Freeway Operations and the Cusp Catastrophe: An Empirical Analysis

DAN S. DILLON AND FRED L. HALL

Previous empirical work has shown that freeway operations sometimes result in discontinuous data and sometimes do not. Catastrophe theory has recently been proposed as one way of understanding the operations of freeways, which can account for the discontinuities in data. Data from the Queen Elizabeth Way in Ontario are used to test statistically how well the cusp catastrophe can replicate freeway operations. The results are extremely promising, with R^2-values generally above 0.7, and with results that make physical sense in terms of behavior of traffic on the roadway.

During periods of heavy traffic, freeway operations are usually characterized by sudden changes in speed as uncongested traffic encounters a bottleneck and becomes subject to stop-and-go conditions. Navin (*1*) has suggested that the cusp catastrophe from catastrophe theory (*2, 3*) can provide a way to model this sudden change. Hall (*4*) pursued this suggestion, using generalized curves based on data from the Queen Elizabeth Way (QEW) in Ontario, Canada, and found quite promising results. No attempt has yet been made, however, to fit actual traffic data to the cusp-catastrophe surface statistically. In this paper that task is accomplished: data on speeds, flows, and occupancies from four separate lane-station combinations on the QEW are used to test catastrophe theory statistically.

In the first section two elements of the background for this analysis, the empirical and the theoretical, are provided. The empirical background draws on earlier studies of the QEW data to show that both continuous and discontinuous data patterns (e.g., for flow-occupancy curves) can result from the same underlying function. In the theoretical discussion the idea behind catastrophe theory, the way the cusp catastrophe was used by both Navin and Hall, and how each of those accords with the data patterns are outlined. In the second section the transformations of the traffic operations variables necessary to have them conform with the catastrophe theory surface are dealt with. This is necessarily an extended treatment, because the range of possible transformations is large, yet only a few conform to both reality and the requirements of catastrophe theory. In the third section the results of the statistical analysis for the four lane-station combinations are presented. The final section is a discussion of some of the key points from the analysis, leading to the conclusions.

Department of Civil Engineering, McMaster University, Hamilton, Ontario L8S 4L7, Canada.

BACKGROUND

Previous empirical work by the authors with these same data forms an important part of the background to the catastrophe theory work, not because they are the only ones to have worked on this problem (which is certainly not the case), but because they have reached different conclusions from the data than others have, and it is those conclusions that have led to the current experiment with the cusp catastrophe. The data to be used here are the same as those used in the previous analyses. Because they have been described in detail elsewhere (*5, 6*), only a brief description is needed here. The data were acquired by the Ontario Ministry of Transportation and Communications through their Freeway Traffic Management System on the QEW in Mississauga (near Toronto). The analyses have focused on 45 days of ideal conditions extracted from a record spanning 8 months. Data are available for the morning peak period for each of three lanes at each of nine stations along the roadway upstream of a major bottleneck. The four sets of data to be used in the current study are from the shoulder and median lanes at Stations 4 (4 km upstream of the bottleneck) and 7 (1.6 km upstream of the bottleneck). Figure 1 shows an example of the data from the Station 4 median lane for the three relationships of interest.

The starting point for the first analysis (*5*) was that most freeway data exhibit a gap or sparseness near capacity. The authors concluded that it is not necessary to postulate discontinuous functions to account for such a gap. Instead, gaps in the data can be accounted for by the specifics of the data collection location with respect to existing bottlenecks. Thus some data sets (collected close to the bottleneck) may show continuous curves, whereas others (considerably upstream of a bottleneck) will have large gaps. In the second analysis (*6*), these initial findings were confirmed at additional lanes and locations within the same data set. Further support was also provided for the idea that the flow-occupancy relationship is best represented by an inverted V.

The transition from uncongested to congested operations is almost always associated with a "jump," or sudden decrease in speed. There is a corresponding sudden increase in speed from the congested to the uncongested regime. These sudden changes in speed occur even though flow and occupancy exhibit a smooth and continuous change. This property—a discrete sudden change in one variable while other related variables are undergoing smooth, continuous change—is the type of behavior that catastrophe theory was developed to explain. Both Navin (*1*) and Hall (*4*) used the cusp catastrophe,

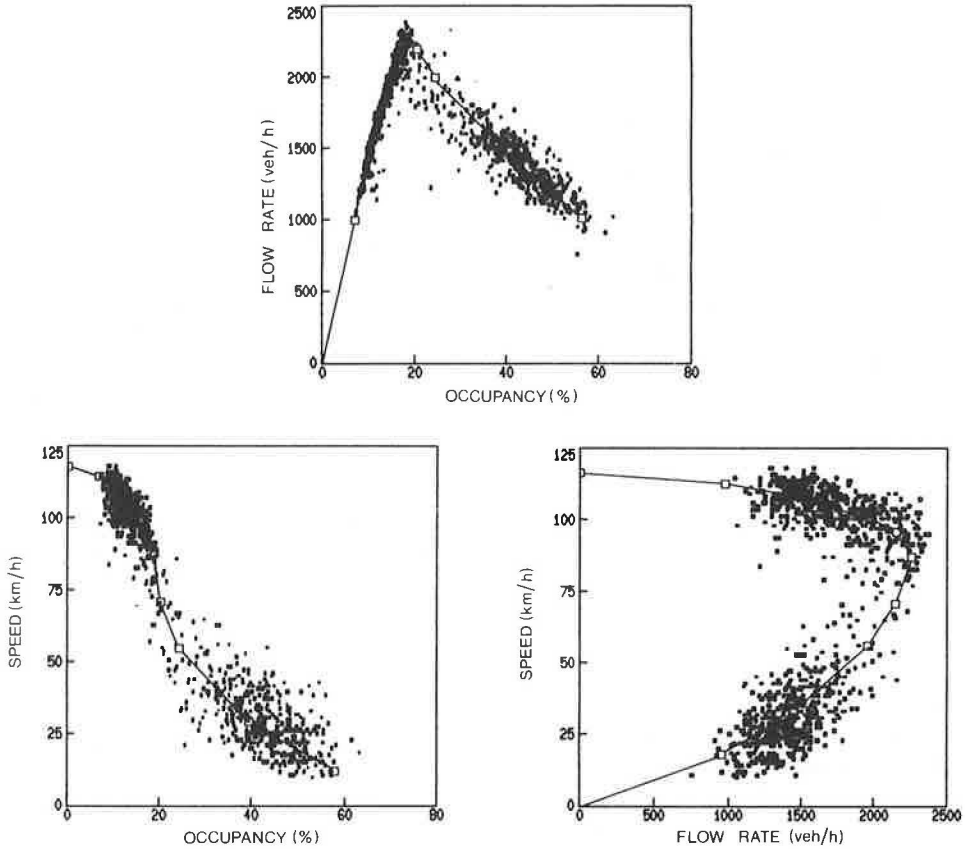

FIGURE 1 Scatter plots of 5-min averaged data from 45 days of ideal conditions, 4 km upstream of a bottleneck (lines represent approximate shapes of curves).

which is one of the seven fundamental catastrophes, so it is the only one to be discussed here.

