# Urban Traffic Network Flow Models

JAMES C. WILLIAMS, HANI S. MAHMASSANI, AND ROBERT HERMAN

**Addressed in this paper are the development and comparative assessment of macroscopic network-level traffic flow models, which describe the behavior and interrelation between traffic variables defined at the network level. These variables include average speed, concentration, flow, the fraction of vehicles stopped in the network, and the two-fluid running time variables. Three alternative sets of interrelated models, each with a different starting postulate, are presented and tested in terms of their performance against a series of microscopic simulation runs corresponding to different concentration levels. In each model system, a different functional form is postulated for either the speed–concentration relation or the fraction of vehicles stopped versus concentration relation. The functional form for the other relation is then derived from the postulated model by invoking the two-fluid theory of town traffic. The models are calibrated and tested using the simulation results. The analysis indicates that the network-level traffic variables are interrelated in a manner similar to that captured by the traffic models established for individual road sections. In particular, a well-known linear speed–concentration model as well as a nonlinear alternative are found to be generally applicable at the network level.**

Macroscopic models of traffic flow in urban networks describe the behavior and interrelation between traffic variables defined at the network level. The development of such network-level models remains in its infancy, especially in comparison to the extensive body of work, conducted over the past three decades, that addresses traffic in individual components of the network, such as arterials or intersections. Network-level characterization of traffic has important practical implications in terms of measuring traffic quality of service, comparing it with other cities, monitoring over time to evaluate the effect of various improvements, identifying deficiencies at the network level, and so on *(1, 2)*.

The most developed network-level traffic modeling approach is based on Herman and Prigogine's two-fluid theory of town traffic *(3, 4)*, which postulates a relation between the speed of moving vehicles and the fraction of running vehicles in a street network. Extensive field studies have been conducted in conjunction with this theory *(2, 4, 5)* supporting the validity of its basic premise and the resulting relation between the average running time per unit distance (the conditional expected trip time per unit distance taken only over moving vehicles) and the average total trip time per unit distance in the network. The sensitivity of the model's parameters to various physical and operational network characteristics has also been explored using simulation experiments *(6)*.

The latter approach actually circumvents what is undoubtedly the major obstacle hindering the development of network-level models, namely, the cost and difficulty of obtaining reliable data at the network level. The feasibility of using microscopic simulation as a tool to investigate network flow relations has been established by the authors in previous work, based on the NETSIM simulation package *(1, 6)*. In addition to the obvious cost and resource considerations, this approach allows the researcher a degree of experimental control that is not practical in actual traffic systems, as well as the ability to explore a wider range of situations than can be observed in field work.

Recently, Ardekani and Herman *(7)* extended the two-fluid modeling framework to include a set of relations between the principal network traffic variables. These relations are derived from a postulated functional relation between the average fraction of vehicles stopped in a network and the prevailing concentration, modifying an earlier form suggested in Herman and Prigogine's original work *(3)*. A speed–concentration model can then be derived from the postulated relation, provided that the two-fluid assumptions hold. Ardekani and Herman also calibrated the parameters of the postulated model, though only limited aerial photographic data were available for this purpose.

From a theoretical standpoint, network-level relations between flow, speed, and concentration cannot be analytically derived from models developed for individual components. On the one hand, there are issues in the very definition of the networkwide averages, as different averaging procedures could be devised. Moreover, the existence of "nice" relations among these quantities cannot, a priori, be taken for granted because the complexity of network interconnections effectively precludes the mathematical derivation of such models from basic principles or established road link and intersection level relations. However, earlier exploratory work by the authors has been encouraging in this regard. Definitions for average speed, concentration, and flow, all at the network level, have been presented, and the results of simulation experiments strongly suggested that these quantities appear to be related in a manner not unlike their counterparts at the individual road or arterial level *(1, 6)*.

The development and comparative assessment of network-level traffic flow models are traced in this paper. Three alternative sets of interrelated models, each with a different starting postulate, are presented and tested in terms of their performance against a series of microscopic simulation runs corresponding to different concentration levels. Each set of models constitutes a system of relations that comprehensively describe the joint behavior of the average speed, flow, and concentration, as well as the average fraction of vehicles stopped in the network, and the two-fluid stopped and running time variables. The next section presents the common theoretical background underlying the derivation of each set of models, given the starting postulate, and reviews the methodological approach followed in the supporting simulation experiments. The three model systems are subsequently presented in turn, and their parameters are calibrated using the results of the simulations. The evaluation of each model system's ability to provide a

J. C. Williams, Department of Civil Engineering, The University of Texas at Arlington, Tex. 76019. H. S. Mahmassani and R. Herman, Department of Civil Engineering, The University of Texas at Austin, Austin, Tex. 786712.

macroscopic characterization of network traffic flow phenomena can thus be performed. Concluding comments are presented in the closing section of the paper.

## THEORETICAL FRAMEWORK AND METHODOLOGICAL APPROACH

In this section the networkwide variable definitions are reviewed; then the principal relations comprising each model system are conceptually described, along with the analytical steps involved in the derivation of the models from a starting functional form relating any two of the variables. This is followed by pertinent information on the simulation experiments that provide the observations needed for the calibration of the various models. An initial presentation of the simulation results is also included to further examine the identity of flow to the product of speed and concentration for the definitions of the networkwide variables adopted here.

### Definition of Networkwide Traffic Variables

The three fundamental traffic variables (speed, concentration, and flow) have been generalized to the network level in previous work *(1)*. For completeness, these definitions are briefly reviewed here. All three are defined as average quantities taken over all vehicles in the network over some observation period $\tau$. Average speed $V$ [in miles per hour (mph)] is then obtained as the ratio of total vehicle miles to total vehicle hours in the network during the period $\tau$. The average concentration $K$ (in vehicles per lane mile) for the same period is the time average of the number of vehicles per unit lane length in the system; it can be simply calculated by dividing the total vehicle hours (during $\tau$) by $\tau L$, where $L$ is the total lane miles of roadway. One of the key advantages of using simulation experiments is that the analyst can maintain a constant concentration level in the network by keeping the number of circulating vehicles constant over the period $\tau$. This strategy, followed in previous work, is also adopted in all the experiments discussed in this paper. The concentration is then simply equal to the known total number of vehicles in the network divided by $L$.

Average network flow $Q$ is interpreted as the average number of vehicles that pass by an average point of the network, and given by $(\Sigma l_i q_i)/(\Sigma l_i)$, where $q_i$ and $l_i$ denote the average flow (during $\tau$) and the length of link $i$, respectively, and the summations are taken over all network links.

