



Optimization of mainline traffic via an adaptive co-ordinated ramp-metering control model with dynamic OD estimation

Tang-Hsien Chang *, Zhung-Yih Li

*Transportation Engineering Division, Department of Civil Engineering, National Taiwan University,
P.O. Box 7-876, Taipei 10617, Taiwan, ROC*

Received 31 August 1998; accepted 2 May 2001

Abstract

Frequently implemented at freeway accesses to streamline traffic, ramp-metering control strategy is often implemented during rush hours in heavily congested areas. This paper presents a novel ramp-metering control model capable of optimizing mainline traffic by providing metering rates for accesses within the control segments. Based on Payne's continuum traffic stream model, a linear dynamic model with a quadratic objective function is constructed for integrated-responsive ramp-metering control. Incorporating on-line origin–destination (OD) estimation of co-ordinated interchanges into the proposed model increases efficiency of the control. In addition, an iterative algorithm is proposed to obtain the optimal solution. Simulation results demonstrate the robustness of the proposed model and its ability to streamline freeway traffic while avoiding traffic congestion. © 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Traffic control; Freeway traffic management; Ramp-metering control; Dynamic OD estimation

1. Introduction

Traffic congestion severely limits road users and traffic officials in many industrialized and developing countries. Freeway congestion in Taiwan is no exception. According to the traffic flow theory, high density of vehicles in a particular section corresponds to a low flow speed and also implies a low volume. Therefore, when high traffic demand occurs, vehicles that continuously approach the mainline must be controlled to ensure that the freeway operates at optimal capacity. Previously, police officers manually executed ramp-metering control whereas traffic signals currently perform the same function. Although applied elsewhere, local ramp control cannot achieve

* Corresponding author. Tel.: +886-2-23631004; fax: +886-2-23631004.
E-mail address: thchang@ce.ntu.edu.tw (T.-H. Chang).

Nomenclature

b_{ij}^k	the transition parameter for y_{ij}^k to y_{ij}^{k+1}
k	discrete time step or stage
l	road section
M	the maximum number of time steps required for vehicles to traverse the entire freeway section
N	the number of road sections
n_l	the number of lanes on section l
q	flow
q_{\max}	the maximum flows or capacity
q_l^k	the flows passing through x_l at time k (veh./lane/step)
$q_{*,l}^k$	the ideal or nominal flow (veh./h)
\tilde{q}_l^k	the difference to the ideal flow on section l at time k (veh./lane/step)
$\mathbf{Q}_1, \mathbf{Q}_2$	the weighting factors for the system and control matrices, respectively
r_l^k	the on-ramp flows at ramp l during k (veh./step)
r_0	the ideal or nominal on-ramp flow rate, ramp capacity
R_l^k	Δtr_l^k
$\mathbf{R}_1(k), \mathbf{R}_2(k)$	covariance matrices with respect to $\mathbf{W}(k)$ and $\mathbf{V}(k)$
s_l^k	the off-ramp flows at interchange l during k (veh./step)
S_l^k	Δts_l^k
$\mathbf{S}(k k-1)$	a predicted k -stage vector on account of $(k-1)$ -stage \mathbf{S}
t	time
Δt	the time duration of one step
u	speed
u_l^k	the actual speed on section l at time k (km/h)
$u_{*,l}^k$	the ideal or nominal speed (km/h)
u_*	speed at q_{\max} (km/h)
\tilde{u}_l^k	the difference to nominal speed on section l at time k
$u_e(\rho_l^k)$	the equilibrium speed under the density condition of ρ_l^k
U_l^k	the proposed on-ramp-metering flow rate at ramp x_l at time $k\Delta t \sim (k+1)\Delta t$
U_{*l}^{k+1}	on-ramp capacity of ramp l during stage $k+1$
\tilde{U}_l^k	the difference between the proposed on-ramp-metering flow rate and the on-ramp capacity at ramp x_l at time $k\Delta t \sim (k+1)\Delta t$
v_j^k, w_{ij}^k	random noises, respectively
$\mathbf{V}(k), \mathbf{W}(k)$	random noise vectors, formed from v_j^k, w_{ij}^k , respectively
x_l	the location of the beginning point of section l
y_{ij}^{k-m}	ratio of the trips of OD pair ij in the total trips from ramp i at time $k-m$
$\tilde{\mathbf{Y}}(k)$	the filtered point of variable $\mathbf{Y}(k)$
$\mathbf{Y}(k k-1)$	a predicted k -stage vector on account of $(k-1)$ -stage \mathbf{Y}
ρ	density
ρ_l^k	the actual density on section l at time k (veh./lane/km)
$\rho_{*,l}^k$	the ideal or nominal density on section l at time k

ρ_*	density at q_{\max} (veh./lane/km)
$\tilde{\rho}_l^k$	the difference to nominal density on section l at time k
$\Delta\rho_{s_l}^k$	the noise for the density on section l by the leaving flows at off-ramp l at time k
$\theta_{ij}^{k-m,k}$	fraction of the arrived trips at k , which are generated for the OD pair ij at $k - m$
τ	the driver reaction time
$\mu_{ij}^{k-m,k}$	mean of travel time from i to j during the period $[k - m, k]$, counted in steps
λ	a preset maximal period in a control to optimum
v	the coefficient representing a driver's prediction to the density of next section

a global optimum. Therefore, this study derives a closed-loop feedback control model to enhance freeway efficiency. The proposed model considers the control of multiple ramps with respect to the co-ordinated interchanges for a corridor. A dynamic origin–destination (OD) estimation technique capable of accurately forecasting flow variation is incorporated in the model, which is adaptively operated and elevates global control optimality.

Open-loop control generally cannot achieve robustness and stability as quickly as closed-loop control (Papageorgiou, 1984). In addition, closed-loop control is preferred in terms of responding to the variations of traffic demand or patterns. While describing the relation of transition among traffic states, Payne's (1971) continuum model is applied as the system dynamic model. Based on the flow situation, the control objective attempts to minimize the difference square between actual traffic and its ideal/nominal value, which is equivalent to maximizing the traffic volumes (Gerlough and Huber, 1975). This approach leads to the linear quadratic (LQ) optimal control problem. In addition, understanding the real-time OD distribution of on-ramp trips would allow us to accurately predict traffic conditions in each section, thereby making the control more precise and efficient. However, obtaining the OD distribution is extremely difficult. Fortunately, several OD estimation methods have been developed in recent years (Papageorgiou, 1990; Bell, 1991; Jason, 1991; Chang and Wu, 1994). Among those, the method proposed by Chang and Wu (1994) is modified and used in our metering control study. The output of our metering control model is a set of optimal on-ramp flow rates for the ramps in a controlled corridor.

While performing a ramp-metering control study, Wattleworth and Berry (1965) applied linear programming method to achieve the control optimum. Among recent investigations having modified Wattleworth and Berry's model to achieve the control objective include those of Wattleworth (1967), Wang and May (1973), Eldor and Adler (1977), Papageorgiou (1980) and Iida et al. (1989). Their systems belong to open-loop types. Later, Papageorgiou et al. (1991) proposed a closed-loop control model, ALINEA, which applies the classic automatic control theory to construct a responsive metering model. ALINEA considers the core of a ramp operation with its downstream occupancy. In general, the ramps in an urban area are relatively close to each other. During peak period, co-ordinated metering control is considered to enhance the traffic throughput. Isaksen and Payne (1973) provided a time-invariant LQ model to serve for co-ordinated ramp control with regulators. However, in extending Isaksen and Payne's (1973) model, Goldstein and Kumar (1982) presented a decentralized control strategy. Although Papageorgiou

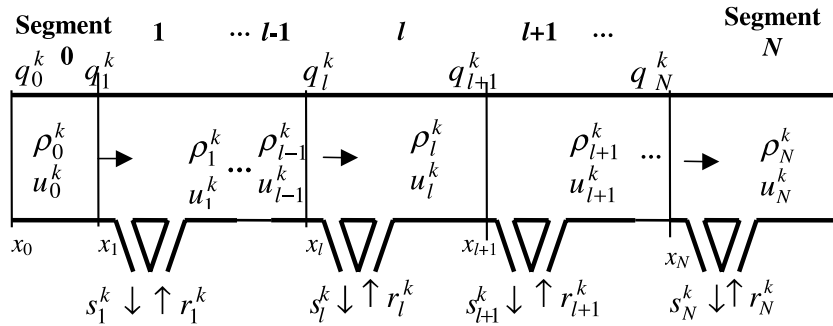


Fig. 1. Freeway layout and relevant variables.

and Mayr (1982) also proposed a nonlinear model for freeway control, the OD ratio is a given constant. While providing a three-layer hierarchical structure for traffic control, Papageorgiou (1984) proposed a dynamic OD concept. However, this model did not describe how to directly achieve dynamic OD estimation. Chang et al. (1993) developed an open-loop DSCOM for ramp-metering control via mathematical programming in which dynamic traffic is estimated through a real-time flow and an OD table. While establishing a linear dynamic flow model for real-time metering control for integrated ramps, Chang et al.'s (1994) subsequent work failed to consider dynamic OD factors. Zhang et al. (1996) achieved nonlinear optimal ramp control but assumed that off-ramp flow is a constant ratio from the mainline. In addition, Zhang and Ritchie (1997) contributed to ramp-metering control via nonlinear artificial neural networks; however, their study only considered a single ramp issue. Mangeas and Haj-Salem (1998) recently proposed a nonlinear closed-loop optimization technique to achieve co-ordinated ramp-metering control. Their work processed traffic assignment on account of the provided splitting rates, and not on a dynamic OD. Therefore, in light of the above developments, this study presents a ramp-metering control model capable of optimizing the mainline traffic by providing metering rates for all accesses within the control area. Particularly, the proposed model adopts the technique of dynamic OD estimation.

