

## ABSTRACT

Title of Thesis: A Q-LEARNING BASED INTEGRATED  
VARIABLE SPEED LIMIT AND HARD  
SHOULDER RUNNING CONTROL TO  
REDUCE TRAVEL TIME AT FREEWAY  
BOTTLENECK

Weiye Zhou, Master of Science, 2019

Thesis Directed By: Professor Lei Zhang, Department of Civil and  
Environmental Engineering

To increase traffic mobility and safety, several types of active traffic management (ATM) strategies, such as variable speed limit (VSL) and hard shoulder running (HSR), are implemented in many countries. While all kinds of ATM strategies show promise in releasing traffic congestion, many studies indicate that stand-alone strategies have very limited capability. This paper proposes an integrated VSL and HSR control strategy based on a reinforcement learning (RL) technique, Q-learning (QL). The proposed strategy bridges a direct connection between the traffic flow data and the ATM control strategies via intensive self-learning processes, thus reduces the need for human knowledge. A typical congested interstate highway, I-270 in Maryland, U.S. is simulated using a dynamic traffic assignment (DTA) model to evaluate the proposed strategy. Simulation results indicated that the integrated strategy outperforms the stand-alone strategies and traditional feedback-based VSL strategy in mitigating congestions and reducing travel time on the freeway corridor.

A Q-LEARNING BASED INTEGRATED VARIABLE SPEED LIMIT AND HARD  
SHOULDER RUNNING CONTROL TO REDUCE TRAVEL TIME AT FREEWAY  
BOTTLENECK

by

Weiyi Zhou

Thesis submitted to the Faculty of the Graduate School of the  
University of Maryland, College Park, in partial fulfillment  
of the requirements for the degree of  
Master of Science  
2019

Advisory Committee:  
Professor Lei Zhang, Chair  
Professor Ali Haghani  
Professor Paul Schonfeld

© Copyright by  
Weiyi Zhou  
2019

## Acknowledgements

This work would not have been possible without the great institution of the University of Maryland. Go Terps! It is truly an honor of a lifetime to meet so many caring, loving, and inspiring colleagues, staff, and professors at this magnificent place.

I am especially indebted to my advisor, Dr. Lei Zhang, at the Civil & Environmental Department at the University. Dr. Zhang's caring attitude, in-depth knowledge, effective teaching, and continuous encouragement, along with financial support through a research assistantship, have enabled me to reach my goal. I would be at a loss without Dr. Zhang's guidance.

I would also like to express my sincere appreciation to Professors Paul Schonfeld and Ali Haghani for serving on my research thesis committee and offering me their valuable comments and encouragement on my research.

I want to thank all my teammates in Dr. Zhang's research group. Thank you for the ideas you have offered, the comments you have shared, and more than anything else, the friendship we have built.

Last, I would like to extend my heartfelt indebtedness to my parents. There are no words that can truly express my appreciation for their love. I love you!

# Table of Contents

ABSTRACT.....	1
Acknowledgements.....	ii
Table of Contents.....	iii
List of Tables .....	v
List of Figures .....	vi
Chapter 1 Introduction .....	1
1.1. <i>Background</i> .....	1
1.2. <i>Active Traffic Management (ATM)</i> .....	2
1.3. <i>Traffic Modeling Approach</i> .....	6
1.4. <i>Objective</i> .....	8
1.5. <i>Paper Organization</i> .....	9
Chapter 2 Literature Review .....	11
2.1 <i>Real World Applications</i> .....	11
2.2. <i>Variable Speed Limit Effects</i> .....	12
2.3. <i>Variable Speed Limit Strategies</i> .....	15
2.3. <i>Coordinated ATM Strategies</i> .....	18
Chapter 3 Methodology and Model Generation .....	22
3.1. <i>Traffic Congestion Causes and Effects</i> .....	22
3.2. <i>VSL &amp; HSR Control Theory</i> .....	24
3.2.1. <i>VSL Control Theory</i> .....	25
3.2.2. <i>HSR Control Theory</i> .....	25
3.3. <i>Basic Q-learning Strategy</i> .....	28
3.4. <i>Coordinated RL-based ATM Algorithm</i> .....	32
3.4.1. <i>State of the QL-Based ATM Strategy</i> .....	34
3.4.2. <i>Actions of the QL-Based ATM Strategy</i> .....	37
3.4.3. <i>Rewards of the QL-Based ATM Strategy</i> .....	39
3.4.4. <i>Parameters of the QL-Based ATM Strategy</i> .....	41
3.5. <i>Study Area Description</i> .....	43
3.6. <i>Simulation Model</i> .....	46

3.6.1. Online Simulation Model.....	46
3.6.2. Model Generation .....	48
3.6.3. Model Calibration and Validation .....	49
3.6.4. ATM Implementations .....	52
3.7. Parameters Setting.....	53
Chapter 4 Simulation Results.....	57
4.1. MOEs .....	57
4.2. Results for Non-control Scenarios .....	58
4.3. Results for Feedback-Based VSL Control Scenario .....	60
4.4. Result for QL-based VSL and QL-based VSL&HSR Control Scenarios .....	61
4.5. Summary of other findings .....	63
Chapter 5 Conclusion.....	65
5.1. Summary of the Research.....	65
5.2 Advantages of the proposed Algorithm.....	66
5.3 Deficiency of the Proposed Algorithm .....	67
5.4 Future Work .....	68
References .....	71

## List of Tables

TABLE 2-1: EXAMPLES OF REAL WORLD VSL APPLICATIONS IN U.S .....	14
TABLE 3-1. LEVEL OF QUEUE DELAY .....	36
TABLE 3-2. LEVEL OF QUEUE DELAY REDUCTION.....	36
TABLE 3-3. LEVEL OF DENSITY .....	36
TABLE 3-4. LEARNING PARAMETERS OF QL ALGORITHM.....	43
TABLE 3-5. COMPONENTS OF DTA SIMULATION MODEL (INITIAL MWCOG MODEL AND SUBAREA MODEL) .....	45
TABLE 3-6. COMPONENTS OF DTA SIMULATION MODEL (INITIAL MWCOG MODEL AND SUBAREA MODEL) .....	49
TABLE 3-7. SUPPLY SIDE PARAMETERS FOR SIMULATION MODEL .....	54
TABLE 3-8. SCENARIO PARAMETERS FOR SIMULATION MODEL.....	54
TABLE 3-9. NUMBER OF VEHICLES FOR DIFFERENT DEMAND TYPES.....	55
TABLE 3-10. VALUE OF TIME (\$) FOR DIFFERENT DEMAND TYPE .....	55
TABLE 4-1. CORRIDOR TRAVEL TIME SUMMARY.....	59
TABLE 4-2. BOTTLENECK SPEED SUMMARY .....	60
TABLE 4-3. SUMMARY OF ALL MOE.....	63

## List of Figures

FIGURE 3-1. (A) FUNDAMENTAL DIAGRAM OF FLOW-DENSITY RELATIONSHIP UNDER UNCONGESTED CONDITION; (B) FUNDAMENTAL DIAGRAM OF FLOW-DENSITY RELATIONSHIP WITH CAPACITY DROP; (C) FUNDAMENTAL DIAGRAM OF FLOW-DENSITY WITH VSL CONTROL; AND (D) FUNDAMENTAL DIAGRAM OF FLOW-DENSITY WITH HSR CONTROL.....	24
FIGURE 3-2. COMPONENTS OF REINFORCEMENT LEARNING. ....	28
FIGURE 3-3: (A) TRADITIONAL ATM ALGORITHM; AND (B) PROPOSED QL-BASED ATM ALGORITHM.....	33
FIGURE 3-4. REWARDS FOR DIFFERENT STATE IN THE QL. ....	40
FIGURE 3-5. TRAFFIC CONDITION AT I-270 SPUR (MORNING PEAK). ....	45
FIGURE 3-6. THE MESOSCOPIC SIMULATION NETWORK AND LOCATIONS OF TRAFFIC COUNTS. ...	48
FIGURE 3-7. CALIBRATION AND VALIDATION PROCESS FOR A DTA MODEL.....	50
FIGURE 3-8. (A) TRAFFIC VOLUME COMPARING BEFORE CALIBRATION; AND (B) AFTER CALIBRATION. (C) CORRIDOR TRAVEL TIME COMPARING BEFORE CALIBRATION; AND (D) AFTER CALIBRATION.....	51
FIGURE 3-9. STUDY AREA AND IMPLEMENTATION OF ATM CONTROL.....	53
FIGURE 4-1. CORRIDOR TRAVEL TIME FOR DIFFERENT SCENARIOS.....	59
FIGURE 4-2. BOTTLENECK SPEED FOR DIFFERENT SCENARIOS .....	60
FIGURE 4-3. SYSTEM TRAVEL FOR DIFFERENT SCENARIOS.....	63



## List of Abbreviations

<b>AADT</b>	<b>Average Annual Daily Traffic</b>
<b>AGBM</b>	<b>Agent-based Travel Behavior Model</b>
<b>AI</b>	<b>Artificial Intelligence</b>
<b>ATIS</b>	<b>Advanced Traffic Information System</b>
<b>ATM</b>	<b>Active Traffic Management</b>
<b>BMC</b>	<b>Baltimore Metropolitan Council</b>
<b>CATT</b>	<b>Center of Advanced Transportation Technology</b>
<b>DOE</b>	<b>Department of Energy</b>
<b>DOT</b>	<b>Department of Transportation</b>
<b>DQL</b>	<b>Deep Q-Learning</b>
<b>DSC</b>	<b>Dynamic Signal Control</b>
<b>DSL</b>	<b>Dynamic Shoulder Lane</b>
<b>DTA</b>	<b>Dynamic Traffic Assignment</b>
<b>DTL</b>	<b>Dynamic Toll Lane</b>
<b>FHWA</b>	<b>Federal Highway Administration</b>
<b>ITMS</b>	<b>Internet Traffic Monitoring System</b>
<b>ICM</b>	<b>Integrated Corridor Management</b>
<b>ITS</b>	<b>Intelligent Transportation System</b>
<b>MAS</b>	<b>Multi-Agent System</b>
<b>MWCOG</b>	<b>Metropolitan Washington Council of Governments</b>
<b>NN</b>	<b>Neural Network</b>
<b>OD</b>	<b>Origin-Destination</b>
<b>QL</b>	<b>Q-Learning</b>
<b>PC</b>	<b>Principle Component</b>
<b>PCA</b>	<b>Principle Component Analysis</b>

<b>RL</b>	<b>Reinforcement Learning</b>
<b>RITIS</b>	<b>Regional Integrated Transportation Information System</b>
<b>RM</b>	<b>Ramp Metering</b>
<b>SHA</b>	<b>State Highway Administration</b>
<b>STT</b>	<b>System Travel Time</b>
<b>SB</b>	<b>Southbound</b>
<b>TAZ</b>	<b>Traffic Analysis Zone</b>
<b>TDM</b>	<b>Transportation Demand Management</b>
<b>TPB</b>	<b>Transportation Planning Board</b>
<b>TTI</b>	<b>Travel Time Index</b>
<b>VSL</b>	<b>Variable Speed Limit</b>
<b>VMT</b>	<b>Vehicle Miles Traveled</b>
<b>VOC</b>	<b>Vehicle Operating Cost</b>
<b>VOT</b>	<b>Value of Time</b>
<b>WISE</b>	<b>Work Zone Impacts and Strategies Estimator</b>
<b>WMSE</b>	<b>Weighted Mean Squared Error</b>
<b>UE</b>	<b>User Equilibrium</b>
<b>UMD</b>	<b>University of Maryland</b>

# Chapter 1 Introduction

## 1.1. Background

In an era of accelerated urbanization around the world, the ability to travel freely is more critical than ever before. Travel is no longer just for commuting to and from work. Travel for other purposes such as entertainment, vacation, leisure activities, running errands, and shopping surpasses work-related trips. While trips through modes such as public transportation (e.g., bus, rail), new transportation network companies (e.g. Uber, Lyft), bicycle, shared bike, and scooters have been increasing, trips carried out through personally owned vehicles are still the predominant method of travel. This phenomenon results in continued pressure on freeway systems, leading to increases in delays and congestion. Highway congestion has truly become a Gordian knot in transportation from both planning and operation standpoints.

To solve the traffic congestion problem, numerous methods and approaches have been researched, developed, and implemented. The traditional method in mitigating traffic congestion from the supply side is to increase the number of travel lanes so as to increase capacity. Widening roads is one of the most direct methods but is often limited by the fiscal appropriation and lengthy construction process. Additionally, the higher capacity of these new lanes through widening is often quickly overtaken by increased demand. This is why roads are widened from two lanes to four lanes initially, and then further expanded to six or eight lanes. Widening roadways on its own has failed to solve the congestion issue.

Recently, more attention has been paid to leveraging traffic operational techniques and approaches to solve traffic congestion. Many researchers state that an efficient operation of existing road infrastructure is the only solution that balances economic benefit and technical performance. Even though the capacity is defined as a constant variable that represents the expected maximum throughput of the roadway, it is claimed that this traditional understanding violates factual conditions. The capacity of a roadway segment should be regarded as a random variable instead of a constant value. These factual evidences favor improving traffic throughput by dynamic traffic control. Several types of dynamic control methods based on traffic flow theory have been developed and deployed in real-world applications. They prove that traffic control is a way to prevent, or at least relieve, traffic congestion, hence improving traffic conditions [1].

### 1.2. Active Traffic Management (ATM)

Among these dynamic control methods, active traffic management (ATM)—such as variable speed limit (VSL), ramp metering (RM), hard shoulder running (HSR), and adaptive traffic signal control (ATSC)—has the ability to manage both recurrent and non-recurrent congestion and has been widely applied in freeway systems. These deployed ATM methods include:

- 1) Variable Speed Limit: variable speed limit (VSL) is one of the relatively new freeway operation methods. VSL improve traffic conditions by posting dynamic speed limits to regulate the traffic flow on mainline. Several VSL

control systems have been developed and implemented in real world applications for different purposes such as safety, mobility, and work zones. A VSL system is composed of multiple traffic sensors, queue warning signs or dynamic guiding signs, commutation systems and online control algorithm. For a VSL application, the traffic sensors collect real-world data such as speed, density, volume and queue delay. The system receives the traffic data and generates optimal solution based on the online control algorithm. And then the system updates the speed limits and posts them on the dynamic message signs to regulate inflow volume from upstream. With the dynamic speed limits based on real traffic conditions, the VSL system can make improvement in safety and mobility [2-3];

- 2) Ramp Metering: Ramp metering is one of the most widely implemented ATM strategies. RM controls the traffic volume merging on to the freeway mainline by installing traffic signals at on-ramps. By regulating inflow volume from one or multiple ramps, the RM strategies can significantly relieve stress on freeway mainline. [4-6];
- 3) Dynamic Queue Warning: dynamic queue warning informs upstream drivers of upcoming traffic conditions based on data collected by real-time traffic sensors. Warning signs are presented several miles upstream to help drivers anticipate the upcoming condition and act upon it. In some real-world applications, DQW strategy is integrated as part of VSL control system to maximize the benefits [7-8].

- 4) Hard Shoulder Running: hard shoulder running, also known as dynamic shoulder lane, temporarily uses the shoulder lane as an additional general purpose lane (GPL) to increase lane capacity during rush hours. Different from other mainline traffic flow control methods, HSR is the only strategy improve traffic congestion by directly increasing roadway capacity. However, HSR is not suggested to be active for long period for safety concern [9].
- 5) Adaptive Traffic Signal Control (ATSC): as one of the most effective traffic signal control methods, the ATSC system allocates green signal time to various vehicle groups in a dynamic way by analyzing real-time data collected by sensors installed on all approaches of an intersection. The ATSC minimizes delays, reduces vehicle hours traveled, and reduces fuel vehicle consumption [10-12].

Early ATM strategies tend to use simple logic-based algorithms. These kind of strategies are easy to implement but fail to provide accurate solutions. Some recent research have adopted various advanced algorithms to enhance the performance. Generally speaking, most recent strategies can be roughly divided into two categories: optimization-based strategies and feedback-based strategies. As the name suggests, optimization-based strategies improve traffic condition by considering traffic operation process as optimization problem. With accurate traffic flow model, car-following model and traffic prediction model, the strategies could provide effective solutions. However, these optimization problem requires powerful computing hardware and accurate models, and it is hard to imagine the investment for real world implementations. Unlike optimization-based strategies, feedback-based algorithms

achieve similar effect without rigorous requirements for model and hardware. Even though they are better in feasibility, the performance of feedback-based strategies are limited due to the deficiency in traffic prediction model. More recently, the reinforcement learning (RL) technique has been integrated in some ATM control strategies to optimize solutions in a rapid way. The problem of RL is initially come up by behaviorist psychology, which focuses on how an agent tend to commutate with the environment and take actions in order to receive maximum cumulative rewards. Applied in a ATM strategy, RL method helps the agent learn how to provide optimal solutions (speed limits, meter rates) that can receive maximum cumulative rewards (minimum queue delay or maximum throughput) for the next few steps. Though the offline training is still time-consuming, the introduction of RL helps the ATM strategies provide accurate and timely response to various traffic conditions.

