Pergamon

Transpn Res.-B. Vol. 30, No. 4, pp. 277–290, 1996

0191-2615/96 $15.00+0.00

0191-2615(95)00011-9

ESTIMATION OF TIME-VARYING ORIGIN–DESTINATION DISTRIBUTIONS WITH DYNAMIC SCREENLINE FLOWS

JIFENG WU* and GANG-LEN CHANG
Department of Civil Engineering, University of Maryland, College Park, Maryland 20742, U.S.A.

(Received 15 February 1994; in revised form 23 August 1995)

Abstract—A variety of methods for dynamic origin–destination (O–D) estimation is available in the transportation literature. Among those, the research direction pursued by Bell (1991)(*Transpn Res.* **25B**, 115—125) and others in the same category offers the best promise for use in practice as it employs only link flow information from existing surveillance systems. Following the same direction, this paper presents an innovative method for estimating the dynamic network O–D matrices with time series of link and screenline flows. The proposed method takes full advantage of available link flow information, and considerably increases the observability of the dynamic interrelations between network O–D patterns and the resulting link flow distributions. Hence, one can estimate the dynamic O–D matrices with much fewer model parameters and less execution burden. With properly selected screenlines and efficient computing algorithms, the proposed model also offers the potential for real-time applications. Copyright © 1996 Elsevier Science Ltd

1. INTRODUCTION

Over the past two decades, the development of an accurate and efficient method for network O–D estimation from link traffic counts has received increasing attention among transportation researchers. Most of such studies in the literature were developed along the following two lines: (1) static O–D estimation, grounded on presumably available prior O–D and assignment model for planning applications; (2) time-dependent O–D analysis, based on time-series of traffic counts for mainly control and operation needs. A comprehensive review of static O–D related research can be found in the articles by Nguyen (1984) and Cascetta and Nguyen (1988).

With respect to the research on dynamic O–D estimation, Cremer and Keller (1981), Cremer and Keller (1984), and Cremer (1983), in a series of pioneering work for identifying intersection flow movements, have first employed relations constructed through time series of traffic counts, and converted the underdetermined static model to an overdetermined dynamic formulation. In subsequent studies, they have worked in parallel with others (Cremer & Keller, 1987; Nihan & Davis, 1987) in developing various recursive and non-recursive algorithms for tracking the time-varying O–D evolution. With such dynamic relations the estimation algorithm needs neither the prior O–D information nor a reliable dynamic assignment model.

Some later developments along the same line, but for use only at isolated intersections or a short freeway segment, can be found in Keller and Ploss (1987), Nihan and Davis (1989), Kessaci *et al.* (1990), Nihan and Hamed (1992) and Yu and Davis (1994). Recently, Bell (1991) has made an extension of the dynamic O–D estimation methods by Cremer and Keller (1987) and Nihan and Davis (1987) to general networks. His first extended method employs a concept similar to Robertson's platoon dispersion model (Robertson, 1969) with a platoon dispersion parameter for estimating the dynamic O–D in small networks, and his second method assumes freely-distributed travel time for computing the effect of travel time variability on network O–D estimation. Both methods

*Present Address: Traffic Management Lab, Viggen Corporation, 6862 Elm Street, Suite 700, McLean, VA 22101, U.S.A.

are, however, effective only for small networks where the travel time for each O–D is shorter than the unit time interval selected for performing estimation.

Another category of studies for attacking the complex dynamic O–D issues focuses on extending the modeling concept from static O–D estimation. For instance, Okutani (1987) formulated a Kalman filtering model for dynamic O–D prediction, using an autoregressive model to capture the temporal interdependency among O–D flows and an incidence matrix for the interrelations between O–D flows and link traffic counts. Its key underlying assumptions are the existence of an accurate dynamic traffic assignment model (DTA) and the available prior time-dependent O–D data for time series model construction. Along the same line, Ashok and Ben-Akiva (1993) have revised Okutani's work with the state variables representing the deviations of O–D flows from prior estimates based on historical data. More recently, Van Aerde *et al.* (1993) have reported the application of a similar logic in their QUEENSOD model based on the identical assumptions. Working in parallel, Cascetta *et al.* (1993) during the same period have proposed two dynamic estimators (i.e. simultaneous estimators and sequential estimators) for network O–D estimation which are also a generalization of those static methods summarized in Cascetta and Nguyen (1988). A comprehensive review of all dynamic O–D estimation literature is available in the article by Wu and Chang (1995).

Note that most dynamic models developed along this direction appear to be able to effectively address the time-dependent O–D estimation issues in urban networks, if their required prior data and assignment information are available. Unfortunately, the acquisition of historical dynamic O–D data and the development of a reliable Dynamic Traffic Assignment (DTA) model are both quite challenging tasks, and may not be solved in the near future. Considering both the data availability and underlying assumptions, it is fair to conclude that the recent research direction by Bell (1991) and others having similar concepts, are relatively promising, as such models for dynamic O–D estimation rely only on direct flow measurements with existing surveillance systems. Further research along the same line should focus on the identification of additional dynamic relations with existing data, a better formulation of the time-dependent interactions between O–D patterns and link flows, and the incorporation of travel time variability, due to behavior differences, on the dynamic O–D distribution.

This paper is to report our research efforts and results in response to such needs. It consists of the following sections. The model formulation for dynamic O–D distributions with the proposed screenline concept is presented in the next section, followed by a discussion of available estimation algorithms in Section 3, and some guidelines for screenline selection in Section 4. A complementary approach for estimation of time-varying trip times is introduced in Section 5. An illustrative example and concluding comments are reported in Section 6 and Section 7, respectively.

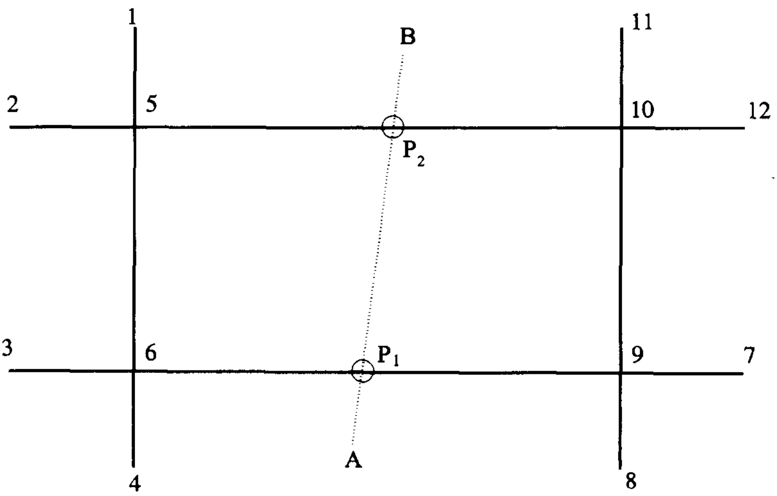

Fig. 1. A graphical illustration of complementary screenlines in a simple network.

