# Optimal Ramp metering with SWARM parameter optimization

Bhargava Rama Chilukuri (Corresponding author)

Georgia Institute of Technology

Department of Civil and Environmental Engineering

Atlanta, GA 30332-0355

Email: bchilukuri3@gatech.edu

Jorge A. Laval, Ph.D.

Georgia Institute of Technology

Department of Civil and Environmental Engineering

Atlanta, GA 30332-0355

Email: jorge.laval@ce.gatech.edu

Angshuman Guin, Ph.D.

Georgia Institute of Technology

Department of Civil and Environmental Engineering

Atlanta, GA 30332-0355

Email: angshuman.guin@ce.gatech.edu

# Abstract

This paper presents a micro-simulation based optimization framework to generate optimal parameter values for the SWARM ramp metering algorithm. The framework includes a micro-simulation model that replicates freeway capacity drop phenomena and a Genetic Algorithm based optimization module that integrates with the micro-simulation. A real-world corridor was simulated with Georgia DOT’s video detection system data and manual counts with the objective to minimize total travel time. The analysis indicated that four parameters are critical for SWARM performance and that they are sensitive to geometric and traffic variables on a corridor. The results indicated that when compared to the SWARM performance with default values, optimal values reduced travel time by over 11.6%.

# Introduction

Ramp meters are used to control the vehicular flow entering the freeway. The primary benefits of the ramp metering come from preventing capacity drop and exit blockage (Papageorgiou and Kotsialos 2002). Other benefits of ramp metering include diversion to alternate routes, enhanced safety, and reduced emissions.

Researchers have developed several ramp metering algorithmsIn most deployments, ramp metering systems have been reported to be successful in reducing congestion and increasing safety. They have increased mainline throughput, lowered congestion and provided significant travel time savings and higher travel time reliability on freeways. However, in the majority of cases, the full potential of ramp metering systems have not been exploited and the associated benefits have not been realized.

Fortunately, recent advances in traffic flow theory and simulation allow us to (i) envision new strategies that can cope with the above drawbacks, and (ii) perform realistic simulations of these strategies.

To date, the main challenge of effective ramp metering strategies is to ensure system-wide coordination. Cassidy [1, 2] identified the potential drawbacks of current strategies and showed in the field that a new class of traffic-responsive strategies can restore high flows in the freeway even after the onset of congestion. Laval and Daganzo [3] proposed a lane-changing theory, and with the help of a “first of its kind” simulation model accurately predicted traffic dynamics under ramp metering. The current study is the first to take advantage of these new strategies to develop optimal ramp metering strategies.

The objective of this study is to develop optimal system-wide freeway ramp metering strategies taking advantage of recent advances in traffic flow theory and simulation. These strategies are to be incorporated within the Georgia Department of Transportation (GDOT) System Wide Adaptive Ramp Metering (SWARM) deployment framework. This framework allows the automated control of a set of freeway entrances in the metropolitan Atlanta region. The success of this deployment, however, is subject to the ramp metering strategies and parameters being supplied to SWARM. The accuracy of the parameters depends heavily on the accuracy of the identification and modeling of the bottlenecks in the network. Thus, specific objectives of this research are:

* Build a framework required to develop optimization strategies that are to be incorporated within the System Wide Adaptive Ramp Metering (SWARM) deployment framework to minimize network travel-time.+
* Develop optimal strategies using the simulation based optimization framework and demonstrate the usefulness of these strategies in an offline fashion

The Phase I project report (TO 02-49; RSCH PROJ 07-22) presented a detailed review of popular ramp metering algorithms. This report presents a review of the SWARM algorithm, its field implementation, Genetic Algorithm (GA), and case studies of GA based optimizations as used by transportation researchers and engineers.

## SWARM

The System Wide Adaptive Ramp Metering (SWARM) control was developed by National Engineering Technology (NET) Corporation [4]. The original development was based on CALTRANS District 7’s ramp metering unit input and recommendations [5]. It was first implemented in Orange County (District 12) and later in Los Angeles and Ventura Countries (District 7) [6]. In 2001 and 2002, SWARM was tested on I-210 and I-405 in California. In 2006, it was again tested on I-210 and is expected to be tested on Routes 10 and 710 in the future. Portland deployed SWARM to replace previous pre-timed algorithms in 2005. Ahn et al. [7] carried out field experiments to compare traffic conditions before and after implementation of SWARM in this deployment. Zhang et al. [8] used a microsimulator called Paramics [9] to evaluate four algorithms including SWARM.

### Algorithm

Like other heuristic coordinated control algorithms, SWARM control operates at two levels: global control and local control, which are named SWARM 1 and SWARM 2 respectively. The two algorithms are independent, with SWARM 1 being a forecasting and system-wide apportioning algorithm using a forecasting methodology. SWARM 2 includes two local traffic-responsive ramp metering algorithms, SWARM 2a and SWARM 2b [7].

In SWARM, a freeway network is divided into continuous segments, whereby each segment contains one bottleneck. There may be one or more on-ramps and off-ramps in one segment. Both the local and global metering rates are computed for every time interval for each ramp, and the more restrictive one out of the two is adopted in the field.

The global control algorithm, SWARM1, forecasts future traffic conditions of the bottleneck based on immediate past traffic data. This algorithm focuses on solutions to prevent real-time density from exceeding saturation levels for each segment. To forecast future density around the bottleneck, SWARM1 performs a linear regression on immediate past traffic data and applies a Kalman filtering process [10] to capture non-linearity. The extent of past data used and time-span into the future forecast, are tunable parameters. Once the future density is obtained, excess density (the portion of future density above the pre-determined threshold density that always refers to saturation level at the bottleneck) can be calculated to get the target density for the next metering cycle using the following formula:

Target Density = (Current Density) - (1/Tcrit )\*Excessive Density

where Tcritis the forecasting time-span.

The volume reduction for each detector station is determined as follows:

Volume Reduction= (Local Density - Target Density)\*number of lanes\*Distance to next ramp

The forecasting logic is shown in Figure 1.



Figure 1: SWARM Control Logic (6)

The volume reduction (positive if local density is greater than target density) is distributed to upstream on-ramps within the system according to the tunable parameters “Intersection Propagation Factor” and “Intrasection Propagation Factor”.

The local control algorithm SWARM 2 operates based on local traffic conditions near each ramp. Essentially it is a local traffic-responsive algorithm. SWARM2a uses a density function to compute metering rates based on headway theory. It attempts to maintain headway at the detector station upstream of the metered ramp by optimizing density to maintain maximum flow. SWARM 2b introduces a concept of storage zone on the on-ramps. The number of vehicles storing within the user defined storage zone will be calculated. Then, SWARM 2b computes metering rates to maintain demand such that a desired level of service is maintained as long as possible. This algorithm needs accurate loop detector data to perform efficiently.

### Field Implementation Case Studies

#### Los Angeles and Ventura Counties’ Experience

CALTRANS (10) compared benefits of new SWARM algorithms with pre-timed and local traffic-responsive ramp metering algorithms carried out before. It was found that the combined strategy (global control and local control) increased mainline speed by 11% during morning peak, cut down travel time by 14%, and reduced freeway delay by 17%. As trade-off, lengths of on-ramp queue at the 9 busiest ramps increased by 40% [11].

#### Portland Experience

Oregon Department of Transportation (ODOT) implemented SWARM on their freeway system in the Portland metropolitan region in May 2005, replacing the previous pre-timed ramp metering. Communication between loop detectors and the traffic management center was found to be a critical problem in implementing SWARM which required a high communication bandwidth. Large amounts of data sometimes caused communication failures and loss of data, which consequently affected the availability of accurate data. During the research, it was observed that the conditions (detector failure rates) were different during the implementation of pre-timed metering when compared to that during SWARM implementation. There was a 2% detection failure rate during pre-timed strategy implementation as compared to almost 10% detection failure rate at some locations, during SWARM implementation.

To collect *before* data, SWARM system was turned off and pre-timed ramp metering was implemented from June 19 to June 23. From June 26 to June 30, SWARM was turned on and *after* data was collected. The study [7] found that under SWARM operation, the VMT increased by 0.8% in morning peak hours. Vehicle hours of travel (VHT) and average travel-time increased by 6.0% and 5.1% respectively. Correspondingly the total freeway delay increased by 34.7%. The significant increase in freeway delay was due to either higher metering rate at on-ramps or diminished flow through the bottleneck. The contribution of diminished flow could not be determined due to the absence of information on bottleneck discharge. However, analysis indicated that traffic flow through the freeway was higher during the morning peak period under SWARM operation. SWARM also admitted slightly higher flows at most on-ramps which decreased travel time on the ramps. Unfortunately, detailed examination of the performance of the parameters was not conducted in this study. It can be postulated that the higher SWARM metering rate might have been an artifact of poor data quality resulting from communication failures. Otherwise, it might have been caused by suboptimal parameter values used in SWARM. Nonetheless, the study did not make any attempts to explain the reasons for the observed behavior.

#### Simulation Experience

The research conducted by Zhang et al. [8] compared the effectiveness of four types of ramp metering algorithms using Paramics: ALINEA, Bottleneck (MBTN), SWARM and Zone. Particularly in the evaluation of SWARM, researchers adopted ALINEA to be the local control algorithm and used an ARX model to replace the linear regression and Kalman filtering process in the original algorithm for forecasting traffic trend. Besides, the study tested two types of time-span, one with one time-step (30 seconds) and the other with five time-steps (150 seconds), which were named MSWARMI and MSWARMV respectively.

Interstate 405 in Orange County, CA, was selected as the study site. This study network had data from typical morning and afternoon peaks and included interchanges. The simulation was conducted for two hours. Three levels were simulated to represent different traffic demand patterns: light traffic, moderate congestion and heavy congestion. Total vehicle travel time (TVTT), which was defined as the sum of all vehicles’ O-D travel times during the simulation time, was adopted as the Measure of Effectiveness (MOE) to evaluate the overall performance.