Many researchers analyzed ATM traffic control methods and stated the benefits of applying these operational methods on freeways. Although various ATM control strategies have been widely implemented, it has gradually been realized that a stand-alone control algorithm has limited ability of improving traffic conditions. Even though RM control can significantly lower the density of the immediate downstream of the controlled on-ramps, it is a hot potato to balance the benefits and drivers' compliance rate. VSL helps release the bottleneck congestion by regulating the inflow volume from upstream. However, several simulation-based evaluation studies demonstrate that variable speed limit fails to make improvements under extreme congestion. Route guidance methods only work in the case of nonrecurring traffic congestion and HSR cannot be utilized all the time without impeding traffic safety [13]. More and more

research studies have demonstrated the limitation of a stand-alone ATM system and the possibility of coordinating two or more ATM-based systems for more benefits [14-16].

### 1.3. Traffic Modeling Approach

To test the reliability of traffic control algorithms, traffic simulation models are developed as an indispensable instrument for transportation planners and traffic engineers. A simulation model should not only represent a traffic network and traffic demand in the real world, but also simulate dynamic traffic conditions. In the past few decades, the traffic-modeling field has made significant progress. Traffic simulator development can be traced back to the 1950s when computers were introduced to universities and research institutions. Limited by the Central Process Unit (CPU) and computer memory, most simulation models were built for only short roadway segments. Additionally, the functions of the traffic simulator were restricted to changing number of lanes and traffic demand. Nowadays, traffic simulators have been revolutionized with powerful computing, multifunctional simulation environment and good visual results [20].

Traffic simulation models can be divided into micro, macro, and meso-scale platforms in terms of their objectives and components. VISSIM, AVENUE, Paramics, Aimsun, and SOMU, etc. are representatives of microscopic simulation models [17, 18]. These models work through pre-defined agents in the system. Complex traffic conditions are visualized by realistic traffic models. The critical advantage of a microscopic model is its efficiency in evaluating complex traffic congestion, intricate



geometric configurations, and system-level impacts of proposed strategies. A microscopic model necessitates detailed and complicated input data and running it is usually time-consuming. The development, calibration, validation, and maintenance are often costly and technically challenging.

Instead of tracking an individual vehicle, macroscopic models simulate a traffic system based on traffic flow theory. Macroscopic models such as cell transmission model (CTM) and TRANSCAD can handle large networks with short simulation time [19]. Compared with microscopic models, macroscopic models need less computational effort and pave the way for integrating multiple control methods such as Kalman Filter and ATM methods. However, limited by available details, specific considerations (e.g., drivers' compliance rate and mixed traffic flow) may be difficult to incorporate in macroscopic models.

As a compromise between micro and macro modeling approaches, mesoscopic models are developed. Mesoscopic models balance between the realism and computational efficiency in demand and supply models, and therefore can handle the non-trivial networks and provide detailed results at the same time. Mesoscopic models are used more and more widely. The most popular usage are dynamic traffic assignment (DTA) models. For example, several DTA models have employed mesoscopic supply simulation, which uses aggregate traffic flow relationships to model individual vehicle movements, and gain computational efficiencies over a time-consuming microscopic simulation. The mesoscopic models are more advantageous in regional traffic analysis,

which is necessary in ATM control studies. It improves precision and increases efficiency, while at the same time, an agent's behavior can still be traced.

#### 1.4. Objective

As the literature on developing and deploying an effective ATM system revealed, a host of issues must be further analyzed and resolved. Obviously, the improvements expected from stand-alone ATM control strategy are limited by external conditions. Traditional VSL-alone algorithms are not reliable enough. With increasing traffic demand, the ability to integrate traffic information with actionable solutions is even more needed. In the meantime, a reinforcement learning technique should also be included as an effective method for optimizing the coordinated ATM control algorithms. However, only a few research efforts have explored the benefits of coordinated ATM algorithms, especially the effectiveness of VSL under reinforcement learning technique. This research work attempts to develop and implement a coordinated ATM control system that can be implemented in regional transportation analysis with shoulder running activated.

The objective of this research is to develop a coordinated dynamic traffic control system that integrates variable speed limit information with hard shoulder running using a reinforcement learning technique. The ultimate goal is to build a model that enables traffic scenario analysis, such as time-of-day, freeway trajectory, future demand assessment, and special event traffic conditions. To analyze the performance of the proposed algorithm, a mesoscopic simulation model based on DTALite is developed, which can dynamically present the traffic improvement.

This research contributes to new information and new approaches in solving traffic congestion in the following ways:

- 1) Reinforcement-learning (RL) technique: a new efficient method that integrates the RL into ATM control strategies to obtain optimal solutions without running into complex calculation.
- 2) VSL integration: a new effective method that integrates VSL control with HSR, enabling the dynamic control system to be more efficient and manageable.
- 3) Reward function of Q-learning: a new formulation on QL reward function is created based on the queue delay instead of the density at a bottleneck. The proposed algorithm provides support for evaluating the possibility of only using queue delay as the key parameter for ATM control.
- 4) DTA simulation model: to the authors' best knowledge, this is the first paper analyzing RL-based ATM control with DTA model. Compared to traditional microscopic analysis, the mesoscopic DTA model requires less computational burdens for regional impact on large-scale network.

### 1.5. Paper Organization

The remainder of this paper is organized as follows. Chapter Two is a literature review of previous studies on VSL, HSR, and other coordinated ATM control strategies. The basic reinforcement learning technique and the QL-based ATM control algorithm are introduced in Chapter 3. Chapter 4 discusses the MOE evaluation under different

scenarios using a case study on I-270, Maryland, United States. Finally, findings and recommendations are summarized in Chapter 5.

## Chapter 2 Literature Review

This chapter is a review of past work on variable speed limit, hard shoulder running, applications of artificial intelligence, and other related researches associated with traffic operational systems.

### 2.1 Real World Applications

A number of variable speed limit control systems have been implemented in the United States since 1960s (Table 2-1). Until now, VSL systems have been widely implemented in many states for various purposes [21-22]. According to a government report in 2015, VSL applications in U.S. are implemented for three primary functions: alleviating recurrent traffic congestion, reducing average speed to ensure traffic safety under severe weather conditions, improving traffic capacity during non-recurrent congestion caused by work zones or incidents [23].

While the United States has installed VSL systems as far back as the 1960s on locations such as the New Jersey Turnpike, the operation systems have experienced an enormous upswing in the last few decades [24]. As the first approach of variable speed limit in U.S., New Jersey Turnpike VSL system was mainly designed for safety concerns. The system monitored the traffic and weather conditions based on the feedback from more than 120 sensors, and then provided decision support for speed limits to improve traffic safety. Another application is on I-90, which has become an important transportation hub for passenger traffic and physical distribution since it opened in 1970s. With the regional economic development and population increasing,

the freeway system was under too much pressure. Variable speed limit was applied on Interstate 90 (I-90) in Washington State, U.S., as a possible transportation system manage and operation (TSMO) type solution to achieve both safety and operation benefits [25]. Based on the analysis by Ulfarsson et al., the speed variation on I-90 was significantly decreased after variable speed limit system started operation [26]. To enhance traffic safety, Abdel-Aty et al. (2006) applied VSL system on I-4 in Orlando, Florida and the performance indicates that the VSL contributes to the reduction of both crash risk and average speed [27]. Recently, Chang et al. developed a multiple-objective VSL control system on MD 100 in Maryland State to increase speed and throughput. The algorithm tends to reduce the speed variance between free flow state and stop-and-go congested state. Furthermore, travel time estimation and drivers' response were also considered in the algorithm to reach better performance [28-29]. A study on MD 100 by Chang et al. (2011) indicates that the VSL strategy could be a possible method in releasing congestion with sudden speed drop [29]. Other applications such as I-66 in Virginia, I-35W and I-494 in Minnesota, and I-255 in Missouri were implemented for different purposes and received visible benefits. Other than the United States, VSL systems have been widely implemented all over the world such as Germany, Australia, and the U.K. [30-31] Studies on these applications also proved the safety and mobility benefits of VSL in reducing speed variation and enhancing traffic throughput [31].

## 2.2. Variable Speed Limit Effects

Congestion on freeway system has become a major problem that leads to capacity, safety, and mobility reduction [32]. Frequent acceleration and deceleration as

a result of high density and low density on congested segment, leading to the increase of unsafety. Studies reveal that drivers are more likely to be involved in the traffic accidents when driving in the traffic with high variations [33]. Historical data collected by the Department of Transportation (DOT) indicates that the incidents rate is significant higher on freeways than urban roads due to wider range of speed [33]. In addition, the incidents rate also increases with higher freeway occupancy [34-36]. Variable speed limit was analyzed by several studies as the possible method to reduce speed variance by decreasing average headway on freeways [35]. Table 2 1. Examples of VSL, DMS and Queue Warning Applications in U.S.

Lee et al. developed a microscopic simulation model to capture drivers' response to speed limit control [36]. The results indicated that the speed deviation was reduced with VSL control, and translate into lower speed variation and incidents rate. Abdel-Aty et al. (2008) proposed a VSL strategy with homogeneous speed zone to explore possible benefits in reducing rear-end and lane-change crash risks [37]. It is proved that VSL could be an effective method in preventing incidents under uncongested conditions and reducing incidents rate under modest congested conditions according to Abdel-Aty et al (2008). However, the safety benefits may be limited when the congestion is severe.

Table 2-1: Examples of Real World VSL Applications in U.S

<b>Real-world VSL Implementations in U.S.</b>						
<b>State</b>	<b>Location</b>	<b>Length (miles)</b>	<b>Status</b>	<b>Authority</b>	<b>Operation Types</b>	<b>Primary Functions</b>
Florida	I-4	10.5	Active	Regulatory	Hybrid	Congestion
Florida	US 27	3	Active	Regulatory	Automated	Congestion
Georgia	I-285	36	Active	Regulatory	Hybrid	Congestion work zones
Minnesota	I-35W	18	Temporarily Deactivated	Advisory	Automated	Congestion
Minnesota	I-94	10	Temporarily Deactivated	Advisory	Automated	Congestion
Nevada	US 395 (Alternate)	5	Active	Regulatory	Automated	Weather
New Jersey	NJ Turnpike	148	Active	Regulatory	Manual	Congestion Weather
Oregon	OR 213	Single Inter- section	Active	Regulatory	Hybrid	Weather
Oregon	OR 217	7	Active	Advisory	Automated	Congestion Weather
Tennessee	I-75	9	Active	Regulatory	Automated	Weather
Virginia	I-66	13	Active	Regulatory	Automated	Congestion Work zones
Virginia	I-95 (Express Lanes)	~10	Active	Regulatory	Manual	Congestion
Washington	I-90	25	Active	Regulatory	Hybrid	Weather
Washington	US 2	23	Active	Regulatory	Automated	Congestion
Washington	I-5	8	Active	Regulatory	Automated	Congestion
Washington	I-90 (Bellevue to Seattle)	10	Active	Regulatory	Automated	Congestion
Washington	SR 520	8	Active	Regulatory	Automated	Congestion

Source: FHWA (Guidelines for the Use of Variable Speed Limits Systems in Wet Weather)

In addition to safety benefits, VSL has received increased interest in relieving traffic congestion. A set of variable speed limit signs placed on freeway segments displaying dynamically controlled speed limits harmonize the speed transition between free-flow segment and congested segment. It has been proved that the capacity is a variable parameter instead of remains constant. Capacity drop at bottleneck reduces the discharging rate, and then results in the nonlinear and discontinuous of speed-volume relation [38-39]. To avoid capacity drop phenomenon or at least reduce the effects,



several VSL strategies have been developed [40]. Properly implemented VSL system can maximize the roadway utilization, so as to decrease average corridor travel time during rush hours. It has been proved that the bottleneck speed and throughput could be improved in an impressive way according to Kwon and Brannan (2007) [41]. Hadiuzzaman et al. (2013) proposed a VSL strategy with model predictive control to avoid capacity drop [42]. The simulation results from VISSIM indicates that the total travel time, total travel delay decreased by 39.0% and 8.0%. The VSL strategy is effective in improving traffic throughput, but not a good choice for travel time reduction according to the macroscopic VSL simulation model taking throughput as the objective function [43]. Other studies point out that the total travel time could be reduced by more than 20%, which contrasts with Alessandri's results (1999) [44-45].

### 2.3. Variable Speed Limit Strategies

Variable speed limit is a commonly used ATM strategy that releases bottleneck congestion by regulating the inflow volume from upstream. VSL enables dynamic changes of posted speed limits in response to different travel conditions to meet the objectives of safety, efficiency, and environmental consistency. Early studies tend to use simple logic-based VSL strategies, in which the speed limits are updated based on some established rules based on volume, density, and throughput. Logic-based VSL strategies are widely adopted in many real-world applications since they are simple, feasible and efficient [46]. These kind of strategies are easy to implement but fail to build close communication with constantly changing traffic conditions. Most logic-based strategies use logic tree to classify the traffic condition into different groups, and then apply speed limits to control traffic. However, the traffic flow is continuous and

ever-changing that could not be simply classify as discrete variables. It is possible for a traffic condition to be wrongly classified and even deteriorate the traffic congestion.

Later some research use several advanced techniques, such as model predictive control, connected vehicle technique, mainline traffic flow control and travel time estimation to intensify the performance of VSL strategies [48-51]. Most of these strategies consider the traffic control as an optimization problem with an objective function that minimizing the total travel time or maximizing the throughput [52-53]. Optimization-based strategies outperform the logic-based approaches in capturing the dynamic traffic conditions and providing effective solutions in general. Even though these strategies provide excellent performance on a theoretical level, the feasibility are deserved to be questioned. The optimization problem usually requires accurate traffic prediction model and complex computation process, which are appropriate for real-world implementations. Additionally, it is very different or even impossible for the optimization-based strategies to provide quick response to the traffic, in terms of the current computing power.

Not restricted by the computing requirement, some other research tend to use feedback-based VSL strategies, which update control variables based on the observed traffic conditions [54-56]. These kind of strategies control the traffic by regulating one or more traffic variables within a certain range. To some extent, the range is similar with the objective function in optimization-based strategies, but can be simply reached using traffic flow theory instead of complex computation. Compared with optimization-based strategies, feedback-based strategies perform similar in accuracy

and efficiency but requires less in hardware. However, most feedback-based strategies fail to timely response to the environment since the strategy are passively motivated by the real-time traffic conditions. In other words, the performance is limited since only the current traffic condition is consider in the strategies. Additionally, the parameters and ranges have to rely more on researchers' knowledge and experience. Without systematic process, human error may affect the performance of feedback-based strategies.

Recently, the reinforcement learning (RL) technique are getting more attention in solving complex optimization problems [57-58]. RL is an area of machine learning concerned with how the software agents ought to take actions in an environment to maximize some notion of cumulative reward. The RL relies on the environment formulated by the Markov decision process (MDP), which refers to a set of sequential decisions under an observable environment. Among several kinds of RL-based methods, Q-learning (QL) is the most popular one with benefits of policy-free and model-free. It can also be viewed as an efficient strategy of asynchronous dynamic programming (DP). In the QL process, the agent is a self-learning machine that receives state information from environment and produces actions at each time step. A reward function is designed to evaluate each state-action pair, and the maximum expected reward is updated and stored in a Q-table. Without implementing a policy, the agent intensifies the policy by using the Q-table as the source to learn and improve by itself. Some more recent research studies have integrated QL method into ATM control strategies such as variable speed limit and hard shoulder running. In a QL-based ATM strategy, the agent communicates with the environment (simulation network) by

receiving states (traffic conditions) and taking actions (updated speed limits, meter rates, queue warning status). A reward function is generated to evaluate the performance (traffic improvement) of each state-action pair. According to the Q-updating function, the agent always learn how to find the actions that can achieve maximum cumulative rewards in the next few steps. Rezaee et al. (2012) introduced a RL-based RM strategy using real-life data from Highway 401 in Toronto [59]. The study compares the RL-based strategy with traditional ALINEA strategy and indicates that RL-based method outperforms the other one in reducing total travel time. Zhao et al. (2011) designed a RL-based RM strategy to relieve both recurrent and non-recurrent congestions [60]. Different from other studies using speed, density as volume, Zhao et al. consider queuing as the effective measurement. Li et al. (2017) developed a RL-based VSL system for freeway recurrent bottlenecks [61]. The results indicate that RL-based strategies are more effective compared with feedback-based strategies. Zhu et al. analyzed the mobility and environmental benefits of a RL-based VSL strategy under stochastic demand [62]. The proposed strategy reduced corridor travel time by 18% and emission consumption by 20%. Even though the QL-based strategies require some time for offline training process, the agent can take optimal actions under various states without complex computation and accurate prediction model.

