

Set of New Traffic-Responsive Ramp-Metering Algorithms and Microscopic Simulation Results

Xiaotian Sun and Roberto Horowitz

A novel switching traffic-responsive ramp-metering controller adapts to different traffic dynamics under different congestion conditions: free-flow or congested. The approach of multirate linear quadratic control with integral action is used to compensate for disturbances and to accommodate the difference between the model sampling time and the metering-rate update interval. In addition, a queue length regulator is designed to prevent the queue from exceeding the ramp storage capacity and yield improved performance over the current ad hoc queue override scheme. Subsequently, a queue length estimator is designed to provide feedback to the queue length regulator with the queue-detector speed data that are available in the field. A local ramp-metering control strategy is proposed to achieve the control goal of reducing the spatial and temporal span of the congestion, while satisfying the on-ramp storage capacity constraints, by using locally available information. Test results on a calibrated microscopic traffic simulator demonstrate the performance and effectiveness of the switching ramp-metering controller, the queue length estimator and regulator, and the overall control strategy. The total vehicle and passenger congestion delays are both reduced by 16%, and the total travel time is improved by 5.6%. As a comparison, simulation results of ALINEA are also presented.

Freeway traffic congestion is a major problem in metropolitan areas. It occurs regularly during commute hours. In addition, non-recurrent congestion often takes place as a result of incidents, road-work, or public events. Congestion causes inefficient operation of freeways, wasting of resources, increased pollution, and intensified driver fatigue.

The 2004 *Urban Mobility Report* finds that "congestion has grown everywhere in areas of all sizes. Congestion occurs during longer portions of the day and delays more travelers and goods than ever before" (1). In that report, the authors calculated that in 2002, congestion cost Americans 3.5 billion h delay and 5.7 billion gal wasted fuel, with an equivalent monetary cost of \$63.2 billion.

On-ramp metering has been widely used as an effective strategy to increase freeway operation efficiency. It has been recommended to the FHWA as the number-one tool with which to address the congestion problem, other than adding more capacity to transportation infrastructures (2). It has been reported that ramp metering was able

to reduce delay by 101 million person hours in 2002, approximately 5% of the congestion delay on freeways where ramp metering was in effect (1).

This paper presents a novel switching ramp-metering controller that uses a different feedback structure, depending on whether the freeway is in a free-flow or a congested mode. It is known that traffic dynamics behave differently under free-flow and congested conditions. This leads to different controllability and observability properties for free-flow versus congested traffic (3, 4). It is thus natural to design different ramp-metering controllers for these different congestion modes.

In addition to the switching mainline traffic-responsive ramp-metering controller, a proportional-integral (PI) regulator used to keep the on-ramp queue below the storage capacity limit and prevent excessively long queues from interfering with surface street traffic is also presented. This PI regulator will have better performance than the queue-override scheme currently used on California freeways. A queue length estimator that uses speed data measured by the queue detector is also designed and implemented. This estimator provides the feedback that is needed by the queue length regulator.

Also proposed is a localized control strategy for the switching metering controller and the queue length regulator to achieve the goal of reducing the spatial and temporal span of the congestion.

Test results of this set of ramp-metering algorithms, as well as those of ALINEA (5), on a calibrated microscopic traffic simulator are presented.

PERFORMANCE MEASURES

In this section, some performance measures are defined for quantitative evaluation of a given freeway segment. All quantities are defined for the period T and the freeway segment L :

TTD_v = total travel distance, which is defined as the sum of the distances traveled by all vehicles in L within T .

TTT_v = total travel time, which is the sum of the time that is spent by all vehicles in L within T . It includes the time spent by vehicles waiting in the on-ramp queues.

TCD_v = total congestion delay, which is the difference between the total travel time and the time that would be spent by all vehicles if there were no congestion.

$$TCD_v = TTT_v - (TTD_v / v_0)$$

where v_0 is the nominal free-flow speed.

Another set of passenger-weighted performance measures can be defined by first collecting the traffic quantities separately for the

Department of Mechanical Engineering, University of California, Berkeley, CA 94720-1742.

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.

low- and high-occupancy vehicle classes and then weighting these quantities by the average passenger number in each vehicle class when calculating the performance measures. This set of passenger-weighted performance measures is total passenger travel distance, TTD_p , total passenger travel time, TTT_p , and total passenger congestion delay, TCD_p .

SWITCHING-MODE MODEL AND CONTROLLABILITY

In previous work (3, 4), the cell transmission model (CTM) (6, 7) was piecewise linearized, and a switching-mode traffic model was derived. Both the switching-mode model and the CTM have been calibrated and tested by using traffic data collected from a segment of I-210 (3, 4, 8). The switching-mode model was able to reproduce traffic behavior as accurately as did the CTM.

Depending on traffic conditions in a freeway section, such as the one shown in Figure 1, the traffic is in a different mode—free-flow or congested. In each mode, the vehicle densities in the cells, denoted by 1 through 4 in the figure, evolve according to a different set of difference equations.

In free-flow mode,

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} (t+1) = \begin{bmatrix} 1 - \frac{v_{f1}T_s}{l_2} & 0 & 0 & 0 \\ \frac{v_{f1}T_s}{l_2} & 1 - \frac{v_{f2}T_s}{l_2} & 0 & 0 \\ 0 & \frac{v_{f2}T_s}{l_3} & 1 - \frac{v_{f3}T_s}{l_3} & 0 \\ 0 & 0 & (1-\beta)\frac{v_{f3}T_s}{l_4} & 1 - \frac{v_{f4}T_s}{l_4} \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} (t) + \begin{bmatrix} 0 \\ \frac{T_s}{l_2} \\ 0 \\ 0 \end{bmatrix} r(t) + \begin{bmatrix} \frac{T_s}{l_1} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} q_{m1} \\ q_{m2} \end{bmatrix} (t) \quad (1)$$

$$= A(1)\rho(t) + B_r(1)r(t) + B_m(1)q_m(t) \quad (2)$$

where

- ρ_i = vehicle density in cell i ;
- q_{m1}, q_{m2} = mainline entering and exiting flows, respectively;
- r = on-ramp flow;
- β = split ratio of the off-ramp flow;

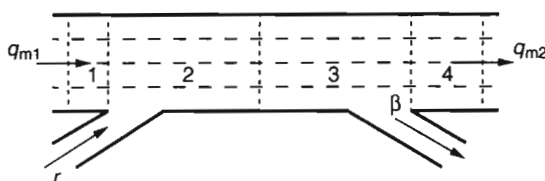


FIGURE 1 Schematic plot of four-cell freeway section with one on ramp and one off ramp.

- l_i = length of cell i ; and
- T_s = sampling time.

