

Piecewise-Linearized Cell Transmission Model and Parameter Calibration Methodology

Laura Muñoz, Xiaotian Sun, Roberto Horowitz, and Luis Alvarez

This paper studies the development of macroscopic freeway traffic models and parameter calibration methodologies that are computationally efficient and suitable for use in real-time traffic monitoring and control applications. Toward the fulfillment of these objectives, a macroscopic traffic model, the switching-mode model (SMM), is presented; it is a piecewise-linearized version of Daganzo's cell transmission model (CTM). The observability and controllability properties of the SMM modes are reviewed because these properties are of fundamental importance in the design of traffic estimators and on-ramp metering controllers. A semiautomated method has been developed for calibrating the CTM and SMM parameters. In this method, a least-squares data-fitting approach is applied to loop detector data to determine free-flow speeds, congestion-wave speeds, and jam densities for specified subsections of a freeway. Bottleneck capacities are estimated from measured mainline and on-ramp flows. The calibration method was tested with loop detector data from Interstate 210 westbound (I-210W) in Southern California. The main traffic data source was the performance measurement system. Parameters were calibrated for a 2-mi (3-km) subsection of I-210W and were tested on both the SMM and CTM, which were shown to perform similarly and to reproduce the general behavior of traffic congestion.

Accurate freeway traffic models are valuable tools for the design and evaluation of traffic management and monitoring strategies. For example, macroscopic and microscopic traffic models can be used to predict the effects of implementing different on-ramp metering algorithms on a freeway (1–3), and a dynamic model can be used as a basis for a freeway traffic state estimator (4, 5). To evaluate a model for possible use in a traffic study, it is worth investigating properties that will affect the usefulness and accuracy of the model, such as the number and types of parameters, possible methods for parameter calibration, computational efficiency, and structural properties such as observability and controllability. To produce an accurate model, it is usually necessary to calibrate the parameters, but this process can be laborious and time-consuming, especially in the case of microscopic models. As discussed by Zhang et al. (6), the topics of traffic model calibration and validation remain an active area of

research, and there is a need for standardized methods of parameter calibration.

To help fulfill the goal of providing an accurate, computationally efficient, and easy-to-calibrate model for the development of freeway traffic control and estimation algorithms, a piecewise-linearized version of Daganzo's macroscopic cell transmission model (CTM) (7, 8), called the switching-mode model (SMM), has been developed (9) and is discussed in this paper. Its linear structure lends the advantage of simplifying control analysis, design, and data-estimation methods. Both the CTM and SMM perform well in describing traffic behavior when tested with data from a 2-mi (3-km) portion of Interstate 210 westbound (I-210W) in Southern California (9, 10). Furthermore, the observability and controllability properties of the SMM, which are of fundamental importance in the design of traffic data estimators and freeway on-ramp control systems, can be derived by standard linear systems techniques, as discussed later in this paper. To facilitate calibration of the CTM and SMM, a semiautomated calibration methodology for estimating their common parameters has been developed (10, 11) and applied to a segment of I-210W that typically endures heavy congestion during the weekday morning commute period.

CELL TRANSMISSION MODEL

The CTM was selected as a starting point for this research because of its analytical simplicity, ease of calibration, and ability to reproduce important traffic behavioral phenomena, such as the backward propagation of congestion waves. The CTM was previously validated for a single freeway link, with no on-ramps or off-ramps, with data from I-880 in California (12). A more complete description of the model can be found elsewhere (7, 8, 13), but the main equations are reviewed in this section.

In the CTM, a freeway is partitioned into a series of cells, as in Figure 1. Here, $\rho_i(k)$ is the density, in vehicles per unit length of freeway, in cell i at time kT_s , where k is the time index and T_s is the discrete time interval. The density evolves according to conservation of vehicles:

$$\rho_i(k+1) = \rho_i(k) + \frac{T_s}{l_i} [q_{i,in}(k) - q_{i,out}(k)] \quad (1)$$

where $q_{i,in}(k)$ and $q_{i,out}(k)$ are, respectively, the total flows, in vehicles per unit time, entering and leaving cell i during the k th time

L. Muñoz, X. Sun, and R. Horowitz, Department of Mechanical Engineering, University of California, Berkeley, CA 94720. L. Alvarez, Institute of Engineering, National Autonomous University of Mexico, México D.F. 04510.

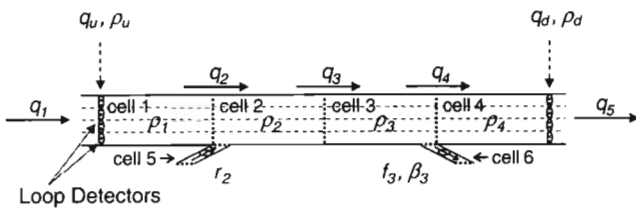


FIGURE 1 Freeway segment: mainline divided into four cells.

interval, $T_s[k, k + 1)$, including flows along the mainline and the on- and off-ramps, and l_i is the length of cell i .

The model parameters include v, w, Q_M , and ρ_j , which are depicted in the trapezoidal fundamental diagram of Figure 2. The parameters can be uniform over all cells, or they can be allowed to vary from cell to cell.

The parameters are defined as follows:

- v = free-flow speed (mph or km/h),
- w = backward congestion wave speed (mph or kph),
- Q_M = maximum allowable flow [vehicles/h (vph)],
- ρ_j = jam density [vehicles/mi (vpm) or vehicles/km (vpk), or vpm or vpk per lane (vpmpl, vpkpl)], and
- ρ_c = critical density (vpm, vpk, vpmpl, or vpkpl).

A trapezoidal fundamental diagram was chosen for the CTM, because this type of diagram is a standard formulation for the model (8) and lends itself easily to the piecewise linearization described in this paper. However, a trapezoidal diagram is not mandatory; it is possible to implement the CTM with other diagram shapes, as discussed by Daganzo (13).

