Stochastic Process Approach to the Estimation of Origin-Destination Parameters from Time Series of Traffic Counts

GARY A. DAVIS AND NANCY L. NIHAN

The origin-destination (OD) matrix gives the volume of traffic from each of a region's origins to each of its destinations and is a fundamental input to transportation planning and network design activities. Because the traditional methods of estimating the OD matrix—surveys and trip generation/distribution modeling—tend to be expensive, cumbersome, and inaccurate, researchers have sought to develop methods for estimating the OD matrix from observations of traffic volumes on the region's road network. For simple linear networks, such as single intersections or freeway sections, OD estimators with desirable statistical properties can be developed using least-squares methods, but for general networks it has not yet been possible to produce consistent estimators of OD parameters using traffic count data alone. It is believed that the link counts on a traffic network are generated by a stochastic process that is parameterized by the means and variances of the separate OD flows. By using a tractable approximation to the traffic-generating process, it is possible to develop both maximum likelihood and method of moments estimators of OD parameters, and the estimators have desirable consistency and asymptotic normality properties. Simulation studies suggest that the maximum likelihood estimator, though efficient in its use of data, is computationally demanding, whereas the method of moments estimator is not computationally demanding but is statistically inefficient.

In transportation modeling, the origin-destination (OD) matrix is an array whose rows index the locations on a network where trips originate and whose columns index the locations where trips terminate. The entry at the intersection of a row and a column gives, for some predetermined time interval, the number of trips between that particular OD pair. The OD matrix is the fundamental summary of a region's demand for travel, so it is an important input to any transportation planning activity. Historically, OD matrices have been estimated using some combination of survey methods and trip generation/distribution modeling, but the data collection needed for these approaches tends to be time-consuming and expensive, and the result is often of unreliable accuracy (*1*). Because the OD matrix can be viewed as an input to a traffic assignment process the outputs of which are the traffic volumes on the network's links, an alternative approach to OD matrix estimation is to start with observed link volumes and somehow "invert" the traffic assignment to obtain the OD matrix. This approach appears increasingly attractive as the proliferation of automatic traffic surveillance and control systems makes automatic traffic count data more readily available. The clearest formal statement of this OD estimation problem to date has been given by Cascetta and Nguyen (*2*), although related work appears in Maher (*3*), Bell (*4*), and Spiess (*5*).

Before considering methods of OD estimation based on traffic count data, it is useful to review desirable properties of parameter estimators. An estimator is said to be consistent if the probability of large discrepancies between it and the true parameter value approaches zero as the amount of data used approaches infinity. Thus for large amounts of data, a consistent estimator is likely to be close to the true parameter value, and consistency can be regarded as a minimal necessary condition for an estimator to be useful. Consistency is a property of point estimators, but in addition to generating a point estimate of a parameter, it is often desirable to be able to compute confidence bounds for the estimate or to test hypotheses concerning parameter values. For a large class of estimators, including many maximum likelihood (ML) estimators, a useful theory of inference can be developed on the basis of the fact that, as the amount of data becomes large, the estimator tends to have a normal distribution. This property is called asymptotic normality. Because a parameter estimator is a random variable, it has a variance, and larger variances indicate greater uncertainty concerning the true parameter value. If the variance of a particular estimator about the true parameter value is lower than that of any other estimator, that estimator is called efficient.

PROBLEM FORMULATION

Imagine indexing the region's OD pairs by the single index $j = 1, \ldots, m$, and let d_j denote the demand for travel between OD Pair j. We can also imagine that counts are available from a total of p of the network's links. Define $\mathbf{y}$, a p-dimensional vector whose kth element y_k denotes the traffic count on the kth link containing a traffic counter, and $\mathbf{q}_j$, a p-dimensional vector whose kth element q_{jk} denotes the probability that a trip between OD Pair j uses Link k. The link use probabilities $\mathbf{q}_j$ are assumed to be constant, consistent with the assumption that the traffic assignment is in equilib-

G. A. Davis, Department of Civil and Mineral Engineering, University of Minnesota, 500 Pillsbury Drive SE, Minneapolis, Minn. 55455. N. L. Nihan, Department of Civil Engineering, FX-10, University of Washington, Seattle, Wash. 98195.

rium. Computation of these probabilities thus requires first solving an equilibrium assignment problem. Following Cascetta and Nguyen (*2*), the relation between the link counts **y** and the OD volumes d_j can then be given as a linear regression of the form

$$\mathbf{y} = \sum_{j=1}^{m} (d_j \mathbf{q}_j) + \mathbf{e} \tag{1}$$

where **e** denotes an error vector accounting for discrepancies between the link counts and their expected values.

