AN ABSTRACT OF THE THESIS OF

Shangjia Dong for the degree of Master of Science in Civil Engineering presented on November 9, 2015.

Title: Stochastic Characterization of Highway Capacity and Its Applications

Abstract approved: _______________________________________________________________

Haizhong Wang

Highway capacity has traditionally been treated as deterministic. In reality, however, capacity can vary from time to time and location to location due to inclement weather, various driving behavior, traffic incident, bottleneck and workzone. This thesis aims to characterize the stochastic variations of highway capacity, and explore its applications.

The stochastic variation of highway capacity is captured through a space-time Autoregressive Integrated Moving Average (STARIMA) model. It is identified following a Seasonal STARIMA model \((0, 0, 2_3) \times (0, 1, 0)_2\), which indicates that the capacities of adjacent locations are spatially-temporally correlated. Hourly capacity patterns further verify the stochastic nature of highway capacity. The goal of this research is to study (1) how to take advantage of second order information, such as capacity variation, and (2) what benefits can be gained from stochastic capacity modeling. The implication of stochastic capacity is investigated through a ramp metering case study. A mean-standard deviation formulation of capacity is proposed to achieve the trade-off between traffic operation efficiency and robustness. Following that, a modified stochastic capacity-constraint ZONE ramp metering scheme embedded cell transmission model (CTM) algorithm is introduced. The numerical experiment suggests that considering second order information would alleviate the bottleneck effect and improve throughput. Monte Carlo simulation further
supports this argument. This study helps verify and characterize the stochastic nature of capacity, validates the benefits of using second order information, and thus enhances the necessity of implementing stochastic capacity in traffic operation.

In addition, this thesis presents a multiple days' highway capacity forecasting model and reveals the importance of incorporating capacity randomness separately in short term forecasting. Empirical observations of capacity with variations suggest that the existing deterministic notion of capacity (i.e., a fixed capacity value of 2400 veh/hr/ln) is inadequate to fully describe capacity fluctuations. The goal of this study is to provide insight on forecasting with inherent stochasticity. A hybrid highway capacity forecasting model based on wavelet dynamic neural time series is proposed. Based on 5-day capacity forecasting results, a 3.07% of Mean Absolute Percentage Error (MAPE) demonstrates improved performance against single models. This is achieved by its ability to capture the details and solve a nonlinear time series problem with a dynamic neural network. Following the results, the study concludes with a discussion on why the individual modeling of stochastic information influences the forecasting performance.
Stochastic Characterization of Highway Capacity and Its Applications

by

Shangjia Dong

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented November 9, 2015
Commencement June 2016
Master of Science thesis of Shangjia Dong presented on November 9, 2015

APPROVED:

________________________
Major Professor, representing Civil Engineering

________________________
Head of the School of Civil & Construction Engineering

________________________
Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

________________________
Shangjia Dong, Author
ACKNOWLEDGEMENTS

I am truly grateful for people who helped me and supported me for the past two years. First, I would like to thank my advisor, Dr. Haizhong Wang. I began working with Dr. Wang in my senior year, and my research path started from there. The academic guidance and life-philosophy he taught me will be beneficial to my whole career.

I also would like to thank my committee members, Dr. David Hurwitz and Dr. Chris A. Bell. The first ever transportation class I took is from Dr. Hurwitz, which laid the foundation of my future research. I was a teaching assistant on Dr. Bell’s highway engineering class. That experience added great value in my graduate school life. Also, I would like to thank Dr. Bell for providing guidance and various opportunities throughout my Master period and help me transit into a Ph.D student. In addition, I would like to acknowledge Dr. Jia Li for his help in building up this thesis.

Thanks to my friends for their consistent support. They accompanied me through the toughest time. The professional meetings, conferences and football games that we attended made this journey an enjoyable and memorable experience in my life.

Finally, I would like to thank my family. They supported me to pursue higher education which leads me here. Their unconditional love and encouragement accompanied me throughout my academic pursuit. Thank you my nephew, Junwang Chen. You added a lot of joyful moments to the family.
# TABLE OF CONTENTS

1. INTRODUCTION ........................................................................... 1
   1.1. Motivation and Contribution ............................................. 4
   1.2. Organization of This Thesis ............................................... 5

2. LITERATURE REVIEW ............................................................... 6
   2.1. Stochastic Capacity .......................................................... 6
   2.2. Capacity Estimation Approaches ....................................... 7
   2.3. Ramp Metering ............................................................... 8
   2.4. Fourier Analysis v.s Wavelet Analysis ................................. 13
   2.5. Short-term Traffic Forecasting .......................................... 14

3. METHODOLOGY AND DATA DESCRIPTION ............................. 16
   3.1. Product Limit Method ...................................................... 16
   3.2. Space-Time ARIMA ....................................................... 19
       3.2.1 Seasonal STARIMA ................................................. 20
       3.2.2 Model Validation ..................................................... 21
   3.3. Wavelet analysis ............................................................ 22
   3.4. Dynamic neural network ............................................... 24
   3.5. Study Area and Data Description ...................................... 25

4. SPATIAL AND TEMPORAL EMPIRICAL ANALYSIS ................ 28
   4.1. STARIMA Experiment ..................................................... 29
   4.2. Seasonal STARIMA Analysis .......................................... 32
   4.3. Discussion ...................................................................... 32

5. APPLICATION I: CORRIDOR LEVEL RAMP METERING ........... 34
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Capacity Mean-Standard Deviation Trade-off</td>
<td>34</td>
</tr>
<tr>
<td>5.2</td>
<td>Weighting Parameter Estimation</td>
<td>37</td>
</tr>
<tr>
<td>5.3</td>
<td>Simulation Study</td>
<td>38</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Simulation Setting</td>
<td>40</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Experiment Results</td>
<td>43</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Discussion</td>
<td>43</td>
</tr>
<tr>
<td>6.</td>
<td>APPLICATION II: HYBRID FORECASTING OVER MULTIPLE DAYS</td>
<td>48</td>
</tr>
<tr>
<td>6.1</td>
<td>Descriptive Statistics of Empirical Data</td>
<td>48</td>
</tr>
<tr>
<td>6.2</td>
<td>Experiment framework</td>
<td>51</td>
</tr>
<tr>
<td>6.3</td>
<td>Results Analysis</td>
<td>52</td>
</tr>
<tr>
<td>6.3.1</td>
<td>Performance measure</td>
<td>52</td>
</tr>
<tr>
<td>6.3.2</td>
<td>Multiple days’ capacity forecasting</td>
<td>52</td>
</tr>
<tr>
<td>7.</td>
<td>CONCLUSION AND FUTURE REMARKS</td>
<td>58</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Three Different Highway Capacity Variation Patterns from Georgia 400 Over a Year</td>
<td>3</td>
</tr>
<tr>
<td>3.1</td>
<td>Distance degree demonstration</td>
<td>19</td>
</tr>
<tr>
<td>3.2</td>
<td>Decomposition Structure of Wavelet Transforms</td>
<td>23</td>
</tr>
<tr>
<td>3.3</td>
<td>Static Neural Network Diagram</td>
<td>24</td>
</tr>
<tr>
<td>3.4</td>
<td>Dynamic Network with One Delay, [a(t) = w_{1,1}f(t) + w_{1,2}f(t - 1)]</td>
<td>25</td>
</tr>
<tr>
<td>3.5</td>
<td>Neural Network Training Framework</td>
<td>25</td>
</tr>
<tr>
<td>3.6</td>
<td>Study Site: GA400 100 Stations</td>
<td>27</td>
</tr>
<tr>
<td>4.1</td>
<td>Nonseasonal STARIMA ACF &amp; PACF</td>
<td>29</td>
</tr>
<tr>
<td>4.2</td>
<td>Seasonal STARIMA ACF and PACF</td>
<td>32</td>
</tr>
<tr>
<td>5.1</td>
<td>Hour-to-hour Capacity</td>
<td>38</td>
</tr>
<tr>
<td>5.2</td>
<td>Cumulative distribution function of capacity at rush hours (station 4000025)</td>
<td>39</td>
</tr>
<tr>
<td>5.3</td>
<td>On-ramp only metering</td>
<td>40</td>
</tr>
<tr>
<td>5.4</td>
<td>Empirical fundamental diagram (station 4000025)</td>
<td>41</td>
</tr>
<tr>
<td>5.5</td>
<td>Comparison of roadway segment density under different capacities</td>
<td>44</td>
</tr>
<tr>
<td>5.6</td>
<td>Throughput comparison between with 1(^{st}) order only and 1(^{st})&amp;2(^{nd}) order information</td>
<td>45</td>
</tr>
<tr>
<td>5.7</td>
<td>Monte Carlo Throughput comparison</td>
<td>45</td>
</tr>
<tr>
<td>5.8</td>
<td>Multiple ramps metering scenario</td>
<td>46</td>
</tr>
<tr>
<td>6.1</td>
<td>Selected Study Site of Detector 4000026 from GA400 Northbound at Atlanta, Georgia (from Google map)</td>
<td>49</td>
</tr>
<tr>
<td>6.2</td>
<td>Descriptive Statistics of Capacity Data</td>
<td>50</td>
</tr>
<tr>
<td>6.3</td>
<td>Wavelet Decomposition at Level 3, [s = a3 + d3 + d2 + d1]</td>
<td>53</td>
</tr>
<tr>
<td>6.4</td>
<td>Neural Network of Approximation A3</td>
<td>54</td>
</tr>
<tr>
<td>6.5</td>
<td>Details of Neural Network: D1</td>
<td>54</td>
</tr>
<tr>
<td>6.6</td>
<td>Details of Neural Network: D2</td>
<td>55</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>6.7</td>
<td>Details of Neural Network: D3</td>
<td>55</td>
</tr>
<tr>
<td>6.8</td>
<td>Prediction at Individual Steps</td>
<td>56</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Summary of Roadway Capacity Estimation Methods</td>
<td>9</td>
</tr>
<tr>
<td>3.1</td>
<td>Product Limit Method Calculation</td>
<td>18</td>
</tr>
<tr>
<td>4.1</td>
<td>Nonseasonal STARIMA Autocorrelations and Partial Autocorrelations</td>
<td>30</td>
</tr>
<tr>
<td>4.2</td>
<td>Seasonal STARIMA Autocorrelations and Partial Autocorrelations</td>
<td>31</td>
</tr>
<tr>
<td>6.1</td>
<td>Forecasting results comparison for a 5-day Horizon</td>
<td>56</td>
</tr>
<tr>
<td>6.2</td>
<td>Performance Measure for a 5-day Forecasting Horizon</td>
<td>57</td>
</tr>
</tbody>
</table>
STOCHASTIC CHARACTERIZATION OF HIGHWAY CAPACITY AND ITS APPLICATIONS

1. INTRODUCTION

In the U.S, 20% of Gross National Product (GNP) is expended on transportation, of which about 85% is spent on highway transportation (passenger and freight). The U.S owns and operates 150 million automobiles and an additional 50 million trucks, with an average of 10,000 miles per year for passenger cars and 50,000 miles per year for trucks on the highway. Car ownership in the U.S reaches 56 per hundred population. These all imply the importance of the transportation system (Lieu et al., 1999), more and more efforts are therefore devoted into transportation research. Traffic flow theory seeks to describe the interactions between the vehicles and their operations (the mobile components) and the infrastructure (the immobile component) in a precise mathematical way (Lieu et al., 1999).