Three different types of intercellular connection are allowed: simple connection, merge, and diverge. In a simple connection, two cells are connected to one another without any intervening on-ramps or off-ramps (e.g., Cells 2 and 3 in Figure 1). Let $i - 1$ be the upstream cell and let i be the downstream cell in the pair. As described by Daganzo (8), $q_i(k)$, the flow entering cell i from the mainline, is determined by taking the minimum of two quantities:

$$q_i(k) = \min[S_{i-1}(k), R_i(k)] \tag{2}$$

$$S_{i-1}(k) = \min(v_{i-1}\rho_{i-1}(k), Q_{M,j-1}) \tag{3}$$

$$R_i(k) = \min\{Q_{M,j}, w_i[\rho_{j,j} - \rho_i(k)]\} \tag{4}$$

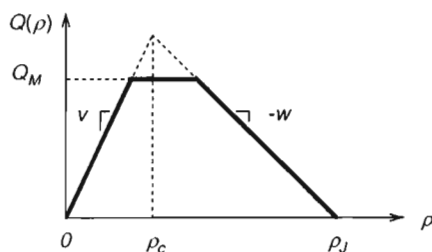


FIGURE 2 Trapezoidal fundamental diagram.

where $S_{i-1}(k)$ is the maximum flow that can be supplied by cell $i - 1$ under free-flow conditions, over the k th time interval, and $R_i(k)$ is the maximum flow that can be received by cell i under congested conditions, over the same time interval.

In the work presented here, the CTM merge and diverge laws of Daganzo (8) were used. A merge and diverge are shown in the freeway segment of Figure 1, where q_2 and r_2 are the flows merging into Cell 2, and q_4 and f_3 are the flows diverging from Cell 3. The diverging flows are defined as $q_4(k) = [1 - \beta_3(k)]q_{3,out}(k)$, and $f_3(k) = \beta_3(k)q_{3,out}(k)$, where $\beta_3(k)$ is the split ratio for the diverge junction (i.e., the fraction of vehicles leaving Cell 3 that exits through the off-ramp during the k th time interval). It is assumed here that the split ratios can be determined externally to the model as functions of time. The version of the CTM used here consists of flow conservation (Equation 1) for each cell, along with the flow specification (e.g., Equation 2) for each intercellular connection. The state variable is $\rho = [\rho_1 \dots \rho_N]^T$ for a freeway partitioned into N cells, and the model inputs include the measured flows entering the freeway at the upstream mainline boundary and at each on-ramp merge location. A complete statement of the boundary conditions can be found elsewhere (10).

For reference, congestion refers to the traffic condition that occurs when vehicle density is high, and drivers proceed at reduced speeds to avoid collisions. Free flow is the opposite condition, where vehicle density is low, and drivers travel at high speeds. The left part of the fundamental diagram of Figure 2, where $Q(\rho) = v\rho$, is an approximation of the typical behavior of free-flow traffic, whereas the right side $[Q(\rho) = w(\rho_j - \rho)]$ is associated with congested traffic. For the case in which the fundamental diagram is triangular instead of trapezoidal, the congestion status of cell i is determined by the cell density being compared with the critical density: if $\rho_i < \rho_{c,i}$, the cell has free-flow status, otherwise $\rho_i \geq \rho_{c,i}$ and the cell has congested status.

SWITCHING-MODE MODEL

To gain additional insight into freeway traffic behavior, and to simplify the control analysis, control design, and data-estimation design methods, a piecewise-linearized version of the CTM, the SMM, has been designed (9, 10). Because the SMM is composed of several linear models, straightforward linear techniques for model analysis and control design can be applied to the individual linear subsystems.

As a first step in specifying the SMM, it is assumed, for simplicity, that a freeway section can be in only one of five modes: (a) free flow–free flow (FF), in which all cells in the section have free-flow status; (b) congestion–congestion (CC), in which all cells in the section have congested status; (c) congestion–free flow (CF), in which the upstream part of the section is congested and the downstream part has free-flow status; (d) free flow–congestion 1 (FC1), in which the upstream part of the section has free-flow status, the downstream part has congested status, and the boundary (i.e., wavefront) separating the two regions is moving downstream; and (e) free flow–congestion 2 (FC2), in which the upstream part of the section has free-flow status, the downstream part has congested status, and the wavefront separating the two regions is moving upstream. This set of modes covers the possible congestion patterns within a section that has, at most, one status transition, referred to here

as a wavefront. The single-wavefront assumption is an approximation that is expected to be acceptable for short freeway sections. To deal more accurately with longer sections, the SMM can be modified to allow multiple wavefronts within a section.

Next, for each mode, the intercellular flows of the CTM in Equation 1 are written as explicit functions of cell densities. For simplicity, a triangular fundamental diagram (which is a special case of the trapezoidal shape) is assumed here, but this assumption is not necessary for development of the SMM. Then, in a segment without merges or diverges, in FF mode, each $q_i(k)$ would be replaced with $v\rho_{i-1}(k)$. If this procedure is carried out for each mode, the CTM conservation equations become linear because of the piecewise linearity of the fundamental diagram. To complete the SMM, a method must be specified to enable mode transitions. This is handled by a set of rules, described later, that determine the current mode, based in part on the most recent measurements at the segment boundaries.

The SMM does not fully replicate the CTM merge and diverge laws described by Daganzo (8). Although the on-ramp entering $[r(k)]$ and off-ramp exiting [e.g., $\beta(k) v_i \rho_i$] flows are represented in the SMM, the ramps are not modeled by cells; hence, traffic densities on the ramps are not included in the model. These simplifications are considered reasonable when (a) the SMM is being used in a calibration or validation study—then, $r(k)$ is set equal to the measured flow merging onto the freeway from the associated on-ramp (this quantity is typically available from field measurements)—and (b) the selected freeway segment contains no off-ramp bottlenecks, or, more generally, there are no situations in which congestion originates on an off-ramp because of insufficient capacity on the off-ramp. A review of the traffic data from the I-210 segment, discussed in part by Muñoz (10) and by Gomes et al. (14), indicates that the segment does not appear to exhibit any off-ramp congestion; hence, the SMM off-ramp flow representation is reasonable. If the SMM is used as a basis for designing ramp-metering controllers, as described by Sun and Horowitz (1, 2), it is advisable also to consider an on-ramp queuing model. See Sun and Horowitz (1, 2) for more detail.

