A Statistical Analysis of Speed Density Hypotheses

JOSEPH S. DRAKE, Research Engineer, Consad Research Corporation;
JOSEPH L. SCHOFER, Graduate Student, Northwestern University; and
ADOLF D. MAY, JR., Associate Professor of Civil Engineering, University of California

•AN UNDERSTANDING of interrelationships among basic characteristics of vehicular traffic flow, such as volume, speed, and density, is of prime importance to the practicing traffic engineer. From the standpoint of design, a knowledge of high-flow-rate characteristics is required for the prediction of highway capacity. Those concerned with traffic operations are faced with the problem of providing an adequate level of service; this calls for an understanding of the entire range of relationships. Development of flow control and ramp metering techniques must be based on these functional interrelationships under high-density conditions. Finally, any efforts toward developing new roadway and vehicular technologies for the purpose of improving flow characteristics will necessarily stem from an understanding of the present relations.

During recent years, a number of hypotheses concerning these interrelationships have been proposed. Some researchers have relied almost completely on the statistical analysis of data for developing functions, while others have begun with a purely theoretical concept, from which relations were derived and later tested.

It is the purpose of this paper to examine currently available hypotheses, using rigorous statistical procedures to test them with a common set of freeway flow data. The degree to which the various functions could be made to replicate the full range of the data was the basis for determining which relations might be better than others, and possibly for deciding that some should be rejected entirely.

Primary reliance was placed on the results of the techniques of statistical analysis. Where these were inadequate, however, sound judgment was applied in order to select the most workable functional relations.

The approach selected was the regression of speed of flow against density, since the other basic relations, volume-density and speed-volume, cannot be transformed into linear functions for all hypotheses. The techniques of linear regression analysis are both simpler and more highly developed than those of the nonlinear analysis. A visual examination of the volume-density and speed-volume relationships derived by conversion from each of the speed-density regression equations was also performed.

SITE AND DATA SELECTION

To investigate the relation between speed and density properly, it is necessary to sample traffic flow characteristics over the range of all possible densities. There are at least two feasible sampling procedures available: the use of fixed time periods, such as a minute or an hour to represent a data point, or the use of a specific number of consecutive vehicles for each point.

Data collected during the fixed time period represent flow characteristics during that period, but the number of vehicles included in the samples varies over the day. During free flow operations a minute sample could contain so few vehicles as to be statistically unreliable. On the other hand, the use of a short time increment, such as a 1-min period, helps to insure that a nearly instantaneous picture of flow conditions is provided. Vehicles in the sample can be no farther apart than 1 min in time, resulting in a limitation in the possible changes in flow characteristics which might take place during the

Paper sponsored by Committee on Characteristics of Traffic Flow and presented at the 45th Annual Meeting.

sampling period. This would imply that the dispersion of phenomena within a sample would be minimal.

The vehicle-based sample insures that each data point represents the same number of vehicles. Over the day, however, the amount of time included in the sample varies considerably. A series of 100 vehicle samples, taken on the Eisenhower Expressway in Chicago during a 24-hr period, covered time spans ranging from approximately 3 to 50 min. The flow characteristics represented by an average over 50 min may not give a good indication of the true relationships among vehicles on the facility. The behavior of vehicles passing a point during a preceding 50-min period may have little or no effect on current operations.

While neither sampling basis is entirely satisfactory, the 1-min time sample is probably best for examining vehicular interactions with respect to density of flow. In considering the freeway as a system generating some distribution of phenomenological output, as represented by vehicular flow characteristics, the regularity of the time-based sample is more satisfying intuitively than the continuously varying output rate of the vehicle-based sample. The detection equipment can be set to "look at" the expressway every minute and "photograph" its operating characteristics, rather than waiting until sufficient vehicles have passed and developing a time average.

A series of 1,224 one-minute observations was recorded with the pilot detection system of the Chicago Area Expressway Surveillance Project. At the end of each minute the minute volume, time-mean-speed, and occupancy were obtained, and density was computed from the volume and time-mean-speed. Data were collected in the middle lane of the three-lane westbound roadway at Harlem Avenue on the Eisenhower Expressway. The observations were made between 1:00 and 6:00 p.m. on four weekday afternoons under dry weather and normal traffic conditions. Thus, many of the data represented peak hour characteristics, while few were associated with the very lowest density range.

It appears appropriate to describe the location of the study site in relation to adjoining portions of the expressway because of its effect on the flow interrelationships. For example, if a location is studied which is not operating at capacity due either to lack of demand or an upstream bottleneck, high density measurements obviously cannot be obtained. On the other hand, if the study site is located just upstream of a severely limiting bottleneck, mid-range density measurements may be limited or almost nonexistent. Hypothetical speed-volume-density relations describing these conditions are shown in Figure 1.

The study site (Harlem Avenue) is located one-half mile upstream from a bottleneck. The capacity of the downstream bottleneck, however, is only slightly less than the capacity of the study site, and therefore density observations can be obtained over a substantial portion of the density range.

A disturbing feature of the original data set was that points seemed to occur in bunches (Fig. 2), particularly in the 20 to 45 and 65 to 84 veh/mi ranges (1). Early attempts at regression analysis with these data indicated that the concentration of points in a few areas resulted in equations which did not adequately represent the entire range of the relationship. Sample densities varied considerably over the range of the independent variable, resulting in a biased estimate of regions represented less than adequately in the sample.

Once the objective of any research project has been defined, a sampling procedure can usually be designed so as to efficiently collect a relevant set of unbiased data. For an investigation of mean free speed—the Y-intercept of the relation—a sample whose mean density is close to zero could be chosen. Conditions around jam concentrations could be studied by sampling only at the lowest congested speeds. Because this analysis was aimed at testing a number of functional relations over the full range of operating characteristics, it was necessary to give equal consideration to all flow conditions which might occur.

One solution to the problem would have been to return to the field and fill in areas which were under-represented in the original sample. The variable over which the sample was being taken was density, in vehicles per mile. An unbiased sampling pro-

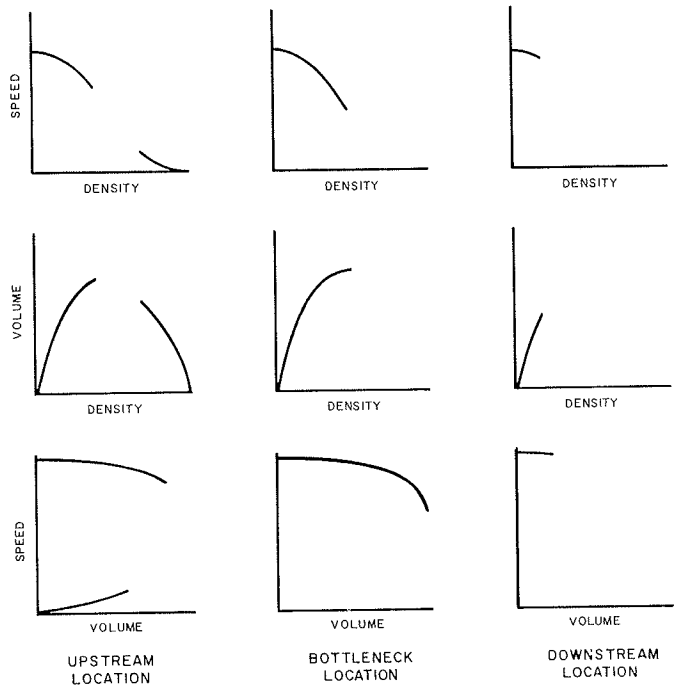


Figure 1. Hypothetical relationships.

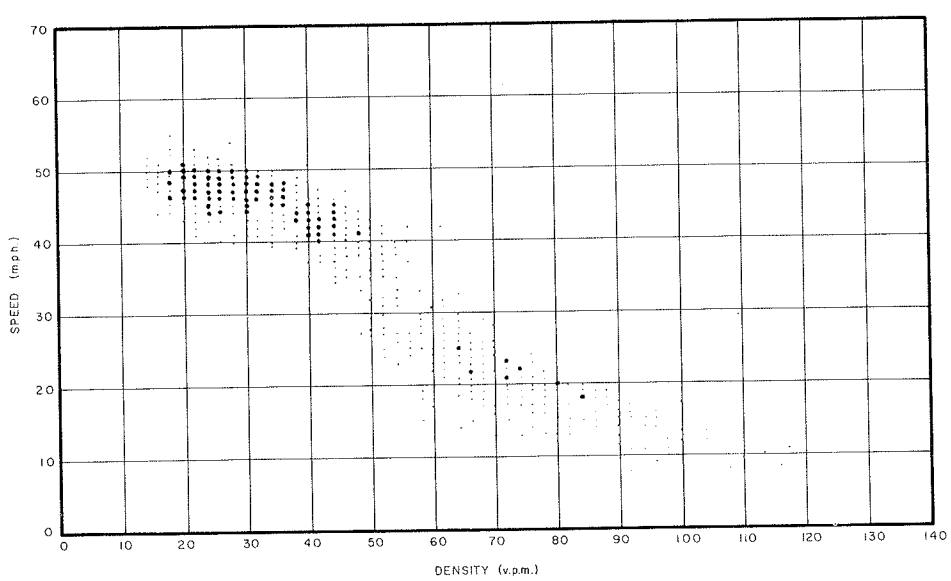


Figure 2. Scatter diagram of total speed-density observations.

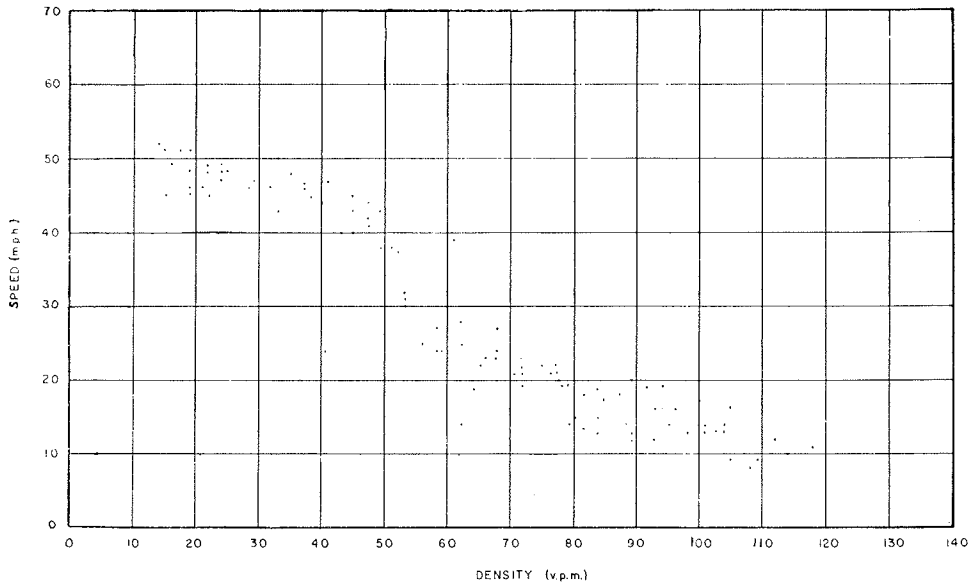


Figure 3. Scatter diagram of selected sample of speed-density observations.

cedure would require that each possible density value should have an equal probability of being in the sample. To collect data by turning on the detectors for a period of 5 hours per day automatically biases the analysis against regions of the relation whose densities rarely occur during the sampling hours.

Because of the additional expense entailed in collecting data under conditions similar to the original sample in order to fill in the sparse areas, another technique was devised. The set of 1,224 one-minute observations, each on a single punch-card, was arranged in the order of increasing density and divided into ranges of approximately 5 veh/mi. The number of observations falling in the most sparse 5-veh/mi range was determined, and a like number of data points was randomly sampled from each of the other ranges. This resulted in a sample of 118 observations, with each 5-veh/mi range containing an equal number of points. The procedure resulted in a considerable degree of uniformity of sample density throughout the range of the independent variable, as shown in Figure 3, particularly when compared with Figure 2. Some characteristics of the sample are as follows:

Lowest density	14.2 veh/mi
Highest density	118.4 veh/mi
Average density	61.9 veh/mi
Lowest speed	8.2 mi/hr
Highest speed	52.3 mi/hr
Average speed	29.6 mi/hr

ANALYSIS OF RELATION BETWEEN TIME MEAN SPEED AND SPACE MEAN SPEED

Data were collected by using motion and presence detectors which measured speed, lane occupancy and volume. Individual speeds were averaged over each minute, resulting in a time-mean-speed value. Density was computed based on the fundamental relation, volume = speed × density.

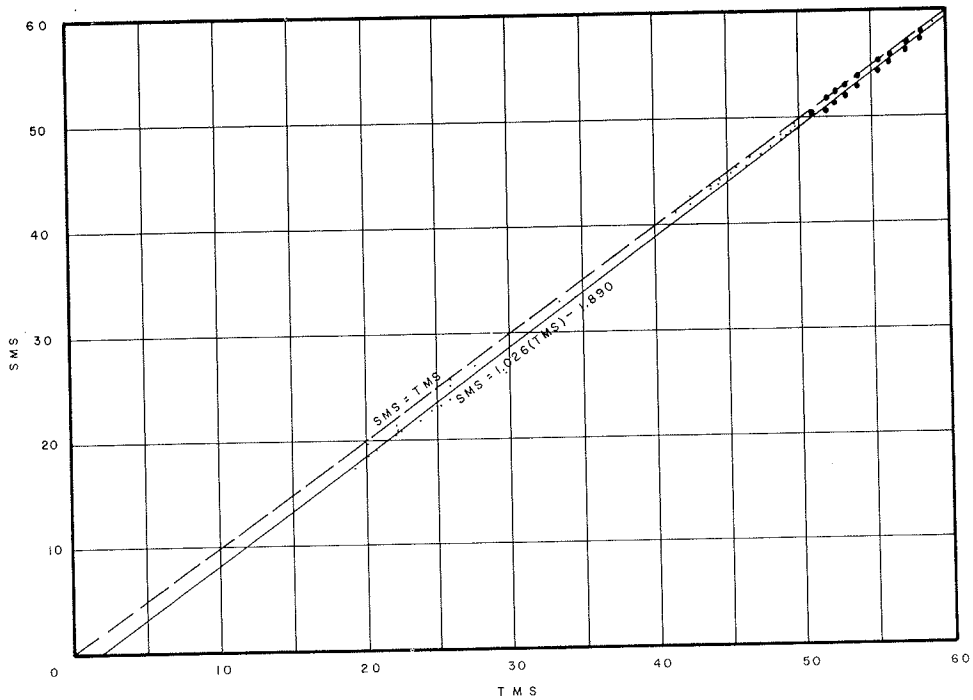


Figure 4. Time-mean-speed and space-mean-speed relationship.

