MERLIN—A Low-Cost Machine for Measuring Road Roughness in Developing Countries

M. A. CUNDILL

The roughness of a road's surface is an important measure of road condition and a key factor in determining vehicle operating costs. A simple roughness measuring machine has been designed especially for use in developing countries. It is called MERLIN (Machine for Evaluating Roughness using Low-cost INstrumentation). The device can be used either for direct measurement or for calibrating response-type instruments such as the vehicle-mounted bump integrator. It consists of a metal frame 1.8 m long with a wheel at the front, a foot at the rear, and a probe midway between them that rests on the road surface. The probe is attached to a moving arm at the other end of which is a pointer that moves over a chart. The machine is placed at successive locations along the road and the positions of the pointer are recorded on the chart to build up a histogram. The width of this histogram can be used to give a good estimate of roughness in terms of the International Roughness Index. Calibration of the device was carried out using computer simulations of its operation on road profiles measured in the 1982 International Road Roughness Experiment. The MERLIN is in use in a number of developing countries. It can usually be made locally at a current cost of typically U.S. $250.

The longitudinal unevenness of a road's surface (normally termed its roughness) is both a good measure of the road's condition and an important determinant of vehicle operating costs and ride quality. Within developing countries, there is particular interest in the effect on vehicle operating costs. A number of studies [e.g., by Hide et al. (*1*), Hide (*2*), the Central Road Research Institute (*3*), and Chesher and Harrison (*4*)] have shown how roughness can influence the cost of vehicle maintenance, the rate of tire wear, and vehicle running speeds (and hence vehicle productivity).

Reliable measurement of road roughness is therefore seen as an important activity in road network management. A variety of machines have been developed to make these measurements, and a number of roughness scales have been established.

However, despite the range of different measuring machines that exists, it was believed that there was a need, particularly within developing countries, for a different type of device. Ideally, it should be simple, inexpensive, require no calibration, and be able to make fairly rapid measurements of reasonable accuracy on one of the standard roughness scales. Such a machine might be used either directly to measure roughness or for calibrating other roughness-measuring equipment, particularly the widely used vehicle-mounted bump integrator.

The standard roughness scale that has been used for many years by the Overseas Unit of the U.K. Transport and Road Research Laboratory (TRRL) in its studies on vehicle operating costs and pavement deterioration is the output of the fifth-wheel bump integrator (BI) towed at 32 km/hr. However, another scale that is now being widely used is the International Roughness Index (IRI) (*5*). This scale, which is derived from road profile data by a fairly complex mathematical procedure, represents the vertical movement of a wheel with respect to a chassis to which it is coupled by an idealized suspension system of specific characteristics. The wheel and chassis are assumed to be traveling along the road at 80 km/hr. As with the BI scale, the IRI scale is measured in terms of units of vertical movement of the wheel per unit length of road, and is normally quoted in meters per kilometer. Traditionally, the BI scale is normally quoted in millimeters per kilometer.

It seemed likely that to make a roughness-measuring machine of the desired performance, the design would have to be a variant of a static-profile measuring machine. In particular, it was believed that a device that could measure the spread of midchord deviations (described in the next section) would offer the most promise.

The problem therefore was twofold:

1. To examine, by computer simulation, the relationships between roughness and the spread of midchord deviations.
2. To design a machine that could reliably make the necessary measurements.

In practice, of course, the two were interrelated, with the simulation depending on the design of the machine and the design of the machine depending on the results of the simulation.

THE SIMULATION

Principle of Operation

The principle of the measurements is as follows. Two feet are rested on the road surface along the wheel track whose roughness is to be measured at a separation L (see Figure 1), and a probe is rested on the road surface midway between them. The vertical displacement is then measured between

Overseas Unit, Transport and Road Research Laboratory, Crowthorne, Berkshire RG11 6AU, United Kingdom.

