Enhanced Dynamic Origin–Destination Matrix Updating with Long-Term Flow Information

Antony Stathopoulos and Theodore Tsekeris

The problem of updating dynamic origin–destination (O-D) matrices by exploiting a long-term time series of link traffic counts in large-scale transportation networks without the need for surveys or census data is investigated. Different time-recursive mechanisms for analysis of these data to enhance the performance of the models currently used to synthesize within-day dynamic O-D matrices are suggested. The efficiency of the proposed procedure is investigated with respect to different formulations and related solution algorithms (based on entropy maximization and generalized least-squares). The impacts of different assumptions on model performance are also examined. These include the length of the time scale in which the flow information is updated and the selection of the weekday for which the flow information is collected. The results of the statistical analysis and the model performance measures demonstrate that the proposed time-recursive procedure for an information updating period of 2 years can produce an improved prior O-D matrix that may significantly enhance the subsequent updating of dynamic O-D matrices corresponding to a series of days of the week.

The updating of origin–destination (O-D) trip matrices by use of link traffic count data, usually referred to as O-D matrix synthesis, can provide valuable information about the demand for trip interchange between zones or nodes (intersections) in the transportation network. The low cost and effort associated with the automated collection of these data have resulted in the widespread use of models for updating O-D matrices, which are extremely useful in transportation planning and traffic operations. But, a prior O-D matrix is typically used as a target or seed to guide the procedure used to solve these models. Comparisons of several dynamic models of O-D matrix synthesis—models that estimate the temporal pattern of O-D demand on the basis of a time series of links counts over successive time intervals in a given study period—have demonstrated that they can produce link flow solutions significantly superior to those obtained by dynamically assigning the prior O-D matrix onto the network (*1*). A review of the models used for the dynamic assignment of the O-D matrix onto the network, known as dynamic traffic assignment models, is provided elsewhere (*2*).

However, a credible prior O-D matrix capable of adequately representing the actual O-D demand in the study period is not available in most urban areas. This is because of the costs involved with regularly conducting travel surveys or collecting socioeconomic or census data to estimate a prior O-D matrix by use of a transportation planning model. This problem can lead to the reduced performance of the O-D matrix updating, particularly as the time between the study period and the period in which the prior O-D matrix was estimated increases. The model performance can be described by the extent to which it can produce an O-D matrix close to the prior one (or the one most likely to exist) while sufficiently replicating the measured link volumes.

The problem of prior O-D demand information has hitherto been investigated only for the case of static (time-uniform) O-D matrix synthesis by use of standard census data (*3*). On the contrary, the corresponding problem for the case of dynamic O-D matrix synthesis has been largely restricted to the treatment of the dynamics underlying the within-day O-D demand information corresponding to previous time intervals and estimated by using some recursive (Bayesian updating) model (*4*). Nonetheless, the latter method entails a considerable computational load that may render it impractical for applications on realistically sized urban networks and related online traffic operations. In addition, such a method does not consider the possible impacts of the long-term systematic changes of prior demand information, for example, because of land use changes, on the dynamic O-D matrix synthesis.

This paper presents a cost-efficient and effective procedure that can be used to enhance the performance of dynamic O-D matrix updating by the use of a long-term time series of link traffic counts as the sole data source. The applicability of the proposed procedure to real-world urban networks is demonstrated by a relevant case study of the central part of the network of the greater Athens area in Greece. The next section presents two alternative formulations and solution procedures for the problem of dynamic O-D matrix synthesis. The dynamic traffic assignment model with which the models used to synthesize dynamic O-D matrices are combined is then described. Different mechanisms for incorporation of the long-term dynamics of O-D information into the suggested models and a description of the case study network are provided. The results of simulation tests undertaken to investigate the impacts of these mechanisms on the model performance measures, including the effects of different time scales, weekdays, and solution procedures, are presented. Moreover, a statistical analysis and interpretation of the results are carried out. Finally, conclusions drawn from the study results are presented.

MODELS OF DYNAMIC O-D MATRIX SYNTHESIS

The problem of dynamic O-D matrix synthesis is first described on the basis of the maximum-entropy (ME) formulation (*5*). Consider a network that includes N traffic zones from and to which vehicular

Department of Transportation Planning and Engineering, School of Civil Engineering, National Technical University of Athens, 5 Iroon Polytechniou, 157 73 Athens, Greece.

Transportation Research Record: Journal of the Transportation Research Board, No. 1882, TRB, National Research Council, Washington, D.C., 2004, pp. 159–166.

trips are allocated. The network is composed of K nodes and M directed links. Let $\mathbf{x}^{w_\ell,\epsilon_k}$ be the (unknown) O-D matrix whose elements $x_{ij}^{\tau_d,w_\ell,\epsilon_k}$ denote the number of vehicular trips between the nth O-D pair departing from origin $i \in N$ during (departure) time interval $\tau_d \in T_d$ and directed toward destination $j \in N$. The total departure period $T_d \subseteq w_\ell$ typically refers to the (morning or afternoon) peak period of some day of the week $w_\ell \in W$ in month (or year) $\epsilon_k \in E$. The set $W = \{w_1, w_2, \ldots, w_\ell, \ldots\}$ is a time series of consecutive days of the week, for example, the Tuesdays within a month (or a season). The set $E = \{\epsilon_1, \epsilon_2, \ldots, \epsilon_k, \ldots\}$ is an equally spaced time series of (not necessarily consecutive) months in a year or years in a decade. Also, consider $M_O \subseteq M$ to be the total number of the observed links that contain a traffic counter, $y_m^{\tau,w_\ell,\epsilon_k}$ the observed traffic count on the mth link, and $\widehat{y}_m^{\tau,w_\ell,\epsilon_k}$ the assigned (estimated) link volume on that link during (count) time interval $\tau \in T \subseteq T_d$. The interval $\tau \geq \tau_d$ may include any interval in which link flow $y_m^{\tau,w_\ell,\epsilon_k}$ has been counted. Then, the problem of the ME-based dynamic O-D matrix synthesis, in terms of estimating the statistically optimal (most probable) trip departure time distribution, can be formulated as follows:

$$\text{minimize} \sum_{i \in N} x_i^{\tau_d,w_\ell,\epsilon_k} \left[\log\left(\frac{x_i^{\tau_d,w_\ell,\epsilon_k}}{\widehat{x}_i^{\tau_d,w_\ell,\epsilon_k}} \right) - 1 \right] \qquad \forall \tau_d \in T_d \tag{1}$$

subject to

$$y_m^{\tau,w_\ell,\epsilon_k} \sum_{i \in N} \sum_{\tau_d \in T} a_{im}^{\tau_d,\tau,w_\ell,\epsilon_k} x_i^{\tau_d,w_\ell,\epsilon_k} \qquad \forall m \in M_O, \forall \tau \in T \tag{2}$$

and

$$x_i^{\tau_d,w_\ell,\epsilon_k} \geq 0 \qquad \forall i \in N, \forall \tau_d \in T_d \tag{3}$$

where $\widehat{x}_i^{\tau_d,w_\ell,\epsilon_k}$ corresponds to the row (origin) sums of the matrix $\widehat{\mathbf{x}}^{w_\ell,\epsilon_k} = \{\widehat{x}_{ij}^{\tau_d,w_\ell,\epsilon_k}\}$. Similarly, $x_i^{\tau_d,w_\ell,\epsilon_k}$ denotes the row sums of $\mathbf{x}^{w_\ell,\epsilon_k}$. The matrix $\widehat{\mathbf{x}}^{w_\ell,\epsilon_k}$ is known as the prior O-D matrix and provides a preliminary estimate for $\mathbf{x}^{w_\ell,\epsilon_k}$. on the basis of an O-D matrix synthesized by using measurements of link volumes during (count) period T of day w_ℓ of some previous month or year ϵ_h, with $h < k$. Another option is to obtain $\widehat{\mathbf{x}}^{w_\ell,\epsilon_k}$ on the basis of a travel survey or the output of a transportation planning model fed socioeconomic or census data corresponding to ϵ_h. The variables $a_{im}^{\tau_d\tau,w_\ell,\epsilon_k}$ are known as link use proportions, and they represent the elements of the assignment matrix A^{w_ℓ,ϵ_k}. Each of these elements determines the proportion of the trip demand $x_i^{\tau_d,w_\ell,\epsilon_k}$ departing from some origin $i \in N$ during τ_d that contributes to traffic flow $y_m^{\tau,w_\ell,\epsilon_k}$ on link m during interval τ. These proportions are calculated here within a dynamic assignment model, whose description is provided below. The right term of Equation 2 represents the estimated flows, $\widehat{y}_m^{\tau,w_\ell,\epsilon_k}$.

Another approach that can be used to synthesize the most probable or least biased dynamic O-D matrix is to seek to minimize the sum-of-squared errors between the measured and the estimated link volumes. Then, the objective function (Equation 1) can be expressed within the following generalized least-squares (GLS) formulation:

$$\text{minimize} \sum_{m \in M_O} \left[\frac{1}{2} \left(y_m^{\tau,w_\ell,\epsilon_k} - \widehat{y}_m^{\tau,w_\ell,\epsilon_k} \right)^2 \right] \qquad \forall \tau \in T \tag{4}$$

Both the ME- and the GLS-based formulations rely on the assumption that the counted flows, $y_m^{\tau,w_\ell,\epsilon_k}$, can be considered accurate with respect to the unknown $x_i^{\tau_d,w_\ell,\epsilon_k}$ values. In addition, the confidence assigned to each of the flow measurements on each link $m \in M_O$ is the same, given that they are obtained from a single datum source, which, in the present case, is an automated traffic-counting system. The same confidence is also assigned to each entry of the prior O-D matrix $\widehat{\mathbf{x}}^{w_\ell,\epsilon_k}$, which was originally obtained for the present study from a travel survey in base year ϵ_1 (see below).

The ME-based problem is solved by using the dynamically extended multiplicative algebraic reconstruction technique, enhanced by a diagonalized Newton-type search, which has been shown to have the most desirable convergence speed and solution accuracy properties in comparison with those of several other available algorithms (*1*). The GLS-based problem is solved by the Jacobi method and is enhanced by a relaxation technique applied in each iteration, as proposed elsewhere (*6*), because this algorithm has been shown to have increased computational efficiency compared with those of other algorithms, such as the standard Jacobi method and the Gauss-Seidel method. The use of departure time variables $x_i^{\tau_d,w_\ell,\epsilon_k}$ and $\widehat{x}_i^{\tau_d,w_\ell,\epsilon_k}$ instead of the corresponding O-D variables $x_{ij}^{\tau_d,w_\ell,\epsilon_k}$ and $\widehat{x}_{ij}^{\tau_d,w_\ell,\epsilon_k}$ to solve both problems results in a reduction of the dimension of the assignment matrix A^{w_ℓ,ϵ_k}, so that addresses the computational burden associated with real-life urban network applications and avoids overspecification of the formulation of the problem. Moreover, the trip departure time is recognized as the principal decision variable that should be used when the evolution of trip demand under dynamic network conditions is considered (*2*).

The technique used to transform trip departures $x_i^{\tau_d,w_\ell,\epsilon_k}$ into equivalent O-D trip flows $x_{ij}^{\tau_d,w_\ell,\epsilon_k}$ is based on the assumption of a fixed distribution of trip destinations from each origin i, according to the prior O-D demand pattern, and is given as follows:

$$x_{ij}^{\tau_d,w_\ell,\epsilon_k} = x_{ij}^{\tau_d,w_\ell,\epsilon_k} \left(\frac{\sum_{\tau_d \in T_d} \widehat{x}_{ij}^{\tau_d,w_\ell,\epsilon_k}}{\sum_{\tau_d \in T_d} \widehat{x}_i^{\tau_d,w_\ell,\epsilon_k}} \right) \qquad \forall i, j \in N, \forall \tau_d \in T_d \tag{5}$$

The transformation process described above does not constrain the trip departures over each time interval τ_d. Hence, the resulting demand is elastic, in the sense that a smaller or larger percentage of the trips originally departing from the origin zones at τ_d may be estimated to optimize the objective function (Equation 1 or 4), given the constraints (Equations 2 and 3). It is also noted that the distribution of trip destinations from each origin i is not constant across a given time scale but varies according to the updated O-D trip flows of the previous time point.