Fig. 1 illustrates the relationship among the variables in this study based on a scheme of controlled segments of a freeway.

2. Dynamic control modeling

According to the traffic characteristics, adopting ramp-metering control strategies to prevent supersaturation over the mainline can well-maintain traffic throughputs. In developing a ramp-control model, a system describing overall traffic conditions should be schematized. The system is built at least with several components including a traffic flow model, input and output variables. A ramp traffic control can be designed as a dynamic operation. Hence, the variables and/or the parameters in the system model should be considered in terms of time. While controlling, the input quantity can then be sequentially adjusted with a limited and programmed value in order to make the system output achieve the target. Herein, the programmed value is obtained with a

decision process on the basis of the designed feedback loop from the output measurement. This procedure involves an automatic control field. Ramp-metering control can be dealt with in the automatic control theory. A dynamic model for ramp-metering control is studied as follows.

2.1. Traffic flow model

Traffic control largely focuses on selecting an available traffic flow model. Among the many flow models that have been proposed include static and dynamic ones. In our ramp-metering control, a dynamic type is required so Payne's (1971) high order continuum model is considered herein. This model considers flow volume q , density ρ , speed u , driver reaction time τ , and an anticipated coefficient v . For operation, a discrete type is employed for convenience. Isaksen and Payne (1973) resolved Payne's original model into difference equations as follows:

Concentration equation

$$\rho_l^{k+1} = \rho_l^k + \Delta t \frac{n_{l-1}q_l^{k+1} - n_lq_{l+1}^{k+1}}{n_l(x_{l+1} - x_l)} + \Delta t \frac{r_l^{k+1} - s_l^{k+1}}{n_l(x_{l+1} - x_l)}. \quad (1)$$

Speed equation

$$u_l^{k+1} = u_l^k - \Delta t u_l^k \frac{u_l^k - u_{l-1}^k}{0.5(x_{l+1} - x_{l-1})} - \frac{\Delta t}{\tau} (u_l^k - u_e(\rho_l^k)) - v \frac{\Delta t}{\tau} \frac{1}{\rho_l^k} \frac{\rho_{l+1}^k - \rho_l^k}{0.5(x_{l+2} - x_l)}. \quad (2)$$

Flow equation

$$q_l^{k+1} = \rho_{l-1}^k u_{l-1}^k. \quad (3)$$

2.2. System state space equation

Ramp-metering control is required only when demand exceeds capacity. When demand is far from capacity, on-ramp flows do not need to be controlled. To maintain the freeway's efficiency during rush hours, the control objective generally involves making the vehicle density and flow speed approach the ideal or nominal values that ultimately maximize the throughput and flow of the controlled freeway. Thus, ramp-metering control is triggered when the on-ramp flows reduce the speed of the mainline below the nominal point. According to the traffic theory, during rush hours, maximizing q requires that ρ and u are at the nominal points, ρ_* and u_* , respectively. For operation, short time interval Δt is adopted for revising the control. In doing so, a linear type model is acceptable for investigating the metering rates. The differences between actual traffic and the nominal condition are denoted as follows:

$$\tilde{\rho}_l^k = \rho_l^k - \rho_{*,l}^k, \quad (4)$$

$$\tilde{u}_l^k = u_l^k - u_{*,l}^k, \quad (5)$$

$$\tilde{q}_l^k = q_l^k - q_{*,l}^k. \quad (6)$$

During rush hours for control, $\rho_{*,l}^k \rightarrow \rho_*$ and $u_{*,l}^k \rightarrow u_*$. Although q , ρ , and u are three components of traffic stream, the control state model always selects speed and density to be the system variables since the relationship between speed and concentration tends to be consistent, in which

speed always declines with an increase of concentration (Highway Capacity Manual, 1994). Substituting Eq. (3) into (1) then, taking Taylor’s expansion of Eqs. (1) and (2) at the nominated points described above and neglecting the high order terms lead to the following:

$$\tilde{\rho}_l^{k+1} = a_{l,l-1}\tilde{\rho}_{l-1}^k + a_{l,l}\tilde{\rho}_l^k + a_{l,N+l-1}\tilde{u}_{l-1}^k + a_{l,N+l}\tilde{u}_l^k + d_l\tilde{R}_l^{k+1}, \tag{7}$$

$$\tilde{u}_l^{k+1} = a_{N+l,l}\tilde{\rho}_l^k + a_{N+l,l+1}\tilde{\rho}_{l+1}^k + a_{N+l,N+l-1}\tilde{u}_{l-1}^k + a_{N+l,N+l}\tilde{u}_l^k, \tag{8}$$

where

$$a_{l,l-1} = \Delta t \frac{n_{l-1}u_*}{n_l(x_{l+1} - x_l)}, \quad a_{l,l} = 1 + \Delta t \frac{-u_*}{x_{l+1} - x_l},$$

$$a_{l,N+l-1} = \Delta t \frac{n_{l-1}\rho_*}{n_l(x_{l+1} - x_l)}, \quad a_{l,N+l} = \Delta t \frac{-\rho_*}{x_{l+1} - x_l},$$

$$a_{N+l,l} = \Delta t \left[\frac{1}{\tau} \frac{\partial u_e(\rho_l^k)}{\partial \rho_l^k} \Big|_{\rho_l^k = \rho_*} + \frac{v}{\tau} \frac{1}{0.5(x_{l+2} - x_l)} \frac{1}{\rho_*} \right],$$

$$a_{N+l,l+1} = \Delta t \frac{-v}{\tau} \frac{1}{0.5(x_{l+2} - x_l)} \frac{1}{\rho_*},$$

$$a_{N+l,N+l-1} = \Delta t \frac{u_*}{0.5(x_{l+1} - x_{l-1})}, \quad a_{N+l,N+l} = 1 + \Delta t \left(\frac{-u_*}{0.5(x_{l+1} - x_{l-1})} - \frac{1}{\tau} \right),$$

$$d_l = \frac{1}{n_l(x_{l+1} - x_l)}, \quad \tilde{R}_l^{k+1} = R_l^{k+1} - S_l^{k+1} - U_{*l}^{k+1}$$

$$R_l^{k+1} = \Delta t r_l^{k+1}, \quad S_l^{k+1} = \Delta t s_l^{k+1},$$

where U_{*l}^{k+1} is the on-ramp capacity of ramp l during stage $k + 1$. For control operation, an appropriate control law or a control gain at stage k should be found for the next stage $k + 1$. Thus, by Eq. (7), the following equation is written:

$$\tilde{\rho}_l^{k+1} = a_{l,l-1}\tilde{\rho}_{l-1}^k + a_{l,l}\tilde{\rho}_l^k + a_{l,N+l-1}\tilde{u}_{l-1}^k + a_{l,N+l}\tilde{u}_l^k - \Delta\rho_l^{k+1} + d_l\tilde{U}_l^k, \tag{9}$$

where

$$\tilde{U}_l^k = U_l^{k+1} - U_{*l}^{k+1} \tag{10}$$

and

$$\Delta\rho_l^{k+1} = d_l S_l^{k+1} \tag{11}$$

occur. In addition, \tilde{U}_l^k is considered to be a virtual control variable. Moreover, U_l^{k+1} is a proposed metering rate for stage $k + 1$.