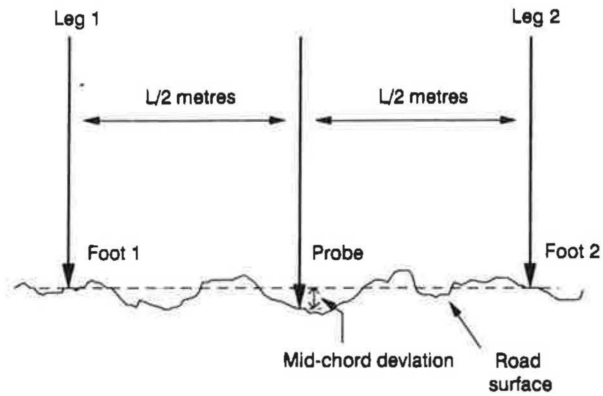

FIGURE 1 Measurement of midchord deviation.

the road surface under the probe and the center point of an imaginary line joining the two points where the road surface is in contact with the two feet. This displacement is the midchord deviation.

If measurements are taken at successive intervals along the road, then the rougher the surface, the greater the variability of the displacements. By plotting the displacements as a histogram, the variability can be estimated by measuring the range of, say, the central 90 percent of the data.

The concept of using the spread of midchord deviations as a means of assessing road roughness is not new. Two roughness indices have been proposed, QI_r and MO, which are each based on the RMS values of two midchord deviations with different base lengths. They are described by Sayers et al. (*5*). The purpose of using two base lengths is to try and match the behavior of response-type road roughness measuring systems (RTRRMSs), which usually have two resonant frequencies, corresponding to the natural vibrations of the wheel and the chassis, respectively. The scales were suggested as standard numerics that could be calculated relatively easily from road profiles measured by rod and level.

However, (for simplicity) the proposed machine would use just one base length, (for speed of operation) it would take its own measurements without the need for rod and level, and (for ease of use) roughness would be derived with little calculation.

The International Road Roughness Experiment

In 1982, a major study, the International Road Roughness Experiment (IRRE), was carried out in Brasilia (*5*) to compare the performance of a number of different road roughness measuring machines and to calibrate their measures to a common scale. As part of this study, the machines were run over a series of test sections 320 m long, for four types of road surface—asphaltic concrete (AC), surface-treated, gravel, and earth.

One of the instruments used in the study was a TRRL Abay beam (*6*). This uses an aluminum beam (3 m in length, supported at each end by adjustable tripods that can be used for leveling. Running along the beam is a sliding carriage that has at its lower end a wheel of 250-mm diameter that is in contact with the road surface. A linear transducer inside the carriage measures the distance between the bottom of the wheel and the beam to the nearest millimeter and this was recorded at 100-mm intervals along the road. By successively relocating the beam along the length of the road section and repeatedly leveling the beam, the recordings provided a continuous sampling of the road profile.

Data from the Abay beam were available for 27 of the test wheel paths. Roughness on the IRI scale was computed from the beam road profile data while roughness on the BI scale was measured by a fifth-wheel bump integrator towed at 32 km/hr. Eight of the paths were on AC roads, five on surface-treated roads, seven on gravel surfaces, and seven on earth surfaces. Roughnesses ranged from 2.44 m/km on the IRI scale (1270 mm/km on the BI scale) for the best AC surface to 15.91 m/km (16 750 mm/km on the BI scale) for the worst earth surface.

Procedure

Given these road profiles, it was possible to carry out a computer simulation of performance. It is assumed that, for ease of operation, the machine mechanically amplifies the displacements by a factor of 10. If the rear foot is placed at a horizontal distance X from the start of the section, then the probe would be at a horizontal distance of $(X + L/2)$ from the start and the front foot at a distance of $(X + L)$. If the corresponding vertical distances at these points are Y_0, Y_1, and Y_2, then the value d, measured by the machine, is given by

$$d = 10 \times [Y_1 - 0.5 \times (Y_2 + Y_0)] \quad (1)$$

Taking measurements at successive positions along the road is simulated by using successively increasing values of X. The values of d are tabulated into different 5-mm ranges to create a histogram, and once 200 observations have been made, the range covered by the central 90 percent of the data points can be measured. This range is defined as D.

For each of the test sections, four simulation runs were carried out. In each run, a measurement was taken every 1.5 m, so that the observations covered virtually the entire test section. In the first run, the starting point was at the beginning of the test section. Subsequent runs started at 0.4, 0.8, and 1.2 m from the beginning. The main analyses are based on the mean of the four resulting values of D.