Another key variable of interest is $f_s$, the average fraction of stopped vehicles over the observation period. It is an important descriptor of the productivity of an urban traffic network, and is suggested by the dichotomization of traffic in the network into moving and stopped vehicles introduced by the two-fluid theory *(3)*. The variation of $f_s$ with the prevailing network concentration $K$ is one of the principal relations considered in this paper, as seen later in this section. Other two-fluid variables of interest here are the average running time $T_r$, stopped time $T_s$, and total trip time $T$, all per unit distance; of course, $T = T_r + T_s$. The main result of the two fluid theory is a model relating $T_r$ to $T$ (or $T_s$), which is invoked in all the following derivations. A detailed presentation of the two-fluid model assumptions and derivations is provided elsewhere *(3, 4, 6)*; the results necessary for the derivations in this paper are presented in the next subsection. Note that one pertinent result was used in the simulations to calculate $f_s$, namely, that $f_s = T_s/T$, derived mathematically by Herman and Ardekani under steady-state conditions *(4)*, and verified numerically in earlier simulation experiments.

### The Traffic Models and Their Interrelation

The model systems of interest can be derived by first specifying a functional relation between any two of the following variables: $V$, $K$, $Q$, and $f_s$. The remaining relations can then be analytically obtained by invoking one or both of the following:

1. The identity $Q = KV$, known to hold for individual roads but not formally established at the network level, has been numerically verified in earlier simulation work for the preceding network-level variable definitions *(6)*. This point is further addressed in this paper for a new set of simulations.
2. The two-fluid model relating $T_r$, $T$, and $T_s$ was defined earlier. However, because $T_r = 1/V_r$ (by definition) and $f_s = T_s/T$, there is a relationship between $f_s$ and $V_r$ (see Equation 4), which ultimately leads to a relation between $f_s$ and $V$. In all three model systems discussed, the two-fluid model is assumed to hold. Comments about the range of applicability of this assumption are also given on the basis of the simulation results.

In each of the three model systems discussed in this paper, a different functional form will be postulated for either the $V$–$K$ relation, or the $f_s$–$K$ relation. The functional form for the other relation is then derived from the postulated model by invoking the two-fluid relation.

The relations composing each model system are

$$V = f(K) \tag{1}$$

$$Q = g(K) \tag{2}$$

and

$$f_s = h(K) \tag{3}$$

In addition, as noted previously, the following are assumed to hold in all cases:

$$Q = KV$$

and

$$V_r = V_m(1 - f_s)^n \tag{4}$$

where

- $V_r$ = the average speed of the moving vehicles in the network over the observation period,
- $V_m$ = a parameter that can be interpreted as the average maximum running speed in the network (without any stopping), and
- $n$ = a parameter that captures the sensitivity of

running speed to the fraction of vehicles stopped in the network and can serve as an indicator of quality of traffic service in a network.

Equation 4 is the fundamental assumption of the two-fluid model, specifying the dependence between the average speed of the moving cars and the fraction of cars that are moving, $f_r = 1 - f_s$.

The functions shown in Equations 1 through 3 are different in each of the three model systems, as they depend on the starting assumption regarding either Equation 1 or Equation 3. Then, because $V = V_s f_r = V_r (1 - f_s)$ (by definition of the corresponding averages), the following is obtained using Equation 4:

$$V = V_m (1 - f_s)^{n+1} \tag{5}$$

Then, if $h(K)$ is given, substituting it for $f_s$ in Equation 5 yields $V = f(K)$, as follows:

$$V = V_m [1 - h(K)]^{n+1} \tag{6}$$

Similarly, if $f(K)$ is given, substituting it for $V$ in Equation 5 allows the derivation of a functional form for $f_s = h(K)$:

$$f(K) = V_m (1 - f_s)^{n+1} \tag{7}$$

which can be rewritten as

$$f_s = 1 - [f(K)/V_m]^{[1/(n+1)]} \tag{8}$$

The function $Q = g(K)$ can then be easily obtained by substituting $f(K)$ into the identity $Q = KV$. Similarly, a relation between $Q$ and $V$ can also be easily derived, but will not be pursued further in this paper. Before the specific functional forms of each model system are discussed, the simulation experiments conducted to calibrate the models are described.

## The Simulation Experiments

The same methodological approach used in earlier work is followed in this paper, and will therefore not be detailed here *(1, 6)*. A series of six simulation runs was performed to evaluate the network flow models of interest. Each of these runs used the same basic network configuration and features, but involved a different vehicular concentration level. As noted earlier, the network is treated as a closed system, with a fixed number of vehicles maintaining a constant concentration throughout the observation (accumulation) period. The circulation of the vehicles is controlled by an elaborate set of detailed microscopic rules (governing car following, queue discharge, lane switching, gap acceptance, and so on) embedded in the NETSIM package *(8, 9)*. Note that some modification of the code was necessary in order to keep track of stopped versus running time in a manner that is consistent with the two-fluid definitions of these variables.

Essentially, each simulation run yields one data point, consisting of values of the desired network-level average quantities taken over the observation period [generally about 15 min of accumulation, following an initial startup and loading period ranging from 5 to 15 min *(1)*]. The basic network configuration is the same as that used in previous work, and consists of 25 nodes, arranged in a 5-node by 5-node square, connected by two-way, four-lane streets forming a regular grid similar to that in the central business district (CBD). Each two-way street is represented by a pair of directed (one-way) links (80 in all). All links used in these simulations are 400 ft long, with no special geometric features such as turning bays or grades. Vehicles are injected into the network via 12 additional entry links placed around the perimeter, three to a side, with each entry connecting a source node (not part of the 25 network nodes) to a noncorner boundary node.

When the desired number of vehicles entered the network, they were not allowed to exit (no sink nodes were specified), thereby maintaining a constant concentration. At the interior intersections 10 percent of the traffic turned left, 15 percent turned right, and the remaining 75 percent continued straight through. When vehicles reached a boundary node, they divided equally between left and right turns, except at the corner nodes where no choice is available. These traffic movements were used throughout the work reported herein.

Each network node corresponds to a signalized intersection, except for the four corner nodes, which are unsignalized because they present no conflicts to the drivers. Timing of signals at the interior nodes followed a two-phase timing scheme with a 50/50 split and no protected turning movements. Three-phase signals were used at the boundary nodes, providing a protected left turn for vehicles reentering the interior of the system, with the green time nearly equally allocated to vehicles leaving the interior of the network and those reentering it from the boundary. Two-way progression at the mean desired speed (35 mph in this case) was provided along the interior arterials by using 40-sec cycle length with single alternate operation.

