SKID RESISTANCE TESTING FROM A STATISTICAL VIEWPOINT

T. D. Gillespie, W. E. Meyer, and R. R. Hegmon,
Pennsylvania Transportation and Traffic Safety Center,
Pennsylvania State University

Standards of minimum skid resistance for highway pavements must be compatible with the limitations of locked-wheel skid testers and skid resistance characteristics of the highways. Variance of skid testing data arises from the testers, test tires, and pavements. The magnitude of each source is illustrated, and some of the causes are explained. The influence of the variance is examined with respect to the confidence of skid test data in repeat and survey modes; the point is made that a large number of tests should be conducted whenever possible. It is concluded that the reduction of variance due to the tester and the test tire, though desirable, will not eliminate the need for statistical methods in skid test data analysis. The way in which minimum skid resistance standards should be defined is discussed as is a proposed method by which statistical uncertainty can be reflected in the interpretation of results from survey testing.

•A UNIFORM method for measuring pavement skid resistance has long been needed. The locked-wheel pavement skid tester has become accepted in the United States as the best current means of skid resistance measurement. The Federal Highway Administration is expected to soon adopt a set of minimum skid resistance requirements applicable to all highways that fall within the purview of the National Highway Safety Program Standards. Individual states will be required to conduct surveys to achieve these standards. As this time approaches, there is need to consider in detail how the standards must be defined to be compatible with the performance limitations of the tester and with the skid resistance characteristics of the highways.

MEASUREMENT STANDARD OF SKID TESTERS

The locked-wheel skid tester is a measuring device that is not entirely precise and accurate. That is illustrated by the use of a micrometer to measure a particular dimension of a machine part. If a large number of measurements are made, they will not all be identical, but will be distributed (usually normally distributed) about some mean that can be calculated from the individual measurements. If, with an infinite number of tests, the mean of the measurements is equal to the "true" dimension as determined from a reference standard, the micrometer is accurate. Whether accurate or not, the spread of the measurements is indicative of its precision. An NCHRP project (1) and the FHWA Area Reference Center Program formally address themselves to the problem of skid tester accuracy. Though accuracy and precision are interrelated, the following discussion will be primarily addressed to the problem of precision.

The precision of a skid tester relates to its ability to reproduce data in a number of repeat tests. Repeat tests here refer to any number of tests on the same pavement and under the same conditions during one continuous effort, aside from day-to-day tester variability, variability of test tires, and other such factors that may also be considered

Publication of this paper sponsored by Committee on Surface Properties-Vehicle Interaction.

in the category of accuracy problems. In repeat tests, the data exhibit either systematic or random variations; the latter are usually distributed normally about a mean. The standard deviation or variance (which is the square of the standard deviation) is used as a measure of the precision; for a normal distribution, 68 percent of the data points fall within ±1 standard deviation of the mean, and 95 percent of the data points fall within ±2 standard deviations.

Because of variability it is only possible to determine the true mean obtained by the tester or the true skid resistance as measured by that tester by taking an infinite number of tests. (The true means obtained by different testers will also differ because of systematic errors that affect their accuracy.) Fortunately probability theory provides an alternate approach, which is more practical and economically acceptable than conducting an infinite number of tests. Using statistical methods, one can draw inferences about the tester's true mean from only a limited number of skid tests.

Any finite number of random tests has a mean value that is probabilistically related to the mean of an infinite number of tests; the closeness and certainty of the relation are determined by the variance of the data and the number of tests. That is, the less variance in the data and the larger the number of tests are, the closer the sample mean is expected to be to the true mean.

One way to express this relation is shown in Figure 1 (2). For a given number of tests, Figure 1 shows the confidence interval of the data mean when the standard deviation of the tester has been estimated. The standard deviation of the tester is obtained from a reasonably large number of similar tests while the measured parameter (i.e., pavement skid resistance) remains constant. The confidence coefficient (0.90) and probability (0.90) used in constructing these curves were selected because they yield practical sample sizes and reasonable ranges that subjectively agree with typical tester behavior. The choice of higher or lower values has a marked effect on the location of the curves and the implied relation, suggesting the need for further research in this area.

