A Ramp Metering Strategy for Rapid Congestion Recovery

Rui Jiang, Edward Chung

Smart Transport Research Centre, Civil Engineering and Built Environment School, Science and Engineering Faculty, Queensland University of Technology, 2 George Street, QLD 4000, Australia

Abstract: *Recurrent congestion caused by high commuter traffic is an irritation to motorway users. Ramp metering (RM) is the most effective motorway control means (M Papageorgiou & Kotsialos, 2002) for significantly reducing motorway congestion. However, given field constraints (e.g. limited ramp space and maximum ramp waiting time), RM cannot eliminate recurrent congestion during the increased long peak hours. This paper, therefore, focuses on rapid congestion recovery to further improve RM systems: that is, to quickly clear congestion in recovery periods. The feasibility of using RM for recovery is analyzed, and a zone recovery strategy (ZRS) for RM is proposed. Note that this study assumes no incident and demand management involved, i.e. no re-routing behavior and strategy considered. This strategy is modeled, calibrated and tested in the northbound model of the Pacific Motorway, Brisbane, Australia in a micro-simulation environment for recurrent congestion scenario, and evaluation results have justified its effectiveness.*

1 INTRODUCTION AND MOTIVATION

Motorway congestion has become a worldwide problem that strongly reduces traffic throughput, fluidity and safety, as well as increasing trip times and environmental pollution. In Australia particularly, the cost of congestion, estimated at around $9.4 billion in 2005, is expected to rise to over $20.4 billion by 2020 according to the Australian Government Department of Transport and Regional Services (Bureau of Transport and Regional Economics Department of Transport and Regional Services, 2007).

In order to tackle motorway congestion, many motorway management tools, such as ramp metering (RM) and variable speed limits (VSL), have been developed to reduce traffic congestion in metropolitan motorway networks. RM is considered to be the most effective tool currently available for motorway congestion management, with its effectiveness already proven by field implementation results (M Papageorgiou & Kotsialos, 2002). In a metered on-ramp, a traffic signal is placed to regulate the rate of vehicles entering motorways. Accordingly, RM is most effective for merging bottleneck. Normally, the metering rate is determined by real-time system conditions from both mainstream and ramp.

The general principle of RM is to temporarily hold ramp traffic to keep total demand at merging area around capacity for managing congestion; in other words, the major objective of the RM algorithm tries to approximately match demand with capacity at merging area to prevent congestion. As long as there is enough space in on-ramps for holding ramp traffic, this objective can be achieved. However, field evaluation results from the literature (Bhouri, Haj-Salem, & Kauppila, 2011; Muhurdarevic et al., 2006; Piotrowicz & Robinson, 1995) indicate that RM systems are successful in delaying the onset of congestion and reducing congestion, but not in eliminating congestion, even with the latest RM strategies (Geroliminis, Srivastava, & Michalopoulos, 2011; Papamichail, Papageorgiou, Vong, & Gaffney, 2010). This is due to the conflict between RM's major objective and the field circumstances in operation. More precisely, field limitations (i.e. maximum ramp waiting time and limited ramp storage space), activated by the expanded peak hours with high traffic demand nowadays, make this impossible in practice.

To further explore these constraints, with short time holding ramp traffic, both mainline traffic and ramp traffic can benefit if no congestion happens. Even for ramp traffic, free flow conditions on the mainline provide them better travel experiences in motorways. However, long time queuing for ramp traffic is unacceptable and inequitable. More importantly, limited ramp storage space in reality would cause ramp queue spillover back to the upstream arterial roads, which could seriously impact upon surface traffic. Consequently, RM systems in the field must increase metering rates at certain points to limit ramp traffic waiting time and to reduce queue spillover. This operation is usually against the objective of keeping the mainline traffic flowing freely, thereby causing flow breakdown.

The expansion of peak hours eventually activates these field limitations. For example, morning peak hours in some motorways of Brisbane have been brought forward two hours, compared with a decade ago (Moore, 2010). With such a long period of high demand, it is impossible to prevent congestion by holding so much ramp traffic.

According to the above analysis, RM can only delay and reduce motorway congestion, but not avoid it, given current field conditions.

Additionally, the traffic conditions at the start and end of peak hours are different. Take a major ramp with high ramp flow as an example. When peak hours start, both ramp and mainstream traffic increases. Ramp meter becomes restrictive for avoiding congestion as long as enough ramp storage is available along the network. Once ramp queue seriously affects upstream arterial roads, ramp meter must increase metering rate, which results in mainline queuing at the merging area. Meanwhile, the mainline queue will propagate upstream because of high mainstream flow. *This mainline queue propagation influences a large area and reduces network efficiency profoundly.* Conversely, at the end of peak hours, mainline queue stops propagating as demand decreasing. *After the mainline traffic condition has recovered, the merging bottleneck will not cause large mainline queuing, even with heavy ramp traffic.* Consequently, the earlier the mainline is recovered, the better the motorway network efficiency is. Moreover, to the best knowledge of the authors, there is no previous RM study taking recovery period as valuable information and investigating a strategy for motorway congestion recovery, although some studies (Geroliminis, et al., 2011; G. Zhang & Wang, 2013) have mentioned RM's impact for clearing congestion. This paper, therefore, focuses on post-congestion strategy targeting rapid recovery to further improve current RM systems implementation for recurrent congestion. Note that this study assumes no incident and demand management involved, i.e. no re-routing behavior and strategy considered.

The remainder of this paper is structured as follows. Section 2 states the concept of recovery for recurrent congestion. In the following section, the feasibility of using RM for recovery is discussed. The proposed recovery strategy is presented in Section 4. The simulation evaluation results are presented and analyzed in Section 5. Section 6 concludes this study.

2 RECOVERY CONCEPT

The aim of this section is 1) to define the recovery concept for recurrent congestion and 2) to explain the benefit of rapid congestion recovery. To achieve this aim, the change of traffic conditions together with the impact of RM operation during peak hours is analyzed. Before the analysis, this section firstly defines the benefits and the costs of RM.

RM uses mainline conditions to regulate ramp traffic entering motorways. This offers two main benefits: one is smooth merging behaviors by breaking large platoons, especially when the one-car-per-green signal principle (only one vehicle can pass the stop-line at each green phase per lane) is enforced, such as in Australia; the other is mainline traffic in free flow condition by delaying ramp traffic. The main disadvantages of RM include queue spillover back to adjacent arterial roads and ramp traffic delays. Obviously, ramp traffic delays and queue spillover can be seen as the costs of RM.

As analyzed in Section 1, motorway congestion is unavoidable due to excessively increased traffic demands during the expanded peak hours. Accordingly, the traffic conditions during peak hours can be divided into three phases as demonstrated in Figure 1:

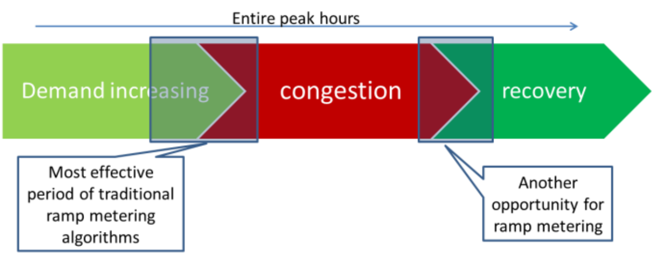

Figure 1 Traffic conditions during peak hours

- Phase 1 – Congestion building up: As peak hours begin, traffic demand starts to increase and congestion is building up.
- Phase 2 – Congestion: As traffic demand keeps increasing and stays high, motorways become congested eventually.
- Phase 3 – Recovery: Naturally, traffic demands reduce when peak hours end, and the traffic will recover from congestion and return to normal conditions.

During Phase 1, the total demands from motorway mainline and ramps do not exceed capacity for a long time, and only some key areas (such as merging area) are temporarily over-flowing; the RM system is able to remain a nearly free flow condition in motorway mainline by holding ramp traffic for a short time and using ramp storage over the network by ramp coordination. The smooth mainline traffic condition not only benefits mainline traffic, but also provides a better opportunity for ramp traffic to use the motorway. In addition, ramp traffic will only be delayed for a short time, as the total demands are around capacity and there is spare space over the network. In other words, the costs of RM can be controlled to be reasonable. Overall, short delays endured by ramp traffic improve system performance, and therefore reward all users in this situation. Phase 1 is when RM is most effective.

