International Journal of Operations Research

International Journal of Operations Research Vol. 8, No. 1, 23–35 (2011)

Optimization of Coordinated Multi-Ramp Metering Control with Simultaneous Perturbation Stochastic Approximation (SPSA)

Steven I. Chien[1,*], Jiangtao Luo[2], and Ching-Jung Ting[3]

[1]Department of Civil and Environmental Engineering, New Jersey Institute of Technology, Newark, New Jersey - 07102, USA

[2]T & M Associates, Middletown, New Jersey - 07748, USA

[3]Department of Industrial Engineering and Management, Yuan Ze University, Chung-Li, Taiwan - 32003

Received July 2010; *Revised November 2010*; *Accepted December 2010*

Abstract— Ramp metering has been recognized as one of viable strategies of freeway traffic management, which may improve travel speed as well as reduce delay. In this study, a dynamic model is developed to optimize metering rates of a series of on-ramps which maximize the total throughput subject to vehicle density, roadway capacity, and queue length constraints with simultaneous perturbation stochastic approximation (SPSA). A calibrated microscopic traffic simulation model is used to evaluate the performance of the developed model. Results show that the total throughput of the studied freeway can be increased without increasing the total delay after implementing optimal metering rates. The developed SPSA has demonstrated itself as an efficient algorithm to search for the optimal solution of a dynamic, multivariate, and non-linear problem.
Keywords— Ramp metering, traffic management, freeway, delay, optimization, simulation

1. INTRODUCTION

Ramp metering has been recognized as one of viable strategies of relieving freeway congestion, which may improve travel speed as well as reduce delay by regulating flow entering the mainline via the ramps. While ramp metering systems have been applied in several U.S. cities (Piotrowicz and Robinson, 1995), more concerns have been focused on evaluating its effectiveness (Kotsialos and Papageorgiou, 2001; Hourdakis and Michalopoulos, 2002; Chu et al., 2004; Zhang and Levinson, 2005).

In this study, a dynamic model is developed to optimize metering rates of a series of on-ramps which maximize the total throughput with simultaneous perturbation stochastic approximation (SPSA). One of the major criticisms of ramp metering has been the delay caused on the ramps because of the queues created by reduced ramp capacity. Therefore, constraints (e.g., meter locations, ramp storage capacities, lower and upper bounds of ramp metering rates) are considered.

Based on collected geometric and traffic data, a microscopic traffic simulation model which contains a 15-mile long freeway segment on Interstate I-80 in New Jersey is developed with a microscopic corridor traffic simulation program, called CORSIM, developed by the Federal Highway Administration. With that various ramp control scenarios can be evaluated, while the performance and benefit of metering control can be quantified.

2. LITERATURE REVIEW

A variety of models have been developed before for optimizing metering rates of freeway ramps. Papageorgiou and Kotsialos (2002) provided an overview on different ramp metering algorithms. Without considering the dynamic nature of traffic, pre-timed linear programming (LP) models were developed to search for metering rates to achieve the optimum objectives [e.g., minimize total travel time (Wattleworth, 1967) and maximize throughput (Chen et al., 1974)]. Such models aimed at sending vehicles onto the freeway subject to the available capacity on the mainline. The advantage of the LP models is the implementation simplicity (Masher et al., 1976), but unable to efficiently adapt to dynamic traffic conditions (Papageorgiou, 1980).

* Corresponding author's email: chien@njit.edu

To capture dynamic traffic condition, local traffic responsive models were developed. With real-time traffic information (e.g., flow, speed, and density), the ramp metering rates can be optimized. In those models, gap-acceptance concept was introduced to break up a platoon of vehicles on the mainline into individual vehicles or a number of smaller groups. Thus, the entering vehicles from on-ramps can smoothly merge into the mainline stream. Because of reduced traffic interruption, the travel speed may be improved. The optimal metering rates were determined based on queue length, wait time and pre-specified critical gap of each on-ramp (Drew, 1967; Wiener et al., 1970). As a local occupancy-based metering control algorithm, ALINEA (Papageorgiou et al., 1991; Hadj-Salem and Papageorgiou, 1995) was developed on the basis of classical feedback control principles and attempted to maximize the throughput via maintaining a desired occupancy downstream of the mainline. The metering rates of controlled ramps were dynamically adjusted but treated as isolated rather than connected ones. Therefore, the interdependency of congestion at ramp junctions as well as the demand change caused by delay at metered ramps was unable to consider. Papamichail et al. (2010) proposed a nonlinear model-predictive hierarchical (three-layer) control approach for coordinated ramp metering of freeway networks. The local feedback control strategy ALINEA was used in the third layer.

The system-wide models were developed to handle a series of ramps of a freeway corridor or network, which attempted to capture the interdependence of traffic conditions among consecutive ramps. Those models often led to hierarchical non-linear optimization programming carried by a centralized computer system to process traffic data, optimize metering rates, and execute the control law through ramp meters in real-time. The equivalent traffic demand was often formulated to represent a queue on a metered ramp, so that vehicles spillback onto local streets can be prevented. A quadratic programming problem was formulated (Yuan and Kreer, 1971) to determine desirable metering rates, considering vehicle conservation, the Greenshields' speed-density, and metering rate constraints. Later, the total travel time was formulated as a quadratic function (Stephanedes and Chang, 1993), which was minimized by a conjugate gradient search method. The system-wide ramp metering model usually involves in rather complex, nonlinear optimization techniques required considerable computation effort to search for the optimal solution. However, the solution of those models might be difficult to implement due to oscillatory metering rates, especially when collected traffic data were frequently corrupted with noise or transmission errors.

Heuristics has been applied in searching for optimal ramp metering rates to adapt to dynamic change in managing traffic on freeways. A number of models were developed previously to achieve local optimum for large-scale freeway networks. In a decentralized metering control model (Goldstein and Kumar, 1982), a series of on-ramps were divided into overlapping groups that were controlled by sub-systems individually to improve efficiency in searching for the optimal solution. An expert fuzzy controller model (Chen et al., 1990) was developed to maximize freeway throughput while minimizing delay on local streets. The controller can infer to the degree true condition to determine metering rates. Unlike the evolutionary programming model (McDonnell, 1995), a fuzzy logic model containing seventeen rules was developed to control entering traffic on multiple on-ramps (Taylor et al., 1998), which was found outperforming the local metering model and the bottleneck model in terms of lower mainline occupancy and higher throughput (Taylor et al., 2000).

Papamichail et al. (2010) presented a traffic-responsive feedback control strategy, heuristic ramp-metering coordination (HERO) that coordinates local ramp-metering actions in freeway networks. HERO has been implemented at six consecutive inbound on-ramps on the Monash Freeway in Melbourne, Australia. Meng and Khoo (2010) proposed a multi-objective optimization model incorporating a modified cell-transmission model (MCTM) that captures dynamic traffic flow pattern with ramp metering operations. The MCTM then is embedded in the Non-dominated Sorting Genetic Algorithm II (NSGA-II) to solve the multi-objective optimization model.

