HyTran: A New Approach for the Combination of Macroscopic and Microscopic Traffic Flow Models

DISSERTATION

submitted to

University of Technology Graz,
Faculty of Civil Engineering

for the degree of
Doktor der technischen Wissenschaften (Dr. techn.)

presented by

Weinan Huang (M.Sc.)

born February 6th, 1984 in P.R. China

Supervisor: Univ.-Prof. Dr.-Ing Martin Fellendorf, University of Technology Graz
Reviewer: Prof. Benjamin Heydecker, University College London

Oral examination: March 26, 2013

Acknowledgements

I would like to thank my supervisor, Prof. Martin Fellendorf, for his continuous support, guidance, and encouragement during my graduate studies at Graz University of Technology. His knowledge and experiences of traffic engineering and simulation inspires my work.

I would like to thank Prof. Benjamin Heydecker from University College London for reviewing my dissertation and providing helpful comments.

I would like to thank Prof. Li Keping and Prof. Sun Jian, both from Tongji University, for their kindly help. I would like to thank Dr. Tobias Pohlmann for sharing his experiences of traffic modeling.

I would like to thank all my colleagues in ISV. You offered so much friendly help that made my living in Graz easier.

I would like to thank my friends, who are living in Graz, Vienna, Stuttgart, Istanbul, Singapore, Beijing, and Nagoya. Thank you so much for being great listeners and encouraging me through all the difficulties.

Above all, my deepest thanks to my parents and my fiancée, thank you for always being by my side.

Graz, 6th, Feb. 2013

Huang Weinan

Abstract

The traffic flow modeling is one of the most challenging domains of traffic engineering research. This topic provides a basic approach for modeling and simulating the behaviors of traffic flows in traffic networks. Based on the simulation tools, a computer-based experimental environment can be established. Furthermore, various analyses, such as network evaluation, signal plan design and evaluation, online traffic state estimation, etc., can be implemented.

In terms of their levels of resolution, the existing traffic flow simulation models can be classified into three categories: macroscopic traffic flow simulation models, microscopic traffic flow simulation models, and mesoscopic traffic flow simulation models. The three classes of models have their own advantages and disadvantages, which determine their fields of application.

The objective of this thesis is to propose an approach for combining the macroscopic traffic flow simulation model and the microscopic traffic flow simulation model. This topic is known as "Hybrid traffic simulation model", and can be considered as an application of Multiple-Resolution Modeling (MRM) in the domain of traffic flow modeling. The main purpose of this approach is to make the two classes of models able to communicate and cooperate under the developed framework. Therefore, a single modeling case is able to utilize the advantages of those two classes of models by simulating different parts of the road network with different models. This approach will improve the simulation tool's feasibility and flexibility.

The consistency maintenance is the essential problem that has to be considered for MRM. In this thesis, the developed integration framework, which is named HyTran, maintains the consistency of the two integrated models. HyTran consists of two parallel frameworks: the Flow Transition based Framework (FTF) and the Data Fusion based Framework (DFF). FTF focuses on transforming the traffic flows between flow representation (macroscopic model) and vehicle representation (microscopic model). Under FTF, different parts of a road network can be modeled by either the macroscopic model or the microscopic model. FTF makes traffic flows able to move across the boundaries between those two models. DFF proposes a different way to implement the model integration. Under DFF, the macroscopic model is used to model the entire road network. In addition, some parts of the network can also be modeled by the microscopic model independently. DFF runs the microscopic model as probing runs; the data obtained from the probing runs is used as a reference for calibrating the macroscopic model. By this method, the accuracy of the macroscopic model is improved. In other words, DFF tries to abstract attributes of a microscopic network and integrate them into the corresponding parts of the macroscopic model. Through this data fusion process, the advanced performance consistency is achieved.

This thesis implements HyTran with Cell Transmission Model and VISSIM (a car-following based microscopic model). In order to improve the performance of FTF, several enhancements on CTM are discussed. The related algorithms, including flow transition algorithms (for FTF) and calibration algorithms (for DFF) are discussed.

Finally, this thesis tests HyTran with simulation examples. The examples show the feasibility and effectiveness of the proposed approach, and the performance of FTF and DFF is compared.

Table of Contents

List of figures

Table of contents

List of tables

List of algorithms

Abbreviations

ACM	Advanced Consistency Maintenance
BPF	Bootstrap Particle Filter
COM	Component Object Model
CTM	Cell Transmission Model
DFF	Data Fusion based Framework
FD	Fundamental Diagram
FTF	Flow Transition based Framework
KWM	Kinematic-Wave Model
MRM	Multi-Resolution Modeling
RMSNE	Root Mean Square Normalized Error
SPSA	Simultaneous Perturbation Stochastic Approximation

1 Introduction

1.1 Motivation

Traffic flow modeling is one of the most challenging domains of traffic engineering research. This research offers a basic approach for describing and modeling the behaviors of traffic flows on traffic networks. Based on those methods, a computer-based experimental environment can be established. Furthermore, various analyses, such as network evaluation, signal control planning design, online traffic state estimation, can be implemented.

To develop more effective models and software tools for traffic system analysis is the perpetual goal for model researchers and commercial companies who offer computer-aided design software to traffic engineers. The first stage of model development focuses on mathematical modeling of traffic flow behavior. One main source of traffic flow models is the established theories for conservation laws, which physicists research and describe idealized flows. In this stage, traffic models describe traffic flow on a rough way, without considering the stochastic attributes and differences in traffic participants. The second stage of model development focuses the individual behavior of traffic participants. The interactions between vehicles are described on one-dimensional spaces. In addition, the pedestrian simulation, existing as an independent topic, opens a door to simulating traffic participants on two-dimensional spaces. Currently, more models extend the one-dimensional vehicle simulation model to two-dimensional one, and are able to be integrated with pedestrian simulation. Another branch of flow modeling research is to make middle leveled model in order to balance the computational complexity and analytical ability.

The applications of traffic flow models are substantial. It works as the foundation of diverse topics in traffic engineering. Traffic design needs traffic flow models to evaluate the strategies, e.g., road channeling, signal control, and ramp metering. Traffic management needs traffic flow models to do the short time state estimation, in order to obtain the real time traffic information. Another hotspot of traffic engineering research, traffic network modeling, which stems from the traditional four-step method for demand estimation, needs travel cost on each modeled route, which is a typical output of traffic flow models. Although network model tends to use a highly simplified method to calculate travel cost and route choice behavior mainly because of the computational complexity of algorithms for route generation, assignment, and so on, are already very high. There is some research on making traffic assignment based on traffic flow models.

The substantial applications of traffic flow modeling methods are one of the main motivations for this thesis. In addition, the idea of Multiple-Resolution Modeling (MRM) and the preceding research on the topic of hybrid traffic simulation inspires this thesis.

Within the domain of system modeling, resolution means "the level of detail at which system components and their behaviors are depicted" (Davis & Bigelow, 1998). Generally, a model has its specific resolution. For example, in a battle field simulation, the basic simulated element can be a single aircraft, and also can be a group of aircrafts that fight as a formation. With different resolution, modeling mechanisms and model performance differs. The modeling resolution has to be determined according to diverse conditions. Compared with a single-resolution modeling, MRM not only describes realistic world more naturally, but also avoids a series of problems that exist in the application of single-resolution modeling. As a new branch of traffic flow modeling methods, hybrid traffic flow modeling improves the feasibility of modeling method by combining models on different

levels of resolution. The introduction of MRM into the domain of traffic flow modeling is possible to solve several basic problems existing in traditional traffic flow modeling approaches.

1.2   The hybrid traffic flow modeling

In traffic flow modeling, the concept of modeling resolution is commonly mentioned as "scope", which refers to the extent of the modeled objects, input details, and output range. According to their levels of resolution, traffic flow models are categorized as three classes: macroscopic models, mesoscopic models, and microscopic models. There will be a basic introduction to those three groups of models presented in section 2.2. In the current section, we only talk about the abstract perspective; therefore we only refer to the models as high resolution models and low resolution models.

In general, a high resolution model describes the simulated objective in more detail. Therefore, users always expect more dependable and accurate simulation results from them than using low resolution models. However, at the same time, the thorough description leads to at least two disadvantages: first, high demand for computational resource; second, high requirements for model building and calibration. Generally speaking, a high resolution model are mathematically more complex than a low resolution one, this means its numerical calculations need more computational resource; in addition, because high resolution models consider more details, broader input information is necessary for model building. Furthermore, because high resolution models describe more detail behaviors of traffic participants, they are more sensitive to change of model parameters and scenarios, it is more difficult to calibrate and validate a microscopic model. Compared with high resolution model, a low resolution model needs neither computational resource as much as a high resolution model needs, nor the input information as detail as a high resolution model needs. Moreover, the calibration work of a low resolution model is not as difficult as that of a high resolution model. Consequently, it cannot offer strong predictability as the high resolution model does. The simulated results are not sensitive to the change of model structure or values of parameters. This means a lack of analytical ability.

Traditionally, models with a specific level of resolution are appropriate for particular purposes. It has to be determined that models on which level of resolution are chosen. However, there is a dilemma about model selection bothering users during applications. A single-resolution model allows only one resolution in a model, no matter it is sufficient or not, no matter it is necessary or not. The main consequence of this attribute is that users always have to choose models according to the most complicated part. A typical example of this dilemma is: when an urban area with several traffic intersections is modeled. If one of those intersections is the one that is complicated and needs careful analysis (named as critical intersection), and the rest of them are stable and concise (named as minor intersections), which type of model should be applied to this example? If a high resolution model is used, a lot of efforts have to be paid to model and calibrate those minor intersections, but without improving the analytical ability of the built model of the critical intersection. However, if a low resolution model is used, it is probably that the developed model is not sensitive enough to the analyzed critical intersection, which means it is not adequate for the desired analysis.

There is cost of the implementing a simulation model. The cost comes from two perspectives: 1) model building, including model development and model calibration; 2) model running, referring to the computational cost for numerical calculating. Essentially, the dilemma of model selection can be

discussed as the relations among model building complexity, computational complexity, and analytical ability.

When applying a simulation project, third conditions have to be met: first, developed models provide sufficient analytical ability that leads to valuable simulation results; second, the input information and data are available for model building and calibration; third, the built model is able to run on a reasonable computer system, which is commonly a single Private Computer. Based on this statement, the essence of model selection dilemma is a question of how to maintain two types of balance: 1) the balance between model building complexity and analytical ability; 2) the balance between computational complexity and analytical ability. In summary, we can also refer them as the balance between implementation cost and analytical ability. Generally speaking, the computational and building complexity increases with the improvement of the model accuracy; in contrast, the computational and building complexity decreases with the deterioration of analytical ability. Traditionally, model developers should consider the modeled objects and choose a model with suitable resolution. For example, a city-wide traffic demand analysis may be performed by a macroscopic model; design of traffic signal control plan may be performed by a microscopic model; a freeway system analysis may be performed by a mesoscopic model. However, it is highly possible that, within a single model, different parts have different requirements for analytical ability; at the same time, different modeled objects may have different requirements for analytical. For example, a macroscopic model may be accurate enough for evaluating freeway system, but may be not sufficient for evaluating a signalized intersection.

One solution to satisfy the requirements for analytical ability is to adopt high resolution model as often as possible. However, its high computational complexity make the simulation runs of a large network a practical problem. In addition, the large amount of data collection and calibration also block this solution practically and make it waste of resource.

Development of mesoscopic model is a possible way to balance implementation cost and analytical ability. However, an awkward situation brought by mesoscopic model is the analytical ability is excess for some parts but insufficient for other parts.

Hybrid traffic flow modeling is a promising approach to utilizing the advantages of high resolution model and low resolution model at the same time. It keeps the two types of balance from a more nature way. The basic idea of this approach is to design a framework under which different models work cooperatively. Specifically, hybrid traffic flow modeling simulates the entire network by a low resolution model; meanwhile, some specific sub-networks, such as a signal controlled intersection or a weaving segment, whose performance is difficult to be described by the low resolution model, are simulated by a high resolution model, in order to model the response of traffic flows to the subtle circumstances. Since models with different levels of resolution are allowed in the same modeling case concurrently, users can choose suitable modeling resolution for each part of the case and avoid the dilemma of resolution selection.

From some perspective, the hybrid modeling may be not only optional but necessary. This means the mixture of different models is unavoidable, considering the realistic road network existing in an urban area. For example, freeway systems (commonly modeled by macroscopic model) connect directly to urban main roads (commonly modeled by microscopic model). This is a fair common situation that happens in metro cities. In addition, the evaluation of emerging ITS related management methods needs extension of analytical ability even for the road types which are traditionally modeled by low resolution models. One good example is the ramp metering control for

freeway system. Without this management strategy, a macroscopic model describes the freeway network well enough. However, if those new management strategies have to be considered, a macroscopic model does not describe system's response to different ramp metering control plans. At the current time, this case needs a microscopic model for the entire object. In summary, a single resolution model cannot meet the growing requirements for modeling tools; therefore a multi-resolution model is becoming necessary.

1.3 Problem statement

Although various methods for describing the attributes of traffic flows are established, traffic flow models derive from two aspects: the individual aspect and the flow aspect. With the first aspect, individual vehicles (or other traffic participants) are modeled. In other words, those models try to describe the behavior of every single vehicle loaded in the simulated network. Therefore, the system's overall performance can be shown by the aggregated behavior of all these individual vehicles. With the flow aspect, traffic flows are described as continuous compressible flow. They follow the dynamic attributes of conservation flows that have been modified for traffic flow. The behaviors of those physical flows are used to estimate the performance of a traffic system. In summary, no matter which aspects are employed, traffic flow models are used to describe the mobility of vehicles (or, traffic participants) in road networks. Various variables are defined for describing those attributes, e.g., travel speed, vehicle density, occupancy, flow rate, and so on. Last but not least, the essence of those two aspects are corresponding to the concepts of levels of resolution, which is commonly referred in the domain of traffic modeling as macroscopic model, mesoscopic model, and microscopic model.

Since hybrid traffic simulation systems combine traffic flow models on various levels of resolution, there are at least two tasks for building it: 1) to choose the models that are integrated, and 2) to design a framework that integrates them. In order to perform those tasks, the related scientific and technological problems have to be solved. Those problems come from two perspectives based on the MRM consideration and the traffic simulation consideration.

From the MRM consideration, two problems can be summarized as following:

- How to select the basic integrated models. MRM approach is always based on the fully developed single resolution model. Therefore, to build a MRM approach should start with model selection. The determination is influenced by the attributes of the chosen models and the target system. There is a dilemma within the process: if we do not choose the models, we cannot determine the integration framework; otherwise, if we do not determine the integration framework, the requirements for the models are not clear enough. Fortunately, the model can be chosen based on some practical considerations, such as the developments progress, ease of use, personal experiences of the researcher, etc. One additional point is that most of the traffic flow models obey the basic principles of traffic flow, such as continuum flow, and basic relations among volume, speed, and density, although only to a certain extent. In addition, some commercial software packages are supposed to be

employed. However, the risk still exists that if the models are not chosen properly, we will face too many difficulties within the process of framework design.

- How to maintain the consistency in the developed system. Consistency maintenance is the well-known and key problem for any MRM research. In order to make two or more independent models cooperate effectively, the difference of those models, in terms of background theories, applications, and attributes, etc., have to be considered carefully. Therefore, a suitable mechanism has to be developed to resolve them. Through those mechanisms, the integrated models share the information smoothly, and one model responses well to another one when it is necessary.

From the traffic engineering perspective, other three problems can be summarized as following:

- Applicable scope for road network. Within a process of developing traffic simulation system, researchers must consider which kind of road network will be modeled by the developed system. In addition, a further question is what scenarios can be modeled by the developed mechanisms. Those considerations in terms of application influence the research significantly.

- Usability for traffic analysis. The most notable accomplishment of the developed framework should be the capability of implementing with traffic modeling. In other words, the developed framework can be used to model a realistic road network and simulate the traffic behaviors on the network, and then, obtain useful results. The developed modeling approaches determine the simulation performance. This approach should be capable of handling typical situations that exist in reality, and should be flexible.

- Model development work flow. When applying a simulation system to a realistic modeling project, users should be able to implement the model step by step smoothly. These steps and needed resource for model building should be intuitive and logical. We should always keep a mind on how to using the data and human resource efficiently and flexibly, try to avoid depending on decisions that are hard to be made by users.

1.4 Objective

The main objective of this thesis is to discuss and propose an approach for integrating a low resolution model (macroscopic model) and a high resolution model (microscopic model) in a single system, by designing an integration framework named "HyTran". HyTran focuses not only on the traditional idea of flow representation transition, but also the data fusion based idea that investigates the difference in dynamic behavior of macroscopic model and microscopic model.

1.5 Contributions

The main contribution of this thesis can be summarized as "has investigated the approach of MRM in the domain of traffic flow modeling". This contribution consists of three parts:

- An approach for the combination of macroscopic and microscopic traffic flow simulation models, that is named "HyTran", is proposed.

