The General Link Transmission Model for Dynamic Network Loading and a Comparison with the DUE Algorithm

Guido Gentile

Dipartimento di Idraulica Trasporti e Strade, Sapienza Università di Roma, Via Eudossiana, 18 – 00184 Roma, Italy.
guido.gentile@uniroma1.it

Abstract: The Continuous Dynamic Network Loading problem is here addressed for given splitting rates, hence allowing for implicit path enumeration. To this aim, a macroscopic flow model for road links based on the Kinematic Wave Theory is coupled with a node model with priority rules at intersections, thus reproducing congested networks including queue spillback. The result is the General Link Transmission Model (GLTM), which extends previous results to the case of any concave fundamental diagram and node topology, without introducing spatial discretization of links into cells. The GLTM is compared with the DUE algorithm in terms of solution accuracy, computation efficiency and memory usage.

Keywords: dynamic traffic assignment, queue spillback, implicit path enumeration with splitting rates, macroscopic flow model with concave fundamental diagram, kinematic wave theory with cumulative flows.

---

1. INTRODUCTION

In the context of within-day Dynamic Traffic Assignment (DTA) the spatial propagation of flows takes a time which depends on the use of the network. The Continuous Dynamic Network Loading (CDNL) problem (see for example Xu et al. (1999)) consists in determining the link flows corresponding to given transport demand and route choices through a performance model yielding travel times as a function of flows, where all such variables are temporal profiles.

Route choices may be defined globally as path probabilities, or approximated locally as splitting rates, either in terms of turn probabilities or in terms of arc conditional probabilities at nodes. Under the assumption that all users of a given class directed to a same destination make the same route choices in probabilistic terms, once they have reached the same node of the network at the same time of the day, regardless the path followed till that point (i.e. they behave according to a sequential approach), splitting rates distinguished by destination can well reproduce path probabilities, thus allowing for implicit path enumeration. It is worth noting that for deterministic and logit route choices the sequential approach is equivalent to the classic joint approach (Gentile and Papola, 2006). If the splitting rates are not destination specific, then the consistency with the o-d demand flows is not guaranteed in correspondence of travel time variations; nevertheless, in some applications this rough approximation is acceptable and allows for the solution of the problem on large networks with a fine-grained time discretization.

The main traffic facts that should inform the performance model are the prevalence of vehicle queues (hypercritical congestion) with their backward propagation, called spillback, specially in the urban context, and the difference in the desired speed among drivers (hypocritical congestion), specially in the extra-urban context.

These phenomena are satisfactorily reproduced through the Theory of Kinematic Waves (KWT), which implements on each link the macroscopic flow paradigm of representing vehicles as a partially compressible mono-dimensional fluid, and through its network extension which requires a node model to propagate flow states among adjacent links involving priority rules. However, many authors addressing DTA encourage the

use of a simplified model reproducing only vertical queues without spillback and hypocritical congestion, where the main state variable becomes the link occupancy (Han and Heydecker, 2006), since its separability in time and space allows for an easier formulation of the problem. The most popular approach to solve the CDNL based on the simplified KWT, where the fundamental diagram has by definition a triangular shape, is the Cell Transmission Model (CTM) proposed by Daganzo (1994, 1995). The CTM, besides the limitation from the modelling point of view regarding the shape of the fundamental diagram, from a computational point of view suffers the spatial discretization of links both in terms of efficiency and accuracy.
Recently, a new approach has been developed by Yperman (2007) to address on a network the simplified KWT without spatial discretization of links into cells, therefore called the Link Transmission Model (LTM), which is based on the solution in terms of the cumulative flows proposed by Newell (1993). Independently, we have proposed a similar approach to solve the CDNL in the case where the fundamental diagram is any concave function (Gentile and Papola, 2007).
In this paper we present the General Link Transmission Model (GLTM), that is the extension of the LTM to any concave fundamental diagram and node topology, and compare it with the DUE algorithm proposed in Gentile et al. (2007) and implemented in Visum (Gentile et al., 2006). The GLTM has been applied to practical instances of the CDNL problem on large real datasets in the context of traffic signal setting (Fusco and Gentile, 2008) and info-mobility (Gentile and Meschini, 2008) proving to be robust and efficient.

2. MACROSCOPIC MODELS

The link performance models aimed at reproducing travel times as a function of link flows under the macroscopic fluid paradigm can be classified into two groups: space-continuous and space-discrete.
Space-continuous models are typically formulated as a system of differential equations in time and space, which is solved through finite difference methods. Such models yield accurate results, but require considerable computing resources, since their algorithmic implementation relies on a dense space discretization; for this reason they are also referred to as point-based. Altogether, they are very effective but poorly efficient. Among them, besides the CTM, we recall METANET, proposed in Messmer and Papageorgiou (1990), which derives from a second order approximation of vehicle trajectories.
Space-discrete models do not require any space discretization, and for this reason are also referred to as link-based. They can be in turn divided into whole link models and wave models.
Whole link models (e.g. Astarita, 1996, and Ran et al., 1997) yield link performances as a function of the space-average density (i.e. the number of vehicles on the link) without considering the propagation of flow states along the link. But this way the representation of hypocritical congestion becomes more and more ineffective as the length of the link increases, so that they may be suitable only to represent hypercritical congestion in terms of vertical queues on short links (Heydecker and Addison, 1998); on the other hand, since spillback is not considered, queues shall not be longer than links. Despite such major deficiencies, these models are widely used in DTA because of their simplicity (e.g. Friesz et al., 1993, and Tong and Wong, 2000).
Wave models, based on the KWT, take (implicitly) into account the propagation of flow states, yielding link performances as a function of the traffic conditions that the vehicle encounters by traveling along the link. They require minimal computing resources, while yielding realistic results both in urban and extra-urban contexts. These models have been first developed for bottlenecks with constant capacity, as in Arnott et al. (1990), Ghali and Smith (1993) and Bellei et al. (2005); that is, when only two speeds may occur on the link: the free-flow speed and the queue speed. Recently, in Gentile et al. (2005) and Gentile et al. (2007), they have been extended to the case of long links and time-varying capacity.

3. REVIEW OF THE KINEMATIC WAVE THEORY BASED ON CUMULATIVE FLOWS

The link is assumed to be a homogenous channel of length $L > 0$, whose physical capacity is reduced at the initial and end points by two time-varying bottlenecks, called respectively the entry capacity and the exit capacity.

The cumulative flow $N(x,\tau)$ is the number of vehicles that passed point $x \in [0, L]$ before time τ. This function is actually discontinuous, but we can consider a smooth approximation that makes it C^2 without changing the essence of the traffic phenomenon. In this case, under the assumption that no vehicle is created or destroyed along the link, the points in the time-space plane such that $N(x,\tau) = n$ represent the trajectory of the n-th vehicle, except if no vehicle is traveling, where the contour line degenerates in an area.
Let $q(x,\tau)$ and $k(x,\tau)$ be the flow and the density of vehicles on point x at time τ, defined respectively as:

$$q(x,\tau) = \partial N(x,\tau) / \partial \tau, \qquad (3.1)$$
$$k(x,\tau) = -\partial N(x,\tau) / \partial x. \qquad (3.2)$$

We now analyze the trajectory of the vehicle travelling at point (x,τ) in the time-space plane. If one considers the cumulative flow $N(x,\tau)$ as the elevation of the point, this is like aiming to determine the contour line passing through (x,τ). Therefore, we are formally seeking a direction $\mathrm{d}x/\mathrm{d}\tau$ in the time-space plane such that:

$$\mathrm{d}N(x,\tau) = \frac{\partial N(x,\tau)}{\partial x} \cdot \mathrm{d}x + \frac{\partial N(x,\tau)}{\partial \tau} \cdot \mathrm{d}\tau = -k(x,\tau) \cdot \mathrm{d}x + q(x,\tau) \cdot \mathrm{d}\tau = 0. \qquad (3.3)$$