To deal with a dynamic control problem in a large-scale network, the computation time required to search for the optimal solution is critical to react existing and in coming traffic in time. Recently, simultaneous perturbation stochastic approximation (SPSA) has dragged attention because of its efficiency and simplicity (Spall, 1998; Ting and Schonfeld, 1998; Kleinman et al., 1998). The procedure to implement SPSA is fairly simple, which starts with an initial guess, which will be iteratively updated until the desired solution is found. SPSA does not rely on direct measurements of the gradient (derivative) of the objective function. Instead, it relies on measurements of the objective function, which avoids the difficulties to obtain the relationship among parameters and decision variables.) SPSA has been applied to optimize metering rates of on-ramps individually with the satisfied results (Chien et al., 2008). This study aims to dynamically optimize coordinated metering rates, considering the joint impact of traffic affected by the controlled ramps.

3. METHODOLOGY

According to the traffic flow theory, high density of vehicle corresponding to a low speed of flow results in a low volume. Therefore, when high traffic demand occurs, vehicles merging to the mainline stream shall be regulated to reduce flow interruptions. To formulate the proposed coordinated multi-ramp metering control model, the following assumptions are made:

(a) The flow on the mainline is steady without serious incidents blocking the traffic, which can represent a steady relationship among flow, speed and density;
(b) Each vehicle passes through a meter separately based on a first-come first-released discipline; and
(c) The average vehicle length is 20 feet, which is used to approximate the storage capacity of the metered ramps.

The objective function developed here is to maximize total throughput subject to constraints of link densities, capacities and the boundaries of metering rates. A general *N*-segment freeway corridor with multiple on-ramps and off-ramps shown in Figure 1 can be formulated as Eq. 1 (Chang et al., 1994). Considering time varying traffic entering the corridor, the metering control period can be divided into a series of equal intervals. The density equation of link i at interval k denoted as $\rho_i(k)$ is formulated as

$$\rho_i(k) = \rho_i(k-1) + [q_{i-1}(k) + \delta_i^{on} R_i(k) - \delta_i^{off} \theta_i(k) Q_i(k) - q_i(k)] \frac{T}{L_i l_i}, \quad \forall i \tag{1}$$

where i : index of links;
k : index of time intervals;
l_i : number of through lanes on link i;
L_i : length of link i (miles);
N : number of links in the network;
$q_i(k)$: traffic volume from link i to link i+ 1 at interval k (vph);
$Q_i(k)$: mean flow rate of link i at interval k (vph);
$R_i(k)$: metering rate of link i at interval k (vph);
T: duration of a time interval (hours);
δ_i^{on} : binary variable (1 if link i is an on-ramp; otherwise, 0);
δ_i^{off} : binary variable (1 if link i is an off-ramp; otherwise, 0); and
$\theta_i(k)$: turning percentage of mainline flow from link i to the off-ramp at interval k (%).

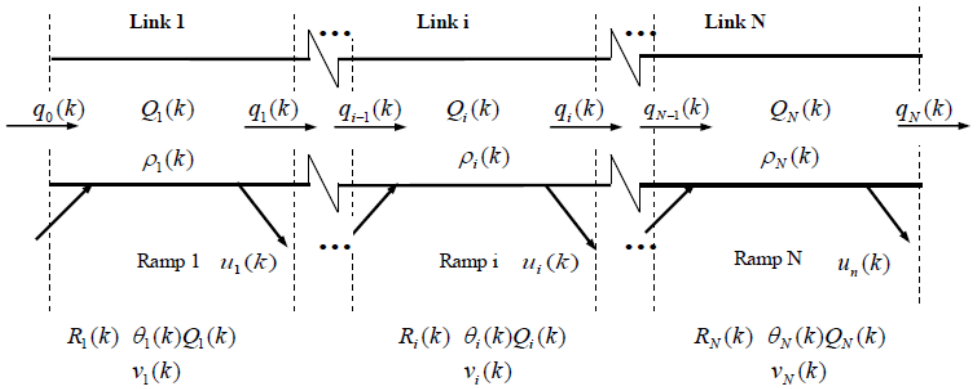

Figure 1. A general freeway corridor with N segments

Eq. 1 shows that the projected mean density $\rho_i(k)$ of link i at interval k is affected by that at interval k - 1 denoted as $\rho_i(k$-1), and the entering and exiting volumes from and to the ramps. The transition flow rate denoted as $q_i(k)$ can thus be formulated as

$$q_i(k) = \alpha_i(k)[1 - \delta_i^{off} \theta_i(k)] Q_i(k) + [1 - \alpha_i(k)][Q_{i+1}(k) - \delta_{i+1}^{on} R_{i+1}(k)], \quad \forall i \tag{2}$$

where $\alpha_i(k)$ is a weighted factor representing the interaction between flows on links i and i + 1 at interval k. An assumed value of 0.5 for $\alpha_i(k)$ indicates the same influence to flows on links i and i + 1, which may simplify the calculation. Note that $\alpha_N(k)$ is equal to 1 on the last link N.

Eq. 3, representing density functions for links 1 through N, can be derived by substituting Eq. 2 into Eq. 1. Thus,

$$\begin{aligned} \rho_i(k) = \rho_i(k-1) &+ \frac{T}{L_i l_i} \alpha_{i-1}(k)[1 - \delta_{i-1}^{off} \theta_{i-1}(k)] Q_{i-1}(k) + \frac{T}{L_i l_i} \{1 - \alpha_{i-1}(k) - \alpha_i(k) - \delta_i^{off}[1 - \alpha_i(k)] \theta_i(k)\} Q_i(k) + \frac{T}{L_i l_i} [\alpha_i(k) - 1] Q_{i+1}(k) \\ &+ \frac{T}{L_i l_i} \alpha_{i-1}(k) \delta_i^{on} R_i(k) + \frac{T}{L_i l_i} [1 - \alpha_i(k)] \delta_{i+1}^{on} R_{i+1}(k), \quad \forall i \end{aligned} \tag{3}$$

where $\alpha_0(k) = \alpha_N(k) = 1$; $Q_0(k) = q_0(k)$; $\delta_0^{on} = 0$; and $\theta_0(k) = 0$. The relationship among speed, flow and density can be expressed by

$$Q_i(k) = \rho_i(k) S_i(k), \quad \forall i \tag{4}$$

where $S_i(k)$, representing the mean speed of link i at interval k, can be collected by loop detectors or other types of road sensors.