In congested mode,

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} (t+1) = \begin{bmatrix} 1 - \frac{w_{c1}T_s}{l_1} & \frac{w_{c2}T_s}{l_1} & 0 & 0 \\ 0 & 1 - \frac{w_{c2}T_s}{l_2} & \frac{w_{c3}T_s}{l_2} & 0 \\ 0 & 0 & 1 - \frac{w_{c3}T_s}{l_3} & \frac{1}{1-\beta} \frac{w_{c4}T_s}{l_3} \\ 0 & 0 & 0 & 1 - \frac{w_{c4}T_s}{l_4} \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} (t) + \begin{bmatrix} \frac{T_s}{l_1} \\ 0 \\ 0 \\ 0 \end{bmatrix} r(t) + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & -\frac{T_s}{l_4} \end{bmatrix} \begin{bmatrix} q_{m1} \\ q_{m2} \end{bmatrix} (t) + \begin{bmatrix} \frac{w_{c1}T_s}{l_1} - \frac{w_{c2}T_s}{l_1} & 0 & 0 \\ 0 & \frac{w_{c2}T_s}{l_2} - \frac{w_{c3}T_s}{l_2} & 0 \\ 0 & 0 & \frac{w_{c3}T_s}{l_3} - \frac{1}{1-\beta} \frac{w_{c4}T_s}{l_3} \\ 0 & 0 & 0 & \frac{w_{c4}T_s}{l_4} \end{bmatrix} \begin{bmatrix} \rho_{j1} \\ \rho_{j2} \\ \rho_{j3} \\ \rho_{j4} \end{bmatrix} \quad (3)$$

$$= A(2)\rho(t) + B_r(2)r(t) + B_m(2)q_m(t) + B_j(2)\rho_j \quad (4)$$

where ρ_{ji} is the jam density (maximum allowable density) in cell i .

The traffic dynamics are quite different in different modes, which is evidenced by the structures of the A matrices in Equations 1 and 3. In free-flow mode, the A matrix is lower bidiagonal, and thus the vehicle densities in the downstream cells are affected by those in the upstream cells, that is, the information travels from upstream to downstream. In congested mode, the A matrix is upper bidiagonal, and the vehicle densities in the upstream cells are affected by those in the downstream cells, that is, the information travels from downstream to upstream.

This observation is critical to the design of an on-ramp metering algorithm because it determines a dynamical system's fundamental properties of controllability and observability. A simple calculation reveals that when a freeway section is in free-flow mode, the on ramp can control the vehicle densities downstream, whereas in congested mode, the on ramp can control the vehicle densities upstream. Therefore, in a different mode, a different feedback structure must be used for the metering controller, as shown in Figure 2.

Although the congestion modes or the cell vehicle densities are not observed or measured directly, an estimator based on a mixture Kalman filter (MKF) can accurately estimate these quantities in real time (9, 10). The mainline vehicle flows and densities, as well as the on-ramp flows, measured at the locations where loop detectors are installed, are used as input and feedback to the Kalman filters. The esti-

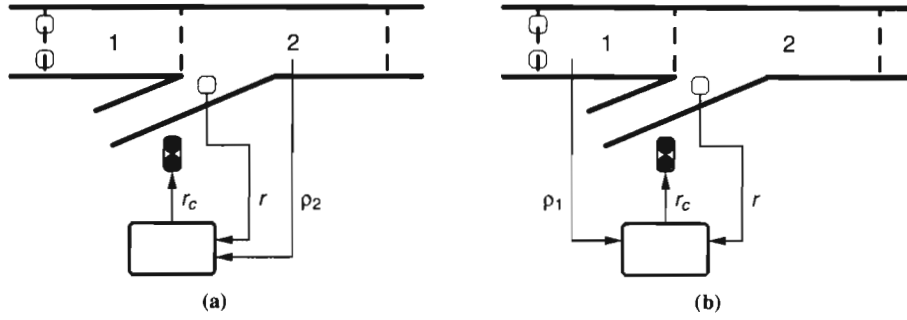


FIGURE 2 Control structures for (a) free-flow mode and (b) congested mode.

mated congestion mode is used to determine the appropriate control structure, and the estimated vehicle densities are used as feedback.

DESIGN OF MULTIRATE LINEAR QUADRATIC CONTROL WITH INTEGRAL ACTION

To compensate for disturbances and to accommodate the difference between the model sampling time and the metering-rate update interval, a multirate linear quadratic control with integral action (multirate LQI) approach (11) was used to synthesize the ramp-metering controller for both congestion modes.

In either congestion mode, the difference equations governing the evolution of the cell vehicle densities can be written as

$$\rho(t+1) = A\rho(t) + B_r r(t) + B_m q_m(t) + B_j \rho_j \quad (5)$$

where

$$\begin{aligned} r(t) &= \text{on-ramp flow,} \\ q_m(t) &= \text{mainline flows, and} \\ \rho_j &= \text{jam densities.} \end{aligned}$$

The desired densities $\bar{\rho}$ have to satisfy the steady-state condition

$$\bar{\rho} = A\bar{\rho} + B_r \bar{r} + B_m \bar{q}_m + B_j \rho_j \quad (6)$$

where \bar{q}_m is the nominal mainline flow. The error dynamics is thus given by

$$\tilde{\rho}(t+1) = A\tilde{\rho}(t) + B_r r(t) - B_r \bar{r} + B_m \bar{q}_m(t) \quad (7)$$

By defining

$$\eta(t+1) = \tilde{\rho}(t+1) - \tilde{\rho}(t) \quad (8)$$

and

$$u(t+1) = r(t+1) - r(t) \quad (9)$$

and considering that in the steady state, $q_m(t)$ varies slowly, one obtains

$$\begin{bmatrix} \tilde{\rho} \\ \eta \end{bmatrix}(t+1) = \begin{bmatrix} I & A \\ 0 & A \end{bmatrix} \begin{bmatrix} \tilde{\rho} \\ \eta \end{bmatrix}(t) + \begin{bmatrix} B_r \\ B_r \end{bmatrix} u(t) \quad (10)$$

which can be written in a more compact form:

$$\underline{\eta}(t+1) = \underline{A}\underline{\eta}(t) + \underline{B}u(t) \quad (11)$$

One would like to synthesize a control law

$$u(t) = -K(t) \begin{bmatrix} \tilde{\rho} \\ \eta \end{bmatrix}(t) \quad (12)$$

that minimizes the cost function

$$J = \frac{1}{2} \sum_{t=0}^{\infty} \begin{bmatrix} \tilde{\rho} \\ \eta \end{bmatrix}^T(t) \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{\rho} \\ \eta \end{bmatrix}(t) + u^T(t) R u(t) \quad (13)$$

$$= \frac{1}{2} \sum_{t=0}^{\infty} \underline{\eta}^T(t) \underline{Q} \underline{\eta}(t) + u^T(t) R u(t) \quad (14)$$

This problem can be solved easily by using the algebraic Riccati equation. This approach is often referred to as the LQI method. It was used by Papageorgiou et al. to design a coordinated ramp-metering control algorithm for a freeway segment with five on ramps (12).

