Robust Optimization of Dynamic Motorway Traffic via Ramp Metering

Andy H. F. Chow, Member, IEEE, and Ying Li

Abstract—This paper presents a robust optimization model for motorway management. The optimization aims to minimize motorway delay via ramp metering with consideration of uncertainties in traffic demand and its characteristics. The robust optimization is formulated as a minimax problem and solved by a two-stage solution procedure. The performances of different control policies are illustrated through working examples with traffic data collected from the M25 motorway in the United Kingdom. Experiments reveal that the robust control provides reliable performance over a range of uncertain scenarios. Results also show that the proposed robust controller is particularly effective during transition periods when congestion has not yet fully developed.

Index Terms—Cell transmission model (CTM), linear program, ramp metering, robust optimization.

I. INTRODUCTION

This paper presents a model-based optimizer for deriving coordinated ramp-metering strategies. Coordinated ramp-metering strategies make use of measurements from a region of motorway to control all metered ramps included therein [1]. A number of coordinated metering approaches have been proposed in the literature. Gomes and Horowitz [2] proposed a deterministic ramp-metering optimizer based on a cell transmission model (CTM) [3] and formulated it as a linear program. Kotsialos and Papageorgiou [4] and Papamichail et al. [1] presented a model-predictive framework for coordinated ramp metering based on a METANET model, which considers capacity drop and transient behavior of traffic congestion. There are also some other rule-based control systems such as ALINEA [5], coordinated control HERO [6], and others (see [7] and [8]). Further details can be referred to [9], which provides a very comprehensive review on the topic of ramp metering. Most of the control systems operate based on a traffic model, and the importance of a reliable model of traffic to effective transport management has been highlighted in [10]. In a traffic model, traffic characteristics are typically represented by a flow–density function, which is known as the fundamental diagram. Fundamental diagrams can be derived by using standard loop detector data, which include measurements of flow (in vehicles per hour), density (in vehicles per kilometer), and speed (in kilometers per hour). As an illustration, Fig. 1 shows a flow–density scatterplot of data collected at a detector station (station 4955A) at the upstream of Junction 15 on the M25 motorway (clockwise) in the United Kingdom. The data were collected over five weekdays from September 24, 2012 (Monday) to September 28, 2012 (Friday) and processed into 5-min averages. It is known that loop detectors can only observe temporal occupancy (i.e., percentage of time that the detector is occupied over each minute interval) but not spatial density, which will be required for calibrating the fundamental diagram. Following [11] and others, we can infer density (ρ, in vehicles per kilometer) from the occupancy measurements over the 1-min time intervals by using the following formula:

\[
\rho = \frac{1000}{L_v + L_d \text{occ}}
\]

where occ is the measured occupancy (in decimal) from the detector, \(L_d\) is the length of the detector (which is taken as 2 m), and \(L_v\) is an estimate of average vehicle length (in meters) passing the location, which is given by the loop-detection system. Fig. 1 reveals that traffic flow, i.e., \(q\), is generally a concave function of density \(\rho\). The maximum value of flow observed, which is known as the capacity flow, is about 7000 veh/h. The critical density \(\rho^*\) associated with the capacity flow is about 80 veh/km. The data with a density value higher than \(\rho^*\) are classified as congested data.

Free-flow data can be easily modeled by simple linear functions and the capacity flow can be regarded as the maximum values of the observed flow [12]. The main challenge of traffic modeling lies on representing the congested data, which exhibit a high level of uncertainty due to the underlying complicated traffic dynamics, including capacity drop, hysteresis, and flow breakdown. Kurzhanskiy and Varaiya [13] proposed modeling the congested data with set-valued functions with which a value
of density can be associated with a range of flow values. This gives rise to an interval estimation of traffic flows in congested situations, and it has been shown that this interval estimation provides important new insight to traffic state estimation [13]. In particular, the critical density \( \rho^c \) is hence associated with a range of flow values from 4000 \( (Q-) \) to 6000 veh/h \((Q+)\), and it implies that the corresponding capacity flow is a random variable lying within that range.