Choice of Base Length

In order to see what value of base length would produce the best estimate of roughness, the operation was simulated with values of L ranging from 0.6 to 3 m. Using the procedure described earlier, linear regressions were derived relating the value of roughness on the two measuring scales to D for different base lengths.

Figure 2 shows the R^2-values for these regressions. As can be seen, good correlations were found. On the IRI scale, the best correlations occur between 1.4 and 2.6 m. The highest value occurs at 1.8 m, so this was chosen as the standard base

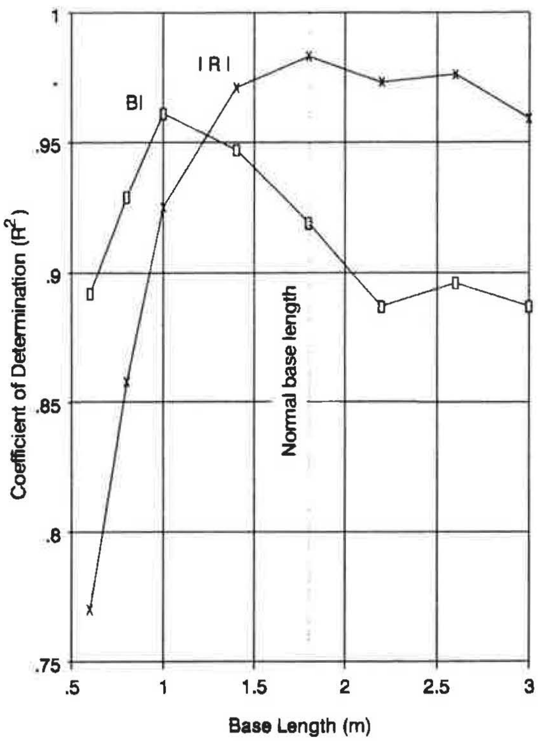

FIGURE 2 Roughness measuring accuracy for different base lengths.

Linear regressions were again carried out between D and roughness, but on this occasion, the percentage of data points was varied. The larger the percentage, the larger the value of D. The resulting values of R^2 are shown in the following table. Of the values tested, 90 percent appeared to be the best choice, with smaller values and higher values giving distinctly poorer correlations.

Percentage of Data Points Used to Determine D	R^2
95	0.932
90	0.983
85	0.966
80	0.923

Simulation Results

A plot of roughness on the IRI scale against the mean value of D (derived from 90 percent of the data points) for each of the test sections is shown in Figure 3. Over the range of roughnesses examined, the points are a good fit to a linear regression passing close to, but not through, the origin. The coefficient of determination is over 0.98. Hence, it appears that a machine working on this principle can be used as a fairly accurate means of measuring roughness on the IRI scale. Table 1 presents the regression coefficients together with their standard errors.

length. Reducing the length below 1.4 m causes a sharp decrease in correlation.

Turning to the results for the BI scale, the answer is quite different. Here the best correlation is more sharply defined and occurs at a base length of 1 m. The degree of correlation is not as good as the best IRI value, but this is to some extent explained by the fact that the BI value was determined from an independent measurement of roughness and not from an analysis of the same profile data. The use of a 1-m base length is an attractive concept, because it would make a considerably more portable measuring device than the 1.8-m version. However, it would be a much poorer predictor of IRI.

The observed behavior in these simulations can be explained on the basis that the BI scale, which relates in these studies to a measuring speed of 32 km/hr, is more sensitive to short-wavelength surface undulations than the IRI scale, which relates to a measuring speed of 80 km/hr. At a length of 1.8 m, the device has a frequency response sufficiently similar to that of the IRI scale to make it correlate well with IRI. As the base length is reduced in length to 1 m, its frequency response becomes more similar to that of the BI scale.

Measurement of Data Spread

Tests were also carried out to find the best way of measuring data spread. In order to determine the central percentage of the data used to derive the value of D that would give the best answers, the performance over the test sections was again simulated. This time, the base length was fixed at 1.8 m and the roughness was measured on the IRI scale.