If the rank of the matrix $[\mathbf{q}_1, \ldots, \mathbf{q}_m]$ is greater than or equal to m, the preceding regression problem is well posed and least-squares estimates of the OD parameters d_j can be readily computed. For networks of realistic size, however, the number of OD elements exceeds the number of links in the network, so that Equation 1 leaves the d_j underdetermined. A natural solution to this problem is to expand Equation 1 by collecting a time series of link counts $\mathbf{y}(1), \mathbf{y}(2), \ldots$ made at times $t = 1, 2, \ldots$ and replacing Equation 1 by an extended version:

$$\mathbf{y}(1) = \sum_{j=1}^{m} d_j \mathbf{q}_j + \mathbf{e}(1)$$

$$\mathbf{y}(2) = \sum_{j=1}^{m} d_j \mathbf{q}_j + \mathbf{e}(2) \tag{2}$$

However, for constant $\mathbf{q}_j$, it is easy to verify that the rank of the extended regression matrix will always equal the rank of $[\mathbf{q}_1, \ldots, \mathbf{q}_m]$, and the identifiability problem encountered above cannot be solved, even with an infinite sequence of observations. Thus in this formulation of the problem, link count data cannot provide a method of consistently estimating the OD parameters. This difficulty has been well recognized, and the focus of research on OD estimation has been on combining link count data with other data (such as from surveys) to provide usable procedures (*2–7*).

The situation for general networks contrasts with that for linear networks, such as single intersections, freeway sections, and transit routes. Here only one route connects each OD pair, and it is possible to obtain counts of the total traffic departing each origin (input counts) and arriving at each destination (output counts). Several papers have established that, given time-series observations of the input and output counts, least-squares–based estimators may be used to estimate the proportions of an origin's traffic that terminate at the various destinations, even when the number of OD parameters exceeds the number of count locations. This approach is originally due to Cremer and Keller (*8*), whereas a recent paper by Nihan and Davis (*9*) gives conditions and a proof for the strong consistency of ordinary least-squares in estimating intersection turning-movement proportions. Thus for linear networks, consistent estimators of OD parameters are readily constructed from least-squares algorithms, whereas for general networks, Equation 1 permits consistent estimation only if the number of counted links exceeds the number of OD pairs.

The source of this discrepancy lies in how the information available in a set of link counts is used. Equation 1 essentially states that the observed traffic counts are equal to the sum of a mean value and an error term, where the mean value vector is a linear function of the OD parameters. Even though the link count mean is consistently estimable from stationary observations, the consistency is not inherited by the estimates of the OD parameters because of the noninvertibility of the function relating the mean link flows to the OD parameters. In linear networks, on the other hand, the statistics of interest are not the mean values of the counts, but the covariance matrix between the input counts and the output counts. This matrix is also consistently estimable, and, given certain conditions on the process generating the input counts, an invertible relationship between the OD parameters and the covariance matrix elements can be established, permitting consistent estimation of the OD parameters. This suggests that if not only the mean values but also the covariance properties of link counts can be expressed as well-behaved functions of the OD parameters, consistent estimators of these parameters requiring only link count data can be constructed. In fact, it has been known for quite some time that the OD flows and the covariance of the link volumes have a well-defined and plausible connection (*10*), but it has been difficult to find an appropriate use for this knowledge. In large part this is because the process that generates traffic counts has nontrivial dynamics, so that a time series of link counts must be viewed as a realization of a stochastic process rather than as the result of random sampling. The temporal dependencies among the link counts often invalidate the use of classical statistical procedures, whereas attempts to develop dynamic, stochastic models of traffic assignment have either been restricted to simple networks (*11*) or have produced intractable models (*12*). However, recent work has investigated this problem in some detail (*13,14*) and established that under conditions similar to those needed to justify the use of stochastic user equilibrium (SUE) traffic assignment methods, a stochastic traffic generation model similar to that used by Cascetta (*12*) can be approximated by a stationary linear stochastic process driven by normally distributed noise as the number of travelers in the system becomes large. The parameters of this process are in turn well-defined functions of the OD parameters, and the approximation can be used to develop both ML and method of moments (MOM) approaches to OD parameter estimation. The approximation model is first presented in some detail, and the way the model is parameterized by the OD parameters is emphasized. OD estimation based on this model is then described.

A STOCHASTIC TRAFFIC MODEL

Before one can develop and validate statistical procedures, one must have an explicit model of the probabilistic mechanisms that are assumed to generate the available data. Equation 1 provides a model for traffic counts, but for the reasons described is not sufficiently rich for developing a useful statistical theory. Fortunately, Equation 1 has been used not so much because of its inherent validity, but because more realistic alternatives are lacking. Development of better alternatives requires more detailed consideration of how traffic counts are generated by the underlying trip generation and assignment processes, and these are poorly understood. One could wait until more detailed knowledge is available (presumably from laboratory studies) and then use it to construct models of traffic generation for actual systems, much as well-established principles of mechanics are used in structural design. A more direct strategy would be to formulate plausible models based on available knowledge with the aim of eventually using real-world systems as one's laboratory, and this is the strategy followed here. Thus, rather than attempting to construct a comprehensive stochastic model of traffic generation, just enough probabilistic structure will be added to the standard assumptions concerning traffic generation so that a tractable stochastic process model of traffic counts can be derived. These assumptions lead to a simple but, from an estimation standpoint, intractable stochastic process model. For large traveling populations, however, it is possible to approximate this intractable model by a more tractable linear time-series model, from which OD estimators are readily constructed. The justification for this approximation uses standard mathematics but is long and technical. Because it is described elsewhere (*13,14*), it is not included here. However, familiarity with the structure of the resulting models helps one to see the straightforward way in which OD parameter estimators are constructed and how the properties of the estimators follow from the properties of the models.

