Examining the Potential of Using Ramp Metering as a Component of an ATMS

BRUCE HELLINGA AND MICHEL VAN AERDE

The current emphasis on utilizing existing transportation infrastructure more efficiently has added impetus to the recent focus on advanced traffic management systems. An advanced traffic management system typically combines existing hardware, software, and traffic engineering expertise to observe and manage transportation systems more effectively. The potential of ramp metering to provide reductions in system delay has been recognized for some time. Simple analytical techniques have been used to demonstrate the magnitude of these benefits. However, these analytical methods can rarely reflect fully all the spatial and temporal dynamics that may exist within integrated freeway and arterial networks. In this paper a network traffic simulation model is used to examine the potential benefits of implementing ramp metering strategies and to quantify how sensitive these benefits are to a number of factors including the metering rate, the timing of the implementation of the metering, and various assumptions regarding driver rerouting behavior. The results of this investigation indicate that, as expected, ramp metering can result in reductions in total travel time, but it may also yield increased net delays if it is not implemented correctly. This investigation indicated that the temporal window of opportunity during which ramp metering can be implemented and be of benefit is surprisingly small. Results for a simple network indicate that under ideal conditions, in which drivers are able to divert their routes, a benefit of as much as a 14 percent reduction in total travel time may be possible. If it is assumed that a capacity loss of 5 percent occurs once the freeway becomes congested, then the benefit of metering may be as great as a 26 percent reduction in total travel time.

With the recent emphasis on advanced traffic management systems, ramp metering is receiving considerable attention as a traffic management strategy. This consideration is not new. Ramp metering has been considered for more than four decades as a traffic management technique. During this time, many forms of ramp metering have been examined, including simple fixed time metering (1) and real time responsive metering (2). Much effort has also been expended on researching methods of optimizing the metering rates of isolated meters (3) and systems of coordinated meters (4).

The earliest works utilized linear programming techniques to determine optimal time-of-day metering rates (5). Most of these optimization strategies assume freeway throughput as the objective function. However, few of the evaluation methods explicitly consider the possibility that route diversion may take place. One recent notable exception is the work carried out by Nsour et al., (6), who utilized the INTRAS simulation model to examine the impacts of ramp metering with and without diversion. Based on their simulation of an 11.2 km section of freeway in California, Nsour et al. concluded that a 10.5 percent reduction in system delay could be obtained under ideal metering and diversion conditions. However, as diversion rates were prespecified, drivers of vehicles did not have the ability to make routing decisions based on currently available estimates of alternative route travel times. Little research has been conducted to identify and quantify the net benefits of ramp metering when realistic route diversion is considered. More importantly, the sensitivity of these benefits to various control parameters has typically not been examined.

PURPOSE OF RAMP METERING

In most ramp metering analyses the intended purpose of utilizing ramp metering is the avoidance of flow breakdown on the freeway. To meet this goal, the capacity of each freeway segment is determined, demands are estimated, and metering rates are imposed such that the freeway operates without congestion. This process is rather straightforward and is described elsewhere in the literature (3). However, the process of quantifying the net benefits of ramp metering is more difficult. Reduced delay is often considered to be the primary benefit of ramp control; nevertheless, impacts on fuel consumption, emissions, and safety may also be components of the net benefit.

POTENTIAL BENEFITS OF RAMP CONTROL

To achieve the goal of ramp control, that of reducing total network travel time, it is necessary to identify, and then seek to satisfy, a number of specific objectives. One such objective might be the reduction of the size of queues on the freeway by controlling ramp access. Another objective might be the improvement of average freeway speed by ensuring that the freeway operates in an uncongested mode. One might even desire a freeway speed that is higher than the speed at capacity, thereby requiring a more restrictive metering rate.

For safety considerations, a potential objective might be the reduction in the variability of vehicle speeds. For delay and throughput considerations, a potential objective might be the avoidance of freeway queues spilling upstream and blocking access to some heavily utilized exit ramps. Ramp control can also be used to avoid capacity reduction effects that occur when flow breaks down or to encourage spatial, temporal, and modal diversions to other roads, times, and modes having lower marginal system costs.