Denoting $v(x,\tau)$ this direction, which is indeed the speed of the vehicle, we have:

$$v(x,\tau) = \frac{q(x,\tau)}{k(x,\tau)}. \qquad (3.4)$$

Based on the Schwarz's theorem we also have the following mass conservation law:

$$\frac{\partial^2 N(x,\tau)}{\partial \tau \cdot \partial x} - \frac{\partial^2 N(x,\tau)}{\partial x \cdot \partial \tau} = \frac{\partial q(x,\tau)}{\partial x} + \frac{\partial k(x,\tau)}{\partial \tau} = 0. \qquad (3.5)$$

To analyze the propagation of traffic states, let us look at the time-space plane for the points in the neighborhood of (x,τ) which are affected by its same flow. If we consider the flow as the elevation of the point, this is like aiming to determine the contour line passing through (x,τ). Therefore, we are formally seeking a direction $\mathrm{d}x/\mathrm{d}\tau$ in the time-space plane such that:

$$\mathrm{d}q(x,\tau) = \frac{\partial q(x,\tau)}{\partial x} \cdot \mathrm{d}x + \frac{\partial q(x,\tau)}{\partial \tau} \cdot \mathrm{d}\tau = -\frac{\partial q(x,\tau)}{\partial \tau} \cdot \mathrm{d}x + \frac{\partial q(x,\tau)}{\partial \tau} \cdot \mathrm{d}\tau = 0. \qquad (3.6)$$

Denoting $w(x,\tau)$ this direction, which is indeed the speed of the propagation wave, based on equation **Errore. L'origine riferimento non è stata trovata.** we have:

$$w(x,\tau) = \frac{\partial q(x,\tau)}{\partial k(x,\tau)}. \qquad (3.7)$$

Moreover, based on the Green's theorem we have the following energy conservation law:

$$N(x_2,\tau_2) - N(x_1,\tau_1) = \int_{C(1-2)} \mathrm{d}N(x,\tau) = \int_{C(1-2)} q(x,\tau) \cdot \mathrm{d}\tau - k(x,\tau) \cdot \mathrm{d}x, \qquad (3.8)$$

where $C(1-2)$ is any curve in the time-space plane that connects point (x_1,τ_1) to (x_2,τ_2); this relation will be the cornerstone of our link model.

The fundamental diagram $q(k)$ is an experimental relation between flow and density that holds for stationary traffic. Classical forms are the triangular shape and the parabolic shape. We will consider any concave function passing through the origin, having a maximum flow $\Phi > 0$, called physical capacity, in correspondence of the critical density $K > 0$, and getting again to zero at the jam density $J > K$. Then, for every flow $q \in [0, \Phi]$ there are two possible densities on the fundamental diagram: the first one, denoted $k^o(q)$, corresponds to a hypocritical state; the second one, denoted $k^+(q)$, corresponds to a hypercritical state. The derivative of the fundamental diagram at $k = 0$ and at $k = J$, which are referred to respectively as the free flow speed $V > 0$ and the jam wave speed $W < 0$, represent the maximum vehicle speed and the minimum wave speed.

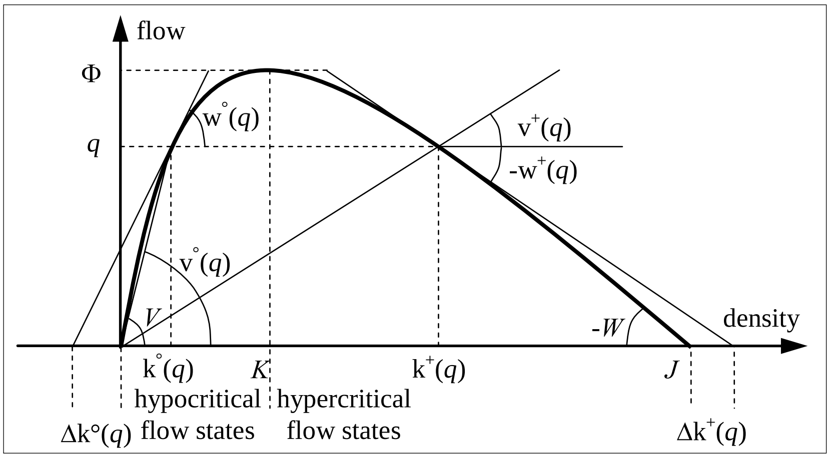

Fig. 1: The fundamental diagram

The KWT is based on the assumption that the fundamental diagram holds also for non-stationary traffic, which implies that vehicles adapt their speed instantaneously with infinite accelerations, thus yielding a first order approximation of trajectories. This is formally expressed as follows:

$$q(x,\tau) = q(k(x,\tau)) \,. \tag{3.9}$$

Using the fundamental diagram we can introduce two functions of the flow yielding the vehicle speed respectively for hypocritical and hypercritical states:

$$v^o(q) = \frac{q}{k^o(q)}, \; v^+(q) = \frac{q}{k^+(q)} \,. \tag{3.10}$$

Since in the KWT the local points such that $\mathrm{d}q(x,\tau) = 0$ are characterized by the same flow states, also $\partial q(x,\tau) / \partial k(x,\tau)$ is constant there, i.e. the local wave $w(x,\tau)$ will keep propagating with the same slope, thus yielding a straight line in the time-space plane. Using the fundamental diagram we can introduce two functions of the flow yielding the wave speed respectively for hypocritical and hypercritical states:

$$w^o(q) = \frac{1}{\mathrm{d}k^o(q)/\mathrm{d}q}, \; \mathrm{w}^+(q) = \frac{1}{\mathrm{d}k^+(q)/\mathrm{d}q} \,. \tag{3.11}$$

To resolve the interactions and the conflicts among vehicle trajectories and among wave trajectories, we will exploit the following two general properties of the fluid paradigm:

a) the First In First Out (FIFO) rule, stating that no overtaking among vehicles can occur on the link;
b) the Newell-Luke Minimum Principle (NLMP), claiming that among all possible states which may affect a given point of a link the one yielding the minimum cumulative flow dominates the others.

Based on the above relations and considerations we can reproduce the traffic pattern at any time and on any point along the link for given boundary conditions. In general, at a given instant the link can be divided into two subsequent segments based on the nature of the prevailing flow states: in the first segment the forward-propagating hypocritical flow states that derive from boundary conditions at the initial point prevail on the backward-propagating hypercritical flow states that derive from boundary conditions at the end point; in the second segment the opposite situation holds, so that vehicles travel in a queue. Concerning the limit cases: when the second segment disappears, there is no queue on the link and we have the so called free flow condition; when the first segment disappears, the queue occupies the whole link and we have the so called spillback condition.
Given that our aim is to use the above link model in a dynamic network loading, we are mainly interested here on how flow states propagate from one terminal point of the link to the other, and whether the boundary condition at a terminal point prevails on the flow state coming from the other terminal point. Exploiting the energy conservation law, we will now address the forward propagation of hypocritical inflows directly to the end point and the backward propagation of hypercritical outflows directly to the initial point, since this allows us to easily formulate the boundary conditions in terms of vertical queue, vertical storage, sending flow and receiving flow, as explained in the following section, without the need of handling the detail of flow states along the link.
Let $F(\tau) = N(0,\tau)$ and $E(\tau) = N(L,\tau)$ be respectively the cumulative inflow and outflow at time τ, i.e. the number of vehicles that passed respectively the initial point and the end point of the link until that instant. By definition, the inflow and outflow rates are then given by:

$$f(\tau) = \frac{\mathrm{d}F(\tau)}{\mathrm{d}\tau}, \tag{3.12}$$

$$e(\tau) = \frac{\mathrm{d}E(\tau)}{\mathrm{d}\tau}. \tag{3.13}$$

The instant $u(\tau) \geq \tau$ when the forward kinematic wave generated at time τ on the initial point of the link by the hypocritical inflow $f(\tau)$ reaches the end point is given by:

$$u(\tau) = \tau + \frac{L}{w^o(f(\tau))}. \tag{3.14}$$

In general, $u(\tau)$ is not invertible, since more than one kinematic wave generated on the initial point may reach the end point at the same time. The integration of $\mathrm{d}N(x,\tau)$ along the kinematic wave from $(x_1,\tau_1) = (0,\tau)$ to $(x_2,\tau_2) = (L,u(\tau))$ is easy since the flow state is constant there, then based on (3.8) we have:

$$\hat{H}(\tau) = F(\tau) + \frac{f(\tau)\cdot L}{w^o(f(\tau))} - L\cdot k^o(f(\tau)) = F(\tau) + L\cdot\Delta k^o(f(\tau)), \tag{3.15}$$

where $\hat{H}(\tau) = N(L,u(\tau))$ if $f(\tau)$ is the prevailing flow state at $(L,u(\tau))$, while function $\Delta k^o(q)$ is implicitly defined in Fig. 1. Based on the NLMP, among all forward kinematic waves that reach the end point at time τ the one yielding the minimum cumulative flow, denoted $H(\tau)$, dominates the others:

$$H(\tau) = \min\{\hat{H}(\sigma):\ u(\sigma) = \tau\}. \tag{3.16}$$

The instant $z(\tau) \geq \tau$ when the backward kinematic wave generated at time τ on the end point of the link by the hypercritical outflow $e(\tau)$ reaches the initial point is given by:

$$z(\tau) = \tau - \frac{L}{w^+(e(\tau))}. \tag{3.17}$$

As above, $z(\tau)$ is not invertible, since more than one kinematic wave generated on the end point may reach the initial point at the same time. Again, the integration of $\mathrm{d}N(x,\tau)$ along the kinematic wave from $(x_1,\tau_1)=(L,\tau)$ to $(x_2,\tau_2)=(0,z(\tau))$ is easy since the flow state is constant there, then based on (3.8) we have:

$$\hat{G}(\tau)=E(\tau)-\frac{e(\tau)\cdot L}{w^{+}(e(\tau))}+L\cdot k^{+}(e(\tau))=E(\tau)+L\cdot J+L\cdot\Delta k^{+}(e(\tau)) \quad (3.18)$$

where $\hat{G}(\tau)=N(0,z(\tau))$ if $e(\tau)$ is the prevailing flow state at $(0,z(\tau))$, while function $\Delta k^{+}(q)$ is implicitly defined in Fig. 1. Based on the NLMP, among all backward kinematic waves that reach the initial point at time τ the one yielding the minimum cumulative flow, denoted $G(\tau)$, dominates the others:

$$G(\tau)=\min\{\hat{G}(\sigma):\ \mathrm{z}(\sigma)=\tau\}. \quad (3.19)$$

It is worth noticing that in (3.16) and (3.19) we did not impose, respectively, the condition that the inflow and the outflow which generates the wave should be hypocritical and hypercritical. In the following we will show that this is not a mistake, indeed, by proving respectively that:

a) given an hypercritical state at the initial point of the link, the forward propagation of the corresponding hypocritical flow can never prevail at the end point;
b) given an hypocritical state at the end point of the link, the backward propagation of the corresponding hypercritical flow can never prevail at the initial point.

We will develop the proof of assertion b), which is less intuitive. Let e be the hypocritical outflow occurring at the generic time τ_2. The backward propagation of the corresponding hypercritical flow based on (3.18) yields:

$$G(\tau_3)=E(\tau_2)+L\cdot J+L\cdot\Delta k^{+}(e),$$

where $\tau_3=\tau_2-L/w^{+}(e)$.
Since the flow state of e is hypocritical, it has been generated at the initial point in such a way that based on (3.15) we have:

$$E(\tau_2)=F(\tau_1)+L\cdot\Delta k^{o}(e),$$

where $\tau_1=\tau_2-L/w^{o}(e)$.
We now show that $F(\tau_3)$ is smaller that $G(\tau_3)$, which implies that the actual inflow at τ_3 is hypocritical, i.e. the back propagation is useless but harmless. Since the inflow during the interval $[\tau_1,\tau_3]$ is by definition not higher than the physical capacity Φ, we shall prove that:

$$F(\tau_3)\le F(\tau_1)+(\tau_3-\tau_1)\cdot\Phi<G(\tau_3).$$

Manipulating the above relations we obtain:

$$\frac{\Phi}{w^{o}(e)}-\frac{\Phi}{w^{+}(e)}<\Delta k^{o}(e)+J+\Delta k^{+}(e),$$

which based on Fig. 1 is patently true.
Applying a similar reasoning one can prove the validity of assertion a).

4. THE LINK MODEL

Let $D(\tau)$ be the cumulative demand at time τ, i.e. the number of vehicles generated (or attracted, if negative) along the link until that instant. The latter is introduced to represent, besides travel demand entering and exiting the network that is modelled by specific dummy links, any adjustment to the number of vehicles that one may force on road links to match external measures in real-time applications. This way we are aware to compromise the assumption that no vehicle is created or destroyed on the link at the price of some modelling inconsistencies, e.g. in this case the contour line of $N(x,\tau)$ doesn't represent anymore vehicle trajectories. Therefore, except for origin links, $D(\tau)$ is intended to be flat with sporadic discontinuities, and can be assumed to be null by most rigorous readers.
The vertical queue at time τ, denoted $S(\tau)$, is defined as:

a) the vehicles entered at the initial point of the link that propagating forward would reach the end point no later than time τ if no queue was present there, represented by $H(\tau)$;
b) minus the vehicles that exited the link no later than time τ, represented by $E(\tau)$;
c) plus the vehicles generated on the link no later than time τ, represented by $D(\tau)$;

$$S(\tau) = H(\tau) - E(\tau) + D(\tau) \, . \tag{4.1}$$

At the end point of the link a forward-propagating hypocritical flow state may actually occur at time τ only when $S(\tau) = 0$, which identifies the free flow condition. Indeed, for $S(\tau) > 0$ we have $E(\tau) < H(\tau) + D(\tau)$, which based on the NLMP means that hypercritical boundary conditions are prevailing, i.e. a queue of vehicles to exit the link is present.
The vertical storage at time τ, denoted $R(\tau)$, is defined as:

a) the storage capacity given by $L \cdot J$, plus the free spaces left by the vehicles at the end point of the link that propagating backward would reach the initial point no later than time τ if a queue was present there, both represented by $G(\tau)$;
b) minus the vehicles that entered the link no later than time τ, represented by $F(\tau)$;
c) minus the vehicles generated on the link no later than time τ, represented by $D(\tau)$;

$$R(\tau) = G(\tau) - F(\tau) - D(\tau) \, . \tag{4.2}$$

At the initial point of the link a back-propagating hypercritical flow state may actually occur at time τ only when $R(\tau) = 0$, which identifies the spillback condition. Indeed, for $R(\tau) > 0$ we have $F(\tau) < G(\tau) - D(\tau)$, which based on the NLMP means that hypocritical boundary conditions are prevailing, i.e. no spillback to enter the link is present since, if there is a queue, it has not reached the initial point yet.
The actual number of vehicles on the link $N(\tau)$ at time τ is given by:

$$N(\tau) = F(\tau) - E(\tau) + D(\tau) \, . \tag{4.3}$$

This is higher than $S(\tau)$, which does not include the vehicles that are travelling forward along the link, and is lower than $L \cdot J - R(\tau)$, which includes free spaces that are travelling backward along the link:

$$L \cdot J - R(\tau) \geq N(\tau) \geq S(\tau) \, . \tag{4.4}$$

In general, the physical capacity Φ of the link is reduced both at the initial point, to reproduce access restrictions (e.g. limited traffic zones), and at the end point, to reproduce intersection regulations (e.g. traffic signals). The entry and exit capacities are modelled here respectively through the time dependent shares $\mu(\tau)$ and $\eta(\tau)$ that multiply the physical capacity.