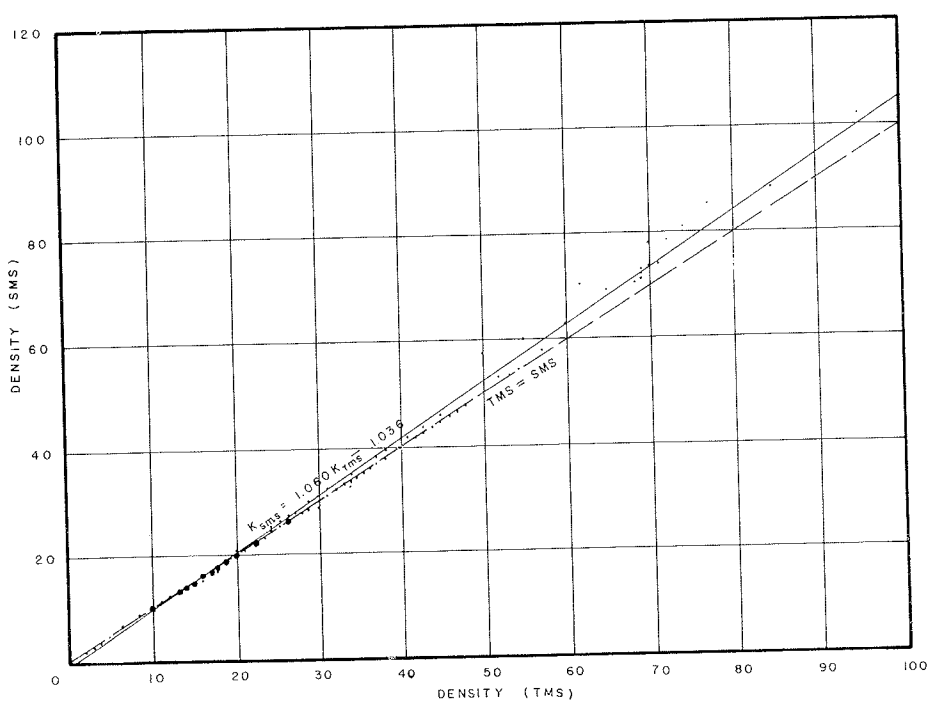


Figure 5. Time-mean-speed density and space-mean-speed density relationship.

Wardrop (2) has shown that this relation holds only for space-mean-speed, the speed computed from the mean of vehicle travel times over a specified distance; he demonstrated that the two speed measurements could be related as follows:

$$TMS = SMS + \frac{(\sigma SMS)^2}{SMS}$$

The use of the collected time-mean-speed data, then, introduces a bias into the values of both speed and density. To investigate the magnitude of this bias and possibly develop a procedure for correcting the data points, a second study was conducted at the same location, in which 224 groups of 100 consecutive vehicles were analyzed. Time-mean-speed, space-mean-speed, standard deviations of both time- and space-mean-speeds, and densities calculated from both speeds for each group were determined.

A regression analysis of time-mean-speed and space-mean-speed resulted in the following equation:

$$SMS = -1.88960 + 1.02619\ (TMS)$$
$$r^2 = 0.99834$$
$$\sigma = 0.38404$$

The nearness of the slope of this relation to 1.00 indicates the close correspondence between the two speed measures (Fig. 4). The maximum difference between the two speed measures occurs at zero speed and is 1.9 mph. As speed increases the difference becomes smaller, and at 72 mph the time-mean-speed and space-mean-speed would be equal. In addition, there was no significant difference in the means of the two speed measurements at the 99 percent confidence level.

A similar investigation of the relation between the densities computed from time-mean-speed and space-mean-speed gave the following regression line:

$$K_{sms} = -1.03638 + 1.06018\ (K_{tms})$$
$$r^2 = 0.99712$$
$$\sigma = 0.94337$$

The two measurements were again found to be quite close, with a maximum difference on the order of 1 veh/mi (Fig. 5). There was no significant difference in the means of the two parameters at the 99 percent confidence level.

Tests of the slopes of both of these equations revealed that they were significantly different from 1.000 at the 95 percent confidence level; i.e., the data indicate the existence of a real difference between space-mean-speed and time-mean-speed, a result not unexpected considering the theoretical work of Wardrop.

In spite of these results, it was felt that any reduction in error achieved through conversion of data from a time-mean-speed to a space-mean-speed basis would be unimportant when compared to the variation introduced by the inaccuracies of the data collection equipment. For example, the range of error in speed measurement over a 1-min sample has been estimated at ± 1 mph; for an individual vehicle this broadens to 4 or 5 mph.

The conclusion from this phase of the analysis was that time-mean-speed and the corresponding density would be used for the investigation, since this would introduce only a minimal amount of error into the results of the study.

PRESENTATION OF HYPOTHESES

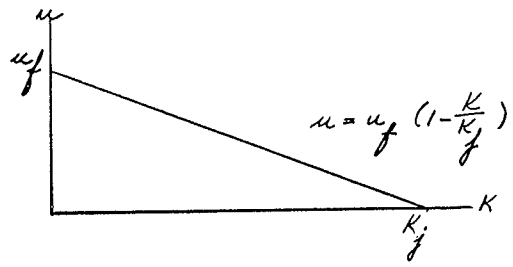
Seven proposed speed-density relationships were selected for examination. Some have no theoretical background, being based primarily on the researcher's observation

of a particular data set. Others have considerable support; four have been shown to be directly related to specific car-following rules (3, 4).

Linear Forms

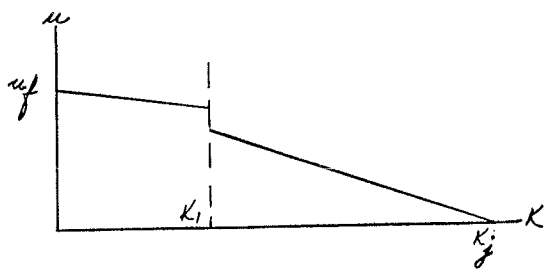
Early work in the analysis of speed-density relationships, notably that by Greenshields (5), was devoted to the investigation of a continuous linear form, as shown below:

HYPOTHESIS I

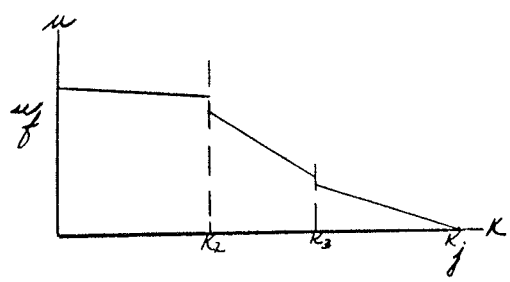


Prompted by the failure of the continuous relation to fit data at all density levels, some researchers (6) have devoted attention to the possibility of the existence of two or three distinct zones, each characterizing a different driver behavior pattern. These hypotheses follow; additional parameters are introduced.

HYPOTHESIS II



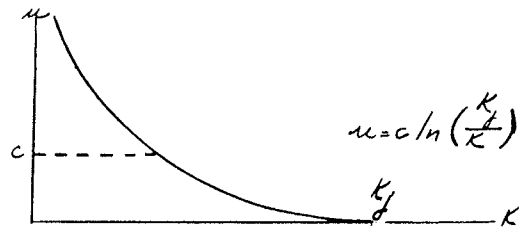
HYPOTHESIS III



The assumption of linearity was adopted because of its simplicity, and although there is no strong theoretical basis, the function may be derived from the car-following rule in which sensitivity of the following vehicle is inversely proportional to the square of the inter-vehicle spacing.

Greenberg's Exponential Curve

In 1959, Greenberg (7) postulated another speed-density form, based on a hydrodynamic analogy. Treating the traffic stream as a perfect fluid, he combined the equations of motion and continuity for one-dimensional compressible flow and arrived at the following form:

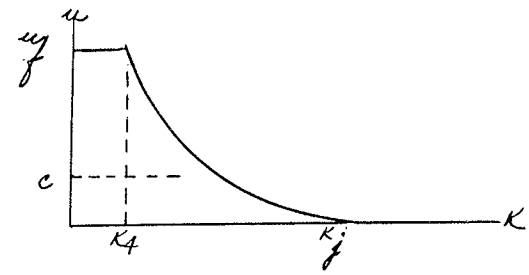


The parameter c is defined as the speed at which volume experiences its maximum with respect to density.

Gazis, Herman and Potts (3) have shown that this relationship can be independently derived from their microscopic car-following theory for the case in which the sensitivity of the following vehicle is inversely proportional to the spacing between vehicles.

Because this equation is not defined at $k = 0$, it will be tested in the following more realistic form; again, another parameter is introduced:

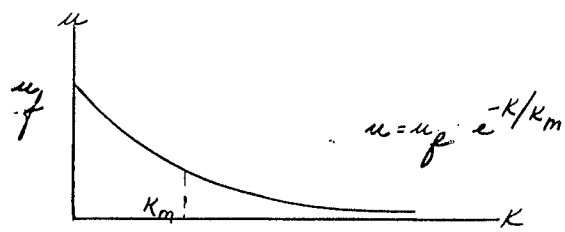
HYPOTHESIS IV



Underwood's Transposed Exponential Curve

Apparently disturbed by the failure of Greenberg's curve to remain finite at zero density, Underwood (8) suggested that perhaps the infinity asymptote should be along the density scale, since almost any jam always has some finite movement:

HYPOTHESIS V

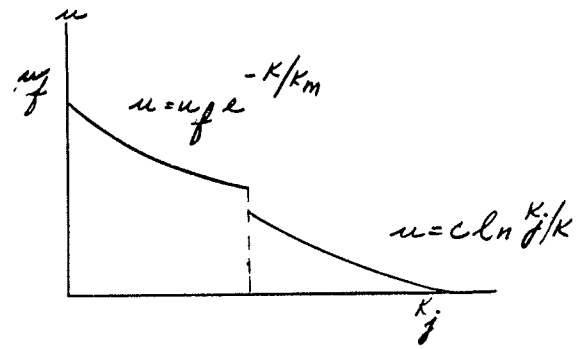


The parameter k_m is defined as the density at which volume experiences its maximum with respect to density. This relation can be traced to the car-following rule where sensitivity is directly proportional to the speed of the following vehicle and inversely proportional to the square of the spacing.

Edie's Discontinuous Exponential Form

The frequent occurrence of a discontinuity in empirical volume-density curves in the vicinity of optimum density and the poor correlation of low-density data with Greenberg's hypothesis led Edie (9) to postulate in 1960 that a different car-following law applies to noncongested operation, namely, that the sensitivity is proportional to velocity and to the reciprocal of the square of spacing between vehicles. Consequently, he derived the following hypothesis:

HYPOTHESIS VI

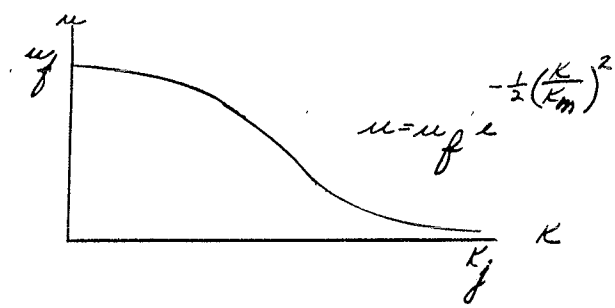


This form is merely a combination of the Greenberg and Underwood relations.

Further Suggestions

May has pointed out that speed-density data from the Eisenhower Expressway in Chicago tend to exhibit concavity at low densities. Guided by these observations, the following bell-shaped curve has been suggested:

HYPOTHESIS VII



This hypothesis has no theoretical foundation.

DISCUSSION OF STATISTICAL ANALYSIS TECHNIQUES

The basic purpose of this research was to make decisions regarding the relative merits of the seven speed-density hypotheses as applied to the study data. The essence of any experimental design lies in the development of rigorous tests which will permit exacting binary decisions to be made. The researcher must rely on his own intuition to decide what is expected of his hypothesis; he must then translate these expectations into testable form.

The attitude taken in the formulation of such tests in this research endeavor was based on the traditional philosophy of rejection. In general, such an approach translates a general hypothesis into several working hypotheses which express the various expectations of the researcher. These working hypotheses are structured in a test framework and are designed to maximize the "falsifiability" of each hypothesis. The failure of any one working hypothesis to meet the researcher's expectations necessarily constitutes rejection of the general hypothesis, and suggests a restudy of the theory underlying that hypothesis.

The use of such an approach was particularly warranted in this study, for two reasons.

1. Traditional applications of regression techniques tend to ignore opportunities for verification. The goodness-of-fit statistic (r^2) is generally quoted, and frequently, a test is conducted to shed some light on the significance of the regression in reducing variation in the dependent variable. Viewed in this limited context, regression analysis

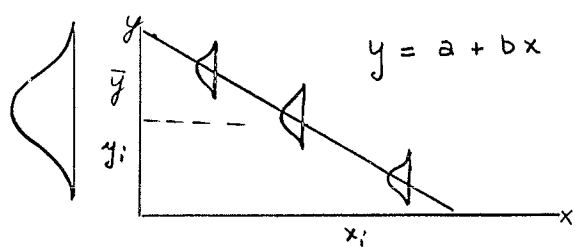
tends to become a self-fulfilling prophecy. The statistical design in this study has been approached with the intention of formulating tests not only of the significance of the regression considered by itself, but also of the calibrated regression model's ability to predict independently derived parameters.

2. The basic purpose here was to compare the seven models with each other. Regression theory presently offers no technique for comparing directly two regression models applied to the same sample. The only alternative, then, is to develop tests which may be applied against each hypothesis without favor, and to insist that the failure of any one hypothesis to meet all such tests necessarily constitutes elimination of that hypothesis from further consideration. This idea is consistent with the notion of comparing hypotheses in terms of their application over the entire domain of density.

Fundamental Concepts of Regression Analysis

Before introducing those statistical concepts which have been developed to cater to the peculiar nature of this research, it is worthwhile to review the traditional setting of regression analysis. Suppose measurements of a variable y are recorded and considerable variation in y about its mean $\bar{y}$ is observed. Suppose that y is influenced by an independent variable x according to some relation; if it is possible to isolate the effects of x on y the dispersion of y might be considerably reduced. Linear regression analysis is a convenient means for accomplishing this desired isolation.

The following diagram, which depicts this situation, provides a convenient review of terminology and concepts which are vital to what follows.