Next, consider the term $\Delta\rho_l^{k+1}$ in Eq. (9). Since S_l^k is historically measured at ramp l at time k , $k = 0, 1, 2, \dots, k$, if S_l^{k+1} can be predicted by a appropriate methodology, the fraction is obtained

$$A_l^k = \frac{S_l^{k+1}}{S_l^k}. \tag{12}$$

According to Eq. (11), we have

$$\Delta\rho_l^{k+1} = d_l S_l^{k+1} = d_l A_l^k S_l^k = A_l^k \Delta\rho_l^k. \tag{13}$$

The next section discusses how to estimate S_l^{k+1} . With ramp-metering control, by incorporating Eqs. (8), (9), and (13), for a N -segmental freeway, $l = 1, 2, \dots, N$, the system state space equation is then expressed as

$$\mathbf{X}(k+1) = \mathbf{A}(k)\mathbf{X}(k) + \mathbf{B}(k)\mathbf{U}(k) + \mathbf{B}'(k) \tag{14}$$

in which

$$\mathbf{X}(k) = [\tilde{\rho}_1^k, \tilde{\rho}_2^k, \dots, \tilde{\rho}_N^k, \tilde{u}_1^k, \tilde{u}_2^k, \dots, \tilde{u}_N^k, \Delta\rho_1^k, \dots, \Delta\rho_N^k]^T,$$

$$\mathbf{X}(k+1) = [\tilde{\rho}_1^{k+1}, \tilde{\rho}_2^{k+1}, \dots, \tilde{\rho}_N^{k+1}, \tilde{u}_1^{k+1}, \tilde{u}_2^{k+1}, \dots, \tilde{u}_N^{k+1}, \Delta\rho_1^{k+1}, \dots, \Delta\rho_N^{k+1}]^T,$$

$$\mathbf{U}(k) = [\tilde{U}_1^k, \tilde{U}_2^k, \dots, \tilde{U}_N^k]^T,$$

$$\mathbf{A}(k) = \begin{bmatrix} \mathbf{A}_{11}(k) & \mathbf{A}_{12}(k) & \mathbf{A}_{13}(k) \\ \mathbf{A}_{21}(k) & \mathbf{A}_{22}(k) & \mathbf{A}_{23}(k) \\ \mathbf{A}_{31}(k) & \mathbf{A}_{32}(k) & \mathbf{A}_{33}(k) \end{bmatrix}, \quad \mathbf{B}(k) = \begin{bmatrix} \mathbf{B}_1(k) \\ \mathbf{B}_2(k) \\ \mathbf{B}_3(k) \end{bmatrix}, \quad \mathbf{B}'(k) = \begin{bmatrix} \mathbf{B}'_1(k) \\ \mathbf{B}'_2(k) \\ \mathbf{B}'_3(k) \end{bmatrix},$$

$$\mathbf{A}_{11}(k) = \begin{bmatrix} a_{1,1} & 0 & 0 & \cdots & 0 & 0 \\ a_{2,1} & a_{2,2} & 0 & \cdots & 0 & 0 \\ 0 & a_{3,2} & a_{3,3} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{N-1,N-1} & 0 \\ 0 & 0 & 0 & \cdots & a_{N,N-1} & a_{N,N} \end{bmatrix}_{N \times N},$$

$$\mathbf{A}_{12}(k) = \begin{bmatrix} a_{1,N+1} & 0 & 0 & \cdots & 0 & 0 \\ a_{2,N+1} & a_{2,N+2} & 0 & \cdots & 0 & 0 \\ 0 & a_{3,N+2} & a_{3,N+3} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{N-1,2N-1} & 0 \\ 0 & 0 & 0 & \cdots & a_{N,2N-1} & a_{N,2N} \end{bmatrix}_{N \times N},$$

$$\mathbf{A}_{13}(k) = \begin{bmatrix} -A_1^k & 0 & 0 & \cdots & 0 & 0 \\ 0 & -A_2^k & 0 & \cdots & 0 & 0 \\ 0 & 0 & -A_3^k & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -A_{N-1}^k & 0 \\ 0 & 0 & 0 & \cdots & 0 & -A_N^k \end{bmatrix}_{N \times N},$$

$$\mathbf{A}_{21}(k) = \begin{bmatrix} a_{N+1,1} & a_{N+1,2} & 0 & \cdots & 0 & 0 \\ 0 & a_{N+2,2} & a_{N+2,3} & \cdots & 0 & 0 \\ 0 & 0 & a_{N+3,3} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{2N-1,N-1} & a_{2N-1,N} \\ 0 & 0 & 0 & \cdots & 0 & a_{2N,N} \end{bmatrix}_{N \times N},$$

$$\mathbf{A}_{22}(k) = \begin{bmatrix} a_{N+1,N+1} & 0 & 0 & \cdots & 0 & 0 \\ a_{N+2,N+1} & a_{N+2,N+2} & 0 & \cdots & 0 & 0 \\ 0 & a_{N+3,N+2} & a_{N+3,N+3} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{2N-1,2N-1} & 0 \\ 0 & 0 & 0 & \cdots & a_{2N,2N-1} & a_{2N,2N} \end{bmatrix}_{N \times N},$$

$$\mathbf{A}_{33}(k) = \begin{bmatrix} A_1^k & 0 & 0 & \cdots & 0 & 0 \\ 0 & A_2^k & 0 & \cdots & 0 & 0 \\ 0 & 0 & A_3^k & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_{N-1}^k & 0 \\ 0 & 0 & 0 & \cdots & 0 & A_N^k \end{bmatrix}_{N \times N} = -\mathbf{A}_{13}(k),$$

$$\mathbf{B}_1(k) = \begin{bmatrix} d_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & d_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & d_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & d_{N-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & d_N \end{bmatrix}_{N \times N},$$

$$\mathbf{B}'_1(k) = \begin{bmatrix} a_{1,0}\tilde{\rho}_0^k + a_{1,N}\tilde{u}_0^k \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{B}'_2(k) = \begin{bmatrix} a_{N+1,N}\tilde{u}_0^k \\ 0 \\ 0 \\ \vdots \\ 0 \\ a_{2N,N+1}\tilde{\rho}_{N+1}^k \end{bmatrix},$$

$$\mathbf{A}_{23}(k) = \mathbf{A}_{31}(k) = \mathbf{A}_{32}(k) = \mathbf{0}_{N \times N}, \quad \mathbf{B}_2(k) = \mathbf{B}_3(k) = \mathbf{B}'_3(k) = \mathbf{0}_{N \times N}.$$

However, Eq. (14) can be transferred into the following standard form:

$$\begin{aligned}
 \mathbf{X}(k+1) &= \mathbf{A}(k)\mathbf{X}(k) + \begin{bmatrix} \mathbf{B}_1(k)\mathbf{U}(k) + \mathbf{B}'_1(k) \\ \mathbf{B}'_2(k) \\ \mathbf{0} \end{bmatrix} \\
 &= \mathbf{A}(k)\mathbf{X}(k) + \begin{bmatrix} \mathbf{B}_1(k)[\mathbf{U}(k) + \mathbf{B}_1^{-1}(k)\mathbf{B}'_1(k)] \\ \mathbf{B}'_2(k) \\ \mathbf{0} \end{bmatrix} \\
 &= \mathbf{A}(k)\mathbf{X}(k) + \begin{bmatrix} \mathbf{B}_1(k) \\ \frac{1}{N}\mathbf{B}'_2(k)\tilde{\mathbf{U}}^{-1}(k) \\ \mathbf{0} \end{bmatrix} \tilde{\mathbf{U}}(k) \\
 &= \mathbf{A}(k)\mathbf{X}(k) + \tilde{\mathbf{B}}(k)\tilde{\mathbf{U}}(k),
 \end{aligned} \tag{15}$$

where

$$\begin{aligned}
 \tilde{\mathbf{U}}(k) &= \left[\tilde{U}_1^k + \frac{a_{1,0}\tilde{\rho}_0^k + a_{1,N}\tilde{u}_0^k}{d_1} \quad \tilde{U}_2^k \quad \cdots \quad \tilde{U}_{N-1}^k \quad \tilde{U}_N^k \right]^T, \\
 \tilde{\mathbf{U}}^{-1}(k) &= \left[\frac{1}{\tilde{U}_1^k + \frac{a_{1,0}\tilde{\rho}_0^k + a_{1,N}\tilde{u}_0^k}{d_1}} \quad \frac{1}{\tilde{U}_2^k} \quad \cdots \quad \frac{1}{\tilde{U}_N^k} \right], \\
 \tilde{\mathbf{B}}(k) &= \begin{bmatrix} \mathbf{B}_1(k) \\ \frac{1}{N}\mathbf{B}'_2(k)\tilde{\mathbf{U}}^{-1}(k) \\ \mathbf{0} \end{bmatrix}.
 \end{aligned}$$

Real on-ramp flow R_l^{k+1} tends towards the provided/allowed flow U_l^k ; in addition, U_l^k must be less than U_{*l}^k during rush hours. Restated, fully on-ramp flows cannot be allowed such that $U_l^k = U_{*l}^k$ during a rush period. Therefore, the inverse of $\tilde{\mathbf{U}}(k)$ exists for $l = 1, 2, \dots, N$. In fact, triggering the ramp-metering control is unnecessary if the output is $U_l^k = U_{*l}^k$.