Consider the freeway segment in Figure 1. Measured aggregate flows and densities at the upstream and downstream mainline detectors are denoted by q_u, ρ_u and q_d, ρ_d . All five modes of the SMM can be summarized as follows:

$$\rho(k+1) = A_s \rho(k) + B_s u(k) + B_{J,s} \rho_J + B_{Q,s} q_M \tag{5}$$

where

$s = 1, 2, 3, 4, 5$ indicates the mode (1, FF; 2, CC; 3, CF; 4, FC1; 5, FC2);

$\rho = [\rho_1 \dots \rho_4]^T$ = the state;

$u = [q_u \ r_2 \ \rho_u]^T$ = flow and density inputs, and r_2 = measured on-ramp flow entering the section, subscripted according to its cell of entry;

$\rho_J = [\rho_{J,1} \ \rho_{J,2} \ \rho_{J,3} \ \rho_{J,4} \ \rho_{J,5}]^T$ = vector of jam densities, and

$q_M = [Q_{M,1} \ Q_{M,2} \ Q_{M,3} \ Q_{M,4}]^T$ = vector of maximum flow rates.

In free-flow mode, the flow across each cell boundary is dictated by upstream conditions. Each cell releases traffic at the free-flow rate (i.e., the total flow exiting cell i is given by $v_i \rho_i$). The state equation is as follows:

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} (k+1) = \begin{bmatrix} 1 - \frac{v_1 T_s}{l_1} & 0 & 0 & 0 \\ \frac{v_1 T_s}{l_2} & 1 - \frac{v_2 T_s}{l_2} & 0 & 0 \\ 0 & \frac{v_2 T_s}{l_3} & 1 - \frac{v_3 T_s}{l_3} & 0 \\ 0 & 0 & [1 - \beta_3(k)] \frac{v_3 T_s}{l_4} & 1 - \frac{v_4 T_s}{l_4} \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} (k) + \begin{bmatrix} \frac{T_s}{l_1} & 0 & 0 \\ 0 & \frac{T_s}{l_2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} q_u \\ r_2 \\ \rho_d \end{bmatrix} (k) = A_1(k) \rho(k) + B_1 u(k) \tag{6}$$

In CC mode, the flow across each cell boundary is dictated by downstream conditions—that is, the total flow entering cell i is given by $w_i(\rho_{J,i} - \rho_i)$. The state equation is as follows:

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} (k+1) = \begin{bmatrix} 1 - \frac{w_1 T_s}{l_1} & \frac{w_2 T_s}{l_1} & 0 & 0 \\ 0 & 1 - \frac{w_2 T_s}{l_2} & \frac{w_3 T_s}{l_2} & 0 \\ 0 & 0 & 1 - \frac{w_3 T_s}{l_3} & \frac{1}{1 - \beta_3(k)} \frac{w_4 T_s}{l_3} \\ 0 & 0 & 0 & 1 - \frac{w_4 T_s}{l_4} \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} (k) + \begin{bmatrix} 0 & \frac{T_s}{l_1} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{w_5 T_s}{l_4} \end{bmatrix} \begin{bmatrix} q_u \\ r_2 \\ \rho_d \end{bmatrix} (k) + \begin{bmatrix} \frac{w_1 T_s}{l_1} - \frac{w_2 T_s}{l_1} & 0 & 0 & 0 \\ 0 & \frac{w_2 T_s}{l_2} - \frac{w_3 T_s}{l_2} & 0 & 0 \\ 0 & 0 & \frac{w_3 T_s}{l_3} - \frac{1}{1 - \beta_3(k)} \frac{w_4 T_s}{l_3} & 0 \\ 0 & 0 & 0 & \frac{w_4 T_s}{l_4} - \frac{w_5 T_s}{l_4} \end{bmatrix} \begin{bmatrix} \rho_{J,1} \\ \rho_{J,2} \\ \rho_{J,3} \\ \rho_{J,4} \\ \rho_{J,5} \end{bmatrix} = A_2(k) \rho(k) + B_2 u(k) + B_{J,2}(k) \rho_J \tag{7}$$

In CF mode, there exists one congestion-to-free-flow transition inside the section. One assumption of the SMM is that the wavefront will always lie on a cell boundary. Cells upstream of the wavefront

accept vehicles at the congested flow rate, and cells downstream of the wavefront release vehicles at the free-flow rate. The wavefront acts as a bottleneck, expelling vehicles at maximum allowed rate Q_M and decoupling the region upstream of the wavefront from the downstream region. When the wavefront is located between Cells 2 and 3, the state equation is as follows:

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} (k+1) = \begin{bmatrix} 1 - \frac{w_1 T_s}{l_1} & \frac{w_2 T_s}{l_1} & 0 & 0 \\ 0 & 1 - \frac{w_2 T_s}{l_2} & 0 & 0 \\ 0 & 0 & 1 - \frac{v_3 T_s}{l_3} & 0 \\ 0 & 0 & (1 - \beta_3(k)) \frac{v_3 T_s}{l_4} & 1 - \frac{v_4 T_s}{l_4} \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} (k) + \begin{bmatrix} 0 & \frac{T_s}{l_1} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} q_u \\ r_2 \\ \rho_d \end{bmatrix} (k) + \begin{bmatrix} \frac{w_1 T_s}{l_1} & -\frac{w_2 T_s}{l_1} & 0 & 0 \\ 0 & \frac{w_2 T_s}{l_2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \rho_{J,1} \\ \rho_{J,2} \\ \rho_{J,3} \\ \rho_{J,4} \\ \rho_{J,5} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{T_s}{l_2} & 0 \\ 0 & 0 & \frac{T_s}{l_3} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} Q_{M,1} \\ Q_{M,2} \\ Q_{M,3} \\ Q_{M,4} \end{bmatrix} = A_3(k)\rho(k) + B_3 u(k) + B_{J,3}(k)\rho_j + B_{Q,3} q_M \quad (8)$$

The wavefront could have occurred between Cells 1 and 2, or between Cells 3 and 4, instead of between Cells 2 and 3; however, the example serves to illustrate the equation structure for this mode.