While Kurzhanskiy and Varraya [13] and a number of others focused on traffic estimation with consideration of the variability in congested data, we have not seen many applications of this stochastic concept to optimization. This paper presents a novel robust optimization framework that incorporates the set-valued fundamental diagrams. In addition to the fundamental diagrams, we also reckon that uncertainty exists on the demand side due to various measurement or estimation errors in collecting the demand flow data (see [14] and [15]). The errors can be associated with the health and quality of the sensors. For example, it is not uncommon to find several percent error in flow measurement due to the configuration of detectors and its alignment with traffic flow. Moreover, for real-time control purposes, we will require some short-term (e.g., 5-min) demand flow prediction, which will induce an additional error [16]. Furthermore, one would expect to see a severe underestimation of demand flows should any associated detector(s) break down or malfunction unexpectedly, which is also common in real-life operations. In this paper, control strategies are derived to minimize average travel delays over these uncertain scenarios in both demand and supply. This paper is organized as follows. Section II reviews the CTM upon which our optimization is formulated. Section III presents the optimization framework. With uncertainty in demand and supply considered, we present a robust optimization that aims to minimize travel delay over a range of uncertain scenarios. The robust optimization is formulated as a minimax problem and solved by a two-stage solution procedure. The performances of different control policies are illustrated and compared through working examples in Section IV, with traffic data collected from the M25 motorway in the United Kingdom. Finally, Section V provides some concluding remarks.

II. Modeling Traffic Flow Dynamics

Considering its credibility and desirable mathematical properties, this paper adopts the CTM to represent traffic dynamics in the optimization framework. The CTM is a finite difference approximation proposed by Daganzo [3] for the widely accepted kinematic wave traffic model (see [17] and [18]). Under the cell-transmission formulation, the road section is discretized into a collection of subsections or cells. Each cell \( i \), \( i = 1, 2, \ldots, I \), can be further associated with an external incoming flow \( r_i(t) \) (e.g., an on-ramp) and an external outgoing flow \( s_i(t) \) (e.g., an off-ramp) at each simulation time step \( t \). In the cell-transmission formulation, traffic dynamics are characterized by flow and density in each cell at each time. The evolution of traffic flow and density is governed by the principles of flow conservation and propagation. Define \( f_i(t) \) to be the traffic outflow from cell \( i \) during time step \( t \) and, hence, \( f_{i-1}(t) \) (outflow from upstream cell \( i-1 \)) will be the inflow to cell \( i \) during the same time \( t \). The density in cell \( i \) at the following time step \( t+1 \) can be then updated by conservation as

\[
\rho_i(t+1) = \rho_i(t) + \frac{\Delta t}{\Delta x_i} \left[ f_{i-1}(t) - f_i(t) + r_i(t) - s_i(t) \right]
\]

where \( \Delta t \) and \( \Delta x_i \) are the lengths of the simulation time step and cell \( i \), respectively. The time step size \( \Delta t \) is set such that \( \Delta t \leq \min (\Delta x_i/v_i) \), which refers to the smallest ratio of cell length to the associated free-flow speed along the section. The preceding condition is known as the Courant–Friedrichs–Lewy condition [19], which is used to ensure the numerical stability and nonnegativity of traffic quantities by constraining the traffic to not travel farther than the length of the cell in one simulation time step. Given the cell density, the cell-transmission rule models the outflow from cell \( i \) within time step \( t \) by a piecewise linear fundamental diagram \( \Phi_i = \Phi (v_i, Q_i, \bar{\rho}_i) \) as

\[
f_i(t) = \min \left\{ v_i \rho_i(t), Q_i, Q_{i+1}, w_{i+1} \left[ \bar{\rho}_{i+1} - \rho_{i+1}(t) \right] \right\}
\]

where \( Q_i \) is the capacity flow at cell \( i \), which corresponds to the maximum flow that can leave cell \( i \), and \( Q_{i+1} \) is the capacity flow at cell \( i+1 \), which corresponds here to the maximum flow that can enter cell \( i+1 \). The inclusion of both capacity flows at adjacent cells is due to the consideration of inhomogeneous sections in which we can have different capacities at different locations. When there is no congestion, the traffic moves from one cell to the next at free-flow speed, i.e., \( v_i \). The variable \( \bar{\rho}_{i+1} \) is the jam density at cell \( i+1 \), with which \( w_{i+1} \), which is the backward shockwave speed specified by the fundamental diagram at the downstream cell \( i+1 \), can be derived as \( w_{i+1} = (Q_i/\rho_{i+1} - \rho_{i+1}) \), where \( \rho_{i+1} = (Q_{i+1}/v_{i+1}) \) is the critical density associated with the capacity flow \( Q_i \). The last term in (3) specifies the available space for incoming traffic at the downstream cell \( i+1 \) during time \( t \). The preceding formulation covers both congested and uncongested regimes.