Note that the total traffic throughput (*TTT*) defined here is the total number of vehicles discharged from freeway exit links over the control intervals. Thus, it can be formulated as

$$TTT = \sum_{k=1}^{K}\left[\sum_{i=1}^{N-1}\delta_i^{off}\theta_i(k)Q_i(k) + Q_N(k)\right]T \tag{5}$$

where K is the last time interval of the control period. By substituting Eq. 4 into Eq. 5, *TTT* can be derived as

$$TTT = T\sum_{k=1}^{K}\left[\sum_{i=1}^{N-1}\delta_i^{off}\theta_i(k)\rho_i(k)S_i(k) + \rho_N(k)S_N(k)\right] \tag{6}$$

4. SOLUTION ALGORITHM – SPSA

Solve a dynamic and large-scale nonlinear optimization problem is challenging, even under the advent of fully deployed advanced surveillance systems (e.g., various types of detectors lead to the improvement in data collection and processing) and developed optimization methods (e.g., linear-quadratic optimization and hierarchical decomposition algorithms). In real-world systems, it is fairly complicated to formulate detailed, explicit relationship between adjustable or controllable system parameters and the resulting system performance. In optimizing nonlinear programming problems, the gradient of the objective function with respect to the decision variables needs to be derived. The optimal solution can be obtained by setting the gradient equal to zero and solving it. However, the gradient might be difficult to derive, depending on the complexity of the objective function.

The SPSA algorithm, a recursive optimization technique for finding local optimizers of linear or nonlinear objective functions, was first introduced and developed by Spall (1992). Based on the measurement of the objective function (not on the gradient of the objective function), SPSA computes the positively and negatively perturbed objective values in each iteration. It is similar to finite difference stochastic approximation (FDSA), only measurements (possibly noisy) of an objective function to form gradient estimates and converge to a local optimum are required. However, unlike FDSA, SPSA requires only two objective function evaluations per gradient estimate, which is a substantial advantage in solving high-dimensional problems.

As discussed earlier, the SPSA algorithm uses objective function measurements to iteratively update system control parameters until the local optimum is yielded. Specifically, let $\lambda \in \mathrm{R}^p$ be a vector whose components represent decision variables to be optimized, such as ramp metering rates and mainline traffic flow in this study. Supposing that $L(\lambda)$ represents the objective function, the goal is to find a root λ_h^* of the gradient $g(\lambda)$ of $L(\lambda)$, where λ should be conducted from

$$g(\lambda) = \frac{\partial L(\lambda)}{\partial \lambda} = 0 \tag{7}$$

Assume that the measurement $y(\lambda)$ of the objective function can be represented by Eq. 8 for any λ. Thus,

$$y(\lambda) = L(\lambda) + noise \tag{8}$$

In the approximation process, SPSA iteratively produces a sequence of estimates (e.g., $\hat{\lambda}_0$, $\hat{\lambda}_1$, $\hat{\lambda}_2$, … , $\hat{\lambda}_{h+1}$) generated by a step procedure discussed below:

Step 0: Initialization

Set counter index h equal to 0. Pick initial guess $\hat{\lambda}_0$ and non-negative coefficients, a, c, β and γ in SPSA gain sequences as shown in Eqs.9 and 10. It should depend on the practical scenario to set initial guess $\hat{\lambda}_0$ as empirical values. Note that $\hat{\lambda}_0$ represents the initial values of a vector with p decision variables, which must satisfy the constraints discussed later in Eqs. 14 - 16. A large a enhances the performance in the later iterations by using a larger step size to search for the solution, while it will be effective to set c as relatively smaller positive number. Recommended values for β and γ are 0.602 and 0.01 (Ting and Schonfeld, 1998), respectively, which affect the convergence speed to yield the optimal solution.

$$a_h = a(h+1)^{-\beta} \tag{9}$$

$$ch = c(h+1) - \gamma \tag{10}$$

Step 1: Generate Simultaneous Perturbation Vector

Generate a random vector denoted as Δ_h with p components, which are independently generated from a zero-mean probability distribution, such as a Bernoulli ±1 distribution with probability of 0.5 for each ±1 outcome. Components generated based on uniform and normal distribution have infinite inverse moments, which and are not recommended, because the solution will not converge.

Step 2: Evaluate Objective Function

Obtain two measurements of the objective function denoted as $L(\lambda)$, based on a simultaneous perturbation offset the current $\hat{\lambda}_h$ (e.g., $y(\hat{\lambda}_h + c_h\Delta_h)$ and $y(\hat{\lambda}_h - c_h\Delta_h)$) at iteration h with c_h and Δ_h obtained from Steps 0 and 1, respectively. Note that $\hat{\lambda}_h$ represents the provisional optimal solution. For example, $\hat{\lambda}_h$ must satisfy the constraints. If not, the boundary values considered in the constraints will apply, and the corresponding values of c_h will be re-calculated.

Step 3: Approximate Gradient

Generate the simultaneous perturbation approximation to the (unknown) p-dimensional gradient denoted as $\hat{g}(\hat{\lambda}_h)$:

$$\hat{g}(\hat{\lambda}_h) = \begin{bmatrix} \dfrac{y(\hat{\lambda}_h + c_h\Delta_h) - y(\hat{\lambda}_h - c_h\Delta_h)}{2c_h\Delta_{h1}} \\ \dfrac{y(\hat{\lambda}_h + c_h\Delta_h) - y(\hat{\lambda}_h - c_h\Delta_h)}{2c_h\Delta_{h2}} \\ \vdots \\ \dfrac{y(\hat{\lambda}_h + c_h\Delta_h) - y(\hat{\lambda}_h - c_h\Delta_h)}{2c_h\Delta_{hp}} \end{bmatrix} \tag{11}$$

where Δ_{hi} is the i^{th} component of Δ_h. The denominators in Eq. 11 will alter the search direction for the optimal solution in next iteration.

Step 4: Update $\hat{\lambda}_h$

Use the standard stochastic approximation form as formulated in Eq. 12 to estimate the solution at iteration $h+1$, denoted as $\hat{\lambda}_{h+1}$. Thus,

$$\hat{\lambda}_{h+1} = \hat{\lambda}_h - a_h\hat{g}(\hat{\lambda}_h) \tag{12}$$

Step 5: Iteration or Termination

Return to Step 1 and increase the counter index from h to $h+1$. Terminate SPSA if the difference between successive iterations is less than a pre-set value that is very small to approximate zero; and the last $\hat{\lambda}_h$ is the estimate of the optimum λ_h^*.