Each of these potential ramp metering objectives may have a different impact on net system benefits. An evaluation of the impact on benefits of many of these potential objectives is usually too complex to be carried out adequately by standard analytical techniques. In this paper we examine a number of these potential objectives and evaluate their impacts on network travel time, using the INTEGRATION simulation model version 1.5c and, where possible, analytical techniques.
BASE CASE: NO METERING

In this section we present the example network used to demonstrate the relative benefits and drawbacks of ramp metering. The network characteristics, origin-destination demands, and speed-flow relationships are provided. An analytical analysis of these base case traffic conditions is conducted. Simulation results reflecting network traffic conditions are examined. Finally, analytical and simulation results are compared.

Example Network

The example network, illustrated in Figure 1, consists of a freeway section that has two identical junctions and a parallel arterial. There are 6 origin-destination zones, 47 nodes, and 64 directional links. All links are 0.5 km in length, except for arterial link 44, which is 5.15 km long. Each freeway link consists of two lanes, whereas all other links consist of a single directional lane.

The network was intentionally made to be simple for two reasons. First, to permit analytical analyses of traffic conditions, the network could not be designed to be too complex. Second, the intent of this study is to examine the impacts of a number of factors on total delay with the express purpose of identifying and illustrating the relative impacts of factors that affect ramp metering. The intent is to quantify the relative impacts, not the absolute ones. Furthermore, this study serves as an initial effort, and, as described in the Conclusions section, further research should be conducted.

In this paper ramp metering at only the second ramp junction (links 53 and 54) is examined. Subsequent research will examine the impacts and implications of metering at both ramp junctions. The ability to use the same network configuration is of benefit, as it will permit results to be compared directly with those described in this paper.

The origin-destination demands initially imposed on the network were the following: 3400 vehicles per hour (vph) from zone 1 to zone 4, 500 vph from zone 1 to zone 3, and 800 vph from zone 6 to zone 4. Initially there are no other demands on the network.

To determine the progression of traffic through the network and to quantify travel time, a single regime speed-flow–density relationship (7) was utilized. The freeway speed-flow relationship is nonparabolic and is characterized by a free speed of 80 km/hr, a capacity of 2000 vph/lane, a speed at capacity of 80 km/hr, and a jam density of 105 vehicles/km/lane.

Analytical Evaluation: No Metering

On the basis of the network characteristics, traffic demands, and specified speed-flow relationships, it is possible to carry out an analytical evaluation of expected traffic conditions by standard shock wave analysis. Figure 2 provides the graphical results of this analysis.

At time 0, traffic begins to enter the network from zone 1 at a rate of 3900 vph and with a speed of 90 km/hr. Approximately 7 min are required for the leading edge of this platoon to reach the on-ramp, 10 km downstream of zone 1.

Because the on-ramp already contributes a flow of 800 vph, there is only 3200 vph of remaining freeway capacity. As there is a demand of 3400 vph, a queue begins to form at an initial rate of 200 vph (200 = 800 + 3400 − 4000).

It must be determined whether the queue forms on the freeway, on the on-ramp, or on both. It is assumed that downstream capacity is apportioned to upstream demand in proportion to the upstream capacity. The on-ramp consists of a single lane with a capacity of 1600 vph. The freeway consists of two lanes, each with a capacity of 2000 vph. Therefore it is assumed that, of the 4000-vph downstream capacity on the freeway, 1143 vph (4000 × 1600/[1600 + 4000]) is apportioned to ramp demand, and the remaining capacity (4000 − 1143 = 2857 vph) is apportioned to upstream freeway demand. In this example, because the ramp demand of 800 vph is less than the ramp’s share of the downstream freeway capacity (1143 vph), none of the 200 vph excess demand is considered to queue on the on-ramp.

