliverables from the survey.

Table 2: Required measurements and parameters compared to what is tested during Accreditation

	Location		Transverse Profile		Rutting		Longitudinal Profile		Ride Quality		Images		Surface defects	
Specification	**D?**	**A?**	**D?**	**A?**	**D?**	**A?**	**D?**	**A?**	**D?**	**A?**	**D?**	**A?**	**D?**	**A?**
Surf_AU	✓	✓			✓	✓			✓	✓				
Surf_AT	✓				✓				✓		✓		✓	
Surf_CA	✓				✓	✓			✓	✓				
Surf_DE	✓	✓	✓			✓	✓			✓	✓		✓	✓
Surf_IE	✓	✓	✓		✓	✓	✓		✓	✓	✓		✓	
Surf_NL1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
Surf_NL2	✓	✓			✓	✓			✓	✓			✓	✓
Surf_NZ	✓	✓		✓	✓	✓			✓	✓				

Specification	Location		Transverse Profile		Rutting		Longitudinal Profile		Ride Quality		Images		Surface defects	
	D?	A?	D?	A?	D?	A?	D?	A?	D?	A?	D?	A?	D?	A?
Surf_SI	✓								✓	✓				
Surf_SE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	(✓)			
Surf_UK1	✓	✓	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓
Surf_UK2	✓	✓		✓	✓	✓		✓	✓	✓			✓	✓
Surf_US	✓	✓			✓			✓	✓	✓	✓		✓	✓

D=Delivered, **A**=tested during Accreditation
(✓) Images are delivered but are not currently used for network assessment of surface deterioration. Quality is not tested for Accreditation but is subjectively tested during the survey contractor procurement process.

3.2.2 What type of data accuracy is tested during Accreditation?

All parameters or measurements subject to Accreditation tests are tested for their accuracy, by comparing them with reference measurements (Table 3). Some of the Accreditation tests also test the system repeatability (3/5 transverse profile tests, 7/10 of the rutting tests, all longitudinal profile tests, 9/12 ride quality parameter tests, 2/5 of the surface deterioration tests, see Table 3). This is achieved by performing multiple (repeat) surveys on the same length of road and either ensuring that all runs fall within an acceptable range of accuracy, or calculating how consistent the repeat measurements are with each other.

In addition to this, one specification (SE) also tests fleet consistency, i.e. how consistent all devices (carrying out the survey contract) are compared with each other. If all survey devices are subject to accuracy testing within the Accreditation tests, then they are likely to be consistent, within a certain tolerance of each other. However, by testing fleet consistency, more stringent requirements can be applied to the consistency of the data delivered by the fleet.

Table 3: Aspects of data accuracy tested during Accreditation

Specification	Transverse Profile		Rutting		Longitudinal Profile		Ride Quality		Surface defects	
	A?	R?	A?	R?	A?	R?	A?	R?	A?	R?
Surf_AU			✓	✓			✓	✓		
Surf_AT										
Surf_CA			✓	✓			✓	✓		
Surf_DE			✓				✓	✓	✓	
Surf_IE			✓				✓			
Surf_NL1	✓	✓†	✓	✓†	✓	✓†	✓	✓†		
Surf_NL2			✓				✓		✓	
Surf_NZ	✓		✓	✓			✓	✓		
Surf_SI							✓			
Surf_SE	✓		✓	✓	✓	✓	✓	✓		

Surf_UK1	✓	✓†	✓	✓	✓	✓†	✓	✓	✓	✓
Surf_UK2	✓	✓†	✓	✓†	✓	✓†	✓	✓†	✓	✓†
Surf_US					✓	✓	✓	✓†	✓	

A=Accuracy, **R**=System repeatability
† Repeat runs are used to check accuracy, which implies repeatability.

3.2.3 Other tests

Equipment calibration
Some of the specifications stated that equipment used should be calibrated according to the manufacturer's instructions. However, none of them described any checks to determine whether this had been carried out, nor to test the robustness of the calibration. This is likely because of the vast differences in equipment used and the need for in-depth knowledge by the specification owner to understand how to check the calibration.

Driving line
It was noted, by some members of the project team, that the driving line, taken by the survey vehicle, played an important part in the accuracy of the data collected, particularly for transverse profiles and rut depths. Whilst this could partially be implied if stringent enough accuracy requirements were placed on the distance measurement, no evidence that a direct assessment of this element of the survey was carried out in any of the Accreditation regimes reviewed.

Survey crew
Similarly, it was noted that the data collected by different survey crews (surveying the same road in the same vehicle) can vary, due to the driving style of the driver, or the approach taken by the operator. Just under half of the specifications reviewed (6) mention assessing the capability of the survey crew, but most do not state how this might be achieved. One specification (SE) uses a scaled down version of the tests for system repeatability to assess the consistency of data collected when different crews survey the same length of road with the same equipment.

Speed effects
The speed of the vehicle can affect the measurements obtained, particularly for longitudinal profile, when surveying with an inertial profiler. Of the five specifications that include accreditation for longitudinal profile, almost all (4) state that data from surveys carried out at different speeds are tested for accuracy. Of the twelve that include accreditation for a ride quality parameter, only 5 explicitly test the accuracy of data collected at different speeds. One of the specifications (UK1) tests the relationship between the ride quality parameter and speed and applies acceptance limits to this. Of the five specifications that include accreditation for transverse profile, three test the accuracy of data collected at different survey speeds. Of the ten performing accreditation on rutting measurements, only 3 specifically test data collected at different speeds.

3.2.4 How Accreditation tests are carried out

How are measurements obtained, during the Accreditation tests?
Most of the specifications do not state how the Accreditation tests are carried out (i.e. on a test track, or on the road network), nor what length is surveyed. Therefore, the expertise of the project group and colleagues has been drawn on to fill this gap.

Since the reference, or golden, devices used to collect transverse and longitudinal profile are mainly slow speed, the data for the tests on these measurements are usually collected over relatively short lengths (<1km) on test tracks, or other private roads, where the traffic can be controlled (Table 4). However, for the parameters i.e. rutting, ride quality, surface deterioration, tests are generally applied to data collected from surveys on the road network, with survey lengths varying between 30 and 3600km (Table 5). For these road network tests, most specifications state that the roads chosen are representative of those found on the network, and display a range of conditions/roughness (Table 6).

Table 4: Length and type of road used during Accreditation testing of measurements

Measurement	Specification	What is surveyed?	Length surveyed
Longitudinal profile	Surf_SE	Sections of test track	3 x 1200m
	Surf_UK1	Test track	~2km
	Surf_UK2	Test track	~2km
	Surf_US	Test track	320m
Transverse profile	Surf_NZ	Test sections	5 x ~500m
	Surf_SE	Sections of test track	3 x 100-200m
	Surf_UK1	Test track	~2km
	Surf_UK2	Test track	~2km

Table 5: Length and type of road used during Accreditation testing of parameters

Parameter	Specification	What is surveyed?	Length surveyed
Ride Quality	Surf_AU	Test sections and road network	Test sections: 5 x 500 m Road network: 32 km
	Surf_CA	Unclear	4 x 750m
	Surf_DE	Road network	30-100km
	Surf_NZ	Test sections	5 x ~500m
	Surf_SE	Road network	3600km
	Surf_UK1	Test track and Road network	Test track: 2km Road network: 75km
	Surf_UK2	Test track and Road network	Test track: 2km Road network: 55km
Rutting	Surf_AU	Test sections and road network	Test sections: 5 x 500 m Road network: 32 km
	Surf_CA	Unclear	4 x 750m
	Surf_DE	Road network	30-100km
	Surf_NZ	Test sections	5 x ~500m
	Surf_SE	Sections of test track	3 x 100-200m
	Surf_UK1	Road network	Road network: 100km
	Surf_UK2	Road network	Road network: 55km

Table 6: Road network sites

Spec'n	If surveys of the road network are included in the tests, are there checks that ensure the roads are representative of what might be found on the network?
Surf_AU	"Sections shall be selected so as to ensure that their surface characteristics (materials, texture, etc.) are representative of the road network(s) to be surveyed"
Surf_CA	"The sites exhibit a representative variety of distress types, range in pavement deterioration, surface types and operating speed with the intent to be representative of the actual survey conditions"
Surf_DE	There is a set of roads included in the test that covers all common types of pavement and pavement failures.
Surf_IE	Not specified
Surf_NL1	Not specified
Surf_NL2	Not specified
Surf_NZ	The validation sites are required to be generally representative of the pavements to be surveyed. Each site must have texture within a set range; have different surfacing, at least one site with curves of <200m radius and a range of gradients.
Surf_SE	There are four routes measured, A - high volume roads, B - middle class volume roads, C - low volume roads, D - a mix of ABC.
Surf_UK1	The Auditor chooses these sites, based on knowledge of the roads and also of the network
Surf_UK2	"The sites include road types that are typical of the local road network in terms of construction, condition and traffic levels. The test sites may include: • Flexible and rigid constructions; • Urban and rural roads; • Single and dual carriageway roads; • Traffic light controlled junctions; • Slip roads; • Roundabouts and • A wide range of typical road geometries."
Surf_US	Not specified

Seasonal effects

Some survey devices are affected by temperature and the season of measurement, for example the TSD and SCRIM (friction measurement device). The review found that none of the Accreditation tests incorporated tests of the effect of temperature or season on the data collected and most Accreditation testing was not carried out over a long enough length of time or at a frequently that was enough to allow this to happen.

3.2.5 Frequency of Accreditation testing

Accreditation testing frequency ranges from every 3 months (UK1) to 12 months, with the most common frequency being every 12 months (Table 7).

Table 7: Frequency of accreditation

Spec'n	Frequency of accreditation
Surf_AU	Nominally 12 months: Any data must be collected within 12 months of accreditation
Surf_CA	The Accreditation is carried out at the start of a survey contract
Surf_DE	The Accreditation lasts for 12 months

Spec'n	**Frequency of accreditation**
Surf_NZ	The Accreditation lasts for 12 months
Surf_SE	An extensive test is carried out every four/five years (when the survey contract is let), with a smaller test yearly.
Surf_UK1	Accreditation is carried out before the contractor can start survey and then 4 Re-Accreditations occur through each survey year.
Surf_UK2	Re-accreditation is required at intervals of no more than 12 months and may be required sooner if maintenance has been carried out on the device
Surf_US	At the start of every survey round

3.2.6 Testing and auditing the data

The body expected to carry out the checks on the data quality for Accreditation was stated in nine of the specifications (Table 8). Independent auditors are only used for four Accreditation regimes (DE, SE, UK1, UK2), with the rest leaving the data checking to the survey contractors. For two of the contractor-check regimes, the specifications state that the data used for the test must be provided to the survey commissioner, to enable independent checking, if required.

Table 8: Who checks the data during Accreditation?

Specification	Survey Contractor	Independent Auditor	Not specified
Surf_AU	✓*		
Surf_CA	✓		
Surf_DE		✓	
Surf_IE			✓
Surf_NL1			✓
Surf_NL2			✓
Surf_NZ	✓	✓	
Surf_SI	✓		
Surf_SE		✓	
Surf_UK1		✓	
Surf_UK2		✓	
Surf_US	✓*		

* Data is also provided to road authority, for checking.

3.2.7 Tests applied to location and distance measurements

The common requirement for the measurement of distance is for the length measured to be within ±0.1% of the actual distance (Table 9). The New Zealand (NZ) specification is much stricter than this, with a requirement of 0.05%. Two specifications (UK1, UK2) allow for a minimum error of 1m for lengths surveyed that are <1km in length. One specifies (UK1) that the measurement should not be affected by Road Geometry, vehicle speed and be unaffected by “changes” in the survey equipment (e.g. from “warming up” of tyres).

For the measurement of location, one specification (NZ) requires that tests that are carried out using routes having good and/or poor GPS signal, whilst another (UK2) states different requirements for test routes with good GPS and for those with poor GPS signal, implying that these are included in the Accreditation tests (Table 10). In practice, this is also done for the UK1 tests but this is not explicitly stated in the specification itself. Most of the specifications require that 95% of the measured locations lie within a set distance (ranging from 1m to 10m) of the actual locations, with some also specifying a requirement for 98% of the data, or a maximum allowed error.

Table 9: Tests applied to distance measurement

Spec'n	Tests and requirements for distance data
Surf_AU	Select a section of road pavement, roughly 1 km in length and measure the 'true' distance using precise geomatic (ground survey) techniques. Survey the section 5 times with the same device. All measurements are required to be within ±0.1% of the true length.
Surf_DE	Location accuracy better than 10 m, checked visually
Surf_NZ	The device surveys a specially chosen route four times. The requirement is that the measured distances are all within ±0.05% of the true distance.
Surf_SE	Distance measurements are just checked by the supplier before the tests are carried out. The supplier will take a risk of failing the tests if the distance calibration is not perfect.
Surf_UK1	Distance shall be measured with an accuracy of ±1m for elapsed distances up to 1km and ±0.1% for distances above 1km. The accuracy should be unaffected by the speed of the equipment, or by the Road Geometry and the measurement must be consistent and stable throughout any period of data collection, being unaffected by changes in the Equipment (for example resulting from "warming up" of vehicle tyres).
Surf_UK2	The UK road network is split into named sections and the requirement for distance is that 95% of the measured Section lengths fall within ±1m (or ±0.1%, whichever is larger) of the lengths measured using the Reference Method.

