Quasi-Continuous Dynamic Traffic Assignment Model

BRUCE N. JANSON AND JUAN ROBLES

Several variants of combined dynamic travel models in discrete time with dynamic user equilibrium or system optimality as the assignment objective have been presented recently. This modeling approach is converted into quasi-continuous time, which enables two key model improvements: (*a*) traffic volumes are spread over time intervals in continuous time, allowing trips to be split among successive time intervals, and (*b*) the first-in first-out ordering of trips between all zone pairs is more precisely maintained. The means by which capacity losses are approximated on upstream links caused by spillback queueing from oversaturated links and accidents are also described. Trips are assumed to have scheduled departure times and variable arrival times, but notational variations allowing other model forms are briefly mentioned. Application of this model to a Denver-area network with comparison of results to observed speeds and volumes is described elsewhere.

Among several key issues addressed by researchers developing dynamic traffic models over the past 20 years, three critical issues are capabilities of the model to (*a*) fractionally split trip flows among multiple paths at any road juncture, (*b*) validly maintain the first-in first-out (FIFO) ordering of trips between all zones pairs, and (*c*) account for spillback queueing effects caused by incidents and oversaturated links. This paper formalizes the inclusion of these three concerns in a dynamic user-equilibrium (DUE) formulation. Robles and Janson (*1*) apply this model to a Denver-area network covering about 100 mi^2(260 km^2) with comparisons of results with observed speeds and volumes.

DUE formulations presented by Janson (*2,3*) use 0-1 variables called node time intervals to track trips across the network in both time and space (i.e., to identify whether trips departing zone *r* in time interval *d* cross node *i* in interval *t*). However, trips departing within a given time interval from any node or zone do not all depart at a single point in time but instead depart as a uniform trip rate over that time interval. This paper improves these earlier DUE formulations by representing time more continuously. Herein, the integer node time intervals are used to compute "trip flow fractions" (i.e., the fraction of trips departing zone *r* in time interval *d* to cross node *i* in time interval *t*). All trip flows are tracked through the network in continuous rather than discrete time. This modeling revision improves its (*a*) spillback queueing effects and dynamic link capacity adjustments, (*b*) FIFO trip ordering between all origin-destination (O-D) pairs, and (*c*) link volume and speed transitions between time intervals.

In DUE, the full analysis period of several hours is sliced into shorter intervals. Each trip has a known departure or arrival time (but not both) and its corresponding trip-end zones. Because travel times are variable, times of departure and arrival cannot both be fixed for any trip. DUE1 assumes known departure times for trips from each zone but only total arrivals to each zone over the full analysis period. DUE2 assumes known arrival times of trips to each zone, but only total departures from each zone over the full analysis period. Janson (*4*) formulates DUE2 similarly to DUE1, except for numerous notational changes to make node time intervals and node-to-destination travel times on the basis of destination zones and arrival times. Janson (*5*) describes a general model including both trip types, and Janson and Robles (*6*) combined route choice with departure or arrival time choice in DUEA.

DUE1 as formulated in this paper with scheduled departures is defined as follows:

> Given a set of zone-to-zone trip tables containing the number of vehicle trips departing from each origin zone in successive time intervals of 1 to 10 min each, and also the destination zone but not the arrival time of each trip, determine the volume of vehicles on each link in each time interval such that, for each O-D pair of zones, no path has a lower travel time than any used path for trips departing within a given time interval.

This DUE condition for fixed departure times, as derived by Janson (*3*), is a temporal generalization of Wardrop's (*7*) condition for static user equilibrium (SUE). Here, the term departing can be replaced by arriving for cases with variable departure times and scheduled arrival times. As an outcome of equal travel time paths, trips between the same O-D pair with the same departure time also have equal arrival times. At equilibrium, each trip arrives at its destination on an equal travel time path that departs from its origin within the same time interval.

DUE1 presented here is quasi-continuous in that link volumes are split fractionally between discrete time intervals from which speeds are calculated. In comparison, Friesz et al. (*8*), Wie (*9*), and Wie et al. (*10*) present optimal control theory formulations of dynamic traffic assignment in continuous time for which the equilibrium condition is that no used path between any two nodes has a higher travel time than any other path at any instant. Path choice according to time-of-departure conditions or en route revisions, or both, according to updated conditions in each time interval lead to models in which complete O-D paths used by trips can have unequal travel times for any given departure time.

Wie (*11*) and Ran et al. (*12,13*) refine and extend optimal control models to include elastic demand and departure time choice in user equilibrium or system optimal forms. Friesz et al. (*14*) formulate the simultaneous route choice and departure time problem in continuous time as a variational inequality. Although proposed solution procedures are prohibitive, these papers address important considerations about the behavioral assumptions of alternative model forms.

Department of Civil Engineering, University of Colorado, Campus Box 113, P.O. Box 173364, Denver, Colo. 80217.

DYNAMIC USER EQUILIBRIUM WITH SCHEDULED DEPARTURES

Dynamic user equilibrium with scheduled departures (DUE1) can be stated equivalently in terms of path flows, but the link flow form shown here does not implicitly assume complete enumeration of all paths between zone pairs. Turn movements at each intersection are represented by separate links at each node. The exact form of each link's impedance function can be specific to the intersection or link type. The O-D trip matrix can be developed from traffic counts or from survey data and trip distribution models. In DUE1 stated by Equations 1 through 4, link lengths are computed on the basis of monotonically nondecreasing impedance functions dependent on each link's volume in each time interval.