  - The Flow Transition based Framework (FTF). Algorithms for flow transformation are developed. In addition, two components are developed for improving the performance of flow transition: method for Mandatory Lane Change (MLC) and IFCTM.

  - The Data Fusion based Framework (DFF). The new approach for model integration, DFF (Data Fusion based Framework), is investigated. In addition, the difference, advantages, and disadvantages are explored, compared with FTF. From the perspective of MRM, DFF is a new approach for consistency maintenance that achieves the advanced consistency.

- The possibility of combine a macroscopic and a microscopic traffic flow is investigated, based on examples of CTM (Cell Transmission Model, a macroscopic traffic flow model) and VISSIM (a simulation system whose kernel is a car-following based microscopic traffic flow model). This thesis discusses the detail stuff of integration according to the attributes of CTM and VISSIM, which has not been reported before.

1.6 Limitations

HyTran tends to provide an effective framework for solving the problem of model integration, however, due to the complexity of this topic, HyTran depends on various assumptions and simplifications, which may limit the effectiveness of it and needs more considerations beyond the contents included in the current thesis. Those limitations come from two categories:

From the theoretical perspective:

- The effectiveness of Data Fusion based Framework (DFF) is discussable if consider the complex scenarios existing in the realistic world. Therefore, DFF may only be applied to several researched scenarios.

- Although paying attention to general problems as far as possible, HyTran highly depends on those two used model, CTM and VISSIM. This is a nature attribute of multiple resolution modeling. Therefore, the effectiveness of using other models may also be limited.

- The efficiency of calibration algorithms for data fusion is probable to be improved.

From the practical perspective:

- The code efficiency. First, the implementation of HyTran is based on the COM interface of VISSIM, which limits the methods which can be used. In addition, calling COM function also reduce the efficiency of the developed code. Third, we do not focus on optimize the code for computational speed or resource, because the main task is to discuss the feasibility of HyTran.

- The codes are implemented with VISSIM and CTM. It cannot guarantee there is no practical problem when using HyTran with other models or software packages.

1.7 Concept clarification

There are several concepts used in the field of traffic modeling. They are widely discussed, but in different publications or reports, slightly different definitions or scope are used, which also bring misunderstanding during communication. Therefore, we set a specific section for clarifying several concepts that tend to bring various interpretations. Some of the contents included in this section will repeat in other chapters.

Objective of macroscopic traffic flow model

Traffic network modeling is one of the most attractive research topics at the current time. The main objective of this research is to describe the situation of traffic assignment in a given road network, considering the route choice behavior of traffic participants. Because of the traffic assignment is also on macroscopic scope, the network modeling is also widely known as "Macroscopic traffic simulation", and is confused with macroscopic traffic flow modeling.

In this thesis, we talk about the difference between macroscopic traffic flow model and microscopic traffic flow model. Therefore, the macroscopic traffic flow model, not including traffic assignment, is discussed. In addition, it is also mentioned as macroscopic model for short. It will be specifically explained if the network modeling related macroscopic model are discussed in some contents.

Macroscopic flow model and macroscopic flow simulation model

Macroscopic flow model is developed to describe the macroscopic performance of a traffic network. As summarized by Williams (1992), there are three categories of this model: travel time models, general network models, and two-fluid models. Generally speaking, macroscopic flow models tries to investigate the relation between macroscopic network variables, e.g. flow rates, density (concentration), speeds, capacity, travel time, etc.. For example, travel time models output travel time contour maps, which describing how much travel time (on average) is needed from the center of a network (such as a central business district) to some other parts of the network with specific distance away from the center. These models provide an overview of average travel time between specific points within a modeled network. Two-fluid model classifies vehicles in urban road networks into two categories (two fluid, moving and stopped), based on several assumptions on traffic flow, this model builds a math relation between stop time per unit distance and trip time, using average minimum trip time per unit distance as a calibrated parameter.

The name of "macroscopic flow model" always makes it confused with "macroscopic flow simulation model". Although those two concepts share the same key word "macroscopic flow", they do not have relation with each other as close as it looks literally. The foundation of macroscopic flow simulation model is not macroscopic flow model described in the last paragraph, but another group

of models which are referred as "continuum flow model". These models will be briefly reviewed in section 2.2. However, the continuum flow models are often referred as macroscopic traffic flow model. This term is reasonable if we consider the resolution of it.

In summary, in this thesis, the term "macroscopic traffic flow model" refers to the models which describe the traffic flow dynamic behavior within road networks, precisely speaking, the continuum flow model.

Model calibration

Model calibration is a well-known process during model applications, which must be involved any simulation based analysis.

One of the widely acceptable definition of calibration can be found in (FHWA, 2004), which says: "Calibration: process where the analyst selects the model parameters that cause the model to best reproduce field-measured local traffic operations conditions". This definition implies that the process's target is to improve the final performance of a model, but not the precision of parameter values. If we suppose that the model mechanism is able to describe anything and any details that happen in the realistic world, it is logical to suppose that more accurate values of parameters lead to better model performance. If this hypothesis is valid, users only need to find the accurate value of each parameter separately. However, this hypothesis is hardly to be valid, because all models come with simplification and idealization, and are only a rough approximation of the real world. This means a better value of a parameter in terms of its definition does not guarantee better general performance. This is the reason why recent research about model calibration pays more attention to target performance of a developed model (measured by some predefined indicators), but not precise values of parameters according to their definition. Although this situation is true, it does not mean to measure values of parameters by their physical definition is not an acceptable solution for model calibration. In fact, to measure the values of the parameters is an effective way to obtain the desired parameters, and it is widely used in the research of model calibration, especially the early stage of a new designed model. For example, after the theoretical description of CTM has been proposed in the middle of 90's of the last century, some validation research based on measuring its parameters according to field data provides persuasive evidence of the feasibility of CTM.

In the current thesis, "calibration" refers to the process of adjusting values of parameters in order to obtain better performance evaluated by the predefined indicators, unless it is specifically clarified. For example, the calibration of CTM discussed in this thesis is to adjust values of parameters (e.g. free flow speed, jam density, etc.) in order to obtain better simulation results in terms of traffic volume and vehicle density.

Model validation

The term, model validation, is also a frequently used on in the domain of traffic simulation. However, model validation is misleading in some circumstance.

For a model inventor, the model validation refers to a process that proofs the "invented" models are theoretically correct (and practical to some extent). Most models are designed with specific goals, which can be described as what objects to be modeled, what features to be captured, etc. Therefore, after the design process, the inventor has to check whether anticipated functions work. This process is always carried out with some simplified examples.

For a model developer, who uses principles of a model to describe a realistic example, the model validation refers to a process that checks whether the calibrated parameters are good enough for

predicting the features of the modeled objects. Compared with calibration, the validation process should utilize the data groups that are not used for calibration, and the performance indicators are also changeable.

In this thesis, we tend to adopt the second definition of model validation.

1.8 Thesis outline

The remainder of this thesis is organized as follows.

Chapter 2 summarizes the related research from three perspectives, including Cell Transmission Model, hybrid traffic simulation research, and calibration methods for traffic simulation model.

Chapter 3 summarizes some basic information about the proposed approach, especially from the perspective of system requirements.

Chapter 4 discusses Cell Transmission Model for the hybrid application. Two extensions, including the slope-CTM and Temporal Fundamental Diagram (TFD), are proposed for improving the effectiveness of flow transition. In addition, the performance of CTM is discussed from the perspective of hybrid application.

Chapter 5 presents the main approach to model integration, which is named "HyTran"; HyTran has two parallel frameworks for model integration: the Flow Transition based Framework (FTF) and the Data Fusion based Framework (DFF). This chapter introduces the two frameworks and their background algorithms.

Chapter 6 evaluates HyTran. At first, this chapter evaluates the effectiveness of FTF and DFF separately, and then, a case study on urban intersection is implemented both with FTF and DFF to compare those two frameworks.

Chapter 7 summarizes the findings of this thesis and draws several conclusions, in addition, proposes directions for further research.

2 Literature review

2.1 Overview

This section summarizes the existing related literatures from four categories: traffic flow models, hybrid traffic simulation models / systems, Cell Transmission model (CTM) and its extensions, and the calibration methods for traffic simulation models, including network modeling and flow modeling.

The review of traffic flow models explains basic background about traffic flow modeling, based on models with single resolution. By this brief review, we tend to show a general situation of traffic flow modeling.

The review of existing hybrid traffic simulation classifies the related works into three categories: flow representation transition, system integration, and model integration. This section presents substantial examples and summarizes their methodologies.

Since CTM is selected as the macroscopic level of the hybrid system for reasons that will be explained in section 3.3, this chapter reviews the literatures about CTM. The summary is organized based on a problem-oriented approach, which means we put forward problems that we want to investigate and summarize how the related literatures handle them.

The reason for which the calibration method for traffic simulation model is reviewed is that the current thesis will propose a new approach for model combination, which focuses on using calibration algorithms as a data fusion interface between high-resolution level and low-resolution level. This strategy is referred as "information integration" in present thesis for the purpose of distinguishing it from the preceding research.

2.2 Traffic simulation model with single resolution

According to their levels of resolution, traffic flow models have been classified into three categories: macroscopic models, mesoscopic models, and microscopic models. The resolution of a model significantly affects its performance, furthermore, determines its application domains. One point worth pointing out is that traffic flow models and traffic simulation models are not identical. Generally speaking, traffic simulation models derive from the corresponding traffic flow models; however, not all the traffic flow models are suitable to work as a foundation of traffic simulation model, which is determined by its internal mechanisms.

Chapter 1 has introduced the idea of hybrid traffic simulation by talking about the advantages and disadvantages of specific models. This section makes a brief summary about the general situation of traffic flow models.

Macroscopic traffic flow model and macroscopic traffic simulation model

As discussed in section 1.7, macroscopic traffic flow models discussed here are continuum flow models, which are the essential foundation of macroscopic traffic simulation models.

The continuum flow model is an application of conservation law with traffic engineering. This model adopts three variables as the basic description of traffic state along any road: mean speed, flow rate, and concentration (or density). The relations between those three variables are used to describe the dynamic of traffic flows on road networks. The key feature is that macroscopic model only capture an aggregated performance of the traffic network, without paying attention to the individual traffic participants and detailed network attributes. One of the most widely used one is the kinematic wave

model developed by Lighthill, Whitham (1955) and Richards (1956). This model is also referred as LWR model. Associated with Riemann problem and its analytical solution (entropy solution) and numerical solution (Godunov's method), LWR model is the most fully research macroscopic traffic model. Several macroscopic simulation models, e.g. CTM, derive from LWR model. Although LWR captures several attributes of traffic flow, the limitations of it are widely discussed according the observed data, including (1) driver differences, (2) shock structure, (3) forward moving wave in queued up traffic, and (4) traffic instability (Zhang, et al., 1992). From this consideration, high order continuum flow models are presented. Those models tend to capture more realistic attributes of traffic flows than LWR model. However, because of their complexity, they are not commonly used as foundation for traffic simulation.

One of the main application domains of the macroscopic model is to work as the base of assignment model, other demand estimation. Macroscopic models are used for traffic state evaluations covering city or even country scale. In the model developing process, not only main roads are modeled, but also the minor roads, because those models should be capable of modeling the traffic assignment results. The output of these models is based on network-wide average values, and works as direct output or more often, the base of other processes.

Microscopic traffic flow model and microscopic traffic simulation model

Microscopic models depict the behaviors of individual vehicles, such as motor vehicles, bicycles, or pedestrians. Those models are born for traffic simulation, because they describe the individual behaviors of vehicles directly and they can show no indicators for traffic flow state without simulating amount of vehicles for a specific time period. The mostly discussed and used traffic flow model for microscopic simulation is the car-following models (a good summary of car-following models can be seen in (Rothery, 1992)). This model is developed for vehicle flows, and associated with a lane changing model; they are capable of describing vehicle flows in road networks effectively. Compared with macroscopic flow models, these models describe more diverse situations, such as signal control, road infrastructure, pedestrian crossing, permitted left-turn, etc. Besides modeling of vehicle flows, microscopic simulation models for pedestrian that describe behavior of crowds are a relative new branch, social force model (Helbing & Molnar, 1995) (Helbing, et al., 2000) and extension on cellular automation model (Nagel & Nelson, 2005) are discussed for this purpose.

When applying a microscopic simulation, based on the simulated individuals, their aggregated performance is measured to show the macroscopic attributes of the simulated objects. Because the vehicles are modeled in detail, the outputs of these models are more sensitive to the change of traffic scenarios. This is considered as the main reason for which microscopic models possess high analytical ability. Generally, microscopic models are used for strategy evaluation for relatively small area, e.g. a ramp-metering strategy, a signal control plan, etc. Users should be careful about the case scope.

Mesoscopic traffic simulation model

A mesoscopic model comes as a compromise between computational complexity and description ability. Its original concept comes from a combination of macroscopic simulation models and simulation microscopic models, in order to simplify the microscopic simulation model as well as keep its analytical ability as much as possible. In contrast to macroscopic simulation model, mesoscopic models are commonly developed from the perspective of traffic simulation. Although without direct origin in traffic flow models, mesoscopic models have been developed according to various theories and their combination. Mesoscopic models do not model individual traffic participants directly, but

recorded them. The dynamic of the model is obtained by implementing macroscopic principles. Although CTM is generally categorized as macroscopic simulation model, its extensions are more often classified into the mesoscopic group if they consider more details than the basic LWR model. Those extended CTM depict the mobility of vehicles according to continuous flow theory associated with the predefined foundational diagram. Another example is queue-server-based models. The queue-server-based models describe each link as a queue-server. The server capacity and queue situation are described by queue theory. These models can be considered as the combination of two-fluid models, continuum models, and practical considerations for traffic simulation. Messo (Burghout, 2004) is a typical queue-server-based mesoscopic traffic flow model. As mentioned before, mesoscopic models are born to balance the needs of computational complexity and analytical ability by using a middle leveled resolution. However, this strategy also means that these models neither have the high analytical ability as the microscopic one nor the low computational complexity as the macroscopic one.

2.3 Hybrid traffic simulation model

The application of MRM in the field of traffic simulation is commonly referred as hybrid traffic simulation, because this application generally covers two levels of resolution. This situation is determined by the attributes of traffic simulation. In this thesis, we also keep this term because it does capture the essence of this topic.

2.3.1 Flow representation transition

The conflicts among different traffic flow modeling methods come from the basic theories on that they are based and the description forms of them.

Macroscopic traffic flow model defines a "flow" through three basic physical variables: speed, density, and volume. Theoretically, a macroscopic model calculates the values of those variables for any point within the modeled network according to its own dynamic principles. Through this, the flow dynamic inside a network is described.

In contrast to the flow presentation, microscopic approaches model the traffic flow from a totally different aspect. Microscopic traffic flow models describe individual vehicles as basic elements. The behavior of vehicles is captured by modeling their response to circumstances. For example, car-following model points out that the speed of a vehicle is affected by its preceding vehicle. This model calculates the speed of individual vehicle by considering its relation with its preceding vehicle. When all the individual vehicles in a road network are simulated, their aggregated performance is an indicator of the traffic flow.

Bourrel (2002) (2003) categories traffic flow models as two classes: Flow Representation models (FR) and Vehicle Representation models (VR). Based on those categories, Bourrel proposes two requirements needed to be satisfied when developing a hybrid model: (1) flow conservation at the boundaries must be ensured; (2) information propagations at interfaces should be correct under both free flow condition and saturated condition.

This research investigates a discretized version of LWR model, Starda model (FR), and summarizes it as equation 2-1. The definition of Starda model is similar to CTM. This model also uses the concept of "cell" to discretize an entire link.

$$\begin{cases} \dfrac{\partial k}{\partial t} + \dfrac{\partial q}{\partial x} = 0 \\ \quad q = kv \\ \quad q = Q(k) \end{cases} \Rightarrow K_i^{t+\Delta t} = K_i^t + \dfrac{\Delta t}{\Delta x}\left(Q_{i-1}^t - Q_i^t\right) \qquad \textbf{2-1}$$

Where K_i^t is the density for cell i at time point t, Q_i^t is the traffic volume coming out of cell i and into cell $i+1$ during time t to $t + \Delta t$.

For the microscopic model (VR), this work cites the vehicle representation from (Newell, 1961): the optimal velocity model. This model describes a function (G) from the space between two adjacent vehicles to the speed (v) of the following vehicle (equation 2-2).

$$v_i(t) = G\big(x_{i-1}(t - \Delta t) - x_i(t - \Delta t)\big) \qquad \textbf{2-2}$$

Where, x_i is the location of the current vehicle, and x_{i-1} is the location of its preceding vehicle. The problem of flow transition at boundaries is discussed under two situations that are mentioned as upstream transition cell and downstream transition cell.