The real control input (the metering rate) is given by

$$r(t) = r(t-1) - K(t) \begin{bmatrix} \tilde{\rho} \\ \eta \end{bmatrix}(t) \quad (15)$$

Because of geometric constraints on the cell lengths, the sampling time of the model is 2 s. However, the control variable $r(t)$ can be updated only every 30 s for the I-210 westbound (I-210W) test site. Thus a multirate approach to the LQI method is necessary.

It is assumed that the actual control input $r(t)$ is updated every p shortest time periods (STPs). STP = 2 s and $p = 15$ in the present problem. The virtual control $u(t)$ as defined in Equation 9 can be injected in the expanded system of Equation 11 only every p STPs, that is, $u(t)$ is nonzero only once every p STPs. Equivalently, instead of letting $u(t)$ be zero, $B(t)$ can be zero and have a periodic dynamic system, Equation 11, in which

$$\underline{B}(t) = \begin{cases} \underline{B} & \text{when } t = np \text{ for some } n \in \mathbb{Z}, \\ 0 & \text{when } t \neq np \text{ for any } n \in \mathbb{Z}. \end{cases} \quad (16)$$

To synthesize the optimal controller, Equation 12, for this periodic system, one must solve the periodic algebraic Riccati equation

$$\underline{P}(t) = \underline{A}^T(t)\underline{P}(t+1)\underline{A}(t) - \underline{A}^T(t)\underline{P}(t+1)\underline{B}(t)[R(t) + \underline{B}^T(t)\underline{P}(t+1)\underline{B}(t)]^{-1}\underline{B}^T(t)\underline{P}(t+1)\underline{A}(t) + \underline{Q}(t) \quad (17)$$

for a periodic solution

$$\underline{P}(t+p) = \underline{P}(t) \quad (18)$$

The optimal control gain will be given by

$$K(t) = [R(t) + \underline{B}^T(t)\underline{P}(t+1)\underline{B}(t)]^{-1}\underline{B}^T(t)\underline{P}(t+1)\underline{A}(t) \quad (19)$$

Note that $K(t)$ is periodic and is nonzero only at $t = np$. It can be computed offline.

Another issue to be addressed is saturation. The districts of the California State Department of Transportation (Caltrans) have established acceptable maximum and minimum metering rates. The "one vehicle per green per lane" policy enforced on the I-210W test site and typical driver and vehicle response times determine the green phase length of the metering signal to be 2 s. Similarly, the minimum length of the red phase is also 2 s. In addition, to avoid the increased probability of signal violation when drivers wait too long, the red phase length is usually limited to 18 or 13 s. Therefore, the maximum metering rate is usually 900 vehicles per hour per metered lane (vphpl), and the minimum metering rate is 180 or 240 vphpl.

In contrast, the realized on-ramp entering flow may be less than the set metering rate, when the arrival rate of the vehicles (the demand) is less than the metering rate and there is not a sufficient number of vehicles waiting in the queue. Therefore, the control input $r(t)$ has to be saturated to not only the established maximum and minimum values but also the number of available vehicles. To increase the response time of the integral control, some sort of antiwindup scheme must be implemented.

Similarly to ALINEA (5), the desired metering rate $r_c(t)$, which is set by the controller, is distinguished from the realized on-ramp flow $r(t)$, which is measured by the entrance loop detector, as illustrated in Figure 3, and modify the control algorithm, Equation 15, to be

$$r'_c(t) = r(t-1) - K(t) \left[\frac{\tilde{\rho}}{\eta} \right] (t) \quad (20)$$

and

$$r_c(t) = \min\{r_{\max}, \max[r_{\min}, r'_c(t)]\} \quad (21)$$

Use of the actual ramp flow $r(t-1)$ in the integral control law, Equation 20, may lead to a nonzero offset in case of biased realization of the ramp flow due to, for example, signal violation (5).

ON-RAMP QUEUE LENGTH REGULATION AND ESTIMATION

Queue Length Regulation

Whenever the on-ramp demand $d(t)$ exceeds the desired metering rate $r_c(t)$, a queue will form. However, the storage capacity of an on-ramp often is quite limited. Without proper control, the vehicle queue length will quickly exceed this capacity, causing the vehicles to spill over into the surface streets and interfere with street traffic.

The current practice to regulate queue length on California freeways is the so-called queue override. A typical loop detector and signal configuration of an on-ramp is shown in Figure 3. In the queue override scheme, the signal controller compares the occupancy measured by the queue detectors to a threshold and determines whether the queue has reached the queue detectors. If the queue detector occupancy is above the threshold, the metering rate is increased by a certain level, for example, 120 vphpl, every metering-rate update time interval (e.g., 30 s). When the queue detector occupancy falls below the threshold, the metering rate is reset to the value determined by the controller. This queue override scheme can be viewed as an integral control with a saturated integrating rate and resetting. It has been noted that this queue override scheme leads to oscillatory behavior and underutilization of on-ramp storage capacities (13-16). Gordon attempted to improve the queue override performance by filtering the occupancy signal and reducing the sampling time interval (13). Ozbay et al. showed through Paramics simulation that an integral ramp queue control sometimes can be effective to improve system performance (14, 15). Smaragdis and Papageorgiou proposed a proportional controller that relies on the on-ramp vehicle demands (16). Unfortunately, real-time on-ramp vehicle demand measurements are not available in the field, and historical demands may have to be used to realize the controller proposed by Smaragdis and Papageorgiou (16). The queue detector may provide an approximate demand measurement before the queue reaches it. After the queue extends beyond the queue detector, it will no longer be able to provide this measurement. The queue length regulator presented here does not require prior knowledge of the demands. It uses a queue length estimation algorithm to implement a PI controller.

The queue length dynamics is modeled as a simple integrator:

$$l(t+1) = l(t) + T_s [d(t) - r(t)] \quad \text{subject to } l \geq 0 \quad (22)$$

where

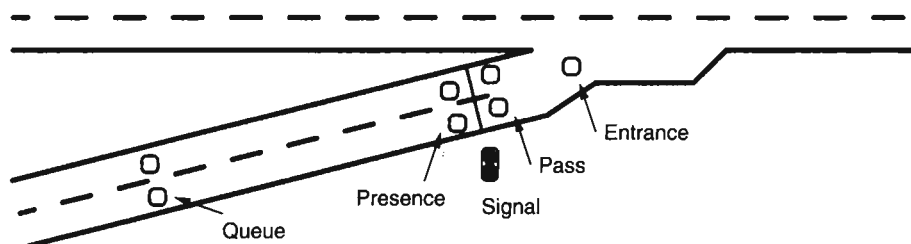


FIGURE 3 Typical on-ramp configuration of loop detectors and signals.