As traffic demands stay over capacity, RM cannot stand for excessive long ramp queues and delay, which means the costs of maintaining free flow mainline traffic are unacceptable. Consequently, congestion will happen eventually as the spare ramp storage space eventually consumes. Once congestion happens, the benefits of RM are reduced due to the capacity drop phenomenon and long mainline queues (M Papageorgiou & Kotsialos, 2002). As the demand is far beyond motorway capacity, RM cannot significantly improve motorway efficiency, but still contribute to traffic safety (Piotrowicz & Robinson, 1995).

When traffic demands reduce at the end of peak hours, the total demand is able to match with capacity, providing another opportunity for RM to benefit the whole system, again with controllable costs. As the mainline demand is decreasing, there would be no more mainline queue accumulation. Once long mainline queue is cleared, queue will not propagate and cause severe congestion due to reduced demand. Consequently, the earlier the mainline queue is cleared and the mainline traffic condition is recovered, the more travel time saving can be achieved. As the ramp demands are reduced, total ramp traffic delays might be able to be managed in an acceptable manner, and the risk of queue spillover is much smaller. Accordingly, the aim of a proper recovery strategy is to accelerate the system recovery and to manage the total ramp costs at the same time.

From the above analysis, it indicates that recovery phase is vital for this study. The definition of recovery phase for recurrent congestion is given as follows:

The recovery phase for recurrent congestion is the phase when total traffic demands start to reduce: the mainline traffic demand reduces and ramp demands do not increase.

3 FEASIBILITY OF USING RAMP METERING FOR RECOVERY

In this study, RM is considered as the motorway management tool for rapid congestion recovery. Consequently, the next question is whether RM can help mainline congestion recovery. As analyzed in Section 2, the primary task to achieve rapid congestion recovery is to clear motorway mainline queues. With RM, the choices are to restrict ramp traffic to a low metering rate or to increase the metering rate for discharging ramp traffic. Obviously, restricting ramp traffic can help discharge mainline queues. Therefore, the basic RM operation at recovery phase is to run the most restrictive metering rate, the restrictive metering control (RMC).

RMC provides two benefits for clearing motorway mainline queues. On one hand, the less traffic coming from the ramp, the more mainline traffic can go through the bottleneck. RMC will only allocate minimum number of ramp traffic entering the motorway, which means the maximum number of mainline traffic can pass the bottleneck.

On the other hand, RMC can increase the throughput at merging area, because it imposes minimum number of merging vehicles and might reduce number of lane changes at other lanes. This can be explained as follows from a microscopic point of view. At the merging area, mainline traffic, especially in the leftmost lane, would be disturbed by merging traffic. As a result, some drivers would apply brakes to allow ramp vehicles entering, which might slow the traffic flow in the leftmost lane. As the speed is slow, the possibility of vehicle lane changing increases and the other lanes would be affected. In other words, minimum number of both compulsory lane changes (merges) and induced lane changes is achieved by RMC. Accordingly, the effect of RMC is similar to the "smoothing effect" caused by high occupancy lane (Cassidy, Jang, & Daganzo, 2010; Menendez & Daganzo, 2007), which suggests that dampening lane changing activity can increase total bottleneck discharge flow rate. According to the above analysis, RMC introduces the minimum disturbances to mainline traffic at merging (given no ramp closure is allowed); therefore, RMC is able to improve the merging throughput at the recovery phase.

Field data analyses reported in the literature (Cassidy & Rudjanakanoknad, 2005; Geroliminis, et al., 2011; Srivastava & Geroliminis, 2013; L. Zhang & Levinson, 2010) also support that RMC has positive impacts on merging throughput. Zhang and Levinson (2010) analysed the data from the well-known Twin Cities RM experiment in Minnesota, USA, and drew the conclusion that RM actually increases active merging bottleneck capacity (2% during the pre-queue transition period and 3% of queue discharge flow rates after breakdown on average). Geroliminis et al. (2011) reported that the total capacity of an active bottleneck (mainline and on-ramp) depends on the ratio of the two flows, and that the capacity is smaller when ramp flows are higher. When congestion happens, Srivastava and Geroliminis (2013) showed that the level of the capacity drop depends on the ratio of mainline to ramp flow. This implies that restricting ramp flow at recovery phase can help mainline queue discharge. Cassidy and Rudjanakanoknad (2005) stated that: “By means of observation and experiment, we show here that metering an on-ramp can recover the higher discharge flow at a merge and thereby increase the merge capacity.”

Field data, from the Birdwood Road on-ramp in the Pacific Motorway, Brisbane, Australia, provides evidence

Table 1
Measured throughput of the Birdwood Road on-ramp.

	2012-12			*2013-12*		
	Upstream mainline	*Ramp*	*Downstream mainline*	*Upstream mainline*	*Ramp*	*Downstream mainline*
Monday	5016	1129	6199	5335	1038	6391
Tuesday	5095	1072	6204	5437	957	6434
Wednesday	5016	1102	6147	5118	960	6094
Thursday	5140	1058	6223	5494	1025	6547
Friday	4721	1128	5881	5226	995	6253
Average	4997.6	1097.8	6130.8	5322	995	6343.8

that the total throughput of an active merging bottleneck depends on the ratio of the two flows (mainline and ramp). The Birdwood Road on-ramp is a major bottleneck during morning peak hours, and is currently under a ramp signal pilot project starting from March 2012 (Queensland Department of Transport and Main Roads, 2014). The mainline motorway has three lanes and an acceleration lane of 120 meters at the Birdwood Road on-ramp. The metering rate has been strictly limited into small ranges due to testing purpose: the range is [1000 veh/h, 1100 veh/h] in 2012, and it eventually expands to [1000 veh/h, 1440 veh/h] in late 2013. According to the signal data, over 80% of the time after activation, ramp signal is running at 1000 veh/h. Consequently, the metered ramp flow rate does not change significantly.

We have collected two weeks data without incident reported (excluding weekends) after the installation of ramp signal: 5/12/2012 to 9/12/2012 and 2/12/2013 to 6/12/2013. The reason to select data after ramp signal installation is to guarantee the Birdwood Road on-ramp is an active bottleneck, which is a bottleneck whose performance is not affected by any bottlenecks occurring downstream. For each day, one-hour (7:00 – 8:00 AM) vehicle count is collected from three detector stations: upstream mainline detector station (located around 150 meters upstream from the merging area), ramp detector after the ramp signal and downstream mainline detector station (located around 280 meters from the end of acceleration lane, this is the measured throughput of the merging bottleneck). The data are illustrated in Table 1.

According to the data, the ramp flow in 2013 has reduced around 100 veh/h compared with in 2012. This ramp flow reduction increases the ratio of mainline to ramp flow. As the ratio is increased, the downstream throughput is also observed to be increased. In order to test whether the given data can establish a positive correlation between mainline/ramp flow ratio and downstream throughput, a regression analysis is carried out (the ratio is the explanatory variable and the downstream throughput is the dependent variable). The regression result is as follows:

$$\widehat{Q_{dn}} = 4802.4 + 289.56\gamma \tag{1}$$

where "$\widehat{Q_{dn}}$" is the estimated downstream throughput, "γ" is the ratio of mainline to ramp flow.

It is learnt that the estimated coefficient of the mainline/ramp flow ratio is 289.56, indicating a positive correlation. The t-test results are shown in Table 2. According to Table 2, the t-statistic is 3.036, and p-value is 0.016. At a 5% significance level, it is then concluded that the positive correlation between the two variables is statistically significant.

In summary, RM is suitable for rapid congestion recovery, especially for merging bottlenecks. Specifically, RMC is selected as the basic RM operation at recovery phase to maximize mainline queue discharge and to recover mainline congestion rapidly.

