QUEUE STORAGE LENGTH DESIGN FOR METERED ON-RAMPS

Guangchuan Yang, Ph.D. Candidate, Zong Tian, Ph.D., P.E.
Department of Civil and Environmental Engineering, University of Nevada, Reno, NV89557

Presentation for the ITE Western District Annual Meeting, June 2017, San Diego, CA

ABSTRACT

This paper proposed an analytical method for queue length modeling at metered on-ramps, and developed a mesoscopic simulation model for queue length estimation under various demand-to-capacity scenarios. For each on-ramp feeding movement, the traffic arrivals were divided into two regimes: the saturated platoon arrival regime and the non-platoon arrival regime. Accordingly, the on-ramp flow arrival profile was descripted; then, the ramp queue generation profile was descripted and queue lengths were estimated based on the input-output method. Simulation results indicated that for under-saturated conditions, the required queue storage length was approximately 6.2 percent of on-ramp demand when demand was less than 500 vphpl, or 4 percent when demand was between 500 and 900 vphpl.

INTRODUCTION

Ramp metering was first implemented in 1963 on the Eisenhower Expressway (Interstate 290) in Chicago, Illinois. Since then it has been systematically deployed in major urban areas in the U.S. It aims to break up platoons of vehicles released from upstream signals, and ensure the total traffic entering the freeway section remains below the operational capacity of the section by using the “access reduction technology”. Numerous studies have proved that ramp metering has the potential of addressing recurring freeway congestion problems *(Cambridge Systematics, Inc., 2001)*; however, there are several challenges involved in the design and operation of ramp metering. These challenges mostly stem from the queue storage length issue (see Figure 1), since the majority of the metered on-ramps in the United States are retrofitted to existing ramps where sufficient queue storage length is not available. When adequate queue storage is not provided, the ramp queue may spillback to the upstream signal and affect the operations of the surface street system *(Tian et al., 2004)*. Indeed, on-ramp queue overspill is one of the most critical reasons that jeopardize the acceptance and implementation of ramp metering. A typical practice is to suspend ramp metering to flush the queue when spillover occurs *(Chaudhary et al., 2004)*; nevertheless, such a queue flush or queue override strategy will go against the intention of ramp metering, and tends to degrade freeway operations. Therefore, an accurate estimation of queue length is of significant importance for optimizing ramp metering design and operations.

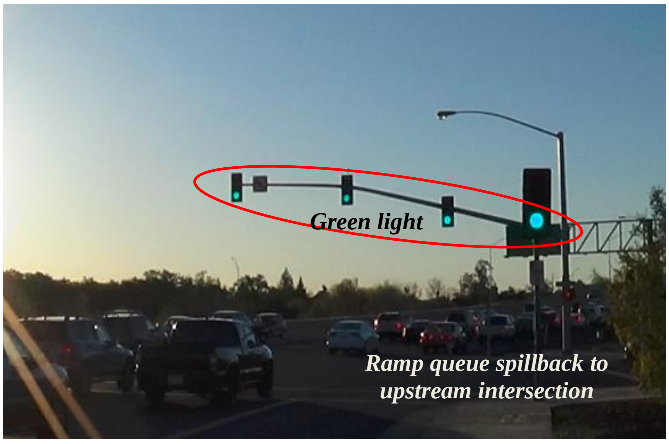


Figure 1 Ramp queue spillback example in Sacramento, CA

In practice, ramp queue storage length designed as a certain percentage of the peak hour on-ramp demand is a widely accepted rule-of-thumb by local agencies *(Wang, 2013)*. The percentage number varies by the queue length estimation methods. In California, it was found that queue storage length as 7 percent of peak hour on-ramp demand would accommodate the majority of existing ramp metering sites. Similarly, the percentage number was determined as 3.3 percent in Nevada, 10 percent in Minnesota, 5 percent for retrofitted ramps and 10 percent for new construction ramps in Wisconsin, and in Texas, a range between 1.7 and 3.1 percent was recommended. The advantage of such guideline recommendations is that they are easy to apply. Nevertheless, the aforementioned guidelines were either estimated based on the queuing theory, which did not take into consideration the varying traffic arrival flow patterns (i.e., uniform arrivals versus random arrivals with vehicle platoons), and normally only provide an average queue estimation, which cannot be directly applied to ramp queue storage length design where the maximum or a percentile queue is generally used; or merely developed based on limited field data from existing ramp metering sites, where the occurrence of queue spillover may make it difficult to observe the true on-ramp demands and queue lengths. With consideration of the challenges involved in the design and operations of ramp metering, this paper aims to develop a reasonable method to estimate the actual queue lengths at metered on-ramps to direct the ramp queue storage length design.