The cusp catastrophe is described as one that minimizes a potential function,

$$V(x) = x^4 + ux^2 + vx \tag{1}$$

The critical points of the cusp catastrophe are defined by the surface:

$$4x^3 + 2ux + v = 0 \tag{2}$$

where x is referred to as the state variable—the one that sometimes exhibits the discontinuous behavior—and u and v are referred to as the control variables. A plot of the resulting partly folded surface is shown in the central part of Figure 2. The projection of the edges of the fold onto the u-v-plane (immediately below the folded surface in Figure 2) forms a cusp, which is the source of the name for this catastrophe. This cusp can be obtained by setting the discriminant of Equation 2 to zero:

$$8u^3 + 27v^2 = 0 \tag{3}$$

If this discriminant is less than zero (a point "inside" the cusp), there are three real roots, or three possible values of x. If it is greater than zero ("outside" the cusp), there is only one real root and therefore only one possible value for x. The similarity between this cusp-shaped projection and the flow-occupancy curves derived in earlier work (*5, 6*) provided the initial rationale for considering the cusp catastrophe after it had been drawn to the authors' attention by discussions with Navin.

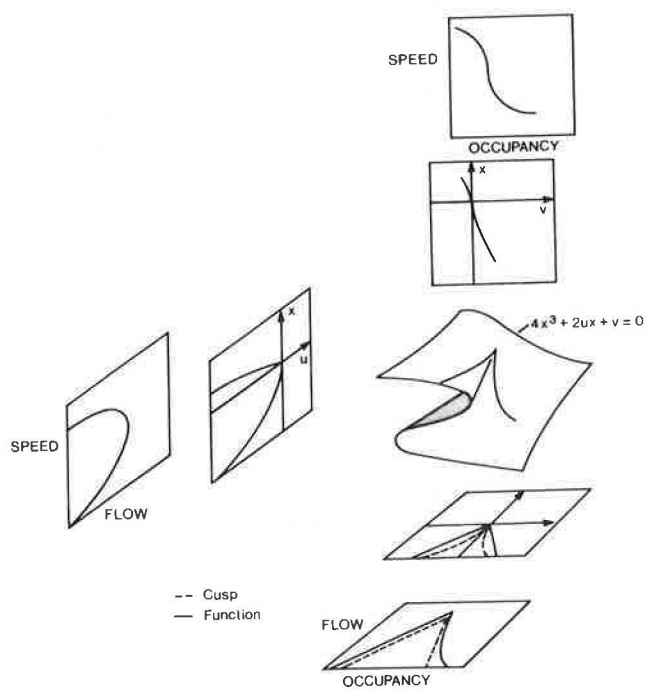

FIGURE 2 Catastrophe theory surface along with *u-v*, *x-v*, and *x-u* projections and corresponding traffic plots.

Surprisingly, this use of the cusp catastrophe is at variance with Navin's by 180 degrees. Because he emphasizes the discontinuity in the data, he draws the flow-concentration curve (for example) so that its high-flow end crosses the open portion of the fold, thus ensuring a discontinuity in all data sets. The use of catastrophe theory here was intended to provide a situation in which there would sometimes be continuous data and other times a gap in the data. Consequently, the authors drew the flow-concentration curve so that the tip of it crosses just above the point on the surface where the fold disappears. For the moment, both versions are merely conjectures, intended to match certain aspects of observed freeway behavior but not yet tested rigorously. One aim of this paper is to provide a more rigorous test of one version of the application of the cusp catastrophe by statistically fitting a curve to the data.

The catastrophe theory variables u, v, and x are not directly equivalent to their corresponding traffic variables but must be first subjected to some transformations. Nor is it self-evident which traffic variable corresponds with which of the three variables describing the catastrophe theory surface. Because of the discontinuities in speed and the similarity between the flow-occupancy plot and the cusp, previous efforts by Hall (*4*) chose to have x correspond with speed, u with flow, and v with occupancy. (Quite independently, Navin had made the same decisions.) It was also decided that the point where the fold disappears [i.e., the origin of the axes, point (0,0,0)] should correspond to operation at capacity. Figure 2 shows the folded surface along with the projection of a possible function onto the three planes and the transformations from these planes into the respective traffic operations plots. The anticipated location of the cusp relative to the function is also shown.

Numerous sets of transformations are possible; two were discussed in the previous paper (*4*). The first set, which serves as the starting point for this paper, is based on a simple linear transformation of speed to x and flow to u. The transformation between v and occupancy was obtained in an ad hoc fashion, using the averaged values of occupancy and v calculated with Equation 2. However, if the data points themselves are used instead, the resulting v-occupancy graph becomes a scatter plot of points that can be fitted to a third transformation. How well this third transformation fits the scatter plot then provides an indication of how well catastrophe theory models freeway operation. That, in essence, is the approach taken in this paper.

TRANSFORMATIONS

In order to determine statistically how well the actual traffic data fit the catastrophe theory surface, it is first necessary to identify the "best" set of transformations between traffic and catastrophe theory variables. The criteria for identifying a best set were established after considerable trial and error in working with possible transformations. However, for clarity the criteria are presented first and then results of the various transformations are discussed.

There are two main criteria for an appropriate set of transformations between traffic variables and the catastrophe theory surface. The first is that there be a discontinuity in speed, because all four data sets to be investigated show one. This requires that the resulting function in x-u-v space occur within that portion of the space where $u < 0$, that is, where the fold occurs.

The second criterion is that at the discontinuity, the physical behavior of the traffic operations be consistent with the mathematical behavior of the cusp catastrophe. This is perhaps best explained by diagrams. Figure 3*a* shows the full surface for the cusp catastrophe, as given by Equation 2. In Figure 3*b* the center fold has been removed, because it in fact corresponds to maxima of the function. Only the upper and lower sheets of the folded area are minima. One possible form of transition from upper to lower surface is sketched on the surfaces in Figure 3*b*: the operations remain on one sheet until it disappears. This is referred to as the perfect delay convention. The most plausible alternative is the Maxwell convention, in which the function takes on its global minimum at all times, resulting in the surface of possible values shown in Figure 3*c*. With this convention, all transitions, up or down, occur along the same plane, that where $v = 0$.

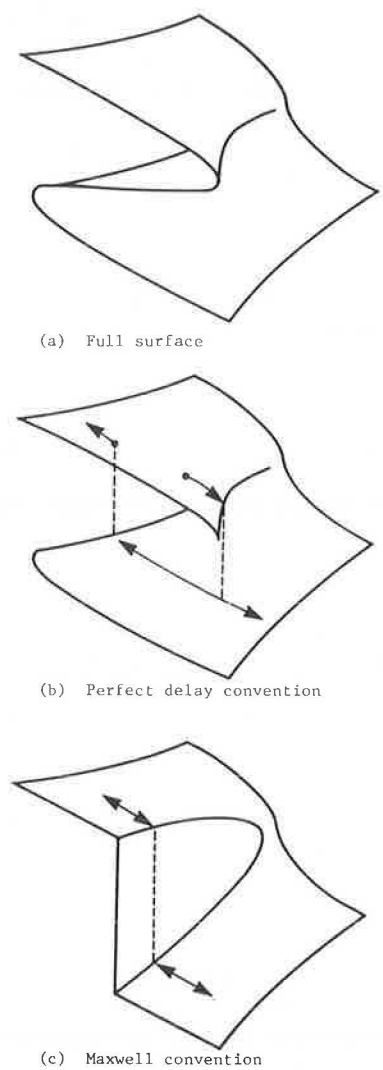

FIGURE 3 Catastrophe theory surface with various delay conventions.