Table 2
Results of t-test.

	coefficients	*Standard Error*	*t-Stat*	*P-value*
Intercept	4802.4	474.63	10.12	0.000
ratio	289.56	95.39	3.036	0.016

4 STRATEGY DEVELOPMENT

This section presents an RM strategy for rapid congestion recovery, named zone recovery strategy (ZRS). In motorway mainline queuing, each queue starts from an active bottleneck and then propagates back to upstream. The tendency of each mainline queue is strongly related to the traffic conditions from the downstream active bottleneck to the next upstream active bottleneck. Therefore, the motorway network can be divided into several zones based on those active bottlenecks. As the major objective of recovery is to clear mainline traffic queues, it is reasonable to consider recovery for every mainline queue in one zone. The definition of zone is introduced first, and is followed by the details of the strategy.

4.1 Zone definition

Zones are determined by the locations of active bottlenecks. Specifically, a zone is defined as a motorway section starting from one active merge bottleneck and finishing at the next upstream bottleneck. An example is shown in Figure 2.

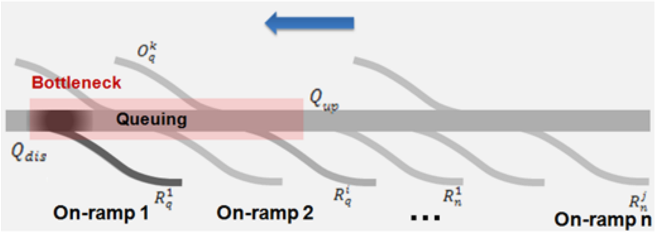

Figure 2 Zone definition and formulation

A zone can be further divided as a queuing area and a non-queuing area. The formulation and symbols of elements in a zone are defined as follows:

- Q_{dis}: the merging bottleneck flow at the active bottleneck;
- Q_{up}: the upstream incoming flow at the end of the mainline queue;
- R_q^i: the ith on-ramp incoming flow in the queuing area;
- R_n^j: the jth on-ramp incoming flow not in the queuing area;
- O_q^k: the ith off-ramp exiting flow in mainline queuing area.

4.2 Strategy framework

At the system level, four logical components are included in the framework to demonstrate the flow of the control. Figure 3 shows the framework of the ZRS. The first step, zone identification, is to divide a motorway into zones based on its mainline queuing condition. Based on the identified zone, the recovery phase is then detected for each zone – the recovery phase identification. As the recovery phase is determined, the special control for rapid congestion recovery purpose using RM is activated. At the same time, the traffic condition is monitored and the recovery can be withdrawn if necessary.

4.3 Zone identification

The objective of zone identification is to identify mainline queues and to formulate each zone. The key for zone identification is to detect mainline queues and to determine the queue head and the queue tail. Unlike queue estimation for work zones (Adeli & Ghosh-Dastidar, 2004; Ghosh-Dastidar & Adeli, 2006; X. Jiang & Adeli, 2004), this study adopts a mainline queue detection algorithm based on monitoring mainline detector measurements (Chung, Rahman, Bevrani, & Jiang, 2011; Queensland Department of Transport and Main Roads, 2008). The queue detection algorithm simply scans all detector measurements from downstream to upstream with a fixed interval (a 1-minute interval in this study), and then determines queues by consecutive congested detectors. Details can be found in the literature (Chung, et al., 2011; Queensland Department of Transport and Main Roads, 2008). This algorithm requires relatively dense detector placement along the whole motorway network; the test-bed used in this study – the northbound Pacific Motorway, Brisbane, Australia – has an average detector spacing of 650 meters and 49 detector stations in total, which makes it suitable for applying the algorithm.

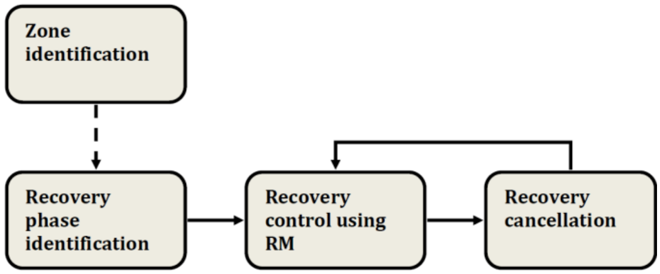

Figure 3 Framework of the ZRS

With the queue detection algorithm, the queue head can be accurately determined, together with a priori knowledge of any fixed bottleneck location. The exact point of the queue tail cannot be given, but the rough location is known. Figure 4 shows an example of mainline queue tail detection.

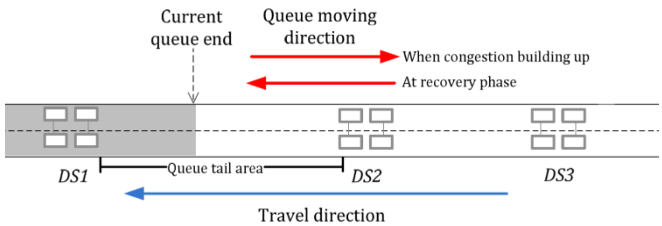

Figure 4 Detection of mainline queue tail

As can be seen from Figure 4, the queue tail area is the information obtained from the detection algorithm; this is accurate enough for monitoring queue changes for the recovery phase identification. Firstly, the queue tail is moving all the time. Secondly, monitoring queue changes is based on detector measures, so the key is to select the proper detector. In the above example, DS2 is the detector used for queue monitoring.

4.4 Recovery phase identification

Recovery phase identification activates the recovery control, and therefore the accuracy of recovery phase identification is critical for the ZRS. Based on the previous defined symbols for a zone, the recovery phase can be formulated as follows:

$$Q_{up}(t) + \sum_i R_q^i(t) - \sum_k O_q^k(t) < Q_{dis} \qquad (2)$$

where "t" is the notion for time stamp and others are the same as previous in Section 4.1. Once Inequality 2 is satisfied, the mainline queue length should reduce at this time interval, which is a necessary condition of the recovery phase. As shown in Figure 4, the queue tail direction is the same as travel direction at the recovery phase. Similar to zone identification, recovery phase identification also requires relatively dense detector placement along the whole motorway network.

In order to reduce the risk of false alarms (defined as mistakenly reporting recovery phase due to short term fluctuation of traffic condition), historical knowledge is incorporated in the component. Basically, the approximate time when the demand will start to reduce is known, so a simple way to incorporate this information is to add a time window in which the recovery phase identification is working. Another method to reduce the impact of short-term fluctuation is to process and project detector data by exponential smoothing using Equation 2, which can filter high frequency noise (fluctuation) in measurement. Additionally, a number of consecutive intervals are introduced to confirm the recovery phase. In this research, a 2-consecutive-interval is applied based on tests.

$$x^{sm}(t+1) = \alpha \cdot x(t) + (1-\alpha) \cdot x^{sm}(t) \qquad (3)$$

where "$x^{sm}(t+1)$" is the projected value by smoothing;
"$x^{sm}(t)$" is the smoothed value;
"α" is the smoothing parameter, 0.3 is chosen in this study based on calibration for data projection accuracy using field data (see details in Appendix A).

Figure 5 shows the flow chart of recovery phase identification. The time window is pre-set based on historical analysis. Once the time is on, then the process of recovery phase identification is activated. In this study, the pre-defined time is 8:30 am for the northbound Pacific Motorway, during morning peaks.