QUEUE LENGTH MODELING METHODOLOGY

The Input-Output Method

The input-output approach, which is also known as the cumulative arrival and departure method *(Newell, 1982)*, was employed by this study for queue length modeling at metered on-ramps. The arrival rate is the on-ramp demand and the departure rate is the metering rate and thereby the queue length is the accumulated difference between the arrival and departure rate over time. Equations (1) through (3) provide a generalized description of the traditional input-output method. Given the on-ramp demand, $V(t)$, and the capacity of the facility, c, the cumulative arrival function, $A(t)$, and the departure rate, $D(t)$, can be determined. Subsequently, the performance measure of queue length, $Q(t)$ can be obtained.

$$\frac{dA}{dt} = V(t) \tag{1}$$

$$\frac{dD}{dt} = \begin{cases} c, \ if \ A(t) > D(t) \\ V(t), \ otherwise \end{cases} \tag{2}$$

$$Q(t) = A(t) - D(t) \tag{3}$$

On-Ramp Flow Arrival Profile

Due to the existence of upstream signals, the on-ramp flow arrives at the ramp meter with unique patterns. The actual on-ramp arrival profiles are mainly influenced by the upstream signal phasing and timing, number of feeding movements, and average arrival rate of each movement.

By the assumption that on-ramp feeding traffic uniformly arrives at the upstream signalized intersection during each cycle, and with knowing the signal timing information, the time period of the two regimes would be estimated deterministically using queuing theory. Figure 2 shows the queuing and discharge diagram at one of the legs of the upstream intersection. For each phase in a cycle, queue clearance time (i.e., G_0), which is also the on-ramp platoon arrival period, could be estimated as:

$$G_0^i = \frac{A_i \times (C - G_i)}{S_i - A_i} \tag{4}$$

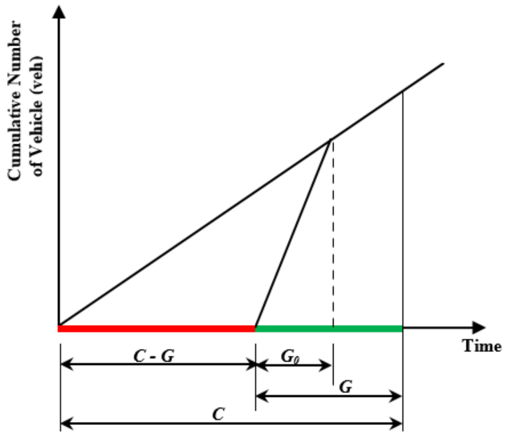


Figure 2 Upstream traffic discharge diagram

Where, G_0^i: queue clearance time for the i^{th} phase, $G_0^i \leq G_i$; C: cycle length of upstream signalized intersection; G_i: effective green time of the i^{th} phase; A_i: uniform arrival flow rate of the i^{th} phase; S_i: saturation flow rate of the i^{th} phase.

By knowing the phasing sequence of an upstream intersection, the saturation flow rate S_i and the average arrival rate A_i of each phase, the on-ramp arrival flow profile during each cycle could be depicted. Figure 3 shows a typical metered on-ramp that has three feeding movements. Each on-ramp feeding movement controlled by the upstream signal will arrive at the ramp meter with two flow regimes: the saturated queue discharge regime and the uniform arrival regime. For Example, during the upstream through movement *(TH)*, the first portion of the flow occurs when the upstream intersection through movement discharges at its saturation flow rate S_{TH}; after the queue of through movement is cleared, the flow rate reduces to its average arrival rate A_{TH}.