Matching the physical behavior of the traffic operations against these two delay conventions produces two acceptable patterns of data once they have been transformed to the catastrophe theory variables. The perfect delay convention requires that all the uncongested data [data whose speed was greater than the speed at capacity (*SPCAP*)] remain on the top surface, to the left of the right-hand cusp on the u-v-plane. The congested data (speed less than *SPCAP*) must be on the lower surface, to the right of the left-hand cusp. Thus the second criterion translates into a specification of the location of the data on the surface. For ease of plotting and comprehension, the projection of the data onto the u-v-plane will be used to identify data locations for the different sets of transformations.

Three transformations are needed for each set: speed to x, flow to u, and occupancy to v. Given Equation 2 and the actual traffic data, any two determine the third one. The choice of the x-speed and u-flow transformations as the initial ones was arbitrary.

The transformations found to give reasonable results in the previous paper were

$$x = \text{speed} - SPCAP \quad (4)$$
$$u = (\text{flow} - \text{capacity})/1{,}000 \quad (5)$$

One way in which those results did not correspond with expectations, however, was that the projection of the cusp onto the u-v-plane was not distinguishable from the negative portion of the u-axis. Closer inspection of those results as part of this work showed that the reason is one of scale. Using the transformations given in Equations 4 and 5, typical values for x ranged from 25 to -70, whereas the corresponding values for u were in the range of 0 to -1.5. To calculate v (Equation 2), the value of x is cubed, whereas u enters only to the first power, so that v is essentially equal to $-4x^3$ and therefore goes as high as 2×10^6. However, the values of v for the cusp depend only on the value of u (Equation 3), and were therefore no greater than 2.

In order to makes the values of v for the data more comparable with those for the cusp, the effect of u on v must be increased or the effect of x must be decreased, or both. Reducing x will greatly reduce the value of v for the data points (because x is cubed in the calculation of v); however, it will not change the cusp. Increasing u will increase the size of the cusp (because, from Equation 3, u is cubed in the calculation of the cusp) but will not have much of an effect on the data points because that calculation is still dominated by x (assuming that x is still given as in Equation 4).

There are essentially two extreme cases between which the proper transformations should lie. In the first case, flows are divided by a very large number, so that u has been eliminated (i.e., $u = 0$): the data would fall on the line $v = -4x^3$. In the second case, the effect of x has been eliminated (by dividing the speed transformation by a very large number) and the data would fall on the negative u-axis. (If $x = 0$, then from Equation 2, v must equal zero as well.) The location of the data on the catastrophe theory surface for these two extreme cases is shown in Figure 4.

The first extreme, setting u equal to zero, is not desirable because it effectively prevents x (and therefore the speed) from exhibiting the discontinuous behavior that was the initial attraction of catastrophe theory. The second extreme, setting x equal to zero, is also unacceptable because that would require the data to lie on the middle surface, which Figure 3 shows is not possible. The original transformations are quite close to the first extreme: to see whether the relationship between the transformed data and the cusp can be made more obvious, other transformations are attempted. In doing so, however, it appears that the data location will migrate across the surface and around the fold, and finally both portions of the data will converge on the line $v = x = 0$. Any of the transformations that place data on the middle fold are inconsistent with the mathematics of catastrophe theory, and therefore will be rejected.

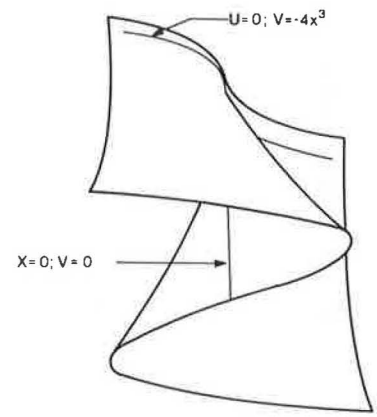

FIGURE 4 Location of traffic data on catastrophe theory surface for extreme transformations.

This migration of data can be seen in the range of u-v-plots shown in Figures 5-13, which contain the results of the first-pass search for better transformations of speed and flow. [In Figures 5-13, circles represent the uncongested data (speed less than *SPCAP*); triangles represent congested data (speed greater than *SPCAP*).] The analysis for this section is based only on the data from the median lane at Station 4 (MED4), because it would be too confusing to attempt all transformations on all four data sets. The basic transformations in Figures 5-13 were assumed to be of the following form:

$$u_1 = \text{flow} - \text{capacity} \quad (6)$$
$$x_1 = \text{speed} - SPCAP \quad (7)$$

From these, nine sets of transformations were obtained by setting $x = x_1/i$ and $u = u_1/j$, where i equals 1, 5, or 10 and j equals 1, 10, or 1,000. Selection of the best of these nine sets of transformations as the area for more detailed search was based on the following reasoning.

Figure 5 corresponds to the transformations given in Equations 4 and 5. As noted earlier, the cusp is indistinguishable from the negative u-axis. The uncongested data appear to be concentrated slightly to the left of the negative u-axis, and the congested data are widely scattered to the right of the axis. Thus this pair of transformations meets the criteria set out earlier, but is perhaps too close to a line for which $u = 0$. The situation is not improved much by dividing x_1 by larger

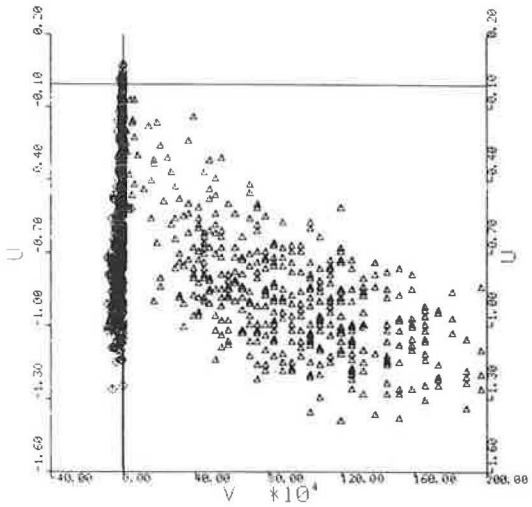

FIGURE 5 Plot of u versus v and cusp projection, MED4: $u = u_1/1{,}000$; $x = x_1/1$.

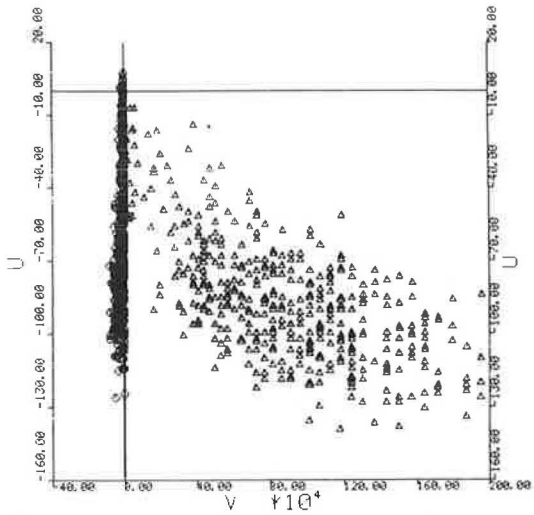

FIGURE 8 Plot of u versus v and cusp projection, MED4: $u = u_1/10$; $x = x_1/1$.

