How do you compare? Correlation and calibration of skid resistance and road surface friction measurement devices

Authors: Roe, P G[1], and Sinhal, R[2]

[1]*TRL Limited, Crowthorne, UK*
[2]*Highways Agency, UK*

ABSTRACT

This paper discusses the issues relating to the calibration and comparison of skid resistance measurement devices, particularly in fleet operation. It reviews UK experience of over fifteen years of controlling a fleet of SCRIM machines in use providing survey data for comparison with skid resistance standards on the national road network. This includes the approach taken to setting and applying acceptance criteria, and briefly compares this with the approach taken in some other European countries in relation to the devices that they use. The paper will also discuss the problems of correlating devices that operate on different principles, drawing on UK experience with GripTester and the Pavement Friction Tester. The paper will also discuss the practical experience gained from the recent HERMES project in Europe conducted by FEHRL (Forum of European National Highway Research Laboratories) that evaluated a proposed standard process for harmonisation of friction measurement devices.

1. INTRODUCTION

Skid resistance measurement began in the UK with the development of the original side-force device based upon a motorcycle sidecar combination. Several of these devices were built in the 1930s and used in early work to assess skid resistance on British roads. The concept was developed into a fifth wheel mounted inside a front-wheel drive car and by the late 1960s, the then Road Research Laboratory (later to become TRL) used the same equipment fitted in more modern vehicles with a small internal water tank. As well as side-way force devices, various other machines were developed, including a small-wheeled braking force trailer that was used to assess skid resistance at higher speeds (Figure 1) and a machine for measuring brake-force coefficient for truck tyres.

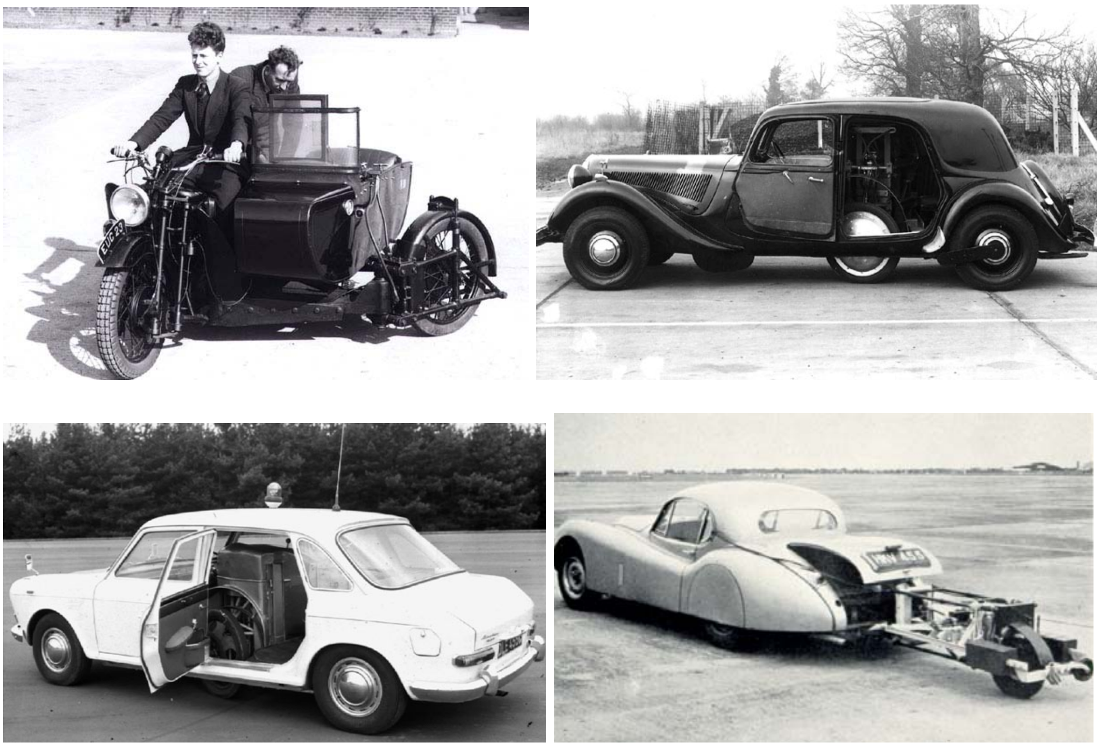

Figure 1 – Early UK skid resistance measurement devices (1930s to 1960s)

All of these devices had their uses in research and much valuable knowledge was gained by studying roads using them. All, however, needed an external water tanker to wet the road, especially if measurements were to be made over long distances (or for repeated runs over shorter distances) and all used analogue technology such as chart recorders to record the data.