Transportation has become more and more important in our daily lives. Unfortunately, the problems it creates, such as congestion, grow increasingly more severe in the meanwhile. Congestion occurs when demand exceeds the capacity. Limited available capacity is not the only reason; it can be caused by several factors such as bad weather, work zones, traffic incidents, poor signal timing, special events, or bottlenecks. The findings show that the major cause of congestion is the inefficient operation of highways during periods of high demand (Chen et al., 2001). It is suggested that intelligently control access to highway through ramp metering is necessary, while such traffic management design requires comprehensive understanding of highway capacity.
Capacity, in general, represents the maximum amount or number that can be contained or accommodated by a certain facility or infrastructure. To transportation engineers, highway capacity is a central concept to planning of either existing or construction of new infrastructure to achieve a proper demand-supply equilibration (Khanal, 2014), transportation facility design, operation, and management. The Highway Capacity Manual (HCM) defines capacity of a road as “the maximum number of vehicles that can pass a given point during a specified period under prevailing roadway, traffic, and control conditions” (2010). The HCM states further that prevailing roadway, traffic, and control conditions define capacity and any change in the prevailing conditions change a facility’s capacity (Khanal, 2014). From this definition, it is obvious that highway capacity is not a constant even for a fixed location due to the variation of prevailing roadway conditions (i.e. geometric and pavement conditions), traffic conditions (i.e. vehicle compositions and heterogeneous driver groups), and control conditions. Therefore, the capacity of a freeway segment changes constantly; the variations in capacity can be even more pronounced for other types of roadway facilities due to uncontrolled access to the roadway facility (Khanal, 2014).

An empirical analysis was conducted using Georgia State Route 400 (GA400) data over a year's observation in 2003. The day-to-day highway capacity was estimated based on a product limit method Minderhoud et al. (1997). Three different capacity variation patterns: concentrated (capacity fluctuates around the mean), half scattered and half concentrated (capacity fluctuations variate along the time) and fully scattered (larger variation exists in capacity), are identified and presented in Figure 1.1. The three different capacity patterns are discovered from three different detectors with same year’s data. Pan et al. (2013) states that “...traffic flow, by nature, is correlated in both spatial and temporal domains due to its dynamics, similar environmental conditions and human behaviors.” The noticeable capacity variations over the temporal and spatial scale motivates this research
to pursue a stochastic characterization of highway capacity and explore the implications of this treatment. The cause of variations in capacity can be attributed to heterogeneous driver behavior and vehicle types, weather, incident/accident, lane closure/work zone.

![Graphs showing three different highway capacity variation patterns from Georgia 400 over a year.](a) Concentrated (b) Mixed (c) Scattered

FIGURE 1.1: Three Different Highway Capacity Variation Patterns from Georgia 400 Over a Year

However, how the stochastic characterization of highway capacity would impact traffic control and management strategies (i.e. ramp metering) is not sufficiently addressed in existing literature. Traditional highway capacity can be interpreted as the mean of the capacity observations over a certain time period which is the first order information of a stochastic process. Alternatively, deterministic highway capacity is generally treated as a threshold, above which traffic breakdown occurs while traffic is smooth if it is below it (Elefteriadou and Lertworawanich, 2003). For example, if the capacity of a location is
2400 vehs/hr/ln, traffic flow is uninterrupted when below capacity and breakdown occurs when it exceeds capacity. The deterministic definition is simple to understand and easy to be implemented in practice but ignores the inherent capacity variations which are shown in Figure 1.1.

1.1. Motivation and Contribution

The impetus to model highway capacity as a stochastic process is largely driven by prevalent variations in empirical capacity observations. Therefore, when freeway capacity is considered to be a constant value, breakdowns become a deterministic phenomenon (Kittelson and Associates, 2010). But researchers have found that breakdowns occur even under controlled conditions. In other words, breakdowns are a stochastic phenomenon. The research question becomes what could engineers or decision-makers gain by knowing how capacity varies over a temporal and spatial domain in terms of designing more robust control and management strategies (i.e. ramp metering control scheme). To better contextualize the motivation, an example is used to illustrate what additional benefits can be gained when the second order component (i.e. the variance of capacity) is incorporated in the design of ramp metering control scheme. For example, there are two parking charge strategies. One is a flat fee of 10 dollars per day, the other is to charge on average of 10 dollars but varies based on the availability in the parking lot. In the first scenario, drivers only need to pay 10 dollars regardless the availability of parking spaces. However in the second scenario, drivers may pay more or less than 10 dollars depending on the availability of the number of parking spaces. A 10 dollar flat fee might lead to “parking failure.” This is applicable to traffic control as well. Empirical observations indicate there are variations in capacity. A robust traffic control strategy requires a stochastic characterization of highway capacity (i.e. mean and variance). A stochastic analysis of highway capacity
is also essential to understand freeway bottleneck breakdowns, congestion dynamics, and prediction of travel time reliability (Jia, 2013). The research question of what additional benefits can be gained when capacity variations are incorporated in the design of the ramp metering scheme largely motivated this study.

Major contributions of this thesis are four-fold: (1) a stochastic characterization of highway capacity through spatial-temporal analysis and hourly pattern inspection; (2) a stochastic formulation of highway capacity with numerical verification of performance through a capacity-constrained ramp-metering control embedded CTM algorithm; (3) a proposed hybrid forecasting model that outperforms single models; (4) an explicit investigation on why hybrid model provides more accurate forecasting results.

1.2. Organization of This Thesis

This thesis begins by presenting a comprehensive literature review on related topics in Chapter 2. Chapter 3. introduces the methods utilized in this thesis, along with the study site and data information. Chapter 4. conducts the stochastic characterization of highway capacity by providing a spatial-temporal analysis on daily capacity. In recognition of the stochastic nature of capacity, Chapter 5. introduces the mean-standard deviation capacity formulation and proposes an stochastic capacity-constraint ramp metering scheme. A modified ZONE ramp metering embedded CTM algorithm, and a case study is followed. Chapter 6. conducts a hybrid highway capacity forecasting over multiple days. Finally, Chapter 7. summarizes the research and concludes the future remarks.
2. LITERATURE REVIEW

2.1. Stochastic Capacity

The conventional definition of capacity according to HCM (2010) has practical limitations when representing the traffic operational impact of freeway system bottlenecks. In order to accurately quantify the capacity benefits of design, operation, and technology, it is essential to investigate the stochastic nature of capacity (Khanal, 2014). Geistefeldt and Brilon (2009) suggests that randomness exists in freeway capacity and is better represented as a random variable. Minderhoud et al. (1997) states that capacity is more a stochastic variate following a distribution function instead of a single capacity value.

Brilon et al. (2005) suggests that if freeway load is at 90% of conventional capacity, the highway will operate at the highest efficiency; this emphasizes that performance measure can be better described by stochastic characterization. “90% of conventional capacity” not only provides high efficiency, but also increase the travel time reliability. There is also potential that using conventional capacity in traffic control would overestimate the system performance, and thus lead to a gap between traffic control design and implementation.

Polus and Pollatschek (2002) also argues that momentary capacity values and the traffic volume observed before breakdown are stochastic in nature and follows the shifted gamma distribution. Lorenz and Elefteriadou (2000) suggests that if freeway capacity is defined when breakdown happens, the definition of capacity should be modified in regard to the probabilistic view of breakdown. Polus and Pollatschek (2002) developed an algorithm to identify the speeds and flow rates which were adopted as a threshold for data analysis to determine capacity. Brilon et al. (2005) used traffic flow counts at a 5-min interval to demonstrate that capacity follows a Weibull distribution with almost a constant
shape parameter. Lorenz and Elefteriadou (2000) conducted extensive analysis of traffic speeds and flow rates which suggested a revised definition of capacity incorporating the stochastic breakdown component: “...the rate of flow (expressed in passenger car per hour per lane and specified for a particular time interval) along a uniform freeway segment corresponding to the expected probability of breakdown deemed acceptable under prevailing traffic and roadway conditions in a specified direction.” Lo and Tung (2003) models link capacity as uniformly distributed random variables. Wu et al. (2010) studied the stochasticity of freeway operational capacity and found that values of freeway operational capacity under different traffic flow conditions generally fit normal distributions. Dixit and Wolshon (2014) concludes that “there exists a consistent and fundamental difference between traffic dynamics under evacuation and those routines under non-emergency periods.” This suggests that traditional capacity estimation is not suitable for extreme condition. Tu et al. (2010) also shows that capacity is stochastic in nature.

2.2. Capacity Estimation Approaches

The following definitions are proposed to distinguish the different meanings of the various roadway capacities (Minderhoud et al., 1997).

Design capacity: A single value represents the maximum traffic volume that pass a cross section of a road with a probability under predefined road and weather condition. It is used for planning and designing roads.

Strategic capacity: A value represents the maximum traffic volume a road section can handle. This value or distribution is derived from observed traffic flow by static capacity models. It is useful for analyzing conditions in road networks.

Operational capacity: A value represents the actual maximum flow rate of the roadway. This capacity value is based on direct empirical capacity methods with dynamic
capacity models. It is valuable for short-term forecasting with which traffic control procedures may be performed.

There are various approaches to describe the capacity of a road. Depending on the time scale, capacity may differ. In this experiment, both day-to-day and hourly capacities are explored using empirical data. They represent roadway capacity under a selected probability that capacity value is lower than a given value. Various methods have been proposed to estimate capacity. Table 2.1 presents a summary with the assumptions and shortfalls of the methods; only methods for uninterrupted roadway capacity estimation are discussed.

According to (Minderhoud et al., 1997), the recommended superiority rank of aforementioned methods is: (a) product limit method, (b) the empirical distribution method and (c) the fundamental diagram method. Geistefeldt and Brilon (2009) finds that capacity estimation based on statistical models for censored data outperforms the direct estimation of breakdown probability for groups of traffic volumes. In addition, the product limit method has superior performance because the theory is well-documented and the calculation generates a capacity distribution instead of a single value. Therefore, the product limit method is adopted as the approach to estimate capacity. Capacity is represented as a distribution instead of treating capacity as a fixed value. When a probability is determined, a capacity value is selected accordingly.