Finally, the boundary conditions of the link model, that allow for combining it together with the node model described in the following section so as to perform the dynamic network loading, are represented by the so called:

a) sending flow, at the end point, i.e. the maximum flow that the link can send to its forward star;
b) receiving flow, at the initial point, i.e. the maximum flow that the link can receive from its backward star.

The sending flow $s(\tau)$ at time τ results from the minimum between:

a) the maximum flow that can exit the link under free flow conditions, which for $\mathrm{d}\tau \to 0$ is given by $\mathrm{d}H(\tau)/\mathrm{d}\tau + \mathrm{d}D(\tau)/\mathrm{d}\tau$, if the vertical queue $S(\tau)$ is null, and tends to infinity otherwise;
b) the exit capacity, given by the current reduction $\eta(\tau)$ of the physical capacity Φ;

$$s(\tau) = \min\left\{\frac{S(\tau)}{\mathrm{d}\tau} + \frac{\mathrm{d}H(\tau)}{\mathrm{d}\tau} + \frac{\mathrm{d}D(\tau)}{\mathrm{d}\tau},\ \eta(\tau)\cdot\Phi\right\}. \qquad (4.5)$$

Analogously, the receiving flow $r(\tau)$ at time τ results therefore from the minimum between:

a) the maximum flow that can enter the link under spillback conditions, which for $\mathrm{d}\tau \to 0$ is given by $\mathrm{d}G(\tau)/\mathrm{d}\tau - \mathrm{d}D(\tau)/\mathrm{d}\tau$, if the vertical storage $R(\tau)$ is null, and tends to infinity otherwise;
b) the entry capacity, given by the current reduction $\mu(\tau)$ of the physical capacity Φ;

$$r(\tau) = \min\left\{\frac{R(\tau)}{\mathrm{d}\tau} + \frac{\mathrm{d}G(\tau)}{\mathrm{d}\tau} - \frac{\mathrm{d}D(\tau)}{\mathrm{d}\tau},\ \mu(\tau)\cdot\Phi\right\}. \qquad (4.6)$$

Under the assumption that the vehicle osmosis represented by $D(\tau)$ is concentrated just before the end point, i.e. vehicles are generated and attracted in the vertical queue, the FIFO rule can be expressed formally as:

$$F(\tau) = E(t(\tau)) - D(t(\tau)), \qquad (4.7)$$

where $t(\tau)$ is the exit time of a vehicle entering the link at time τ. On this basis, once the cumulative inflow and outflow temporal profiles are known, the exit time temporal profile can be easily determined as:

$$t(\tau) = \max\left\{\tau + \frac{L}{V},\ \min\{\sigma:\ F(\tau) = E(\sigma) - D(\sigma)\}\right\}, \qquad (4.8)$$

where the max min operator is needed to handle the case of null inflows.

5. THE NODE MODEL

The road network is modeled as usual in terms of an oriented graph $G = (X, A)$, where X is the set of nodes, each representing an intersection where roads merge and diverge, and $A \subseteq X \times X$ is the set of links, each representing a road that connects two intersections. The initial node of a link $a \in A$ is denoted by $TL(a)$ and referred to as the tail, while its final node is denoted by $HD(a)$ and referred to as the head. The backward star of $x \in X$ is denoted by $BS(x) = \{b \in A:\ y = TL(b)\}$, while its forward star is denoted by $FS(x) = \{a \in A:\ y = HD(a)\}$. In this section, any references to time is omitted in the notation, since all relations informing the node model regard a single instant, while the reference to the link, which has been omitted so far, is introduced.

The link model presented in the previous section provides the main input for the node model, that are the sending and receiving flows. In turn, the output of the node model are the inflow and outflow rates, that constitute the main input for the link model.
In a merging $x \in N$, where no routing may occur, the problem is to split the receiving flow r_b of the link $b \equiv FS(x)$ available at time τ among the links belonging to its backward star, whose outflows compete to get through the intersection. In principle, we assume that the receiving flow is partitioned proportionally to the priority of each link $a \in BS(x)$, defined by $\varphi_{ab} \cdot \eta_a \cdot \Phi_a$ where φ_{ab} is the priority coefficient of turn ab. This way it may happen that for some link c the turn flow f_{cb} is lower than the share of receiving flow assigned to it, so that only a lesser portion of the latter is actually exploited. Let ω_{cb} be 1 for such links and 0 for the others. The rest of the receiving flow $r_b - \sum_{c\in BS(x)} f_{cb} \cdot \omega_{cb}$ shall then be partitioned among the links that are in spillback from link b. On this basis, the receiving flow r_{ab} of turn ab is given by:

$$\xi_b = \frac{r_b - \sum_{c\in BS(x)} f_{cb} \cdot \omega_{cb}}{\sum_{c\in BS(x)} \varphi_{cb} \cdot \eta_b \cdot \Phi_c \cdot (1-\omega_{cb})}, \tag{5.1}$$

$$r_{ab} = \xi_b \cdot (\varphi_{ab} \cdot \eta_a \cdot \Phi_a), \tag{5.2}$$

$$\omega_{ab} = 1, \text{ if } f_{ab} < r_{ab}; \;\; \omega_{ab} = 0 \text{ otherwise}. \tag{5.3}$$

The partition set expressed by the coefficients ω_{ab} can be easily proved to be unique, and it can be simply obtained by iteratively adding to an initially empty set each link $a \in BS(x)$ such that $f_{ab} / (\varphi_{ab} \cdot \eta_a \cdot \Phi_a) < \xi_b$, as in Gentile et al. (2007). Based on(5.1), doing this ξ_b increases, so that (5.2) makes sense also for the links that are not in spillback from link b, thus enhancing the overall continuity of the model. Whenever all the coefficients ω_{ab} become equal to 1, then no link is in spillback and we can set $r_{ab} = \eta_a \cdot \Phi_a$.
Path choice is represented here by the splitting rate p_{ab}, expressing the probability that the next link of the path is $b \in FS(x)$ for vehicles coming from link $a \in BS(x)$, so that the sending flow s_{ab} of turn ab is given by:

$$s_{ab} = s_a \cdot p_{ab}. \tag{5.4}$$

In a diversion $x \in N$, where routing takes place, the problem is to determine at the generic time τ the most severe reduction, if any, to the sending flow s_{ab} from link $a \equiv BS(x)$ among those produced by the receiving flow r_{ab} of each link $b \in FS(x)$ and by the turn capacity Φ_{ab}. In order to ensure the FIFO rule applied to the vehicles exiting from link a, the share of sending flow ρ_a that actually gets through is the same for all links $b \in FS(x)$:

$$\rho_a = \frac{f_{ab}}{s_{ab}} = \frac{e_a}{s_a}, \tag{5.5}$$

$$\rho_a = \min\left\{1, \frac{\Phi_{ab}}{s_{ab}}, \frac{r_{ab}}{s_{ab}} : b \in FS(x), s_{ab} > 0\right\}. \tag{5.6}$$

When considering a generic node $x \in N$ with both mergings and diversions, the above relations shall hold jointly. Finally the resulting inflows and outflows are simply given as follows:

$$f_b = \sum_{a\in BS(x)} f_{ab}, \tag{5.7}$$

$$e_a = \sum_{b\in FS(x)} f_{ab}. \tag{5.8}$$

In the particular case where node $x \in N$ works like several separate mergings, i.e. when:

$$f_{ab} > 0 \Rightarrow f_{ac} = 0, \ \ \forall a \in BS(x), \ \ \forall b \in FS(x), \ \ \forall c \in FS(x), \ \text{ with } \ b \neq c \ ,$$

we introduce the hypothesis that drivers do not occupy the intersection if they can't cross it due to the presence of a queue on their successive link, but wait until the necessary space becomes available. Indeed, our model is not capable of addressing the deterioration of performances due to a misusage of the intersection capacity.