2. MODEL FORMULATION

Consider an urban road network, as shown in Fig. 1, in which trip origins and destinations are assumed to be at network nodes. A screenline is defined as a hypothetical directed line or curve that intersects with a set of links, and divides the network into a left-hand-side (LHS) and a right-hand-side (RHS) subnetwork. The screenline flow is defined as the sum of flows in those links intersecting with the screenline from the LHS to the RHS. Note that each identified screenline may be coupled by a screenline with opposing flows, if all network links crossed by the screenline have two directional flows. For instance, the dotted line in Fig. 1 represents a pair of two screenlines: AB and BA. Screenline AB intersects with links {5–10} and {6–9}, and its flow is the sum of flows at P_2 and P_1 on these two links. In contrast, the flow rate for screenline BA is the sum of flows at P_1 and P_2 but on links {9–6} and {10–5}.

The notation for model presentation hereafter is summarized below:

N: set of the network nodes
A: set of the network arcs
L: set of screenlines selected for the network
l: screenline index, $l \in L$
l': index of the screenline with opposing flows for screenline l
O_l: set of origin nodes at the LHS of screenline l
D_l: set of destination nodes at the RHS of screenline l
SL_l: set of intersections on screenline l and its crossing links
T: length of a unit time interval
k: time interval index
$V_l(k)$: vehicle trips through screenline l during interval k
$q_i(k)$: the number of vehicle trips generated from origin node i during interval k
$x_{ij}(k)$: O–D vehicle trips from origin node i to destination node j during interval k
$b_{ij}(k)$: proportion of demand $q_i(k)$ heading to destination node j.
$t_{ilj}(k)$: average travel time for trips from $i \in O_l$ to $j \in D_l$ to reach screenline l during interval k
$t_{ilj}^-(k)$: the maximum integer less than or equal to $t_{ilj}(k)/T$
$t_{ilj}^+(k)$: $= t_{ilj}^-(k) + 1$
$T_{ipj}(k)$: the shortest travel time from $i \in O_l$ to $j \in D_l$ through $p \in SL_l$ during interval k
σ_{ilj}: the integer portion of the average travel time deviation from its mean value $t_{ilj}(k)$ measured by the number of time intervals.

As in most studies for dynamic O–D estimation (e.g. Bell, 1991), our proposed model concentrates on estimating the O–D split proportions $\{b_{ij}\}$, which are assumed to remain at the same level over a certain number of consecutive intervals. The time interval index k is omitted hereafter for convenience of presentation. The interrelations between the time-varying O–D flows, $\{x_{ij}(k)\}$, the total generation $\{q_i(k)\}$, and its split proportions $\{b_{ij}\}$ can be stated as:

$$x_{ij}(k) = q_i(k) b_{ij}, \quad \forall i,j. \tag{1}$$

Naturally, the variables, $\{b_{ij}\}$, are subject to the following two constraints:

$$\Sigma_j b_{ij} = 1 \quad \forall i \tag{2}$$

$$b_{ij} \geq 0 \quad \forall i,j. \tag{3}$$

Consider an O–D pair (i, j) crossing a screenline l, where $i \in O_l$, $j \in D_l$. Certainly, without any route choice information, the travel time, $t_{ilj}(k)$, and the crossing locations ($p \in SL_l$) used by such O–D trips are unknown. However, as the travel time from node i to any $p \in SL_l$ without incidents should lie within a certain range, one may assume that it follows a normal distribution with its mean arriving time to the screenline during time interval k being denoted as $t_{ilj}(k)$. With such an assumption, it leads to the result that most trips from origin i, arriving at the screenline during interval k, should depart within the interval $k - t_{ilj}^+(k)$ to $k - t_{ilj}^-(k)$. Thus, given a sufficiently long time interval, T, the screenline flows, $V_l(k)$, should contain O–D trips mainly from

$$x_{ij}[k - t^+_{ilj}(k)] \text{and} x_{ij}[k - t^-_{ilj}(k)], \text{for all} i \in O_l, j \in D_l.$$

This relation may best be presented with the following expression:

$$V_l(k) \approx \sum_{i \in O_l} \sum_{j \in D_l} \frac{1}{2} [x_{ij}(k - t^+_{ilj}(k)) + x_{ij}(k - t^-_{ilj}(k))] \quad l \in L. \tag{4}$$

Note that the screenline flow, $V_l(k)$, may contain some fraction of those trips departing beyond the interval $k - t_{ilj}{}^+(k)$ and $k - t_{ilj}{}^-(k)$. By the same token, some vehicles from $x_{ij}[k - t_{ilj}{}^+(k)]$ and $x_{ij}[k - t_{ilj}{}^-(k)]$ may also contribute to the screenline volumes $V_l(t)$, for $t < k$ or $t > k$. To account for such facts in Eqn 4, one may incorporate time lag factor either in the input side (the O–D flows) only, or in both the input and output (the exit link or screenline flows) sides. Mathematical expressions for each of these two methods are presented below.