We begin by treating the OD flows not as constants but as random variables, and in particular assume that $d_j(t)$, the flow between OD Pair j on Day t is a binomial random variable with parameters n_j and p_j. The outcomes for each OD pair and for each day are assumed to be independent, and the means and variances of the OD flows are thus given by

$$\bar{d}_j = n_j p_j$$

$$\sigma_j^2 = n_j p_j (1 - p_j) \tag{3}$$

The values of the OD parameters $\bar{d}_j$ and σ_j^2 are what we desire to estimate from link count observations. (If the OD flows are treated as constants rather than random variables, we simply deal with the special case where $\sigma_j^2 = 0, j = 1, \ldots, m$.) Now assume we have a total of n links in our network, and let

$\mathbf{x}(t)$ = the n-dimensional vector whose kth element $x_k(t)$ gives the traffic volume on Link k on Day t;

$c_k(\mathbf{x})$ = a differentiable function that gives the cost of traversing Link k as a function of the traffic volume vector $\mathbf{x}$;

$g_k(t)$ = the traveling population's anticipated cost of traversing Link k on Day t;

$p_{jr}[\mathbf{g}(t)]$ = differentiable functions giving the probability that a traveler between OD Pair j uses the rth route connecting j, as a function of the current anticipated cost vector $\mathbf{g}(t)$; and

δ_{jrk} = 1 if Link k lies on Route r connecting j and 0 otherwise.

Then the underlying traffic generation model, Model A, can be expressed in recursive form as follows:

0. Given initial anticipated costs $\mathbf{g}(0)$ and link volumes $\mathbf{x}(0)$, let $t = 1$.
1. Generate the $d_j(t)$ as binomial outcomes with parameters n_j and p_j, $j = 1, \ldots, m$.
2. Let $g_k(t) = (1 - \alpha)g_k(t - 1) + \alpha c_k[\mathbf{x}(t - 1)]$, $0 \le \alpha \le 1$.
3. Generate the route flows $d_{jr}(t)$ as multinomial outcomes with parameters $[d_j(t), p_{jr}(\mathbf{g}(t)]$.
4. Let $x_k(t) = \Sigma_j \delta_{jrk} d_{jr}(t)$, $k = 1, \ldots, n$.
5. Let $t = t + 1$ and go to Step 1.

The process described in Model A is a first-order Markov process. The recursion in Step 2 allows the anticipated cost to be a weighted average of the actual historical link costs, with α controlling the relative importance of recent costs. Adjustment mechanisms of this sort have appeared elsewhere (*11,12*), whereas the idea that route selection is a multinomial process appears to date back to Daganzo (*10*). Model A is easy to simulate, at least for small networks, but the convolutional nature of the link volumes expressed in Step 4 makes derivation of the probability distribution of the link volumes a difficult practical problem. Similar problems have been encountered in statistical mechanics, population biology, mathematical sociology, and so forth (*15,16*). They have often been successfully dealt with by using tractable approximations that become increasingly accurate as the size of the population increases. Intuitively, the approximation of Model A can be based on the fact that, conditional on what has happened on Day $t - 1$, the link volumes on Day t are the result of a large number of independent, individual route choice decisions, so that analogs of the Strong Law of Large Numbers and the Central Limit Theorem ought to apply. This intuition is given formal substance elsewhere (*13,14*). The result is that for large traveling populations, and in the vicinity of a stable user equilibrium assignment, the link count process generated by Model A can be approximated by a vector-valued autoregressive moving average process. In many cases the moving average component of this approximation can be neglected, giving the following first-order, vector autoregressive [VAR(1)] model, Model B:

$$\mathbf{x}(t) - \bar{\mathbf{x}} = \mathbf{F}[\mathbf{x}(t - 1) - \bar{\mathbf{x}}] + \mathbf{a}(t) \tag{4}$$

In Model B, $\bar{\mathbf{x}}$ is a stochastic user equilibrium assignment satisfying

$$\bar{\mathbf{x}} = \sum_{j=1}^{m} \mathrm{d}_j \mathbf{q}_j[\mathbf{c}(\bar{\mathbf{x}})] \tag{5}$$

$\mathbf{F}$ is a weighted Jacobian matrix of the right hand side of Equation 5:

$$\mathbf{F} = \alpha \sum_{j=1}^{m} \bar{d}_j \left[\frac{\partial \mathbf{q}_j[\mathbf{c}(\mathbf{x})]}{\partial \mathbf{x}} \right]_{\bar{\mathbf{x}}} \tag{6}$$

The $\mathbf{a}(t)$ are independent, identically distributed normal random vectors with mean vector equal to $\mathbf{0}$ and covariance matrix $\mathbf{Q}$ given by

$$\mathbf{Q} = \sum_{j=1}^{m} (\bar{d}_j[\bar{\mathbf{Q}}_j - \mathbf{q}_j\mathbf{q}_j^T] + \sigma_j^2 \mathbf{q}_j \mathbf{q}_j^T) \tag{7}$$