2.3. Ramp Metering

Ramp metering is widely implemented in freeway access control. It helps to reduce delays by managing on ramp influx. It also improves safety, increases the travel time reliability, and reduces emission (Zhang and Levinson, 2010). Ramp metering strategies are summarized below (Papageorgiou and Kotsialos, 2000).
## TABLE 2.1: Summary of Roadway Capacity Estimation Methods

<table>
<thead>
<tr>
<th>Methods</th>
<th>Assumption/Requirement</th>
<th>Advantage/Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Headway Model</td>
<td>Distribution of constrained drivers (followers) at capacity level can be compared with the distribution at an intensity below capacity.</td>
<td>Only headways at one cross section of an arterial observed at an intensity below capacity are needed.</td>
</tr>
<tr>
<td>Bimodal Distribution</td>
<td>Two separate distributions represent the compound distribution of the observed flow rates.</td>
<td>That traffic demand can also be represented with a Gaussian-type distribution is doubtful and depends on the chosen observation period.</td>
</tr>
<tr>
<td>Method</td>
<td>Capacity estimated with a normal, Gaussian type distribution can be accepted without much resistance.</td>
<td>The number of capacity observations strongly affects the reliability of the calculated capacity value. Choosing the average value is rather arbitrary.</td>
</tr>
<tr>
<td>Selected Maxima Method</td>
<td>The road capacity is equal to the selected traffic flow maxima observed during the total observation period.</td>
<td>Observations for all sampling intervals are independently (flow rates between sampling intervals are not related) and identically (all counts are elements of the same distribution function) distributed. It implies, the mean flow rate during the observation period must be constant. Predicted capacity value strongly depends on the duration of the averaging interval.</td>
</tr>
<tr>
<td>Expected Extreme Value Method</td>
<td>Traffic volumes conform to a theoretical model such as the Poisson process.</td>
<td>Estimated capacity strongly depends on the duration of the averaging interval.</td>
</tr>
<tr>
<td>Asymptotic Method</td>
<td>Traffic volume observations for all averaging intervals are independently and identically distributed.</td>
<td>Estimated capacity strongly depends on the duration of the averaging interval.</td>
</tr>
<tr>
<td>Methods</td>
<td>Assumption/Requirement</td>
<td>Advantage/Disadvantage</td>
</tr>
<tr>
<td>-------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Product Limit Method</td>
<td>Requires more than a single day’s volume and speed measurement.</td>
<td>There is no information about the quality of the estimated capacity value.</td>
</tr>
<tr>
<td></td>
<td>A bottleneck location should be chosen to be sure of the capacity state of the road whenever congestion is detected upstream.</td>
<td></td>
</tr>
<tr>
<td>Fundamental Diagram Method</td>
<td>A mathematical model that fits the observed data.</td>
<td>The parameters of the chosen model need to be obtained for each location anew. Sufficient data over a broad range of intensities are needed to make a reliable curve fitting possible.</td>
</tr>
<tr>
<td>Knowledge is required of the actual prevailing capacities of road sections.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online Procedure</td>
<td>The relationship between intensity and occupancy may be adapted with a scaling factor only to fit the intensity-occupancy curve under various prevailing road, weather, and traffic conditions.</td>
<td>Determination of the critical occupancy in the capacity-estimation procedure is in question.</td>
</tr>
</tbody>
</table>
**Fixed-time strategy:** It is off-line for particular times-of-day. Constant historical demands are used instead of real-time measurements, which can be a drawback. Due to the absence of real-time measurements, fixed-time ramp metering strategies can lead to an overload of the mainline flow or an underutilization of the freeway.

**Reactive ramp metering strategies:** These are employed at a tactical level based on real-time measurements.

- **Local ramp metering:** This makes use of traffic measurements in the vicinity of the ramp to calculate suitable ramp metering values and includes demand-capacity strategy, occupancy strategy, ALINEA (Papageorgiou et al., 1991), neural networks and the ZONE algorithm etc (Demiral and Celikoglu, 2011). The trials show that ALINEA is superior to other local strategies and the no-control case with regards to total time spent, total traveled distance, mean speed, and mean daily congestion duration.

- **Multi-variable regulator strategy:** This strategies tries to hold freeway traffic conditions near desired values. Instead of performing independently for each ramp as local ramp metering, multivariable regulators make use of all available mainline measurements to simultaneously calculate the ramp volume values for all controlled ramps. The representative method is METALINE, a generalization and extension of ALINEA. It requires a rather complicated design procedure, but show no advantages over ALINEA under recurrent congestion. Due to comprehensive information, METALINE performs better than ALINEA in case of non-recurrent congestion.

**Nonlinear optimal ramp metering strategy:** This strategy calculates optimal and fair set values from a proactive, strategic point of view in real time. It considers current traffic states both on freeway and on-ramps, demand predictions over a
sufficiently long time horizon, limited storage capacity of the on-ramps, nonlinear traffic flow dynamics, etc.

**Integrated freeway network traffic control:** Designed and implemented independently, most of the control strategies fail to exploit synergistic effects. Suitable extension of the optimal control approaches presented above can lead to integrated freeway network control.

Replacing the deterministic capacity by a probabilistic one that changes dynamically based on real-time traffic conditions and varying probabilities of risk, Wu et al. (2010) improves the Minnesota ZONE metering algorithm by converting the capacity-constraint into a chance-constraint algorithm. Cassidy and Rudjanakanoknad (2005) shows that “metering an on-ramp can recover the higher discharge flow at a merge and thereby increase the merge capacity.” Zhang and Levinson (2010) finds that meters increase the bottleneck capacity by “postponing and sometimes eliminating bottleneck activations.” Papamichail et al. (2010) presents a nonlinear model-predictive hierarchical control approach for coordinated ramp metering of freeway networks. Gomes and Horowitz (2006) proposes an asymmetric cell transmission model (ACTM) to address on-ramp metering control problems. It resembles the cell transmission model but two allocation parameters are used instead of one in the case of merging flows. Simulation results suggest a 17.3% delay will be reduced when queue constraints are enforced. Smaragdis et al. (2004) develops an extension of ALINEA that enables automatic tracking of the critical occupancy in the aim of mainline flow maximization. The proposed AD-ALINEA strategy is valuable in the case that the critical occupancy cannot be estimated beforehand or is subject to real-time change. An upstream-measurement based strategy AU-ALINEA was also proposed. Both strategies exhibited good performance in a stochastic macroscopic simulation environment. For easy implementation and desirable results of local ramp metering strategy, a mean-standard deviation capacity-constraint ZONE algorithm is proposed to study the
stochastic capacity impact on ramp metering in this research.

2.4. Fourier Analysis v.s Wavelet Analysis

Fourier analysis is a mathematical method used for transforming the signal from time domain to a frequency domain. However, it has limitations that when the signal is transformed to the frequency domain, temporal information would be lost. Looking at a signal’s Fourier transformation, it is difficult to assess when a particular event happened. In addition, the precision of the information from Fourier analysis is determined by a fixed time window. Wavelet analysis allows the use of long time intervals for more precise low frequency information, and shorter intervals for high frequency information (Misiti et al., 2014).

When a signal is stationary, Fourier analysis can serve as a powerful tool to investigate the details of a signal. However, common signals contain several non-stationary components, such as trend, cycle, sudden change, breakdown points, and self-similarity. While these components are dominant in the signal, Fourier analysis is inadequate to detect them. Therefore, techniques like wavelet analysis become a nature alternative to investigate the original signal.

Wavelet analysis can reveal aspects of data that other techniques missed, i.e, discontinuity. One major advantage of wavelets is its ability to perform local analysis, that is, to analyze one segment of a larger signal. It can also compress and denoise a signal without appreciable degradation (Misiti et al., 2014).

In this research, Wavelet Analysis was chosen primarily because the empirical capacity data is nonstationary with cycle and trend signals, and the temporal information is of interest.


2.5. Short-term Traffic Forecasting

According to Giebel and Kariniotakis (2009), there are five timescales for prediction model: ultra short-term range (in seconds), very short-term range (1-10 minutes to an hour), short-term range (0 to 6 or 8 hours), medium-term range (0 hour to 7 days ahead), and long-term range (weeks). Various models have been used for short-term traffic flow forecasting (Chang et al., 2011), such as Historical Average methods, Time Series methods (i.e., Autoregressive Integrated Moving Average(ARIMA)) (Voort et al., 1996; Lee and Fambro, 1999; Koutroumanidis et al., 2009), State-Space methods (Stathopoulos and Karlaftis, 2003), Non-parametric methods (Artificial Neural Networks (ANN) (Smith and Demetsky, 1994; Chang and Su, 1995; Innamma, 2005; Taylor et al., 2007), Non-parametric regression (Smith et al., 2002), Support Vector Machine (SVM) (Castro-Neto et al., 2009)) and Hybrid methods. Wang et al. (2015) summarized three categories of forecasting models: Statistical models (Holt-Winters, Box-Jenkins, Multivariate State Space, Nonlinear Dynamics), Hybrid models (Empirical Mode Decomposition (EMD)-ANN(Liu et al., 2012), EMD-ARIMA (Wang et al., 2015), EMD-SVM (Bao et al., 2012), ARIMA-SVM (Pai and Lin, 2005; Zhu and Wei, 2013), ARIMA-EGARCH-ANN (Kumar and Thenmozhi, 2012)), and AI-based methods (ANN model, Kalman Filter method (Okutani and Stephanedes, 1984; Xie et al., 2007), SVM model, Fuzzy Logic (Zhang and Ye, 2008), Neuron-Genetic (Abdulhai et al., 2002)).