$l(t)$ = queue length (number of vehicles),
 $d(t)$ = demand (flow entering the queue), and
 $r(t)$ = on-ramp flow entering the mainline (flow leaving the queue).

As previously discussed, the currently used queue override scheme can be viewed as an integral control with a saturated integrating rate and resetting. A simple analysis shows that the integral control alone on an integrator is not asymptotically stable and a proportional action is needed, that is,

$$r_c(z) = \left(k_p + \frac{k_i}{z-1} \right) \tilde{l}(z) \quad (23)$$

which is written in transfer function form. In Equation 23, $\tilde{l}(z)$ is the queue length error. In the design of the PI queue length regulator, $d(t)$ is treated as an external disturbance and $r(t)$ as the control variable. The closed-loop sensitivity function from the disturbance to the error is

$$\frac{\tilde{l}(z)}{d(z)} = \frac{T_s(z-1)}{(z-1)^2 - k_p T_s z + (k_i - k_p) T_s} \quad (24)$$

Proper PI gains k_p and k_i can be selected with the root locus method. Furthermore, the antiwindup and saturating mechanisms in Equations 20 and 21 need to be implemented in the queue length regulator, too.

Queue Length Estimation

Although it has a more stable response than the queue override scheme, the PI regulator described earlier needs the current queue length as its feedback, which unfortunately is not available in the field. A suitable estimator must be designed by using available information, such as the vehicle speed measured by the queue detector.

Assume the following simplified driving behavior model for a vehicle approaching the end of the queue: the vehicle decelerates at a constant rate, $-a$, from its cruising speed to a target speed v_0 at the position where the distance from the end of the queue is s . Also assume a uniform effective vehicle length g . Let l_0 be the number of vehicle spaces from the stop line to the queue detector and $v(t)$ be the vehicle speed measured by the queue detector. See Figure 4.

A straightforward kinematic calculation yields

$$g[l_0 - l(t)] - s = \frac{v(t)^2 - v_0^2}{2a} \quad (25)$$

where $l(t)$ is the current queue length, in number of vehicles. From Equation 25, one obtains

$$gl(t) = gl_0 - s + \frac{v_0^2}{2a} - \frac{v(t)^2}{2a} = c_0 - c_2 v(t)^2 \quad (26)$$

To determine the coefficients c_0 and c_2 in Equation 26, curve fitting was performed on the $gl(t)$ and $v(t)$ data collected by using the VISSIM microscopic traffic simulator (17). Figure 5 shows a typical scatter plot of queue lengths versus speeds. A few points must be noted:

- When the queue is shorter than a certain length, the approaching vehicles pass the queue detector at the drivers' desired cruising speeds, which are independent of the queue length. This phenomenon corresponds to the data points at the lower right corner of the scatter plot.
- When the queue is longer than gl_0 , that is, the queue has extended beyond the queue detector, the measured speed is also a constant, which is related to the queue discharging rate and the vehicle lengths and also is independent of the queue length. This phenomenon corresponds to the data points at the upper left corner of the scatter plot.
- There are many outliers among the data points. Therefore, the usual least-squares curve fitting method, which is biased toward outliers, is not suitable.

For these reasons, the data points whose speeds are below v_{\min} or above v_{\max} and those whose queue lengths are below l_{\min} or above l_{\max} are neglected in the curve fitting. These values were determined by visual inspection of the scatter plots.

To increase robustness to outliers, the least-median-of-squares curve fitting method was used (18), instead of the usual least (sum of) squares. The fitted curve is also shown in Figure 5. After the l - v curve is fitted for each on ramp, the difference between the actual and desired queue length, which is used as the feedback to the regulator, Equation 23, is estimated as

$$\tilde{l}(t) = \begin{cases} [c_0 - gl_0 - c_2 v(t)^2] / g & \text{if } v(t) \geq v_{\min} \\ -kc_2 [v(t)^2 - v_{\min}^2] / g & \text{if } v(t) < v_{\min} \end{cases} \quad (27)$$

where k is a tuning parameter.

When $v(t) < v_{\min}$, the end of the queue is very close to or beyond the queue detector, and the speed $v(t)$ measured by the queue detector is a constant, which is roughly gr_c . Therefore, Equation 27 can be thought of as saturating \tilde{l} to $-kc_2[(gr_c)^2 - v_{\min}^2]/g$, which is larger when the metering rate r_c is lower. This has a desirable effect on the regulator: the metering rate r_c will be increased more aggressively when there is more room for this increase and more slowly when r_c is close to its maximum value. In addition, this saturation value can be further tuned by changing the value of k .

The coefficients c_0 and c_2 identified by the least-median-of-squares fitting are very close to the nominal values predicted by the actual distance between the stop line and the queue detector and a nominal vehicle deceleration of 2.5 m/s². Therefore, when the queue length measurements are unavailable through any means to perform a curve fitting, these nominal values can be used in the queue length estimation.

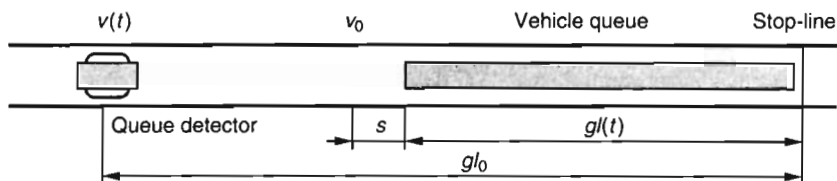


FIGURE 4 Schematics for on-ramp queue length estimation.