- Upper Problem (UP)

Minimize

$$\sum_{ij\in K}\sum_{t\in T}\int_0^{x_{ij}^t} f_{ij}^t(w)\,dw \tag{1}$$

subject to

$$x_{ij}^t = \sum_{r\in Z}\sum_{d\leq t} v_{rij}^d\ \phi_{ri}^{dt} \qquad \text{for all } ij\in K,\ t\in T \tag{2}$$

$$q_{rn}^d = \sum_{t\geq d}\left[\sum_{in\in K} v_{rin}^d\,\phi_{ri}^{dt} - \sum_{nj\in K} v_{rnj}^d\ \phi_{rn}^{dt}\right] \qquad \text{for all } n\in N,\ r\in Z,\ d\in T \tag{3}$$

$$v_{rij}^d\ \phi_{ri}^{dt} \geq 0 \qquad \text{for all } r\in Z,\ ij\in K,\ d\in T,\ t\in T \tag{4}$$

$$\Delta b_{ri}^d = b_{ri}^d - b_{ri}^{d-1} \qquad \text{for all } r\in Z,\ i\in N,\ d\in T, \text{ and } b_{ri}^0 = b_{ri}^1 - \delta t \tag{5}$$

and, for all $r \in Z$, $i \in N$, $d \in T$, $t \in T$ in Equations 6a through 6c

$$\phi_{ri}^{dt-\text{k}} = \left(\min\{1, [\text{b}_{ri}^d - (\text{t}-1)\Delta\text{t}]/\Delta b_{ri}^d\}\right)\alpha_{ri}^{dt} \qquad \text{for k} = 0 \tag{6a}$$

$$\phi_{ri}^{dt-\text{k}} = [\ \Delta t/\Delta b_{ri}^d]\alpha_{ri}^{dt} \qquad \text{for all } k > 0 \text{ for which } b_{ri}^{d-1} - (t-1-k)\Delta t \leq 0 \tag{6b}$$

$$\phi_{ri}^{dt-\text{k}} = \left(\max\{0, [\Delta t(t-k) - b_{ri}^{d-1}]/\Delta b_{ri}^d\}\right)\alpha_{ri}^{dt} \qquad \text{for min } k \text{ for which } b_{ri}^{d-1} - (t-1-k)\Delta t > 0 \tag{6c}$$

where all $\{\alpha_{ri}^{dt}\}$ and $\{b_{ri}^d\}$ are optimal for:

- (Lower problem):

Maximize

$$\sum_{r\in Z}\sum_{i\in N}\sum_{d\in T} b_{ri}^d \tag{7}$$

subject to

$$\alpha_{ri}^{dt} = (0,1) \qquad \text{for all } r\in Z,\ i\in N,\ d\in T,\ t\in T \tag{8a}$$

$$\sum_{t\in T}\alpha_{ri}^{dt} = 1 \qquad \text{for all } r\in Z,\ i\in N,\ d\in T,\ t\in T \tag{8b}$$

$$[b_{ri}^d - t\Delta t]\ \alpha_{ri}^{dt} \leq 0 \qquad \text{for all } r\in Z,\ i\in N,\ d\in T,\ t\in T \tag{9a}$$

$$[b_{ri}^d - (t-1)\Delta t]\ \alpha_{ri}^{dt} \geq 0 \qquad \text{for all } r\in Z,\ i\in N,\ d\in T,\ t\in T \tag{9b}$$

$$b_{rr}^d = d\Delta t \qquad \text{for all } r\in Z,\ d\in T \tag{10}$$

$$b_{ri}^d = \max[e_{ri}^d, b_{ri}^{d-1} + h\Delta t] \qquad \text{for all } r\in Z,\ i\in N,\ d\in T, \text{ and } b_{ri}^0 = b_{ri}^1 - \Delta t \tag{11}$$

$$\theta_{ri}^{dt} = [(b_{ri}^d - (t-1)\Delta t)/\Delta t]\ \alpha_{ri}^{dt} \qquad \text{for all } r\in Z,\ i\in N,\ d\in T,\ t\in T \tag{12}$$

$$g_{rij}^{dt} = [\theta_{ri}^{dt} f_{ij}^t(x_{ij}^t) + (1-\theta_{ri}^{dt}) f_{ij}^p(x_{ij}^p)]\ \alpha_{ri}^{dt} \qquad \text{for all } r\in Z,\ ij\in K,\ d\in T,\ t\in T,\ p = t-1 \tag{13}$$

$$\left(e_{rj}^d - \max\{b_{ri}^d, (t-1)\Delta t + \Delta f_{ij}^{tp}\}\right)\alpha_{ri}^{dt} \leq g_{rij}^{dt}\ \alpha_{ri}^{dt} \qquad \text{for all } r\in Z,\ ij\in K,\ d\in T,\ t\in T,\ p=t-1,\ \Delta f_{ij}^{tp} = f_{ij}^p(x_{ij}^p) - g_{rij}^{dt} \tag{14}$$

where

N = set of all nodes;

Z = set of all zones (i.e., trip-end nodes);

K = set of all links (directed arcs);

Δt = duration of each time interval (same for all t);

T = set of all time intervals in the full analysis period (e.g., 18 intervals of 10 min each for 3-hr peak-period assignment);

x_{ij}^t = number of vehicle trips between all zone pairs assigned to link ij in time interval t (variable);

v_{rij}^d = number of vehicle trips departing zone r in time interval d assigned to link ij at some time (variable);

$f_{ij}^t(x_{ij}^t)$ = average travel impedance on link ij in time interval t (variable);

q_{rn}^d = number of vehicle trips from zone r to node n departing in time interval d via any path; 0 for any node $n \notin Z$ (variable);

e_{ri}^d = time (including $d\Delta t$) at which last trip departing zone r in time interval d crosses node i via its shortest path less FIFO delay time at node i (variable);

b_{ri}^d = time(including $d\Delta t$) at which last trip departing zone r in time interval d crosses node i via its shortest path (variable);

α_{ri}^{dt} = 0–1 variable indicating whether last trip departing zone r in time interval d crosses node i in time interval t (henceforth called a "node time interval") (0 = no; 1 = yes) (variable);

ϕ_{ri}^{dt} = fraction of all trips departing zone r in time interval d to cross node i in time interval t (henceforth called a trip flow fraction) (variable);

θ_{ri}^{dt} = fraction of time interval Δt into time interval t that last trip departing zone r in time interval d crosses node i (variable);

g_{rij}^{dt} = "average" travel time on link (i,j) of last trip departing zone r in time interval d adjusted for time into interval t versus $t-1$ that this trip enters link (variable); and

h = minimum fraction of time interval that trips departing zone r in time interval d must follow trips departing in time interval $d - 1$.