Table 10: Tests applied to location measurement

Spec'n	Tests and requirements for location data
Surf_IE	The minimum requirements for GPS data are: • National Grid coordinates derived from the GPS are provided over no less than 950m in any 1 km length. • National Grid Co-ordinates to be provided to a coverage requirement of at least 99% of the total length surveyed. • 95% of the measured positions in any 1 km length shall be within a horizontal error of 1m or better from the true position. • 95% of the measured positions in any 1 km length shall be within a vertical error of 2m or better from the true position. • The horizontal error between the measured and the true position never to exceed 10m. • The vertical (altitude) error between the measured and true position never to exceed 20m.
Surf_NL1	95% of the hectometre markers should be recorded within 10m (longitudinally) of their actual position
Surf_NL2	95% of GPS measurements should be within 5m of the true position, in the X and Y directions.
Surf_NZ	The contractor identifies one or two routes, having good and/or poor GPS signal. 10 surveys are carried out over these routes over 2 days, preferably at different times of the day, and 10 measurement locations tested. The surveys are carried out at 3 different speeds, representative of normal survey speed. The GPS and inertial systems are acceptable if: 1) GPS co-ords are within 1.0m horizontal and 5.0m vertical of the reference values

Spec'n	Tests and requirements for location data
	2) 98% of all differentially corrected coordinate values from repeat runs are within 1.0m horizontal and 5.0m vertical of each other 3) the positional error of the location given by the inertial navigation system operating in unaided mode (i.e. no GPS signal) for 10 minutes should be no more than 50m.
Surf_SE	The requirement for coordinates is that 98% < 10 m and 95% < 5m from true position.
Surf_UK1	National Grid Co-ordinates must be provided to an accuracy such that: • At least 90% of the measured positions lie within a horizontal error of 2m or better from the True position • At least 95% of the measured positions lie within a horizontal error of 4m or better from the True position • The horizontal error between the measured position and the True position shall never exceed 20m.
Surf_UK2	The UK road network is split into named sections and the requirement for distance is: • 95% of the measured National Grid Co-ordinates are within ±2m of the National Grid Co-ordinates measured using the Reference Method for those Sections having better than 70% availability of the signal used by the equipment for the calculation of National Grid Co-ordinates • 95% of the measured National Grid Co-ordinates are within ±10m of the National Grid Co-ordinates measured using the Reference Method for those Sections having less than 70% availability of the signal used by the equipment for the calculation of National Grid Co-ordinates. • All of the measured National Grid Co-ordinates are within ±50m of the National Grid Co-ordinates measured using the Reference Method. • 95% of the measured Altitudes are within ±5m of the Altitudes measured using the Reference Method for those Sections having better than 70% availability of the signal used by the equipment for the calculation of National Grid Co-ordinates • 95% of the measured Altitudes are within ±10m of the Altitudes measured using the Reference Method for those Sections having less than 70% availability of the signal used by the equipment for the calculation of National Grid Co-ordinates • All of the measured Altitudes are within ±50m of the Altitudes measured using the Reference Method.

3.2.8 Tests applied to Transverse Profile measurements

Test procedures

Only five of the specifications test the accuracy of transverse profile data, with only four detailing how this is achieved (Table 11). Of these, all test transverse profiles that have been collected during a traffic speed survey of a site, and then compare these with reference profiles. Two of the specifications test transverse profiles that have been collected whilst the measurement device is static. All four specifications test the transverse profiles collected at different survey speeds, enabling the effect of measurement speed to be assessed (Table 12).No tests are carried out to determine the effect of temperature, season etc. on the consistency of the data (Table 13).

Reference data

Reference data for transverse profiles is delivered by a number of different devices, most capable of measuring the profile at 100mm transverse spacing, across the whole lane width (Table 14). No details are given in the specifications as to how often the data is updated, nor how the quality of the reference data is checked but this information has been provided by the partners for 3 of the specifications. The reference data is updated either every 2 years, or every time an accreditation test occurs and the quality of the data is ensured by checking it against the data from other measurement devices (Table 14).

Length of assessment

The most common approach, in terms of the length over which transverse profiles are assessed, is to assess individual transverse profiles, with only one case of an average transverse profile, calculated over a 1m length, being used (Table 15).

Assessing system repeatability

None of the specifications explicitly test the repeatability of the transverse profile data. However, all four require that repeat surveys are performed on the test sites and these are all assessed for accuracy, thus implying a level of repeatability.

How is the data assessed?

All specifications require a percentage of the measured transverse profile points to lie within '*x*'mm of the reference profile. The actual percentage and range applied varies, with only the two UK specifications having the same requirement (Table 16). One of the specifications also calculates the correlation coefficient between the statically measured profile and the reference, to ensure that the shape of the profiles is similar, whilst another uses subjective visual assessment to check this.

Table 11: What surveys and tests are carried out for transverse profile?

Specification	What do the tests entail?
Surf_NL1	What is tested and how is not specified
Surf_NZ	Survey of test sections plus static tests
Surf_SE	Survey of a test site, then comparison of the measured transverse profile with the reference
Surf_UK1	The device surveys a test site (~2km), chosen by the Accreditation Tester and also one or more lengths laid out with features of a known Transverse Profile. The transverse profiles from these lengths are visually inspected, to ensure that they look "sensible". Static tests on artificial profiles are also carried out.
Surf_UK2	The device surveys a test site (~2km), chosen by the Accreditation Tester and also one or more lengths laid out with features of a known Transverse Profile. The transverse profiles from these lengths are visually inspected, to ensure that they look "sensible". In addition a static test is carried out on a flat surface and also artificial profiles.

Table 12: Speed of surveys

Specification	Is the vehicle tested for all operating speeds (e.g. 20, 40, 60, 80km/h)?
Surf_NL1	No details given
Surf_NZ	A minimum of 4 survey speeds are required per survey, evenly spaced from the minimum survey speed to the maximum survey speed.
Surf_SE	No details given in the specification but in practice the test track sections are measured five times in three different speeds, 30, 50 and 70 km/h.
Surf_UK1	A number of test surveys are carried out at constant survey speed, at a range of speeds, in addition to test surveys carried out under conditions of deceleration on the test site.
Surf_UK2	The test surveys are carried out at a number of different speeds (constant speed per survey) and also under conditions of deceleration.

Table 13: Timeframe of surveys and tests

Specification	Is data collected from different days/months compared or tested, or only data collected on the same day?
Surf_NL1	No details given
Surf_NZ	Not specified but appears to be all in one day.
Surf_SE	The test takes place in one week where all suppliers measure at the same occasion.
Surf_UK1 and Surf_UK2	The Accreditation test usually takes place on one day, or a number of days very close in time.

Table 14: Reference data for transverse profile

Specification	What is used for the reference (is there a "golden device")?	How often is the reference data updated?	How is the quality of the reference data confirmed?
Surf_NL1	"Reference transverse profiles are measured in accordance with Netherlands standard"	No details given	No details given
Surf_NZ	A calibrated reference device (e.g. a transverse profile logger, rod and level, etc.). The device must be static and be capable of measuring the TP at 100mm centres over each of the validation sections to a vertical accuracy of 0.5mm and a TP length of at least 3.3m.	No details given	No details given
Surf_SE	No details are given in the spec but the reference is VTIXPS (Cross Profile Scanner, a line laser device with cross slope capability)	Each time a test is carried out	Calibration and checking against VTI's RST device
Surf_UK1 and Surf_UK2	Static: Artificial Profiles Dynamic: Transverse profile measured over the road edge, by HARRIS2 profiler	No details given but in practice, this is about every 2 years.	No details given but in practice, the reference data is checked against the last set of reference data and also compared with all SCANNER and TRACS devices.

Table 15: What length is the transverse profile assessed over?

Specification	What reporting length is the measurement assessed over?
Surf_NL1	No details given
Surf_NZ	Individual transverse profiles
Surf_SE	1m (the average transverse profile for each 1m length is assessed)
Surf_UK1	0.1m (individual transverse profiles)
Surf_UK2	0.1m (individual transverse profiles)

Table 16: Data accuracy requirements

Specification	What are the requirements for the data?
Surf_NL1	95% of cross sections are within ± 1mm of the reference
Surf_NZ	**Static**: The regression analysis must give a correlation coefficient r^2 ≥0.9 when the static transverse profiles are regressed against the reference for each site and >0.925 for all sites combined. 90% of the static transverse profile heights are within 2mm of the reference transverse profile heights. **Dynamic**: 90% of the dynamic transverse profile heights are within 2mm of the reference.
Surf_SE	≥85% of transverse profile points will be ≤0.5mm different from the reference
Surf_UK1	• 95% of the differences between the measured Transverse Profile points and the Reference Profile points fall within ±1.5mm. • The visual examination of the transverse profile confirms that the Equipment is not adversely affected by the measurement of transverse profile over the road edge. • For Transverse Profile measured over a nominally flat surface, 65% of the differences between the measured Transverse Profile, and a linear fit to the measured Enhanced Transverse Profile, fall below ±0.3mm. • For Transverse Profile measured over a nominally flat surface, all of the differences between the measured Transverse Profile, and a linear fit to the measured Enhanced Transverse Profile, fall below ±1.0mm.
Surf_UK2	• 95% of the differences between the measured Transverse Profile points and the Reference Profile points fall within ±1.5mm. • The visual examination of the transverse profile confirms that the Equipment is not adversely affected by the measurement of transverse profile over the road edge.

3.2.9 Tests applied to Road Marking profile

One of the specifications required a road marking profile to be delivered alongside the transverse profiles, so that transverse profile measurements, collected on road markings, could be excluded from the rutting calculation. This road marking profile has a value of 0 (no road marking) or 1 (road marking) for each transverse profile point delivered. The accuracy of the road marking profile is tested during Accreditation by obtaining survey data from a ~2km site and assessing it over 10m lengths. A 50mm square grid is placed over the profile and grid tiles containing road markings are classified as “road marking grid tiles” and grid tiles not containing road markings are classified as “none road marking grid tiles”. This is then compared to a Reference Road Marking Grid, usually obtained by manual analysis of downward facing images, collected concurrently with the road marking profile. The requirements for the data are that:

- At least 97% of the road marking grid tiles in the Reference Road Marking Grid shall be reported as road marking grid tiles in the Measured Road Marking Grid, in either the same grid tile, or immediately adjacent grid tiles.
- Tiles reported as road marking grid tiles in the Measured Road Marking Grid, for which there is no corresponding road marking grid tile in the Reference Road Marking Grid shall be known as false positives.
- False positives lying within 200mm of a road marking in the Reference Road Marking Grid shall be ignored.
- Any false positives lying to the left of the Reference nearside road marking or to the right of the Reference offside road marking shall be ignored.
- For the remaining false positives:

- No more than 0.5% of the 10m lengths shall contain false positives covering more than 25% of their length.
- No more than 2% of the 10m lengths shall contain false positives covering more than 1% of their length.

3.2.10 Tests applied to Rutting parameters

Test procedures

Eleven specifications test the accuracy of rutting parameters (Table 17), with seven including a description of how the tests are performed. Two specifications require the rutting parameters to be assessed with static tests, as well as surveys over short (<1km) and long (>1km) sites. One only requires surveys of short sites, with another requiring both surveys of short and long sites and the remaining 3 requiring surveys of long sites only (Table 17 and Table 20). Six specifications require repeat surveys to be carried out.

No tests are carried out to determine the effect of temperature, season etc. on the consistency of the data and only three of the specifications require data to be collected at different survey speeds, enabling the effect of measurement speed to be assessed (Table 18).

Reference data

Reference data for rutting parameters is delivered by a number of different devices, with static or slow speed equipment used for the static tests and the short site surveys. Independent profilers are often used to provide data for the longer network surveys (Table 19). Only two specifications state how often the reference data is updated, with none describing how the reference data is itself checked but this information has been provided by the partners for 3 of the specifications. The reference data is updated either every year, or every time an accreditation test occurs and the quality of the data is ensured by checking it against the data from other measurement devices (Table 19).

Length of assessment

The most common approach is to assess the rutting parameters over the reporting length used, with only two exceptions to this (Table 20).

Assessing data accuracy

Most of the specifications require a percentage of the measured rutting parameters to lie within xmm of the reference parameter values (Table 21). The actual percentage and range applied varies, but a common approach is to require 95-100% of the values to lie within 3mm. Two of the specifications (AU and NZ) apply a more statistical approach to assessment.

Assessing system repeatability

Six of the eight specifications require that repeat surveys are carried out on the network routes surveyed. Of these, five test system repeatability; with the sixth implying the repeatability from testing the accuracy of repeat surveys (Table 22). Different approaches are used to test the repeatability, with two specifications using statistical analyses, one calculating the standard deviation of the repeat runs, and two calculating the error between repeat runs (Table 22).

Table 17: Testing rutting parameters

Specification	What do the tests entail?
Surf_AU	There are two test methods implemented Method 1 - survey of test sections plus static tests Method 2 -survey of 'Roughness Calibration Loop' (including 5 repeat surveys).
Surf_CA	Four test sites are used, each 500 metres in length with a 250 metre lead-in. The sites exhibit a representative variety of distress types, range in pavement deterioration, surface types and operating speed with the intent to be representative of the actual survey conditions. If possible, the sites are also located within close proximity to one another to enable the automated surveys to be conducted in sequence (i.e. site 1, site 2, site 3, site 1, site 2, etc.). They are surveyed five times each
Surf_DE	The device surveys a number of road sections
Surf_IE	The specification does not say how the rutting parameters should be assessed, however, the device must have a SCANNER certificate (and so will have met the requirements for the UK2 specification)
Surf_NL1	No details are given
Surf_NL2	No details are given
Surf_NZ	Survey of test sections plus static tests. The test sections are surveyed 5 times.
Surf_SE	Survey of a test track, then comparison of four rutting parameters with those calculated from the reference data. Repeat surveys are performed on the site.
Surf_UK1	The device surveys a number of network routes that include a range of rut depths (currently 4 sites covering ~100km in total). Repeat surveys are performed on these sites.
Surf_UK2	The device surveys a number of network routes that include a range of rut depths (currently 2 sites covering ~55km in total). Repeat surveys are performed on these sites.
Surf_US	The specification does not state how testing should take place, nor how accurate/repeatable the data should be. It just says that the contractor must provide the precision and bias of the equipment.