The time scale of the problem, that is, the distance between ϵ_h and ϵ_k, may give rise to different interpretations about the loss of O-D information conveyed into the model. This O-D information loss describes the extent to which $\widehat{\mathbf{x}}^{w_\ell,\epsilon_k}$ should be transformed (increased, decreased, or restructured) so that when the resultant dynamic O-D matrix $\mathbf{x}^{w_\ell,\epsilon_k}$ is dynamically assigned onto the network it produces traffic flows close to the measured ones (the appropriate mathematical expressions are provided below). When this time scale corresponds to a few days or weeks, the O-D information loss can probably be attributed to random, local, and day-specific variations in the customary travel patterns of users, for example, because of incidents or weather conditions. In the case in which the time scale is several months, the O-D information loss may be attributed to more systematic changes in the activity patterns of users and seasonal variations in demand.

In the case in which the time scale extends from a few years to decades, the O-D information loss is associated with longer-term decisions of users, such as those concerning the workplace and housing location, permanent trends in traveler population compositions, and land use changes. In particular, the analysis of long-term time series of flow information, such as those extending over a decade (see below), may be useful for addressing the practical problem of the scarcity of O-D information and the cost and difficulty of conducting travel surveys, which are usually updated at periods of more than one decade.

DYNAMIC TRAFFIC ASSIGNMENT PROCEDURE

The present dynamic assignment model is based on a macroscopic simulation procedure, originally developed previously to enable the estimation of time-dependent link use proportions $a_{im}^{\tau_d\tau,w_\ell,\epsilon_k}$ (*7, 8*). This procedure provides a plausible tool, in terms of mathematical complexity, as well as an efficient and tractable tool, in terms of computational cost, for calculation of the dynamic traffic loading conditions over realistic large-scale networks. Furthermore, it requires a relatively small amount of data as input. It basically uses the same data used by a standard static assignment model, although it enables the updating of $a_{im}^{\tau_d\tau,w_\ell,\epsilon_k}$ proportions endogenously and has the single requirement that the link volumes and travel times of the previous count interval, $\tau - 1$, be stored. It is based on the reactive or instantaneous link travel cost definition. Namely, each traveler departing from a particular location at any instant chooses the shortest path to his or her destination on the basis of the travel time perceived according to the currently prevailing traffic conditions. Thus, the procedure calculates at the beginning of each interval τ the time-dependent minimum-cost path from each origin i and then assigns each packet of simulated vehicles departing from the corresponding origin i to the best route to reach the intended destination, j. More detailed information on the formulation and solution of the model is provided elsewhere (*1*).

TREATMENT OF LONG-TERM DYNAMICS

This section describes several alternative time-recursive mechanisms that can be used to incorporate the dynamic O-D matrix estimate corresponding to period T_d on day w_ℓ of some past month (or year) ϵ_h in the synthesis of the dynamic O-D matrix corresponding to period T_d on day w_ℓ of the current month (or year) ϵ_k. The first mechanism, referred to here as the time-recursive estimation procedure, can be expressed as follows:

$$\hat{x}_{ij}^{\tau_d,w_\ell,\epsilon_k} = \eta^{\tau_d,w_\ell,\epsilon_k} \sum_{\tau_d \in T_d} \left(\sum_{i \in N} x_{ij}^{\tau_d,w_\ell,\epsilon_{k-1}} \right) \tag{6}$$

where $x_{ij}^{\tau_d,w_\ell,\epsilon_{k-1}}$ denotes the estimated dynamic O-D trip flows corresponding to period T_d on day w_ℓ of month (or year) ϵ_{k-1}, that is, the month (or year) immediately before ϵ_k in the set E. The factor $\eta^{\tau_d,w_\ell,\epsilon_k}$ refers to the temporal split proportions that partition the estimated $\mathbf{x}^{w_\ell,\epsilon_{k-1}}$ matrix (corresponding to the previous point in time) into a number of partial matrices equal to the total number of count intervals, τ. These proportions are calculated as follows:

$$\eta^{\tau_d,w_\ell,\epsilon_k} = \frac{\sum_{m \in M_O} y_m^{\tau_d,w_\ell,\epsilon_k}}{\sum_{\tau_d \in T} \sum_{m \in M_O} y_m^{\tau_d,w_\ell,\epsilon_k}} \qquad \forall \tau_d \in T \subseteq T_d \tag{7}$$

The relationship in Equation 7 implies that $\Sigma_{\tau_d \in T}\, \eta^{\tau_d,w_\ell,\epsilon_k}$ is equal to 1 and refers to those link volumes that have been generated by trips departing during time interval τ_d, which was assumed above to be equal to the count interval τ ($\tau_d = \tau$). For those trips departing at $\tau_d \notin T$, which were generated before the count period, the corresponding link flows can be assumed to be equal to the flows measured during the first interval τ. It should be noted that this partitioned O-D matrix based on $\eta^{\tau_d,w_\ell,\epsilon_k}$ factors is used only as the initial solution to guide the search of the estimation procedure used to update the dynamic O-D matrix on the basis of the current flow information. The use of these proportions is justified by the lack of costly information based on disaggregated data, regularly updated O-D surveys, or census data.