6. PROBLEM FORMULATION AND SOLUTION ALGORITHM

Based on (4.5), (4.1), (3.14-16), (3.12) and (4.6), (4.2), (3.17-19), (3.13) we can formalize the proposed link model as the following functional, for each link $a \in A$ and time τ :

$$(s_a(\tau), r_a(\tau)) = \Gamma_a^\tau(f_a(\sigma), \ e(\sigma): \ \sigma < \tau) \ . \qquad (6.1)$$

The link model is separable in space but non-separable in time.
Based on (5.2-8) we can formalize the proposed node model as the following functional, for each node $x \in X$ and time τ :

$$(f_b(\tau), e_a(\tau): \ a \in BS(x), \ b \in FS(x)) = \Psi_x^\tau(s_a(\tau), \ r_b(\tau): \ b \in FS(x)) \ . \qquad (6.2)$$

The node model is separable in time but non-separable in space.
Given the above link and node models, we can formulate the Continuous Dynamic Network Loading problem as the system of differential equations (6.1-2), which can be solved in chronological order.
To address numerically the problem we shall discretize the time horizon into intervals of equal duration $\Delta\tau$ (e.g. 10 sec) and introduce a look-ahead time window of duration T (e.g. 30 min) that satisfy, respectively, the following conditions:

a) no kinematic wave generated at any time τ on a link terminal (initial or end point) can reach the other terminal before time $\tau + \Delta\tau$;
b) no kinematic wave generated at any time τ on a link terminal can reach the other terminal after time $\tau + T$.

Condition a) reflects the theoretical need to compute the sending and receiving flows of a given time interval based on the inflows and outflows already computed for the previous time intervals.
Condition b) is consistent with the practical need of setting an upper bound in time to the direct effects of any propagating flow state.
It is convenient to define the entire period of simulation as a multiple of T . Let then λ be the number of time intervals that make up the look-ahead time window: $T = \lambda \cdot \Delta\tau$, and let ν be the number of time windows that make up the simulation period. While analyzing the *m*-th look-ahead time window, the *i*-th time instant is:

$$\tau_i = T_0 + (m-1) \cdot T + i \cdot \Delta\tau \ , \qquad (6.3)$$

where T_0 is the initial instant of the simulation, with $m \in [1, \nu]$ and $i \in [0, 2\lambda]$. Note that we need to handle two consecutive look-ahead windows at the time, because whilst the first one of them is being analyzed the generated kinematic waves may reach the opposite terminal point during the next one. A fictitious time instant $\tau_{2\lambda+1} = \infty$ is useful to avoid some checks in the propagation algorithm. Results for each link and turn can be memorized for pre-specified times.
To satisfy condition b) and enhance model flexibility it is useful to define separately the hypocritical and hypercritical branches of the fundamental diagram, while it is necessary to introduce a lower bound ϖ on the absolute value of the wave speeds. For example, in the numerical applications we considered the following polynomial form, defined for $k \in [\kappa, K]$:

$$q(k)=\frac{\Phi}{\upsilon}\cdot\left(1-\left(1-(k-\kappa)\cdot\Omega\cdot\upsilon\cdot\frac{1-\gamma}{\Phi}\right)^{1/_{1-\gamma}}\right), \qquad (6.4)$$

where $|\Omega|$ is the maximum wave speed, κ is the corresponding density, γ is a shape parameter, and $\upsilon=1-(\varpi/|\Omega|)^{1/\gamma}$. In particular, it is:

$\Omega=V>0,\ \ \kappa=0$, for hypocritical flows;
$\Omega=W<0,\ \ \kappa=J$, for hypercritical flows.

The resulting functions that express the density and the wave speed in terms of the flow are:

$$k(q)=\kappa+\Phi/\Omega/\upsilon/(1-\gamma)\cdot(1-(1-\upsilon\cdot q/\Phi)^{1-\gamma}) \qquad (6.5)$$

$$w(q)=\Omega\cdot\left(1-\upsilon\cdot\frac{q}{\Phi}\right)^{\gamma}. \qquad (6.6)$$

The critical density $K=k(\Phi)$ for hypocritical flows must be not greater than for hypercritical flows, otherwise the two branches would overlap; to avoid faults, in this case we shall decrease Φ accordingly. For $\gamma=1/2$ and $\gamma\to 0$, the fundamental diagram assumes respectively a parabolic and a trapezoidal shape; the latter reduces to the classical case of simplified kinematic waves.
To force the satisfaction of conditions a) and b), respectively, the length of too short links is suitably increased and too long links are divided into a suitable number of equally long segments.
We now present the solution algorithm, which is based on the numerical approximation that during each time interval $\tau\in(\tau_{i-1},\tau_i]$ flow rates are assumed to be constant, i.e. $y_a(\tau)=y_a^i$, so that cumulative flows are linear, i.e.:

$$Y_a(\tau)=Y_a^{i-1}+\frac{(\tau-\tau_{i-1})\cdot(Y_a^i-Y_a^{i-1})}{\tau_i-\tau_{i-1}}.$$

function GLTM
for each $a\in A$ * initialization
    $F_a^{\,0}=0$, $E_a^{\,0}=-L_a\cdot f_a^{\,0}\cdot \mathrm{v}_a°(f_a^{\,0}),\ D_a^{\,0}=S_a^{\,0}$
    for $i=0$ to $2\cdot\lambda$: $H_a^{\,i}=\infty$, $G_a^{\,i}=\infty$
    $j_a=0$, $u_a=\mathrm{t}_0+L_a/\mathrm{w}_a°\left(f_a^{\,0}\right)$, $H_a=F_a^{\,0}+L_a\cdot\Delta\mathrm{k}_a°(f_a^{\,0})$
    loop until $\tau_{ja}>u_a$: $H_a^{\,ja}=H_a-f_a^{\,0}\cdot(u_a-\tau_{ja})$, $j_a=j_a+1$
    $h_a=0$, $z_a=\mathrm{t}_0-L_a/\mathrm{w}_a^{\,+}\left(f_a^{\,0}\right),\ G_a=E_a^{\,0}+L_a\times J_a+L_a\cdot\Delta\mathrm{k}_a^{\,+}(e_a^{\,0})$
    loop until $\tau_{ha}>z_a$: $G_a^{\,ha}=G_a-e_a^{\,0}\cdot(z_a-\tau_{ha}),\ h_a=h_a+1$
for $m=1$ to υ * for each look-ahead window in chronological order
    for $i=1$ to λ * for each time interval in chronological order
        for each $a\in A$
            * compute sending flows
            $s_a^{\,i}=\max\{0,\ \min\{\left(H_a^{\,i}-E_a^{\,i-1}+D_a^{\,i}\right)/\Delta\tau,\ \eta_a^{\,i}\cdot\Phi_a\}\}$
            * compute receiving flows
            $r_a^{\,i}=\max\{0,\ \min\{\left(G_a^{\,i}-F_a^{\,i-1}-D_a^{\,i}\right)/\Delta\tau,\ \mu_a^{\,i}\cdot\Phi_a\}\}$
        for each $n\in N$ call NP * compute each node model
        for each $a\in A$
            call FP , call BP * propagate waves
            $F_a^{\,i}=F_a^{\,i-1}+f_a^{\,i}\cdot\Delta\tau,\ E_a^{\,i}=E_a^{\,i-1}+e_a^{\,i}\cdot\Delta\tau$ * cumulative flow update