For light congestion level (Level 1), MSWARMI achieved the best performance, 2.5% reduction of TVTT from the no-control case. Statistical tests showed that the TVTTs of the four control algorithms were lower than no-control and there were no statistically significant differences between performance of ALINEA, MBTN, and Zone algorithms. However MSWARMV did not reduce TVTT. For moderate traffic demand (Level 2), all five algorithms significantly reduced TVTT from no-control. MSWARMI reduced about 3.3% of TVTT. But its performance was poorer than MBTN and Zone algorithms, which obtained about 5% reduction each. For heavy traffic demand (Level 3), MSWRMI reduced about 6.4% of TVTT from no-control. Statistical tests indicated that there were no significant differences among the ALINEA, MBTN, MSWARMI, and Zone algorithms. However, MSWARMV was poorer than the other four algorithms. It reduced TVTT by only about 3.5%. These simulation results indicated that ramp metering might be more effective under heavy traffic demand. As traffic demand increases, ramp metering tends to be more effective in reducing system travel time.

Unfortunately, the study did not test the sensitivity of SWARM algorithm comprehensively. MSWARMI often performed better than MSWARMV under the three scenarios examined in this study. These findings indicate that forecast size is a critical factor in SWARM algorithm, since it affects the accuracy of prediction which in turn affects the performance of SWARM.

## Genetic Algorithm

Genetic Algorithms (GA) belong to a class of evolutionary algorithms that mimic natural selection process. They are popularly used for searching optimal solutions for combinatorial problems. Though these algorithms were used in some elementary form since 1950, they were made popular through the studies of John Holland at University of Michigan during 1970s [12]. Their reduced susceptibility to getting stuck with local minima compared to other solution searching algorithms such as gradient search methods made them very attractive. However, these algorithms could be computationally expensive depending on the shape of the fitness landscape.

### Algorithm

GA can be broadly described as starting with an initial solution subset, evaluating its performances, and selecting the subsequent subsets based on the performance of current solution set. Thus, this algorithm can be represented as shown in Figure 2.

Create initial population set (Generation i=0)

Evaluate the objective function for Generation i

Select the best performing individuals and create Generation i+1

Replace Generation i with Generation i+1

Figure 2: GA Logic

Generally, the initial population set is selected from the solution set using a random process. However, the size of population set is a user defined parameter. Literature qualitatively states that for larger solution space, larger population set is desired. Once the population set of generation I is evaluated, it’s best performing individuals are used to determine the next generation using two process; crossover and mutation as shown in Figure 3.

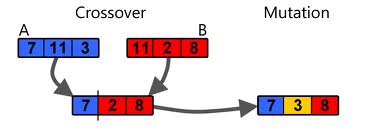


Figure 3: Crossover and Mutation

Crossover is a process of selecting two or more well performing individuals and producing a new individual by selecting the parameter values from the parents. Thus, the crossover mechanism guarantees preserving the good features of the parent. Mutation generates a parameter value that is from neither of the parents. Thus mutation mechanism helps avoid getting trapped in local minima and randomly generates new solutions. The GA process is generally stopped based on a termination criterion or after arriving at a satisfactory solution.

There are a number of specific attributes of GAs that give them an edge over other traditional optimization techniques. These are:

1. GA works with a subset of solution space (not a single point) to avoid being trapped at a local optimum.
2. Working on a pool of solutions allows GA to have a parallel implementation architecture.
3. GA does not need the derivative of the objective function to determine the optimum. It only needs that the objective function be able to be evaluated for any individual point from the solution space.
4. Abu-Lebdeh and Benekohal [13, 14] showed that as the size of the solution space increases exponentially, the time requirements for the GA grow only linearly.

GAs have limitations as well:

1. Due to the randomness inherent to GA and the dependence of GA parameters on the type of problem, analytical analysis of GA is not possible.
2. GA is slower than methods that work with derivatives.
3. Global optima is guaranteed only with appropriate parameters.
4. Terminal criteria is not the same for all GA problems. Some of the commonly used criteria are number of generations, generational rate of improvement or computational time.

Note that the speed of convergence to the global minima/maxima is primarily dependent on the logic used for selecting the subsequent subsets based on the performances of the current subset (crossover and mutation). More details on how GA is implemented in this study is described in section 3.3. The next section describes some of the GA applications in traffic engineering.

### Field Implementation Case Studies

Researchers have used GA for limited traffic engineering applications. Most of the GA applications have been in the area of traffic signal control of isolated intersections and network of intersections [15, 16]. Researchers have implemented GA for both offline and real-time estimation of control parameters: split, cycle, and offset. One should note that in combinatorial problems of multiple variables with large allowable range of values, the solution space will become very large making it challenging to converge to global optima in a reasonable amount of time. Therefore, researchers have often defined rational boundaries for the allowable parameter values to make the problem manageable. Table 1 shows the GA parameters as used by various researchers.

Some researchers have used GA for short term traffic state prediction [17, 18, and 19]. They used GA in conjunction with neural networks to predict state variables such as flow and occupancy values on freeway sections. They have also used these methods for short-term travel time prediction based on time series data from neighboring sections. Other researchers have used GA for Dynamic Traffic Assignment [20] and traffic safety [21]. Thus, GA method has been used for solving other traffic problems, but was never used in the area of SWARM ramp metering. This is the first study to implement simulation based optimization method for estimating optimal parameters for SWARM ramp metering algorithm.

Table 1: GA Parameters from literature

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Authors | Population Size | Crossover | Mutation | Generations |
| Sadek et al. | 30 | 0.25 | 0.02 | 1000 |
| Ceylan and Bell | 40 | 0.5 | 0.02 | 100 |
| Khosravi et al. | 10 | 0.7 | - | 40 |
| Girianna and Benekohal | 500 | 0.8 | - | 500 |
| Baher et al. | 300 | - | - | 30 |
| Liu et al. | 50 | 0.8 | 0.05 | 1000 |

Table 1 shows that there is no one set of GA parameter values that work for all traffic related applications. Therefore, researchers explored different sets of parameters to ensure GA converges to global optima.

# Optimization Framework

The simulation based optimization framework for determining optimal SWARM parameter values is shown in Figure 4.

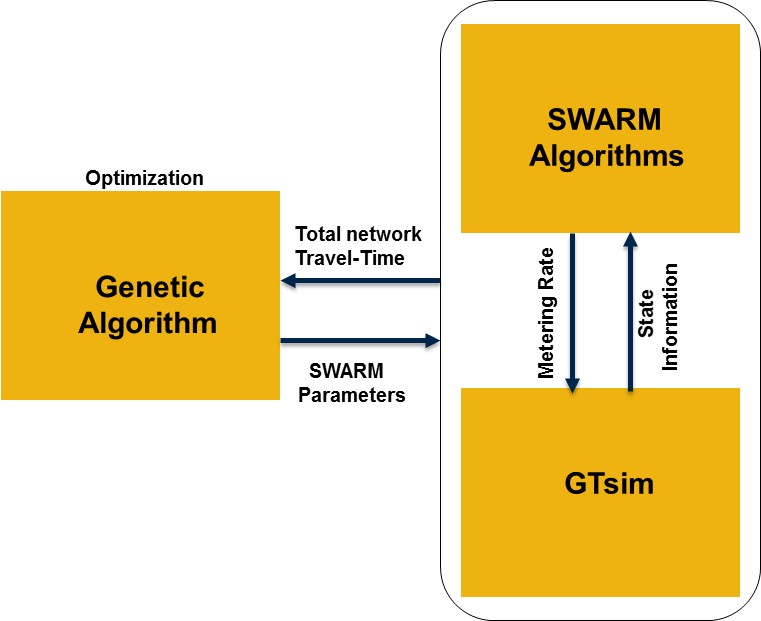


Figure 4: Optimization Framework

Two main components of this framework are the GA based optimizer and GTsim-SWARM module. The optimizer will provide a set of parameter values that are utilized by the GTsim-SWARM module to estimate the total travel time that will be sent back to the optimizer. The GTsim application will provide continuous state information to the SWARM algorithm that calculates metering rates based on the state information and the parameters provided by the GA based optimizer. The sections below describe GA based optimizer, GTsim, and SWARM algorithm as implemented in this study.

## GTsim Application

It is well known that the existing off-the-shelf traffic simulation software are deficient in simulating congested traffic dynamics on freeways. This is mainly due to insufficient inbuilt models and limited flexibility to allow users to change the models. Georgia Tech has developed a micro-simulation application, called hereafter GTsim, which includes the latest advancements in lane changing models that are capable of explaining congestion dynamics such as capacity drop [3]. GTsim was built in JAVA to perform faster than real-time simulation. For example, a 3-hr simulation of a 5 mile corridor during peak congestion could be simulated in under 3 minutes on a 2.66 GHz processor with 2GB RAM.

Figure 5 shows the process flow for the GTsim application. The classes shown on the left form the critical modules of the application. The methods shown in the center are some of the critical functionalities of the modules on the left. Lastly, the supporting classes provide some additional functionality for the modules.

The framework for GTsim is setup such that every time the simulation is initiated, it creates an instance of simulation which calls various modules during every time-step to update each vehicle’s velocity, location, and animation. Thus this framework allows writing new methods for any future models to replace the old models and facilitate easy implementation of new models in the simulation.

## SWARM Application

The SWARM high level design documentation provided by GDOT (73) consisted of a family of algorithms. Each algorithm performs a specific task whose results are used by other algorithms. Thus, the SWARM application is developed as a supplementary module that integrates all the different interdependent algorithms as shown in Figure 10. The rest of this section will describe some of the important modules of the SWARM application:



Figure 10: Process Flow for the SWARM Application

**Minimum Rate Computation**

Minimum metering rate for each ramp within the segment is calculated dynamically for every 15 minutes time period using the 15 minute average ramp volumes and ramp length. This algorithm requires a tunable parameter that represents the average space occupied (required) per vehicle in feet that is used in the minimum rate computation.

**Saturation Density Generation**

Saturation Density (SD) represents the operational capacity of a vehicle detector station. SD is a key parameter required to be determined for the all vehicle detector stations in each segment. The SWARM algorithm calculates the metering rates for each ramp such that saturation density is maintained at the all vehicle detector stations and bottlenecks. SD is computed by measuring the density of the vehicle detector station when the station volume is the greatest. Over a large sample this computed value of Saturation Density represents the maximum value of the parabola formed, with the y-axis being Volume and the x-axis being Density. Smaller values represent congested conditions (or low volume free flow conditions) and larger values represent near capacity conditions.