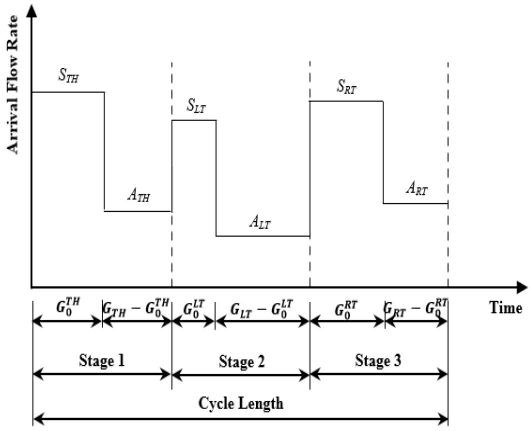


Figure 3 On-ramp flow arrival profile

Queue Generations at Metered On-Ramps

A metered on-ramp is a standard queuing system; a vehicle queue is formed behind a ramp metering signal when the vehicle arrival rate exceeds the ramp metering rate. In consistency with Figure 4, for each feeding movement, the platoon arrival rate equals its saturation flow rate; after that, on-ramp traffic is assumed to arrive at the ramp at the average arrival rate. The platoon arrival period of each phase (G_0^i) is estimated using equation (4), and the queue profile is generated through the input-output method. In this study, only fixed-time ramp metering strategy is considered. The maximum queue length during each phase or each cycle can then be identified, as illustrated in Figure 5. These maximum queue lengths would be used for producing a percentile queue length during a long period (e.g., 95th percentile queue length during one hour), which would eventually be used to direct metered on-ramp queue storage design.

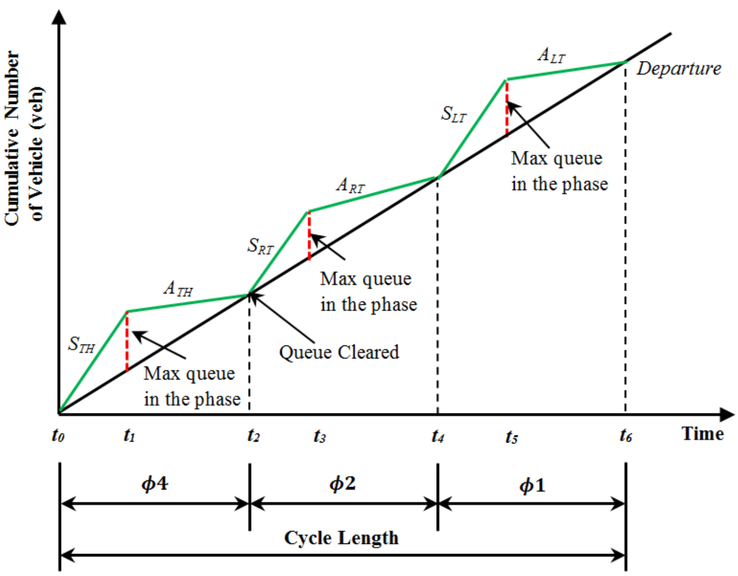


Figure 4 Queue generation profile at metered on-ramp

Modeling versus Observation

To validate the accuracy of the proposed queue length modeling method, the phase-by-phase arrival and departure traffic data, as well as the actual queue lengths were collected at several representative metered on-ramp in California. Figure 5 illustrates the camera layout for field data collection. Three video cameras were placed at different locations of the ramp to capture the on-ramp demands, metering rate, and queue lengths, respectively. Eventually, these data were manually extracted from the video clips.

The field collected phase-by-phase traffic data were used for estimating ramp queues based on the proposed queue length modeling method. The modeling results were compared to the field observed queue lengths to verify the derivations between the two, as illustrated in Figure 6.