Measurements from loop detectors allow fundamental diagrams to be estimated for the corresponding road sections. For the piecewise linear fundamental diagrams in the CTM, we can divide and calibrate the fundamental diagram in three components: free-flow curve, capacity, and congested curve [12]. The associated critical and jam densities can be derived accordingly. To determine the free-flow curve, we can first extract data in the free-flow state by setting a speed threshold, which is regarded as the 85th percentile of all measured speeds. A data point will be classified as free flow if its associated speed is higher than \( v_{i+1} \). To determine the free-flow curve, we can first extract data in the free-flow state by setting a speed threshold, which is regarded as the 85th percentile of all measured speeds. A data point will be classified as free flow if its associated speed is higher than \( v_{i+1} \).
mentioned before, the critical density $\rho_c^i$ at cell $i$ is determined as $\rho_c^i = (Q_i/v_i)$. A data point is regarded as congested if the associated density is higher than this critical density $\rho_c^i$.

The original CTM adopts single-valued fundamental diagrams, which poses difficulty in modeling the congested data. A number of studies have been done to incorporate the variability shown in the congested traffic data. Sumalee et al. [20] proposed a stochastic CTM, which requires predetermination of various traffic state transitions and associated probabilities, which may not be easy to determine and implement in practice. This paper adopts an alternative approach proposed by Kurzhanskiy and Varaiya [13], in which fundamental diagrams in a congested region are considered to be set valued. Consequently, capacity $Q_i$ of cell $i$ is regarded as a random variable lying within a given range $[Q_i^-, Q_i^+]$. Likewise, the jam density $\bar{\rho}_i$ is also a random variable lying in $[\bar{\rho}_i^-, \bar{\rho}_i^+]$. Given $Q_i$ and $\bar{\rho}_i$, the shockwave speed $w_i$ can be derived accordingly, and it will be also a random variable following $Q_i$ and $\rho_i$. With the ranges of $Q_i$, $\bar{\rho}_i$, and $w_i$ specified, an interval estimate of flow associated with a density value in the congested region can be derived by running iterative CTM simulations with the lower and upper bounds of the parameters [13]. Obviously, this “set-valued” CTM can be reduced to the original CTM with zero uncertainty considered.

III. OPTIMIZATION MODELS

Section III-A starts with reviewing the CTM-based deterministic optimization adopted by [2], [21], and [22]. Section III-B then presents a novel robust optimization with consideration of uncertain demand and set-valued fundamental diagrams.

A. Deterministic Optimization

The CTM-based motorway optimization can be formulated as

$$
\min_c Z = \sum_{i=1}^I \sum_{t=1}^T d_i(t) + \sum_{j=1}^J \sum_{t=1}^T l_j(t) \Delta t
$$

subject to

$$
\begin{align*}
\rho_i(t+1) &= \rho_i(t) + \frac{\Delta t}{x_i} \times [f_i-(t) - f_i(t) + r_i(t) - s_i(t)], \quad \forall i, t \\
f_i(t) &\leq v_i \rho_i(t), \quad \forall i, t \\
f_i(t) &\leq Q_i, \quad \forall i, t \\
f_i(t) &\leq Q_i+1, \quad \forall i, t \\
f_i(t) &\leq w_{i+1} [\bar{\rho}_{i+1} - \rho_i(t)], \quad \forall i, t \\
l_j(t+1) &= l_j(t) + [\lambda_j(t) - r_j(t)] \Delta t, \quad j = 1, 2, \ldots, J, \quad \forall t \\
l_j(t) &\leq l_{j,max}, \quad \forall t.
\end{align*}
$$