2.1. *Treat time lag factors in the input side*

Let $$\{k - t^+_{ilj}(k) - m, k - t^-_{ilj}(k) + m \;|m = 0, 1, ..., \sigma_{ilj}\}$$

represent the set of time intervals encompassing the distribution of departure times of all vehicles arriving at the screenline during interval k; where σ_{ilj} is the integer portion of average travel time deviation of $t_{ilj}(k)$. By denoting $\alpha_{lmij}{}^-$ and $\alpha_{lmij}{}^+$, respectively, as the fraction of flows from $x_{ij}(k - t_{ilj}{}^+(k) - m)$ and $x_{ij}(k - t_{ilj}{}^-(k) + m)$, contributing to $V_l(k)$, one can describe the interrelations between screenline and O–D flows as follows:

$$V_l(k) = \sum_{i \in O_l} \sum_{j \in D_l} \sum_{m=0}^{\sigma_{ilj}} [x_{ij}(k - t^+_{ilj}(k) - m)\alpha^-_{lmij} + x_{ij}(k - t^-_{ilj}(k) + m)\alpha^+_{lmij}] \tag{5}$$

$$= \sum_{i \in O_l} \sum_{j \in D_l} \sum_{m=0}^{\sigma_{ilj}} [q_i(k - t^+_{ilj}(k) - m)\alpha^-_{lmij} + q_{ij}(k - t^+_{ilj}(k) + m)\alpha^+_{lmij}] b_{ij}.$$

This formulation is actually a generalization of our previous work on a single freeway section (Chang & Wu, 1994) which turns out to be a nonlinear model. Although the Extended Kalman filter approach may be employed for recursive estimation of both the O–D parameters, $\{b_{ij}\}$, and the additional parameters, $\alpha_{lmij}{}^-$ and $\alpha_{lmij}{}^+$, this approach may not be efficient in large network applications due to the required parameter size and its nonlinear nature. Hence, it will not be pursued in this study.

2.2. *Treat time lag factors in both the input and output sides*

With respect to a given time interval k, Eqn 5 presents only those input flows contributing to the output flow, $V_l(k)$. Conceivably, vehicles from $x_{ij}[k - t_{ilj}{}^+(k)]$ and $x_{ij}[k - t_{ilj}{}^-(k)]$ may also contribute to screenline volumes $V_l(t)$, for $t < k$ and $t > k$. To express the total screenline volume, resulting from vehicles departing during these two time intervals, Eqn 4 can be stated as:

$$\sum_{i \in O_l} \sum_{j \in D_l} \sum_{m=0}^{\sigma_{ilj}} [x_{ij}(k - t^+_{ilj}(k) - m) + x_{ij}(k - t^-_{ilj}(k) + m)] \tag{6}$$

$$= V_l(k) + \sum_{m=1}^{M} [\beta^-_m V_l(k - m) + \beta^+_m V_l(k + m)] \quad l \in L$$

where

$$M = \max\{\sigma_{ilj} \mid i \in O_l, j \in D_l, l \in L\} + 1$$

and β_m^- and β_m^+ are additional parameters, representing weighting coefficients, and are independent of the selected screenline l. Obviously, these weighting parameters are subject to the following constraints

$$0 \le \beta_m^-, \beta_m^+ \le 1 \qquad m = 1, 2, ..., M \tag{7}$$

because those flows contributing to the screenline volume are centered at time interval k, and thus $V_l(t)$ for $t = k$ should not be given greater weight than $V_l(k)$ in Eqn 6. By defining

$$z_{ijl}(k) = \sum_{m=0}^{\sigma_{ilj}} [q_i(k - t_{ilj}^+(k) - m) + q_i(k - t_{ilj}^-(k) + m)]$$

then Eqn 6 can be naturally stated as

$$\sum_{i \in O_l} \sum_{j \in D_l} z_{ijl}(k) \cdot b_{ij} - \sum_{m=1}^{M} V_l(k-m) \cdot \beta_m^- - \sum_{m=1}^{M} V_l(k+m) \cdot \beta_m^+ = V_l(k) \;\; l \in L. \tag{8a}$$

To compress the notation, we further define

$$\mathbf{Z}_l(k) = \{z_{ijl}(k) | \, i \in O_l, j \in D_l; -V_l(k-m) | \, m = 1, 2, ..., M; -V_l(k+m) | \; m = 1, 2, ..., M\}$$

as a column vector; and

$$\mathbf{b} = \{b_{ij} | \; i \in O_l, j \in D_l; \beta_m^- | \; m = 1, 2, ..., M; \beta_m^+ | \; m = 1, 2, ..., M\}$$

as the corresponding parameter vector. Eqn (8a) can then be written in the following concise form

$$V_l(k) = \mathbf{b}^T \mathbf{Z}_l(k) l \in L. \tag{8b}$$

Note that Eqns 8a and 8b presents a fundamental observation equation for estimating O–D parameters b_{ij} and the other model parameters (β_m^-, β_m^+). It is both an enhancement and an extension of the linear system models proposed by Bell (1991), as the constraints established with only exit flows are a subset of the dynamic relations represented in Eqns 8a and 8b. Since each exit link corresponds to a unique network screenline, the number of observation equations in Eqns 8a and 8b can be substantially larger than the total network exits, as one can select as many screenlines as necessary. Eqns 8a and 8b is also an improved version of the model presented by Bell (1991) as it captures the fact that the travel time needed for most vehicles of an identical O–D during the same time interval should not vary substantially. Consequently, the number of additionally required parameters are greatly reduced.

Note that the reduction in the number of parameters for estimation is significant only for roadway networks of realistic size on which the time lag factors should not be neglected. With respect to small networks where the travel time needed to traverse the network is shorter than the unit time interval selected for performing estimation, it is not advantageous to employ our proposed model, and Bell's method will be applicable. More specifically, one may consider, for instance, a particular screenline l that corresponds to a single exit link. Assuming that the longest travel time from the farthest origin node of the $|O_l|$ origins to the given exit is P unit intervals, then the number of parameters needed with Eqns 8a and 8b is $|O_l| + 2M$, but with Bell's method it will be $P \cdot |O_l|$. As P increases proportionately with the network size, but not the parameter M (the variance of travel times for an identical O–D pair), it is clear that for a network of realistic size, the following relation will hold:

$$| \, O_l \, | + 2M << P \cdot | \, O_l \, | \, .$$

One can thus see the substantial reduction in parameters by applying Eqns 8a and 8b.

Also note that the above formulations are derived on the assumption that no trip may go through the screenline more than once. When this condition does not hold, the screenline volume, $V_l(k)$, computed with Eqns 8a and 8b could be under estimated. For instance, some trips from node 1 to 4 (see Fig. 1) may take a path as 1–5–10–9–6–4 that crosses screenline AB twice. The screenline flow computed with Eqns 8a and 8b does not include those trips having their origins and destinations on the same side of the screenline. To take into account such scenarios, one should recognize that a trip must cross that particular screenline either zero or an even number of times if its origin and destination are located at the same side of a screenline.