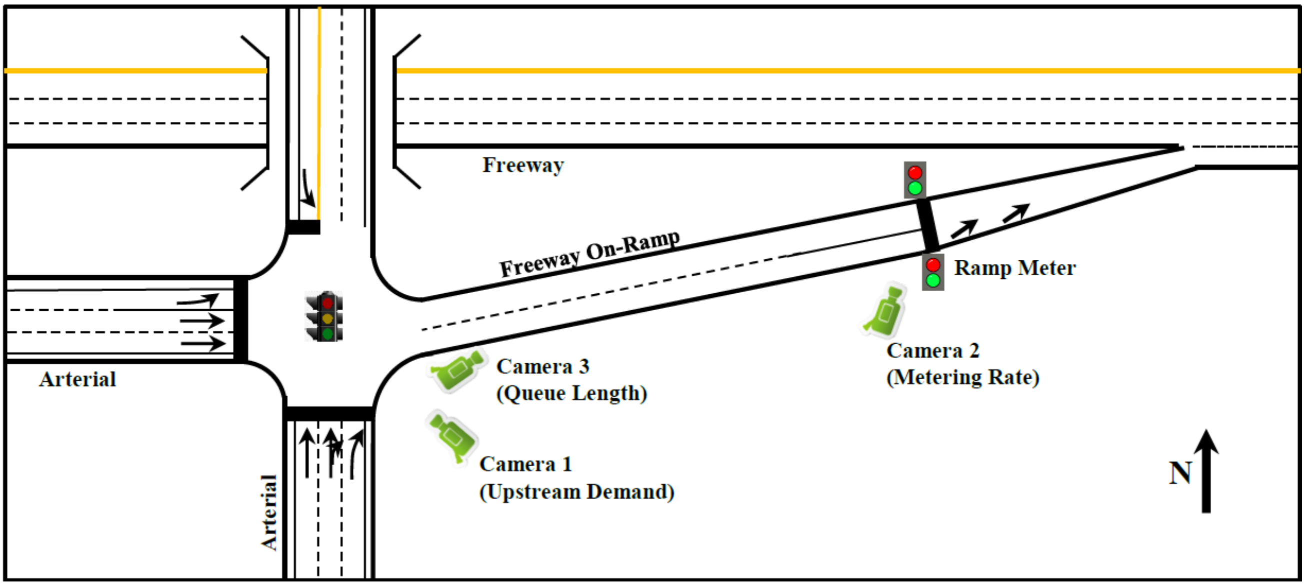


Figure 5 Camera layout for queue length data collection at representative ramp metering location

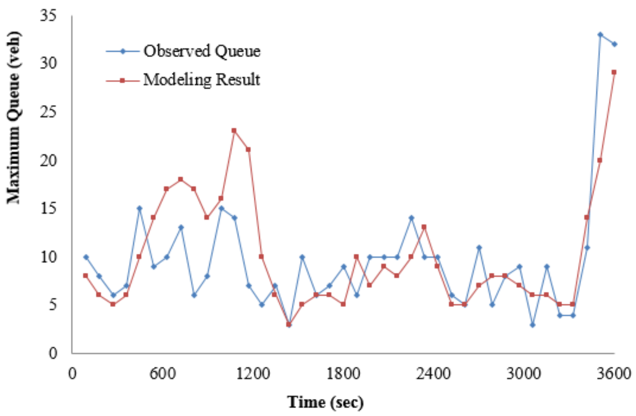


Figure 6 Queue length modeling results versus field observed queue profile

The comparison shows that the proposed queue length estimation model can properly capture the observed queue profile. The deviations mostly stemmed from the random nature of the parameters involved in the modeling. For instance, it was found that drivers usually turn right on a red signal whenever there is a safe gap. With the un-expected right-turn-on-red vehicles, the on-ramp arrival flow profile of each feeding movement is changed. Right-turn-on-red tends to exacerbate the maximum queue length in a cycle, as it adds additional on-ramp demands to one particular feeding movement.

MESOSCOPIC QUEUE LENGTH SIMULATION MODEL

For a queue length study, generally speaking a field measurement method provides a more accurate assessment of actual conditions; however, the data availability (e.g., various combination of demand and capacity scenarios) and quality issues (e.g., randomness of traffic flow, and measurement error when queue spillovers) may limit the usefulness of the measured data. In these situations, a simulation method may be more feasible for queue storage design, since a simulation model has the ability of rapid modeling traffic performance for various scenarios.

The aforementioned queue length modeling algorithm was built into a mesoscopic queue length simulation model to estimate queue lengths under various on-ramp demands and metering rates. The mesoscopic simulation model was developed using the C# programming language; it can generate random traffic volume numbers to capture the randomness of on-ramp arrival flow. The model contains three modules: the on-ramp demand simulation module, the metering rate simulation module, and the queue length simulation module. The flow chart as presented in Figure 9 illustrates the queue length simulation process. Using this model, regression equations can be developed based on a large number of simulation runs. In turn, summary tables and charts are generated for quick estimation of queue length for a given demand versus capacity scenario.