The upstream transition cell solves the problem of transition from flow representation model to vehicle representation model. Therefore, the task for the transition cell is to generate vehicles in the downstream cell based on the upstream demand flow. The demand from upstream cell can be expressed as $D(t) \cdot \Delta t$, where $D(t)$ is the demand flow rate calculated according to upstream cell state, and Δt is the time interval. A tentative generation times for vehicles during the time interval Δt can be calculated by using a uniformly distributed generation gap. In addition, the minimum space between two generated vehicles is $s_{min} = 1/k_{max}$, and the total number of vehicles that are expected to be generated in a single time step is $D(t) \cdot \Delta t$. The minimum space between two generated vehicles is used to calculate the supply ability of the downstream link; this is to say, the system measures the distance between the boundary and the latest generated vehicle, if s_{min} is not satisfied, the generation will be delayed.

The downstream transition cell solves the problem of transition from vehicle representation cell to flow representation cell. Therefore, the task for the transition cell is to determine an equilibrium flow rate based on the space headway of the upstream cell with vehicle representation. The transition cell calculates the minimum exit time gap (which is the inverse of maximum number of supplied vehicles) based on the density of itself. Afterward, the transition cell determines whether a vehicle (in upstream cell) is allowed to pass the boundary by comparing the minimum exit time gap with the time gap between the current vehicle and the last passed vehicle.

This research compares the simulated traffic volumes from hybrid model and single Strada, under three different situations: stationary situation, single state change situation, and multiple state change situation. This research draws a conclusion that the mechanism described above translates boundary conditions from one model to the other correctly under both congested and free conditions.

Leclercq (2007) presents a hybrid LWR model, and focuses on implementing an interface between the model boundaries (microscopic to Macroscopic and Macroscopic to microscopic), which ensures traffic waves can propagate effectively from macroscopic model to microscopic model, and inversely. In this research, not only the typical LWR model on macroscopic resolution is employed, but also a LWR based microscopic model is used. The macroscopic LWR can also be described by equation 2-1. The microscopic LWR model is described as a relation between concentration (or distance between adjacent vehicles) and speed, it can be shown in equation 2-3.

$$x_i^{t+\Delta t} = x_i^t + V_e^*(x_{i-1}^t - x_i^t)\Delta t \qquad \textbf{2-3}$$

Where, i is the index of a vehicle, x is the location of it, $V_e^*(s)$ is a function from s to average speed (v), where $s = x_{i-1}^t - x_i^t$ is the front-to-front distance of the adjacent vehicles. At the boundaries, the flow state of microscopic model can be transformed into flow rate, based on the basic relation $= kv = v/s$, which matches the description of traffic flow in macroscopic LWR model.

For the microscopic to macroscopic situation, the demand volume from upstream microscopic cell is:

$$D(x, t \to t + \Delta t) = \begin{cases} \dfrac{V_e^*(s_l^t)}{s_l^t} & if \ s_l^t \geq s_c \\ Q_c & if \ s_l^t < s_c \end{cases} \qquad \textbf{2-4}$$

Where, Q_c is the critical volume of a cell (cell capacity); K_c is the critical density; s_c is the critical spacing, which leads to $s_c = 1/K_c$; s_l^t is the front-to-front distance of vehicles at the downstream end of the microscopic link time t.

The supply volume from downstream macroscopic cell is:

$$S(x, t \to t + \Delta t) = \begin{cases} Q_c & if \ K \leq K_c \\ K \cdot V_e(K) & if \ K > K_c \end{cases} \qquad \textbf{2-5}$$

For the macroscopic to microscopic boundaries, the demand volume from upstream macroscopic cell is:

$$D(x, t \to t + \Delta t) = \begin{cases} K \cdot V_e(k_l) & if \ K \leq K_c \\ Q_c & if \ K > K_c \end{cases} \qquad \textbf{2-6}$$

The supply volume from downstream microscopic cell is:

$$S(x, t \to t + \Delta t) = \begin{cases} Q_c & if \ s_r \geq s_c \\ \dfrac{V_e^*(s_r^t)}{s_r^t} & if \ s_r < s_c \end{cases} \qquad \textbf{2-7}$$

Where, s_r is the vehicle distance at the downstream end of the microscopic link.

Therefore, according to equation 2-4, 2-5, 2-6, and 2-7, the traffic flow can be described as equation 2-8 according the relation of demand and supply.

$$Q(x, t \to t + \Delta t) = min\big(D(x, t \to t + \Delta t), S(x, t \to t + \Delta t)\big) \qquad \textbf{2-8}$$

This research looks into four different situations to evaluate the proposed approach, including 1) supply variations just downstream of the microscopic to macroscopic interface; 2) a shock wave coming from the downstream macroscopic link; 3) demand variations just upstream of the macroscopic to microscopic interface; 4) a shockwave coming from a downstream microscopic link. The density profile (for macroscopic link) and vehicle trajectories (for microscopic link) are used to show the simulation results of the hybrid system. The paper draws a conclusion that the developed interface ensures correct traffic wave propagation, and the interface does not induce unexpected delays, wave speed modification or oscillations.

Magne (2000) presents an approach to combining a microscopic simulator SITRA-B (using a car-following-based microscopic model "Helly") and SIMRES (based on typical macroscopic model). The hybrid system is named MICMAC. Before talking about the combination approach, this paper discusses the compatibility between these selected models. It declares that if the microscopic model has the same macroscopic characteristics as the macroscopic model, it ensures that the two models are compatible. In addition, this paper presents five detail constraints for model.

(1). $q = kv$;

(2). $q|_{k=k_m} = q_m$, where k_m is the density corresponding to the maximum traffic flow volume rate;

(3). $\frac{\partial q}{\partial k}\big|_{k=0} = v_l$

(4). $q|_{k=0} = 0$, where, v_l is the speed limit;

(5). $q|_{k=k_{jam}} = 0$, where, k_{jam} is the jam density;

Where k_m is the desity corresponding to maximum traffic volume rate; V_l is the maximum speed limit, which describes the speed limit for a low density flow; k_{jam} is the jam density.

The Helly model is a linear car-following model which can be described in equation 2-9.

$$\ddot{x}_{i+1}(t + \Delta t) = C_1 \cdot [\dot{x}_i(t) - \dot{x}_{i+1}(t)] + C_2[x_i(t) - x_{i+1}(t) - d] \qquad \textbf{2-9}$$

Where, $x_i(t)$ is the position of vehicle i at time t, $\dot{x}_i(t)$ is the speed of vehicle i at time t, $\ddot{x}_i(t)$ is the acceleration of vehicle i at time t; d is the desired distance headway. C_1 is the velocity control parameter, C_2 is the distance headway control parameter, and d is the desired space headway. Noting that Helly model is a linear car-following model, which is simpler than most of other nonlinear car-following models that are widely used by microscopic traffic simulation system, such as the psycho-physical car-following model used in VISSIM.

This paper points out that the original Helly model does not satisfy the constraints (1) and (3) listed above. In order to resolve this conflicts, a new defined desired headway (d) is adopted as a modification for SITRA-B (using Helly). According to the adapted microscopic model, the macroscopic model (SIMRES) is described by speed-density equation 2-10 and flow-density equation 2-11.

$$v = \begin{cases} v_l & if \ k < k_0 \\ \dfrac{\left(\dfrac{1}{k} - l\right)}{\lambda + \dfrac{1}{k} - l} & if \ k \geq k_0 \ and \ k \leq k_{jam} \end{cases} \qquad \textbf{2-10}$$

$$q = \begin{cases} k v_l & if \ k < k_0 \\ \dfrac{(1 - kl)v_{max}}{\lambda + \dfrac{1}{k} - l} & if \ k \geq k_0 \ and \ k \leq k_{jam} \end{cases} \qquad \textbf{2-11}$$

The parameters are also designed in order to satisfy condition (1)-(5), where, $\lambda = l(1/\alpha - 1)^2$, $\alpha = l k_m$ and $v_{max} = q_m \cdot l/\alpha^2$. v_{max} is the maximum speed when speed is not limited by v_l. l is the distance between two stationary vehicles. k_0 is the density at which density can be determined by density.

Based on the model definitions, MICMAC discusses interface between macroscopic model and microscopic model as following.

For the macroscopic to microscopic transition, vehicles should be generated based on the upstream vehicle density. A Poisson distribution using the corresponding upstream flow rate as mean value is adopted to obtain the number of vehicles to be generated in microscopic downstream link. In addition, the speeds of the generated vehicles are obtained by using a normal truncated distribution, whose mean value is determined by the corresponding average speed of the upstream link. Furthermore, a binomial distribution is used to determine on which lane the generated vehicles should be placed. Finally, a virtual input cell is placed at the upstream end of the microscopic link.

The interface uses the concentration values of this virtual input cell as feedback to the upstream macroscopic link, which explains upstream propagation of congestion.

For the microscopic to macroscopic transition, the output flow of the upstream microscopic link is counted and transmitted to the downstream macroscopic link. In addition, the system put a virtual microscopic vehicle preceding the last microscopic vehicle in microscopic model, and the virtual vehicle's speed and distance to its following vehicle are calculated by the downstream macroscopic cell conditions. By this method, the traffic state on the downstream influences the upstream link.

This research evaluates performance of MICMAC by comparing the simulation results of MICMAC and single SIMRES, using average cell speed and density as indicators. The results show that MICMAC results do not have significant difference to SIMRES results, which means the internal consistency of MICMAC is well maintained.

2.3.2   System integration

The examples involved in this category focus on combine two developed simulation systems together, in order to improve their analytical ability. Generally, those examples do not require modification on the existing systems; the combination is done by information communication and sharing by additional program modules.

DynusT (Dynamic Urban Systems for Transportation) provides an off-line integration function to convey the dynamic traffic assignment results obtained from DynusT to a VISSIM model. A software package DVC (DynusT to VISSIM Conversion), coming as a component of DynusT, can convert a sub-area in DynusT model (including network, routes, and assigned flows) to VISSIM model. At the meanwhile, compared with the manually process, DVC ensures the compatibility and consistency between the DynusT model and the VISSIM model. The main purpose of this strategy is to use DynusT to perform Dynamic Traffic Assignment (DTA), and at the same time, to use VISSIM to perform detailed analysis. Because those two systems have to run separately, the user has to make sure those two models share same traffic flow dynamics. This condition is met by performing calibration on both two models with field data.

VISSIM and VISUM, as two main software programs produced by PTV Group (Karlsruhe, Germany), provide a series of functions for sharing and exchanging information between them. Originally, VISSIM is a car-following-based microscopic model; VISUM is a macroscopic system focused on demand modeling and networking modeling. In addition, the network data and demand data can be shared by those two programs. For example, the network model built by VISUM can be exported as VISSIM model automatically, and the traffic assignment results obtained from VISUM can also be used as input of VISSIM. Although this process is off-line, the application of this function is noticeably extensive. By those functions, PTV tries to integrate the work flows of network modeling and microscopic infrastructure evaluation.

2.3.3   Model integration

Model integration refers to research on model level and models cooperate in real time. The integrated models are possible to be two highly-developed independent, and also possible to be especially designed for the hybrid traffic simulation research.

In contrast to system integration, the researchers of model integration do not limit the integration work to the level of static off-line information sharing, but evolve to the stage of on-line cooperation. Technically, the integrated systems are running parallel. In order to make this evolution, this research

has to consider more details of the integrated models, and may have to modify the models from its internal structure, rather than only control the output and input of those independent models. In other words, the model integration offers more seamless communication between models. A model integration based approach always involves a flow representation transition method as a sub-component. In addition, it needs to consider more problems both from theoretical and practical perspectives.

Jayakrishnan (2007) presents an example of integrating the microscopic simulator Paramics and the macroscopic simulator DYNASMART, in order to improve the performance of the model when evaluating routing schemes in ATMIS (Advanced Traffic Management and Information Systems). In contrast to a typical hybrid traffic modeling approach, this work does not need a flow transition method, because Paramics and DYNASMART do not interact through boundaries in this example. The integration works as following: a fully developed Paramics model is given. In addition, an equivalent abstract network is constructed in DYNASMART. The main purpose of the abstract network is to incorporate the path dynamics features of DYNASMART by utilizing its ability of describing route changing behavior. In other words, DYNASMART works as an external route decision module for Paramics, and both pre-trip route choice and en-trip route choice can be realized.

Poschinger (Poschinger, et al., 2000) (Kates & Poschinger, 2000) discusses the model-coupling interface between two traffic flow models: PELOPS (microscopic) and SIMONE (Macroscopic). This work considers the integration as a problem of data aggregation/disaggregation, and focuses on the disaggregation part, which was considered to be not investigated enough.

The integration algorithm, which is a control loop similar to an Euler-forwards differencing scheme, is adopted. This algorithm use the state of upstream link at the time step t (demand) to predict the dynamic variables of downstream link for time period from t to $t + \Delta t$. Between boundaries of two models, an overlapping area is placed to implement the flow transition.

The aggregation process used in this research is achieved by introducing a variable speed limit on the upstream microscopic link (in the overlapping area), which is calculated according to traffic state of the downstream macroscopic link; For the disaggregation process, a disaggregator is used to solve the problem of generating vehicles on microscopic link according to the limited macroscopic level information. Stochastic attributes of the generation process is also considered, including initial speed distribution and time headway distribution. The arrival time series are determined by adding an additionally randomized process on the deterministic procedure in which arriving gap equals n/T, where, n is the number of arriving vehicles estimated according macroscopic model and T is the time interval. The initial speed is dependent on arrival time by applying an integrated moving average process ($IMA(1,1)$), which makes the generated speed neither completely deterministic nor completely random.

This research evaluates the proposed approach using generated vehicle speed, time gap, and time to collision (TTC) as indicators. And it shows the proposed method produces data fitting the real data.

Burghout (2004) presents a framework (MiMe) for meso-microscopic model integration, and implements it with a mesoscopic model (Mezzo) and a microscopic simulator MITSIMLab. Mezzo is a queue-server-based mesoscopic traffic flow model. MiMe proposes and focuses on 3 main consistency problems:

- Consistency in route choice and network representation;
- Consistency of traffic dynamics at meso-microscopic boundaries;

- Consistency in traffic performance for mesoscopic and microscopic sub-models.

In addition, MiMe also considers the technological problems about information communication between those two program packages.

The consistency in route choice and network representation is maintained by virtual links (Figure 1). As shown in this figure, microscopic sub-networks are located within a bigger mesoscopic network. For the network definition, it is assumed that any origin or destination node in the microscopic area should be a boundary node that connected to the mesoscopic network. In mesoscopic level, the microscopic model is simplified as several virtual links (Figure 1 (b)) that connect the entering nodes and the exiting nodes of microscopic network. The travel time on mesoscopic virtual links are directly obtained from the corresponding microscopic network. By the virtual links, mesoscopic level includes a complete network, which makes the traffic assignment feasible on mesoscopic level. The similar idea of virtual links can also exit in microscopic network for en-route choice. Microscopic virtual links simplifies the mesoscopic network as shown in Figure 1 (c), by which a complete simulated network on microscopic level exists. Therefore, microscopic level can also perform an assignment process by itself (optional). In summary, by mesoscopic virtual link and microscopic virtual link, the route choice can be represented consistently in both mesoscopic and microscopic level.