When the queue spills back upstream 0.5 km, direct access to the off-ramp is blocked. The flow that can be accommodated upstream of the off-ramp is a function of the downstream capacity flow (3200 vph) and of the number of vehicles that will flow onto the off-ramp. The 500-vph demand attempting access to the off-ramp constitutes 12.8 percent (500/3900 × 100) of the total freeway flow. Therefore it can be expected that a flow of 3670 vph (3200/[1 − 0.128]) can be accommodated upstream of the off-ramp. From this point the queue grows at an accelerated rate of 230 vph (3900 − 3670). The queue continues to grow until the demand is stopped after 1 hr. The maximum length of the queue is computed to be 7.5 km, and the total system travel time is estimated to be approximately 855 vehicle-hr.

Simulation Results: No Metering

The INTEGRATION simulation model is a microscopic traffic simulation model capable of modeling integrated networks, various traffic control devices, and advanced route guidance systems (8,9). The INTEGRATION model has been used to model a number of hypothetical and real networks (10,11) and is suited for use in evaluating the effectiveness of traffic control devices, including ramp meters. For the base case, no traffic control devices were modeled, and routes were prespecified such that all traffic utilized the freeway.
The model initiated simulation at time 0 with an empty network. Although demands were active for only 1 hr, the network was simulated for 80 min to permit all vehicles to reach their destinations. The model was configured to output average speed and volume every 5 min for each link in the network. At the end of each completed simulation run the total network travel time was computed.

Figure 3 shows the temporal and spatial variation in average 5-min freeway speed estimated by the simulation model. The significant region of speeds less than the speed at capacity (80 km/hr) is indicative of the congestion that occurs upstream of the on-ramp. It is also evident that after period 12 (1 hr into the simulation) the estimated speeds at the upstream end of the freeway section (low link numbers) return to the free speed value. This recovery occurs because the inflow of new demand has ceased and the network is simply emptying any vehicles that are already on the network.

Total network travel time incurred by the 4700 vehicles was 817.6 hr, or an average of 10.4 min per vehicle.

Comparison of Analytical and Simulation Results

Before using the simulation model to carry out sensitivity analyses, it is instructive to compare model results with analytical solutions. Temporal and spatial variation in speed estimated by the simulation model (Figure 3) and analytically computed (Figure 2) can be qualitatively and quantitatively compared. Some interpolation of simulation results is necessary, as simulation results are requested from the model only at 5-min intervals. Both the analytical and simulation results indicate a triangular region of congestion. The analytical solution indicates that congestion begins at approximately 7 min and ends at time 70 min. The maximum length of queue is 7.5 km. The simulation results indicate that congestion exists after 15 min but does not yet exist after 10 min. Inasmuch as simulation results are produced at 5-min intervals, it is necessary to interpolate to estimate more precisely when congestion occurred. On the basis of Figure 3, the time at which congestion occurs is estimated to be 14 min. Congestion ends at time 67 min, and the maximum length of queue is approximately 5.7 km.

These results indicate that the simulation model predicts a shorter period during which traffic flow is congested than does the analytical solution. The main cause for this discrepancy is the different manner in which the initial flow is assumed to traverse the empty network.

The analytical solution assumes that a platoon of traffic, having some constant volume and remaining within either the congested or the uncongested flow regime, travels as a homogeneous unit at some constant speed. Figure 2 indicates that the shock wave begins at time 0 and distance 0. This shock wave has a constant speed and represents the boundary between a regime that has a flow of 3900 vph (density of 43 vehicles/km) and one that has no flow.

In reality, the speed of a vehicle is determined more microscopically by the level of freedom that the driver of the vehicle has to maneuver. The presence of upstream vehicles generally does not affect the speed of downstream vehicles. Therefore it would be expected that the first vehicle to depart zone 1 would do so at approximately free speed. Subsequent vehicles would travel marginally slower as each additional vehicle increased the impedance of upstream vehicles. The result, then, is platoon dispersion, as vehicles that are first to enter the empty network have a higher speed than those entering later, even though all vehicles enter at a constant rate of 3900 vph. As this platoon travels the 10 km to the on-ramp, this dispersion effect is magnified such that the flow rate of the downstream end of the platoon is much lower than 3900 vph. In fact, during the simulation, the flow on link 19, upstream of the off-ramp, did not reach 3900 vph until period four, 20 min into the simulation.