The diagram illustrates the manner in which the allowance for variation in x reduced the variation in y. Such allowance reduced the difference between an individual observation, y_i, and the mean, $\bar{y}$, to the difference between y_i and precited Y_i. The equation $Y = a + bx$ is fitted to the data according to the criterion of minimizing the sum of the squares of deviations about the regression line. More formally, the following equations are solved for the parameters a and b:

$$\frac{\partial}{\partial a}\left[\sum_{i=1}^{n} (y_i - Y_i)^2\right] = 0$$

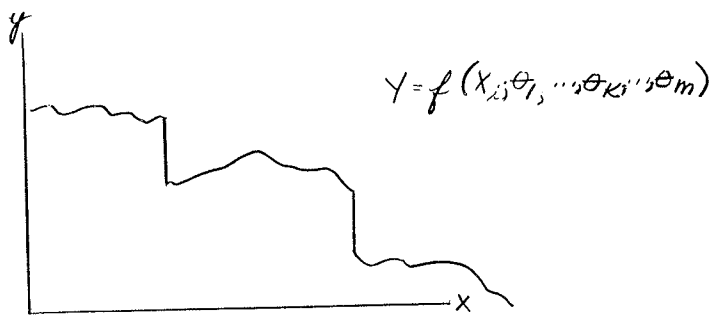
$$\frac{\partial}{\partial b}\left[\sum_{i=1}^{n} (y_i - Y_i)^2\right] = 0$$

In this fashion, the sum of squares (of deviations) about the mean is decomposed into two components: the sum of squares due to the regression, and the resulting sum of squares about the regression. In more formal terms,

$$\Sigma(y_i - \bar{y})^2 = \Sigma(Y_i - \bar{y})^2 + \Sigma(y_i - Y_i)^2$$

Composite Statistics for Discontinuous Regressions

The following diagram shows the same type of problem in more general terms. Some general relationship y = f(x) is to be fit to a set of observations. The relationship is not susceptible to textbook approaches, because of its complexity. Not only is it nonlinear, but it also suffers discontinuities. For the purely nonlinear case, with no discontinuities, it is frequently feasible to perform some transformation upon y and/or x which reduces the relationship to linear form. Given discontinuities, however, two fundamental questions arise: Can a discontinuous relationship be described in terms of single statistics? For example, can one r^2 value be quoted for the entire relationship? And, can one decide whether or not discontinuities exist in the true relationship between y and x?



Ideally, these questions might be answered by fitting the function to data according to the least squares criterion described for the linear case. This would be accomplished by minimizing the expression

$$\sum_{i=1}^{n} [y_i - f(X_i; \theta_1, \ldots, \theta_k, \ldots, \theta_m)]^2$$

with respect to each parameter θ_k.

Even for very simple functions, such techniques lead to mathematics too cumbersome to handle. The alternative approach would be to treat each continuous regime in its own right, and then treat the result as a whole. More specifically, the relationship between Y and X for each regime would be fitted (after transformations, if necessary) to its associated data according to its own least squares criterion. The results may then be reassembled, and single composite statistics may be computed for the entire discontinuous regression in terms of the values of y_i, Y_i, and $\bar{y}$. This approach assumes that minimizing the sum of squares about the regression line for each regime is equivalent to the more desirable but insurmountable task of minimizing the sum of squares about the entire discontinuous regression. Such an approach was applied in this research.

Testing of Multi-Regime Hypotheses

Having devised a method for quoting composite statistics for discontinuous regression models, the question regarding how one might decide whether the true relationship between y and x is discontinuous or not still remains. The approach to this problem was derived from earlier work by Quandt (10) on the treatment of two-regime linear regression hypotheses. He discusses two problems: (a) the estimation of the parameters of a two-regime hypothesis, and (b) the possible methods for testing whether the data do in fact obey two separate regimes. The treatment of the first problem, i.e., locating the breakpoint between regimes, precludes the exercise of any tests.

The location of the breakpoint t for two linear regimes is ascertained by maximizing the likelihood function of the entire sample, formed by multiplying together the (normal) frequency functions of the individual error variances, σ_1 and σ_2. Upon application of conventional maximization procedures for obtaining best estimates of parameters a_1, a_2, b_1, b_2 and σ_1 and σ_2, the following expression results:

$$L(t) = -(\tfrac{1}{2} + \ln\sqrt{2\pi})\, T - t \ln\hat{\sigma}_1 - (T - t) \ln\hat{\sigma}_2$$

where T is the total number of observations, t is a discrete variable representing the serial rank of each observation (with respect to x), and σ_1 and σ_2 are functions of t:

$$\hat{\sigma}_1^{\,2} = \frac{1}{t} \sum_{i=1}^{t} (y_i - \hat{a}_1\ x_i - \hat{b}_1)^2$$

$$\hat{\sigma}_2^{\,2} = \frac{1}{T-t} \sum_{i=t+1}^{T} (y_i - \hat{a}_2\ x_i - \hat{b}_2)^2$$

The problem then, is to find the value of t which maximizes L(t). Because t is a discrete variable, and anticipating several local maxima, the only feasible method for determining an optimal t, t*, is a systematic search procedure. This requires the performance of a complete regression analysis on each regime, for every selected value of the breaking point, t.

Having established the value of t*, Quandt (11) examined various alternatives for testing the hypothesis that the data obey two separate regimes. For this study, his F-test recommendation was adopted. Essentially, the resulting regression equation for regime I is applied to observations for $i \leq t$, and then to observations for $i > t$. The quantities

$$\eta_{I-I} = \frac{1}{\sigma^2} \sum_{i=1}^{t} (a_1 + b_1\ x_i - y_i)^2$$

$$\eta_{I-II} = \frac{1}{\sigma^2} \sum_{i=t+1}^{T} (a_1 + b_1\ x_i - y_i)^2$$

are independently distributed as χ^2 with t - 1 and T - t - 1 degrees of freedom, respectively. Their ratio therefore follows the F-distribution with the given degrees of freedom. A similar ratio is obtained by applying the resulting regression equation for regime II to observations in each regime, and forming the appropriate F-ratio in terms of $\eta_{II-1} + \eta_{II-2}$. If either F-ratio exceeds F-critical, the hypothesis of one continuous regime will be rejected. The power of this test depends on how close the true breakpoint t* is to the endpoints of the sample.

To meet the demands of hypothesis III, Quandt's work was extended to the case of three regimes, with the following likelihood function resulting:

$$L(t) = -(\tfrac{1}{2} + \ln\sqrt{2\pi}\ T - r \ln\hat{\sigma}_1 - (s - r) \ln\hat{\sigma}_2 - (T - s) \ln\hat{\sigma}_3$$

The appropriate tests were formulated by applying each of the three regression equations to its own regime and comparing it to the remaining two regimes. Six F-ratios

were developed; the failure of any one ratio to exceed F-critical constituted the basis for rejection of the three-regime hypothesis.

Tests for All Hypotheses

Regression analysis offers several statistical tests which may be exercised according to the purpose of the investigation. Two such tests were adopted for this study: the F-test for significance of the regression and the t-test for determining whether the slope differs from zero. For continuous regression models, these two tests are equivalent. For discontinuous hypotheses, however, the t-test was applied to the slope of each regime, while the F-test was formulated in terms of the entire regression, according to the methods discussed previously for quoting composite statistics for multi-regime hypotheses.

The use of such tests in themselves does not constitute much verification power. To promote independent verification, the predicted values of mean free speed for each hypothesis were tested for significant difference from an independent measurement of that parameter.

ANALYSES OF RESULTS

Break-Point Analysis

The first stage of the analysis was the establishment of the break-points for the four discontinuous hypotheses according to the technique described. For the three hypotheses which required only one break-point, separate regression analyses were performed for each of the possible ways in which the observations could be broken into two groups, using a minimum separation of 5 veh/mi between alternative break-points. The resulting likelihood functions are shown in Figures 6, 7 and 8, for hypotheses II, IV, and VI, respectively.

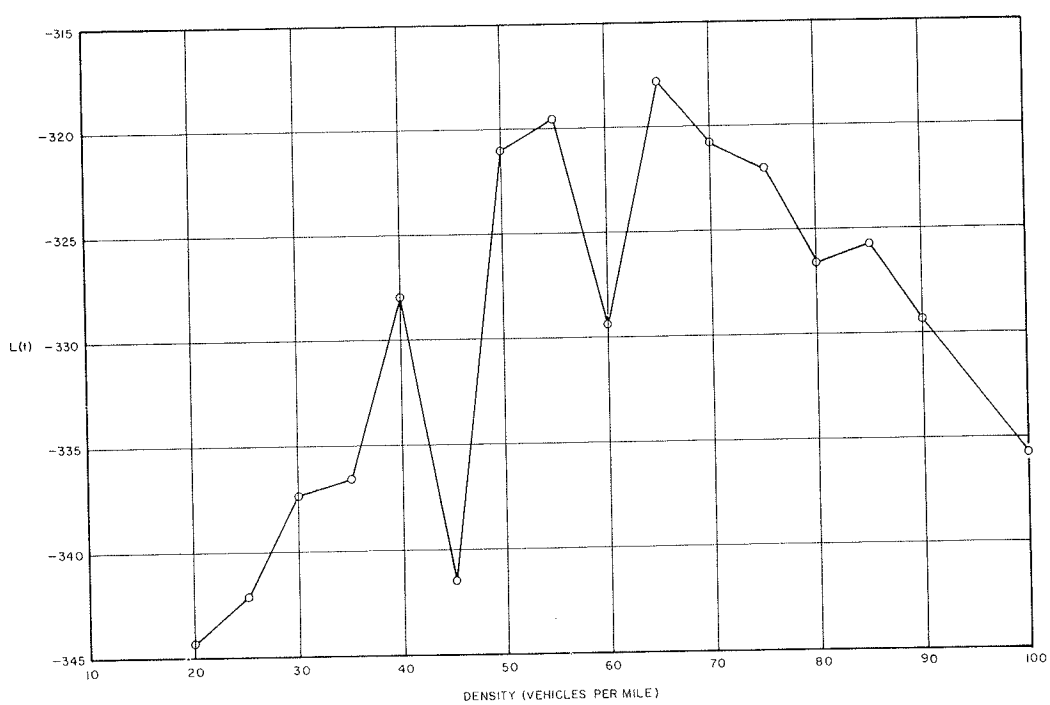


Figure 6. Likelihood function for hypothesis II (2 Regime Linear).

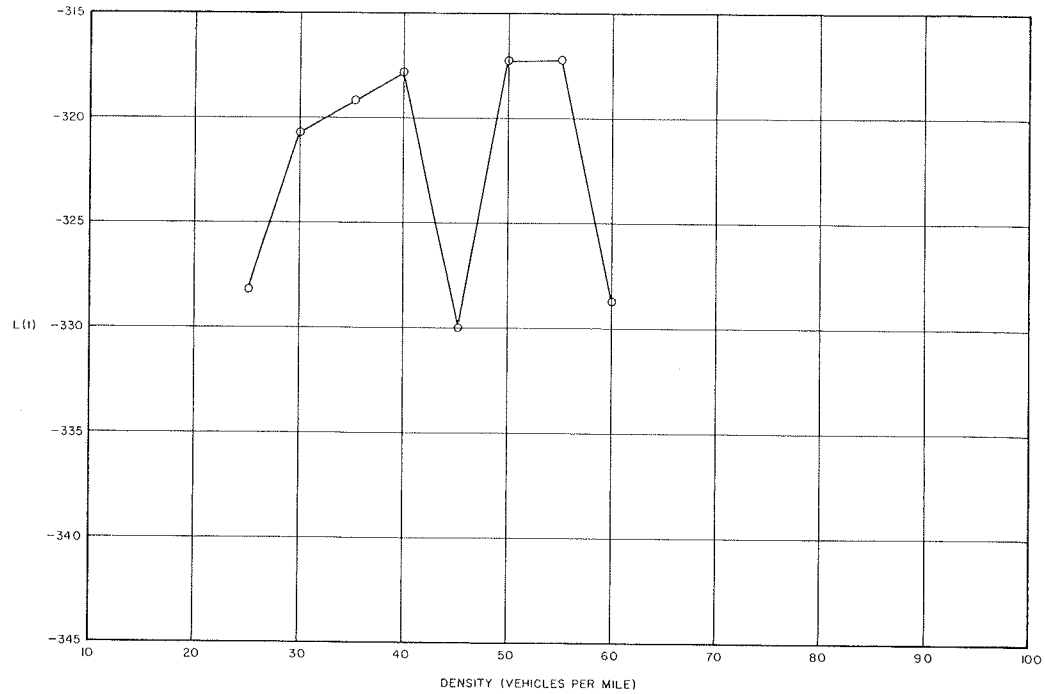


Figure 7. Likelihood function for hypothesis IV (Greenberg).

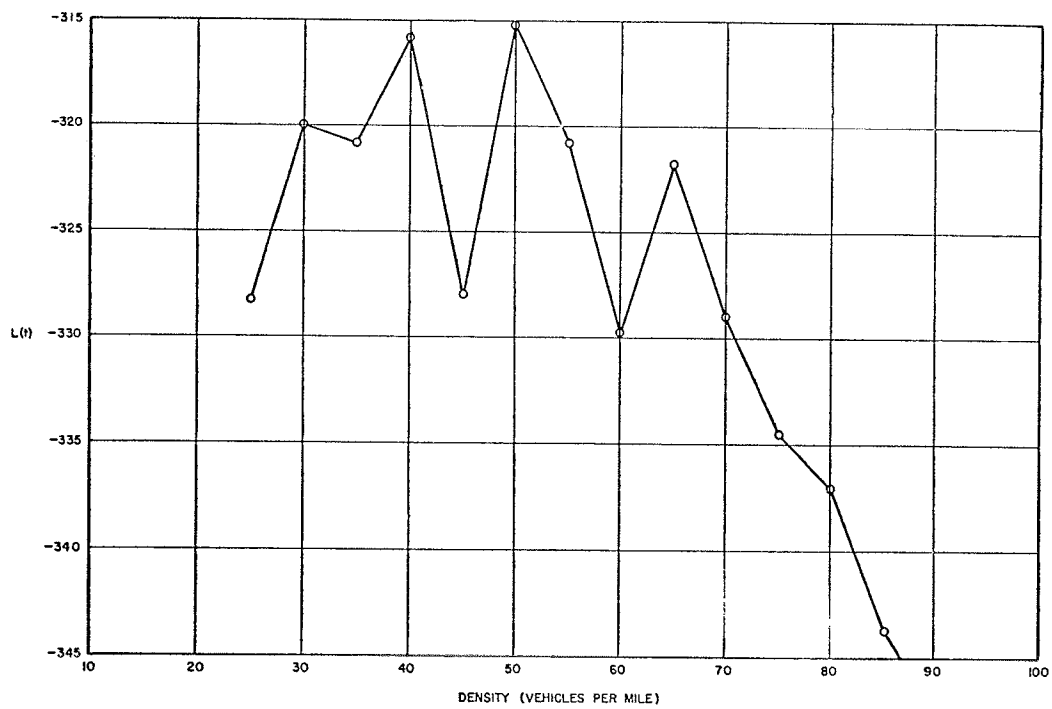


Figure 8. Likelihood function for hypothesis VI (Edie).