### 2.3. Coordinated ATM Strategies

Meanwhile, with increasing traffic demand, requirements of VSL control strategies are more stringent and complex. The disadvantages of stand-alone strategies have been noticed [62-63]. Abdel-Aty et al. (2008) suggested that VSLs work better at lower demand level [37]. Grumert and Tapani (2012) pointed out that VSL was less

effective than ramp metering [65]. Grumert's study also indicated that VSL has poor performance compared with RM. To maximize the benefits, some research explored the possibility of integrating two or more ATM strategies together [66]. One such system is the integrated RM and VSL control. Ramp metering aims at improving the mainline traffic flow condition by appropriately regulating the inflow from on-ramps to mainline [67]. Due to the limitation of ramp capacity and user endurance, ramp metering alone cannot make much improvement on network travel time [42, 68]. According to behavior researches, drivers' responses to RM control could be affected by unbalanced psychology, meaning that drivers may drive more aggressively after entering a freeway mainline to make up for the delay experienced on the ramps. Similarly, VSL reduces density and increases throughput at bottlenecks by holding more vehicles at upstream. However, when volume exceeds capacity on a freeway, the VSL algorithms have minimum value in reducing speed variance. Integrated the advantages of the two strategies, Abdel-Aty and Dhindsa (2007) proposed a coordinated strategy that significantly outperformed the stand-alone strategy in reducing crash possibility and improving traffic mobility [69]. A Genetic-Fuzzy feedback-based strategy integrated RM and VSL was introduced by Ghods et al. (2007) [70]. The integrated algorithm achieved 5.1% reduction of total travel time, which was significantly better than stand-alone scenario.

As literature revealed, a host of issues need to be further analyzed and resolved. Obviously, the improvements expected from stand-alone ATM control strategy are limited. Traditional VSL-alone algorithms are not reliable enough. The ability to integrate traffic information with actionable solutions is more needed. In the case that

VSL is not efficient under high volume conditions, a question worthy of further study is the possibility to integrate HSR to directly increase roadway capacity. Hard shoulder Running has been adopted in many studies and real-world practices for its effectiveness in reducing traffic congestion [71-72]. The effects of the HSR on freeway capacity and traffic flow characteristics were analyzed by Geistefeldt (2013) based on data collected from freeways in Germany [9]. With the implementation of Hard shoulder running, the freeway congestion is relieved by increasing traffic hourly throughput by 1000 vehicles. Samoili et al. (2013) developed a short-time prediction model to evaluate the network performance of HSR [73]. The HSR could increase the roadway capacity by 10% and maintain the traffic speed at a stable level. In addition, the short-term prediction model predicted that more than 20% volume on the left-lane would be attracted to the shoulder lane. Ma et al. (2016) analyzed the benefits of using HSR to improve traffic efficiency of nonrecurring traffic incidents [72]. The study provided several suggestions on the length of the shoulder opened upstream of an incident, the length of the shoulder opened downstream of incidents, and the opening duration of the shoulder. The study concluded that HSR could improve traffic condition by reducing an average delay up to 80% and increasing traffic throughput up to 40%.

Even though HSR has significant performance, the safety impact has been a national controversy in recent years [73-77]. Geistefeldt (2011) analyzed the impacts of shoulder lane operation on traffic efficiency and safety using Brilon's method on several highways in Germany. It was found that the HSR contributed to a significant improvement in traffic capacity by 25%, but failed to improve traffic safety. Chapoton and Dumont (2015) studied the HSR system in Switzerland and emphasized the

concern of the impact of HSR on drivers' behavior. The study analyzed people's response to HSR configuration and indicated that drivers' attitudes presented great polarization [73]. Since HSR should not be active for long period, the coordination between HSR and other strategies may provide both safety and mobility benefits.

## Chapter 3 Methodology and Model Generation

In this study, a coordinated ATM control strategy is proposed and tested based on the reinforcement learning (RL) technique. To better demonstrate the methodology, the concept of basic RL and QL algorithms are first introduced in this chapter. Next, the proposed coordinated QL-Based ATM concepts are presented with some practical implementation principles. The rest of this chapter includes:

- (1) Traffic congestion causes and effects
- (2) VSL & HSR control theory
- (3) Basic Q-learning theory
- (4) Coordinated QL-based ATM control algorithm
- (5) Study area description
- (6) Simulation model
- (7) Parameters setting

### *3.1. Traffic Congestion Causes and Effects*

There are several reasons for traffic congestion. The leading reason is the imbalance between traffic demand and roadway capacity. Traffic congestion usually happens when inflow demand exceeds the road capacity, resulting in saturation or oversaturation. Traffic congestion includes recurring congestion and non-recurring congestion.

- 1) Recurring congestion: recurring congestion refers to continuous traffic congestion formulated at a relatively fixed roadway segment with special



trajectories such as diverging area, merging area, on-ramps, off-ramps, curves, and strong grade during peak hours. Previous studies suggest that recurring congestion is often seen as a capacity problem and is logically combated by increasing roadway capacity.

- 2) Non-recurring congestion: different from recurring congestion, trajectory and demand are not the leading causes of non-recurring congestion. The reasons of non-recurring congestion include crashes and incidents, work zones, heavy weather conditions, big events, and others influence factors (polices or unexpected foreign object) .

A direct way to judge congestion is comparing inflow traffic  $q_{inflow}$  from upstream and the capacity  $c_{downstream}$  at the downstream. If  $q_{inflow}$  is less than or equal to  $c_{downstream}$  , the roadway segment is close to free flow status. In a case where the inflow  $q_{inflow}$  is higher than the downstream capacity  $c_{downstream}$ , the bottleneck is activated, blocking inflow vehicles. The head of the queue is always at the bottleneck while the tail spills back as long as the inflow is high. When congestion happens, there are two effects: capacity drop and blocking of ramps:

- 1) Capacity drop: several studies on roadway capacity indicate that the capacity is not a constant variable that equals to the expected maximum throughput of the roadway as shown in Figure 3-1(a). The capacity decreases by about 20% at a bottleneck due to speed reduction. Capacity drop happens because drivers need time to accelerate from low speed to normal speed when leaving a bottleneck [55].

- 2) Blocking of ramps: another negative effect of active bottlenecks is the blockage of ramps due to the spill-back of a queue. Blocking off-ramps also adds pressure on the mainline, and then resulting in the paralysis of traffic system.

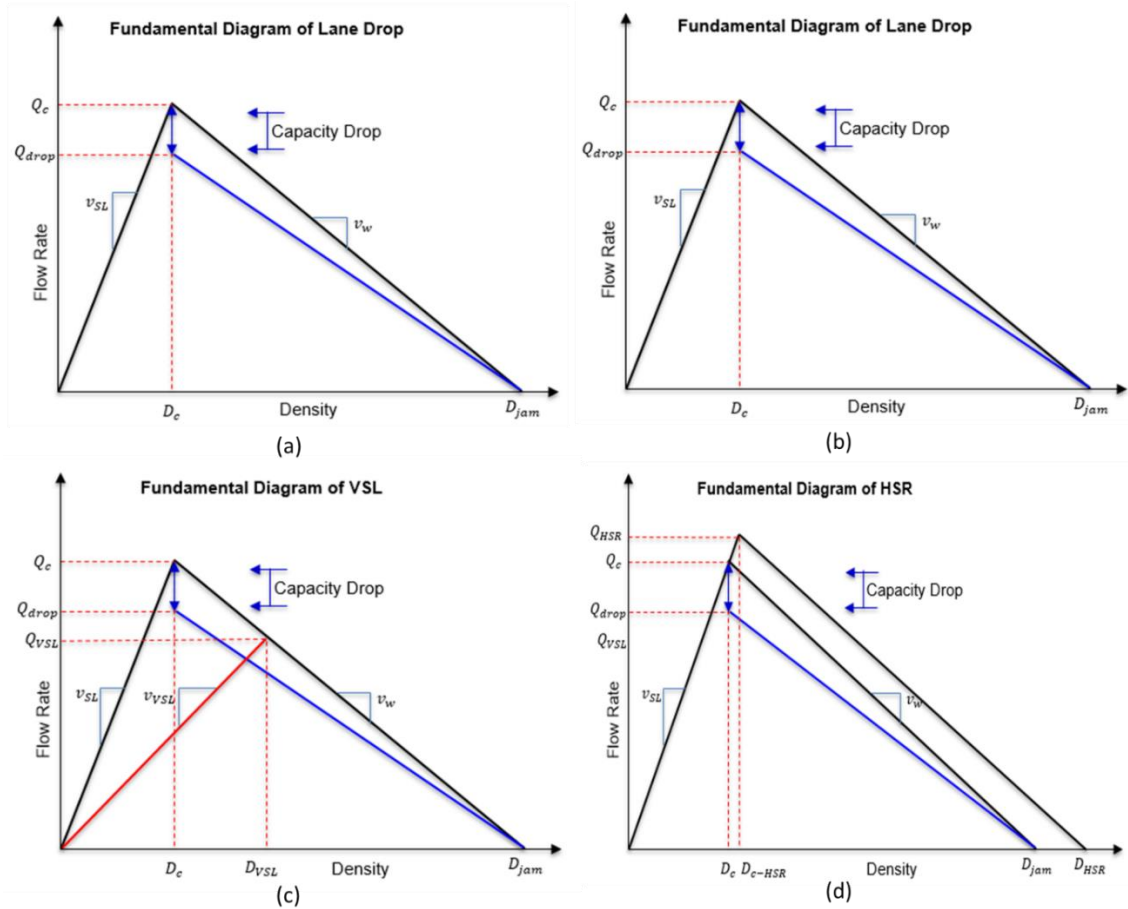


Figure 3-1. (a) Fundamental diagram of flow-density relationship under uncongested condition; (b) fundamental diagram of flow-density relationship with capacity drop; (c) fundamental diagram of flow-density with VSL control; and (d) fundamental diagram of flow-density with HSR control.

### 3.2. VSL & HSR Control Theory

As indicated by previous studies, insufficient roadway capacity is the main reason for both recurring congestion and non-recurring congestion. MTFC strategies such as VSL and Rm tend to maximize the bottleneck throughput by regulating the

inflow volume from upstream. Unlike theirs, however, HSR improves the throughput by increasing the temporary capacity of bottleneck and downstream.

### 3.2.1. VSL Control Theory

The idea of VSL is to regulate upstream traffic volume with appropriate controls to avoid capacity drop at a bottleneck. For example, the free flow speed on the freeway is  $v_{ffs}$ . If no VSL control is applied to this segment, the traffic flow at a bottleneck is  $q_{ffs}$ . However, using VSL, the outflow  $q_{vsl}$  from the VSL control segment is controlled less than or equal to the bottleneck capacity ( $q_{vsl}$  is controlled less than  $c_{downstream}$  in most situations due to the capacity drop). The congestion on the freeway could certainly not be avoided since  $q_{inflow}$  is always higher than  $c_{downstream}$  even if the VSL is applied.

As shown in Figure 3-1(c), the red line represents the flow rate under VSL control. It is obvious that with a lower posted speed limit, the critical density increases while the maximum flow rate  $Q_{VSL}$  decreases. As long as  $Q_{VSL} < Q_{drop}$ , the congestion releases. In addition, a higher density at the VSL control segment proves lower flow rate since more vehicles are stored in the segment.

### 3.2.2. HSR Control Theory

Although VSL can release traffic congestion to some degree by avoiding the capacity drop and increasing outflow at the bottleneck, the congestion on the freeway cannot be avoided since  $q_{inflow}$  is higher than  $c_{downstream}$ . To remedy the defects of VSL, HSR is introduced by temporarily increasing  $c_{downstream}$ . Under normal

conditions, an additional lane provides enough capacity to handle inflow volume  $q_{inflow}$  from upstream.

As shown in Figure 3-1(d), the maximum flow rate at a bottleneck location increases when HSR control is active. With the same free-flow speed and wave speed, the  $Q_{HSR}$  is significantly higher than  $Q_c$ . For a three-lane freeway, the temporary capacity can increase by 22% using a shoulder lane as general-purpose lane, which can greatly relieve severe traffic congestion.

Recently, HSR has been widely implemented on many highways in the U.S., such as I-595 Reversible Express Lanes, I-66 between Merrifield, Virginia and Washington D.C., I-35W in Minneapolis, and I-110/I-10 Metro Express Lanes in Los Angeles, California. Judging by their performance, HSR is one possible strategy for addressing congestion and reliability issues within the transportation system, and is particularly cost-effective where widening roads is infeasible, undesirable, or cost prohibitive. HSR exists in many different forms, but they are all designed as designating the left or right shoulder lane as a normal travel lane during certain times of the day.

### 3.3. Basic Q-Learning Algorithm

#### 3.3.1. Reinforcement Learning Algorithm

Machine learning has rapidly developed in the past few decades. As a self-learning system that can extract information and develop knowledge, machine-learning algorithms prove to be more efficient than traditional optimization algorithms. Machine

learning has a variety of learning methods that generally fall into three categories: supervised learning, unsupervised learning and reinforcement learning. Supervised learning uses training examples to learn how to classify the inputs while unsupervised learning methods form the concepts by themselves. Different from these two categories, a reinforcement learning method uses a reward function to tell the agent how the action performs.

As one of the effective learning methods for complex relationships, reinforcement learning technique (RL) is a potential method for addressing optimization problems. A RL method works by formatting the optimization problem into Markov chain decision process, which refers to a set of sequential decisions under an observable environment. For ant Markov chain decision process, the time-dependent character indicates that the future state is completely independent of the past states or actions, as long as the current state is given. This kind of relationship can be described in a mathematical equation as follows:

$$P(S_{t+1}|S_t) = P(S_{t+1}|S_1, \dots, S_t) \quad (1)$$

where  $S_t$  refers to the current state at time  $t$  and  $S_{t+1}$  refers to the state of next timestamp.

A RL method consists of three parts, agent, reward function, and environment. The agent is a self-learning machine that receives its state information from its environment and produces actions at each time step. The performance of a state-action pair is evaluated by the reward function, which is the source for the agent to learn and

improve itself. As shown in Figure 3-2, at each time step, the agent perceives the state of its environment and takes an action to transfer the system from its current state to a new state. A reward calculated by the reward function is posted to the agent to evaluate the quality of the transition. After sufficient iterations, the agent traverses all state-action pairs and learns how to find a sequence of optimal actions that yields the maximum cumulative reward over the time period. For a successful RL process, the cumulative reward an agent received at each iteration will converge to a relatively stable level, which is an indication that the RL has completed its training.

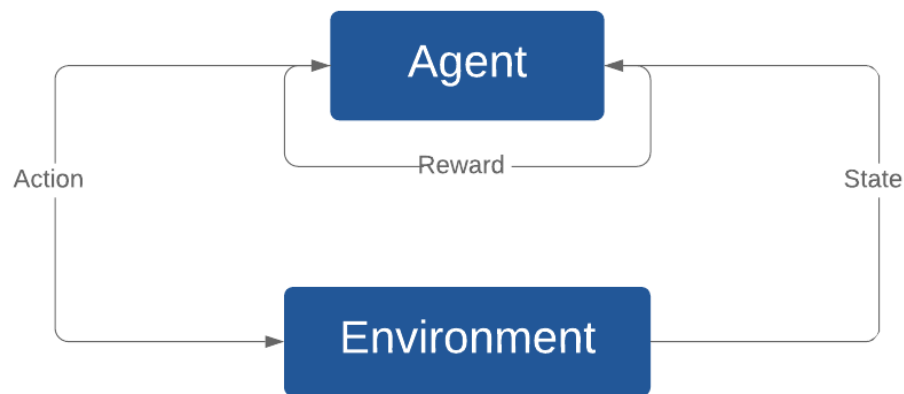


Figure 3-2.Components of Reinforcement Learning.

### 3.3. Basic Q-learning Strategy

Q-learning is a model-free and policy-free RL technique [78]. It can also be viewed as an efficient strategy of asynchronous dynamic programming (DP). The Q-learning method provides an agent with decision-making capability regarding optimal solutions by traversing the entire set of states and actions without modeling the

environment. Q-learning is the most popular RL technique, with reliable performance in many fields of engineering [79].

QL consists of the state set  $S$ , the action set  $A$ , and the reward function  $R$ . At each step of the learning process, the QL agent receives state information from the environment and chooses an action. Usually the agent randomly chooses actions in the incipient stage of the process and tends to choose a particular action after the convergence. The agent takes an action to transfer the environment from current state to a new state. In the QL method, a reward function is determined and assigned to each state-action pair to evaluate the performance of the action. After multiple iterations of learning, the agent learns how to take an action from the current state to maximize the possible rewards in the next steps. Also, in an ideal world, for any given state, the agent could select a sequence of actions that maximize the cumulative rewards. The relation between the states set  $S$ , the action set  $A$ , and the reward function  $R$  can be described as:

$$Q : S * A \rightarrow R \quad (2)$$

where  $Q$  refers to the Q-value that represents the quality of each state-action pair. For any QL problem, the ultimate goal is learning a policy  $\pi$  for an agent operating in an environment with stochastic actions and rewards, and to do so without a model. For each possible policy  $\pi$ , the value the agent can adopt is:

$$V^\pi(S) = R_t + \gamma R_{t+1} + \gamma^2 R_{t+2} + \dots \quad (3)$$

where  $V^\pi(S)$  refers to the possible cumulative rewards under policy  $\pi$  as for the current state  $S$ . The Q-function using the Bellman equation could be represented as:

$$Q^\pi(s_t, a_t) = \underline{E}[R_{t+1} + \gamma R_{t+1} + \gamma^2 R_{t+1} + \dots | s_t, a_t] \quad (4)$$

Therefore, in terms of the value function, the learning task can be reformulated to learn the optimal policy  $\pi^*$  such that:

$$\pi^* = \underset{\forall \pi}{\operatorname{argmax}} V^*(s), (\forall S) \quad (5)$$

A one-step look-ahead search can be performed from any state to determine the optimal policy using:

$$\pi^* = \underset{\forall \pi}{\operatorname{argmax}} [r(s, a) + \gamma V^* \delta(s, a)] \quad (6)$$

where  $r(s, a)$  is the current reward and  $\gamma V^* \delta(s, a)$  is the reward for the sequence of next steps.  $\delta$  refers to state transition function. For any infinite MDP that could possibly be solved using QL, the agent's ultimate goal is to maximize the total cumulative rewards:

$$\sum_{i=0}^{\infty} \gamma^i R_t \quad (7)$$

where  $R_t$  is the reward at time step  $t$ , and  $\gamma^i$  is the discount factor that defines the relative importance of the current rewards and those earned earlier ( $0 \leq \gamma \leq 1$ ).

