The Network Kinematic Waves Model: A Simplified Approach to Network Traffic

Daiheng Ni, John D. Leonard II, and Billy M. Williams

Daiheng Ni, Ph.D.
Department of Civil and Environmental Engineering
University of Massachusetts Amherst
Marston Hall 219
130 Natural Resources Road
Amherst, MA 01003-5205
Phone: (413)-545-5408
Fax: (413)-545-2840
ni@ecs.umass.edu

John D. Leonard II, Ph.D.
School of Civil & Environmental Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0355
Phone: (404) 894-2360
Fax: (404) 894-2278
john.leonard@ce.gatech.edu

Billy M. Williams, Ph.D.
Department of Civil Engineering,
North Carolina State University,
Raleigh, NC 27695-7908.
Phone: 919-515-7813
billy_williams@ncsu.edu
ABSTRACT

Flow of traffic on freeways and limited access highways can be represented as a series of kinematic waves. Solutions to these systems of equations become problematic under congested traffic flow conditions, and under complicated (real-world) networks. A simplified theory of kinematics waves (KWaves) was previously proposed. Simplifying elements includes translation of the problem to moving coordinate system, adoption of triangular speed-density relationships, and adoption of restrictive constraints at the on- and off-ramps. However, these simplifying assumptions preclude application of this technique to most practical situations. By directly addressing the limitations of the original theory, this paper proposes a simplified kinematic waves model for network traffic (N-KWaves). Several key constraints of the original theory are relaxed. For example, the original merge model, which gives full priority to on-ramp traffic, is relaxed and replaced with a capacity-based weighted queuing (CBWFQ) merge model. The original diverge model, which blocks upstream traffic as a whole when a downstream queue exceeds the diverge, is also relaxed and replaced with a contribution-based weighted splitting (CBWS) diverge model. Based on the above, the original theory is reformulated and extended to address network traffic. Central to the N-KWaves model is a 5-step computational procedure based on a generic building block. It is assumed that a freeway network can be represented by the combination of some special cases of the generic building block. An empirical field study showed satisfactory results. The N-KWaves model is best suited for modeling traffic operation in a regional freeway network and has a strong connection to Intelligent Transportation Systems (ITS).

Keywords: Macroscopic model, Kinematic waves, Freeway networks, Traffic simulation, ITS.
THE NEEDS FOR TRAFFIC SIMULATION IN ITS

Proactive Intelligent Transportation Systems (ITS) requires predictive input from traffic simulation. The roles that traffic simulation plays in ITS can be described as follows.

First, ITS calls for a tool to link discrete observations and generate a full picture. For example, many states have deployed automated traffic surveillance systems which constantly monitor their regional transportation networks. However, field observations are disjunct and a traffic manager may find it helpful to have a full picture covering the whole area including both observed and unobserved locations. Compared with mathematical/statistical interpolation/extrapolation techniques, traffic simulation provides a sounder ground and a more productive means to infer the unknown’s from the known’s. For another example, a traffic manager may be better equipped to make decision if the links in the system are color-coded indicating their current states (such as flow, speed, and density). As a result, the traffic manager can easily identify bottlenecks, congestion, emergency vehicle routing, etc. Traffic simulation can offer help in this regard, too, because system- and link-wide measures of effectiveness (MOEs) are standard output of virtually all traffic simulation models. Combined with proper visualization technique, the above graphical indication is easy to realize.

Second, ITS calls for an on-line tool that functions as a real-time simulation engine. For example, the following automated system may be attractive to many traffic managers. Field data collected by an automated traffic surveillance system are fed into an on-line simulation engine. The simulation engine evaluates system performance and computes MOEs in real time. Outputs are sent to a wall map for live update or Internet for public consumption. By this means, traffic managers are able to “preview” system operation, identify potential problems in advance, and initiate actions to prevent the problems or minimize their impacts.
Third, ITS calls for a real-time simulation tool that allows insertion of new data at any time. For example, traffic managers will have better chances to make educated decisions if they are enabled to test assumptions and compare alternatives such as incident management alternatives, ramp metering strategies, and traffic diversion plans. Whenever a problem is identified and actions have to be taken, traffic managers can always see the effect of the proposed actions by pausing or rolling-back the simulation, modifying system input, inserting new data, and viewing the result immediately.

Unfortunately, conventional microscopic simulation models are not ideal to fulfill these tasks. This is because microscopic models involve lengthy setup and off-line update. It is therefore desirable to develop a traffic simulation model that can potentially support the above functionalities for ITS. This is one of the major motivations of this paper.

In this paper, we develop a simplified kinematic waves model for network traffic (abbreviated as N-KWaves thereafter). This paper is organized as follows. First, we present N-KWaves in a historical context, highlighting its relation and connection to existing works. Then, we develop the N-KWaves model following a logical order. Next, we perform empirical test on the model and show test results. Following that, we discuss model implementation and its applications in ITS. Finally, we summarize the findings and draw conclusions.

**HISTORICAL CONTEXT**

This paper proposes the N-KWaves model which is an extension of Newell’s simplified theory of kinematic waves (abbreviated as “KWaves” thereafter; Newell 1993a, 1993b, and 1993c). KWaves is a special case of the so-called LWR model (Lighthill and Whitham 1955;
Richards 1956). LWR stems from hydrodynamics. Figure 1 illustrates the above chain of relation and the position of N-KWaves in this historic context. In this figure, we also show relevant analytical models (represented by boxes with bold borders) and procedural/discrete models (represented by boxes with thin borders) in the family of continuum flow models in relative chronological order. Starting from hydrodynamics, a family of continuum flow models is derived based on the law of conservation and assumptions on flow-speed-density relationship. In the family are first-order models (listed in the left pane) such as LWR model (Lighthill and Whitham 1955; Richards 1956), Bick and Newell (1960), and Munjal and Pipes (1971) and high-order models (listed in the right pane) such as Payne (1971) and Whitham (1974), Phillips (1979), Kühne (1984, 1989), Kerner and Konhäuser (1993), Michalopoulos, et al (1993), Zhang (1998), and Treiber, et al. (1999). Payne (1971) and Whitham (1974) are also derived from Prigogine (1961) which does not seem to have a connection with the conservation law. As far as solution procedures to the continuum models are concerned, Shock Wave is given by the LWR model. KRONOS (Michalopoulos 1984), KWaves (Newell 1993a, 1993b, and 1993c), and CTM (Daganzo 1994, 1995a, 1995b, 1999) are derived from the LWR model. FREFLO (Payne 1979) is derived from Payne (1971) and Whitham (1974). FREQ (May, et al 1991; May 1998) and CORQ (Yagar 1975) do not seem to have any connection with other models or solutions, so they are just placed somewhere in the middle of the figure.