In both FC modes, one free-flow-to-congestion transition exists inside the section. Unlike the previous mode, the state matrices change depending on the direction of motion of the wavefront. In FC1, the wavefront moves downstream. Assuming, for example, that the wavefront is between Cells 2 and 3, the state equation for this mode is as follows:

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} (k+1) = \begin{bmatrix} 1 - \frac{v_1 T_s}{l_1} & 0 & 0 & 0 \\ \frac{v_1 T_s}{l_2} & 1 - \frac{v_2 T_s}{l_2} & 0 & 0 \\ 0 & \frac{v_2 T_s}{l_3} & 1 - \frac{1}{1 - \beta_3(k)} \frac{w_4 T_s}{l_3} & 0 \\ 0 & 0 & 0 & 1 - \frac{w_4 T_s}{l_4} \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} (k)$$

$$\begin{bmatrix} \frac{T_s}{l_1} & 0 & 0 \\ 0 & \frac{T_s}{l_2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{w_5 T_s}{l_4} \end{bmatrix} \begin{bmatrix} q_u \\ r_2 \\ \rho_d \end{bmatrix} (k) + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{1 - \beta_3(k)} \frac{w_4 T_s}{l_3} \\ 0 & 0 & 0 & \frac{w_4 T_s}{l_4} \end{bmatrix} \begin{bmatrix} \rho_{J,1} \\ \rho_{J,2} \\ \rho_{J,3} \\ \rho_{J,4} \\ \rho_{J,5} \end{bmatrix} = A_4(k)\rho(k) + B_4 u(k) + B_{J,4}(k)\rho_j \quad (9)$$

For FC2, the wave moves upstream. Again assuming that the wavefront is between Cells 2 and 3, this mode differs from the previous case in that, because of the dominance of the congested flow rate at the wavefront boundary, the tridiagonal row is now the second instead of the third row, and more terms appear in $B_{J,5}$:

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} (k+1) = \begin{bmatrix} 1 - \frac{v_1 T_s}{l_1} & 0 & 0 & 0 \\ \frac{v_1 T_s}{l_2} & 1 & \frac{w_3 T_s}{l_2} & 0 \\ 0 & 0 & 1 - \frac{w_3 T_s}{l_3} & \frac{1}{1 - \beta_3(k)} \frac{w_4 T_s}{l_3} \\ 0 & 0 & 0 & 1 - \frac{w_4 T_s}{l_4} \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} (k) + \begin{bmatrix} \frac{T_s}{l_1} & 0 & 0 \\ 0 & \frac{T_s}{l_2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{w_5 T_s}{l_4} \end{bmatrix} \begin{bmatrix} q_u \\ r_2 \\ \rho_d \end{bmatrix} (k) + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{w_3 T_s}{l_2} & 0 \\ 0 & 0 & \frac{w_3 T_s}{l_3} & -\frac{1}{1 - \beta_3(k)} \frac{w_4 T_s}{l_3} \\ 0 & 0 & 0 & \frac{w_4 T_s}{l_4} \end{bmatrix} \begin{bmatrix} \rho_{J,1} \\ \rho_{J,2} \\ \rho_{J,3} \\ \rho_{J,4} \\ \rho_{J,5} \end{bmatrix} = A_5(k)\rho(k) + B_5 u(k) + B_{J,5}(k)\rho_j \quad (10)$$

At each time step, the SMM determines its mode on the basis of the measured mainline boundary data and the congestion status of

the cells in the section. If both ρ_u and ρ_d have free-flow status (i.e., both densities are below ρ_c), the FF mode is selected; if both of these densities are congested (i.e., both densities are at or above ρ_c), the CC mode is selected. If ρ_u and ρ_d are of opposite status, then the SMM performs a search over the ρ_i to determine whether there is a status transition inside the section. This wavefront search consists of searching through the cells, in order, looking for the first status transition between adjacent cells (10).

The quality of density estimates can be improved with output feedback; for example, Sun et al. (4, 5) applied a mixture Kalman filter (MKF) based estimator to two of the SMM modes (FF and CC). At each time step, the estimator determines which mode is stochastically most likely (i.e., whether the FF or CC model is a better fit for each freeway section). Through its estimation of the congestion mode, the MKF-based estimator alleviates the need to specify a set of logical mode-selection rules, such as those described here and by Muñoz et al. (9, 10).

OBSERVABILITY AND CONTROLLABILITY

The observability and controllability properties of the SMM modes are presented here, because those properties are of fundamental importance for the design of data estimators and ramp-metering control systems. Loosely, controllability answers the question of whether a control input (e.g., a regulated on-ramp flow) can affect the model state (in this case, traffic density). Observability answers the question of whether the measurements recorded at a particular detector station (along with knowledge of the system inputs) can be used to reconstruct the initial system state. Brogan (15) provides a more thorough discussion of these properties.

Table 1 summarizes the observability and controllability for each SMM mode. These results were derived by standard linear systems techniques [see Muñoz (10) for more detail]. In the first and second columns, “upstream cells” and “downstream cells” give the status of cells both upstream and downstream of the congestion wavefront. If there is no such wavefront, both sets of cells have the same status. The third column indicates which of the two mainline boundary measurements, if either, can be used to make each SMM mode observable. The fourth column states whether each SMM mode is controllable from an on-ramp at the upstream end of the section or from an on-ramp at the downstream end of the section.