The optimization problem seeks the optimal control policy $c$ to be implemented over time $t = 1, 2, \ldots, T$ and cells $i = 1, 2, \ldots, I$, which minimizes the total system delay $Z$ of the motorway. The first term in $Z$ is the total mainline delay, where $d_i(t) = \rho_i(t) \Delta t - f_i(t) \Delta t/v_i$ is the delay in cell $i$ at time $t$. The objective function $Z$ also includes the total on-ramp delay in the second term, where $l_j$ is the queue sizes on all ramps $j$. The constraint set [(5)--(9)] is equivalent to the CTM, as shown by [2], [21], [22], and others. Constraints (6) and (7) can be regarded as the (demand) limitations on flow under a free-flow condition, whereas constraints (8) and (9) can be regarded as the (supply) limitations on flow under congested conditions. The exit flows $s_i(t)$ are assumed to be given. Constraint (10) governs the evolution of queues $l_j(t)$ on boundary links $j = 1, 2, \ldots, J$, where $J$ is the total number of source links (e.g., on-ramps). The variable $\lambda_j(t)$ denotes the traffic demand that wants to enter the system through link $j$ during time interval $t$ and $r_j(t)$ is the actual demand that enters the system. Finally, the last constraint (11) is added to specify the maximum queue length on the ramps such that an unacceptably long queue on the sources will not be obtained as an optimal solution. The control policy $c$ can be realized through metering the on-ramp inflows $r$, which will be the focus of this paper, or through mainline speed control [23].

B. Robust Optimization

The deterministic optimization is extended to a robust formulation that incorporates uncertain demands and set-valued fundamental diagrams. The robust optimization is formulated as a minimax problem [15] in which a likelihood set, i.e., $\Omega$, is defined to specify the region of uncertainty in demand flows and fundamental diagrams. The general idea of the robust optimization here is that, given the predefined likelihood set $\Omega$, we first seek the possible combination of demands and fundamental diagrams that would give the worst system performance (e.g., highest travel delays). Then, given this set of demands and fundamental diagrams, we seek the corresponding optimal control strategy (e.g., ramp metering) that will give us the best system performance (e.g., lowest travel delays) in this worst case scenario. Let $\lambda = \lambda_j(t)$ be the collection of all demand flows from the on-ramps $j$ at time $t$. We also denote $\Phi$ as the set of fundamental diagrams of all cells $i \in I$. Unlike [20], we do not need to specify the probability distributions of these demand flows and fundamental diagrams, whereas we use the likelihood set $\Omega$ to specify the respective possible upper and lower bounds of these variables. The corresponding size and geometry of $\Omega$ can be defined based on on-site measurements and engineers’ judgement. The likelihood set $\Omega$ can be either box constrained or ellipsoidal [24]. A box-constrained $\Omega$ can be derived as follows. We define the intervals

$$
\Omega_{\lambda_j} = [\lambda_j^{\min}, \lambda_j^{\max}]
$$

for all $\lambda_j$ on the on-ramps $j = 1, 2, \ldots, J$, where $\lambda_j^{\min}$ and $\lambda_j^{\max}$ represent the minimum and maximum possible values of $\lambda_j$, respectively, and

$$
\Omega_{\Phi_i} = [\Phi_i^{\min}, \Phi_i^{\max}]
$$

for all fundamental diagrams $\Phi_i$ on cells $i = 1, 2, \ldots, I$, where $\Phi_i^{\min}$ and $\Phi_i^{\max}$ represent the two fundamental diagrams that will give the minimum and maximum possible values of flow, respectively, based on a given value of density. The collection of all intervals shown in (12) and (13) gives a box-constrained likelihood set

$$
\Omega_0 = \Omega_{\lambda_1} \times \Omega_{\lambda_2} \times \cdots \Omega_{\lambda_J} \times \Omega_{\Phi_1} \times \Omega_{\Phi_2} \times \cdots \times \Omega_{\Phi_I}.
$$
Adopting the box-constrained set \( \Omega_b \) (14) in the robust optimization will lead to a worst scenario in which we will have the highest demand flows from all on-ramps and lowest discharge flows everywhere along the mainline at all times. Nevertheless, such a scenario will be too conservative for engineering design purposes [24].