As defined earlier, the $\mathbf{q}_j$ appearing in Equation 7 denote the vectors of link use probabilities, whose elements are given by

$$q_{jk} = \sum_r (\delta_{jrk}) p_{jr}[\mathbf{c}(\bar{\mathbf{x}})] \quad (8)$$

The matrices $\mathbf{Q}_j$ appearing in Equation 7 are defined so that the *kl*th element $Q_{j,kl}$ gives the probability that a trip between OD Pair *j* uses both Links *k* and *l*:

$$\bar{Q}_{j,kl} = \sum_r (\delta_{jrk})(\delta_{jrl}) p_{jr}[\mathbf{c}(\bar{\mathbf{x}})] \quad (9)$$

The advantages of replacing Model A with Model B derive from the fact that the VAR(1) process is arguably the most tractable and well understood of Markov processes. VAR(1) processes are parameterized by the quantities $\bar{\mathbf{x}}$, **F**, and **Q** and, for a set of observations $\mathbf{x}(t)$, $t = 1, \ldots, N$, the ML estimates of $\bar{\mathbf{x}}$, **F**, and **Q** are easily calculated using least-squares methods (*17*). We are not interested in all VAR(1) processes, however, but only those that can result from underlying traffic assignment processes, and thus must restrict our attention to VAR(1) processes whose parameters satisfy Equations 5 through 7. Because Equation 5 expresses the mean vector **x** as an implicit function of the OD flow means, the question arises as to whether the relationship between the VAR parameters $\bar{\mathbf{x}}$, **F**, and **Q** and the OD parameters $\bar{d}_j$ and σ_j^2 is well defined. If the VAR parameters are differentiable functions of the OD parameters, we can consider taking the derivative of the VAR likelihood function with respect to the OD parameters and solving for the OD parameter values making these derivatives equal to zero, producing ML estimates. If no such functions exist, the estimation task is much more difficult. Fortunately, it can be shown (*13*, chapter 4) that if (*a*) the link cost functions $c_k(.)$ and the route choice probability functions $p_{jr}(.)$ are continuously differentiable and (*b*) the rank of the matrix $\mathbf{I} - (1/\alpha)\mathbf{F}$ is *n* when the d_j are the true values of the OD parameters, then at least in a neighborhood of the true values of $\bar{d}_j$, σ_j^2, the VAR parameters are continuously differentiable functions of the OD parameters. Although an explicit representation of this function cannot be given, implicit differentiation can be used to obtain the necessary derivatives. It can also be readily verified that when the link cost functions have the BPR form, while the route choice probabilities are given by the multinomial logit formula, conditions (*a*) and (*b*) are satisfied (*13*). This well-behaved relationship between the OD parameters and the VAR parameters is the basis of OD estimation.

Essentially Model B claims that if one had time-series data consisting of, say, morning peak-period traffic volumes for all links in a network collected over several months, the data should be accurately modeled as a VAR(1) process. The mean of the data set should be equal to the SUE assignment and the regression and covariance matrices, respectively, should be equal to the **F** and **Q** given in Equations 6 and 7. If traffic counts are only available from a subset of the network links, the time-series model describing these partial counts will no longer necessarily be VAR(1) but will still be a stationary linear model whose parameters are at least in theory computable from $\bar{\mathbf{x}}$, **F**, and **Q** (*18*). Thus Model B makes some strong claims concerning the nature of traffic count data that are, at least in principle, testable from data collected by automatic surveillance and control systems. A limited amount of published empirical work suggests that Model B is not implausible. For instance, Cascetta (*12*) provides evidence that an SUE assignment is as reasonable a forecast of traffic volumes as the more standard DUE assignment, whereas in earlier work (*19,20*) we found that once seasonal trends have been accounted for, stationary linear models provide reasonable representations of freeway volume counts. However, more adequate empirical testing of traffic assignment methods requires accurate estimation of OD parameters, because otherwise one cannot distinguish between poor fit caused by a poor model and poor fit caused by ignorance of the actual OD patterns (*21,22*). By treating OD estimation as a problem in system identification, these aspects—model estimation and model validation—can be integrated.

ESTIMATION METHODS

Before turning to the construction of estimators for the OD parameters, it is necessary to resolve two technical issues. The first concerns the placement of the traffic counters generating the available data. In most practical cases, not all network links contain traffic counters, although the advent of automatic surveillance and control systems makes the availability of traffic counts more widespread. The estimation theory for VAR(1) processes, however, assumes that, subject to a limitation to be discussed shortly, observations from all links are available. This assumption allows the estimation theory for Markov processes developed by Billingsley (*23*) to be applied to this problem, resulting in relatively straightforward estimation methods whose asymptotic properties are readily established. On the other hand, when we assume that only a subset of the links have counters, it is well known from systems theory that the stochastic process describing the observations will no longer be Markov, even though the underlying process is. Computation of the likelihood function for this case requires employment of the Kalman filter, and establishing the asymptotic properties of the resulting estimators is an unsolved problem. Thus in this paper we concentrate on the case where a full set of traffic counts is available, saving the partial count case for later research.