Similarly, if a trip with its origin and destination at different sides of a screenline, it must go through that screenline for an odd number of times. Thus, if the length of a unit time interval is sufficiently long, or the time difference between the first and the last crossings of the screenline is relatively short, then the under-counted volumes for $V_l(k)$ and $V_{l'}(k)$ with Eqns 8a and 8b, should be approximately equal. Hence, the actual difference between the pair of screenline volumes, $V_l(k)$ and $V_{l'}(k)$, shall be approximately equal to the difference between their estimated results with Eqn 8a and 8b. As such, the following relation can be constructed to represent a generalization of Eqn 8a and 8b:

$$V_l(k) - V_{l'}(k) = \mathbf{b}_l^T \mathbf{Z}_l(k) - \mathbf{b}_{l'}^T \mathbf{Z}_{l'}(k) \quad l \in L. \tag{9}$$

3. ESTIMATION ALGORITHMS

Given a time series of traffic counts $\{q_i(k)\}$, $\{V_i(k)\}$ and travel times $\{t_{ilj}(k)\}$ for all i, j, and l, one can compute $\{Z_l(k)\}$ and establish a series of linear equations for the O–D flow parameters $\{b_{ij}\}$ with Eqn 8a and 8b or Eqn 9. To account for observation noises, the set of linear equations may allow to include stochastic noise terms. Furthermore, as the O–D flow parameters $\{b_{ij}\}$ may vary or evolve over time, it is quite common to represent its time-varying nature as either a random walk process, or some type of known relations (e.g. ARIMA model) if data are available. A standard state-space dynamic model with time-varying parameters can thus be formulated and solved with existing recursive estimation algorithms.

Note that the proposed model has very similar properties as those reported by Bell (1991), but with fewer parameters and more constraints. Therefore, all recursive estimation algorithms reviewed and developed by Bell (1991) are applicable for this study, including both Least-Square-based and Kalman-Filtering-based methods. A detailed description regarding those algorithms is available in the literature (e.g. Bell, 1991; Chang & Wu, 1994), and is not the focus of this paper.

We now turn to address the selection of screenlines and the estimation of dynamic travel times in a network as both are critical to successful applications of the proposed model.

4. SCREENLINE SELECTION

Certainly, the selection of a non-dependent set of screenlines may vary with the network topology and geometry. Although rigorous procedures or algorithms for generating an optimal set of network screenlines remain to be developed, some general guidelines are presented below:

A special screenline (i.e. cordon-line) must first be specified for each origin node, so that those trips, $q_i(k)$, originated from each node i, can be obtained. The LHS part of the network, divided with the selected cordon line, must consist of only the subject node while the RHS part of the network will include all other nodes. Care should be exercised in determining the $q_i(k)$ under the following scenarios:

If an origin node i is not a transition node (i.e. all the trips going through this node originate from this node), $q_i(k)$ is just the volume counts at the cordon line.

If an origin node i is also a transition node but not a destination node (i.e. no vehicles going through node i have their destinations at node i), $q_i(k)$ is the difference between the cordon line (i.e. a screenline) volume and its coupled screenline containing the opposing traffic.

If an origin node i is both a destination node and a transition node, then $q_i(k)$ can not be determined only with the cordon volumes. Hence, one may create a pseudo node to distinguish trips using node i as an origin from those using it as a destination, if necessary.

A cordon line should also be constructed with respect to each destination node j, because such special screenline volumes (i.e. the exit volumes) contain the most direct O–D information for that destination node.

All selected screenlines should be mutually independent, so as to truly increase the system observability.

A screenline should be so designed to minimize its length and the number of intersections with network links. As the O–D flows and their travel times through such screenlines are likely to be distributed in a relatively short interval as discussed in Section 2, it will result in a substantial reduction in the required parameters $\{\beta_m^-, \beta_m^+\}$ for Eqns 8a and 8b and Eqn 9.

Screenlines should be selected to minimize multiple passing trips. Ideally, if there are only a few such trips, one can apply the simplified model (Eqns 8a and 8b) for estimation. Otherwise, its general form, Eqn 9, is more applicable to account for the presence of multiple-crossing trips.

In principle, one may use as many independent screenlines as possible to set up more independent constraints and enhance the system observability.

5. TRAVEL TIME ESTIMATION

As noted in Eqns 8a and 8b and Eqn 9, a reliable estimation of such time-dependent travel times is critical to the effectiveness of the proposed dynamic O–D model. One can certainly take advantage of the emerging surveillance technologies and the information available in advanced traffic management systems (ATMS). Research regarding the application of such information, however, is not the focus of this study. Instead, as a by-product of our research effort on this subject, we have developed the following procedures for trip-time approximation. The resulting accuracy may be inadequate for detailed traffic operational analyses, but shall be sufficient for the macroscopic modeling of network flows.

5.1. Compute travel time for each link, and each intersection turning flows for interval k
Note that there exists a vast body of literature on estimating link travel times and intersection delays (e.g. Glough and Huber, 1975; Takaba *et al.*, 1991; Dailey, 1993 and Ran *et al.*, 1993). A comprehensive review of link travel time estimation with surveillance detector data is available in the article by Sisiopiku and Rouphail (1994).

5.2. Compute the shortest travel time from node i to node j during time interval k (denoted as $s_{ij}(k)$)

Given the time-dependent link travel time, one can employ the dynamic shortest path algorithm developed recently by Mahmassani *et al.* (1993) to compute the approximate time-varying shortest paths for each O–D pair, namely, the shortest travel time $[s_{ij}(k)]$ from origin i to destination j during time interval k, as well as the shortest trip time $[s_{ip}(k)]$ from origin node i to $p \in SL_l$ during time interval k.