Table 18: Conditions for data collection

Specification	Is data collected from different days/months compared or tested, or only data collected on the same day?	Is the vehicle tested for all operating speeds (e.g. 20, 40, 60, 80km/h)?
Surf_AU	No details given but the tests appear to be all in one day.	A minimum of 3 survey speeds are required for Method 1: one near the bottom, mid-range and top of the profilometer's specified operating speed range (nominally 100km/h). Survey speed not specified for Method 2.
Surf_CA	No details given	Not specified
Surf_DE	This is dependent on the number of sections to be surveyed. However, the whole test takes 2 or 3 days.	No, just traffic speed
Surf_IE	No details given	No details given
Surf_NL1	No details given	No details given

Surf_NL2	No details given	No details given
Surf_NZ	No details given but the tests appear to be all in one day.	A minimum of 4 survey speeds are required per survey, evenly spaced from the minimum survey speed to the maximum survey speed.
Surf_SE	No details given but the test takes place in one week where all suppliers measure at the same occasion.	Not specified but the test track sections are measured five times at three different speeds, 30, 50 and 70 km/h.
Surf_UK1	No details given but the data for the Accreditation test are usually collected within 1 month	The network routes are surveyed at normal traffic speed.
Surf_UK2	No details given but the data for the Accreditation test are usually within 1 month	The network routes are surveyed at normal traffic speed.

Table 19: Reference data used for rutting parameter accreditation

Specification	What is used for the reference (is there a "golden device")?	How often is the reference data updated?	How is the quality of the reference data confirmed?
Surf_AU	Method 1 - A calibrated reference device (e.g. a transverse profile logger, rod and level, etc.). Method 2 -five repeat runs by an independent laser profilometer.	No details given	No details given
Surf_CA	Use of a straight edge and wedge, at 10m intervals along the survey site.	Every time a round of surveys is carried out	No details given
Surf_DE	Golden device from BASt	Once a year	Golden device is considered true
Surf_IE	No details given	No details given	No details given
Surf_NL1	The reference data is provided by a class 1T31111 instrument according to EN 13036:6*	No details given	No details given
Surf_NL2	No details given	No details given	No details given
Surf_NZ	2.0m straight edge and wedge and/or the reference device used for measurement of transverse profile, if the wheel path rut depths obtained from this device match that of a straight edge.	No details given	No details given
Surf_SE	The reference is VTIXPS (Cross Profile Scanner, a line laser device with cross slope capability)	No details given but each time a test is carried out	Calibration and checking against VTI's RST device
Surf_UK1 and Surf_UK2	An independent reference profilometer.	No details given but in practice, this is yearly.	The reference equipment is accredited and the quality of the data confirmed by checking it against the last set of reference data and also comparing it with all UK survey devices (this is also not stated in the spec).

* i.e.: travelled distance accuracy ≤0.05%, transversal resolution >0.5mm but ≤1.5mm, acquisition sampling interval ≤75mm, acquisition repetition interval ≤1m, reporting repetition interval ≤5m, transverse gradient measurement accuracy ≤±0.15%

Table 20: Length of survey sites and parameter assessment length

Specification	If surveys are carried out as part of the tests, what distance is surveyed?	What reporting length is the parameter assessed over?
Surf_AU	Method 1 - five test sections of road pavement, each 500 m long Method 2 - 32 km long 'Roughness Calibration Loop'	Method 1 - 1m Method 2 - 50m
Surf_CA	4 sites of 750m (incl. 250m run-in)	The data is provided at 50m spacing and this is averaged to provide data for 500m (i.e. for a whole site)
Surf_DE	30 - 100 km	100m
Surf_IE	No details given	No details given
Surf_NL1	No details given	100m
Surf_NL2	No details given	100m
Surf_NZ	Five test sections of road pavement, each 500m long	10m
Surf_SE	3 sections on the test track, with a length of between 100 and 200m	20m
Surf_UK1	The network routes total about 100km	10m
Surf_UK2	The network routes total about 55km	50m

Table 21: Accuracy requirements for rutting parameters

Spec'n	What are the accuracy requirements for the data?
Surf_AU	**Method 1, Static**: Each of the measured rut depths must be within ±1 mm or ±10%, whichever is the greater, of the reference rut depths. **Method 1 Dynamic**: For each of the three speeds, group the rut depth data measured by the multi-laser profilometer into a single set of data, totalling 500 records. Using least squares regression, a line of best fit between the two sets of results should be identified for each speed in the form: RutDepthBase = A.RutDepthProfilometer+B where RutDepthBase =Rut depth calculated from the base reference measurements, RutDepthProfilometer =Rut depth calculated from the operational laser profilometer, A =regression equation slope B =regression equation intercept The coefficient of determination, r^2, for the regression must also be determined. The requirements are 0.90 ≤ A ≤ 1.10 –2.5 ≤ B ≤ 2.5 mm r2 ≥ 0.900 for an automated device or 0.85 ≤ A ≤ 1.15 –3.0 ≤ B ≤ 3.0 mm r2 ≥ 0.80 for manual reference methods (e.g. straight edge) Do the same for all of the Rut Depth data measured by the multi-laser profilometer (1500 records). The requirements are: 0.925 ≤ A ≤ 1.075 –2.0 ≤ B ≤ 2.0 mm r2 ≥ 0.925 for automated reference device or 0.90 ≤ A ≤ 1.10 –2.5 ≤ B ≤ 2.5 mm r2 ≥ 0.85 for manual reference methods. **Method 2**: Calculate the average rut depth for each 100m from each run. Using least squares regression, a line of best fit between the average 100 m results from the measured Rut Depths and the reference data set should be determined, as well as the coefficient of determination of the line (r^2). The requirement is that r^2≥0.90. The average percentage difference is determined, which is the overall average of the percentage differences for each 100 m section between the average of the five runs of the profilometer being tested and the corresponding reference data. The requirement is that average percentage difference ≤ ±10%

Spec'n	What are the accuracy requirements for the data?
Surf_CA	The average values for each site should be within 3mm of reference
Surf_DE	The required reproducibility is 1.39 mm for a single measurement and 1.25 mm for a double measurement (DIN 55350-13)
Surf_IE	The accuracy of the rut depth measurements shall be such that the difference between the measured rut depth and the true rut depth is less than 3 mm for 95% of readings.
Surf_NL1	95% of the measurements should lie within 3mm of the reference
Surf_NL2	95% of the average values per 100m may not deviate by more than ±4mm from the reference
Surf_NZ	Static: Each of the measured rut depths must be within ±1 mm or ±10%, whichever is the greater, of the reference rut depths. The regression analysis must give a correlation coefficient r^2 ≥0.9 when the static transverse profiles are regressed against the reference for each site and >0.925 for all sites combined. Dynamic: The data is dealt with in a similar manner to Australia, with the requirements being 0.90 ≤ A ≤ 1.10 –2.5 ≤ B ≤ 2.5 mm r2 ≥ 0.900 for an automated device or 0.925 ≤ A ≤ 1.075 –0.75 ≤ B ≤ 0.75 mm r2 ≥ 0.925 for manual reference methods for each wheelpath, when the mean of the runs at each speed is regressed against the reference rut depth. Similarly when considering all data (1500 records): 0.925 ≤ A ≤ 1.075 –2.0 ≤ B ≤ 2.0 mm r^2 ≥ 0.925 for automated reference device or 0.90 ≤ A ≤ 1.10 –2.5 ≤ B ≤ 2.5 mm r^2 ≥ 0.85 for manual reference methods. The difference in the mean rut depth obtained for any combination of two different speeds must agree to better than ±2mm for any 20m segment.
Surf_SE	≥80% of lengths assessed have an absolute difference from the reference of ≤1.0mm if the reference rut is ≤7.5mm or ≤(1.0+0.05*(refRut-7.5)) otherwise, for two of the rutting parameters.
Surf_UK1	At least 65% of the differences between the Rut Depth in each Wheelpath calculated from the measured Enhanced Transverse Profile and the Rut Depth in each Wheelpath calculated from the True Transverse Profile fall within ±1.5mm. At least 95% of the differences between the Rut Depth in each Wheelpath calculated from the measured Enhanced Transverse Profile and the Rut Depth in each Wheelpath calculated from the True Transverse Profile fall within ±2.75mm. The differences between the Rut Depth in each Wheelpath calculated from the measured Enhanced Transverse Profile and the Rut Depth in each Wheelpath calculated from the True Transverse Profile shall never exceed 8mm, or 50% of the magnitude of the True Rut Depth, whichever is the greater.
Surf_UK2	The test is passed if all the following criteria are met: The rut depths calculated by the Contractor must match those calculated by the Accreditation Tester to 2 decimal places 65% of the differences between the measured maximum Rut Depths in each Wheelpath, and the Reference maximum Rut Depth in each Wheelpath fall within ±1.5mm. 95% of the differences between the measured maximum Rut Depths in each Wheelpath, and the Reference maximum Rut Depth in each Wheelpath fall within ±3.0mm. All of the differences between the measured maximum Rut Depths in each Wheelpath, and the Reference maximum Rut Depth in each Wheelpath are less than ±10.0mm, or 50% of the magnitude of the Reference Rut Depth, whichever is the greater.

Table 22: System repeatability requirements for rutting parameters

Spec'n	What are the system repeatability requirements for the data?
Surf_AU	The standard deviation, Snw, and the coefficient of variation is determined (i.e. the standard deviation expressed as a percentage of the mean), Snw%, in each wheelpath for each 100 m segment for each series of repeat measurements and reported to the nearest 0.1%. The requirement is that 90% of all standard deviation values are ≤ 1 mm or that 90% of all coefficient of variation values are ≤ 10%. The average segment standard deviation, S, and the average of the segment coefficients of variation, S% (reported to the nearest 0.1mm) are determined. The requirement is that either S≤1 mm or S%≤7%. Using least squares regression, the coefficient of determination, r^2, is calculated, when the individual rutting values for each wheelpath and segment (dependent variable, y) are regressed against the mean values for that wheelpath and segment (independent variable, x). The requirement is that all of these values are ≥0.90. Calculate the bias error between the comparison data set and the reference data set. The requirement is that the bias error is ≤ 5%.
Surf_CA	The standard deviation between the five runs should be within ±3mm
Surf_DE	Not assessed
Surf_IE	Not assessed
Surf_NL1	Not assessed
Surf_NL2	Not assessed
Surf_NZ	The coefficient of variation is determined in each wheelpath for each 10m segment for each series of repeat measurements and reported to the nearest 0.1%. The coefficient of variation, at each speed, should be <5% or each segment value should be within 2mm of the mean values, whichever is the greatest, for each series of 5 repeat measurements. The average of the segment standard deviations at each speed is required to be ≤3% or 0.8mm, whichever is the greater. The coefficient of determination, r^2, is calculated, when the individual rutting values for each wheelpath and segment (dependent variable, y) are regressed against the mean values for that wheelpath and segment (independent variable, x). The requirement is that all of these values are ≥0.95. The bias error from the two sets of five repeat surveys at each speed should be ≤ 2% or the difference between the sets of five repeat runs is within 2mm of the mean.
Surf_SE	On the test track, the 75th percentile of absolute differences between the runs should be <0.5mm for RutMax17, RutMax15, RutLeft17 and RutLeft15. On the network routes, measurement error ≤0.4mm for all 4 parameters relative measurement error ≤6% for RutLeft15, ≤5% for all 3 other parameters.
Surf_UK1	• At least 65% of the differences between the repeat measurements fall within ±1.0mm. • At least 95% of the differences between the repeat measurements fall within ±1.75mm.
Surf_UK2	Not explicitly tested but implied by assessing the accuracy of repeat runs.

3.2.11 Tests applied to Longitudinal Profile measurements

Test procedures

Only six of the specifications test the accuracy of longitudinal profile data, with only five detailing how this is achieved (Table 23). All five specifications test longitudinal profiles that have been collected during a traffic speed survey of a site (usually a test track), and then compare these with reference profiles.

Four specifications require that the longitudinal profiles collected at different survey speeds are assessed, enabling the effect of measurement speed to be assessed (Table 24). No tests are carried out to determine the effect of temperature, season etc. on the consistency of the data (Table 25).

Reference data
Reference data for longitudinal profiles is delivered by a number of different devices, most being slow speed devices (Table 26). No details are given in the specifications as to how often the data is updated, nor how the quality of the reference data is checked but this information has been provided by the partners for 4 of the specifications. The reference data is updated every 2 to 5 years and the quality of the data is ensured by checking it against the data from other measurement devices (Table 26).

Assessing accuracy of longitudinal profile
A common approach to assessing the accuracy of longitudinal profile is to compare the heights of individual profile points with the corresponding points in the reference data. A limit on the percentage of points greater than a certain tolerance is either set for the whole survey site, or for the usual reporting length for the ride quality parameter (Table 27). In addition, three of the specifications also require the correlation between the measured and reference profiles to be assessed. This ensures that the measured profile remains in phase with the reference.

Assessing system repeatability
Only two of the specifications explicitly set out requirements for system repeatability, with repeatability being implied from testing the accuracy of repeat runs in a further two of the specifications (Table 28). For those that do test, one considers the difference in the profile heights, whilst the other considers the standard deviation of the points, for each measurement location.

Table 23: What surveys and tests are carried out for longitudinal profile?

