ANALYTICAL PROBLEMS ENCOUNTERED IN THE CORRELATION OF SUBJECTIVE RESPONSE AND PAVEMENT POWER SPECTRAL DENSITY FUNCTIONS

L. F. Holbrook and J. R. Darlington, Michigan Department of Transportation

It is argued and demonstrated that, when human subjective response to road roughness is functionally related through multiple regression to power spectral density frequencies of the road profile, highly unreliable estimates of frequency coefficients result. Hence, one will be misled in assuming that such roads are especially detrimental to ride. The problem, generally designated "multicollinearity," is caused by extremely high intercorrelation of many of these frequencies. This follows from the mathematical treatment required in power spectral density analysis as well as from the inherent nature of road profiles. Nor is the situation any better if frequency selection procedures such as stepwise multiple regression are used in an attempt to capture only the most important frequencies. The presence of high multicollinearity between frequencies makes trivial the statistical selection and rejection criteria and thereby allows sampling error to essentially determine which frequencies are selected. A proposed solution to this problem is taken from the econometrics literature and applied to a small sample of subjective ride data for illustrative purposes only. The conventional full multiple regression estimates of frequency coefficients give totally unreasonable results, and the proposed solution gives results consistent with known automobile pass-band characteristics.

•THE GENERAL MOTORS rapid-travel profilometer is currently the only distortion-free system for profile measurement (1, 2). This does not imply that all wavelengths can be measured, and in practice the device is accurate only for wavelengths extending from 3 in. to approximately 200 ft. This implies that profiles are measured with respect to a linear reference that is no longer than 200 ft. Moreover, the position of this reference is arbitrary and slowly varying as the profile is traversed. This produces the seeming paradox of repeat runs on the same profile appearing different when plotted. Again however, this is merely a consequence of measuring the profile with respect to a linear but slowly varying, arbitrarily positioned reference.

Normally, this situation causes no problems in frequency domain analysis of a single wheelpath profile because, as stated above, the profile is entirely accurate within a given band. A problem does arise, however, in measuring the difference in elevation between inner and outer wheelpaths. This signal, known as the roll component, may have a strong bearing on ride quality as determined by subjective response of passengers in the vehicle. Moreover, the problem is not solved by profilometers with dual wheelpath measuring systems. The separate wheelpath measuring elements are completely independent and produce profiles measured with respect to arbitrary and independent references. In this respect, the dual wheelpath systems provide no improvement over measurement of each wheelpath separately with single wheelpath units. Unfortunately, a process called tipping, which inserts the arbitrary reference of one

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

profile into another, thereby permitting crude comparisons, cannot be used because the error introduced is generally similar to most roll components.

In view of the possible importance of roll components in ride quality, it is natural to seek some method of utilizing the separate wheelpath profiles. Although the roll component signal cannot be obtained directly, each wheelpath profile yields a power spectral density (PSD) function that is not affected by the reference differences. It is then possible to combine power between inner and outer wheelpaths at each frequency to provide a power function relating the 2 signals. Three obvious combinations of power at each frequency are (a) the average power between lanes $(I_i + O_i)$, where I_i equals power in the inner lane at frequency I_i and O_i equals power in the outer lane; (b) the absolute difference in power between lanes $|I_i - O_i|$; and (c) the product of $(I_i + O_i) \times |I_i - O_i|$.

Interpretation of function a is straightforward, but functions b and c deserve some comment. If power in each wheelpath is similar, the difference will be small and in theory will not indicate the presence or absence of roll component. It is safe to assume, however, that a small roll component is present because transverse finishing produces parallel roughness components spanning the entire lane. It is also probable that large differences in power between paths imply a high roll component. If this proves to be the case, function b may correlate with subjective response. Function c, which is average power multiplied by difference in power, expresses the interaction between average and roll power. This may prove to be a sensitive measure because rough pavements, particularly flexible, have a high roll component.