2.3. Performance function (objective function)

The control is set to maximize the mainline traffic volumes along the freeway in a given incoming period λ . Based on the traffic theory (Gerlough and Huber, 1975), maximizing the traffic volumes at time t is equivalent to minimizing the deviation of the real flow from the ideal flow. Therefore, the performance function is provided as follows:

$$\text{MIN : } J = \sum_{\xi=0}^{\lambda-1} [\tilde{\mathbf{q}}^T(k + \xi)\mathbf{Q}_1\tilde{\mathbf{q}}(k + \xi) + \tilde{\mathbf{U}}^T(k + \xi)\mathbf{Q}_2\tilde{\mathbf{U}}(k + \xi)], \tag{16}$$

where

$$\begin{aligned} \tilde{\mathbf{q}}(k + \xi) &= \mathbf{q}(k + \xi) - \mathbf{q}_*(k + \xi), \\ \mathbf{q}(k + \xi) &= \Delta t [n_1 q_1^{k+\xi} \quad n_2 q_2^{k+\xi} \quad n_3 q_3^{k+\xi} \quad \cdots \quad n_{N-1} q_{N-1}^{k+\xi} \quad n_N q_N^{k+\xi}]^T, \\ \mathbf{q}_*(k + \xi) &= \Delta t [n_1 q_{*,1}^{k+\xi} \quad n_2 q_{*,2}^{k+\xi} \quad n_3 q_{*,3}^{k+\xi} \quad \cdots \quad n_{N-1} q_{*,N-1}^{k+\xi} \quad n_N q_{*,N}^{k+\xi}]^T. \end{aligned}$$

In addition, \mathbf{Q}_1 and \mathbf{Q}_2 are the weighting factors for the system and control matrices, respectively. Linearizing Eq.(3) and incorporating Eqs. (4)–(6) for each $l, l = 1, 2, 3, \dots, N$ lead to

$$\tilde{\mathbf{q}}(k + \xi) = \mathbf{C}\mathbf{X}(k + \xi), \tag{17}$$

where

$$\mathbf{C} = [\mathbf{C}_1 \quad \mathbf{C}_2 \quad \mathbf{C}_3],$$

$$\mathbf{C}_1 = \Delta t u_* \begin{bmatrix} n_1 & 0 & \cdots & 0 & 0 \\ 0 & n_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & n_{N-1} & 0 \\ 0 & 0 & \cdots & 0 & n_N \end{bmatrix}_{N \times N}, \quad \mathbf{C}_2 = \Delta t \rho_* \begin{bmatrix} n_1 & 0 & \cdots & 0 & 0 \\ 0 & n_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & n_{N-1} & 0 \\ 0 & 0 & \cdots & 0 & n_N \end{bmatrix}_{N \times N},$$

\mathbf{C}_3 is an $N \times N$ zero matrix.

Consequently, Eq. (16) can be transferred into

$$\text{MIN} : J = \sum_{\xi=0}^{\lambda-1} [\mathbf{X}^T(k + \xi) \mathbf{Q}'_1 \mathbf{X}(k + \xi) + \tilde{\mathbf{U}}^T(k + \xi) \mathbf{Q}_2 \tilde{\mathbf{U}}(k + \xi)], \tag{18}$$

where $\mathbf{Q}'_1 = \mathbf{C}^T \mathbf{Q}_1 \mathbf{C}$.

3. Solution of the presented model

3.1. Off-ramp flow prediction via a dynamic OD estimation

In recalling Eqs. (12) and (13), completing and solving the system equation at step k initially involves predicting $S_j^{k+1}, j = l$, i.e., the next step off-ramp flow. Many methods are available to forecast S_j^{k+1} such as time series, fuzzy neural networks and Kalman filters. Herein, although the Kalman filter method is applied, a dynamic OD estimation is involved when considering the off-ramp flow prediction herein in order to improve the precision of freeway control. Details of the approach are as follows.

According to Chang and Wu (1994), each off-ramp flow S_j^k can be expressed by all predecessors of upstream on-ramp flows $R_i^{k-m}, m = 0, 1, 2, \dots, M (M < k)$, with the factors of OD ratio y_{ij}^{k-m} and travel time distributed fraction $\theta_{ij}^{k-m,k}$ (a fraction of the vehicles that arrive at ramp j during interval k , and they entered the freeway at ramp i during $k - m$)

$$S_j^k = \sum_{i=0}^{j-1} \sum_{m=0}^M (R_i^{k-m} y_{ij}^{k-m} \theta_{ij}^{k-m,k}). \tag{19}$$

However, for $0 \leq i < j \leq N$, the OD ratios and travel time distributed fractions must satisfy:

$$\sum_{j=i+1}^N y_{ij}^k = 1, \tag{20}$$

$$\sum_{m=0}^M \theta_{ij}^{k,k+m} = 1. \tag{21}$$

Chang and Wu (1994) simplified m such that it would only occur in two stages, t_{ij}^{k+} and t_{ij}^{k-} . They also assumed that $y_{ij}^{k-t_{ij}^{k+}} \approx y_{ij}^{k-t_{ij}^{k-}} \approx \bar{y}_{ij}^k$, implying that

$$S_j^k = \sum_{i=0}^{j-1} \left[\left(R_i^{k-t_{ij}^{k+}} \theta_{ij}^{k-t_{ij}^{k+},k} + R_i^{k-t_{ij}^{k-}} \theta_{ij}^{k-t_{ij}^{k-},k} \right) \bar{y}_{ij}^k \right]. \tag{22}$$

The above simplification and assumption are unrealistic if the travel distance in the controlled freeway section is not short. Herein, m is extended to constitute a probability distribution function. In general, vehicle arrival events are investigated as a Poisson distribution when the traffic is at a free flow condition, as a binomial distribution during congested traffic (Gerlough and Huber, 1975). A Poisson distribution (of course, a binomial distribution can replace it if necessary) is utilized hereinafter for stage m to count the fraction $\theta_{ij}^{k-m,k}$. Thus, Eq. (19) yields

$$S_j^k = \sum_{i=0}^{j-1} \sum_{m=0}^M \left\{ R_i^{k-m} y_{ij}^{k-m} \frac{(\mu_{ij}^{k-m,k})^m \exp(-\mu_{ij}^{k-m,k})}{m!} \right\}, \tag{23}$$

where $\mu_{ij}^{k-m,k}$ denotes the mean of travel time from i to j during the period $[k-m, k]$, which is counted in steps. Eq. (23) can be rewritten in two parts

$$S_j^k = \sum_{i=0}^{j-1} (R_i^k e^{-\mu_{ij}^{k,k}}) y_{ij}^k + \sum_{i=0}^{j-1} \sum_{m=1}^M \left\{ R_i^{k-m} y_{ij}^{k-m} \frac{(\mu_{ij}^{k-m,k})^m \exp(-\mu_{ij}^{k-m,k})}{m!} \right\} \tag{24a}$$

$$= \sum_{i=0}^{j-1} H_{ij}^{k-} y_{ij}^k + \sum_{i=0}^{j-1} H_{ij}^{k+} \tag{24b}$$

in which

$$H_{ij}^{k-} = R_i^k e^{-\mu_{ij}^{k,k}},$$

$$H_{ij}^{k+} = \sum_{m=1}^M \left\{ R_i^{k-m} y_{ij}^{k-m} \frac{(\mu_{ij}^{k-m,k})^m \exp(-\mu_{ij}^{k-m,k})}{m!} \right\}.$$

Assume that each on/off ramp is equipped with flow detectors capable of supplying the real-time data of flows. When at the end of time k , R_i^k and S_j^k as well as R_i^{k-m} are known for $0 \leq i \leq j \leq N$ and $m = 0, 1, 2, \dots, M$. In addition, the travel time of each segment is calculated since the flow q , the speed u , and the density ρ of each segment are measured from time to time. In doing so, $\mu_{ij}^{k-m,k}$ can be summed up from the calculated travel time of all segments between i and j . However, y_{ij}^k is difficult for any equipment to obtain. In addition, y_{ij}^k must be identified by a mathematical calibration method, in which Kalman filter method is utilized herein. If y_{ij}^k is calibrated on-line at the

end of time k , for $m < k$, y_{ij}^{k-m} is known, implying that H_{ij}^{k-} and H_{ij}^{k+} are constants. Thus, Eq. (24a) and (24b) is only a function of y_{ij}^k . Moreover, because y_{ij}^k is possibly varied in a sequential manner assume that y_{ij}^k constitutes a one-step Markovian random walk process for each OD pair such that

$$y_{ij}^{k+1} = b_{ij}^k y_{ij}^k + w_{ij}^k, \tag{25}$$

where w_{ij}^k is an independent event and sequence $\{w_{ij}^k, k = 0, 1, 2, \dots\}$ constitutes a white noise with respect to zero mean, $E(w_{ij}^k) = 0$, and covariance $E(w_{ij}^k w_{ij}^{k'}) = \omega_{ij}^k \delta_{kk'}$ for $0 \leq i < j \leq N$. $\delta_{kk'}$ is the Kronecker delta. b_{ij}^k is an unknown parameter to be determined simultaneously with y_{ij}^k . In a short period at k and for approaching the system stability, the space equation defined in an OD pair between i and j is taken as

$$\begin{bmatrix} y_{ij}^{k+1} \\ b_{ij}^{k+1} \end{bmatrix} = \begin{bmatrix} b_{ij}^k y_{ij}^k \\ b_{ij}^k \end{bmatrix} + \begin{bmatrix} w_{ij}^k \\ 0 \end{bmatrix}. \tag{26}$$

By incorporating Eq. (24a) and (24a), the observation function is set to be

$$S_j^k = \sum_{i=0}^{j-1} H_{ij}^{k-} y_{ij}^k + \sum_{i=0}^{j-1} H_{ij}^{k+} + v_j^k, \tag{27}$$

where v_j^k is also independent and $\{v_j^k, k = 0, 1, 2, \dots\}$ constitutes a white noise with respect to $E(v_j^k) = 0$, and $E(v_j^k v_j^{k'}) = v_j^k \delta_{kk'}$.