(a) The entire network   (b) Meso virtual links   (c) Micro virtual links
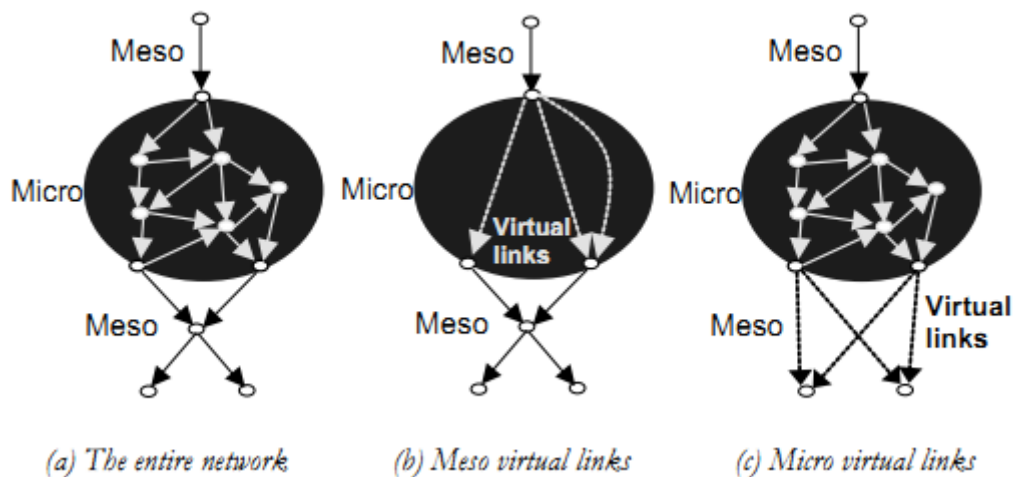
Figure 1 Virtual links for route choice and network representation consistency (Burghout, 2004)

Dummy nodes, which are placed on existing links, maintain the consistency of traffic dynamics at meso-microscopic boundaries. A dummy splits a link into two segments: microscopic segment and mesoscopic segment; in other words, the dummy node works as the boundary between mesoscopic and microscopic models. MiMe performs flow transition at dummy nodes. For boundaries from mesoscopic to microscopic, MiMe generates vehicles in microscopic entrance, at the same time, assign them the initial values, including lane, time-headway, speed, and acceleration. These values are determined in accordance with the traffic situation upstream (mesoscopic) and downstream (microscopic) of the boundary. For the boundaries from microscopic to mesoscopic, MiMe creates a virtual vehicle, whose attributes are determined by the downstream (mesoscopic) traffic state, before the last real vehicle. By this method, the traffic state in the downstream mesoscopic link can affect the state of vehicles in upstream link. The idea of flow transition is similar to the methods discussed in (Bourrel & Lesort, 2003) and (Magne, et al., 2000). In addition, Burghout focuses on the initial speed generation problem for the boundaries from macroscopic to microscopic. Four different

methods are proposed and tested, and the three regime method is chosen. The three regime method determines the initial speed of generated vehicle as equation 2-12.

$$v = \begin{cases} v_{front} & if \ t_1 < t_h \leq t_2 \\ \alpha v_{desired} + (1 - \alpha)v_{front} & if \ t_2 < t_h \\ v_{desired} & if \ t_h > t_3 \end{cases} \qquad \textbf{2-12}$$

Where, $\alpha = (t_h - t_2)/(t_3 - t_2)$; t_1, t_2, t_3 are empirical threshold; t_h is time headway $[s]$; v is the generated speed $[km/h]$; v_{front} is the speed of the leading vehicle $[km/h]$; $v_{desired}$ is the desired speed $[km/h]$;

For the boundaries from microscopic to mesoscopic, virtual vehicles are placed just ahead the exiting vehicles to ensure these vehicles exit microscopic link with proper speeds (Burghout, 2004). However, this method is modified in (Burghout & Wahlstedt, 2007), because using VISSIM instead of MITSIMLab as the microscopic component. The speed of an existing vehicle is calculated as following:

$$v = \begin{cases} v_0 & if \ x > x_{lookahead} \ \ or \ v_0 \leq v_M \\ v_M & if \ x < x_{critical} \ \ and \ v_0 > v_M \\ \alpha v_0 + (1 - \alpha)v_M & if \ x_{critical} < x < x_{lookahead}, and \ v_0 > v_M \end{cases} \qquad \textbf{2-13}$$

Where, $\alpha = x/x_{lookahead}$; v is the new speed of vehicle; v_0 is the initial speed of the vehicle; v_M is the speed corresponding to the upstream mesoscopic link; x is the distance of vehicle to the exit point; $x_{lookahead}$ is the distance of a vehicle to the exit point from where the speed of the vehicle is supposed to be influenced by the downstream mesoscopic link; $x_{critical}$ is the distance of a vehicle to the exit point from where the speed of the vehicle is supposed to be determined by the downstream mesoscopic link.

MiMe is evaluated by both an artificial example and a realistic example. The artificial example shows the basic performance of boundary interface with an idealized link. The simulation results show that both queue forming and dissipating process are well simulated crossing two types of boundaries. A realistic road network, Brunnsviken network, is modeled by MiMe. In this model, the north part of the network is modeled by Mezzo and south part of it modeled by MITSIMLab. In addition, a MITSIM model and a Mezzo model are also built for comparison. 15 min traffic counts collected from 12 control points in the network are compared with field data. The results show the microscopic model has the best performance, and MiMe results are slightly better than Mezzo.

Sewall (2011) proposes a real-time approach for modeling large-scale area by a hybrid model, which combines an agent-based microscopic model and Aw-Rascle-Zhang (ARZ) macroscopic model. By this approach, a large-scale area is modeled as several sub-areas. Each of them is modeled by either microscopic model or macroscopic model.

This paper discusses the methods of conversion of agent-based model to continuum model and of continuum regions to agent-based model.

For the conversion of agent-based model to continuum model, an abstract expression of vehicle are generated from agent-based model (equation 2-14)

$$S_i(x) = H(x - p_i + l_i) - H(x - p_i) \qquad \textbf{2-14}$$

Where, $H(x)$ is the Heaviside function. Therefore the density is defined as (which is different to the widely accepted definition in the society of traffic engineering research):

$$k_k = \frac{1}{\Delta x} \int_{k\Delta x}^{(k+1)\Delta x} D(x)dx \qquad \textbf{2-15}$$

Where,

$$D(x) = \sum_{0}^{n-1} S_i(x)$$

2-16

The calculated density is used to calculate related average speed. After the process (calculating density and average speed), a virtual macroscopic link is described, which can work as an upstream link of its downstream macroscopic link.

The conversion of the continuum model to the agent-based model is accomplished by using an inhomogeneous Poisson processes to generate individual cars on target links. The basic idea of this method is to implement a Poisson processes, which is originally used for describing time distribution of events happening, to a spatial distribution of "events happening", i.e., the location.

Algorithm 1 INSTANTIATE-VEHICLES

INSTANTIATE-VEHICLES($\rho[n], \Delta x$)

> // $\rho[n]$ — an array of n density values,
> // Δx — the length of each grid cell

1   $p = []$
2   $\Lambda_{\text{last}}, i, \sigma = 0$
3   **while** true
4      $U = $ UNIFORM-RANDOM-NUMBER$((0, 1])$
5      $\Lambda_{\text{cand}} = \Lambda_{\text{last}} - \ln U$
6      $\sigma_{\text{cand}} = \sigma$
7   **while** $i < n$ and $\sigma_{\text{cand}} + \frac{1}{l}\rho[i]\Delta x < \Lambda_{\text{cand}}$
8          $\sigma_{\text{cand}} = \sigma_{\text{cand}} + \frac{1}{l}\rho[i]\Delta x$
9          $i = i + 1$
10     $p_{\text{cand}} = \frac{(\Lambda_{\text{cand}} - \sigma_{\text{cand}})}{\frac{1}{l}\rho[i]} + i\Delta x$
11   **if** $p_{\text{cand}} > n\Delta x$
12   **return** p
13   **if** $p_{\text{cand}} + l > p[-1]$
14         $\Lambda_{\text{last}} = \Lambda_{\text{cand}}$
15         $\sigma = \sigma_{\text{cand}}$
16         $p = p + [p_{\text{cand}}]$

An algorithm for vehicle instantiation from continuum data

Figure 2 Instantiate-vehicle generation algorithm (Sewall, et al., 2011)

The input of this algorithm is a piece-wise linear density value for each segment with the length of Δx. The output of this algorithm, a vector p, represents location of all the generated vehicles. A uniform random variables U is used to generate an exponentially-distributed random variable.

Based on the methods described above, the flow from the continuum model is converted to the agent-based model by "flux capacitors"; additionally, the inverse transition is done by flow averaging. The basic idea of those two methods are to use virtual links to make conversion (using methods described above, and the converted vehicles (or flow) can flow into its downstream link).

This research compares the simulation frames per second under four conditions: 100% continuum modeling, hybrid modeling with 1% agent, hybrid modeling with 10% agent, and 100% agent modeling. The result show that the hybrid model save large amount of computational resources. The simulation results for one data collection point are compared with real-world data, in terms of 15-seconds vehicle number and average speed. The results match real-world data well.

2.3.4 Summary

As summarized in this section, the research on hybrid traffic simulation can be categories as flow representation transition, system integration, and model integration.

The literature that belongs to the class of flow representation transition focuses on solving the problems of boundary performance from the mathematical perspective. The findings of them are widely implemented by the following research. However, that research is limited in the theoretical stage and ignores diverse practical problems. This is to say, they are not ready for applications for realistic simulation analysis.

The research categorized as system integration comes from different sources. They are based on developed simulation tools, especially the commercial software tools. The original idea of the integration is to improve the analytical ability of those tools by combining their functions. Therefore, they have to face more various problems, such as network consistency, route expression, etc. Their solid foundations (the developed and well coded simulation software) make them have diverse output and far closer to practical application than the class discussed before, however, the function is limited because they exchange information off-line.

Model integration is the state of the art of hybrid traffic simulation. The performance and flexibility are improved by their foundations of developed basic models. This research integrates theoretical findings of flow representation transition, and also considers practical applications as discussed by system integration. Because the integrated models exchanges information on-line, it improves the analytical ability of the developed system by integrating the benefits of independent models. Because of the large amount of coding and validation work, they are not fully developed and investigated with real cases.

However, the model integration still focuses on utilizing data of those independent models in an intuitive and uncomplicated way, without digging deep into the relation between model dynamic attributes and the difference of those models. In other words, it ignores the unique value that belongs to a model with a specific level of resolution.

2.4 Cell Transmission Model (CTM)

This chapter reviews the existing works on CTM by a problem-oriented structure. Six problems are discussed in this chapter including: cell length, road representation, traffic signal, permitted left-turn, dynamic behavior of CTM, and disadvantages inherited from LWR model. Those problems are important if CTM is used as the macroscopic level of a hybrid traffic simulation system.

2.4.1 Introduction to CTM

Cell transmission model (CTM) is a discrete and simplified version of the LWR model (also called Kinematic-wave model) (Lighthill & Whitham, 1955). It is first presented by Daganzo (1994) (1995) base on the Godunov's method. Because of its advantages of low requirements for computational resource and ease of coding, it has been investigated for diverse applications, including freeway modeling (Lee & Ozbay, 2009) (Munoz, et al., 2006), urban road modeling (Almasri, 2006) (Pohlmann, 2010), traffic assignment (Ishak, 2006), etc. Besides applications, various extensions of CTM have been investigated, such as lagged CTM (Daganzo, 1999) (Szeto, 2008), stochastic CTM (Sumalee, et al., 2011) (Alecsandru, 2006), etc.

As a typical flow representation-based model, LWR model describes traffic participants as "flow", which is quantified by three basic variables: (average) speed ($[m/s]\ or\ [km/h]$), vehicle density ($[veh/m]or\ [veh/km]$), and traffic volume rate ($[veh/s]\ or\ [veh/h]$). The principle of LWR model can be describes as two items:

1. The conservation law (equation 2-17). This equation describes a basic attribute of a flow: for any point in the space, at any time interval, the volume flowing into the point equals that flowing out of it. Or it can be summarized as the "conservation of mass". Theoretically, this equation can be deduced from 3 aspects (Zhang, et al., 1992), based on application of Stokes' theorem.

2. Fundamental diagram among volume rate, vehicle density, and average speed (equation 2-18). Those equations explain the basic attributes based on the field data. Function Q describes the relation between density (k) and volume (q), which is also known as fundamental diagram.

$$\frac{\partial k}{\partial t} + \frac{\partial q}{\partial x} = 0 \qquad \textbf{2-17}$$

$$\begin{cases} q = Q(k) \\ \quad q = kv \end{cases} \qquad \textbf{2-18}$$

In LWR, the conservation relation is maintained, because equation 2-17 is deduced from equation (2-19), which comes from a strict description of conservation attribution:

$$\frac{d}{dt}\iiint kdV = -\iint k\vec{v}\cdot\vec{n}dA$$

$$\Rightarrow \iiint \left\{\frac{\partial k}{\partial t} + div(k\vec{v})\right\}dV = 0$$

$$\Rightarrow \frac{\partial k}{\partial t} + div(q) = 0 \qquad \textbf{2-19}$$

Where, V is any three-manifold in the space; A is the boundary of V; $\vec{n}$ is the unit normal vector field to A; $div(q)$ is the divergence of q. This derivation is based on the Divergence Theorem. Under a scalar situation, in which the flow (q) is one-dimensional, equation 2-19 reduces to equation 2-17.

The equation system expressed by equation 2-17 and 2-18 is possible to be solved both by analytical methods and numerical methods. The solution of them is deeply discussed in the field of physics and applied mathematics. Theoretically speaking, the difficulties in solving those equations are determined by the form of equation 2-18 ($Q(k)$).

CTM is an application of Godunov's method in continuum (traffic) flow, which means that as $\Delta t \to 0\ and\ \Delta x \to 0$, the solution of CTM converges to the analytical solution of the conservation system. The main idea of CTM is to split a link into several cells with equally length, a traffic flow propagates from one cell to its following one.

However, as a simulation model, CTM is different to original Godunov's method from several perspectives. One of the main differences is that CTM has a cell length (L) and related time interval (Δt), which cannot be considered as infinitesimal. Therefore, the classic CTM is based on an assumption: within a cell, the density of vehicles is evenly distributed. Another basic assumption of classic CTM is about the cell length. The cell length is equal to the length over which a vehicle can run

at the free flow speed during one time interval, e.g. $L = v\Delta t$. Where, L is cell length, v is free flow speed, and Δt is the simulation cycle. Considering the cell length and simulation cycle that are generally adopted, CTM does not offer an accurate solution but an approximation. Based on this fact, CTM are more suited to be considered as a traffic simulation model that is based on the demand / supply relation.

Generally, CTM uses a piece-wise linear fundamental diagram for a simplified equation 2-18 ($Q(k)$). A typical example of it is shown in Figure 3 (the basic diagram is from (Daganzo, 1995), and the colorful mark is made by the current thesis).
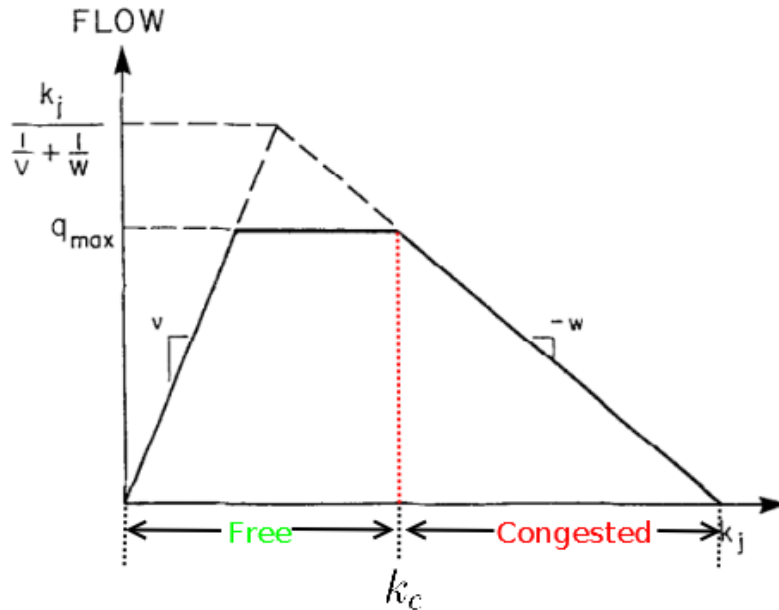
Fig. 1. The equation of state of the cell-transmission model.

Figure 3 A trapezoidal Fundamental Diagram (Daganzo, 1995)

The basic update equation of CTM is (Daganzo, 1994):

$$q_{i-1,i}(t) = min\left\{n_{i-1}(t), Q_{max,i}(t), \frac{\omega}{v}[N_i(t) - n_i(t)]\right\} \qquad \textbf{2-20}$$

In addition, the equation can be summarized from the perspective of demand (D) and supply (S), by equation 2-21, 2-22, and 2-23.

$$D_{i-1} = min\{n_{i-1}(t), Q_{max,i}(t)\} \qquad \textbf{2-21}$$

$$S_i = min\left\{\frac{\omega}{v}[N_i(t) - n_i(t)]\right\} \qquad \textbf{2-22}$$

$$q_{i-1,i}(t) = min\{D_{i-1}, S_i\} \qquad \textbf{2-23}$$

As a discrete version of LWR model, the original CTM inherits typical disadvantages of LWR (which will be discussed in section 2.4.7); in addition, the mechanism of CTM has its own limitations. All those observed limitations and disadvantages lead to diverse research of enhancements on original CTM.

The disadvantages of and limitations on CTM come from the following perspectives:

1. Uniform cell length. In classic CTM, the free flow speed and time interval determine the cell length. In order to satisfy the stable condition of Godunov's method and utilize the most accurate level of Godunov's method (minimize viscosity of the method), the cell length is

generally set to $v\Delta t$. But, this setting brings two questions for applications. First, it is difficult to divide any link into several equal cells for realistic roads. Because the cell length is determined before model building, it is impossible to guarantee all lengths of links are an integral multiple of the given cell length; second, the range of free flow speed is limited. On the requirement of CFL condition, the cell length is the maximum limitation of vehicle traveling distance with free flow speed in one time interval, which means if the cell length is decided, the maximum free flow speed is also decided. This situation presents difficulties to calibration process.

2. Specific simulating objects and lack of expression for diverse infrastructures. The original CTM is designed for highway simulation. Generally, the CTM has the capabilities to simulate an ideal traffic flow on typical freeways, but not the capabilities to simulate complex infrastructure environments and operational methods. For example, the original CTM only defines diverging and merging cells, but does not include the expression of traffic signals, permitted left-turn, influences of pedestrians, etc.

3. Deterministic model. The original CTM is a deterministic model; this is to say, a specific input will lead to a unique output. The stochastic attributes of traffic behaviors cannot be described by CTM.