THE PROBLEM

As measurement of road profiles with rapid travel profilometry techniques becomes more popular, it is natural to expect that profile PSD will be used to predict human subjective response to road roughness (3). If problems in the measurement of subjective response (SR) can be overcome (4, 5), further problems will arise if conventional multiple regression techniques are used to estimate the β parameters in the linear formulation

$$SR = \gamma + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_N X_N \tag{1}$$

where X_i is some form of intensity (usually log variance) of the respective profile frequencies f_i. When the number of frequencies (the N regressors) in Eq. 1 is large, say, 3 or more, a high degree of intercorrelation among them will usually be present. One reason for this is the mathematical smoothing induced by the PSD analysis. To understand why this must be so, we will examine briefly the PSD analysis used in the present study. For this analysis, 4 statistical decisions are important:

1. The analog profile signal from inner and outer wheelpaths is filtered to eliminate all wavelengths outside the band of 2 to 50 ft;
2. The filtered signals are sampled every 6 in. providing 4 points per cycle of the highest frequency present;
3. Twenty-five ordinates are computed providing 13 "independent" estimates of the power spectrum (estimates are spaced 0 to 0.02 cycle/ft apart, and the resolution bandwidth is 0.04 cycle/ft; and
4. A Hanning spectral window is used to smooth the final estimates.

These considerations imply a theoretical correlation among PSD estimates for broadband white noise (6). Adjacent ordinates are correlated about 0.6 if the true spectrum is flat. In addition, 2 other conditions may increase the correlation among ordinates by unpredictable amounts.

1. If a signal from a nonflat power spectrum has a great deal of power in a narrow frequency band and very little in an adjacent band, estimates of power in the weak band will be too high. This occurs because a finite data sample implies a data window that transforms through analysis into a spectral window with nonvertical skirts and side lobes. This simply means that the PSD resolution filter attenuates but does not completely eliminate power in adjacent frequencies. Thus, the resolution filter may not attenuate

power at an adjacent frequency enough to keep it well below power at the frequency being examined. This problem is significant only when the true spectrum has a very steep roll-off of perhaps 40 dB per decade or a narrow band of power perhaps 40 dB above adjacent bands.

2. Consider the case in which the PSD for a number of roads have the same general shape but simply move up or down as a unit depending on general roughness. In this case power in a given band rises or falls along with power from an adjacent band. This would generate the correlation between an ordinate and its neighbors. That the PSD often have similar shapes but different general power levels can be seen from the examination of a number of power spectra. Indeed, neighboring frequencies may be so highly correlated that no useful information is supplied by one that is not supplied by its close neighbors.

High independent variable intercorrelation is frequently encountered in multiple regressions with large numbers of regressors and has been termed "multicollinearity" (7, 8, 9). Multicollinearity in the limit where 2 or more regressors are perfectly correlated aborts the multiple regression estimation of β because the X^TX matrix has 2 or more identical columns and cannot be inverted as required by the multiple regression procedure. Short of this, high multicollinearity induced by either high correlation between 2 variables or moderately high correlation among all variables causes the β estimates to be extremely unreliable. If β is used as a measure of the relative importance of profile frequencies, high multicollinearity will almost certainly lead to erroneous inferences. A second sample will give radically different β weights and consequently inconsistent designation of those frequencies that most seriously affect riding quality.

It might be thought that some of the variable selection procedures such as stepwise multiple regression might solve the problem (10). Unfortunately, this is not the case—these procedures are seriously influenced by multicollinearity (11). What will generally happen with the forward selection procedures is that the variables that first enter the equation do so on the basis of relatively high correlations with the dependent variable. However, PSD frequencies considered as regressors are very highly intercorrelated, and those chosen initially by the selection procedure would be only insignificantly more correlated with SR than neighboring frequencies. However, these latter frequencies will never be selected by the procedure because, by virtue of high intercorrelation with the initial set, their unique relation with SR will not be large enough to permit their inclusion in the regression equation. In other words, when high multicollinearity is present and forward variable selection procedures are used, variables selected early in the procedure drastically militate against the inclusion of potentially important remaining variables. But the initial set chosen by the procedure, being only marginally more correlated with SR than other variables, suggests that it would not be selected initially in subsequent samples. Therefore, if we want to identify important PSD frequencies, we can expect procedures such as stepwise regression to pick a different set every time we process new data.