To derive a more practical control strategy, one can use an ellipsoidal likelihood set [24]. We define \( \Sigma = (\lambda, \Phi) \), where \( \Sigma \in \mathbb{R}^{I+J} \), to be the collection of all \( \lambda \) and \( \Phi \). The ellipsoidal likelihood region \( \Omega_e \) is defined as a subset of \( \Sigma \) that satisfies

\[
I + J \sum_{s=1}^{I+J} \left( \frac{\Sigma_{s \text{max}}^2 - \Sigma_{s \text{min}}^2}{2} \right)^{-2} (\Sigma_s - \Sigma_0^0)^2 \leq \theta^2 \tag{15}
\]

where \( \Sigma_{s \text{max}}, s = 1, 2, \ldots, I + J \), is an element in \( \Sigma \), which can refer to \( \lambda_j \) or \( \Phi_i \); \( \Sigma_0^0 \) is the nominal value of \( \Sigma_s \) (which, for example, can be the mean demand flow from an on-ramp \( j \) or the average capacity flow at a location \( i \) on the mainline); \( \Sigma_{s \text{min}} \) and \( \Sigma_{s \text{max}} \) represent the corresponding minimum and maximum possible values of \( \Sigma_s \), respectively. The parameter \( \theta \in [0, 1] \) reflects the degree of uncertainty taken into account in optimization. This parameter \( \theta \) can be regarded as a tradeoff between efficiency and robustness. The larger \( \theta \) is, the more preference is given to consideration of robustness. On the other hand, \( \theta = 0 \) reduces the robust optimization into conventional deterministic optimization in which no uncertainty is considered.

Given \( \Omega \), the robust optimization can be now formulated as a minimax problem, i.e.,

\[
\min_{c} \max_{\lambda, \Phi} Z(c, \lambda, \Phi) \tag{16}
\]

subject to

\[
\lambda, \Phi \in \Omega. \tag{17}
\]

This optimization, of course, is also subject to traffic dynamics constraints (5)–(10), as specified by the CTM, in addition to constraint (17). There is no straightforward solution algorithm for solving the preceding optimization problem. It is conventional to decompose and solve this minimax problem in two stages.

1) Given a set of demands \( \lambda \) and fundamental diagrams \( \Phi \), we seek a control policy \( c \) to minimize the total delay \( Z \).

2) Given the control policy \( c \) determined in Step 1, we seek a set of demands \( \lambda \) and fundamental diagrams \( \Phi \) within the likelihood set \( \Omega \) such that the total delay \( Z \) is maximized.

Step 1 is essentially the same optimization problem as in Section III-A. Step 2 involves maximization of the linear total delay function \( Z \) over a constraint set \( \Omega \). It can be easily verified that both box-constrained and ellipsoidal \( \Omega \) likelihood sets are convex in \( \lambda \) and \( \Phi \). The likelihood set constraint can hence be augmented into the objective function through a Lagrangian multiplier. Consequently, the maximization problem in Step 2 becomes a convex (indeed quadratic) optimization subject to linear constraints, which can be solved by standard gradient search algorithms. The two-step procedure above will be run iteratively, and the process will converge to an optimal control policy \( c^* \), which minimizes delay \( Z \) under the worst scenario realized in \( \Omega \) (see [5] and [25]).

IV. NUMERICAL EXAMPLES

We select a 10-km section of the orbital M25 motorway (clockwise) between Junctions 13 and 15. The section covers three on-ramps (one at Junction 14 and two at Junction 15) and two off-ramps. The road section is one of the busiest sections in England. The motorway stretch contains 20 detector stations with a spacing of 500 m. We focus on the duration between 17:00 and 20:00, which are the busiest hours. The road section is discretized into 20 cells, which are configured such that the upstream and downstream boundaries of each cell will be coincided with the location of the associated detector station. The on-ramps are located at Cells 10 (Junction 14), 17 (Junction 15a), and 18 (Junction 15b); the off-ramps are located at Cells 5 (Junction 14) and 15 (Junction 15). Each cell is characterized by a piecewise linear fundamental diagram, which is calibrated by the flow and density data from the associated detectors. A CTM is calibrated with successive linear regressions for the fundamental diagrams [12]. The calibrated CTM gives root-mean-square errors of 9.8% and 12.4%, respectively, when we compare the estimated flows and densities against the measured values.

