# INVESTIGATION OF FLOW-DENSITY DISCONTINUITY AND DUAL-MODE TRAFFIC BEHAVIOR 

Barry D. Hillegas, Donald G. Houghton, and Patrick J. Athol, School of Engineering, University of Pittsburgh


#### Abstract

This investigation of traffic behavior was based on an unusually strong data set: The data were taken from a two-lane expressway with only one ramp operation in the 7 -mile $(11-\mathrm{km})$ length. Truck and bus traffic was not allowed to operate on the two-lane expressway, and the data set spanned a 7hour period and reflected all phases of traffic behavior. Autocovariance and cross-covariance time series analyses were applied to traffic-stream occupancy. The autocovariance functions indicated random flow-density behavior for occupancy less than 15 percent (free-flow behavior). The autocovariance functions for higher occupancies indicated varying degrees of Markovian behavior. Cross-covariance analysis indicated that, under freeflow conditions, disturbances in the traffic stream were propagated with the flow of traffic at nearly the free-flow traffic speed. Analysis of flowdensity behavior yielded distinct and discontinuous ranges of linear and nonlinear behavior. Further investigation through multivariate discriminant analysis indicated that, although density was the more important parameter, a flow-density criterion function was superior to a simple density criterion function. Furthermore, such a flow-density criterion function would change over time because of differences in the breakdown and recovery processes.


- UNDERSTANDING the behavior and interrelationships of traffic parameters, especially flow and density, is fundamental to techniques of traffic flow control. Various models have been advanced. The early models assume the well-known continuous smooth parabolic relationship between flow and density. However, Edie (1), after observing a number of data sets from the Port Authority of New York and New Jersey files relating to traffic flow in tunnels, noted that the empirical flow-density plots could be represented by two curves, one for the uncongested or free-flow state and one for the congested state. He proposed a distinct discontinuity in the region of maximum flow (Fig. 1) and showed that two curves fit the data better than a single curve. Drake, Schoffer, and May (2), in a statistical comparison of several hypotheses that describe stream flow characteristics, found that the Edie hypothesis yielded a comparatively low value for the standard error of estimate.

Athol (3) showed the flow-density relationship to be discontinuous with a linear trend to maximum volume for free flow and a breakdown to a flow less than maximum volume in congestion (Fig. 1).

According to Athol, congestion results from a driver behavioral response, i.e., drivers have a threshold tolerance of other vehicles. Once this threshold is exceeded, a reaction sets in that results in less effective individual driving and lower speeds. This interpretation of the onset of congestion is compatible with Edie's theory of discontinuity.

In a different approach, Mika, Kreer, and Yuan(4) investigated data from the John C. Lodge Freeway in Detroit. They categorized flow into two distinct modes-a steady

[^0]state with a parabolic relationship between flow and density and an oscillatory mode in which speed and density exhibit out-of-phase periodicities when plotted as a function of time. They found that the transition between these two modes of behavior is near the maximum flow value. If the steady-flow mode can be interpreted as free-flow operation and the oscillatory mode can be construed as congested operation, the work of Mika, Kreer, and Yuan is further evidence that the flow-density curve involves a discontinuity, or at least instability, about the maximum flow value.

This paper describes a further investigation of the flow-density relationship and dualmode behavior.

## DATA, SYSTEM, AND INSTRUMENTATION

Data used in this investigation were collected by the Expressway Surveillance Project in Oak Park, Illinois. The John F. Kennedy Expressway reversible lane section was the monitored segment of expressway. It is used to relieve the traffic load on the adjacent freeway system during peak periods. The freeway system configuration is shown in Figure 2. During the morning peak period, the reversible lanes are available to traffic flowing southeast toward Chicago. The roadway is open to northwest-bound, outward-flowing traffic in the afternoon and early evening. The slip ramp is open only to afternoon traffic.

There are seven monitoring stations along the 7 -mile ( $11-\mathrm{km}$ ) length of the highway. As shown in Figure 2, the slip ramp is situated between Stations 3 and 4. Afternoon traffic flows from Station 7 towards Station 1. The monitors are $6-\mathrm{ft}(1.8-\mathrm{m})$ square electromagnetic coils centered in the pavement of each $12-\mathrm{ft}(3.6-\mathrm{m})$ lane.