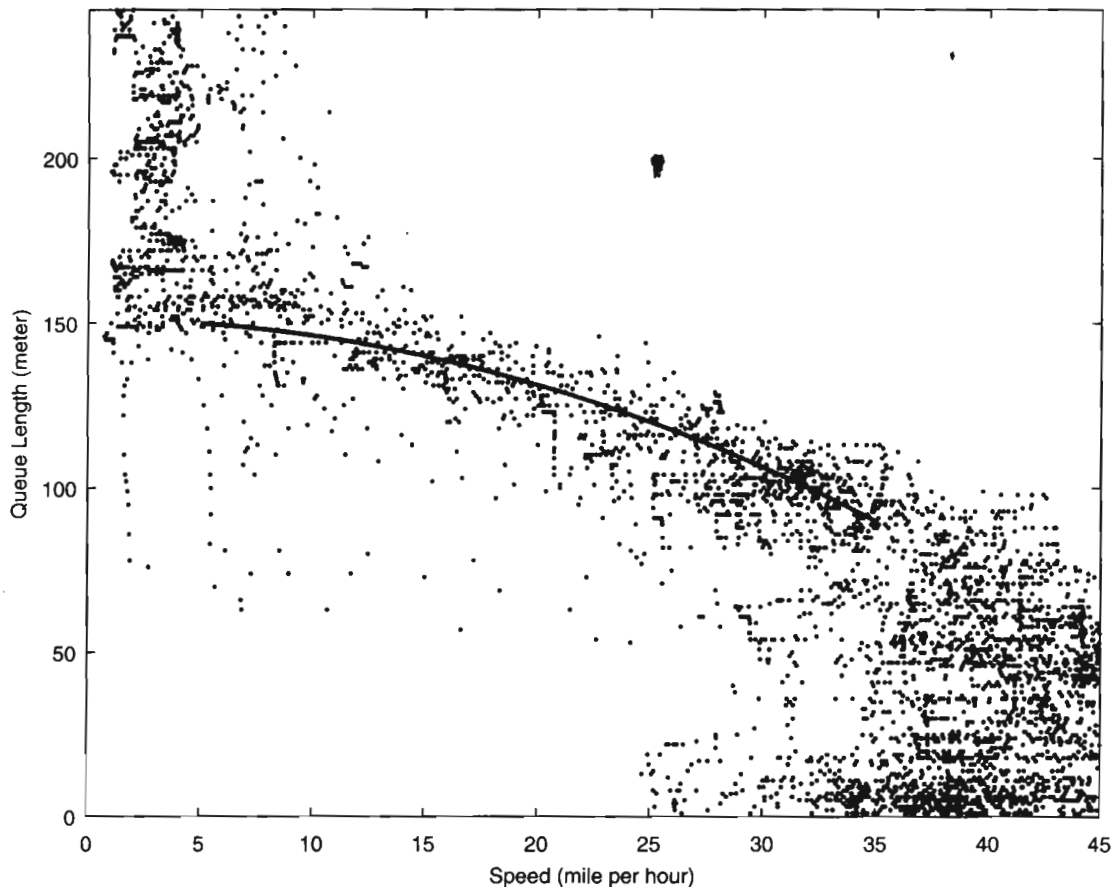


FIGURE 5 Scatter plot of queue lengths versus queue detector speeds and least-median-of-squares curve fit for on ramp.

LOCALIZED RAMP-METERING STRATEGY

In addition to the switching LQI mainline traffic-responsive metering controller and the queue regulator, a control strategy is still needed to achieve the best freeway performance. For example, the desired mainline density profiles and the desired queue lengths need to be defined. Gomes and Horowitz formulated a linear program that minimizes the so-called generalized total travel time with explicit queue length constraints (19, 20). This linear program is based on a modified version of the cell transmission model and produces a set of mainline densities, time-of-day ramp-metering rates, and queue lengths as the solution, by using the predicted traffic demands. These quantities can be used as the desired profiles in the switching controller and queue length regulator.

However, a localized ramp-metering strategy is also desired for other reasons, including reduced algorithmic complexity, lower computational requirements, and higher robustness to changing traffic conditions, such as unpredicted demands.

The main goal for freeway traffic control is to limit the spatial and temporal span of the congestion, that is, to slow congestion propagation in the upstream direction and to speed up any congestion wavefront that is moving downstream, in other words, to make the wavefront propagation speed v_w larger.

From the kinematic-wave traffic theory, for the congestion wavefront and freeway fundamental diagram (flow-density relationship) shown in Figures 6a and 6b, the propagation speed of the congestion wavefront is

$$v_w = -\frac{q_2 - q_1}{\rho_2 - \rho_1} \quad (28)$$

where q_1 and q_2 are the flows at Points 1 and 2 and ρ_1 and ρ_2 are the densities at Points 1 and 2.

It can be seen from the foregoing analysis that the local ramp-metering control should move the congested point (Point 2) to the left on the fundamental diagram, that is, decrease the vehicle density. But, if there is no local congestion, vehicles can be allowed to enter the mainline as fast as possible, so long as they do not induce congestion. Therefore, the set point of the feedback ramp-metering controller, in both the free-flow and the congested mode, is chosen to be the critical density ρ_c , that is, the density at which congestion is about to form.

To utilize all the available storage capacity on the ramp, it is natural to choose the set point for the queue length regulator to be the maximum allowed queue length. A long on-ramp queue is also desired for the purpose of deterring short-distance travelers from using the freeway, thus making freeway capacity available to long-distance travelers.

However, the objectives of the queue length regulator and the mainline traffic-responsive metering controller conflict with each other: when the mainline traffic is light, the ramp-metering controller allows vehicles to enter the mainline as fast as possible, with a short queue resulting; at the same time, the queue length regulator is trying to accumulate vehicles so as to form a long queue. When the mainline traffic is heavy, the ramp-metering controller allows

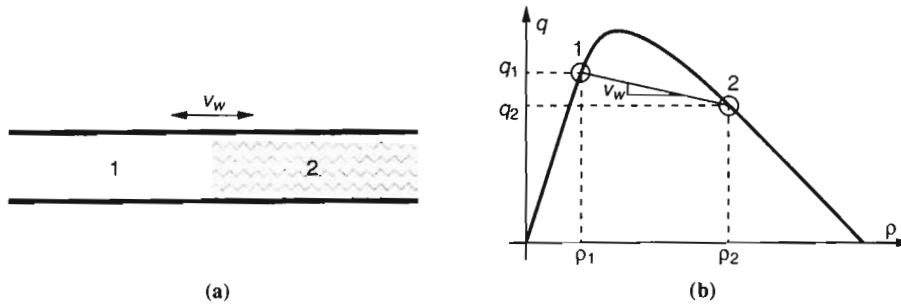


FIGURE 6 Kinematic-wave traffic theory: (a) congestion wavefront propagation and (b) fundamental diagram.

vehicles to enter only at a slow rate, and the queue extends quickly beyond the limit; the queue length regulator will have to increase the metering rate to dissipate the queue until it reaches an acceptable length.

After consideration of these factors, it was decided to set the metering rate at the larger of the two values determined by the mainline traffic-responsive metering controller and the queue length regulator. Smaragdakis and Papageorgiou proposed the same formula (16).

TEST SITE

A segment of I-210W in Pasadena, California, as shown in Figure 7, was selected as the test bed for new ramp-metering algorithms. It is approximately 14 mi long, from Vernon Avenue (Milepost 39.159) to Fair Oaks Avenue (Milepost 25.4), with 20 metered on-ramps, one uncontrolled freeway-to-freeway connector (I-605), and 18 off ramps. In previous efforts, microscopic (22) and macroscopic (4, 8) traffic simulation tools were calibrated to this test site.