For a nondeterministic environment, the QL methods usually follow the following steps:

- 1) Initialize the Q-Table: an initial Q-table with  $n$  columns and  $m$  rows should first be set, where  $n$  refers to the number of possible actions and  $m$  represents the



number of possible states. Since there are no reward records at the beginning, values for all cells are initialized as 0;

- 2) Choose and take an action: the agent chooses an action for the current state based on the Q-table. Based on the exploration and exploitation theory in reinforcement learning, the epsilon greedy strategy is introduced to speed up the learning process. The epsilon greedy strategy takes higher epsilon rates at the beginning stage to allow the agent to explore the environment by randomly choosing actions. The reason for this is because at the initial phase of the learning process, the agent has less knowledge on policy and needs to traverse as many state-action pairs as possible to learn. After multiple rounds of exploration, the agent traverses enough state-action pairs and receives sets of reward records. As the learning process moving forward, the epsilon rate decreases and the agent begins to exploit the environment. During the process of exploration, the agent progressively becomes more confident in estimating the Q-values.
- 3) Evaluate the state-action pair: instead of simply traversing state-action pairs, the agent updates the Q-table based on the token actions and observed rewards. For a non-deterministic environment, the Q-value is updated through Q-function:

$$\text{New } Q(s, a) = Q(s, a) + \alpha [R(s, a) + \gamma \max_{a'} Q'(s', a') - Q(s, a)] \quad (8)$$

where  $Q(s, a)$  refers to the Q value of the current adopted reward for state-action pair  $(s, a)$ .  $\text{New } Q(s, a)$  represents the new Q value for the state-action

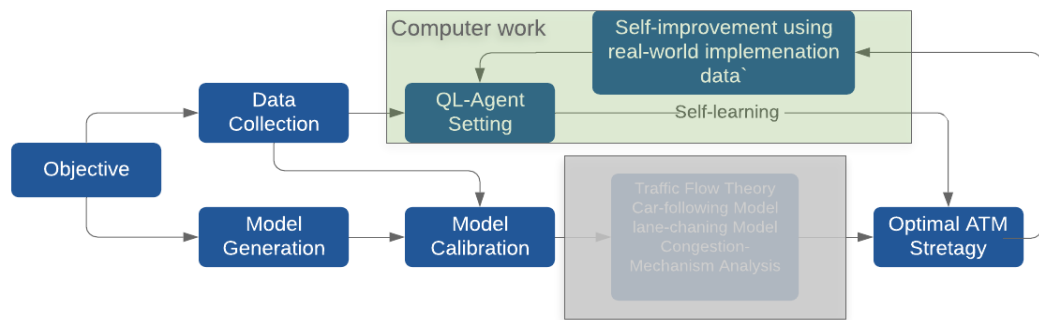
pair  $(s, a)$ .  $R(s, a)$  is the reward for taking that action at that state, and  $\gamma \max_{a'} Q'(s', a')$  refers to the maximum expected reward given the new state  $(s')$  and all possible actions at that new state. Learning rate and discount rate are represented by  $\alpha$  and  $\gamma$ , respectively.

Theoretically speaking, if the state and action are defined properly, the learning process can reach convergence after enough iterations of learning. Sometimes the learning rate and discount rate also influence the converging speed. When the learning process reaches convergence, the agent has traversed plenty of state-action pairs and received the corresponding rewards. The final product of the QL process is an updated Q-table, from which the optimal action (the action with the largest Q-value) for a state can be found. Now the agent can be used for the optimal control according to its knowledge.

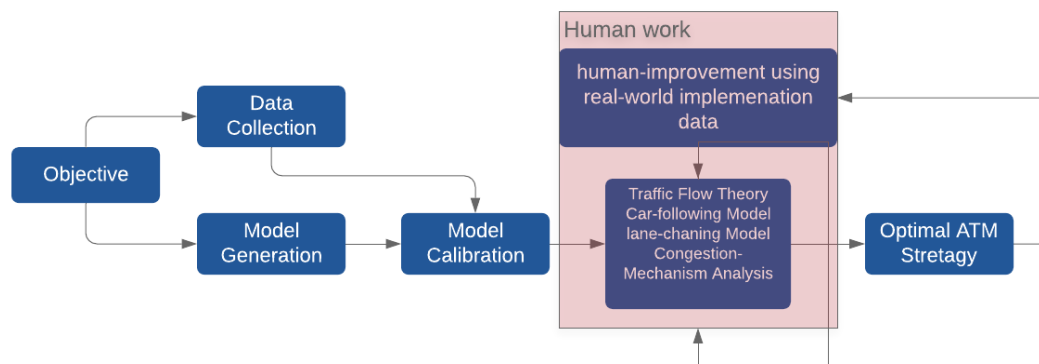
#### 3.4. Coordinated RL-based ATM Algorithm

As one of the most popular reinforcement methods, the Q-learning technique has offered a promising performance in dealing with the requirement of complex optimization. Recently more and more researchers have incorporate the QL method into traffic operation strategies. In the QL-based ATM strategies, the agent communicates with the environment by receiving states (traffic condition) and taking actions (updated traffic controls). A reward function is generated to evaluate the performance (traffic improvement) of each state-action pair. Even though the QL-based strategies require some time for offline training process, the agent can take optimal actions under various states without complex computation and accurate prediction

model. In this study, a QL-based ATM control strategy incorporating the variable speed limit and hard shoulder running strategies is proposed and tested to relieve traffic congestion. The operations of the proposed ATM strategy and traditional strategy are presented in Figure 3-3:



(a)



(b)

Figure 3-3: (a) Traditional ATM Algorithm; and (b) Proposed QL-Based ATM Algorithm

As indicated by the flowchart, the QL-based strategy increases the efficiency by replacing manual efforts with computation work. The proposed coordinated strategy

is composed of three parts: a QL-based offline agent, an online ATM control simulator, and an offline model set. The following parts of the section outline the specification of the proposed strategy.

### 3.4.1. State of the QL-Based ATM Strategy

State is the consolidation of horal, spatial, and material information. A state reflects a step change in the environment. The more detailed information a state provides, the more accurate the solution is, but the more time is needed for the learning process. Balancing learning time and solutions' accuracy is a common consideration when using machine learning methods. Typically, learning time increases exponentially as the number of state variables increase.

The objective function of most ATM strategies is minimizing total travel time in the system, which can be represented by the inflow and outflow traffic in the network:

$$TTT = q_{in}^0 - q_{out}^0 + q_{in}^1 - q_{out}^1 + \dots + q_{in}^n - q_{out}^n \quad (9)$$

Where  $q_{in}^k$  refers to the inflow during time interval k and  $q_{out}^k$  means the outflow during time interval k. If we consider  $q_{in}^k$  and  $q_{out}^k$  together as the discharging rate in the system at time interval k, TTT can be reformulated as Equation 11:

$$TTT = Q^0 + Q^1 + \dots + Q^k = \sum_{\kappa=0}^K Q^k \quad (10)$$

$$TTT = \sum_{k=1}^K [Q^0 + \sum_{\kappa=0}^{k-1} q_{in}^{\kappa} - \sum_{\kappa=0}^{k-1} q_{out}^{\kappa}] \quad (11)$$

Comparing the system travel time during one period, the cumulated outflow travel time equals to the cumulated outflow travel time under uncongested condition plus the queue delay caused by traffic congestion. Now the system travel time can be transformed to a new format using queue delay. For the given system, the initial demand and inflow are unchangeable. That is to say, any strategies that aims at decreasing the cumulated queue delay tend to decrease system travel time.

In contrast to previous studies that applied traditional variables such as speed, density, and volume, the current study takes the queue delay as the variable in the objective function. Although the link-based volume, speed and density well represent the average and the standard deviation of traffic in each link, they could not provide an overall view of the network. For some optimization-based strategies, bottleneck condition was mostly used as the only state variable. To minimize the negative effect on upstream VSL control section, the average queue delay of the entire corridor was designed as one state variable. Additionally, the posted speed limits were also included.

It should be mentioned that even though queue delay is a discrete variable, the states may have higher dimensions. Based on the analysis of a non-control scenario, the queue delay has the range from 0 to 215 per lane per time interval. To decrease the training time of the agent, this study classifies the queue delay and queue delay reduction into five levels (Table 3-1) and thirteen levels (Table 3-2).

Table 3-1. Level of Queue delay

<b>Level</b>	<b>Range</b>
1	$N_Q \leq 30$
2	$30 < N_Q \leq 60$
3	$60 < N_Q \leq 120$
4	$120 < N_Q \leq 150$
5	$150 < N_Q$

Table 3-2. Level of Queue Delay Reduction

<b>Level</b>	<b>Range</b>
-3	$\Delta_Q \leq -75$
-2	$-75 < \Delta_Q \leq -35$
-1	$-35 < \Delta_Q \leq 0$
0	$0 < \Delta_Q \leq 20$
1	$20 < \Delta_Q \leq 35$
2	$35 < \Delta_Q \leq 75$
3	$75 < \Delta_Q$

*Similar to the queue delay, density also meets the same condition. It is obvious that density is a continuous variable that should also be aggregated into discrete dimensions. The result of non-control scenario indicates that the density varies from 0 to 176 vehicles/mile/lane. Therefore, the density used in this study is categorized into 7 levels (Table 3-3). Table 3-3. Level of Density*

<b>Level</b>	<b>Range</b>
1	$D \leq 35$
2	$35 < D \leq 75$
3	$75 < D \leq 100$
4	$100 < D \leq 125$
5	$125 < D \leq 150$
6	$150 < D \leq 165$
7	$165 < D$

Above all, the proposed strategy takes the level of queue delay and queue delay reduction combined with the density of VSL control segment as three state variables.

### 3.4.2. Actions of the QL-Based ATM Strategy

Actions are defined as actionable activities in creating or resolving a problem in a follow-up step: what to do next. In this study, an action refers to the control of variable speed limit and hard shoulder running. The control period should be set to ensure that the effect (or reward) after the action taken can be perceived by the agent. In our study, the authors tested two control periods for the QL-based strategy which are 5 minutes and 10 minutes. In other words, the QL-based agent is able to post new speed limits or update shoulder lane status after 5 minutes and 10 minutes from the previous action. Some restrictions may need to be set for practical considerations such as to ensure drivers are not confused due to frequently changing speed limits. For example, an ideal posted speed limit could be either whole numbers (e.g., integers) or numbers with decimal points. In practice, control speed limit displayed on a dynamic message sign does not have decimal digits to reduce driver confusion [80]. Based on real world experience, the actions should follow the following rules:

- 1) Control speed limits are discrete and are multipliers of five (e.g., 5, 10, 15, 20, et al.);
- 2) Control speed limits on a freeway corridor should never exceed the maximum speed limit or the free flow speed;
- 3) Control speed limits should always be higher than one or several lower bound speed limits, even if the corresponding congestion is severe. This is to avoid confusing drivers and to increase traffic flow efficiency;

- 4) Control speed limits should not change with significant difference. The posting of significant speed differences tends to cause sudden acceleration or deceleration, which may worsen the existing congestion.

Hard shoulder running improves traffic congestion by increasing the capacity at bottlenecks and corresponding upstream roadway segments. The activation or deactivation of a shoulder lane requires the coordination of speed limit signs, notice light signs, relief zones, and variable message signs (VMS). In the proposed strategy, actions include:

- 1) Increasing the control speeds by the same amount (5 mph) on VSL controlled segments;
- 2) decreasing the control speeds by the same amount (5 mph) on VSL controlled segments;
- 3) Increasing the control speeds by the same amount (10 mph) on VSL controlled segments;
- 4) Decreasing the control speeds by the same amount (10 mph) on VSL controlled segments;
- 5) Increasing the control speeds by different amounts (5 mph or 10 mph) on VSL controlled segments;
- 6) Activating shoulder lane;
- 7) Deactivating shoulder lane.



### 3.4.3. Rewards of the QL-Based ATM Strategy

In terms of previous studies on ATM strategies, the judgment for the performance could be the reduction of travel time, the reduction of total delay, and/or the improvement of throughput. As just discussed, the objective of the proposed strategy is to minimize network travel time, that is, to minimize the cumulated queue delay in the network. In the simulation model, the number of queued vehicles on each link at each time interval was recorded and updated in real time. Refer to previous studies, any control that manage to increase the early exit flows of the freeway section will lead to a decrease in the total travel time. Therefore, in addition to total network queue delay, bottleneck queue delay was included as part of the reward function. To increase learning efficiency, the reward function uses an exponential distribution function, which allows the agent to get relatively large reward or penalty when the queue changes largely. In the study, the reward function is designed as:

$$R(s) = -\alpha(Q_k^s - Q_{k-1}^{s'})|(Q_k^s - Q_{k-1}^{s'})| - \beta Q(b)_k^s \quad (13)$$

Where  $Q_k^s$  represents the cumulated queue delay in the network at state  $s$  during time interval  $k$ , and time step  $k$ .  $Q(b)_k^s$  refers to the queue delay at the bottleneck location at state  $s$  during time interval  $k$ . The parameter  $\alpha$  and  $\beta$  are introduced to determine the magnitude of the reward. The exponential distribution function helps the learning process reach the convergence at an accelerated speed. With the designed reward function, the agent tends to learn how to provide optimal sequence of actions for any given state. The pattern of reward distribution is shown in Figure 3-4. As the

figure presents, the green dots represent the positive rewards while red ones refer to negative rewards. The size of the dots represent the absolute value of the penalty.

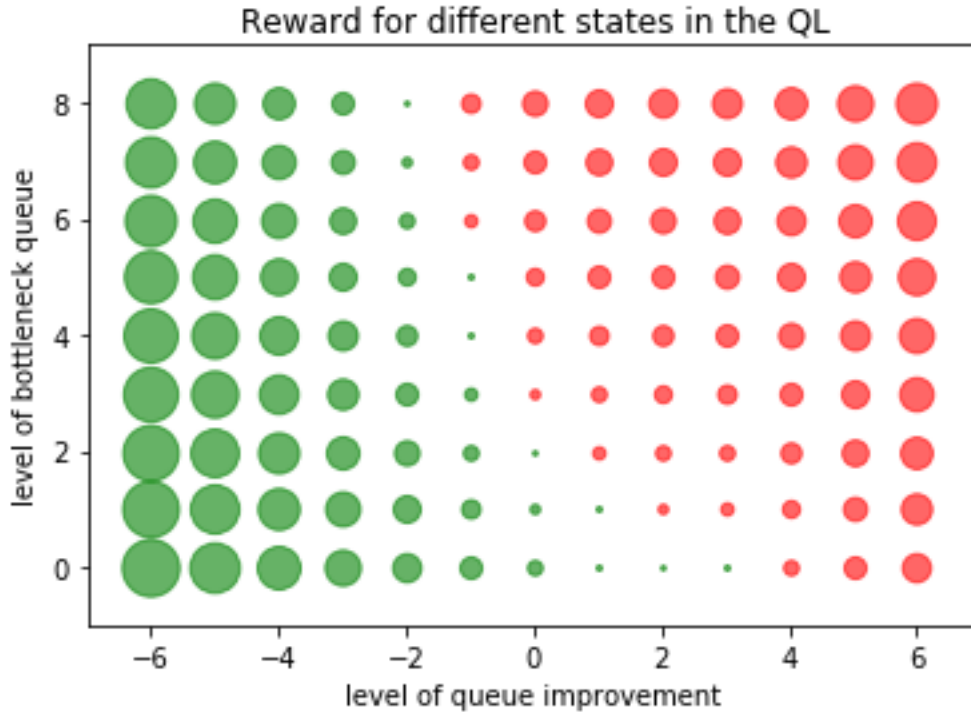


Figure 3-4. Rewards for different state in the QL.