Several endeavors are identified which enrich the KWaves model. Leonard (1997) coded the model into software GTWaves, which bridged the model and its application. Hurdle and Son (Son 1996; Hurdle and Son 2000) conducted a validation of the model and tested its accuracy as well as the adequacy of its underlying assumption, the triangular flow-density relation, with real data collected from freeways in the San Francisco Bay Area. The test results supported the
validity of the model, and showed that the model works best under over-saturated conditions.
Banks (2000) proposed an on-ramp queuing model by constraining ramp departure by arrival, capacity, and metering rate. This model implicitly assumes that ramp traffic is always dictated by forward waves and backward waves never back onto the ramp, an assumption that is always true for meter-controlled on-ramps.

MODEL DEVELOPMENT

In this section, we present the development of the N-KWaves model. The presentation follows a logical order. First we review and summarize the KWaves model which is the starting point of the N-KWaves model. After this, we discuss a merge model and a diverge model that are necessary for N-KWaves to address network traffic. Then, we represent a freeway network by means of links and nodes and define associated notations to be used in subsequent discussion. After the above preparation, the N-KWaves model is formulated in a manner to facilitate computerized implementation. At the end of this section, we briefly relate the N-KWaves model to Daganzo’s CTM model and discuss our contributions.

Newell’s Original Work – The KWaves Model

In KWaves, Newell (1993a, 1993b, and 1993c) worked on a special case of the LWR model. Newell showed that, if the underlying flow-density relationship can be simplified as a triangular one, the propagation of kinematic waves (disturbances in traffic states) on a freeway can be solved by simply translating cumulative curves of vehicle counts. In light of this,
upstream demand will arrive at a downstream location after a trip time assuming free flow speed. Downstream congestion will spill back to an upstream location after some time delay. At any point of the freeway, the actual number of vehicles that are allowed to pass is the minimum of multiple solutions including those determined by upstream demand, local capacity, and downstream congestion. To address multiple-destination flows, Newell made three additional assumptions. The first deals with ramp entering traffic, the second mainline traffic, and the third ramp exiting traffic. The first assumption gives full priority to ramp entering traffic so that it can bypass any queue and experience no delay at the merge. The second assumes that mainline traffic experiences the same travel time within a homogeneous section regardless of destination. The third stipulates that exiting traffic can always exit if it has successfully arrived at the diverge.

Newell’s work, though simple and elegant, left two issues unresolved. First, it is proposed in an analytical form and facilitates manual solution, i.e. it is not suited for computerized implementation which is desirable in practical use. Given the few studies (Son 1996; Hurdle and Son 2000; Leonard 1997, 1998; Banks 2000) which have implemented the KWaves model, there is still a lack of well-documented procedure in public domain that is reproducible by interested parties. To serve such a purpose, we developed a 5-step procedure which summarizes the core idea of the KWaves model and can be easily implemented in computer. Limited by space, we skip the details of the procedure in this paper and we encourage interested readers to refer to Chapter 3 of Ni (2004) for a thorough treatment. Further, we will follow the 5-step procedure later when we formulate the N-KWaves model. Second, Newell’s presentation is confined in freeway mainline only. Though KWaves model allows traffic entering from on-ramps and exiting via off-ramps, Newell’s assumptions regarding on- and off-ramps
essentially preclude queuing on ramps. As a result, the KWaves model is incapable of modeling traffic in a network. Such a limitation greatly limits the practical appeal of the KWaves model considering that addressing traffic problems from a network-wide perspective is a more practical situation. Such a consideration is another major motivation of this paper. We believe that Newell’s simple model, if enabled network capability, will provide the profession yet another efficient tool to solve real world traffic problems.

**Merge and Diverge Models**

The key issue toward network applications of the KWaves model is its lack of ramp queuing models. To relax the assumptions in the KWaves model and allow modeling network traffic, we proposed a capacity-based weighted queuing (CBWFQ) merge model (Ni and Leonard 2005) to address queuing on on-ramps and a contribution-based weighted splitting (CBWS) diverge model (Ni and Leonard 2004) to address queuing on off-ramps. The paragraphs that follow briefly introduce the two models. Interested readers are encouraged to refer to the above literature as well as Ni (2004) for detailed information.

When two streams of traffic flow into one, a merge forms. Traffic on the two upstream branches competes for resources (e.g. capacity) at the downstream branch. A wealth of literature of merge models has been identified and documented in (Ni and Leonard 2005). These merge models, either too general or over-simplified, are not suited for our needs. The CBWFQ merge model is developed with the following considerations in mind: it should be realistic in capturing real world situation, it should be simple to implement, and it should be extensible at virtually no extra cost. At a merge, we deal with a demand-supply framework in the CBWFQ merge model.
We impose a limit on each upstream branch and the limit is the capacity of the downstream branch shared by each upstream branch proportionally to its capacity. Any demand at an upstream branch exceeding its limit will be considered only after satisfying the demand at the other upstream branch. In light of this, both upstream branches will discharge at their respective limits if demands at both branches exceed their respective limits. Otherwise, there must be at least one upstream branch whose demand is less than its limit. We satisfy this demand and allocate the remaining downstream capacity to satisfy the demand at the other upstream branch. If, however, demands at both upstream branches are less than their respective limits, both demands are satisfied simultaneously. The CBWFQ merge model was formulated in (Ni and Leonard 2005) where we showed that the model satisfies the considerations mentioned above. Similar to the CBWFQ merge model is the simplest distribution scheme proposed by Jin and Zhang (2003), but the later is formulated under different assumptions.