The equipment detects traffic flow in units of vehicles per 20 seconds and occupancy in units of percentage of the 20 -second sampling period ( $\times 100$ ) in which an automobile occupied the coil area. [Volume may be converted to vehicles por hour hy multiplying by 180. Occupancy may be converted to appropriate density in vehicles per mile by multiplying by 2.8 (assuming $18.75 \mathrm{ft}(5.72 \mathrm{~m})=$ average automobile length)(3).] The data were collected from 1:04 to 8:01 p.m. on a day of "clean" conditions-dry pavement, fair weather, and average weekday volume.

In considering this investigation, it is very important to realize the strength of the data set. First, there was less than 1 percent detectable error (the difference in total occupancy at each station over the 7 -hour sampling period) in the data. Second, the monitored highway segment was 7 miles long with only one ramp in operation. No trucks or buses were allowed to operate in the reversible lanes. Thus, the data had a minimum disturbance because of ramp activity and nonautomobile traffic, and this is important in analyzing driver behavior. This minimization of noise would hopefully enable us to see stream flow characteristics much more clearly. Third, the data were collected over a 7 -hour time period that included free flow, transitional buildup to congestion, congested behavior, transitional decay to frec flow, and frec flow again. Undoubtedly, any strong trends in the data would affect the relationship between adjacent data points. However, the measurements are on a sufficiently small time scale compared with the time scale of a trend that such trending effects would be rather small.

TIME SERIES ANALYSIS:
A CASE FOR DUAL-MODE BEHAVIOR

## Consideration of Entire 7-Mile Section by Autocovariance

Consideration of the wide range of traffic stream behavioral characteristics might lead to classification of two ranges of stream behavior: (a) the state in which there is little or no vehicle interaction affecting stream behavior, and (b) the state in which there is some varying intensity of vehicle interaction that affects stream behavior. It seems intuitive that the state of little vehicle interaction might be characterized as random behavior. With regard to the state of high vehicle interaction, one might propose an autoregressive model in which the behavior of the traffic stream for a given time interval is influenced by the behavior of the previous time intervals (5).

In general, we can allow $Z_{t}$ (the event or observation at time $t$ ) to be influenced by all previous events. For a discreet autoregressive time series,

$$
Z_{t}=\alpha_{1} Z_{t-1}+\alpha_{2} Z_{t-2}+\ldots+\alpha_{j} Z_{t-j}+\epsilon_{t}
$$

where $\epsilon_{t}$ is a random variable with zero mean and uncorrelated values for $p \neq q$. Then, a random time series is defined by

$$
Z_{t}=\epsilon_{t} \quad\left(\text { i.e., } \alpha_{1}=\alpha_{2}=\ldots=\alpha_{1}=0\right)
$$

For a discreet time series, such as our data represent, one can investigate the dependence of $Z_{t}$ on previous events by examining the autocovariance function for the series (5, 6).

The $\bar{c} o \overline{v a r i a n c e}$ of two variables measures the degree to which the two variables vary together. If the two variables are not independent, then their covariance is different from 0 (7). The autocovariance coefficient $\gamma_{k}$ at time lag $k$ measures the covariance between two values $Z_{t}$ and $Z_{t-k}$, a distance $k$ apart. The plot of $\gamma$ versus the time lag $k$ is the autocovariance function of the process (5). Box and Jenkins conclude that the most satisfactory K th lag autocovariance is $\mathrm{C}_{\mathrm{k}}^{-}$where

$$
C_{k}=\frac{1}{N} \sum_{t=1}^{N-k}\left(Z_{t}-\bar{Z}\right)\left(Z_{t+k}-\bar{Z}\right), k=0,1,2, \ldots, K
$$

At this point we can say that the autocavariance function of a random process would appear similar to that shown in Figure 3, whereas a process of some autoregressive nature would appear similar to that shown in Figure 4.