Certainly, the assumption within the analytical approach that vehicle speeds are based on the macroscopic flow rate of upstream vehicles is less realistic. However, because of the additional complexity, it is very difficult to capture this effect of dispersion in an analytical methodology.

On the basis of this comparison, suitable explanations exist for the discrepancies between the model and analytical results. These
IMPACTS OF RAMP METERING

Having determined expected traffic conditions when no ramp control is in place, we are interested in determining the impact that ramp metering controls may have. Operational advanced traffic management systems utilize a wide range of ramp metering control strategies, from fixed metering rates to more complex ramp metering control strategies in which metering rates are determined on-line as a function of the freeway traffic conditions, the minimum and maximum metering rates, and queue spillback constraints. However, in this paper it is assumed that a fixed-rate time-of-day metering control is to be used. Because rates are fixed, no consideration is given to rate modification as the result of queue spillback. Before evaluation, the metering rate and the time period during which metering should take place must be determined.

Typically, ramp metering rates are set such that maximum utilization of the freeway is achieved without the occurrence of congestion. Analytically, it is quite clear that the ramp demand of 800 vph exceeds the available freeway capacity by 200 vph. Therefore a metering rate of 600 vph, or one vehicle every 6 sec, could be used. To avoid incurring unnecessary delay, metering should not begin until the traffic flow on the freeway reaches 3400 vph. Furthermore, metering should continue only as long as the freeway flow remains at this level. For this analysis it is assumed that no spatial, temporal, or modal diversion occurs.

The INTEGRATION model's ability to represent traffic signals was utilized to emulate fixed-time ramp meters. On the basis of results for the premetering case, ramp metering controls were initiated at 800 sec, as this is the time at which flow on the freeway (at the on-ramp) reaches 3400 vph. Metering continued until time 4000 sec, at which time the flow on the freeway (at the on-ramp) dropped to zero.

Figure 4 illustrates the temporal variation in average 5-min speed for three freeway links adjacent to the on- and off-ramps. It is clear from this figure that speeds never fall below 80 km/hr, the speed at
capacity. This indicates that this section of freeway never operates in the congested regime of the speed–flow relationship. The effect of ramp control can clearly be seen when the link speeds depicted in Figure 4 are compared with those provided in Figure 3.

Total network travel time was found to be 814.4 vehicle-hr, a travel time reduction of only 0.39 percent over that in the pre-metering case.

Because the computed benefit was rather small compared with benefits reported in the literature, it was decided to investigate those factors that might affect these benefits. This investigation is described in the next two sections.

SENSITIVITY ANALYSIS: NO DIVERSION

To identify those factors that might have a significant impact on the benefits of ramp metering, a series of sensitivity analyses was performed. In each case the base condition is the metering condition discussed in the previous section.

Four factors that affect ramp metering benefits were examined: the timing of the ramp control, the metering rate, the capacity drop effects, and the origin-destination (O-D) demands. Each of these is discussed in turn in the following sub-sections.

Effect of Timing of Implementation

The time of implementation of ramp control was found to have a significant impact on the estimated benefits of ramp metering. Figure 5 illustrates the variation in total network travel time with changes in the implementation time of the ramp metering. For this evaluation only the time at which metering began was altered; all other conditions remained unchanged from those discussed in the previous section, including the duration of the period for which metering was in effect. Figure 5 indicates that, for the effective conditions here, initiating ramp metering just 2 min earlier than optimal can negate any metering benefits.

Beginning metering later than optimal does not have such a significant effect. In fact, one would expect that if metering were begun approximately 1 hr after the optimal time, then all demands would have already passed the ramp and the result would be the same as for the pre-metering situation.

The implication of Figure 5 is that fixed metering plans, which invoke metering at pre-specified times of day independently of actual main-line flows, may cause a net increase in total delay if metering begins before the freeway flows reach capacity.