**SWARM 1 Local Minimum and Maximum Metering Rates**

SWARM 1 algorithm determines metering rates in a multi-step approach. The first step is to determine the minimum and maximum rates for each ramp. However, the minimum and maximum rates for a given ramp vary depending on the type of mode applied at the ramp; i.e., SWARM1 or SWARM2, or SWARM1 and SWARM2. The range of rates at a given ramp is always bounded by the Absolute Minimum Rate and the Absolute Maximum Rate.

**SWARM 1 Forecasting and Required Density**

All the entrance ramps of a segment that are located upstream of the bottleneck are controlled by SWARM as a Bottleneck Set. The algorithm continuously monitors the density at the bottleneck and uses a linear regression approach to estimate future conditions which is called forecasting. The 20-second density values are used to estimate the density at a time in the future. How far into the future (Tcrit) the algorithm will estimate and how many historic data points are to be used (Forecast Size) for the regression analyses are tunable parameters. If the forecasted density exceeds the saturation density before the Tcrit minutes, then the required reduction in the density is estimated. A new forecasting is made every 20 seconds and a new density reduction is computed.

The following four different types of Densities are used by SWARM to determine the metering rates at all the ramps:

Current Density (CD): Represents the density at the vehicle detector station every 20 seconds.

Saturation Density (SD): The saturation density values are computed for each vehicle detector station and are used as target density values attempted to be attained by the algorithm.

Forecast Density (FD): The future density is forecasted every 20 seconds for each designated Bottleneck and is called the Forecast Density (FD).

Required Density (RD): This is determined based on the CD and SD for the bottleneck location and is a key value for apportionment. RD represents the maximum density that is aspired at each vehicle detector station and is always between SD and SD/2. RD is computed as follows:

* If CD <= SD and FD <=SD then RD=SD
* If CD <= SD and FD > SD then RD = SD – (FD-SD)/Forecast Lead Time
* If CD > SD and FD <=SD then RD = SD
* If CD > SD and FD > SD then RD = SD – (FD-SD)

Once the RD value is determined for the bottleneck in a given segment, the rates for all the metered ramps are computed independent of all constraints. The initial metering rates computed for each ramp can be negative or positive and are computed as follows:

Desired Metering Rate = AbsMax - (CD x N – RD x Nbn) x D – Excess

Where

AbsMax = The Absolute Maximum Metering Rate at the ramp

N = The number of lanes at the ramp

Nbn = The number of lanes at the bottleneck

D = The distance to the next downstream ramp

Excess = the number of vehicles that were not handled from downstream ramps. Excess can be positive; vehicles were not handled by the downstream ramp or negative; extra room is available between the downstream ramp and the current ramp.

The Final Metering Rate for each ramp is computed by enforcing the boundary conditions of Local Minimum and Local Maximum rates.

**SWARM 1 Apportionment**

In the SWARM 1 algorithm the metering rates are computed starting at the furthest downstream ramp of the segment and proceeds upstream until the furthest upstream ramp is reached. Excess is the value that is used to allow SWARM 1 to adapt to areas of congestion by propagating upstream those vehicles that cannot be handled at each ramp in those areas. Excess starts at zero at the first ramp at the furthest downstream ramp. After the final SWARM 1 metering rate is computed a new value of Excess is computed to be used at the next upstream ramp as follows:

Excess = (Final Metering Rate – Desired Metering Rate) x Distance between ramps. ()

**SWARM 2 Metering Rate**

SWARM2 algorithm uses data from the vehicle detector stations located adjacent and downstream of the metered ramp. The average density is used to lookup the metering rate from a local table. However since the rate table was not available during this study, ALINEA was implemented in GTsim to determine local metering rates. Local metering rates are used to ensure avoid any local merge bottlenecks at each ramp location. This metering rate for each ramp is bounded by the local minimum and the absolute maximum rates. Finally if SWARM1 and SWARM 2 modes are chosen along a section, the most restrictive of the SWARM1 Final Metering Rate and SWARM2 metering rates is chosen for each ramp.

## Genetic Algorithm

Since this is the first implementation of GA for optimization of SWARM parameters, the team tried different variants of GA in terms of methods and parameter values to determine the best methodology and GA parameter values for the current problem.

When the solution space is large, GA parameter values play an important role in converging to the global optimum. Inappropriate values will result in GA converging to local optima and not reach global optima in a reasonable amount of time. When the termination criterion is based on convergence and not an explicit terminal value, validating the minima obtained by GA becomes challenging. One of the ways to differentiate local minima from global minima is to choose a different initial population and GA parameter values to check if it converges to the same minima. The following subsections describe the GA components and the parameters used in this study.

**Population**

A GA starts its search from a population of possible solutions. How this initial population is created and its size is largely dependent on the problem type. Literature shows that a randomly created initial population is not a bad start unless the user has some prior knowledge of the solution location in the search space. Based on trial and error different population sizes were used depending on the purpose. For determining impact parameters (defined in chapter 4) a population size of 600 was used. For evaluating optimal parameters for the impact parameters, a population size of 150 was used.

**Tournament Size**

Tournament size is used for creating a tournament pool from which new individuals are created for the next generation. Typically the tournament pool is created by random sampling from the population. Sometimes, Roulette-wheel selection is used to accelerate convergence to the minimum. In Roulette-wheel selection tournament pool is created by prioritizing better performing individuals (better performing individuals have higher probability for selection). After implementing Roulette-wheel selection in the initial stages of this study it was found that GA was quickly converging to local minima and did not attain global minima. Therefore, the Roulette-wheel selection was discarded for the later research. In this study a tournament size of 30 was used for determining impact parameters and a tournament size of 10 was used for optimal value selection.

**Cross Over**

Crossover is used to create new individuals from parents. What pieces from the parents are transferred to the “child” is decided by crossover probability. In this study a crossover probability of 0.5 was used to give equal chance to both parents.

**Mutation**

Mutation is used to randomly explore new areas in the solution space. The probability of mutation gives the likelihood that a given bit of an individual will be mutated. In this study a mutation probability of 0.15 was used for determining impact parameters and a value of 0.4 for optimal value selection. A higher value for mutation was used to ensure GA explores the entire solution space.

**Termination criteria**

The termination criterion is critical for determining when the GA iterations will be stopped. Since the target minima is not known for the current problem, incremental improvement in performance over successive generations is used to determine when the GA generations would be stopped. For determining impact parameters, we ran GA for 75 generations. However, while determining the optimal values for impact parameters, it was found that GA converged to global minima after 42 generations. Therefore, for all subsequent analyses, at least 50 generations were used to determine the optimal values.

# Methodology

## Study corridor

The I-285 EB/SB corridor between GA-400 and I-85 was selected for this study. This chapter describes different types of data needed for this study and data collection methodology used in this study.

Based on the data analysis performed during the Phase I of this project (TO 02-49; RSCH PROJ 07-22), a 6.5 miles long EB/SB I-285 segment between GA-400 and I-85 was selected for this study. The study corridor has the following five entry locations (called origins for the OD terminology) to feed traffic to the network:

* Upstream Freeway
* Peachtree Dunwoody Road
* Ashford Dunwoody Road
* North Peachtree Road
* Peachtree Industrial Pkwy

Note that the Bufford Hwy on-ramp initially merges with the SB I-85 off-ramp and later connects to WB I-285 outside the boundaries of the study area. Therefore, this on-ramp was not considered as an origin in this study.

The corridor has the following eight exit locations (called destinations) as shown in figure 11:

* Ashford Dunwoody Road
* Chamblee Dunwoody Road
* SB Peachtree Industrial Pkwy
* NB Peachtree Industrial Pkwy
* Buford Hwy
* SB I-85 connector
* NB I-85 connector
* Downstream Freeway

Based on the congestion characteristics analysis performed as a part of the Phase I project, this analyses focus on the congested evening peak period. Within the 6.5 -mile study corridor, there are 21 GDOT detection stations that collect velocity, volume and occupancy data. All the four on-ramps on the corridor were metered and the volume data at these ramps were polled out over the TACTICS framework. However, initial observation of the TACTICS data and the data obtained from GDOT’s video detection system (hereafter referred to as the NaviGAtor data) revealed that these systems were recording unreasonable ramp volumes. To verify this, three ramps (two off-ramps and one on-ramp) from the test corridor were video recorded during evening peak period and the volumes were manually counted to compare against the NaviGAtor and TACTICS data.

Results indicated that there are significant discrepancies in vehicle count observations and the data reported by these stations. The quality of the data was not suitable for use in a micro-simulation based travel-time prediction framework. Therefore, based on discussions with GDOT, the research team proceeded with manually collecting data for the rest of the study while GDOT worked towards improving the quality and availability of data from the TACTICS framework. GDOT’s PTZ cameras were used to record all the entry and exit points on the corridor and the videos were processed manually to extract the time-series of traffic volumes on the study corridor. The next section describes the traffic volume data collection procedure in more detail.