RESULTS

The switching LQI mainline traffic-responsive metering controller and the queue length regulator were implemented and interfaced with the VISSIM microscopic traffic simulation model that has been calibrated to the I-210W test segment (22). The localized control strategy described earlier was used. Figure 8 shows the congestion patterns, as determined by the MKF traffic state estimator (9, 10), before and after ramp metering. In the plots, darker gray

indicates congested mode, and lighter gray indicates free-flow mode. The vertical axis is time, from 5:30 to 11:00 a.m. The horizontal axis is the milepost along the freeway, and the traffic travels from left to right. It can be seen that the localized ramp-metering strategy was able to reduce the congestion, in both the spatial span and the time duration.

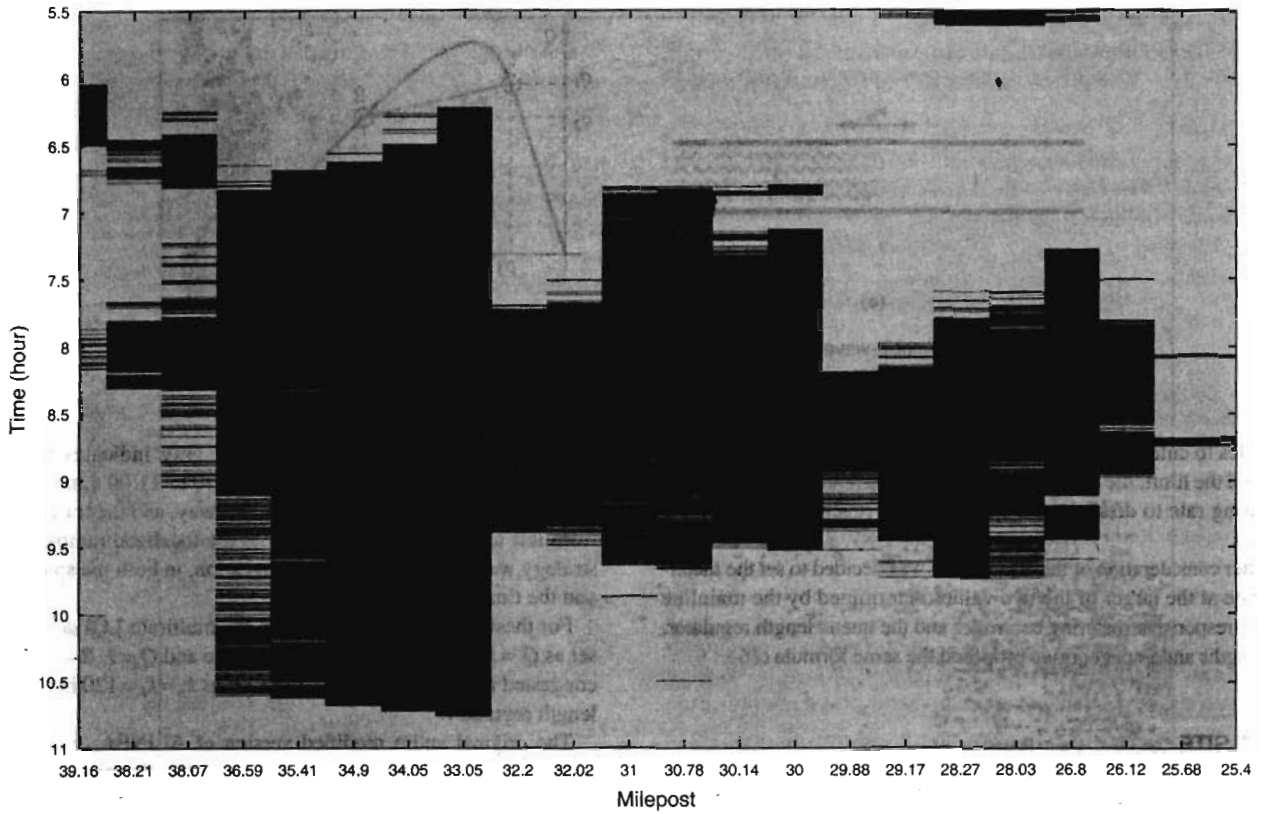
For these tests, the parameters in the multirate LQI design were set as $Q = 1, R = 5$ for the free-flow mode and $Q = 1, R = 20$ for the congested mode, and the gains were set as $k_p = k_r = 120$ in the queue length regulator.

The original and a modified version of ALINEA (5) were also implemented and combined with the queue length estimator and regulator. In this modified ALINEA, the occupancy data measured upstream to the on ramps was used instead of those measured downstream in the original ALINEA, because of the loop-detector configuration on California freeways. Gomes showed, by using the calibrated I-210W VISSIM model, that this modified ALINEA can achieve comparable, sometimes even better, performance, when compared to the original ALINEA (24). Optimal ALINEA gain (7,000) and set point (27% for upstream ALINEA and 21% for downstream ALINEA) found by Gomes were used in the simulations (23). The same gain and set point are used for all metered on ramps.

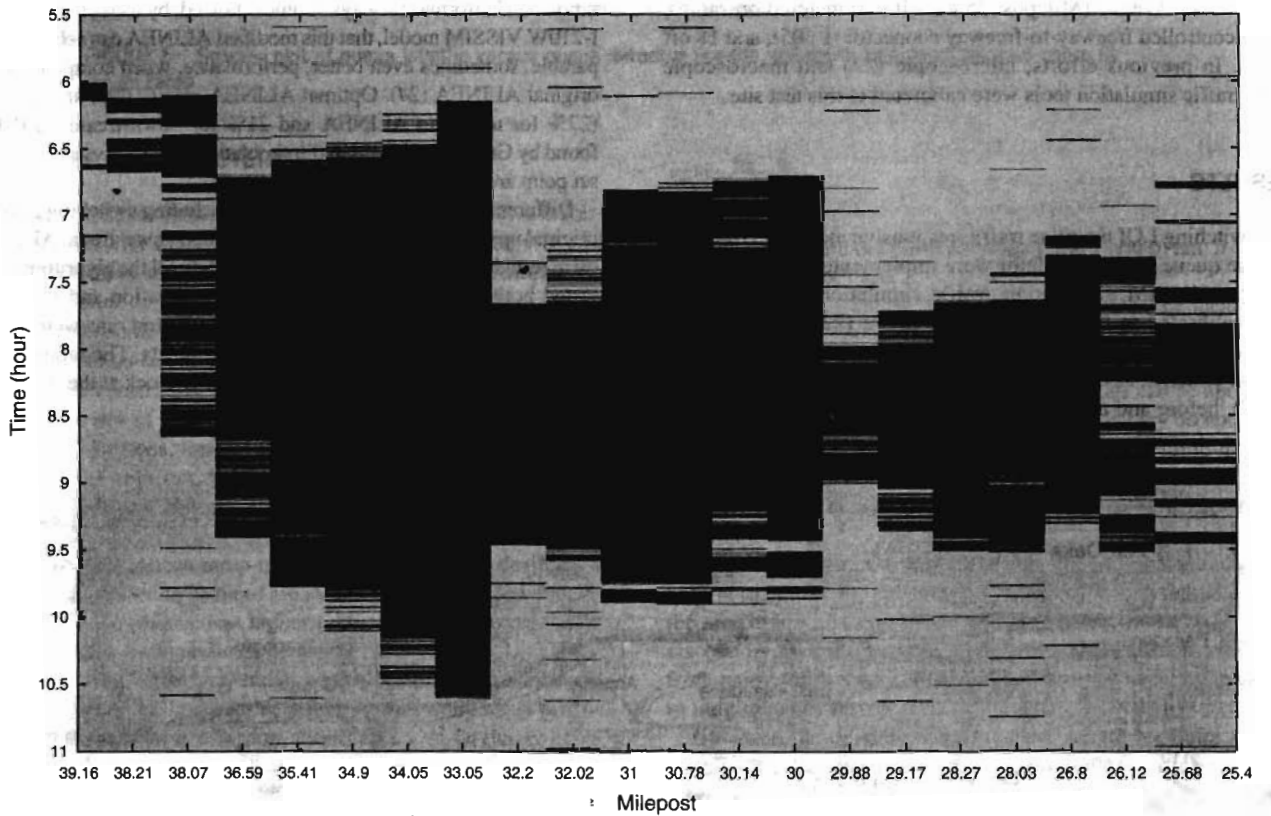
Different ramp-metering algorithms, including switching LQI, the original upstream ALINEA, and the modified downstream ALINEA, were tested with the I-210W VISSIM model. All the algorithms were tested both with and without the queue estimation and regulation algorithms. Under each scenario, eight simulation runs were carried out, with eight different VISSIM random seeds. The random seed was chosen to be the second of the computer clock at the time when it was changed, to ensure its randomness.