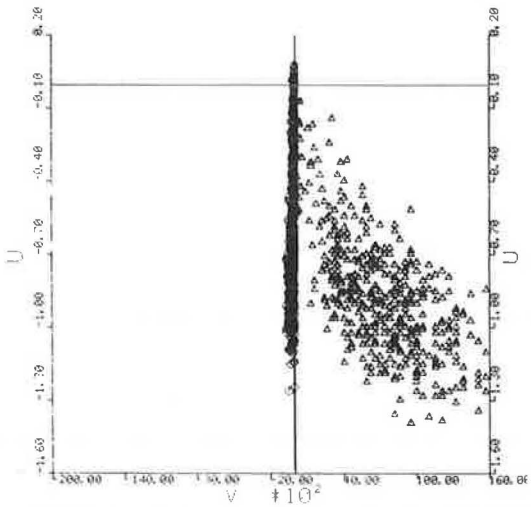

FIGURE 6 Plot of u versus v and cusp projection, MED4: $u = u_1/1{,}000$; $x = x_1/5$.

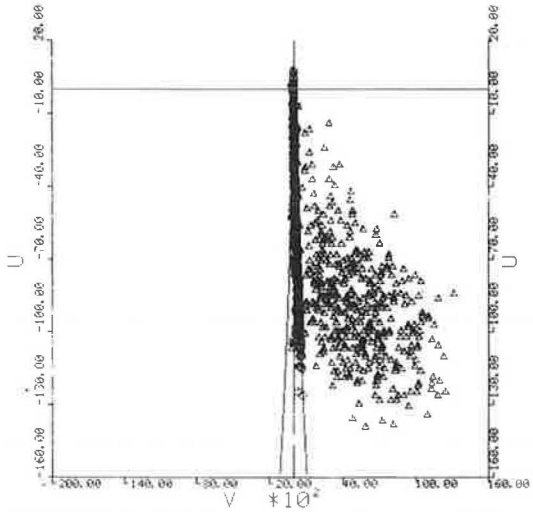

FIGURE 9 Plot of u versus v and cusp projection, MED4: $u = u_1/10$; $x = x_1/5$.

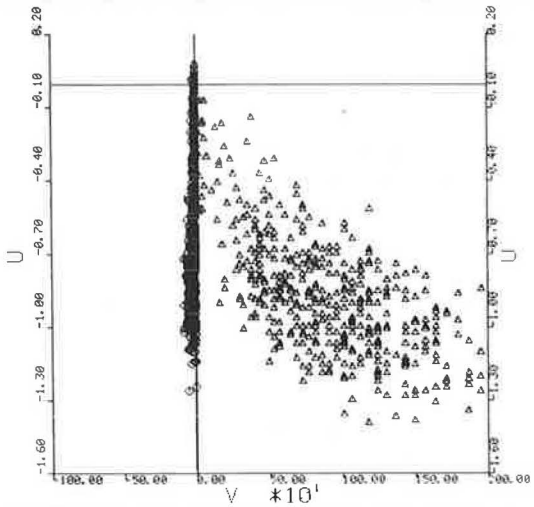

FIGURE 7 Plot of u versus v and cusp projection, MED4: $u = u_1/1{,}000$; $x = x_1/10$.

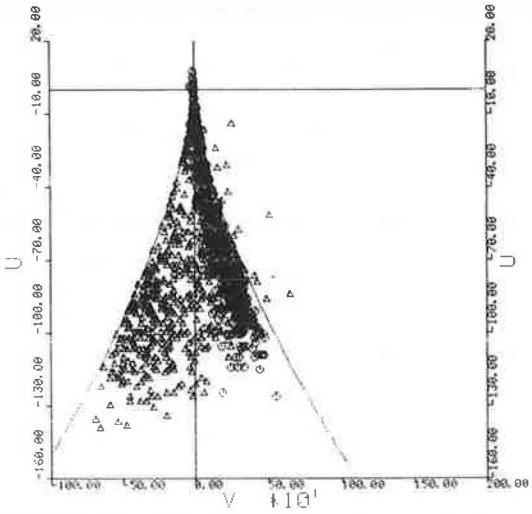

FIGURE 10 Plot of u versus v and cusp projection, MED4: $u = u_1/10$; $x = x_1/10$.

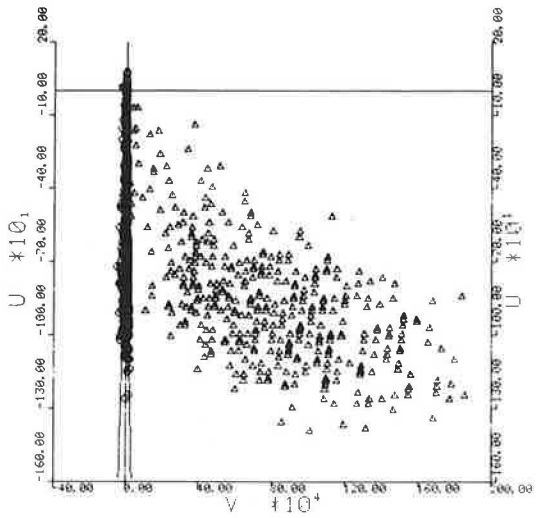

FIGURE 11 Plot of u versus v and cusp projection, MED4: $u = u_1/1$; $x = x_1/1$.

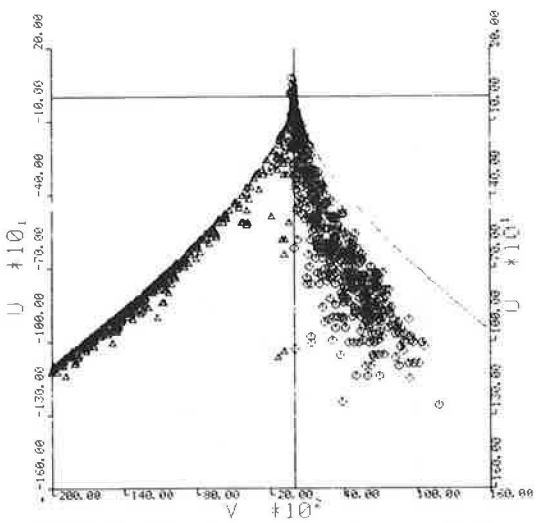

FIGURE 12 Plot of u versus v and cusp projection, MED4: $u = u_1/1$; $x = x_1/5$.

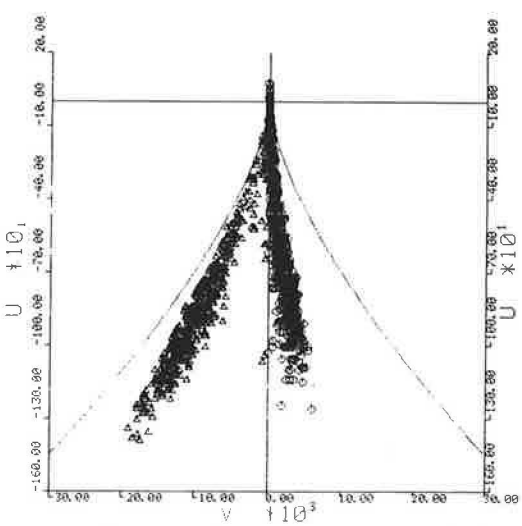

FIGURE 13 Plot of u versus v and cusp projection, MED4: $u = u_1/1$; $x = x_1/10$.

numbers, thereby decreasing the importance of x (Figures 6 and 7).

As u_1 is divided by smaller numbers, such that the effect of u is increased (Figures 8 and 11), there appears to be little effect on the relative locations of the congested and uncongested data and only a minimal effect on the size of the cusp. However, closer examination of the uncongested data (Figures 14-16) shows that increasing the effect of u while keeping x constant causes the uncongested data to migrate, first into the cusp and then across the negative u-axis, and finally to cluster along the right-hand cusp boundary.