Similar comments apply to the variable rejection part of stepwise multiple regression: Variables included later may through intercorrelation "rob" earlier variables of their contribution and thereby cause them to be rejected at a later state. Finally, even if a few important frequencies are repeatedly selected by these procedures, their coefficients will vary considerably from sample to sample. This can be seen in the formula for the variance of the $\hat{\beta}_2$ coefficient for the case of only 3 regressors in Eq. 1:

$$\mathrm{Var}(\hat{\beta}_2) = \frac{\sigma^2}{(1 - r_{23}^2) \sum\limits_{i=1}^{K} (X_i - \overline{X})^2} \qquad (2)$$

As the correlation between X_2 and X_3 (r_{23}) increases, Var ($\hat{\beta}_2$) also increases until in the limit, when $r_{23} = \pm 1.0$, Var ($\hat{\beta}_2$) $= \infty$.

ANALYSIS PROCEDURE

To simplify the illustration of problems encountered with road frequency multicollinearity, we will use only 1 of the first of the 3 power combinations discussed and define a roughness measure S_i as $(O_i + I_i)$, where I_i = logarithm of squared amplitude for the ith frequency band of the inside wheel track, and O_i = logarithm of squared amplitude for the ith frequency band of the outside wheel track. S can be thought of as a measure of "average" intensity for the respective frequencies. If S is standardized to s, we will be able to compare the multiple regression coefficients, $\hat{\beta}_i$, directly. If, under ordinary multiple regression specifications, we were to regress subjective response SR on s for each frequency band, we would have

$$\hat{SR} = \hat{\gamma} + \hat{\beta}_0 s_0 + \hat{B}_1 s_1 + \ldots + \hat{\beta}_N s_N \tag{3}$$

which requires the estimation of N + 2 parameters. Suppose now that we require $\hat{\beta}$ to conform to some reasonable function, remembering that it would be unlikely that β would jump around as capriciously as the ordinary multiple regression estimates do (12, 13). Not knowing a priori the form of this function, we should initially use only very general functions such as polynomials. Suppose we specify a kth degree polynomial together with the condition that $\beta_0 = 0$, i.e., that the weight corresponding to the frequency f_0 is 0 (Fig. 1). The equation functionally relating β_i to i will be

$$\beta_i = \alpha_1 i + \alpha_2 i^2 + \ldots + \alpha_K i^K \tag{4}$$

Substituting Eq. 4 into Eq. 3, we have

$$\begin{aligned} SR = \gamma &+ (\alpha_1 + \alpha_2 + \ldots + \alpha_K)s_1 + (2\alpha_1 + 2^2\alpha_2 + \ldots + 2^K\alpha_K)s_2 \\ &+ \ldots + (N\alpha_1 + N^2\alpha_2 + \ldots + N^K\alpha_K)s_N \end{aligned} \tag{5}$$

Factoring α_j, we have

$$\begin{aligned} SR = \gamma &+ \alpha_1(s_1 + 2s_2 + \ldots + Ns_N) \\ &+ \alpha_2(s_1 + 2^2 s_2 + \ldots + N^2 s_N) + \ldots \\ &+ \alpha_K(s_1 + 2^K s_2 + \ldots + N^K s_N) \end{aligned} \tag{6}$$

Or by defining the terms in s as z, we have

$$SR = \gamma + \alpha_1 z_1 + \alpha_2 z_2 + \ldots + \alpha_K z_K \tag{7}$$

Because we will choose $K \ll N$, parameter estimation by ordinary multiple regression will not be nearly so subject to multicollinearity problems in Eq. 7 as in Eq. 1. Therefore, we can estimate γ_j conventionally and by virtue of Eq. 4 estimate β in turn.

Two problems remain.