When one stream of traffic splits into two, a diverge forms. Queue(s) on either or both of the downstream branches can spill back onto the upstream branch, constraining discharge at the diverge point. Similar to the case of merge model, we reviewed literature and identified the need to develop a new diverge model (Ni and Leonard 2004). We developed the CBWS diverge model based on our field observations that a queue spilled back from one of the downstream branches may or may not affect traffic at the upstream branch as a whole. Our treatment is to let a downstream queue spill and extend backwards (if it does) to the corresponding portion of upstream lanes. Upstream traffic on this portion of lanes discharges at a lower rate and runs at a slower speed. In contrast to this is traffic on the other portion of the upstream lanes where the traffic is not affected by the downstream queue and can maintain higher speed and discharge rate. There is obviously an interface between the two streams of traffic at the upstream branch.
Two effects work on the interface: “inpatient drivers” who do not want to wait in queue any more and squeeze out joining the higher-speed traffic stream and “last-minute decision-makers” who want to squeeze in at the head of the queue at the cost of other drivers in queue. Therefore we apply a friction at the interface to model the effects. Depending on how strong the friction is, the operation of the two traffic streams at the upstream branch varies from flowing independently to being constrained as a whole. Suppose, after some initial treatments, we have determined the total outflow of the upstream branch and the supplies at both downstream branches, the CBWS diverge model distributes the total outflow of the upstream branch to each downstream branch proportionally to its supply. The CBWS merge model was formulated in (Ni and Leonard 2004) where we showed the validity of the above treatment.

Network Representation and Notation

Represented by a link-node structure, a freeway network consists of the following three basic building blocks: (i) a “mainline” block where one upstream link feeds traffic to one downstream link, as in part A of Figure 2; (ii) a “merge” block where two upstream links feed traffic to one downstream link, as in part B of Figure 2; (iii) a “diverge” block where one upstream link feeds traffic to two downstream links, as in part C of Figure 2. To facilitate computerized implementation, it is desirable to have a generic building block, as in part D of Figure 2, which encompasses the three basic building blocks as its special cases. It is reasonable to assume that a freeway network can be decomposed into a series of basic building blocks. Modeling the freeway network involves traversing these building blocks which are special cases of the generic one. Therefore, the following discussion is based on the generic building block.
In part D of Figure 2, node $x_i$ connects 4 links, some of which may or may not exist depending on local situations. Referenced from node $x_j$, the two incoming links are $(x_i, x_j)$ and $(x_j, x_i)$ and the two outgoing links are $(x_i, x_n)$ and $(x_i, x_m)$. The following notations are used in subsequent discussion:

$\tau$ denotes time increment which should be less than the time needed for a vehicle to traverse the shortest link at free flow speed.

$x$ denotes a node and its subscription indexes the node. For example, $x_i$ denotes the current node, $x_i$ and $x_j$ denote its 2 upstream nodes, $x_n$ and $x_m$ denote its two downstream nodes, $x_\ell$ denotes any further downstream nodes of $x_i$ via $x_n$, and $x_s$ denotes any further downstream nodes of $x_i$ via $x_m$. Nodes are sorted and indexed such that all potential origins of a node bear lower indices than the index of the node and all potential destinations of the node bear higher indices than the index of the node. On the other hand, a node remembers its adjacent nodes as well as its potential destinations. However, no implication is made on the relative order of $x_i$ and $x_j$, nor of $x_n$ and $x_m$.

$A$ and $D$ denote cumulative number of vehicles arrive and depart, respectively. Considering that a node may have multiple upstream nodes and/or downstream nodes, additional information, such as origin and destination, has to be supplied to clarify the $A$’s and $D$’s. For example, $A_{i,n}(x_i^-, t)$ denotes the cumulative arrival count past the left of
node $x_j$ at time $t$ originated from node $x_i$ and destined for node $x_n$ and beyond, as
the dotted thin line with arrow in part D of Figure 2. $D_{i,n}(x_j^-,t)$ denotes the
cumulative departure count past the right of node $x_i$ at time $t$ originated from node
$x_i$ and destined for node $x_n$ and beyond.

$a$ and $d$ have similar meaning to $A$ and $D$, respectively. The only difference is that the
lowercase letters denote vehicle counts resulted since last time step, as oppose to
cumulative counts.

$Q$ denotes capacity. For example, $Q_{il}$ denotes the capacity of link $(x_i, x_j)$.

$k$ denotes jam density. For example, $k_{il}$ denotes the capacity of link $(x_i, x_j)$.

$v$ denotes forward wave propagation speed (i.e., free flow speed under the assumption
of triangular flow-density relationship). For example, $v_{il}$ denotes the forward wave
propagation speed of link $(x_i, x_j)$.

$u$ denotes backward wave propagation speed (it is a constant under the assumption of
triangular flow-density relationship). For example, $u_{il}$ denotes the backward wave
propagation speed of link $(x_i, x_j)$.

$n$ denotes the number of lanes of a link. For example, $n_{il}$ denotes the number of lanes
of link $(x_i, x_j)$. Notice that, when $n$ appears in subscription, it means the index of a
node, as defined above.
\( l \) denotes the length of a link. For example, \( l_{ji} \) denotes the length of link \((x_i, x_j)\).

Again, when \( l \) appears in subscription, it means the index of a node, as defined above.

**The Computational Procedure**

In this subsection, we formulate the N-KWaves model. To facilitate computerized implementation, we do not follow an analytical form as in the KWaves model. Instead, we present the N-KWaves model in a computational procedure. A basic 5-step procedure has been developed to summarize Newell’s work on KWaves. Interested readers are encouraged to refer to Chapter 3 of (Ni 2005) for details. The procedure presented below follows the same steps as outlined in the basic procedure, but in an extended form.