FIGURE 7 Map of test segment of I-210W, from Vernon Avenue to Fair Oaks Avenue, Pasadena, California. [Composed with U.S. Census Bureau TIGER/Line data (21)].



(a)



(b)

FIGURE 8 Congestion modes for I-210W test segment: (a) without ramp metering and (b) with proposed ramp-metering algorithms. Light gray indicates free-flow, dark gray indicates congested.

TABLE 1 Performance Measures for I-210 Test Segment

Mainline	Queue	TTD _v (10 ³ mi)	TTT _v (10 ³ h)	TCD _v (10 ³ h)	TTD _p (10 ⁶ m)	TTT _p (10 ³ h)	TCD _p (10 ³ h)
None	None	973	24.0	8.52	1.32	31.7	10.8
Switching LQI	Q/R	972	22.6	7.19	1.32	29.9	9.0
Improvement		—	5.6%	16%	—	5.5%	16%
U/S ALINEA	Q/R	974	23.5	8.01	1.32	31.0	10.1
Improvement		—	2.1%	5.9%	—	2.1%	6.3%
D/S ALINEA	Q/R	973	23.3	7.82	1.32	30.7	9.8
Improvement		—	2.9%	8.2%	—	2.9%	8.6%
Switching LQI	None	974	22.3	6.81	1.32	29.3	8.3
Improvement		—	7.1%	20%	—	7.7%	23%
U/S ALINEA	None	973	22.3	6.87	1.32	29.3	8.39
Improvement		—	6.9%	19%	—	7.5%	22%
D/S ALINEA	None	974	22.5	7.04	1.32	29.5	8.59
Improvement		—	6.2%	17%	—	6.9%	20%

Q/R: queue estimation and regulation; U/S: upstream; D/S: downstream

Some of the performance measures for this freeway segment are listed in Table 1. The listed numbers are the averages from the eight simulation runs for each scenario. In calculating these quantities, the average passenger numbers per one low- and one high-occupancy vehicle are assumed to be 1.2 and 2.5, respectively, and the nominal free-flow speed v_0 is 63 mph.

Under all the scenarios, the freeway segment served almost the same amount of demand, as measured by TTD, or TTD_p. Ramp metering was able to reduce the congestion under all the metered scenarios. For example, with the switching LQI mainline control and queue length regulation, the total vehicle delay (also known as congestion delay) TCD_v was reduced by 16%, whereas with the switching LQI mainline control only, TCD_v was reduced by 20%.

When only the switching mainline traffic-responsive metering was used, without activating the queue length regulator, on-ramp queues can be accumulated to arbitrary lengths, sometimes hundreds of vehicles. In this case, almost all the congestion on the mainline was eliminated. Another interesting phenomenon in this case is that the relative improvements in passenger-weighted performance measures were greater than those in vehicle performance measures. This is because many of the metered on-ramps on this freeway segment have designated lanes for HOVs to bypass the long queues.

It can also be seen from the numbers in Table 1 that the switching control algorithm outperforms both the original and the modified ALINEAs, whether the queue length estimator and regulator was active or not. The switching LQI controller distinguishes from the ALINEA algorithm in two aspects, which contribute to its better performance: the switching LQI controller uses different feedback structures that are suitable to the different traffic dynamics under the free-flow or congested conditions, and the proportional action in the LQI controller provides a faster reaction than the integral action alone.

Individual mainline travel times as functions of the start time for different ramp-metering scenarios are plotted in Figure 9. It can be observed from the figure that under all the metered scenarios, the onset of the congestion was slower and the dissipation of the congestion was faster than the unmetered scenario. In particular, the switching mainline control with queue regulation was able to dissipate congestion faster than the upstream and downstream ALINEAs. This

is more obvious in Figure 9b, where the individual travel times are plotted for the scenarios when the queue length constraints were not enforced. The switching mainline control was able to achieve a lower maximum travel time, and it dissipated the congestion much faster than the downstream and upstream ALINEAs did.

CONCLUSIONS

This paper presented a novel switching traffic-responsive ramp-metering controller that adapts to the different traffic dynamics under free-flow or congested conditions. The approach of multirate LQI was used to compensate for disturbances and accommodate the difference in the model sampling time and the metering rate update interval. In addition, a PI queue length regulator was designed to prevent the on-ramp queue from exceeding the storage capacity and yield improved performance over the current ad hoc queue override scheme. A queue length estimator was also designed to provide feedback to the queue length regulator, by using the queue-detector speed data that are available in the field. A localized strategy was proposed to achieve the control goal of reducing the spatial and temporal extent of the congestion, by using locally available information.

Test results on the calibrated VISSIM I-210W microscopic model demonstrated the performance and effectiveness of the switching ramp-metering controller, the queue length estimator and regulator, and the overall control strategy. The total vehicle and passenger delays were both reduced by 16%, and the total vehicle time was improved by 5.6%. As a comparison, simulation results of ALINEA were also presented. The switching mainline traffic-responsive control was able to outperform ALINEA, when both algorithms were combined with the same queue length estimator and regulator.