for each $a \in A$ * switch to the next the look-ahead window

$F_a^0 = F_a^\lambda,\ E_a^0 = E_a^\lambda,\ D_a^0 = D_a^\lambda$

$j_a = j_a - \lambda,\ u_a = u_a - \lambda \cdot \Delta\tau,\ h_a = h_a - \lambda,\ z_a = z_a - \lambda \cdot \Delta\tau$

for $i = 0$ to λ: $H_a^i = H_a^{\lambda+i},\ H_a^{\lambda+i} = \infty,\ G_a^i = G_a^{\lambda+i},\ G_a^{\lambda+i} = \infty$

call ET * compute exit times

In the initialization, S_a^0 is the initial vertical queue and f_a^0 is the initial hypocritical flow. Consistently with the formulation where the sending and receiving flows depend on previous inflows and outflows, in the above calculation the vertical queue and the vertical storage are evaluated at the initial instant τ_{i-1} of the *i*-th time interval currently processed, thus yielding:

$$S_a^{i-1} + H_a^i - H_a^{i-1} + D_a^i - D_a^{i-1} = H_a^{i-1} - E_a^{i-1} + D_a^{i-1} + H_a^i - H_a^{i-1} + D_a^i - D_a^{i-1} = H_a^i - E_a^{i-1} + D_a^i,$$

$$R_a^{i-1} + G_a^i - G_a^{i-1} - D_a^i + D_a^{i-1} = G_a^{i-1} - F_a^{i-1} - D_a^{i-1} + G_a^i - G_a^{i-1} - D_a^i + D_a^{i-1} = G_a^i - F_a^{i-1} - D_a^i$$

where $H_a^i - H_a^{i-1}$ is the number of vehicles reaching the end point during the *i*-th time interval, if no queue is occurring, and $G_a^i - G_a^{i-1}$ is the vehicle space available to enter the link during the *i*-th time interval, if spillback is occurring. Based on condition a), no vehicle that entered the link during the *i*-th time interval can exit during the same interval, and no vehicle space freed during the *i*-th time interval at the end point of the link is available during the same interval at the initial point, therefore, H_a^i and G_a^i are already known at this moment of the computation, despite f_a^i and e_a^i are still to be calculated.

The numerical solution of equations (3.14-16) and (3.17-19) can be easily addressed under the assumption that the inflows and outflows, respectively, are constant in each time interval, and that the resulting discontinuities yield a fan of kinematic waves. In this case, to the constant flow q_i during the interval $(\tau_{i-1}, \tau_i]$ at one terminal point corresponds a linear cumulative flow at the other terminal point, that is a segment in the time-vehicles plane between the points

$$(\tau^0 = \mathrm{t}(\tau_{i-1},\ q_{i-1}),\ Q^0 = \mathrm{Q}(\tau_{i-1},\ q_{i-1})) - (t^1 = \mathrm{t}(\tau_i,\ q_{i-1}),\ Q^1 = \mathrm{Q}(\tau_i,\ q_{i-1})),$$

where functions $\mathrm{t}(\tau,\ q)$ and $\mathrm{Q}(\tau,\ q)$ express (3.14) or (3.17) and (3.15) or (3.18), respectively, for the two cases. If we connect these segments through additional segments between the points $(t^1, Q^1) - (t^2 = \mathrm{t}(\tau_i, q_i), Q^2 = \mathrm{Q}(\mathrm{t}_i,\ q_i))$, then $H(\tau)$ or $G(\tau)$ can be obtained as the minimum number of vehicles among the values taken at time τ by the segments that are defined at such instant. It is worth pointing out that connecting the segments through straight lines implies an approximation, since the points $(\mathrm{t}(\tau_i,\ q_i), \mathrm{Q}(\tau_i,\ q_i))$ and for $q \in [q_i,\ q_{i+1}]$ form actually a curve in the time-vehicles plane.

For the generic time interval i, the following algorithm scans first the additional segment between the points $(t^0, Q^0) - (t^1,\ Q^1)$ seeking for any τ_j such that $t^0 \le \tau_j \le t^1$, and then the segment between the points $(t^1, Q^1) - (t^2,\ Q^2)$ seeking for any τ_j such that $t^1 \le \tau_j \le t^2$. While scanning, if the value taken at τ_j by the segment under analysis is lower than the current estimate of the cumulative flow $H(\tau_j)$ or $G(\tau_j)$, respectively, then the latter is updated to the former; this way, at the end of the two procedures $H(\tau)$ and $H(\tau)$ are at each τ the lower envelope of all above segments; elsewhere a piecewise linear approximation is considered.

*function FP   * forward propagation of the inflow*

$j = j_a,\ t^0 = u_a,\ Q^0 = H_a$

$t^1 = \tau_{i-1} + L_a / \mathrm{w}_a°\left(f_a^{\ i} \right),\ t^2 = \tau_i + L_a / \mathrm{w}_a°\left(f_a^{\ i} \right)$

$Q^1 = F_a^{\ i-1} + L_a \cdot \Delta \mathrm{k}_a°(f_a^{\ i})\ ,\ Q^2 = F_a^{\ i-1} + f_a^{\ i} \cdot \Delta\tau + L_a \cdot \Delta \mathrm{k}_a°(f_a^{\ i})$

if $t^1 \le t^0$ then

loop until $\tau_j \ge t^1$

$Q = Q^0 + (\tau_j - t^0) \cdot \left(Q^1 - Q^0 \right) / \left(t^1 - t^0 \right),$ if $H_a^{\ j} > Q$ then $H_a^{\ j} = Q,\ j = j+1$

else

$j = j-1$

loop until $\tau_j \le t^1$

$Q = Q^0 + (\tau_j - t^0) \cdot \left(Q^1 - Q^0 \right) / \left(t^1 - t^0 \right),$ if $H_a^{\ j} > Q$ then $H_a^{\ j} = Q,\ j = j-1$

$j = j+1$

loop until $\tau_j \ge t^2$

$Q = Q^1 + (\tau_j - t^1) \cdot \left(Q^2 - Q^1 \right) / \left(t^2 - t^1 \right),$ if $H_a^{\ j} > Q$ then $H_a^{\ j} = Q,\ j = j+1$

$j_a = j\ ,\ u_a = t^2,\ H_a = Q^2$

*function BP * backward propagation of the outflow*

$j = h_a,\ t^0 = z_a,\ Q^0 = G_a$

$t^1 = \tau_{i-1} - L_a / \mathrm{w}_a^{\ +}\left(e_a^{\ i} \right),\ t^2 = \tau_i - L_a / \mathrm{w}_a^{\ +}\left(e_a^{\ i} \right)$

$Q^1 = E_a^{\ i-1} + L_a \cdot J_a + L_a \cdot \Delta \mathrm{k}_a^{\ +}(e_a^{\ i}),\ Q^2 = E_a^{\ i-1} + e_a^{\ i} \cdot \Delta\tau + L_a \cdot J_a + L_a \cdot \Delta \mathrm{k}_a^{\ +}(e_a^{\ i})$ if $t^1 \le t^0$ then

loop until $\tau_j \ge t^1$

$Q = Q^0 + (\tau_j - t^0) \cdot \left(Q^1 - Q^0 \right) / \left(t^1 - t^0 \right),$ if $G_a^{\ j} > Q$ then $G_a^{\ j} = Q,\ j = j+1$

else

$j = j-1$

loop until $\tau_j \le t^1$

$Q = Q^0 + (\tau_j - t^0) \cdot \left(Q^1 - Q^0 \right) / \left(t^1 - t^0 \right),$ if $G_a^{\ j} > Q$ then $G_a^{\ j} = Q,\ j = j-1$