For the controlled section with N interchanges of a freeway, based on Eqs. (26) and (17) the entire off-ramp flow prediction space equation is written as

$$\mathbf{Y}(k+1) = \phi(\mathbf{Y}(k), k) + \Gamma(k)\mathbf{W}(k) \tag{28}$$

with the observation function

$$\mathbf{S}(k) = \varphi(\mathbf{Y}(k), k) + \mathbf{V}(k), \tag{29}$$

where

$$\begin{aligned} \mathbf{Y}(k) &= [y_{01}^k \quad y_{02}^k \quad \cdots \quad y_{(N-1)N}^k \quad b_{01}^k \quad b_{02}^k \quad \cdots \quad b_{(N-1)N}^k]^T, \\ \phi(\mathbf{Y}(k), k) &= [b_{01}^k y_{01}^k \quad b_{02}^k y_{02}^k \quad \cdots \quad b_{(N-1)N}^k y_{(N-1)N}^k \quad b_{01}^k \quad b_{02}^k \quad \cdots \quad b_{(N-1)N}^k]^T, \\ \Gamma(k) &= \begin{bmatrix} \mathbf{I}_N \\ \mathbf{0}_N \end{bmatrix}, \end{aligned}$$

where \mathbf{I}_N is an $N \times N$ identity matrix and $\mathbf{0}_N$ is an $N \times N$ zero matrix

$$\begin{aligned} \mathbf{W}(k) &= [w_{01}^k \quad w_{02}^k \quad \cdots \quad w_{(N-1)N}^k]^T, \\ \mathbf{S}(k) &= [S_1^k \quad S_2^k \quad \cdots \quad S_N^k]^T, \\ \varphi(\mathbf{Y}(k), k) &= \begin{bmatrix} \sum_{i=0}^0 H_{i1}^{k-} y_{i1}^k \\ \sum_{i=0}^1 H_{i2}^{k-} y_{i2}^k \\ \vdots \\ \sum_{i=0}^{N-1} H_{iN}^{k-} y_{iN}^k \end{bmatrix} + \begin{bmatrix} \sum_{i=0}^0 H_{i1}^{k+} \\ \sum_{i=0}^1 H_{i2}^{k+} \\ \vdots \\ \sum_{i=0}^{N-1} H_{iN}^{k+} \end{bmatrix}, \\ \mathbf{V}(k) &= [v_1^k \quad v_2^k \quad \cdots \quad v_N^k]^T. \end{aligned}$$

Since Eq. (28) is actually nonlinear, to calibrate $\mathbf{Y}(k)$ the method of extended Kalman filter (EKF) is applied. Consider the linearity of Eq. (28) at the filtered point $\hat{\mathbf{Y}}(k)$

$$\mathbf{Y}(k+1) = \phi(\hat{\mathbf{Y}}(k), k) + \frac{\partial \phi}{\partial \mathbf{Y}} \Big|_{\mathbf{Y}(k)=\hat{\mathbf{Y}}(k)} [\mathbf{Y}(k) - \hat{\mathbf{Y}}(k)] + \Gamma(k)\mathbf{W}(k), \quad (30)$$

which yields to

$$\mathbf{Y}(k+1) = \Phi\hat{\mathbf{Y}}(k)\mathbf{Y}(k) + \Gamma(k)\mathbf{W}(k) + \mathbf{u}(k), \quad (31)$$

where the transition matrix is

$$\Phi(\hat{\mathbf{Y}}(k)) = \begin{bmatrix} \hat{b}_{01}^k & 0 & \cdots & 0 & \hat{y}_{01}^k & 0 & \cdots & 0 \\ 0 & \hat{b}_{02}^k & \cdots & 0 & 0 & \hat{y}_{02}^k & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \hat{b}_{(N-1)N}^k & 0 & 0 & \cdots & \hat{y}_{(N-1)N}^k \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{bmatrix},$$

$$\mathbf{u}(k) = -[\hat{b}_{01}^k \hat{y}_{01}^k \quad \hat{b}_{02}^k \hat{y}_{02}^k \quad \cdots \quad \hat{b}_{(N-1)N}^k \hat{y}_{(N-1)N}^k \quad 0 \quad 0 \quad \cdots \quad 0]^T.$$

In addition, Eq. (29) can be rewritten as

$$\mathbf{S}(k) = \Psi(\hat{\mathbf{Y}}(k))\mathbf{Y}(k) + \mathbf{V}(k) + \mathbf{z}(k), \quad (32)$$

where

$$\Psi(\hat{\mathbf{Y}}(k)) = \begin{bmatrix} H_{01}^{k-} & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ H_{02}^{k-} & H_{12}^{k-} & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & 0 \\ H_{0N}^{k-} & H_{1N}^{k-} & \cdots & H_{(N-1)N}^{k-} & 0 & 0 & \cdots & 0 \end{bmatrix},$$

$$\mathbf{z}(k) = \left[\sum_{i=0}^0 H_{i1}^{k+} \quad \sum_{i=0}^1 H_{i2}^{k+} \quad \cdots \quad \sum_{i=0}^{N-1} H_{iN}^{k+} \right]^T.$$

According to EKF (Saridis, 1994), combining Eqs. (31) and (32) leads to:

$$\hat{\mathbf{Y}}(k|k-1) = \phi(\hat{\mathbf{Y}}(k-1), k-1), \quad (33)$$

$$\mathbf{P}(k|k-1) = \Phi(\hat{\mathbf{Y}}(k|k-1))\mathbf{P}(k-1)\Phi^T(\hat{\mathbf{Y}}(k|k-1)) + \Gamma(k-1)\mathbf{R}_1(k-1)\Gamma^T(k-1), \quad (34)$$

$$\mathbf{K}(k) = \mathbf{P}(k|k-1)\Psi^T(\hat{\mathbf{Y}}(k|k-1))[\Psi(\hat{\mathbf{Y}}(k|k-1))\mathbf{P}(k|k-1)\Psi^T(\hat{\mathbf{Y}}(k|k-1)) + \mathbf{R}_2(k)]^{-1}, \quad (35)$$

$$\hat{\mathbf{Y}}(k) = \hat{\mathbf{Y}}(k|k-1) + \mathbf{K}(k)[\mathbf{S}(k) - \phi(\hat{\mathbf{Y}}(k|k-1), k)], \quad (36)$$

$$\mathbf{P}(k) = [\mathbf{I} - \mathbf{K}(k)\Psi(\hat{\mathbf{Y}}(k|k-1))]\mathbf{P}(k|k-1), \quad (37)$$

where $\mathbf{R}_1(k)$ and $\mathbf{R}_2(k)$ denote the covariance matrices with respect to $\mathbf{W}(k)$ and $\mathbf{V}(k)$, respectively. Given the initial values of $\hat{\mathbf{Y}}(0)$ and $\mathbf{P}(0)$, the variance of $\hat{\mathbf{Y}}(0)$, by Eqs. (33)–(37) sequentially,

$\hat{\mathbf{Y}}(k)$ is iteratively calculated. Obviously, Eq. (33) has prediction capability. Replacing $\mathbf{Y}(k)$ by $\hat{\mathbf{Y}}(k+1|k)$ in Eq. (29) and/or substituting the further step result by Eq. (33) into (29) lead to

$$\hat{\mathbf{S}}(k+1|k) = \varphi(\hat{\mathbf{Y}}(k+1|k), k+1). \quad (38)$$

Eq. (38) is utilized for basis of predicting off-ramp flows.

3.2. Control law and searching procedure

After the off-ramp flow is predicted for each interchange of the controlled segments, the parameters of the system equation (Eq. (15)) and the objective function (Eq. (18)) are determined. According to the optimal control theory (Saridis, 1994), the feedback control law results in

$$\tilde{\mathbf{U}}(k) = -\mathbf{Q}_2'^{-1}(k)\tilde{\mathbf{B}}^T(k)\mathbf{G}(k+1)[\mathbf{I} + \tilde{\mathbf{B}}(k)\mathbf{Q}_2'^{-1}(k)\tilde{\mathbf{B}}^T(k)\mathbf{G}(k+1)]^{-1}\mathbf{A}(k)\mathbf{X}(k), \quad (39)$$

where $\mathbf{G}(\xi)$, $\xi = 1, 2, \dots, f$ (at final), should be satisfied with

$$\mathbf{G}(\xi) = \mathbf{A}^T(\xi)\mathbf{G}(\xi+1)[\mathbf{I} + \tilde{\mathbf{B}}(\xi)\mathbf{Q}_2'^{-1}(\xi)\tilde{\mathbf{B}}^T(\xi)\mathbf{G}(\xi+1)]^{-1}\mathbf{A}(\xi) + \mathbf{Q}_1'(\xi) \quad (40)$$

and in this study, at final target we set

$$\mathbf{G}(f) = \mathbf{Q}_1'(f). \quad (41)$$

Then, first in Eq. (40) setting $\xi+1=f$, and substituting Eq. (41) into Eq. (40), $\mathbf{G}(\xi)$, $\xi = 1, 2, \dots, f$, is calculated. The corresponding $\tilde{\mathbf{U}}(k)$, $k = 0, 1, 2, \dots, f-1$, are also obtained from Eq. (39).

Obviously, the search for the control law is quite elaborate. Fig. 2 depicts the following procedure:

Step 1: At step $k=0$, initiate the number of controlled segments and the length of each segment. Set the mean and variance of $\hat{\mathbf{Y}}(0)$ and $\mathbf{P}(0)$ based on the historical data. Give \mathbf{Q}_1 , \mathbf{Q}_2 , \mathbf{R}_1 and \mathbf{R}_2 .