Equation 2 defines total flow on link ij in time interval t to be the sum of flows departing any zone r in any time interval $d \leq t$ using link ij in time interval t to formulate the objective function as given by Equation 1. Conservation of flow Equation 3 constrains inflow minus outflow at each node and zone in each time interval to sum to the proper trip departure totals in each time interval between each O-D pair, and Equation 4 requires all link volumes to be nonnegative. DUE1 requires nonlinear mixed-integer constraints with "node time intervals" and "trip flow fractions," indicating the time intervals in which trips from each origin cross each node so as to ensure temporally continuous trip paths and to spread the trips more continuously in time over these intervals.

A node time interval α_{ri}^{dt} differs from a trip flow fraction ϕ_{ri}^{dt} as follows. A node time interval is a 0 – 1 variable that indicates the time interval in which trips departing zone r in time interval d "last cross" node i. Each node time interval acts as an "if-then" operator to activate or deactivate certain constraints as needed. A node time interval applies to the last vehicle of each departure time interval, with the time of the last vehicle departing zone r in time interval d to cross node i, via its shortest path given by b_{ri}^{d}. The difference between node i crossing times of "last" vehicles departing in successive time intervals, defined as Δb_{ri}^{d} in Equation 5, is used in Equation 6*a* through *c* to determine the temporal spread of dispersion of trips crossing node i from the same origin.

In Figure 1 explained later, trips v_{rij}^{d} departing zone r in time interval d using link (i,j) are uniformly spread temporally over the vertical difference between the times when "last" vehicles in these two streams pass a given node. In contrast to previous papers, the trip variable v_{rij}^{d} does not have a time superscript t because it is always joined by the variable ϕ_{ri}^{dt}. Arrays $\{\alpha_{ri}^{dt}\}$ and $\{b_{ri}^{d}\}$ used in Equations 5 and 6 are passed in from the lower problem. These equations are not solved simultaneously with the upper problem, because they provide only exogenous inputs to the upper problem and have no variables that vary while solving the upper problem.

One may ask whether shorter time intervals could be used instead of trip flow fractions. Shorter time intervals add to the computational burden of the problem for several reasons. First, more time intervals to span the same analysis period require many more calculations and storage of floating point values than using trip flow fractions. Second, FIFO trip ordering in this formulation is best maintained if link lengths remain well below the time interval length as explained shortly. Using shorter time intervals may require dividing links into shorter links, thus increasing the computational burden of the problem in all three dimensions (nodes, links, and time intervals).

Equations 8 and 9 compute node time intervals that define temporally continuous trip paths with the zone-to-node travel times. Equation 8a defines each α_{ri}^{dt} term to be 0 or 1, which defines the last time interval t in which arcs incident from node i are used by trips departing zone r in time interval d. For trips departing zone r in time interval d, any link ij incident from node i can only be "last" used (if at all) in time interval t, and first used in the last time interval for trips departing in the previous interval d-1. Equation 8*a* allows only one interval t in which trips departing zone r in interval d can last cross node i.

According to Equations 8 and 9, links are traversed within time intervals that trip paths cross their tail nodes. For 5-min intervals, Interval 1 begins at 0, Interval 2 begins at 5 min, Interval 3 begins at 10 min, and so on. If the travel time from zone r to node i is within $t\Delta t$, then α_{ri}^{dt} is 1 because these trips must cross node i in time interval t. Because of tracking last vehicles, if any path crosses a node at the exact start of a time interval (to the degree of floating point precision being used), then the solution algorithm sets $\alpha_{ri}^{dt} = 1$, but all trips from that origin for that departure interval will be assigned to the link over previous time intervals.

In Equation 10, b_{rr}^{d} (equal to the start time of the "last" vehicle departing zone r in time interval d) is set to $d\Delta t$ to correctly set the clock to the end of each time interval and also prevents the (LP) maximization from having an infinite solution. A second subtle change from previous papers is that b_{ri}^{d} includes $d\Delta t$ so that it rep-

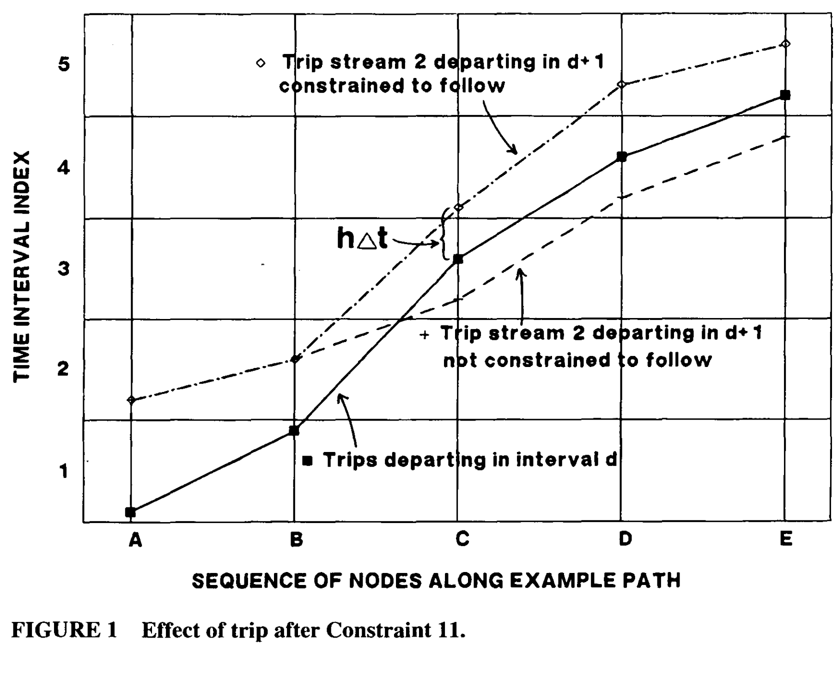

FIGURE 1 Effect of trip after Constraint 11.

resents "clock time" from the start of the entire analysis period rather than travel time from the time of departure. Using "clock time" enables a clearer accounting of time in this quasi-continuous model. Although α_{ri}^{dt} can never equal 1 when $t=1$, trips departing in interval 1 are uniformly distributed over the previous time span $b_{ri}^{1} - \Delta t$ at each node of the network such that some trip volumes are still assigned in interval 1.