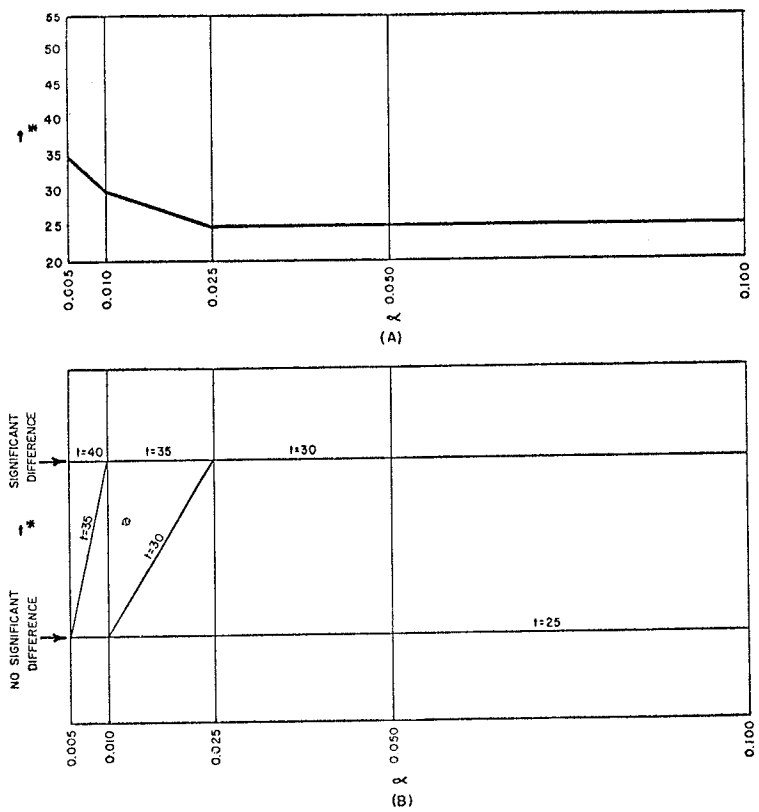


Figure 9. Comparsion of slope for low-density regime of hypothesis IV to zero for various confidence levels.

The likelihood functions for the Edie relationship and the two linear regimes hypothesis behaved quite similarly. While the optimal break-points of these two functions differed appreciably (65 veh/mi for hypothesis II and 50 veh/mi for hypothesis VI), both likelihood functions showed two local peaks in the middle range of density, and these peaks did not differ much in value for each hypothesis. Furthermore, an abrupt drop in the function occurred between these peaks, suggesting that L(t) was quite sensitive to t.

This apparent sensitivity had no important bearing on the use of the likelihood function analysis because the optimal density value was the sole matter of interest. Had the analysis investigated L(t) more accurately, e.g., for intervals of 2.5 veh/mi, it is conceivable that the maximum values of L(t) might have been found at such intermediate points, and that these maximum values might have been considerably greater than the values derived in Figures 6 and 8. The goal of this phase of the investigation was not to determine the global maximum value of L(t), but merely to establish the proper break-point. The resulting location of this point must be interpreted with a maximum tolerance of ± 2.5 veh/mi.

The break-point analyses for hypotheses II and VI differed in one important respect. The Edie hypothesis required a different transformation for each of its two regimes in order to permit the use of linear regression analysis. The low-density regime was regressed in terms of $Y = \ln u$ and $x = k$, while the high-density regime was regressed in terms of $Y = u$ and $x = \ln k$. The resulting error variances were dimensionally incompatible and could not be used directly in computing L(t). To avoid this problem the low-density regression line (for each value of t) was converted to its corresponding nonlinear form, and the error variance (in terms of u) was computed in terms of the difference between the observations and the calibrated curvilinear relationship. Such an approach lacked

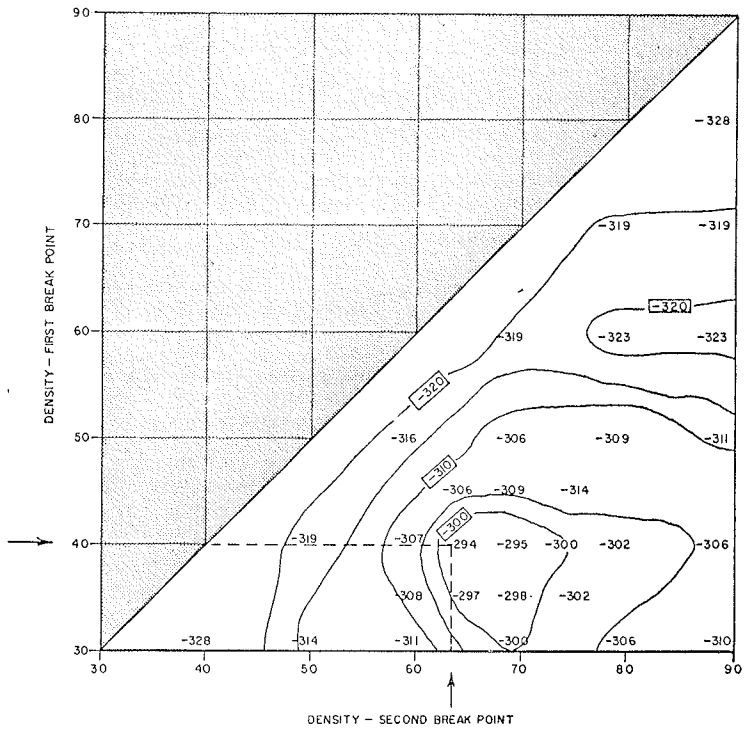


Figure 10. Likelihood function for hypothesis III (3 Regime Linear).

some rigor, since the regression line was determined not by minimizing the sum of squares about the curvilinear form (the mathematics of which are quite cumbersome), but by a minimization of the squared residuals in the transformed relationship.

Hypothesis IV (the modified Greenberg form) was treated in a similar fashion, except that the condition of zero slope was imposed on the low-density regime. The likelihood function given by the break-point analysis, again evaluated at 5-veh/mi intervals, is shown in Figure 7. Three local maxima were observed (k = 40, 50, 55), with the global maximum occurring at 55 veh/mi. The differences between the various peaks were slight.

In the process of computing separate regression lines for each 5-veh/mi increment, t-values were computed to test the slope of the low-density regime vs zero. Figure 9a shows the results of these tests at five confidence levels. For a break-point $t \geq 40$ veh/mi, the slope was significantly greater than zero for all levels, while for t = 25 veh/mi there was no significant difference at any level tested. For t = 30 veh/mi there was no significant difference at levels 0.005 and 0.010.

Figure 9b was derived from these results and shows for any given confidence level, the maximum value of t for which the slope was not significantly greater than zero. For the levels considered, the curve could not be greater than 35 veh/mi.

To establish the break-point value for this hypothesis, the likelihood function was considered together with the interpretation of Figure 9b, which provided a basis for imposing a "hat" of zero slope on the Greenberg hypothesis. Considering the values of t which yielded the peak values of L(t), and choosing a confidence level of 0.005, the optimal break-point was selected to be 35 veh/mi. At any greater value, the slope became significantly greater than zero. At any lower value, the likelihood function fell appreciably below its maxima. The result of this compromise indicated that "significant" interaction between vehicles began at an average spacing of about 150 ft. (While the original purpose of quoting results for various confidence levels was to provide the

reader freedom to exercise his own intuition, the peculiar nature of this analysis forced a decision at the 0.005 level.)

The location of the optimal break-points for hypothesis III (three linear regimes) was accomplished in the same manner as the analysis of hypothesis II, except that it required an extension to three dimensions. Denoting the two break-points by r and s (r separating the low- and middle-density regimes), L(r, s) was computed for combinations of r and s for 10-veh/mi increments in either variable. Then r and s were varied in 5-veh/mi increments in the vicinity of the coordinates (40, 70) where the maximum appeared to be located. The analysis in terms of this finer subgrid yielded the (approximately) optimal break-points or r = 40 veh/mi and s = 65 veh/mi (Fig. 10).

The "spiked" nature of the two-dimensional likelihood functions shown in Figures 6, 7 and 8 suggests that the decision to use a 5-veh/mi subgrid only in one section of the (r, s) plane might have been a dangerous one. It is conceivable that an abrupt increase might have been discovered anywhere on the (r, s) plane had a finer analysis grid been used throughout. The surface represented by the L(r, s) values was characterized by very slight undulations, however, implying that no such spikes may have existed. Furthermore, the fact that the three-dimensional analysis resulted in the same break-point for its high-density regime as did the two-dimensional analysis for hypothesis II lends credence to the strategy selected.

Tests for Distinctly Separate Regimes

Having established the break-points for the four discontinuous hypotheses, the F-tests were conducted to compare each such "calibrated" hypothesis to the supposition of strict continuity over the entire density range. (Quandt points out that the F-test used here is sensitive to the location of the true break-points with respect to the extreme ends of the data; the hypothesis of continuity becomes less susceptible to rejection as the optimal break-point approaches either extreme of the range of the density observations.)

Table 1 gives the results of these tests. At the 0.025 and 0.010 confidence levels, the two linear regimes of hypothesis II did not differ significantly from the supposition that the regression line for the low-density regime prevailed over the entire range of density. Tests for all other hypotheses revealed significant differences of varying magnitudes, indicating that the assumption of separate regimes was in fact justified. In view of the rejection policy adopted for this research, hypothesis II was eliminated from further consideration, although it was subjected to the remaining tests.

TABLE 1

RESULTS OF F-TESTS FOR DISTINCTLY SEPARATE REGIMES

Hypothesis	Test	Calculated F	F-Critical	α
II	I on II	1.555	1.414	0.100
			1.552	0.050
			1.696[a]	0.025
			1.876[a]	0.010
	II on I	33.52	1.876	0.010
III	I on II	33.10	2.34	0.010
	II on I	5.388	2.36	0.010
	II on III	3.477	2.22	0.010
	III on II	64.07	2.06	0.010
IV	I on II	529.7	2.25	0.010
	II on I	3.744	2.00	0.010
VI	I on II	25.46	1.92	0.010
	II on I	3.729	1.84	0.010

[a]At the 0.025 and 0.010 confidence levels, hypothesis II's two linear regimes do not differ significantly from one linear regime. All other tests for all other hypotheses reveal significant differences of varying magnitudes, indicating that separate regimes as suggested by each of these hypotheses in fact exist.

TABLE 2

SUMMARY OF REGRESSION ANALYSES

Hypothesis	Equation	Coefficient of Determination (r^2)	Standard Error (S_e)	F-Ratio Test of Significance-Values of F	Slope vs Zero Values of t	U_f Prediction Values of t	Multi-Regime vs Single	Mean Free Speed (U_f)	Jam Density (k_j)	Optimum Density (k_m)	Optimum Speed (c)	Maximum Flow (q_{max})
I Greenshields	$U=58.6 - 0.468k$	.896	4.648	1005	31.8	1.57	—	58.6	125	62.5	29.3	1830
II 2-regime linear	$U=60.9 - 0.515k$ ($k \le 65$) $U=40 - 0.265k$ ($k \ge 65$)	.685	4.158	250	I 11.2 II 11.7	7.46	a	60.9	151	59.2	30.4	1800
III 3-regime linear	$U=50 - 0.098k$ ($k \le 40$) $U=81.4 - 0.913k$ ($40 \le k \le 65$) $U=40.0 - 0.265k$ ($k \ge 65$)	.590	3.556	167	I 6.9 II 35.8 III 11.7	22.6	b	50.0	151	44.6	40.7	1815
IV Modified Greenberg	$U=48.0$ ($k \le 35$) $U=32.8 \ln \frac{145.5}{k}$ ($k \ge 35$)	.866	3.867	745	I 2.1 II 28.8	26.3	b	48.0	146	53.7	32.8	1760
V Underwood	$U=76.8e^{\frac{-k}{56.9}}$	.901	5.076	1050	32.4	40.4	—	76.8	—	56.9	28.3	1610
VI Edie	$U=54.9e^{\frac{-k}{163.9}}$ ($k \le 50$) $U=26.8 \ln \frac{162.5}{k}$ ($k \ge 50$)	.681	3.550	245	I 5.3 II 15.4	8.59	b	54.9	162	50.0	40.5	2025
VII Bell Curve	$U=48.6e^{-0.00013k^2}$	.884	4.571	872	29.5	21.2	—	48.6	—	62.0	29.5	1830

[a]No significant difference.
[b]Significant difference.

6.85 2.36 2.33
Critical values at .01 level.

Tests for Significance of Entire Regression

Table 2 gives the F- and t-values for the several tests performed on each hypothesis. The fifth column gives the F-values (essentially a measure of the ratio of explained to unexplained variance) for significance of the entire regression in terms of the extent to which the ratio exceeded unity. The critical values of F for each of the confidence levels considered were:

Confidence Level	F-Critical
0.10	2.75
0.05	3.92
0.025	5.15
0.01	6.85

All hypotheses were highly significant.

Tests for Slope Greater Than Zero

The regression equations estimated for each hypothesis were tested to determine whether or not the slopes were significantly different from zero. Nonlinear hypotheses were so tested in the context of the appropriate transformation to linearity. Each regime of the discontinuous relationships was tested separately. For entirely continuous hypotheses, this test was redundant to the F-test. The sixth column in Table 2 gives the values of t obtained in this test. Critical t-values for the levels of confidence considered were:

Confidence Level	t-Critical
0.10	1.29
0.05	1.66
0.025	1.98
0.01	2.36
0.005	2.62

The only hypothesis to exhibit a slope not significantly different from zero was the modified Greenberg equation (low-density regime). However, this condition was forced on the hypothesis as described previously.

Prediction of Mean Free Speed

An independent analysis of mean free speed for the study section provided a possible means for independent verification of the "calibrated" speed-density regressions. The ability of each hypothesis to predict the (lane 2) mean free speed for the facility was examined through the use of a t-test for significant difference between the predicted values and the independently measured value. The substudy, which consisted of sixteen 100-vehicle samples all having average densities of less than 10 veh/mi, yielded an estimated mean free speed of 57.9 mph with a standard deviation of 4.9 mph.

All hypotheses except the Greenshields relationship failed this test (see t-values in column seven, Table 2). The critical t-values were as follows:

Confidence Level	t-Critical
0.10	1.28
0.05	1.64
0.025	1.96
0.01	2.33
0.005	2.58

Hypothesis I exhibited a significant difference only at the 0.05, 0.025, etc., confidence levels. Ignoring the critical t-values, the actual t-values suggested that the Edie rela-

TABLE 3

STATISTICS FOR INDIVIDUAL REGIMES OF DISCONTINUOUS HYPOTHESES

Hypothesis	Regime	r^2	s_e
II	I	0.679	5.343
	II	0.717	2.373
III	I	0.125	2.122
	II	0.561	6.066
	III	0.717	2.373
IV	I	0.093	2.292
	II	0.874	4.214
VI	I	0.299	3.858
	II	0.767	3.424

tionship and the two linear regimes hypothesis (previously rejected) were considerably less inaccurate than the other hypotheses (except Greenshields) on a relative basis.

In one respect, an examination of mean free speed as performed here is little more than a "curious extrapolation," since no car-following rule is intended to apply at very low densities. Indeed, the breakpoint analysis for hypothesis IV suggested that no such (statistically significant) interaction occurred at densities less than 35 veh/mi. Since in the final analysis some type of car-following rules theoretically underlies, or could probably be discovered to underlie all hypotheses tested, the results of these mean free speed tests were not surprising. The very poor performance of the modified Greenberg hypothesis in predicting mean free speed, however, was quite disappointing since exacting procedures were employed to force a condition of "no interaction" on its low-density regime.

A truly comprehensive methodology would have imposed similar "hats" on all other hypotheses. The Greenberg equation was so modified primarily to permit a finite mean free speed prediction for purposes of comparative testing. Had a zero-slope regime been forced on other hypotheses, perhaps a variety of mean free speed predictions would have resulted, depending on the behavior of the likelihood function for each hypothesis. The complexity of the likelihood function analysis increased considerably with number of regimes, and time did not permit such extensions.