The SPSA algorithm is very general and can be applied in many different situations to optimize many different kinds of objective functions. For instance, $L(\lambda)$ could be the function dealt with total throughput, while λ could represent the optimal metering rates at different ramps in a freeway corridor or network considering various objectives (e.g., minimize delay or maximize throughput). The constraints can be imposed by adding penalty functions in the objective function. The flexibility of the algorithm stems from the fact that only objective function measurements are required, instead of full objective function or gradient information. The objective function measurements required by the SPSA algorithm can come from a real system as well as from a computer simulation from a real world probabilistic system. The efficiency of SPSA in solving high dimensional problems, especially when evaluating the objective function is expensive or time-consuming, has been discussed by Spall (1992).

5. OPTIMIZATION

In order to integrate SPSA algorithm and the ramp metering control model, the equivalent objective function of Eq. 6 can be formulated as Eq. 13. Through this conversion, the objective total throughput can be maximized by minimizing $L(\lambda)$:

$$L(\lambda) = Z - T\sum_{k=1}^{K}\left[\sum_{i=1}^{N-1}\delta_i^{off}\theta_i(k)\rho_i(k)S_i(k) + \rho_N(k)S_N(k)\right] \tag{13}$$

where Z represents a big number. By substituting Eq. 3 into Eq. 13, the decision variables are time varying metering rates $R_i(k)$, where $\boldsymbol{\lambda}$ represents a vector $R_i(k)$ for metered ramp i at time interval k. In Eqs. 3, 4, and 13, the values of geometry related parameters [e.g., Z, N, δ_i^{on}, δ_i^{off}, $\alpha_i(k)$, T, L_i, and l_i] can be collected from the study site, while the traffic values of related parameters [e.g., $Q_i(k)$, $\theta_i(k)$, $\rho_i(k\text{-}1)$, $q_0(k)$ and $S_i(k)$] can be calculated from the data collected form detectors or other types of surveillance systems that can detect traffic volume, speed, and density information.

One of the major criticisms of ramp metering has been delay caused on the ramps because of the queues created by ramp control strategies. Therefore, constraints (e.g., meter locations, ramp storage capacities, lower and upper bounds of ramp metering rates) are considered in this study while optimizing the metering rate. Considering the density and capacity of the link and the boundaries to define the feasible range of metering rages, the objective function formulated in Eq. 13 should be minimized subject to a set of constraints formulated in Eqs. 14, 15, and 16:

$$0 \le \rho_i(k) \le \rho^{\max}, \ \forall i,k \tag{14}$$

$$0 \le Q_i(k) \le Q_i^{\max}, \ \forall i,k \tag{15}$$

$$R_i^{\min} \le R_i(k) \le R_i^{\max}, \ \forall i,k \tag{16}$$

Eq. 14 defines that the density of link i at interval k should be positive and less than the maximum density $\rho^{\max}$. Similarly, the flow of link i at any interval k in Eq. 15 must be positive and less than the link capacity Q_i^{max}. The constraint in Eq. 16 defines that the range of feasible metering rates, where R_i^{min} and R_i^{max} represent the lower and upper boundaries of metering rates, respectively.

Assume that vehicles approach a ramp meter with a mean arrival rate m, the relation between the queue storage capacity L_q (number of vehicles) and the metering rate $R_i(k)$ can be expressed by an M/M/1 queuing model as shown in Eq. 17.

$$L_q = \frac{\tau m}{R_i(k) - m} \tag{17}$$

where τ is the ratio of vehicle arrival rate m and ramp service rate $R_i(k)$. Note that, the ramp service rate is the ramp metering rate during control interval. In this study, the service rate is equivalent to the ramp metering rate.

The maximum storage capacity of a metered ramp can be determined by the total lane-miles of the ramp divided by the average vehicle length (e.g., 20 feet). The minimum metering rate R_i^{min} guarantees that the queuing length L_q will not exceed the storage capacity. Thus, the queuing vehicles on the metered ramp will not spillback to the local street. R_i^{min} can be derived from Eq. 17 as

$$R_i^{\min} = \left(\frac{L_q + \tau}{L_q} \right) m \tag{18}$$

According to previous study (Chien and Luo, 2008), the maximum metering rate R_i^{max} was suggested 900 vph (4.0 seconds/vehicle), considering driver's reaction and operation time and the time consumed for vehicle acceleration. The feasible range of metering rate can thus be determined.

In this study, SPSA is integrated into a microscopic traffic simulation model as shown in Figure 2, in which real-time metering dates will be optimized corresponding to time-varying traffic volumes fed into a freeway network. Therefore, the optimal ramp metering rates can be dynamically found by SPSA, considering dynamic traffic conditions and capacity constraints (e.g. on-ramp volumes, mainline capacity and the boundaries of feasible metering rates, etc.). The input parameters for optimizing metering rates will be automatically fed by the simulation model.

6. SIMULATION ANALYSIS

A 12-mile segment of eastbound I-80 in New Jersey is selected to evaluate the proposed dynamic metering control model, which contains seven on-ramps and five off-ramps supporting the entry and exit flows. The daily traffic is over 100,000 vehicles. The mainline speed limit is 65 mph, and the speed limits on ramps range from 20 to 40 mph.

6.1 Network Modeling

The network is coded along the mainline considering critical points such as interchanges, potential ramp meter locations, curvature/superelevation change, and entry or exit points. In the network modeling as shown in Figure 3, each link consists of two nodes, which represents one direction of a freeway segment.

The data describing freeway geometric, traffic movements and freeway features (e.g., location of ramps and warning signs) are input into CORSIM. Traffic operations during a peak hour (7:00am-8: 00am) on the eastbound of I-80 are simulated, while

discrepancies are reduced by calibrating parameters (e.g. car-following sensitivity factor, lane change time and percentage of cooperative driver) in the simulation model (Chien et al., 2005).