Having made the case for a full set of traffic counts, we next note that the link flows on a network contain linear dependencies because of conservation of flow requirements. This means that the link flow vector **x** can be partitioned into $\mathbf{x}_1$ and $\mathbf{x}_2$, with a linear relationship $\mathbf{x}_2 = B\mathbf{x}_1$ existing between $\mathbf{x}_1$ and $\mathbf{x}_2$. The linear dependence causes the covariance matrix **Q** to be singular, which in turn causes both practical and theoretical problems in OD estimation (*13*). This is easily solved by working with the independent set of link counts $\mathbf{x}_1$ and deleting the appropriate rows and columns from **Q**. Computation of the corresponding matrix **F** is readily done using the chain rule for vector-values functions. Thus from here on it is assumed that we have available a full set of linearly independent counts, and the quantities $\bar{\mathbf{x}}$, **F**, and **Q** are computed for this linearly independent set.

With these cautions noted, let

$$\hat{\mathbf{x}}(t) = \bar{\mathbf{x}} + \mathbf{F}[\mathbf{x}(t-1) - \bar{\mathbf{x}}] \quad (10)$$

denote the predicted value for $\mathbf{x}(t)$. ML estimates for the model parameters are then obtained by minimizing the scaled log-likelihood function

$$L = \log|\mathbf{Q}| + (1/N) \sum_{t=1}^{N} [\mathbf{x}(t) - \hat{\mathbf{x}}(t)]^T \mathbf{Q}^{-1} [\mathbf{x}(t) - \hat{\mathbf{x}}(t)] \quad (11)$$

with respect to the parameters of interest. For the VAR parameters $\bar{\mathbf{x}}$, $\mathbf{F}$, and $\mathbf{Q}$, the ML estimates have a closed-form solution, which can be readily computed using standard regression methods. It is also well known that these estimates are consistent, asymptotically normally distributed, and asymptotically efficient (*17,23*). For the OD parameters $\bar{d}_j$ and σ_j^2, one can still take derivatives of Equation 11 using the chain rule and implicit differentiation, set the derivatives equal to zero, and then in principle solve the resulting equations for the ML estimates. The likelihood equations will in this case not admit a closed-form solution, so that numerical solution methods must be employed. However, given some technical conditions, it can still be established that Billingsley's results on ML estimation for Markov chains apply to this situation and that the solutions to the likelihood equations are consistent, asymptotically normally distributed, and asymptotically efficient (*13*). Thus the classical results concerning ML estimators apply to the ML estimates of the OD parameters, though with some numerical difficulties in actually computing these estimates.

Alternatively, Equations 5 through 7 defining the VAR parameters can be written in a vectorized form:

$$\bar{\mathbf{x}} = \sum_{j=1}^{m} \bar{d}_j \mathbf{q}_j$$

$$v(\mathbf{F}) = \alpha \sum_{j=1}^{m} \bar{d}_j v\left(\left[\frac{\partial \mathbf{q}_j}{\partial \mathbf{x}}\right]\right)$$

$$vh(\mathbf{Q}) = \sum_{j=1}^{m} [\bar{d}_j vh(\tilde{\mathbf{Q}}_j - \mathbf{q}_j\mathbf{q}_j^T) + \sigma_j^2 vh(\mathbf{q}_j\mathbf{q}_j^T)] \qquad (12)$$

where $v(.)$ denotes the vector operator (the stacking of the columns of a matrix on top of each other) and $vh(.)$ denotes the vector half-operator (the stacking of the columns of the lower diagonal of a symmetric matrix on top of each other). Given estimates of $\bar{\mathbf{x}}$, $\mathbf{F}$, and $\mathbf{Q}$ and knowledge of the parameter α, Equation 12 defines a system of linear equations with the OD parameters as the unknowns. When the number of individual paramaters in $\bar{\mathbf{x}}$, $\mathbf{F}$, and $\mathbf{Q}$ exceeds the number of OD parameters, these equations can be solved using the Moore-Penrose pseudoinverse (or equivalently, ordinary least-squares) to obtain estimates of the OD parameters. Because the ML estimates for the VAR parameters are also the MOM estimators, the second approach gives a MOM estimator of the OD parameters. Under the conditions that the link cost and route choice functions are continuously differentiable, the function relating the MOM OD estimates to the MOM VAR estimates will also be continuously differentiable, so that the consistency and asymptotic normality shown by the VAR estimators will be inherited by the MOM OD estimators (*24*, p. 388). The asymptotic efficiency of the VAR estimators, however, is not guaranteed to transfer to the MOM OD estimators.