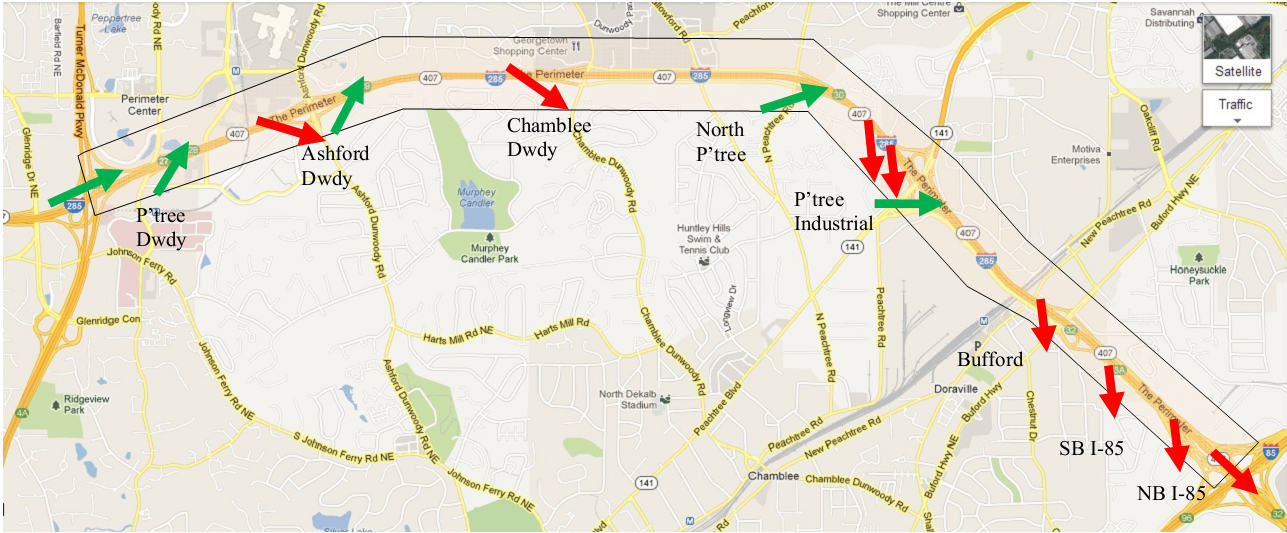


Figure 11: Case study corridor (courtesy www.google.com)

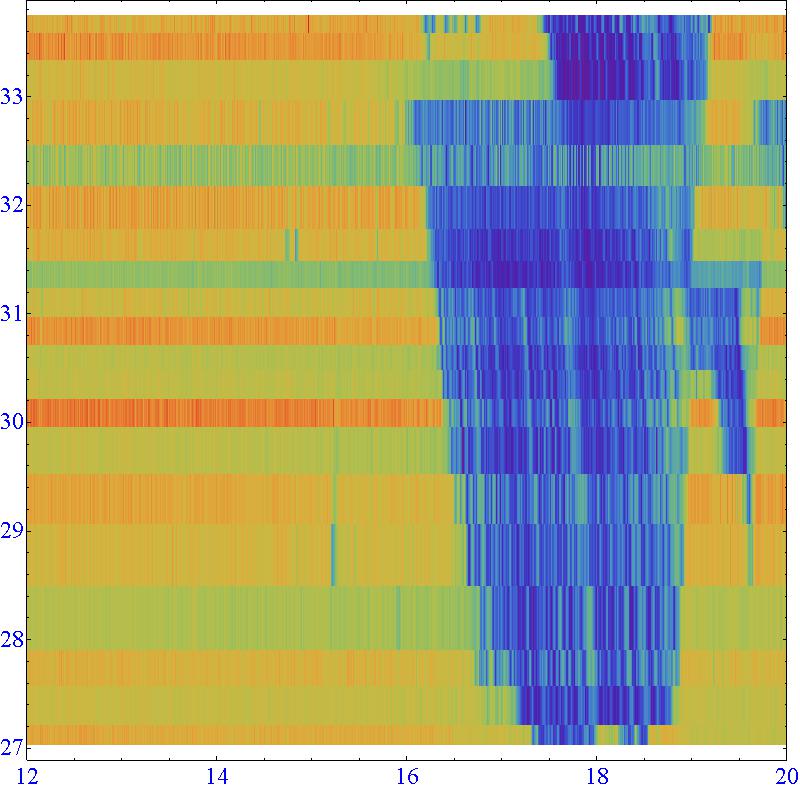
## Traffic Volume Data Collection

For this study thirteen PTZ cameras were used to record the five entry locations and eight exit locations from 14:00 to 20:00 on Wednesday, March 7, 2012. Time-series of traffic volumes were manually extracted from the videos using the tablet based traffic counting application developed at GaTech [24]. Oblique curves were plotted and extensively examined to verify the quality of the data. To minimize chances of data loss one of the researchers was stationed at the GDOT’s Traffic Management Center during the data collection period to ensure that the PTZ cameras maintained the view required for the counts. While most of the cameras recorded for the 6 hours, one of the cameras only recorded from 15:00 to 17:00. Therefore, this study was performed for only the duration that all six required cameras were continuously available.

It was observed that during the study duration the corridor was initially free flowing and the starts to get congested as shown in Figure 12a. The figure showed that the corridor got congested initially due to spillback from the NB I-85 off-ramp and later due to weaving and merging at various locations. By the end of the study duration, the congestion on the freeway encompasses all the ramps in the study corridor.

It was observed that the Peachtree Industrial Pkwy on-ramp dumped a lot of traffic on the freeway (2-lane ramp). However, due to its proximity to the I-85 interchange, the queue spill back from I-85 ramps disrupted the smooth merging of the ramp traffic and resulted in a queue downstream of the ramp meter that sometimes spilled to the ramp meter upstream. Moreover, huge inflows from the ramp and close proximity of the off-ramp downstream (Buford Hwy off is 0.5 miles downstream) significantly increased lane changing activity in this vicinity.

The study corridor was modeled in GTsim and was calibrated and validated using the Bluetooth travel time data as a part of the earlier project Travel Time Estimation and Prediction (GDOT research project 10-01; TO 02-60). The same simulation model was used for this project.



15 17 19

NB I-85

SB I-85

Bufford

PtreeInd

CbleeD

AshDdy

PtreeD

Figure 12: a) Time-space speed plot of the VDS data

The corridor was simulated for 90 minutes (15:00 to 16:30) that included congestion build-up and completely congested conditions. The OD flows were automatically updated every 5 minutes. The simulation continuously stored the flow and velocity information at detector stations prescribed in the model. SWARM algorithm modules use the SWARM parameters sent by the GA optimizer and model state information every 20 seconds to determine metering rate at every on-ramp. The SWARM metering rates are implemented in the simulation model. The Total travel time, defined as the sum of all vehicles’ OD travel times during the simulation time, was adopted as the Measure of Effectiveness (MOE) to evaluate the performance of a set of SWARM parameters.

## Critical Parameters

The SWARM system uses more than 20 parameters to control the operation and performance of the algorithm based on the specific deployment. While some of the parameters are required for data processing and data smoothing, other parameters are used in the core SWARM algorithms. The “status threshold” parameters that are related to quality of the real-time data are excluded from this study. Tables 2 show a list of 14 parameters examined in this study.

MaxPoints: This is used by the data smoothing algorithm to calculate the average flow and speed at a detector station for a desired time period.

Min/Max Vol, Min/Max Speed, Min/Max Den: These parameters are used as the thresholds to update the saturation density of the bottleneck.

Saturation Density: This is used to calculate “excess” flow forecasted at the bottleneck.

Q: Bottleneck saturation density is updated using a weighted average of “old” saturation density and “new” density that satisfies the Min/Max Volume, Min/Max Speed, and Min/Max Density thresholds.

Table 2: SWARM Parameters studied for optimization in this study

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Algorithm** | **Variable Name** | **Description** | **Range** | **Default Value** |
| Data Smoothing | MaxPoints | Sample size of points to average | 3-15 | 6 |
| Saturation Density Calculation | Min Vol | Minimum volume threshold | 4-6 veh | 5 veh |
| Max Vol | Maximum volume threshold | 8-12 veh | 10 veh |
| Min Speed | Minimum speed threshold | 40-56 kph | 48 kph |
| Max Speed | Maximum speed threshold | 58-81 kph | 72 kph |
| Min Density | Minimum density threshold | 9-28 vpkm | 22 vpkm |
| Max Density | Maximum density threshold | 28-56 vpkm | 50 vpkm |
| Saturation Density | Initial saturation density | 22-41vpkm | 55 vpkm |
| Q | Sat. Density Smoothing parameter | 0-1 | 0.02 |
| SWARM1 Forecasting and Required Density | DenSampSize | Number of polls in the past to use for the forecast | 3-51 | 30 |
| ForecastLead Time | Number of polls into the future to forecast | 3-42 | 15 |
| SWARM1 Apportionment | Intersection propagation factor | Propagate “excess” between ramps | 0-1 | 0.85 |
| SWARM2 algorithm | MaxUpRate | Maximum rate at which the metering rate can increase per lane | 20-900 vph | 300 vph |
| MaxDownRate | Maximum rate at which the metering rate can decrease per lane | 20-900 vph | 60 vph |

DenSampSize: SWARM algorithms use linear regression to forecast density at the bottleneck. Historic data is used for the prediction model. This parameter describes the number of historic data points that needs to be used for forecasting. This is an important parameter since an excessively large sample size will result in smoothing out the short-term patterns, and an excessively small sample size will magnify the impact of local fluctuations.

ForecastLeadTime: The forecasting model estimates the density at the bottleneck and determines the metering rate based on the predicted traffic congestion levels. The amount of time into the future that the algorithm needs to forecast is an important parameter, since an excessively short time period into future will result in frequent rate changes, and an excessively long time period will be inefficient.

Intersection propagation and Intrasection propagation factors: Once the density at the bottleneck is forecasted and required density is determined, SWARM calculates metering rate starting at the downstream ramp closest to the bottleneck. Excess values are calculated for each ramp and are transferred to the upstream ramp using the Intersection propagation factor. Similarly, excess is transferred across section using the Intrasection propagation factor.

UpRate and DownRate: The metering rate determined every 20 seconds by SWARM algorithms are bounded by these values.

## Strategy Evaluation Plan

The strategy development and evaluation is an iterative process. Table 3 shows the number of allowable values of each of the 14 parameters evaluated in this study which corresponds to a total of 212,468,465,664,000,000 combinations of parameter values.

Table 3: SWARM Parameters and the number of allowable values for each parameter

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Parameter | maxpoints | DenSampSize | ForecastLeadTime | UpRate | DownRate | InterSectionPropagationFactor | MinVol |
| Range | 3-15 | 3-51 | 3-42 | 20-900 | 20-900 | 0-1 | 4-6 |
| #of values | 13 | 49 | 40 | 16 | 16 | 20 | 3 |
| Parameter | MaxVol | MinSpeed | MaxSpeed | MinDen | MaxDen | q | SatDen |
| Range | 8-12 | 48-65 | 56-81 | 9-28 | 28-56 | 0-1 | 22-41 |
| #of values | 5 | 18 | 26 | 20 | 29 | 20 | 20 |

Determining the optimal set of parameter values from this huge set is extremely time consuming even with any sophisticated search algorithm. Therefore, parameters that have a significant influence on the SWARM performance, here after called impact parameters, have to be identified and separated from the rest (here after called no-impact parameters).