Nevertheless, if this graph is accepted for purposes of illustration, a simplified though approximate interpretation is that, for a given set of test data, one may be 90 percent confident that the true mean of the tester is within the indicated interval of the data sample mean.

Thus, it is demonstrated that the locked-wheel tester can provide only an estimate of skid resistance. This statement applies to any measuring device; however, it carries particular significance with respect to the practical employment of locked-wheel skid testers.

Table 1 gives data taken by 6 testers in 10 repeat tests at the same location on a pavement during a skid-tester calibration and correlation study (3). At other speeds (30 and 60 mph) the standard deviations for the individual testers changed significantly, although the data spread over all testers remained at about the same magnitude. The standard deviation of each tester was estimated by the standard deviation given in Table 1 on the basis of 10 tests. Figure 1 shows that the best tester is 90 percent confident of obtaining a mean skid number within 0.6 SN of its true mean; the poorest tester is only within 2.4 SN.

The data given in Table 1 were assembled from the tests of 6 typical testers operated under closely controlled conditions. The average standard deviation for all 6 testers is approximately 2 SN and may be considered to be representative of typical testers in use today. Applying this standard deviation to Figure 1, one can see that at the stipulated confidence level a typical tester can only be expected to measure a mean skid number within 2.5 SN of its true value for 5 tests and within 1.5 SN for 10 tests. On a percentage basis, these figures translate into roughly 6 percent and 4 percent for 5 and 10 tests respectively on pavements such as those on which Table 1 data are based. Practically, these values may be taken as an indication of the accuracy of a tester in determining the skid resistance of a particular pavement location relative to the number of tests that are made.

According to the data shown in Figure 1, the best skid resistance data are obtained when a large number of tests are made with a tester having a low standard deviation. The key question then becomes, What are the sources of data variance, and how can they be controlled?

The answer is that variance of skid test data has many sources that can be grouped into 3 major sources: pavement, test tire, and tester. Within each of these, the variance includes influences from sources such as the operators, test procedures, and data evaluation methods. Figure 2 shows data from several series of tests that illustrate the magnitude of the variance from each of the 3 major sources. On a specific pavement site 50 tests were run with an ASTM E 249 test tire and a commerical steel-belted radial tire. The reduction in standard deviation between the first 2 series of tests was attributed to the greater uniformity of the steel-belted test tire. Tests with the ASTM tire on a textured steel plate yielded a reduction in standard deviation, suggesting a reduction in pavement variability. In this case, the downward drift of the data with successive tests was accommodated by using a sloping "best fit" control line because the metallic surface was visibly burnishing. Finally, the steel-belted tire was tested on the textured steel surface and yielded an even smaller standard deviation. Regarding this last test condition, it may be concluded that both the ASTM test tire and the pavement (even in repeat tests at the same location) each roughly double the spread of the test results.

The tests described illustrate the nature of data variance and the typical magnitudes of the influence from pavement, test tire, and tester. To expect all users to run such extensive tests with locked-wheel skid testers is both impractical and unnecessary. Because all survey work must be done with the ASTM test tire, the standard deviation of interest is that of the tester and the tire combined.

To obtain an estimate of the standard deviation for a tester with test tire would require at least 50 tests to be conducted on the most uniform available pavement sites. The pavements should have adequate side slope to prevent water accumulation; however, complete drying of the site between tests is impractical because the time for tests would become excessive. The testing rate should not exceed 1 test per minute. If a data drift is observed that can logically be attributed to changes in test conditions (and can be substantiated by control tests), a best fit straight line may be applied to the data as was done in the cited example (Fig. 2).

The tests should be conducted over a variety of surfaces and speeds representative of those found in routine testing. If the standard deviations obtained are clearly and repeatably functions of speed or friction level, they may be applied accordingly.