The network concentrations in these runs are approximately 10, 20, 40, 60, 80, and 100 vehicles per lane-mile (the actual values, determined by simulation input practicalities, are 9.90, 19.80, 41.58, 61.38, 81.18, and 100.65 vehicles per lane-mile), representing a range from very light to extremely heavy traffic. For comparison purposes, the highest concentration observed in related field work *(7)* was approximately 30 vehicles per lane-mile in the Austin CBD during peak traffic conditions. For all practical purposes, network concentrations in excess of 50 vehicles per lane-mile can be considered as very high, and are rather unlikely in actual operations. Therefore, the simulation results permit an exploration of network behavior over a wide spectrum of concentrations, including conditions near saturation, which may not be easy to observe in actual street networks.

The basic network-level results of interest from these runs are presented in Figures 1–4. The variation of speed and flow with concentration is shown in Figures 1 and 2, respectively; the corresponding variation of speed with flow in these runs is shown in Figure 3; and the variation of the fraction of vehicles stopped with concentration is shown in Figure 4. The general similarity of the $K$–$V$–$Q$ patterns to their characteristic counterparts for individual road sections is striking; this observation forms the basis for two of the suggested models explored in the following sections. The results of the simulations are first examined to verify that the $Q = KV$ identity holds for the network flow under consideration.

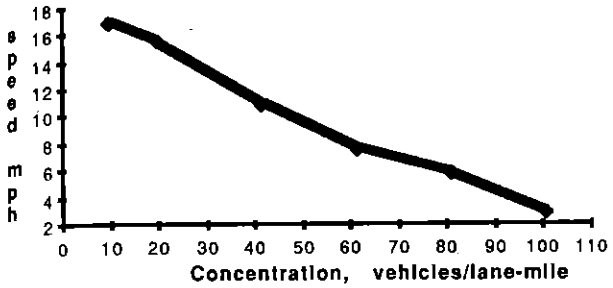


FIGURE 1 Plot of speed versus concentration for simulation experiments.

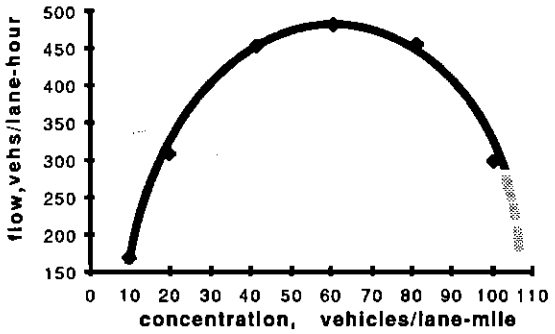


FIGURE 2 Plot of flow versus concentration for simulation experiments.

## The $Q = KV$ Identity

In earlier simulation experiments, the basic traffic flow identity relating speed, flow, and concentration was verified at the network level as well, for the proposed definition of the networkwide averages *(1)*. This identity and the procedure for measuring the variables are briefly addressed here for the new simulations, particularly because this identity plays a very important role in the development of the network flow model systems. The average flow $Q$ (in vehicles per lane-hour) for each run is estimated by averaging, over the accumulation period, the minute-by-minute average flows taken over all links, based on the number of vehicles discharged from the downstream end of each link (each minute). Each link's discharge is first divided by two to obtain the flow on a per-lane basis. Note with regard to the expression given earlier in this section for $Q$ that the link length $l_i$ is the same across all links in these experiments. The other network variables, $K$ and $V$, are obtained as described earlier in this section. The product $KV$ can then be compared with the independently estimated $Q$.

The results of this comparison for the six runs are given in Table 1, clearly illustrating the closeness of the two quantities. The small numerical discrepancies are mostly due to the 1-min discretization used in estimating $Q$, whereas $V$ is determined from semicontinuously accumulated quantities. The largest differences occur at the high concentration levels, where network operation is least stable. These results, coupled with earlier

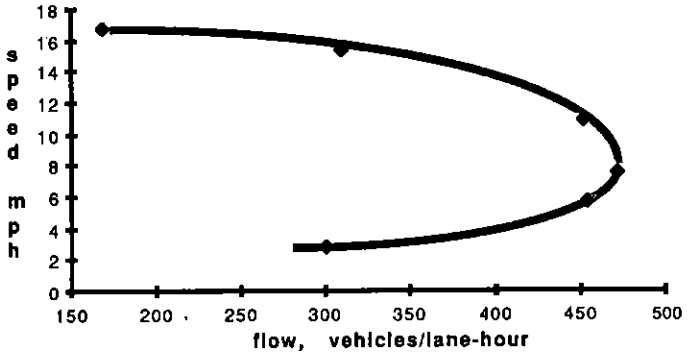


FIGURE 3 Plot of speed versus flow for simulation experiments.

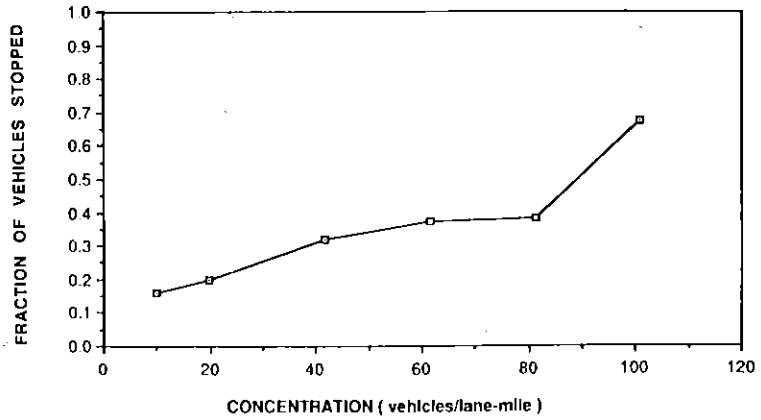


FIGURE 4 Plot of fraction of vehicles stopped ($f_s$) versus concentration for simulation results.

TABLE 1 RESULTS OF THE $Q$, $KV$ COMPARISON

| $K$ (vehicles per lane-mile) | $V$ (miles per hour) | $KV$ (vehicles per lane-hour) | $Q$ (vehicles per lane-hour) | Percent Difference $[(KV - Q)/Q]$ |
|---|---|---|---|---|
| 9.90 | 16.836 | 166.7 | 168.6 | −1.1 |
| 19.80 | 15.418 | 305.3 | 309.6 | −1.4 |
| 41.58 | 10.904 | 453.4 | 453.0 | 0.1 |
| 61.38 | 7.592 | 466.0 | 473.1 | −1.5 |
| 81.18 | 5.751 | 466.9 | 454.5 | 2.7 |
| 100.65 | 2.881 | 290.0 | 300.3 | −3.4 |

similar results *(1)*, further confirm the appropriateness of $Q = KV$ for the networkwide averages defined in this work.