In the late 1960s, it became clear that if skid resistance was to be taken seriously, it would need to be monitored on a regular basis and equipment was needed that would make this possible. This led to the development of SCRIM (Sideway-Force Coefficient Routine Investigation Machine) as a tool for monitoring skid resistance on a network scale (Hosking and Woodford, 1976). The side-force coefficient principle was adopted, to allow for continuous skid resistance measurement, and this was combined with a large-capacity on-board water supply and electronic data recording, initially using a simple printer and, later, punched paper-tape.

The prototype machine was evaluated alongside the older side-force cars. A license to build machines commercially was granted to WDM Limited by TRL (then TRRL) and the first production machines were built in 1971, two for TRL and one for the UK Department of Transport. In the early 1970s further machines were purchased by a number of UK Local Authorities, and WDM Limited themselves built machines that could be made available to provide a survey service. With a greater number of machines in service and more widespread use, the importance of regular checking of the equipment became apparent.

This paper discussed the use of correlation trials in the UK and the experience and issues arising from them. It then explores some wider issues relating to use of different devices and fleets, including the potential for harmonisation and the possibilities for a more consistent approach

2. BRIEF HISTORY OF THE UK SCRIM CORRELATION TRIALS

2.1 THE EARLY YEARS

The UK SCRIM Correlation trials have evolved from an informal arrangement in the 1970s and 1980s through to the regular “compulsory” exercise that is now held in April each year.

Initially, a “SCRIM Users Group” was formed, under the auspices of the County Surveyors Society (CSS), as a forum to discuss experience and develop ideas. The Department of Transport, through its Road Surface and Strength Testing Unit (RSSTU) which was based at the TRRL site in Crowthorne, worked closely with TRRL by bringing the machines together periodically for assessment.

During this period, much experience was gained and shared regarding the way in which SCRIM should be used and its operation. This, in turn, led to a number of incremental improvements to the design of the equipment to make it more reliable and robust in service. However, when, in 1987, the then Department of Transport planned the introduction of standards for skid resistance on in-service roads, it became a requirement that the network would be monitored using SCRIM and that the SCRIM fleet would need to be cross-checked and approved before machines could be used in this role.

This led to the introduction of formal correlation trials that all machines would need to attend and “pass” every year. At first these were arranged and co-ordinated by the RSSTU. The trials took place over a week during which time each machine underwent a detailed inspection of all aspects of its set-up and operation. Experience of these exercises led to the development of simple equipment and procedures for static calibration of the machines that were eventually owned and used by all operators and which are still used today.

Those machines that had satisfactorily passed the inspection stages went forward to a programme of running trials. In these, the fleet of machines carried out a number of repeat passes on a range of surfaces on the TRL track and on local roads. In the light of the results of the running tests, individual machines were approved for use on the trunk road network. An important emphasis at the trials at that time was that no machine should leave without passing, so much effort was spent in trying to identify and rectify problems. However, the criteria by which the machines were actually judged were not formally defined.

2.2 THE PRESENT TRIAL FORMAT

In 1995, re-organisation within central government led to the disbanding of the old RSSTU and the responsibility for organising and running the correlation trials was transferred to TRL, working under contract to what was now the Highways Agency. The HA remained the organisation approving the use of SCRIMS, in the light of their performance at the trials.

At first, the same broad approach to the trials as had been was taken but, because the skidding standards were now well established and the ability to satisfy large commercial survey contracts depended upon the outcomes of the trial, it was decided that the exercise should become more focussed and carried out as efficiently as possible. As a result the trials were re-structured to take a general form which has been refined but has remained essentially similar for the last nine years.

The trial is planned to take place during one week in mid-April, with machines usually arriving direct from their winter service for checking prior to the start of the UK test season, which runs from 1 May to 30 September. Occasionally, there may be additional trials at the end of the test season. Specific tasks are carried out on each day:

- **Day 1** is reserved for the TRL team to prepare the test track, sweeping the test sections and marking out the test lines. If it has not been done earlier, this day may include runs with the TRL SCRIM to pre-condition the tyres that are to be used in the main exercise.