Probably the simplest form of autoregressive behavior is Markovian. A Markovian time series is defined by

$$
Z_{t}=\alpha_{1} Z_{t-1}+\epsilon_{t}
$$

According to Kendall and Stuart (8),

$$
\alpha_{1}=0
$$

for Markovian behavior, where $\rho$ is the correlation coefficient between $Z_{t}$ and $Z_{t-1}$.
One statistical test of the randomness or nonrandomness of a time series can be performed using serial correlation (6). The serial correlation coefficient of order $1, r_{1}$, is defined by

$$
r_{1}=\frac{\operatorname{Cov}\left(Z_{4}, Z_{1+1}\right)}{\sqrt{\operatorname{Var}\left(Z_{1}\right) \operatorname{Var}\left(Z_{1+1}\right)}}
$$

so that

$$
r_{1}=\frac{C_{k}}{\sigma^{2}}
$$

where $\sigma^{2}$ is the variance. Serial correlation $r_{1}$ is the sample estimate of autocorrelation $\rho$, so that verification of the hypothesis

$$
r_{1}=0
$$

means the series is random, and verification of the hypothesis

$$
\mathrm{r}_{1} \neq 0, \mathrm{r}_{2}=\mathrm{r}_{1}^{2}, \mathrm{r}_{3}=\mathrm{r}_{1}^{3}, \ldots
$$

Figure 1. Discontinuous flow-density curves.


DENSITY

Figure 2. John F. Kennedy Expressway reversible lane section


Figure 3. Autocovariance function of random time series.

Figure 4. Autocovariance function of autoregressive time series.


Figure 5. Autocovariance function of occupancy for Station 1, Lane 2, 2:00 to 3:00 p.m.


Figure 6. Autocovariance function of occupancy for Station 1, Lane 2, 3:00 to 4:00 p.m.


Figure 7. Autocovariance function of occupancy for Station 1, Lane 2, 4:00 to 5:00 p.m.

would mean that the series is Markovian, for example.
The autocovariance functions of occupancy with up to 40 time lags of 20 seconds each were generated for all seven stations for each 1-hour interval from 2:00 to 8:00 p.m. [An autocovariance and power spectral analysis package known as BMD 62 T of the Biomedical Computer Programs (9) was used. The method of calculation agrees with that recommended by Box and Jenkins (5).] Occupancy (or density in vehicles per mile-note the 2.8 scalar multiplier) was used because of its unique nature in defining the traffic stream. Lane 2 data were used to attempt to minimize noise, because Lane 1 was being fed by the slip ramp. Figures 5, 6, and 7 show some of the autocovariance functions of occupancy for Station 1, Lane 2, over 1-hour intervals from 2:00 to 5:00 p.m.

As shown in Figure 5, Station 1 appears to behave randomly from 2:00 to 3:00 p.m. Similar behavior was observed from 7:00 to 8:00 p.m. For the hours of 3:00 to 5:00 p.m. however, Station 1 appears to have varying degrees of autoregressive behavior as shown by Figures 6 and 7. When similar characteristics were observed from 5:00 to 7:00 p.m., an interesting point arose: The apparent autoregressive behavior is not so pronounced during the time of highest occupancy-28.8 percent from 4:00 to 5:00 p.m., and 25.4 percent from 5:00 to 6:00 p.m.-as it is during the transitional times from 3:00 to 4:00 and 6:00 to 7:00 p.m. with respective occupancies of 15.3 percent and 16.3 percent. This might be due to one of two things:

1. Because congestion (the condition of suboptimal flow with high-density, low-speed behavior) may occur downstream of Station 1 and back up through Stations 1, 2, and 3 during peak periods, a well-behaved Markovian autocovariance function may be most apparent only while there is very high vehicle interaction and congestion occurring in the immediate vicinity of the detector. This conjecture is supported in that further analysis showed that Stations 2 and 3 from 4:00 to 5:00 p.m. had a very random autocovariance function, yet their mean occupancies were 31.8 percent and 46.0 percent, and mean flows were 1,410 and $1,470 \mathrm{vph}$ respectively (Table 1). Also from 3:00 to 4:00 p.m. the autocovariance functions for Stations 2 and 3 were similar to that for Station 1.
2. A higher order autoregressive process or integrated autoregressive moving average process may be occurring that cannot be analyzed using only autocovariance techniques.

Although Stations 1, 2, 3, and 4 show varying degrees of autoregressive behavior at different times, Stations 5, 6, and 7 are continuously random in behavior. The autocovariance functions for these three stations from 2:00 through 8:00 p.m. were all quite similar to that shown in Figure 8 for Station 7, Lane 2, 4:00 to 5:00 p.m.