The standard deviation thus obtained is still only an estimate and applies to the combination of tester and tire. The actual deviations should be kept to a minimum by maintaining test tires in good condition and discarding those that have been dry skidded or have become visually nonuniform. Similarly, operators have an influence on the data, especially through the procedures used. Every effort should be made to maintain consistency of performance through accurate speed holding, controlling lateral placement of the tester, and consistently evaluating skid records. Good tester design and high quality equipment are essential in keeping variability to a minimum. Consistent water system performance and low drift in the gain and zero of the instrumentation are important in this respect.

PAVEMENT FRICTION VARIANCE

The discussion thus far has dealt with skid-tester variability and its influence on the confidence interval of the skid resistance determined by sampling at just one pavement site. However, the friction of a pavement varies laterally, longitudinally, and temporally. The last factor affects the ultimate level of significance in a skid resistance survey, but for now the interest may be confined to the first 2 factors.

Figure 3 shows typical skid number profiles across 2 typical pavements. The profiles were obtained by averaging the skid numbers from 10 longitudinal pavement sites when the lateral position was closely maintained with a tracking device. Though not shown by the average, typical variations as large as 20 percent were observed at individual longitudinal locations when the tester moved into or out of the wheel track. For this reason, close control of lateral position in skid resistance testing is important. A skilled and experienced driver can keep variations due to this source to a minimum but cannot eliminate them because, as shown by the graph, a 10-in. deviation

Figure 1. Statistical relation of number of skid tests and confidence interval associated with data mean.

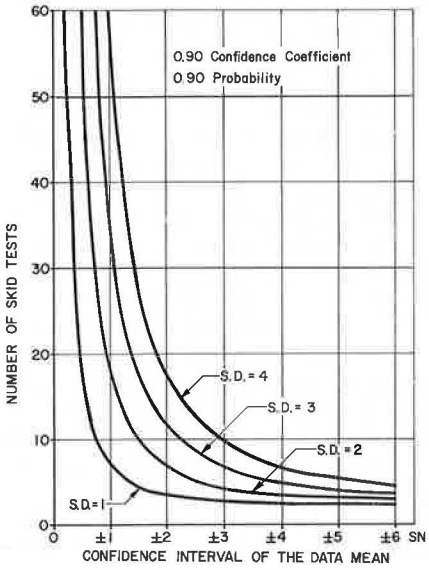


Figure 2. Repeatability tests with 2 tires and 2 surfaces.

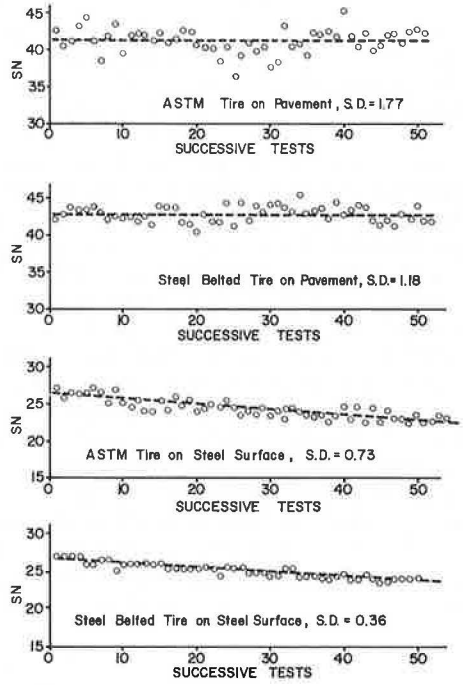


Table 1. Repeat data for 6 testers at test speed of 40 mph.

	Tester					
Test	1	2	3	4	5	6
1	48.2	55.7	47.9	46.5	46.0	40.0
2	48.2	48.7	44.8	42.0	46.0	40.0
3	49.3	49.7	45.8	43.5	52.0	39.0
4	47.1	51.7	49.9	42.0	46.0	40.5
5	49.3	48.7	49.4	46.5	42.0	39.5
6	52.4	51.7	44.3	44.5	49.0	41.0
7	48.2	48.7	47.4	42.0	44.0	41.0
8	50.2	49.7	46.3	48.5	46.0	41.5
9	49.1	49.7	44.3	50.0	48.0	40.5
10	—	50.7	44.8	—	46.0	39.0
Mean	49.11	50.50	46.49	45.05	46.50	40.20
Standard deviation	1.525	2.149	2.074	2.99	2.718	0.856

from the intended path can produce a 2 to 3 percent variation. Because of the problem of wheel track detection and tester control, lateral placement errors are an ever-present cause of variance in survey testing data.