The rest of this chapter summarizes and discusses several enhancements on CTM, or solutions for practical applications. The concerned points include non-uniform cell length, road network representation, signal expression, and road network representation. The stochastic CTM is not an interesting point of the current thesis.

2.4.2 Non-uniform cell length

Last section states the disadvantages of using uniform cell length. There are at least four advantages brought by non-uniform cell lengths, which makes its application more flexible.

- Better representation of road networks;

- Better adaptability by allowing more wider free flow speed;

- Fewer cells for the same network.

- The changeable simulation cycle.

Daganzo (1994) (1995) presents the original CTM with a uniform cell length, however, the non-uniform cell length can also be implemented without any modification on CTM. For cases of non-uniform cell lengths, only the minimum cell length is fixed for model stability. Daganzo (1999) mentions the cell length problem within the new extension of CTM, the lagged CTM, but also by considering the stability of lagged CTM. Munoz (2004) presents an example of CTM's application with non-uniform cell lengths. The minimum, mean, and maximum used in this research are 293m, 551m, and 1224m. In this application, the vehicle density is used instead of vehicle load number to describe cell state, which makes non-uniform cell lengths work with the original CTM update principles.

Munoz's method does not solve the cell length problem from the theoretical aspect but an approximation of original CTM by using the assumption of evenly vehicle density distribution. In theory, the accuracy of CTM can be improved by implying smaller cell length (together with corresponding simulation cycle), because of the basic principle of Godunov's method. Sumalee (2011) proposes a stochastic CTM and also allows the non-uniform cell lengths by using vehicle density instead of cell occupancies.

Flötteröd (2005) introduces a new concept into CTM: resource. Resource means the possible transport vehicles between two adjacent cells. This is to say, the system update process can be defined as a resource assumption process. This system does not have a fixed simulation cycle. The system should be updated at the time points on which at least one resource point exhausts. In other words, this example allows the cell length longer than the minimum cell length.

From the theoretical perspective, if the stable condition is satisfied, no odd results emerge, such as minus density values. With regard to the accuracy, long cell length does increase the viscosity of the model, however, the exact results can only be obtained when $\Delta x \to 0 \ and \ \Delta t \to 0$, which also means the convergence does not break if Δx and Δt are equivalent infinitesimals.

Lee (1996) discusses the "adjustable cell length" problem. Cells are classified into two categories: big cells and small cells. The length of a big cell should be an integral multiple of the length of a small cell. The simulation time interval is decided by the length of small cell. This method can also be considered as a combination of several basic cells (small cells) into one big cell. In addition, Lee researched the traffic flowing both from big cell to small cell and inverse one. For the traffic flow from big cell to small cell, the main idea is that the total potential output volume of upstream big cell should be divided into several parts evenly. For the traffic flow from small cell to big cell, the available acceptance of the downstream big cell should be divided into several parts too. A similar but more detail method is used by Alecsandru (2006). This thesis divides a cell into several sub-cells; all sub-cells share a common predefined length except for the last one of each cell. This last sub-cell is always shorter than the common one. By this design, a cell can have arbitrary length. The simulation cycle is decided by the length of common sub-cell. This means that the update principle of this model is much more complex than the classic CTM. In this model, it is possible for a vehicle, which is located in the second last sub-cell, to go into its downstream cell (the first sub-cell of downstream cell) passing through the last sub-cell (without stay in the last sub-cell for one cycle) in a single simulation cycle. This approach develops a special update principle of the last sub-cell, by considering the extra vehicles that get across the last sub-cell in a single simulation cycle.

2.4.3 Road network representation

The original CTM derives from discrete numerical for solving LWR model, which does not include a consideration on the road network representation. Daganzo (1995) proposes the description of diverging and merging cells in order to modeling necessary road elements for freeway, but they are limited to 2-legged situation. This is sufficient for freeway system modeling, because most of the on/off-ramp can be modeled by the 2-legged cells. However, since the urban roads need to be modeled, a 3-legged cell or even the multi-legged cell is needed.

In (Daganzo, 1995), the diverging situation is shown in Figure 4. For the diverging cell, constant parameters (β_1, β_2), which represents turning percentages, are defined. The following relation has to be kept:

$$\beta_1 + \beta_2 = 1$$

Therefore, the traffic volume between the cells can be summarized as equation 2-24.

$$\begin{cases} q(t) = \min\{D(t), S_1(t)/\beta_1, S_2(t)/\beta_2\} \\ q_1(t) = q(t) \cdot \beta_1 \\ q_2(t) = q(t) \cdot \beta_2 \end{cases}$$

2-24
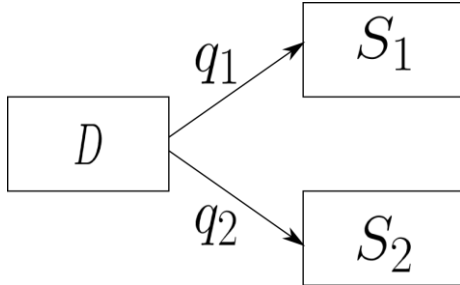
Figure 4 Diverging cell　　　　　　**Figure 5 Merging cell**
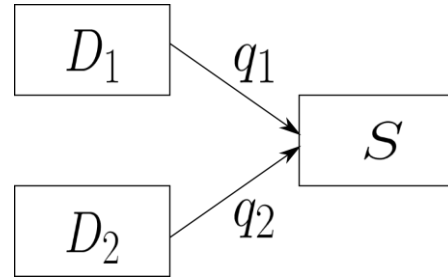
The merge is shown in Figure 5. The key point of this solution is to introduce a priority parameter (p_1, p_2), which is defined as

$$p_1 + p_2 = 1$$

This parameter describes the nature attribute of the modeled merge cell. p_1 and p_2 are predefined positive decimals. Therefore, the volume q, q_1, q_2 can be formulated as equation 2-25 and 2-30

$$\begin{cases} q_1(t) = D_1(t) \\ q_2(t) = D_2(t) \qquad if\ D_1(t) + D_2(t) \leq S(t) \\ q(t) = q_1(t) + q_2(t) \end{cases}$$

2-25

$$\begin{cases} q_1(t) = mid\{D_1(t), S(t) - D_2(t), p_1 S(t)\} \\ q_2(t) = mid\{D_2(t), S(t) - D_1(t), p_2 S(t)\} \qquad if\ S < D_1 + D_2 \\ q(t) = q_1(t) + q_2(t) \end{cases}$$

2-26

Pohlmann (2010) presents an extension on diverging and merging cell for 3-legged situation, which is designed for the intersection modeling. This approach inherits the method in classic definition, and extends it into three-dimensions. For a 3-legged diverge,

$$\begin{cases} q(t) = \min\{D(t), S_1(t)/\beta_1, S_2(t)/\beta_2, S_3(t)/\beta_3\} \\ q_1(t) = q(t) \cdot \beta_1 \\ q_2(t) = q(t) \cdot \beta_2 \\ q_3(t) = q(t) \cdot \beta_3 \end{cases}$$

2-27

For a 3-legged merge, we have:

$$\begin{cases} q_1(t) = D_1(t) \\ q_2(t) = D_2(t) \\ q_3(t) = D_3(t) \qquad if\ D_1(t) + D_2(t) + D_3(t) \leq S(t) \\ q(t) = q_1(t) + q_2(t) + q_3(t) \end{cases}$$

2-28

$$\lambda_{min} = \min\{D_1(t)/p_1, D_2(t)/p_2, D_3(t)/p_3\}$$

2-29

$$\begin{cases} q_1(t) = p_1 S(t) \\ q_2(t) = p_2 S(t) \\ q_3(t) = p_3 S(t) \qquad if\ \lambda_{min} \geq S(t) \\ q(t) = q_1(t) + q_2(t) + q_3(t) \end{cases}$$

2-30

$$\begin{cases} q_i(t) = D_i(t) \\ q_k(t) = mid\left\{D_k(t), \left(S(t) - D_1(t)\right)\dfrac{p_k}{p_k + p_j}, S(t) - D_i(t) - D_j(t)\right\} \\ q_j(t) = mid\left\{D_j(t), \left(S(t) - D_1(t)\right)\dfrac{p_j}{p_k + p_j}, S(t) - D_i(t) - D_k(t)\right\} \\ q(t) = q_1(t) + q_2(t) + q_3(t) \end{cases} \quad if \ \lambda_{min} < S(t) \qquad \textbf{2-31}$$

Which upstream leg is selected as q_i is determined by whether it has the minimum value of λ, e.g. $\arg\min\{D_1(t)/p_1, D_2(t)/p_2, D_2(t)/p_3\}$.

All the approaches for merging and diverging stated above are based on the constant turning percentages.

Flötteröd (2011)presents a general intersection connector which can work with CTM. In this research, an incremental node model (INM) is used to describe an intersection. INM have several receiving and sending legs, which can be used to model the entrance and exit link of an intersection. The relation (constrains) among receiving, sending, and the attributes of the intersection is described as a optimization problem. The modeled problem can be solved by the incremental transfer principle of Daganzo (1995).

2.4.4 Expression of Traffic signal

In order to simulate an urban network, it is impossible to avoid signal controlled intersection modeling. CTM only describes two signal signs: red and green. Therefore, when applying CTM for intersection modeling, the original signal plan has to be transformed to a plan expressed by effective green and effective red. In other words, the displayed green time or red time cannot be used directly.

Almasri (2006) and Pohlmann (2010) share similar methods for describing traffic signals in CTM. In those two examples, red, amber and red/amber sign time is modeled as effective red, and the rest signal time is modeled as effective green. In terms of practice, Almasri (2006) sets the outflows of critical cells to be zero if their corresponding signal head shows red, and the regular outflows are used if the corresponding signal head shows green. Pohlmann (2010) does the modification with same cells but on demand volume of cells instead of traffic volume between cells. .

2.4.5 Permitted left-turn

Pohlmann (2010) summarizes the method for modeling permitted left-turns, and presents a simplified method for CTM based on HBS.
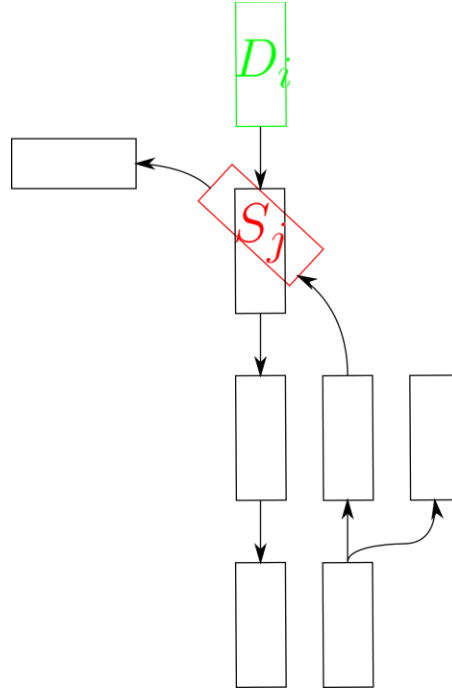
Figure 6 Permitted left-turn

Figure 6 shows an example of permitted left-turn. The basic idea of this method is to amend the supply volume (S_j) of the turning cell (Cell j), depending on the demand volume (D_i) of cell i, which is the related cell in the direct move. The relation between the amended supply volume and related demand volume is described in equation 2-32.

$$S_j(t) = \frac{1}{t_f} \cdot \left(T - t_0 \cdot D_i(t)\right)$$
2-32

Where, t_f is additional headway accepted by following vehicles; $t_0 = t_g - t_f/2$; t_g is shortest headway accepted by leading vehicle in queue. T is simulation cycle.

Flötteröd (2011) considers the permitted-left turn by adding an extra constrain within the proposed INM (Incremental node model). The constrain can be described as equation 2-33.

$$\Delta_{minor} = \frac{1}{t_f} e^{-q_{main}\left(t_g - \frac{t_f}{2}\right)}$$
2-33

Where, Δ_{minor} is the inflow volume of the left-turn flow (the minor one); q_{main} is the inflow of the related direct flow (the main one); t_f and t_g also come from German HBS (FGSV, 2001), and have the same definitions to the last examples.

All the two methods are based on the recommendation from HBS, which is modified to work with the specific mechanism of CTM and INM. The basic values of those two parameters (t_g and t_f) also comes from HBS.

2.4.6 Dynamic behavior of CTM

Although there are large amount of applications of CTM, the dynamic behavior of it has not been fully investigated. Gomes (2008) presents an analysis on dynamic behavior of CTM.

The dynamic analysis of CTM is based on the structure of equilibriums. Given the traffic demand for each cell (cell 1 to N, traffic flow moves from cell N to cell 1) is expressed by a vector $r = (r_0, \dots, r_N)$, the density of all cells is expressed by a vector $k(t) = \left(k_0(t), \dots, k_N(t)\right)$, and the flow rate is

expressed by vector $f(t) = (f_0(t), \ldots, f_i(t))$. If a demand r satisfies $0 \le f_i \le F_i (0 \le i \le N)$, r is said to be feasible; if a demand r satisfies $0 \le f_i < F_i (0 \le i \le N)$, r is said to be strictly feasible.

The CTM principle can be expressed as equation 2-34, 2-35, 2-36, and 2-37. In this model, traffic flow moves from cell N to cell 1.

$$k_i(t + \Delta t) = k_i(k) - f_i(t)/\bar{\beta}_i + f_{i+1}(t) + r_i(t), 0 \le i \le N-1 \qquad \textbf{2-34}$$

$$f_i(t) = f_i(k(t)) = \min\{\bar{\beta}_i v_i k_i(t), \omega_{i-1}[\bar{k}_{i-1} - k_{i-1}(t)], F_i\} \quad 1 \le i \le N \qquad \textbf{2-35}$$

$$f_0(t) = f_0(k(t)) = \min\{\bar{\beta}_0 v_0 k_0(t), F_0\} \qquad \textbf{2-36}$$

$$k_N(t + \Delta t) = k_N(t) - f_N(t) + r_N(t) \qquad \textbf{2-37}$$

Where, $f_i(t)$ is the traffic volume from cell i-1 to cell i-1 in time period t to $t + \Delta t$.

An equilibrium refers to a state $k = (k_0, \ldots, k_{N-1})$ for a feasible demand r, if the constant state $k(t) \equiv k$ is a solution of equation 2-34, 2-35, and 2-36. In addition, if is guaranteed that a feasible demand r has a unique uncongested equilibrium. The main conclusions of this research can be summarized as two theorems.

This paper proposes several useful conclusions about the dynamic performance of CTM based on strictly mathematical deduction. We cite three of them as following:

Theorem 1: Let r be a feasible demand, f be the induced equilibrium flow, and $E(r)$ the equilibrium set. If r is strictly feasible, $E(r)$ consists of the unique uncongested equilibrium k^u. Otherwise, partition the freeway into segments $S^0, \ldots, S^K$ corresponding to the bottleneck sections $0 \le I_1 < \cdots < I_K \le N-1$. Then $E(r)$ is the direct product:

$$E(r) = E^0(r) \times \cdots \times E^K(r) \qquad \textbf{2-38}$$

Each $E^k(r)$ decomposes as the union:

$$E^0(r) = \{k^{u,0}\}, \qquad E^t(t) = \bigcup_{j \in S^k} E_j^t(r), t \ge \Delta t \qquad \textbf{2-39}$$

Each $E_j^t(r)$ is the union of two connected line segments, given by the "close form" expression. $E^t(r)$ is the union of $|S^k|$ connected straight line segments. Consecutive sets $E_j^t(r)$ and $E_{j+1}^t(r)$ have exactly one point in common, and they are ordered: if $k^t \in E_j^t(r)$ and $k'^t \in E_{j+1}^t$, then $k^t \le k'^t$. In particular, the most congested equilibrium in $E^k(r)$ is $k^{con,t}$ with components $k_i^{con}, i \in S_k$. Every $k^t \in E^k(r)$ lie between the uncongested equilibrium $k^{u,t}$ and $k^{con,t}$, i.e., $k^{u,t} \le k^t \le k^{con,t}$. Hence for all $k \in E(r)$, $k^u \le k \le k^{con}$, in which the most congested equilibrium is $k^{con} = (k^{u,0}, k^{con,1}, \cdots, k^{con,T})$. Lastly, $E(r)$ froms a connected, topologicall closed surface of dimension K in the N-dimensional state space.

Theorem 2: the CTM model is a convergent system, i.e. every trajectory converges to some equilibrium in $E(r)$.

Theorem 3: every trajectory lies between $\{\hat{k}(t)\}$ and $\{\bar{k}(t)\}$; the trajectory starting with the empty freeway converges to the uncongested equilibrium k^u; the trajectory starting with the completely jammed freeway converges to the most congested equilibrium k^{con}.

Those three theorems predict the behavior of CTM. It can be considered as two categories: strictly feasible demand or feasible demand. If the demand is strictly feasible, the model converge to a unique uncongested stable state; if the demand is only feasible, several queues will form upstream of blockage points, and the final state is determined by demand and boundary condition together.