The second mechanism, referred to as the smoothing time-recursive estimation procedure, aims to achieve a more stable transition from ϵ_{k-1} to ϵ_k by avoiding possible large variations among the estimated past and present O-D trip flows. This is done endogenously by weighing past estimates with weights that decrease exponentially with time, as follows:

$$\begin{aligned}\hat{x}_{ij}^{\tau_d,w_\ell,\epsilon_k} = \eta^{\tau_d,w_\ell,\epsilon_k} \Bigg[\lambda \sum_{\tau_d \in T_d} \left(\sum_{i \in N} x_{ij}^{\tau_d,w_\ell,\epsilon_{k-1}} \right) \\ + (1 - \lambda) \sum_{\tau_d \in T_d} \left(\sum_{i \in N} \hat{x}_{ij}^{\tau_d,w_\ell,\epsilon_{k-1}} \right) \Bigg]\end{aligned} \tag{8}$$

where $\hat{x}_{ij}^{\tau_d,w_\ell,\epsilon_{k-1}}$ denote the prior dynamic O-D trip flows corresponding to period T_d on day w_ℓ of month (or year) ϵ_{k-1}. The weight parameter $0 < \lambda < 1$, usually known as a smoothing factor, determines the relative contributions of $\hat{x}_{ij}^{\tau_d,w_\ell,\epsilon_{k-1}}$ and $x_{ij}^{\tau_d,w_\ell,\epsilon_{k-1}}$ to the current prior O-D trip pattern. In all simulation tests and for all time scales considered in the study (see below), the optimal value of the λ factor, in terms of minimizing the discrepancy between the measured and the estimated link flows (see Equation 13 below), was experimentally determined to range from 0.30 to 0.35. This output may suggest a considerable influence of the past resultant dynamic O-D matrix on the current one.

The relationship in Equation 8 implies the existence of a flat trend in the temporal evolution of the aggregate dynamic O-D trip flows. To account for the effects of significant (systematic) changes in the distribution of dynamic O-D trip flows between ϵ_{k-1} and ϵ_k, a normalization process is followed to suitably adjust (increase or decrease) the total production of trips to the present level of prevailing traffic congestion. This approach, referred to as normalized smoothing time-recursive estimation procedure, is expressed here as follows:

$$\hat{x}'^{\tau_d,w_\ell,\epsilon_k}_{ij} = \phi^{\tau_d,w_\ell,\epsilon_k} \hat{x}_{ij}^{\tau_d,w_\ell,\epsilon_k} \tag{9}$$

where $\hat{x}'^{\tau_d,w_\ell,\epsilon_k}_{ij}$ is a modified prior O-D matrix corresponding to ϵ_k and is estimated on the basis of the prior O-D matrix $\hat{x}_{ij}^{\tau_d,w_\ell,\epsilon_k}$, as this is given by the relationship in Equation 8. The factor $\phi^{\tau_d,w_\ell,\epsilon_k}$, called the size correction factor, is calculated at the first iteration of the solution procedure at each count interval τ by the following formula:

$$\phi^{\tau_d,w_\ell,\epsilon_k} = \frac{1}{L} \left[\sum_{m \in M_O} \left(\frac{y_m^{\tau_d,w_\ell,\epsilon_k}}{\sum_{i \in N} a_{im}^{\tau_d,\tau,w_\ell,\epsilon_k} \hat{x}_i^{\tau_d,w_\ell,\epsilon_k}} \right) \right] \tau_d \in T \subseteq T_d \tag{10}$$

where L is the number of links for which flow measurements are available. The definition of time intervals τ_d in which the method described above is implemented is the same as that described for the relationship in Equation 7. The effect of each of the time-recursive mechanisms described above on the performance of model updating of dynamic O-D matrices can be determined by calculating three appropriate measures. The first one refers to the measurement of the O-D information loss, as defined above, and it is given by the mean absolute relative error of trip departure flows ($MARE_D$; in percent):

$$MARE_D = 100 \times \left(\frac{1}{F} \sum_{\tau_d \in T_d} \sum_{i \in N^+} \frac{\left| x_i^{\tau_d, w_\ell, \epsilon_k} \; \widehat{x}_i^{\tau_d, w_\ell, \epsilon_k} \right|}{\widehat{x}_i^{\tau_d, w_\ell, \epsilon_k}} \right) \quad (11)$$

where F denotes the number of origin zones in which trip departure flows are generated during T_d. The second measure concerning the O-D information loss refers to the relative root mean square error between the prior trip flows (used as an initial solution) and the estimated O-D trip flows ($RRMSE_{\text{O-D}}$; in percent):

$$RRMSE_{\text{O-D}} = 100 \times \left[\sqrt{\frac{\sum_{\tau_d \in T_d} \sum_{i,j \in N} \left(x_{ij}^{\tau_d, w_\ell, \epsilon_k} - \widehat{x}_{ij}^{\tau_d, w_\ell, \epsilon_k} \right)^2}{P}} \left(\frac{\sum_{\tau_d \in T_d} \sum_{i,j \in N} \widehat{x}_{ij}^{\tau_d, w_\ell, \epsilon_k}}{P} \right)^{-1} \right] \quad (12)$$

where P denotes the number of O-D pairs traversed by trip flows. The third performance measure refers to the metric difference between the measured link flows and the corresponding estimated flows, obtained after the convergence of the given solution algorithm. This measure is given by the relative root mean square error of link flows ($RRMSE_{\text{LINK}}$; in percent):

$$RRMSE_{\text{LINK}} = 100 \times \left[\sqrt{\frac{\sum_{m \in M_O} \left(y_m^{\tau, w_\ell, \epsilon_k} - \widehat{y}_m^{\tau, w_\ell, \epsilon_k} \right)^2}{L}} \left(\frac{\sum_{m \in M_O} y_m^{\tau, w_\ell, \epsilon_k}}{L} \right)^{-1} \right] \quad (13)$$

The relationships in Equations 11 to 13 imply that the smaller the values of $RRMSE_{\text{LINK}}$, $RRMSE_{\text{O-D}}$, and $MARE_D$ are, the better the model performance of the dynamic O-D matrix updating is.