Figure 4 shows the skid number profile measured longitudinally along 3 miles of a 2-lane highway that has few access points. The tests were run so as to eventually obtain 20 evenly spaced data points per mile. The pavement was homogeneous in age, design, and construction method. The data suggest that the pavement is frictionally homogeneous because no noticeable trend appears even though the data vary over a 30 percent range. On the other hand, the homogeneity would not have been apparent had only 5 or 10 consecutive tests been made, for example, between test numbers 25 and 35.

The element of risk or chance in skid testing can be illustrated by closer examination of the data shown in Figure 4. Table 2 gives the means, standard deviations, and confidence intervals for the case in which either 3, 5, 15, or 64 tests are conducted over the test distance. For the cases with 3, 5, and 15 tests, the data points are selected at uniform intervals from the sequence of 64 points to simulate the way data might be acquired in routine testing. Neither the means nor the standard deviations show a clear trend but are statistically a result of pure chance. The fact that the standard deviation for the case of 15 data points is less than that for 64 is a result of chance and demonstrates the risk involved in estimating the standard deviation from a limited amount of data. When, however, the statistical significance of the data is considered and the confidence interval is determined (by using Fig. 1), the importance of a large number of tests becomes obvious.

The data given in Table 2 were obtained with a locked-wheel tester using an ASTM test tire so that the calculated standard deviations include the composite of tester-tire and pavement variability. Assuming the variances do not interact, the standard deviations combine according to the relation

$$s_c^2 = s_T^2 + s_p^2$$

where

s_c = composite standard deviation,
s_T = tester-tire standard deviation, and
s_p = pavement standard deviation.

Using 2.0 SN as a conservative estimate of the tester-tire standard deviation, the pavement standard deviation can be computed and is given in the last column in Table 2. From the set of 64 tests for which the statistical confidence is highest, the pavement contribution to the test data variance is greater than that of the tester-tire combination. Thus, even when testers and test tires have been improved to the point where they become insignificant as sources of variance, the statistical problem in skid resistance surveying will remain because of the nonhomogeneous nature of the pavements.

From the standpoint of homogeneity, the pavement used as an example is typical and far better than many encountered in skid resistance surveys. Yet, to obtain an estimate of pavement skid resistance that has a precision of 3 or 4 percent (in this case less than a 2 SN confidence interval) requires far more than 5 tests specified as minimum in ASTM Method E 274-70. On the other hand, there is a practical limit to the number of tests that can be accomplished with a single tester in routine use. Figure 5 shows the mean skid number, standard deviation, and confidence interval versus the number of tests for tests conducted according to the ASTM Method E 274-70 and at the rate of 4 tests/min. The mean skid number, standard deviation, and confidence interval values are cumulative from the first to the n th test so that the values at the n th test are the result that would have been obtained if testing had been discontinued at that point.

The 4-tests/min rate is high enough that tire heating occurred and caused a downward trend in the skid resistance data. A similar effect might just have well been observed on a pavement that is nonhomogeneous. The trend shows up as a downward drift in the mean skid number and an eventual upward drift in the standard deviation. The confidence interval itself starts high and settles down, in this case a little faster than

expected because of a chance low point in the standard deviation. In theory, the confidence interval will continue to decrease with increasing number of tests such that the data mean is expected to approach the true mean of the tester. In this case, as in others that will occur in practice, systematic changes in test conditions take place in the course of the testing and, thus, prevent the confidence interval from approaching zero. A total of 10 or 20 tests appears to represent the point of diminishing returns. A data drift as shown in Figure 5 might also arise from pavement nonhomogeneity, and the more complicated problem of defining homogeneous test sections would be raised.