- **Day 2** is the day on which machines arrive and undergo initial static inspections. An inspection checklist covering most general features of the machine is provided for crews to complete before arrival, completion of which is a pre-requisite of joining in the trial. Inspections by TRL on the arrival day are confined to checks on static wheel weight and water flow rates and direction. Each machine has its own rig to carry out static calibrations of the side-force measurement. While the other inspections are being carried out, the load-cell that is attached to each rig to measure the applied force during the calibration sequence is itself separately calibrated, by measuring the electrical output response from known applied loads using calibrated static weights.

- **Day 3** is the main running trial day. All machines carry out a sequence of passes over the test sections. An initial run by all machines in convoy allows crews to familiarise themselves with the layout of the test sections, warms-up the equipment and conditions the track ready for the main runs. This is followed by a static calibration check by all crews on their own machines. The main test runs then begin. Machines take turns to run with a number of “standard” test tyres. These have been chosen ahead of the trial from the same batch and with resilience as close as practical to the centre of the permitted tolerance range. The

machines operate in pre-planned, randomised groups until each machine makes three replicate passes with each of at least three of the “standard” tyres. The TRL SCRIM makes additional passes with a “control” tyre to monitor the general condition of the test track through the day.
Data from each group of runs is processed in real time so that early warning of machines that may be failing is given and steps can be taken to identify and rectify any problems.

- **Day 4** is used for any additional runs that are needed, usually to cross-check machines that have had difficulties on the previous day and have needed attention. Two or three other machines may be retained to provide cross-referencing with the previous day for these tests. Additional tests to provide extra data or address other matters of interest may be incorporated into this day.

- **Day 5** is usually available as a contingency against extremes of weather disrupting the programme and for final clearing up after the exercise.

On completion of the running trials it is usually possible to tell the crews whether they have “passed” the trial or not. Highways Agency will subsequently formally advise the operating organisation of their acceptance to work on the trunk road network. Occasionally, machines may be unable to attend or may “fail” the trial due to a problem that cannot be rectified on the day. In such cases, a re-test at a mini-trial may be necessary, organised at the expense of the operators involved.

3. THE ASSESSMENT CRITERIA

3.1 GENERAL PRINCIPLES

In trials of this type, it is important that the reasons for carrying them out are clear and that the criteria that are to be used to make any judgments are both well established and transparent to all who need to know. This is made easier in the case of the UK trials because of their long history, in which a spirit of co-operation among the various operating organisations has always been maintained.

It is not possible to define an absolute value for skid resistance. Rather, at any particular time, the “correct” result can only be estimated and arguably the best estimate for any particular type of measurement device would be the average value given by all machines of that type. For UK SCRIMs, that would be the average of all acceptable machines in the UK fleet.

It is recognised that, as with any measurement system, there will inevitably be small differences between devices and that there are a great many factors that affect skid resistance measurements in the field. Acceptance criteria, therefore, must take this into account, providing a definition of what the value for comparison will be and, recognising that there will be inevitable variations, what the tolerances should be.

The purpose of the UK correlation trials is to give confidence that the fleet provides results that are as consistent as possible, with individual machines as close as is practical to the overall mean. On this basis, the result from any one machine on any particular occasion can be assumed to be representative of the fleet mean. The acceptance criteria must be defined in order to achieve this. Therefore, the general principle that has been applied in the UK is that the total variance in readings is limited so that when the machines are used to compare measured skid resistance with Investigatory Level in support of the Trunk Road Skidding Standards, there is only an acceptable risk of incorrect classification of a site from a single run.