However, the final state can also be expressed as an equilibrium, and for given demand, the equilibrium set can be found.

The convergence attributes described in Theorem 2 and 3 is determined mainly by the attributes of the strictly monotone map (g). Given $k(t + \Delta t) = g(k(t))$ as the principle of CTM, it is guaranteed that: $x < y \Rightarrow g(x) < g(y)$. This method is addressed as monotone method.

At least 2 useful points can be derived from this research:

1. The cells do not have to share a unique FD. With different values of FD, the dynamic behaviors described in this paper still exist.

2. Capacity of bottleneck determines the final state of the modeled network. Therefore, the capacity of each link (or specific part of a link) is the key points to be cared for the user.

2.4.7 Disadvantages inherited from LWR model

From the aspect of mathematic, the CTM is a numerical method for solving the partial differential equation of LWR model (or conservation law equations). The validation about CTM shows that it captures the basic attributes of traffic flow. However, there are several features which are observed from realistic data cannot be predicted by this model. And those points are inherited from LWR model.

Zhang (1992) summaries the deficiencies of LWR model as the following four points:

(1) Driver differences. LWR model is not adequate for modeling light traffic in which passing behavior is frequent.

(2) Shock structure. In LWR, the shock has no width. However, in realistic world, a shock is not a sharp front but an area, because vehicles cannot change their speed during infinite time

(3) Only one family of waves. LWR only describes a family of waves that travel against the traffic flow.

(4) Traffic instability. LWR model maintains the stability of traffic flow. When disturbance emerges in a flow, LWR model leads to a stable state quickly. This means the vehicles has to show a "right" behavior forever. However, this situation is not true in the realistic world. A typical unstable phenomenon, the "stop and go" situation, is observed broadly and lasts for as long as hours.

Nagel (2005) discusses the empirical agreement and disagreements with KWM. The general conclusion of this paper includes two main points. On one hand, KWM model captures many effects of realistic attributes of traffic flows, not fully but to some specific extent. On another hand, several observed phenomenon cannot be explained well by KWM. Several points discussed in this paper are interesting for us.

The main empirical disagreement with KWM relate to the congested regime. As seen in Figure 7, the data points in the free-flow regime shows a well fit to linear function, however, the data points in the congested regime are widely scattered. A reasonable hypothesis of this reason is that in congested regime, the interaction between vehicles dominates the density increase in terms of affecting the

traffic flow. This phenomenon gives rise to several disagreement points of KWM. Several phenomena, including unstable flow, spontaneous breakdown out of nothing, two-capacity phenomenon, wide jams (wide jam moving regions), and uncorrelated flow (stop and go situation), which can be observed from empirical data, cannot be described by KWM.

The empirical agreements with KWM are not fully but to some extent. For example, when describing an acceleration flow, which relates to both free regime and congested regime, such as a flow start up before a stop line, KWM has a rough description about the discharging flow rate, because a real queue discharging flow rate does not show a steady maximum flow rate as KWM predicated. Another example is the no-width shock waves.
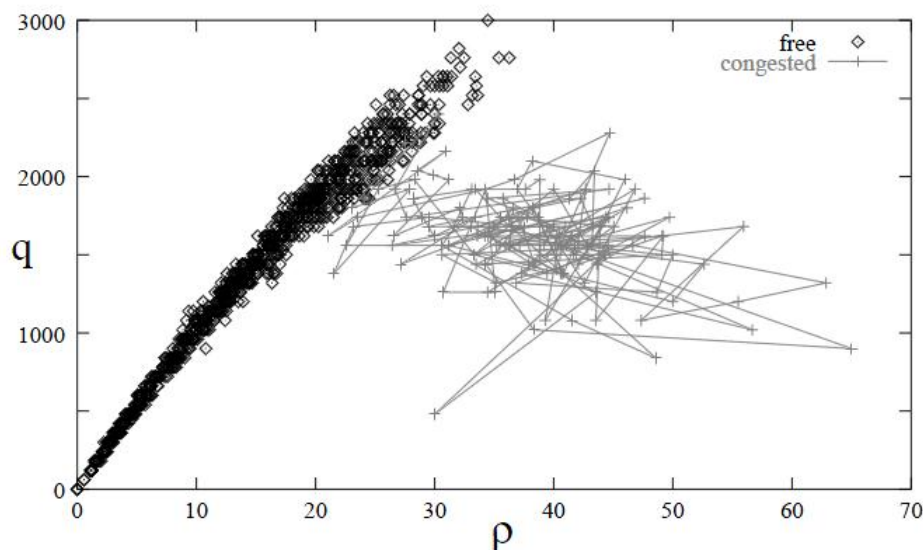
Figure 7 Empirical flow-density data

source: (Nagel & Nelson, 2005), recorded on the German freeway A43; Density (x axis) is in [veh/km/lane], flow (y axis) is in [veh/h/lane]

2.4.8 Summary

The CTM has been adopted for modeling diverse objectives, including freeway system and urban roads. In addition, the tasks of those application are also various, including state estimation, dynamic traffic assignment, signal evaluation, etc. The large amount of applications reported is a solid evidence of the feasibility and flexibility of CTM. Associated with the ease of code of CTM, CTM is one of the most promising macroscopic / mesoscopic traffic flow modeling approach.

2.5   Calibration algorithm for traffic simulation model

2.5.1 Overview

Model calibration and validation are two key processes within a traffic system modeling project. For engineers and researchers, those two steps are even more important than model building, because these two steps directly determine whether the simulation model is an accurate representation of the system under analysis. Generally speaking, calibration is the process that searches the suitable values for model parameters; and the validation is the process that verifies whether the calibrated parameter values produce simulation results that represents the reality.

The calibration and validation of traffic simulation model are structured as four stages by (Barcelo, 2010) as: (1) error checking; (2) capacity calibration; (3) route choice calibration; (4) performance validation. Error checking is the step that checks for the sensible errors that are introduced into the model by mistake. Capacity calibration is the process that focuses on the parameters that influence the traffic capacity, in order to produce the expected traffic capacities at the control points. For a traffic flow model, the capacity calibration is the main process that is must to be performed. Route choice calibration should be carried out if the traffic assignment is included in the model results; this process tunes the parameters of route choice and tries to produce better simulated flow paths on links. Performance validation is the final process, which uses selected performance indicators, such as travel times, queue lengths, traffic volumes, and vehicle densities, to check whether the calibrated parameters are valid; in addition, some fine-tuning adjustments are also needed within the validation process. The four steps are idealized strategies for model calibration; however, practical modeling processes are probable to face diverse problems and may need repetitions of those steps.

There are two categories of model calibration: the first category is the measurement based method, which measure the values of parameters according to their definitions. The second category is the performance based method, which evaluates the performance of a model by predefined indicators and try to improve the model performance by tuning model parameters. Considering the discrepancy between realistic world and simulation model, the second group of methods tends to offer more robust and dependable solution, since the target of simulation is to get performance indicators but not values of parameters, and the measurement based method does not guarantee a better performance directly due to the mechanisms of simulation models.

In this section, the literatures are summarized from two perspectives: the calibration of CTM and the calibration of other models. Each example is summarized focusing on three perspectives: the used data group, the selected indicators, and adopted algorithms.

2.5.2 Calibration of CTM

The main task of CTM calibrating is to calibrate its fundamental diagram. The simplified fundamental diagram can be depicted by 4 factors: free flow speed (v), backward wave speed (ω), cell capacity (q_m), and jam density (k_j). Based on the literature review about CTM, different cells can have their own values of parameters if necessary. Besides the model structure, which should be determined during the model development process, the values of parameters determines the performance of the developed model significantly; therefore, those four parameters are also involved in the calibrated parameters. Certainly, different applications may have other specific calibrated parameters, such as the case including traffic assignment problems also take the parameters of route choice as calibrated parameters.

Generally speaking, there are two classes of methods for CTM calibration. The first class is based on parameter definitions, which means the calibration methods try to "measure" the parameters by their physical definitions. The second class is based on performance optimization. These methods treat the built model as a black box. During calibration process, users tune values of parameters in order to obtain better consistency between simulated performance and field performance. Typical traffic variables, i.e., traffic volume, vehicle density, are used as the performance indicators.

Lin (1995) reports a validation research on CTM for freeway links. The task of this work is to investigate the feasibility of original CTM (this work is done when CTM is just presented). This research includes the parameter calibration process that is mainly based on parameter definitions.

This research calibrates the parameters of classic CTM by the data obtained from freeway segments of a real case (I-880 in California). There are four stations located on a 2.1km long freeway; the distance between two adjacent stations is about 0.5km. This freeway link consists of 5 lanes, each of them are equipped with loop detectors. Data stations collected occupancy (%,[-]) and vehicle count ([veh/h]) at various sample intervals from 1 second to 60 seconds. The capacity and free flow speed are determined by examining all available data sets from 5:00 am to 6:00 am. During that time period, the free-flow traffic dominates other scenarios. Based on this data sub-set, suitable values for capacity and free flow speed are chosen empirically according to their definitions. The backward wave speed is measured by specifically chosen data sets, which have to be identified as under the congestion condition, and the congestion situation should result from a capacity recovery of a temporal downstream bottleneck, under which the wave speed is considered to equal the defined backward wave speed. Once the target data sets are picked out, the wave speed can be calculated by measure the time difference between the time points of flow regime change at two adjacent stations. The jam density is determined by the intersection point of FD curve with x-axis, because the triangular FD is adopted here (critical density and backward wave speed are already decided). In addition to parameter calibration, this research evaluates CTM by modeling a 1km-long sub link of the same freeway mentioned above. Besides the calibrated backward wave speed, the validation process tested the backward wave speed to be $1/1, 1/2, 1/3, …, 1/8$ of the free flow speed. The comparison of traffic flows shows that the coefficient with value between $[1/2, 1/6]$ produces best performance, and this range also fits the measured backward wave speed. Traffic flow volume and occupancy are used as indicators of model performance for validation, and CTM shows an acceptable performance for modeling freeway system.

Munoz (2006) reports a calibration process for the switching-model CTM. The Switching-model CTM is an extension of CTM by categorizing the classic CTM as five sub-modes. Each of the modes still follows the basic principles of CTM, but is transferred to a linearized model. Therefore, the switching-model CTM has the same parameters definition to CTM, but can have independent parameter values for different mode. A 23km long test freeway segments are divided into 41 cells in this paper; in reality, there are 22 detector stations located on the segments. Each station has loop detectors for each lane, and offers the traffic volume ([veh]) and occupancy (%, [-]) at time intervals of 30 seconds. A trapezoid FD is employed in this example. Free-flow speed is determined by implementing least-squares fit on the flow-density data which are over period 5:00 am to 6:00 am, based on the same reason mentioned in the last example. Cell capacities are considered as two groups: first, nominal cell capacity, which means the ideal capacity without bottleneck, is determined by choosing a value larger than the maximum observed volume data; second, bottleneck capacity, which is the capacity that is affected by traffic flows, is determined by specific data groups that obtained from the bottleneck points. The backward congestion wave speed (ω) and jam density (k_J) are determined by implementing a least-squares fit on the flow-density data with several constrain conditions on the parameter values. The calibrated parameters shows spatial variation, e.g., $v \in [94,109]km/h$, $Q_M \in [1488, 2300] \ vphpl, \omega \in [22,31] \ km/h$. In addition, time-varying parameter adjustments are also adopted for backward wave speed. Partial Total Travel Time (TTT) and cell densities are used as the indicator of the model performance for validation.

Dervisoglu (2009) measures basic parameters of FD, including free flow speed, backward wave speed, and capacity, by investigating field source of 5-mins density-volume data on a freeway link. The capacity is determined by the maximum observed traffic volume, and the variation of this value is discussed (that leads to the stochastic capacity described in (Muralidharan, et al., 2011)). In addition,

the free wave speed is determined by performing constrained least-squares fit on the volume-density data. After the critical density is obtained, backward wave speed is also determined by performing a constrained least-squares fit on the volume-density data. The intersection point of the congested FD curve and the x-axis is considered as the location of jam density. Besides the normal cell capacity, the capacity drop phenomenon at bottleneck is also considered in this research. The dropped capacity is obtained by performing a pure least-squares fit on the volume-flow data points under the congested regime. Speed, Vehicle Miles Traveled (VMT) and Vehicle Hours Traveled (VHT) are used to show the performance of the calibration results.

Lee (2008) presents a calibration approach based on Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm. Be different to the measurement-based calibration method, this research is optimization-based. By implementing a CTM with trapezoid fundamental diagram, free flow speed and jam density is selected as calibrated parameters. Backward wave speed is set to be equal to the free wave speed, and cell capacities are predefined based on experiences. An objective function (equation 2-40) is defined to measure the fitness of selected parameters by comparing observed and simulated densities.

$$L = \sum_{Interval} \sum_{Detector} \sum_{Lane} \frac{\|k_{real} - k_{sim}\|^2}{k_{real}} \qquad \textbf{2-40}$$

The CTM is used to model a freeway section with 2-lanes (one-way). This 1.6 km length freeway section is divided into 20 cells. The results show that the error of simulated density data reduces from 35.72% to 12.37% by the optimization process.

2.5.3 Calibration of traffic simulation models

There are numerous research about the calibration work of traffic modeling. This section summarizes the calibration approaches that have been used for diverse models, especially about the online calibration. This summary focuses on which simulation tools are used, the calibration based on which data set, which indicators are used to evaluate the parameter performance, what parameters are chosen to be calibrated, and which algorithm or mathematic model is used for describing the calibration problem and solving it.

Ma (2007) presents an approach for parameters calibration of micro simulation model by SPSA (simultaneous Perturbation Stochastic Approximation), and compares the performance of this algorithm with a typical GA (Genetic Algorithm). The calibrated parameters in this research are the driving behavior parameters involved in car-following mode and lane-changing model. This approach treats the parameters calibration process as an optimization process. The optimal objective used in this research is link capacities. And the objective function is defined as the:

$$F = \sum_{i=1}^{M} [GEH(Cap_i) + A \cdot GEH(Occ_i)]$$

Where, $GEH(X)$ is a function defined as:

$$GEH(X) = \sqrt{\frac{(X_p - X_m)^2}{(X_p + X_m)/2}}$$

Where, X_p is the estimated value of item X, and X_m is the field value of item X. In function F, Cap_i is the capacity of all general purpose lanes in one direction on which the data collection location i is located; Occ_i is occupancy of a link at data collection location i.

The field data is collected on highway SR-99 in the city of Sacramento, California. Traffic measurements data, including flow rates and occupancies, are collected by PeMS loop detectors, at time intervals of 30 seconds. The data are collected over one hour from 2:30-3:30 PM. For this research, the model is built by Paramics. Three global parameters are calibrated, including mean target headway, mean reaction time, and driver aggressiveness; Six local parameters are calibrated, including link headway factor, link reaction factor, sign-posing, ramp headway factor, minimum ramp time, and ramp awareness distance.

The optimal results show that both GA and SPSA are able to obtain stable solution, in addition, GA get slice better solution, however, SPSA can save more computational resources.

Lee and Ozbay (2009) propose a method to combine the SPSA and Bayesian sampling, in order to obtain the distribution of model parameters. Paramics is used as the modeling tool. The mean target headway and mean reaction time are selected as calibrated parameters. Traffic volume and speed data is obtained for the Paramics model, in order to calculate the objective function:

$$F = \sum_{lane} \sum_{time} \frac{|Q_{real} - Q_{sim}|}{Q_{real}} + \frac{|V_{real} - V_{sim}|}{V_{real}}$$

Where, Q is traffic volume, and V is vehicle speed.

In this research, a freeway network (5 lanes one-way road) is modeled. The field data is obtained from the database of the modeled freeway segment. Those data are collected from 6:00 am to 10:00 am, for about one month (16 days are used), and aggregated into 5-min counts.

Before calibration process, this research estimates the demand input by a three-order linear regression base on the collected demand volume data. Two parameters, the mean target headway and the mean reaction time, are calibrated. In addition, during the calibration process, not only one group of values are obtained, but several groups according to different demand input sample (generated from the estimated distribution). Simulation results (volume) obtained from 100 samples of demand input and their corresponding optimized parameters are used to evaluate the designed method, by comparing the observed data distribution with the sampled simulation data using a typical Kolmogorov-Smirnov (K-S) test.