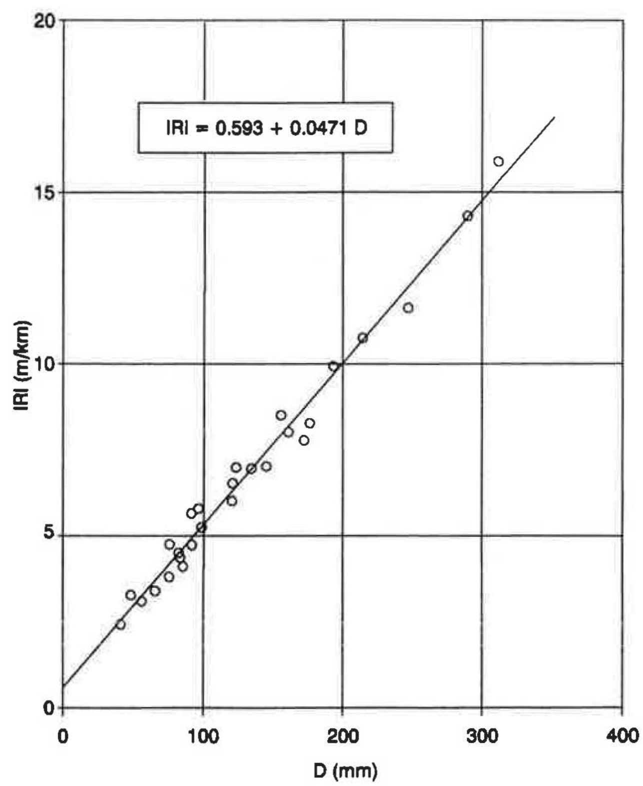

FIGURE 3 Relationship between IRI and *D*.

Figure 4 shows a plot equivalent to that shown in Figure 3, but this time measuring roughness on the BI scale. Once again, the points can be fitted to a linear regression passing close to the origin. However, the fit to the line is not as good as in the previous instance, and the coefficient of determination is reduced to just under 0.92. As mentioned earlier, this will be partly because of the fact that the BI value was derived from an independent measurement of roughness and not from an analysis of the same profile data.

The points shown in the figure distinguish between the different types of road surface, and on closer examination it can be seen that there are consistent differences between them. For example, all the results for gravel roads lie below the regression line. The analysis can therefore be improved by considering the different surface types separately, and the result of doing so is indicated in Table 1, which presents the regression coefficients. The coefficient of determination ranges from 0.914 on AC surfaces to 0.987 on surface-treated sections.

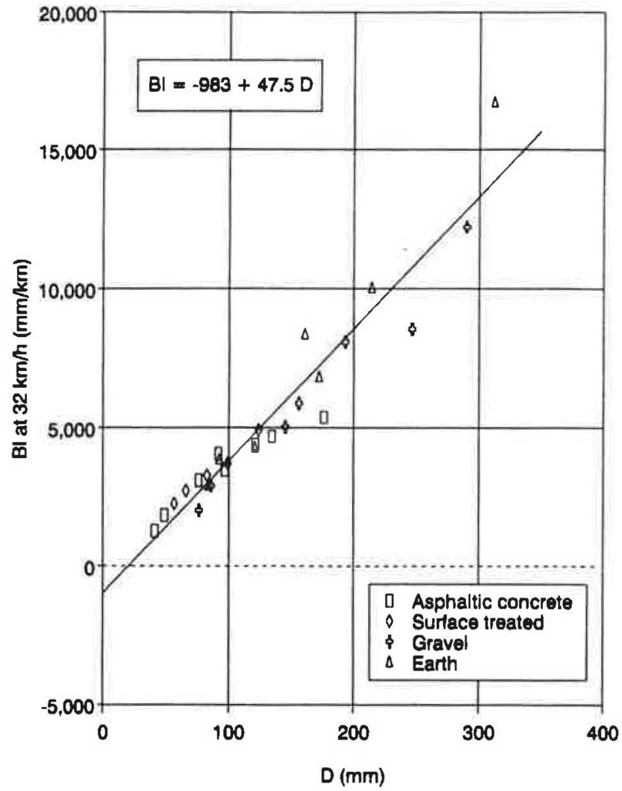

FIGURE 4 Relationship between BI and *D*.