Modeling a freeway network starts with representing the network in a link-node structure and constructing a lattice which has all the nodes on its horizontal axis and time steps with increment \( \tau \) on its vertical axis. The algorithm keeps track of the cumulative arrival and departure counts at each lattice point iteratively and progressively such that, at each time step, all nodes, from the upstream end to the downstream end, are evaluated one by one. Then system time advances one step and the above process repeats. The algorithm stops after all lattice points have been traversed. The following procedure summarizes how the algorithm works at lattice point \((x_j, t)\) assuming that all previous lattice points have been evaluated.
Step 1

The goal of this step is to determine the cumulative departure counts to the right of $x_l$. In the following discussion, “upstream”, “downstream”, “left”, and “right” are all relative to $x_l$ and $x_l$ will be skipped if doing so will not cause any confusion. There are two downstream links at the right of $x_l$, $(x_l, x_n)$ and $(x_l, x_m)$, so the cumulative departure counts to be determined are $D_{l,n}(x_l^+, t)$ and $D_{l,m}(x_l^+, t)$. $D_{l,n}(x_l^+, t)$ is constrained by the following:

**Upstream arrival**

According to Newell, the cumulative arrival count to the left of $x_l$ originated from $x_l$ destined for $x_n$ and beyond, $A_{i,n}(x_l^-, t)$, can be obtained by translating the cumulative departure curve to the right of $x_l$ destined for $x_n$ and beyond, $D_{i,n}(x_l^+, t)$, by a link travel time, $T_{il} = l_{il} / v_{il}$, assuming free flow speed. The same applies to $A_{j,n}(x_l^-, t)$. Therefore the upstream arrival, $A_{l,n}(x_l^+, t)$, is the sum of the above two, i.e.,

$$A_{l,n}(x_l^+, t) = A_{i,n}(x_l^-, t) + A_{j,n}(x_l^-, t) = D_{i,n}(x_l^+, t - l_{il} / v_{il}) + D_{j,n}(x_l^+, t - l_{jl} / v_{jl})$$

**Right capacity**

$D_{l,n}(x_l^+, t)$ is also constrained by the local capacity to the right of $x_l$, $Q_{ln}$. In cumulative terms, $D_{l,n}(x_l^+, t)$ is constrained by:

$$D_{l,n}(x_l^+, t - \tau) + \tau Q_{ln}$$
**Downstream queue**

According to Newell, $D_{l,n}(x_i^+, t)$ is also constrained by:

$$D_{l,n}(x_i^-, t - l_{in} / u_{in}) + l_{in} k_{in}$$

Newell’s minimum principle stipulates that $D_{l,n}(x_i^+, t)$ takes the minimum of the above constraints, i.e.,

$$D_{l,n}(x_i^+, t) = \min \{ A_{i,l}(x_i^+, t), D_{l,n}(x_i^+, t - \tau) + \tau Q_{lm}, D_{l,n}(x_n^-, t - l_{in} / u_{in}) + l_{in} k_{in} \}$$

In a similar fashion, the cumulative departure count to the right of $x_i$ originated from $x_i$ destined for $x_m$ and beyond can be determined as:

$$D_{l,n}(x_i^+, t) = \min \{ A_{i,m}(x_i^+, t), D_{l,m}(x_i^+, t - \tau) + \tau Q_{lm}, D_{l,m}(x_m^-, t - l_{lm} / u_{lm}) + l_{lm} k_{lm} \}$$

**Step 2**

The goal of this step is to determine the cumulative departure counts to the left of $x_i$.

There are two upstream links at the left of $x_i$, $(x_i, x_j)$ and $(x_j, x_i)$, so the cumulative departure counts to be determined are $D_{i,j}(x_i^-, t)$, and $D_{j,i}(x_j^-, t)$. $D_{i,j}(x_i^-, t)$ is constrained by the following:

**Upstream arrival**

$$A_{i,j}(x_i^-, t) = D_{i,j}(x_i^+, t - l_{il} / v_{il})$$
Left capacity

\[ D_{i,l}(x_i^-, t - \tau) + \tau \times Q_{jl} \]

Downstream departure

\[ D_{i,l}(x_i^+, t) + D_{l,m}(x_i^+, t) \]

Similarly, \( D_{j,l}(x_j^-, t) \) is constrained by the following:

Upstream arrival

\[ A_{j,l}(x_j^-, t) = D_{j,l}(x_j^+, t - l_{jl} / v_{jl}) \]

Left capacity

\[ D_{j,l}(x_j^-, t - \tau) + \tau \times Q_{jl} \]

Downstream departure

\[ D_{j,n}(x_j^+, t) + D_{j,m}(x_j^+, t) \]

In the above expressions, the unknown terms are listed below in a matrix form:

\[
\begin{bmatrix}
D_{i,n}(x_i^+, t) & D_{i,m}(x_i^+, t) \\
D_{j,n}(x_j^+, t) & D_{j,m}(x_j^+, t)
\end{bmatrix}
\] (I)

Notice that the first column of matrix I sums up to \( D_{i,n}(x_i^+, t) \) which is known from step 1, i.e.,

\[ D_{i,n}(x_i^+, t) + D_{j,n}(x_j^+, t) = D_{i,n}(x_i^+, t) \]

Similarly, the sum of the second column is also known from step 1, i.e.,
\[ D_{i,m}(x_i^+, t) + D_{j,m}(x_j^+, t) = D_{i,m}(x_i^+, t) \]

Now, the question is how to split the known terms \( D_{i,n}(x_i^+, t) \) and \( D_{i,m}(x_i^+, t) \). This is essentially to distribute a downstream supply to its upstream links. The left hand side terms can be determined by applying the CBWFQ merge model as follows. Let

\begin{align*}
S_{il}^n & \text{ be the demand of link } (x_i, x_i) \text{ waiting to be served by link } (x_i, x_n) : \\
S_{il}^n & = \min \left\{ a_{i,n}(x_i^-, t), \tau Q_{il} \right\} \text{ where } a_{i,n}(x_i^-, t) = A_{i,n}(x_i^-, t) - D_{i,n}(x_i^-, t - \tau) \\
S_{jl}^n & \text{ be the demand of link } (x_j, x_i) \text{ waiting to be served by link } (x_i, x_n) : \\
S_{jl}^n & = \min \left\{ a_{j,n}(x_i^-, t), \tau Q_{jl} \right\} \text{ where } a_{j,n}(x_i^-, t) = A_{j,n}(x_i^-, t) - D_{j,n}(x_i^-, t - \tau) \\
R_{ln} & \text{ be the supply of link } (x_i, x_n) : \\
R_{ln} & = D_{i,n}(x_i^+, t) - D_{i,n}(x_i^+, t - \tau) \\
\Delta_{in}^i & \text{ be link } (x_j, x_i) \text{’s share of downstream supply, i.e.,:} \\
\Delta_{in}^i & = R_{ln} \times \frac{Q_{il}}{Q_{il} + Q_{jl}} \\
\Delta_{in}^j & \text{ be link } (x_j, x_i) \text{’s share downstream supply, i.e.,:} \\
\Delta_{in}^j & = R_{ln} \times \frac{Q_{jl}}{Q_{il} + Q_{jl}}
\end{align*}