## Station 4, a Special Case

The behavior of Station 4 appears to be unique. This is probably due to the effect of the merging slip ramp traffic that enters the roadway between Stations 3 and 4. Lane 2 shows lower maximum volumes and densities for Stations 1, 2, 5, 6, and 7, but higher maximum volumes and densities for Stations 3 and 4. Generally, then, Lane 2 is used less than Lane 1 except in the area where the on-ramp is added to Lane 1 (Stations 3 and 4), where drivers tend to move into Lane 2 to avoid merging disturbances. Once merging is completed, however, these drivers tend to move back into Lane 1.

This driver reaction to anticipated merging disturbance may show itself in two ways. First, it might make drivers more conscious of vehicle interaction in this particular area. This increased awareness seems to be indicated by the autocovariance function. Figure 9 shows the occupancy autocovariance function for Station 4 from 4:00 to 5:00 p.m. Although the autocovariance function for 5:00 to 6:00 p.m. is nearly identical, the function during the other time intervals is random. The driver reaction to anticipated merge behavior may also increase dual-mode behavior.

Cross Covariance, a Measure of Simultaneous Similar Behavior
The cross-covariance function is the covariance between two time series, and it analyzes the behavior of two points of the traffic stream over time. A high crosscovariance value or peak in the function indicates that the two points under analysis

Table 1. Mean operating characteristics.

| Station | 2:00 to 3:00 p.m. |  |  |  | 4:00 to 5:00 p.m. |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Percent <br> Occu- <br> pancy | Density (vehicle/ mile) | Flow (vph) | Speed <br> (mph) | Percent Occupancy | Density (vehicle/ mile) | Flow (vph) | Speed <br> (mph) |
| 1 | 8.9 | 25.0 | 1,250 | 50.0 | 28.8 | 80.6 | 1,430 | 17.7 |
| 2 | 7.3 | 20.4 | 1,200 | 58.8 | 31.8 | 89.0 | 1,410 | 15.8 |
| 3 | 12.9 | 36.1 | 1,510 | 41.8 | 46.0 | 128.8 | 1,470 | 11.4 |
| 4 | 6.1 | 17.2 | 1,100 | 64.0 | 11.6 | 32.5 | 1,330 | 40.9 |
| 5 | 4.7 | 13.2 | 801 | 60.6 | 6.3 | 17.6 | 1,020 | 58.0 |
| 6 | 5.9 | 16.5 | 833 | 50.5 | 7.4 | 20.8 | 1,000 | 48.1 |
| 7 | 7.6 | 21.1 | 1,050 | 49.8 | 9.8 | 27.6 | 1,320 | 47.8 |
| Slip <br> ramp |  |  | 899 |  |  |  | 603 |  |

Note: Mean speeds were calculated from flow/occupancy using $\bar{q} v p h / \bar{k}$ vehicles per mile. This will be accurate only for linear q-k behavior; it ignores nonlinear and discontinuous considerations. However, it should roughly reflect speed ranges in nonlinear hehavior.

Figure 8. Autocovariance function of occupancy for Station 7, Lane 2, 4:00 to 5:00 p.m.


Figure 9. Autocovariance function of occupancy for Station 4, Lane 2, 4:00 to 5:00 p.m.


Table 2. Free-flow and congested variances.

| Station <br> n | 2:00 to 3:00 p.m. |  |  |  | 4:00 to 5:00 p.m. |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Variance <br> at <br> Station $n$ | Maximum Cross Covariance of Station n With Station 1 | -Time <br> Lag <br> (second) | Wave <br> Speed <br> $(m p h)^{2}$ | Variance <br> at <br> Station n | Maximum Cross Covariance of Station n With Station 1 | +Time <br> Lag <br> (second) | Wave Speed (mph) |
| 1 | 14.24 |  | 0 |  | 76.61 |  | 0 |  |
| 2 | 8.24 | 5.90 | 80 | 45.0 | 72.26 | 14.98 | 440 | 8.2 |
| 3 | 21.25 | 4.62 | 160 | 45.0 | 150.65 | 24.36 | 760 | 9.5 |
| 4 | 9.21 | 3.35 | 240 | 45.0 | 116.94 | -19.20 | 620 | 17.4 |
| 5 | 4.56 | 2.01 | 320 | 45.0 | 7.48 | 4.32 | 680 | 21.2 |
| 8 | b. 12 | 2.12 | 380 | 47.4 | 8.42 | -4.48 | 220 | 81.8 |
| 7 | 8.05 | 2.10 | 420 | 51.4 | 9.26 | -4.85 | 60 | 360.0 |

[^1]experience similar behavior at some distance apart in time equal to a lag value, k . The cross-covariance function is not necessarily symmetrical about zero as was the autocovariance function (5).