Effect of Metering Rate

We examined the effect of the actual metering rate by varying the ramp signal cycle length within consecutive runs of the simulation model. Six metering rates, ranging from one vehicle/8 sec (450 vph) to one vehicle/3 sec (1200 vph), were evaluated. Figure 6 illustrates the impact of metering rate on total network travel time. Clearly, metering rates that are more restrictive than necessary to prevent flow breakdown result in rather significant increases in total travel time. Under these conditions the additional delay incurred by traffic utilizing the on-ramp far outweighs the travel time savings experienced by freeway users as the result of the slightly higher freeway speed.

Interestingly, a metering rate of 720 vph results in a marginally lower total travel time than for all other rates. It must be remembered that capacity drop effects are not considered in this analysis. Inasmuch as there is no additional penalty incurred when flow breakdown occurs, total travel time is minimized when queuing occurs on both the freeway and the arterial.
Effect of Capacity Drop

There has been some debate over whether freeway capacity is reduced after flow breakdown occurs (12, 13). It is not our intent in this paper to add to this debate. Rather, our intent is to examine the impact that this phenomenon might have on modifying the potential benefits of ramp metering.

Proponents of the capacity drop concept indicate that, once flow breakdown occurs on the freeway, the capacity is reduced from precongested conditions to a lower congested value and is not restored to the precongested conditions until the freeway flow is again uncongested. It is not clear, however, what the potential magnitude of this capacity loss is. It has been stated that capacity reductions of as much as 25 percent may be possible (14).

This capacity loss was replicated in the INTEGRATION model by introduction of an incident to reduce the capacity of the freeway immediately upstream of the on-ramp. This capacity reduction was implemented at the onset of congestion and remained in effect until the freeway became uncongested.

To explore the potential effects of capacity reduction, we selected a modest reduction of 5 percent. Without metering, the effective capacity of the freeway immediately upstream of the on-ramp is

![Image of Figure 5](image-url)

**FIGURE 5** Effect of ramp control initialization time on total network travel time when diversion is not considered.

![Image of Figure 6](image-url)

**FIGURE 6** Effect of ramp metering rate on total network travel time when diversion is not considered.
The discussion so far has not considered the diversion of vehicles. In reality, if alternative routes are available a nontrivial diversion may occur. In this section we investigate the impact of two alternative diversion strategies, namely, a user optimal and a system optimal diversion.

Effect of O-D Demand

The absolute magnitude of metering benefits is also known to depend on the characteristics of the network in question, the operation of the ramp controls, the O-D demands on the network, and the availability and the quality of alternative routes.

O-D demands can have significant effects, particularly on ramp metering benefits, when queues that form on the freeway when metering does not exist spill back upstream and block access to upstream off-ramps.

To illustrate this, we consider that the capacity of the example freeway section has been increased from 4000 to 5000 vph upstream of the off-ramp. We consider also that demands from zone 1 to 3 have increased from 500 to 1500 vph, while all other characteristics of the freeway remain unchanged. The simulation of these conditions without ramp control results in a total travel time of 967.6 vehicle-hr. When these conditions are simulated again with the ramp flow metered at a rate of 600 vph, in the same manner as described in the initial metering model, the total travel time is estimated to be 958.3 vehicle-hr, which represents a reduction in travel time of 0.94 percent from the nonmetering case. This reduction of 0.94 percent can be compared with the 0.39-percent reduction obtained earlier. Clearly, inasmuch as the two models were identical except for the flow from zone 1 to 3, the additional benefits result from the fact that, with metering, the flow utilizing the off-ramp is not impeded. As in this model this flow is three times as large, the benefits are also much larger.

SENSITIVITY ANALYSIS: DIVERSION

The discussion so far has not considered the diversion of vehicles. In reality, if alternative routes are available a nontrivial diversion may occur. In this section we investigate the impact of two alternative diversion strategies, namely, a user optimal and a system optimal diversion.

Effect of User Optimal Diversion

In general, it is considered that individual drivers tend to choose routes that minimize their own travel times (15). In accordance with this behavior, it can be expected that drivers faced with extensive delays caused by the metering of a ramp will seek alternative routes that will result in a lower travel time cost.