Separating impact parameters from no-impact parameters was done using GA. Since the solution space was very large, an initial population of 600 was assumed and a high mutation rate of 0.15 was used and the GA was run for 75 generations that converged to a minima. It was observed that out of the 45,000 individuals evaluated, 28,713 individuals (parameter value combinations) produced the same travel time. Then, those 28,713 individuals were examined for the repeating parameter values. It was observed that only 4 parameters had one value each for all the 28,713 individuals and the other 10 parameters accepted all possible values. This indicates that the following 4 parameters can be categorized as the impact parameters:

1. UpRate,
2. DownRate,
3. SatDen,
4. InterSectionPropagationFactor

These 4 impact parameters were optimized as described in the next chapter to determine their optimal values. However, the above hypothesis was validated by performing a sensitivity analysis of the no-impact parameters as explained in the later part of the next chapter.

# Results and Discussion

This chapter describes the results of the GA based analysis and the performance of optimal parameters compared to other strategies.

## Optimal Parameter Values

Based on all possible values for the 4 impact parameters, the solution space was still very large (102,400). Therefore, the optimal values of the 4 impact parameters were determined using a different set of GA parameters. The values of no-impact variables were set to be the default parameters. A population of 150 and a mutation rate of 0.4 were used and the GA was run for 90 generations. The 15th percentile values of all the generations are plotted in Figure 13 that shows the convergence of GA over generations.

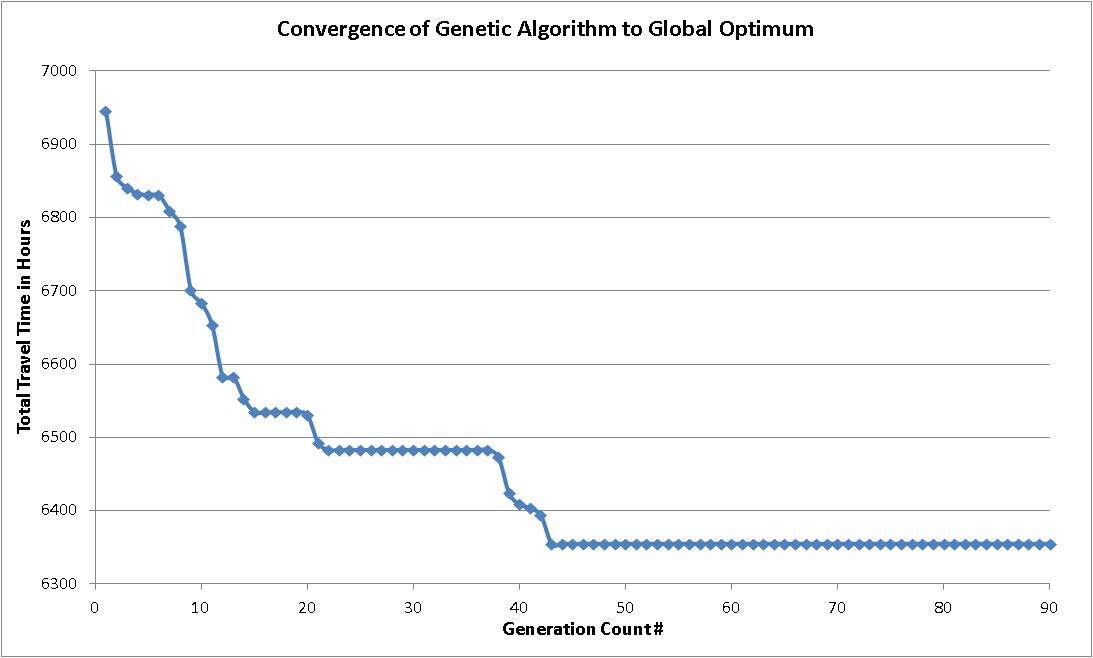


Figure 13: GA Convergence to global optimal over successive generations

It can be seen from figure 13 that after 42 generations, the algorithm converged to a minima. It was found that there were two sets of parameter values as shown in Table 4 that resulted in the same total travel time.

To confirm that it is the global minima, all the combinations of allowable values of impact parameters were simulated. It was found that the minima obtained from GA after 42 generations was indeed the global minima. Thus, it was confirmed that the GA parameters used above will converge to global minima.

Table 4: Optimal SWARM Parameter values with queue flush

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Parameter | maxpoints | DenSampSize | ForecastLeadTime | UpRate | DownRate | InterSectionPropagationFactor | MinVol |
| Set 1 | 6 | 30 | 15 | 780 | 720 | 0 | 5 |
| Set 2 | 6 | 30 | 15 | 780 | 720 | 0.1 | 5 |
| Parameter | MaxVol | MinSpeed | MaxSpeed | MinDen | MaxDen | q | SatDen |
| Set 1 | 10 | 48 | 72 | 22 | 50 | 0.02 | 22 |
| Set 2 | 10 | 48 | 72 | 22 | 50 | 0.02 | 22 |

Inter-section Propagation Factor:

Note that the only difference between the two sets is the InterSectionPropagationFactor that determines how excess demand is distributed between ramps. This is implemented in SWARM algorithm by raising the InterSectionPropagationFactor to the power of distance between the ramps. With distance between the a congested ramp merge and the upstream free flowing ramp being as high as 3 miles (between Ashford Dunwoody ramp and North Peachtree onramp) low values of InterSectionPropagationFactor ( in this case 0 and 0.1) may have similar impact.

Optimization function dependence:

The optimal parameter values could be obtained for any objective function. Next, GA was used to find optimal parameters to minimize freeway travel time. It was found that the optimal parameters in Table 4 hold true and resulted in minimal freeway travel time. However, when the objective was changed to minimizing the total number of on-ramp queue-flushes at all the four on-ramps, it was found that optimal parameter values were different. Thus, it is evident that the optimal values for the impact parameters are dependent on the desired objective function.

Sensitivity Analysis:

Only 4 parameters need to be optimized based on our earlier finding that 10 parameters do not impact the performance. This was verified as follows. Multiple simulation runs were performed starting with the optimal values for four impact parameters and default values for no-impact parameters, and changing one parameter value at a time. While one parameter value is changed other parameter values are set to be optimal values or default values as appropriate. Thus, a total of 206 runs were performed to check the sensitivity of non-optimized parameters. As expected, it was found that that the total travel time remained same for all the runs. This showed that default values perform adequately for the 10 non-optimized parameters.

Queue Flush:

In the rest of the report the results of the minimal total travel time are examined. Figure 14 shows the results of the 45th generation. It can be seen that optimal values performed much better than the default values for all the three measures shown in the plot. In this study, an on-ramp queue flush strategy was implemented (using actual thresholds from the field) to replicate real-world constraints. Queue-flush resulted in continuously clearing the queue on the ramp limiting the variation in ramp travel time across various strategies (ranges between 39-43 veh-hrs).

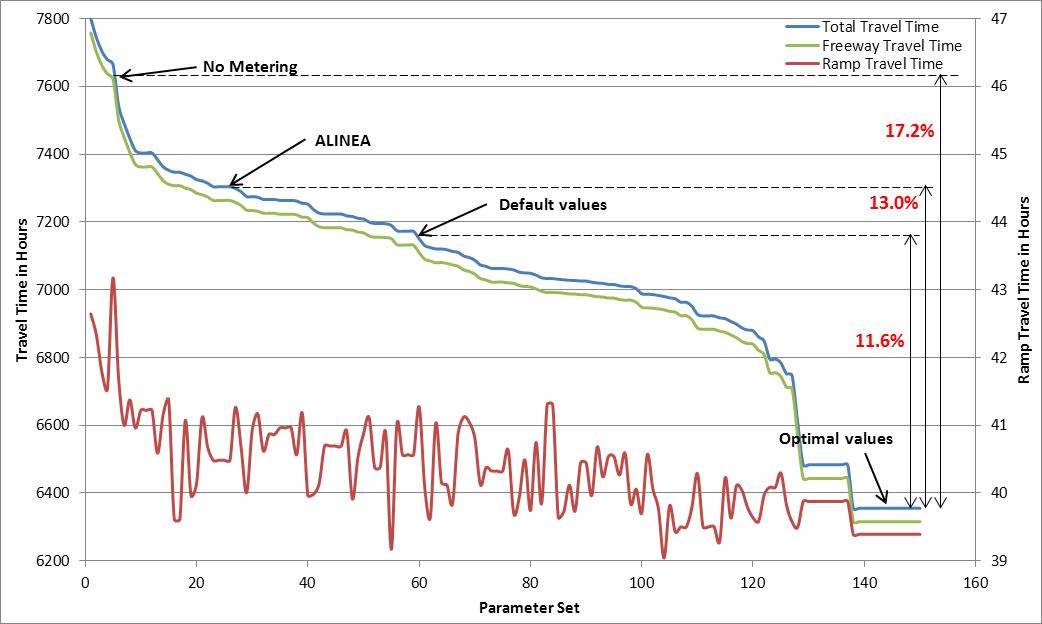


Figure 14: Sample results of the 45th generation of GA optimization

Comparison with other metering strategies:

Table 5 shows a comparison of performance of various strategies. The results validate that both ALINEA and SWARM improve both system and freeway travel times compared to a no-metering scenario. However, due to the queue flush mechanism implemented in this study, ramp travel times do not increase significantly in spite of ramp metering. The results indicate that with the optimal parameters, the system travel time benefits could be as much as 17%.

Table 5: Travel time reduction (veh. hrs) comparison of SWARM (with queue flush) with other strategies

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Comparison with No-Metering | | | Comparison with ALINEA | | | Comparison of SWARM with default values | | |
|  | System | Ramp | Freeway | System | Ramp | Freeway | System | Ramp | Freeway |
| ALINEA | 375  (4.9%) | -14.2  (-54.3%) | 389  (5.1%) | - | - | - | - | - | - |
| SWARM w/default parameters | 491  (6.4%) | -13.7  (-52.2%) | 505  (6.6%) | 116 (1.6%) | 0.5 (1.3%) | 115 (1.6%) | - | - | - |
| SWARM w/optimal parameters | 1324  (17.2%) | -13.2  (-50.1%) | 1337  (17.5%) | 949  (13.0%) | 1.1  (2.7%) | 948  (13.1%) | 833  (11.6%) | 0.6  (1.4%) | 832  (11.7%) |

The results indicate that SWARM performs better than ALINEA even with default values. The total travel time reduction could be as much as 13%. Finally, optimal parameters result in travel time reduction of almost 12% compared to default parameters, in this case study.