A. Deterministic Optimization

The deterministic optimization seeks an optimal ramp metering strategy that minimizes the total delay in which fundamental diagrams are considered to be single valued and there is no uncertainty in the measured demand. The size of the simulation time step, i.e., \( \Delta t \), is set to be 15 s, which gives the total number of time steps \( T = 720 \) for the 3-h (17:00–20:00) planning horizon. The length \( \Delta x \) of each cell is 500 m, which is the same as the detector spacing. The maximum allowable queue length \( I_{\text{max}} \) on all ramps is 60 veh. The optimization is solved by IBM ILOG CPLEX Optimization Studio V12.5 running on a desktop computer with Intel Core i5-2400 3.1-GHz processor, 4-GB random access memory, and Windows 7 64-bit operating system. The optimization problem takes a computational time of 4 min to be solved. The optimization reduces the mainline delay from the original 2145 veh-hr (with zero ramp delays) to 1776 veh-hr with the ramp queues bounded below a reasonable \( I_{\text{max}} = 60 \) veh. For the readers’ further information, the associated ramps’ delay is 332.6 veh-hr with metering, which gives a total system delay (i.e., mainline + ramps) of 2109 veh-hr, which is smaller than the original 2145 veh-hr. Fig. 2 compares the density contours under the no-control and metered scenarios, whereas the ramp queues are not allowed to exceed 60 veh. The layout of the road section, including the locations of the ramps, is shown on the left of the plots. The lighter color in the plot below implies a reduction in traffic density under metering. We also highlight the reduction of a congestion spillover around the off-ramp (Cell 5) at Junction 14 (marked by the rectangle in the figure). This indicates an overall system-wide benefit from ramp metering.

B. Robust Optimization

Now, we present the results of robust optimization in which we consider whether there is an uncertainty of \( \pm 5\% \) associated...
with the demand flows measured at each ramp \( j \). This gives \( \lambda_{j}^{\text{min}}(t) = 0.95\hat{\lambda}_{j}(t) \) and \( \lambda_{j}^{\text{max}}(t) = 1.05\hat{\lambda}_{j}(t) \) for all \( \lambda_{j} \) and times \( t \). The notation \( \hat{\lambda}_{j}(t) \) is the measured value of \( \lambda_{j}(t) \) at time \( t \) as reported from the detectors. The fundamental diagrams \( \Phi \) are specified by the free-flow speed \( v \), capacity \( Q \), and jam density \( \rho \), with which other parameters such as shockwave speed \( w \) and critical density \( \rho^{*} \) can be determined accordingly. As suggested by empirical observations, the free-flow speed \( v \) is considered to be deterministic. To incorporate the set valueness \( \epsilon \) of the fundamental diagrams, we assume that capacity \( Q \) and \( \rho \) are associated with a \( \pm 5\% \) estimation error at all locations \( i \) along the mainline. Hence, we have \( Q_{i}^{\text{min}} = 0.95Q_{i} \) and \( Q_{i}^{\text{max}} = 1.05Q_{i} \); \( \hat{\rho}_{i}^{\text{min}} = 0.95\hat{\rho}_{i} \) and \( \hat{\rho}_{i}^{\text{max}} = 1.05\hat{\rho}_{i} \) for all cells \( i \). The notations \( Q_{i} \) and \( \hat{\rho}_{i} \) denote the calibrated capacity and the jam density at cell \( i \) with the detector data, respectively.

Given the bounds on \( \lambda, Q, \) and \( \rho \), we construct an ellipsoidal likelihood set \( \Omega_{\epsilon} \) for the robust optimization problem following (15), where the design parameter \( \theta \) is taken as one. The robust optimization problem can be reduced to the deterministic one presented in Section IV-A, with \( \theta \) set to be zero. The robust ramp metering is derived by solving the minimax problem with the two-stage algorithm as presented in Section III-B. To measure the convergence of the two-stage solution algorithm, we define \( \epsilon = (\|c_{n} - c_{n-1}\|/\|c_{n-1}\|) \), where \( \| \cdot \| \) is the Euclidean norm and \( c \) is the vector containing all control variables (i.e., the on-ramp flows) computed at iteration \( n \) in the two-stage algorithm. The two-stage algorithm reaches a steady state (i.e., converges) when \( \epsilon = 0 \). It is observed that our two-stage solution procedure reaches an \( \epsilon = 0.003 \) after four iterations.