When the effects of both x and u are varied simultaneously by dividing x_1 by larger numbers and u_1 by smaller ones, relative to Equations 4 and 5, the migration of the congested data becomes apparent (Figures 8-10). These same figures show migration of the uncongested data onto the center surface. In Figure 9, the concentration of data along the right-hand cusp corresponds to uncongested data that have migrated as far to the right as possible. In Figure 10, the uncongested data have moved away from the cusp again.

Data inside the cusp could fall on any one of the three surfaces, and it is not clear from the u-v-plots alone which one of the surfaces they are on. In order to clarify where the data appear on the surface, a short BASIC program was written that plotted the folded catastrophe theory surface and allowed the user to move a cursor along the surface. The cursor's coordinates in the x, u, v-space were displayed so that the user could position in cursor at appropriate data point locations. Using this interactive plotting routine, it was determined that the uncongested data in Figure 10 fall on the middle surface.

This same procedure was used to confirm that the migration of the uncongested data continues still further toward the $v = 0$ line, on the middle surface, in Figures 12 and 13. The congested data have also continued their migration in these two figures, first appearing to be trapped against the left-hand boundary of the cusp, then moving away from it again,

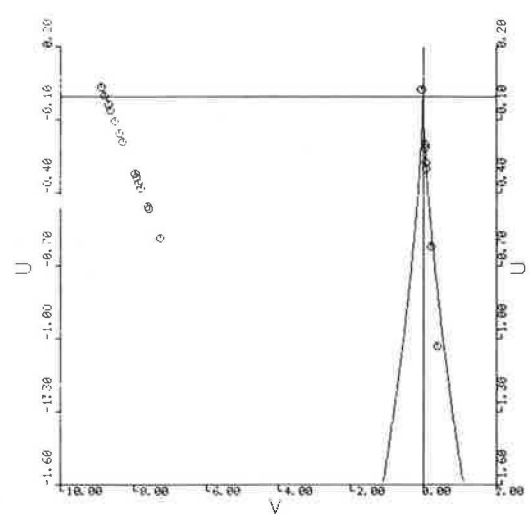

FIGURE 14 Enlargement of u-v-plot (uncongested data only) and cusp projection, MED4: $u = u_1/1{,}000$; $x = x_1/1$.

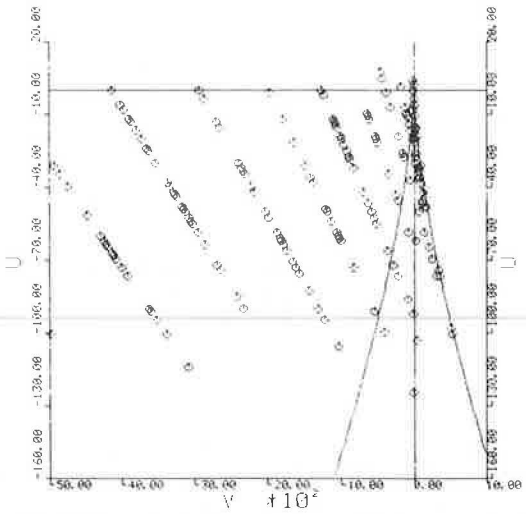

FIGURE 15 Enlargement of *u*-*v*-plot (uncongested data only) and cusp projection, MED4: $u = u_1/10$; $x = x_1/1$.

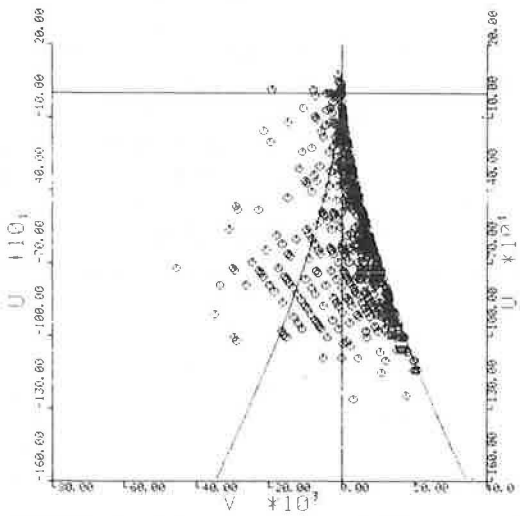

FIGURE 16 Enlargement of *u*-*v*-plot (uncongested data only) and cusp projection, MED4: $u = u_1/1$; $x = x_1/1$.

but on the middle surface. There are also some data on the middle surface in Figures 9 and 11.

The conclusion from this first series of transformations, then, is that only Figure 8 contains reasonable results, and further inspection of transformations between it and Figure 5 is warranted. From this further inspection, the following transformations were determined to be the best (given the time available):

$$u = (\text{flow} - \text{capacity})/100 \quad (8)$$

$$x = \text{speed} - SPCAP \quad (9)$$

The overall data pattern for the best transformation found is shown in Figure 17. When the cusp and the uncongested data in its vicinity are inspected at a much larger scale (Figure 18), it is found that only 11 data points (out of 478 for uncongested operations) occur across the plane that would define the Maxwell convention. (By definition, none can be out of place for the perfect delay convention.) This was deemed to be acceptable for empirical research, so these are the transformations that are used to estimate the statistical fit.

STATISTICAL ANALYSIS

The purpose of this section is to determine statistically how well the transformations in Equations 8 and 9 fit the data to the catastrophe theory surface. For each data point, the value of v, obtained by using Equations 8, 9, and 2, was plotted against the corresponding occupancy (*OCC*); thus the *v*-*OCC* plot was obtained for each of the four lane-station combinations (Figures 19-24). In Figures 19-24 and Table 1, the median and shoulder lanes of Stations 4 and 7 are referred to as MED4, SHL4, MED7, and SHL7, respectively. Both median lanes combined are referred to as MED and both shoulder lanes combined as SHL.

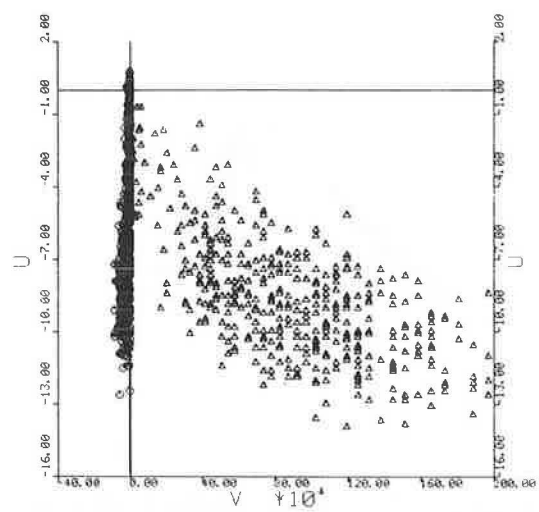

FIGURE 17 Plot of *u* versus *v* and cusp projection, MED4: $u = u_1/100$; $x = x_1/1$.

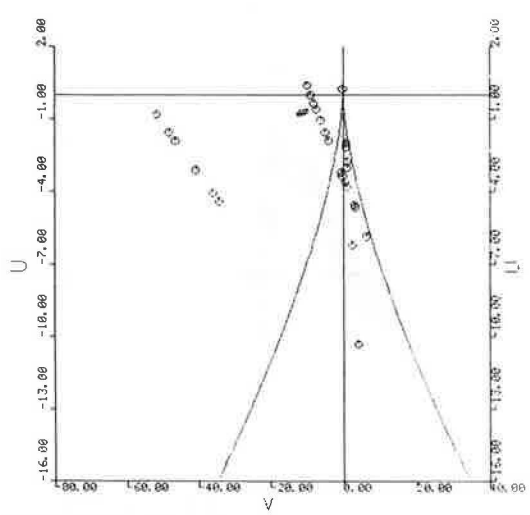

FIGURE 18 Plot of *u* versus *v* and cusp projection, MED4: $u = u_1/100$; $x = x_1/1$ (uncongested data only).