Equations 11 through 14 impose FIFO trip ordering between all O-D pairs according to their travel times in successive time intervals as explained next. Vehicles are assumed to make only one-for-one (or zero-sum) exchanges of traffic positions along any link, which is acceptable and expected in aggregate traffic models.

Equation 11 is a vehicle following constraint that prevents later trips from "getting too close" to trips departing earlier from the same zone in successive time intervals so as to prevent these trips from bunching. The value h is the fraction of a time interval that the last trip departing from zone r in time interval d must follow the last trip departing from zone r in interval $d - 1$. Note that when solving for b_{ri}^{d} on the left side of Equation 11, b_{ri}^{d-1} on the right side is held fixed.

Figure 1 illustrates the effect of Constraint 11 on two trip streams (1 and 2) traversing the same series of nodes and departing from the same zone in intervals d and $d + 1$, respectively. The node sequence (A,B,C,D,E) denotes a series of links along the trip path. A trip stream consists of trips with the same origin and departure time interval. At Node A, Trip Stream 2 is $0.8\Delta t$ behind Trip Stream 1. The two trip streams traverse Node A in different time intervals and thus have different travel times for Link A,B. At Node B, the separation between the trip streams has decreased to $0.6\Delta t$. The two trip streams traverse Node B in different time intervals and have different travel times for Link B,C. Without Constraint 11, Stream 2 would pass Stream 1 and traverse Node C earlier. With Constraint 11, Stream 2 is forced to traverse Node C at least $h\Delta t$ behind Stream 1, where $h = 0.5$ in this figure.

Allowing that many other paths include the node sequence (A,B,C,D,E), it can be deduced from Figure 1 that Constraint 11 is less likely to be binding in networks with shorter arc lengths relative to Δt. Hence, this representation works best for networks in which most arc lengths have free-flow travel times less than 20 percent of the interval duration and in which loadings on the network do not cause arc lengths to exceed $h\Delta t$. Example runs revealed the solution algorithm explained later to converge more easily if time-varying travel demands do not cause arc travel times to exceed these bounds.

Although Constraint 11 prevents successive trip streams from passing each other, the exact specification of h can be improved in further research. Reasonable values of h lie between $0.3\Delta t$ and $0.7\Delta t$, but the exact value of h depends on traffic densities of arcs incident to a node in each time interval. However, the value of h must lie between 0 and 1. If $h = 0$, a trailing trip stream can completely overlay (but not overtake) a leading trip stream so that the two streams become coincident, which is not realistic. If $h = 1$, then trailing trips can never partly "gain ground on" leading trips, and later departing trips can never have lower travel times than earlier departing trips. Because Constraint 11 applies to trips departing in intervals $d + 1$ and $d + 2$, trips departing from zone r in interval $d + 2$ must follow trips departing in interval d by at least $2h\Delta t$ at any node. Constraint 11 also applies to all nodes in the network regardless of whether any trips departing from zone r in intervals d and $d + 1$ actually cross node i.

Because Equation 11 does not insure FIFO trip ordering between all O-D pairs, Equations 12 through 14 are also required. Equations 12 and 13 compute an average travel time on link (i,j) of the "last" trip departing zone r in time interval d adjusted for the time into interval t versus $t - 1$ that this trip enters the link. Equations 12 and 13 "smooth out" speed transitions between time intervals in a "quasi-continuous" manner so that vehicle speeds do not abruptly change if they enter links just split seconds before or after a time interval change. When finding shortest paths in the upper problem, equations 12 and 13 are also used to calculate link travel times based on when trips enter links so as to be consistent with how paths are found in the lower problem.

Even without Equation 14, this improvement to the model eliminates FIFO violations for trips between all O-D pairs unless a link's travel time exceeds a full time interval (and even for most of those cases) as required by the following inequality.

$$\theta_{ri}^{dt} f_{ij}^{t}(x_{ij}^{t} + (1 - \theta_{ri}^{dt}) f_{ij}^{p}(x_{ij}^{p}) + \theta_{ri}^{dt}\Delta t \geq f_{ij}^{p}(x_{ij}^{p})$$
$$\text{for all } r \in Z,\ ij \in K, d \in T,\ t \in T, p = t - 1 \quad (15)$$

which simplifies to

$$f_{ij}^{t}(x_{ij}^{t}) + \Delta t \geq f_{ij}^{p}(x_{ij}^{p}) \qquad \text{(q.e.d.)} \quad (16)$$

If no link travel times exceed Δt, then the left side of Equation 14 could be written more simply as $(e_{ij} - b_{ri}^{d})\,\alpha_{ri}^{dt}$, but Equation 14 is written as shown in DUE1 to prevent FIFO violations in cases where link travel times exceed Δt if needed.

Figure 2 illustrates the effect of Constraint 14 on two trip streams (1 and 2) between any two zones traversing the same series of nodes departing in two successive time intervals. Without Constraint 14, Trip Stream 2 could depart from Node C in Time Interval 3 at a faster speed than Trip Stream 1 and pass stream 1, which departs from node C in Interval 2 at a slower speed. With Constraint 14, Trip Stream 2 will depart from Node C such that it reaches Node D no earlier than Trip Stream 1. Kaufman and Smith (*15*) show that FIFO adjustments such as Constraint 14 are easily added to shortest-path label-correcting algorithms (but not label-setting algorithms) so long as labels are properly updated when it occurs.