THE MERLIN

General Description

The device that has been developed to take the measurements is called MERLIN (Machine for Evaluating Roughness using Low-cost INstrumentation).

TABLE 1 RESULTS OF REGRESSION ANALYSES (ROUGHNESS = $A_0 + A_1 \cdot D$)

Roughness scale	Surface type	A_0	A_1	R^2	N
IRI (m/km)	All	0.593 (0.185)	0.0471 (0.0012)	0.983	27
BI (mm/km)	All	-983 (423)	47.5 (2.8)	0.918	27
BI (mm/km)	AC	574 (401)	29.9 (3.7)	0.914	8
BI (mm/km)	ST	132 (220)	37.8 (2.5)	0.987	5
BI (mm/km)	GR	-1134 (676)	44.0 (3.6)	0.967	7
BI (mm/km)	EA	-2230 (797)	59.4 (4.4)	0.973	7

Notes: Bracketed values are one standard error
AC = Asphaltic concrete
ST = Surface treated
GR = Gravel
EA = Earth

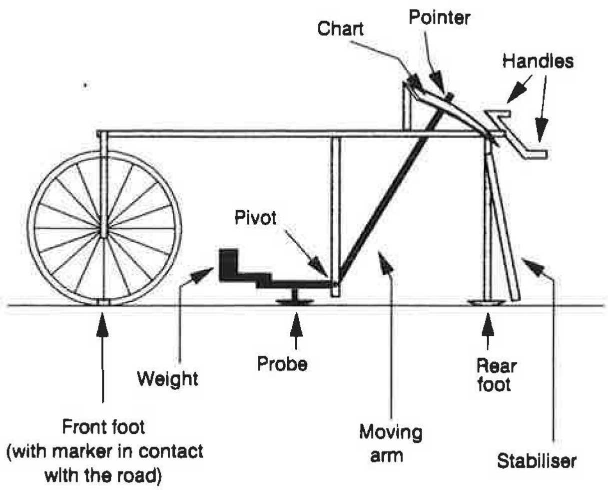

FIGURE 5 Sketch of the MERLIN.

Figure 5 shows a sketch of the device. For ease of operation, a wheel is used as the front leg; the rear leg is a rigid metal rod. On one side of the rear leg is a shorter stabilizing leg that prevents the device from falling over when taking a reading. Projecting behind the rear leg are two handles, so that the device looks in some ways like a long and slender wheelbarrow.

The probe is attached to a moving arm, pivoted close to the probe. The arm is weighted so that the probe moves downward, either until it reaches the road surface or the arm reaches the limit of its traverse. At the other end of the arm is attached a pointer that moves over the prepared data chart. The position of the parts is such that a movement of the probe of 1 mm will move the pointer by 1 cm. The chart consists of a series of columns, each 5 mm wide, and divided into boxes (see Figure 6).

Any variation in the radius of the wheel will result in a variation in the length of the front leg and this will give rise to unwanted movement of the probe. In order to overcome this unwanted movement, a mark is painted on the rim of the wheel and all measurements are taken with the mark at its closest proximity to the road. The wheel is then in its normal position.

Method of Use

In order to measure the roughness over a stretch of road, 200 observations are made at regular intervals. At each observation, the machine is rested on the road with the wheel in its normal position and the rear foot, probe, and stabilizer in contact with the road surface. The position of the pointer on the chart is then recorded with a cross in the appropriate column and, to keep a record of the total number of observations made, a cross is also recorded in the tally box on the chart.

The handles of the MERLIN are then raised so that only the wheel remains in contact with the road and it is moved forward to the next sample position, where the process is repeated. The spacing between the sample positions is not critical but readings must always be taken with the wheel in the normal position and so a spacing of one wheel circumference is the most convenient in practice.