3.1.1 Gathering appropriate data

The purpose of the trials is to assess the SCRIM fleet as a whole and the performance of individual machines within the fleet. The first function of the trials, therefore, is to collect data that can provide appropriate information. The trial format is therefore arranged so that it is the performance of the machines themselves that is assessed. To achieve this, as many factors as possible are controlled so that their effects are minimised, or at least randomised, so that all machines are similarly affected. Strategies used to achieve this include:

- All machines make a similar number of tests.
- All machines test the same surfaces
- Surfaces are chosen to test the machine over a range of skid resistance levels.
- All machines carry out three repeat tests on each surface in any one set of measurements.
- All machines make at least three sets of measurements.
- All machines use the same test tyres (or a subset of at least three from a pool of four or five standard tyres).
- Running order is randomised during the day and an individual machine’s measurements are spread through the day in case track conditions change.
- All machines operate at a constant speed (50km/h).
- The test line on each surface is clearly identified and the path followed by individual drivers is audited from time to time during the day.
- Data are processed independently by the TRL team – no processing by the machine crews is necessary, although they will advise the team of any out-of-course events that may have influenced the results.

3.1.2 Reference sections

As commented above, it is not possible to establish an absolute measurement of skid resistance against which the machines can be judged. For this exercise, therefore, a set of “reference sections” are defined and the performance of the SCRIM fleet is judged on the basis of the average measurements that they make on these surfaces.

The overall mean SCRIM Reading for all the standard tyres on the reference sections is used as a reference level for each machine and the grand mean of all machines provides the reference level for the fleet.

Historically, the “reference sections” were a group of samples from a 0.5km length of hot-rolled asphalt surfacing. This type of surfacing was typical of most of the trunk road network at that time. However, because the surfacings on the TRL track were not trafficked, the general level of skid resistance was relatively high. Other sections were included in the trials that covered a wider range of skid resistance and these were used for comparative purposes only.

Changes to the TRL track in recent years have meant that the full length of the historic sections reference sections no longer available. The trials now use a different set of sections that include a range of surfacings with an average skid resistance level closer to the typical level found on much of the network. Again, additional sections provide a further range of surfacings for additional comparisons.

One section has very low microtexture and virtually no texture depth. although it is not included in the “reference sections” set, the section is tested by the machines in order to verify that they respond to the very low skid resistance, which is markedly below anything normally encountered on the network. However, at such a low skid resistance level, small differences between measurements appear relatively much greater than on the other surfaces. Also, physical factors, such as hydraulic drag, that do not occur on the textured surfaces, can sometimes influence the results.

It is important to stress that the reference sections are simply those that provide the average skid resistance level against which the machines are compared. They are not reference surfacings in an absolute sense and the measured values typically cary from year to year as a result of seasonal variation. In fact, the average value may vary from day to day during the trials in certain weather conditions. However, these sections do provide a general indication that the overall level of skid resistance recorded from year to year is broadly the same.

3.2 SPECIFIC CRITERIA

3.2.1 Machine repeatability

Individual machines must be able to provide consistent results from one run to the next when nothing else has changed. There will inevitably be random variations between runs, partly due to the machine itself and partly due to the road, particularly when the driver deviates from the ideal test line. Therefore, the first criterion that is used to assess the machines provides a check on the repeatability of each machine and also provided a means to identify outlying groups of measurements that might occur on individual test runs during the trial. At the trials, units of SCRIM Reading (SR) are used. This is the value recorded by the machine every ten metres and is equivalent to the sideway-force coefficient (SFC) multiplied by 100.

- *The between-run standard deviation on any individual section for any individual machine and tyre should be 3.0 or less.*

Where this figure is exceeded, the data are examined for evidence of variation in test line and if necessary individual runs on that section may be excluded from subsequent analysis. If a machine consistently records data with unacceptable between-run variation, the operation of the machine is regarded as unacceptable. If necessary, the test line that is followed by the driver(s) will be reviewed and consideration given to making repeat runs to bring this factor under control.

3.2.2 Overall fleet variability

It is important that the variability of the fleet is as small as is practical to minimise the risk of wrongly classifying a length of road when comparing measurements with Investigatory Levels. This criterion provides a check on the reproducibility of the fleet, so that each approved machine gives results consistent with the rest of the fleet during normal surveys.

- *The maximum standard deviation between the overall machine means is 2.6*

The standard deviation will be influenced by any outlying machines. If necessary, outlying machines will be rejected in order to reduce the standard deviation to an acceptable level.

When the between-machine standard deviation exceeds the maximum permitted level it will be necessary to identify outlying machines.