$j = j+1$

loop until $\tau_j \ge t^2$

$Q = Q^1 + (\tau_j - t^1) \cdot \left(Q^2 - Q^1 \right) / \left(t^2 - t^1 \right),$ if $G_a^{\ j} > Q$ then $G_a^{\ j} = Q,\ j = j+1$

$h_a = j,\ z_a = t^2,\ G_a = Q^2$

The node model is implemented by means of two loops, the external one cycling until the receiving flows are fully exploited, the internal one cycling until the receiving flows still available are properly shared. At the end of each external loop all the flow variables are reduced to take into account the vehicles that already got through the node. The interesting feature of this approach is that at most $|FS(x)|$ external loops and $|BS(x)|$ internal loops are needed, since at least one receiving flow is fully exploited and one sending flow exits the spillback set, respectively.

*function NP * node propagation of flows*

* initialization of flows and auxiliary variables

for each $b \in FS\left(x \right)$

$f_b^{\ i} = 0$

$r_b = r_b^{\ i}$ * receiving flow to be shared among the BS

for each $a \in BS(x)$

$e_a^{\ i} = 0$

$s_a = s_a^{\ i}$ * sending flow to be served by the FS

for each $b \in FS(x)$

$f_{ab}^{\ i} = 0$

$s_{ab} = s_a^{\ i} \cdot p_{ab}^{\ i}$ * the turn sending flow depends on the slipping rate

$\Phi_{ab} = \Phi_{ab}^{\ i}$ * the turn capacity

$\varphi_a = \varphi_{ab}^{\ i} \cdot \eta_a^{\ i} \cdot \Phi_a, \psi_a = \eta_a^{\ i} \cdot \Phi_a$

$LFS = 0$ * external loop, until the receiving flows are fully exploited

loop until $LFS = 1$

$LFS = 1$

for each $b \in FS\left(x\right)$

* r_b can remain slightly positive for numerical approximations

if $r_b < 1/\infty$ then

for each $a \in BS(x)$

* link a gets closed, since link b cannot accept any more flow from it

if $s_a > 0$ and $s_{ab} > 0$ then $s_a = 0$

else

$r = r_b$ * initialize the receiving flow still available

$\varphi = 0,\ \psi = 0$ * initialize the total priority

for each $a \in BS(x)$

if $s_a > 0$ then

if $s_{ab} > 0$ then

$r_{ab} = 0$ * initialize the share of receiving flow

* initialize the sending flow to be still served

$q_a = \min\{s_{ab},\ \Phi_{ab}\}$

$\omega_a = 0$ * add link a to the spillback set

$\varphi = \varphi + \varphi_a,\ \psi = \psi + \psi_a$ * add link a to the total priority

else

$\omega_a = 1$

else

$\omega_a = 1$

* internal loop, until the receiving flow still available is properly shared

$LBS = 0$

loop until $LBS = 1$ or $\psi = 0$

$LBS = 1$

$\theta = \varphi$

if $\theta > 0$ then $\xi = r / \varphi$ else $\xi = r / \psi$

for each $a \in BS(x)$

if $\omega_a = 0$ then

* the share is proportional to priorities

if $\theta > 0$ then $\Delta r = \xi \times \varphi_a$ else $\Delta r = \xi \cdot \psi_a$

if $\Delta r \geq q_a$ then

$r = r - q_a$

$r_{ab} = r_{ab} + \Delta r$

$q_a = 0$

$\omega_a = 1$ * remove link a from the spillback set

* remove link a from the total priority
$\varphi = \varphi - \varphi_a, \ \psi = \psi - \psi_a$
* loop again, since $\Delta r - q_a$ is unused
if $\Delta r > q_a$ then $LBS = 0$
else
$r = r - \Delta r$
$r_{ab} = r_{ab} + \Delta r$
$q_a = q_a - \Delta r$
for each $a \in BS(x)$
if $s_a > 0$ then
$\rho_a = \min\{1, \ \Phi_{ab} / s_{ab}, \ r_{ab} / s_{ab} : \ b \in FS(x), \ s_{ab} > 0\}$
if $\rho_a = 1$ then
* close link a, since its remaining sending flows can be served
$s_a = 0$
else
* loop again, since on some $b \in FS(x)$ there is still receiving flow
$LFS = 0$
if $\rho_a > 0$ then
for each $b \in FS(x)$
if $s_{ab} > 0$ then * update the receiving and sending flows
$r_b = r_b - \rho_a \cdot s_{ab}$
$s_{ab} = s_{ab} - \rho_a \cdot s_{ab}$
$\Phi_{ab} = \Phi_{ab} - \rho_a \cdot s_{ab}$
$f_{ab}{}^i = f_{ab}{}^i + \rho_a \cdot s_{ab}$
else
$\rho_a = 1$
for each $a \in BS(x)$
for each $b \in FS(x)$
$f_b{}^i = f_b{}^i + f_{ab}{}^i$
$e_a{}^i = e_a{}^i + f_{ab}{}^i$

The ψ variables are introduced only to substitute the priorities when all those remaining in the spillback set are zero, which may happen in some particular cases, e.g. in roundabouts.
It is very interesting from the point of view of computational resources that, to implement the above algorithms, the only vectors that must actually have a temporal dimension in memory are H_a^i and G_a^i; indeed $f_a{}^i, e_a{}^i, F_a{}^i, E_a{}^i, s_a{}^i, r_a{}^i$ are needed only at the current time, while $D_a{}^i, \eta_a{}^i, \mu_a{}^i, \Phi_{ab}{}^i, \varphi_{ab}{}^i, p_{ab}{}^i$ are input of the model and can be retrieved on the fly from given temporal profiles.
Once the whole simulation is completed, the exit times can be computed based on the cumulative inflows and outflows that have been memorized through the following algorithm.

*function ET * compute exit times*
for each $a \in A$
$t_a{}^0 = \tau_0 + L_a / V_a$
$j = 1$
for $i = 1$ to $\lambda \cdot \nu$

until $E_a^{\ j} - D_a^{\ j} \geq F_a^{\ i}$ do $j = j+1$

if $F_a^{\ i} = F_a^{\ i-1}$ then

$t_a^{\ i} = \max\{\tau_i + L_a / V_a ,\ t_a^{\ i-1}\}$

else

$$t_a^{\ i} = \tau_{j-1} + \left(F_a^{\ i} - E_a^{\ j-1} + D_a^{\ j-1}\right) \cdot (\tau_j - \tau_{j-1}) / \left(E_a^{\ j} - D_a^{\ j} - E_a^{\ j-1} + D_a^{\ j-1}\right)$$

The cycle aims at finding, for each instant τ_i in chronological order, the earliest instant τ_j such that $E_a^{\ j-1} < F_a^{\ i} \leq E_a^{\ j}$. Since the outflow is by definition constant during the interval (τ_{j-1} , τ_j], the cumulative outflow increases linearly with slope $\left(E_a^{\ j} - E_a^{\ j-1}\right) / (\tau_j - \tau_{j-1})$. Therefore, in the general case where $F_a^{\ i} > F_a^{\ i-1}$, the exit time t_a^i results from a simple proportion. In the particular case where no vehicle enters the link in the interval $(\tau_{i-1}, \tau_i]$, the exit time t_a^i may be undetermined; it is thus set by definition as the maximum between the free flow exit time, given by $\tau_i + L_a / V_a$, and the exit time t_a^{i-1}.
We now summarize the main features of the proposed algorithm. In terms of memory usage, which is the critical issue, the GLTM only requires to store two link vectors for each instant of two look-ahead time windows. Moreover, the duration of the latter can be freely chosen by the modeler, since too long links are automatically split into segments. However, if one reduces too much the look-ahead time window, no memory gain is obtained since the number of link segments increases accordingly, while numerical noise is added to the simulation. In terms of computation efficiency, the GLTM has a linear complexity, given by the product between the number of links segments and the number of simulation intervals. Just to give an idea of its performances, the procedure runs on the whole Regione Emilia-Romagna (2,400 nodes and 6,800 links) for the whole day (14,400 time intervals of 6 sec each) in only 2 min.
The GLTM can well be implemented with splitting rates specific for each destination, but this in not really convenient, since the equilibrium issue is more efficiently addressed through the DUE model, as explained in the next section.