If all cells have free-flow status, densities are observable with a downstream measurement, whereas, in congested mode, they are observable with an upstream measurement. If no downstream measurement is available when cells are in free-flow mode, or if there is no upstream measurement when cells are congested, the system is

unobservable. This is related to the wave (information) propagation directions on a freeway in different congestion modes. When a freeway section is in free-flow mode, information propagates downstream at speed v , which is the vehicle traveling speed. Therefore, to be able to estimate the cell densities, the downstream density measurement is needed. When the freeway is in congestion, information propagates upstream at speed w , and an upstream measurement is needed to estimate densities.

The controllability results can be derived similarly to the observability results. Generally, a section in free-flow mode is controllable from an on-ramp at its upstream end, whereas a congested section can be controlled from an on-ramp at its downstream end. These controllability results have motivated the design of a traffic-responsive, linear-quadratic-integral-based local ramp-metering regulator that switches between using upstream and downstream densities for feedback, depending on whether the freeway section surrounding the regulated on-ramp has congested or free-flow status. This regulator has been observed to reduce total travel times significantly and to perform favorably in simulation (1–3) with regulators such as ALINEA (16) and SWARM (17).

CTM AND SMM CALIBRATION METHOD

In this section, a methodology is described for tuning the CTM and SMM parameters to reproduce observed freeway traffic behavior. The calibration method was previously tested (on the CTM only) with data from a 14-mi (23-km) stretch of I-210W (11).

Freeway Representation and Traffic Data

Figure 3 is a schematic of the freeway section used to test the CTM and SMM. It is a subsection of I-210W, approximately 2 mi (3 km) long, with four mainline lanes, two on-ramps, two off-ramps, three mainline loop-detector stations labeled Myrtle (ML 34.05), Huntington (ML 33.05), Santa Anita (ML 32.20), and additional detector stations on each ramp. ML stands for mainline, and the numbers (e.g., 34.05) are the absolute postmile indices of the detector stations (postmiles are a measurement of distance, in miles, along the highway). The mainline segment is partitioned into eight cells, with lengths (0.088 0.375 0.375 0.192 0.088 0.276 0.276 0.246) mi. The full I-210 testbed extends beyond this segment; it is 14 mi (23 km) long, from ML 39.16 to ML 25.40. A 41-cell partition has been developed for the full testbed (10, 11).

The performance measurement system (PeMS) (18), developed by researchers at the University of California, Berkeley, was the main source of traffic data used in this study. Each freeway loop

TABLE 1 Observability and Controllability for Different SMM Modes

Upstream Cell	Downstream Cell	Observable with	Controllable from
Free-flow	Free-flow	Downstream measurement	Upstream on-ramp
Congested	Congested	Upstream measurement	Downstream on-ramp
Congested	Free-flow	Upstream and downstream measurement	Not controllable
Free-flow	Congested 1	Unobservable	Upstream and downstream on-ramps
Free-flow	Congested 2	Unobservable	Upstream and downstream on-ramps

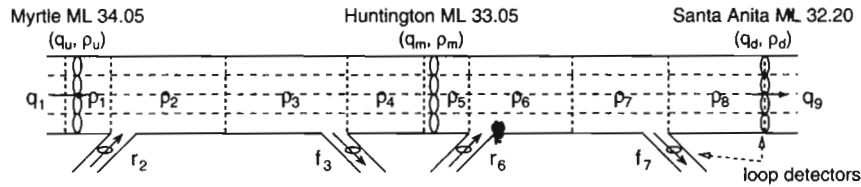


FIGURE 3 Segment of I-210W divided into cells.

detector provides measurements of volume (vehicles per time step) and percent occupancy every 30 s. PeMS aggregates these data, and makes them available online, for freeways throughout California. For mainline detectors, densities (vpm) were computed for each lane using density = occupancy/g factor, where the g factor is the effective vehicle length, in miles, for that detector. For single loop-detector freeways such as I-210, PeMS provides g factors calculated according to the PeMS algorithm (19).

Calibration Methodology

The main steps of the calibration procedure are as follows:

1. Free-flow parameter calibration. The free-flow traffic velocities, v_i , are determined by performing a least-squares fit on flow versus density data over the period 5:00 to 6:00 a.m. For the I-210 section, traffic typically flows freely during this period. For the j th detector station (numbered in the direction of traffic flow), v_j is the least-squares solution to the equation $\Phi_j v_j = Y_j$, where Φ_j and Y_j are column vectors that, respectively, contain densities and flows measured over the specified time interval. The free-flow speed v_j is assigned to the cell containing detector j , and free-flow speeds are computed for cells that do not contain detectors by linear interpolation.

2. Bottleneck identification. Bottleneck locations are identified by examining space-time contour plots of measured traffic densities or speeds and determining the locations of fixed spatial boundaries that divide the freeway into an upstream congested region and a downstream free-flow region.

3. Non-bottleneck capacity selection. A set of nominal $Q_{M,i}$ are assigned to cells that are not located at bottlenecks. It is not advisable to set $Q_{M,i}$ equal to the maximum observed flow at each detector-equipped cell, because this will most likely result in underestimating the true capacity of the freeway. Typically, nominal $Q_{M,i}$ must be chosen to be larger than the maximum observed flows [usually $\geq 2,000$ vph per lane (vphpl)] in each region of the freeway.