Next, the three model systems are presented and discussed on the basis of their calibrated performance against the observations generated by the simulations. The first is from Ardekani and Herman *(7)*, which starts with a postulated functional form for the $f_s = h(K)$ function. The other two are based on functional forms for the $V$–$K$ relation $f(K)$; these forms are well-known for this relation on individual road segments, and their appropriateness at the network level will be examined.

## MODEL SYSTEM 1 RESULTS

Model system 1 was derived by Ardekani and Herman *(7)*, who started with the following postulated functional form for $h(K)$, which specifies the $f_s$–$K$ relation:

$$f_s = f_{s,min} + (1 - f_{s,min})\ (K/K_j)^{\pi} \tag{9}$$

where

- $f_{s,min}$ = a parameter intended as the minimum fraction of vehicles stopped in the network;
- $K_j$ = a parameter intended as the "jam" concentration at which the network is effectively saturated; and
- $\pi$ = a parameter that determines the fraction of vehicles stopped at a given partial concentration $(K/K_j)$, and could serve as a measure of the quality of service in a network.

The principal assumptions of this relation are reflected in its boundary conditions. In particular, it recognizes that in an urban street network, even under very low concentrations, there is some nonzero $f_{s,min}$ fraction of stoppage that is inevitable in the network (unless traffic control is fully responsive, unlike any currently in operation, or drivers do not obey traffic laws), as suggested by simulation results (see Figure 4) and field studies *(7)*. On the other hand, $f_s$ goes to 1, meaning that all vehicles are stopped, as $K$ goes to $K_j$. Thus, Equation 9 states that $f_s$ is an increasing function of $K$ that varies in the range from $f_{s,min}$ to 1 as $K$ goes from zero to $K_j$.

Following the steps presented in the previous section, $V = f(K)$ can then be found by substituting Equation 9 into Equation 6, yielding

$$V = V_m (1 - f_{s,min})^{n+1}\ [1 - (K/K_j)^{\pi}]^{n+1} \tag{10}$$

The flow-concentration model is then easily found using $Q = KV$, as shown in the previous section:

$$Q = KV_m (1 - f_{s,min})^{n+1}\ [1 - (K/K_j)^{\pi}]^{n+1} \tag{11}$$

The estimation of parameters of Equation 9 can be performed by rewriting the equation as $f_s = a + bK^{\pi}$, where $a$ equals $f_{s,min}$ and $b$ equals $(1 - f_{s,min})/K_j^{\pi}$. Nonlinear least squares estimates can then be obtained for $a$, $b$, and $\pi$, from which the original parameters can be recovered, yielding for the set of observations generated from the simulation experiments: $\hat{f}_{s,min} = 0.187$, $\hat{K}_j = 134.12$ vehicles per lane-mile, and $\hat{\pi} = 2.08$. Figure 5 depicts the curve corresponding to the estimated parameter values, and also includes the points observed in the simulation experiments for comparison purposes. It appears that the estimated model somewhat overestimates $f_{s,min}$. More seriously, between (the practically meaningful) concentrations of 20 and 80 vehicles per lane-mile, it predicts the opposite concavity to that suggested by the observed values.

In order to examine the ability of this model to describe the $V$–$K$ pattern, by comparing the resulting Equation 10 with the observed simulation results, it is necessary to estimate the two-fluid parameters $n$ and $T_m$. These can be found by performing a least-squares estimation of the linear regression equation *(4, 6)*:

$$\ln T_r = [1/(n + 1)]\ \ln T_m + [n/(n + 1)]\ \ln T \tag{12}$$

With all six data points, the two-fluid parameter estimates were found to be $\hat{n} = 1.051$ and $\hat{T}_m = 2.692$ min/mile. However, it should be noted that, in this analysis, the two-fluid model is assumed to hold at very high concentrations, which have simply never been encountered in the extensive and successful field validation studies. Moreover, the ability of the simulation package, which is clearly in agreement with the two-fluid assumptions over the range of practically meaningful concentrations, to reliably simulate what happens at very high concentrations, where conditions are inherently unstable, cannot be taken for granted. The two-fluid model assumes the $\ln T_r$ versus $\ln T$ trend to be linear; however, the plot of the six observed points (Figure 6) suggests that the sixth point (representing the highest concentration examined) deviates from the linear trend established by the other five points. Therefore, $T_m$ and $n$ were reestimated, omitting the sixth point, yielding 1.809 min/mile and 2.349, respectively.

The resulting curves for the speed–concentration relation are shown in Figure 7, where the curve labeled Method 1 refers to the case where the estimated values of $n$ and $T_m$ are based on all six data points, and Method 2 refers to the case in which the sixth point was omitted from the estimation data. Both curves exhibit the same basic shape, although several predictable differences can be seen. In particular, because of the underlying calibration, the Method 1 curve can be expected to provide a better fit to the observed values at very high concentrations (greater than approximately 70 vehicles per lane-mile), while the Method 2 curve performs better in the lower-to-medium concentration ranges. However, neither curve fits the observed