1. The order of the polynomial chosen to govern β is arbitrary. Or, in other words, How large should K be to ensure faithful representation of the population of β? For most applications, we might be satisfied with K = 3 or 4; however, unless we have considerable information about β, we can never be sure that important peaks and valleys in the β function are blurred by polynomials of low order. One course of action would be to let the selection procedures such as stepwise multiple regression determine K. For example, one could compute the various z_j in Eq. 7 for, say, K = 10. If all the z_j were regressed on SR, $\hat{\beta}$ would follow a tenth-order polynomial. This would be excessive for most applications; however, stepwise regression could statistically select those z_j that proved important enough to significantly reduce the residuals. The α_j corresponding to the z_j selected would then estimate the β_i by virtue of Eq. 4. Because we do not care which z_j are selected (unlike the case with the s_i), many combinations of z_j would probably suffice to estimate the function. We must not

be too permissive in the stepwise variable rejection procedure—too high an order of polynomial will tend to give the same unreliable results as an ordinary multiple regression on Eq. 1.

2. Not all the information governing the β_i distribution is used. It can be fairly assumed that no β_i should be negative. If this were not the case, the implication would be that power in these wavelengths improves ride. There are several ways in which this constraint on the β_i can be incorporated into the estimation procedure. One is to specify the β_i exponentially:

$$\beta_i = i\epsilon^{\sum_{k=1}^{K} \alpha_k i^k} \tag{8}$$

for K = 1, 2, ... This formulation forces the β distribution through 0 and requires all $\hat{\beta}_i$ to be positive. Moreover, the polynomial in the exponent can be of any desired order. Unfortunately, stepwise procedures cannot be used to determine this order because ordinary linear least squares procedures are not applicable. However, the α_k can be estimated by nonlinear computer search procedures.

EXAMPLE WITH FIELD DATA

As an illustration of the problems encountered with multicollinearity, consider the following example using the SR data and PSD described in detail by Holbrook in an earlier report (4). Fourteen test roads rated on roughness at 30 to 50 mph by a panel of 96 subjects using graphic rating scales were profiled with the General Motors rapid travel profilometer. This allowed the computation of frequency spectra for the range 0.02 to 0.50 cycle/ft. The degree of sample correlation found between pairs of PSD ordinates is shown in Figure 2 for the case of the 0.22 frequency. Correlation with adjacent ordinates is extremely high: +0.9990 and 0.9819. Throughout most of the frequency range, correlations with the 0.02 ordinate are 0.9000 or higher. With the multicollinearity problem as acute as this, ordinary multiple regression should provide very poor estimates of the relative importance of the respective frequencies. That this is so is shown by the $\hat{\beta}_i$ coefficients for regressions of SR on even and odd frequencies in Figure 3. Interpretation of these weights is difficult particularly in view of the large number of negative signs and the fact that the weights from the even frequency analysis are considerably different from those of the odd weight analysis. It would seem unlikely that negative weights are realistic when one considers that they imply that high amplitudes in the associated frequencies induce a better ride! Clearly we cannot depend on this traditional procedure to detect the important frequency ranges. Because the polynomial lag procedure might provide a better estimate of the β_i weight distribution, it was applied to the same data by using the stepwise procedure with Eq. 4. A $\hat{\beta}_i$ distribution was found that peaked at about 7-ft waves. However, for shorter waves, the curve became unstable and actually went negative. This was due to poorer correlations of PSD ordinates with SR in this wavelength region—possibly due to tape deck vibration in the test vehicle.

Equation 8 was then used because it disallows negative $\hat{\beta}_i$. Polynomials of orders 1 through 4 were used as shown in Figure 4. It appears that at least a second-order polynomial is necessary, although little change in the $\hat{\beta}_i$ distribution resulted from increasing the order beyond 2 (see inset of Fig. 4). One would infer from Figure 4 that 5- to 10-ft waves are of special importance in the determination of riding quality if vehicle speed were held constant throughout the test series. This was not true for these data (vehicle speed ranged from 30 to 50 mph); therefore, these results should be augmented with more extensive and better controlled experimental data. Notice also that no information concerning important wavelength ranges can be obtained from conventional multiple regression (Fig. 3).

Problems with this particular set of experimental data notwithstanding, if these results are taken as valid, the importance of this wavelength range can be rationalized as follows: Shake table tests of a typical vehicle resulted in maximum reactive force from tires when input frequencies were near 15 cps (14, 15, 16). This is the frequency range generated by 6- to 8-ft waves in a vehicle traveling at typical highway speeds

Figure 1. Kth degree polynomial.