Constraint 14 does not entirely replace the need for Constraint 11. Constraint 14 allows trips between different O-D pairs to become concurrent while sharing the same path, whereas Constraint 11 ensures a minimum separation of "last" vehicles departing from the same zone in successive time intervals. If Constraint 11 is removed from the problem, then trips from the same zone can "bunch" together in overly dense flows. As mentioned earlier, additional research will lead to better treatment of this bunching problem.

Albeit counter-intuitive, the maximization of zone-to-node travel times in subproblem (LP) is the correct determination of node time intervals and shortest-path travel times subject to arc lengths plus FIFO delay in Equations 10 through 14. Using the mechanical analogy of Minty (*16*, p. 724), Bertsekas (*17*) defines this formulation as the dual shortest-path problem according to the min-path/max-tension theorem defining this primal-dual relationship.

OPTIMALITY CONDITIONS OF DUE1

Although DUE is nonconvex over the domain of feasible node time intervals for all trip departures to all destinations, DUE1 is convex with a unique global optimum for any given set of fixed node time intervals. The optimality conditions of DUE1 stated in the paper's

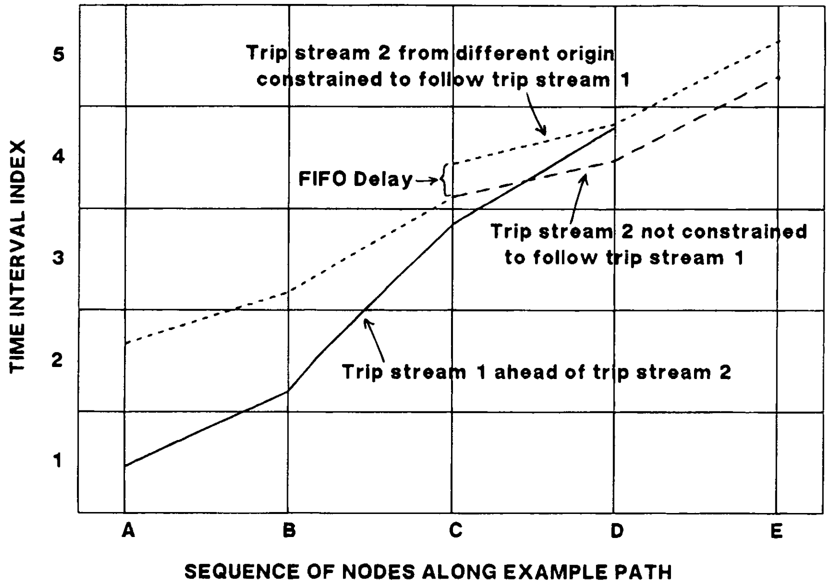

FIGURE 2 Effect of Constraint 14 on FIFO order of all O-D trips.

first section can be derived from (UP) for a given set of node time intervals as given by an optimal solution to (LP). Subproblem (LP) is a shortest-path linear program for which there exists an optimal solution for any given solution to (UP). Any set of node time intervals resulting from (LP) defines a directed network for which (UP) is a convex nonlinear program with a global optimum solution. Because node time intervals resulting from (LP) are uniquely determined by a given set of link volumes resulting from (UP), they can be assumed to be known in the derivation of optimality conditions for DUE1. The optimality conditions of DUE1 are next derived from (UP) for a given set of node time intervals to which all temporally continuous trip paths in the optimal solution must conform.

Equation 17 is the Lagrangian of (UP) with fixed-integer node time intervals. Over the domain of variable integer values, Equation 17 is nonconvex, and there are many local optima that are inferior to the global optimum. For a given set of fixed node time intervals, the bordered Hessian matrix of Equation 17 is positive definite, which means that there is a unique global optimum with no local optima (*18*). The bordered Hessian matrix of Equation 17 is only positive definite so long as each impedance function is a monotonically nondecreasing function of flow on link ij in time interval t alone, as was stipulated earlier.

$$
\begin{aligned}
L(X,V,\lambda,\mu,\tau) = & \sum_{ij\epsilon K}\sum_{t\epsilon T}\int_0^{x_{ij}^t} f_{ij}^t(w)\,dw \\
& -\sum_{ij\epsilon K}\sum_{t\epsilon T}\gamma_{ij}^t\left[x_{ij}^t - \sum_{r\epsilon Z}\sum_{d\leq t} v_{rij}^d\ \phi_{ri}^{dt}\right] \\
& +\sum_{r\epsilon Z}\sum_{n\epsilon N}\sum_{d\epsilon T}\mu_{rn}^d \\
& \times\left(q_{rn}^d - \sum_{t\geq d}\left[\sum_{in\epsilon K} v_{rin}^d\ \phi_{ri}^{dt} - \sum_{nj\epsilon K} v_{rnj}^d\ \phi_{rn}^{dt}\right]\right) \\
& +\sum_{r\epsilon Z}\sum_{ij\epsilon K}\sum_{d\epsilon T}\sum_{t\epsilon T}\tau_{rij}^{dt}\,(-v_{rij}^d\phi_{ri}^{dt}) \qquad (17)
\end{aligned}
$$

The optimality conditions are given by Equation 18 through 20.

$$\partial L/\partial x_{ij}^t - < f_{ij}^t(x_{ij}^t = \lambda_{ij}^t \qquad \text{for all } ij \epsilon K,\ t \epsilon T \qquad (18)$$

$$\partial L/\partial v_{rij}^d - < \mu_{rj}^d - \mu_{ri}^d)\ \phi_{ri}^{dt} = (\lambda_{ij}^t - \tau_{rij}^{dt})\ \phi_{ri}^{dt} \qquad \text{for all r } \epsilon Z,\ ij \epsilon K,\ d \epsilon T,\ t \epsilon T \qquad (19)$$

$$\partial L/\partial \tau_{rij}^{dt} - > \tau_{rij}^d\ v_{rij}^d\ \phi_{ri}^{dt} = 0,\ (\tau_{rij}^{dt} \geq 0) \qquad \text{for all } r \epsilon Z,\ ij \epsilon K,\ d \epsilon T,\ t \epsilon T \qquad (20)$$

where τ_{rij}^{dt} equals 0 if $v_{rij}^d\ \phi_{ri}^{dt} > 0$, nonnegative otherwise; equals impedance difference from node i to node j via a used path versus by link ij (used or unused) in time interval t for trips departing from zone r in time interval d.