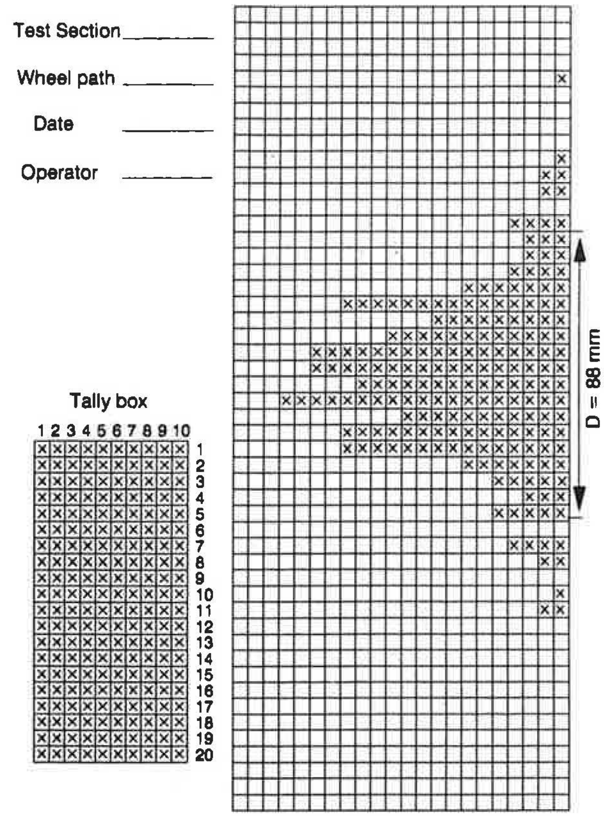

FIGURE 6 Specimen result.

When the 200 observations have been completed, the chart is removed from the MERLIN. The position midway between the 10th and 11th cross, counting the number of crosses in from one end of the distribution (e.g., from the top right in Figure 6), is marked on the chart below the columns. The procedure is repeated for the other end of the distribution. It may be necessary to interpolate between column boundaries, as shown by the lower mark of the example. The spacing between the two marks, D, is then measured in millimeters. This result is the roughness on the MERLIN scale. Road roughness, in terms of IRI or as measured by a towed fifth-wheel bump integrator, can then determined using one of the equations presented in Table 1.

Practical Details

Figures 7 and 8 show the MERLIN. For ease of manufacture, the main beam, the central and rear legs, the moving arm, the stabilizer, and the handles are all made from steel tubing of square cross section, 25 × 25 mm, with wall thickness of 1.5 mm. Joints are welded where possible, though the stabilizer and handles are fixed by bolts so that they can be removed for easier transportation. In order to strengthen the

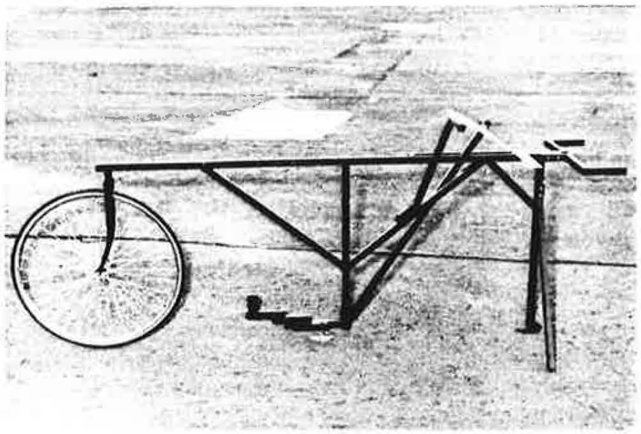

FIGURE 7 The MERLIN.

joints between the main beam and the legs, additional struts are used. The wheel is mounted in a pair of bicycle front forks and the tire has a fairly smooth tread pattern.

In order to reduce sensitivity to road surface microtexture, the probe and the rear foot are both 12 mm wide and rounded in the plane of the wheel track to a radius of 100 mm. The rounding also tends to keep the point of contact with the road in the same vertical line. The pivot is made from a bicycle wheel hub and the arm between the pivot and the weight is stepped to avoid grounding on rough roads.

The chart holder is made from metal sheet and is curved so that the chart is close to the pointer over its range of movement. In order to protect the arm from unwanted sideways movement, a guide is fixed to the side of the main beam, retaining the arm close to the beam. One end of this guide acts as a stop when the machine is raised by its handles.