There are, of course, numerous issues regarding perceived versus real costs, quality of available information, and bias toward certain roadway types. These concerns, though they are sometimes important, are not examined here.

For this model, vehicles received traffic information every 2 sec. Because, in practice, perfect information is rarely available, a 5-percent error was introduced into the information before it was provided to drivers.

The initial ramp metering model presented above was simulated with all drivers traveling from zone 1 to 4 receiving network information and able to divert. The resulting total travel time was 718.1 vehicle-hr, a reduction of 12.17 percent of that for the base premetering case.

To check that drivers routed themselves according to user optimal criteria, the average travel times for the two alternative routes from zone 1 to 4 were computed. Traversal of the ramp route required, on average, 7.1 min, whereas the arterial route required 6.9 min. As both routes have approximately the same average travel time, it can be concluded that the vehicles were diverted in accordance with user optimal behavior.

Effect of System Optimal Diversion

In the previous subsection we examined the effect of user optimal diversion, which is the way in which individuals are considered to behave at present. If, however, drivers were routed such that system optimal routings could be achieved, total system travel time would be further minimized.

Figure 7 illustrates the proportion of vehicles from zone 1 to 4 that use the arterial route and the associated system cost in terms of total travel time. As indicated, the system optimal diversion rate indicates that 100 percent of the vehicles that would normally use the controlled ramp should divert to the arterial. In this case the total travel time is only 701.4 vehicle-hr, representing a 14.21-percent reduction in system travel time compared with that for the base premetering case.

Drivers do not select system optimal routes unless they are forced to do so, so this analysis would be difficult to implement in practice. However, it serves as a convenient estimate of the upper limit on the benefits that one could achieve through the implementation of ramp metering in this example network.

CONCLUSIONS

A number of factors were shown to have a significant impact on the net benefits of ramp metering. Specifically, the effects of several of these factors, including O-D demands, metering rates, initiation time of metering, capacity drop, and diversion strategies, were examined.

These effects were quantified through the application of a simulation model for a small example network. Figure 8 provides a summary of these results. This analysis indicated that benefits of as much as a 26-percent reduction in total travel time may be obtained if metering is carried out efficiently while drivers are routed in a system optimal manner and that a 5-percent reduction in capacity occurs when the freeway becomes congested.
In the absence of any capacity reduction during congestion, benefits in the range of 12–14 percent can be obtained if drivers are permitted to divert to alternative routes.

In the absence of alternative diversion routes and a reduction in capacity during congestion, ramp metering was shown to be a potentially inefficient means of reducing total travel time.

It must be noted that travel time reductions as the result of the implementation of ramp metering strategies are highly network dependent. The presence and quality of potential diversion routes, the prevailing origin-destination patterns, and the physical locations of alternative routes all affect the premetering traffic conditions and dictate the benefits that might be obtained through the use of ramp metering.

In this paper we have examined the relative benefits of several control parameters through the use of a simple example network. The net benefits computed from this examination may not be applicable to more general networks.
Examining the impact of ramp metering by using analytical techniques for even simple networks can be rather difficult. This level of difficulty rises rapidly when traffic conditions and control strategies become more representative of actual field conditions. Furthermore, analytical techniques rely on simplifying assumptions regarding traffic behavior that limit their range of applicability.

The INTEGRATION model was found to be a robust evaluation tool that can be used objectively to quantify expected benefits of different ramp metering models under a variety of routing and controlled conditions, something that is difficult to do by using analytical techniques.

In this paper we have evaluated factors that affect ramp metering strictly in terms of reductions in total travel time. Because fuel consumption, emissions, and safety are also significant attributes of net benefits, effort should be undertaken to incorporate these factors into the evaluation.

Having shown that the INTEGRATION simulation model is able adequately to reflect traffic behavior and network control devices, we can then use it to evaluate the impact of ramp metering for various conditions on a more representative network.

REFERENCES


Publication of this paper sponsored by Committee on Transportation System Management.