Specification	What do the tests entail?
Surf_NL1	No details given
Surf_SI	Test track survey
Surf_SE	Each device surveys 3 lengths (~1200m each) on a test track. Repeat surveys are performed.
Surf_UK1	The device surveys a test site (~2km), repeat surveys are performed.
Surf_UK2	The device surveys a test site (~2km), repeat surveys are performed.
Surf_US	Survey of a 320m site with minimal variation in the transverse pavement profile. Repeat surveys are performed.

Table 24: Speed of surveys

Specification	Is the vehicle tested for all operating speeds (e.g. 20, 40, 60, 80km/h)?
Surf_NL1	No details given
Surf_SI	Profiler can measure at least at speeds 30 km/h to 120 km/h without affect the measurement results on the test track.
Surf_SE	Not specified but the test track sections are measured five times in three different speeds, 30, 50 and 70 km/h.
Surf_UK1	A number of test surveys are carried out at constant survey speed, at a range of speeds, in addition to test surveys carried out under conditions of deceleration on the test site.
Surf_UK2	A number of test surveys are carried out at constant survey speed, at a range of speeds, in addition to test surveys carried out under conditions of deceleration on the test site.
Surf_US	No details given

Table 25: Timeframe of surveys and tests

Specification	Is data collected from different days/months compared or tested, or only data collected on the same day?
Surf_NL1	No details given
Surf_SI	Data collected on different days.
Surf_SE	The measurements take two days in good conditions. (in May or June normally)
Surf_UK1	The Accreditation test usually takes place on one day, or a number of days very close in time.
Surf_UK2	The Accreditation test usually takes place on one day, or a number of days very close in time.
Surf_US	No details given

Table 26: Reference data for longitudinal profile

Specification	What is used for the reference (is there a "golden device")?	How often is the reference data updated?	How is the quality of the reference data confirmed?
Surf_NL1	The reference data is provided by a class 1L1221 instrument according to EN 13036:6 i.e.: · travelled distance accuracy ≤0.05% · vertical resolution ≤100mm · acquisition sampling interval >50mm but ≤125mm · reporting sampling interval >100mm but ≤250mm · Large wavelength cutoff ≥100m	No details given	No details given
Surf_SI	Geodetically determined longitudinal profile.	Test track is chosen in the way which supports updating of data in about 3 to 5 years.	Geodetical measurements are taken as being good enough to accept the reference data without hindrance.

Specification	What is used for the reference (is there a "golden device")?	How often is the reference data updated?	How is the quality of the reference data confirmed?
Surf_SE	Not specified but is a Total station combined with VTIs Primal (gives a "true" reference profile every 4 mm).	Each time the test is carried out (usually every 4-5 years but can be yearly)	It is calibrated and checked against VTI's RST vehicle.
Surf_UK1 and Surf_UK2	The ARRB Walking Profiler and/or Artificial Profile (characterised using micrometer and/or rod and level) are used for the reference data for longitudinal profile from the test site.	The specification does not state how often the reference data is updated but this is about every 2 years.	The quality of the data for the test site is confirmed by checking it against the last set of reference data and against all devices surveying the test site (there are >10 devices used for surveying UK roads). This is not stated in the spec.
Surf_US	An accepted reference pavement profile measuring method (for example, rod and level)	No details given	No details given

Table 27: Data accuracy requirements for longitudinal profile

Specification	Length for assessment	What are the accuracy requirements?
Surf_NL1	1m	95% of the measurements are to be within ±1mm
Surf_SI	200m	Correlation coefficient R^2 >=0.92
Surf_SE	0.1m/20m	The measured LP should be within 0.4mm of the reference for >90% of the points.
Surf_UK1	0.1m/10m	The longitudinal profile is filtered to attenuate wavelength features of greater than 3m, 10m and 30m in length. For these three filtered datasets: • 95% of the differences between the measured Longitudinal Profile and the Reference Profile fall within ±2.0mm for the 3m filter, ±4.0mm for the 10m filter and ±6.0mm for the 30m filter. • 95% of the cross-correlation coefficients equal or exceed 0.75 for the 3m filter and 0.85 for the 10m and 30m filters. The data from tests for the effects of speed and deceleration are used to determine the range of speeds and accelerations for which the equipment delivers acceptable data, using these criteria.

Specification	Length for assessment	What are the accuracy requirements?
Surf_UK2	0.1m/10m	The longitudinal profile is filtered to attenuate wavelength features of greater than 3m and 10m in length. For these two filtered datasets: • 95% of the differences between the measured Longitudinal Profile and the Reference Profile fall within ±2.0mm for the 3m filter and ±4.0mm for the 10m filter • 95% of the cross-correlation coefficients equal or exceed 0.75 for the 3m filter and 0.85 for the 10m filter. The data from tests for the effects of speed and deceleration are used to determine the range of speeds and accelerations for which the equipment delivers acceptable data, using these criteria.
Surf_US	0.3m	The bias in the measurement of longitudinal profile is the average of the absolute value of the individual biases for each measurement point. The bias should not exceed 2.5mm.

Table 28: System repeatability requirements for longitudinal profile

Specification	Length for assessment	What are the repeatability requirements?
Surf_SE	0.1m	>90% of the points from the repeat run should be within 0.4mm of the original run
Surf_UK1	0.1m/10m	The repeatability is implied from the accuracy assessment of repeat runs
Surf_UK2	0.1m/10m	The repeatability is implied from the accuracy assessment of repeat runs
Surf_US	0.3m	Calculate the standard deviation of profile heights for all measurement points. The mean standard deviation for the 10 surveys of the whole site must not exceed 0.76mm

3.2.12 Tests applied to Ride Quality parameters

Test procedures

Eleven specifications test the accuracy of ride quality parameters (Table 29), with nine including a description of how the tests are performed. One specification only requires the ride quality parameters to be assessed using surveys over short (<1km) sites, with another requiring only surveys over long (>1km) sites. The remaining 7 specifications require surveys of both short (usually test tracks) and long sites (Table 29 and Table 32). Eight specifications require repeat surveys to be carried out.

As with the other measurements/parameters, no tests are carried out to determine the effect of temperature, season etc. on the consistency of the data but five of the specifications require data to be collected at different survey speeds, enabling the effect of measurement speed to be assessed (Table 30).

Reference data

Reference data for ride quality parameters is delivered by a number of different devices, with slow speed equipment used for the short site surveys and independent profilers used to provide data for the longer network surveys (Table 31). Only two specifications state how often the reference data is updated, with none describing how the reference data is itself checked but this information has been provided by the partners for 3 of the specifications.

The reference data is updated either every year, or every time an accreditation test occurs and the quality of the data is ensured by checking it against the data from other measurement devices (Table 31).

Length of assessment
The most common approach is to assess the ride quality parameters over the reporting length used, with only two exceptions to this (Table 20).

Assessing data accuracy
Most of the specifications require a percentage of the measured ride quality parameters to lie within a set tolerance of the reference parameter values (Table 33). The actual percentage and range applied are specific to the parameter being tested and also the individual specification. Two of the specifications (AU and NZ) apply a more statistical approach to assessment.

Assessing system repeatability
Eight of the specifications require repeat surveys to be carried out on network routes. Of these, five directly test system repeatability; with the others implying the repeatability from testing the accuracy of repeat surveys (Table 34). Different approaches are used to test the repeatability. Two specifications use statistical analyses, one calculates the standard deviation of the repeat runs, and two calculate the error between repeat runs (Table 34).

Table 29: Testing ride quality parameters

Specification	What do the tests entail?
Surf_AU	Roughness Method 1 - survey of test sections Roughness Method 2 - survey of ‘Roughness Calibration Loop’ Plus 5 repeat surveys of 10 km site.
Surf_CA	Four test sites are used, each 500m long with a 250 metre lead-in. The sites exhibit a representative variety of distress types, range in pavement deterioration, surface types and operating speed with the intent to be representative of the actual survey conditions. Five repeat surveys are performed on each site.
Surf_DE	The device has to measure a number of road sections, including repeat surveys.
Surf_IE	The specification states that the Service Provider has to establish a satisfactory correlation between the IRI values derived from their equipment and IRI values derived from previous survey devices (ARAN, RSP). How to carry out this assessment, what is tested and what is deemed to be satisfactory is not specified.
Surf_NL1	No details given
Surf_NL2	No details given
Surf_NZ	Survey of test sections, plus 5 repeat surveys of 10 km site
Surf_SE	Survey of a test track and comparison to IRI calculated from reference data, plus five repeat surveys on network routes.
Surf_UK1	The equipment is used to survey a primary test site (~2km) (test track) and also some routes on the road network (currently 2 sites covering ~55km in total). Repeat surveys are performed on all sites
Surf_UK2	The equipment is used to survey a primary test site (~2km) (test track) and also some routes on the road network (currently 2 sites covering ~55km in total). Repeat surveys are performed on all sites
Surf_US	Survey of the trial sections in the Baton Rouge area and possibly other road sections. Three repeat surveys are performed on each site.

Table 30: Conditions for data collection

Specification	Is data collected from different days/months compared or tested?	Is the vehicle tested for all operating speeds (e.g. 20, 40, 60, 80km/h)?
Surf_AU	No details are given but the tests appear to be performed all in one day.	Each test section is surveyed at several test speeds: A minimum of three test speeds are selected; one near the bottom, mid-range and top (nominally 100 km/h) of the profilometer's specified operating speed range.
Surf_CA	No details are given in the specification	No details are given
Surf_DE	No details are given but in practice, it will depend on number of sections, with the whole test taking 2 or 3 days.	No
Surf_IE	No details are given in the specification	No details are given
Surf_NL1	No details are given in the specification	No details are given
Surf_NL2	No details are given in the specification	No details are given
Surf_NZ	Data is collected over 2 consecutive days: 5 runs on the first day, at least 3 different speeds each, repeated on the second day	Each test section is surveyed at several test speeds: A minimum of three test speeds are selected; one near the bottom, mid-range and top (nominally 100 km/h) of the profilometer's specified operating speed range.
Surf_SE	No details are given but the measurements take two days in good conditions. (in May or June normally)	Not specified but the test track sections are measured five times at three different speeds, 30, 50 and 70 km/h.
Surf_UK1 and Surf_UK2	Not details given but all surveys of the test site usually take place on one day, or a number of days very close in time. Surveys of the network routes may take place at any time but generally repeat runs are completed on the same day.	A range of (constant) speeds are used when surveying the test track (20, 40, 60, 80km/h) and the network routes are surveyed at normal traffic speed.
Surf_US	No details are given in the specification	No details are given

Table 31: Reference data used for ride quality parameter accreditation

Specification	What is used for the reference (is there a "golden device")?	How often is the reference data updated?	How is the quality of the reference data confirmed?
Surf_AU	Method 1 - Walking Profiler or a surveyor's staff and level. Method 2 - five repeat runs by an independent laser profilometer.	No details given	No details given
Surf_CA	Class 1 profiler	Every time a round of surveys is carried out	No details given
Surf_DE	Golden device from BASt	Once a year	Golden device is considered true

Specification	What is used for the reference (is there a "golden device")?	How often is the reference data updated?	How is the quality of the reference data confirmed?
Surf_IE	No details given	No details given	No details given
Surf_NL1	The reference data is provided by a class 1L1221 instrument according to EN 13036:6*	No details given	No details given
Surf_NL2	No details given	No details given	No details given
Surf_NZ	ARRB Walking Profiler or a surveyor's staff and level.	Not specified	Not specified
Surf_SE	Not specified but is a Total station combined with VTIs Primal (gives a "true" reference proifile every 4 mm).	When needed	Calibration and checking against VTI's RST device
Surf_UK1 and Surf_UK2	The ARRB Walking Profiler is used for the reference data for LPV on the primary test site and a reference profilometer is used as the reference on the network routes.	The specification does not state how often the reference data is updated but this is about every 2 years on the primary site and yearly on the network routes.	No details given but the quality of the data is confirmed by checking it against the last set of reference data and also comparing it with the all devices undergoing SCANNER or TRACS accreditation. The reference profilometer undergoes all of the tests carried out on the test site and is also compared to the other devices.
Surf_US	The DOTD's "South Dakota" style laser profiler (inertial profiler) and/or a Class I profiling instrument.	No details given	No details given

*i.e. travelled distance accuracy ≤0.05%, vertical resolution ≤100mm, acquisition sampling interval >50mm but ≤125mm, reporting sampling interval >100mm but ≤250mm, Large wavelength cutoff ≥100m

Table 32: Length of survey sites and parameter assessment length

Specification	If surveys are carried out as part of the tests, what distance is surveyed?	What reporting length is the parameter assessed over?
Surf_AU	There are two test methods applied. Method 1 - five test sections of road pavement, each 500 m long Method 2 - 32 km long 'Roughness Calibration Loop' Plus 10km for repeat surveys	50m
Surf_CA	4 sites of 750m (incl. 250m run-in)	The data is provided at 50m spacing and this is averaged to provide data at

Specification	If surveys are carried out as part of the tests, what distance is surveyed?	What reporting length is the parameter assessed over?
		500m spacing (i.e. for a whole site)
Surf_DE	30 - 100 km	100 m reporting length
Surf_IE	No details given	No details given
Surf_NL1	No details given	100m
Surf_NL2	No details given	100m
Surf_NZ	Five test sections of road pavement, each 500 m long Plus 10km for repeat surveys	100m
Surf_SE	~3600m in total	20m
Surf_UK1	About 2km primary test site + 55km of network routes	10m on test site and network routes
Surf_UK2	About 2km primary test site + 55km of network routes	10m on primary test site and 50m on network routes
Surf_US	No details given	0.1mile (160.9m)