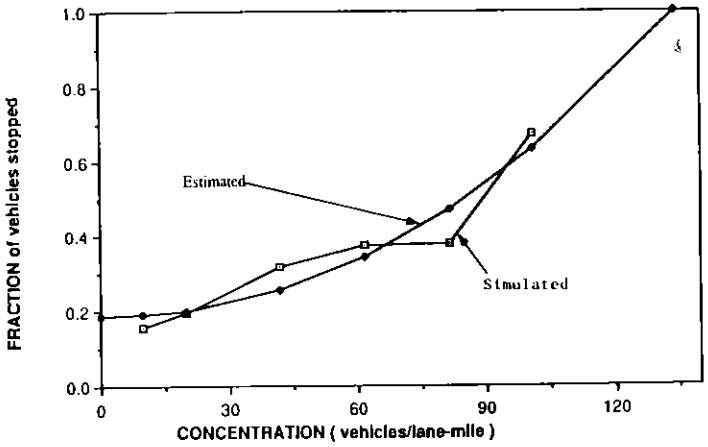


**FIGURE 5** Comparison of Ardekani and Herman's $f_s$–$K$ model with observed simulation results.

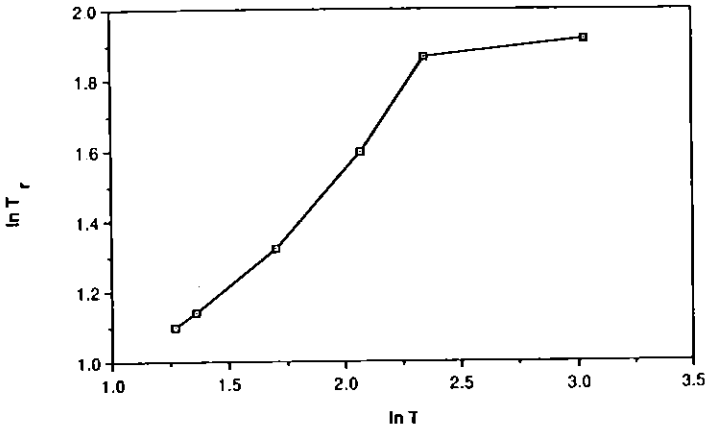


**FIGURE 6** Plot of $\ln T_r$ $\ln T$ for simulation experiments.

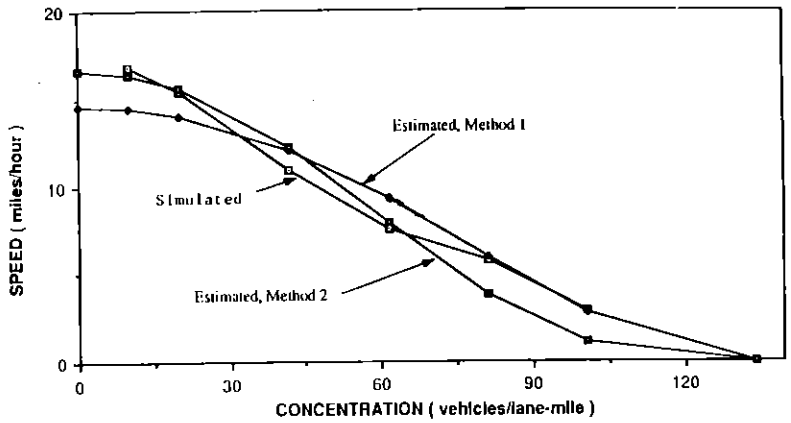


**FIGURE 7** Comparison of Ardekani and Herman's $V$–$K$ model with observed simulation results.

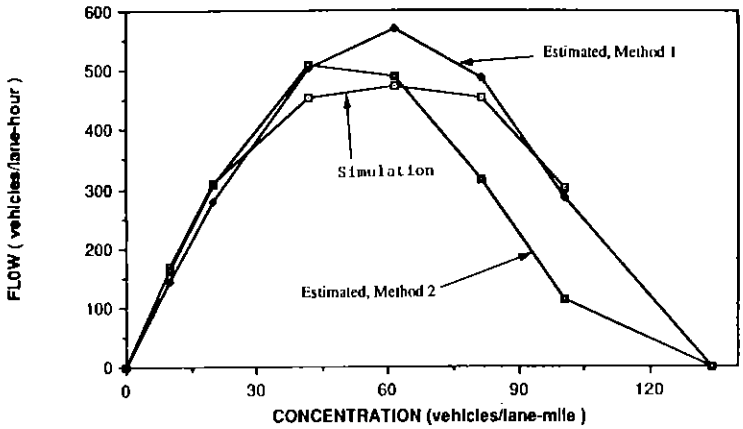


**FIGURE 8 Comparison of Ardekani and Herman's $Q$–$K$ model with observed simulation results.**

data particularly well: at most, the model overpredicts speed at medium concentrations (30 to 60 vehicles per lane-mile), but underpredicts speed at lower concentrations. These discrepancies correspond to those encountered in connection with the plot of the fraction of vehicles stopped versus concentration in Figure 5, as discussed earlier. Finally, the flow–concentration relation (Equation 11) is shown in Figure 8 for both sets of two-fluid parameter estimates, along with the observed values. Neither curve fits the data particularly well, exhibiting the same kind of problems seen with the $V$–$K$ relation, especially in the medium concentration range. The next two model systems, which directly specify the $V$–$K$ equation, perform considerably better in this regard as will be observed.

## MODEL SYSTEM 2: LINEAR $V$–$K$

The starting point for this set of models, as well as for the one discussed in the next section, is a postulated functional form for $f(K)$, which specifies the network-level speed–concentration relation. As mentioned earlier, the plots in Figures 1 through 3 remarkably evoke the patterns encountered for individual road facilities. Two well-known forms used in conjunction with individual facilities are therefore calibrated and evaluated for the network-level variables. The first, which forms the basis of this model system, is Greenshields' widely used linear speed-concentration relation, namely *(10)*:

$$V = V_f(1 - K/K_j) \tag{13}$$

where $V_f$ and $K_j$ are parameters to be estimated and interpreted as the mean free speed (experienced when interference from other vehicles in the traffic stream is virtually nonexistent) and the "jam" concentration, respectively, defined earlier.

Note that the $V_f$ parameter is distinct from the two-fluid parameter $V_m$, the average maximum running speed (equal to $1/T_m$), which according to Equation 4, occurs when $f_s$ equals zero, that is, when no vehicles are stopped in the network. As stated in the previous section, this condition does not occur in existing urban traffic systems over any meaningful length of time, even at very low concentrations, because of the presence of the traffic control system. Therefore, because the average travel time corresponding to the free mean speed will exceed the minimum running time $T_m$ usually by some nonzero stopped time, $V_f \leq V_m$ will always exist, and, in most cases, $V_f < V_m$. Note that in Ardekani and Herman's model system discussed in the previous section, $V_f = V_m(1 - f_{s,min})^{n+1}$ (by setting $K = 0$ in Equation 10).

The functional form for $h(K)$, the $f_s$–$K$ relation that is compatible with the foregoing linear $V$–$K$ model and the two-fluid model, is found following the steps presented earlier. In other words, substituting Equation 13 into Equation 8 yields

$$f_s = 1 - [(V_f/V_m)(1 - K/K_j)]^{1/(n+1)} \tag{14}$$

The boundary conditions of this function are similar to those of Ardekani and Herman's function because $f_s = 1 - (V_f/V_m)^{1/(n+1)}$ for $K = 0$, and $f_s = 0$ for $K = K_j$. However, for the former, the minimum fraction of vehicles stopped, $f_{s,min}$, is explicitly stated in terms of the speed parameters $V_f$ and $V_m$. For instance, it becomes clear in this expression that if $V_f = V_m$, then $f_s = 0$ at $K = 0$.