The last part of Equation 17 ensures nonnegative link flows and results in a third optimality condition given by Equation 20, which requires τ_{rij}^{dt} to be 0 if any trips departing zone r in time interval d are assigned to link ij in time interval t and nonnegative otherwise. According to Equation 18, the optimal solution has a unique equilibrium impedance for each link in each time interval. According to Equations 19 through, 20, for any given pair of nodes, all used paths for a given origin and departure time must have equal impedances, and any unused path between these nodes cannot have a lower impedance.

Optimality conditions for dynamic user equilibrium can be stated similarly to Wardrop's (*7*) statement of necessary conditions for static user equilibrium. Let v_{rij}^d be the equilibrium flow on link ij in interval t of trips departing zone r in interval d, and λ_{ij}^t be the equilibrium impedance of link ij in interval t. Also, for trips departing zone r in interval d, let μ_{rj}^d be the equilibrium impedance of all used paths to node i. At equilibrium, according to Equations 21 and 22, all paths from zone r to node j used by trips departing in interval d have impedance μ_{rj}^d and no unused path for this same (r,j,d) combination has a lower impedance.

$$\mu_{ri}^d + \lambda_{ij}^t = \mu_{rj}^d \qquad \text{if } v_{rij}^d\ \phi_{ri}^{dt} > 0$$

and

$$\tau^{dt}_{rij} = 0 \qquad \text{for all } r \in Z,\ ij \in K,\ d \in T,\ t \in T \tag{21}$$

$$\mu^d_{ri} + \lambda^t_{ij} = \mu^d_{rj} \qquad \text{if } v^d_{rij}\ \ \phi^{dt}_{ri} = 0$$

and

$$\tau^{dt}_{rij} \geq 0 \qquad \text{for all } r \in Z,\ ij \in K,\ d \in T,\ t \in T \tag{22}$$

CONVERGENT DYNAMIC TRAFFIC ASSIGNMENT ALGORITHM

Whereas SUE can be solved efficiently by linear combination methods for nonlinear programs with all linear constraints [e.g., Frank-Wolfe (F-W) and PARTAN], these methods can easily create temporally discontinuous flows if applied directly to DUE1. Instead, the two subproblems of DUE1 are solved successively by a convergent dynamic algorithm (CDA). As indicated in Figure 3, CDA first solves (UP) with fixed node time intervals using the F-W method of linear combinations (or a similar technique), and then solves (LP) (which is a linear program) to update all node time intervals for the next F-W solution of (UP). Adjustments to link capacities are made between the upper and lower problems to account for spillback queueing effects or signal timing changes as explained in the next section.

The CDA algorithm terminates when fewer than an acceptable number of node time intervals change from one (LP) solution to the next. With fixed node time intervals, subproblem (UP) is solved without fixing which links are used. Only time intervals in which links are used by trips are fixed, depending on their trip origins and departure times. To clarify the CDA algorithm, the following steps are performed successively to solve subproblems (UP) and LP to near convergence.

1. Input all network data, temporal trip departure matrixes, and initial link flows. Initial link flows are optional, and can be set to 0, but SUE link flows reduced to the chosen time interval duration may be good starting values. Calculate initial node time intervals by solving (LP) with initial link flows. Set iteration counter $n = 0$.
2. Increment iteration counter $n = n + 1$.
3. (UP) Minimize Equation 1 subject to Equations 2 through 4, where all x^t_{ij} are variable and all $\{\phi^{dt}_{ri}\}$ are fixed as computed with Equations 5 and 6 according to the optimal values of $\{b^d_{ri}\}$ and $\{\alpha^{dt}_{ri}\}$ from solving (LP).
4. Adjust link capacities, which alters the impedance function $f^t_{ij}(x^t_{ij})$, for links and time intervals affected by spillback queues or other events.
5. (LP) Maximize Equation 7 subject to equations 6 through 14, where all $\{\alpha^{dt}_{ri}\}$ and $\{b^d_{ri}\}$ are variable and all x^t_{ij} are fixed to their optimal values from (UP).
6. Sum NDIFFS = total number of node time interval differences between iterations $n - 1$ and n. Compare each $\{\alpha^{dt}_{ri}\}^{n-1}$. If NDIFFS $\leq$ small percent of all node time intervals [$Z(N - 1)T$], then stop. Otherwise, return to Step 2.

CDA converges toward a dynamic user-equilibrium solution for the following reasons. First, if node time intervals corresponding to the true equilibrium are known, then solving (UP) will reproduce the equilibrium link volumes from which these node time intervals can be calculated. That convergence proof follows from the fact that any set of node time intervals resulting from (LP) defines a directed network for which (UP) is a convex nonlinear program for which a global optimum exists. Second, given node time intervals that do not correspond to a true dynamic equilibrium, then solving (UP) with the F-W algorithm will produce link volumes that shift the node time intervals toward their correct values. For example, if a node time interval is too early, then solving (UP) will assign more traffic to paths leading up to that node such that the node time interval is shifted later when recalculated in (LP). Oppositely, if a node time interval is too late, then solving (UP) will assign less traffic to paths leading up to that node such that the node time interval is shifted earlier when recalculated in (LP). Thus, CDA converges to a set of node time intervals that, when used to assign trips to the network in solving (UP), result in temporal link volumes that give rise to the same node time intervals when recalculated in (LP).