MONTE CARLO EVALUATION OF ESTIMATORS

Both the ML and the MOM estimators assume the VAR(1) Model B is an accurate description of the process generating the link volumes. Model B is in turn derived as a large-population approximation to the Markov process described in Model A. With infinite populations and infinite amounts of data, the asymptotic theory characterizes the properties of these estimators. However, for a given finite population and finite data sequences, the asymptotic properties may not be operative, and simulation studies are required to verify the practical usefulness of the asymptotic results. We have already described some asymptotic properties of the ML and MOM estimators based on large-population and large-sample approximations. To gain some appreciation of how the ML and MOM estimators would perform on finite data sets generated by a Model A process, we conducted the following simulation study. A FORTRAN program, SIMFLO, was written that could simulate the Model A process for networks small enough that enumeration of the network routes was feasible. (This was necessary so that the multinomial simulation described in Step 3 of Model A could be performed.) SIMFLO assumes that the link costs are given by the BPR functions with the form

$$c_k(x_k) = a_k\left[1 + b_k\left(\frac{x_k}{f_k}\right)^4\right] \qquad (13)$$

The route choice probabilities are given by the logit formula

$$p_{jr} = \frac{e^{-\Theta c_{jr}}}{\sum_s e^{-\Theta c_{js}}} \qquad (14)$$

Two FORTRAN estimation programs were also written. Program MARKOD computes numerical ML estimates of the OD parameters using the quasi-Newton routine E04JBF, obtained from the NAG subroutine library (*25*). Program MOMOD computes MOM estimates using standard least-squares procedures. Although we conducted tests with several hypothetical networks, space limitations permit the description of only one simulation experiment. However, the results of all experiments were consistent with the results that we now present. Figure 1 shows a 14-link network with three origin nodes (O1, O2, and O3) and three destination nodes (D1, D2, and D3), for a total of nine OD pairs. If one imagines collecting weekday peak-period volume data for a network, a time series of length 150 would correspond to 7 to 8 months of observations, and this was considered a reasonable upper bound on the duration over which OD patterns might be taken as being constant. From such a time series, one could obtain one set of estimated OD parameters, so to sample the statistical properties of the OD estimators, 50 separate time series, each of length 150, were generated using SIMFLO. The 50 time-series were then input to MARKOD and MOMOD to obtain a sample, of size 50, of the ML and MOM estimates. Although in principle it is possible to treat the weighting parameter α and the logit parameter Θ as unknown (MARKOD permits this option), the relationship between $\bar{\mathbf{x}}$, $\mathbf{F}$, and $\mathbf{Q}$ and the unknown parameters shown in Equation 12 is no longer linear, and computation of the MOM estimators will be more difficult. Because the primary goal of this study is assessment of the estimates of the OD parameters $\bar{d}_j$ and σ_j^2, α and Θ will be treated as known a priori. Estimation using MARKOD was done on the Cyber 180/855 computer at the University of Washington, and estimation using

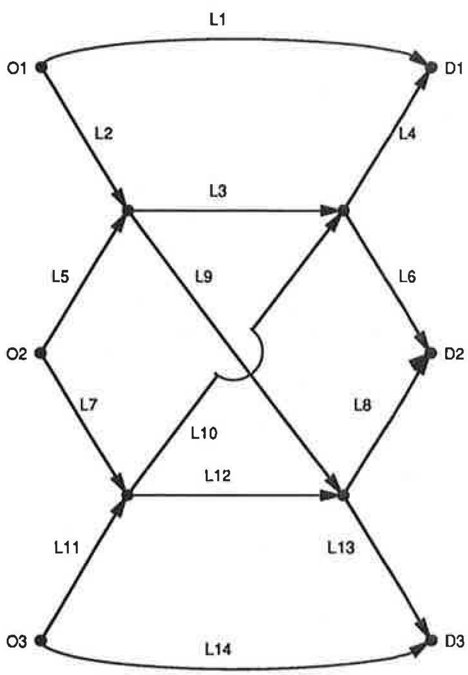

FIGURE 1 Diagram of 14-link network used in simulation study.

MOMOD was done on the Cyber 170/845 computer at the University of Minnesota.

Table 1 gives statistical information concerning the properties of the ML estimator. The column labeled "True" gives each OD parameter's true value; the column labeled "Mean" gives the average, across the 50 simulations, of the estimated values produced by the ML method; and the column labeled "S.D." gives the standard deviation of these estimates about their average. The column labeled "t" gives a computed *t*-statistic testing the hypothesis that the average estimate equals its corresponding true value, which is a test for bias in the estimator. The column labeled "r" gives the normal-score correlation test for normality (*26*), which tests the asymptotic normality of the estimates. Table 2 gives similar information for the MOM estimator. One asterisk indicates statistical significance of a test at the .05 level, whereas two asterisks indicate statistical significance at the .01 level.

From the "r" columns, it can be seen that both the ML and the MOM methods produce estimates that tend to be normally distributed, indicating that a data series of length 150 appears sufficient to obtain asymptotic normality. Looking at the "t" columns, we see an undeniable tendency for both estimators to generate biased estimates, indicating that with data sequences of length 150 there is still some gap, on the average, between the estimate and the true value. Inspection of the "True" and "Mean" columns in both tables indicates that for the ML estimator the biases tend to be less than 10 percent of the true values, whereas for MOM the biases are somewhat worse. A comparison of the "S.D." columns in Tables 1 and 2 indicates that the ML estimates have markedly less variance than the MOM estimates, a result consistent with the asymptotic efficiency property of the ML estimator. The variability of the MOM estimates of the OD variances is in fact so large that it calls into question the practical value of such estimates.