Toledo (2003) presents a case study applying calibration and validation of a microscopic traffic simulation model. This research employs MITSIMLab, the model object is a urban-freeway network in the Brunnsviken area in the north of Stockholm, Sweden. Because only aggregate data was collected through loop detectors, this data is used to calibrate general parameters in the simulator. The calibration process is based on an optimization method. This paper points out that the reason that the individual models are not calibrated separately is the difficulties to isolate the contribution of individual models to the overall error (if only aggregated data are available). Therefore, an objective function, which measures the discrepancy between simulated data and field data, are designed. Because of the internal structure of MITSIMLab, the calibration cannot be done by a single step, but split into several sub-steps. The calibrated parameter set consists of three groups: driving behavior parameters, route choice parameters, and OD flow parameters. When the calibration process starts, driving behavior parameters are initialized on a small sub-network abstracted from the Brunnsviken network, because within this sub-network, the located sensor can offer accurate OD flows at one minute intervals. The point measured vehicle speed is used as the performance indicator. By that, path choice set is generated by MITSIMLab. Based on the initialized parameters and path choice set, OD estimation is done by optimizing corresponding objective function, which is based on deviation

between estimated and field OD flows. The output of this step is OD flows. Based on the OD flows and route choice parameters, the driving behavior parameters can be calibrated for the entire network using the same algorithm for the initialize step. After the calibration process, the calibrated parameters are validated by comparisons of traffic flows, traffic times, and queue lengths.

Molina (2005) presents a MCMC (Markov Chain Monte Carlo) method based input estimation for a microscopic network simulation model with CORSIM. The main purpose of this research is to estimate the distributions of input volume (λ) and turning percentage (T) according to the limited observation data, using a full Bayesian analysis method. In contrast to a typical application of calibration of traffic simulation model, this research does not calibrate the traffic behavior parameters embedded in CORSIM system, but the input of the model. Three types of data are used for calibration. Observers placed on the streets entering the simulated network count the vehicle numbers entering the network over one-hour period (C^D), the true value of this item is N^D; video cameras placed at the intersections in the simulated network record the traffic flow situation over one-hour period, the video footages are used to determine the traffic counts for each direction accurately (C^V), assuming this value is accuracy ($C^V = N^V$); Observers placed at the intersections (which are not equipped with cameras) count the numbers of vehicles for each direction (C^T), assuming this value is accuracy ($C^T = \widetilde{N}^T$). The target of a typical Bayesian Analysis is to obtain the posterior distribution of studied variables, based on the likelihood function, and prior distribution of the studied variables. In this research, likelihood function is predefined according to the attributes of CORSIM model, which does not have an analytical expression. Therefore, it can only be estimated by running CORSIM repeatedly. In order to reduce the computational cost of model running, a simulator approximation of CORSIM is developed. According to Bayes' theorem, the posterior distribution is propositional to the product of prior possibility and likelihood value. When those two possibilities can be calculated, the original input estimation problem is transformed to a typical Bayesian analysis problem. In this research, a MCMC analysis based on Metropolis-Hastings scheme is used to solve it.

Antoniou (2007) presents an online calibration approach for a DTA system. In this research the calibration problem is formulated as a nonlinear state-space model, therefore the calibration problem becomes a state estimation problem. The vector of state consists of the parameters needing calibration, including OD flows, segment capacities, and fundamental diagram parameters (speed-density relationship). In total, the state vector has 80 dimensions. The state transition function is formulated as an autoregressive process, which describes the relation between historical states to the current state; the simulator works as the measurement function (from system state to observations). The basic setting of the model makes it a non-linear system. Because the classic Kalman Filter (KF) is designed for linear systems, the modified KFs have to be employed. Three typical filters are tested: the Extended KF (EKF), the Limiting EKF, and the unscented KF. The modeled network is a freeway with eight on-ramps and seven off-ramps, which means there is no route choice problem. Data of speed, density, and vehicle count are obtained from ten sensors, and used as observations. Vehicles count is offered at 15-min intervals; speed and density are offered at 1-min intervals. Two different strategies are adopted for calibration: mixed off-line / on-line approach and jointly on-line approach. In mixed off-line/on-line approach, an off-line calibration is performed to get speed-density relationship parameters, which is based on a nonlinear regression for the predefined function parameters. And only the demand parameters are calibrated on-line. However, the jointly on-line approach calibrates all the parameters by the state estimation method. Final results show that jointly on-line calibration offers better performance in terms of accuracy.

2.5.4 Summary

In this section, the works about calibrating CTM and calibrating traffic simulation models are reviewed.

For CTM calibration, several works use the measurement based method to "measure" the parameter values directly, especially the works at the early stage of CTM development. However, some optimization based methods are also applied to calibration of CTM. This idea matches the definition of model calibration better.

For the traffic simulation model calibration (also including network modeling cases), performance based model calibration are more extensively discussed. In order to improve the performance of developed model, the suitable values for parameters need to be found. To solve this problem, two methods are used: the optimization method, and Bayesian analysis method. In addition, the boundary between model parameters and model inputs are fading. Any factor, if it influences the model performance but is not strongly fixed or known, can be taken as calibrated variables. Under this condition, dimension of calibration problem is possible to be high. And, those problems are commonly nonlinear. Therefore, the used algorithm for solving them should be capable of dealing with nonlinear high-dimensional problem.

3 System requirements

In this thesis, we propose an approach to Hybrid Traffic Simulation, which is named "HyTran". Based on the literature review, this chapter proposes a basic description of HyTran, focusing on requirements for model integration.

3.1 What's new in HyTran

Generally speaking, there are two main tasks for the developed integration framework: first, to make the integrated models to cooperate in the same system; second, to build a strategy for sharing information between the integrated models effectively. As summarized in the last chapter, the existing research pays more attention to making the models able to run under a unified system, but less to exploring the strategy for sharing information. In other words, the existing research aims at transmitting information from one model to another, but not on how does one model utilize the information obtained from another model. This is summarized in the last chapter as flow representation transition, system integration, and model integration. In those three types of integration framework, if a sub-network is modeled by microscopic level, the simulation results or model performance obtained from the microscopic model will be used directly in the macroscopic level; therefore, the microscopic level run lasts for the entire simulation period in order to offer results continuously. This means the internal utilization of information is considered in the most intuitive and simplified way. In contrast to the existing research, HyTran tends to focus both on information transition and information utilization. In summary, the proposed framework evolves from the model integration to information integration by maintaining the advanced performance consistency that will be explained in chapter 5 .

The proposed framework in this chapter is named as "HyTran", which stands for "Hybrid Transport molding". Different to the preceding research, HyTran presents two options for integration framework. The first option is based on traffic flow representation transition, which is addressed as Flow Transition based Framework (FTF); and the second option is based on the idea of data fusion by performing parameter calibration. This framework is addressed as the Data Fusion based Framework (DFF).

Although we will pay some attention to DFF, it does not mean DFF is a better solution than FTF or we have any preference about using DFF. The main purpose of proposing HyTran is to offer a new possibility for Hybrid traffic flow modeling, and investigate its effectiveness and feasibility. Furthermore, we tend to investigate the difference of dynamic attributes between macroscopic model and microscopic model.

3.2 Starting point: the application-oriented discussion

HyTran tends to build an application-oriented framework. Therefore, before stating the considerations on the integration framework, we discuss the possible application scenarios in terms of the function of roads. A summary is shown in Table 1. The figure cited in this table comes from the website of TSS (Aimsun), the figures representing similar concepts are also done by other companies such as PTV Group and Caliper. This figure shows relationship among models with those different levels of resolution visually.

From the perspective of road functions, in a traffic network simulation project, the freeway networks and urban main roads are two main objects to be modeled. The main reason is that these two types

of roads attract and serve most of the traffic demands, in other words, the performance of those networks determines the traffic state significantly, and it is what the traffic engineering concerns. All the traffic engineering tasks, such as traffic assignment, evaluation of level of service (LOS), signal control plan design and evaluation, etc., have a direct relation with the traffic state on those networks. Because of the importance of those roads, it needs to be considered by any model, i.e., macroscopic model, mesoscopic model, and microscopic model. In terms of macroscopic model, to model these roads can be considered as one of the basic function of macroscopic model. Another basic function of macroscopic model is to work as the base of assignment model by offering basic data for a link cost function, which means macroscopic model must involve all the links that needs to be considered. If a link is not modeled in macroscopic level, assignment model will not be able to assign any traffic demand on it. At the meanwhile, if a link's traffic volume is low, it is not necessary to model it by microscopic model, because for these types of roads the traffic situation is relatively constant, and a macroscopic level estimation of traffic volume or travel time is already representative enough.

The main function of microscopic model is also to model freeway networks and urban main roads. However, limited by its building and computational complexity, it cannot be used to cover areas as large as a macroscopic or a mesoscopic model can. Generally speaking, two reasons give rise to an application of microscopic modeling: first, infrastructures or scenarios cannot be modeled sufficiently well by macroscopic model due to its complexity, such as actuated signal control, complex weaving behaviors, or pedestrian influence; second, traffic demand on this networks are close to the service capacity, therefore any modification on infrastructure, e.g. , change of signal plan, change of lane length, etc., leads to significant change of the traffic state. On the contrary, it can also be inferred from those two reasons that it is not necessary to model every part of a large network by microscopic models.

From the perspective of simulation model, mesoscopic models are a compromise between macroscopic models and microscopic models. In other words, the main purpose of this model is to simulate a larger area with affordable implementation cost. From this aspect, hybrid simulation and mesoscopic model can also be considered as two parallel solutions, which means macroscopic model and mesoscopic model are not necessary to exist at the same time. A hybrid model needs and only needs two main levels:

1) A low-resolution level (macroscopic model, or a mesoscopic model), which describes the entire network; it works as a skeleton of the simulated network.

2) A high-resolution level (microscopic model), which describes some key elements of the networks, i.e., complex intersections, weaving/merging/diverging segments on freeway, ramp metering, pedestrians. This level is used to enhance the accuracy and analysis ability of the entire system, by introducing a more detailed description of road elements, control strategy, and vehicle behaviors.

Table 1 Road functions and related simulation model

	Typical application	functions		Motorways (highways/ expressways/ freeways)	Urban roads	Country roads
			Traveler	Motor vehicles	Motor vehicles, Bicycle, pedestrians (street cross)	Motor vehicles, Bicycles, pedestrians
			Function	Connect cities; skeleton of a city	Connect areas in a city	Roads serve in country and connect the countries with other place
			features	No or seldom signal control; hard isolation between different direction; high speed limit; few route choices	Signal control, mixed traffic flow, low speed limit, more choices	No signal control, seldom lane divide, mixed traffic flow
Macroscopic model	City-wide road network simulation / analysis	Demand assignment; Aggregated behavior		× (complete)	× (complete)	× (if necessary)
Mesoscopic model	Motorway system simulation	Short time traffic state estimation; Traffic assignment		× (complete)	× (Main Roads only)	-
Microscopic model	Urban streets with signal control; highway network with ramps	Analysis of road design, signal plan, etc.; control strategy analysis		× (complete or only subarea)	× (subarea or partly)	-

3.3 Model selection

The design of model integration framework highly depends on the attributes of the specific models that are chosen to be integrated. Therefore, it is necessary to determine which models are used as early as possible. However, there is a risk that some potential problems are possible to emerge after the models have been determined. The risk is reduced by carefully review on existing research and the flexibility of the HyTran.

In the later part of this thesis, we try to discuss the general problems about continuum flow model and car following based microscopic model, however, when the detailed information is necessary for description or discussion, the attributes of the two selected models will be mentioned.

Selection of Macroscopic level

In HyTran, CTM is selected as the macroscopic level, and VISSIM (a car following based microscopic model) is selected as the microscopic level.

The related work about CTM has been reviewed in the last chapter. The reason for which the CTM is selected is based on three considerations:

1. The CTM derives from LWR models, which do not possess an essential conflicts with the microscopic models (this will be discussed in section 5.1);

2. The CTM is easy to be implemented because of its concise structure and principles, which makes it a flexible and feasible candidate for HyTran;

3. The CTM is a deeply discussed model. The research about it, from both theoretical and practical perspectives, is well reported.

Selection of Microscopic level

Although there are diverse microscopic traffic simulation models reported, the most used models come from the group of cooperation of car-following model and lane change model. In other words, car-following models are almost the only option for microscopic level. In addition, the mechanisms of microscopic models are much more complex than macroscopic models, especially when considering model coding and execution efficiency of the code. In order to overcome this disadvantage, we consider using a mature commercial software program as the microscopic level. The selected program has to meet the following requirements:

1. The program should be based on a well-developed and widely confirmed theoretical model. The feasibly of it has been proved and validated through practical implementations.

2. The program should have fully functions for user-defined secondary development; therefore HyTran can be realized based on it.

Considering those two requirements, VISSIM is selected as the basic component of microscopic level. VISSIM is a microscopic traffic simulation software tools. Its core algorithm is based on the cooperation of psycho-physical car following model (Wiedemann 74 and Wiedemann 99) and a rule-based algorithm for lateral movements (lane change model). The psycho-physical model is highly developed car-following model. Same to the general idea of car-following models, psycho-physical model also determines the moving state of the following car by considering its relation with its leading car. However, this model categorizes the drivers' behavior to four modes: free driving,

approaching, following, and braking. For each mode, specific mathematical description of driving behaviors is made. In addition, the random attributes of traffic participants, such as the individual characteristics of drivers, are also considered in the model. Specifically, Wiedemann 74 is suitable for urban traffic and merging / waving areas, and Wiedemann 99 is more suitable for freeway without merging/weaving areas. Wiedemann 74 has 3 parameters, and Wiedemann 99 has 9 parameters. In terms of lateral movement, VISSIM categorizes lateral movements into two types: necessary lane change and free lane change. Necessary lane change is aroused by a vehicle's assigned route, which a vehicle has to move to a proper lane in order to keep its route. Free lane change is aroused by better running condition in the adjacent lane that is attractive for a vehicle. For both two types, similar mechanism for lane change is adopted. A vehicle, which has desire of changing to the adjacent lane, has to check the distance with its potential trailing vehicle (in the target lane) and to decide whether to do or not. Changing the value of related parameters can show the different intensity of its desire.

Besides the two basic core models, VISSIM also considers the tactical driving behavior, including anticipated driving at conflict areas, cooperative emerging. These two points complete the description for driver behavior that are not covered by pure car-following model and lane-changing model.

In both longitudinal and lateral movement, random attributes of traffic participants are considered by introducing key parameters not by a specific value, but a desired distribution. This feature makes VISSIM a stochastic model.

4 Cell Transmission Model: discussion and preparation for low resolution level

The consistency problem is essential for multiple resolution modeling. In this chapter, we discuss the application of CTM for MRM. In order to improve the performance of model consistency, two points of enhancement are proposed on the original CTM. Those enhancements improve the performance of flow transition interface. The proposed method will be evaluated together with the boundary interface of HyTran.

4.1 Overview

This section presents the basic equations of CTM in detail.
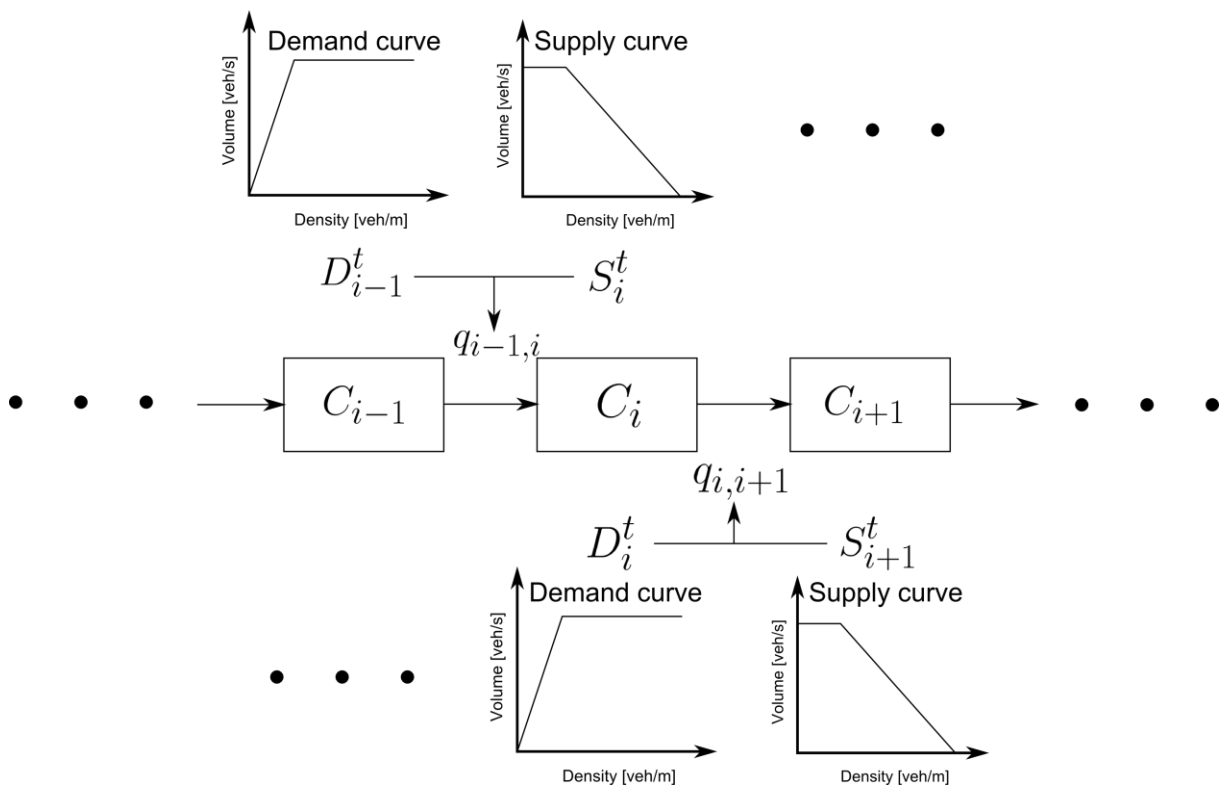
D_i^t is the demand volume rate of cell i during time interval t to $t + \Delta t$; S_i^t is the supply volume rate of cell i during time interval t to $t + \Delta t$; $q_{i,i+1}$ is the flow rate from cell i to cell $i + 1$ during time t to $t + \Delta t$; Δt is the simulation cycle;

Figure 8 Basic structure of CTM

As shown in D_i^t is the demand volume rate of cell i during time interval t to $t + \Delta t$; S_i^t is the supply volume rate of cell i during time interval t to $t + \Delta t$; $q_{i,i+1}$ is the flow rate from cell i to cell $i + 1$ during time t to $t + \Delta t$; Δt is the simulation cycle;

Figure 8, the traffic volume between two adjacent cells. The demand volume of upstream cell is:

$$D_i^t = \min\{v_i^t \cdot K_i^t, Q_{m,i}\}$$

And supply volume of downstream cell is:

$$S_{i+1}^t = \min\{\omega_{i+1}(K_{jam,i+1} - K_{i+1}^t), Q_{m,i+1}\}$$

Where, v_i^t is the free-flow speed; ω_i^t is the backward wave speed; K_i^t is the density of cell i; $Q_{m,i}^t$ is the capacity of cell i. The demand curve and supply curse come from predefined fundamental diagram (FD).