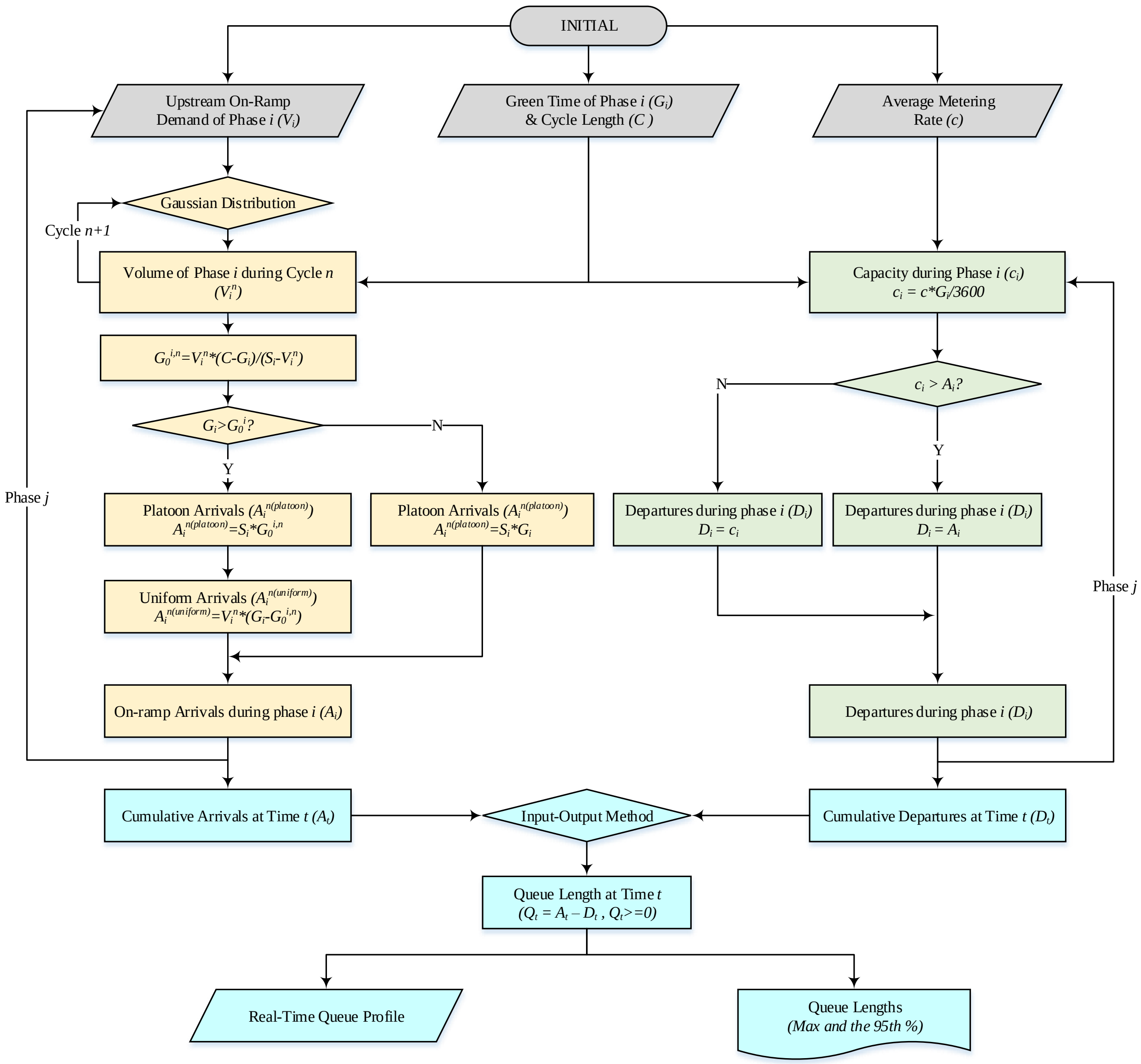


Figure 8 Mesoscopic queue length simulation flow chart

Users need to input the general simulation parameters including: average hourly on-ramp demand of each feeding movement, saturation flow rate of each on-ramp feeding movement departing from the upstream signal, average hourly metering rate, and signal timing information of upstream signal. The on-ramp demand modeling module reads the average hourly on-ramp demand input of each feeding movement, and randomly distributes the total demand to each cycle. For a feeding movement, traffic flow arrival at the upstream signalized intersection is assumed to follow the Gaussian distribution. Based on the collected cycle-by-cycle traffic arrivals at the study ramp metering site, for each feeding movement, the generated traffic arrivals in a cycle was assumed to range between 0.5 and 1.5 times the average arrival rate, and the summation of all the arrivals equal to the total on-ramp demand. Then, the metering rate modeling module reads the metering rate input (a fixed-time metering strategy was assumed in this paper) and the queue length modeling module simulates the queue lengths at the metered on-ramp. The default simulation time is one hour. The cumulative arrivals and departures at time *t* could be determined; accordingly the queue length at time *t* could be calculated through the input-output method. Finally, the simulation model can generate the queue versus time profile and output the maximum and the 95th percentile queue for each simulation. The user interface of the developed mesoscopic simulation model is shown in Figure 9.