4. Bottleneck capacity determination. Consider a freeway portion divided into two consecutive cells, 1 and 2, where an active bottleneck exists between the two cells; hence, the upstream cell (Cell 1) is congested, whereas the downstream cell (Cell 2) remains in free-flow status. Further assume that an on-ramp (with merging flow r_2 entering Cell 2) enters between the two cells. It can be shown (10) that the bottleneck capacity in this situation is represented by $Q_{M,2} = q_2 + r_2$, where q_2 is the modeled flow entering Cell 2 from the mainline (10, 11). Assuming both q_2 and r_2 are measurable, these quantities are used to estimate the bottleneck flow rate, with the default method (assuming no faulty or missing data) being

$$\hat{Q}_{M,2} = \text{mean}_{k \in K_M} [q_2(k) + r_2(k)]$$

K_M corresponds to the half-hour time interval ending at $\arg \max [q_2(k) + r_2(k)]$. Determination of the causes and evolution of bottlenecks and estimation of the ensuing capacities are active areas of research. For a recent example, see Bertini and Cassidy (20).

5. Congestion parameter calibration. w_i and $\rho_{j,i}$ are estimated by performing a constrained least-squares fit on flow versus density measurements. First, the critical density is estimated for each detector: $\hat{\rho}_{c,j} = \max_k [q_{d_j}(k)]/v_j$, where q_{d_j} is the flow measured at detector j . The $[\rho_{d_j}(k), q_{d_j}(k)]$ data are sorted so that only congested pairs are used in the estimation. Let $\kappa = \{k_1 \dots k_{N_c}\}$ denote the set of all k for which $\rho_{d_j}(k) > \hat{\rho}_{c,j}$. For a cell in a congested section, it can be shown that the CTM conservation equation (or alternatively the CC mode of the SMM) can be rewritten in the form $\Phi_j [w_j w_j \rho_{j,i}]^T = Y_j$, where

$$\Phi_j^T = \begin{bmatrix} -\rho_{d_j}(k_1) & \dots & -\rho_{d_j}(k_{N_c}) \\ 1 & \dots & 1 \end{bmatrix}, Y_j = \begin{bmatrix} q_{d_j}(k_1) + \frac{l_j}{T_s} \Delta \rho_{d_j}(k_1) \\ \vdots \\ q_{d_j}(k_{N_c}) + \frac{l_j}{T_s} \Delta \rho_{d_j}(k_{N_c}) \end{bmatrix}$$

and $q_{d_j}(k)$ is used in place of the flow exiting the cell containing detector j , and ρ_{d_j} values are used in place of the difference terms in the conservation equation—that is, $\Delta \rho_{d_j}(k) = \rho_{d_j}(k+1) - \rho_{d_j}(k)$ (10, 11). Then, $[w_j w_j \rho_{j,i}]^T$ is the least-squares solution to $\Phi_j [w_j w_j \rho_{j,i}]^T = Y_j$, subject to the constraint

$$Q_{M,j} \leq \frac{v_j w_j \rho_{j,i}}{v_j + w_j}$$

which can be written as linear in $[w_j w_j \rho_{j,i}]^T$. The constraint is included to prevent the solution $[w_j w_j \rho_{j,i}]^T$ from limiting the maximum possible flow in cell j below the $Q_{M,j}$ determined in previous steps. Currently, only values of w_j that fall within a physically reasonable range, $10 \leq w_j \leq 20$ mph (16 and 32 km/h), are retained. If the constrained least-squares solution does not yield w , the acceptable range for a particular detector cell n , this cell is assigned the w_j of the nearest downstream neighbor with a w inside the range. The corresponding $\rho_{j,n}$ is found by solving the equality case of the constraint; w_i and $\rho_{j,i}$ are determined for nondetector cells through linear interpolation.

6. Time-varying parameter adjustments. If necessary, temporary parameter changes (e.g., reduction of $Q_{M,i}$ in a region) can be applied to reproduce the effect of an incident. Additionally, with w_i in the midmorning time range being reduced, when traffic is still congested but beginning its recovery back to free-flow mode, the effect of

flow-density hysteresis can be approximated. Additional detail on the calibration method can be found elsewhere (10, 11).

RESULTS AND DISCUSSION

For comparison with results presented later in this section, the validation work covered by Muñoz et al. (9, 10) is reviewed briefly. Muñoz et al. (9, 10) validated the CTM and SMM on the same 2-mi (3-km) segment of I-210W indicated in Figure 3. A single set of hand-tuned parameters [$v=63$ mph (101 km/h), $Q_M=2,000$ vphpl, $\rho_f=172$ vpmpl (107 vpkpl)] was selected, and the models were simulated with boundary data from five different weekdays over the period 5 a.m. to 12 noon. Approximate flow hysteresis was induced in the models by w being reduced from 14.26 mph (22.96 km/h) to 12.5 mph (20.13 km/h) at 9 a.m. The CTM and SMM were shown to perform similarly to one another and to reproduce density and flow measurements with mean absolute percentage errors of approximately 13% and 4%, respectively (9, 10). However, because hand-tuning the parameters is time-consuming, a standardized calibration procedure was desired; this led to development of the calibration methodology described earlier.

To test this calibration methodology, the procedure was applied to I-210 testbed data collected on April 25, 2001, and yielded parameters for the segment of Figure 3. The resulting calibrated parameters showed little spatial variation: 60 to 62 mph (97 to 100 km/h) for v , 2,000 vphpl for Q_M , 15 mph (24 km/h) for w , and 168 to 169 vpmpl (104 to 105 vpkpl) for ρ_f . These parameters are similar to the hand-tuned values used by Muñoz et al. (9, 10).

The desired next step was to simulate the CTM and SMM with the newly calibrated parameters and boundary data from April 25, 2001. Additional data processing was first conducted to produce suitable boundary data for the simulations. To compensate for two nonfunctioning loops (one in Lane 2 of Myrtle station and another in one of the off-ramp lanes at Santa Anita), scaling-corrected

aggregates were prepared with the mean of the densities (or flows) over the functioning loops at each station being taken and then multiplied by the total number of loops. These reconstructed data were not used in the calibration procedure; specifically, faulty mainline detector stations were left out of the estimation of v_i , w_i , and $\rho_{j,i}$, which were then interpolated, with the methods described previously, from the nearest available good values. Furthermore, the cell lengths chosen for the 2-mi segment prohibited a simulation time step as large as 30 s; thus, a zeroth-order interpolation was applied to the PeMS data to yield data with $T_s=5$ s, which was chosen as the time step for both models. To counteract noise in the PeMS 30-s data, a first-order Butterworth lowpass filter with cutoff frequency $0.01T_s^{-1}$ Hz was applied to the data using a zero-phase forward-and-reverse filtering technique. The reader is referred to Muñoz et al. (9, 10) for additional details on data processing and split ratio estimation.