The CBWFQ merge model determines upstream departure based on the following rules:

Case 1: when \( S_{il}^n \leq \Delta_{in}^i \) and \( S_{jl}^n \leq \Delta_{in}^j \), the solution is:

\[ d_{i,n}(x_i^+, t) = S_{il}^n \text{ and } d_{j,n}(x_i^+, t) = S_{jl}^n \]

Case 2: when \( S_{il}^n \leq \Delta_{in}^i \) and \( S_{jl}^n > \Delta_{in}^j \), the solution is:
\[ d_{i,n}(x^+_i, t) = S^u_{ni} \text{ and } d_{j,n}(x^+_i, t) = \min\{S^u_{nj}, R_{ln} - d_{i,n}(x^+_i, t)\} \]

Case 3: when \( S^u_{ni} > \Delta^i_{ln} \) and \( S^u_{nj} \leq \Delta^i_{ln} \), the solution is:

\[ d_{i,n}(x^+_i, t) = S^u_{ni} \text{ and } d_{j,n}(x^+_i, t) = \min\{S^u_{nj}, R_{ln} - d_{i,n}(x^+_i, t)\} \]

Case 4: when \( S^u_{ni} > \Delta^i_{ln} \) and \( S^u_{nj} > \Delta^i_{ln} \), the solution is:

\[ d_{i,n}(x^+_i, t) = \Delta^i_{ln} \text{ and } d_{j,n}(x^+_i, t) = \Delta^i_{ln} \]

Therefore, the first column of matrix I is computed as:

\[ D_{i,n}(x^+_i, t) = D_{i,n}(x^+_i, t - \tau) + d_{i,n}(x^+_i, t) \]
\[ D_{j,n}(x^+_i, t) = D_{j,n}(x^+_i, t - \tau) + d_{j,n}(x^+_i, t) \]

In a similar fashion, the second column of matrix I can be determined. Therefore,

\[ D_{i,j}(x^-_i, t) \] is determined as:

\[ D_{i,j}(x^-_i, t) = \min\{A_{i,j}(x^-_i, t), D_{i,j}(x^+_i, t - \tau) + \tau \times Q_{ij}, D_{i,n}(x^+_i, t) + D_{i,m}(x^+_i, t)\} \]

and, similarly, \( D_{j,i}(x^-_i, t) \) is determined as:

\[ D_{j,i}(x^-_i, t) = \min\{A_{j,i}(x^-_i, t), D_{j,i}(x^+_i, t - \tau) + \tau \times Q_{ji}, D_{j,n}(x^+_i, t) + D_{j,m}(x^+_i, t)\} \]

Obviously,

\[ D_{i,n}(x^-_i, t) + D_{i,m}(x^-_i, t) = D_{i,j}(x^-_i, t) \text{ and } D_{j,n}(x^-_i, t) + D_{j,m}(x^-_i, t) = D_{j,i}(x^-_i, t) \]

Put the left hand side terms, which are unknown, in matrix form:

\[
\begin{bmatrix}
D_{i,n}(x^-_i, t) & D_{i,m}(x^-_i, t) \\
D_{j,n}(x^-_i, t) & D_{j,m}(x^-_i, t)
\end{bmatrix}
\]

(II)

It is easy to identify that matrix I may or may not be equal to matrix II. Determining the first row of matrix II is equivalent to splitting an upstream departure count among two
downstream diverging branches and this can be accomplished by applying the CBWS diverge model as follows.

\[
D_{i,n}(x_i^-, t) = D_{i,n}(x_i^-, t - \tau) + d_{i,n}(x_i^-, t) \quad \text{where} \quad d_{i,n}(x_i^-, t) = S \times \frac{d_{i,n}(x_i^+, t)}{d_{i,n}(x_i^+, t) + d_{i,m}(x_i^+, t)}
\]

\[
D_{i,m}(x_i^-, t) = D_{i,m}(x_i^-, t - \tau) + d_{i,m}(x_i^-, t) \quad \text{where} \quad d_{i,m}(x_i^-, t) = S \times \frac{d_{i,m}(x_i^+, t)}{d_{i,n}(x_i^+, t) + d_{i,m}(x_i^+, t)}
\]

where \( S \) denotes the outflow at link \((x_i, x_j)\):

\[
S = D_{i,j}(x_i^-, t) - D_{i,j}(x_i^-, t - \tau)
\]

Similarly, the second row of matrix II can be computed. Note that an element in matrix I and its corresponding element in matrix II refer to the same thing but are the results of viewing traffic from different sides of node \(x_i\). Here, we first determine elements in matrix I by viewing the traffic from the right side of node \(x_i\). Based on the initial results, we update them by viewing the traffic from the left side of node \(x_i\). As a result, matrix II represents the latest information and it overwrites matrix I.

**Step 3**

The goal of this step is to determine the travel times for vehicles on links \((x_i, x_j)\) and \((x_j, x_i)\) decomposed by destinations, i.e., the following matrix:

\[
\begin{pmatrix}
T_{ii}^n(t) & T_{ij}^n(t) \\
T_{ji}^n(t) & T_{jj}^n(t)
\end{pmatrix}
\]
where $T^n_{il}(t)$ denotes the travel time experienced by vehicles on link $(x_i, x_j)$ destined for node $x_n$ and beyond. $T^n_{il}(t)$ is determined by comparing curve pair $D_{i,o}(x_i^+, t)$ vs. $D_{i,o}(x_i^-, t)$ on a time-cumulative count diagram such that, starting at the point determined by current time and the value of $D_{i,o}(x_i^+, t)$ at this time, the curve $D_{i,o}(x_i^+, t)$ is traced back to an earlier time $t'$ when $D_{i,o}(x_i^+, t') = D_{i,o}(x_i^-, t)$. Then, link travel time, $T^n_{il}(t)$, is determined as:

$$T^n_{il}(t) = t - t'$$

A similar procedure applies to other link travel times $T^n_{il}(t)$, $T^n_{jl}(t)$, and $T^n_{jl}(t)$.