It is worth noting that the time series and non-parametric methods have been more extensively investigated in literature. Lippi et al. (2013) conducted reviews with respect to different statistical and machine learning techniques on traffic forecasting, and two new support vector regression (SVR) models incorporating seasonal kernels were presented. The result shows SARIMA model with a Kalman filter has relatively better performance on average, while the proposed seasonal SVR is appropriate for the most congested condi-
tions. Park et al. (2007) provided a test of RBF (Radial Basis Function) neural network in forecasting freeway traffic volumes and demonstrated a better performance over Taylor series, Exponential Smoothing Method (ESM), double exponential smoothing method, and back propagation neural network. However, 35% of the forecasted traffic volumes exhibited 10% errors and thus developing a more reliable traffic-forecasting model was suggested as future work. Cortez et al. (2012) presented three methods, a novel neural network ensemble approach and two adapted time series model, ARIMA and Holt-Winters. Both methods were adopted to forecast the traffic in TCP/IP based network. It suggests neural ensemble model achieved better results for 5 min and hourly prediction, while Holt-Winters is appropriate for daily forecasts. Dia (2001) discussed an object-oriented neural network model for predicting traffic conditions. R. Chrobok and Schreckenberg (2004) examined two forecasting methods and implied the constant model provides a good prediction for short horizons while the heuristics is better for longer intervals. A method combining both models was proposed. Chan and Singh (2011) pursued a neural network based on an exponential smoothing method. Regarding to the test error, the proposed approach achieved more than an 8% rate of improvement relative to the neural network developed based on original traffic flow data.
3. METHODOLOGY AND DATA DESCRIPTION

Considering the advantages and limitations of each capacity estimation method, product limit method is adopted in this research. While most of the studies neglected the fact that highway capacity is spatially and temporally correlated, a space-time ARIMA analysis is employed. Furthermore, in order to conduct a precise forecasting, a hybrid model that combines Wavelet Transform and Dynamic Neural Time series is proposed to conduct multiple days’ forecasting on basic segments from Georgia State route 400 (GA400). Wavelet analysis is used to decompose the capacity series into different sub-series with varying frequencies. These sub-series serve as the input of the forecasting method. Dynamic Neural Time Series combines the time series concept and training mechanism of dynamic neural network, provides accurate forecasting results. Single model like ARIMA and Dynamic Neural Time Series are employed to compare with the hybrid model proposed and accuracy measures are calculated to demonstrate the forecasting performance. A brief introduction of the methods involved in this experiment is given as follows.

3.1. Product Limit Method

Minderhoud et al. (1997) introduced the following formulation for the derivation of freeway capacity distribution function. It is based on statistical methods and commonly used for lifetime or failure data analysis.

\[ F(q) = \text{Prob}(q_c \leq q) \]  
\[ (3.1) \]

More specifically, it could be written as

\[ F(q) = \frac{N_c}{N} \]  
\[ (3.2) \]
where

\[ F(q) = \text{cumulative distribution function of capacity}; \]
\[ q_c = \text{capacity value}; \]
\[ q_i = \text{observed traffic volume at interval } i; \]
\[ N_c = \text{number of observation elements } i \text{ from } C \text{ with } q_i \text{ less than or equal to } q; \]
\[ N = \text{total number of observations } i \text{ in set } C; \] and
\[ C = \text{set of observed congested flow measurement.} \]

After specifying a \( F(q) \), for example, \( F(q) = 0.5 \), a single capacity value \( q_c \) is estimated. A nonparametric approach, product limit method, is used to estimate the capacity distribution function. Function \( G(q) \) is defined as the probability that the capacity value \( q_c \) is higher than a given value \( q \). The function \( F(q) \) is then defined by \( 1 - G(q) \). The general expression of the product limit method is presented below (Minderhoud et al., 1997) (Brilon et al., 2005).

\[
G(q) = \text{Prob}(q_c > q) \tag{3.3}
\]
\[
G(q) = 1 - \prod_{q_i} \frac{k_{q_i} - 1}{k_{q_i}}; i \in C \tag{3.4}
\]

where

\[ K_q = \text{number of observation elements } l \text{ in set } S \text{ with volume } q_i \text{ larger than or equal to } q; \]
\[ C = \text{set of observed congested flow volume}; \]
\[ Q = \text{set of observed free-flow volume}; \]
\[ S = Q \cup C \text{ and } S \text{ is set of all observations } l. \]

A simple example is presented to understand the calculation of product limit method. As shown in Table 3.1, 15-min interval is used to calculate the capacity. Derived density data are used to classify the observations into set \( \{ C \} \) or set \( \{ Q \} \). Critical density is \( k_c \).

In the last column, \( F(q) = 1 - G(q) \) was derived. In this example, defining capacity at \( F(q) = 0.5 \) (the median) would result in a road capacity of 4200 veh/hr per two lanes. If only observations were used, an average of 4025 veh/hr per two lanes would be resulted.
### TABLE 3.1: Product Limit Method Calculation

<table>
<thead>
<tr>
<th>Interval i</th>
<th>$q_i$ [veh/h]</th>
<th>Set</th>
<th>Order $j$</th>
<th>$K_q, q_i \in {C}$</th>
<th>$G(q)$</th>
<th>$F(q)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>16:00-16:15</td>
<td>3200</td>
<td>C</td>
<td>3</td>
<td>6</td>
<td>$5/6=0.83$</td>
<td>0.17</td>
</tr>
<tr>
<td>16:15-16:30</td>
<td>2600</td>
<td>Q</td>
<td>1</td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>16:30-16:45</td>
<td>3000</td>
<td>Q</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16:45-17:00</td>
<td>4000</td>
<td>C</td>
<td>4</td>
<td>5</td>
<td>$5/6 \cdot 4/5=0.67$</td>
<td>0.33</td>
</tr>
<tr>
<td>17:00-17:15</td>
<td>4200</td>
<td>Q</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17:15-17:30</td>
<td>4400</td>
<td>C</td>
<td>7</td>
<td>2</td>
<td>$5/6 \cdot 4/5 \cdot 1/2=0.33$</td>
<td>0.67</td>
</tr>
<tr>
<td>17:30-17:45</td>
<td>4500</td>
<td>C</td>
<td>8</td>
<td>1</td>
<td>$5/6 \cdot 4/5 \cdot 1/2 \cdot 0/1=0$</td>
<td>1</td>
</tr>
<tr>
<td>17:45-18:00</td>
<td>4100</td>
<td>Q</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Capacity = $(4000+4400)/2=4200$

However, this method requires data collected from a bottleneck, empirical dataset may not meet this criteria in the reality. In this thesis, therefore, product limit method combined with a worst 15-min interval calculation according to (HCM, 2010) are adopted to estimate the freeway capacity. When congestion was detected, product limit method would be utilized to estimated highway capacity. When there are no congestions, the worst 15-min interval is employed to estimate a capacity. The raw GA400 data was extracted at 20 seconds intervals from the video. In a 15-min interval calculation, capacity per day per lane was calculated by the following procedure: every 15-min traffic flow was calculated starting from 12:00am (Flow rate = Auto flow + Van flow + Truck flow × PCE_e); the maximum output is selected and multiplied by 4, which produces the capacity of the day.
3.2. Space-Time ARIMA

The space-time autoregressive integrated moving average (STARIMA) model is characterized as a weighted linear combination of observations and innovations at time \( t \) and location \( i \). The mechanism for this expression is the hierarchical ordering of the neighbors of each site and corresponding sequence of weighting matrices (Kamarianakis and Prastacos, 2005).

\[
W_1 = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 \\
1/2 & 0 & 1/2 & 0 & 0 \\
0 & 1/2 & 0 & 1/2 & 0 \\
0 & 0 & 1/2 & 0 & 1/2 \\
0 & 0 & 0 & 1 & 0
\end{bmatrix} \quad W_2 = \begin{bmatrix}
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
1/2 & 0 & 0 & 0 & 1/2 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0
\end{bmatrix}
\]

If \( Z_t \) is a \( N \times 1 \) vector of observations at time \( t \) at \( N \) locations, then the STARIMA model is of the form

\[
Z_t = \sum_{k=1}^{p} \sum_{l=0}^{\lambda_k} \phi_{kl} W_l Z_{t-k} - \sum_{k=1}^{q} \sum_{l=0}^{m_k} \theta_{kl} W_l a_{t-k} + a_t \tag{3.5}
\]

where \( p \) is the autoregressive order, \( q \) stands for the moving average order, \( \lambda_k \) is the spatial order of the \( k \)th autoregressive term, \( m_k \) refers to the order of the \( k \)th moving average term, \( \phi_{kl} \) and \( \theta_{kl} \) represent parameters need to be estimated and \( W_l \) is the \( N \times N \) matrix for spatial order \( l \), and \( a_t \) is the random normally distributed innovation or disturbance vector at time \( t \). In this research, detectors are represented as the data points as in

![FIGURE 3.1: Distance degree demonstration](image)
Figure 3.1. The detectors are placed along the freeway mainline, a simple weight scheme is therefore adopted without complex structure. Also, they are nearly evenly located, which enables us to assign the same weight on different detectors. As aforementioned, $Z_t$ is input of the model, it represents the data collected from different detectors or regions at various time steps.

3.2.1 Seasonal STARIMA

Traffic flow has seasonality, a seasonal component should be included (Kamarianakis and Prastacos, 2005). Seasonal STARIMA model is then formulated as

$$
\Phi_{P,\Lambda}(B^S)\phi_{p,\lambda}(B)\nabla_S^D\nabla^d Z_t = \Theta_{Q,M}(B^S)\theta_{q,m}(B)\alpha_t
$$

(3.6)

where

$$
\Phi_{P,\Lambda}(B^S) = I - \sum_{k=1}^{P} \sum_{l=0}^{\Lambda_k} \Phi_{kl}W_lB^{klS}
$$

(3.7)

$$
\phi_{p,\lambda}(B) = I - \sum_{k=1}^{p} \sum_{l=0}^{\lambda_k} \phi_{kl}W_lB^k
$$

(3.8)

$$
\Theta_{Q,M}(B^S) = I - \sum_{k=1}^{Q} \sum_{l=0}^{M_k} \Theta_{kl}W_lB^{klS}
$$

(3.9)

$$
\theta_{q,m}(B) = I - \sum_{k=1}^{q} \sum_{l=0}^{m_k} \theta_{kl}W_lB^k
$$

(3.10)

$\Phi_{kl}$ and $\phi_{kl}$ are the seasonal and nonseasonal autoregressive parameters at temporal lag $k$ and spatial lag $l$, accordingly; $\Theta_{kl}$ and $\theta_{kl}$ are the seasonal and nonseasonal moving average parameters at temporal lag $k$ and spatial lag $l$ similarly. $P$ and $p$ refer to the seasonal and nonseasonal autoregressive orders; $Q$ and $q$ represent the seasonal and nonseasonal moving average orders. $\Lambda_k$ and $\lambda_k$ are the seasonal and nonseasonal spatial orders for the $k$th autoregressive term; $M_k$ and $m_k$ are the seasonal and nonseasonal spatial orders for the $k$th moving average term. $D$ and $d$ refer to the required number of seasonal and nonseasonal differences, respectively, where $\nabla_S^D$ and $\nabla^d$ are the seasonal and
nonseasonal difference operators, for instance, \( \frac{D}{S} = (I - B^S)^D \) and \( d = (I - B)^d \). \( \alpha_t \) is a normally distributed random error vector at time \( t \) subject to

\[
E\{\alpha_t\} = 0 \tag{3.11}
\]

\[
E\{\alpha_t^{\prime}\alpha_{t+s}\} = \begin{cases} G & s = 0 \\ 0 & s \neq 0 \end{cases} \tag{3.12}
\]

\[
E\{Z_t\} \alpha_{t+s}^{\prime} = 0, s > 0 \tag{3.13}
\]

Equation 3.6 represents a seasonal multiplicative STARIMA model of order \((p_\lambda, d, q_m)\times(P_\Lambda, D, Q_M)_S\) (Kamarianakis and Prastacos, 2005).