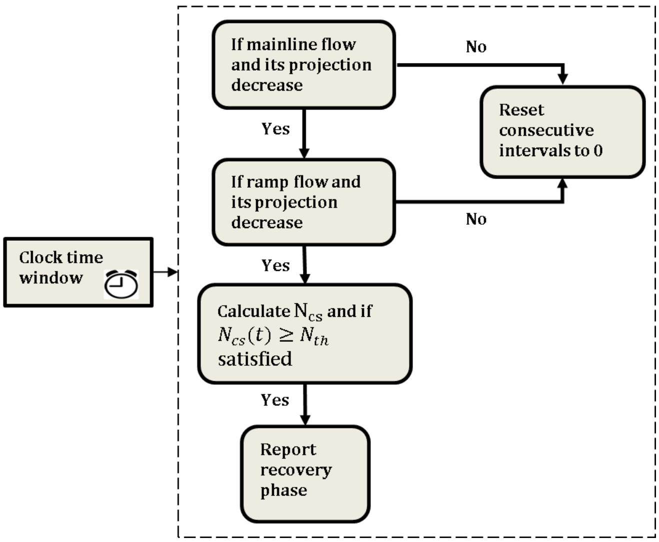

Figure 5 Flowchart of recovery phase identification

The process logic has three steps for checking the recovery phase as shown in the dashed rectangle in Figure 5. The first step is to monitor the change of the mainline incoming flow, ΔQ_{up}^{sm}, at the end of the queue and to project it for one more interval. The projection is used to check its changing trend, which is calculated also by simple exponential smoothing by Equation 2 above. If the flow and its projection are both reducing, the first condition is satisfied. In the second step, ramp arrival flows in the mainline queuing area are monitored and projected. Similarly, if both the flow and the projection are not increasing, the second condition is considered to be satisfied. The last step is to check the consecutive intervals. If the two conditions are satisfied, this interval is considered to be an effective interval. Once the consecutive effective interval is over the threshold, it would be acknowledged as an active recovery phase and the recovery control of the ZRS will be activated. The three steps are formulated as follows:

$$\begin{cases} \Delta Q_{up}^{sm}(t) < 0 \\ \Delta Q_{up}^{sm}(t+1) < 0 \end{cases} \qquad (4)$$

$$\begin{cases} \Delta \sum_i R_q^{i,sm}(t) \leq 0 \\ \Delta \sum_i R_q^{i,sm}(t+1) \leq 0 \end{cases} \qquad (5)$$

$$N_{cs}(t) \geq N_{th} \qquad (6)$$

where "Δ" indicates the change over last interval ($\Delta x(t) = x(t) - x(t-1)$);
"N_{cs}" is the number of consecutive intervals;
"N_{th}" is the threshold for N_{cs} (2 in this study).

4.5 Two-phase recovery control

As discussed in Section 3, RMC is selected as the basic RM operation for the recovery control. RMC can accelerate the discharge of the mainline queue by reducing ramp traffic and increasing merging bottleneck throughput. Besides, RMC naturally increases ramp traffic costs, including queue spillover and ramp traffic travel delay. In the recovery phase, temporarily increased ramp traffic cost can trade for the accelerated recovery of the whole network, and ultimately ramp traffic can benefit from a recovered mainline traffic. Therefore, the concept of the recovery control is stated as follows:

At the very beginning of the recovery phase, the recovery control ignores ramp costs for a short period from a system point of view, and applies RMC to achieve an increased mainline queue discharge rate. A reactive control based on both mainline and ramp conditions would then be activated to avoid unnecessary and excessive ramp costs.

The proposed recovery control is designed to be a two-phase control: a compulsory control phase followed by a reactive control phase.

4.5.1 Compulsory control

The objective of compulsory control is to accelerate the mainline queue discharge. In a zone, there are two groups of on-ramps: those in the queuing area and those not in the queuing area. Different strategies are designed for the two groups:

- Ramps in the queuing area will run the RMC. Operating RMC at these ramps would accelerate the discharge of mainline queue.

- The other ramps will run the local RM algorithm. Before activation of the ZRS, the system is running a coordinated RM (CRM) developed by STRC (R. Jiang, Lee, & Chung, 2013). The CRM algorithm is based on a master-slave mechanism to assign ramp coordination group: master ramp is a ramp with its own ramp queuing problem with high mainline flow, and slave ramps are to assist the downstream master ramp. Consequently, the ramps in the non-queuing area are highly likely to be slave ramps. In this case, they might already keep a certain length of on-ramp queue. In addition, mainline demand has reduced significantly, and these ramps are away from the mainline queue. Also, the demands of these ramps are usually not very high. Accordingly, even though these ramps increase their metering rates through the local RM, i.e. ALINEA (M. Papageorgiou, Hadj-Salem, & Middelham, 1997), the mainline queue is still likely to keep reducing. As the downstream ramps in the queuing area (highly likely to be masters) are forced to run RMC, it is an opportunity for the slaves to clear their own ramp queues. The benefit of clearing upstream ramp queues is that after the compulsory control phase, upstream ramps are available to become slaves with plenty of on-ramp storage space.

The most important parameter for compulsory control is the length of the compulsory period. If the length is over ten minutes long, it will cause unnecessary ramp costs; if it is as short as one or two minutes, it will not have any impact. Based on simulation tests (see details in Appendix B), a 3-minute period for the compulsory control phase is used in the study.

4.5.2 Reactive control

As compulsory control would make ramps in the queuing area suffer from the RMC, it is not reasonable to run it for too long, such as over ten minutes. After the compulsory control phase, the objective is to regain control of the total ramp costs. Again, different strategies are designed for the two ramp groups:

- Ramps in the queuing area will run the cost constrained additive increase minimal release (CC-AIMR) algorithm, a localized RM algorithm. The reason of using only local information is that even in coordinated RM, these ramps would be the master running the local RM anyway. The details of the CC-AIMR algorithm are presented in Section 4.5.3.
- The other ramps in the zone will run the coordinated RM. As these ramps already get refreshed in terms of their ramp queues, they are now able to contribute more to the system.

Figure 6 demonstrates the two-phase recovery control.

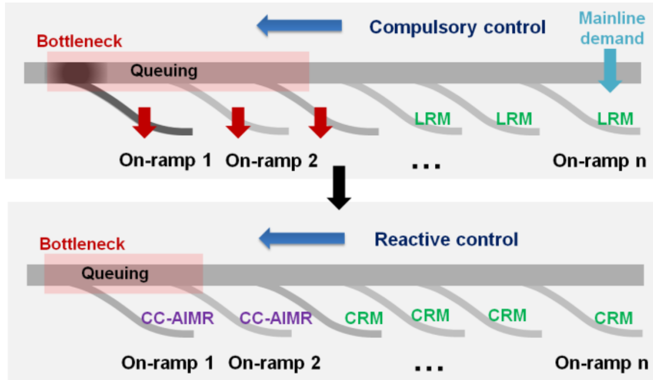

Figure 6 Two-phase recovery control

4.5.3 Cost constrained additive increase minimal release algorithm (CC-AIMR)

In the CC-AIMR, the cost constraint part is to impose a constraint for limiting the maximum queue spillover time, while the AIMR part is inspired by the TCP congestion control algorithm (additive increase and multiplicative decrease, AIMD algorithm) (V. Jacobson, 1988). Using cost constraint will bring local ramp costs into consideration. Besides, AIMR algorithm is designed from the TCP congestion control principle, which has a quick recovery response to congestion. In the CC-AIMR algorithm, a metering rate is selected from the AIMR algorithm and a local queue management algorithm (R. Jiang, Chung, & Lee, 2012). The flow chart is shown in Figure 7. The first step is to check if ramp queue spilt over and to record accumulative queue spillover. The second step is to compare the accumulative queue spillover with the pre-defined maximum acceptable spillover: if the accumulative queue spillover is smaller than the constraint, the metering rate from AIMR will be applied; otherwise, the metering rate from the local queue management algorithm will be applied. In this study, the maximum of accumulated spillover time is 4 minutes according to simulation tests (see details in Appendix C). Details of AIMR algorithm are introduced below.

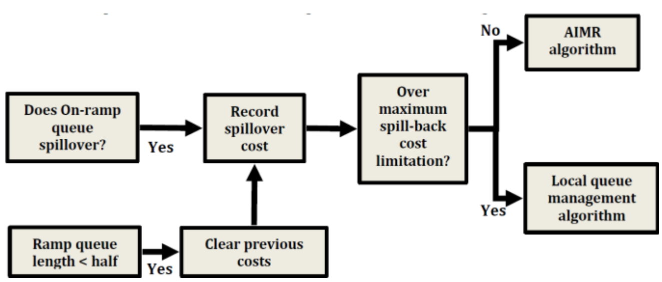

Figure 7 Flowchart of the CC-AIMR algorithm

As noted, the AIMR algorithm originates from the AIMD (additive increase multiplicative decrease) algorithm within TCP congestion control. The basic logic of the prototype algorithm is similar to the AIMD algorithm in TCP control.