Step 2: Obtain the current speed and flow on freeway and on on/off ramp of each segment. Calculate travel time on each segment as well as each OD pair.

Step 3: Calculate ϕ , Φ , Ψ and φ . Execute $\mathbf{Y}(k)$ calibration and prediction via Kalman filter method.

Step 4: Execute off-ramp flow prediction via

$$\hat{\mathbf{S}}(k+\xi|k) = \varphi(\hat{\mathbf{Y}}(k+\xi|k), k+\xi) \quad (42)$$

for $\xi = 0, 1, 2, \dots, \lambda$. Eq. (42) is an extension of Eq. (38).

Step 5: Based on Step 4, the predicted $A_l^{k+\xi}$ on account of the data at stage k , i.e., $A_l^{k+\xi|k}$ is calculated for all l .

Step 6: Based on Steps 2 and 5, the predicted matrix $\mathbf{A}(k+\xi|k)$ is obtained.

Step 7: Since the control matrix $\tilde{\mathbf{B}}(k)$ defined in Eq. (15) is expressed in terms of $\tilde{\mathbf{U}}^{-1}(k)$, the inverse of control variable vector $\tilde{\mathbf{U}}(k)$ which will be solved from a given $\tilde{\mathbf{B}}(k)$, $\tilde{\mathbf{U}}(k)$ should be assumed in advance. Then, confirm whether the results are equal to the assumed value. According to the definition of $\tilde{\mathbf{U}}(k)$ in Eqs. (10), (14) and (15), $\tilde{\mathbf{U}}(k)$ is located in the range approximately between zero and ramp capacity. Thus, take a constant $c_j = 0$ for $j = 1, 2, \dots, N$, for the initial condition in the following trial and error loop.

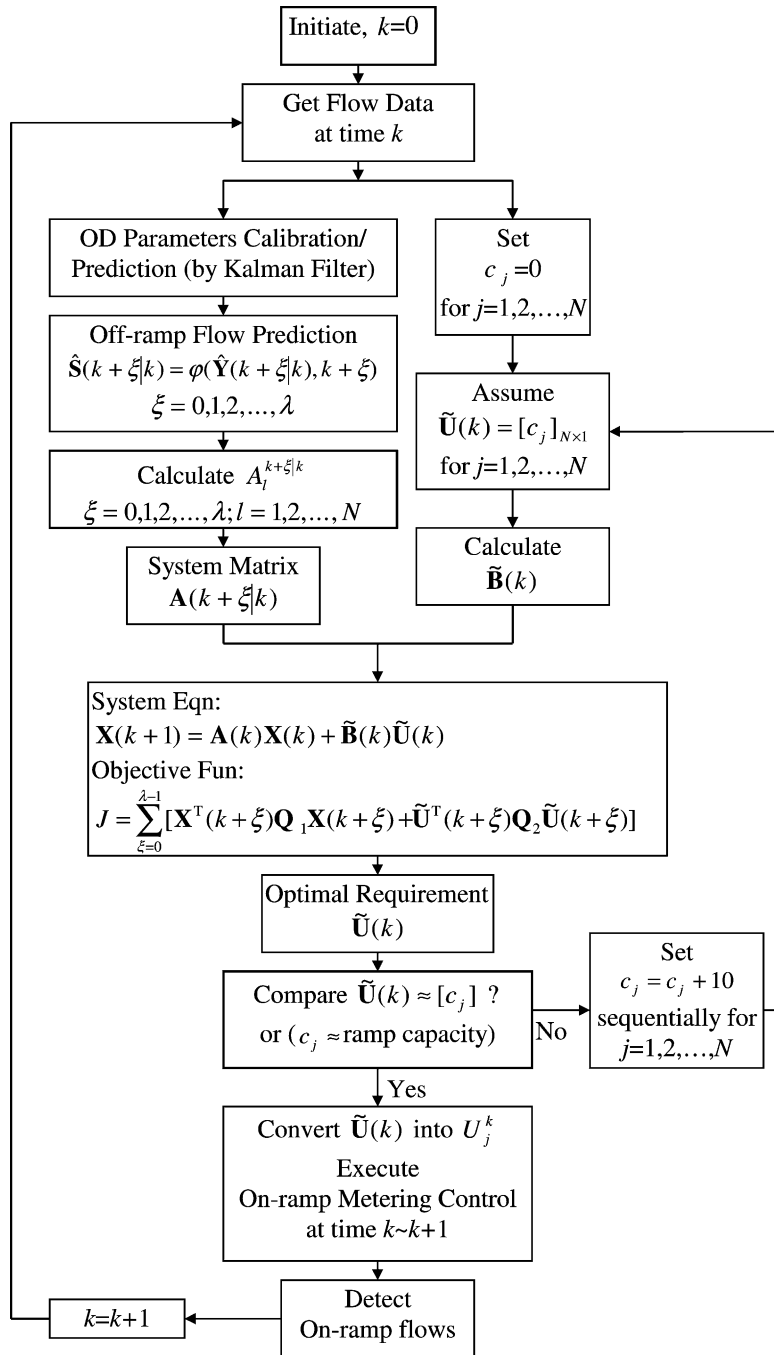


Fig. 2. Control law searching flow-chart.

Step 8: Put $\tilde{\mathbf{U}}(k) = [c_1 \ c_2 \ \cdots \ c_N]^T$, thus beginning the trial and error iteration.

Step 9: Calculate $\tilde{\mathbf{B}}(k)$.

Step 10: Based on Steps 6 and 9, the system equation of Eq. (15) with the objective function is constructed.

Step 11: According to the optimal control theory, a fresh $\tilde{\mathbf{U}}(k)$ is obtained.

Step 12: However, the fresh $\tilde{\mathbf{U}}(k)$ must equal $[c_1 \ c_2 \ \cdots \ c_N]^T$ as assumed in Step 8. If “No”, go to the next step, Step 13; if “Yes” (practically operated by that their mean square error is smaller than an acceptable small constant), go to Step 14.

Step 13: Set $c_j = c_j + 10$ sequentially for an interchange in an iteration, $j = 1, 2, \dots, N$. Then return to Step 8.

Step 14: $\tilde{\mathbf{U}}(k)$ is settled and converted into U_l^k , $l = 1, 2, \dots, N$, for the execution of on-ramp flow control during k to $k + 1$.

Steps 15 and 16: Meanwhile, during the execution, the real on-ramp flows should obviously be detected for next optimal searching. Set $k = k + 1$ and return to Step 2.

4. Verification via a numerical case

For application, the proposed model should be verified and evaluated. Herein, simulation is performed to execute the verification and evaluation procedures. A freeway with five segments is involved for the demonstrative case. To understand the control efficiency and compare demand with real on-ramp flows, the case is associated with the scenario of a varied demand in terms of simulated time t as

$$\begin{aligned} \text{Demand} &= 100 + 15t, & 0 < t \leq 10 \\ &= 400 + 15t, & 10 < t \leq 20 \\ &= 850, & 20 < t \leq 50 \\ &= 350 - 10t, & 50 < t \leq 70. \end{aligned} \tag{43}$$

Obviously, the demand initiates at 100 veh./h and increases with a slope 15 veh./h until 20 min of the simulation. However, the demand has a jump of 150 veh./h at around 10 min. Between 20 and 50 min, the demand is assumed to be a constant 850 veh./h. Beyond 50 min, the demand reduces owing to a sudden drop of 500 veh./h and then decreases with a slope of 10 veh./h along the simulation time. The demand function is uniformly provided for all ramps and, then, the differences of responsive controls of ramps are investigated. Applying such scenario for a metering control is relatively severe to real situations. The running time of the simulation is continuous for 70 min. The traffic is described by Eqs. (1)–(3). According to Haynes (1965), for a three-lane freeway

$$q = 65.5\rho - 0.179\rho^2 - 80 \tag{44}$$

in which the density is denoted in vehicles per mile. According to Eq. (44), the following form is utilized for the term of $u_e(\rho)$ in Eq. (2) and for identifying the nominal points:

$$u_e = q/\rho = 105.39 - 1.39\rho - \frac{26.67}{\rho}, \tag{45}$$

where the unit of the density is converted into vehicles per kilometer per lane. Eq. (45) represents the real world traffic on the simulation and analysis. Based on Eq. (45), the maximal flow is easily verified to be 1976 veh./h, which corresponds to $u_* = 52$ km/h and $\rho_* = 38$ veh./km. Assume that all segments uniformly have three lanes. The scanning interval Δt of the simulation is set for 1 min. $\nu = 19.2$ (km²/h), $\tau = 0.0083$ (h) are provided for completing Eq. (2) (Goldstein and Kumar, 1982). The capacity of each ramp is defined at $r_0 = 1000$ veh./h.