Delay Savings:

Figure 15 shows the delay savings time-series of optimal values and ALINEA. If the existing system implemented in the field is comparable to ALINEA, delay savings of over 8 minutes will be possible with optimal values. This is significant savings considering that the corridor is only 6 miles long and the average travel time observed on the corridor is around 25 minutes. Notice that the delay savings increase as the corridor gets more and more congested before the flow break down. The 1-minute smoothed data shows sustained benefits of over 6 minutes as the corridor gets congested. However, once the corridor is completely congested, the delay benefits reduce indicating that ramp metering benefits are more prominent during congestion build-up phase when the ramp metering avoid/postpone capacity drop.

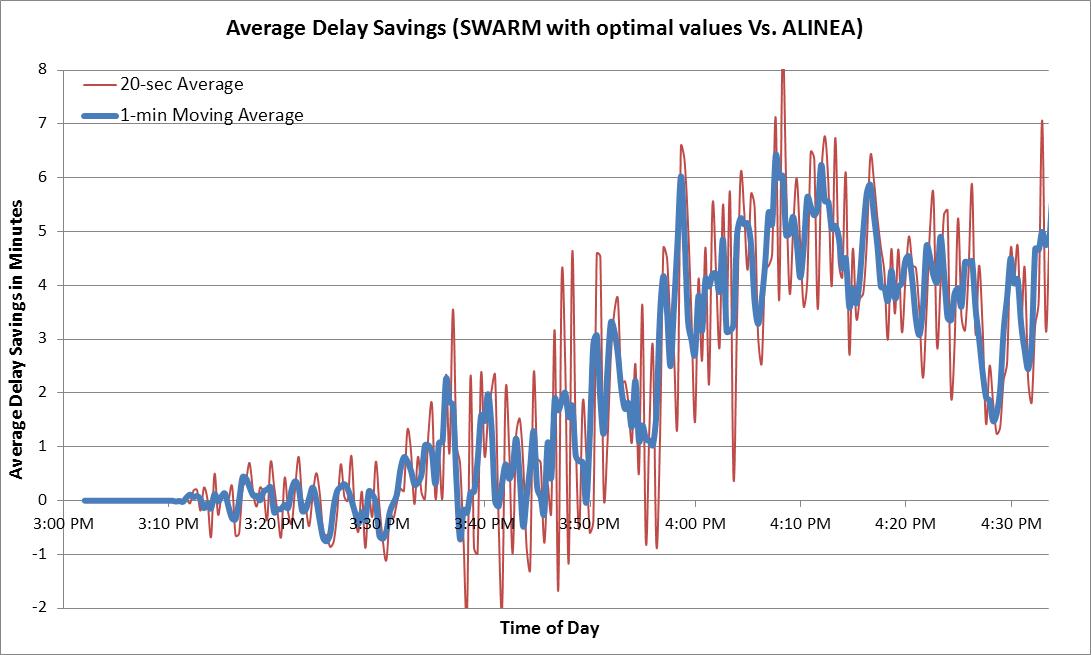


Figure 15: Delay savings with optimal values compared to ALINEA

Figure 16 shows the delay savings time-series of optimal values and default values. The results show delay savings of over 7 minutes will be possible with optimal values. Similar to comparison with ALINEA earlier, in this case too optimal parameters are more influential during congestion build-up phase compared to congested conditions.

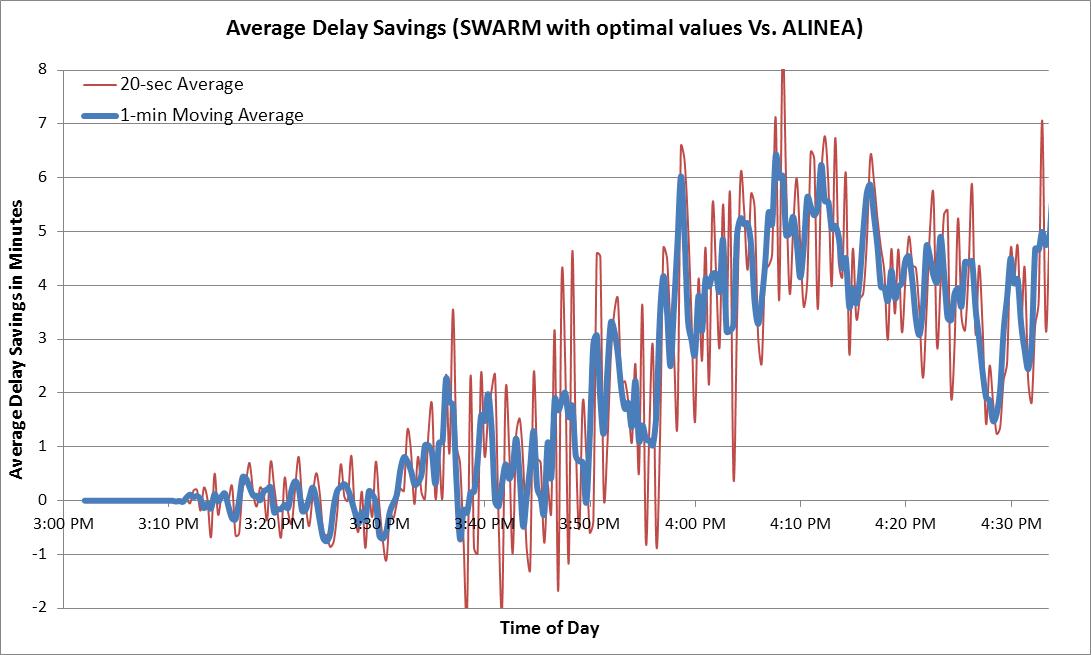


Figure 16: Delay savings with SWARM using optimal values compared to SWARM using default values

Figure 17 shows the average travel time time-series of different strategies. The lines in the Figure indicate that the average travel time values with optimal parameters deviates from the rest just before the corridor gets congested. The deviation is significant indicating that optimal values greatly reduce the travel time, by as much as 20%.

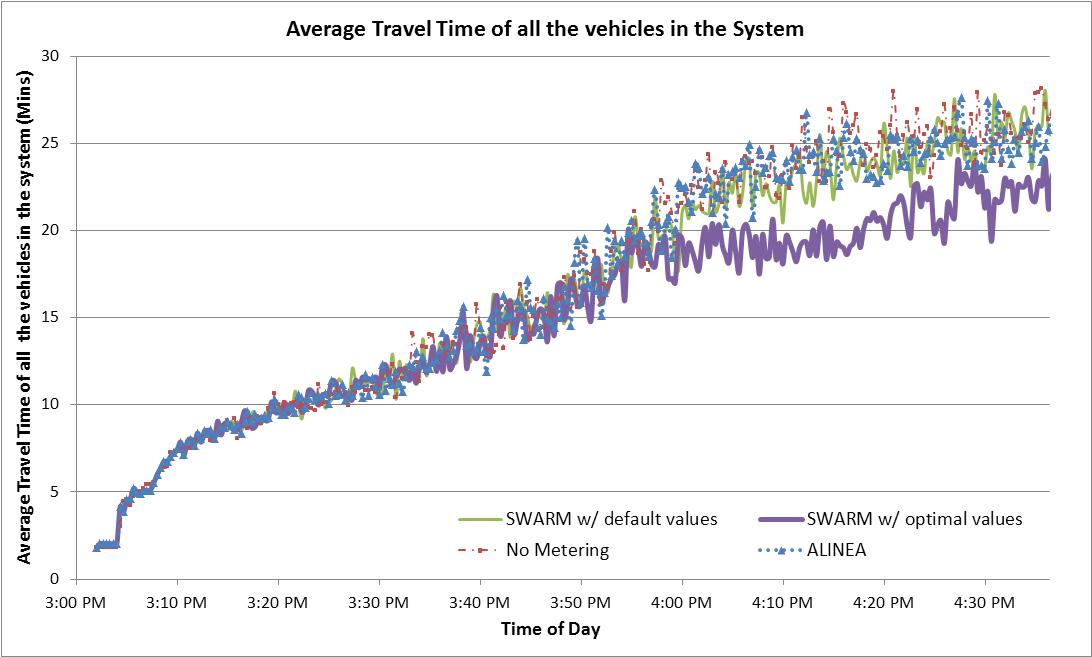


Figure 17: Average Travel Time with different strategies

Queue Flush:

It is expected that the number of flushes at a metered ramp will be inversely proportional to the storage area and directly proportional to the difference between the cumulative arrivals (or inflow rate) and departures (or metering rate if meter not blocked and ramp lane departure flow rate otherwise) at the meter. Table 6 shows a summary of queue-flush with optimal and default values.

Table 6: Queue-flush Analysis

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Average Queue clearance duration (min) | Optimal Values | | | | Default Values | | | |
| # of Flushes | # of times meter blocked by d/s queue | Max flush duration (min) | Total flush duration (min) | # of Flushes | # of times meter blocked by d/s queue | Max flush duration (min) | Total flush duration (min) |
| P'tree Dunwoody (storage: 480 ft.) | 1.33 | 30 | 0 | 1.67 | 38.92 | 30 | 0 | 1.67 | 33 |
| Ashford Dwdy (storage: 680 ft.) | 1.67 | 21 | 0 | 2.00 | 31.96 | 14 | 0 | 1.68 | 21.32 |
| N.Ptree (storage: 710 ft.) | 1.67 | 14 | 1 | 27.67 | 48.32 | 8 | 1 | 43.66 | 54 |
| Ptree Industrial (storage: 1000 ft.) | 2.33 | 6 | 2 | 91.33\* | 107.36 | 6 | 1 | 102.68 | 111.66 |

\*first instance was 7.67 minutes long and the second was 91.33 minutes long.