In addition to VSL, HSR is also an integral component of the proposed strategy. The cooperation of all strategies contribute to utilizing resources. As one of the two controls, HSR plays a distinctive role in improving the performance for its rapid supplement of road capacity. While HSR is effective, it is not encouraged be active all the time for safety concern. Under normal condition, a freeway with multiple lanes usually has two shoulders in each direction. Closure of any shoulder lane increases the difficulty for a needy vehicle to relocate to the nearest shoulder when such a need arises. Also, if a needy vehicle fails to move to the shoulder lane, it will result in blocking regular traffic lanes and causing unusual congestions. Meanwhile, the operation of HSR

is costly, which requires expensive guiding signals, warning signs, cameras, speed limit signs and VMS. In addition, the occupancy of shoulder lanes for any prolonged period increases the burden of emergency service. To deal with this issue, this study enforces HSR to be activated only when the VSL alone is not effective. Several studies suggest that VSL only provides good performance in less serious conditions but fails to provide reliable control under over-saturated conditions. Therefore, the coordinated strategy designed for this study should focus on finding the time point to activate HSR, which is when the VSL loses effectiveness. To help the agent learn when to activate HSR, a penalty function is integrated in the reward function considering the safety and traffic throughput impact. After multiple tests, the penalty function is designed as:

$$P(s) = \frac{\theta}{[(Q_k^s) * (Q(b)_k^s)]^\varepsilon} \quad (14)$$

The penalty function indicates that the HSR only activates when both the corridor queue delay and bottleneck queue delay are high.

#### 3.4.4. Parameters of the QL-Based ATM Strategy

In the QL process, several parameters need to be properly defined since these variables greatly affect the performance of the strategy.

- 1) Learning rate: learning rate refers to the learning speed with range 0 ~ 1 that controls how quickly the agent communicates with the environment. A setting of 0 means that Q-values are never updated, hence nothing is learned, and while a higher value means that learning occurs quickly. Learning rate is a tricky variable in a QL process and should be properly defined. If the learning rate is too small, the learning process completes complex computation too frequently.

The heavy computation burden prevents the learning process from reaching convergence. However, a large value may result in the large variance of each iteration, and the agent may not find optimal solution.

- 2) Discount factor: the discount factor defines the relative importance of the current rewards and the rewards earned in the previous steps. In other word, the discount factor decides the ‘sight’ of the agent. A discount factor close to zero forces the agent to be ‘short-sighted’ by only considering the reward in the immediate future. In contrast, a factor closer to one makes the agent forecasts the cumulative rewards that expects to get. With a higher value, the agent tends to explore the higher cumulative rewards for the next few steps during the learning process.
- 3) Epsilon-greedy: the epsilon-greedy decides the rates of exploitation and exploration. A high epsilon-greedy value means that the agent tends to explore more while a low value means that the agent chooses an action based on the estimated value. For the learning process, the agent should explore more at the beginning stage to traverse as many action-state pairs as possible. However, a continuous higher exploration rate may prevents the agent learning from the accumulative results. In the current study, the epsilon-greedy is set as a high value at the beginning and decreases systematically.
- 4) Learning iteration: the current study undergoes many iterations to ensure the QL algorithm has enough time to converge.

The detailed setting of learning parameters is summarized in Table 3-4.

Table 3-4. Learning Parameters of QL Algorithm

Learning parameter	Value
Learning iterations	150~500
Learning rate	0.005
Reward decay	0.99
Discount factor	0.9
Start epsilon-greedy	0.9

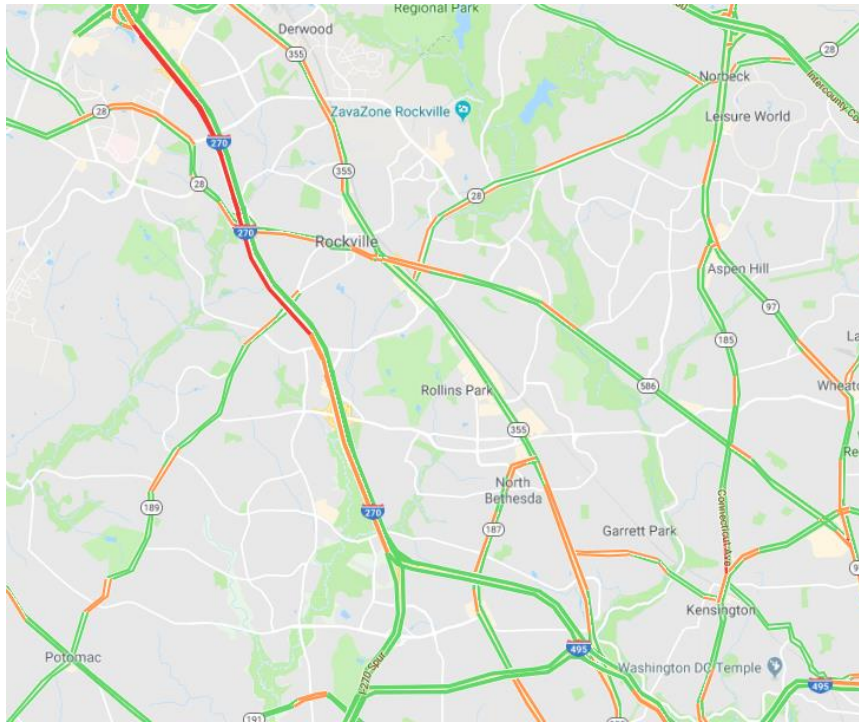
3.5. Study Area Description

Within any metropolitan area, traffic congestion associated with urban freeways is typically the worst. To evaluate the feasibility and benefits of the proposed algorithm, the I-270 corridor in suburban Maryland outside of Washington, D.C. is selected as the experimental freeway segment, with a large study area including the I-495 beltway from America Legion Memorial Bridge to Woodrow Wilson Memorial Bridge, and MD-295 from US-50 to I-695. The model also covers all on-ramps, off-ramps and a significant number of local roads.

I-270, one of the most important freeways connecting Washington D.C. to Frederick, MD, experiences heavy congestion during morning and evening peak periods on any weekday. While free flow travel time is only about 29 minutes, travel time reaches as high as 67 minutes and 52 minutes for the AM and PM peak periods, respectively. Based on historical data from RITIS, weekday peak-period speed drops quickly from 65 mph to 30 mph. In several bottleneck areas, the speed was as low as 10 mph, which caused further congestion spill back. There are two main reasons for

traffic congestion: high traffic demand causing recurrent congestion and incident blockage, causing a sudden capacity drop that leads to non-recurrent congestion. RITIS Data indicates that more than 150 incidents occurred during the peak period on I-270 in 2016. These incidents resulted in an average 24.6 minutes delay during peak hour.

In this paper, the model is built to represent the real-world diverging bottleneck at the I-270 spur (Figure 3-5). At the diverging point, the five-lane I-270 mainline splits into two three-lane freeways that lead to the counter-clockwise direction and clockwise-direction of I-495, respectively. Although the total capacity of the two branches satisfies the needs of inflow from upstream, what matters more is the misdistribution of traffic demand. The counter-clockwise direction of 495 has higher demands from the I-270 spur than the clockwise direction. Another reason is the frequent lane changing at the diverging point, since some drivers take advantage of the lighter traffic of one side and merge into the high-demand flow at the last moment. The frequent lane changing close to the diverging point also results in the sharp deceleration and increases the spill back of traffic congestion. Additionally, the American Legion Memorial Bridge at downstream is one of the segments with high incidence of congestion, which also affects the I-270 diverging section.



Source Google Map (<https://www.google.com/maps>)

Figure 3-5. Traffic condition at I-270 spur (morning peak).

Table 3-5. Components of DTA Simulation Model (Initial MWCOG Model and Subarea Model)

	<b>Large-initial model</b>	<b>Subarea model</b>
Number of nodes	23589	9860
Number of links	47511	28287
Number of zones	3722	2030
Number of HOV links	158	107
Number of agents	4925741	2434299
Number of demand types	24	24
Number of vehicle types	5	5

### 3.6. Simulation Model

#### 3.6.1. Online Simulation Model

Most ATM algorithms are integrated with traffic simulators, including macroscopic, microscopic, and mesoscopic models. Taking advantages of the mesoscopic model, which can perform complicated simulations within a reasonable period, DTALite, a dynamic traffic assignment (DTA) model written with open-source mesoscopic scale, was adopted in the study [81]. DTALite integrates many advanced simulation functions and offers a visualized user interface named Network Explorer for Traffic Analysis (NEXTA). The attractive functions model the traffic impacts of road reconstructions, special traffic conditions, active traffic managements, tolling systems, and general evolution of demand. The efficient running mechanism of DTALite also supports the free combination of multiple scenarios, which helps decision-makers find optimal solutions to improve traffic operation systems. The crucial components of DTALite software are:

- 1) Agent-based traffic assignment: reassigns agents in the network to reach the user equilibrium. The user equilibrium is usually evaluated by the travel time index (TTI) which is a ratio of average simulated travel divided by the free flow travel time.
- 2) Origin destination matrix estimation (ODME): adjust the OD matrix by minimizing the absolute error between observed traffic counts data and simulated results. The error of traffic counts is usually calculated using weighted mean squared error.



Research based on the DTALite simulation model proved the model's ability in simulating dynamic traffic operation algorithms [81-82]. Reasons for adopting DTALite as the simulation model in this study are listed below.

- 1) Difference compared to macroscopic and microscopic model: the DTALite is a mesoscopic model and can provide detailed simulation results with short running time and low memory need.
- 2) Built-in simulation functions: the DTALite has built-in functions to simulate ATM strategies such as variable speed limit and ramp metering using lane capacity and dynamic speed control.
- 3) Rigorous traffic queuing model and build-in parallel computing capability: the queuing model and parallel computing capability can speed up the analysis process through multi-core CPU hardware [82].
- 4) Real-time information: during the customized setting of updating time interval, the DTALite provides real-time system-level and link-level statistical outputs that describe time-dependent network performance, such as volume, density, speed, number of queued vehicles, and bottleneck locations. These time-dependent outputs help the proposed algorithm make decision to support dynamic controlling;
- 5) Agent-based inputs: agent-based input allows DTALite to analyze the route choice changes under ATM control [83];
- 6) NEXTA: the interface of DTA model provides a visual animation of vehicles running on the links for a large-scale network over the simulation period.

### 3.6.2. Model Generation

For the analysis of regional impact, this study covers the Washington Metropolitan Area, which encompasses Maryland, Washington D.C., and Northern Virginia. The base-year (2015) Metropolitan Washington Council of Governments (MWCOCG) travel demand model is used as the seed input for model developing. The geocoded MWCOCG model in a GIS environment is converted into DTA network format. Preparatory volume calibration and origin-destination matrix estimation (ODME) process are applied to ensure the demand is consistent with observed data. Instead of using the large network, the MWCOCG model was cut into smaller subarea model for time-consuming concern. The subarea network that contains 2,030 TAZ zones, 9,860 nodes, and 28,287 links, is displayed in Figure 3-6.

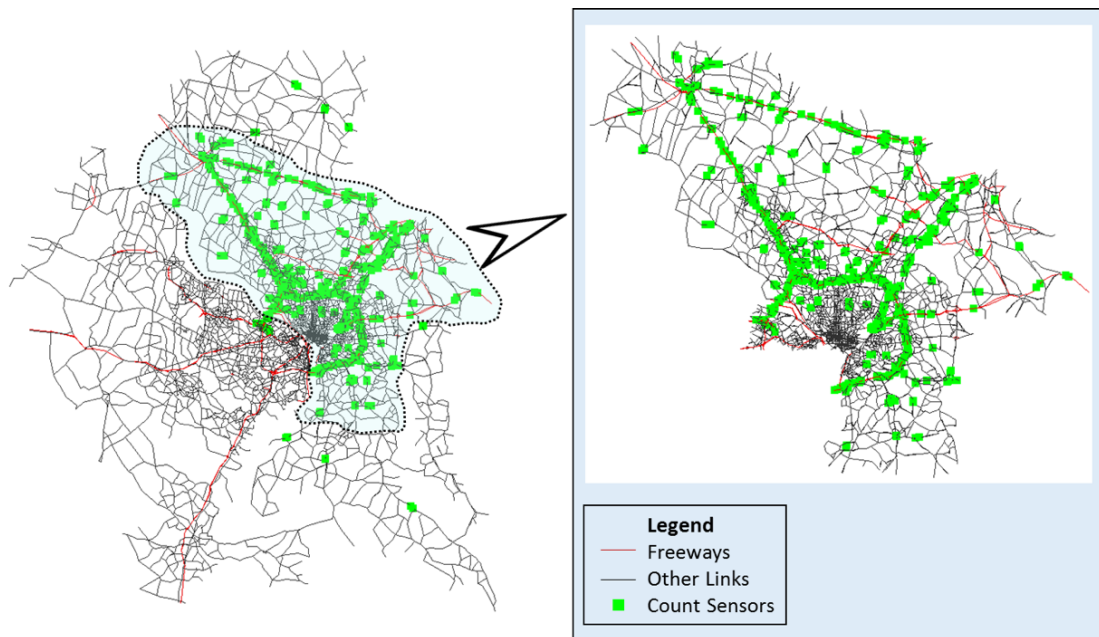


Figure 3-6. The Mesoscopic Simulation Network and Locations of Traffic Counts.

Table 3-6. Components of DTA Simulation Model (Initial MWCOG Model and Subarea Model)

	<b>Large-initial model</b>	<b>Subarea model</b>
Number of nodes	23589	9860
Number of links	47511	28287
Number of zones	3722	2030
Number of HOV links	158	107
Number of agents	4925741	2434299
Number of demand types	24	24
Number of vehicle types	5	5

### 3.6.3. Model Calibration and Validation

Model calibration and validation process are critical for building a consistent and reliable model before integrating any control algorithms. Several key parameters such as volume, travel time, and speed are calibrated in the study. Link-based volume is calibrated based on the time-dependent (e.g., hourly) link volume data collected by real-world detectors. Traffic data from 179 sensors used in the Regional Integrated Transportation Information System (RITIS) and State Highway Administration (SHA) Internet Traffic Monitoring System (I-TMS) from 2015 are utilized (Figure 3-6). In the DTALite model, a build-in calibration function, known as ODME, is used as the approach to adjust input OD. Since the automated adjustment may not be reliable when the corridor is congested, the authors adopt a manual process to adjust the demand side to reduce the gap between observed and simulated counts.

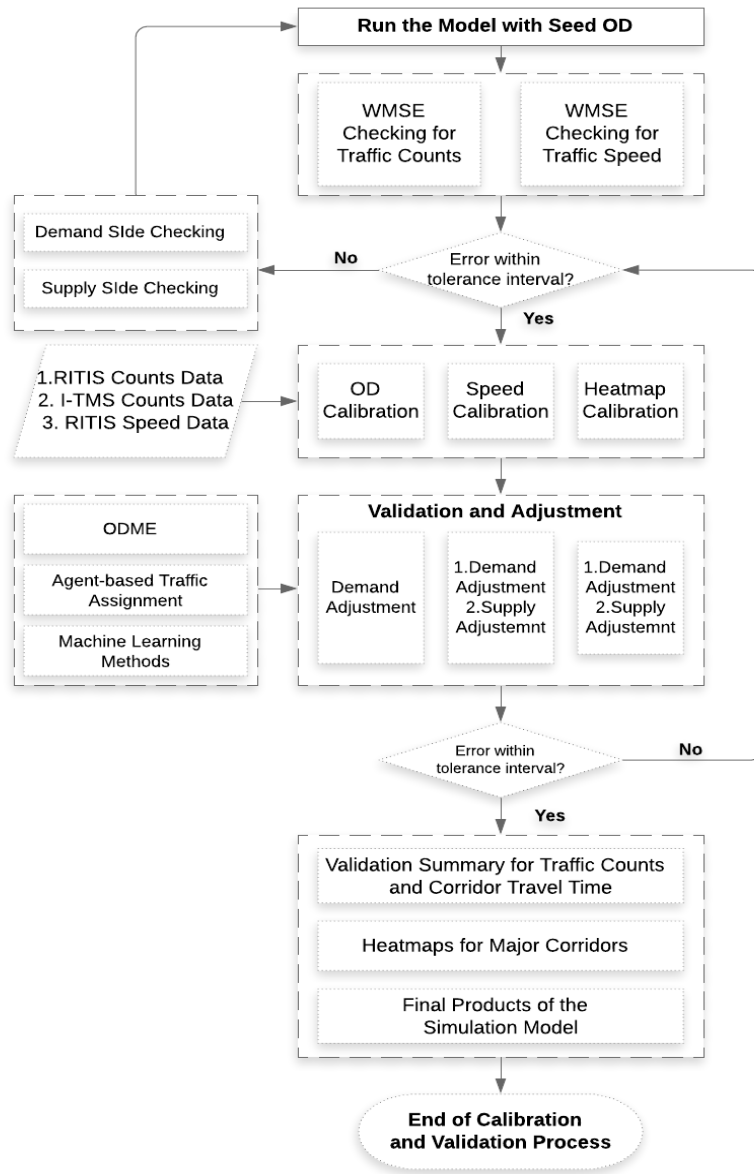


Figure 3-7. Calibration and Validation Process for A DTA Model

Beyond volume, corridor travel time must also be consistent with real-world data. In this study, observed travel time data from the RITIS website is used as the reference for corridor travel time calibration. Average 15-minutes of travel time on

each freeway corridor is calibrated against simulated results. In addition to the demand side calibration, the supply side parameters such as lane capacity and jam density are also adjusted. Figure 4 demonstrates the calibration and validation process.