The probe is attached to the moving arm by a threaded rod that allows both vertical and lateral adjustment. The position of the probe must be set so that the pointer is close to the middle of the chart when the probe displacement is zero, otherwise the histogram will not be central. Also, if the traverse of the probe does not pass centrally through the line joining the bottom of the tire and the rear foot, then it will be found that when the machine is tilted from side to side, the pointer moves. When correctly adjusted, leaning the machine over to one side so that the stabilizer rests on the road has little effect on the position of the pointer.

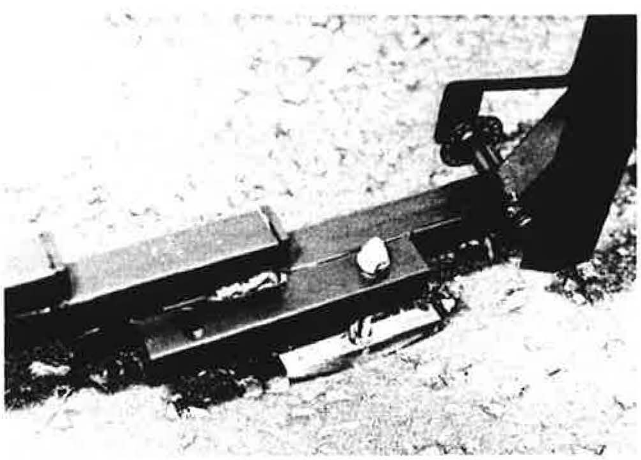

FIGURE 8 The probe and moving arm.

The only calibration required for the instrument is to check the mechanical amplification of the arm. This is normally done simply by measuring the movement of the pointer when a block of known thickness is placed under the probe.

When making measurements on a rough road, care has to be taken to ensure that fewer than 10 readings are at each limit of the arm's movement. If this is not the case, the probe can be moved to an alternative position that is twice as far from the pivot. This procedure reduces the mechanical amplification of the arm to 5 and halves the width of the distribution.

ACCURACY OF MEASUREMENT

As a simple check on the performance of the MERLIN on AC roads, the MERLIN and the Abay beam were used on a series of test sections of the TRRL test track. Roughness values on the MERLIN scale are shown plotted in Figure 9 against roughness on the BI scale as computed using the RMSD procedure with the Abay beam (*6*). The graph also shows the MERLIN-BI relationship for AC roads as presented in Table 1. Each point represents the mean of four MERLIN measurements. Although the check is by no means comprehensive, it does lend strong support to the calibration relationships derived from the simulation.

Whether using the MERLIN for calibrating other instruments or for direct measurement of roughness, two considerations about accuracy have to be borne in mind. The first consideration is that the MERLIN roughness for a road section is derived from a sample of observations and so is subject to a random sampling error. This error can be reduced by repeat observations on the same section. The second consideration is that there are systematic differences between the roughness scales that can only be reduced by repeat observations on different road sections.

Undulations in the road surfaces can be considered as surface waves with a spectrum of spatial frequencies. These spatial waves are converted into vertical oscillations of the wheel of a vehicle, the conversion factor depending on the vehicle's speed. The IRI, BI, and MERLIN scales and any RTRRMS being calibrated, all have different sensitivities to different spatial frequencies so they will correlate uniquely with each other only for surfaces with the same spectrum of spatial waves (spectral signature). In practice, surfaces will have different spectral signatures, though there are broad similarities, especially between the signatures of individual surface types. Hence, the relationship between the scales will not be unique and this gives rise to the systematic differences mentioned earlier.

If roughness is being measured directly on the MERLIN scale, then there are no systematic errors to contend with and the error falls with the reciprocal of the square root of the number of observations. A single measurement should have an RMS residual error of 8 percent, and taking the mean of four observations should reduce the error to 4 percent. When