### 3.2.2 Model Validation

The choice of spatial order is critical. It is important that \( l \) be at least as large as the maximum spatial order of any hypothesized model. However, larger \( l \) generally requires higher computing power. The size of the system should be considered in the decision making process. Larger system would warrant a higher \( l \), for example 3 or 4, but for moderate systems \( l = 2 \) will suffice (Pfeifer and Deutrch, 1980). In order to ensure the comprehensiveness of the experiment, a spatial order 3 is utilized to explore the Spatial and Temporal correlation of highway capacities at varying locations.

According to Pfeifer and Deutrch (1980), “Equally scaled weights are chosen not only because they are useful in their own right, especially toward modeling regularly spaced homogeneous spatial systems, but also because they can serve as a pattern for more general weighting schemes”. In this study, equal weights are assigned to each detector to formulate a general stochastic capacity model.

To identify space-time models, a good way is “to combine \( N^2 \) possible cross-covariance between all possible pairs of sites in a logical manner consistent with the forms associated with the proposed model class” (Pfeifer and Deutrch, 1980). The result is space-time autocorrelation function (ACF). It represents the covariance between points lagged both in
space and time. Despite each model in STARIMA family has a unique space-time ACF, but the selection of an appropriate candidate model is still challenging. This problem could be solved by another measure, space-time partial autocorrelation function (PACF).

3.3. Wavelet analysis

The Wavelet transform can be divided into two categories: continuous wavelet transform (CWT) and discrete wavelet transform (DWT). The continuous wavelet transform $W(a, b)$ of signal $f(x)$ with a mother wavelet $\phi(x)$ is defined as (Misiti et al., 2014; Catalao et al., 2011):

$$W(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(x) \phi\left(\frac{x - b}{a}\right) dx$$  \hspace{1cm} (3.14)

Where, $a$ denotes a scale parameter and controls the spread of wavelet; $b$ is a translation parameter and determines its central position. $W(a, b)$ shows how well the original signal $f(x)$ matches the translated mother wavelet. Through continuously scaling and translating of $\phi(x)$, a fair amount of redundant information is generated by CWT. If scales and positions are chosen based on powers of the two, so-called dyadic scales and positions, then the analysis will be more efficient and just as accurate. It is known as discrete wavelet transform (DWT), and defined as:

$$W(m, n) = 2^{(-m/2)} \sum_{t=0}^{T-1} f(t) \phi\left(\frac{t - n \cdot 2^m}{2^m}\right)$$  \hspace{1cm} (3.15)

Where $T$ denotes the length of a signal $f(t)$; $t$ represents the discrete time index; and the scaling and translation parameters become the functions of the $m$ and $n \ a = 2^m, \ b = n \cdot 2^m$

An efficient way to implement this scheme using filters developed by Mallat (1989) is used here. This practical filtering algorithm yields a fast wavelet transform. By successively decomposing the approximations, a multilevel decomposition tree is achieved by breaking down the original signal into lower resolution components as shown in Figure 3.2 (Misiti et al., 2014). “Approximation” and “Details” from a given signal are
obtained from multi-resolution. The approximations are at the high scale, low frequency components of the signal, while the details are at the low scale, high frequency components (Misiti et al., 2014).

The order of the Daubechies functions represents the number of vanishing moments of the wavelet function. The wavelets are built based on the function

\[ \tau(t) = \sqrt{2} \sum_{n}^{N} l_n \tau(2t - n) \]  

(3.16)

Where \( \tau(t) \) is the scale function and \( l_n \) denotes the low frequency filter coefficients. The mother wavelet \( \phi(t) \) is defined by

\[ \tau(t) = \sqrt{2} \sum_{n}^{N} h_n \tau(2t - n) \]  

(3.17)

where \( h_n \) is the high frequency filter coefficients.

Considering the nonlinear features in capacity series, a wavelet function of Daubechies of order 3 (db3) is adopted as the mother wavelet \( \phi(t) \) in this experiment, because it encodes 3-polynomials, i.e. constant, linear and quadratic signal components. The Daubechies wavelet db3 offers an appropriate trade-off between wavelength and smoothness, providing a good forecasting performance. A3, D1, D2 and D3 as shown in Figure 3.2 will be used as inputs for forecasting method.
3.4. Dynamic neural network

Neural networks are composed of simple elements, which are inspired by biological nervous system in parallel. The network is adjusted, based on comparisons of the output and the target, until the network output match the target. The essence of neural network is that such parameters can be altered so that the network demonstrates desired behavior. Therefore, the network is trained by adjusting the weight or bias parameters, or the network itself fine-tunes the parameters to achieve the results (Demuth and Beale, 2014).

An illustration of static neural network is shown in Figure 3.3, the solid vertical bar \( p \) denotes the input vector; \( R \) is the dimensions of the \( p \). These inputs multiply the weight matrix \( W \), while by multiplying a scalar bias \( b \) by a constant, as is shown in the diagram. Then the sum of bias and the product \( Wp \) is passed to the transfer function \( f \) to get the neuron’s output \( a \). Commonly used transfer functions are Hard-Limit Transfer Function, Linear Transfer Function and Log-Sigmoid Transfer Function.

Forecasting is one kind of dynamic filtering, in which past values of one or more time series are used to predict future values. Dynamic neural network includes tapped delay lines used for nonlinear filtering and forecasting. Here, forecasting is considered as a Nonlinear Autoregressive (NAR) problem, that is to predict series \( y(t) \) given \( d \) past values.
of $y(t)$.

$$y(t) = f(y(t - 1), ..., y(t - d))$$  \hfill (3.18)

In a dynamic neural network, as is shown in Figure 3.4 (Demuth and Beale, 2014), feedback and delay are added to better train the network. With the updated information, the output would reach closer to the target. Also, dynamic network requires that the input be placed in sequence, and the network produce a cell array containing sequential outputs, which matches our expectation of forecasting with time information. The trained neural network of approximation and details of this study are shown in Figure 3.5.

In a dynamic neural network, as is shown in Figure 3.4 (Demuth and Beale, 2014), feedback and delay are added to better train the network. With the updated information, the output would reach closer to the target. Also, dynamic network requires that the input be placed in sequence, and the network produce a cell array containing sequential outputs, which matches our expectation of forecasting with time information. The trained neural network of approximation and details of this study are shown in Figure 3.5.

3.5. Study Area and Data Description

This research uses data that were collected by Georgia Department of Transportation ITS program on state route GA400 through the whole year of 2003. The data interval is 20 seconds. This route is the North/South freeway entering/exiting city of Atlanta. The
data were collected over 20 miles long corridor with 100 detectors, in which 78 detectors were on basic freeway and 22 of them were installed at On/Off ramps, as shown in Figure 3.6.

The time interval for capacity estimation is important and impacts the resulting capacity values. Different time intervals are adopted in other researches, e.g. Brilon et al. (2005) adopt a 5-min interval. The time interval suggested by Transportation Research Board (2000) for capacity observation is 15 min (Polus and Pollatschek, 2002). Minderhoud et al. (1997) suggests “15-min interval appears to be a good compromise because the independence of the observation between the averaging intervals can be defended, local fluctuation are smoothed out, and the maximum traffic volume can hold for more than the duration of the interval ”. Thus, we utilized a 15-min interval to measure the daily capacity. The data investigated in this thesis were obtained from the basic freeway segments.
FIGURE 3.6: Study Site: GA400 100 Stations.
4. SPATIAL AND TEMPORAL EMPIRICAL ANALYSIS

Day-to-day capacity is vital to traffic controls such as lane closure, work zone. On a daily basis, capacity change can be caused by factors such as weather conditions, which can be accounted for as a temporal effect. Spatial information like geometrical effects should also be considered. However, how temporal and spatial effects interact is not explicitly explored. A spatial-temporal analysis is thus conducted to characterize the stochasticity in capacity.

As Pfeifer and Deutrch (1980) characterized, “… Processes amenable to modeling via this class are characterized by a single random variable observed at N fixed sites in space wherein the dependencies between the N time series are systematically related to the location of the sites. A hierarchical series of $N \times N$ weighting matrices specified by the model builder prior to analyze the data is the basic mechanism for incorporating the relevant physical characteristics of the system into the model form. Each of the N time series is simultaneously modeled as a linear combination of past observations and disturbances at neighbor sites. Just as univariate ARIMA models reflect the basic idea that the recent past exerts more influence than the distant past, so STARIMA models reflect (through the specification of the weighting matrices) the idea that near sites exert more influence in each other than distant ones.” Therefore, STARIMA model expresses each observation at time $t$ and location $i$ as a weighted linear combination of observations and innovations both in space and time. The mechanism for this expression is the hierarchical ordering of the neighbors of each site and corresponding sequence of weighting matrices (Kamarianakis and Prastacos, 2005).
4.1. STARIMA Experiment

The experiment is conducted on northbound of GA400. There are 38 detectors along the mainline, in order to keep the continuity of the data, 10 detectors are deleted for the missing value. 28 detectors’ whole year weekdays’ data are employed in this study. The STARIMA modeling is applied to the capacity series, the autocorrelations and partial autocorrelations are presented as in Table 4.1.