CASE STUDY NETWORK

The present case study is for a real urban-scale road network, that of the central part of the greater Athens area network in Greece. It was modeled by using 151 assignment nodes, 256 assignment links, and a total of 1,936 O-D pairs; in other words, the assigned O-D matrix is 44 × 44. The original prior (seed) O-D matrix was obtained from O-D travel surveys and corresponds to the base year (ϵ_1) 1989. It represents the whole morning period (the count period T extends from 6:00 a.m. to 9:00 a.m.) and includes a total of about 35,000 vehicles to be loaded onto the network per hour. A description of the procedures involved in the automated traffic count data collection, which refers to about 10% of the total network links, is provided elsewhere (*1*). The link count data used in the study were collected on various weekdays (see below) in one decade (1991 to 2001). The solution algorithms of the ME- and GLS-based models were implemented in FORTRAN and run on a computer with a Pentium III processor. For all periods, the convergence of both algorithms was reached within a few seconds of central processing unit time after about 200 iterations for the ME-based model and 800 iterations for the GLS-based model.

In addition, any substantial changes concerning link characteristics and new roads and streets have been incorporated into the network data used at each time point in the present analysis. This treatment enables consideration of the impact of any major change in the road infrastructure on the dynamic O-D matrix updated across different months or years. Furthermore, attention has been given to ensure the compatibility or homogeneity of the link flow information and the arrangements of the O-D zones (network centroids) among the different study periods.

COMPUTATIONAL EXPERIENCE

Effects of Time-Recursive Mechanisms

Table 1 provides a comparison of the ME-based model performance measures obtained by implementation of the different time-recursive mechanisms to derive the prior O-D matrix in each period. This comparison includes three different time points within a decade: November 11, 1991; November 11, 1996; and November 12, 2001. These dates correspond to the second Monday of November of each of the years. Use of the smoothing time-recursive estimation procedure entails significant improvements in $RRMSE_{\text{LINK}}$ in comparison with the performance of the simple time-recursive procedure. However, this results in an increase in the average amount of O-D information lost, particularly in terms of $MARE_D$. On the contrary, the normalized smoothing time-recursive procedure demonstrates the largest improvements in all performance measures—$RRMSE_{\text{LINK}}$, $RRMSE_{\text{O-D}}$, and $MARE_D$—in terms of their averages as well as their period-specific (for each year) values.

TABLE 1 ME-Based Model Performance Measures for Different Methods of Treating Long-Term Flow Dynamics

Date	$RRMSE_{\text{LINK}}$ (%)	$RRMSE_{\text{O-D}}$ (%)	$MARE_D$ (%)
Time-recursive estimation procedure			
11/11/1991	19.65	279.71	2.50
11/11/1996	28.39	310.55	2.78
11/12/2001	18.46	185.47	2.18
Average	*22.17*	*258.58*	*2.49*
Smoothing time-recursive estimation procedure			
11/11/1991	19.65	279.71	2.50
11/11/1996	24.88	361.92	3.86
11/12/2001	2.37	136.33	1.70
Average	*15.63*	*259.32*	*2.69*
Normalized smoothing time-recursive estimation procedure			
11/11/1991	19.65	279.71	2.50
11/11/1996	23.86	242.29	2.08
11/12/2001	2.37	113.17	1.54
Average	*15.29*	*211.72*	*2.04*

Effect of Time Scale on Model Performance

Table 2 demonstrates the ME-based model performance measures obtained by implementing the normalized smoothing time-recursive procedure, which performed best in comparison with the other mechanisms, across different time scales. These include a 4-year time scale, in which the traffic flow information is updated every 2 years (i.e., on November 11, 1996; November 9, 1998; and November 13, 2000), and a 10-month time scale, in which the traffic flow information is updated every 5 months (i.e., on February 14, 2000; June 12, 2000; and November 13, 2000). These dates also correspond to the second Monday of November of the corresponding years or to the second Monday of each month examined in 2000. The prior O-D matrix referring to November 11, 1996, was based on the prior and resultant dynamic O-D matrices corresponding to November 11, 1991. The prior O-D matrix referring to February 14, 2000, was based on the prior and resultant dynamic O-D matrices corresponding to November 9, 1998. The results underline the significant impact of the frequency of updating of the flow information on the model performance measures. In general, $RRMSE_{\text{LINK}}$ follows an increasing trend as the time scale decreases from 10 years to 10 months. The results concerning the measures $MARE_D$ and $RRMSE_{\text{O-D}}$ are mixed, with the smallest average value for both of them occurring across the 4-year time scale.

TABLE 3 ME-Based Model Performance Measures Based on Long-Term Flow Information for Different Weekdays

Date	$RRMSE_{\text{LINK}}$ (%)	$RRMSE_{\text{O-D}}$ (%)	$MARE_D$ (%)
Flow information updated every 5 years			
11/5/1991	22.10	261.02	2.32
11/5/1996	25.11	207.13	2.14
11/6/2001	3.11	109.64	1.46
Average	*16.77*	*192.60*	*1.97*
Flow information updated every 2 years			
11/5/1996	25.11	207.13	2.14
11/3/1998	23.44	162.73	1.84
11/7/2000	19.88	132.81	1.61
Average	*22.81*	*167.56*	*1.86*
Flow information updated every 5 months			
2/1/2000	21.09	184.82	1.89
6/6/2000	32.31	195.45	2.18
11/7/2000	16.63	191.67	2.04
Average	*23.35*	*190.65*	*2.04*

Comparison of Different Days of Week

Table 3 shows the ME-based model performance measures obtained by implementing the normalized smoothing time-recursive procedure across three time scales—the 10-year, 4-year, and 10-month time scales—on a weekday. The weekday is the first Tuesday of November of the corresponding years or to the first Tuesday of each month examined within 2000. The temporal trends in $RRMSE_{\text{LINK}}$ presented in Table 3 for Tuesdays are the same as those for Mondays presented in Tables 1 and 2 across the different time scales. The corresponding average magnitudes of $RRMSE_{\text{LINK}}$ in the 2 weekdays are found to be close to each other, particularly for the 4-year time scale. $RRMSE_{\text{LINK}}$ remained below 25% for all weekdays and time scales examined. $RRMSE_{\text{O-D}}$ and $MARE_D$ for Tuesdays (Table 3) have the same trends as those for Mondays (Tables 1 and 2) across the different time scales. The average magnitudes of $RRMSE_{\text{O-D}}$ and $MARE_D$ are smaller on Tuesday than on Monday, while the differences in $RRMSE_{\text{O-D}}$ and $MARE_D$ among the two weekdays (Mondays and Tuesdays) are the smallest across the 4-year time scale. These results may suggest that variations in travel demand across successive days of the week do not have a considerable impact on the overall performance of the model, particularly in terms of the reproducibility of the measured link flows.