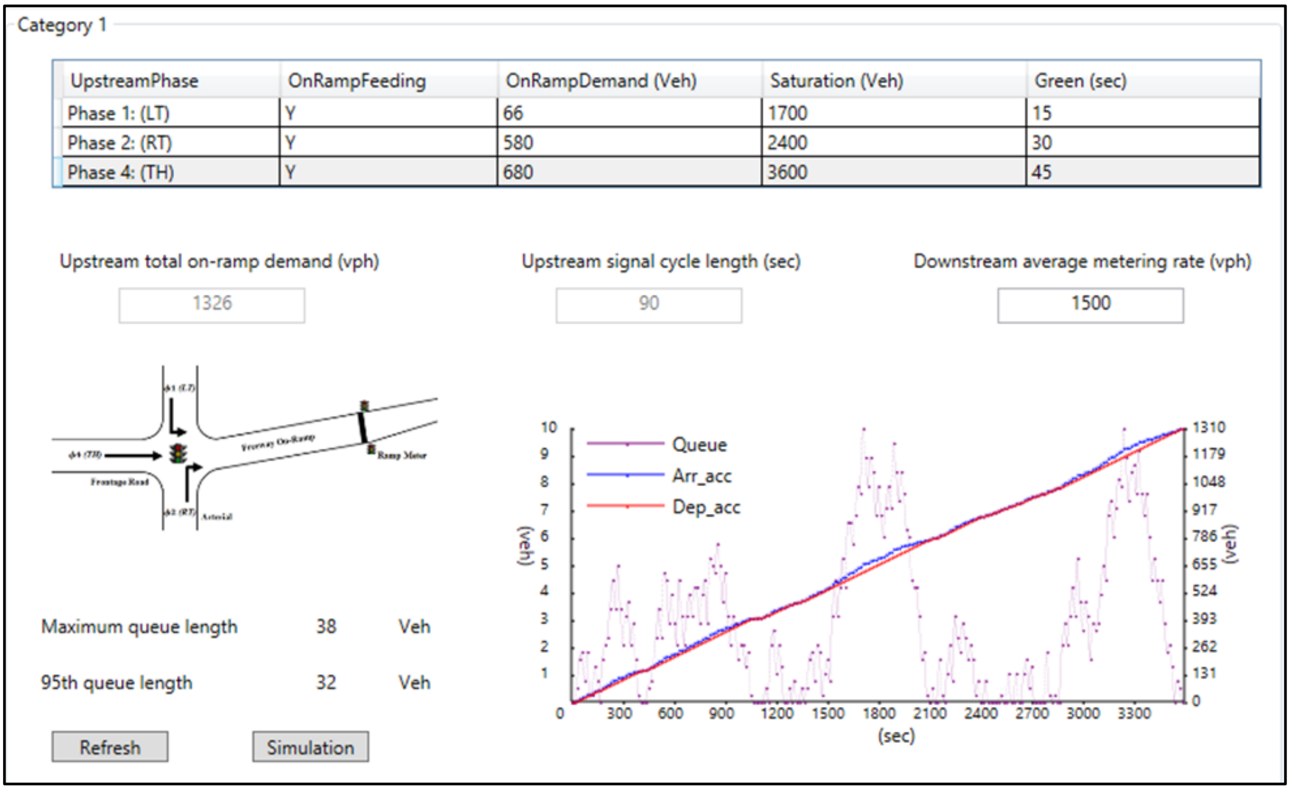
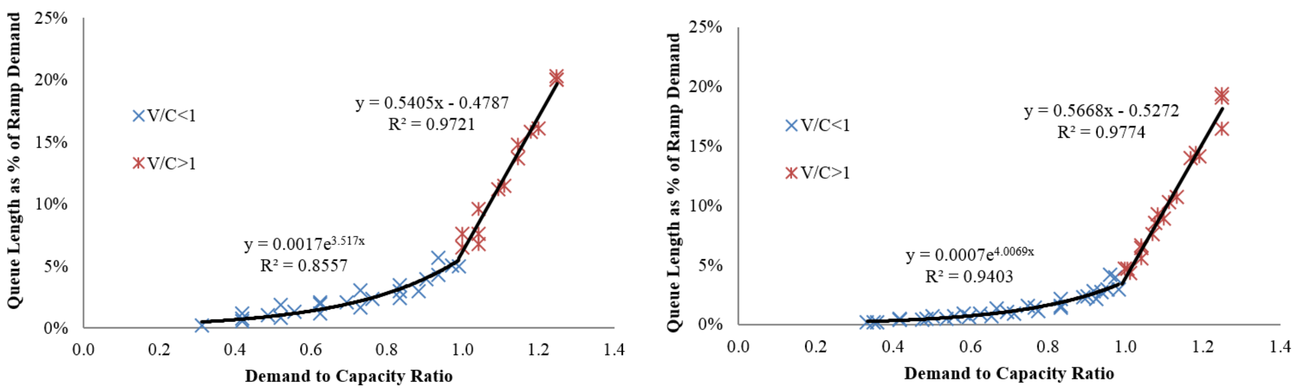