Figs. 3–7 summarize the simulation results. The simulation corresponds to an initial condition of which all mainline segments are unsaturated in traffic at around 1400–1700 veh./h. The flow as well as the speed of uncontrolled segment S0, the upstream, is assumed to remain constant. Figs. 3–5 illustrate the flow, speed and density of five segments of the controlled freeway, respectively. Fig. 6 contains six curves, one demand and five on-ramp flow curves. Fig. 7 displays the queue lengths of each ramp. According to Figs. 3 and 5, owing to that the on-ramp flow in each ramp increases in coincident with the simulated time during the first stage (Fig. 6), the flow as well as the density of each segment of the freeway increases. Meanwhile, the speed decreases (Fig. 4). However, according to Fig. 3, every segment flow gradually tends to approach 1976 veh./h/lane, i.e., the saturation point. In the middle stage, each curve remains flat, never increasing again. The traffic is controlled at the exact saturated condition. While saturated, the speed tends to 52 km/h and density 38 veh./km, the actually nominal values. Because the speed does not fall below the nominal point and ensure the flow at capacity, the control is clearly robust. This observation implies that the proposed model can maintain traffic at a high performance, which also avoids freeway congestion during peak hours.

Fig. 6 obviously reveals the differences between demand and real/allowed on-ramp flow. During the first stage, the allowed on-ramp flow coincides with demand. However, during the middle stage, the on-ramp flows are restricted owing to that the mainline is saturated. The allowed-

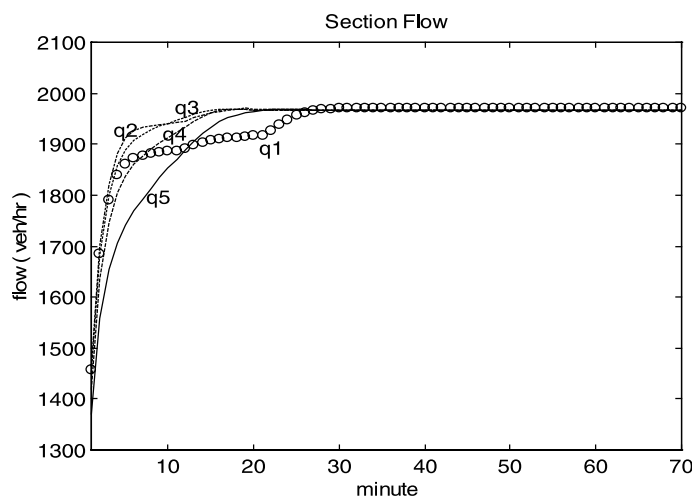


Fig. 3. Mainline flows of the controlled segments.

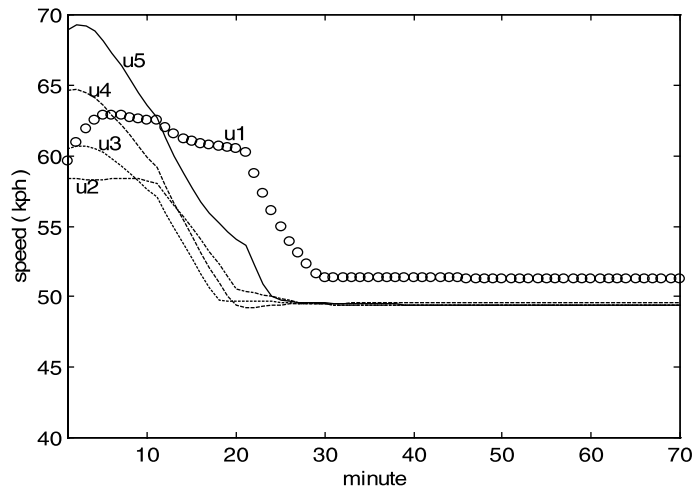


Fig. 4. Mainline speeds of the controlled segments.

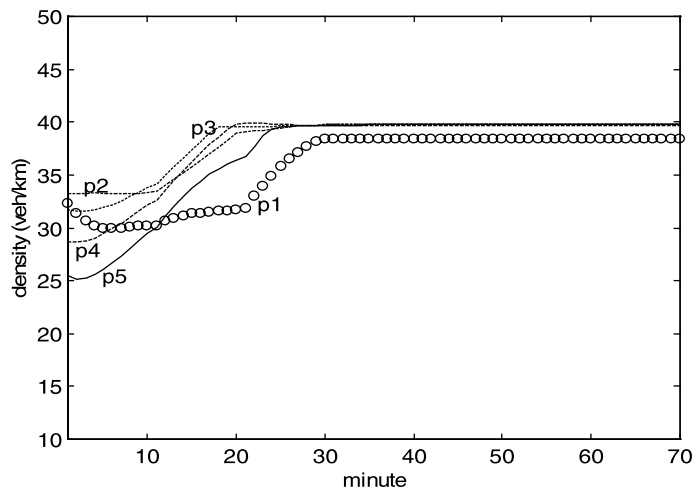


Fig. 5. Mainline density of the controlled segments.

pouring volumes decline. This is owing to that the demand exceeds supply, subsequently causing queuing on ramps. Fig. 7 illustrates the variation of the queue length for each ramp. During the final stage, the allowed on-ramp flows are gradually flat owing to the dispersion of the freeway traffic and the decrease of demand. Nevertheless, the dispersion for each segment has a time lag after the decrease of demand. Evidently, Ramp 1 reveals that its on-ramp flow initially exceeds the demand increment at around 50 min, and its traffic returns to normal approximately at the final stage, in which the allowed quantity matches the demand again, i.e., queue length vanishes (Fig. 7). On Ramp 2, the on-ramp flow is roughly equivalent to the demand increments, but it does not break even between supply and demand by the end of simulation. The queues of Ramps 3, 4, and

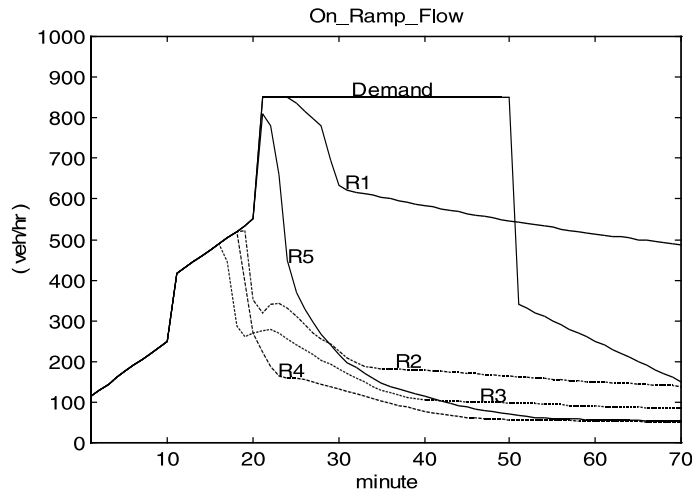


Fig. 6. On-ramp demand and allowed flows.

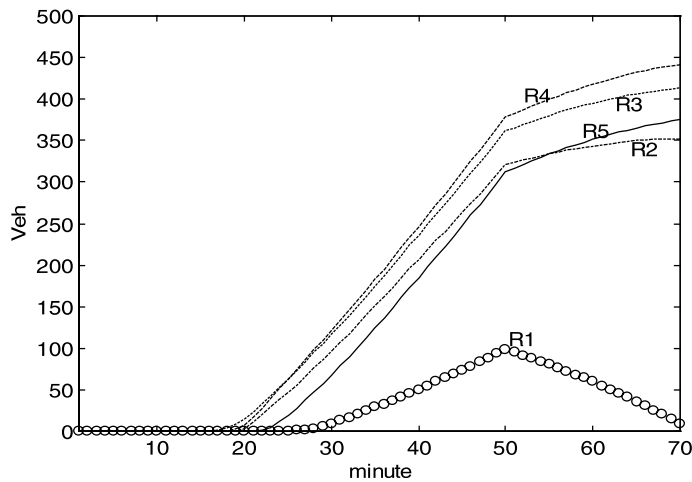


Fig. 7. Queue lengths at on-ramps.

5 are being reduced from 50 min because the curves do not incline again. Until 70 min, queuing still persists. Obviously, these ramps require more recovery time. However, the queue lengths occurred at Ramps 2, 3, 4, and 5 are too large due to the high demand scenario given for the simulation and no queue constraints set for the ramps in the model.

This work also attempts to understand the efficiency of the proposed model by conducting two additional simulations of one for no on-ramp flow control and the other for the control with a fixed OD table. The case of ‘no-control’ is used in order to reveal the traffic problem during rush hours and clarify the role of ramp control. ‘Fixed OD’ case is to manifest the power of the control with ‘dynamic OD estimation’. These two simulations are also handled under the same traffic situation as in the previous case. The fixed OD table is taken from a highly reliable condition, i.e.,

Table 1
Mainline traffic comparison between ‘with control’ and ‘no control’

Segment number	Segment 1 (S1)	Segment 2 (S2)	Segment 3 (S3)	Segment 4 (S4)	Segment 5 (S5)
Length (km)	1.70	1.90	1.90	2.20	2.00
Control with dynamic OD	5870 ^a	5894 ^a	5897 ^a	5841 ^a	5881 ^a
Control with fixed OD	53.56 ^b	50.33 ^b	49.98 ^b	50.25 ^b	50.97 ^b
No Control	5889 ^a	5889 ^a	5897 ^a	5894 ^a	5881 ^a
	53.19 ^b	49.90 ^b	49.88 ^b	50.23 ^b	50.96 ^b
	5860 ^a	3494 ^a	4229 ^a	4749 ^a	5015 ^a
	40.44 ^b	28.50 ^b	29.21 ^b	36.87 ^b	41.39 ^b
Gain (%) ^c	0.2	68.7	39.4	23.0	17.3

^a Flows in veh/h.