![Nonseasonal STARIMA ACF & PACF](image)

**FIGURE 4.1:** Nonseasonal STARIMA ACF & PACF

Figure 4.1 suggests that the space-time autocorrelations is decaying temporally but not spatially. Hence, the first order seasonal difference is needed for transformation. Table 4.2 shows the pattern of autocorrelations and partial autocorrelations.
### TABLE 4.1: Nonseasonal STARIMA Autocorrelations and Partial Autocorrelations

<table>
<thead>
<tr>
<th>Spatial lag (l)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time lag (k)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.293</td>
<td>0.201</td>
<td>0.205</td>
<td>0.228</td>
<td>1.109</td>
<td>-0.662</td>
<td>-0.558</td>
<td>-0.544</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.111</td>
<td>0.020</td>
<td>0.024</td>
<td>0.050</td>
<td>0.429</td>
<td>-0.340</td>
<td>-0.288</td>
<td>-0.331</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.075</td>
<td>-0.011</td>
<td>-0.002</td>
<td>0.020</td>
<td>0.586</td>
<td>-0.467</td>
<td>-0.384</td>
<td>-0.424</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.073</td>
<td>0.001</td>
<td>0.008</td>
<td>0.032</td>
<td>0.527</td>
<td>-0.391</td>
<td>-0.332</td>
<td>-0.363</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.099</td>
<td>0.019</td>
<td>0.025</td>
<td>0.044</td>
<td>0.595</td>
<td>-0.428</td>
<td>-0.357</td>
<td>-0.416</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.075</td>
<td>0.003</td>
<td>0.007</td>
<td>0.029</td>
<td>0.553</td>
<td>-0.414</td>
<td>-0.353</td>
<td>-0.385</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.005</td>
<td>-0.052</td>
<td>-0.047</td>
<td>-0.035</td>
<td>0.421</td>
<td>-0.360</td>
<td>-0.292</td>
<td>-0.389</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.040</td>
<td>-0.024</td>
<td>-0.019</td>
<td>-0.003</td>
<td>0.640</td>
<td>-0.487</td>
<td>-0.414</td>
<td>-0.435</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.053</td>
<td>-0.007</td>
<td>-0.001</td>
<td>0.010</td>
<td>0.525</td>
<td>-0.389</td>
<td>-0.315</td>
<td>-0.395</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.032</td>
<td>-0.027</td>
<td>-0.018</td>
<td>-0.012</td>
<td>0.434</td>
<td>-0.363</td>
<td>-0.297</td>
<td>-0.373</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.082</td>
<td>0.025</td>
<td>0.026</td>
<td>0.044</td>
<td>0.738</td>
<td>-0.510</td>
<td>-0.441</td>
<td>-0.453</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.018</td>
<td>-0.042</td>
<td>-0.041</td>
<td>-0.028</td>
<td>0.360</td>
<td>-0.311</td>
<td>-0.261</td>
<td>-0.337</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.020</td>
<td>-0.036</td>
<td>-0.034</td>
<td>-0.024</td>
<td>0.585</td>
<td>-0.474</td>
<td>-0.405</td>
<td>-0.439</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.086</td>
<td>0.034</td>
<td>0.036</td>
<td>0.045</td>
<td>0.602</td>
<td>-0.409</td>
<td>-0.341</td>
<td>-0.387</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.064</td>
<td>0.020</td>
<td>0.025</td>
<td>0.041</td>
<td>0.734</td>
<td>-0.575</td>
<td>-0.497</td>
<td>-0.508</td>
<td></td>
</tr>
<tr>
<td>Spatial lag (l)</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.126</td>
<td>0.120</td>
<td>0.118</td>
<td>0.119</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-0.455</td>
<td>-0.360</td>
<td>-0.341</td>
<td>-0.328</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-0.130</td>
<td>-0.127</td>
<td>-0.124</td>
<td>-0.123</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-0.019</td>
<td>-0.009</td>
<td>-0.006</td>
<td>-0.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.072</td>
<td>0.064</td>
<td>0.063</td>
<td>0.064</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.020</td>
<td>0.018</td>
<td>0.014</td>
<td>0.016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>-0.085</td>
<td>-0.070</td>
<td>-0.072</td>
<td>-0.076</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>-0.012</td>
<td>-0.012</td>
<td>-0.013</td>
<td>-0.010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.015</td>
<td>0.011</td>
<td>0.014</td>
<td>0.010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.005</td>
<td>0.009</td>
<td>0.016</td>
<td>0.005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.045</td>
<td>0.044</td>
<td>0.043</td>
<td>0.052</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>-0.052</td>
<td>-0.057</td>
<td>-0.063</td>
<td>-0.056</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>-0.056</td>
<td>-0.062</td>
<td>-0.065</td>
<td>-0.072</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.031</td>
<td>0.029</td>
<td>0.031</td>
<td>0.026</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.032</td>
<td>0.036</td>
<td>0.037</td>
<td>0.042</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time lag (k)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0.342</td>
<td>-0.061</td>
<td>-0.057</td>
<td>-0.025</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-0.748</td>
<td>-0.009</td>
<td>0.009</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.190</td>
<td>-0.002</td>
<td>-0.013</td>
<td>0.004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-0.439</td>
<td>-0.083</td>
<td>-0.052</td>
<td>-0.026</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.101</td>
<td>0.055</td>
<td>0.037</td>
<td>0.031</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-0.245</td>
<td>-0.108</td>
<td>-0.078</td>
<td>-0.034</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-0.047</td>
<td>0.086</td>
<td>0.058</td>
<td>0.027</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>-0.124</td>
<td>-0.110</td>
<td>-0.077</td>
<td>-0.028</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>-0.021</td>
<td>0.063</td>
<td>0.037</td>
<td>0.013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-0.179</td>
<td>-0.046</td>
<td>-0.011</td>
<td>-0.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.123</td>
<td>-0.014</td>
<td>-0.037</td>
<td>-0.016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>-0.279</td>
<td>0.036</td>
<td>0.051</td>
<td>0.033</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.083</td>
<td>-0.019</td>
<td>-0.044</td>
<td>-0.025</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>-0.127</td>
<td>0.014</td>
<td>0.027</td>
<td>0.014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.030</td>
<td>-0.002</td>
<td>-0.018</td>
<td>-0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 4.2: Seasonal STARIMA Autocorrelations and Partial Autocorrelations**
4.2. Seasonal STARIMA Analysis

![Seasonal STARIMA ACF and PACF](image)

FIGURE 4.2: Seasonal STARIMA ACF and PACF

Table 4.2 shows the pattern of autocorrelations and partial autocorrelations. Observed from Figure 4.2, the space-time autocorrelation is significant in time lag 2 and cuts off at spatial lag 3; the partial autocorrelations decays both in time and space. This is characterized by the space-time MA factor. Therefore, the candidate space-time ARIMA model would be Seasonal STARIMA \((0, 0, 2_3) \times (0, 1, 0)_2\). It implies that an observed station’s performance could be severely influenced by the stations that locate within three degrees.

4.3. Discussion

The result suggests that a short-term seasonality lies in highway capacity. This can be explained from a psychological perspective. A high highway capacity indicates a higher possibility of traffic congestion. Drivers would change travel schedules or routes based on previous days’ experiences which lead to the capacity drop and vice versa. Accident should be taken into account in modeling the stochastic highway capacity. However, due to the limitation of the data, no statistical evidence can be presented in this stage. This
empirical study is conducted with weekday data, traffic flow is thus mainly consisted of work commutes. Therefore, weather conditions have a minor influence on capacity except for the case of inclement weather. Also, as leisure travel is mostly concentrated on weekends, weekday data can eliminate some of the exogenous factors of uncertainty.
5. APPLICATION I: CORRIDOR LEVEL RAMP METERING

How to incorporate the stochastic capacity in real-time traffic control and evaluate what additional benefits can be gained are the central arguments which motivates this case study. Ramp metering is chosen as an application because ramps are the joints within the transportation network. More importantly, the ramps are where the merging and diverging traffic occurs. Travelers often observe long queues at on-ramps and the queue spill-back phenomenon particularly in peak hours. Therefore, ramps could be vulnerable depending on the performance of the ramp-metering control scheme. To fulfill this end, a mean-standard deviation capacity-constraint ZONE algorithm embedded cell transmission model (CTM) is proposed.

5.1. Capacity Mean-Standard Deviation Trade-off

Generally, ramp metering is a throughput maximization problem given a deterministic capacity constraint (Wu et al., 2010). Zone algorithms aim to maximize freeway throughput while balancing traffic entering and leaving the zone Lau. (1996). In Minnesota’s ZONE metering algorithm, capacity constraint is applied to regulate the input flow (Lau., 1996).

\[(A + E - X) \leq (S + B)\]  \hspace{1cm} (5.1)

where \(A\) is the upstream mainline volume; \(E\) is the summation of entrance ramp flow; \(X\) is the summation of exit ramp flows; \(S\) is the vehicle storage on freeway, and \(B\) is the downstream bottleneck capacity.

In equation 5.1, \(B_{z,t}\) is a fixed number. However, capacity is a stochastic concept, without considering the inherent randomness, the traffic control scheme designs are far from optimal. For example, if capacity is overestimated, the roadway would be underuti-
lized. If the capacity were actually lower than expected, traffic congestion would occur. The mean capacity is normally used to measure a system’s performance, but it is inadequate to use in designing a traffic control scheme. Generally, the mean is understood as first order information, and variability is considered as second order information. Thus, how to take this second order information (variability) into account with stochastic capacity is crucial. In order to avoid being overaggressive (without considering variability of capacity at all) or conservative with traffic control, a trade-off needs to be pursued between efficiency and robustness.

Yin (2008) presents three models to determine the robust optimal signal timings. The scenario-based mean-variance optimization is relatively easier to implement due to the simple formulation which presents a reasonable approach to achieve a trade-off between robustness and efficiency. Nagengast et al. (2011) later introduces another mean-variance model to balance the benefits and returns. Enlightened by Nagengast et al. (2011) and Yin (2008), a modified mean-standard deviation scheme is proposed based on the empirical characteristics of highway capacity.

\[ B = \mu + \gamma \sigma \]  \hspace{1cm} (5.2)

where \( B \) is the equivalent downstream bottleneck capacity, \( \mu \) denotes the mean of the empirical capacity observation, \( \sigma \) represents the standard deviation of the capacity, and \( \gamma \) is a weighting parameter represents the designer’s attitude towards the risk. \( \gamma < 0 \) represents a risk averse strategy; \( \gamma = 0 \) indicates it is risk neutral; while \( \gamma > 0 \) is risk seeking. According to Chebyshev’s inequity (Kvanli et al., 2005), 75% of the values lie within two standard deviations, that is \( \mu \pm 2\sigma \). It is reasonable to assume that the probability of extreme traffic conditions happening as fairly low. Thus \( \gamma \) falls in the range of \([-2, 2]\).

Assuming the distribution of the hour-to-hour capacity in the specific location is known, the deterministic capacity is \( q_c \). If the probability that real capacity \( q \) less than \( q_c \)
is larger than 0.5, then the chances that the empirical capacity exceeds the predetermined capacity is less likely to happen. Then, weighting parameter $\gamma$ should represent a risk averse attitude. Similarly, if the probability that real capacity $q$ less than $q_c$ and is less than 0.5, the chance that the empirical capacity exceeds the predetermined capacity is more likely to happen. The weighting parameter $\gamma$ would then represent a risk taking attitude.

The reason it was called “attitude towards risk” is that even though the probability is less than 0.5, there is still a chance of occurring. Hence, selecting $\gamma$ is also taking the risk into account.