TABLE 2 ME-Based Model Performance Measures for Different Time Scales of Long-Term Flow Information

Date	$RRMSE_{\text{LINK}}$ (%)	$RRMSE_{\text{O-D}}$ (%)	$MARE_D$ (%)
Flow information updated every 2 years			
11/11/1996	23.86	242.29	2.08
11/9/1998	24.39	172.17	1.89
11/13/2000	18.49	136.17	1.75
Average	*22.25*	*183.54*	*1.90*
Flow information updated every 5 months			
2/14/2000	20.98	179.18	1.93
6/12/2000	34.80	263.00	2.38
11/13/2000	17.19	183.02	1.97
Average	*24.32*	*208.40*	*2.09*

Comparison of ME- and GLS-Based Models

Table 4 shows the various performance measures obtained by implementation of the GLS-based model to update dynamic O-D matrices. The normalized smoothing time-recursive procedure for estimation of the prior O-D matrix in each time period was also used for the three different time scales. The $RRMSE_{\text{LINK}}$ produced by the GLS-based model is greater than that produced by the ME-based model for the lowest frequency of flow information updating (5 years). On the contrary, for higher frequencies of flow information updating (i.e., 2 years and 5 months) the value of $RRMSE_{\text{LINK}}$ from the GLS-based model is smaller than that from the ME-based model. Despite the small magnitude of the differences in the values of $RRMSE_{\text{LINK}}$, the two models demonstrate considerably different performances in terms of the measures of $RRMSE_{\text{O-D}}$ and $MARE_D$, in favor of the ME-based model. This difference is larger, in relative terms, for $MARE_D$ than for $RRMSE_{\text{O-D}}$. Thus, the overall performance of the ME-based model can be considered superior to that of the GLS-based model across the different time scales, since the latter model leads to significantly greater O-D information loss, whereas it produces link flow solution results basically similar to those of the former model.

But the temporal evolution of the average magnitude of the GLS-based model performance measures indicates the same trends across

TABLE 4 GLS-Based Model Performance Measures Based on Long-Term Flow Information for Different Time Scales

Date	$RRMSE_{\text{LINK}}$ (%)	$RRMSE_{\text{O-D}}$ (%)	$MARE_D$ (%)
Flow information updated every 5 years			
11/11/1991	23.10	272.67	3.16
11/11/1996	21.33	278.33	3.55
11/12/2001	4.27	170.56	2.64
Average	*16.23*	*240.52*	*3.11*
Flow information updated every 2 years			
11/11/1996	21.33	202.72	3.00
11/9/1998	22.02	207.12	3.02
11/13/2000	18.23	178.29	2.91
Average	*20.53*	*196.04*	*2.97*
Flow information updated every 5 months			
2/14/2000	19.37	278.33	3.55
6/12/2000	33.38	219.31	3.12
11/13/2000	17.14	189.79	2.77
Average	*23.30*	*229.14*	*3.14*

the different time scales with respect to those corresponding to the ME-based model. Figure 1 illustrates the percent improvements in the magnitudes of the various performance measures calculated after implementation of the proposed time-recursive estimation procedure for each model in comparison with the corresponding magnitudes calculated when the base-year O-D matrix was dynamically loaded onto the network at the period of the last point in time of each time scale. For the base-year O-D matrix, the same prior trip departure time pattern was adopted with respect to that calculated on the basis of the implementation of the proposed time-recursive procedure for each study period. In addition to the original prior O-D matrix (see above), referred to as the normal seed O-D matrix, an alternative seed matrix is also used. The elements of the alternative matrix, referred to as the poor seed O-D matrix, are derived by random perturbation of the corresponding elements of the normal seed within the range 0.5 to 1.5.

Implementation of the proposed time-recursive procedure provides significant improvements in all model performance measures. These improvements are considerably larger for the normal seed and exceed 40% for $RRMSE_{\text{LINK}}$ (Figure 1*a*), 50% for $RRMSE_{\text{O-D}}$ (Figure 1*b*), and 25% for $MARE_D$ (Figure 1*c*). Use of the poor seed results in reductions in the improvements of all model performance measures in comparison with those obtained by use of the normal seed. The largest improvement reductions correspond to the 10-year time scale, that is, when a low frequency of flow information updating, such as 5 years, is used. This outcome suggests that the extent to which the quality of the seed O-D matrix affects the ability of the proposed time-recursive procedure to update the dynamic O-D matrix depends on the frequency with which the flow information is updated. More specifically, the smaller the time scale is, the less sensitive the results of the estimation procedure are with respect to the quality of the seed matrix. It should also be mentioned that the increase in the amount of information used in the time series analysis of the traffic counts cannot ensure a continuous improvement of model performance. This behavior can be explained by the existence of several exogenous factors other than