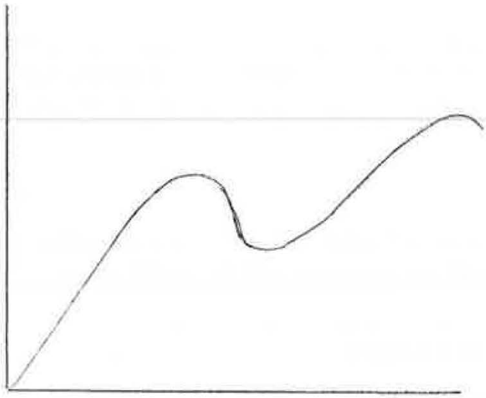


Figure 2. Correlation of frequency 0.22 with neighboring frequencies.

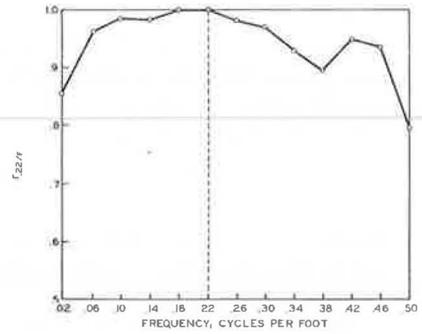


Figure 3. Frequency weights determined by using Eq. 1.

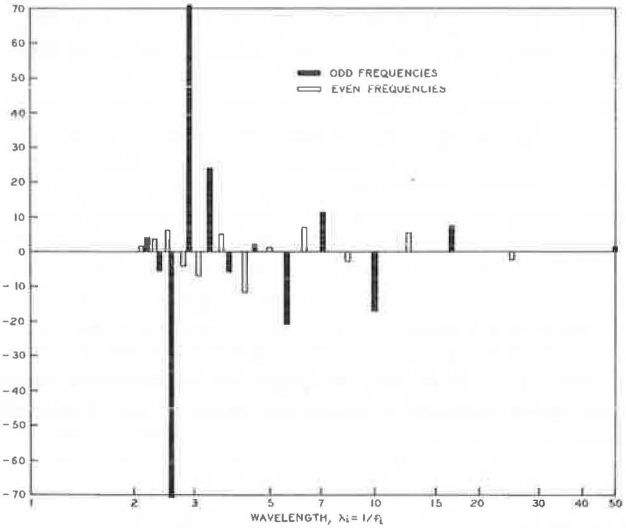
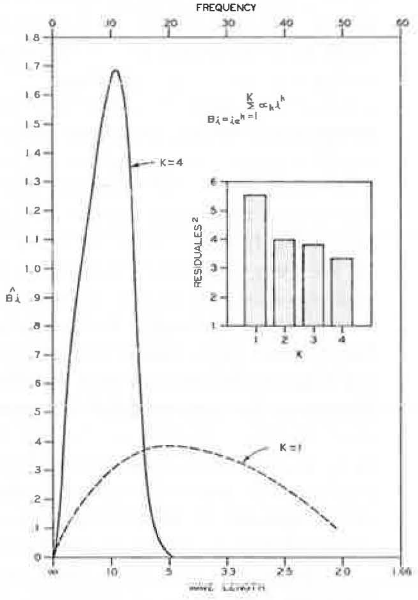


Figure 4. Frequency weights determined by using Eq. 8.



near 70 mph. Moreover, it is reasonable to assume that most automobiles have a similar response function despite differences in weight or suspension system. This is because the major determinant of vehicle response, the ratio of suspension spring constant to vehicle mass, remains generally constant over a wide range of vehicles. In addition, the 6- to 8-ft wavelength band is of sufficient width to accommodate some change in the mass to spring constant ratio.

An important implication of these findings concerns the design of a standardized ride-quality measuring system. The 6- to 8-ft wavelength band or even the 2- to 8-ft band can be easily measured by uncomplicated instruments. Such a device would be much simpler than the GM profilometer and would involve a single accelerometer and simple processing circuits. The statistic returned would be a single number representing average power in the 5- to 10-ft band. It would be possible to continuously display this statistic during a profile run to locate areas of excessive roughness. Although details of this system are not presented here, the authors will supply information on its design.