5.3. Compute the constrained shortest path travel time for each O–D pair

Given the time-varying shortest path information for each O–D pair, the shortest travel time from origin i to destination j through a particular location $p \in SL_l$, $T_{ipj}(k)$, can be approximated as

$$T_{ipj}(k) = s_{pj}(k) + \lambda s_{ip}[k - s_{pj}^-(k)] + (1-\lambda) s_{ip}[k - s_{pj}^+(k)] \tag{10}$$

where $s_{pj}^-(k)$ is the largest integer less than or equal to $s_{pj}(k)/T$;

$$s_{pj}^+(k) = s_{pj}^-(k) + 1; \text{and}$$

$$\lambda = s_{pj}^+(k) - s_{pj}(k)/T.$$

5.4. Compute the mean travel time $t_{ilj}(k)$

Theoretically, the mean travel time $t_{ilj}(k)$ can be expressed with a weighted average of all $s_{ip}(k)$, i.e. $\sum_p \theta_p s_{ip}(k)$, with the weighting factor θ_p being proportional to the fraction of trips from i through location p to j. Without any route choice information, one of the alternatives is to apply the commonly used logit model for approximating the trip distribution among all available routes, based mainly on the travel time, $T_{ipj}(k)$. Such weighting factors can thus be represented as:

$$\theta_p = \frac{\exp[-\mu T_{ipj}(k)]}{\sum_{r \in SL_l} \exp[-\mu T_{irj}(k)]}$$

where μ is a logit model parameter. Consequently, the mean travel time can be estimated as follows:

$$t_{ilj}(k) = \sum_{p \in SL_l} s_{ip}(k) \frac{\exp[-\mu T_{ipj}(k)]}{\sum_{r \in SL_l} \exp[-\mu T_{irj}(k)]}.$$

6. AN ILLUSTRATIVE EXAMPLE

This section presents an application of the proposed model with simulated network flow data. Note that this example remains mainly illustrative in nature, as a rigorous model evaluation requires the use of either field data or a large-scale simulated network under various traffic scenarios. Unfortunately, none of the existing network traffic simulation models are capable of simulating the network flow movement at a microscopic level, based on the input dynamic O–D flow patterns. The most accessible and reliable traffic models, TRAF-NETSIM and FRESIM (FHWA, 1992) also lack the ability to take given dynamic O–D matrices. The design of a numerical example with realistic dynamic relations between time-varying O–D and link flows thus becomes a challenging research issue. As an extensive evaluation with simulation experiments is not the focus of this study, we have managed to manipulate FRESIM (that can take dynamic O–D matrices for only a single freeway segment) to create a simple freeway network. The numerical results from the proposed example, however, are sufficient for illustrating the application procedures and for ensuring the promise of the proposed model. A further investigation on its potential application in large-scale networks, involving research on both network-decomposition and parallel computing, is beyond the scope of this paper, and is available in Chang and Tao (1994).

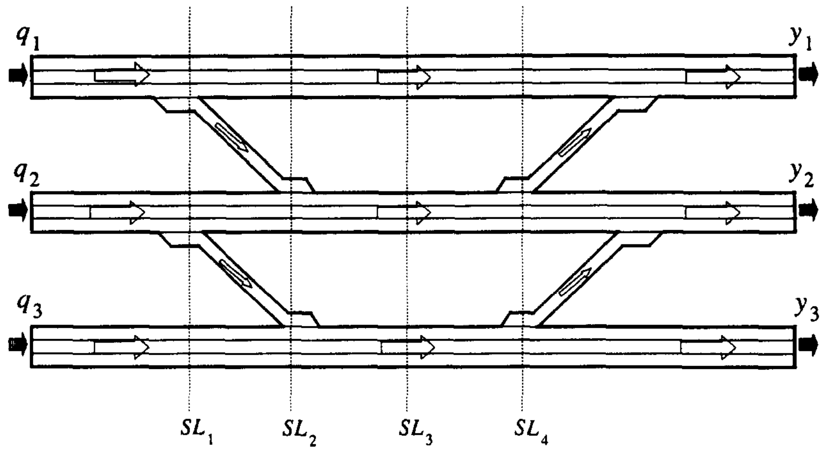

Fig. 2. A hypothetical freeway network for numerical test.

6.1. Example network and data set generation

Figure 2 illustrates a freeway network, consisting of freeway segments and four ramp links. There are three entry streams, q_1, q_2, q_3 and three exit streams, y_1, y_2, y_3. Under the given network structure, entry flow q_1 cannot reach exit y_3; nor can any entry flow q_3 go to exit y_1. Thus, there exists a total of the following seven O–D parameters:

$$b_{11}, b_{12}; b_{21}, b_{22}, b_{23}; b_{32}, b_{33}.$$

Some guidelines employed in design of this numerical example are summarized below:

The time-varying link travel time has been taken directly from the FRESIM simulation output. The dynamic travel time for each node pair is computed with Eqn 10.

Since FRESIM can handle up to 19 time periods with different O–D flows for each simulation run, a total of 19 sets of entry volumes $\{q_1(k), q_2(k), q_3(k)\}$, $k = 1, 2, \ldots, 19$ representing 19 time intervals, are randomly generated and shown in Table 1.

The following O–D flow parameters, b_{ij}, are input originally to FRESIM:

$$b_{11} = 0.8, b_{12} = 0.2; b_{21} = 0.15, b_{22} = 0.7, b_{23} = 0.15; b_{32} = 0.2, b_{33} = 0.8$$

The resulting O–D flow fractions, after being selected randomly during each small simulation interval (5 min) by FRESIM, are listed in Table 2, which were identified through printing out the FRESIM inner variables.

Four screenlines, SL_1 to SL_4, as shown with the dotted lines in Fig. 2, have been taken for numerical evaluation. Each screenline volume equals the sum of the three associated link flows.

Based on the simulation results, the actual travel time deviation parameters have been identified to be very small, relative to the unit time interval duration (5 min). Hence, $m_{ilj} = 0$, for all i, l and j; and thus $M = 1$. Due to the presence of a time lag ($M = 1$), the observation equation for time interval 19 cannot be constructed without the screenline flow data at the 20th time interval. Thus, only 18 time intervals of traffic data are used in the model estimation.