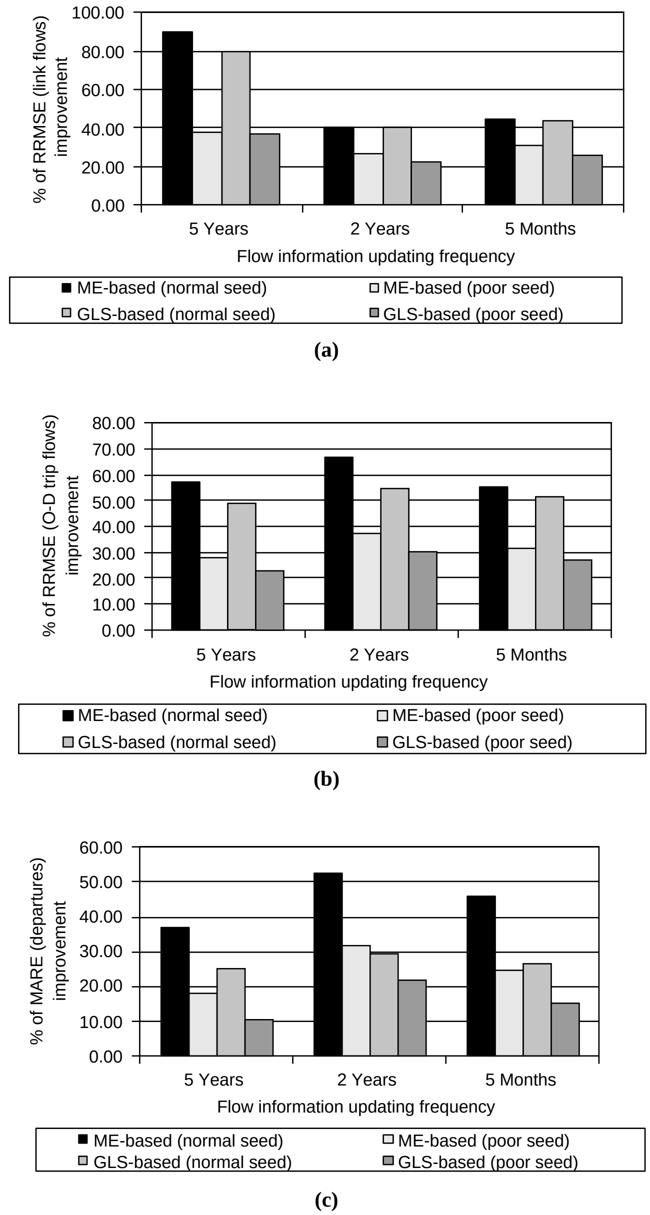

FIGURE 1 Improvements in ME- and GLS-based model performance measures compared with estimates that resulted from the base-year O-D information by using the normal and poor seed matrices: (*a*) percent improvement of $RRMSE_{\text{LINK}}$; (*b*) percent improvement of $RRMSE_{\text{O-D}}$; (*c*) percent improvement of $MARE_D$.

traffic congestion that influence the long-term evolution of trip demand (see above).

The ME-based model shows the largest improvements in the estimates based on the base-year O-D matrix, as reflected in all performance measures, in comparison with the improvements produced by the GLS-based model, by use of both the normal and the poor seed matrices. The largest differences in performance improvement between the two models are observed in $MARE_D$, whereas the smallest differences are observed in $RRMSE_{\text{LINK}}$. The longest period of flow information updating, 5 years, is associated with the largest reduction in $RRMSE_{\text{LINK}}$ with respect to the $RRMSE_{\text{LINK}}$ produced by dynamically assigning the base-year prior O-D matrix on November

12, 2001. On the contrary, the shorter periods of flow information updating—and, in particular, that of 2 years—are associated with the greatest reductions in O-D information loss, as expressed by $RRMSE_{\text{O-D}}$ and $MARE_D$.

The better performance of the ME-based model can be interpreted by the structure of its objective function (Equation 1) in comparison with structure of the objective function of the GLS-based model (Equation 4). More specifically, the solution procedure followed by the ME-based model relies on the ratio between the estimated and prior trip departure rates and, hence, the proportion of the counted link flows corresponding to the current ϵ_k and the previous ϵ_{k-1} points in time. On the contrary, the solution procedure followed by the GLS-based model relies on the absolute difference between the counted flows and the estimated flows, which depends on the current prior O-D matrix and the counted flows of the previous point in time. Given that the distance between two successive time points may correspond to several months or years, the GLS-based model cannot appropriately consider significant changes in traffic flows on low-volume links. On the other hand, the ME-based model intrinsically considers the rates of longer-term changes in traffic flows on links traversed by volumes of different magnitudes. In this way, the ME-based model can provide a more realistic representation of the impacts of the long-term evolution of traffic congestion on the updated dynamic O-D matrix than the GLS-based model can.

Statistical Analysis of Long-Term Dynamic O-D Matrices

The statistical analysis includes the calculation of the lower and upper 95th percentile confidence intervals (referred to here as *LCI* and *UCI*, respectively) for $MARE_D$ between two successive points in time (months or years) within a given time scale on the basis of the *t*-distributions, as described previously (*9*). The estimated values of *LCI* and *UCI* are shown in Table 5. The differences in $MARE_D$ values for the periods studied within the time scales of 10 years and 10 months were found to be statistically significant since the $MARE_D$ values for the time scales of 10 years and 10 months were found to be either lower OR higher than the corresponding *LCI*s and *UCI*s. On the contrary, the $MARE_D$ values for the periods within the 4-year time scale were not found to be significantly different from each other, since they were within the interval defined by the corresponding *LCI* and *UCI* limits. In addition, the *LCI*s and *UCI*s corresponding to the periods within the 4-year time scale were generally found to be smaller than those corresponding to the periods within the 10-year and 10-month time scales.

The significant differences in $MARE_D$ values for an information updating period of 5 years verify that the cells of the updated dynamic O-D matrix represent the result of long-term changes that have systematic and permanent impacts on trip distribution. Nonetheless, the insignificant differences in $MARE_D$ for smaller updating periods (2 years) probably indicate that the changes that occurred within such time scales are not adequate to result in time-persistent and systematic effects on trip distribution. The significant differences in $MARE_D$ values across the 10-month time scale indicate that, despite the relatively short period of information updating (i.e., 5 months), seasonal traffic variations underlying different months of the year can have a systematic impacts on the loss of information incorporated in the prior dynamic O-D matrix.