- *Any machine that is 7.8 or more (i.e. three times a standard deviation of 2.6) from the all-machine mean will be rejected outright.*

- *Any machine that is between 5.2 and 7.8 (i.e. between two and three times a standard deviation of 2.6) from the mean will be subject to further investigation in the context of the overall distribution and performance on the full range of surfacings.*

These statistical criteria are based on probability levels which mean, for example, that there is a one in twenty chance of two standard deviations being exceed simply due to random variation. With fewer than twenty machines in the fleet, it is possible that one of them may be identified as an outlier due to random effects rather than a systematic error. This is even more likely to occur when the variation between most machines is small. Therefore, if the between-machine standard deviation is markedly below the required criterion, then the two or three-standard deviation rules may be relaxed so that machines are not unfairly rejected due to random errors in a small population.

4. EXPERIENCE AT THE TRIALS

4.1 TRIAL LOGISISTICS

The trial logistics are now well established with the result that most crews and the key members of the TRL team are familiar with what happens. Prior to the running trials each crew is given a diagrammatic timetable that shows the sequence of the groups of runs which they will be joining, which tyre they are expecting to use, their position in the running order and approximately when the run should take place. Thus the teams are ready to collect their tyre from the last machine to be using it (Figure 2, Figure 3) and the gap between a wheel finishing a sequence of runs on one machine and starting on the next is short.

It has been found to be valuable to have one marshal directly responsible for coordinating the test runs, making sure that each machine is ready at the right time and to make decisions for dealing with out-of-course events. Other aspects, such as making provision for machines to refill with water without disrupting the programme are also planned in advance and are the responsibility of the track marshal.

The track marshal usually is supported by a senior member of the team who maintains an overview of progress and the emerging results, ready to take strategic decisions if machines are appear to be failing or the trial needs to be suspended during bad weather.

Figure 2 Machines and crews wait for their turn to run with a “standard” tyre at the 2004 trial

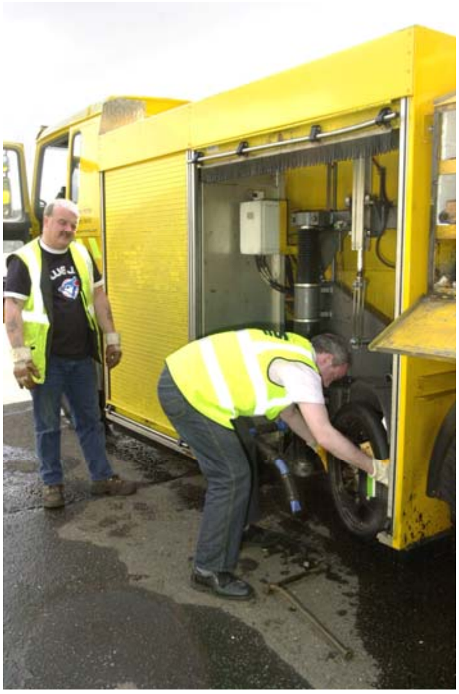

Figure 3 Fitting a test wheel ready to run

As a result of these preparations, the testing normally proceeds smoothly. Inevitably, there will be occasions when problems arise, from trivial causes, such as when an operator forgets to start his recorder and therefore needs to repeat a pass, or a more serious matter such as a fault in the recorder electronics or a major mechanical failure on one of the trucks.

4.2 REVIEWING THE DATA AND APPLYING THE CRITERIA

The statistical approach to assessing machines at the trials has now been in regular use for several years. Generally, it works well, but there are occasions when difficulties may arise. The greatest problems occur when there are two or more machines that are not performing satisfactorily but are on opposite sides of the distribution – reading "high" or "low" in respect to the overall mean.

In these circumstances the deviations from the mean may be such that it is obvious that both are at fault. However, the situation can also occur in which rejecting either machine brings the overall standard deviation within the required limits, but there is no obvious reason to choose either. It is in these situations that the performance on individual sections as well as the overall average on the reference sections must be brought into consideration.

An example of a straightforward situation is illustrated in Figure 4. This is a summary graph showing the average results on the different test sections from the first running day of a recent trial. The set of "reference sections" includes all sections except 24 and 26. In this graph, Machine 1 is a clear outlier. In this case, a mechanical problem was identified and a replacement part fitted which led to follow-up measurements on the reserve day that confirmed that the affected machine was performing acceptably.