Table 33: Accuracy requirements for ride quality parameters

Spec'n	What are the accuracy requirements for the data?
Surf_AU	**Method 1**: For each of the three speeds, group the IRI data measured by the inertial profilometer into a single set of data, totalling 125 records. Using least squares regression, a line of best fit between the two sets of results should be identified for each speed in the form: IRIBase = A.IRIProfilometer+B where IRIBase =Lane IRI calculated from the base reference measurements (i.e. either ARRB Walking Profiler, or staff and level), IRIProfilometer =Lane IRI calculated from the operational laser profilometer, A =regression equation slope B =regression equation intercept The coefficient of determination, r^2, for the regression must also be determined. The requirements are $0.95 \le A \le 1.05$ $-0.25 \le B \le 0.25$ m/km, $r^2 \ge 0.950$ Do the same for all of the IRI data measured by the inertial profilometer (375 records). The requirements are: $0.97 \le A \le 1.03$ $-0.25 \le B \le 0.25$ m/km, $r^2 \ge 0.975$. **Method 2**: Calculate the average roughness value for each 100m section, then calculate a line of best fit, using least squares regression. The requirement is that r2 ≥0.95 (see Method 1) Calculate the average percentage difference. This is the overall average of the percentage differences for each 100 m section between the average of the five runs of the profilometer being tested and the corresponding reference data. It is obtained as follows: 100*sum((rAverage-rReference)/rReference)/n, where n =the total number of 100 m sections in the analysis. rReference roughness of the 100 m section from the reference dataset rAverage roughness of the five repeat runs for the 100 m section. Requirement is that average percentage difference ≤ ±5%.
Surf_CA	The average values for each site should be within 10% of the reference value
Surf_DE	Three AUN parameters are calculated 1 cm^3, 3 cm^3, 10 cm^3. The requirements are that all 1 cm^3 AUN measurements are within 0.58, all 3 cm^3 AUN measurements are within

Spec'n	What are the accuracy requirements for the data?
	1.50, and all 10 cm^3 AUN measurements are within 4.29.
Surf_IE	No details given
Surf_NL1	95% of the measurements should lie within 0.5% of the reference
Surf_NL2	95% of the average values per 100m may not deviate by more than ±0.50m/km from the reference
Surf_NZ	The same method as Australia Method 1 is implemented but with the following requirements: For each speed, $0.98 \leq A \leq 1.02$ $-0.25 \leq B \leq 0.25$ m/km, $r^2 \geq 0.95$ For all measurements $0.98 \leq A \leq 1.02$ $-0.25 \leq B \leq 0.25$ m/km, $r^2 \geq 0.975$. This should be done for both the lane IRI and also the individual wheel path IRIs.
Surf_SE	75% of the 20m length should have an absolute error (between the measured and reference data) of ≤0.35mm/m if the reference ≤2.0mm/m or 0.35+0.1*(refIRI-2.0) otherwise
Surf_UK1	Three eLPV parameters are calculated: 3m, 10m and 30m eLPV. For these: • 65% of the errors between the measured LPV and the Reference LPV fall within ±0.3mm^2 for 3m eLPV, ±0.35mm^2 for the 10m eLPV and ±0.45mm^2 for the 30m eLPV • 95% of the errors between the measured LPV and the Reference LPV fall within ±0.6mm^2 for 3m eLPV, ±0.7mm^2 for the 10m eLPV and ±0.9mm^2 for the 30m eLPV. The data from tests for the effects of speed and deceleration are used to determine the range of speeds and accelerations for which the equipment delivers acceptable data, using these criteria.
Surf_UK2	Two LPV parameters are calculated: 3m and 10m LPV. For these: • 65% of the errors between the measured LPV and the Reference LPV fall within ±0.3mm^2 for 3m LPV and ±0.35mm^2 for the 10m LPV • 95% of the errors between the measured LPV and the Reference LPV fall within ±0.6mm^2 for 3m LPV and ±0.7mm^2 for the 10m LPV. The data from tests for the effects of speed and deceleration are used to determine the range of speeds and accelerations for which the equipment delivers acceptable data, using these criteria.
Surf_US	Max. error of 5% bias or 20 inches/mile (0.32mm/m)(whichever is less)

Table 34: System repeatability requirements for ride quality parameters

Spec'n	What are the system repeatability requirements for the data?
Surf_AU	Determine the coefficient of variation (i.e. the standard deviation expressed as a percentage of the mean), Sn%, for each 100 m segment for each series of repeat measurements. The requirement is that when 95% of coefficient of variations are ≤5%. Determine the average of the segment coefficients of variation, S%. The requirement is that S% should be ≤3%. Using least squares regression, determine the coefficient of determination, r2, when the individual roughness values for each segment (dependent variable, y) are regressed against the mean values for that segment (independent variable, x). The requirement is that all of these values ≥ 0.95. Calculate the bias error between the comparison data set and the reference data set. The requirement is that the bias error is ≤ 1%.
Surf_CA	0.1 mm/m standard deviation for five runs
Surf_DE	Repeatability is implied by assessing the accuracy of repeat surveys
Surf_IE	No details given

Spec'n	What are the system repeatability requirements for the data?
Surf_NL1	Not assessed
Surf_NL2	Not assessed
Surf_NZ	The difference in mean roughness obtained for any combination of two different survey speeds must be within 0.1 IRI for 68%, and within 0.2 for 95% of the 100m segments. The coefficient of variation, Sn%, for each 100 m segment for each series of repeat measurements must be ≤5% for each speed, for 95% of 100m lengths. Determine the average of the segment coefficients of variation, S%. The requirement is that S% should be ≤3% for each speed. Using least squares regression, determine the coefficient of determination, r^2, when the individual roughness values for each segment (dependent variable, y) are regressed against the mean values for that segment (independent variable, x). The requirement is that all of these values ≥ 0.95 for each segment at each speed. Calculate the bias error between the comparison data set and the reference data set. The requirement is that the bias error from the two sets of five repeats at each speed is ≤ 2%.
Surf_SE	For IRIright on the test track, 75th percentile of errors <0.2mm/m For IRIright and IRIleft on the network routes, relative errors ≤5%, measurement errors ≤0.12mm/m Xij is the mean of valid measurements for run i and 400m length j X.j is the mean of the Xij for each 400m length j X.. is the mean of all X.j for all 400m lengths MSEj is the variance of the Xij for each length j MSEtot is the mean of the MSEj measurement error =sqrt(MSEtot) relative error = sqrt(MSEtot)/X..
Surf_UK1	Three eLPV parameters are calculated: 3m, 10m and 30m eLPV. For these: • 65% of the errors between the measured LPV from repeat surveys fall within ±0.3mm2 for 3m eLPV, ±0.35mm2 for the 10m eLPV and ±0.40mm2 for the 30m eLPV • 95% of the errors between the measured LPV from repeat surveys fall within ±0.55mm2 for 3m eLPV, ±0.65mm2 for the 10m eLPV and ±0.75mm2 for the 30m eLPV.
Surf_UK2	Repeatability is not directly assessed – it is implied by assessing accuracy of repeat runs
Surf_US	Repeatability is not directly assessed – it is implied by assessing accuracy of repeat runs

3.2.13 Tests applied to Surface Defect measurements

Test procedures

Of the eight specifications requiring a measure of surface defects (e.g. cracking), six define accreditation tests (Table 35). For all specifications, for which detail is given, the surface defect parameters are assessed using surveys over long (>1km) sites (Table 38). Three specifications require repeat surveys to be carried out. As with the other measurements/parameters, no tests are carried out to determine the effect of temperature, season etc. on the consistency of the data, nor are there any tests for the effect of speed (Table 36).

Reference data

Reference data for surface defect parameters is delivered either by manual analysis of downward facing images, or an on-site visual survey (Table 37). None of the specifications state how often the reference data is updated, nor describe how the reference data is itself checked but this information has been provided by the partners for 3 of the specifications (Table 37).

Length of assessment
The most common approach is to assess the surface defect parameters over the reporting length used, with only one exception to this (Table 38).

Assessing data accuracy
Where the surface deterioration data has been produced automatically (i.e. with software) most of the specifications require the amount of measured surface deterioration to lie within the same category, or a very close category to the reference. (Table 39), with one specification requiring the contractor to calibrate the automatic system, using specified sites. One specification allows delivery of surface deterioration data via visual assessment of downward facing images (US). For this, the contractor is expected to calibrate all raters used for manual analysis, to ensure consistent results, using specified sites.

Assessing system repeatability
Three of the eight specifications require repeat surveys on network routes. Of these, only one directly tests system repeatability; with the others implying the repeatability from testing the accuracy of repeat surveys (Table 40).

Table 35: Testing surface defect parameters

Specification	What do the tests entail?
Surf_DE	Take pictures of a target board with specific patterns for determining level of detail captured, survey certain road sections and do evaluation of these section according to catalogue of defects (cracking and patching on asphalt; cracks, corner breaks, edge damage on concrete)
Surf_IE	Same tests as UK2
Surf_NL2	The level of precision of cracking is based on assessment per 100m and tested against reference values. The levels of cracking severity and volume are defined in CROW publication 146a and b (Dec 2011).
Surf_UK1	The device surveys a number of network routes that include a range of surface deterioration types and prevalence (this currently consists of 10 routes, ~120km).
Surf_UK2	The device surveys a number of network routes that include a range of cracking types and prevalence (this currently consists of 10 routes covering ~120km).
Surf_US	The Consultant shall use the Louisiana version of Protocols for Automated Distress Data Collection, the Distress Identification for Long-Term Pavement Performance Project Manual (SHRP-338), with the appropriate changes and adaptations, and the Manual for the Identification of Pavement Distresses for Louisiana. The DOTD shall test and verify the consistency of several quantified processed data. The contractor is expected to calibrate all raters used for manual analysis, to ensure consistent results, using specified sites. Or, if using automatic detection, the contractor is expected to calibrate the automatic system, using specified sites.

Table 36: Conditions for data collection

Specification	Is data collected from different days/months compared or tested?	Is the vehicle tested for all operating speeds (e.g. 20, 40, 60, 80km/h)?
Surf_DE	Depends on number of sections, whole test takes 2 or 3 days.	No
Surf_IE	Same as UK2	Same as UK2
Surf_NL2	No details given	No details given
Surf_UK1	Generally the data is collected over a short period of time	The network routes are surveyed at normal traffic speed.

Specification	Is data collected from different days/months compared or tested?	Is the vehicle tested for all operating speeds (e.g. 20, 40, 60, 80km/h)?
Surf_UK2	Generally the data is collected over a short period of time	The network routes are surveyed at normal traffic speed.
Surf_US	No details given	No details given

Table 37: Reference data used for surface defect parameter accreditation

Specification	What is used for the reference (is there a "golden device")?	How often is the reference data updated?	How is the quality of the reference data confirmed?
Surf_DE	Golden device from BASt	No details given	Some cracks on the sections are measured in-situ by hand
Surf_IE	Same as UK2	Same as UK2	Same as UK2
Surf_NL2	Detailed visual survey	No details given	No details given
Surf_UK1	Manual assessment of digital images	Every 2 years	A number of lengths are randomly chosen on the sites and the surface deterioration is independently checked
Surf_UK2	Manual assessment of digital images	Every 2 years	A number of lengths are randomly chosen on the sites and the surface deterioration is independently checked
Surf_US	No details given	No details given	No details given

*i.e. travelled distance accuracy ≤0.05%, vertical resolution ≤100mm, acquisition sampling interval >50mm but ≤125mm, reporting sampling interval >100mm but ≤250mm, Large wavelength cutoff ≥100m

Table 38: Length of survey sites and parameter assessment length

Specification	If surveys are carried out as part of the tests, what distance is surveyed?	What reporting length is the parameter assessed over?
Surf_DE	30 - 100 km	100 m
Surf_IE	Same as UK2	Same as UK2
Surf_NL2	No detail given	No detail given
Surf_UK1	The network routes currently total ~120km for an accreditation and ~25km for a re-accreditation	10m
Surf_UK2	The network routes currently total ~120km	50m
Surf_US	No detail given	No detail given

Table 39: Accuracy requirements for surface defect parameters

Spec'n	What are the accuracy requirements for the data?
Surf_DE	The average percentage of measured cracking must be within 0.5 of the reference percentage of cracking and the standard deviation of all repeat runs must be within 0.5. This is also applied to the other surface defects reported.
Surf_IE	Same as UK2
Surf_NL2	On bituminous pavements: 90% of the assessed 100m road sections must meet the following requirements: • The difference between the "volume" category for the measured value and that for the reference value may not exceed 1. • The aggregated length of light, moderate and severe cracking in the measured values may not differ by more than 20m from the reference. On concrete pavements: 90% of the assessed road sections must meet the following requirement: • The difference in "volume" category between the measured data and the reference may not exceed 1. Transverse cracks on bituminous: 90% of the assessed 100m road sections must meet the requirement below, whereby the average of the volume of transverse cracks in the measurements and in the reference values for the relevant section is called "DWSm". • The difference in volume between the number of metres of transverse cracks measured and that of the reference may not exceed 20% of "DWSm" plus 20 m.
Surf_UK1	A high level of surface deterioration should be reported for at least 70% of the lengths where the reference has reported high levels. A moderate level of surface deterioration should be reported for at least 45% of the lengths where the reference has reported moderate levels. A low level of cracking should be reported for at least 80% of the lengths where the reference has reported low levels. Also, the % of area of non-Surface Deterioration features reported as Surface Deterioration must be less than 10%.
Surf_UK2	A high level of cracking should be reported for at least 65% of the lengths where the reference has reported high levels. A low level of cracking should be reported for at least 85% of the lengths where the reference has reported low levels. The overall level of agreement between the level of cracking, for low, moderate and high cracking combined, is at least 75%.
Surf_US	The DOTD shall test and verify the consistency of several quantified processed data. The contractor is expected to calibrate all raters used for manual analysis, to ensure consistent results, using specified sites. Or, if using automatic detection, the contractor is expected to calibrate the automatic system, using specified sites.