On the basis of the results presented above, the prior dynamic O-D matrix estimated for Monday, February 14, 2000 (w_ℓ), based on the prior and the resultant dynamic O-D matrices corresponding to November 9, 1998, was used to update the dynamic O-D matrices corresponding to the morning period of the following two Mondays: $w_{\ell+1}$ = February 21, 2000, and $w_{\ell+2}$ = February 28, 2000, of the same month (February). Statistical analysis of the $MARE_D$ values corresponding to these three Mondays indicates that their differences are not significantly different from each other. Table 5 provides the corresponding *LCI* and *UCI* values of $MARE_D$ for each morning period examined in February. Table 6 shows the performance measures of the dynamic O-D matrices updated on Mondays of February 2000. These results probably verify the random nature of the variations between the cells of dynamic O-D matrices of the same day of the week and month or, possibly, season.

Furthermore, Table 6 presents the model performance measures corresponding to the second and third Tuesdays of the same month: $w_{\ell+1}$ = February 8, 2000, and $w_{\ell+1}$ = February 15, 2000. The latter estimates were obtained by dynamically assigning the prior dynamic O-D matrix estimated for Tuesday (w_ℓ = February 1, 2000) on the basis of the prior and resultant dynamic O-D matrices corresponding to November 3, 1998. All performance measures across and between the two weekdays (i.e., Monday and Tuesday) present the smallest differences in comparison with the results for all other (longer) time scales. These results signify that when the proposed

TABLE 5 *LCI*s and *UCI*s of Differences in $MARE_D$ Between Periods of Different Time Scales

Date	*LCI*	*UCI*
11/11/1991–11/11/1996	2.11	2.44
11/11/1996–11/12/2001	1.64	1.98
11/11/1996–11/9/1998	1.79	2.18
11/9/1998–11/13/2000	1.58	2.09
2/14/2000–6/12/2000	1.95	2.34
6/12/2000–11/13/2000	1.99	2.36
2/14/2000–2/21/2000	1.82	2.06
2/21/2000–2/28/2000	1.85	2.04

TABLE 6 ME-Based Model Performance Measures for Morning Periods of Days of Successive Weeks of Month

Date	$RRMSE_{\text{LINK}}$ (%)	$RRMSE_{\text{O-D}}$ (%)	$MARE_D$ (%)
Mondays			
2/21/2000	22.89	186.42	1.95
2/28/2000	21.52	177.77	1.88
Tuesdays			
2/8/2000	22.44	182.34	1.90
2/15/2000	21.73	188.15	1.98

time-recursive procedure is implemented within a suitable period of information updating, it can produce an improved prior O-D matrix that may significantly enhance the subsequent updating of dynamic O-D matrices corresponding to a series of days of the week.

CONCLUSIONS

The potential of the normalized smoothing time-recursive estimation procedure to provide an improved initial solution in the form of a prior dynamic O-D matrix was demonstrated. Therefore, this procedure would increase the levels of performance of various models of dynamic O-D matrix synthesis. The smallest reductions in O-D information loss were obtained by implementation of the proposed time-recursive procedure by use of a low frequency of flow information updating, such as 5 years, or a high frequency of flow information updating, such as 5 months. The greatest reduction was obtained when an information updating frequency of 2 years was used. Moreover, adoption of the longest period of information updating—5 years—was found to result in the greatest improvements in the flow solution accuracy (in terms of updating of O-D matrices) that, when dynamically assigned to the network, would produce time-dependent link flows close to the measured ones. Use of the ME- and GLS-based models led to different magnitudes of improvement in the related performance measures, with use of the former model resulting in the greatest benefits.

The effect of the day of the week on the performance measures was not found to be considerable. However, application of the proposed time-recursive procedure across different months of the year was shown to affect the model's performance significantly. This result may suggest the importance of developing seasonal O-D matrices over past count periods or for the base year, for example, by conducting travel surveys across different seasons (e.g., winter and summer) of the year. These seasonal O-D matrices should be selectively updated over regular time horizons, for example, every 1 or 2 years, to capture factors that have significant impacts on the temporal evolution of trip distributions in urban areas, such as changes in housing or job location.

REFERENCES

1. Tsekeris, T., and A. Stathopoulos. Real-Time Dynamic Origin-Destination Matrix Adjustment with Simulated and Actual Link Flows in Urban Networks. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1857,* TRB, National Research Council, Washington, D.C., 2003, pp. 117–127.
2. Ziliaskopoulos, A. K., and S. Peeta. Review of Dynamic Traffic Assignment Models. *Networks and Spatial Economics,* Vol. 1, 2001, pp. 233–267.
3. Sivanandan, R., D. Nanda, and E. D. Arnold, Jr. Method to Enhance Performance of Synthetic Origin-Destination Trip-Table Estimation Models. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1676,* TRB, National Research Council, Washington, D.C., 1999, pp. 124–135.
4. Chang, G. L., and J. Wu. Recursive Estimation of Time-Varying Origin–Destination Flows from Traffic Counts in Freeway Corridors. *Transportation Research,* Vol. 28B, 1994, pp. 141–160.
5. Willumsen, L. G. Estimating Time-Dependent Trip Matrices from Traffic Counts. *Proc., 9th International Symposium on Transportation and Traffic Theory* (J. Volmüller and R. Hamerslag, eds.), VNU Science Press, Utrecht, Netherlands, 1984, pp. 397–411.
6. Van Aerde, M. *QUEENSOD Release 2.0. User's Guide: Estimating Dynamic Origin–Destination Traffic Demands from Link Flow Counts.* Transportation Systems Research Group, Queen's University, and M. Van Aerde and Associates, Ltd., Kingston, Ontario, Canada, 1997.
7. Stathopoulos, A., J. Polak, A. Tillis, S. Mitropoulos, and B. Ryan. Parking Management and Control. In *Advanced Telematics in Road Transport, Vol. 1. Proc., DRIVE Conference,* Brussels, Belgium, Elsevier, Oxford, United Kingdom, 1991, pp. 752–777.
8. *Parking Management, Control and Information System. Final Report.* The PARCMAN Consortium, Athens, Greece, 1992.
9. Dixon, M. P., and L. R. Rilett. Real-Time OD Estimation Using Automatic Vehicle Identification and Traffic Count Data. *Computer-Aided Civil and Infrastructure Engineering,* Vol. 17, No. 1, 2002, pp. 7–21.