CONCLUSIONS

From the discussion of the application of statistical methods to the problem of variability in skid testing and some typical examples of actual tests, certain conclusions can be drawn about the limitations of using the locked-wheel method for skid resistance measurement.

1. The skid tester-test tire variance is of the same order or somewhat smaller than the variance of typical pavements.
2. Even if skid tester and test tire variances could be eliminated, the necessity of statistical analysis of skid test data remains because of the variance in pavements themselves.
3. The optimum number of tests is a compromise between precision and practicality. The point of diminishing returns appears to be about 10 to 20 tests per test section.
4. The practical limit of precision in skid testing appears to be typically 3 to 4 percent and to depend primarily on pavement homogeneity and obtaining a statistically significant number of tests.
5. Precision skid resistance data in short pavement sections or at discrete locations cannot be obtained without multiple passes through the test section to achieve statistical validity of the data.

SETTING MINIMUM SKID RESISTANCE STANDARDS

In the establishment of minimum skid resistance standards, 2 factors must be recognized: Skid resistance cannot be determined to an arbitrarily high degree of precision, and the definition of requirements in terms that necessitate discrimination of skid resistance at discrete pavement locations will greatly increase the effort involved in testing.

In the final analysis the standards must recognize traffic demands created by factors such as grades, curves, speed limits, and intersections and must likewise allow for skid resistance determination wherever possible by averaging from a large number of tests.

The suggested method is to define skid resistance standards by categories based on the design standards of each test section as measured by the criteria for design of curves, grades, shoulders, speed, and other factors that affect traffic friction requirements. The acceptability of a given highway would then be determined by comparison of the mean skid resistance, reduced by a weighted portion of the associated confidence interval, and the minimum requirement.

Methods that require discrete minimum friction levels in individual curves or other traffic situations have been postulated. To include the degree of refinement that recognizes singular locations of low skid resistance level or high traffic demand will require adoption of repeat testing methods and will increase the effort required in survey testing by an order of magnitude.

APPLICATION TO SKID RESISTANCE SURVEYS

Extensive effort is required to ensure high statistical confidence in all skid resistance surveys. A logical alternative is to conduct 2 surveys: a general survey to identify highway sections suspected of failing to meet minimum requirements and a precision survey of the suspected sections to obtain the precision data necessary for final

Figure 3. Average lateral profiles for 2 highways.

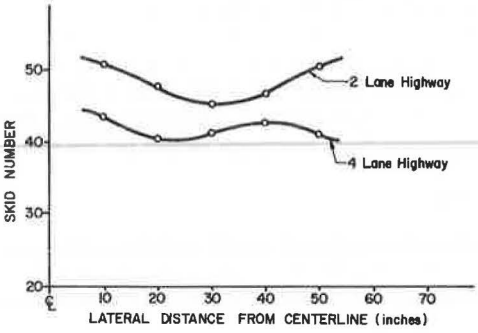


Figure 4. Longitudinal skid number profile on 2-lane highway.

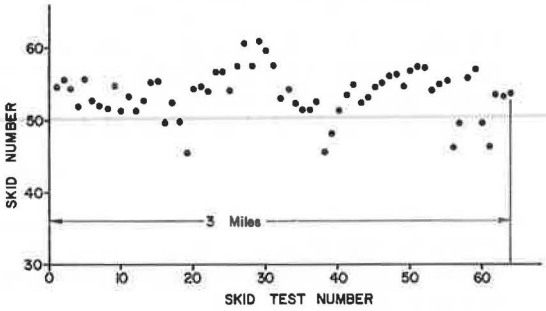


Table 2. Statistical characteristics of data of different sizes.

	Skid Number			
Number of Tests	Mean	Standard Deviation	Confidence Interval	Pavement Standard Deviation
3	53.0	6.00	>±6	5.65
5	52.7	4.50	±6	4.03
15	54.5	2.77	±1.6	1.91
64	53.9	3.39	±0.8	2.74

Figure 5. Data drift obtained at 4 tests/min.