Similar to the scenario-based mean-standard deviation optimization in Yin (2008), a set of capacities $\Omega = 1, 2, 3, ..., K$ is introduced to represent the stochasticity of highway capacity. We assume that within one day the capacities in a specific location are independent. For each capacity $C_k \in \Omega$, the probability of occurrence is $p_k$. Although criticized by the additional efforts to specify the scenarios, computational experiments in Yin (2008) demonstrate that “relatively small number of scenarios will be able to produce near-optimal policies.”

\[
\mu = \sum_{k \in \Omega} p_k C_k \quad (5.3a)
\]
\[
\sigma = \sqrt{\sum_{k \in \Omega} p_k (C_k - \sum_{k \in \Omega} p_k C_k)^2} \quad (5.3b)
\]

In this study, a simplified scenario is designed to prove the concept. Only the upstream mainline volume($A$) and the volume from local access metered ramps ($M$) are considered (only one on-ramp for this experiment). Combined with the mean-standard deviation trade-off scheme, the modified stochastic capacity-constraint ZONE algorithm is presented
as follows.

\[
\begin{align*}
Max & \quad A + M \\
A + M & \leq B \\
B &= \sum_{k \in \Omega} p_k C_k + \gamma \sqrt{\sum_{k \in \Omega} p_k (C'_k - \sum_{k \in \Omega} p_k C_k)^2} \\
M & \geq M_{\min} \\
M & \leq M_{\max}
\end{align*}
\] (5.4)

where \( \gamma \) is a weighting parameter, \(-2 \leq \gamma \leq 2\). The first component is the mean of the capacity across all capacity scenarios while the second represents the standard deviation of the capacity. The parameter \( \gamma \) reflects the amount of capacity variation engineer wants to include, also called as “attitude towards risk.”

5.2. Weighting Parameter Estimation

Deploying the capacity estimation methodology proposed in Section 2.2., an hour-to-hour capacity can be estimated. After excluding invalid data points, hour-to-hour capacities in June’s weekdays are plotted in Figure 5.1. Horizontally, hour-to-hour capacity within one day changes dramatically. Vertically, capacity varies from day to day in the same hour. Utilizing the mean capacity will lead to missing important variation details, resulting in a decrease of network reliability. In such a case, using deterministic capacity in traffic control design will significantly deteriorate the system’s efficiency. Therefore, variations in different hours’ capacity should be characterized to increase system robustness. Further, how to employ the second order information, such as capacity variation we gained from history data in the traffic operation is worth pursuing.

In this study, rush hours 7am, 8am, 9am, 4pm, 5pm and 6pm are selected to demonstrate the idea. For instance, the whole year’s capacity at 8am can be estimated and collected; the capacity series then can be empirically formulated into a cumulative distribution function. If the traditional capacity is defined as 2000 veh/hr/lane, then the
FIGURE 5.1: Hour-to-hour Capacity

probability that capacity at 7am is less than the 2000 veh/hr/lane is 0.2, which means
the capacity at 7am will more likely to exceed 2000 veh/hr/lane. Therefore, designer’s
attitude towards risk will be positive. In terms of the value, it can be decided depends
on the probability. For example, at station 4000025, the mean of capacity at 5pm is
1999 veh/hr/lane, and the standard deviation is 305 veh/hr/lane. Figure 5.2(e) implies
that the probability that capacity less than 1999veh/hr/lane is 0.35. Therefore, γ should
represent a risk seeking strategy by taking a positive value or an interval leans to risk
seeking, such as [−1, 2]. Similarly, the weighting parameter γ at 8am, 4pm, and 5pm can
be determined.

5.3. Simulation Study

In order to discover how stochastic capacity would affect traffic operation, a ramp
metering study was designed and evaluated numerically. A combined CTM and modified
FIGURE 5.2: Cumulative distribution function of capacity at rush hours (station 4000025)
stochastic capacity-constraint ZONE algorithm was proposed to capture the performance discrepancies between consider $1^{\text{st}}$ & $2^{\text{nd}}$ order (mean and standard deviation) information and $1^{\text{st}}$ order (mean) information only.

5.3.1 Simulation Setting

Figure 5.3 illustrates an on-ramp only metering scheme. The simulation is conducted with the following settings. The simulation environment is a 1 km long mainline with a 300 meters long on-ramp. The on-ramp is in the middle of the mainline. This simplification reduces the equation 5.1 into three parameters: $A$ is the upstream mainline volume, $M$ is the local access metered ramps, and $B$ is freeway capacity at downstream detector. As observed from empirical speed-density and flow-density plot at station 4000025 in Figure 5.4, critical density at this location ranges from 25 veh/km to 35 veh/km and speed under capacity ranges from 60 km/h to 80 km/h. This also indicates the capacity varies in an interval instead of a fixed value. In this case study, 5pm’s capacity information is used in the simulation and an interval of $[-1, 2]$ for $\gamma$ is adopted.

![Bottleneck Effect Occurs Here](image)

FIGURE 5.3: On-ramp only metering

A combined cell transmission model and modified capacity-constraint ZONE algorithm is designed to numerically evaluate the impacts of stochastic capacity on downstream bottleneck throughput. The algorithm is presented in Algorithm 1.
FIGURE 5.4: Empirical fundamental diagram (station 4000025)
Algorithm 1 Combined Cell Transmission Model and capacity-constraint ZONE algorithm

1: *Initialization*: Given the initial conditions (as mentioned above).

2: *Defining Functions*:

- **Supply**: represents number of vehicles that could be accommodated by next cell, \( \min(\text{cell capacity}, \text{available space in the cell}) \)

- **Demand**: represents number of vehicles that desire to enter into next cell, \( \min(\text{cell capacity}, \text{current traffic volume}) \)

3: *Cell Transmission*: For both ramp and mainline, \( \text{flux} = \min(\text{supply}, \text{demand}) \).

4: *Ramp Metering*: At merge location

- \( \text{mainline flux} = \min(\text{supply}, \text{demand}) \)

- \( \text{ramp flux} = \min(\text{minimum metering rate}, \text{supply} - \text{mainline flux}) \)

5: *Simulation Results*: For both ramp and mainline, vehicles in cell = initial flow + \( \text{(in flux} - \text{out flux)} \times \text{simulation time step} \).
5.3.2 Experiment Results

To demonstrate the impacts of stochastic capacity on ramp metering performance, density in each road segment is recorded, and throughputs are collected. As presented from Figure 5.5(a) 5.5(b), a sudden increase of density occurs at the 25th segment. This is caused by ramp merge behavior. As the simulation moves on, the density begins to change. The shock-wave can be captured in Figure 5.5(c) 5.5(d). The differences are also revealed. Under the stochastic capacity constraint system (both first and second order information included), traffic density tends to be lower than the deterministic one (first order information only). This is because when the traffic flow rate varies, capacity is affected. A deterministic capacity system would fail to adapt to traffic’s dynamic nature. Traffic congestion is thus resulted. The density change also indicates queue formation on mainline traffic. Figure 5.5 demonstrates that ramp metering has better performance under a stochastic capacity constraint system.

As ramp metering is designed to maximize the throughput of highway, throughput is deployed to further improve performance. In Figure 5.6, throughput is recorded along the 400 iterations. When both 1st & 2nd order information (variation) is considered, the throughput is larger than only 1st order information involved in general. Although there are cases in which stochastic capacity would lead to lower throughput, this is acceptable because pursing high benefits would definitely require taking additional risks.

In Figure 5.7, a Monte Carlo simulation with 100 runs is conducted to validate stochastic capacity’s superior performance. In each run, 400 iterations are collected and the mean is calculated. The throughput comparison clearly suggests that considering the 2nd order information, standard deviation in this case, is beneficial to traffic operation.

5.3.3 Discussion

System efficiency and robustness have always been the focus of designers. In the traffic operations field, robustness means a system’s capability to operate under different
FIGURE 5.5: Comparison of roadway segment density under different capacities

(a) Mainline traffic density under capacity with 1st order information only
(b) Mainline traffic density under capacity with 1st & 2nd order information
(c) Mainline traffic density under capacity with 1st order information only (heat map)
(d) Mainline traffic density under capacity with 1st & 2nd order information (heat map)
FIGURE 5.6: Throughput comparison between with 1\textsuperscript{st} order only and 1\textsuperscript{st} & 2\textsuperscript{nd} order information

FIGURE 5.7: Monte Carlo Throughput comparison
statuses. Traffic flow patterns change over time so it is critical that the traffic control scheme accommodate various situations. Advanced techniques allow us to acquire the transportation history data, while how to benefit from gained additional information is the focus of this research. The proposed mean-standard deviation formulation allows the system to run at its efficient condition, but also adds the robustness based on the weighting parameter $\gamma$. A trade-off between efficiency and robustness is thus achieved.

Capacity-constraint ramp metering highly relies on the selected capacity value. If a predetermined capacity is higher than the real capacity, the freeway would be congested by discharging more vehicles from the ramp; the freeway would be underutilized if a lower capacity is used. Knowing the probability that capacity will exceed the predetermined value will help engineers decide if a risk adverse or risk seeking strategy should be taken. The capability that $\gamma$ is able to adapt to both engineers’ judgment and empirical experience allows the proposed stochastic-capacity constraint ramp metering algorithm to provide superior performance over deterministic capacity. This is validated by using system’s second order information which better informs the decision-maker.

![FIGURE 5.8: Multiple ramps metering scenario](image)

In this case, only one on-ramp was considered in the simulation for a proof of concept, but the model can be generalized to multiple ramps metering case. As the Figure 5.8 shows, a segment of highway with multiple on-ramps is several on-ramps connect together. The $n^{th}$ zone’s throughput will be $n + 1^{th}$ zone’s input flow. Wang et al. (2014) suggested that there is dependency exists in different entrance-ramp flows and different mainline flows. Thus, in order to consider multiple on-ramps scenario, correlation between
ramp flows and mainline flows should be considered, it will be another topic for the future research.
6. APPLICATION II: HYBRID FORECASTING OVER MULTIPLE DAYS

An accurate prediction would help enhance the efficiency of freeway operation and management through the Intelligent Transportation Systems framework. With the accurate capacity prediction, traffic operation decision-makers can utilize the prediction results to deploy various traffic control strategies, such as ramp metering control, to increase operational efficiency of the existing transportation network. Advanced forecasting technique provides us a good alternative to pursue. However, challenges yet remains in modeling a reliable forecasting model. The crux of accurately forecasting a nonlinear system has not been clearly understood. Suggested by Chang et al. (2011), any kind of the single method can barely satisfy the forecasting accuracy needs. Therefore, novel algorithm and proper combination of methods are encouraged for the short-term forecasting. In this study, we proposed a hybrid Wavelet transform and dynamic neural time series to forecast multiple days’ capacity.