When the feedback variable still indicates a congested state in the merging area, the metering rate goes directly back to its minimum; when the feedback variable shows a recovery state, the metering rate increases additively. Therefore, this algorithm is called the additive increase minimum release (AIMR) algorithm. Compared with the AIMD algorithm, the feedback variable here is generated by the loop detector data in real time. In this research project, speed at merging area is chosen as the feedback variable, because it is the direct detector measure to check the occurrence of traffic flow breakdown. There are two operations in the AIMR algorithm:

- AI operation, which is the logic to increase metering rate additively, see Equation 7.

$$r(t+1) = r(t) + \mu \tag{7}$$

where "r" is metering rate;

"μ" is a metering rate increasing constant, and 180 veh/h/lane is used in this study.

- MR operation, which is the logic to set the metering rate to its minimum.

The AIMR algorithm has three steps for every calculation interval (CI, the time interval to update the metering rate, a 1-minute interval is used in this study). There are three steps in the process:

a. Get detector data.
b. Update and record state of the CI. There are three states:
   - Congested state (state C): If the speed of merge area drops below a pre-defined threshold (45km/h is used in this study), this CI will be counted as state C.
   - Recovered state (state R): If the speed of merge area recovers to a pre-defined threshold (55km/h is used in this study), this CI will be counted as state R.
   - Middle state (state M): If the CI state is neither state R nor state C, it will be counted as state M.
c. Determine metering rate for next CI:
   - If consecutive recovered CI satisfies a pre-defined threshold (2 used in this study), the metering rate for the next CI will increase by the pre-defined increment, "μ".
   - If consecutive congested CI satisfies a pre-defined threshold (2 used in this study), the new metering rate will be set back to the minimum metering rate.
   - Otherwise, the metering rate will keep the same as in the previous CI.

4.6 Recovery cancellation

Recovery cancellation component allows withdrawing the recovery control for a zone or not. There are three reasons for the cancellation:

- False alarm: if the queue keeps propagating back to upstream again, this would be considered as a false alarm, and the recovery is cancelled.
- Success of recovery: once the mainline queue is cleared, it is considered a successful recovery and the recovery action is then finished.
- Timeout of recovery: After a certain period of recovery control, the mainline may still be queuing, but the traffic condition should be improved. Consequently, the normal RM should be capable of handling the improved condition. This is the timeout condition. In this research project, the timeout is set as 30-minute, the maximum running period for the recovery control.

5 SIMULATION EVALUATION

This section evaluates the performances of the proposed ZRS at the recovery phase of the recurrent congestion; that is, the morning peak hours. The modeling platform used in the investigation is Aimsun 6.1. The proposed strategy was implemented using the Application Programming Interface (API) functions provided by Aimsun.

5.1 Simulation test-bed and simulation settings

The Pacific Motorway northbound (inbound) model was used for this study. The test-bed network is approximately a 30-km motorway (M3) from Logan City to the Brisbane CBD, as displayed in Figure 8. This motorway section serves a large volume (approximately 130,000 veh/day) of commuter traffic in the morning peak hours, leading to heavy recurrent congestion. There are 16 on-ramps and 17 off-ramps along the network. There are three active bottlenecks during the morning congestion:1) the Gateway Motorway Interchange bottleneck is the most severe bottleneck caused by large off-ramp flow; 2) the Birdwood Road on-ramp causes congestion by large ramp flow; and 3) the Stanley Street on-ramp causes severe congestion by both large mainline and ramp flows. Figure 9 shows the three bottlenecks in a speed contour (red color indicates low speed while green one indicates high speed) of the Pacific Motorway northbound during morning peak.

The simulation network used in this study was edited by Queensland Department of Transport & Main Roads, and model parameters calibrated by the STRC (Rahman, et al., 2011). The calibration process included two steps: the calibration was conducted first at a disaggregate level using the real dataset of individual vehicles at Vulture Street from the Queensland Department of Transport and Main Roads; and then the parameters were adjusted at a network-level by comparison of overall simulated traffic situation (contours of flow, speed and occupancy) with the reality. In Table 3, the key model parameters calibrated are listed, and more calibrated parameters can be found in the literature (Rahman, et al., 2011). For validation, field data were collected for the Gateway point from the same source, for the same period.

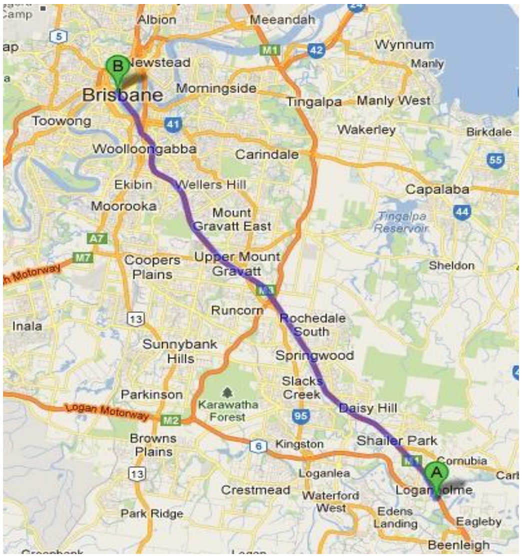

Figure 8 Pacific Motorway, Brisbane, Australia

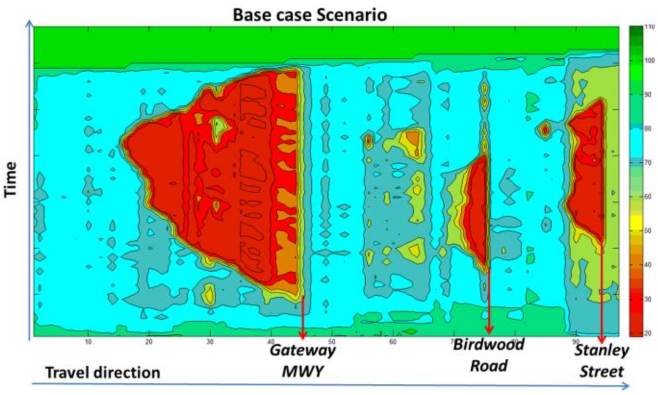

Figure 9 Mainline speed contour of the Pacific MWY

The complete scenario to depict the real traffic demand on the network was developed in terms of traffic state

Table 3
Key calibrated parameters (Rahman, Bevrani, & Chung, 2011).

	Parameter name	*Value*
Global parameters	Reaction Time	0.9 s
	Simulation Step	0.45 s
	Percent Overtake	98%
	Percent Recover	99%
Vehicle parameters	Maximum Desired Speed	110 km/h
	Maximum Acceleration	6 m/s2

according to PTDS (Public Transport Data Source) database. The selected case day, 15 March 2010, was a regular business day (Monday) with major educational institutions running, with good weather (no rain) and with no incidents reported. A detailed comparison between simulated data and real data can be found in the literature (Rahman, et al., 2011). The complete scenario was conducted for a period of 17-hour with time intervals of 15 minutes from 5:00 am to 10:00 pm. According to the whole day volume contour, the morning peak period was determined as a 5-hour period from 5am to 10am, when the northbound (inbound) motorway witnessed high levels of recurrent congestion. As mentioned in the abstract, no diverge behavior is considered, so one more hour without any given demand is added to the end of the simulation period for clearing all the generated traffic. Consequently, the total simulation period is 6 hours.