Finally, the flow–concentration relation is again obtained by applying $Q = KV$:

$$Q = V_f(K - K^2/K_j) \tag{15}$$

Least squares estimates of the model parameters can be easily obtained because Equation 13 is linear in $K$, and can be rewritten as $V = b_0 + b_1 K$. The original parameters can then be recovered from the estimated values as $V_f = b_0$ and $K_j = -V_f/b_1$. With the observed simulation data of Figure 1, the estimated parameters are $\hat{V}_f = 18.02$ mph and $\hat{K}_j = 116.3$ vehicles per lane-mile. The resulting calibrated functions, along with the observed values, are plotted in Figures 9 and 10, for Equations 13 and 15, respectively, indicating in both cases a rather close fit to the data. Moreover, the mean free speed does not appear to be underpredicted; however, because Equation 13 ignores the slight nonlinear trend exhibited by the $K$–$V$ data, the speed is slightly overpredicted in the middle concentration range, resulting in an overprediction of the flow in the same range (Figure 10).

Note that there is no need to exclude the sixth data point (for the highest $K$) as in Method 2 in the previous section when estimating the parameters of Equation 13 because the postulated $K$–$V$ model and resulting $Q$–$K$ relation appear to be

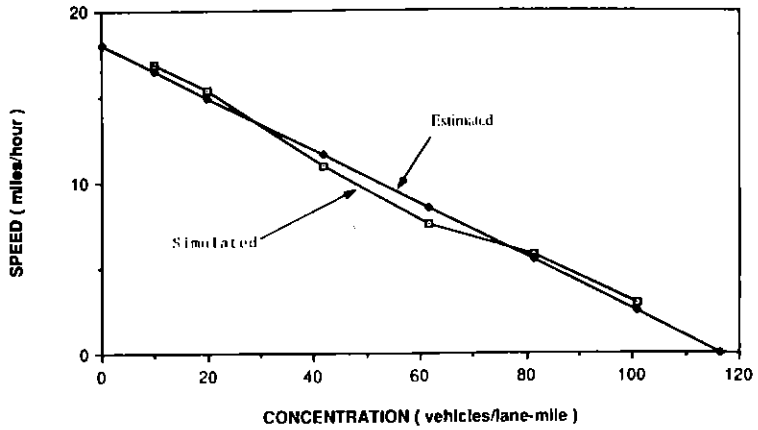


**FIGURE 9 Comparison of linear $V$–$K$ model with observed simulation results.**

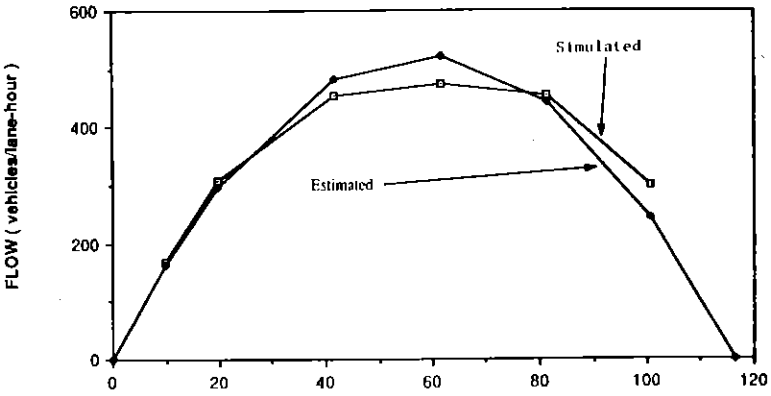


**FIGURE 10 Comparison of $Q$–$K$ model derived from linear $V$–$K$ relation with observed simulation results.**

consistent with the pattern exhibited by the data, including the range corresponding to that high concentration point. The main justification for excluding that point in the previous section is that the discrepancy appears with regard to the two-fluid model assumption's applicability at higher concentrations. Because the $V$–$K$ function has been specified directly here, this assumption has no direct bearing on the parameter estimates presented so far in this section. However, this is not the case when Equation 14 is evaluated for the $f_s$–$K$ relation because this function's parameters are not estimated directly on the $f_s - K$ data but from the two-fluid and $V$–$K$ models separately (just as the parameters of Ardekani and Herman's $V$–$K$ model were not directly estimated using the corresponding $V$–$K$ data).

Evaluation of the $f_s$–$K$ function in this model system requires the estimation of the two-fluid parameters, $n$ and $T_m$. Those found by Methods 1 and 2 in the previous section are both used here. The results are plotted, along with the observed data, in Figure 11. If Method 2 is used in the two-fluid estimation, the resulting $f_s$–$K$ curve provides a much better fit to the data, up to a concentration of about 80 vehicles per lane-mile, than that obtained by Method 1. As with Ardekani and Herman's model, this model's concavity is opposite to that exhibited by the data in the middle range of concentrations. What is interesting, however, is that this model's fit to the $f_s$–$K$ data, with which it was not directly calibrated, is almost as good as that obtained with Equation 9, which was directly calibrated using that data. Actually, neither model seems capable of reflecting the concavity of the data over the middle range of concentrations, or the apparent inflection point before the higher concentration values.

## MODEL SYSTEM 3: NONLINEAR $V$–$K$

The second postulated functional form for the $V$–$K$ model is a nonlinear bell-shaped function originally proposed by Drake et al. for arterials *(10, 11)*, which could capture the apparent shape of the speed–concentration data in Figure 1:

$$V = V_f \exp[-\alpha\,(K/K_m)^d] \tag{16}$$

where $V_f$ (previously defined), $K_m$, $\alpha$, and $d$ are parameters to be estimated. The resulting expressions for the $f_s$–$K$ and $Q$–$K$ models are then obtained as shown previously, yielding

$$f_s = 1 - \{(V_f/V_m)\exp[-\alpha\,(K/K_m)^d]\}^{1/(n+1)} \tag{17}$$

and

$$Q = KV_f \exp[-\alpha\,(K/K_m)^d] \tag{18}$$

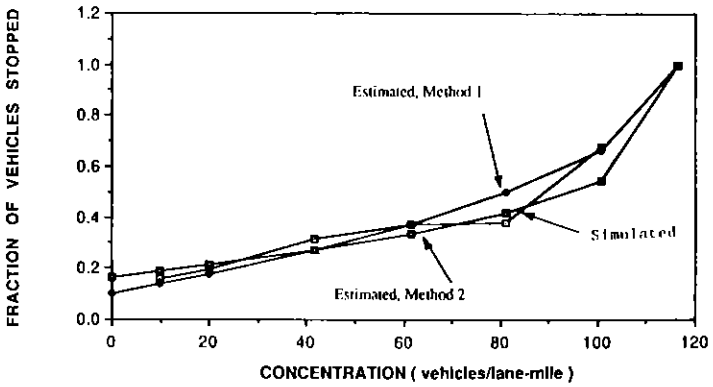


**FIGURE 11 Comparison of $f_s$–$K$ model derived from linear $V$–$K$ relation with observed simulation results.**

It can be shown, by solving $dQ/dK = 0$, that $K_m$ is the concentration at which maximum flow occurs in the network.