6.1. Descriptive Statistics of Empirical Data

The data used in this study were obtained from the basic freeway segments. The detector recorded the traffic information at 20 seconds intervals. The time interval for capacity estimation is important and impacts the resulting capacity values. Different time intervals are adopted in other researches, e.g, Brilon et al. Brilon et al. (2005) adopt a 5-min interval. The time interval suggested by Transportation Research Board (2000) for capacity observation is 15-min (Polus and Pollatschek, 2002). This research therefore uses 15-min interval to derive the daily capacity. In this case, data at detector 4000026 were utilized as shown in Figure 6.1. The descriptive statistics are presented in Figure 6.2.
FIGURE 6.1: Selected Study Site of Detector 4000026 from GA400 Northbound at Atlanta, Georgia (from Google map)
FIGURE 6.2: Descriptive Statistics of Capacity Data
6.2. Experiment framework

The wavelet-neural time series framework involves three steps:

1. Through a wavelet function, the original capacity series is decomposed into a smooth series (approximation) and a set of series with finer details;

2. The decomposed smooth series and details are used as forecasting inputs, which are trained by dynamic neural network and forecasted independently;

3. The extended smooth and detailed series are reconstructed to obtain the forecasting results;

4. The results are compared with the forecasting accuracy test dataset and the prediction from other models, which then uses performance measures to demonstrate the forecasting power.

In step 2, a data structure is designed to better calibrate the network. The previous 11 months are used as modeling set and the first week of December are kept as forecasting accuracy comparison set. In dynamic neural time series development process, 70% data used for training, 15% data for validation, and 15% data for testing.

- **Training set**: This set is used to adjust the weights on the neural network according to its error;

- **Validation set**: These are utilized to measure network generalization, and to halt training when generalization stops improving, namely, minimizing the overall fitting.

- **Testing set**: This data set provides an independent measure of network performance during and after a training, testing the final solution to confirm the forecasting power of the network.
6.3. Results Analysis

In this section, we present a case study from a detector 4000062 on GA400. The stochastic feature of the capacity was investigated through time series and wavelet transform. We then proposed a hybrid model using the decompositions generated from wavelet analysis as the input of neural time series to forecast a 5-day capacity. The results will be compared against with other hybrid models and single models, and scientific measures of performance are employed to demonstrate the capability of different forecasting techniques.

6.3.1 Performance measure

In order to quantify the forecasting performance, Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) are adopted for each method to evaluate their accuracy.

\[
MAE = \frac{1}{N} \sum_{t=1}^{N} |x_t - \hat{x}_t| \quad (6.1)
\]

\[
MAPE = \frac{1}{N} \sum_{t=1}^{N} \left| \frac{x_t - \hat{x}_t}{x_t} \right| \quad (6.2)
\]

\[
RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^{N} (x_t - \hat{x}_t)^2} \quad (6.3)
\]

where \(x_t\) is the empirical value at the \(t^{th}\) time interval, \(\hat{x}_t\) is the forecasted value at the \(t^{th}\) time interval, and \(N\) is the number of forecasting steps.

6.3.2 Multiple days’ capacity forecasting

In Figure 6.3, a wavelet transform was conducted. \(a_3\) represents the approximation and \(d_1, d_2, d_3\) are the details. As observed in the figure, they are in different frequency. Take sound track as an example, \(a_3\) would be the main stream of the sound track, and \(d_1, d_2, d_3\) would be the detailed information. Different frequencies are in charge of different
tunes. The decomposed components are all served as the input of the proposed forecasting model. In dynamic neural time series forecasting model, the Bayesian Regularization training algorithm is chosen because of its superior performance in generalization for difficult, small or noisy datasets. The results are presented in Figure 6.4, 6.5, 6.6, 6.7.

The forecasted results are compared against other methods and the numerical comparison of three performance measures are presented in Table 6.1. The results indicate the proposed hybrid Wavelet Dynamic Neural Time series model is superior to single models (ARIMA).

According to the Table 6.2, the proposed hybrid wavelet-dynamic neural time series model yields the lowest MAPE of 3.07% among all three models. The single model dynamic neural time series has better performance than ARIMA. It indicates the fact that learning scheme also improves the prediction accuracy. In addition, as shown in
FIGURE 6.4: Neural Network of Approximation A3.

FIGURE 6.5: Details of Neural Network: D1
FIGURE 6.6: Details of Neural Network: D2

FIGURE 6.7: Details of Neural Network: D3
TABLE 6.1: Forecasting results comparison for a 5-day Horizon

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empirical</td>
<td>2308</td>
<td>2227</td>
<td>2202</td>
<td>2262</td>
<td>2030</td>
</tr>
<tr>
<td>ARIMA</td>
<td>2181</td>
<td>2184</td>
<td>2176</td>
<td>2176</td>
<td>2175</td>
</tr>
<tr>
<td>Dynamic Neural Time Series</td>
<td>2186</td>
<td>2187</td>
<td>2186</td>
<td>2186</td>
<td>2186</td>
</tr>
<tr>
<td>Wavelet + Dynamic Neural Time Series</td>
<td>2358</td>
<td>2272</td>
<td>2261</td>
<td>2169</td>
<td>2119</td>
</tr>
</tbody>
</table>

Figure 6.8, only wavelet-dynamic neural time series captured the change in the capacity. Single models, such as ARIMA and dynamic neural time series, fail to reflect the variation.

FIGURE 6.8: Prediction at Individual Steps

The experiment conducted above proved hybrid model has a better forecasting performance than single models. But what factor leads to its superior performance is not explicitly explored. In the experiment design, dynamic neural time series is applied to both original capacity series and decomposed capacity. The forecasting results showed distinct difference. It is safe to assert that wavelet decomposition helps improving the forecasting accuracy. This could be explained by superposition theory. Similar to a signal, capacity
series is also composed by sub-series with different frequencies. When conducting the forecasting experiment with the combined series, the network is trained as they inherit the same frequency while they are not instead. By decomposing the capacity series as shown in Figure 6.3, different frequencies are revealed. Predicting the different sequences separately allows the neural network includes their inherent properties and thus incorporates the detailed stochastic information in capacity series. This is the main reason that resulted in the deteriorated performance of single models.

In addition, there are lots of decomposition methods developed other than wavelet transformation, such as Hilbert-Huang transformation (HHT). The main feature of HHT is the empirical mode decomposition (EMD). The main difference between HHT and wavelet transform is that “the elementary wavelet function is derived from the signal itself and is adaptive” (Ding et al., 2007). They all can be used in decomposing the non-linear signals but hold different advantages. EMD method does not require to choose wavelet basis and decomposition scales, but it is time consuming, while the wavelet decomposition has high speed of decomposition. They are all proved to have a high forecasting accuracy performance (Long et al., 2011).

TABLE 6.2: Performance Measure for a 5-day Forecasting Horizon

<table>
<thead>
<tr>
<th>Model</th>
<th>Model Type</th>
<th>MAE</th>
<th>MAPE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Model</td>
<td>ARIMA</td>
<td>85.4</td>
<td>3.91%</td>
<td>97.03</td>
</tr>
<tr>
<td></td>
<td>Dynamic Neural Time Series</td>
<td>82.13</td>
<td>3.78%</td>
<td>96.86</td>
</tr>
<tr>
<td>Hybrid Model</td>
<td>Wavelet + Dynamic Neural Time Series</td>
<td>67.18</td>
<td>3.07%</td>
<td>70.13</td>
</tr>
</tbody>
</table>

series is also composed by sub-series with different frequencies. When conducting the forecasting experiment with the combined series, the network is trained as they inherit the same frequency while they are not instead. By decomposing the capacity series as shown in Figure 6.3, different frequencies are revealed. Predicting the different sequences separately allows the neural network includes their inherent properties and thus incorporates the detailed stochastic information in capacity series. This is the main reason that resulted in the deteriorated performance of single models.

In addition, there are lots of decomposition methods developed other than wavelet transformation, such as Hilbert-Huang transformation (HHT). The main feature of HHT is the empirical mode decomposition (EMD). The main difference between HHT and wavelet transform is that “the elementary wavelet function is derived from the signal itself and is adaptive” (Ding et al., 2007). They all can be used in decomposing the non-linear signals but hold different advantages. EMD method does not require to choose wavelet basis and decomposition scales, but it is time consuming, while the wavelet decomposition has high speed of decomposition. They are all proved to have a high forecasting accuracy performance (Long et al., 2011).
7. CONCLUSION AND FUTURE REMARKS

Stochastic capacity as a concept has been raised for decades but the causes of stochasticity and the implications of the stochastic characterization have not been explicitly explored. This thesis presents a stochastic characterization of highway capacity through studying the inherent spatial-temporal correlation in daily capacities and variations presented in estimated hourly capacities. The numerical results show that the spatial-temporal variations in capacity can be captured through a seasonal STARIMA model \((0,0,23) \times (0,1,0)_2\). This presents a strong spatial correlation between detectors within three distance degrees, and temporal correlation in two steps of time. The hourly capacity sequence further demonstrates the capacity’s stochastic nature.

In addition, this research proposed a stochastic capacity-constraint ZONE ramp metering embedded CTM algorithm. The capacity is formulated into a mean-standard deviation form to included both first and second order information, and achieves a balance between system efficiency and robustness. The weighting parameter is based on the cumulative capacity probability and decision-maker’s attitude towards the risk. A ramp metering case study is conducted, and the numerical results show that when additional information is included, congestion will be alleviated and throughput is increased. The Monte Carlo simulation further validates the significance of deploying the second order information in a stochastic system.

Since roadway capacity is essential to traffic control such as ramp metering design, signal control, and traffic delay, knowing the capacity in advance will help engineers enhance transportation system’s efficiency and reliability. Prior research on stochasticity of highway capacity concentrates on what distribution the empirical capacity follows, for example Weibull or Gaussian distribution without investigating the temporal feature underlying it. The year of observations considered in this thesis allows us to examine the
capacity from a time series perspective. This thesis further proposed a hybrid forecasting model: wavelet-dynamic neural time series and conducted a capacity forecasting study through a case study using empirical data collected from Georgia State Route 400.

The forecasting results show that the hybrid model: wavelet-dynamic neural time series possesses superior forecasting performance with the MAPE of 3.07% compare to ARIMA’s 2.91% and dynamic neural time series’ 3.91%. The experiment design reveals that the decomposition of capacity series enhances the forecasting accuracy. This is because the details of capacity series inherit different frequencies. Training the network separately allows more accurate stochastic information to be incorporated. Comparing the forecasting results of ARIMA and dynamic neural time series, it suggests that a learning technique also improves the forecasting performance. It presented a leading effort on capacity forecasting and an analysis on the factors that lead to the high accuracy performance. It would be interesting to investigate additional forecasting models to further enhance the conclusion. Different decomposition methods could also be explored.
BIBLIOGRAPHY