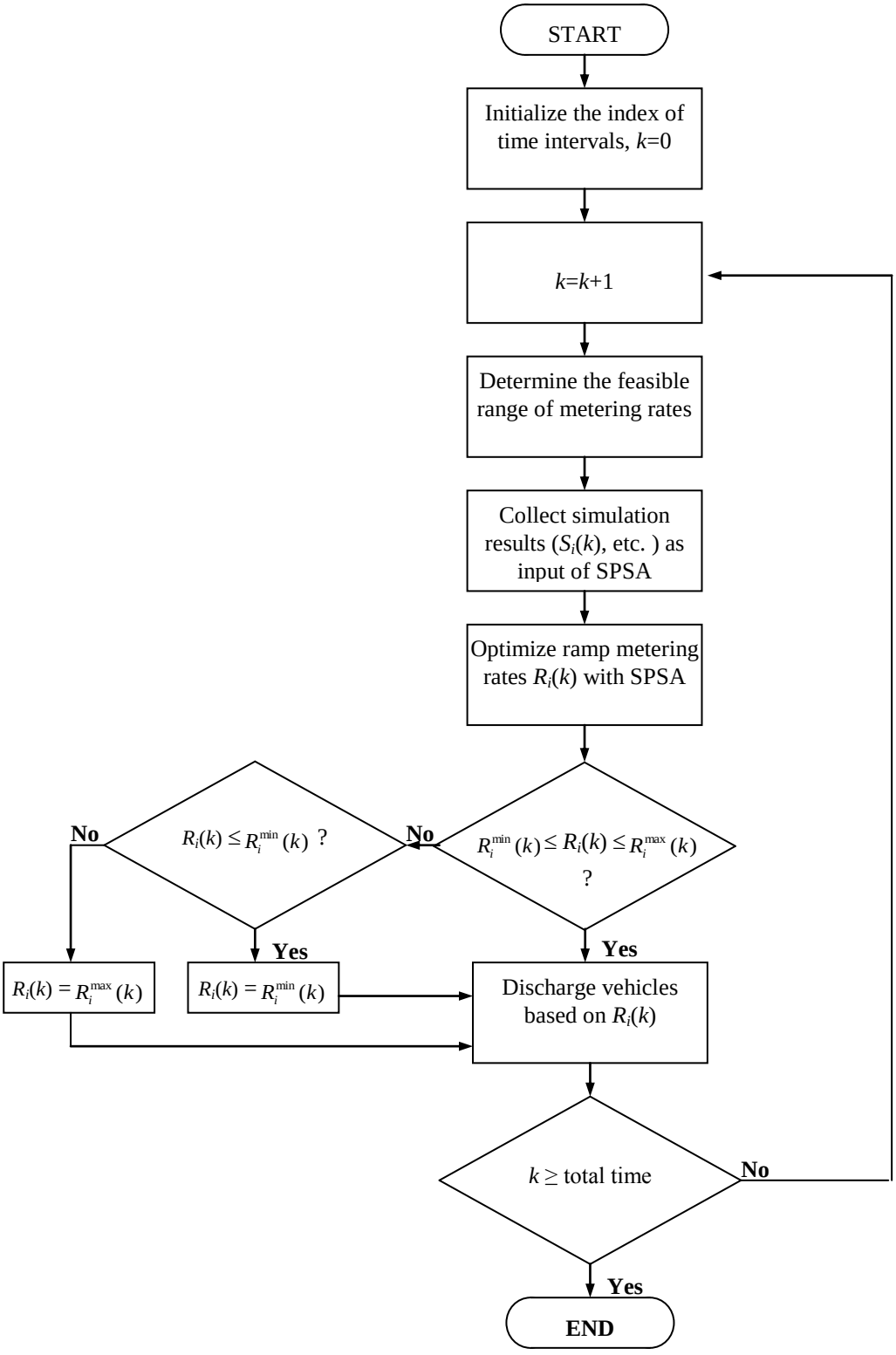

Figure 2. The real-time metering control system

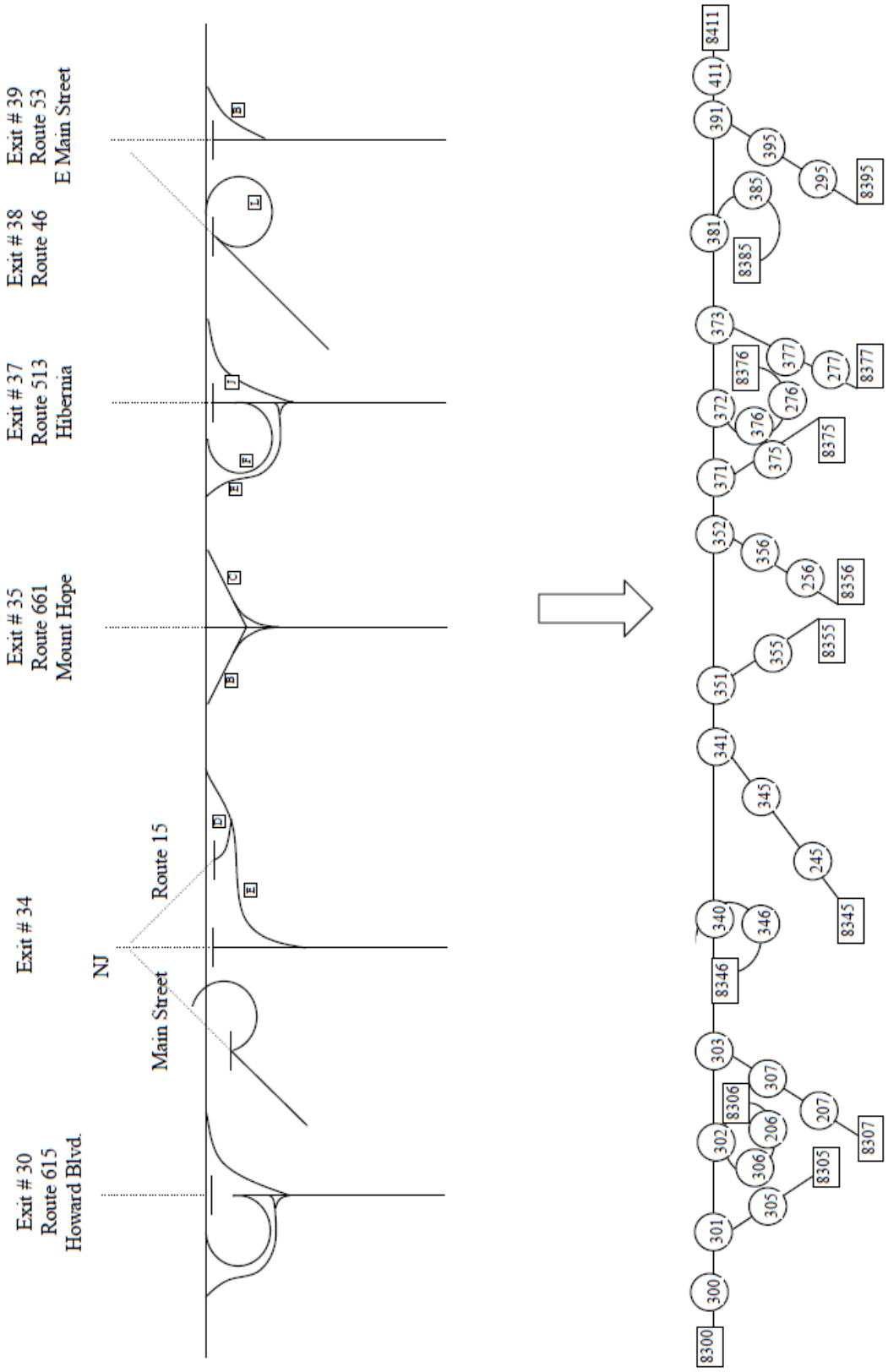

Figure 3. Link – node diagram of the study site

The calibrated model is validated through comparing simulation outputs with the field counterparts. The statistical analysis is conducted by calculating mean absolute percentage errors (MAPE) and root mean square errors (RMSE) between the field and the simulated traffic volumes of various links on the mainline. The results show that the simulation model can accurately reflect the traffic speed on this segment.