**Step 4**

The goal of this step is to determine the cumulative departure counts to the left of node $x_i$ destined for further downstream destinations, such as $x_r$ and $x_s$, and beyond, i.e.,

$$D_{i,r}(x_i^-, t) = D_{i,s}(x_i^-, t)$$

$$D_{j,r}(x_i^-, t) = D_{j,s}(x_i^-, t)$$

The cumulative departure count to the left of $x_i$ originated from $x_i$ destined for $x_r$ and beyond, $D_{i,r}(x_i^-, t)$, can be obtained by simply translating $D_{i,r}(x_i^+, t)$ to the right by a link travel time, $T^n_{il}(t)$, i.e.,

$$D_{i,r}(x_i^-, t) = D_{i,r}(x_i^+, t - T^n_{il}(t))$$
Similarly,

\[ D_{i,s}(x_i^-, t) = D_{i,s}(x_i^+, t - T_{jl}^m(t)) \]

\[ D_{j,s}(x_j^-, t) = D_{j,s}(x_j^+, t - T_{jl}^n(t)) \]

\[ D_{j,s}(x_i^-, t) = D_{j,s}(x_i^+, t - T_{jl}^m(t)) \]

**Step 5**

The goal of this step is to determine the cumulative departure counts to the right of node \( x_i \) destined for further downstream destinations, such as \( x_r \) and \( x_s \), and beyond, i.e.,

\[ D_{i,r}(x_i^+, t) \text{ and } D_{i,s}(x_i^+, t) \]

The cumulative departure count past the right of \( x_i \) originated from \( x_i \) destined for \( x_r \) and beyond, \( D_{i,r}(x_i^+, t) \), is simply the sum of cumulative departure curves past the left of \( x_i \) originated from all upstream nodes destined for \( x_r \) and beyond, i.e.,

\[ D_{i,r}(x_i^+, t) = D_{i,r}(x_i^-, t) + D_{j,r}(x_i^-, t) \]

Similarly,

\[ D_{i,s}(x_i^+, t) = D_{i,s}(x_i^-, t) + D_{j,s}(x_i^-, t) \]

This concludes the processing procedure at a lattice point of the time-space diagram and the algorithm is ready to proceed to the next lattice point. Note that the above discussion assumes
an explicit upstream/downstream structure which may or may not hold in a traffic network. In case of a loop, one has to break it by carefully defining boundary conditions at the break point.

**Contributions and Relation with CTM**

On discussing KWaves/N-KWaves, one can not avoid comparing it with CTM (Daganzo 1994, 1995a, 1995b, 1999). Indeed, the two models are very similar in that they both are derived from the same source, i.e. LWR model (Lighthill and Whitham 1955; Richards 1956), and they both address the same set of problems, i.e. network traffic queuing. Meanwhile, they are different in many ways. Readers are referred to Ni (2004) for a thorough treatment of the similarities and differences. Limited by space, we conceptually highlight two of the differences here and we believe they might be of primary concern to users when they choose among a set of alternative models. First, CTM is essentially cell-based. A cell is a segment of highway in which vehicles are received, stored, and sent. CTM is most accurate if the length of a cell equals to the distance traversed by a vehicle at free flow speed in one time increment, as recommended in Daganzo (1994). This poses a challenge to computation resources. For example, CTM needs to process over 1000 cells to simulate a freeway of 20 miles long with free flow speed 70 mph and time increment of 1 second. In contrast, N-KWaves is based on links and nodes. A node is a point along a freeway where roadway geometric properties or traffic characteristics change. The accuracy of N-KWaves is not sensitive to link length. Therefore, for the same problem, N-KWaves may need to process only 20 nodes if interchange density is about 1 ramp per mile. Second, CTM eventually decomposes time-varying origin-destination (O-D) table to time-varying turn percentages at diverges. As a result, traffic may be routed through a network in a
different path than specified in the O-D table. On the other hand, under congested condition, the actual turn percentage at a diverge is constrained by the congestion and may not be the same as derived from the O-D table. In contrast, N-KWaves moves traffic through the network based on full O-D paths according to the O-D table, so N-KWaves does not encounter the above problems.

We have not yet conducted numerical comparison between the two models. From a model development perspective, we felt it is of our primary interest to exam how close our model approximates the real world system. The relative strength and accuracy between our model and some other models are of secondary interest at this point.

In summary of the model development, our contributions come from the following: (i) we derived a discrete version of KWaves, i.e. the 5-step procedure, in public domain to facilitate computerized implementation; (ii) we developed a CBWFQ merge model to allow queuing at on-ramps; (iii) we developed a CBWS diverge model to allow queuing at off-ramps; (iv) we extended the KWaves model and enabled it to address network traffic.

EMPIRICAL STUDY AND RESULTS

To validate the N-KWaves model, an empirical test is carried out to compare model simulation against field observation. Both qualitative and quantitative measures are employed to assess the performance of the model.
The Study Site

ITS data collected on GA 400 by Georgia NAVIGATOR system is used to provide empirical data for testing the model. The automated traffic surveillance system covers the section of GA 400 between I-285 (to the south) and Old Milton Parkway (to the north), a stretch of road of approximately 12.56 miles and this serves as our study site. Though GA 400 does not provide completely uninterrupted flow along all its way, the section under coverage does.

Figure 3 illustrates the study site which has four on-ramps and four off-ramps. Table 1 lists the geometry and traffic characteristics data of this site. There are two bottlenecks at this site, one at the downstream of node 4001128 which is not shown in the figure and the other at node 6104. The first bottleneck is caused by insufficient capacity at somewhere downstream of node 4001128 and is simulated by adding a dummy link after node 4001128. The second bottleneck is caused by high exit volume at link 6104-4006104 as well as limited capacity at link 6104-5104.

Test Results

Data for the test was collected on Friday, Oct. 11, 2002 from 05:50:00 to 19:20:00. Origin-destination flows are synthesized from observed link traffic counts based on the recursive predictions error (RPE) estimator proposed by Nihan and Davis (1987).