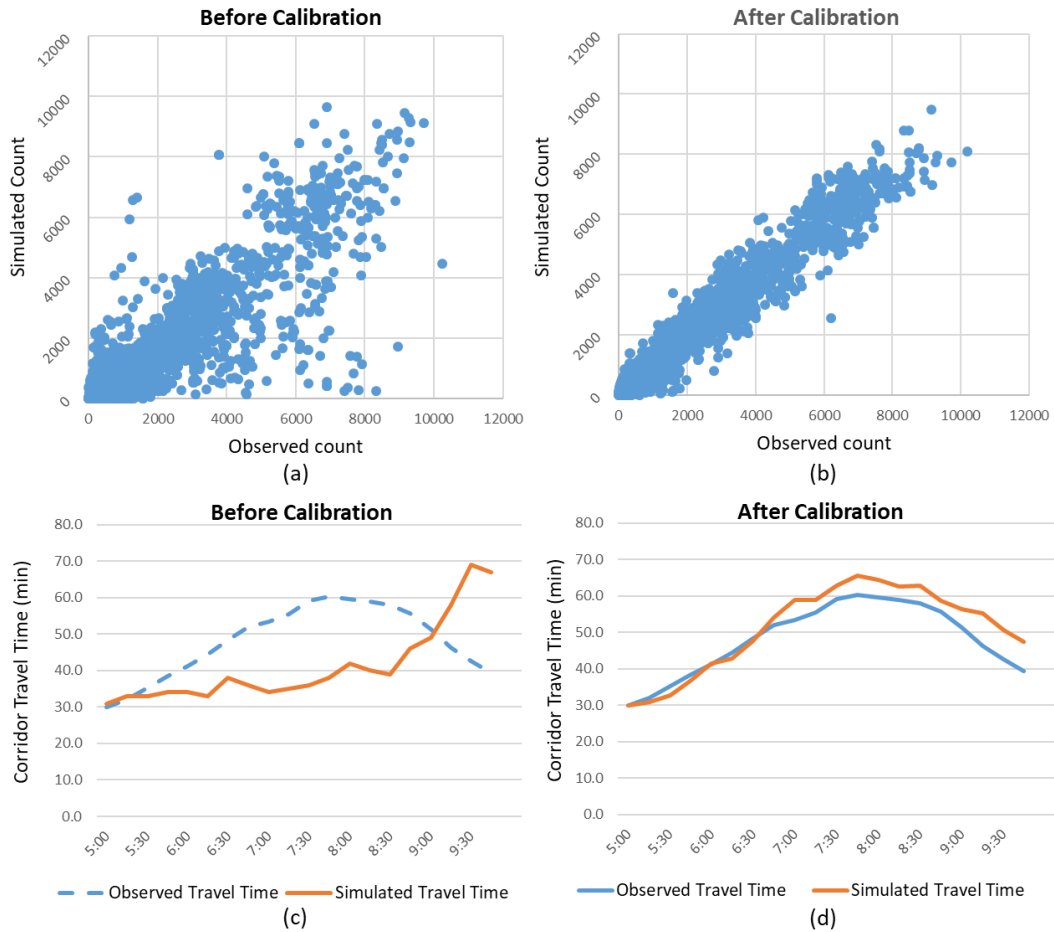


Figure 3-8. (a) Traffic Volume Comparing before Calibration; and (b) after Calibration. (c) Corridor Travel Time Comparing before Calibration; and (d) after Calibration

For validation, weighted mean squared error (WMSE) is used to validate the corridor travel time using the following calculation method:

$$WMSE = \sqrt{\frac{\sum_{i=1}^N \sum_{t=1}^T (y_{i,t}^* - y_{i,t})^2}{\sum_{i=1}^N \sum_{t=1}^T y_{i,t}^2}}$$

$y_{i,t}^*$  and  $y_{i,t}$  denote the observed and simulated value, respectively, at each link  $i$  during time interval  $t$ .  $N$  is the total number of sensors, and  $T$  is the total number of time intervals. For most simulation-based analysis, WMSE is accepted below 15 or 20% in practice for model calibration. As shown in Figure 5, validation results indicate that the WMSE decrease to 7.91% from 32.74% for volume, and to 10% from 40% for corridor travel time after calibration.

#### 3.6.4. ATM Implementations

A single-direction freeway of a total length of 5.5 miles on I-270 Southbound is coded as the ATM control segment, as shown in Figure 3-9. The mainline in this segment has two types of lanes that represent the general-purpose lanes (GPL) and high-occupancy lanes (HOV), respectively. In this study, ATM control is only applied on GPL. A 4.25-mile segment upstream to the diverging point is included in the VSL control segment. The 0.75-mile segment close to the diverging point is the acceleration section without speed limit control, which helps vehicles accelerate to free flow speed. The acceleration section can remain critical density and the outflow keeps close to bottleneck density. At further upstream, a 3.5-mile segment is controlled by VSL. The advisory speed limits in the VSL control section mainly follow a declining curve to prevent the downstream congestion from growing too fast and blocking the ramps. The 3.5-mile VSL control section was divided evenly into 7

parts, the lengths of which are expected to be similar but not exactly the same.

Considering real-world applications, dynamic message signs are placed along the road to present the current speed limit, active warnings, and prohibitions.

The sudden capacity drop at the diverging area creates the bottleneck and the congestion spills back to upstream with the increase of demand. To directly provide additional capacity, 1-mile segment at bottleneck is designed as HSR control section, including a 0.25-mile buffer zone at upstream and 1-mile convertible zone at downstream. After this part of HSR section, the shoulder lane will change to a special-occupied section that can only be used by vehicles leaving the freeway using the closest off-ramp. This design helps vehicles smoothly drive out of the fully-occupied section and avoid the conflict with off-ramp and on-ramp.

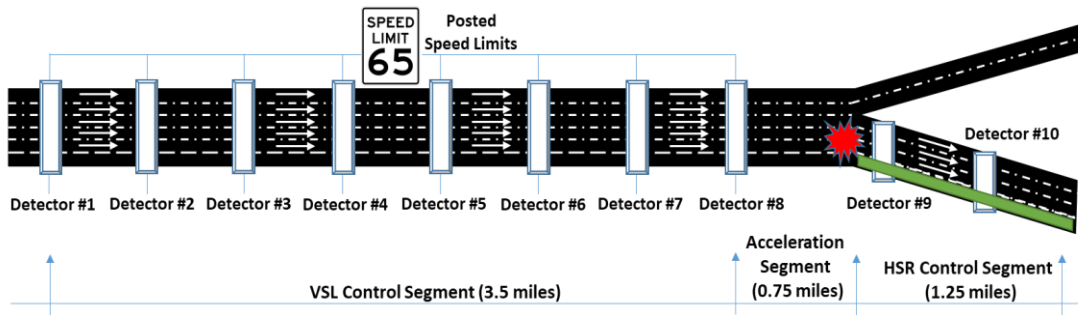


Figure 3-9. Study Area and Implementation of ATM Control

### 3.7. Parameters Setting

The simulation covers the morning peak period from 5:00 am to 11:00 am. The first two hours is a warm-up period that leads the system to a steady state and is excluded from analysis. The supply side parameters such as capacity, free flow speed, and length are kept in consistent with the MWCOG model. Other parameters are set

based on the previous experience of traffic simulation. Other defined supply side parameters and simulation configurations are shown in Table 3-6 and Table 3-7:

Table 3-7. Supply Side Parameters for Simulation Model

<b>Traffic parameter</b>	<b>Value</b>
Free flow speed (miles/hour)	65
Lane capacity (Freeway) (vehicles/mile/lane)	2000
Lane capacity (Ramp) (vehicles/hour/lane)	1500
Wave speed (miles/hour)	12
Critical density (vehicles/mile/lane)	30.7
Jam density (vehicles/mile/lane)	180
Percentage of dropped capacity (%)	10
Discharge flow after capacity drop (vehicles/mile/lane)	1800

Table 3-8. Scenario Parameters for Simulation Model

<b>Setting</b>	<b>Value</b>
Warm period	4:00 am ~ 6:00 am
Cool period	9:00 am ~ 11:00 am
Simulation period	6:00 am ~ 9:00 am
Signal control representation	Continuous flow with link capacity constraint
Traffic flow model	Newell's cumulative flow count model
Routing method	Without OD demand estimation
Number of learning iterations	120
Number of assignments per iteration	20



The simulation model considers six driver classes: SOV, HOV2, HOV3, APV, COM, and TRK, of which the percentage is calculated based on MWCOG 2015 base demand. The ratios of each demand type are listed in Table 3-8.

Table 3-9. Number of Vehicles for Different Demand Types

	<b>SOV</b>	<b>HOV</b>	<b>COM&amp;APV</b>	<b>TRK</b>
<b># of Vehicles</b>	1545170	72147	33906	117258
<b>Ratio</b>	63.8%	29.9%	1.4%	4.9%

The value of time (VOT) is also considered based on the SHA for Maryland I-270 and I-495 Level 2 Traffic and Revenue Study. The study estimates the VOT on a county basis using each county’s average median household income, the average number of hours worked per household per year, and by considering the breakdown of trip purposes and applying perception weighting factors for the trip purposes. Some regions and demand types are combined in the study to decrease the dimensions. Table 3-9 presents the VOT estimation for each demand type of each region in the MWCOG study area using the introduced approach.

Table 3-10. Value of Time (\$) for Different Demand Type

	<b>Region 1</b>	<b>Region 2</b>	<b>Region 3</b>	<b>Region 4</b>	<b>Region 5</b>	<b>Region 6</b>
<b>SOV</b>	15.6	18.0	18.9	19.8	21.6	24.0
<b>HOV</b>	18.7	21.6	22.7	23.8	25.9`	28.8
<b>COM&amp;APV</b>	18.7	21.6	22.7	23.8	25.9	28.8
<b>TRK</b>	57.7	66.6	69.9	73.3	79.9	88.8

In the simulation model, there are default detectors on each link to provide real-time traffic conditions to the reinforcement-learning agent. The real-time data includes inflow, outflow, speed, density, and number of queued vehicles. This setting is the ideal situation for simulation models, while the arrangement of detectors needs deeper analysis for field implementation in the future. In this study, the model provides one-minute interval traffic information, which presents the best performance after several experiments.

Various scenarios are simulated in this study. That is, different combinations of techniques are tested for different time intervals: 5 minutes, and 10 minutes. The scenarios include:

- 1) No control scenario: the traffic condition is the same with real world.
- 2) QL-based VSL: only variable speed limit control could be activated during the simulation.
- 3) QL-based HSR: only hard shoulder running control could be activated during the simulation.
- 4) QL-based VSL&HSR: both variable speed limit control and hard shoulder running control could be activated during the simulation.

## Chapter 4 Simulation Results

This study develops a coordinated ATM control method that integrates the variable speed limit and hard shoulder running. To fully validate the advantages of coordinated control strategy and the integrated reinforcement learning techniques, the authors test and compare several scenarios. This section summarizes the simulation results of different scenarios.

### 4.1. MOEs

Measures of effectiveness, also known as MOE, are the measures selected to quantify if the results accomplish the objectives or conformity of expected results. For some model-based experiments, MOE are especially crucial to evaluate the proposed algorithm and provide effective suggestions for model improvement. In this study, corridor travel time, bottleneck speed and system travel time are defined as three major measures. Additionally, other performances such as queue delay on ramp and density are also considered.

- 1) Corridor travel time (ATM control segment): corridor travel time is the direct way to evaluate the improvement of the entire corridor. Also, what mostly attracts users may not be the density or speed improvement at the bottleneck location but the reduction of estimate travel time on the corridor.
- 2) Bottleneck speed and density: bottleneck speed and density directly present the influence of ATM control on bottleneck location.

- 3) System travel time: system travel time is the more representative measure to present what affects the proposed algorithm have on the entire system. The ATM control on I-270 not only influences the bottleneck location, but also affects other corridors.
- 4) Ramp queue delay: Even though the proposed strategy controls freeway mainline, ramps are inevitably affected. Excessive delay on ramps may kill drivers' patience and decrease compliance rate.

As a benchmark, the author also tests a feedback-based VSL strategy on the same network to compare with the proposed algorithm. Details of the feedback-based strategy are discussed in a research paper of Zhang et al [83].

#### 4.2. Results for Non-control Scenarios

The fully calibrated model without ATM controls is the mapping of the actual traffic condition. As presented by the blue curve in Figure 4-2, the speed at I-270 diverging point suddenly decreases around 5:50 a.m. Within a very short time, the speed drops from 60 to 40 mph and then 10 mph. Meanwhile, the volume at I-270 mainline reaches 7,500 vehicles/hour, of which approximately half are assigned on to the I-270 Spur toward I-495 CCW direction to Virginia. The demand on I-270 Spur increased to almost 1600 vehicles/hour/lane, which is close to the roadway capacity, stays at this level, and remains for 2.5 hours. The insufficient roadway capacity could not satisfy the continuous increasing traffic demand, and then gradually form the bottleneck. Additionally, the sudden acceleration and deceleration among high-density traffic also contributes to the congestion. During this period, the density at bottleneck

stays around jam density (180 vehicles/mile/lane) and the spilling back queue keeps that of entire corridor at high level. Several upstream links (S4 to S7) of active bottleneck are highly affected by the shockwave and the speed dropping, resulting in the corridor travel time increases to 18 minutes, more than twice the free flow travel time (Figure 4-1). It can also be observed that the bottleneck speed decreases in staggering speed, that is to say, the queue also spills back at express speed. This is just more proof that a control strategy with good prediction model and quick response is of the essence in releasing congestion.

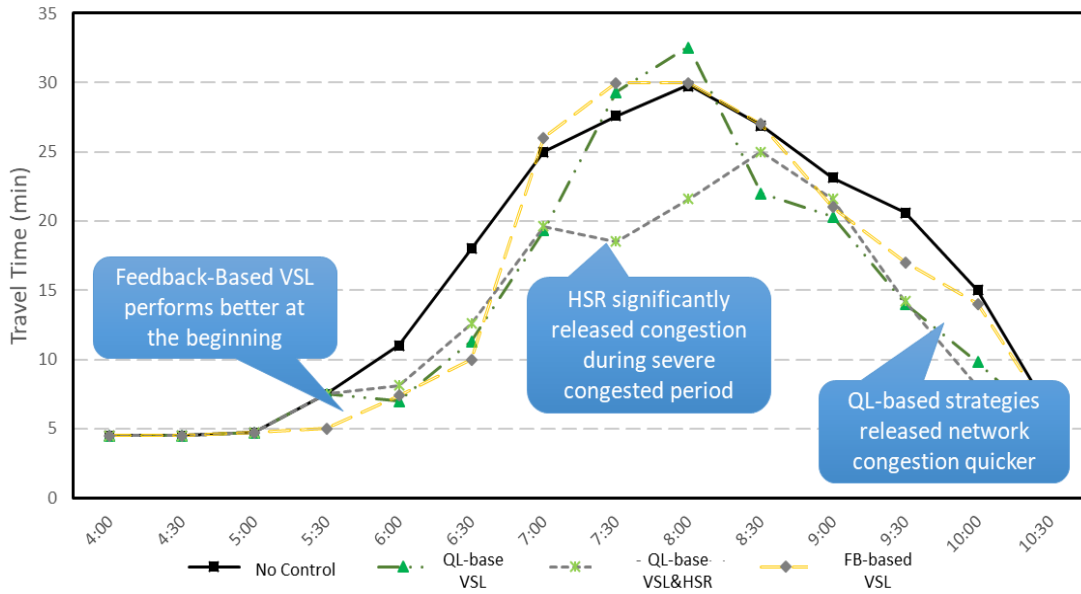


Figure 4-1. Corridor Travel Time for Different Scenarios

Table 4-1. Corridor Travel Time Summary

Control Scenarios	Corridor Travel Time (min)	Improvement (%)
No Control	21.91	/
Single VSL Control	18.40	-15.98
VSL&HSR Control	16.01	-26.94
Feedback Control	20.09	-8.20