It should be observed that two hypotheses (V and VI) were formulated to include mean free speed explicitly as a parameter. In view of the observations discussed above, interpretation of this parameter as mean free speed per se in any accurate sense implies perhaps too rigorous a faith in such relationships.

In summary, from the standpoint of verifying theory, the test for mean free speed prediction was not meaningful. An improvement might have been to impose "hats" on all hypotheses. From the standpoint of practice, if one is interested only in predicting mean free speed, but is restricted to measurements at higher ranges of density, the Greenshields form would appear to be most appropriate.

Measures of Goodness-of-Fit

Table 2 gives the values of the composite coefficient of determination (r^2) and the composite standard error of estimate (s_e) for all hypotheses. The quotation of such composite statistics for discontinuous hypotheses was intended to facilitate comparisons among all relationships considered. (Values of r^2 and s_e for individual regimes of each discontinuous relationship are given in Table 3.)

Table 2 gives the consistent tradeoff between r^2 and s_e: the higher the r^2 value, the higher the standard error. Because of this consistent tradeoff, these statistics taken together at face value gave no meaningful comparative information. One interesting result, however, was the tendency for these two composite statistics to decrease with the number of regimes hypothesized.

INTERPRETATION

Results

Although the tests described followed quite religiously the rejection philosophy underlying the entire research effort, they failed to give very meaningful results after all. It should be emphasized that this failure in no way implies that similar rigor should not be applied in other research endeavors. To develop a rigorous test structure within which binary conclusions might be permitted is one matter; the tests described served

this criterion well. But the structure itself was not sufficient to permit meaningful comparisons, simply because most hypotheses behaved either extremely well or extremely poorly and consistently so for each test. There were always considerable variations among the F- and t-values for each hypothesis. But compared to critical values for conventional confidence limits, these differences were negligible. The tests could have exhibited considerable falsifiability had critical values been applied for levels of confidence on the order of 10^{-5}. The rejection of a hypothesis at such a confidence level presents a difficult problem in interpretation, for the difference between 10^{-5} and 10^{-6} is not likely to be as intuitively evident to the researcher as that between 0.01 and 0.005.

While it might appear from these results that classical tools of statistical inference are not necessarily the key to successful experimentation, it should be realized that more meaningful tests could have been developed had it not been for the fact that independent estimates of parameters other than mean free speed were not available. Reference is made specifically to jam density and maximum flow. Measuring mean free speed in an independent fashion was (and is, in general) a much simpler task than acquiring independent statistical estimates of either k_j or q_m. The only way to estimate jam density would have been to measure the average density over a mile-long section of lane 2 (or some sufficiently great fraction of that distance) when the average flow rate was negligible, and to conduct such measurements with frequency sufficient to yield a meaningful sample size (such an analysis of the facility studied here would have required advanced knowledge of the occurrence of accidents). Similarly, an independent estimate of maximum flow is not easy to determine.

Because the various hypotheses endured these tests with so little differentiation, there remained considerable latitude for judgment on more directly intuitive grounds. While the reader is encouraged toward self-interpretation, it seems appropriate to point out the more obvious deficiencies in the various hypotheses outside of the contest of a strict test structure:

1. The value of mean free speed (76.8 mph) predicted by the Underwood curve was considerably high, particularly since this relationship was developed partly to permit a finite estimate of this parameter (as opposed to the original Greenberg relationship).
2. Previous experience with the operation of the study location and inspection of speed-volume diagrams indicated an optimum speed in the vicinity of at least 40 mph. Except for the relationship advanced by Edie and the three linear regimes hypothesis, the various values for optimum speed were rather low.
3. The jam density value (125 veh/mi) predicted by the Greenshields hypothesis was extremely low.

Considering these observations in a "rejection" context, it became evident that the Edie form (and perhaps the three linear regimes hypothesis, ignoring its relatively poor prediction of mean free speed) warranted further attention. Three further observations were of interest in this connection:

1. These two hypotheses yielded the highest estimates for maximum flow (1,845 and 2,025 veh/hr for hypotheses III and VI, respectively);
2. They furthermore yielded the highest (finite) estimates for jam density (151 and 162.5 veh/mi for III and VI, respectively); and
3. Although they resulted in the two lowest values for r^2, they yielded the two lowest values for the standard error of estimate.

The fact that these two hypotheses gave the highest estimates for maximum flow is significant. That a flow of 1,845 to 2,025 veh/hr is possible on the middle lane of a three-lane urban freeway is not an unreasonable contention. Indeed, within some upper bound (say, 2,100 veh/hr), a high value is most appealing since this parameter estimates the maximum flow which might possibly be sustained over a meaningful time period.

With regard to these two maximum flow values, the maximum in each case was not a local maximum, but rather a boundary value of the low-density regime. In particular, the extension of the freeflow curve of hypothesis VI yielded an "optimal" density (of interest as a parameter only) of 163.0 veh/mi. This extremely high value, which

actually exceeded the jam density value given by the forced-flow curve, indicated a definite discrepancy in comparison to Edie's results for Lincoln Tunnel data (9). Allowing the two regimes to overlap, he determined a break-point "range" of 75 to 100 veh/mi (presumably by inspection), and an optimum density of 90 veh/mi. While it is generally agreed that optimum density is lower for a freeway than for a tunnel, the parameter representing the corresponding local maximum location in this research was almost twice as high as the value determined by Edie.

Inspection of the sample observations (Fig. 2) shows that the form of each regime was somewhat sensitive to the location of the break-point. Consequently it would not be surprising to find, using a similar likelihood function analysis, that k_m for some other urban freeway were considerably lower than 164 veh/mi. This is particularly true since the study location was immediately upstream from a section with slightly lower capacity. For the data given in this research, however, a very vigorous methodology was exercised in locating the break-point; the 163.9 figure tests upon quite sound statistical procedures. To have assumed an overlap between regimes might have led to considerably different results, but his assumption could not have been treated in any fashion except by inspection for "best fit." Whether an overlap is desirable from a theoretical standpoint is debatable.

The second observation listed—that hypotheses III and VI also predicted the highest (finite) values for jam density—is also significant. In general, the (finite) k_j values predicted by all hypotheses seemed quite low. In terms of extremes, it is conceivable that a density of roughly 300 veh/mi is potentially measurable (bumper-to-bumper conditions for one mile). More realistically, it is generally agreed that the jam density for a freeway lane is less than for a tunnel lane. But even the value estimated by hypothesis VI (162.5 veh/mi) was far below Edie's result for the Lincoln Tunnel (265 veh/mi). Of course, neither value has been checked against an independent estimate; thus, whether the tunnel value is too high or the freeway values are too low is yet to be resolved. (At k = 265 veh/mi, both hypotheses V and VII predicted a value of speed which was not significantly greater than zero.)

Finally, with regard to the third observation listed, the fact that hypotheses III and VI yielded the two lowest values for the standard error of estimate was somewhat appealing. After all, the end purpose of a regression analysis is to predict; this point of view would tend to favor the use of s_e over r^2 as a figure-of-merit.

In summary, the results tended to support these two hypotheses above all others considered (according to the interpretations). From the standpoint of logical theoretical consistency, the Edie hypothesis certainly excels in comparison to the three linear regimes alternative. From the less elegant standpoint of application, however, all hypotheses (except the two linear regimes) performed well enough to warrant continued use.

The converted speed-volume and volume-density plots of the regression lines of the various hypotheses are shown in Figures 11 through 31 in the Appendix. In these forms, the three linear regimes, and particularly the Edie formulation, are somewhat more satisfying in the maximum flow range.

Methodology

Several novel techniques were employed in this research. Of interest is the computation of representative composite statistics for discontinuous and/or nonlinear relationships. These techniques should be considered for continued use and improvement. The state of regression theory has lagged the need for treating complicated relationships such as those analyzed in this research. While the techniques used here suffered from certain limitations, they seem to offer the best available means for conducting such a comparative analysis.

The most important result of this research, as far as experimental techniques are concerned, was the successful implementation of the break-point analysis suggested by Quandt. The break-points resulting from these analyses agreed quite well with visual inspections of the sample scatter diagram. Such analyses were time-consuming only because of the magnitude of the particular study. The times of computer runs and

complexity of analyses were not at all overburdening. The only disadvantage of the technique was that the large scope of the study did not permit a very precise exploration of the likelihood functions; consequently, the results based on such break-point locations were not as accurate as might be desired.

The validity of the break-point technique was checked by applying it to a set of data for which three linear regimes had been hypothesized and located by a factor analysis scheme (12). The resulting break-points given by the likelihood function analysis were virtually identical to the results of the factor analysis.

Finally, it seems appropriate to emphasize one point: the question of establishing a test structure to permit meaningful statistical inference. Establishing a test structure in terms of confidence limits is a necessary but insufficient condition for providing a meaningful yes-or-no decision. The sufficient condition is that the test be designed to give such binary decision power within the range of intuitively comprehensible confidence limits.

CONCLUSIONS

The important results regarding the comparison of alternative hypotheses are listed below. It should be emphasized that these conclusions reflect the authors' intuitions to some degree.

1. The Greenshields hypothesis prediction of mean-free-speed was much superior to that of all other hypotheses. The importance of this prediction is slight from a theoretical standpoint since no following-rule is presumed to apply at such low-density conditions.
2. The supposition of two linear regimes was shown to be insignificantly different from one linear regime.
3. Although its performance on the mean-free-speed test was relatively poor, the three linear regimes hypothesis gave relatively good estimates of optimum speed, maximum flow and jam density. Its standard error was the second lowest of the seven hypotheses; its coefficient of determination was lowest.
4. The modified Greenberg hypothesis exhibited only fair performance with respect to all parameters examined. Its values for maximum flow, mean-free-speed, and optimum speed were slightly low, but not distinctly so. The break-point analysis for this form implied that the threshold of interaction between vehicles on the facility was about 35 vehicles/lane-mile, a spacing of 150 ft.
5. Except for having the highest r^2 value, the Underwood form gave poor results. While no statistical test offered such a distinction, the extremely high value of mean-free-speed was quite disturbing, particularly since this relationship was developed partly to permit a finite estimate of this parameter. Furthermore, both the optimum speed and maximum volume values were low.
6. The Edie formulation gave the best estimates of the fundamental parameters. While its r^2 was the second lowest, its standard error was the lowest of all hypotheses. The analysis of this research yielded a somewhat different form than that given in Edie's own work, in that the optimum density value was found to occur at the boundary between the free flow and congested regimes, rather than at the apogee of the free flow curve.
7. In no respect was the bell curve outstanding in its predictive ability. A relatively low estimate of mean-free-speed was anticipated by its very formulation. However, the actual mean-free-speed of the facility appeared to be considerably higher than predicted by this hypothesis.

From a methodological point of view, the following results of this study were of interest:

1. A maximum likelihood technique for locating optimal break-points in multi-regime regression analyses was implemented successfully.
2. Novel techniques of estimating goodness-of-fit for nonlinear and/or discontinuous regressions were developed and used.
3. The original intention of this research was to approach the problem of comparing alternative hypotheses by means of a rigorous structure of falsifiable tests. In the final

analysis, however, almost all conclusions were based on intuition alone since the statistical tests provided little decision power after all. Establishing a test structure in terms of confidence limits is a necessary but insufficient condition for providing a meaningful yes-or-no decision. The sufficient condition is that the test be designed to give such binary decision power within the range of conventional confidence limits.

ACKNOWLEDGMENT

This paper is one of a series of reports of the Chicago Area Expressway Surveillance Project sponsored by the Illinois Division of Highways in cooperation with the Bureau of Public Roads, Cook County, and the city of Chicago. The cooperation received from Northwestern University in carrying out this study is greatly appreciated.

REFERENCES

1. May, A. D., Jr. Discussion of "Freeway Level of Service as Influenced by Volume and Capacity Characteristics," by D. R. Drew and C. J. Keese. Highway Research Record 99, pp. 39-43, 1965.
2. Wardrop, Charles. Some Theoretical Aspects of Road Traffic Research. Proc. Inst. of Civil Eng. (London), Vol. 1, pp. 325-362, 1952.
3. Gazis, D. C., Herman, R., and Potts, R. Car-Following Theory of Steady-State Traffic Flow. Operations Research, Vol. 7, pp. 499-595, 1959.
4. Gazis, D. C., Herman, R., and Rothery, R. W. Nonlinear Follow-the-Leader Models of Traffic Flow. Operations Research, Vol. 9, pp. 545-567, 1961.
5. Greenshields, B. D. A Study in Highway Capacity. HRB Proc., Vol. 14, p. 468, 1935.
6. Ellis, R. H. Analysis of Linear Relationships in Speed-Density and Speed-Occupancy Curves. (unpublished report) Northwestern University, Dec. 1964.
7. Greenberg, H. An Analysis of Traffic Flow. Operations Research, Vol. 7, pp. 78-85, 1959.
8. Underwood, R. T. Speed Volume, and Density Relationships: Quality and Theory of Traffic Flow. Yale Bureau of Highway Traffic, pp. 141-188, 1961.
9. Edie, L. C. Car Following and Steady-State Theory for Non-Congested Traffic. Operations Research, Vol. 9, pp. 66-76, 1961.
10. Quandt, R. E. The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes. Journal of the American Statistical Association, Vol. 53, pp. 873-880, 1958.
11. Quandt, R. E. Tests of the Hypothesis That a Linear Regression System Obeys Two Separate Regimes. Journal of the American Statistical Association, Vol. 25, pp. 324-330, 1960.
12. Drake, J. S. Test for a Hierarchy in Southwest Iowa: A Methodological Exercise. (unpublished report) Northwestern University, March 1965.

Appendix

GRAPHS OF HYPOTHESES

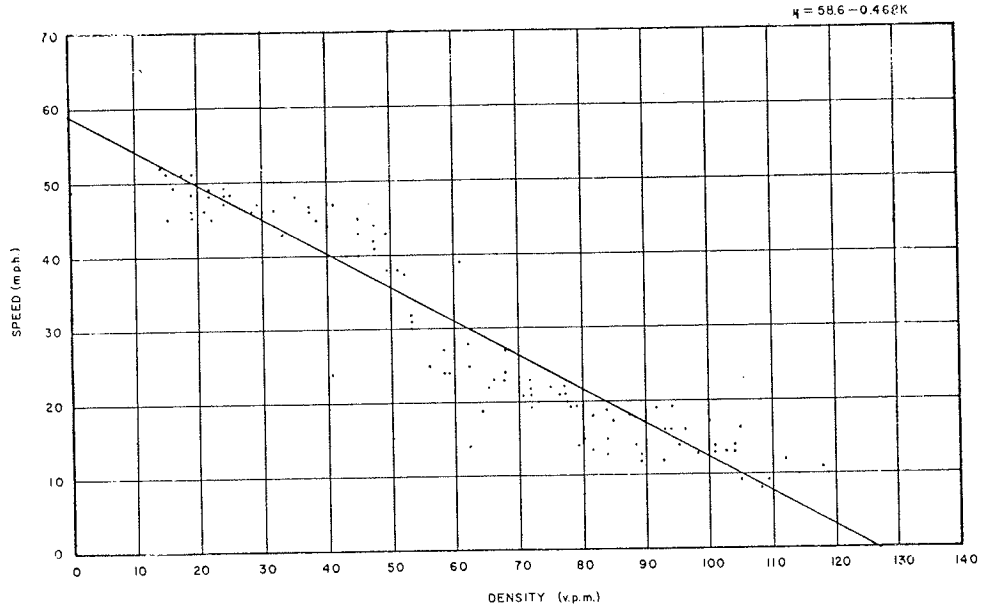


Figure 11. Speed-density relationship—Greenshields hypothesis.