The calibrated parameters were implemented in Matlab simulations of both the CTM and SMM, with the same hysteresis conditions as Muñoz et al. (9, 10). The simulated densities and flows in the middle of the section (near postmile 33.05) were compared with the measured values at that location (i.e., ρ_m and q_m), as indicated in Figure 4. The densities and flows predicted by the models are similar to one another and agree reasonably well with the measured values, although the models tended to overestimate density during the peak congestion hours. The mean percentage error is defined as

$$E_{\text{MPE}} = \frac{1}{M} \sum_{k=1}^M \left| \frac{x(k) - \hat{x}(k)}{x(k)} \right|$$

where $x(k)$ represents the measured value (flow or density) at the middle detector station, and $\hat{x}(k)$ is the model-predicted value. The CTM and SMM density errors were 13% and 14%, respectively, and the flow errors were 4% and 5% over the simulated time period. These results closely resemble those of the validation tests of Muñoz et al.

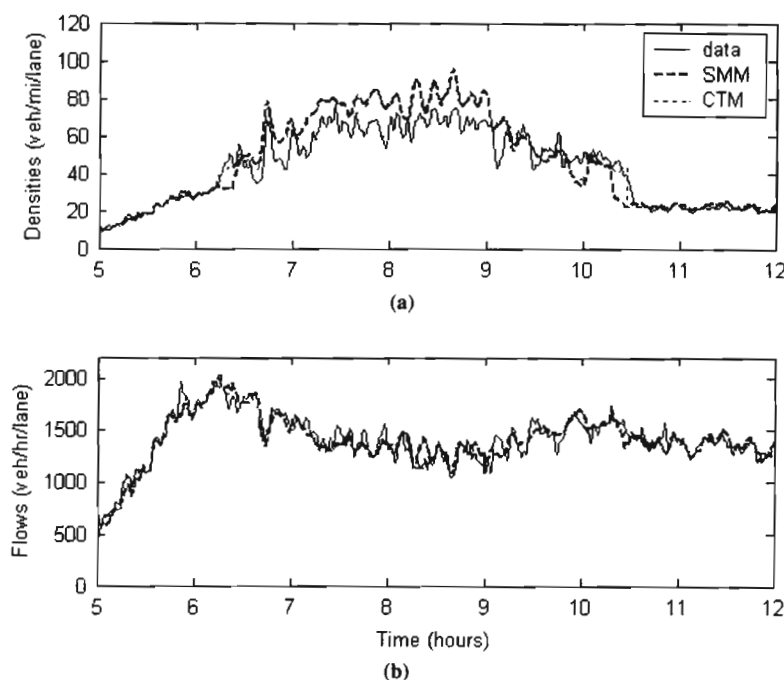


FIGURE 4 Measured and simulated mainline (a) densities and (b) flows for detector station at postmile 33.05, on April 25, 2001.

(9, 10); this is unsurprising, because the calibration method produced parameters similar to the hand-tuned ones.

The automated steps of the calibration method have been determined to have low computational cost. The most time-consuming automated step of the calibration process is the constrained linear least-squares estimation of congestion parameters for each cell, which can be carried out for the 14-mi (23-km) portion of I-210W, partitioned into 41 cells, in less than 1 min for 7 h of 5-min average measurement data on a 1.0-GHz, 256-megabyte Pentium III computer.

The single-wavefront assumption of the SMM means that the SMM is expected to be suitable for short freeway sections; the model requires modification to handle longer sections, which are more likely to contain multiple wavefronts simultaneously. However, this issue can be avoided in the design of a closed-loop estimator. In this case, a long segment can be broken into a series of shorter segments, and a series of SMMs can be put together to produce state estimates at unmonitored locations, using feedback from any available nonfaulty mainline station (4, 5).

Muñoz et al. (10) made a preliminary comparison between the CTM and SMM estimation errors and results reported in other work such as that of Chen et al. (21). This analysis indicates that the CTM and SMM may yield density and flow estimates that are similar to, or better than, those derived with certain other approaches. A complete comparison with other methods, under normalized test conditions, is the subject of ongoing work.

CONCLUSIONS

In this paper, a macroscopic freeway traffic model, the SMM, was presented. This model is a piecewise-linearized version of Daganzo's CTM (7, 8), which was found to be computationally efficient and well suited for implementation in real-time control, estimation, and traffic monitoring applications (5). The observability and controllability properties of the individual modes of the SMM were stated, because they are of fundamental importance in the design of data estimators and ramp-metering control systems. It was explained that the free-flow traffic mode is observable from a downstream measurement and controllable from an upstream on-ramp and that the congested mode is observable from an upstream measurement and controllable from a downstream on-ramp.

The following points summarize the main disadvantages and advantages of the CTM relative to the SMM, and vice versa.

Cell Transmission Model

- Disadvantages. It is a nonlinear model; hence, the model-analysis methods and estimator and controller designs will be more complex, and potentially slower performing in real-time applications, than those for a linear model.
- Advantages. The congestion status of a cell (free-flow or congested) can be determined automatically by comparison of $\rho_i(k)$ with $\rho_{c,i}$. In addition, unlike the SMM, the model lacks an assumption of a single wavefront per section.

Switching-Mode Model

- Disadvantages. Congestion-mode selection rules, such as those described previously, must be specified for the model. In addition,

there is the assumption of single wavefront per section, also discussed previously.