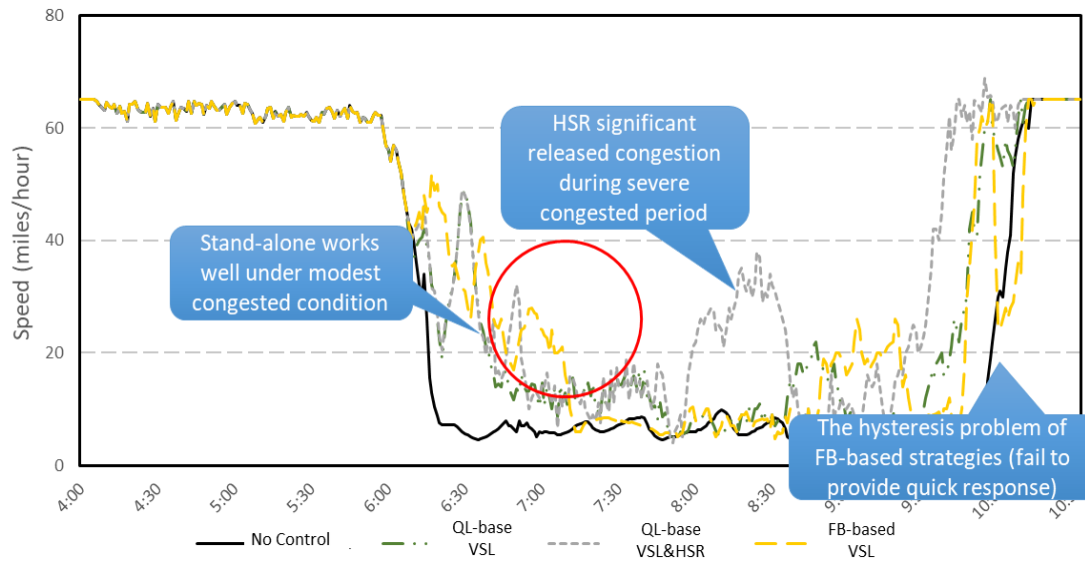


Figure 4-2. Bottleneck Speed for Different Scenarios

Table 4-2. Bottleneck Speed Summary

Control Scenarios	Bottleneck Speed (mph)	Improvement (%)
No Control	16.20	/
Single VSL Control	23.57	+45.59
VSL&HSR Control	29.32	+80.98
Feedback Control	23.02	+42.10

#### 4.3. Results for Feedback-Based VSL Control Scenario

The feedback-based VSL control strategy is tested as the benchmark to evaluate the performance of the proposed strategy, as shown by the orange curve in Figure 4-2. Different from non-control scenario, the bottleneck speed fluctuates several times from 5:50 to 6:15 a.m. instead of keeping falling to the bottom. Even though the traffic condition is still not optimistic for the conversation, implementation of the feedback-based strategy has kept traffic speed at bottleneck above 20 mph, and density below 150 vehicles/mile/lane during most of the time. It can also be observed from the figures

that the congestion dissipates faster than the non-control scenario, which is a strong demonstration that VSL control can increase traffic throughput. Toward the end of congested period, the orange curve has a significant falling down after recovering to free flow speed. This is attribute to the quick increase of posted speed limit, which is an inevitable result of lacking the benefits from traffic state prediction. A feedback-based strategy works based on the feedback from the environment, to some extent, it is a reactive operation that response to the traffic condition passively. Even though the traffic condition could be improved, the performance is limited by the feedback nature. This explains the feedback-based strategies has limitation in providing quick response especially to the rapidly changing traffic conditions. Extends to the ATM control section, although the feedback-based control has not completely eliminated speed dropping and spilling back queue, the shockwave has been suppressed to a great extent. In addition, Figure 4-1 indicates that feedback-based control benefits corridor travel time to a certain extent, average 8.2% lower than non-control scenario.

#### 4.4. Result for QL-based VSL and QL-based VSL&HSR Control Scenarios

The grey curve and yellow curve in Figure 4-2 represent the performance of QL-based VSL and QL-based VSL&HSR strategy, respectively. To begin with, let us compare the QL-based VSL strategy and the feedback-based VSL strategy. As far as the bottleneck speed is concerned, the QL-based strategy and feedback-based strategy differs in the trend but not markedly in the substance of average speed during entire congested period. However, we may still discover from Figure 4-1 that the curve of speed changed gently under the QL-based strategy. It is still fluctuating under the control, but could hardly see the sudden decrease from free flow speed which is similar

to the end of feedback-based scenario. This strongly confirms that the introduction of QL method redeems the limitation of traffic prediction in feedback-based strategy tremendously. Even though these two strategies present no significant difference in bottleneck speed improvement, the QL-based scenario outperforms the other one in corridor travel time reduction. All of these are profit from adopting cumulative queue delay on the entire corridor in the states and reward function of QL process. With the powerful self-learning ability, the QL-based strategies can better learn how to balance the performance of entire corridor instead of sacrificing upstream VSL control segment for bigger improvement at bottleneck.

Compared with stand-alone strategy, coordinated strategy drives the traffic system to improve on all fronts. There is no significant difference during the first 1 hour. However, the bottleneck speed experiences a sharp increase up to 40 mph from 8:00 to 8:30 a.m. This attributes to the direct increase of roadway capacity by the activation of hard shoulder running. After 8:30 a.m., the agent predicts there is a tendency for the congestion to moderate, so it deactivates shoulder lane control and only uses variable speed limit control. During the entire simulation, the average speed increases 29.3 mph, almost 80% higher than non-control scenario and 40% higher than VSL-alone scenarios. Benefits from the coordinated ATM control, the congestion dissipated 10 minutes earlier than stand-alone strategies. For the corridor travel time, hard shoulder running helps it decrease from 30 minutes to 20 minutes.



Table 4-3. Summary of All MOE

	I-270 SB Freeway Mainline				I-270 SB Ramps	
	Corridor Travel Time (min.)	Speed (mph)	Density (vpmpl)	Queue Delay (veh)	Travel Time (min.)	Queue Delay (veh*min)
No Control	21.90	39.15	76.25	27.5	0.58	3.89
QL-VSL Control	18.4	44.62	67.99	19.34	0.55	3.31
	(-15.98%)	(+13.97%)	(-10.83%)	(-29.67%)	(-5.17%)	(-14.9%)
QL-VSL&HSR Control	16.0	48.85	61.88	16.39	0.50	2.66
	(-26.94%)	(+24.77%)	(-18.84%)	(-40.4%)	(-13.79%)	(-31.61%)
FB-VSL Control	20.1	42.29	70.92	21.02	0.68	5.04
	(-8.20%)	(+8.14%)	(-6.99%)	(-23.56%)	(+17.24%)	(+29.56%)

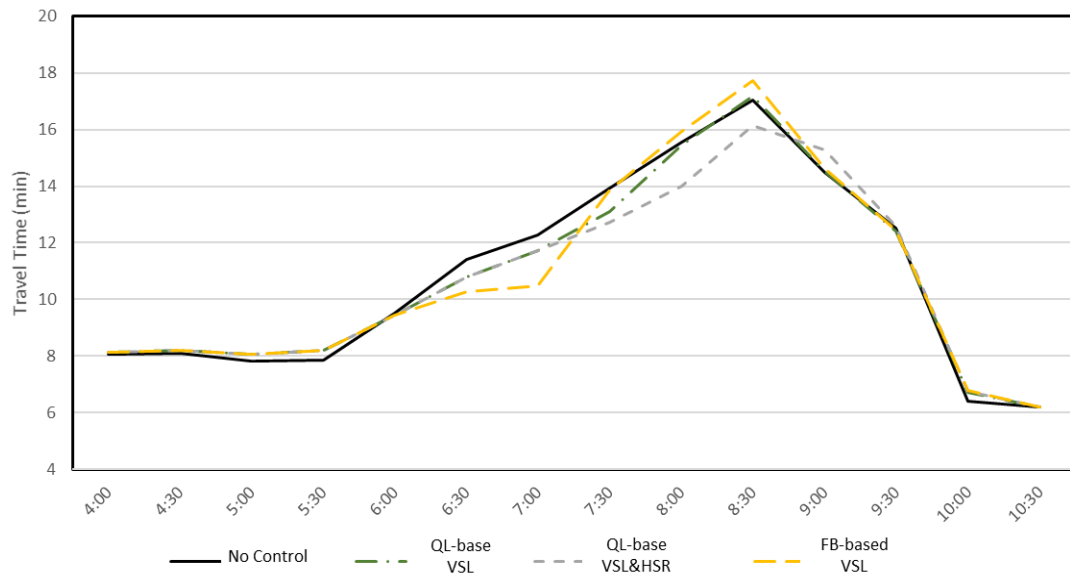


Figure 4-3. System Travel for Different Scenarios

4.5. Summary of other findings

As shown in Table 4-3, the study compares other MOEs such as density and queue delay. To distinguish the impact on freeway mainline and ramps, the results of I-270 freeway mainline and freeway ramps are summarized separately. It proves that ATM

control performs well in releasing traffic congestion on freeways. An in-depth comparison of flow speeds between the stand-alone VSL strategy and the coordinated ATM strategy indicates that the HSR brings great improvement in improving corridor travel time. There is another interesting concern that the traffic performance deteriorates significantly under the feedback-based ATM control, especially the queue delay on ramps, while the QL-based strategies show considerable improvement.

## Chapter 5 Conclusion

### 5.1. Summary of the Research

This study proposes a coordinated ATM control algorithm to release freeway congestion using reinforcement-learning technique. The algorithm takes advantage of the VSL and HSR to achieve better performance in traffic improvement. Benefited from the contribution of reinforcement learning technique, the proposed algorithm works efficiently by reducing the burden of time-consuming optimization calculation.

The proposed control framework has been applied to a network, where the 5.5-mile segment is controlled by the variable speed limit and hard shoulder running strategies. Considering both the detailed information about individual agents and speed of model services, DTALite, a mesoscopic dynamic traffic analysis (DTA) model is adopted to simulate the traffic improvement under the proposed strategy. The DTA model is adjusted, calibrated and validated using real-world traffic data to capture the realistic traffic conditions. Five scenarios are compared based on 5-minutes-VSL control, 10-minutes-VSL control, 5-minutes-coordinated control, 10-minutes-coordinated control, and no control. The simulation results indicate that the proposed strategy could improve the traffic flow condition by reducing the corridor travel time up to 27%. By comparing the results of the proposed coordinated strategy with that of the stand-alone VSL strategy, coordinated strategy outperforms the other one. By applying random traffic flow in the network, the model set correctly captured 90% (23 of 25 tests) of the congestion scenarios and provided optimal controls immediately. As a benchmark, the proposed QL-based strategies are compared to a feedback-based VSL

strategy. The feedback-based VSL strategy reduces the corridor travel time by 8.6% and increases bottleneck speed by 42.1%. The results suggest that QL-based strategy method redeems the limitation of traffic prediction in feedback-based strategy tremendously. Additionally, the QL-based strategy better balances the performance at bottleneck and upstream ATM control segment.

### 5.2 Advantages of the proposed Algorithm

The simulation model shows the advantages of the proposed strategy in several aspects. First, the coordinated ATM strategy provides the proof in solving the long-standing problem of stand-alone ATM control strategies, which is regarded as unreliable in complex traffic congestion scenarios. The effectiveness of any stand-alone ATM control strategy like variable speed limit or ramp metering is limited by the size of flowing traffic. Too heavy a traffic often results in the quickly reach of a roadway's capacity, which leaves limited room for manipulation. As the only method based on the capacity improvement, the introduction of hard shoulder running can work with the variable speed limit to guide the traffic to flow through a bottleneck.

Second, adopting the queue delay as the key parameter to evaluate the traffic condition with the reward function provides a possible method to evaluate the performance of the algorithm. Past reinforcement learning algorithms took density, speed and volume as the parameters to define the states and reward functions. However, the traditional variables are insufficient to provide the overall evaluation of the traffic condition of the network. Additionally, the corridor queue delay helps the agent learn how to balance the performance of entire corridor instead of sacrificing one part.

Therefore, this study takes the queue delay as the key parameter to replace the traditional variables.

Furthermore, the adoption of reinforcement learning technique minimizes the burden of complex optimization process, which existing with many traditional ATM algorithms. The Q-Learning agent is trained in an offline scheme. After the agent learns how to get the optimal strategy for various traffic conditions, the controlling system does not need to perform heavy computing which enables real-time decision-making. In addition, through the continuous learning function, the RL-based ATM algorithm has the capability of predicting traffic state transitions and acts in a proactive control scheme.

### 5.3 Deficiency of the Proposed Algorithm

While the proposed strategy performs well in improving traffic flow under congested condition, there are still areas that require further research. First, while the reinforcement learning technique can provide an optimal action for any given state immediately, the model requires long period of training. Defects in any link will result in a great reduction of traffic control effects. In addition, some studies indicate that the effects of ATM control, especially VSL, are affected by traffic flow characteristics at different types of freeway bottlenecks. For example, the VSL strategy should also consider the unique variable for a particular type of bottleneck caused by diverging, merging, on-ramp, off-ramp, or even different number of lanes. The one-to-one correspondence between simulation models and scenarios means a series of models need to be built, trained and tested for different traffic conditions.

Secondly, some assumptions used in the algorithm may reduce the accuracy of the performance. In this study, it is assumed that all drivers would fully obey the posted speed limits while it is known this is not the case in the real world. The dynamic speed limit signs are contributing mainly under free-flow traffic condition. And when the situation is on the contrary, for example the heavy congested or the inclement weather conditions, the speed limit signs may not work well since drivers are easily influenced by the surrounding drivers and weather conditions. Therefore, the simulation results based on this assumption may not reflect the performance in real world applications.

In addition, this study assumes perfect input data with true data reliability. In the real world, most of the ATM operation systems use loop detectors to collect real-time data to support the dynamic control algorithm and it is a significant challenge for a loop detector to deliver continuous and reliable data. For the record: the data quality is limited due to the huge operating and maintaining expense, and the on blemish of the inaccuracy of loop detectors.

#### 5.4 Future Work

Although the proposed strategy presents great benefit in improving freeway traffic congestion, the performance loss caused by the miss-considering factors or hard constraints are still examined. To improve the strategy, several topics for future studies include the following:

- 1) As discussed in the weakness for the proposed strategy, issues related to training time can be further improved. Regarding the model training time, more efficient coding and computational method should be tried to reduce the learning time.

The network of parallel computing, and the networks of distributed appliances could also be introduced to the computational process.

- 2) Like discussed in the previous section, the simulation results are based on the theoretical framework presented in the current study. Not surprisingly, the performance of real-world applications is not what the simulation model presents, but far different. Notably, traffic improvement under the proposed ATM strategy is relatively depended on the drivers' compliance rate. In this study, the drivers are assumed to fully follow the posted speed limits, which is impossible in real world conditions.
- 3) To extend the issue indicated above, the effects of drivers' compliance with posted speed limits should be considered for evaluating the operational effects of ATM control strategies. How to design the implementation of ATM system to increase drivers' compliance ratio, and how to minimize the influence of abnormal driving behavior are two important aspects, which significantly strengthen the performance. On this matter, some field studies are suggested to have better understanding of drivers' behavior. For simulation purpose, a distribution function based on behavior models such as stochastic equations should be developed to represent true behavior.
- 4) About the imperfection of data assumption as related to the reliability issue with loop detectors, while other technology can be deployed, for simulation purposes, intermittent real data from loop detector complemented by historical data may be modeled as a single input data for the model run.

- 5) Current variable speed limit and hard shoulder implementations follow simple rules, which may not be the optimal installation method. About the optimized installation of these two strategies, deeper analysis is suggested to coordinate with microscopic simulation models. Traffic data from a VSL-equipped freeway are also needed to provide an enhanced understanding of the impact of different ATM implementations.



## References

- [1] Müller, E. R., R. C. Carlson, W. Kraus, and M. Papageorgiou. Micro simulation Analysis of Practical Aspects of Traffic Control With Variable Speed Limits. *IEEE Transactions on Intelligent Transportation Systems*, 2015. 16(1):512-523.
- [2] Breton, P., A. Hegyi, B. De Schutter and H. Hellendoorn. Shock Wave Elimination or Reduction by Optimal Coordination of Variable Speed Limits. *IEEE 5th International Conference on Intelligent Transportation Systems*, Singapore, 2002. 225-230.
- [3] Hegyi, A., Bart De Schutter and J. Hellendoorn. Optimal Coordination of Variable Speed Limits to Suppress Shock Waves. *IEEE Transactions on Intelligent Transportation Systems*, 2005. 6(1):102-112.
- [4] Papageorgiou. M., H. Hadj-Salem, and F. Middelham. ALINEA Local Ramp Metering Summary of Field Results. *Transportation Research Record: Journal of the Transportation Research Board*, 1997. 1603(1):90-98.
- [5] Smaragdis, E. and M. Papageorgiou. A Series of New Local Ramp Metering Strategies. *Transportation Research Record: Journal of the Transportation Research Board*, 2003. 1856(1):74-86.
- [6] Papamichail, I., E. Kosmatopoulos, and I. Chrysoulakis. HERO Coordinated Ramp Metering Implemented at the Monash Freeway. *Transportation Research Record: Journal of the Transportation Research Board*, 2010. 2178(1):10-20.
- [7] Khazraeian, S., M. Hadi, and Y. Xiao. Safety Impacts of Queue Warning in a Connected Vehicle Environment. *Transportation Research Record: Journal of the Transportation Research Board*, 2017. 2621(1):31-37.
- [8] Khan, A.M. Intelligent Infrastructure-based Queue-end Warning System for Avoiding Rear Impacts. *IET Intelligent Transport Systems*, 2007. 1(2):138-143.
- [9] Geistefeldt, J Operational Experience with Temporary Hard Shoulder Running in Germany. *Transportation Research Record: Journal of the Transportation Research Board*, 2012. 2278(1):67-73.
- [10] Chiu. S. (1992) Adaptive Traffic Signal Control Using Fuzzy Logic. *The First IEEE Regional Conference on Aerospace Control Systems*, Westlake Village, CA, USA, 1993. 122-126.
- [11] Cai, C., C. K. Wong, and B. G. Heydecker. Adaptive Traffic Signal Control Using Approximate Dynamic Programming. *Transportation Research Part C: Emerging Technologies*, 2009. 17(5):456-474.