![Mainline density contour plots. (Top) No control. (Bottom) Metered.](image1)

**Fig. 2.** Mainline density contour plots. (Top) No control. (Bottom) Metered.

The performances of different control strategies are compared over 11 levels of nominal demand values \( \lambda \), as shown in Fig. 3. On the \( x \)-axis, “100%” is the situation where we are simulating with all \( \lambda \) as measured by detectors, “95%” refers to a situation in which all \( \lambda \) are scaled down by 5%, “105%” is the situation where \( \lambda \) are all scaled up by 5%, and so on. A demand multiplier less than 100% represents the situation in which the actual demand is being overestimated (i.e., the actual demand is lower than the design demand) and vice versa. Under each scenario, we randomly generate 300 combinations of corresponding demands and fundamental diagrams within the region “\( \Omega_{\epsilon} \)”.

The ramp metering strategies are then simulated over these 11 \( \times \) 300 = 3300 scenarios and the results, including total delays and ramp delays, are summarized in Fig. 3. In the figure, the middle line in each box represents the median delays and the heights of the box are the interquartile range. The experiment also includes the well-established ALINEA control policy [5] as a benchmark for comparison. The ALINEA strategy operates as an integral controller (I-controller), which regulates the ramp flow \( r_{j}(t) \) according to the measurement or the estimate of mainline density \( \rho_{i}(t) \) in the neighborhood of ramp \( j \) as

\[
r_{j}(t) = r_{j}(t - \Delta c) + K (\hat{\rho} - \rho_{i}(t))
\]

(18)
where $\Delta c$ is the control period governing the frequency of updating the ramp-metering strategy; $\hat{\rho}$ is a target value of the mainline density, which is typically set as the associated critical density value on the mainline; and $K$ is the control gain, and we set it to be 100 km / lane/h following [6]. It is noted that previous studies show that ALINEA is robust where the control gain $K$ does not need to be fine-tuned and is applicable to a wide range of scenarios with the same value (see [26] and [27]). Fig. 3 reveals that the robust control generally outperforms the deterministic one in particular when the demand is appropriately estimated (i.e., when demand multiplier is 100% or 105%). Moreover, the ramp delays under robust control are lower than those under deterministic control, because an extra buffer on ramps is introduced when the robust ramp metering policy is calculated and, hence, ramp delays grow less rapidly than in the traditional deterministic control with the extra buffer. This reveals the advantage of robust control for incorporating potential uncertainties in the overall system and protecting ramps in motorway operations. It is also found that ALINEA is generally outperformed by the deterministic and robust controller, because ALINEA is not designed as a system optimizer but a stabilizer trying to maintain the mainline density at the target value $\hat{\rho}$ and hence maximize the mainline throughput [27]. Nevertheless, this could be done at the expense of ramps and hence may imply a higher overall system delay than the other two strategies, as observed in our experiment. However, we also observe that the discrepancy between ALINEA and the optimal control strategies would reduce when the actual demand gets significantly higher than the expected ones, for example, when demand multipliers are greater than 102%. We believe that this finding is sensible as this reveals the loss of efficiency of the optimal control when the model estimation errors become too large. Moreover, this also reveals the robustness of ALINEA even with its simple structure. In addition to the three strategies discussed above, we also compute a “deterministic” strategy by running a deterministic optimization with all demand flows multiplied by 1.05 and all capacities and jam densities multiplied by 0.95. This indeed represents the traditional “safety factor” based engineering design approach [28] in which we simply scale up or down relevant design parameters without running the robust optimization. Note that here we need to disable the ramp queue constraint when calculating this deterministic scheme as the optimization will become infeasible with increased demands and capacities due to insufficient space in the system for storing all the queues. The deterministic strategy is shown to be very robust as it shows little variations in terms of the box sizes due to the huge buffer created from optimization. However, this deterministic strategy will be too conservative and counterproductive. As shown, the deterministic strategy performs significantly worse than all the other strategies by shutting down ramps unnecessarily long.