Table 40: System repeatability requirements for surface defect parameters

Spec'n	What are the system repeatability requirements for the data?
Surf_DE	Same as accuracy (Table 39)
Surf_NL2	Same as UK2
Surf_UK1	The same tests are carried out for the repeatability of data as for reproducibility, using data from repeat surveys. The requirements for repeatability of the data are the same as those for reproducibility.
Surf_UK2	System repeatability is not explicitly tested but is implied by testing the accuracy of repeat runs.

3.2.14 Tests applied to Images

Of the five specifications requiring delivery of downward facing images, only two have tests for the quality of the images delivered. Both of these require the device to survey a special mat or board, which has specific patterns on it for determining the level of detail provided by the images and also the quality of the images (Table 41). In addition, one of the specifications includes subjective assessment of images collected during site surveys and also quantitative tests on the level of contrast and focus of the images.

Table 41: Testing images

Specification	What do the tests entail?
Surf_DE	Take pictures of a target board with specific patterns for determining level of detail captured
Surf_UK1	Accreditation testing includes: • Manual visual assessment of the Downward Facing Images on selected test sites. • Controlled tests carried out on a test mat. • “Network” tests of the contrast and focus of the Downward Facing Images collected on the Accreditation network sites. Re-Accreditation Tests include: • All or part of the tests carried out for Accreditation testing. • Tests on the consistency of illumination of the Downward Facing Images. • “Network” tests of the contrast and focus of the Downward Facing Images collected on the re- Accreditation network sites

3.3 Review of current surface condition equipment Quality Assurance testing

The findings of the reviews have been summarised in the following subsections, for Quality Assurance (QA) regimes applied to the surface condition measurements and parameters identified in Work Package 1.

3.3.1 What parameters and measurements are tested during QA?

Three (IE, NE1, NZ) of the fourteen surface condition specifications (Table 1) do not specify requirements for Quality Assurance testing of the data. One specification (SI) does require it but leaves it to the contractor to propose an appropriate regime, whilst another (Surf_US2) requires survey devices to go through a quality assurance process on selected segments of the network but contains no further details. Therefore, the nine remaining specifications have been considered in the review.

Location referencing

Despite all specifications requiring delivery of either distance and/or position measurements, only five include QA tests for the accuracy of location referencing.

Transverse and longitudinal profile measurements/parameters

For the transverse and longitudinal profile measurements/parameters, six of the specifications require what is delivered to be tested during QA. Three specifications (DE,UK1, UK2) test the calculated parameters, not the profiles that are delivered. A further specification (SE) tests everything delivered, apart from transverse profile. It is interesting to note that only one specification (SE) tested longitudinal profile during the QA tests, whereas none tested transverse profile. This is summarised in Table 42.

Images and surface deterioration

For the four specifications that require delivery of images (in order to perform surface deterioration analysis), two test the quality of the images (UK1, US) during QA. In a further specification (UK2) tests are required if the images are used to obtain cracking data, although the images themselves are not a deliverable (UK2). It is thought that the relative lack of testing is because such tests are difficult to devise, carry out and also assess quantitatively. The majority (4 out of 6) of specifications requiring delivery of a surface deterioration parameter included tests for this in QA (Table 42).

Thus, the most common practice appears to be that the QA tests only assess the parameters that are deliverables (or derived from deliverables) from the survey.

Table 42: Required measurements and parameters compared to what is tested during QA

	Location		Transv. Profile		Rutting		Long'l Profile		Ride Quality		Images		Surface defects	
Specification	**D**	**QA**	**D**	**QA**	**D**	**QA**	**D**	**QA**	**D**	**QA**	**D**	**QA**	**D**	**QA**
Surf_AU	✓				✓	✓			✓	✓				
Surf_AT	✓				✓	✓			✓	✓	✓		✓	
Surf_CA	✓				✓	✓			✓	✓				
Surf_DE	✓		✓			✓	✓			✓	✓		✓	✓
Surf_NL2	✓				✓	✓			✓	✓			✓	✓
Surf_SE	✓	✓	✓		✓	✓	✓	✓	✓	✓	(✓)			
Surf_UK1	✓	✓	✓			✓	✓			✓	✓	✓	✓	✓
Surf_UK2	✓	✓	✓		✓	✓			✓	✓		✓	✓	✓
Surf_US	✓	✓			✓	✓			✓	✓	✓	✓	✓	

D=Is it Delivered in the survey?, **QA**=is it tested under a QA regime?

(✓) Images are delivered but are not currently used for network assessment of surface deterioration.

3.3.2 What type of data accuracy is tested during QA?

All parameters or measurements that are assessed under a QA regime are tested for system repeatability (Table 43). This is achieved by performing multiple (repeat) surveys on the same length of road and comparing the data. Most of the QA tests also test the accuracy (6/9 of the rutting tests, 6/9 ride quality parameter tests, 4/5 of the surface deterioration tests, see Table 43). Only one specification (SE) tests longitudinal profile directly under a QA regime (both accuracy and system repeatability), whilst none test transverse profile directly

Table 43: Aspects of data accuracy tested during QA

	Transverse Profile		Rutting		Longitudinal Profile		Ride Quality		Surface defects	
Specification	**A?**	**R?**	**A?**	**R?**	**A?**	**R?**	**A?**	**R?**	**A?**	**R?**
Surf_AU				✓				✓		
Surf_AT				✓				✓		
Surf_CA			✓	✓			✓	✓		
Surf_DE			✓	✓†			✓	✓†	✓	✓†
Surf_NL2				✓				✓		✓
Surf_SE			✓	✓	✓	✓	✓	✓		
Surf_UK1			✓	✓			✓	✓	✓	✓
Surf_UK2			✓	✓			✓	✓	✓	✓
Surf_US			✓	✓			✓	✓†	✓	✓

A=Accuracy, **R**=System repeatability
† Repeat runs are used to check accuracy, which implies repeatability.

3.3.3 How QA tests are carried out

How are measurements obtained, during the QA tests?
Most of the specifications do not state how the QA tests are carried out (i.e. on a test track, or on the road network), nor what length is surveyed. Therefore, the expertise of the project group and colleagues has been drawn on to fill this gap and assist in completing Table 44. In general the QA tests are carried out on the road network. Most specifications state that the roads chosen will be representative of those found on the network, and display a range of conditions/roughness.

Table 44: Length and type of road used during QA testing

Measurement	Spec'n	What is surveyed?	Length surveyed
Longitudinal profile	Surf_SE	Road network	Repeat surveys of 5% of each area (Each area is 200-1000km)
Ride Quality and Rutting	Surf_AU	Road network	10 km
	Surf_AT	Road network	Minimum of 2km
	Surf_CA	Unclear	4 x 750m
	Surf_DE	Road network	Minimum of 2km
	Surf_NL2	Road network	Repeat surveys on 5% (minimum of 8km)
	Surf_SE	Road network	Repeat surveys of 5% of each area (Each area is 200-1000km) for ride quality For every 1000 km production 40 km must be re-measured for rutting

Measurement	Spec'n	What is surveyed?	Length surveyed
Ride Quality and Rutting	Surf_UK1	Road network	20-30km for the accuracy tests For system repeatability: Calibration site >400m long, preferably >1km. Primary sites are a minimum of 10km Secondary sites are at least 400m, preferably >1km. Daily sites are >400m, preferably >1km.
	Surf_UK2	Road network	20-30km for the accuracy tests For system repeatability: Calibration site >400m long, preferably >1km. Primary sites are a minimum of 10km Secondary sites are at least 400m, preferably >1km. Daily sites are >400m, preferably >1km.

Seasonal effects

Some survey devices are affected by temperature and the season of measurement, for example the TSD and SCRIM (friction measurement device). The review found that, as with the Accreditation tests, none of the QA tests incorporated tests of the effect of temperature or season on the data collected, although two (UK1 and UK2) include repeat surveys on the same sites throughout the year.

3.3.4 Frequency of QA testing

Most specifications did not state how often QA testing should take place. However, for the UK specifications, the frequency ranges from daily and weekly, for system repeatability tests, through to every 3 months for accuracy tests.

3.3.5 Testing and auditing the data

Independent auditors are only used for four QA regimes (DE, SE, UK1, UK2) (Table 45), whilst the survey commissioner checks the data for another two (CA,NL2), with the rest leaving the data checking to the survey contractors. For two of the contractor-checked regimes, the specifications state that the data used for the test must be provided to the survey commissioner, to enable independent checking, if required.

Table 45: Who checks the data during QA?

Specification	Survey Contractor	Independent Auditor	Survey commissioner
Surf_AU	✓*		
Surf_AT	✓		
Surf_CA	✓		✓
Surf_DE		✓	
Surf_NL2			✓
Surf_SE	✓	✓	
Surf_UK1	✓†	✓	
Surf_UK2	✓†	✓	
Surf_US	✓*		

* Data is also provided to road authority, for checking.

† For the daily and weekly tests but Independent Auditor can request to see this data at any time.

3.3.6 QA tests applied to location referencing and distance measurement

Only five of the specifications include QA tests for distance measurements (Table 46). Two just check the measured length against the known lengths of sections (SI, US) and the survey commissioner is able to check the difference. However, there are no details as to what action is taken if too large a discrepancy is found, nor if the measurements are found to be inconsistent. The remaining specifications consider the difference between the measured distance and the actual distance, with only two specifications (UK1, UK2) stating the actual requirements for accuracy. These requirements vary, depending on the type of site being surveyed.

Only two specifications include QA tests for location referencing (Table 47), both comparing the measured positions with the actual positions and calculating the horizontal distance between them. Different requirements are imposed for this horizontal error depending on the sites being surveyed.

Table 46: Tests applied to distance measurement

Spec'n	Tests and requirements for distance data
Surf_SI	Left to the contractor, client checks with his official data of section lengths.
Surf_SE	There are limits for the differences between measured distance and the distance according the national reference system. The coordinates are not checked.
Surf_UK1	The length of the site should be within ±1m or ±0.1% for calibration checks, ±3m or ±0.1% for monthly and weekly checks, ±5m or ±0.1% for daily checks and reproducibility tests.
Surf_UK2	The length of each section should be ±3 m for lengths <1km or ±0.1% otherwise.
Surf_US	The Consultant shall calibrate the DMI (Distance Measuring Instrument) using sites provided by the DOTD. The Consultant must provide all findings, inclusive of the calibration number before the calibration process, the calibration number after the calibration process, location of the calibration site, length of the calibration site, and length of calibration site as measured by the DMI before and after calibration, and list any discrepancies found during the calibration process. Any discrepancies that are found shall be reported to DOTD with the corrective action taken with a detailed explanation. The calibration of the DMI shall be performed and reported weekly to DOTD.

Table 47: Tests applied to location measurement

Spec'n	Tests and requirements for location data
Surf_UK1	The OSGR coordinate of section start points should be within 2m for calibration checks, 4m for monthly and weekly checks, 5m for daily checks and reproducibility tests.
Surf_UK2	95% of GPS measurements should be within 3m of actual and all should be within 10m.

3.3.7 QA tests applied to Transverse Profile measurements

None of the specifications include a test for the accuracy of raw transverse profile within the QA regime. This is likely because it is assumed that issues with this measurement will become apparent when assessing the rutting parameter, derived from the transverse profile.

3.3.8 QA tests applied to Rutting parameters

Test procedures

Nine specifications test the accuracy of rutting within the QA regime (Table 48), with five including a description of how the tests are performed. All data for assessment is collected during surveys of the road network, with varying lengths being surveyed and different frequency of testing. The routes are surveyed at the same speed as the surrounding traffic. Four specifications require the data to be tested for both accuracy and repeatability, whilst three only require testing of system repeatability, with the remaining specification requiring only testing of the accuracy.

Reference data

For most specifications, the QA test surveys cover fairly long routes on the road network (Table 48). For these, reference data for rutting is provided by independent profilometers (DE, UK1, UK2) or a different profilometer operated by the same or another contractor (SE) (Table 49). No survey length details were given for one specification (CA) but the reference used for this is a straight edge and wedge. The reference data is collected very close in time to (within a month of) the original measurement for all specifications.

Length of assessment

Three specifications (DE, UK1, UK2) assess the rutting parameters over the reporting length used, with two assessing over longer lengths than used for reporting (Table 49).

Assessing data accuracy

All of the specifications require all or a percentage of the measured rutting parameters to lie within *x*mm of the reference parameter values (Table 50). The actual percentage and range applied varies. Two of the specifications (DE, SE) apply further statistical tests to the data.

Assessing system repeatability

Five of the seven specifications have requirements for the difference between rutting parameter values delivered for repeat runs, with two (DE, SE) applying further statistical tests to the data. The two remaining specifications (AU, CA) apply only statistical analyses to the data (Table 51).