- Advantages. It yields similar performance (in terms of accuracy and computation time) as the CTM but with the advantage of possessing a linear-mode structure. Linear analysis methods and controller and estimator design techniques can be applied to the individual modes.

Both of the aforementioned disadvantages of the SMM can be mitigated with the closed-loop estimation approach described by Sun et al. (4, 5). The SMM has been found to be well suited for use in a closed-loop estimator, with feedback from available measurements to estimate densities at unmonitored locations.

A procedure for calibrating the CTM and SMM parameters was summarized. Parameters were calibrated for a short [2-mi (3-km)] subsection of I-210W and tested on both the SMM and CTM, which were shown to perform similarly and to have reasonable agreement with observed values. A main benefit of the overall calibration method is that it provides a well-defined, partially automated procedure for using loop-detector data to estimate free-flow speeds, congestion parameters, and bottleneck capacities for these models.

ACKNOWLEDGMENT

This research was supported by Partners for Advanced Transit and Highways (PATH).

REFERENCES

1. Sun, X., and R. Horowitz. Set of New Traffic-Responsive Ramp-Metering Algorithms and Microscopic Simulation Results. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1959, Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 9–18.
2. Sun, X., and R. Horowitz. A Localized Switching Ramp-Metering Controller with a Queue Length Regulator for Congested Freeways. *Proceedings of the 2005 American Control Conference*, Portland, Ore., June 8–10, 2005, pp. 2141–2146.
3. Sun, X., and R. Horowitz. Localized Switching Ramp-Metering Control with Queue Length Estimation and Regulation and Microscopic Simulation Results. *Proc., 16th IFAC World Congress*, Prague, Czech Republic, 2005.
4. Sun, X., L. Muñoz, and R. Horowitz. Highway Traffic State Estimation Using Improved Mixture Kalman Filters for Effective Ramp Metering Control. *Proc., 42nd IEEE Conference on Decision and Control*, Maui, Hawaii, Dec. 9–12, 2003, pp. 6333–6338.
5. Sun, X., L. Muñoz, and R. Horowitz. Mixture Kalman Filter Based Highway Congestion Mode and Vehicle Density Estimator and its Application. *Proceedings of the 2004 American Control Conference*, Boston, Mass., June 30 to July 2, 2004, pp. 2098–2103.
6. Zhang, L., J. Colyar, J. Halkias, and Y. Zhang. Validation and Calibration of CORSIM Traffic Simulation Model at Signalized Intersections. Presented at 83rd Annual Meeting of the Transportation Research Board, Washington, D.C., 2004.
7. Daganzo, C. F. The Cell Transmission Model: A Dynamic Representation of Highway Traffic Consistent with the Hydrodynamic Theory. *Transportation Research B*, Vol. 28, No. 4, 1994, pp. 269–287.
8. Daganzo, C. F. The Cell Transmission Model, Part II: Network Traffic. *Transportation Research B*, Vol. 29, No. 2, 1995, pp. 79–93.
9. Muñoz, L., X. Sun, R. Horowitz, and L. Alvarez. Traffic Density Estimation with the Cell Transmission Model. *Proceedings of the 2003 American Control Conference*, Denver, Colo., June 2003, pp. 3750–3755.

10. Muñoz, L. *Macroscopic Modeling and Identification of Freeway Traffic Flow*. PhD dissertation, University of California, Berkeley, 2004.
11. Muñoz, L., X. Sun, D. Sun, G. Gomes, and R. Horowitz. Methodological Calibration of the Cell Transmission Model. *Proceedings of the 2004 American Control Conference*, Boston, Mass., June 30 to July 2, 2004, pp. 798–803.
12. Lin, W.-H., and D. Ahanotu. *Validating the Basic Cell Transmission Model on a Single Freeway Link*. PATH Technical Note 95-3. Institute of Transportation Studies, University of California, Berkeley, 1994.
13. Daganzo, C. F. A Finite Difference Approximation of the Kinematic Wave Model of Traffic Flow. *Transportation Research B*, Vol. 29, No. 4, 1995, pp. 261–276.
14. Gomes, G., A. May, and R. Horowitz. Congested Freeway Microsimulation Model Using VISSIM. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1876, Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 71–81.
15. Brogan, W. L. *Modern Control Theory*, 3rd ed. Prentice Hall, Englewood Cliffs, N.J. 1991.
16. Papageorgiou, M., H. Hadj-Salem, and J.-M. Blosseville. ALINEA: A Local Feedback Control Law for On-Ramp Metering. In *Transportation Research Record 1320*, TRB, National Research Council, Washington, D.C., 1991, pp. 58–64.
17. *System Wide Adaptive Ramp Metering High Level Design*. Technical Report. National Engineering Technology Corp., La Mirada, Calif., June 1996.
18. *Freeway Performance Measurement System (PeMS)*. The PeMS Group, Berkeley, Calif. www.pems.eecs.berkeley.edu. Accessed Oct. 2002.
19. Jia, Z., C. Chen, B. Coifman, and P. Varaiya. The PeMS Algorithms for Accurate, Real-Time Estimates of g-Factors and Speeds from Single-Loop Detectors. In *2001 IEEE Intelligent Transportation Systems Conference Proceedings*, Oakland, Calif., Aug. 2001, pp. 536–541.
20. Bertini, R. L., and M. J. Cassidy. Some Observed Queue Discharge Features at a Freeway Bottleneck Downstream of a Merge. *Transportation Research A*, Vol. 36, 2002, pp. 683–697.
21. Chen, C., J. Kwon, J. Rice, A. Skabardonis, and P. Varaiya. Detecting Errors and Imputing Missing Data for Single-Loop Surveillance Systems. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1855, Transportation Research Board of the National Academies, Washington, D.C., 2003, pp. 160–167.

The Traffic Flow Theory and Characteristics Committee sponsored publication of this paper.