5.2 Test scenarios and performance indicates

Three test scenarios are tested in both test-beds to demonstrate the effectiveness of the ZRS:

- Base case (BC) scenario assumes no RM control.
- Coordinated RM scenario (CRM) operates the coordinated RM control system developed by STRC (R. Jiang, et al., 2013). According to the best knowledge of the authors, coordinated RM systems represent the current practice. In the field implemented RM systems, some, including the Zone algorithm (Stephanedes, 1994), the Helper ramp algorithm (Lipp, Corcoran, & Hickman, 1991), SWARM (system wide adaptive ramp metering) (Paesani, Kerr, Perovich, & Khosravi, 1997) and the Bottleneck algorithm (L. Jacobson, Henry, & Mehyar, 1989), were developed in the 1980s and 1990s, so they are outdated and would not be appropriate representatives of the latest state of practice. Since 2000, two well-known field implementations of coordinated RM reported in the literature are HERO (heuristic ramp metering coordination) (Papamichail, et al., 2010) and SZM (stratified zone metering) (Geroliminis, et al., 2011). For HERO, a commercial system, there is not enough detail from the literature to fully model it. For SZM, the logic is much more complicated, and not enough detail can be found in the literature. Therefore, STRC developed a CRM strategy based on master-slave mechanism as the benchmark.
- ZRS scenario operates the proposed recovery strategy after 8:30 am. Otherwise, it is the same as the CRM scenario.

A total of four performance indicators are used in the evaluation. Although the whole simulation period is 6 hours, the performance indicators are only collected for the recovery period (after 8:30 am). The definitions are as follows:

- Total Travel Time (TTT): the most widely used efficiency indicator at a system level for RM. It is calculated by summing up all the individual vehicle travel times in the network. The unit of TTT is veh·h.

- Average mainline traffic delay (MTD): this indicator gives a sense of the coordination benefit. The northbound Pacific Motorway is divided into 31 sections based on the location of metered ramps. For each section, individual vehicle travel time within the section is collected and aggregated into the average section travel time. The sum of average section travel times is the entire motorway travel time. The free flow travel time for the entire motorway is also calculated assuming 80 km/h as the free flow speed. Finally, MTD is defined as the difference between the actual mainline traffic travel time and the free flow travel time. The unit of this indicator is sec/trip.
- Total queue spillover time (TQST): the sum of the total time for each on-ramp when ramp queue spills over to upstream arterials. In this study, the queue spillover is defined as 1-min time occupancy of the ramp entrance detector is over 70%.
- Average ramp traffic delay (RTD): the way to calculate RTD is slightly different from MTD. Firstly, the aggregated travel time for each ramp is calculated by collecting individual vehicle travel times in ramp. The ramp travel time is collected from the ramp entrance to the downstream merge area. The free flow speed for this section is assumed at 70 km/h. The delay for each ramp is defined as the difference between the actual ramp travel time and the free flow travel time. To consider that the ramp traffic volume varies by each location, the average RTD is calculated using the following equation. The unit of this indicator is sec/veh.

$$RTD = \sum_i RTD_i \cdot Q_i \Big/ \sum_i Q_i \quad (8)$$

where "RTD_i" is the ramp traffic delay for ramp i; and,
"Q_i" is the total volume of ramp i.

The first two indicators are used to measure the benefit of RM, while the other two are used to measure the ramp costs.

5.3 Results and analyses

In order to reduce the impact of random seed in micro-simulation, 30 replications were simulated and the results were collected. Figure 10 illustrates the iterative average of TTT of the BC scenario as more replications are simulated. It can be seen that the iterative average remains stable when the number of replications is over 25. This implies that the minimum number of runs should be larger than 25, and 30 runs in this study is an appropriate selection. Table 4 summarizes the average results from 30 replications of the four performance indicators. Also, root mean square error (RMSE see Equation 9) is provided in the bracket of each average result.

Table 4
Simulation results summary.

	BC(no RM)	*CRM*	*ZRS*
TTT (veh·h)	4406.7(313.5)	3264.7(106.2)	3203.3(98.9)
MTD (sec/trip)	858.8(146)	242.4(44.8)	212.3(42.8)
TQST (minute)	90.2(10.4)	244.3(31.8)	246.6(26.4)
RTD (sec/veh)	59.4(8.2)	153.3(11.9)	148.0(13.7)

$$RMSE = \sqrt{\frac{1}{n}\sum_n (Observation - Estimation)^2} \quad (9)$$

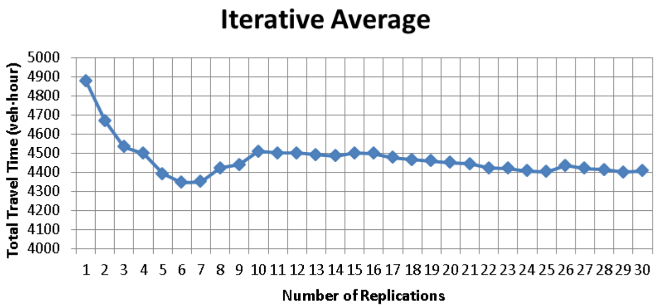

Figure 10 Iterative average of TTT of the BC scenario

When comparing the BC scenario with other two, it clearly shows the benefits and the costs of applying RM. On the one hand, the overall system efficiency, represented by TTT, has been improved significantly, 25.9% by the CRM scenario. The comparison of MTD between the BC and the CRM scenario indicates mainline traffic condition benefits dramatically: MTD reduces from 858.8 sec/trip in the BC scenario to 242.4 sec/trip in the CRM scenario. On the other hand, ramp traffic experiences long delays at on-ramps: averagely increased 93.9 sec/veh by the CRM scenario. Moreover, TQST is much higher once activating RM, which is a natural result for RM systems. Noted, the CRM scenario does improve system efficiency at recovery phase over the BC scenario, because the CRM scenario reduces overall congestion level and results in an easier recovery compared with the BC scenario.

The focus here is to compare the CRM and the ZRS scenario. In particular, the percentage reduction of TTT is 1.9% (from 3264.7 to 3203.3 veh·h). For MTD, the improvement is remarkable, with a 12.4% reduction from 242.4 to 212.3 sec/trip. This indicates that the quicker recovery provides a much better mainline traffic condition. Even for ramp traffic, the average delay in recovery period, RTD, is decreased by 3.5%. However, TQST has increased slightly from 244.3 to 246.6 minutes, showing the marginal costs paid for the rapid recovery. All these results imply that the proposed ZRS can further improve RM systems at the recovery phase, and verify the design principle of the ZRS: at the very beginning of the recovery phase, running RMC for a short time can assist the mainline congestion recovery, and then running CC-AIMR for managing ramp queue spillover; with a recovered motorway network, not only mainline traffic but also ramp traffic will benefit (RTD decreases with the ZRS). Noted that the main bottleneck for the test-bed is the weaving bottleneck at the Gateway Motorway interchange, which indicates the recovery concept is effective for a network in which the non-merging bottleneck dominates.

Table 5 and Table 6 show ramp queue spillover and ramp traffic travel time during the recovery period, for individual ramps.

Table 5
Individual ramp spillover results.

	CRM	*ZRS*
Beenleigh Road	0.0	0.0
Grandis Street	0.1	0.0
Murrays Road	0.0	0.0
Centenary Road	0.0	0.0
Loganlea Road	3.3	3.2
Service Road	3.6	9.2
Fitzgerald Avenue	0.0	0.0
Sports Drive	113.6	109.9
Logan Road	1.0	0.3
Kessels Road	0.0	0.0
Mains Road	1.5	1.4
Birdwood Road	14.0	19.9
Duke Street	0.0	0.0
Stanley Street	104.5	99.5
Alice Street	2.6	3.1
Ann Street	0.0	0.0

Individual ramp results provide inside views from which to examine the ZRS control algorithm. Based on Table 5, the increased queue spillover time is distributed on the Service Road on-ramp (an increase of 5.6 minutes), the Birdwood Road on-ramp (a 5.9-minute increase) and the Alice Street on-ramp (an increase of 0.5 minutes). The first two ramps are high-demand ramps in the two mainline queuing areas: the Gateway Motorway interchange bottleneck and the Birdwood Road merging bottleneck (see Figure 9). In the ZRS scenario, these ramps are forced to operate RMC in the compulsory control phase, thereby causing more queue spillover. The situation for the Alice Street ramp is different. It is right at the downstream of the Stanley Street on-ramp (the CBD bottleneck, see Figure 9). As the Stanley Street on-ramp is operating RMC at recovery phase, increased flow makes for denser mainline traffic at the Alice Street on-ramp. Consequently, the local RM system holds more ramp traffic for mainline priority, resulting in slightly more queue spillover.