6.2 Data Processing

In Eq. 1, the density $\rho_i(k)$ of link i at interval k can be determined by $\rho_i(k\text{-}1)$ and the entry and exit flows. As suggested by Highway Capacity Manual (2000), $\rho_i(k)$ should not exceed the maximum value 45 veh/ln-mi, if the free flow speed is 65 mph. If $\rho_i(k)$ is greater than 45 veh/ln-mi, the suggested ramp metering rate $R_i(k)$ should be set as default value of $R_i^{\min}(k)$.

The SPSA algorithm is developed to maximize the total throughput achieved by optimized ramp metering rates. The input data, including $\theta_i(k)$, $\rho_i(k\text{-}1)$, $q_0(k)$, $S_i(k)$, $Q_i(k)$ and $R_i(k)$, are initialized at the beginning of the first time interval, which will be updated in real-time during the course of the evaluation period. While optimizing the metering rates, parameters $\boldsymbol{a}$ and $\boldsymbol{c}$ in SPSA are modified iteratively in Eqs. 9 and 10.

As discussed earlier, the optimal metering rate in each interval is derived based on the data (simulation output) collected from the previous time interval. Reducing the length of time interval can increase the precision of the control by accurately capturing time varying demand and density. However, the computation effort increases as well. In real world, the appropriate length of time interval should be determined to achieve optimal control. In this study, the time interval of 3 minutes is applied.

The input of the proposed model, including link volume, accumulated vehicle-mile, and average delay of the network, are collected during simulation and fed into the SPSA algorithm. Since CORSIM only generates accumulated statistics, additional calculation is performed to obtain the MOEs (e.g., volume, vehicle-mile, and delay) statistics for individual intervals. A macro program has been developed with Microsoft Excel to fast retrieve accumulative simulation statistics generated in each time interval. The time varying total delay and total throughput can thus be derived for benefit assessment. Note that the total delay is obtained from the total vehicle-miles travel multiplied by the average delay in minute per vehicle-mile, while the total throughput is the sum of the exit volumes in each interval.

The benefit assessment of the dynamic ramp metering control with SPSA is conducted by comparing total throughput and delay with and without the proposed dynamic metering control model.

6.3 Performance Analysis

In order to evaluate the benefit of the developed model, the time-varying traffic flows over 16 time intervals, entering from the mainline entry link and on-ramps are summarized in Table 1. To observe the impact of entering flow to the performance of the proposed control, the flow of link 300-301 increases 200 vph by interval, while the flows entering the on-ramps are fixed. Two scenarios are designed for conducting a comparative analysis.

Table 1. Traffic Demand for the Network

Time Interval*	Entry Flow	Node 306	Node 307	Node 345	Node 356	Node 376	Node 377	Node 395
	vph							
1	2960	360	489	2159	819	400	1216	430
2	3160							
3	3360							
4	3560							
5	3760							
6	3960							
7	4160							
8	4360							
9	4560							
10	4760							
11	4960							
12	5160							
13	5360							
14	5560							
15	5760							
16	5960							

*: The duration of each time interval is 3 minutes.

- Scenario A (Individual Control): The optimal metering rate of each ramp was optimized individually without coordination. The metered ramps are Nodes 306, 307, 356, 376, and 395, where the corresponding optimal metering rates over the study intervals are summarized in Table 2. Note that the durations of time intervals are identical and is 3 minutes.
- Scenario B (Coordinated Control): The coordinated metering rates of controlled ramps (same as those discussed in Scenario A) are jointly optimized with SPSA. The optimal metering rates over the study intervals were found slightly different from those derived in Scenario A (Table 3).

Table 2. Optimal Metering Rates for Individual Ramp Control

Time Interval*	**Node 306**	**Node 307**	**Node 356**	**Node 376**	**Node 395**
	Metering rate (vph)				
1	705	794	895	737	766
2	705	794	895	737	766
3	705	794	895	737	766
4	720	815	900	753	783
5	735	836	900	771	802
6	751	858	900	789	822
7	767	881	900	808	842
8	779	881	900	820	842
9	779	881	900	820	842
10	779	881	900	820	842
11	779	881	900	820	842
12	779	881	900	820	842
13	779	881	900	820	842
14	779	881	886	819	774
15	779	881	860	758	761
16	779	881	835	744	748

*: The duration of each time interval is 3 minutes.

Table 3. Jointly Optimal Ramp Metering Rates

Time Interval*	**Node 306**	**Node 307**	**Node 356**	**Node 376**	**Node 395**
	Metering Rate, vph (**Headway**, seconds)				
1	650 (5.54)	766 (4.70)	900 (4.00)	753 (4.78)	799 (4.51)
2	650 (5.54)	766 (4.70)	900 (4.00)	753 (4.78)	799 (4.51)
3	650 (5.54)	766 (4.70)	900 (4.00)	753 (4.78)	799 (4.51)
4	664 (5.42)	766 (4.70)	900 (4.00)	769 (4.68)	816 (4.41)
5	680 (5.29)	766 (4.70)	900 (4.00)	787 (4.57)	835 (4.31)
6	696 (5.17)	776 (4.64)	900 (4.00)	805 (4.47)	855 (4.21)
7	712 (5.06)	799 (4.51)	900 (4.00)	824 (4.37)	875 (4.11)
8	724 (4.97)	799 (4.51)	900 (4.00)	824 (4.37)	875 (4.11)
9	723 (4.98)	799 (4.51)	900 (4.00)	824 (4.37)	875 (4.11)
10	719 (5.01)	754 (4.77)	900 (4.00)	900 (4.00)	889 (4.05)
11	713 (5.05)	747 (4.82)	900 (4.00)	900 (4.00)	900 (4.00)
12	709 (5.08)	742 (4.85)	900 (4.00)	900 (4.00)	900 (4.00)
13	705 (5.11)	737 (4.88)	891 (4.04)	900 (4.00)	900 (4.00)
14	693 (5.19)	721 (4.99)	865 (4.16)	900 (4.00)	900 (4.00)
15	681 (5.29)	705 (5.11)	839 (4.29)	867 (4.15)	900 (4.00)
16	668 (5.39)	689 (5.22)	814 (4.42)	887 (4.06)	900 (4.00)

*: The duration of each time interval is 3 minutes

Considering the existing traffic entry flow of 4,160 vph on mainline, the total throughput before the implementation of metering control is 442 veh/3-min, while that under Scenarios A and B are 478 and 482 veh/3-min (Table 4), respectively. Apparently, considering the improvement of throughput, Scenario B (Coordinated Control) outperformed Scenario A

(Individual Control), albeit the delays under both scenarios (625.3 and 687.6 veh-min) are higher than no control (623.9 veh-min).