Traffic density is used as the target variable to evaluate model performance and the comparison is made between observed density and simulated density. Figure 4 compares simulated (dashed lines) and observed (solid lines) congested regions in time-space domain. A
density level of 45 vehicles/mile/lane is used to delineate congested and uncongested regions. There are two peaks in the test, one in the morning and the other in the afternoon. The morning queue appears approximately between 7:00 and 8:30 and extends from the downstream end backwards up to node 6103. At about 6:50, the morning queue starts to build up passing node 4001128. About 40 minutes later, the queue reaches node 6103. The queue then starts to dissipate and reaches somewhere at the downstream of node 5103 at 8:15. In the next 15 minutes, the queue swiftly shrinks to node 4001128 and disappears before 9:05. The afternoon peak is much smaller in scale. It appears approximately from 17:15 to 17:55 and spans 3 links between nodes 5102 and 6104. There are, of course, some slight discrepancies in Figure 4 between the simulated and the observed lines. For example, the simulated morning peak builds up a little earlier and shrinks a little later than the observed one. On the other hand, the simulated afternoon peak fails to capture the minor portion of the observed afternoon peak. Given these, the figure shows a good agreement between simulated and observed congestion regions in general. Figure 5 shows simulated vs. observed flow-density relationships. The main purpose is to check where the model faithfully reproduces traffic stream characteristics at the study site. Visual inspection shows that the simulated relationship approximates the observed well. Figure 6 plots the frequency of simulation errors which involves 2592 samples. Each sample consists of an observed density value and a simulated density value during a 5-minute interval. Notice that the simulation errors are tightly centered around zero, indicating a good fit between the simulation and the observation.

Quantitative evaluation of model performance is carried out based on simultaneous statistical inference technique (Ni, et al 2004). Basically the technique tests the following hypothesis simultaneously so that both the accuracy and the precision of the model are assessed:
Hypothesis 1: the model is unbiased.

Hypothesis 2: modeling error is reasonably small.

Statistical test result is listed in Table 2. To account for autocorrelation, we adopted the batch means technique (Goldsman and Tokol 2000) using a batch size of 81 which results in 32 batches. At a significance level of 0.05, paired student t-test shows that there is insufficient evidence to reject the null hypothesis. This means that the means of the simulated result and the observed result are not statistically different. On the other hand, the variance of the simulation error ($\varepsilon$) is reasonably small. Interpreted in another way, this translates to a 95% confidence interval of (-0.09, 0.09) × 100% for percentage error. It has to be pointed out that this test pools three sources of error, i.e., the simulation error, the observation error, and the O-D estimation error. Since we are unable to separate them, all the errors are accumulated and have to be explained by simulation error. Given this, the above result is still quite satisfactory.

IMPLEMENTATION AND APPLICATIONS IN ITS

The N-KWaves model is well-suited to address the needs for traffic simulation in ITS, as outlined at the beginning of this paper. First, being a macroscopic traffic simulation model, it is an ideal tool to model traffic operation in a regional freeway network. By loading traffic demands (generated from point observations by traffic surveillance systems) into the network, the model is able to evaluate how the facility performs under the demands and output system- and link-wide MOEs. Assisted by proper visualization techniques, traffic managers are able to
view the full picture, identify problems, and develop solutions. Second, being simple in algorithm (having only five steps and involving simple operations such as addition, subtraction, and comparison), the model is light-weight and computationally efficient. This enables the model to perform faster than real time. If implemented with proper technologies such as Java and XML, the model is able to run on-line in real time. One can establish an automated system by taking as input data collected by automated surveillance system, running the model on-line as a simulation engine, and outputting simulation results instantly. Third, the nature of the model is such that it scans nodes from upstream to downstream, and then time advances one step and the above process repeats. Therefore, it is convenient to pause the simulation at any time, modify the system and/or insert new data, and continue the simulation. This enables traffic managers to test assumptions, compare alternatives, and implement control strategies. Taking real-time input from automated traffic surveillance systems and assisted by near-term traffic demand predictor, the K-Waves model functions as a simulation engine sitting at the core of ITS. The model can serve Advanced Traffic Management Systems (ATMS) by providing real-time updates on system- and link-wide MOEs and allowing testing assumptions and comparing alternatives. The model can also serve Advanced Traveler Information Systems (ATIS) by publishing point-to-point travel times, indicating mobility in the system, and suggesting alternative routes.

Currently, the N-KWaves has been implemented using Java and XML technologies. Practical application is underway after more comprehensive field tests.
CONCLUSION

As an extension to Newell’s work, this paper developed an N-KWaves model to address network traffic. Simulation of network traffic is made possible by the capacity-based weighted queuing (CBWFQ) merge model and the contribution-based weighted splitting (CBWS) diverge model. The former enables merging traffic to compete for downstream supply and applies a capacity-weighted distribution only when both merging branches are dictated by backward waves. The latter allows queues from either or both of the diverging branches to spill back and the queues affect their corresponding part of upstream traffic, instead of affecting upstream traffic as a whole.

Empirical test results showed that the model performs well even with the presence of error of other sources, such as O-D estimation error. Visual comparison shows a good agreement between simulation and observation. The model performs well in terms of delineating time-space region of congestion which is of the major interest to traffic engineers. Quantitative assessment reveals that the model is accurate with reasonably precision.

The proposed model is very simple to implement and computationally efficient. In addition, it provides the potential to run on-line in real-time and to allow insertion of new data. These properties render the model well-suited for applications in ITS.

ACKNOWLEDGMENT

The authors are grateful to an anonymous referee of this journal whose insightful comments and constructive suggestions greatly improved the quality of this paper.
REFERENCES


Tables and Figures

Tables

Table 1 Geometry and traffic characteristics data for the site

Table 2 Statistical test result

Figures

Figure 1 Taxonomy of continuum flow models

Figure 2 Generic building block of freeway networks

Figure 3 Sketch of the study site (GA 400 Southbound)

Figure 4 Comparison of simulated and observed congested regions

Figure 5 Comparison of observed and simulated flow-density relationships (5-minute bins)

Figure 6 Frequency of simulation errors
Table 1 Geometry and traffic characteristics data for the site