Table 48: Testing rutting parameters

Specification	What do the tests entail?
Surf_AU	It is not specified that repeatability of data is tested outside of accreditation testing, however, the text "The bias check included in this method is used to determine whether there is a systematic drift in a profilometer's measurements over time" would suggest that this is an ongoing test and thus should be included in QA. The test includes a survey of a clearly defined section of a lane of 10 km total length exhibiting a significant range of rutting at the 100 m segment level. The route is surveyed 5 times.
Surf_AT	Repeat measurements on a ≥2 km long section
Surf_CA	No details given but the tests will only be performed for larger surveys
Surf_DE	A minimum length of 2 km has to be measured by both contractor and BASt
Surf_NL2	No details are given but there are requirements for repeatability tested during QA
Surf_SE	The QA tests involve the same device resurveying 5% of the area that they have just surveyed and also a different device, with a different crew surveying 5% of the area.
Surf_UK1	Auditor's Repeat Surveys (ARS) are carried out throughout the year: The Auditor will choose a route recently surveyed by the contractor, during routine network surveys, and survey that route with the reference device. This reference data is then compared to the measured data. Repeat survey data is collected at regular intervals on a calibration site, on primary sites, on secondary sites and on daily sites. The contractor compares the current data with data collected previously on the site, to check repeatability.
Surf_UK2	Auditor's Repeat Surveys (ARS) are carried out throughout the year: The Auditor will

Specification	What do the tests entail?
	choose a route recently surveyed by the contractor, during routine network surveys, and survey that route with the reference device. This reference data is then compared to the measured data. Repeat survey data is collected at regular intervals on a calibration site, on daily sites, on primary sites and on secondary sites. The contractor compares the current data with data collected previously on the site, to check repeatability.
Surf_US	The specification does not state how testing should take place, nor how accurate or repeatable the data should be. It just says that the contractor must provide the precision and bias of the equipment.

Table 49: Reference data used for rutting parameter QA and parameter assessment length

Specification	What is used for the reference?	What reporting length is the parameter assessed over?
Surf_AU	N/A	100m
Surf_AT	N/A	50m
Surf_CA	Use of a straight edge and wedge, at 10m intervals along the survey site.	500m (calculated from 50m averages)
Surf_DE	Golden device from BASt	100 m
Surf_NL2	N/A	400m
Surf_SE	A different device with a different crew	400 m
Surf_UK1	An independent reference profilometer.	10m
Surf_UK2	An independent reference profilometer.	10m
Surf_US	N/A	No details given

Table 50: QA Accuracy requirements for rutting parameters

Spec'n	What are the accuracy requirements for the data?
Surf_CA	All 500m values should be within 3mm of reference
Surf_DE	Δ mean rut depth < 1,0; standard deviation of mean rut depth < 2.5 mm
Surf_SE	Correlation >=0.9 Systematic error 0.5 mm Within interval < 0.5 + 0.5*Rut
Surf_UK1	95% of the errors between the measured rut depths and the Reference rut depths fall within ±2.5mm. All errors are within ±10mm or 50% of the Reference rut depth.
Surf_UK2	95% of the errors between the measured rut depths and the Reference rut depths fall within ±3mm. All errors are within ±10mm.

Table 51: QA System repeatability requirements for rutting parameters

Spec'n	What are the system repeatability requirements for the data?
Surf_AU	The standard deviation, Snw, and the coefficient of variation is determined (i.e. the standard deviation expressed as a percentage of the mean), Snw%, in each wheelpath for each 100 m segment for each series of repeat measurements and reported to the nearest 0.1%. The requirement is that 90% of all standard deviation values are ≤ 1 mm or that 90% of all values for the coefficient of variation are ≤ 10%. The average segment standard deviation,S, and the average of the segment coefficients of variation, S% (reported to the nearest 0.1mm) are determined. The requirement is that either S≤1 mm or S%≤7%. Using least squares regression, the coefficient of determination, r2, is calculated, when the individual rutting values for each wheelpath and segment (dependent variable, y) are regressed against the mean values for that wheelpath and segment (independent variable, x). The requirement is that all of these values are ≥0.90. Calculate the bias error between the comparison data set and the reference data set. The requirement is that the bias error is ≤ 5%.
Surf_AT	Δ of rut depth avg <= 0.2 mm; 2xStddev of delta rut depth <= 1.0 mm
Surf_CA	The standard deviation between the five runs should be within ±3mm
Surf_NL2	If the average rut depth (for the two runs) is ≤15mm, then at least 80% of the repeat measurements must be within 2mm of the original measurement. If the average is >15, then 80% must be within 0.1*(5+avg value)
Surf_SE	Correlation >=0.9 Systematic error 0.5 mm Within interval < 0.5 + 0.5*Rut
Surf_UK1	95% of the errors between the measured rut depths and the original rut depths fall within ±2.5mm. All errors are within ±10mm.
Surf_UK2	95% of the errors between the measured rut depths and the original rut depths fall within ±3mm. All errors are within ±10mm.

3.3.9 QA tests applied to Longitudinal Profile measurements

Only one specification (SE) includes a test for the accuracy of longitudinal profile within the QA regime. As with transverse profile, it could be expected that any issues with this measure would be apparent when assessing the ride quality parameters.

In this specification, the quality of the longitudinal profile data is checked by resurveying 5 % of a measurement area with a different device and different crew, within a month of the original measurement. The length of the road network in each measurement area varies between 200 and 1000km.

The repeatability of the data is also checked by re-surveying 40km out of every 1000km surveyed with the same device and crew.

For either the repeat survey data from a different or the same device, the correlation of the longitudinal profile data over 400m lengths is calculated and also the quota of standard deviations over each 400m length. The requirements for both accuracy and system repeatability are that 90% of correlations >0.8 and 90% of standard deviation quotas >0.8.

3.3.10 QA tests applied to Ride Quality parameters

Test procedures

Nine specifications test the ride quality parameters during QA (Table 52), with six including a description of how the tests are performed, with all data for the tests being provided by traffic speed surveys on the road network. Four specifications only require the system repeatability to be tested, whilst another (DE) only requires the accuracy to be tested. The remaining four test both accuracy and repeatability, although one (US) does not test repeatability directly – it is implied from testing the accuracy of repeat runs.

Reference data
Reference data for ride quality parameters is provided by either independent profilers or a different profilometer operated by the same or another contractor (SE) (Table 53).

Length of assessment
Six specifications (AU, AT, DE, UK1, UK2, US) assess the rutting parameters over the reporting length used, with three assessing over longer lengths than used for reporting (Table 53).

Assessing data accuracy
All of the specifications require all or a percentage of the measured ride quality parameters to lie within *a* set range of the reference parameter values (Table 50). Two of the specifications (DE, SE) apply further statistical tests to the data.

Assessing system repeatability
Five of the seven specifications have requirements for the difference between rutting parameter values delivered for repeat runs, with two (DE, SE) applying further statistical tests to the data. The two remaining specifications (AU, CA) apply only statistical analyses to the data (Table 55).

Table 52: Testing ride quality parameters

Specification	What do the tests entail?
Surf_AU	It is not specified that repeatability of data is tested outside of accreditation testing, however, the text "The bias check included in this method is used to determine whether there is a systematic drift in a profilometer's measurements over time" would suggest that this is an ongoing test and thus should be included in QA.
Surf_AT	Data accuracy is not tested within QA but repeat surveys are carried out on a ≥2km long section
Surf_CA	No details given, other than repeatability tests are only performed for larger surveys
Surf_DE	A minimum length of 2 km has to be measured by both contractor and BASt
Surf_NL2	No details given but requirements for system repeatability are given.
Surf_SE	Data accuracy is tested by letting a different device with a different crew resurvey at least 5 % of a measurement area. System repeatability is tested by surveys of the same lengths as for accuracy but using the same device
Surf_UK1	To test data accuracy, the Auditor will choose a route recently surveyed by the contractor, during routine network surveys, and survey that route with the reference device. For system repeatability, data is collected at regular intervals on a calibration site, on primary sites, on secondary sites and on daily sites.
Surf_UK2	To test data accuracy, the Auditor will choose a route recently surveyed by the contractor, during routine network surveys, and survey that route with the reference device. For system repeatability, data is collected at regular intervals on a calibration site, on primary sites, on secondary sites and on daily sites.
Surf_US	Reference data is collected on selected sections within the district that the survey contractor has recently surveyed. Repeatability is implied from the accuracy tests: 3 runs are completed for each site

Table 53: Reference data used for ride quality parameter QA and parameter assessment length

Specification	What is used for the reference (is there a "golden device")?	What reporting length is the parameter assessed over?
Surf_AU	N/A	100m
Surf_AT	N/A	50 m
Surf_CA	Class 1 profiler	500m (calculated from 50m averages)
Surf_DE	Golden device operated by BASt	100m
Surf_NL2	N/A	400m
Surf_SE	A different device with a different crew	400 m
Surf_UK1	Reference profilometer	10m
Surf_UK2	Reference profilometer	10m
Surf_US	The DOTD's "South Dakota" style laser profiler (inertial profiler) and/or a Class I profiling instrument.	0.1mile (160.9m)

Table 54: QA Accuracy requirements for ride quality parameters

Spec'n	What are the accuracy requirements for the data?
Surf_CA	The average values for each site should be within 10% of the reference value
Surf_DE	deltaAUN <0,3 or <0,6 depending on AUN value, sigmaAUN <0,6 or 1,0
Surf_SE	Correlation >=0.9 Systematic error 0.1 mm/m Within interval < 0.1 + 0.1*IRI
Surf_UK1	Three eLPV parameters are calculated: 3m, 10m and 30m LPV. For these: 95% of the errors between the measured LPV and the Reference eLPV fall within ±0.6mm^2 for 3m eLPV, ±0.7mm^2 for the 10m eLPV and ±0.9mm^2 for the 30m eLPV .
Surf_UK2	Two LPV parameters are calculated: 3m and 10m LPV. For these: 65% of the errors between the measured LPV and the Reference LPV fall within ±0.3mm^2 for 3m LPV and ±0.35mm^2 for the 10m LPV, with 95% within ±0.6mm^2 for 3m LPV and ±0.7mm^2 for the 10m LPV.
Surf_US	Maximum error of 5% bias or 20 inches/mile (0.32mm/m), whichever is less

Table 55: QA System repeatability requirements for ride quality parameters

Spec'n	What are the system repeatability requirements for the data?
Surf_AU	Determine the coefficient of variation (i.e. the standard deviation expressed as a percentage of the mean), Sn%, for each 100 m segment for each series of repeat measurements. The requirement is that when 95% of coefficient of variations are less than or equal to 5%. Determine the average of the segment coefficients of variation, S%. The requirement is that S% should be ≤3%. Using least squares regression, determine the coefficient of determination, r2, when the individual roughness values for each segment (dependent variable, y) are regressed against the mean values for that segment (independent variable, x).The requirement is that all of these values ≥ 0.95. Calculate the bias error between the comparison data set and the reference data set The requirement is that the bias error is ≤ 1%.

Spec'n	What are the system repeatability requirements for the data?
Surf_AT	delta IRI_avg <= 0.05; 2xstddev of delta IRI <= 0.5
Surf_CA	The standard deviation between the five runs should be within ±3mm
Surf_NL2	If the average IRI (for the two runs) is ≤1.5mm/m, then at least 80% of the repeat measurements must be within 0.25 of the original measurement. If the average is >1.5, then 80% must be within 0.1*(1+avg value)
Surf_SE	Correlation >=0.9 Systematic error 0.1 mm/m Within intervall < 0.1 + 0.1*IRI
Surf_UK1	Three eLPV parameters are calculated: 3m, 10m and 30m LPV. For these: 95% of the errors between the measured LPV and the Reference eLPV fall within ±0.6mm^2 for 3m eLPV, ±0.7mm^2 for the 10m eLPV and ±0.9mm^2 for the 30m eLPV .
Surf_UK2	Two LPV parameters are calculated: 3m and 10m LPV. For these: 65% of the errors between the measured LPV and the Reference LPV fall within ±0.3mm^2 for 3m LPV and ±0.35mm^2 for the 10m LPV, with 95% within ±0.6mm^2 for 3m LPV and ±0.7mm^2 for the 10m LPV.

3.3.11 QA tests applied to Surface Deterioration measurements

Test procedures

Of the eight specifications requiring delivery of a measure of surface deterioration (e.g. cracking), five have QA tests for these measures (Table 56). The tests are carried out at traffic speed on the road network. One specification (NL2) only tests the system repeatability, whilst another two just test accuracy (DE, US), with the remaining two testing both repeatability and accuracy.

Reference data

Only three specifications state how reference data is obtained for the QA tests. For two, manual analysis of downward facing images is used. This may possibly be the case for the third specification also but not enough detail was given to be able to confirm this (Table 57).

Length of assessment

Two of the specifications assess the surface deterioration parameters over the reporting length used, with two using longer lengths (Table 57). There are no details given for the other specification.

Assessing data accuracy

For two of the specifications, the requirement is that the amount of measured surface deterioration to lie within the same category, or a very close category to the reference (Table 58). One specification compares the amount of measured cracking with the reference.

Assessing system repeatability

Three of the five specifications require that repeat surveys are carried out on the network routes surveyed (Table 59). Two of the specifications require delivery of the percentage of the surface affected by surface deterioration and, for these, the requirement for system repeatability is that 65% of these repeat values lie within a set tolerance. For the other specification, the data is delivered as a condition category and the repeatability requirements are applied to these categories.

Table 56: Testing surface deterioration measurements

Specification	What do the tests entail?
Surf_DE	A minimum length of 2 km has to be measured by both contractor and BASt
Surf_NL2	No details given but there are requirements for repeatability of cracking data
Surf_UK1	For data accuracy test, the Auditor will choose a route recently surveyed by the contractor, during routine network surveys, and survey that route with the reference device. For system repeatability, data is collected at regular intervals on a calibration site, on primary sites, on secondary sites and on daily sites. The contractor compares the current data with data collected previously on the site.
Surf_UK2	For data accuracy test, the Auditor will choose a route recently surveyed by the contractor, during routine network surveys, and survey that route with the reference device. For system repeatability, data is collected at regular intervals on a calibration site, on primary sites, on secondary sites and on daily sites. The contractor compares the current data with data collected previously on the site.
Surf_US	For each district surveyed, the surface distress data is checked before the DOTD accepts it.