As shown Figure 4, the throughputs under Scenarios A and B vary over time. With coordinated control (from the 3rd to 9th intervals), the accumulated total throughputs of Scenarios A and B increase by 6.85 % and 8.07%, respectively. Consider the overall performance for the whole simulation period (from the 1st to 16th intervals), the accumulated total throughputs of Scenarios A and B increase by 0.77 % and 1.16%, respectively.

Table 4. Total Throughputs and Total Delay for Multiple Ramp Control

Time Interval	**Entry Flow** (vph)	**Total Throughputs** (veh/3min)			**Total Delay** (veh-min)		
		Before	**Scenario A**	**Scenario B**	**Before**	**Scenario A**	**Scenario B**
1	2960	467	422	432	614.5	629.0	603.7
2	3160	492	452	431	619.7	636.1	641.6
3	3360	437	462	449	560.3	616.5	623.7
4	3560	448	441	475	655.0	652.2	638.2
5	3760	438	472	519	624.0	642.0	610.3
6	3960	422	480	447	527.2	595.0	599.3
7	**4160**	**442 (0%)**	**478 (8.14%)**	**482 (9.05%)**	**623.9**	**625.3**	**687.6**
8	4360	448	468	487	742.5	737.7	715.0
9	4560	450	497	477	836.0	810.5	719.8
10	4760	482	473	494	981.4	853.4	874.4
11	4960	498	467	461	973.0	1000.3	859.2
12	5160	512	505	454	1170.2	1003.3	1049.3
13	5360	509	506	553	1278.9	1093.3	1089.3
14	5560	532	525	516	1338.6	1288.6	1184.9
15	5760	537	513	543	1469.0	1369.0	1364.9
16	5960	542	556	529	1554.1	1627.7	1480.0
Total (1-16)		7656	7715	7745	14568.2	14179.9	13741.3
%		0.00	0.77	1.16	0.00	-2.66	-5.68
Total (3-9)		**3085**	**3296**	**3334**	4568.8	4679.2	4594.0
%		**0.00**	**6.85**	**8.07**	0.00	2.41	0.55
Total (8-15)		**-**	**-**	**-**	**10343.6**	**9783.8**	**9336.9**
%		**-**	**-**	**-**	**0.00**	**-5.41**	**-9.73**

*: The duration of each time Interval is 3 minutes.

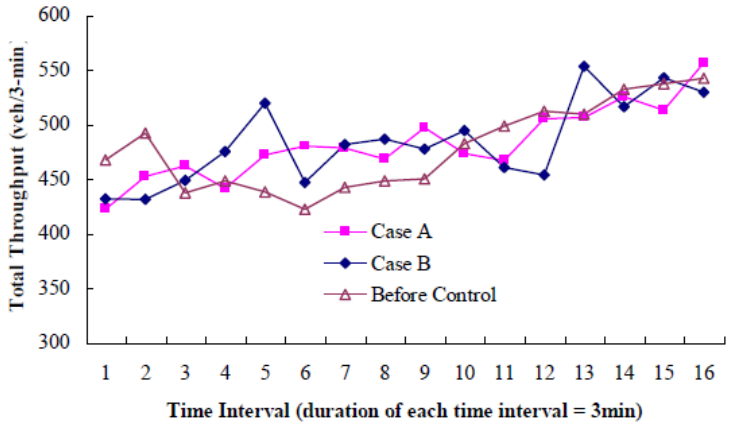

Figure 4. Total Throughput over Time

Figure 5 indicates the relationship between the total delay and time. It was found that without metering control, the accumulated total delay from the 8th to 15th time interval was 10,343.6 vehicle-minutes, while that under Scenarios A and B are 9783.8 and 9336.9 veh-min (Table 4), respectively. Consider the overall performance for the whole simulation period (from the 1st to 16th intervals), the accumulated total delays of Scenarios A and B decreased by 2.66 % and 5.68%, respectively.

In general, the optimal coordinated control outperformed the optimal individual control in increased throughput and reduced delay; and both controls were better than no control.

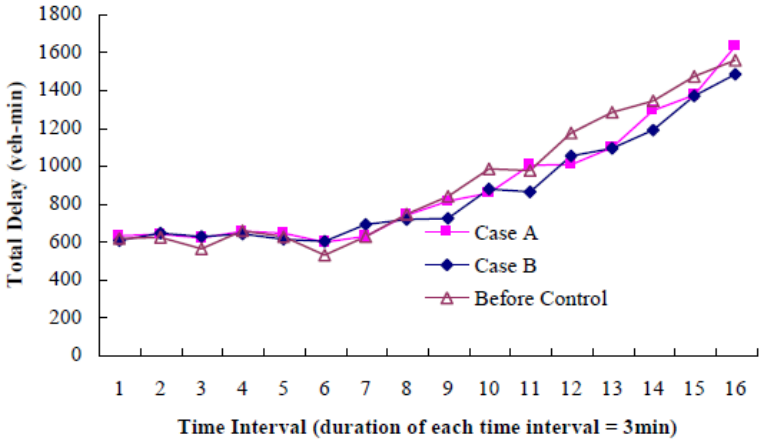

Figure 5. Total Delay over Time

7. CONCLUSION

Unrestricted flows emitted from on-ramps may result in bottlenecks and attendant congestion on freeways specifically during peak periods. When the traffic flow on the mainline exceeds its capacity, congestion and queue are formed. A sound metering control system is intended to mitigate congestion, which has been developed in this study, by increasing throughput with regulated flows entering from ramps. In order to optimize dynamic metering rate, the time-varying relationship among upstream demand, downstream capacity, and entering/exiting volumes on ramps was developed.

The objective function of the coordinated multi-ramp metering control problem is total throughput, which is a constrained discrete-time non-linear optimal control problem. SPSA is successfully applied to search for the optimal solutions (metering rates) in a series time intervals, which maximize total throughput. A simulation approach (with CORSIM) is applied to evaluate the performance of the developed model, which produce emulated real-time information to feed SPSA to search for the optimal solution.

Two scenarios - individual and coordinated multi-ramp metering controls were considered, subjected to actual ramp storage space for vehicle queues. The results of the study are very promising, which demonstrate the efficiency and general applicability of SPSA and the performance achieved by optimal coordinated multi-ramp metering control. The issues illustrated below will be considered as potential extensions of this study:

- Simulation results shows that achieving the maximum total throughput and the minimum total delay can not be achieved simultaneously. The tradeoff analysis between increased throughput and associated delay would be further investigated.
- In the developed model, the objective was to maximize total throughput. Other objective functions, such as total delay, fuel consumption, vehicle emissions, or the combination of them, may be considered.
- With metering control, either individual or coordinated, the behavior of demand due to increasing delay on metered ramps shall be investigated.