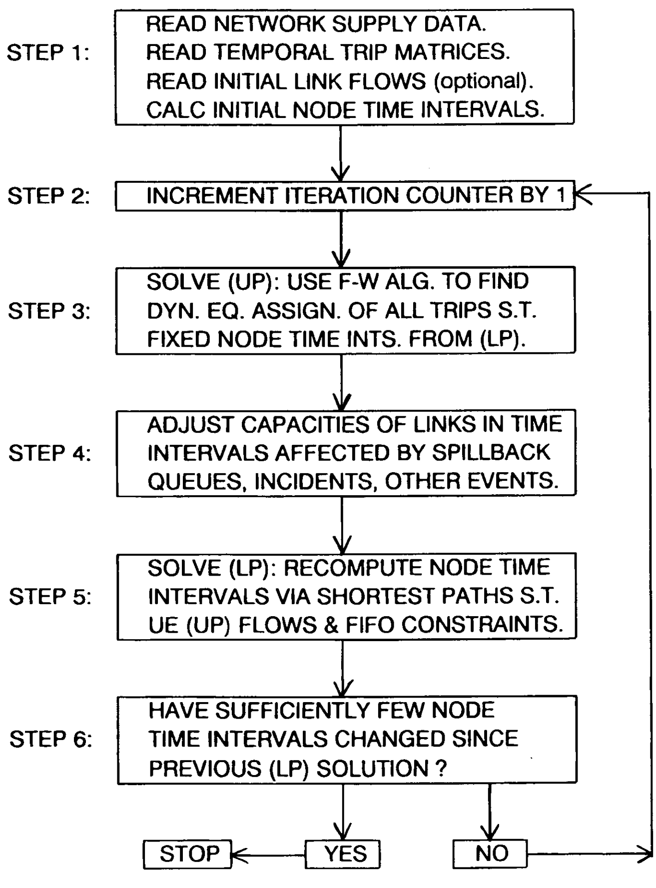

FIGURE 3 Steps of CDA solution algorithm.

DYNAMIC LINK CAPACITY ADJUSTMENTS FOR SPILLBACK QUEUES

A key feature of this modeling approach is the adjustments of link capacities input exogenously or generated endogenously. Exogenous changes (e.g., accidents, weather effects, signal timing changes, or time-of-day road restrictions for special events or construction) can be input to the program for specific links and times of day when they occur (expected or unexpected). Endogenous link capacity changes occur when spillback queues reduce the capacities of upstream links, or when integrated algorithms for ramp metering and signal timing make adjustments. When accidents or recurrent congestion create oversaturated conditions, then upstream link capacities are reduced in time intervals affected by spillback queues using wave propagation speeds to determine the timing of upstream effects. These capacity losses create further upstream effects to the extent and duration of the oversaturated condition. For example, if an accident occurs at 7:30 a.m. and blocks the right lane of a three-lane freeway, then the capacity of that link is reduced by roughly 50 percent from 7:30 a.m. until the estimated clearance time of the accident.

A subroutine called QUECAP is executed between the upper and lower problems of DUE1 to recognize exogenous link capacity changes, and toe compute endogenous link capacity changes. Three essential steps of QUECAP are to (*a*) track the locations of multiple spillback queues in a network, (*b*) weight the effects of multiple spillback queues when jointly affecting inflows to any node, and (*c*) adjust the capacities of inflow links to each node in proportion to the fractions of flows affected by each queue. Endogenous link capacity changes stabilize with successive solutions to the two subproblems of DUE1 and are made only between subproblem solutions so that the upper problem remains convex for each given set of supply specifications.

The queue propagation and capacity adjustment procedure is explained next. First note that all nodes are configured such that every node has only one outflow link (a merge node) or one inflow link (a diverge node). No node has multiple inflow and outflow links (i.e., no merge/diverge nodes). Intersections always have turn movement links, and weave sections always have connecting links between the merge and diverge nodes. Capacity adjustment steps starting from original unadjusted capacities are as follows:

1. Increment the pass number from 1 until convergence is achieved,
2. Increment the time interval from 1 to *T*,
3. Increment the node number from 1 to *N*,
4. For each outflow link from a node, compute the cumulative queue equal to all excess flow above capacity through the current time interval. Compute the link length fraction occupied by queueing during the current interval on the basis of the cumulative queue and wave speed. Exact times and locations where accidents initiate queueing can be specified to the program. Spillback from an oversaturated bottleneck link not caused by an accident is assumed to begin at the start of the time interval and at the entry to the link. The link length fraction occupied by a queue is computed by comparing the cumulative queue on the link to what the link can absorb as the vehicle stream compresses to higher density. Wave speed (with upstream being the negative direction) is computed as

$$\text{wave speed} = (\text{flow}^2 - \text{flow}^1) / (\text{density}^2 - \text{density}^1)$$

where

flow^1 = flow before queue (below original but above adjusted capacity),
flow^2 = flow inside queue (assumed to equal the adjusted capacity),
speed^1 = link length / travel time for flow^1 with original capacity,
speed^2 = link length / travel time for flow^2 with adjusted capacity,
$\text{density}^1 = \text{flow} / \text{speed}^1$ = density before queue (less dense), and
$\text{density}^2 = \text{flow}^2 / \text{speed}^2$ = density inside queue (more dense).

5. The link length fraction occupied by a queue equals the portion of link length (perhaps all) covered by the queue over successive time intervals. Inflow links are unaffected until a queue extends beyond the tail node of an outflow link, and only a fraction of the time interval will be affected when this first occurs. The affected time interval fraction equals (the interval start time)—(time that queue reaches link's tail) divided by the time interval duration Δt.
6. Compute the weighted volume-to-capacity (*V/C*) ratio of the outflow links from this node, which is weighted by each outflow link's volume and fraction of the time interval affected. A detailed explanation of this computation is not possible within the brevity of this paper.
7. Adjust the capacity of each inflow link so that its *V/C* ratio equals the weighted *V/C* ratio of the outflow links just found. Return to Step 3 until all nodes are processed; then return to Step 2 until all time intervals are processed. Then, if the capacity of any link in any time interval changes by more than an acceptable percent, return to Step 1 for another pass or, else, stop.