If we have two time series

$$
X_{t-n}, \ldots, X_{t-2}, X_{t-1}, X_{t}, X_{t+1}, X_{t+2}, \ldots, X_{t+n}
$$

and

$$
Y_{t-n}, \ldots, Y_{t-2}, Y_{t-1}, Y_{t}, Y_{t+1}, Y_{t+2}, \ldots, Y_{t+n}
$$

then $C_{x y}(k)$, the estimate of the cross-covariance coefficient at lag $K$, is

$$
C_{x y}(k)=\frac{1}{n} \sum_{t=1}^{n-K}\left(X_{t}-\bar{X}\right)\left(Y_{t+k}-\bar{Y}\right) \text { for } k=0,1,2, \ldots, K
$$

and

$$
C_{x y}(k)=\frac{1}{n} \sum_{t=1}^{n+K}\left(Y_{t}-\bar{Y}\right)\left(X_{t-k}-\bar{X}\right) \text { for } k=0,-1,-2, \ldots, K
$$

where $\bar{X}$ and $\bar{Y}$ are the means of the $X$ series and $Y$ series (5).
Cross covariance may be used to determine the propagation of a pattern within the traffic stream or the similarity in the traffic pattern in different lanes at the same points along the roadway. At a lag of zero for all stations, a definite peak in the crosscovariance function between Lanes 1 and 2 was found for each station. This indicates that the traffic pattern of the two lanes was similar over time at the same points along the roadway.

Table 2 gives certain cross covariance of occupancy values for Lane 2, 2:00 to 3:00 and $4: 00$ to $5: 00 \mathrm{p} . \mathrm{m}$. The maximum absolute value of cross covariance was selected as shown. Cross covariance with negative time lags was used in the free-flow period from 2:00 to $3: 00$ p.m. to obtain the approximate mean wave speed for the propagation of a disturbance downstream. The cross covariance with positive time lags was used for the $4: 00$ to $5: 00$ p.m. period in the hope that it would yield cross-covariance peaks reflecting the speed of propagation of a disturbance upstream in the case of congested behavior.

In the free-flow behavior of 2:00 to 3:00 p.m., it seems clear that disturbances in the traffic stream move through the freeway with the flow of traffic in a predictable manner at roughly the free-flow traffic speed. However, in congested operation, with regard to Table 2 and the cross-covariance analysis, there seemed to be no order in the transmission of disturbances.

If a wave is traveling at 10 mph ( $16 \mathrm{~km} /$ hour)-a little faster than Table 2 gives for the wave speed from Station 1 to 2 -then in 880 seconds the wave would only travel about 2.5 miles ( 4 km ). In other words, if the speed of 8 or $9 \mathrm{mph}(12.8$ to $14.5 \mathrm{~km} / \mathrm{hour}$ ) is realistic, our analysis would not detect the disturbance upstream of Station 2 because our maximum time lag was only 800 seconds.

## TRAFFIC-STREAM BEHAVIOR <br> WITHIN FLOW-DENSITY PARAMETER RANGES

For this analysis, flow-density curves were generated with the data, which were broken into hourly segments such that they were divided according to traffic condition, i.e., free flow, congested, and transitional. The analysis yielded three results: (a) a range of distinct linear behavior as shown in Figure 10, (b) a nonlinear range of be-
havior as shown in Figure 11, and (c) combined linear and nonlinear behavior as shown in Figure 12. [Scattergrams of Figs. 10, 11, and 12 were produced by FAKAD, a data analysis package developed by K. McDonald at the University of Essex. The origin of the axes is set to $(0,0)$. The length of each axis represents an equivalent range of the variables in terms of their standard deviations above and below the mean. The mark 0 on the scattergram indicates one observation at that point, 1 indicates two observations, and 9 indicates ten or more.]

Although the transition effects can make identification of the states difficult, the behavior did recognizably shift in time from one range to the other. Table 3 gives a breakdown of the station behavior. For Station 1, the shifts occurred at approximately 3:40 p.m. and 6:32 p.m. For Station 4, the shifts occurred at approximately 4:40 and 4:52 p.m.

From the time series analysis and flow-density behaviors, the investigation narrowed, seeking parameters that could distinguish between the two operational states.