8. ACKNOWLEDGEMENT

The authors would like to thank the Guest Editors and the anonymous referees for their valuable comments and constructive suggestions.

REFERENCES

1. Chang, G., Wu, J. and Lieu, H. (1994). A Real-Time Incident-responsive System for Corridor Control: A Modeling Framework and Preliminary Results. Presented at the 73rd Annual Meeting, Transportation Research Board.
2. Chen, C. I., Cruz, J. B. Jr. and Paquet, J. G. (1974). Entrance Ramp Control for Travel-Rate Maximization in Expressways. *Transportation Research*, 8: 503-508.
3. Chen, L. L., May, A. D. and Auslander, D. M. (1990). Freeway Ramp Control Using Fuzzy Set Theory for Inexact Reasoning. *Transportation Research Part A*, 24: 15-25.
4. Chien, S. I. and Luo, J. (2008). Optimization of Dynamic Ramp Metering Control with Simultaneous Perturbation Stochastic Approximation. *International Journal of Control and Intelligent Systems*, 36: 57-63.

5. Chu, L., Liu, H., Recker, W. and Zhang, M. (2004). Performance Evaluation of Adaptive Ramp-metering Algorithm Using Microscopic Traffic Simulation Model, *ASCE Journal of Transportation Engineering*, 130: 330-338.
6. Drew, D. R. (1967). Gap Acceptance Characteristics for Ramp-Freeway surveillance and Control. *Highway Research Board*, 157: 108-143.
7. Goldstein, N. B. and Kumar, K. S. P. (1982). A Decentralized Control Strategy for Freeway Regulation. *Transportation Research Part B*, 16: 279-290.
8. Hadj-Salem, H. and Papageorgiou, M. (1995). Ramp Metering Impact on Urban Corridor Traffic, Field Results. *Transportation Research Part A*, 29: 303-319.
9. Hourdakis, J. and Michalopoulos, P. (2002). Evaluation of Ramp Control Effectiveness in Two Twin Cities Freeways. *Transportation Research Record*, 1811: 21-29.
10. Kleinman, N. L., Hill, S. D. and Ilenda, V. A. (1997). SPSA/ SIMMOD Optimization of Air Traffic Delay Cost. *Proceedings of the ASCE International Conference on Airport Modeling and Simulation*, pp. 45-63.
11. Kotsialos, A. and Papageorgiou, M. (2001). Efficiency Versus Fairness in Network-wide Ramp Metering. *Proceedings of IEEE Intelligent Transportation Systems Conference*, pp. 1190-1195.
12. Masher, D. P., Ross, D. W., Wong, P. J., Tuan, P., Zeidler, H. M. and Petracek, S. (1975). *Guidelines for Design and Operation of Ramp Control Systems*. Stanford Research Institute, Menid Park, CA, USA.
13. McDonnell, J. R., Fogel, D. B., Rindt, C. R., Recker, W. and Fogel, L. J. (1995). Using Evolutionary Programming to Control Metering Rates on Freeway Ramps. In: J. Biethan and V. Nissen (Editors), *Evolutionary Algorithms for Management Applications,* Springer, Berlin, pp. 305-327.
14. Meng, Q and Khoo H. L. (2010). A Pareto-optimization Approach for a Fair Ramp Metering. *Transportation Research Part C*, 18: 489-506.
15. Papageorgiou, M. (1980). A New Approach to Time-of-Day Control Based on a Dynamic Freeway Traffic Model. *Transportation Research Part B*, 14: 349-360.
16. Papageorgiou, M., Hadj-Salem, H. and Blosseville, J. M. (1991). ALINEA: A Local Feedback Control Law for On-Ramp Metering. *Transportation Research Record*, 1320: 58-64.
17. Papageorgiou, M. and Kotsialos, A. (2002). Freeway Ramp Metering: An Overview. *IEEE Transactions on Intelligent Transportation Systems*, 3: 271-281.
18. Papamichail, I., Kotsialos, A., Margonis, I. and Papageorgiou, M. (2010). Coordinated Ramp Metering for Freeway Networks – A Model-predictive Hierarchical Control Approach. *Transportation Research Part C*, 18: 311-331.
19. Papamichail, I., Papageorgiou, M., Vong, V. and Gaffney, J. (2010). Heuristic Ramp-Metering Coordination Strategy Implemented at Monash Freeway, Australia. *Transportation Research Record*, 2178: 10-20.
20. Piotrowicz, G. and Robinson, J. (1995). Ramp Metering Status in North America, *FHWA DOT-T-*95-17.
21. Spall, J. C. (1992). Multivariate Stochastic Approximation Using a Simultaneous Perturbation Gradient Approximation. *IEEE Transactions on Automatic Control*, 37: 332-341.
22. Spall, J. C. (1997). An Overview of the Simultaneous Perturbation Method for Efficient Optimization. *Proceedings of the ASCE International Conference on Airport Modeling and Simulation*, pp. 141-153.
23. Stephanedes, Y. J. and Chang, K. K. (1993). Optimal Control of Freeway Corridors. *ASCE Journal of Transportation Engineering*, 119: 504-515.
24. Taylor, C., Meldrum, D. and Jacobson, L. (1998). Fuzzy Ramp Metering: Design Overview and Simulation Results. *Transportation Research Record*, 1634: 10-18.
25. Taylor, C. and Meldrum, D. (2000). Evaluation of a Fuzzy Logic Ramp Metering Algorithm: A Comparative Study among Three Ramp Metering Algorithms Used in the Greater Seattle Area. Research Report, Research Project Agreement No. T9903, Task 84.
26. Ting, C. J. and Schonfeld, P. (1998). Optimization through Simulation of Waterway Transportation Investments. *Transportation Research Record*, 1620: 11-16.
27. Wattleworth, J. A. (1967). Peak-Period Analysis and Control of a Freeway System. *Highway Research Record*, 157: 1-11.
28. Wiener, R., Pignataro, L. J. and Yagoda, H. N. (1970). A Discrete Markov Renewal Model of a Gap-Acceptance Entrance Ramp Controller for Expressways. *Transportation Research*, 4: 151-161.
29. Yuan, L. S. and Kreer, J. B. (1971). Adjustment of Freeway Ramp Metering Rates to Balance Entrance Ramp Queues. *Transportation Research*, 5: 127-133.
30. Zhang, L. and Levinson, D. (2005). Balancing Efficiency and Equity of Ramp Meters, *ASCE Journal of Transportation Engineering*, 130: 477-481.