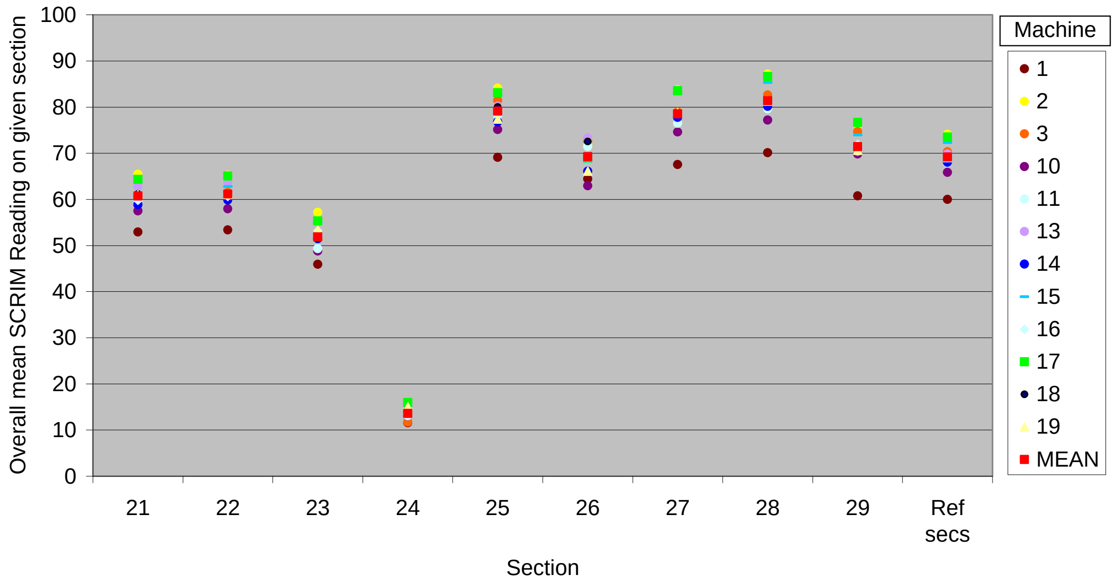

Figure 4 Summary of results from first running day at a recent UK trial

4.3 OTHER EXPERIENCE

A secondary role for the correlation trials is to identify aspects of the operation of SCRIM that would benefit from further improvements. Similarly, practice at the trials is also improved and this can bring advantages. A good example of this is the way in which static vertical load is measured. Historically, this was achieved by "suspending" the test wheel assembly from a cradle attached to a load-cell to measure the load. This worked, but was difficult to carry out routinely and had the disadvantage that the load was not measured with the test wheel fully lowered to the ground.

Taking advantage of improving technology, the process at the trials was changed to use an electronic weighpad under the test wheel so that the vertical load could be measured directly under the tyre. To use this, the SCRIMs were driven on to lifting ramps that enabled the top of the weighpad to be at the normal road level. This revealed some variations in static weight that had previously been undetected and, combined with new developments to SCRIM described in another paper at this conference (Roe and Sinhal, 2005), eventually enabled the cause of the problem to be identified and eliminated.

5. CORRELATION WITH OTHER DEVICES IN UK

Although SCRIM is the standard device for monitoring skid resistance of the trunk road network and many local roads, other devices are also used. The GripTester is used by some local authorities, either for network monitoring or for investigative purposes and the potential use of the device for its purposes is currently being assessed by the Highways Agency. The Pavement Friction Tester, owned by the Highways Agency and operated by TRL on their behalf, is used primarily as a research tool, particularly where locked-wheel friction measurements at different speeds are required.

There are no formal arrangements to correlate these devices but nevertheless, some attempts at comparison and correlation have been made over the years.

5.1 GRIPTESTER

Activities with GripTester in the UK have been co-ordinated through a user group, initially linked with the County Surveyors Society, reflecting the fact that most owners were local authorities. In the early and mid-1990s the GripTester User Group organised occasional correlation trials. These served a similar purpose to the early SCRIM trials in that they provided a means for identifying areas for improvements and developments to the machines, with a particular focus on their reproducibility.

Associated with these trials have been attempts to correlate the device with SCRIM. Trials in the early 1990s in, which four SCRIMs were compared with four GripTesters on the SCRIM trials test sections, demonstrated that there was a strong linear correlation between the GripTester and SCRIM for testing in a straight line at a constant speed (Figure 5).