Table 1. The time-series of entry volumes generated randomly for FRESIM simulation

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
$q_1(k)$	476	441	418	476	489	497	496	481	456	471	486	499	475	461	473	425	440	449	498
$q_2(k)$	178	155	201	130	204	212	161	130	155	151	208	216	223	136	127	135	130	155	290
$q_3(k)$	413	434	438	395	391	356	407	414	418	380	414	375	376	458	412	437	382	453	380

Table 2. The actual O–D split parameters used in FRESIM simulation

k	$b_{11}(k)$	$b_{12}(k)$	$b_{21}(k)$	$b_{22}(k)$	$b_{23}(k)$	$b_{32}(k)$	$b_{33}(k)$
1	0.806	0.194	0.176	0.693	0.131	0.205	0.795
2	0.784	0.216	0.097	0.766	0.136	0.181	0.819
3	0.822	0.178	0.135	0.710	0.155	0.161	0.839
4	0.808	0.192	0.186	0.674	0.140	0.176	0.824
5	0.759	0.241	0.148	0.690	0.163	0.208	0.792
6	0.788	0.212	0.152	0.687	0.161	0.201	0.799
7	0.814	0.186	0.162	0.688	0.150	0.178	0.822
8	0.785	0.215	0.163	0.682	0.155	0.211	0.789
9	0.771	0.229	0.182	0.695	0.123	0.221	0.779
10	0.808	0.192	0.180	0.687	0.133	0.191	0.809
11	0.800	0.200	0.180	0.684	0.136	0.214	0.786
12	0.795	0.205	0.154	0.716	0.130	0.198	0.802
13	0.810	0.190	0.144	0.721	0.135	0.240	0.760
14	0.817	0.183	0.141	0.652	0.207	0.167	0.833
15	0.811	0.189	0.135	0.738	0.127	0.242	0.758
16	0.813	0.187	0.127	0.701	0.172	0.161	0.839
17	0.811	0.190	0.186	0.643	0.170	0.208	0.792
18	0.837	0.163	0.169	0.708	0.123	0.231	0.769
19	0.801	0.199	0.160	0.705	0.135	0.223	0.777

6.2. *Experimental design*

The following six scenarios have been used for estimating the O–D flow parameters, based on the constraints in Eqns 8a and 8b and the selected screenlines. Note that Eqns 8a and 8b is sufficient to use in this example due to the use of a one-way network structure.

Base scenario: To construct the observation equations based solely on the three exit volumes, $y_1(k)$, $y_2(k)$, $y_3(k)$.

Scenario 1: To construct the observation equations, based on the three exit volumes, $y_1(k)$, $y_2(k)$, $y_3(k)$, and the first screenline flow, $V_1(k)$.

Scenario 2: To construct the observation equations, based on the three exit volumes, $y_1(k)$, $y_2(k)$, $y_3(k)$, and the first two screenline flows, $V_1(k)$ and $V_2(k)$.

Scenario 3: To construct the observation equations, based on the three exit volumes, $y_1(k)$, $y_2(k)$, $y_3(k)$, and the first three screenline flows, $V_1(k)$, $V_2(k)$, $V_3(k)$.

Scenario 4: To construct the observation equations, based on the three exit volumes, $y_1(k)$, $y_2(k)$, $y_3(k)$, and the first four screenline flows, $V_1(k)$, $V_2(k)$, $V_3(k)$, $V_4(k)$.

6.3. *Estimation results*

The sequential Kalman filtering algorithm (Chui & Chen, 1991) has been employed in parameter estimation for all the six scenarios. The truncation and normalization procedures by Nihan and Davis (1987) has also been adopted to satisfy the natural constraints for the O–D split proportions. Other recursive estimation algorithms may also be used and should yield similar results (Grillenzoni, 1994).

The recursive Kalman-filtering is started with the following arbitrarily selected initial parameters:

$$b_{11}(0) = b_{12}(0) = b_{32}(0) = b_{33}(0) = 0.5$$

$$b_{21}(0) = b_{22}(0) = b_{23}(0) = 0.333$$

All other Kalman filter parameters (i.e. the variance–covariance matrices of the random terms) are given with identical values for the five proposed scenarios. Though none of the five proposed scenarios can yield the set of O–D parameters identical to those specified originally, the estimated results have proved the potential of extending the dynamic screenline modeling concept to large-scale network applications. In fact, as shown with the rooted-mean-squared (RMS) errors in Table 3, the use of screenline flow data in addition to the exit flows has substantially improved the estimation results. For instance, the RMS error has reduced from 0.11 down to 0.05, when four screenline flows are used in model construction and estimation.

Note that these RMS results were computed from the 5th time interval to remove the initial bias due to the required model warming-up time. To further present the estimation results, the dynamic pattern of all O–D parameters and their estimated values are depicted in Fig. 3. It is obvious that over the entire period, the base scenario, which utilized only the three exit flows, has yielded the worst results, while Scenario 4 which utilized the set of four screenline flows along with the three exit flows has generated the relatively best estimates. Also note that although the proposed model is grounded on the work by Bell