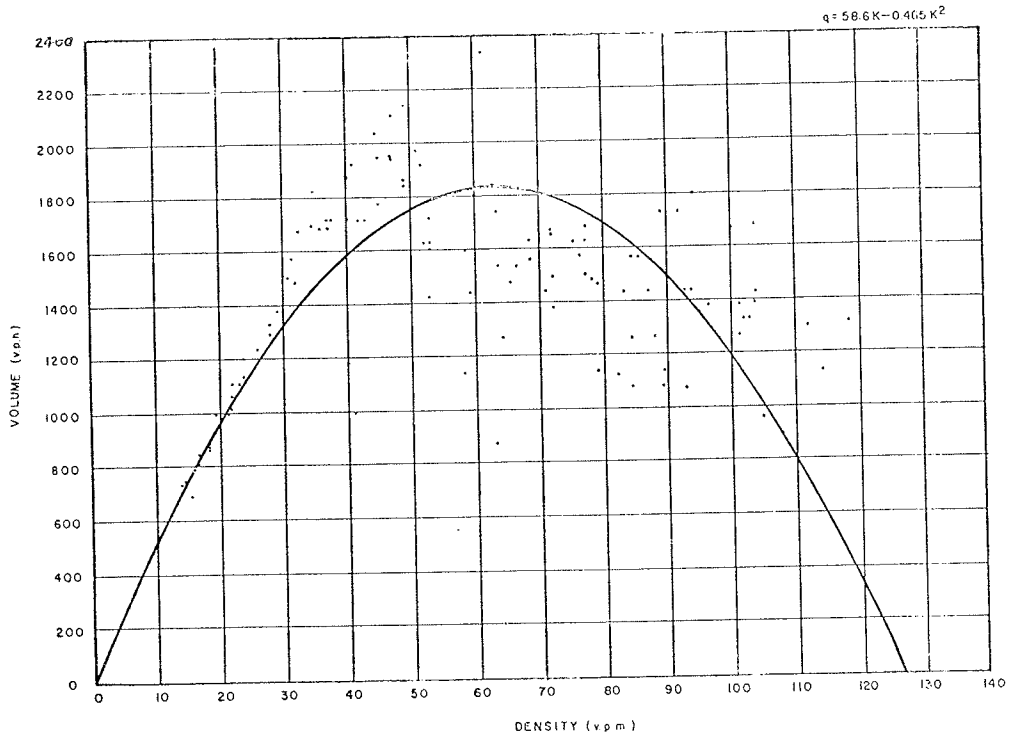


Figure 12. Volume-density relationship—Greenshields hypothesis.

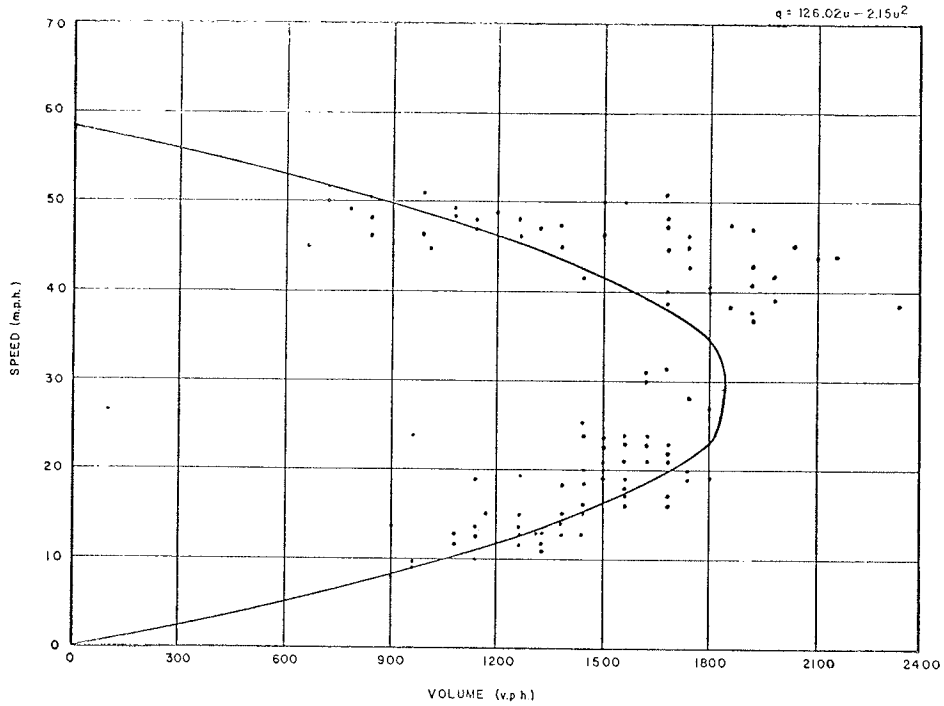


Figure 13. Speed-volume relationship—Greenshields hypothesis.

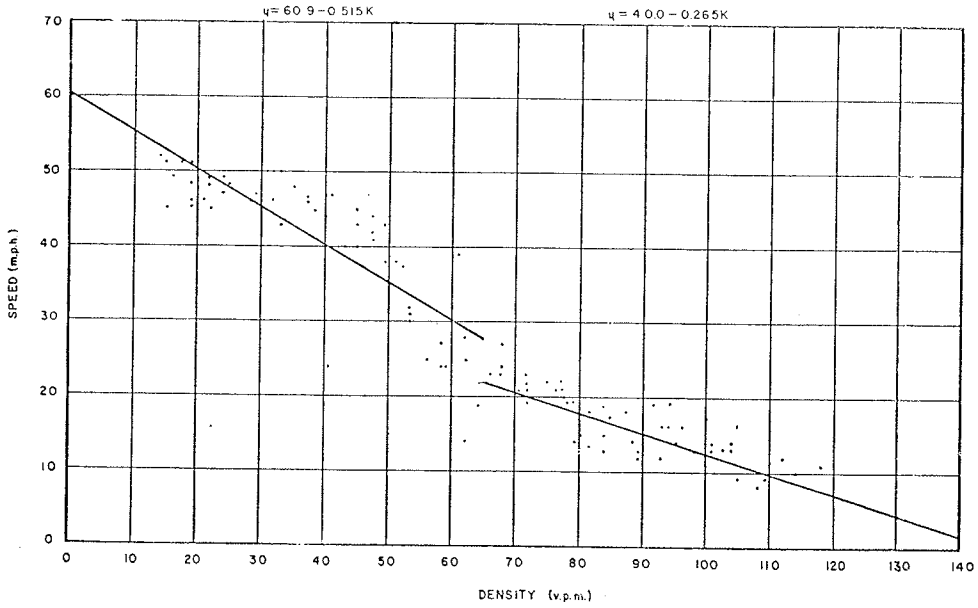


Figure 14. Speed-density relationship—linear, two regime hypothesis.

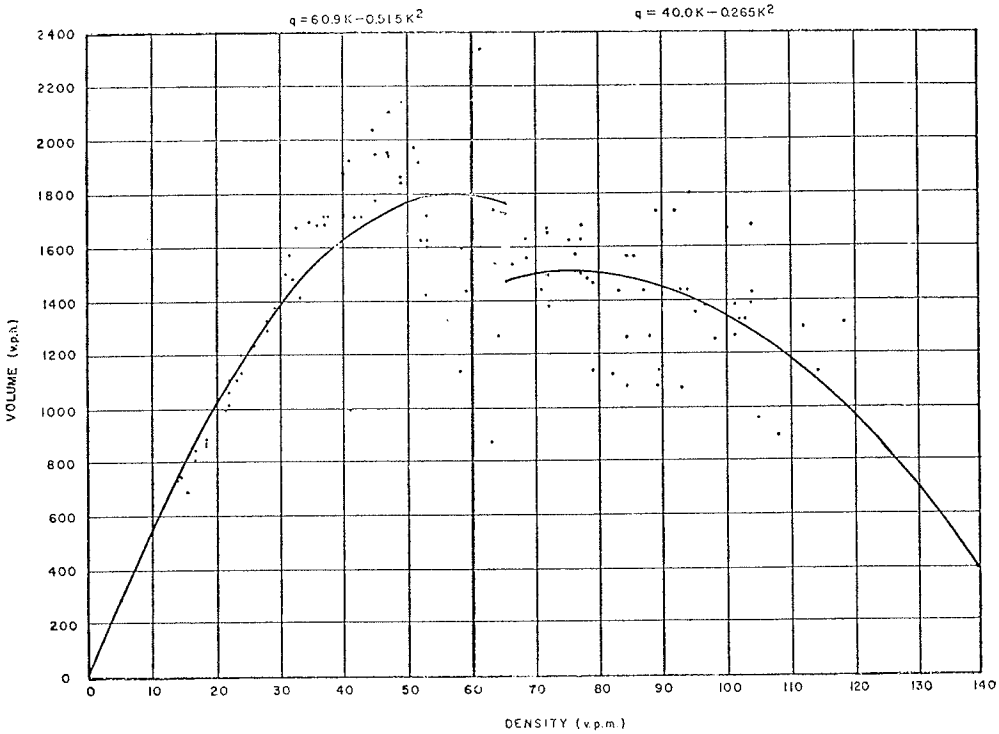


Figure 15. Volume-density relationship—linear, two regime hypothesis.

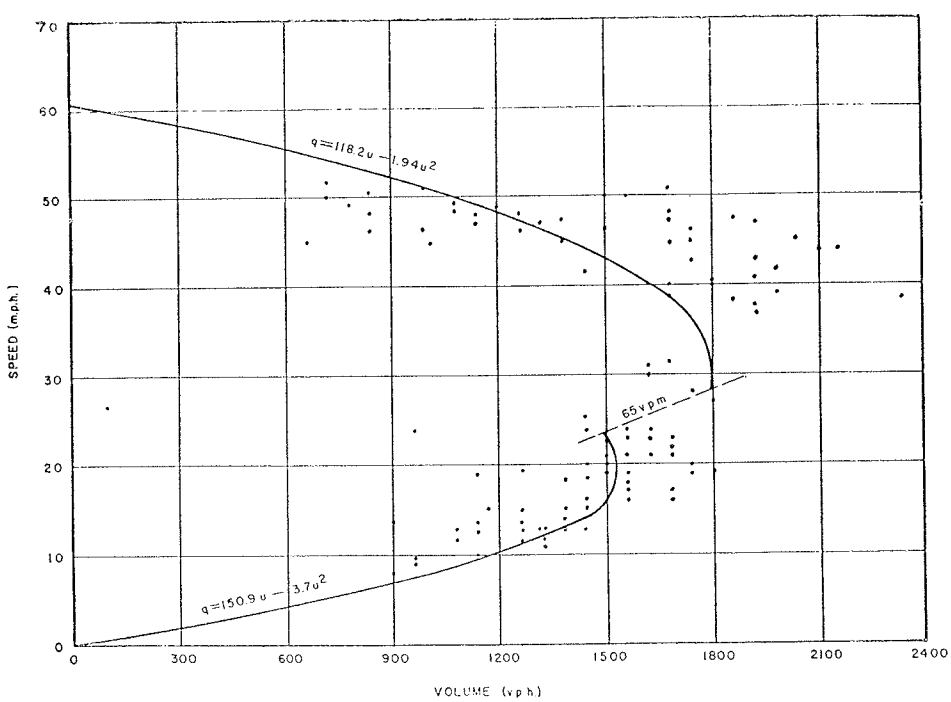


Figure 16. Speed-volume relationship—linear, two regime hypothesis.

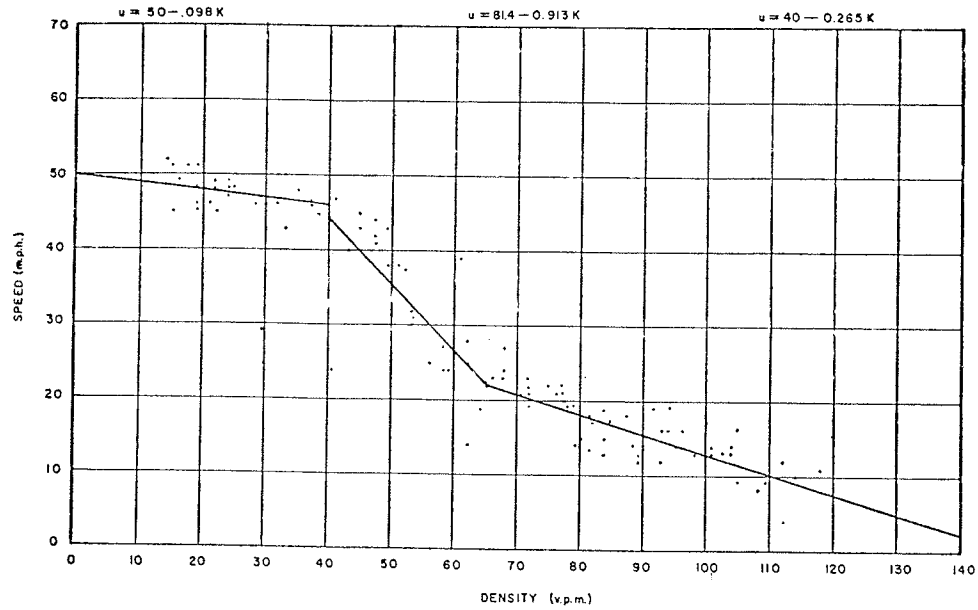


Figure 17. Speed-density relationship—linear, three regime hypothesis.