However, although the correlation was good, there was inevitably some scatter in the measurements. Since that time the GripTester has undergone further development. The standard tyre has changed and there have been technical changes to improve its precision and its performance in particular situations such as on curves. An informal trial in 2004 showed that there is still a reasonable correlation between the new versions of the device and present SCRIM.

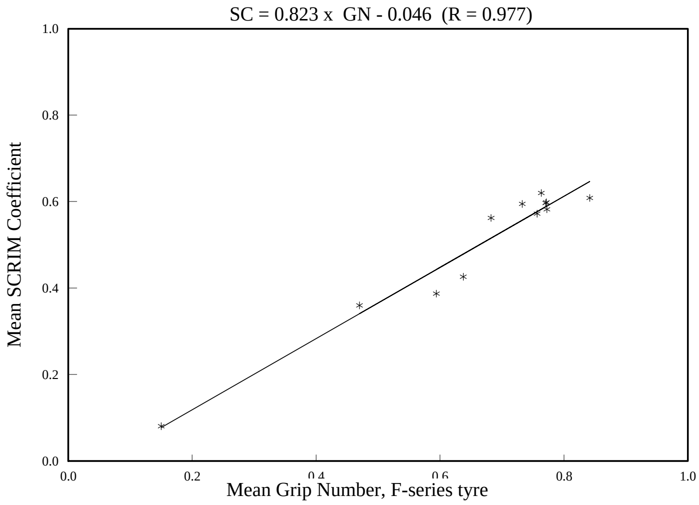

Figure 5 Comparison of SCRIM in GripTester in the 1990s

5.2 PAVEMENT FRICTION TESTER

The Pavement Friction Tester (PFT, Figure 6), a version of the ASTM Friction Trailer was used in a major UK study of the effects of texture depth and speed on skid resistance (Roe, Parry and Viner, 1998). As part of that work the device worked alongside SCRIM on a wide range of sites. It was found that there was a reasonable correlation between the measurements made with the PFT at 20km/h and those made by SCRIM at 50km/h, when the equivalent slip-speed of the SCRIM tyre is approximately 17km/h (Figure 7). Again, as had been found with the GripTester, although a linear relationship could be deduced, there was wide scatter that would limit the value of applying a generalised correlation equation in a specific situation.

Figure 6 The Pavement Friction Tester

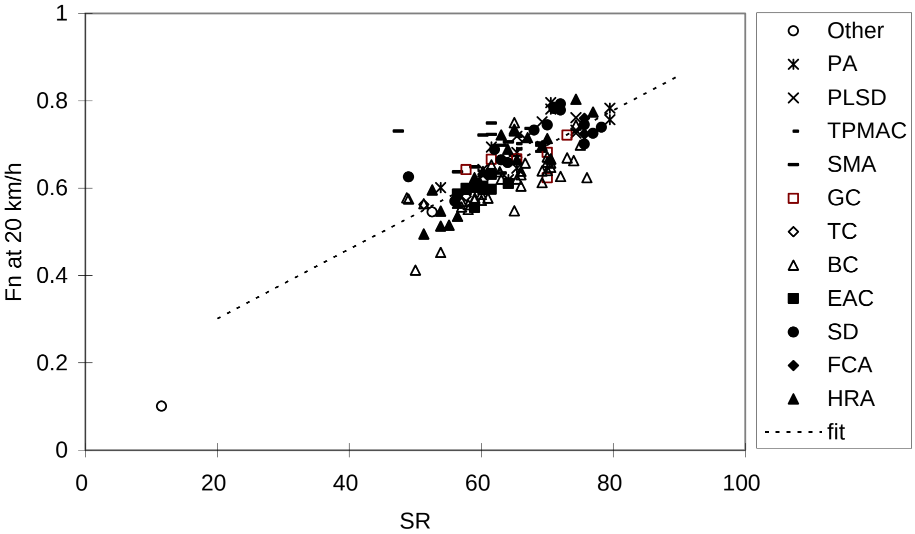

Figure 7 Comparison of measurements from PFT at 20km/h with SCRIM at 50km/h on a range of surfacing types (after TRL 367)

6. THE PROBLEM OF HARMONISATION

So far, this paper has concentrated on the issue of controlling the consistency of a relatively large fleet of skid resistance measurement devices *of the same type*. Experience in the UK has shown that close control over the devices can be maintained but some variations between machines are inevitable. SCRIM, however, is only one of many different types of device used around the world to measure skid resistance, and there is inevitably a desire to find a means of harmonising the measurements from the different devices that will allow results and standards developed with one to be compared with results and standards with another.