- [12] Priemer, C. and B. Friedrich. A Decentralized Adaptive Traffic Signal Control Using V2I Communication Data. *2th International IEEE Conference on Intelligent Transportation Systems*, St. Louis, MO, 2009. 1-6.
- [13] Guerrieri, M., and R. Mauro. Capacity and safety analysis of hard-shoulder running (HSR). A motorway case study. *Transportation Research Part A: Policy and Practice*, 2016. 92:162-183
- [14] Hegyi, A., B. D. Schutter., and H. Hellendoorn. Model predictive control for optimal coordination of ramp metering and variable speed limits. *Transportation Research Part C: Emerging Technologies*, 2005. 13(3):185-209.
- [15] Hegyi, A., B. D. Schutter., H. Hellendoorn, and T.van den Boom. Optimal coordination of ramp metering and variable speed control-an MPC approach. *Proceedings of the 2002 American Control Conference (IEEE Cat. No.CH37301)*, Anchorage, AK, USA, 2002. 5:3600-3605.
- [16] Schelling, I., A. Hegyi, and S. P. Hoogendoorn, S.P. SPECIALIST-RM — Integrated variable speed limit control and ramp metering based on shock wave theory. *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Washington, DC, 2011. 2154-2159.
- [17] Gomes, G., A. May, and R. Horowitz. Congested Freeway Microsimulation Model Using VISSIM. *Transportation Research Record: Journal of the Transportation Research Board*, 2004. 1876(1):71-81.
- [18] Fernandes, P. and U. Nunes. Platooning of Autonomous Vehicles with Inter-Vehicle Communications in SUMO Traffic Simulator. *3th International IEEE Conference on Intelligent Transportation Systems*, Funchal, 2010. 1313-1318.
- [19] Faganzo, C. F. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological*, 1994. 28(4):269-287.
- [20] Acosta, A. F., J. E. E. Oviedo, and J. J. E. Oviedo. TraCI4Matlab: Enabling the Integration of the SUMO Road Traffic Simulator and Matlab® Through a Software Re-engineering Process. In: *Modeling Mobility with Open Data* (Behrisch M., Weber M.). Lecture Notes in Mobility. Springer, Cham, 2015. pp. 155-170.
- [21] Zarean, M., M. D. Robinson, and D. Warren. APPLICATIONS OF VARIABLE SPEED LIMIT SYSTEMS TO ENHANCE SAFETY. *Proceedings of the 7<sup>th</sup> World Congress on Intelligent Systems*, Turin, Italy, 2000.
- [22] Abdel-Aty, M., K. M. Haleem, R. J. Cunningham, and V. V. Gayah. Application of Variable Speed Limits and Ramp Metering to Improve Safety and Efficiency of Freeways, 2019.

- [23] Katz, B., J. Ma, and H. Rigdon. Synthesis of Variable Speed Limit Signs. Report of Federal Highway Administration, 2017.
- [24] Hadiuzzaman, M. and T. Z. Qiu. Cell Transmission Model Based Variable Speed Limit Control for Freeways. *Canadian Journal of Civil Engineering*, 2013. 40(1):46-45.
- [25] I-90: Four Lakes to Idaho State Line Operations Study Existing Conditions Final Report, Washington State Department of Transportation, 2017.
- [26] Ulfarsson, G. F., V. N. Shankar, and P. Vu. The effect of variable message and speed limit signs on mean speeds and speed deviations. *International Journal of Vehicle Information and Communication Systems*, 2005. 1(1/2):69-87.
- [27] Abdel-Aty, M., J. Dilmore, and L. Hsia. Applying Variable Speed Limits and the Potential for Crash Migration. *Transportation Research Record: Journal of the Transportation Research Board*, 2006 1953(1):21–30.
- [28] Hellinga, B. and M. Mandelzy. Impact of Driver Compliance on the Safety and Operational Impacts of Freeway Variable Speed Limit Systems. *Journal of Transportation Engineering*, 2011. 137(4).
- [29] Chang, G. L., S. Y. Park, and J. Paracha. Intelligent Transportation System Field Demonstration: Integration of Variable Speed Limit Control and Travel Time Estimation for a Recurrently Congested Highway. *Transportation Research Record: Journal of the Transportation Research Board*, 2011. 2243(1):55-56
- [30] Allaby, P., B. Hellinga and M. Bullock. Variable Speed Limits: Safety and Operational Impacts of a Candidate Control Strategy for Freeway Applications. *IEEE Transactions on Intelligent Transportation Systems*, 2007. 8(4):671-680.
- [31] Hoogen, E. and S. Smulder. Control by variable speed signs: results of the Dutch experiment. Seventh International Conference on Road Traffic Monitoring and Control, London, UK, 1994. 145-149.
- [32] Papamichail, I. and A. Kotsialos. Freeway ramp metering: An overview. *IEEE Transactions on Intelligent Transportation Systems*, 2002. 3(4):271-281.
- [33] US Department of Transportation. Traffic safety facts 2002: a compilation of motor vehicle crash data from the fatality analysis reporting system and the general estimate system. 2004.
- [34] Zahabi, S. A. H., J. Strauss, and K. Manaugh. Estimating Potential Effect of Speed Limits, Built Environment, and Other Factors on Severity of Pedestrian and Cyclist Injuries in Crashes. *Transportation Research Record: Journal of the Transportation Research Board*, 2011. 2247(1):81-90.

- [35] Abdel-Aty, M., J. Dimore, and A. Dhindsa. Evaluation of Variable Speed Limits for Real-time Freeway Safety Improvement. *Accident Analysis & Prevention*, 2006. 38(2):335-345.
- [36] Lee, C., B. Hellinga, and F. Sacomanno. Assessing Safety Benefits of Variable Speed Limits. *Transportation Research Record: Journal of the Transportation Research Board*, 2004. 1897(1):183-190.
- [37] Abdel-Aty, M., R. J. Cunningham, and V. V. Gayah. Dynamic Variable Speed Limit Strategies for Real-Time Crash Risk Reduction on Freeways. *Transportation Research Record: Journal of the Transportation Research Board*, 2008. 2078(1):108-116.
- [38] Li, J. and H. M. Zhang. Modeling space-time in homogeneities with the kinematic wave theory. *Transportation Research Part B: Methodological*, 2013. 54:113-125.
- [39] Du, S. and S. Razavi. Variable Speed Limit for Freeway Work Zone with Capacity Drop Using Discrete-Time Sliding Mode Control. *Journal of Computing in Civil Engineering*, 2019. 33(2).
- [40] Jin, H.Y. AND W. L. Jin. Control of A Lane-drop Bottleneck through Variable Speed Limits. *Transportation Research Part C: Emerging Technologies*, 2015. 58(Part C):568-584.
- [41] Kwon, E. D.Brannan, and K. Shouman. Development and Field Evaluation of Variable Advisory Speed Limit System for Work Zones. *Transportation Research Record: Journal of the Transportation Research Board*, 2007. 2015(1):12-18.
- [42] Hadiuzzaman, M., Q. Z. Qiu, and X.Y. Liu. Variable Speed Limit Control Design for Relieving Congestion Caused by Active Bottlenecks. *Journal of Transportation Engineering*, 2013. 139(4).
- [43] Alessandri, A., A. D. Febbraro, A. Ferrara, and E. Punta. Nonlinear Optimization for Freeway Control Using Variable-Speed Signaling. *IEEE Transactions on Vehicular Technology*, 1999. 48(6):145-149.
- [44] Hegyi, A., De Schutter, B., & Hellendoorn, J. (2005). Optimal Coordination of Variable Speed Limits to Suppress Shock Waves. *IEEE Transactions on Intelligent Transportation Systems*, 2005. 6(1): 102–112.
- [45] Khondaker, B.and L. Katten, L. Variable Speed Limit: A Microscopic Analysis in a Connected Vehicle Environment. *Transportation Research Part C: Emerging Technologies*, 2015. 58(Part A):146-159.

- [46] Allaby, P., B.Hellinga, and M.Bullock. Variable Speed Limits: Safety and Operational Impacts of a Candidate Control Strategy for Freeway Applications. *IEEE Transactions on Intelligent Transportation Systems*, 2007. 8(4):671-680.
- [47] Carlson, C.R., I. Papamichail, M. Papageorgiou, and A. Messmer. Optimal mainstream traffic flow control of large-scale motorway networks. *Transportation Research Part C: Emerging Technologies*, 2010. 18(2):193-212.
- [48] Han, Y., D. Chen, and S. Ahn. Variable speed limit control at fixed freeway bottlenecks using connected vehicles. *Transportation Research Part B: Methodological*, 2017. 98:113-134.
- [49] Lee, J. and B. B. Park. Evaluation of Variable Speed Limit under Connected Vehicle environment. *International Conference on Connected Vehicles and Expo (ICCVE)*, Las Vegas, NV, 2013. 966-967.
- [50] Hegyi, A., B. D. Schutter, and H. Hellendoorn. Model predictive control for optimal coordination of ramp metering and variable speed limits. *Transportation Research Part C: Emerging Technologies*, 2005. 13(3):185-209.
- [51] Chang, G. L., S. Y. Park, and J. Paracha. Intelligent Transportation System Field Demonstration: Integration of Variable Speed Limit Control and Travel Time Estimation for a Recurrently Congested Highway. *Transportation Research Record: Journal of the Transportation Research Board*, 2011. 2243(1):55-56
- [52] Alessandri, A., A. Di Febbraro, A. Ferrara, and E. Punta. Optimal Control of Freeways via Speed Signaling and Ramp Metering. *IFAC Proceeding Volumes*, 1997. 30(8):1019-1024.
- [53] Yang, Y., H. Lu, Y. Yin, and H. Yang. Optimization of Variable Speed Limits for Efficient, Safe, and Sustainable Mobility. *Transportation Research Record: Journal of the Transportation Research Board*, 2013. 2333(1):37-45.
- [54] Carlson, R.C., I. Paramichail, and M. Papageorgiou. Local Feedback-Based Mainstream Traffic Flow Control on Motorways Using Variable Speed Limits. *IEEE Transactions on Intelligent Transportation Systems*, 2011. 12(4): 1261-1276. 16(2): 610-621.
- [55] Iordanidou, G., C. Roncoli, I. Papamichail and M. Papageorgiou, Feedback-Based Mainstream Traffic Flow Control for Multiple Bottlenecks on Motorways. *IEEE Transactions on Intelligent Transportation Systems*, 2015. 16(2):610-621.
- [56] Iordanidou, G., C. Roncoli, I. Papamichail and M. Papageorgiou. Feedback-Based Mainstream Traffic Flow Control for Multiple Bottlenecks on Motorways. *IEEE Transactions on Intelligent Transportation Systems*, 2015. 16(2):610-621.

- [57] Rezaee, K., B. Abdulhai, and H. Abdelgawad. Application of Reinforcement Learning with Continuous State Space to Ramp Metering in Real-world Conditions. *15th International IEEE Conference on Intelligent Transportation Systems*, Anchorage, AK, 2012. pp. 1590-1595.
- [58] Mahadevan, S. Average Reward Reinforcement Learning: Foundations, algorithms, and empirical results. *Machine Learning*, 1996. 22(1-3): 159-195.
- [59] Rezaee, K., B. Abdulhai, and H. Abdelgawad. Application of reinforcement learning with continuous state space to ramp metering in real-world conditions. *15th International IEEE Conference on Intelligent Transportation Systems*, Anchorage, AK, 2012. 1590-1595.
- [60] Zhao, D., X. Bai, F. Wang, J. Xu and W. Yu. DHP Method for Ramp Metering of Freeway Traffic. *IEEE Transactions on Intelligent Transportation Systems*, 2011. 12(4):990-999.
- [61] Li, Z., P. Liu, C. Xu, H. Duan, and W. Wang. Reinforcement Learning-Based Variable Speed Limit Control Strategy to Reduce Traffic Congestion at Freeway Recurrent Bottlenecks. *IEEE Transactions on Intelligent Transportation Systems*, 2017. 18(11): 3204-3217.
- [62] Zhu, F., and S. V. Ukkusuri. Accounting for dynamic speed limit control in a stochastic traffic environment: A reinforcement learning approach. *Transportation Research Part C: Emerging Technologies*, 2014. 41:30-47.
- [63] Oh, S., and H. Yeo. Microscopic Analysis on the Causal Factors of Capacity Drop in Highway Merging Sections. Presented at *91rd Annual Meeting of the Transportation Research Board*, Washington, D.C., 2012.
- [64] Ben-Akiva, M. E., S. Gao, Z. Wei, and Y. Wen. A dynamic Traffic Assignment Model for Highly Congested Urban Networks. *Transportation Research Part C: Emerging Technologies*, 2012. 24:62-82.
- [65] Grumert, E., A. Tapani, and X. Ma. Characteristics of Variable Speed Limit Systems. *European Transport Research Review*, 2018.
- [66] Ghods, A.H., A. R. Kian and M. Tabibi. Adaptive freeway ramp metering and variable speed limit control: a genetic-fuzzy approach. *IEEE Intelligent Transportation Systems Magazine*, 2009. 1(1):27-36.
- [67] Lu, X.-Y., P. Varaiya, R. Horowitz, D. Su, and S. E. Shladover. Novel Freeway Traffic Control with Variable Speed Limit and Coordinated Ramp Metering. *Transportation Research Record: Journal of the Transportation Research Board*, 2011. 2229(1):55–65.

- [68] Papamichail, L., K. Kampitaki, M. Papageorgiou, and A. Messmer. Integrated Ramp Metering and Variable Speed Limit Control of Motorway Traffic Flow. *The International Federation of Automatic Control*, 2008. 41(2):14084-14089.
- [69] Abdel-Aty, M., and A. Dhindsa. Coordinated Use of Variable Speed Limits and Ramp Metering for Improving Safety on Congested Freeways. Presented at *Transportation Research Board*, 2007.
- [70] Ghods, H., L. Fu., and A.R. Kian. An Efficient Optimization Approach to Real-Time Coordinated and Integrated Freeway Traffic Control. *IEEE Transactions on Intelligent Transportation Systems*, 2010. 11(4):873-884.
- [71] Haj-Salem, H., N. Farhi, and J. Lebacque. Combining Ramp Metering and Hard Shoulder Strategies: Field Evaluation Results on the Ile de France Motorway Network. *Transportation Research Procedia*, 2014. 1002-1010.
- [72] Ma, J., J. Hu, D. K. Hale, and J. Bared. Dynamic Hard Shoulder Running for Traffic Incident Management. *Transportation Research Record*, 2016. 2554(1):120–128.
- [73] Samoili, S., D. Efthymiou, C. Antoniou, and A.G. Dumont. Investigation of Lane Flow Distribution on Hard Shoulder Running Freeways. *Transportation Research Record: Journal of the Transportation Research Board*, 2011. 2396(1):133–142.
- [74] Aron, M., R. Seidowsky, and S. Cohen. Safety Impact of Using the Hard Shoulder During Congested Traffic. The Case of A Managed Lane Operation on A French Urban Motorway. *Transportation Research Part C: Emerging Technologies*, 2013. 28:168-180.
- [75] Jade, C. and D. Andre-Gilles. Effects of Hard Shoulder Lanes Configuration on Driver's Behavior. Presented at *5th International Symposium Highway Geometric Design*, Vancouver, BC, Canada, 2015.
- [76] Coffey, S., Park, S. (2018). Operational Evaluation of Part-Time Shoulder Use for Interstate 476 in the State of Pennsylvania. *Advances in Civil Engineering*. 2018.
- [77] Michalaki, P., M. Quddus, and D. Pitfield. A Sensor-based System for Monitoring Hard-shoulder Incursions: Review of Technologies and Selection Criteria. *5th International Conference on Transportation and Traffic Engineering*, 2016.
- [78] Sutton, R.S. and A. G. Barto, Reinforcement Learning: An Introduction. Cambridge, MA: MIT Press. 1998.
- [79] Watkins, C. and P. Dayan, Q-Learning. *Machine Learning*. 1992. 8(3), 279-292.

- [80] Islam, M.T., M. Hadiuzzaman, J. Fang, and T.Z. Qiu. Assessing Mobility and Safety Impacts of a Variable Speed Limit Control Strategy. *Transportation Research Record: Journal of the Transportation Research Board*, 2013. 2364(1):1-11.
- [81] Zhou, X. and J. Taylor. DTALite: A Queue-Based Mesoscopic Traffic Simulator for Fast Model Evaluation and Calibration. *Cogent Engineering*, 2014. 1(1).
- [82] Xiong, C., X. Zhou, and L. Zhang. AgBM-DTALite: An Integrated Modelling System of Agent-based Travel Behaviour and Transportation Network Dynamics. *Travel Behavior and Society*, 2018. (12):141-150.
- [83] Zhang, J., H. Chang, and P. A. Ioannou. A Simple Roadway Control System for Freeway Traffic. *2006 American Control Conference*, Minneapolis, MN, 2006.