Table 3. The RMS errors of the estimation results

Scenario	Base	1	2	3	4
RMS error	0.1110	0.0906	0.0821	0.0605	0.0518

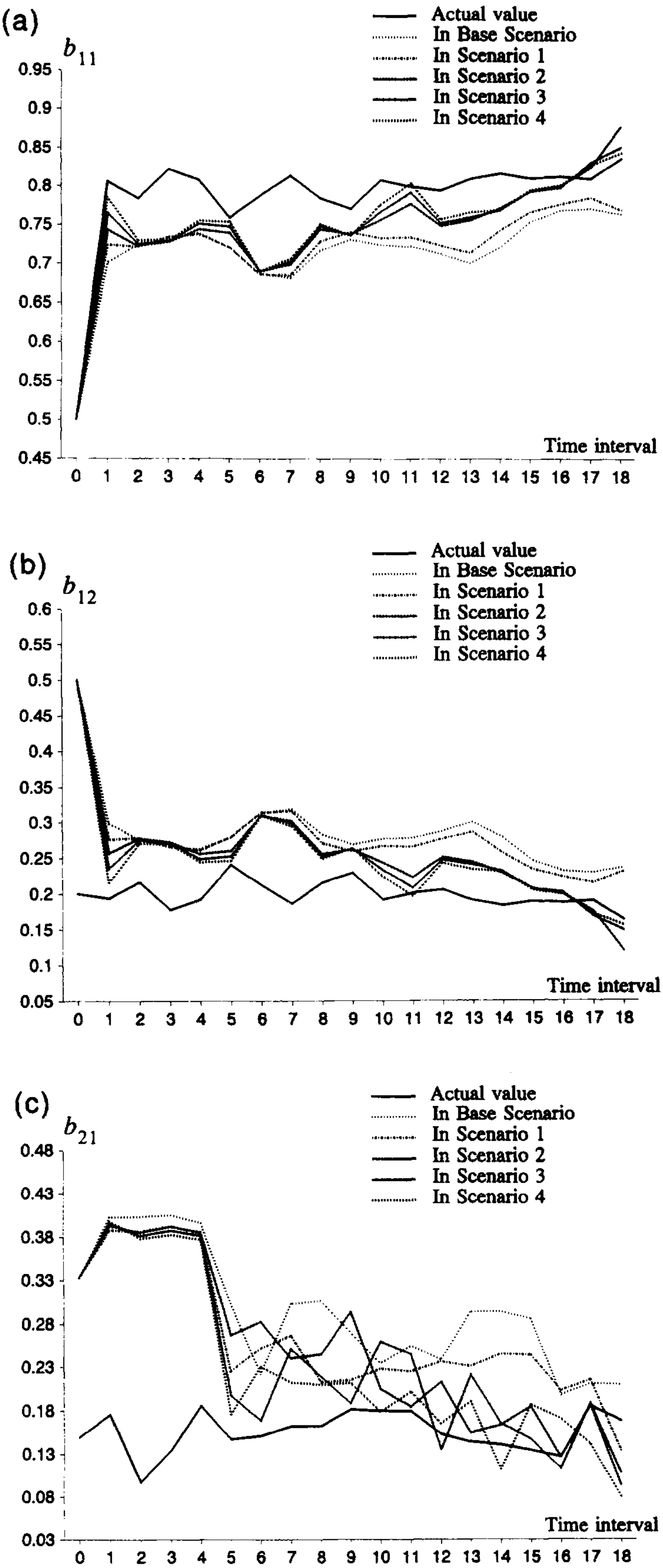

Fig. 3(a)–(c). *Caption on p. 289.*

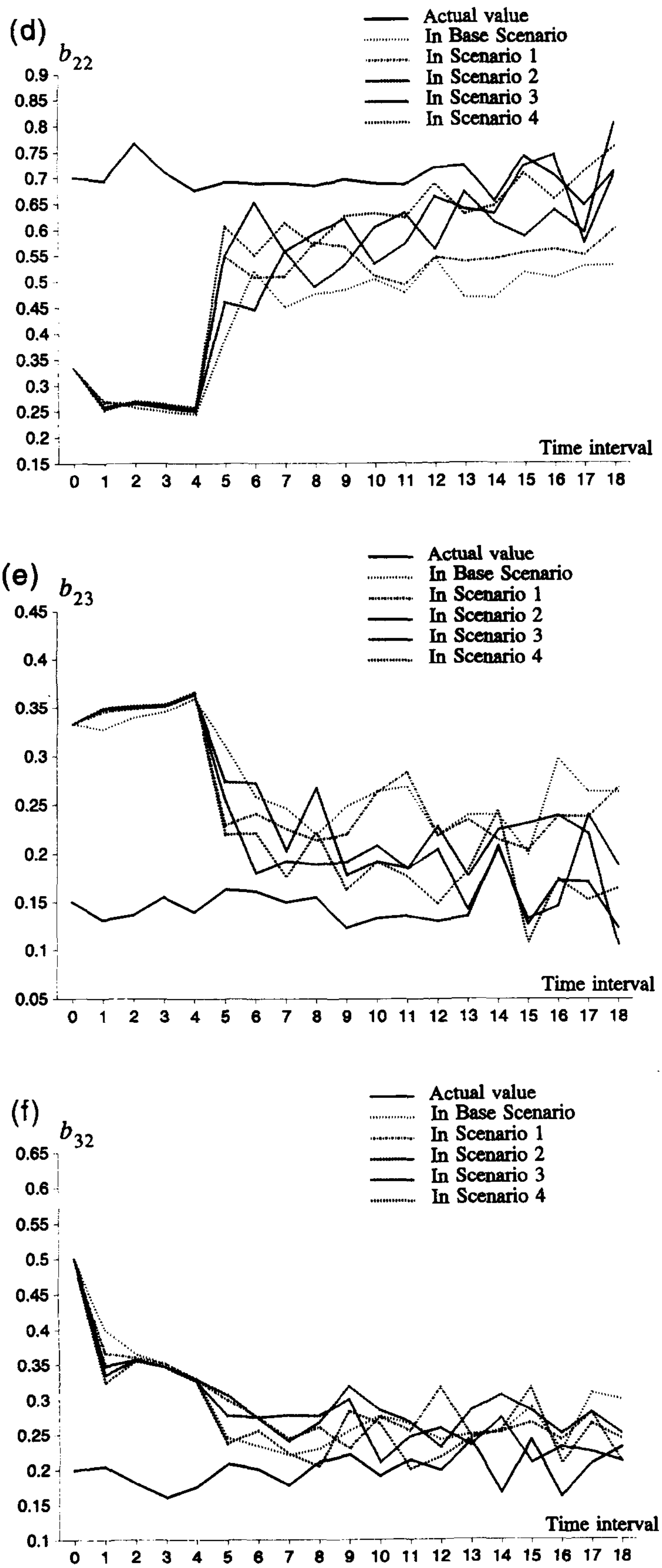

Fig. 3(d)–(f). *Caption on p. 289.*

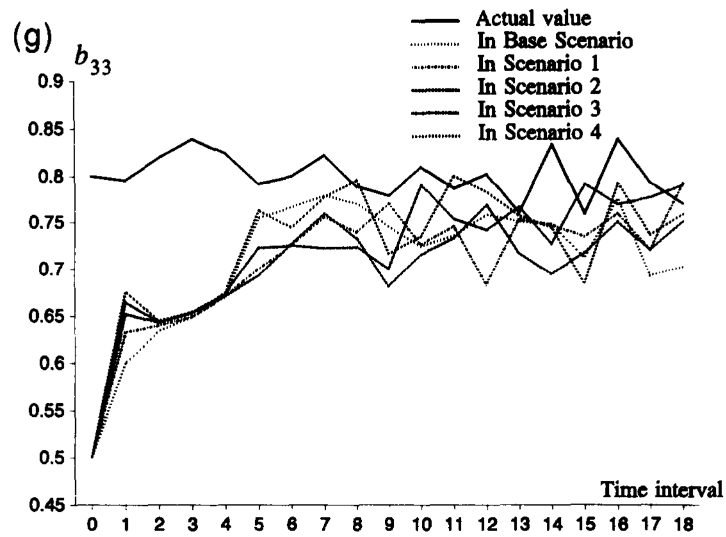

Fig. 3. A graphical representation of the estimation results for all O–D parameters.

(1991), both the dynamic O–D and flow relations as well as the system observability have been significantly enhanced due to the inclusion of dynamic screenline flows and travel time variability.