The results in Table 6 indicate that the number of flushes at the Peachtree Dunwoody and Peachtree Industrial ramps are the same for both default and optimal parameter values. Interestingly, the first and last flush happened around the same time, but in between the flushes happened at different times. At the Ashford Dunwoody and North Peachtree ramps, optimal parameters produced lower metering rates, that resulted in frequent spillback and subequently more number of flushes. Also, it was observed that at these two ramps, the first flush occurred much earlier into the simulation with optimal parameters (13 and 26 minutes) compared to the default parameters (16 and 27.66 minutes). However, the timing of last flush was same for both scenarios.

At the North Peachtree ramp, both optimal and default parameters resulted in the ramp meter being blocked by spillback from downstream freeway at around 44 minutes from the start of simulation. However, larger inflows from upstream ramps with default parameters meant that the North Peachtree ramp was congested for a longer duration (44 minutes versus 28 minutes) with default values.

At the Peachtree Industrial ramp, both optimal parameters and default parameters resulted in the ramp meter being blocked by spillback from downstream freeway at around 22 minutes from the start of the simulation. With default parameters, the blockage lasted for 103 minutes, from 22 to 125 minutes into the simulation. However, with optimal parametes the blockage started at 21 minutes into simulation, but lasted only 8 minutes. Then the blockage cleared for 2 minutes before blockage again reached the ramp and lasted for 91 minutes. This shows that the optimal parameters have a significant impact on the freeway operation.

Metering Rate Variations:

Figures 18 shows a comparison of the metering rates with ALINEA, default parameters and optimal parameters at different ramps. It can be seen that ALINEA occationally changed metering rate and made slow transition in changing metering rates. Reactive to that SWARM with default parameters was more sensitive to traffic conditions and changed the metering more frequently. However, the limits on uprate (300 vph) and downrate (60 vph) restricted the amount of change made at any time. However, SWARM with large uprate (780 vph) and downrate (720 vph) thresholds was much more sensitive to traffic conditions and changed metering rate frequently to obtain maximum benefits. However, note that the uprate and downrate values are the thresholds and does not mean a metering rate change of that magnitude during successive intervals.

Figure 18 also shows that both ALINEA and SWARM (default and optimal) resort to minimum metering rate at some time during the study period (that rate remain active for the rest of the simulation duration). This shows that none of the algorithms evaluated in this study can prevent congestion on the corridor. They will only manage the ramp flows to minimize total travel time.

Note that any simulation based optimization has noise associated with the random numbers used in the simulation that needs to be addressed. In traffic related applications, additionally traffic data variation also creates noise. The next two sections will evaluate the sensitivity of optimal parameters to the random number variation and traffic flow variation.

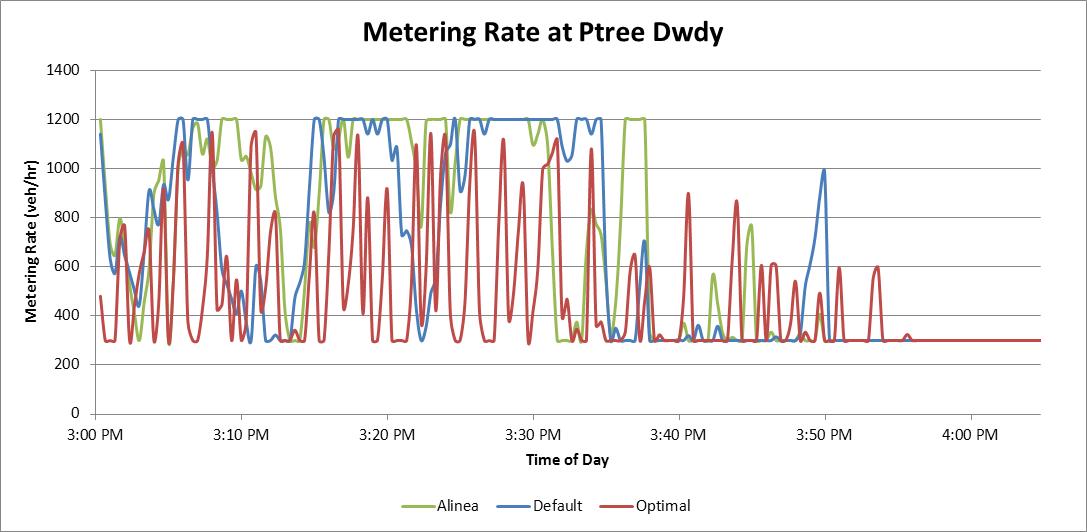
## Sensitivity to random number variation

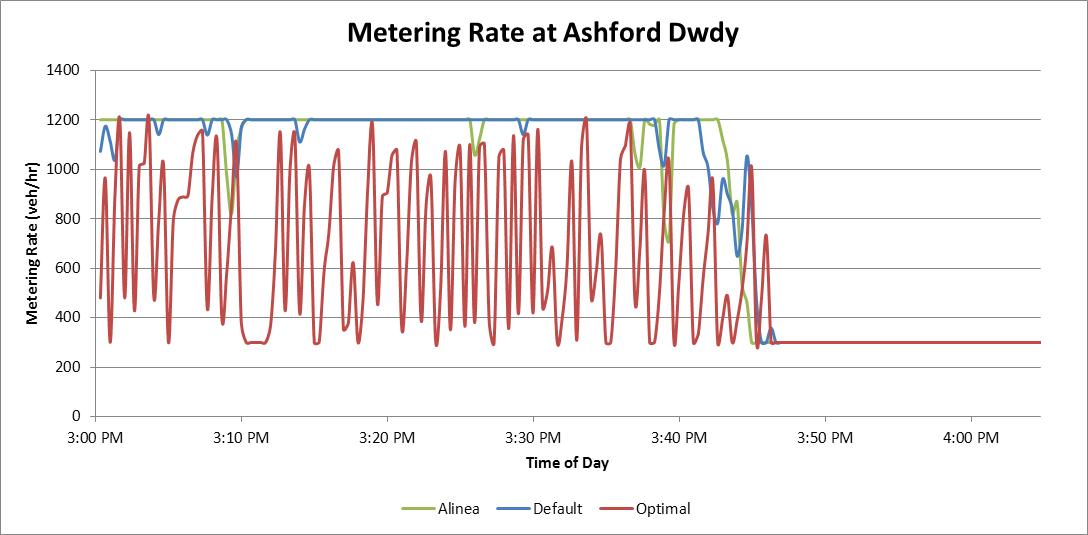
Just like other micro-simulation models, GTsim also uses several random numbers for vehicle destination determination, mandatory and discretionary lane changes, determine vehicle acceleration/deceleration etc. To verify if the global optimal values still hold true for different set of random numbers, sensitivity analysis was performed.

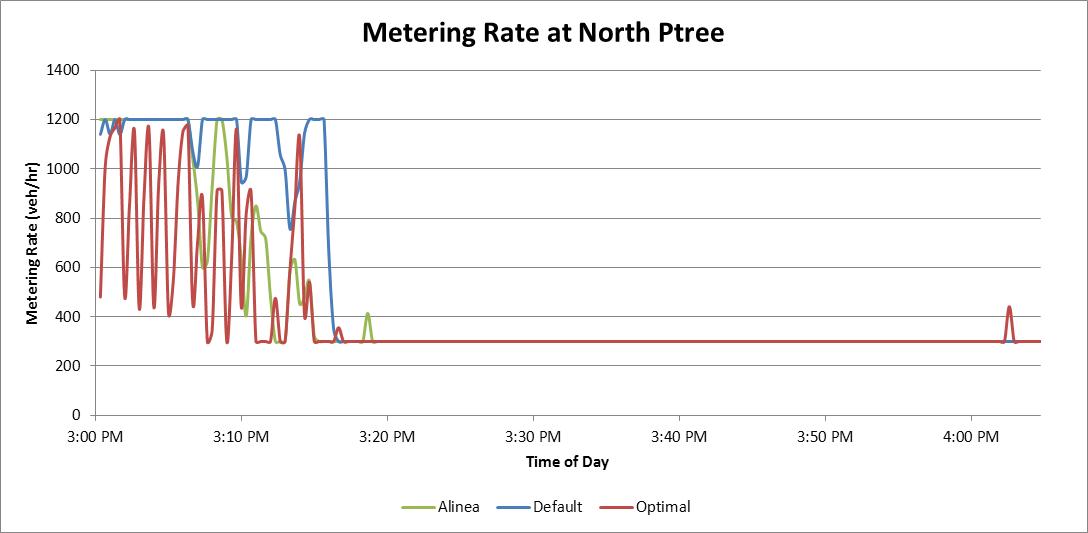
The random number seeds were changed and GA based optimization was performed using the GA parameter values used earlier to obtain global minima. After 50 iterations, it was found that the global optimal parameters still performed the best compared to other parameter combinations. However, the total travel time values varied in the range of + 0.3%. Therefore, the global optimum obtained is impervious to change in the random number seed and the results are stable.

## Sensitivity to traffic flow variation

The sensitivity of the optimal parameters to the day-to-day flow variations is also studied. To verify if the global optimal values still hold true for variations in OD flows, sensitivity analysis was performed. The OD flows are changed randomly by + 15% and GA based optimization was performed with the GA parameters used earlier. The GA was run for 50 generations and the results indicated that for some flow conditions GA converged to a different set of optimal parameter values. However, it was observed that the performance of the optimal values obtained earlier is comparable. If one were to use the earlier optimal values for various flow conditions, the elasticity of the performance degradation for flow variation was found to be 0.3% for every 1% change in flow.