When checking the individual ramp traffic travel time in Table 6, only slight increases in percentage are found for the Service Road on-ramp (6.3%) and the Mains Road on-ramp (less than 0.1%). At both the Birdwood Road on-ramp and the Alice Street on-ramp, the average ramp traffic

Table 6
Individual ramp traffic travel time results.

	CRM	*ZRS*
Beenleigh Road	44.5	43.8
Grandis Street	35.5	35.3
Murrays Road	45.2	40.6
Centenary Road	38.2	38.1
Loganlea Road	52.1	51.2
Service Road	105.1	111.7
Fitzgerald Avenue	32.5	31.8
Sports Drive	396.1	382.5
Logan Road	142.3	130.7
Kessels Road	57.6	47.6
Mains Road	197.2	197.3
Birdwood Road	614.6	614.2
Duke Street	84.5	81.7
Stanley Street	108.3	103.1
Alice Street	348.8	328.3
Ann Street	63.2	58.5

travel time reduces simply because the mainline queue is cleared in a short time and then ramp traffic can be quickly discharged. This indicates that the short period “pains” (maximum 5.9 minutes more spillover at the Birdwood Road on-ramp) benefit the whole system while not increasing individual ramp travel time much. Another conclusion that can be drawn from the individual ramp results is that the most effective ramp control for recovery is to meter the closet ramp upstream of the bottleneck, which is similar to the conclusion from Zhang and Levinson (2004).

Figure 11 illustrates the speed contour for the recovery phase, in which the color scheme demonstrate the speed (red color indicates low speed while green one indicates high speed). As can be seen, the time duration of mainline queues in the Gateway Motorway area and the City area has been clearly reduced, which indicates a quick recovery of the mainline traffic achieved by the ZRS. Note that the speed contours in Figure 11 are samples; all the replications produced similar speed contour patterns.

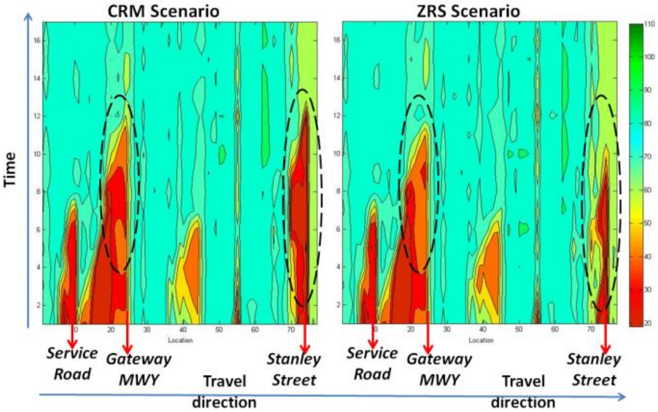

Figure 11 Mainline speed contour – CRM vs. ZRS

6 CONCLUSION REMARKS

This study investigates the concept of the post-congestion management for motorways in the recovery period of recurrent congestion, and developed a zone recovery strategy (ZRS) using RM for rapid congestion recovery. The study begins by defining the recovery concept, and an investigation for the feasibility of using RM for recovery is conducted. A field data analysis shows that the total throughput of an active merging bottleneck depends on the ratio of the two flows (mainline and ramp). Then, the ZRS with four components is proposed. Especially for the recovery control, a two-phase control algorithm is designed to rapidly recover the mainline congestion and to control the ramp costs (queue spillover and ramp travel delay) at the same time. The strategy is then simulated for evaluation in a microscopic simulation model with fixed demand (no diverge behavior considered, which is important for field implementation of any traffic control). Evaluation results concluded the effectiveness of the ZRS. The following conclusions are drawn from this study:

- Recovery phase for recurrent congestion is the phase when total traffic demands naturally reduce. This is another opportunity for RM to benefit the whole system with reasonable costs, because the total demand is able to be matched with network capacity.
- RM is an appropriate control tool for rapid congestion recovery, especially for merging bottlenecks. Specifically, RMC is the basic RM operation at recovery phase to maximize mainline queue discharge, because RMC will maximize the proportion of mainline traffic through bottlenecks.
- The proposed ZRS provides a better mainline congestion recovery at the recovery phase with the two-phase control algorithm: that is, running RMC for a short time to accelerate mainline traffic recovery and ultimately benefiting the ramp traffic. Individual ramp analyses reveal that the quicker recovery is achieved by the "sacrifice" of the on-ramps at the major bottlenecks.
- Two future works, if addressed, could further improve this study. The first one is to consider diverge behavior in the simulation tests of the strategy. For any field traffic control, the diverge behavior caused by the control implementation will change the dynamics of the traffic network profoundly. Consequently, a comprehensive evaluation should include the consideration of diverge behaviors. Secondly, the other future work is to include VSL for better integrated motorway management. VSL can directly impact mainline traffic flow, which has the potential to limit mainline flow to merging area (Carlson, Papamichail, Papageorgiou, & Messmer, 2010; Hegyi, De Schutter, & Hellendoorn, 2005). By doing so, more ramp traffic can be discharged to alleviate ramp overflow limitation, and thereby increasing the effective range of RM.

ACKNOWLEDGMENTS

This research was supported by the members of the Smart Transport Research Centre (STRC). The authors wish to acknowledge the support and contribution of the STRC members.

REFERENCES

Adeli, H., & Ghosh-Dastidar, S. (2004). Mesoscopic-wavelet freeway work zone flow and congestion feature extraction model. *Journal of Transportation Engineering, 130*(1), 94-103.

Bhouri, N., Haj-Salem, H., & Kauppila, J. (2011). Isolated versus coordinated ramp metering: Field evaluation results of travel time reliability and traffic impact. *Transportation Research Part C: Emerging Technologies*.

Bureau of Transport and Regional Economics Department of Transport and Regional Services. (2007). *Estimating urban traffic and congestion cost trends for Australian cities*. Canberra: Retrieved from http://www.bitre.gov.au/publications/2007/files/wp_071.pdf.

Carlson, R. C., Papamichail, I., Papageorgiou, M., & Messmer, A. (2010). Optimal motorway traffic flow control involving variable speed limits and ramp metering. *Transportation Science, 44*(2), 238-253.

Cassidy, M. J., Jang, K., & Daganzo, C. F. (2010). The smoothing effect of carpool lanes on freeway bottlenecks. *Transportation Research Part A: Policy and Practice, 44*(2), 65-75.

Cassidy, M. J., & Rudjanakanoknad, J. (2005). Increasing the capacity of an isolated merge by metering its on-ramp. *Transportation Research Part B: Methodological, 39*(10), 896-913.

Chung, E., Rahman, M., Bevrani, K., & Jiang, R. (2011). Evaluation of Queensland Department of Transport And Main Roads Managed Motorway. Brisbane: Smart Transport Research Centre, QUT.

Geroliminis, N., Srivastava, A., & Michalopoulos, P. (2011). A dynamic-zone-based coordinated ramp-metering algorithm with queue constraints for Minnesota's freeways. *Intelligent Transportation Systems, IEEE Transactions on, 12*(4), 1576-1586.

Ghosh-Dastidar, S., & Adeli, H. (2006). Neural network-wavelet microsimulation model for delay and queue length estimation at freeway work zones. *Journal of Transportation Engineering, 132*(4), 331-341.

Hegyi, A., De Schutter, B., & Hellendoorn, H. (2005). Model predictive control for optimal coordination of ramp metering and variable speed limits. *Transportation Research Part C: Emerging Technologies*, *13*(3), 185-209.

Jacobson, L., Henry, K., & Mehyar, O. (1989). Real-time metering algorithm for centralized control. *Transportation Research Board*(1232), 17-26.

Jacobson, V. (1988). *Congestion avoidance and control*.

Jiang, R., Chung, E., & Lee, J. B. (2012). Local On-ramp Queue Management Strategy with Mainline Speed Recovery. *Procedia-Social and Behavioral Sciences*, *43*, 201-209.

Jiang, R., Lee, J., & Chung, E. (2013). *A Multihierarchical Motorway Ramp Signalling Strategy*. Paper presented at the Australasian Transport Research Forum 2013, Brisbane.