7. COMPARISON BETWEEN THE GLTM AND DUE

These two models share the same theoretical assumptions, but are based on a diametric vision of the problem: the first one solves the CDNL for each temporal layer in chronological order, but requires to satisfy the acyclic rule on time discretization, i.e., any vehicle entering a link in a given interval will exit no earlier than the next interval; the second one considers the reciprocal relations between link temporal profiles of flows and travel times, but requires iterating until convergence of a fixed point problem is reached. Specifically, any iteration of DUE has a complexity equal to the complete run of the GLTM divided by the scale factor of interval duration. Moreover, DUE in presence of path alternatives achieves the equilibrium jointly with the solution to the resulting CDNL problem.
By aggregating on large time intervals the results of the GLTM, which is a more robust model, we can determine the level of approximation to the solution of the CLDN provided by the DUE algorithm, which is a more efficient model when we are interested in the within-day evolution of traffic every say 10 minutes rather than in a detailed description of macroscopic flows second-by-second. The comparison between the two methods is carried out based on the solution to the CDNL problem for an extra-urban corridor of 10 km (the Venezia-Mestre belt, in Italy), where no path choice is possible, so that no equilibrium issue arises. The GLTM is first applied with time intervals of 1 sec.

Table 1: Standard deviation of DUE with respect to the GLTM.

$\Delta\tau$ (sec) *aggregation*	N (veh) *deviation*	t (sec) *deviation*
6	6.62	15.10
60	7.71	15.48
600	7.84	16.00

Then the DUE is applied to the same dataset for increasing time intervals of 6, 60 and 600 sec, with the aim of computing for each experiment the distribution of the absolute differences in terms of inflow and travel time for every link and time interval (see Table 1). This way it is possible to evaluate the amount of numerical noise induced by the linear approximations introduced into the DUE model for long time intervals.
The differences are contained into an acceptable range, but more important, the effect of linearization for long time intervals is very small, which provides a valid justification for the use of DUE when only aggregate results are needed.

8. CONCLUSIONS

Although the applications of the GLTM presented in the current paper are out of the scope of the work at hand, it's worth knowing that this assignment method has been employed successfully as a simulation engine in the solution of a signal synchronization problem based on a genetic algorithm, which by its nature requires many fast runs of the black box, and to extend in space and time the traffic measured by probe vehicles in a travel time estimation problem, which requires short term predictions on large congested networks.
The model proved to be flexible, reliable and easy to calibrate, as well as efficient in the use of memory and CPU. It is then recommended when route choice is not crucial or elastic, otherwise DUE is preferable, since it is specifically designed for the equilibrium problem. However, also in the latter case, it is very useful to run the GLTM at the end of DUE based on the equilibrium splitting rates, so as to achieve a fine grained solution of the resulting CDNL.

REFERENCES

Arnott R., De Palma A., Lindsey R. (1990), Departure time and route choice for the morning commute, *Transportation Research B* **24**, 209-228.

Astarita V. (1996), A continuous time link model for dynamic network loading based on travel time function, in *Proceedings of the 13th International Symposium on the Theory of Traffic Flow*, Lyon, 87-102.

Bellei G., Gentile G., Papola N. (2005), A within-day dynamic traffic assignment model for urban road networks, *Transportation Research Part B* **39**, 1-29.

Daganzo C.F. (1994), The cell transmission model: a dynamic representation of highway traffic consistent with hydrodynamic theory, *Transportation Research B* **28**, 269-287.

Daganzo C.F. (1995), The cell transmission model, part II: network traffic, *Transportation Research B* **29**, 79-93.

Friesz T.L., Bernstein D., Smith T.E., Tobin R.L., Wie B.W. (1993), A variational inequality formulation of the dynamic network user equilibrium problem, *Operations Research* **41**, 179-191.

Fusco G., Gentile G. (2008), Synchronization of urban arteries through plan selection, research project financed by ATAC S.p.A., available in Italian as an internal report.

Gentile G., Meschini L. (2008), The Optima Project, industrial research financed by Regione Umbria, available in Italian as an internal report.

Gentile G., Meschini L., Papola N. (2005), Macroscopic arc performance models with capacity constraints for within-day dynamic traffic assignment, *Transportation Research B* **39**, 319-338.

Gentile G., Meschini L., Papola N. (2007), Spillback congestion in dynamic traffic assignment: a macroscopic flow model with time-varying bottlenecks, *Transportation Research B* **41**, 1114-1138.

Gentile G., Nöekel K., Meschini L. (2006), Time and space discretization in dynamic traffic assignment models, in *Proceedings of the First International Symposium on Dynamic Traffic Assignment – DTA 2006*, Leeds, England.

Gentile G., Papola A. (2006), An alternative approach to route choice simulation: the sequential models, in *Proceedings of the European Transport Conference 2006 – ETC 2006*, Strasbourg, France.

Gentile G., Papola N. (2007), Section 7.5: Dynamic traffic assignment with non separable link cost functions and queue spillovers, to appear in Cascetta E. *Transportation systems engineering: theory and methods*.

Ghali M.O., Smith M.J. (1993), Traffic assignment, traffic control and road pricing, in *Transportation and Traffic Theory*, ed. C.F. Daganzo, Elsevier, Amsterdam, The Netherlands, 147-169.

Han S-J., Heydecker B.G. (2006) Consistent objective and solutions of dynamic user equilibrium models, *Transportation Research B* 40, 16-34.

Heydecker B.G., Addison J.D. (1998) Analysis of traffic models for dynamic equilibrium traffic assignment, in *Transportation networks: recent methodological advances*, ed. Bell M.G.H., Pergamon, Oxford, UK, 35-49.

Messmer A., Papageorgiou M. (1990), METANET: a macroscopic simulation program for motorway networks, *Traffic Engineering & Control* **31**, 466-470.

Newell G.F. (1993), A simplified theory of kinematic waves in highway traffic, part I: general theory; part II: queuing at freeway bottlenecks; part III: multi-destination flows, *Transportation Research B* **27**, 281-313.

Ran B., Rouphail N.M., Tarko A., Boyce D.E. (1997), Toward a class of link travel time functions for dynamic assignment models on signalised networks, *Transportation Research B* **31**, 277-290.

Tong C.O., Wong S.C. (2000), A predictive dynamic traffic assignment model in congested capacity-constrained road networks, *Transportation Research B* **34**, 625-644.

Xu Y.W., Wu J.H., Florian M., Marcotte P., Zhu L.H. (1999), Advances in the continuous dynamic network loading problem, *Transportation Science* **33**, 341-353.

Yperman I. (2007), The Link Transmission Model for dynamic network loading, PhD Thesis, Katholieke Universiteit Leuven, Belgium.