Figure 9 User interface of the developed mesoscopic queue length simulation model

QUEUE LENGTH SIMULATION

Queue Length as a Percent Number of On-Ramp Demand

Various scenarios were created for different combinations of on-ramp demands and metering rates. For each scenario, 20 simulations were performed to obtain the mean of the simulated 95th percentile queue lengths. To provide a generalized queue storage length design recommendation, the simulated queue length (i.e., absolute queue length number of vehicles) was converted to queue length as a percent number of on-ramp demand. Based on engineering judgement, the simulation scenarios were classified into two categories: low metering rate conditions (metering rate less than 500 vphpl) and high metering rate conditions (metering rate between 500 vphpl and 900 vphpl). Based on large number of simulation runs, the scatter plots of queue length as percentage of demand under various demand-to-capacity ratios were presented for the two metering rate conditions, as illustrated in Figure 11.



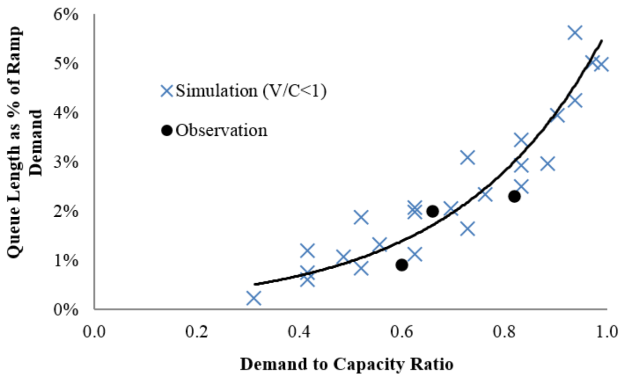
(a) Low demand scenario ***(b) High demand scenario***

Figure 10 Queue length as percentage of ramp demand

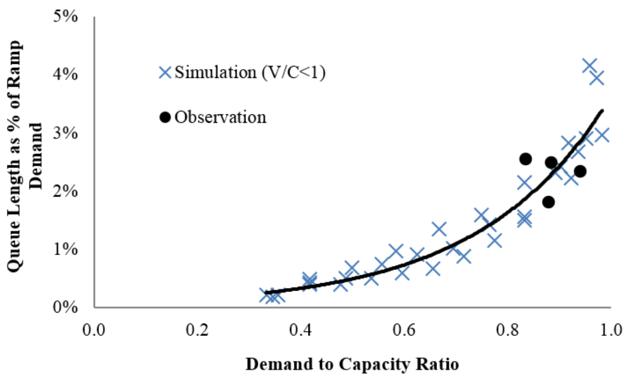
Simulation results show that the general queue length profiles of the two metering rate conditions are similar. Regression equations were developed to figure out the relationship between queue length and demand-to-capacity ratio. It was found that for under-saturated scenarios, the queueing process at a metered on-ramp tends to be more stochastic; ramp queue is mainly caused by the vehicle platoons released from the upstream intersection, and the exponential function could best capture the relationship between queue length and demand-to-capacity ratio. While for over-saturated scenarios, the simulated queue length tends to increase linearly with demand-to-capacity ratio.

Simulation versus Observations

To further verify the accuracy of the developed mesoscopic queue length simulation model, a total number of seven sets of queue length data were collected at four existing metered on-ramps that have different demands, metering rates, and signal timing schemes. For each observation, cycle-by-cycle on-ramp demand and the maximum queue length of each cycle was documented. Based on these maximum queue lengths, the 95th percentile queue length during the peak hour period was determined. Field observed queue length data points are compared to the generated queue length profiles. As the demand-to-capacity ratio of the field observations are less than one, only under-saturated scenarios are presented, as shown in Figure 12.