## Development of a Criterion Function

The simplest criterion function is density. In this case, some density value k will distinguish free flow from congested behavior, i.e., the criterion function would be $\mathrm{Z}^{*}=\mathrm{k}$.

An advance on this ultrasimplistic approach would be a criterion function of flow and density variables. Multivariate discriminant analysis can provide a linear form of such

Figure 10. Flow-density curve for Station 1, Lane 1, 2:00 to 3:00 p.m.


Density k Veh. /Mi.

Figure 11. Flow-density curve for Station 1, Lane 1, 4:00 to 5:00 p.m.


Density k Veh. / Mi.

Figure 12. Flow-density curve for Station 1, Lane 1, 3:00 to 4:00 D.m.


Density k Veh. / Mi.

Table 3. Station flow-density behavior.

| Time Period, <br> p.m. | Linear <br> Behavior, <br> Station No. | Nonlinear <br> Behavior, <br> Station No. | Combined, <br> Station No. |
| :--- | :--- | :--- | :--- |
| 2:00-3:00 | All stations |  |  |
| $3: 00-4: 00$ | $3,4,5,6,7$ |  | 1,2 |
| 4:00-5:00 | $5,6,7$ | $1,2,3$ | 4 |
| $5: 00-6: 00$ | $5,6,7$ | $1,2,3$ | 4 |
| $6: 00-7: 00$ | $3,4,5,6,7$ |  | 1,2 |
| $7: 00-8: 00$ | All stations |  |  |

Figure 13. Criterion function for combined linear and nonlinear behavior.

a relationship (10). In this case, $Z^{*}$, the value of the discriminant function that best separates the two states, is

$$
\mathrm{Z}^{*}=\mathrm{aq}+\mathrm{bk}
$$

where $a$ and $b$ are constants, $q$ is a given flow value, and $k$ is a given density value.
This type of criterion function would appear similar to that shown in Figure 13.

## Evaluation of the Density Criterion Function

Analysis of the data for Station 1, Lane 1, from 3:00 to 4:00 p.m. revealed that, for uncongested data, average density

$$
\overline{\mathrm{k}}_{\mathrm{unc}}=39.4
$$

and the standard deviation

$$
\text { s.d. }\left(\mathrm{k}_{\mathrm{unc}}\right)=11.99
$$

For the congested data,

$$
\overline{\mathrm{k}}_{\mathrm{con}}=78.11
$$

and

$$
\text { s.d. }\left(\mathrm{k}_{\mathrm{con}}\right)=27.24
$$

These values were determined after the data points were assigned to a congested or uncongested classification after individual analysis.

If the density readings are assumed to be normally distributed, the point at which there is an equal probability that an observation is a member of each group is at a density of 51.3 vehicles/mile. At this point, the probability of misclassifying a data point is 0.1635 . Thus, the density criterion function is $\mathrm{k}^{*}=51.3$.

This criterion was applied to the 3:00 to 4:00 p.m. data set for Station 1, Lane 1, and was found to classify four congested points and eight uncongested points incorrectly. This compares with nine and twenty errors respectively that would be expected with normally distributed data and suggests the data are skewed away from the transition region.

The discriminant analysis was done using BMDØ2T (9). All congested data but one had $Z$ values greater than -0.002 and all uncongested data had $Z$ values less than $\mathbf{- 0 . 0 0 2}$. Thus, a suitable criterion function for this particular segment of the data is $\mathrm{Z}^{*}=-0.002$. This discriminant function coefficients a and $b$ were 0.00002 and -0.00082 respectively, giving a criterion function of

$$
Z^{*}=0.0002=0.00002 q-0.00082 \mathrm{k}
$$

which can be rearranged to the form

$$
q=41 k-100
$$

This function correctly allocates all but one point. An F statistic value of 131.7 for 2,177 was obtained that is significant at the 99 percent confidence level. The point at which there is a probability of misclassification into either group is at $Z=0.0015$. This occurs at 1.13 standard deviations from each mean. Setting $Z=0.0015$ results in nine misclassifications, compared to one misclassification when the empirically derived $\mathbf{Z}=-0.002$ is used. This probably means that $\mathbf{Z}^{*}$ for each group is not normally distributed.