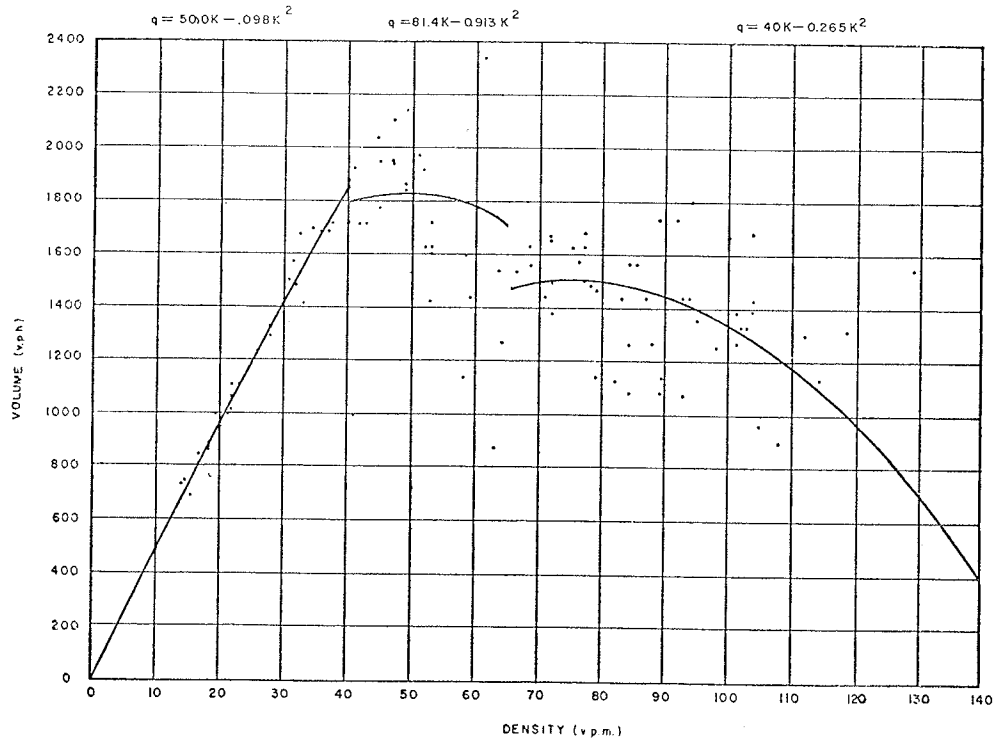


Figure 18. Volume-density relationship—linear, three regime hypothesis.

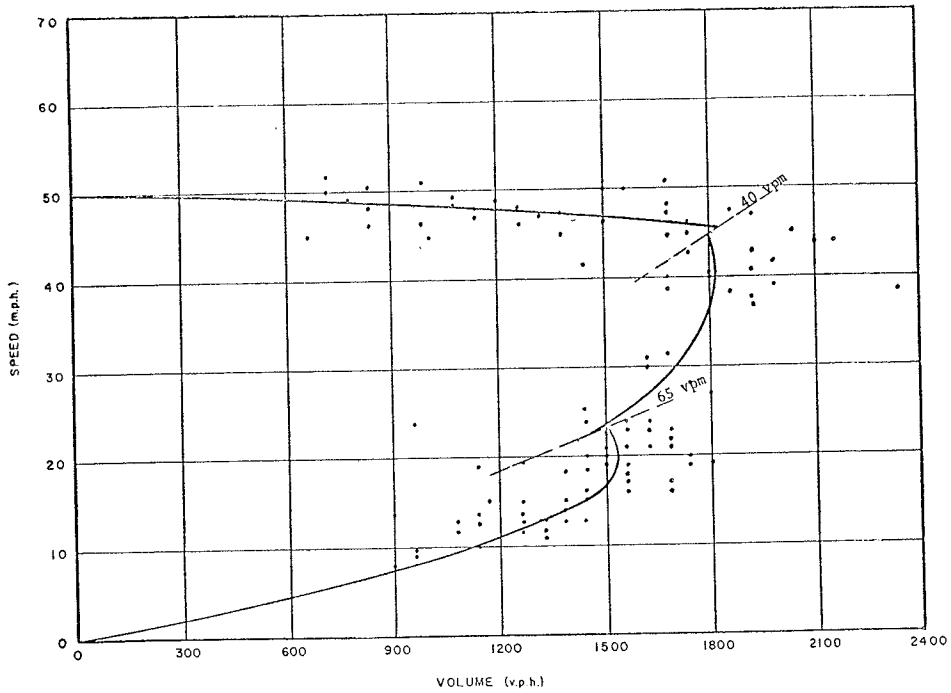


Figure 19. Speed-volume relationship—linear, three regime hypothesis.

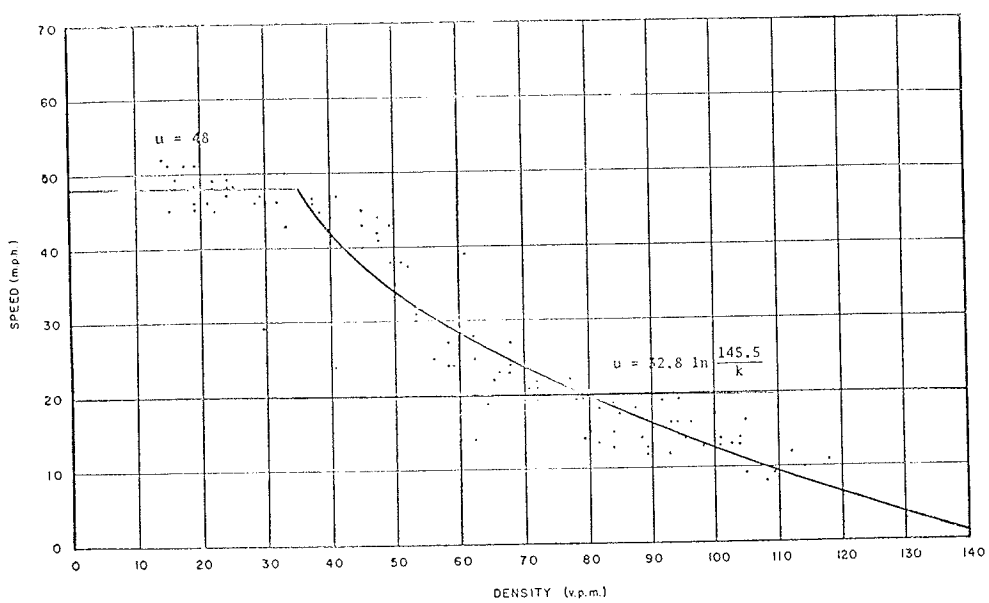


Figure 20. Speed-density relationship—modified Greenberg hypothesis.

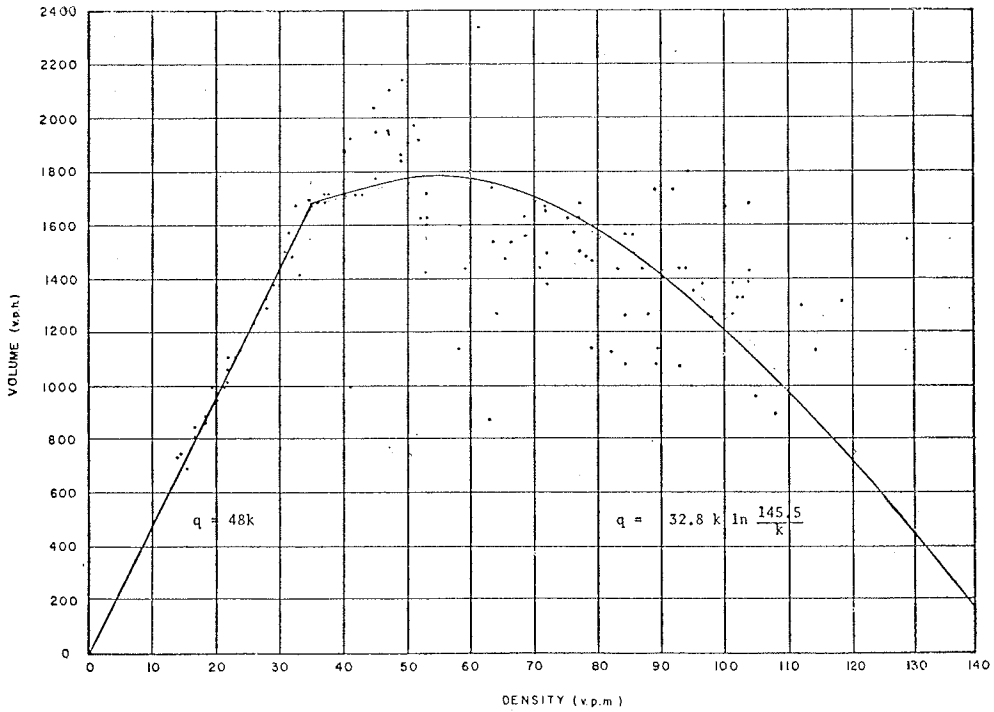


Figure 21. Volume-density relationship—modified Greenberg hypothesis.

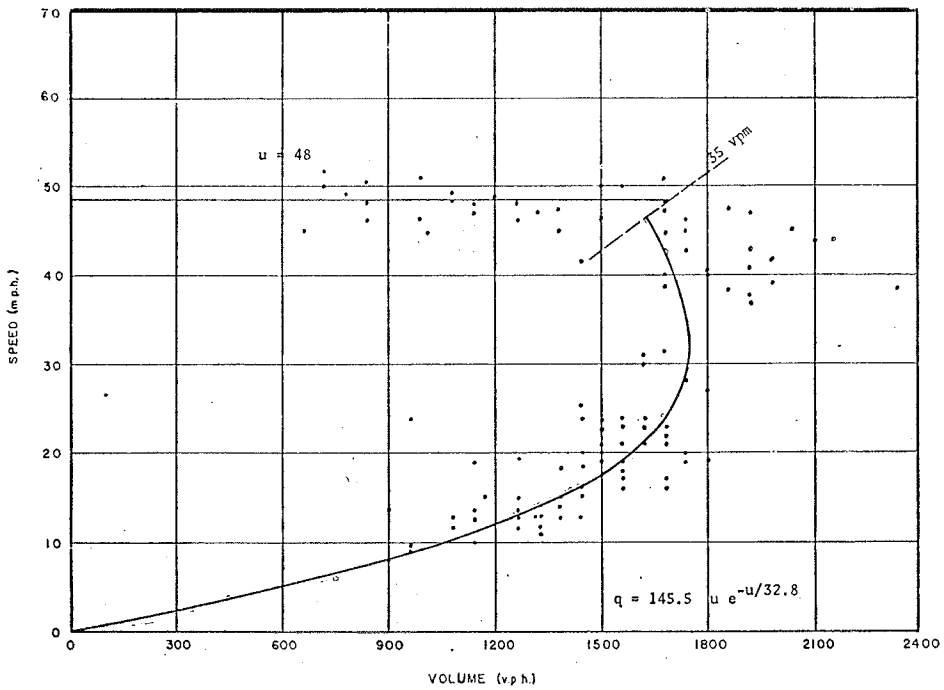


Figure 22. Speed-volume relationship—modified Greenberg hypothesis.

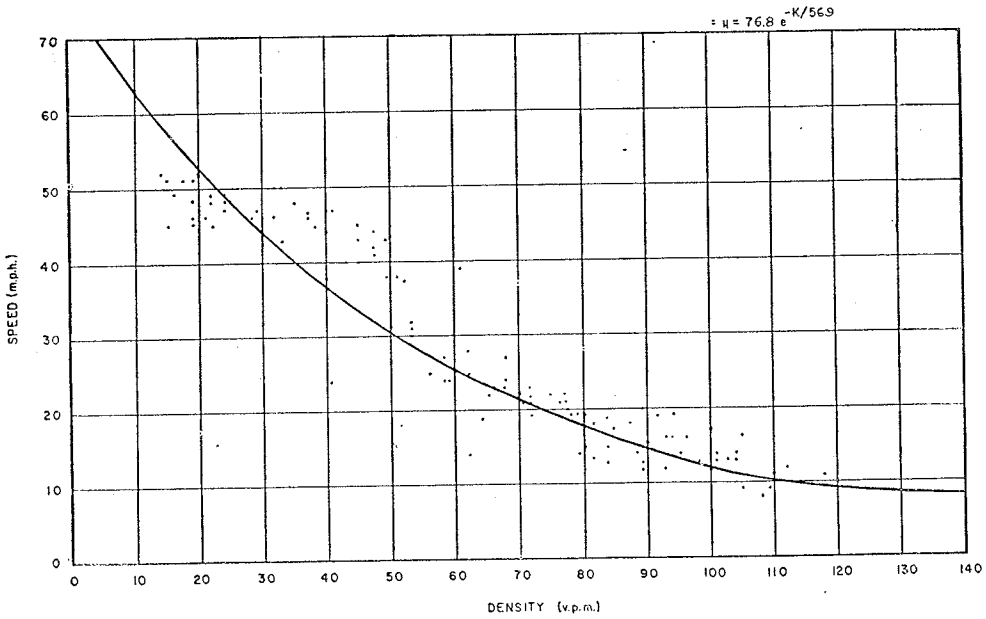


Figure 23. Speed-density relationship—Underwood hypothesis.

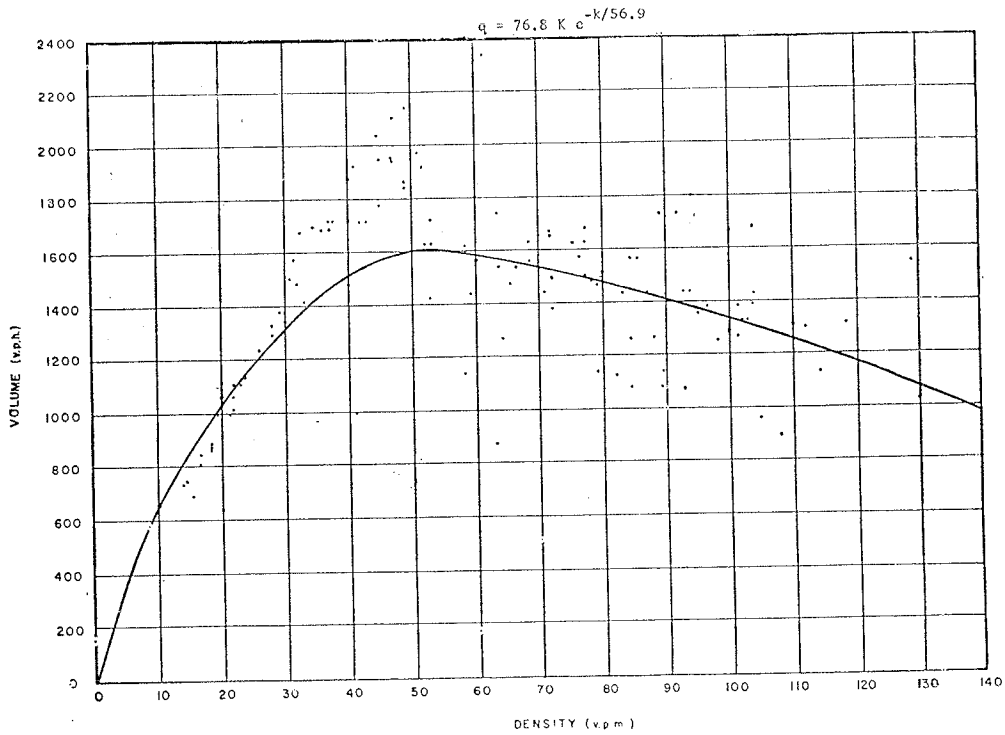


Figure 24. Volume-density relationship—Underwood hypothesis.