Estimation of the parameters of Equation 16 can be accomplished by nonlinear least squares after some manipulation. Taking the natural logarithms of both sides of the equation, and rearranging the last term on the right-hand side (RHS) yields

$$\ln V = 1n\ V_f - (\alpha/K_m^d)K^d \tag{19}$$

which is of the form $\ln V = c_0 + c_1 K^d$. Nonlinear least squares estimates can then be found for $c_0$, $c_1$, and $d$, and it is then possible to recover $V_f = \exp(c_0)$. But, because $c_1 = (\alpha/K_m^d)$, and only $c_1$ and $d$ are known, it is not possible to obtain unique values of $\alpha$ and $K_m$. However, this is not particularly problematic from the standpoint of the model's performance because only the value of $c_1$ is necessary. With the observed simulation results, the parameter estimates obtained by the foregoing technique are $\hat{d} = 1.49$, $\hat{V}_f = 17.95$ mi/hour, and $\hat{c}_1 = 1.83 \times 10^{-3}$. In order to gain an idea of the approximate magnitude of $\alpha$, the property that $K_m$ represents, the concentration at maximum flow, can be used. Noting from Figure 2 that the maximum flow seems to occur for a concentration between 60 and 65 vehicles per lane-mile, a range of $\alpha$ can be calculated given the estimated values of $c_1$ and $d$, yielding a range from 0.81 (for $K_m = 60$) to 0.91 (for $K_m = 65$).

In Figure 12 it can be seen that this model provides a very close fit to the observed data throughout its range; it also provides a reasonable estimate of the free mean speed. The resulting flow–concentration model, given by Equation 18, is plotted in Figure 13, along with the observed points. Unlike the relation derived from the linear $V$–$K$ model, Equation 18 is not symmetric, and exhibits asymptotic decay of flow as the concentration increases. The model does, however, provide a close fit to the observed data (the one near miss being at $K = 80$ vehicles per lane-mile, apparent in both Figures 12 and 13).

The estimated values of the two-fluid parameters are again used to examine the descriptive ability of the $f_s$–$K$ relation derived in this model system (Equation 17). This function is plotted in Figure 14 for the two sets of two-fluid parameter estimates presented earlier. The curve with the parameters found by Method 2 (excluding the last data point) fits the corresponding five points in Figure 14 much better than that with parameters found by using the Method 1 estimates because the latter seem to be noticeably influenced by the highest concentration point. Unlike the previous two model systems discussed in this paper, maximum $f_s$ (which, for plausible signs of the parameters, tends to 1 in the limit) is approached asymptotically. Whether this behavior is correct or not is difficult to determine, given that it requires loading the network to very unrealistic levels. Moreover, the performance of the $f_s$–$K$

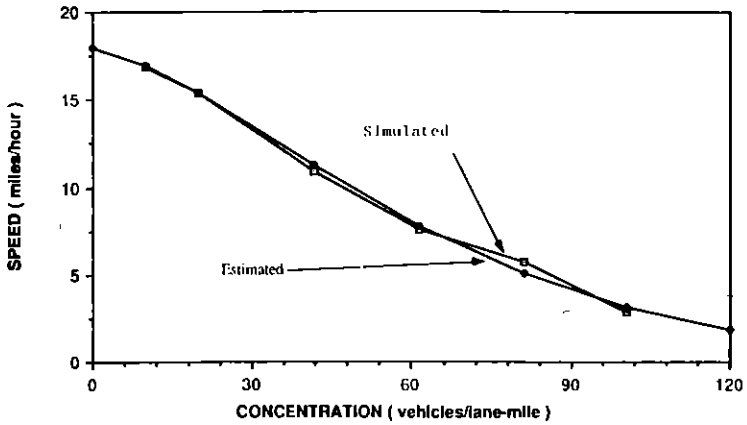


**FIGURE 12 Comparison of bell-shaped $V$–$K$ model with observed simulation results.**

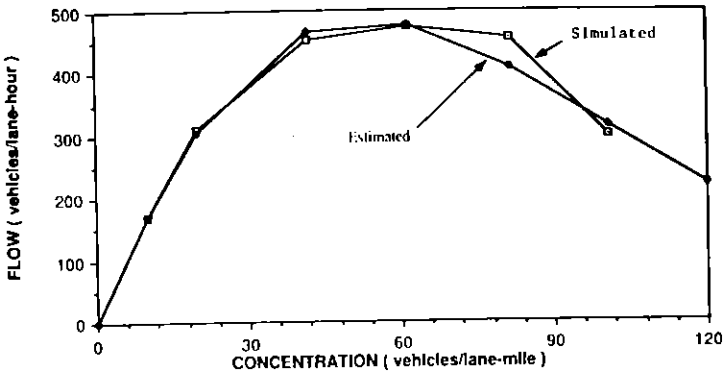


**FIGURE 13 Comparison of $Q$–$K$ model derived from bell-shaped $V$–$K$ relation with observed simulation results.**

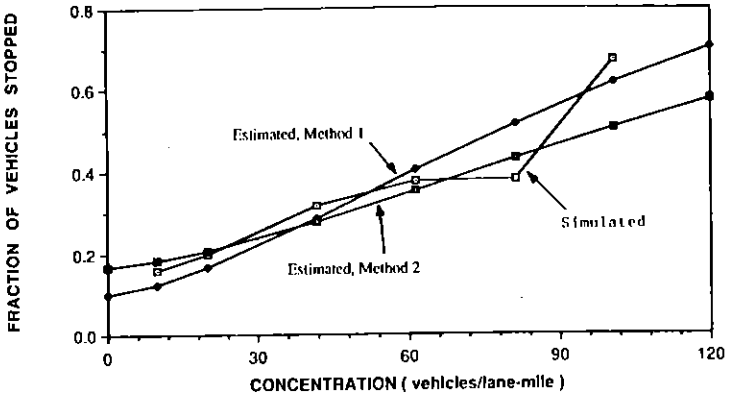


**FIGURE 14 Comparison of $f_s$–$K$ model derived from bell-shaped $V$–$K$ relation with observed simulation results.**

model does not depend only on the postulated $V$–$K$ function, but also on the two-fluid model that is invoked in its derivation. As seen earlier, the simulation data appeared to significantly depart from the two-fluid trend at higher concentrations. More importantly, Equation 17 seems to provide as good a fit as any of the other models, and in some cases better, for the range of concentrations that are likely to be encountered in actual urban traffic networks.