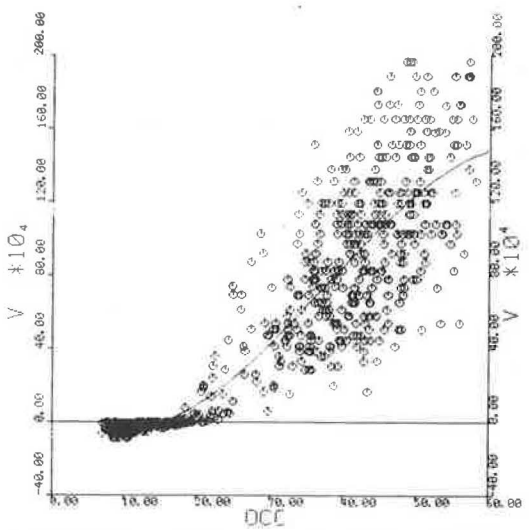

FIGURE 19 ***v-OCC*** **plot and best-fit line: MED4.**

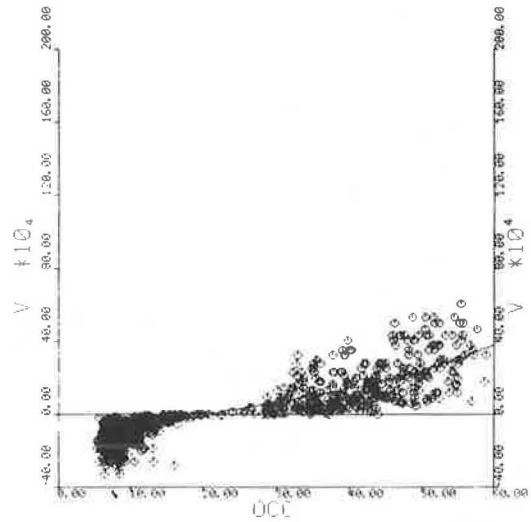

FIGURE 20 ***v-OCC*** **plot and best-fit line: SHL4.**

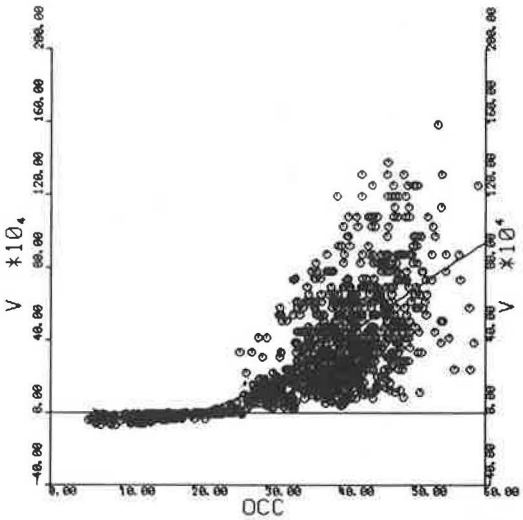

FIGURE 21 ***v-OCC*** **plot and best-fit line: MED7.**

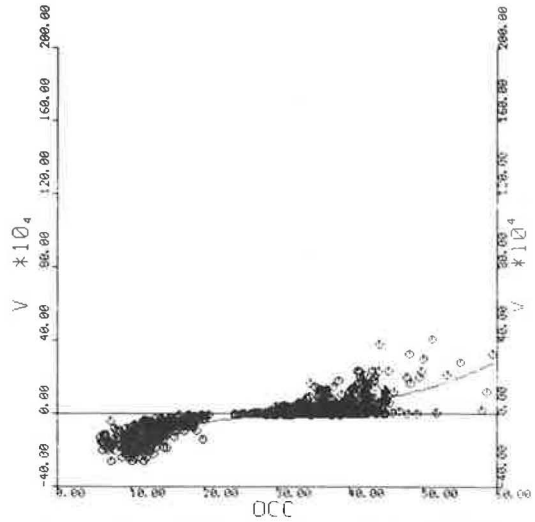

FIGURE 22 ***v-OCC*** **plot and best-fit line: SHL7.**

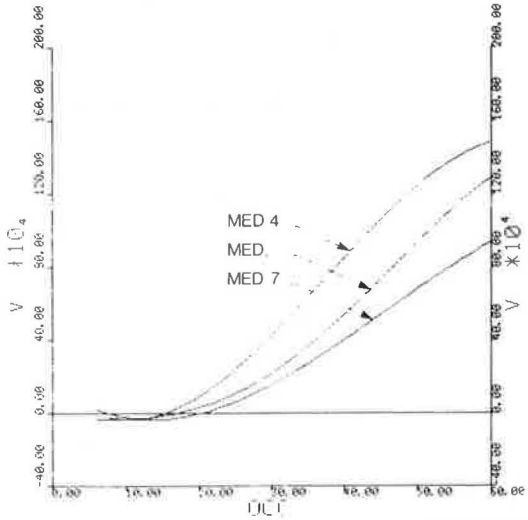

FIGURE 23 ***v-OCC*** **plot and best-fit line: MED4, MED7, and MED.**

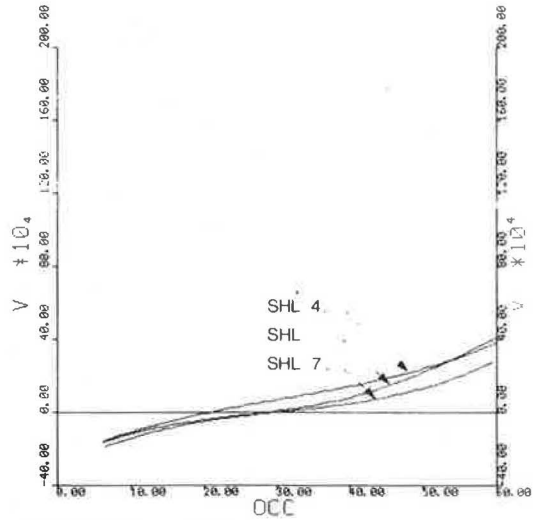

FIGURE 24 ***v-OCC*** **plot and best-fit line: SHL4, SHL7, and SHL.**

TABLE 1 RESULTS OF STATISTICAL ANALYSIS

Lane-Station Combination	No. of Data Points	R^{2a}
MED4	950	0.839
MED7	1,096	0.586
MED	2,046	0.634
SHL4	882	0.785
SHL4[b]	882	0.786
SHL7	1,117	0.730
SHL	1,999	0.723
SHL[b]	1,999	0.728

[a]For the cubic equation:
$v = a_0 + a_1(OCC)^1 + a_2(OCC)^2 + a_3(OCC)^3$
[b]Results for the quatric equation:
$v = a_0 + a_1(OCC)^1 + a_2(OCC)^2 + a_3(OCC)^3 + a_4(OCC)^4$

The main task was to determine what function, if any, could be used to describe the relationship between v and OCC and how well this function fit the data statistically. Numerous different types of functions were contemplated, but it was finally decided that a simple polynomial in OCC can best describe v. In other words, the general form of the relationship would be

$$v = a_0 + a_1(OCC)^1 + a_2(OCC)^2 + \ldots + a_n(OCC)^n \quad (10)$$

A polynomial regression routine (routine RLFOR of the International Mathematical and Statistical Libraries, Inc.) was utilized to determine the best polynomial equation to fit the data for each lane-station combination. The routine was also used to determine the maximum degree of the polynomial, that is, what the value of n is in Equation 10. A sequential, or stepwise, procedure utilizing partial F-values was used to determine how many terms the polynomial should have. It was found that a cubic ($n = 3$) was sufficient in explaining the v-OCC relationship for nearly all of the lane-station combinations. (The shoulder lane of Station 4 and both shoulder lanes combined were explained better by a quatric ($n = 4$) in OCC; however, for consistency, the cubic was used throughout.) A summary of the results for the four separate lane-station combinations as well as both median lanes and both shoulder lanes can be found in Table 1. The best-fit lines along with the v-OCC scatter plots for the four lane-station combinations are shown in Figures 19-24.