ACKNOWLEDGMENT

This work was supported by the California Partners for Advanced Transit and Highways.

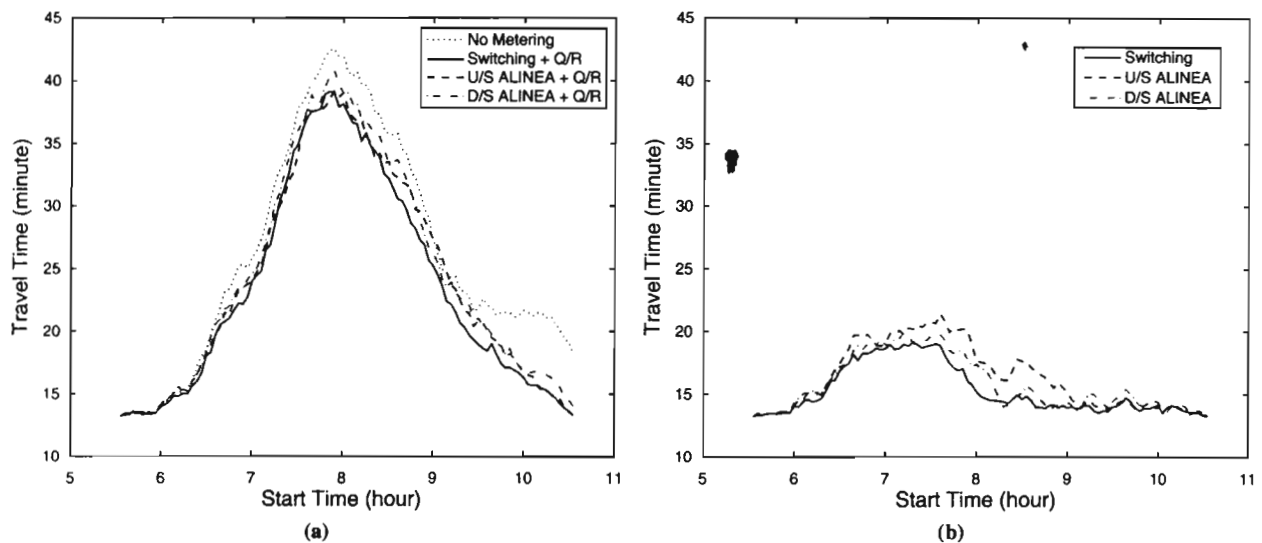


FIGURE 9 Individual mainline travel times as functions of start time for different ramp-metering scenarios: (a) queue regulation active and (b) queue regulation inactive.

REFERENCES

- Schrank, D., and T. Lomax. *The 2004 Urban Mobility Report*. Texas Transportation Institute, Austin, Sept. 2004. mobility.tamu.edu/. Accessed Feb. 2005.
- Cambridge Systematics and Texas Transportation Institute. *Traffic Congestion and Reliability: Linking Solutions to Problems*. FHWA, U.S. Department of Transportation, July 19, 2004. www.ops.fhwa.dot.gov/congestion_report/. Accessed Feb. 2005.
- Muñoz, L., X. Sun, R. Horowitz, and L. Alvarez. Traffic Density Estimation with the Cell Transmission Model. *Proc., 2003 American Control Conference*, Denver, Colo., 2003, pp. 3750–3755.
- Muñoz, L., X. Sun, R. Horowitz, and L. Alvarez. Piecewise-Linearized Cell Transmission Model and Parameter Calibration Methodology. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1965, Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 183–191.
- 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.
- 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, Aug. 1994, pp. 269–287.
- Daganzo, C. F. The Cell Transmission Model, Part II: Network Traffic. *Transportation Research B*, Vol. 29, No. 2, April 1995, pp. 79–93.
- Muñoz, L., X. Sun, D. Sun, G. Gomes, and R. Horowitz. Methodological Calibration of the Cell Transmission Model. *Proc., 2004 American Control Conference*, Boston, Mass., 2004, pp. 798–803.
- 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, 2003, pp. 6333–6338.
- Sun, X., L. Muñoz, and R. Horowitz. Mixture Kalman Filter Based Highway Congestion Mode and Vehicle Density Estimator and Its Application. *Proc., 2004 American Control Conference*, Boston, Mass., 2004, pp. 2098–2103.
- Sun, X., and R. Horowitz. A Localized Switching Ramp-Metering Controller with a Queue Length Regulator for Congested Freeways. *Proc., 2005 American Control Conference*, Portland, Ore., 2005, pp. 2141–2146.
- Papageorgiou, M., J.-M. Blosseville, and H. Hadj-Salem. Modeling and Real-Time Control of Traffic Flow on the Southern Part of Boulevard Périphérique in Paris: Part II: Coordinated On-Ramp Metering. *Transportation Research A*, Vol. 24, No. 5, Sept. 1990, pp. 361–370.
- Gordon, R. L. Algorithm for Controlling Spillback from Ramp Meters. In *Transportation Research Record 1554*, TRB, National Research Council, Washington, D.C., 1996, pp. 162–171.
- Ozbay, K., and P. Kachroo. *Feedback Ramp Metering for Intelligent Transportation System*. Kluwer Academic Publishers, New York, 2003.
- Ozbay, K., I. Yasar, and P. Kachroo. Comprehensive Evaluation of Feedback-Based Freeway Ramp-Metering Strategy by Using Microscopic Simulation: Taking Ramp Queues into Account. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1867, Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 89–96.
- Smaragdīs, E., and M. Papageorgiou. Series of New Local Ramp Metering Strategies. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1856, Transportation Research Board of the National Academies, Washington, D.C., 2003, pp. 74–86.
- VISSIM. PTV AG. 2004. www.english.ptv.de/cgi-bin/traffic/traf_vissim.pl. Accessed Feb. 2005.
- Rousseeuw, P. J. Least Median of Squares Regression. *Journal of the American Statistical Association*, Vol. 79, No. 388, 1984, pp. 871–880.
- Gomes, G., and R. Horowitz. Globally Optimal Solutions to the Onramp Metering Problem, Part I. *Proc., Seventh International IEEE Conference on Intelligent Transportation Systems*, Washington, D.C., 2004, pp. 509–514.
- Gomes, G., and R. Horowitz. Globally Optimal Solutions to the Onramp Metering Problem, Part II. *Proc., Seventh International IEEE Conference on Intelligent Transportation Systems*, Washington, D.C., 2004, pp. 515–520.
- TIGER: Topologically Integrated Geographic Encoding and Referencing system. U.S. Census Bureau. Jan. 2005. www.census.gov/geo/www/tiger/. Accessed Feb. 2005.
- 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.
- Gomes, G. C. *Optimization and Microsimulation of On-ramp Metering for Congested Freeways*. PhD dissertation. University of California, Berkeley, 2004. Accessed Feb. 2005.