<table>
<thead>
<tr>
<th>Link</th>
<th>Up Node</th>
<th>Down Node</th>
<th>Length (mi)</th>
<th># of Lanes</th>
<th>Link Type</th>
<th>FFS (mi/h)</th>
<th>Capacity (veh/h/ln)</th>
<th>Jam Density (veh/mi/ln)</th>
<th>BWS (mi/h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4001104</td>
<td>6102</td>
<td>0.4</td>
<td>2</td>
<td>Mainline</td>
<td>56</td>
<td>2200</td>
<td>180</td>
<td>15.6</td>
</tr>
<tr>
<td>2</td>
<td>6102</td>
<td>5102</td>
<td>1.1</td>
<td>3</td>
<td>Mainline</td>
<td>57</td>
<td>2200</td>
<td>180</td>
<td>15.6</td>
</tr>
<tr>
<td>3</td>
<td>6103</td>
<td>5103</td>
<td>0.9</td>
<td>3</td>
<td>Mainline</td>
<td>65</td>
<td>2100</td>
<td>180</td>
<td>14.2</td>
</tr>
<tr>
<td>4</td>
<td>5103</td>
<td>6104</td>
<td>0.88</td>
<td>3</td>
<td>Mainline</td>
<td>61</td>
<td>2200</td>
<td>180</td>
<td>15.3</td>
</tr>
<tr>
<td>5</td>
<td>6104</td>
<td>5104</td>
<td>0.5</td>
<td>3</td>
<td>Mainline</td>
<td>60</td>
<td>1900</td>
<td>180</td>
<td>12.8</td>
</tr>
<tr>
<td>6</td>
<td>5104</td>
<td>6105</td>
<td>2.9</td>
<td>4</td>
<td>Mainline</td>
<td>64</td>
<td>2100</td>
<td>180</td>
<td>14.3</td>
</tr>
<tr>
<td>7</td>
<td>6105</td>
<td>5105</td>
<td>0.5</td>
<td>4</td>
<td>Mainline</td>
<td>60</td>
<td>2100</td>
<td>180</td>
<td>14.5</td>
</tr>
<tr>
<td>8</td>
<td>5105</td>
<td>4001128</td>
<td>0.5</td>
<td>4</td>
<td>Mainline</td>
<td>56</td>
<td>2200</td>
<td>180</td>
<td>15.6</td>
</tr>
<tr>
<td>9</td>
<td>4001102</td>
<td>5102</td>
<td>0.5</td>
<td>1</td>
<td>Off-ramp</td>
<td>22</td>
<td>1000</td>
<td>180</td>
<td>7.4</td>
</tr>
<tr>
<td>10</td>
<td>4006102</td>
<td>4001102</td>
<td>0.5</td>
<td>1</td>
<td>Off-ramp</td>
<td>22</td>
<td>1000</td>
<td>180</td>
<td>7.7</td>
</tr>
<tr>
<td>11</td>
<td>4005102</td>
<td>5102</td>
<td>0.5</td>
<td>1</td>
<td>Off-ramp</td>
<td>20</td>
<td>1000</td>
<td>180</td>
<td>7.4</td>
</tr>
<tr>
<td>12</td>
<td>4006103</td>
<td>4006103</td>
<td>0.5</td>
<td>1</td>
<td>Off-ramp</td>
<td>20</td>
<td>1000</td>
<td>180</td>
<td>7.7</td>
</tr>
<tr>
<td>13</td>
<td>4005103</td>
<td>5103</td>
<td>0.5</td>
<td>1</td>
<td>On-ramp</td>
<td>22</td>
<td>1250</td>
<td>180</td>
<td>10.1</td>
</tr>
<tr>
<td>14</td>
<td>4006104</td>
<td>4006104</td>
<td>0.5</td>
<td>1</td>
<td>Off-ramp</td>
<td>40</td>
<td>1200</td>
<td>180</td>
<td>8.0</td>
</tr>
<tr>
<td>15</td>
<td>4005104</td>
<td>5104</td>
<td>0.5</td>
<td>2</td>
<td>On-ramp</td>
<td>40</td>
<td>2000</td>
<td>180</td>
<td>15.4</td>
</tr>
<tr>
<td>16</td>
<td>4006105</td>
<td>4006105</td>
<td>0.5</td>
<td>1</td>
<td>Off-ramp</td>
<td>20</td>
<td>1500</td>
<td>180</td>
<td>14.3</td>
</tr>
<tr>
<td>17</td>
<td>4005105</td>
<td>5105</td>
<td>0.5</td>
<td>1</td>
<td>On-ramp</td>
<td>20</td>
<td>1700</td>
<td>180</td>
<td>17.9</td>
</tr>
</tbody>
</table>

Note:
Up / Down node – upstream / downstream node.
FFS – Free flow speed, also wave forward propagation seed
BWS – Wave backward propagation speed, determined by FFS, capacity, and jam density.
Table 2 Statistical test result

<table>
<thead>
<tr>
<th>Item</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Batch size ((m))</td>
<td>81</td>
</tr>
<tr>
<td>Batches ((b))</td>
<td>32</td>
</tr>
<tr>
<td>Total samples ((n))</td>
<td>2592</td>
</tr>
<tr>
<td>Significance level ((\alpha))</td>
<td>0.05</td>
</tr>
<tr>
<td>Grand mean ((\bar{W}))</td>
<td>-0.000937</td>
</tr>
<tr>
<td>Variance ((\hat{V}_\theta))</td>
<td>0.003034</td>
</tr>
<tr>
<td>T-test Statistic ((t_0))</td>
<td>-0.096263</td>
</tr>
<tr>
<td>Critical Value ((t_{0.025,31}))</td>
<td>2.04227</td>
</tr>
<tr>
<td>T-test result</td>
<td>Fail to reject (H_0), i.e., (E[\bar{W}_i] = 0).</td>
</tr>
<tr>
<td>Small number ((\varepsilon))</td>
<td>0.0021</td>
</tr>
<tr>
<td>Critical (\chi^2) value ((\chi^2_{0.05,31}))</td>
<td>45.0</td>
</tr>
<tr>
<td>95% confidence interval for percentage modeling error</td>
<td>((-0.09, 0.09) \times 100%)</td>
</tr>
</tbody>
</table>
Figure 1 Taxonomy of continuum flow models
Figure 2 Generic building block of freeway networks
Figure 3 Sketch of the study site (GA 400 Southbound)
Figure 4 Comparison of simulated and observed congested regions
Figure 5 Comparison of observed and simulated flow-density relationships (5-minute bins)
Figure 6 Frequency of simulation errors