The results are promising. Most of the lane-station combinations had R^2-values of more than 0.7. Also, from the plots of the best-fit lines, it appears that the lines cross the $v = 0$ axis near the point of critical occupancy.

However, these results need to be interpreted with considerable caution. The v-OCC plots show that the data are heteroscedastic (especially in the median lanes): there is a larger variance for the higher occupancies (congested regime) than for the lower ones (uncongested regime). The presence of heteroscedasticity means that the regression coefficients may be biased estimates. There are two potential explanations for the occurrence of this problem. The first rests on the lack of precision of the measurements underlying v, which is highly dependent on x, which in turn is calculated solely from speed. The measurement of speed is imprecise in the congested regime because of the stop-and-go nature of the traffic and the 5-min time intervals used for the data acquisition.

The second explanation follows from the observation that the heteroscedasticity is not as apparent in the shoulder lanes as it is in the median lanes. For the median lanes, $SPCAP$ was estimated to be approximately 90 km/hr, which is clearly closer to speed during uncongested operation (typically 100 to 115 km/hr) than that during congested operation (20 to 40 km/hr). This difference between observed speed and $SPCAP$ is cubed in the calculation of v, thus leading to larger variances for larger values of x and v. For the shoulder lanes, $SPCAP$ was estimated to be about 60 km/hr, which is more centrally located between speeds in uncongested operation (typically 70 to 90 km/hr) and those in the congested regime (30 to 50 km/hr). Therefore, the variance in the congested and uncongested regimes would be more nearly equal, as Figures 19-24 confirm.

Because $SPCAP$ could not be identified with any certainty (it occurs where the gap in the data does), it seemed appropriate to investigate briefly the consequences of using other values for this key variable. Again, the median lane at Station 4 was used, in part because it shows the most scatter in Figures 19-24, in part because it was the one investigated in detail to develop the initial transformations. Critical speeds of 75, 80, and 85 km/hr were used, and plots of the u-v and v-OCC diagrams were created (Figures 25-30). These show that the use of a lower value for $SPCAP$ causes the uncongested data to move farther to the left of the cusp (Figures 25, 27, and 29). As is apparent from the v-OCC plots in Figures 26, 28, and 30, the disparity in variance between congested (negative v) and uncongested regimes has improved slightly by decreasing $SPCAP$. However, a lower R^2 was obtained for the lower critical speeds. This is most likely caused by the fact that lowering $SPCAP$ while lowering the variance in the congested data also increases the variance for uncongested operations. Because there are more uncongested than congested data from this lane-station combination and because the uncongested data appear to be more clustered (in terms of occupancy), any increase in the variance of the uncongested data will have an adverse effect on the fit. Consequently, the original estimates of $SPCAP$ have been retained.

DISCUSSION AND CONCLUSIONS

The overall results of this effort to apply catastrophe theory to traffic operations data appear to be very promising. The data can be transformed in such a manner that they conform to both catastrophe theory and the physical reality of traffic operations. The statistical analysis appears to be promising in that a generally good fit was obtained in regressing the v-OCC data to a polynomial in OCC for the four lane-station combinations studied. There are, however, problems that have arisen in this first effort to fit data to theory. Four main ones are worthy of note. Two have been raised already, but not resolved: the heteroscedasticity of the data for the v-OCC plot and the identification of $SPCAP$. The other two are more basic, and have not yet been raised. One relates to the generality of results with respect to different lanes and

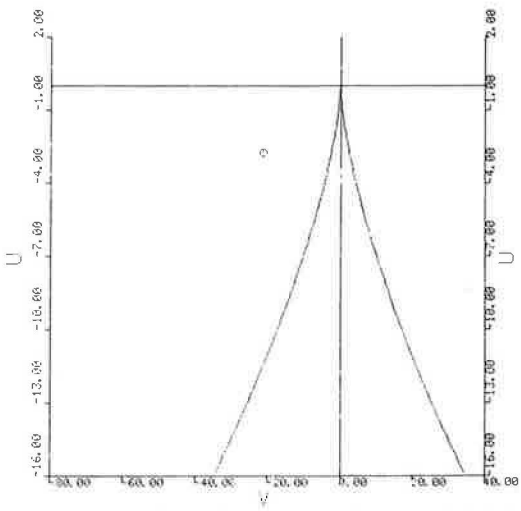

FIGURE 25 Plot of *u* versus *v* (uncongested data only) with cusp projection for best MED4 transformation: *SPCAP* = 75 km/hr.

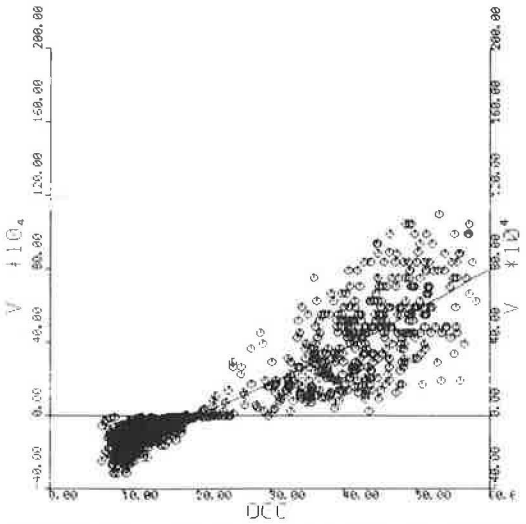

FIGURE 26 *v-OCC* plot (all data) for best MED4 transformation: *SPCAP* = 75 km/hr.

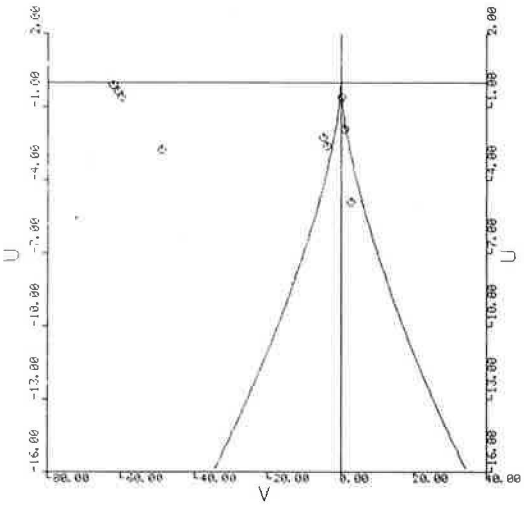

FIGURE 27 Plot of *u* versus *v* (uncongested data only) with cusp projection for best MED4 transformation: *SPCAP* = 80 km/hr.