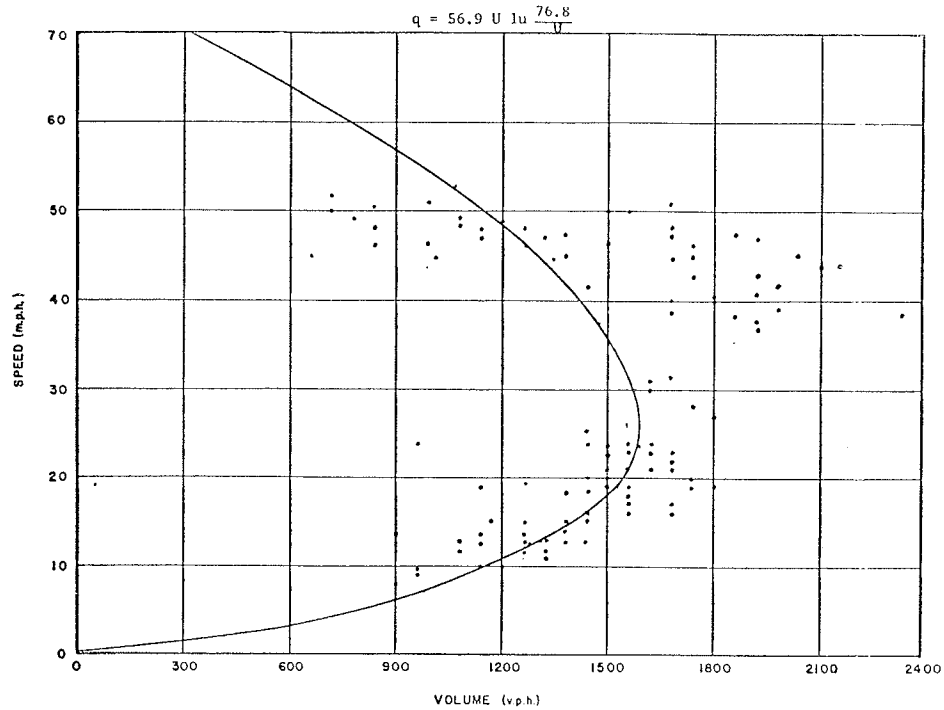


Figure 25. Speed-volume relationship—Underwood hypothesis.

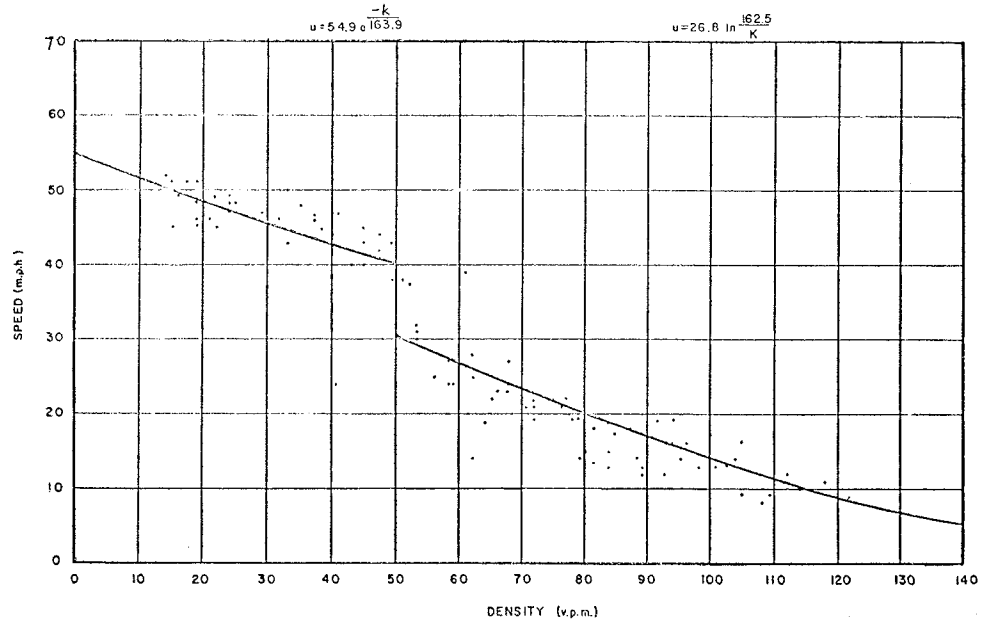


Figure 26. Speed-density relationship—Edie hypothesis.

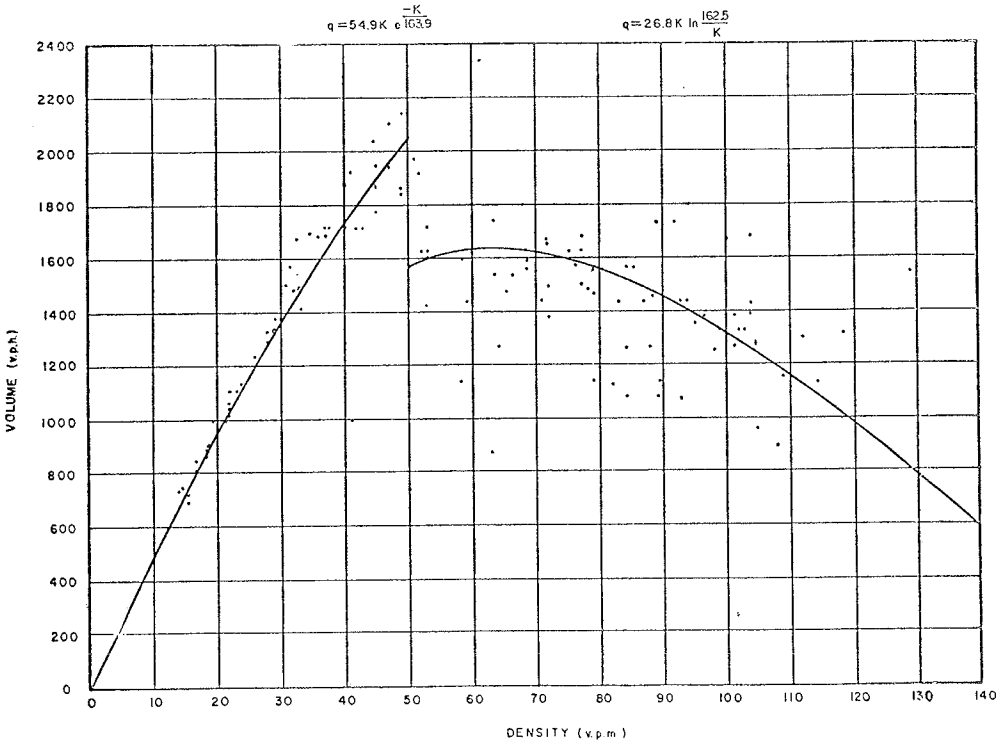


Figure 27. Volume-density relationship—Edie hypothesis.

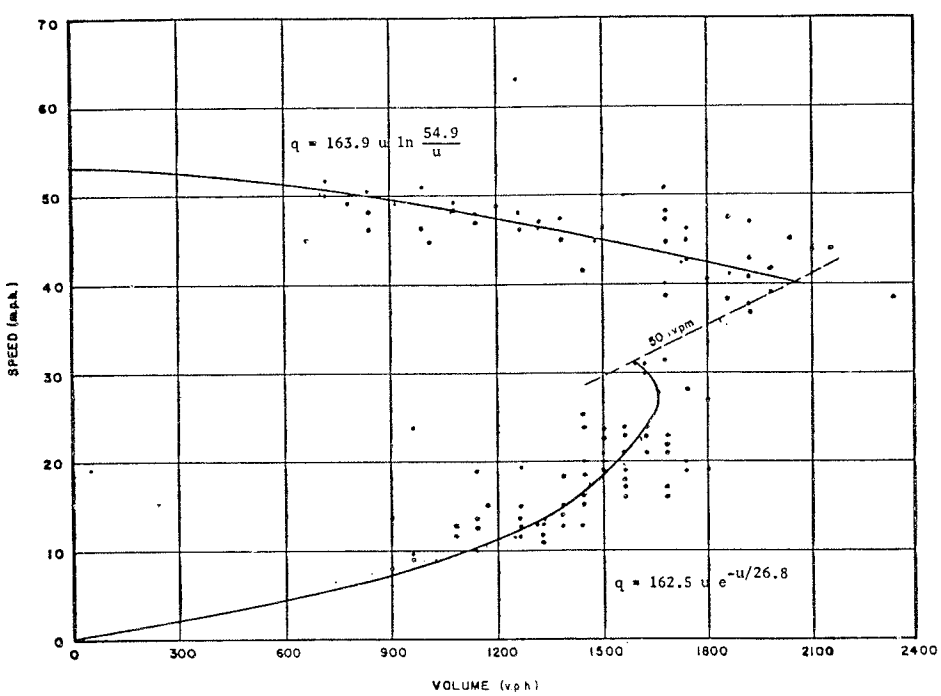


Figure 28. Speed-volume relationship—Edie hypothesis.

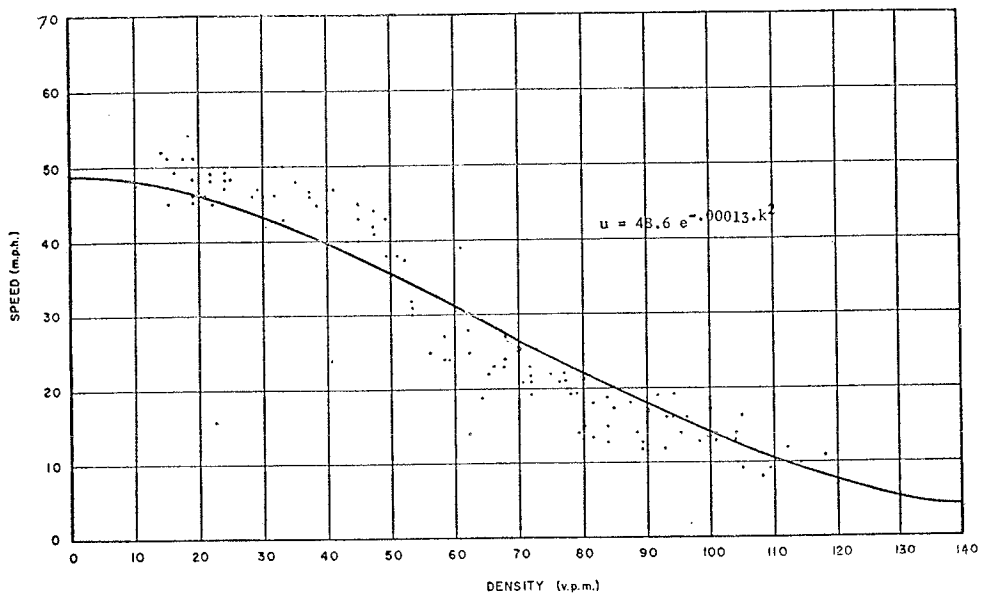


Figure 29. Speed-density relationship—bell curve hypothesis.

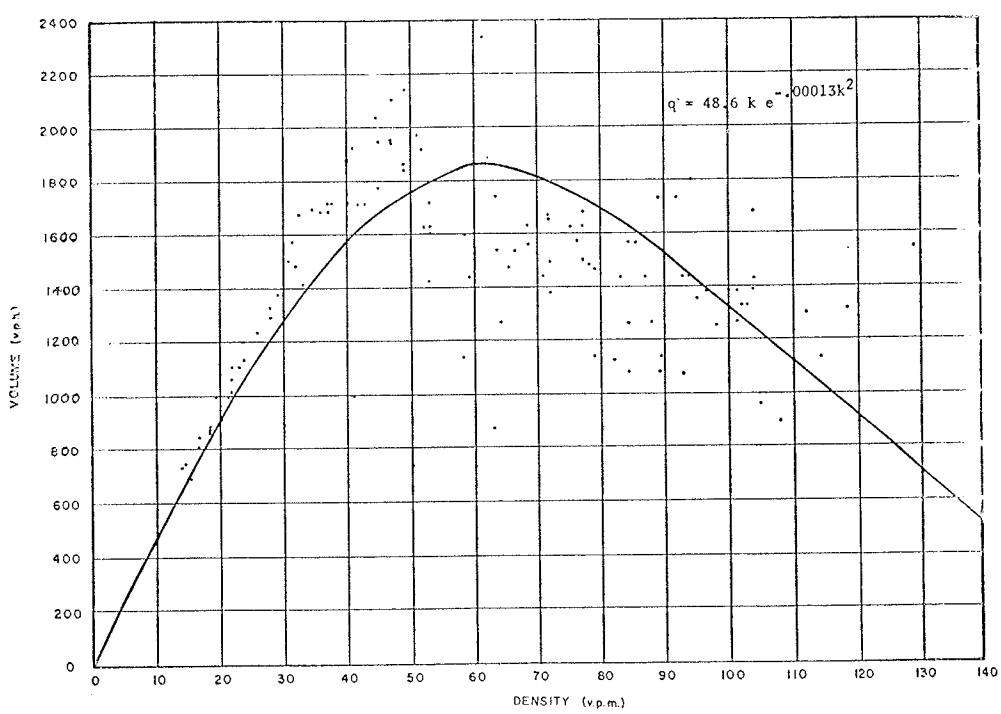


Figure 30. Volume-density relationship—bell curve hypothesis.

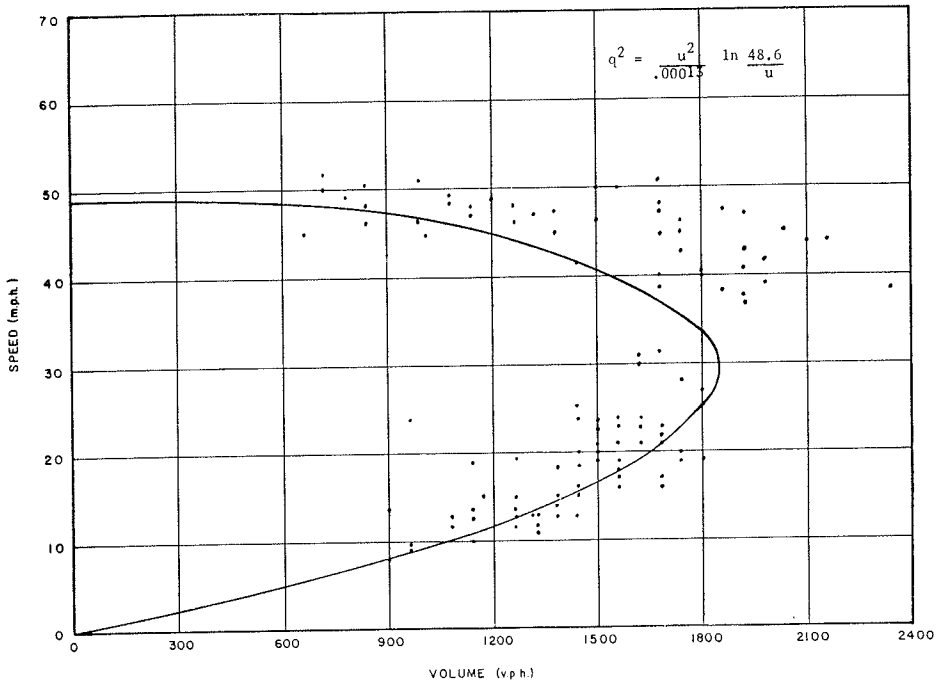


Figure 31. Speed-volume relationship—bell curve hypothesis.