To capture the effects of multiple queues spilling back from several places in the network, and queues spilling back farther than one link in any time interval, multiple passes of the above steps are performed until all capacities in all time intervals do not change significantly. Bounding rules are added to prevent endogenously adjusted capacities from becoming too small (which generates a warning), and to prevent queues from dissipating too quickly. The rule here is that a link's capacity during dissipation cannot exceed the average of the above calculation and its reduced capacity in the prior interval. Adjusted capacities are then returned to the upper problem. Hence, both flows and capacities are time dependent in the travel time functions of the DUE1 objective function. This method of adjusting capacities has performed well in these applications but can be improved with further refinements.

SUMMARY AND CONCLUSIONS

This paper converts the DUE modeling approach presented in previous papers by the authors into quasi-continuous time, which enabled three key model improvements: (*a*) inclusion of spillback queueing effects and dynamic link capacity adjustments, (*b*) more accurate FIFO trip ordering between all O-D pairs, and (*c*) better link volume and speed transitions between time intervals. This paper also describes how capacity losses are approximated on upstream links because of spillback queueing from oversaturated links.

This dynamic traffic modeling approach has been applied to several networks of realistic size, detail, and complexity. In addition to the paper by Robles and Janson (*1*) mentioned at the start, Darjardi and Janson (*19*) apply it to the Colorado ski region of rural arterials and I-70 linking Denver and surrounding communities to ski resorts. The CDA found very good solutions to each of these applications without convergence difficulties. These diverse applications show the model to be a flexible analysis tool and CDA to be a robustly convergent solution algorithm for these types of dynamic traffic assignment problems.

REFERENCES

1. Robles, J., and B. N. Janson. Dynamic Traffic Modelling of the I-25/ HOV Corridor Southeast of Denver. In *Transportation Research Record,* in press.
2. Janson, B. N. Dynamic Traffic Assignment for Urban Road Networks. *Transportation Research,* Vol. 25B, Nos. 2/3, 1991, pp. 143–161.
3. Janson, B. N. A Convergent Algorithm for Dynamic Traffic Assignment. In *Transportation Research Record 1328,* TRB, National Research Council, Washington, D.C., 1991, pp. 69–80.
4. Janson, B. N. Network Design Effects of Dynamic Traffic Assignment. *Journal of Transportation Engineering,* ASCE, Vol. 121, No. 1, 1995, pp. 1–13.
5. Janson, B. N. A General Dynamic Assignment Model and Solution Algorithm. Presented at 39th North American Meeting of the International Regional Science Association, Chicago, Ill., Nov. 1992.
6. Janson, B. N., and J. Robles. Dynamic Traffic Assignment with Arrival Time Costs. In *Transportation and Traffic Theory* (C.F. Daganzo, ed.), Elsevier Science Publishers, Amsterdam, The Netherlands, 1993, pp. 127–146.
7. Wardrop, J. G. Some Theoretical Aspects of Road Traffic Research. *Proc., Institution of Civil Engineers,* Vol. 1, Part 2, 1952, pp. 325–378.
8. Friesz, T. L., J. Luque, R. L. Tobin, and B. Wie, Dynamic Network Traffic Assignment Considered As a Continuous Time Optimal Control Problem. *Operations Research,* Vol. 37, No. 6, 1989, pp. 893–901.
9. Wie, B. An Application of Optimal Control Theory to Dynamic User-Equilibrium Traffic Assignment. In *Transportation Research Record 1251,* TRB, National Research Council, Washington, D.C., 1989, pp. 66–73.
10. Wie, B., T. L. Friesz, and R. L. Tobin. Dynamic User-Optimal Traffic Assignment on Congested Multi-Destination Networks. *Transportation Research,* Vol. 24B, 1990, pp. 431–442.
11. Wie, B. Dynamic Analysis of User Optimized Network Flows with Elastic Demand. In *Transportation Research Record 1328,* TRB, National Research Council, Washington, D.C., 1991, pp. 81–87.
12. Ran, B., D. E. Boyce, and L. J. LeBlanc. A New Class of Instantaneous Dynamic User-Optimal Traffic Assignment Models. *Operations Research,* Vol. 41, No. 1, 1993, pp. 192–202.
13. Ran, B., and D. E. Boyce. Dynamic Urban Transportation Network Models: Theory and Implications for IVHS. In *Lecture Notes in Economics and Mathematical Systems.* Springer-Verlag, Berlin, (in press).
14. Friesz, T. L., D. Bernstein. T. E. Smith, R. L. Tobin, and B. Wie. A Variational Inequality Formulation of the Dynamic Network User-Equilibrium Problem. *Operations Research,* Vol. 41, 1993, pp. 179–191.
15. Kaufman, D. E., and R. L. Smith. Fastest Paths in Time-Dependent Networks for IVHS Applications. *IVHS Journal,* Vol. 1, No. 1, 1993, pp. 1–12.
16. Minty, G. J. A Comment on the Shortest Path Problem. *Operations Research,* Vol. 5, 1957.
17. Bertsekas, D. P. *Linear Network Optimization: Algorithms and Codes.* MIT Press, Cambridge, Mass., 1991.
18. Sheffi, Y. *Urban Transportation Networks.* Prentice-Hall, Englewood Cliffs, N.J., 1985.
19. Darjadi, P., and B. N. Janson. *Dynamic Modelling of Winter Weekend Ski Traffic in Colorado.* Colorado Department of Transportation, Denver, 1994.

Publication of this paper sponsored by Committee on Transportation Supply Analysis.