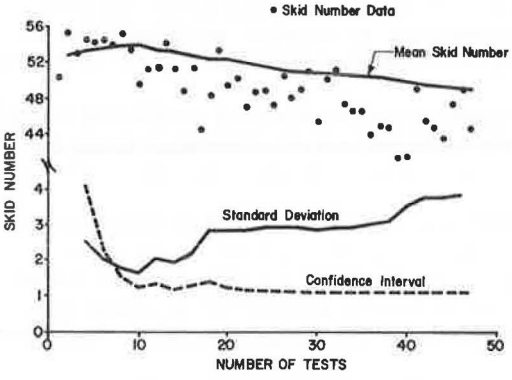
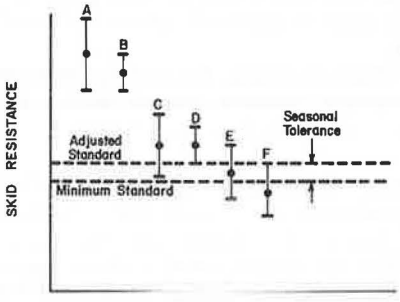


Figure 6. Comparison of skid resistance data with minimum standards.



judgment on compliance with requirements and to determine the maintenance action required.

Judgment of individual pavement sections might be done in both types of surveys in the fashion shown in Figure 6. Rather than consider only the skid number means one should, it is suggested, consider the mean minus one-half the confidence interval. The skid resistances of surfaces A and B are clearly above the minimum requirement and may be judged adequate even though they differ in mean and confidence interval. Even when the confidence interval is large because of a low number of data points, such as on surface A, further testing is not required. Surfaces E and F have means that are respectively above and below the minimum requirement. Surface F is clearly substandard, and surface E should likewise be considered substandard because the confidence interval indicates that portions of it are substandard. If the confidence interval in this latter case is large because of a low number of test data, a precision survey is called for. The larger number of data points may possibly reduce the confidence interval enough so that the mean minus one-half the interval will no longer fall below the minimum standard.

Surfaces C and D are more difficult to judge because of the closeness of their skid number means to the minimum requirement, and the judgment must take into account other phenomena in skid resistance testing. For instance, temperature and seasonal effects are known to influence skid resistance, and the effects are variable throughout the country. If it is postulated that surface must be adequate at all times, a seasonal tolerance based on empirical knowledge of such effects in the local geographic area can be added to the minimum and used to adjust the requirement. As illustrated, this would make surfaces E and F clearly substandard.

In the general survey, the suggested criterion for identifying suspected substandard surfaces is whether the skid number mean less one-half its confidence interval falls below the adjusted requirement. For such surfaces, precision surveys should then be conducted during that portion of the test season when the minimum skid resistance prevails.

A frequent second objective in skid resistance surveys is predicting when highways will reach substandard skid resistance levels, and statistical methods can also be used for that. The use of data means with confidence intervals can be implemented in existing procedures, thereby placing the predictions on a more rational basis because the data variances can be transformed into variance on the predictions as well.

In this paper no attempt has been made to set statistical criteria for methods in the use of locked-wheel skid testers or for decisions on pavement skid resistance adequacy. The criteria used in the discussions illustrated how and why pavement skid resistance is a statistical quantity. The valid use of skid resistance data requires statistical methods, and it is hoped that its importance and some potential methods may be recognized.

ACKNOWLEDGMENTS

The data used for illustration in this paper were obtained in connection with the conduct of the NCHRP Project 1-12(2).

REFERENCES

1. Meyer, W. E., Hegmon, R. R., and Gillespie, T. D. Locked-Wheel Pavement Skid Tester Correlation and Calibration Techniques. Pennsylvania Transportation and Traffic Safety Center, Pennsylvania State Univ., May 1973, 360 pp.
2. Guenther, W. C. Concepts of Statistical Inference. McGraw-Hill, 1965, App. D-11.
3. Neill, A. H., Jr., and Boyd, P. L. Calibration Study Should Eliminate Variations in Skid-Testing Trailers. Highway Research News, No. 45, Autumn 1971, pp. 31-36.