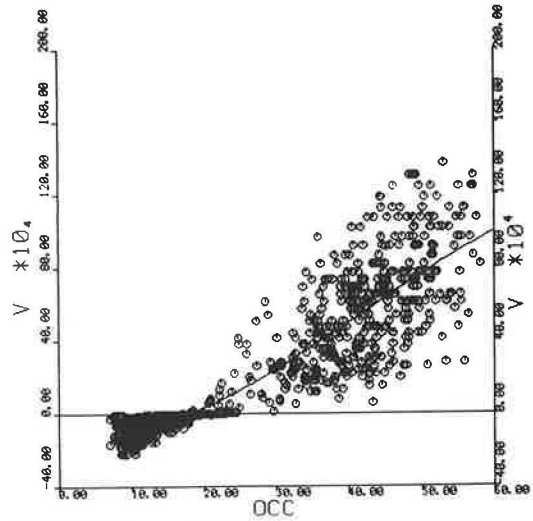

FIGURE 28 *v-OCC* plot (all data) for best MED4 transformation: *SPCAP* = 80 km/hr.

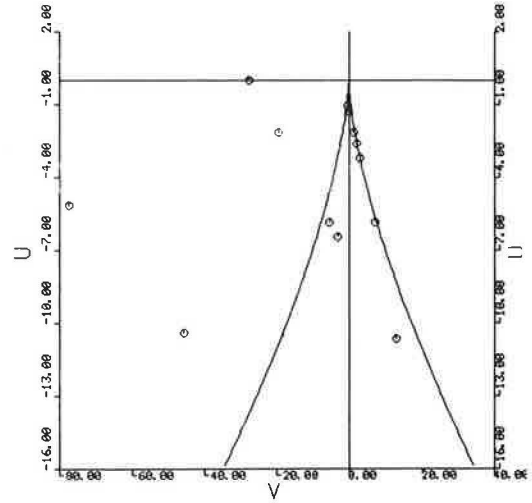

FIGURE 29 Plot of *u* versus *v* (uncongested data only) with cusp projection for best MED4 transformation: *SPCAP* = 85 km/hr.

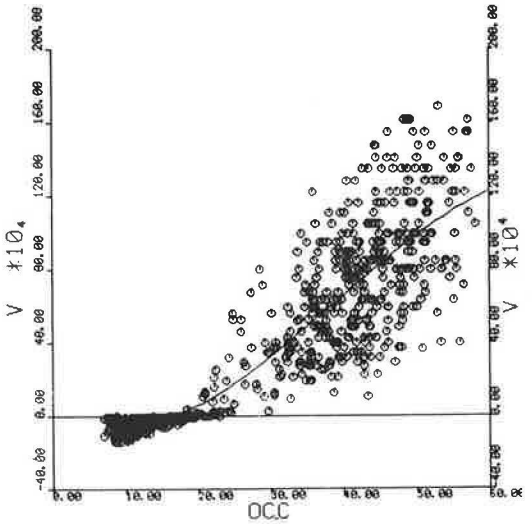

FIGURE 30 *v-OCC* plot (all data) for best MED4 transformation: *SPCAP* = 85 km/hr.

stations—should one be trying to fit each separately, or expect the same parameters to apply to many places? The second addresses the mechanics of conducting the transformations—what might happen, for example, if v were calculated first, rather than x?

There was clearly some heteroscedasticity in the data, especially for the median lanes. There are several standard ways to overcome this, which time did not permit to be included in this paper. Perhaps the polynomial should not be regressed in terms of OCC, but rather in terms of $1/OCC$ or $\exp(OCC)$ (*7*). Alternatively, as mentioned in the previous section, the selection of $SPCAP$ may affect the variance of the error, although the results (Figures 25-30) did not provide very strong support for that approach.

Identification of $SPCAP$ for this paper was essentially estimated from the speed-flow plots. Clearly this identification affects the goodness of the fit, and therefore should be included in the optimizing procedures.

Yet before that is done, a decision should be made regarding the desired or expected generality of the v-OCC relationship, or indeed of all of the transformations. For these analyses, the same general transformation was used for all lanes and locations, but different values were allowed in each case for flow and $SPCAP$. Is that reasonable, or would it be better to select a single value that can serve for all locations? Likewise, should the third transformation determined statistically be the same for all lane-station combinations or different in each case? Alternatively, there could be one set of transformations for median lanes and a different set for shoulder lanes regardless of location.

The solution of this issue is not entirely a matter for curvefitting and statistical analysis, although they can help to resolve it. Earlier work (*6*) demonstrated that there are distinct differences in the operational characteristics of the shoulder and median lanes, particularly in the values of the flows and $SPCAP$ but possibly also in the general shape of the curves. This finding leads one to expect that a common transformation would not be desirable, but that there should be separate transformations for the shoulder and the median lanes. However, the two median-lane results shown here in Table 1 and Figure 8 are not particularly similar. Whether they would be improved with some other value of $SPCAP$ appears unlikely.

The final issue to be raised is the mechanics of conducting the transformations. It is possible to begin with transformations of flow to u and occupancy to v, and from them to calculate x. Then the final step would be to fit a transformation between x and speed. A few preliminary efforts in that direction suggest that the behavior of the data with respect to the cusp can be much better controlled, but that calculation of x (as a cube root) is more difficult to control. This approach may or may not help the heteroscedasticity problem as well. Time and space limitations precluded the extension of the work at this time, but it appears potentially both interesting and valuable. Among other things, it would permit the prediction of speeds on the basis of flow and occupancy, which would be of considerable practical importance to those freeway systems with single rather than paired detectors at each station.

In this paper it has been demonstrated that catastrophe theory can describe freeway operations with a reasonable amount of statistical precision. However, it has also been shown that there are numerous areas where future work could be carried out. The most important point is that the cusp catastrophe provides a valuable new way to understand freeway operations as they move in and out of congested conditions, and therefore may provide a better tool for management of freeway systems.

ACKNOWLEDGMENTS

Margot Gunter provided considerable assistance in the early stages of these analyses, for which the authors are grateful. The cooperation of the Ontario Ministry of Transportation and Communications in providing the data is greatly appreciated as is the financial support of the Natural Sciences and Engineering Research Council of Canada, both through their Operating Grant and Summer Undergraduate Research Award.

REFERENCES

1. F. P. D. Navin. Traffic Congestion Catastrophes. *Transportation Planning and Technology*, Vol. 11, 1986, pp. 19-25.
2. R. Thom. *Structural Stability and Morphogenesis* (English translation by D. H. Fowler of *Stabilité Structurelle et Morphogenèse*). Benjamin, Reading, Mass., 1975.
3. E. C. Zeeman. *Catastrophe Theory*. Addison-Wesley, Reading, Mass., 1977.
4. F. Hall. An Interpretation of Speed-Flow-Concentration Relationships Using Catastrophe Theory. *Transportation Research*, Vol. 21A, No. 3, 1987, pp. 191-201.
5. F. L. Hall, B. L. Allen, and M. A. Gunter. Empirical Analysis of Freeway Flow-Density Relationships. *Transportation Research A*, Vol. 20, 1986, pp. 197-210.
6. F. L. Hall and M. A. Gunter. Transitions in the Speed-Flow Relationship. In *Transportation Research Record 1091*, TRB, National Research Council, Washington, D.C., 1986, pp. 18-21.
7. R. J. Freund and P. D. Minton. *Regression Methods—A Tool for Data Analysis*. Marcel Dekker Inc., New York, 1979.