Finally, we also compare the performances of deterministic and robust control under different time periods: 16:30–17:00 and 17:30–18:00. The period 16:30–17:00 is when congestion starts to build up where the road network is relatively empty, whereas the period 17:30–18:00 is the midst of the peak when the road network is jammed with traffic. It is found that the robust control policy is more effective during the period 16:30–17:00 in terms of reducing delay variance. When the congestion is just starting to develop. During 16:30–17:00, the medians of delay associated with deterministic and robust control are 16.3 and 19.1 veh-hr, respectively; the interquartile ranges of delays are [10.2, 46.2] (deterministic) and [16.5, 30.8] (robust), respectively. During 17:30–18:00, the medians of delay associated with deterministic and robust control are 109.1 and 108.9 veh-hr, respectively; the interquartile ranges of delays are [89.4, 150.1] (deterministic) and [90.1, 138.6] (robust), respectively. We believe that this is because, during peak period (17:30–18:00), the consideration of extra buffer does not help significantly as the road networks are already running close to their physical capacity. This observation suggests that robust control is more useful in managing transient traffic before congestion fully develops.

V. CONCLUSION

This paper has presented a novel robust optimization model for designing motorway ramp metering strategies with consideration of demand and supply uncertainties. The model is formulated as a minimax problem and solved with a two-stage solution algorithm. Given a specific set of demands and fundamental diagrams, we first seek an optimal ramp-metering strategy that minimizes the total system delay. The problem is formulated as a linear program based on CTM. Given the ramp-metering strategy, the next stage seeks a set of demands and fundamental diagrams within a likelihood region ($\Omega$) that maximizes the system delay. The specification of $\Omega$ represents a tradeoff between efficiency and robustness. A small $\Omega$ represents a situation in which the demands and fundamental diagrams are well estimated with little uncertainties involved. With this small $\Omega$, one will get a ramp-metering strategy that will result in a low expected delay. Nevertheless, such a ramp-metering strategy can be vulnerable to unexpected perturbations in demand and fundamental diagrams, as shown in our experiments. On the other hand, if one adopts a large $\Omega$, the corresponding ramp-metering strategy can be very robust while it can be highly ineffective. Following [24], we adopt an ellipsoidal $\Omega$, which is effective in balancing the efficiency and robustness. We compare four different ramp-metering strategies: deterministic, robust, a “safety-factor” based deterministic design (deterministic $^*$), and the rule-based ALINEA strategies. The results show that robust control is effective in managing the development of both mainline and ramp delays with consideration of potential uncertainties encountered over a range of scenarios. The results also show that the robust control is more effective during transition periods than periods when congestion has already fully developed. We have to admit that the scale of the case study adopted herein is rather limited. Further studies include testing the proposed methodology on larger scale such as the entire M25 motorway or national-wide motorway networks. It will be also interesting to implement the proposed robust control algorithm on a microsimulation test bed to mimic the effect of real-world applications. We are also extending the proposed optimization framework to a dynamic rolling horizon platform for real-time operations. Finally, we do not consider the route choice behavior of travelers, and we leave this dynamic traffic assignment component for future study.
ACKNOWLEDGMENT

The authors would like to thank the editor and all anonymous referees for their constructive comments, the U.K. Highways Agency for providing the loop detector data through Motorway Incident Detection and Automatic Signalling, and the IBM Academic Initiative for providing the software and license for using the CPLEX Optimization Studio.

REFERENCES


Andy H. F. Chow (M’09) received the B.Eng. and M.Phil. degrees in civil engineering from The Hong Kong University of Science and Technology, Hong Kong, in 2000 and 2002, respectively, and the Ph.D. degree in transport studies from University College London (UCL), London, U.K., in 2007. He received a Gordon Newell Memorial Dissertation Prize in 2008 for his doctoral dissertation on optimal control of dynamic transport networks. He is currently a Lecturer (Assistant Professor) in transport studies with UCL. Before joining UCL, he was a Postdoctoral Researcher with the University of California Berkeley under the California PATH program during 2007–2010. His research interests are in developing tools for analyzing and managing transport systems. Mr. Chow received a Gordon Newell Memorial Dissertation Prize in 2008 for his doctoral dissertation on optimal control of dynamic transport networks completed at UCL.

Ying Li received the B.Eng. degree in automation from Chang’an University, Xi’an, China, in 2010 and the M.Sc. degree in transport studies from Imperial College London, London, U.K., in 2011. She is a Doctoral Research Student in transport studies with University College London, London. Her research is on robust traffic modeling and optimization.