TABLE 1 ESTIMATION SUMMARY FOR ML METHOD

Parameter	True	Mean	S.D.	t	r
d_1	450.	449.2	2.1	2.76**	.991
d_2	300.	300.1	23.5	0.03	.985
d_3	301.	301.	23.9	0.00	.989
d_4	352.	362.3	25.4	-2.86**	.996
d_5	490.	471.7	8.4	15.5**	.995
d_6	150.	157.8	24.5	-2.23*	.988
d_7	100.	90.9	24.9	2.60*	.994
d_8	200.	217.1	25.1	-4.89**	.993
d_9	700.	691.9	3.3	17.23**	.978
σ_1^2	45.	47.0	10.6	-1.33	.994
σ_2^2	75.	75.4	12.8	-0.22	.994
σ_3^2	42.1	44.3	9.7	-1.59	.992
σ_4^2	42.2	40.1	9.9	1.07	.989
σ_5^2	147.	141.3	22.4	1.78	.991
σ_6^2	37.5	38.2	10.9	-0.47	.994
σ_7^2	50.	49.6	10.8	0.27	.990
σ_8^2	40.	39.6	12.0	0.24	.973*
σ_9^2	210.	202.0	27.5	2.05*	.988

CPU seconds/estimation = 2504
(Cyber 180/855)

TABLE 2 ESTIMATION SUMMARY FOR MOM

Parameter	True	Mean	S.D.	t	r
d_1	450.	449.99	25.87	0.003	.987
d_2	300.	286.07	54.98	1.790	.985
d_3	301.	315.54	45.15	-2.270*	.973
d_4	352.	341.97	61.27	1.158	.987
d_5	490.	504.80	99.10	-1.057	.996
d_6	150.	128.00	70.58	2.204*	.986
d_7	100.	104.95	45.61	-0.760	.990
d_8	200.	189.44	63.75	1.170	.993
d_9	700.	716.89	60.03	-1.989	.994
σ_1^2	45.	28.31	34.40	3.427**	.993
σ_2^2	75.	67.63	32.29	1.612	.986
σ_3^2	42.1	34.41	23.83	2.280*	.988
σ_4^2	42.2	30.80	45.22	1.781	.988
σ_5^2	147.	126.23	31.08	4.720**	.990
σ_6^2	37.5	34.47	45.79	0.467	.993
σ_7^2	50.	45.38	23.02	1.417	.995
σ_8^2	40.	39.77	26.88	0.060	.984
σ_9^2	210.	187.56	52.80	3.000**	.988

CPU seconds/estimation = 0.89
(Cyber 170/845)

Finally, at the bottom of Tables 1 and 2 we display the average CPU time required for the two estimation methods. Whereas the MOM method is very efficient from a numerical standpoint, the ML procedure proved to be computationally burdensome, suggesting that ML estimation for networks of realistic size may prove difficult even on a supercomputer.

SUMMARY AND CONCLUSION

We have considered here the problem of estimating the parameters describing the OD demand on a traffic network from time-series observations of the network's link volumes. The existence of such methods, coupled with the availability of automatic traffic-counting technology, holds the promise of being able to obtain timely, inexpensive estimates of existing travel demand and of being able to detect and track changes in demand over time or in response to transportation initiatives. Because the dynamics of the processes generating traffic cannot be neglected, we argue that OD estimation is more properly seen as statistical inference on stochastic processes rather than an example of classical statistical procedures. After developing a tractable stochastic process model that is a plausible approximation of the traffic count process, we use the model to develop both ML and MOM estimators of the process's OD parameters. Both estimators have desirable consistency and asymptotic normality properties, but only the ML estimator can claim asymptotic statistical efficiency. Simulation studies indicate that the ML estimator is superior from a statistical standpoint, but the numerical labor needed to compute ML estimates makes application to networks of realistic size problematical. On the other hand, the MOM estimates, though clearly inferior statistically, are numerically efficient enough to be considered for networks of realistic size. At this point, the trade-off between statistical and numerical efficiency is unforgiving. Thus, although it is possible to begin developing a statistical theory for making inferences about OD parameters from traffic count data, numerical issues stand between the theory and its general usefulness.

The obvious research need then is to develop methods that preserve the statistical efficiency of the ML method but reduce its computational burden. The minimization procedure implemented in Program MARKOD can be viewed as a variant of the reduced-gradient algorithm, and it has been recently reported that reduced-gradient algorithms tend to be numerically slower than competitors such as sequential quadratic programming (SQP) (*27*). A promising line of investigation would be to reformulate the ML estimation problem as an SQP and use a state-of-the-art SQP code to solve it. We are currently pursuing this course. Another need is to develop methods that do not require a full set of traffic counts. Numerical computation of ML estimators for this case is relatively straightforward (although some limited numerical experiments indicate that the CPU demands tend to be greater than those for the full-count case). What is problematic is demonstrating that these estimators have desirable asymptotic properties, such as consistency and asymptotic normality. Because the numerical difficulties shown by the full-count case also tend to be shown by the partial count case, the main obstacle to practical use is the lack of a numerically efficient computation method.