## CONCLUDING COMMENTS

The most important conclusion from the work presented in this paper is that it is possible to characterize traffic flow in urban street networks using relatively simple macroscopic models relating the principal networkwide traffic variables. These relations extend beyond the established two-fluid model to include average network speed, concentration, flow, and the fraction of vehicles stopped in the network. Furthermore, it is remarkable that these relations appear to be not unlike those that have been established at the individual facility level. As illustrated in this paper, the characteristic shape of the fundamental traffic relations encountered for highways and arterials, and on which traffic engineering procedures have been built, seem to be present at the network level as well, despite the complex interactions that take place in urban street networks.

In terms of the relative performance of the model systems presented in this paper, all three seem to provide a more or less reasonable approximation of the patterns exhibited by the simulated data, at least over certain ranges of concentration. Model Systems 2 and 3 provided a much better fit to the speed–concentration data (compared to the derived model in the first system) over the full range of observations; however, this is to be expected because these model systems were calibrated using that data. Generally, the relation between the fraction of vehicles stopped and concentration remains the most problematic in terms of finding a model that is sufficiently convincing over the full concentration spectrum. If one limits the comparisons to concentrations up to about 80 vehicles per lane-mile, which already exceeds realistic values in actual networks, then the set of models based on the nonlinear (bell shaped) $V$–$K$ relation is probably the best performer overall, especially in terms of capturing the shape of the $V$–$K$ pattern. However, the linear approximation for that same relation, on which the second set of models was based, has the advantage of familiarity and ease of comparability.

Still focusing on the range of less than 80 vehicles per lane-mile, the results in this paper provide another demonstration of the two-fluid model's validity. This model served as the principal theoretical bridge between the postulated function (generally selected to fit one particular relation) and the derived function in each of the three model systems. In all three cases,

the derived function performed remarkably well against the observed data, even though it was not directly calibrated using that data. However, the behavior of various network variables at very high concentrations remains to be understood. Based on the simulation results presented here, the applicability of the derived functions at these levels decreased markedly in all three cases, suggesting that the two-fluid assumptions might not be directly extended to these high concentration levels. For instance, it is possible that the relation between $\ln T_r$ and $\ln T$ may be nonlinear over the full spectrum of concentrations, although it can be treated as effectively linear over most of the practically meaningful range of concentrations.

Naturally, because the results presented here are based on a limited set of simulated experiments, they must be interpreted in an exploratory sense aimed at stimulating further inquiry into this problem area. Nevertheless, the same general patterns described here were also observed in a large number of other simulations conducted by the authors as part of an ongoing study of the sensitivity of network relations to various network characteristics. These simulations were still performed for relatively small networks, under controlled conditions, and subject to the microscopic rules embedded in the NETSIM package. However, these simulations have been instrumental in supporting the derivation of network-level traffic flow relations and the investigation of their properties. Although simulations in larger, more elaborate networks will undoubtedly provide useful information to advance knowledge on this topic, it is essential to obtain networkwide data on the operation of actual urban traffic systems. The cost and the scale of the problem may appear discouraging; however, technological developments in remote-sensing, telecommunications, and optoelectronics, among others, offer challenging opportunities for learning about the workings of traffic in urban areas.

## ACKNOWLEDGMENTS

This paper is based on research funded in part by the General Motors Research Laboratories. The authors have benefited from many insightful discussions with Siamak Ardekani on this general topic, as well as from the beneficial comments of Richard Rothery. The authors remain solely responsible for the contents of the paper.

## REFERENCES

1. H. S. Mahmassani, J. C. Williams, and R. Herman. Investigation of Network-Level Traffic Flow Relationships: Some Simulation Results. In *Transportation Research Record 971*, TRB, National Research Council, Washington, D.C., 1984, pp. 121–130.
2. S. Ardekani and R. Herman. *Quality of Traffic Service*. Research Report 304-1. Center for Transportation Research, The University of Texas at Austin, 1982.
3. R. Herman and I. Prigogine. A Two-Fluid Approach to Town Traffic. *Science*, Vol. 204, 1979, pp. 148–151.
4. R. Herman and S. Ardekani. Characterizing Traffic Conditions in Urban Areas. *Transportation Science*, 1984, Vol. 18, pp. 101–139.
5. S. Ardekani, V. Torres-Verdin, and R. Herman. The Two-Fluid Model and the Quality of Traffic in Mexico City (in Spanish). *Revisto Ingenieria Civil*, Calegio de Ingenierus Civiles de Mexico, Jan./Feb. 1985.
6. J. C. Williams, H. S. Mahmassani, and R. Herman. Analysis of Traffic Network Flow Relations and Two-Fluid Model Parameter Sensitivity. In *Transportation Research Record 1005*, TRB, National Research Council, Washington, D.C., 1985, pp. 95–106.
7. S. Ardekani and R. Herman. Urban Network-wide Variables and Their Relations. *Transportation Science*, Vol. 21, No. 1, 1987.
8. Peat, Marwick, Mitchell and Company. *Network Flow Simulation for Urban Traffic Control System—Phase II, Volumes 1–5*. FHWA, U.S. Department of Transportation, Washington, D.C., 1973.
9. *Traffic Network Analysis with NETSIM—A User Guide*. Implementation Package FHWA-IP-80-3. FHWA, U.S. Department of Transportation, 1980.
10. D. L. Gerlough and M. J. Huber. *Special Report 165: Traffic Flow Theory: A Monograph*. TRB, National Research Council, Washington, D.C., 1975.
11. J. Drake, J. Schofer, and A. D. May. A Statistical Analysis of Speed-Density Hypotheses. *Proc., Third International Symposium on Theory of Traffic Flow*, American Elsevier, New York, N.Y., 1967.

*Publication of this paper sponsored by Committee on Traffic Flow Theory and Characteristics.*