Table 57: Reference data used for surface deterioration parameter QA and parameter assessment length

Specification	What is used for the reference (is there a "golden device")?	What reporting length is the parameter assessed over?
Surf_DE	Golden device operated by BASt	100 m
Surf_NL2	N/A	400m
Surf_UK1	Manual analysis of digital images	10m
Surf_UK2	Manual analysis of digital images	50m
Surf_US	No details given	No details given

Table 58: Accuracy requirements for surface deterioration parameters

Spec'n	What are the accuracy requirements for the data?
Surf_DE	Delta longitudinal and transversal cracks total length < 0,5, sigma 0,5; same for corner breaks and edge damage
Surf_UK1	A high level of surface deterioration should be reported for at least 60% of the lengths where the reference has reported high levels. A moderate level of surface deterioration should be reported for at least 35% of the lengths where the reference has reported moderate levels. A low level of cracking should be reported for at least 70% of the lengths where the reference has reported low levels. Also, the % of area of non-Surface Deterioration features reported as Surface Deterioration must be less than 20%.
Surf_UK2	A high level of cracking should be reported for at least 65% of the lengths where the reference has reported high levels. A low level of cracking should be reported for at least 85% of the lengths where the reference has reported low levels. The overall level of agreement between the level of cracking, for low, moderate and high cracking combined, is at least 75%.
Surf_US	No details given

Table 59: System repeatability requirements for surface deterioration parameters

Spec'n	What are the system repeatability requirements for the data?
Surf_NL2	On **bituminous** pavements: 90% of the assessed road sections must meet the following requirements: • The difference between the "volume" category between two runs must not exceed 1. 90% of the assessed road sections must meet the requirement below, whereby the average of the volume of the transverse cracks in the production measurements and repeat measurements on the relevant section is called "DWSm": • The difference in volume between the number of metres of transverse cracks of the production measurements and the repeat measurements may not exceed 20% of "DWSm" plus 20 m • If the "severity" categories of the two runs are different then, the difference between the "volume" category between the two runs must not exceed 1.4 • The aggregated volume of light, moderate and severe cracking for the two runs must be within 20m. On **concrete** pavements: 90% of the assessed road sections must meet the following requirements: • The difference between the "volume" category for the two runs must not exceed 1.4 • If the "severity" category of the two runs differs then the difference between the "volume" categories must not exceed 1.4.
Surf_UK1	For primary checks, 65% of the difference in the reported levels of surface deterioration between the two survey runs must lie within 0.28
Surf_UK2	For the calibration and primary sites, the difference in the reported levels of cracking (reported as a %) between the two survey runs, is calculated. 65% of these differences should be within ±0.1

3.3.12 *QA tests applied to Images*

Of the four specifications requiring delivery of downward facing images, only two apply QA tests on image quality (Table 60). Tests of images are also required if used to obtain cracking data, for another specification, for which images are not a deliverable (UK2).

One specification (UK1) requires the contractor to quantitatively assess images collected throughout the survey year on the same site for focus, illumination, contrast. The specification requires that these remain consistent. Another (UK2) requires subjective assessment of the images, whilst the last specification (US) requires a daily measurement of the footprint of the image to ensure that it stays the same throughout the duration of the contract.

Table 60: Testing images

Specification	What do the tests entail?
Surf_UK1	The first time that a Primary Site is surveyed in each survey year, the Consultant will provide the Downward Facing Images collected on the Downward Facing Image Reference Length for that site to the Auditor. The Auditor will carry out a manual assessment of the images to confirm their quality (similar to Accreditation tests). If necessary, the Consultant will be requested to re-survey the site until images of good quality are delivered. The first time that a primary Site the site is surveyed the distribution of greyscale values of the images collected on the Downward Facing Image Reference Length will be calculated, along with the mean value and standard deviation of values, to obtain the reference intensity distribution for the Primary Check Site. For each subsequent Primary Check the Consultant will obtain the distribution of greyscale values found in the images over the Downward Facing Image Reference Length, and the mean greyscale value and standard deviation of greyscale values, and will ensure that the requirements for consistency of illumination (used in Accreditation) are met. For each subsequent Primary Check the Consultant will also ensure that the requirements for image contrast and focus are met via “Network” tests. A procedure is carried out by the Consultant that enables the operator of the Equipment to carry out Secondary or Daily checks on the downward facing images. The procedure for Secondary checks includes an element of quantitative assessment of image quality. The Consultant is expected to demonstrate how their procedure will ensure consistency in the collection of downward facing images to the Auditor. Approval by the Auditor is required before the method is deemed acceptable for use as a method for carrying out Secondary or Daily Checks.
Surf_UK2	The survey contractor is required to perform weekly and monthly checks on the system recording the cracking e.g., if they are using an image system, they are required to check that lighting levels and evenness are acceptable, that the images collected are of high enough quality and resolution for cracks to be detected, and to check that the crack detection software is working correctly.
Surf_US	The Consultant's pavement view camera image shall be measured with a visual measurement instrument (supplied by the Consultant) under DOTD supervision to determine the actual footprint (length and width) of the image; this shall be performed for each Data Collection Vehicle prior to data collection. That footprint image must be maintained for the duration of the contract. The Consultant shall be required to verify daily that the footprint is the same as the previous day. Such verification shall be documented (i.e., results from tests are recorded and any corrective action taken explained in detail) and reported to DOTD on a weekly basis and summarized in the final report.

3.4 Review of current structural condition equipment testing

Only one specification for Accreditation and QA of high-speed structural survey equipment was identified: The TRASS survey, used to define TSD surveys in the UK.

3.4.1 Accreditation

There is only one TSD device available in the UK, and, unlike the survey equipment used for surface condition surveys, the TSD is owned by the Highways Agency (the survey commissioner), not the survey contractor. Thus, the Accreditation and QA regimes include tests for the survey contractor’s ability to use, maintain and operate the TSD in a way that will produce consistent data, equipment examinations, as well tests for the data. A Technical Advisor is used to provide support to the survey contractor in understanding the TSD and an independent Auditor used to ensure the quality of the data delivered.

Before undertaking the Survey, the Technical Advisor provides training for the survey contractor in driving the equipment, operation of the equipment, the operation of the equipment processing/post processing software and in troubleshooting equipment technical problems. Once this training has taken place, the Consultant is approved as an Accredited Equipment Operator but the Auditor can undertake checks from time to time to ensure the ongoing competency of the survey team.

At regular points during the contract, the TSD is then subject to an equipment examination, carried out by the Technical Advisor. This examination checks the visual condition of the vehicle, the visual condition of the measurement systems, the operational status of the vehicle and the operational status of the measurement systems. This is used to monitor any damage and/or deterioration to the equipment.

Before undertaking surveys, and during the survey year, the equipment is subject to Accreditation tests, which are overseen by the Auditor. This includes checking of equipment performance via tests on one or more test tracks and/or appropriate local road networks. The performance of the equipment is assessed against reference data, provided either by a reference device or collected during previous TSD measurements collected on the same site. Table 61 summarises the accreditation data to be tested and the assessment criteria.
The Accreditation is valid for one month.

Table 61: Summary of UK TSD accreditation tests

Parameter		**Values to be compared with reference**	**Reference method (Site tests)**	**Tolerance, 95% limits**	
				Site tests	**Network tests[2]**
Location referencing	Distance travelled	Measured lengths of Sections	Calibrated measuring wheel and Steel Tape	±1m or ±0.1% of the section length[1]	±3m or ±0.1% of the section length[1]
	National grid coordinates of section start points	National Grid coordinates of section start points	Static GPS and/or optical survey	3m	4m
	National grid coordinates of positions of	National Grid coordinates of positions of	Static GPS and/or optical survey	4m	6m

	moving vehicle	moving vehicle			
Longitudinal profile	Enhanced LPV	3m, 10m eLPV	ARRB Walking profiler and/or previous accredited equipment survey	±0.6mm^2 (3m), ±0.7mm^2 (10m)	±0.6mm^2 (3m), ±0.7mm^2 (10m)
Deflection slope	Deflection slope	Deflection slope	Previous accredited equipment survey	±0.050	±0.050

1. Whichever is the larger
2. The reference for the Network tests shall be a survey carried out using the Accredited equipment

3.4.2 Quality Assurance

The UK approach to QA testing of TSD surveys is based on four levels of checking using three types of site:

- Checks on Primary Site
- Checks on Secondary Sites
- Daily checks.

The survey contractor establishes the required sites (Table 62). The sites are then surveyed within a set time period and at regular intervals throughout the period of surveying (Table 63). The survey contractor checks the data to ensure it meets the requirements given in Table 64 and provides the results to the Auditor for independent checking.

Table 62: QA test site description

Site type	Minimum length	Minimum number of sites	Characteristics
Primary	10-20km	1	The majority of the site shall be without sharp bends or extremes of profile unevenness. It shall contain features that are representative of the survey network.
Secondary	N/A	As provided by the Auditor	N/A – network site
Daily	400m	As required	Reasonably flat and without sharp bends or extremes of profile unevenness. Free from isolated surface defects.

Table 63: QA test site requirements

Site type	Location referencing method	When should reference data be collected?	Frequency of testing
Primary	Grid Reference Co-ordinates will be established for the start and the end of the site and all Section start points within the site. The distance shall be measured between each Section Start Point.	No more than 7 days following an Accreditation or re-test	Three surveys every week
Secondary	Grid Reference Co-ordinates will be established for the start and the end of the site. The length of Site shall be measured.	Auditor extracts from Survey Data as appropriate	Whenever the contractor is surveying nearby
Daily	Grid Reference Co-ordinates will be established for the start and the end of	Start/End of the survey day	Surveyed in the same direction both at the

	the site. The length of Site shall be measured.		start and the end of the day.

Table 64: QA test site data requirements

Measured Parameter	Tolerance for Primary Site	Tolerance for Secondary and Daily Sites	Tolerance All checks (maximum)
Location referencing – length of each section	±3m or ±0.1%, whichever is greater	±6m or ±0.1%, whichever is greater	N/A
Location referencing – National Grid coordinate	4m	7.5m	N/A
Measured Parameter	Tolerance - Primary Site (90% limits)	Tolerance - Secondary and Daily Sites (75% limits)	Tolerance All checks (maximum)
Location referencing – National Grid coordinate	±6m	7.5m	20m
Longitudinal profile – 3m eLPV	±0.6mm^2	±0.6mm^2	N/A
Longitudinal profile – 10m eLPV	±0.7mm^2	±0.7mm^2	N/A
Deflection slope	±0.050	±0.050	±0.200

4 Definitions

The following subsections list the technical terms to be used, along with the definitions of the terms as they will be used within the HiSPEQ project.

Accreditation
Accreditation is a process that is usually implemented at the start of a survey regime. The aim of the process is to demonstrate whether a high-speed survey device is capable of delivering the data required by the survey, and to the level of accuracy required.

Accuracy
A system's accuracy is how closely its measured data reproduces reference data.

Derived Parameter
A parameter, such as rutting, IRI or area of cracking, calculated from the measured raw profile or image data.

Filter
In the context of this work a Filter can be considered to be a mathematical transformation applied to measured profile data to remove features (wavelengths) considered undesirable in the calculation of a derived parameter.

Fleet consistency
Fleet consistency is how closely matched data is, when collected by multiple (different) devices surveying the same route.

Golden device
A golden device is an ideal example of a device against which all later devices are tested and judged. The term "golden" is used to describe the precision of the device to standard specifications.

IRI
International Roughness Index, a parameter calculated from longitudinal profile data

Longitudinal Profile
The longitudinal profile is a measure of the shape of a road surface, in a single line parallel with the direction of travel. It is usually measured in the wheel paths and sometimes measured in the middle of the lane.

Quality Assurance (QA)
Quality Assurance (QA) is the process that is implemented during the course of a survey regime, to ensure that the data quality has remained at an acceptable level.

Ride quality
Ride quality is indicated by a parameter or parameters that are derived from the longitudinal profile. These parameters attempt to quantify the level of comfort or discomfort that road users will experience when driving down the road.

Rutting
Rutting is the permanent deformation of pavement layers which can accumulate over time. It is limited to asphalt roads, and can be indicative of pavement failure. There are two types of rutting that can develop on a road: Surface course rutting and structural rutting. Surface course rutting only occurs in the top ~50mm of the pavement and is caused by the surface course mixture being displaced by vehicle wheels, usually during hot weather. Structural rutting is the result of excessive consolidation of the pavement along the wheel path due to either reduction of the air voids in the surface layers, or the permanent deformation of the

base or subgrade. It is this type of rutting that causes most concern to road engineers, since it is most indicative of pavement failure.

Surface deterioration
In general surface deterioration includes any deterioration in the condition of the road surface, for example, cracking, fretting/ravelling, pot holes, fatting up, bleeding. Since the HiSPEQ project is concerned with measuring the durability of a pavement, we will only be considering defects that either effect the structural integrity of the surface layers, or might allow water ingress into the lower layers of the pavement i.e. cracking and pot holes (fretting tends to develop into cracking or potholes when it gets severe enough to effect the structural integrity of the surface layers).

System repeatability
The repeatability of a system is how precise an individual device is i.e. how closely matched data is when collected by the device during two separate surveys.

Transverse Profile
The transverse profile is a measure of the shape of a road surface, in a single line running across the lane, perpendicular to vehicle travel.

5 Acknowledgements

The research presented in this report/paper/deliverable was carried out as part of the CEDR Transnational Road Research Programme Call 2013. The funding for the research was provided by the national road administrations of Denmark, Germany, Ireland, Netherlands, UK and Slovenia.

6 References

ISO 13473-1:2004: “Characterization of pavement texture by use of surface profiles. Part 1: Determination of Mean Profile Depth”.