^b Mean speed in km/h.

^c Gain = (control flow – no_control)/no_control × 100%.

Table 2
Delay at ramps due to control or no control

Ramp number	Ramp 1 (R1)	Ramp 2 (R2)	Ramp 3 (R3)	Ramp 4 (R4)	Ramp 5 (R5)
Dynamic OD	3.79 ^a	49.54 ^a	77.40 ^a	97.19 ^a	52.48 ^a
Fixed OD	3.20 ^a	70.39 ^a	101.65 ^a	113.98 ^a	59.39 ^a
No control	2.24 ^a	36.69 ^a	25.52 ^a	29.21 ^a	27.57 ^a

^a Average stopping delay in seconds per vehicle.

the mean state of the dynamic OD estimation. Table 1 compares ‘control with dynamic OD estimation’ by the proposed model, ‘control with a fixed OD table’ and ‘no control’ in the mainline traffic flows, which is summarized from the simulation results based on the final single hour data from 10 to 70 min. According to this table, ‘control’ is obviously more effective than without any control policy in the mainline traffic operations. ‘Control’ maintains smooth flows of traffic. ‘No control’ leads to congestion, low throughput, and unstable speed of each segment. Comparing the control with the dynamic OD estimation to which with the fixed OD policy reveals almost no difference at the mainline traffic. This appears to be attributed to that the fixed OD table is selected from the mean of the dynamic estimation. Notably, the difference is significant in average delays during vehicles at on-ramps. Table 2 summarizes this data. The control with the fixed OD policy inefficiently manages all ramps, which lacks flexibility harmonizing overall on-ramp operations. Conversely, the dynamic OD methodology can achieve global efficiency. It reduces some on-ramp flows at Ramp 1 to satisfy the demand at downstream ramps. Synthetically, a reasonable OD table also produces efficient mainline control. Otherwise, imprecise OD data would worsen control. Currently, obtaining a precise OD table is extremely difficult. Ramp-metering control with dynamic OD estimation seems a viable solution. Without unexpected, ‘no control’ lets vehicles save time at on-ramps. But in the point of total travel time spent, ‘control’ is advantageous. Table 3 depicts the average travel time in the simulation period. The indicated time is summed up with the time that the vehicle spent at an origin on-ramp and which spent in its traveling segments to a destination (represented by an interchange). This table reveals that ‘ramp

Table 3
Average travel time in minutes (time spent at ramp and segments)

Origin	Destination			
	Interchange 2 (R2)	Interchange 3 (R3)	Interchange 4 (R4)	Interchange 5 (R5)
Ramp 1 (R1)	1.967 ^a	4.232 ^a	6.513 ^a	9.140 ^a
	1.971 ^b	4.256 ^b	6.541 ^b	9.169 ^b
	2.559 ^c	6.559 ^c	10.462 ^c	14.042 ^c
Ramp 2 (R2)	–	3.091 ^a	5.372 ^a	7.999 ^a
	–	3.458 ^b	5.743 ^b	8.371 ^b
	–	4.612 ^c	8.515 ^c	12.095 ^c
Ramp 3 (R3)	–	–	3.571 ^a	6.198 ^a
	–	–	3.979 ^b	6.607 ^b
	–	–	4.328 ^c	7.908 ^c
Ramp 4 (R4)	–	–	–	4.247 ^a
	–	–	–	4.528 ^b
	–	–	–	4.040 ^c

^a Average time spent while ramp control with dynamic OD.

^b Average time spent while ramp control with fixed OD.

^c Average time spent without ramp control.

control' during rush hours has benefit to travelers, although there is a bit of interference at on-ramp. The longer distance a vehicle travels, the more time the traveler can save.

5. Conclusion

Traffic flow theory stipulates that ramp-metering control is a prerequisite for improving freeway traffic efficiency. While an efficient control model allows for a high performance operation, an inefficient one does not. Owing to available techniques, a dynamic operation is required. Based on Payne's continuum traffic stream model, this work derives a co-ordinated-responsive and closed-loop system model to achieve efficient on-ramp flow control. The proposed model is a linear-quadratic type with on-line OD estimation, capable of improving control efficiency. The Kalman filter method is essential for on-line OD estimation. However, the proposed model is quite elaborate. An iterative algorithm is provided for searching for the optimal results.

Simulation results indicate that the proposed model is robust. Verification and evaluation by an illustrative case reveals the efficiency of the proposed model in terms of freeway access control. With such control, freeway traffic eliminates congestion.

Despite its merits, the proposed model has its limitations. Results presented herein indicate that to implement this model, historical upstream approaching flows, the flows on each segment, on-ramp as well as off-ramp are required during the initial stage. During which, flows should be detected in real-time to produce a new control law for the following control. The law that represented by allowed on-ramp flow rates is generally converted into on/off timings executed by traffic signals. Such an integrated control should be implemented through either a master controller or a co-ordinated traffic control center. Moreover, the constraint of queue lengths on

ramps is absent considering in this study. The presented model should be improved in the future work.

References

- Bell, M.G.H., 1991. The real time estimation of original-destination flows in the presence of platoon dispersion. *Transportation Research B* 25, 115–125.
- Chang, G.L., Ho, P.K., Wei, C.H., 1993. A dynamic system-optimum control model for commuting traffic corridors. *Transportation Research C* 1 (1), 3–22.
- Chang, G.L., Wu, J., 1994. Recursive estimation of time-varying origin–destination flows from traffic counts in freeway corridors. *Transportation Research B* 28 (2), 141–160.
- Chang, G.L., Wu, J., Cohen, S.L., 1994. Integrated real-time ramp metering modeling for nonrecurrent congestion: framework and preliminary results. *Transportation Research Record* 1446, 56–65.
- Eldor, M., Adler, H., 1977. Post-optimality analysis methodology for freeway on-ramp control. *Transportation Research Record* 644, 51–53.
- Gerlough, D.L., Huber, M.J., 1975. *Traffic Flow Theory*. Special Report 165, Transportation Research Board, National Research Council, Washington, DC, 17–21.
- Goldstein, N.B., Kumar, K.S.P., 1982. A decentralized control strategy for freeway regulation. *Transportation Research B* 16 (4), 279–290.
- Haynes, J.J., 1965. Some considerations of vehicular density on urban freeways. *Highway Research Record* 99, 59–80. *Highway Capacity Manual (1994)*, Transportation Research Board, National Research Council, Washington, DC.
- Iida, Y., Hasegawa, T., Asakura, Y., Shao, C.F., 1989. A formulation of on ramp traffic control system with route guidance for urban expressway. In: *Proceedings of IFAC Control Computers, Communications in Transportation*, Paris, France.
- Isaksen, L., Payne, H.J., 1973. Suboptimal control of linear systems by augmentation with application to freeway traffic regulation. *IEEE Transactions on Automatic Control* AC18 (3), 210–219.
- Jason, B.N., 1991. Dynamic traffic assignment for urban road networks. *Transportation Research B* 25, 143–161.
- Mangeas, M., Haj-Salem, H., 1998. Application of nonlinear optimisation technique for coordinated ramp metering control. In: *5th World Congress on ITS*, Seoul, Korea, paper number 2128.
- Papageorgiou, M., 1980. A new approach to time-of-day control based on a dynamic freeway traffic model. *Transportation Research B* 14 (5), 349–360.
- Papageorgiou, M., Mayr, R., 1982. Optimal decomposition methods applied to motorway traffic control. *International Journal of Control* 35 (2), 269–280.
- Papageorgiou, M., 1984. Multilayer control system design applied to freeway traffic. *IEEE Transactions on Automatic Control* AC29 (6), 482–490.
- Papageorgiou, M., 1990. Dynamic modeling assignment, and route guidance in traffic networks. *Transportation Research B* 24, 471–495.
- Papageorgiou, M., Habib, H.S., Blosseville, J.M., 1991. ALINEA: a local feedback control law for on-ramp metering. *Transportation Research Record* 1320, 58–64.
- Payne, H.J., 1971. Models of freeway traffic and control. *Simulation Council Proceedings* 1, 51–61.
- Saridis, G.N., 1994. *Stochastic Processes, Estimation, and Control*. Wiley, NY 10158-0012.
- Wattleworth, J.A., Berry, D.S., 1965. Peak-period analysis and control of a freeway system-some theoretical investigations. *Highway Research Record* 89, 11–25.
- Wattleworth, J.A., 1967. Peak-period analysis and control of a freeway system. *Highway Research Record* 157, 1–21.
- Wang, J.J., May, A.D., 1973. Computer model for optimal freeway on-ramp control. *Transportation Research Record* 469, 16–25.
- Zhang, H., Ritchie, S.G., Recker, W.W., 1996. Some general results on the optimal ramp control problem. *Transportation Research C* 4 (2), 51–69.
- Zhang, H.M., Ritchie, S.G., 1997. Freeway ramp metering artificial neural networks. *Transportation Research C* 5 (5), 273–286.