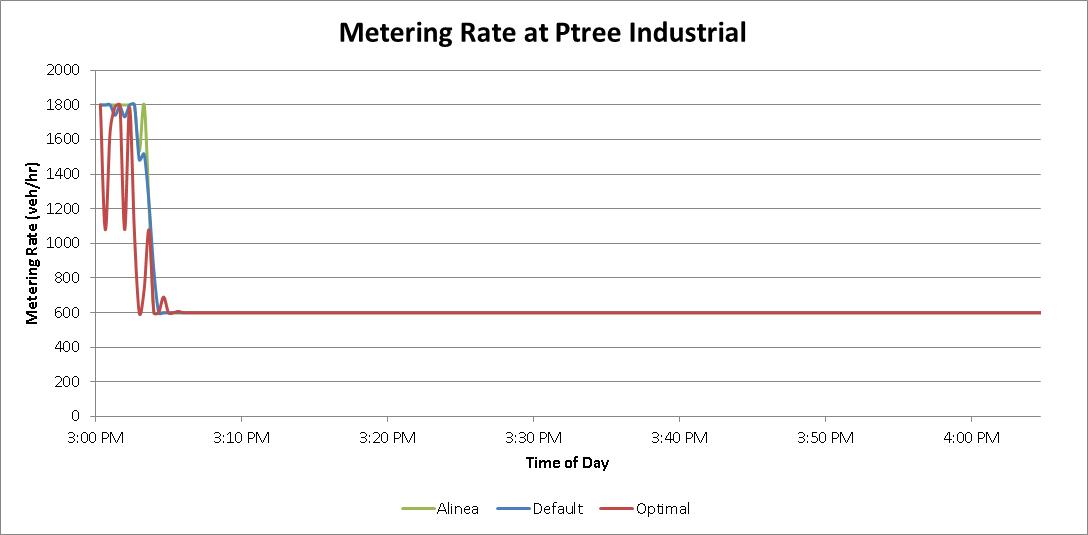


Figure 18: Comparison of metering schema at different ramps

# Conclusions and Recommendations

This study used a simulation based optimization framework to determine optimal parameter values the GDOT’s SWARM system to minimize the total travel time. As a part of this framework, the microsimulation model (GTsim), SWARM algorithms module, and a GA based optimization module were integrated. It was found that out of the 14 parameters evaluated in this study 4 parameters, Up rate, Down rate, Intersection propagation factor, and saturation density, were found to impact the performance of SWARM significantly. For other parameters, the default values performed satisfactorily. The optimal values derived from this study were found to reduce travel times by more than 17% compared to no-metering scenario. When compared to the SWARM performance default values, optimal values resulted in a reduction of travel time by over 11%. The delay savings of the optimal values was over 5 minutes which corresponds to about 20% compared to default values. These savings are significant considering that the corridor gets more or less completely congested by the end of the congestion period irrespective of what metering algorithm is used.

The optimal parameter values derived in this case study are location sensitive and need to be optimized for other location. However, the optimization framework developed in this study can be seamlessly used to generate optimal parameters for other locations.

Considering that all ramps operate at minimum rate after certain time regardless of the metering strategy used, it can be stated that most of the benefits of parameter optimization is realized during the congestion build-up phase. Moreover, it was observed at the Peachtree Industrial on-ramp and the North Peachtree on-ramps that downstream freeway queue blocked the ramp meters preventing flows smaller than metering rates to enter the freeway. This indicates that it may be beneficial to possibly turn-off the ramp meters at these locations once the congestion sets in. More analyses are needed in this regard.

Given the significance of travel time reduction and delay savings with the use of SWARM versus local ramp metering algorithm such as ALINEA, it is highly recommended that SWARM is implemented in the GDOT system. In addition, given the benefits of using the optimal parameters versus a default set of parameters, it is recommended that the optimization framework developed in this study be used for optimizing the parameters specific to each segment of the network to maximize the benefits of ramp metering.

Future research:

While this project has successfully identified optimal parameter values for SWARM algorithm, further research is necessary to continue to develop optimal strategies for efficient freeway operation. Some of the research directions are as follows:

* Optimize with NaviGAtor Data: Valid traffic data streams from the current ramp-metering system will enable identifying the realistic boundaries of the flow variations on the corridor. This data will help perform accurate sensitivity analysis of SWARM parameters. Also, NaviGAtor data will enable optimization of data quality related thresholds for the SWARM algorithms.
* Field implement SWARM on a test corridor: GDOT should consider field implementation of SWARM on a test corridor to address the deployment related and other practical issues. This will also enable a better understanding of SWARM’s field performance and help GDOT be prepared for system-wide deployment of SWARM.
* Optimize current ramp metering system: Using the optimization framework and the applications developed in this study, the operation of the current ramp metering system could be optimized for operation without the availability of a communication framework. New research will develop optimal Time-of-day schedules, threshold parameters and their values, metering rate table, and a better on-ramp queue flush methodology.
* Alternative Ramp Metering Strategies: Since it was found that some of the ramps operated at minimum metering rate irrespective of the ramp metering algorithm used, it may be beneficial to evaluate non-conventional and innovative ramp metering strategies such as meter shut-off after congestion sets. New research will evaluate the benefits of alternative/innovative ramp metering strategies and develop guidelines for their implementation.
* Ramp metering strategy evaluation tool: The GTsim microsimulation application developed in this study can be used as the backbone for developing a tool that will help GDOT personnel evaluate different ramp metering strategies such as changing metering rates, threshold parameters, and threshold values at various on-ramps on a corridor. GDOT personnel will be able to visualize the anticipated flows and queues on the freeways and ramps, effects of metering rates on ramp performance, flow at the merge locations, interaction of varied metering rates across different ramps on corridor performance, etc. This tool will be especially useful to GDOT personnel to efficiently manage freeways during incidents and special events.

# References

1. Cassidy, M.J. (2003) Freeway on-ramp metering, delay savings and the diverge bottleneck. Transpn Res. Rec. 1856, pp. 1-5.
2. Cassidy, M.J. and Rudjanakanoknad (2005) Increasing the capacity of an isolated merge by metering its on-ramp. Transportation Research B, 39 (10), pp. 896-913.
3. J.A. Laval and C.F. Daganzo. (2006) Lane-changing in traffic streams. Transportation Research Part B, 40 (3), pp. 251-264.
4. G. Paesani, J. Kerr, P. Perovich, and E. Khosravi (1997) ―System Wide Adaptive Ramp Metering In Southern California”, ITS America 7th Annual Meeting, June 1997
5. Report 3: System wide adaptive ramp metering algorithm - high level design, Final Report, Prepared by NET for Caltrans and FHWA, March 1996.
6. CALTRANS (office of freeway operations-Ramp Metering Branch), (2007). 2005/06 Ramp Metering Annual Report-District 7 Los Angeles and Ventura Countries.
7. Ahn, S, Bertini, RL, Auffray, B, Ross, JH, & Eshel, O. (2007). Evaluating Benefits of Systemwide Adaptive Ramp-Metering Strategy in Portland, Oregon. Transportation Research Record: Journal of the Transportation Research Board, 2012(-1), 47-56.
8. Zhang, M., Kim, T., Nie, X., Jin, W., Chu, L., Recker, W. (2001) Evaluation of On-ramp Control Algorithms. California PATH Research Report UCB-ITS-PRR-2001-36, Berkeley, CA.
9. Aitken, S. (2000). Quadstone Paramics Modeller User Guide. Scotland, UK.
10. Kalman, R.E. (1960). "A new approach to linear filtering and prediction problems". Journal of Basic Engineering 82 (1): 35–45.
11. Pham, H, Jreij, W, Otani, C, Kalkatechi, H, Toorawa, I, Kao, J, . . . Benitez-Lopez, R. (2002). SWARM Study Final Report on W/B Foothill Freeway (W/B LA-210). Research Report to California Department of Transportation.
12. Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence. Holland, John H. Oxford, England: U Michigan Press. (1975).
13. Abu-Lebdeh, G., and R. F. Benekohal. Convergence Variability and Population Sizing in Micro-Genetic Algorithms. Computer-Aided Civil and Infrastructure Engineering, Vol. 14, 1999, pp. 321–334.
14. Abu-Lebdeh, G., and R. F. Benekohal. Genetic Algorithms for Traffic Signal Control and Queue Management of Oversaturated Two-Way Arterials. In Transportation Research Record: Journal of the Transportation Research Board, No. 1727. TRB, National Research Council, Washington, D.C., 2000, pp. 61–67.
15. Ceylan and Bell, 2004 Traffic signal timing optimisation based on genetic al0067orithm approach, including drivers' routing Transportation Research B, 38B (4) (2004), pp. 329–342
16. Using Genetic Algorithms to Design Signal Coordination for Oversaturated Networks.  Girianna, Montty; Benekohal, Rahim F. Journal of Intelligent Transportation Systems. Apr2004, Vol. 8 Issue 2, p117-129. 13p.
17. Short-Term Traffic Flow Prediction Using Neuro-Genetic Algorithms.  Abdulhai, Baher; Porwal, Himanshu; Recker, Will. ITS Journal - Intelligent Transportation Systems Journal. Jan2002, Vol. 7 Issue 1, p3-41. 39p
18. Mingzhe Liu, Ruili Wang, Jiansheng Wu, Ray Kemp, "A Genetic-Algorithm-Based Neural Network Approach for Short-Term Traffic Flow Forecasting", Lecture Notes in Computer Science, Vol 3498, pp. 965-977, May 2005
19. A genetic algorithm-based method for improving quality of travel time prediction intervals, Khosravi, Abbas; Mazloumi, Ehsan; Nahavandi, Saeid; Creighton, Doug; Van Lint, J.W.C. Transportation Research: Part C. Dec2011, Vol. 19 Issue 6, p1364-1376. 13p.
20. Sadek, A. W., B. L. Smith, and M. J. Demetsky. Dynamic Traffic Assignment: A Genetic Algorithms Approach. In Transportation Research Record 1588, TRB, National Research Council, Washington, D.C., 1997, pp. 95-103.
21. Kunt, Mehmet Metin; Aghayan, Iman; Noii, Nima. Prediction for traffic accident severity: comparing the artificial neural network, genetic algorithm, combined genetic algorithm and pattern search methods.   Transport (16484142). Dec2011, Vol. 26 Issue 4, p353-366. 14p.
22. Newell, G.F. A simplified car-following theory: a lower order model, Transportation Research Part B: Methodological, Volume 36, Issue 3, March 2002, Pages 195-205.
23. Guin, A., M. Hunter, and R. Guensler. Analysis of Reduction in Effective Capacities on High-Occupancy-Vehicle Lanes Related to Traffic Behaviour, Transportation Research Board: Journal of the Transportation Research Board, No. 2065, Washington, D.C., 2008. pp 47-53.
24. Toth C., Suh W., Elango V., Sadana R, Guin A., Hunter M., Guensler R. (2013), Tablet-Based Traffic Counting Application Designed to Minimize Human Error, Transportation Research Board Annual Conference, Washington, D.C.