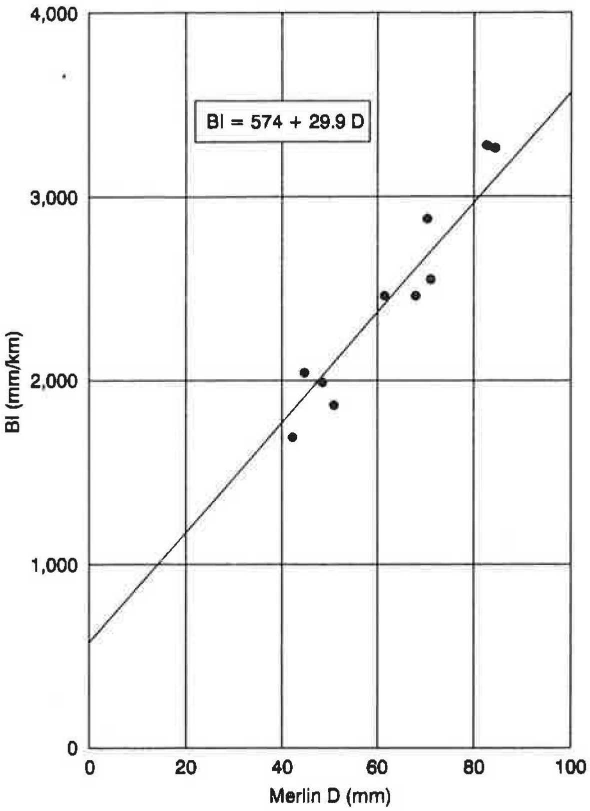

FIGURE 9 Calibration check on AC.

trying to measure roughness on the other scales, the systematic errors mean that repeat measurements do not produce as much improvement in accuracy.

Calibrating an RTRRMS on one of the standard roughness scales at a larger number of sites is better than making many repeat measurements at the same site. Moreover, particularly if working on the BI scale, these sites should have similar surfaces to those on which the RTRRMS is to be used. A number of other practical considerations should be borne in mind when measuring or calibrating and a useful guide is provided by Sayers et al. (7).

DISCUSSION OF RESULTS

The intention behind the MERLIN was that the device should be easy to use and reasonably accurate and yet able to be manufactured and maintained with the limited resources available in developing countries. Field experience so far has been satisfactory. A number of MERLINs have been made at TRRL and shipped overseas; other units have been made in developing countries from drawings provided by TRRL. To date, MERLINs have been used in 11 developing countries in South America, Africa, and Asia; in six of these countries the equipment was made locally at current prices of typically around U.S. $250.

One disadvantage of the device is that it is quite large and not easily transported within a vehicle. A shorter machine would be more convenient, but a reduction in the base length would lead to a poorer correlation with the IRI scale. Alternatively, a more portable design could be considered using a structure that dismantles. Although this is a possibility, it has been avoided because of the need to retain rigidity. Although the device is simple, it is able to measure displacements to better than a millimeter and this ability could easily be compromised by unwanted flexing of the structure.

Over the years, a number of road roughness scales have been proposed. Now the MERLIN scale, which correlates well with the IRI scale, can also be considered for applications such as identifying road maintenance intervention levels.

ACKNOWLEDGMENTS

This work forms part of the program of research of the Overseas Unit of TRRL. It is published by permission of the director of TRRL.

REFERENCES

1. H. Hide et al. *The Kenya Road Transport Cost Study: Research on Vehicle Operating Costs.* TRRL Laboratory Report 672. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1975.
2. H. Hide. *Vehicle Operating Costs in the Caribbean: Results of a Survey of Vehicle Operators.* TRRL Laboratory Report 1031. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1982.
3. *Road User Cost Study in India: Final Report.* Central Road Research Institute, New Delhi, India, 1982.
4. A. Chesher and R. Harrison. *Vehicle Operating Costs: Evidence from Developing Countries.* Johns Hopkins University Press, Baltimore, Md., and London, U.K., 1987.
5. W. S. Sayers et al. *The International Road Roughness Experiment: Establishing Correlation and a Calibration Standard for Measurements.* Technical Paper Number 45. World Bank, Washington, D.C., 1986.
6. S. W. Abaynayaka. Calibrating and Standardizing Road Roughness Measurements Made with Response Type Instruments. In *ENPC, International Conference on Roads and Development*, Presses de l'Ecole Nationale des Ponts et Chaussées, Paris, May 1984, pp. 13–18.
7. W. S. Sayers et al. *Guidelines for Conducting and Calibrating Road Roughness Measurements.* Technical Paper Number 46. World Bank, Washington, D.C., 1986.

Views expressed herein are not necessarily those of the Overseas Development Administration.