7. CONCLUSIONS

We have fully recognized that the entire research on the dynamic network O–D estimation remains in its infancy. Certainly, our presented results represent only one of the many required steps towards the development of a reliable and efficient model for use in large-scale network O–D estimation. Some critical issues to be overcome include:

— selection of the optimal set of screenlines for a given network, and the estimation of time-varying trip times under various congestion levels;

— the necessary decomposition of a large network into more tractable inter-connected subnetworks;

— the potential of using parallel computing hardware and software for improving the operational efficiency and for real-time applications.

The recent advancements in surveillance technologies and advanced traveller information systems may also provide some useful dynamic network information resources to increase system observability and estimation accuracy.

Acknowledgment—The authors are grateful to three anonymous referees and the Associate Editor for their constructive comments.

REFERENCES

Ashok K. and Ben-Akiva M. E. (1993) Dynamic origin–destination matrix estimation and prediction for real-time traffic management systems. *Proceedings of 12th International Symposium of Transportation and Traffic Theory. Berkeley, CA.*

Bell M. G. H. (1991) The real time estimation of origin–destination flows in the presence of platoon dispersion. *Transpn Res.* **25B**, 115–125.

Cascetta E., Inaudi D. and Marquis G. (1993) Dynamic estimators of origin–destination matrices using traffic counts. *Transpn Sci.* **27**, 363–373.

Cascetta E. and Nguyen S. (1988) A unified framework for estimating or updating origin/destination matrices from traffic counts. *Transpn Res.* **22B**, 437–455

Chang G.-L and Tao X. (1994) Estimation of large-scale network O–D flows with decomposition and parallel computing methods. *Working Paper*, Department of Civil Engineering, University of Maryland at College Park.

Chang G.-L and Wu J. (1994) Recursive estimation of time-varying O–D flows from traffic counts in freeway corridors. *Transpn Res.* **28B**, 141–160.
Chui C. K. and Chen G. (1991) *Kalman Filtering with Real-Time Applications.* Springer-Verlag, NY.
Cremer M. (1983) Determining the time dependent trip distribution in a complex intersection for traffic responsive control. *Preprint of the 4th Int. IFAC/IFIP/IFORS Conf. on Control in Transportation Systems*, Baden-Baden, Germany.
Cremer M. and Keller H. (1981) Dynamic identification of O–D flows from traffic counts at complex intersections. *Proc. 8th Int. Symp. on Transp. and Traffic Theory*, University of Toronto Press, Toronto, Canada.
Cremer M. and Keller H. (1987) A new class of dynamic methods for the identification of origin–destination flows. *Transpn Res.* **21B**, 117–132.
Cremer M. and H. Keller H. (1984) A Systems dynamics approach to the estimation of entry and exit O–D flows. *Proc. 9th Int. Symp. on Transp. and Traffic Theory* (J. Volmuller and R. Hamerslag, Eds). VUN Science Press, Utrecht, The Netherlands
Dailey D. J. (1993) Travel time estimation using correlation techniques. *Transpn Res.* **27B**, 97–107
FHWA (1992) *TRAF User Reference Guide.* Publication No. FHWA-RD-92-060. U.S. Department of Transportation, Federal Highway Administration, McLean, VA.
Glough D. L. and Huber M. J. (1975) *Traffic Flow Theory.* TRB Special Report 165. Washington, D.C.
Grillenzoni C. (1994) Optimal recursive estimation of dynamic models. *J. Am. Stat. Assoc., Applic. Case Studies* **89**, 777–787.
Kessaci A. *et al.* (1990) Online estimation of turning movements and saturation flows in PRODYN. *Proceedings of the 6th IFAC/IFIP/IFORS Symposium on Transportation*, France, Paris, September.
Keller H. and Ploss G.(1987) Real-time identification of O–D network flow from counts for urban traffic control. Transportation and Traffic Theory (Gartner N. H. and Wilson N. H. M., Eds, pp. 267–284). Elsevier Science Publishing Co. Inc., NY.
Mahmassani H. S., Hu T.-Y, Peeta S. and Ziliaskopoulos (1993) *Development and testing of dynamic traffic assignment and simulation procedures for ATIS/ATMS applications.* Technical Report DTFH61-90-R-00074-FG. Center for Transportation Studies, The University of Texas at Austin.
Nihan N. L. and Davis G. A. (1987) Recursive estimation of origin–destination matrices from input/output counts. *Transpn. Res.* **21B**, 149–163
Nihan N. L. and Davis G. A. (1989) Application of prediction-error minimization and maximum likelihood to estimate intersection O–D matrices from traffic counts. *Transpn Sci.* **23B**, 77–90
Nihan N. L. and Hamed M. M. (1992) Fixed-point approach to estimating freeway origin–destination matrices and the effect of erroneous data on estimate precision. *Transp. Res.* Record 1357, 18–28.
Nguyen S. (1984) Estimating origin–destination matrices from observed flows. *Transportation Planning Methods* (Florian M., Ed., pp. 363–380). North Holland, Amsterdam.
Okutani I. (1987) The Kalman filtering approaches in some transportation and traffic problems. *Transportation and Traffic Theory* (Gartner N. H. and Wilson N. H. M., Eds., pp. 363–380), Elsevier Science Publishing Co., Inc.
Ran B., Rouphail N. M., Tarko A. and Boyce D. E. (1992) Toward a class of link travel time functions for dynamic network models. *Paper submitted for publication in Transportation Research.*
Robertson D. I. (1969) *TRANSYT: A Traffic Network Study Tool.* RRL Report LR253.
Sisiopiku V. and Rouphail N. M. (1994) Towards the use of detector output for arterial link travel time estimation: A literature review. *TRB Preprint No. 940160.* Washington, D.C.
Van Aerde M., Hellinga B., Mackinnon G. and Wang J. (1993) QEENSOD: A method for estimating time varying origin–destination demands for freeway corridors/networks. *Paper presented at the 72nd annual meeting of Transportation Research Board*, Washington D.C., January.
Wu J. and Chang G.-L. (1994) A review of existing approaches for estimation of dynamic O–D matrices. *Working Paper*, Department of Civil Engineering, University of Maryland, College Park, MD.
Yu P. and Davis G. A. (1994) Estimating freeway origin–destination patterns using automatic traffic counts. *TRB Preprint No. 940112.* Washington, D.C.