(a) Low demand scenario



(b) High demand scenario

Figure 11 Field observed queue length as percent of on-ramp demand versus simulation result

Comparison results show that the majority of observations are close to the developed queue length profile, which indicates that the mesoscopic queue length simulation model provided accurate queue length estimations for metered diamond interchanges. The deviations were mostly from the random nature of on-ramp flow arrive profile. Since each observation or simulation is an individual event, thus queue length at a specific time point may be affected by a variety of factors (e.g., an unexpected vehicle platoon entering the ramp in a short period tends to produce a longer queue than normal condition). While the simulation method aims to use multiple runs to eliminate this randomness at aggregate level, and provide a universal queue storage length design guidance.

QUEUE STORAGE LENGTH DESIGN

For design purpose, it is necessary to keep the demand less than capacity; therefore, the corresponding on-ramp demands at a two-lane ramp could be determined as: low demand scenario when demand is less than 1000 vph, and high demand scenario when demand is between 1000 and 1800 vph. Then, with the generated regression equations, the queue length as a percent number of on-ramp demand are calculated for each demand-to-capacity situation, as summarized in Table 1.

Simulation results show that at metered on-ramps and under the most challenging situation (i.e., $V/C = 1$), queue length is approximate 6.2 percent of on-ramp demand for low demand scenario ($V < 1000\ vph$), or 4 percent for high demand scenario ($1000 < V < 1800\ vph$). The percent numbers are much lower than these numbers when the demand-to-capacity ratio is below 0.8. This indicates that when the demand-to-capacity ratio is greater than 0.8, queue length tends to increase at a more significant manner. These percent numbers could be used for metered diamond interchange queue storage length design. The first step is to estimate the on-ramp demand based on the empirical data; after that, an acceptable demand-to-capacity ratio should be determined according to either state-level highway design manual or engineering judgement. Then, queue length as a percent number of on-ramp demand could be identified through Table 3 in the revised manuscript. Finally, the required queue storage length could be designed as the percent number multiplied by the on-ramp demand.

CONCLUSION

In real world conditions, on-ramp feeding traffic flow varies cycle by cycle; hence, simply using an average peak hour demand cannot provide accurate queue length estimation. Based on mesoscopic simulation, it was found that for under-saturated scenarios, queue length shows an exponential increasing trend with demand-to-capacity ratio; while for over-saturated scenarios, the queue length tends to increase linearly with demand-to-capacity ratio. Finally, this paper recommends that the maximum required queue storage length at a metered on-ramp was approximately 6.2 percent of on-ramp demand when demand was less than 1000 vph (500 vphpl), or 4 percent when demand was between 1000 and 1800 vph (500 and 900 vphpl).

Table 1 Queue length as percent of on-ramp demand under various demand-to-capacity ratio situations

V/C Ratio	Queue Length as Percent of Ramp Demand	
	Low Demand Scenario	High Demand Scenario
0.4	0.7%	0.3%
0.5	1.0%	0.5%
0.6	1.4%	0.8%
0.7	2.0%	1.2%
0.8	2.8%	1.7%
0.9	4.0%	2.6%
1.0	6.2%	4.0%

ACKNOWLEDGEMENTS

This study was sponsored by the California Department of Transportation (Caltrans). The authors thank Dr. Zhongren Wang of Caltrans Headquarter for his constructive comments, and Dr. Ben Wang of UNR for help with coding the mesoscopic simulation model.

REFERENCES

Cambridge Systematics, Inc. (2001). *Twin Cities Ramp Meter Evaluation – Final Report*. Available: http://www.camsys.com/pubs/ramp_meter_final_report.pdf

Chaudhary, N.A., Tian, Z., Messer, C.J. and Chu, C.L. (2004). *Ramp Metering Algorithms and Approaches for Texas*. Report FHWA/TX-05/0-4629-1. Texas Transportation Institute, College Station, Texas. http://tti.tamu.edu/documents/0-4629-1.pdf

Newell, G. F. (1982). *Applications of Queuing Theory*, 2nd ed. Chapman and Hall, New York.

Tian, Z., C. Messer, and K. Balke. (2004). Modeling the Impact of Ramp Metering Queues on Diamond Interchange Operations, *Transportation Research Record*, No. 1867, pp. 172-182.

Wang, Z.R. (2013). Queue Storage Design for Metered On-Ramps. *International Journal of Transportation Science and Technology*, Vol.2(1), pp. 47-64.