The same discriminant analysis was also done by using normalized variables, i.e.,

$$
z_{11}=\frac{q_{1}-\bar{q}_{1}}{\text { s.d. }(q)}
$$

instead of flow, and

$$
z_{21}=\frac{k_{1}-\bar{k}_{1}}{\text { s.d. }(q)}
$$

instead of density, for $\mathrm{i}=1,2, \ldots, 180$ (for each data point).
The discriminant function for Station 1, Lane 1, using normalized variables for the time period 3:00 to 4:00 p.m. was

$$
Z^{*}=0.00935 z_{1}-0.01536 z_{2}
$$

Thus, density seemed more important than flow in determining whether the freeway was congested or uncongested by a factor of about 1.65 . The difference in sign value of $z_{1}$ and $z_{2}$ meant that a high flow would give a high value of $Z$, and thus uncongested operation was more likely, whereas a high density would give a low value of Z , and thus congested operation was more likely.

## CONCLUSIONS

The strength of this further investigation of flow-density behavior was the data set: the data had a minimum of disturbance due to ramp activity and nonautomobile traffic.

Time-series analysis of traffic-stream occupancy was applied in two ways. Autocovariance reflected the behavior of a single point in the traffic stream over time. The autocovariance indicated random flow-density behavior for occupancy less than anproximately 15 percent and varying significant degrees of autoregressive behavior for higher occupancies. The different forms of observed autocovariance functions for various states of traffic-stream occupancy suggest potential use of this technique for detecting behavior under controlled conditions.

Cross covariance analyzed the behavior of two points of the traffic stream over time. This analysis indicated that, under free-flow conditions, disturbances in the traffic stream were propagated with the flow of traffic in a predictable manner at roughly the free-flow traffic speed. Cross-covariance results were not conclusive for congested conditions.

Analysis of flow-density behavior yielded distinct and discontinuous ranges of linear and nonlinear behavior. The existence of these two distinct states of behavior was previously indicated by time-series analysis, particularly the autocovariance functions. Further investigation of criteria to distinguish between the two states of behavior indicated that, although density was the more important parameter, a flow-density criterion function was superior to a simple density criterion function. Multivariate discriminant analysis further suggested that a flow-density criterion function would change over time because of differences in the recovery and breakdown processes.

## ACKNOWLEDGMENT

Appreciation is extended to A. G. R. Bullen, P. Doreian, N. Hummon, and J. Vilale for their suggestions. C. Tsai, M. Zubi, and N. Rattay contributed to background data manipulation. Many thanks are extended to Joseph McDermott, Project Director, Chicago Expressway Surveillance Project.

## REFERENCES

1. Edie, L. C. Car-Following and Steady-State Theory for Noncongested Traffic. Operations Research, Vol. 9, 1961, pp. 66-77.
2. Drake, J., Schoffer, J., and May, A. A Statistical Analysis of Speed-Density Hypotheses. In Vehicular Traffic Science (Edie, L. C., Herman, R., and Rothery, R., eds.), American Elsevier, New York, 1967.
3. Athol, P. Interdependence of Certain Operational Characteristics Within a Moving Traffic Stream. Highway Research Record 72, 1965, pp. 58-87.
4. Mika, H. S., Kreer, J. B., and Yuan, L. S. Dual Mode Behavior of Freeway Traffic. Highway Research Record 279, 1969, pp. 1-12.
5. Box, G. E. P., and Jenkins, G. M. Time Series Analysis-Forecasting and Control. Holden-Day, San Francisco, 1970.
6. Yamane, T. Statistics: An Introductory Analysis, 2nd Ed. Harper and Row, New York, 1967, pp. 870-874.
7. Marascuilo, L. A. Statistical Methods for Behavioral Science Research. McGrawHill, New York, 1971.
8. Kendall, M. A., and Stuart, A. The Advanced Theory of Statistics, Volume 3: Design and Analysis, and Time-Series. Hafner Press, 1966, p. 405.
9. Dixon, W. J. Biomedical Computer Programs, 2nd Ed., No. 2. Univ. of California Press, Berkeley, 1971.
10. Snedecor, G. W., and Cochran, W. G. Statistical Methods, 6th Ed. Iowa State Univ., Ames, 1967.

[^0]:    Publication of this paper sponsored by Committee on Freeway Operations and Committee on Traffic Flow Theory and Characteristics.

[^1]:    ${ }^{\text {a }}$ Mean wave speed is determined by distance/time, assuming 1 mile ( 1.6 km ) between stations,