Jiang, X., & Adeli, H. (2004). Object ‐ Oriented Model for Freeway Work Zone Capacity and Queue Delay Estimation. *Computer ‐ Aided Civil and Infrastructure Engineering, 19*(2), 144-156.

Lipp, L., Corcoran, L., & Hickman, G. (1991). Benefits of central computer control for Denver ramp-metering system. *Transportation Research Record*(1320), 3-6.

Menendez, M., & Daganzo, C. F. (2007). Effects of HOV lanes on freeway bottlenecks. *Transportation Research Part B: Methodological, 41*(8), 809-822.

Moore, T. (2010, November 16, 2010). Early bird gets the traffic jam, *Brisbane Times*.

Muhurdarevic, Z., Condie, H., Ben-Shabat, E., Ernhofer, O., Hadj-Salem, H., & Papamichail, I. (2006). *Ramp Metering in Europe - State-of-Art and Beyond*. Paper presented at the 5th International Conference on Traffic and Transportation Studies (ICTTS), Xi'an, China. http://www2.napier.ac.uk/euramp/papers/ICTTS_2006.pdf

Paesani, G., Kerr, J., Perovich, P., & Khosravi, E. (1997). *System Wide Adaptive Ramp Metering In Southern California*. Paper presented at the ITS America 7th Annual Meeting.

Papageorgiou, M., Hadj-Salem, H., & Middelham, F. (1997). ALINEA Local Ramp Metering: Summary of Field Results. *Transportation Research Record: Journal of the Transportation Research Board, 1603*(-1), 90-98. doi: 10.3141/1603-12

Papageorgiou, M., & Kotsialos, A. (2002). Freeway ramp metering: an overview. *Intelligent Transportation Systems, IEEE Transactions on, 3*(4), 271-281.

Papamichail, I., Papageorgiou, M., Vong, V., & Gaffney, J. (2010). Heuristic ramp-metering coordination strategy implemented at monash freeway, australia. *Transportation Research Record: Journal of the Transportation Research Board, 2178*(1), 10-20.

Piotrowicz, G., & Robinson, J. (1995). *Ramp Metering Status in North America. 1995 Update*. Retrieved from http://isddc.dot.gov/OLPFiles/RSPA/006815.pdf.

Queensland Department of Transport and Main Roads. (2008). Variable speed limits system - Concept of Operations. Brisbane.

Queensland Department of Transport and Main Roads. (2014). Managed Motorways, from http://www.tmr.qld.gov.au/Travel-and-transport/Road-and-traffic-info/Managed-Motorways.aspx

Rahman, M. M., Bevrani, K., & Chung, E. (2011). *Issues and concerns of microscopic calibration process at different network levels: Case study of Pacific Motorway*. Paper presented at the Proceedings of the Australasian Transport Research Forum 2011.

Srivastava, A., & Geroliminis, N. (2013). Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model. *Transportation Research Part C: Emerging Technologies, 30*(0), 161-177. doi: http://dx.doi.org/10.1016/j.trc.2013.02.006

Stephanedes, Y. (1994). *Implementation of on-line zone control strategies for optimal rampmetering in the Minneapolis Ring Road*. Paper presented at the Proceedings of 7th International Conference on Road Traffic Monitoring and Control.

Zhang, G., & Wang, Y. (2013). Optimizing Coordinated Ramp Metering: A Preemptive Hierarchical Control Approach. *Computer-Aided Civil and Infrastructure Engineering, 28*(1), 22-37. doi: 10.1111/j.1467-8667.2012.00764.x

Zhang, L., & Levinson, D. (2004). Optimal freeway ramp control without origin–destination information. *Transportation Research Part B: Methodological, 38*(10), 869-887.

Zhang, L., & Levinson, D. (2010). Ramp metering and freeway bottleneck capacity. *Transportation Research Part A: Policy and Practice, 44*(4), 218-235.

APPENDIX A: CALIBRATION OF SMOOTHING PARAMETER

The smoothing parameter is tested and selected based on field data. Three on-ramps with ramp signal installed are used for data collection. The three on-ramps are Logan Road ramp, Mains Road ramp and Sports Drive ramp on the Pacific Motorway Northbound, Brisbane, Australia. The

Table 7
Calibration results of smoothing parameter.

Ramp	*a*	*0.2*	*0.3*	*0.4*	*0.5*	*0.6*
Logan Road	RMSE	7.36	7.25	7.47	7.48	7.76
	ARE	11.93%	11.69%	12.10%	12.13%	12.56%
Mains Road	RMSE	11.81	11.81	12.77	11.81	14.56
	ARE	9.46%	9.40%	10.23%	9.54%	11.83%
Sports Drive	RMSE	9.38	9.33	9.50	9.51	11.26
	ARE	9.67%	9.52%	9.65%	9.88%	11.51%
Average	RMSE	9.52	9.46	9.91	9.60	11.19
	ARE	10.35%	10.20%	10.66%	10.52%	11.97%

raw data is 1-minute vehicle count from the ramp entrance loop detector and the exit loop detector. RMSE (see Equation 8) and mean percentage error (MPE) are used to evaluate the accuracy:

$$MPE = \frac{\frac{1}{n}\sum_n |Observation - Estimation|}{\frac{1}{n}\sum_n Observation} \times 100\% \qquad (9)$$

Five values are tested, and the results are listed in Table 7. According to the results, 0.3 is chosen in this study.

APPENDIX B: CALIBRATION OF COMPULSORY CONTROL LENGTH

The compulsory control length is calibrated by simulation data. The performance indicators used are the same as in simulation evaluation (see Section 5.2). Five values are tested, and the results from 30 replications are illustrated in Table 8.

Table 8
Calibration results of compulsory control length.

	3-minute	*4-minute*	*5-minute*	*6-minute*	*7-minute*
TTT	3205(97.5)	3196(94.1)	3192.7(102.3)	3186.7(103.2)	3182(99.4)
MTD	213.3(42.6)	209.6(40.9)	208.4(43.2)	205.4(43.8)	204(43.1)
TQST	246(26.7)	255.5(36.6)	266.1(35)	265.5(29.3)	286.3(30.4)
RTD	148.2(13)	148.2(13.4)	146.7(13.6)	148.2(11.7)	153.4(14.6)

It can be seen from Table 8 that as the compulsory control length increases, the overall system efficiency is improved (reduced TTT) and MTD reduces. This indicates that longer compulsory control accelerates the mainline recovery. On the contrast, increasing compulsory control length increases TQST. When comparing RTD results with the CRM scenario in Table 4, only 7-minute scenario has higher RTD than the CRM scenario, which means ramp traffic do not benefit from quick recovery if the compulsory control phase is unnecessarily long. Overall, 3-minute compulsory control is selected in this study, because almost no additional spillover is caused by the ZRS.

APPENDIX C: CALIBRATION OF MAXIMUM SPILLOVER TIME IN CC-AIMR ALGORITHM

A calibration based on simulation data is conducted to see the impact of maximum spillover time in CC-AIMR algorithm. The performance indicators used are the same as in simulation evaluation (see Section 5.2). Five values are tested, and the results from 30 replications are illustrated in Table 9. According to the results, this parameter has minor impact on the system performance. Specifically, this parameter does not have much impacts on MTD and RTD, and affects TTT and TQST limitedly. Consequently, 4-minute is selected because of better TTT performance and similar TQST performance compared with 5-minute.

Table 9
Calibration results of maximum spillover time in CC-AIMR algorithm.

	3-minute	*4-minute*	*5-minute*	*6-minute*	*7-minute*
TTT	3203.1(99)	3203.3(98.9)	3205(97.5)	3206.6(100.6)	3207.9(100)
MTD	212 (43)	212.3(42.8)	213.3(42.6)	214.5(43.4)	214.8(43.4)
TQST	252.9(23.8)	246.6(26.4)	246(26.7)	253.5(27)	250.7(26.6)
RTD	149.2(13.3)	148.0(13.7)	148.2(13)	149.7(14.3)	148.8(14.2)