CONCLUSION

Because of high intercorrelations among amplitudes or road profile spectra, conventional multiple regression techniques should not be used to correlate frequency bands with subjective response to pavement roughness. In particular, one would get misleading estimates of the relative importance of the various bands as far as human response to road roughness is concerned. It would, therefore, become very unlikely that one would be able to separate the effects of the several frequencies on subjective response. Stepwise regression procedures merely exacerbate the problem because the statistics such as partial F-tests and partial correlation coefficients used in these procedures will reject important frequencies highly correlated with frequencies already included as independent variables. Consequently, these procedures discriminate against frequencies less correlated with subjective response than those frequencies selected earlier because of only slightly higher correlation. In short, with conventional regression analysis, one would expect to get drastically different estimates of the frequency coefficients from sample to sample. The method of overcoming the problem put forth in this paper is to impose restrictions on the frequency coefficients. An obvious restriction is the requirement that the coefficients must lie on a polynomial of specified order. The order can be arbitrarily set by the investigator (particularly if he has some prior knowledge of the coefficient distribution), or it can be determined by a selection procedure such as stepwise multiple regression. The coefficients of the polynomial are considerably fewer in number than those of the conventional regression and are, therefore, more reliably estimated. These parameters are then used with the specified polynomial to estimate the frequency coefficients originally sought.

Sample data provide an example that shows how conventional multiple regression fails to produce a reasonable distribution of the frequency coefficients. Application of the polynomial procedures to these same data provides an initial coefficient distribution that peaks at about 8 cycles/ft. Further studies based on more and better data are needed to establish the validity of this distribution.

REFERENCES

1. Spangler, E. B., and Kelly, W. J. GMR Road Profilometer—A Method for Measuring Road Profile. Highway Research Record 121, 1965, pp. 27-54.
2. Darlington, J. R. Evaluation and Application Study of the General Motors Corporation Rapid Travel Profilometer. Michigan Department of State Highways, Rept. R-731.
3. Walker, R. S., and Hudson, W. R. Practical Uses of Spectral Analysis With Surface Dynamics Road Profilometer. Highway Research Record 362, 1971, p. 113.
4. Holbrook, L. F. Prediction of Subjective Response to Road Roughness by Use of the Rapid Travel Profilometer. Highway Research Record 291, 1969, pp. 212-226.
5. Hutchinson, G. B. Principles of Subjective Rating Scale Construction. Highway Research Record 46, 1964.

6. Blackman, R. B. Data Smoothing and Prediction. Addison-Wesley, 1965, pp. 136-152.
7. Goldberger, A. S. Econometric Theory. John Wiley and Sons, 1964, pp. 192-194.
8. Goldberger, A. S. Topics in Regression Analysis. Macmillan, 1968, pp. 132-133.
9. Johnston, J. Econometric Methods. McGraw-Hill, 1963, pp. 201-207.
10. Draper, N., and Smith, H. Applied Regression Analysis. John Wiley and Sons, 1966, pp. 163-173.
11. Wonnacott, R. J., and Wonnacott, T. H. Econometrics. John Wiley and Sons, 1970, pp. 309-312.
12. Almon, S. The Distributed Lag Between Capital Appropriations and Expenditures. Econometrica, Vol. 33, Jan. 1965, pp. 178-197.
13. Kmenta, J. Elements of Econometrics. Macmillan, 1971, pp. 441-495.
14. Quinn, B. E., and DeVries, T. W. Highway Characteristics as Related to Vehicle Performance. HRB Bull. 250, 1960, pp. 20-36.
15. Quinn, B. E., and DeVries, T. W. Effects of Pavement Condition on Dynamic Vehicle Reactions. HRB Bull. 328, 1962, pp. 24-31.
16. Quinn, B. E., and VanWyk, R. A Method for Introducing Dynamic Vehicle Loads Into Design of Highways. Highway Research Record 40, 1961, pp. 18-20.