ACKNOWLEDGMENTS

This research was supported in part by the Valle Foundation through a Scandinavian Study Scholarship at the Royal Institute of Technology, Stockholm, Sweden awarded to Gary Davis and in part by the National Science Foundation through a grant awarded to Nancy Nihan. The authors would like to thank Mr. Konstantinos Koutsoukos for programming and running some of the simulation results.

REFERENCES

1. S. Atkins. Transportation Planning Models—What the Papers Say. *Traffic Engineering and Control*, Vol. 27, 1986, pp. 460–467.
2. E. Cascetta and S. Nguyen. A Unified Framework for Estimating or Updating Origin/Destination Matrices from Traffic Counts. *Transportation Research*, Vol. 22B, 1988, pp. 437–455.
3. M. Maher. Inferences on Trip Matrices from Observations on Link Volumes: A Bayesian Statistical Approach. *Transportation Research*, Vol. 17B, 1983, pp. 435–447.
4. M. Bell. The Estimation of an Origin-Destination Matrix from Traffic Counts. *Transportation Science*, Vol. 17, 1983, pp. 198–217.
5. H. Spiess. A Maximum Likelihood Model for Estimating Origin-Destination Matrices. *Transportation Research*, Vol. 21B, 1987, pp. 395–412.
6. S. McNeil and C. Hendrickson. A Regression Formulation of the Matrix Estimation Problem. *Transportation Science*, Vol. 19, 1985, pp. 278–292.
7. E. Cascetta. Estimation of Trip Matrices from Traffic Counts and Survey Data: A Generalized Least-Squares Estimator. *Transportation Research*, Vol. 18B, 1984, pp. 289–299.
8. M. Cremer and H. Keller. Dynamic Identification of Flows from Traffic Counts at Complex Intersections. In *Eighth International Symposium on Transportation and Traffic Theory* (V. F. Hurdle et al., eds.), University of Toronto Press, Toronto, 1983, pp. 121–142.
9. N. L. Nihan and G. A. Davis. Application of Prediction-Error Minimization and Maximum Likelihood To Estimate Intersection O-D Matrices from Traffic Counts. *Transportation Science*, Vol. 23, 1989, pp. 77–90.
10. C. F. Daganzo. Some Statistical Problems in Connection with Traffic Assignment. *Transportation Research*, Vol. 11, 1977, pp. 385–389.
11. J. L. Horowitz. The Stability of Stochastic Equilibria in a Two-Link Transportation Network. *Transportation Research*, Vol. 18B, 1984, pp. 13–28.
12. E. Cascetta. A Stochastic Process Approach to the Analysis of Temporal Dynamics in Transportation Networks. *Transportation Research*, Vol. 23B, 1989, pp. 1–17.
13. G. A. Davis. *A Stochastic, Dynamic Model of Traffic Generation and its Application to the Maximum Likelihood Estimation of Origin-Destination Parameters*. Doctoral dissertation. University of Washington, Seattle, 1989.
14. G. A. Davis and N. L. Nihan. A Dynamic, Stochastic Approach to Traffic Assignment. Submitted to *Operations Research*, 1990.
15. Karmeshu and R. K. Pathria. A Stochastic Model for Highway Traffic. *Transportation Research*, Vol. 15B, 1981, pp. 285–294.
16. T. G. Kurtz. Strong Approximation Theorems for Density Dependent Markov Chains. *Stochastic Processes and Their Applications*, Vol. 6, 1978, pp. 223–240.
17. T. W. Anderson. *The Statistical Analysis of Time Series*. John Wiley and Sons, New York, 1971.
18. P. E. Caines. *Linear Stochastic Systems*. John Wiley and Sons, New York, 1988.
19. N. L. Nihan and G. A. Davis. Estimating the Impacts of Ramp-Control Programs. In *Transportation Research Record 957*, TRB, National Research Council, Washington, D.C., 1984, pp. 31–32.
20. G. A. Davis and N. L. Nihan. Using Time-Series Designs to Estimate Changes in Freeway Level of Service, Despite Missing Data. *Transportation Research*, Vol. 18A, 1984, pp. 431–438.
21. M. Florian. An Introduction to Network Models Used in Transportation Planning. In *Transportation Planning Models* (M. Florian, ed.), Elsevier, Amsterdam, 1984, pp. 137–152.
22. Y. J. Stephanedes, I. Z. Argiopoulos, and P. Michalopoulos. On-line Traffic Assignment for Optimal Freeway Corridor Control. *ASCE Journal of Transportation Engineering*, Vol. 116, 1990, pp. 744–755.
23. P. Billingsley. *Statistical Inference for Markov Processes*. University of Chicago Press, Chicago, 1961.
24. C. R. Rao. *Linear Statistical Inference and Its Applications*, 2nd ed. John Wiley and Sons, New York, 1973.
25. *NAG Fortran Library Manual*, Mark 9, NAG (USA), 1976.
26. J. J. Filliben. The Probability Plot Correlation Coefficient Test for Normality. *Technometrics*, Vol. 17, 1975, p. 111.
27. P. Gill, W. Murray, M. Saunders, and M. Wright. Constrained Nonlinear Programming. In *Optimization*, Vol. 1 (G. L. Nemhauser et al., eds.), Elsevier Science Publishers, Amsterdam, 1989.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.