A major difficulty is that, although all skid resistance testers use the same basic idea – slide some rubber on a wet road and measure the forces or work done – the implementations are many. There are three basic principles:

- Rubber slider
- Angled test wheel (side force)
- Braked in-line test wheel (brake force)

However, each of these has different variants such as: pendulum or rotating sliders; different wheel angles, loads and tyre sizes; locked-wheel, fixed-slip or variable slip. Add to these the other test methods to assess road surface friction, such as drag sleds and decelerometers, used by police forces and the picture becomes even more complex.

All of these devices respond differently to the road surface depending upon the conditions in which they are used and particular features of the road, especially the texture depth.

A major attempt at harmonisation was the PIARC International Experiment in 1992 from which the concept of the International Friction Index (IFI) emerged. A major weakness, however, was the difficulty of having sufficient confidence in precision of the results when compared through the IFI to make them of practical use. In Europe, these concerns have subsequently led to a proposal for a harmonised index based on a similar principle but focussed on the specific devices used in Europe.

This so-called European Friction Index, and means of calibrating different devices to it, has recently been evaluated in a major study carried out through FEHRL (Forum of European National Highway Research Laboratories); the full final report of the HERMES experiment, as it was known, has yet to be published. However, experience of the work showed that while series of calibration tests could be arranged, there are still significant obstacles to be overcome before measurements from different devices can be used interchangeably with similar confidence to measurements from different members of a single fleet.

7. CLOSING DISCUSSION

When one device is to be used to make a measurement, whatever answer it gives is assumed to be “correct”, by definition, although there will be a certain amount of possibly unknown error. The main issue is to ensure that it is correctly calibrated. When there are two examples of the same device, even though they are correctly calibrated there will always be a problem of deciding which one is correct when they differ. When three or more are used the problem becomes even more complex. This is particularly the case when measuring skid resistance since there is no absolute value against which devices can be compared. The question *“How do you compare?”* then becomes particularly pertinent.

Until such time as an absolute reference exists (such as a very carefully controlled device that itself has been calibrated against a surface or surfaces with known, stable, reproducible friction properties), choices for comparison are limited.

There are two potential approaches that might be taken:

- To make sure that the machines measure as closely to one another as is practicable; to understand the precision of the measurements that they make and to allow for that in setting the standards that will utilise the measurements, the principle followed in the UK.

- To establish detailed correlations between the machines and an established standard (such as a “golden” device that is “correct” by definition) and provide each machine with its own “calibration factor” to relate it to the reference machine, the principle employed in some other countries, such as Germany.

Which approach is preferred will be influenced by many factors, including the way the data are to be used, for example, whether the test is to be applied as an acceptance test or as part of an investigation.

Experience in the UK has shown that a relatively large fleet of devices can be controlled within practical limits that make the data they produce useable within the context of an asset maintenance strategy. Regular trials also reveal problems with machines that might otherwise pass undetected, another argument in favour of such exercises.

The UK fleet also provides a pool against which other SCRIMs may be compared; the machine used in New Zealand is normally checked against representatives of the UK fleet after its annual service and three years ago a machine from Slovenia joined the UK trial for assessment.

Other countries may have different experiences but wherever skid resistance is to be assessed, the importance of reliable comparison between devices must not be ignored.

8. REFERENCES

Roe P G and R Sinhal (2005). Recent developments to the SCRIM measurement technique in the UK. International Conference on Surface Friction of Roads and Runways, Christchurch, New Zealand, May 2005.

Roe P G, A R Parry and H E Viner (1998). High and low-speed skidding resistance: the influence of texture depth. TRL Report TRL367, TRL Limited, Wokingham, Berkshire.

9. ACKNOWLEDGEMENTS

The work in this paper was carried out in the Infrastructure division of TRL Limited as part of projects carried out for the Highways Agency.

Paper reference PA/INF/4522/05