The traffic volume rate between cell i and cell $i + 1$ is:

$$q_{i,i+1}^t = \min\{D_i^t, S_{i+1}^t\}$$

Therefore, the vehicle density of a cell can be updated as:

$$K_i^{t+\Delta t} = K_i^t + q_{i-1,i}^t \cdot \frac{T}{l_i} - q_{i,i+1}^t \cdot \frac{T}{l_i}$$

Where, l_i is the length of cell i; T is the clock interval of the model. In classic CTM, the vehicle density within a cell is assumed to be evenly distributed, in other words, $k_i(t)$ represents the average vehicle density in cell i at time t.

Another important principle is about the relation among $v\ (v > \omega)$, T, and l. Based on the CFL condition of Godunov's method (Leclercq, 2007), the following condition must be satisfied for all cells, in order to maintain the stability of the numerical method (CFL condition, equation 4-1):

$$l \geq vT \qquad\qquad\qquad \textbf{4-1}$$

This condition can be explained as within a single time step, vehicles can at most move to the next cell, e.g. if three adjacent cells are marked as cell i, cell $i + 1$, and cell $i + 2$, within a single time step, vehicles that are located in cell i can move at most to cell $i + 1$, but not cell $i + 2$.

Based on the definition for normal cells, we use the methods reviewed in section 2.4.3 for diverting cell, merging cell, and intersections; we use the methods reviewed in section 2.4.4 for signal modeling; we use the methods reviewed in section 2.4.5 for modeling permitted left-turn.

Based on the contribution of Gomes (2008), the cells do not have to share same FDs for getting stable solutions. Based on the summary of LeVeque (1992), this principle is capable of handling diverse forms of FD, including concave one, non-concave one (with a convex tail for the congested part), piecewise linear one.

In this chapter, two enhancements on original CTM are proposed, including slope-CTM and Temporal Fundamental Diagram (TFD). In summary, the enhanced CTM is named "IFCTM", which stands for CTM for Interrupted Flow.

4.2   Original points for IFCTM

4.2.1   Cell length problem

As mentioned in section 2, several extensions have been discussed for improving the analysis ability of CTM. However, no attention has been paid on the accuracy problem introduced by cell length.

The cell length is determined the original definition of CTM, which equals the distance at which a vehicle can travel during one simulation cycle at free flow speed. As an application of Godunov's method, CTM determines cell length by the predefined free flow speed in fundamental diagram. This principle assures that the method is numerically stable, and also makes the model to take full advantage of the accuracy that can be anticipated on Godunov's method.

If a cell length that is shorter than the defined length, CFL condition (equation 4-1) is violated. Theoretically speaking, this cell length will break the convergence of Godunov's method. From the

traffic engineering's point of view, if a shorter cell is used, within a single simulation cycle, a vehicle is possible to skip across a cell without staying in it, but this phenomenon cannot be described by CTM.

If a cell that is longer than the defined length, it does not destroy the convergence of Godunov's method as the shorter cell dose. However, the viscosity phenomenon will emerge. This phenomenon can be explained with a numerical example with FD shown in Figure 9. From the numerical perspective, viscosity can be shown by the results of an impulse with the scalar conservation problem. In this numerical example, we use a FD in which: $v = \omega = 1$, $q_m = 1$, $k_c = 1$, $k_{jam} = 2$.
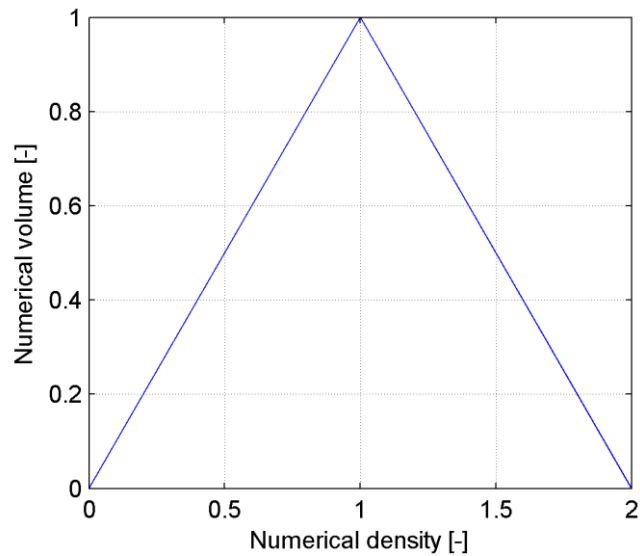
Figure 9 Idealized FD for numerical example

Where: $v = \omega = 1$, $k_c = 1$, $k_{jam} = 2$, q_m=1

The impulse input and constant input solution by CTM with cell lengths equal to 1 and 1.5 are shown in Figure 10. All examples run at time interval of 1s. For the impulse input, the initial state of the first cell is set to $1/l_{cell}$. When $l_{cell} = 1$, the impulse does not smear out during the propagation (sub figure (b)); when $l_{cell} = 1.5$, the impulse peak smears out during the propagation (sub figure (b)). For the constant input, not only the initial state is set to be same to impulse, but for the entire simulation process, for each cycle, we input 1 unit "vehicle" into the first cell. As shown in sub figure (d), when $l_{cell} = 1$, the wave front keeps sharp for the entire simulation period, however, when $l_{cell} = 1.5$, the front smears out. Sub figure (e) and (f) show the accumulated volume passing the point x=4 during the entire simulation process. For the impulse input situation, when $l_{cell} = 1.5$, the accumulated volume line starts to rise up earlier but reaches the final stable value (1) later; for the constant input situation, when $l_{cell} = 1.5$, the line also starts to rise earlier, but it tends to maintain a constant error with the line obtained by $l_{cell} = 1$. This means the relative error reduces as the total volume grows.

The same phenomenon also can be observed when simulating a congested flow with queue formation. In addition, even if the viscosity phenomenon can be avoid by adopting a cell length that equals vt_c, the same viscosity phenomenon cannot be avoid with the queue forming situation. The main reason is that for a traffic flow, backward wave speed is always slower than free flow speed. For example, according to (Lin & Ahanotu, 1995), a reasonable value of backward wave speed should be estimated between $\left[\frac{v}{6}, \frac{v}{2}\right]$. This relation means that vt_c is the minimum acceptable cell length for applications. The lagged CTM improves the accuracy by debilitating the viscosity phenomenon for backward wave propagation. The main idea is to use the density values several simulation cycles

earlier (Daganzo, 1999). However, this approach does not guarantee a non-negative density and a density that is less than the predefined jam density (Szeto, 2008). Because the traffic flow rate calculated by the lagged density may be higher than the maximum flow rate that can offered by the cell at the current time step.

Table 2 lists several examples of practical application of CTM. E1 and E5 follows the basic definition of CTM, using $l_{cell} = vT$; E1 use $\omega < v$, but E5 uses $\omega = v$. E2 and E3 uses adjustable cell length by applying enhancement on original CTM, and $v < \omega$; E3 use principles same to original CTM, but using $l_{cell} > vT$ Those examples show several common situations in CTM application. First, when applying original CTM, $v < \omega$ is widely adopted, this decision also agree with its basic definition (especially the definition of ω). The ω related viscosity is generally ignored; second, some of the application tend to have cells that are much more longer than the minimum cell length, if this goal is possessed, extra enhancements on CTM are made. Third, although the viscosity phenomenon exists, the influence of it is not fatal. In addition, considering the practical modeling work, it is difficult to make all cell length equal vT. Generally, a cell length that is slightly longer than vT is fairly acceptable for application. In the current thesis, if necessary, we allow cell length that is not longer than 1.5 times of the minimum cell length.
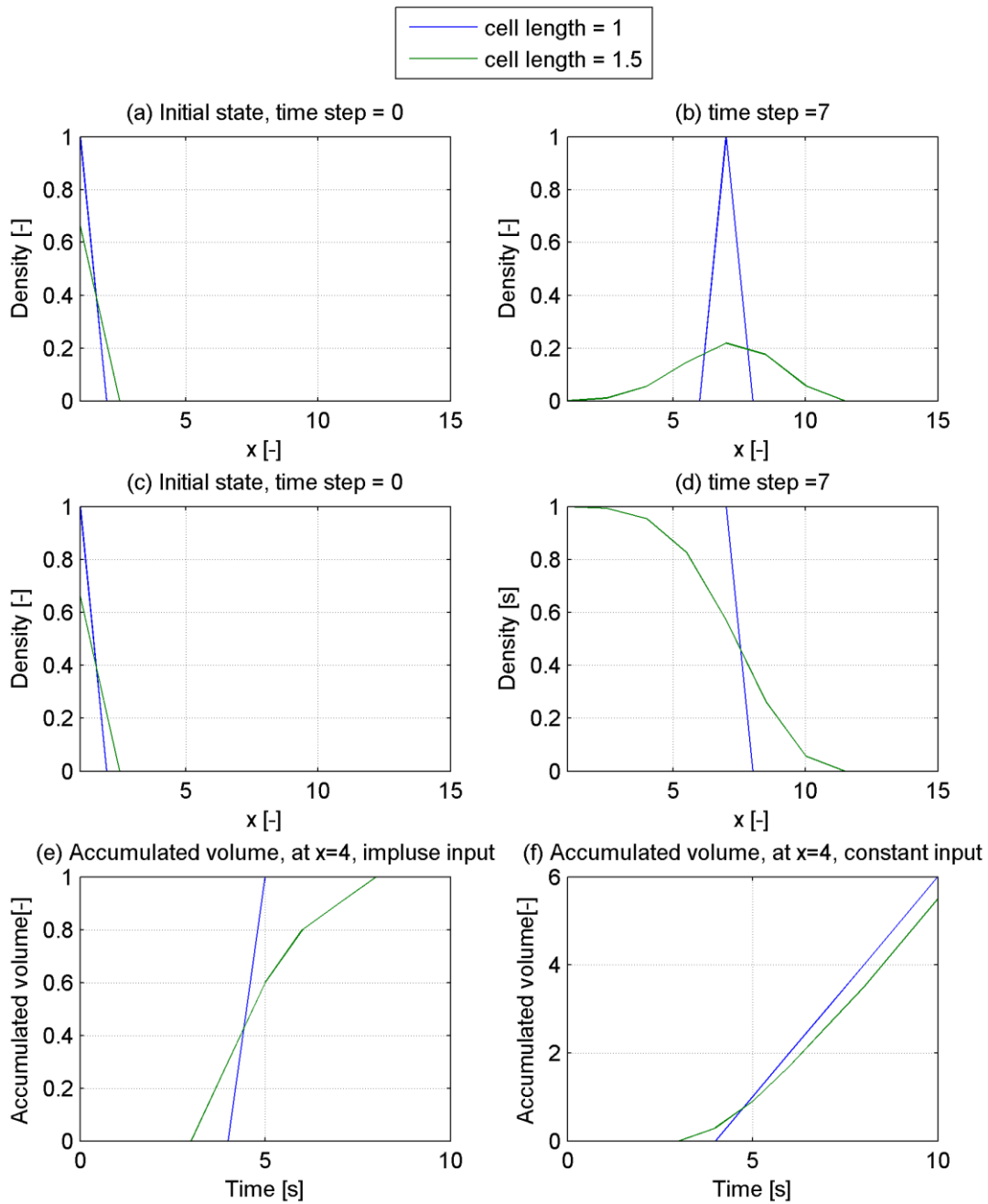
Figure 10 Viscosity for free flow

(a) Initial state for an impulse input; (b) density state at 7th step corresponding to (a); (c) Initial state for a constant input; (d) density state at 7th step corresponding to (c); (e) Accumulated volume at x=4 for impulse input; (f) accumulated volume at x=4 for constant input;

Table 2 CTM application examples: from the perspective of cell length determination

	E1	E2	E3	E4	E5
Example	(Lin & Ahanotu, 1995)	(Lee, 1996)	(Munoz, et al., 2006)	(Ishak, 2006)	(Pohlmann, 2010)
Object	freeway	Urban roads / freeway	freeway	freeway	Urban roads

Cell length type	Fixed	Adjustable by model enhancement	Adjustable	Adjustable by model enhancement	Fixed, determined by free flow speed
Cell length [m]	60 (determined by $v \cdot T$)	adjustable	293 (min) 551(mean) 1224 (max)	Diverse, minimum value about 106	13.89/8.33 (determined by vt_c)
Capacity [veh/h/lane]	1900	1800	Diverse, based on calibration (1488~2300)	2250	1800
Free flow speed [m/s]	29.2	11.2/15.7/20.1	Diverse, based on specific calibration (26.8~27.7)	27.8	13.89/8.33
Backward wave speed [m/s]	4.44~6.11	-	Diverse, based on specific calibration (5.8~8.8)	13.9	=free flow speed
Simulation cycle [s]	2	2	10	4	1
Jam density [veh/m/lane]	0.12	-	Diverse, based on specific calibration (0.089~0.117)	diverse	0.16667

4.2.2 Assumption of evenly distributed vehicle density

From the simulation perspective, CTM depends on a basic assumption that vehicles inside a cell are even distributed. This point may be considered as a reason for which the cell length brings viscosity problem. However, this is a long-time ignored weak point of CTM. One possible reason is for the original Godunov's method, the accuracy of the algorithm can be improved by using shorter space interval, associated with shorter simulation cycle (satisfying the CFL condition). Because when the cell length becomes infinitesimal, the results can be guaranteed to converge to the exact solution. No matter viscosity exists or not, it does not affect the algorithm's convergence. However, in traffic simulation, the adopted cell length cannot be too short.

$$l_{cell} : with\ avg.\ density\ k$$

$$vT : with\ avg.\ density\ k^d$$

$$\omega T : with\ avg.\ density\ k^u$$
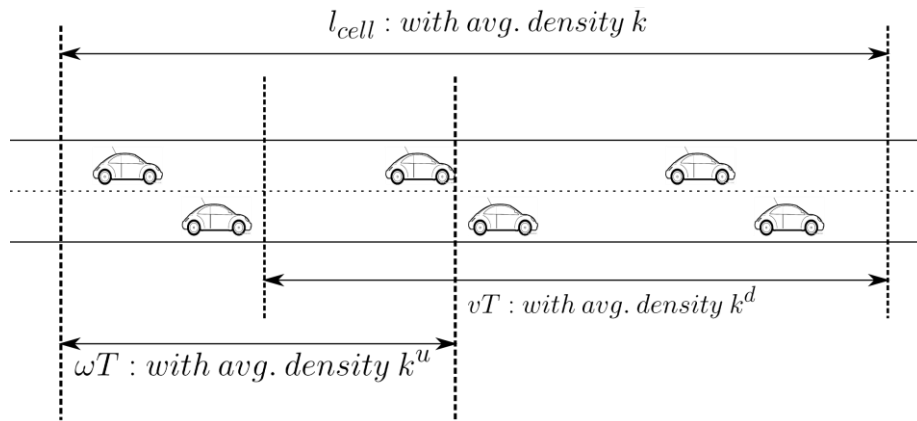
$$Stable\ situation : k^u \approx \bar{k} \approx k^d$$

Figure 11 Evenly vehicle density

$$l_{cell} : with\ avg.\ density\ \bar{k}$$

$$vT : with\ avg.\ density\ k^d$$

$$\omega T : with\ avg.\ density\ k^u$$
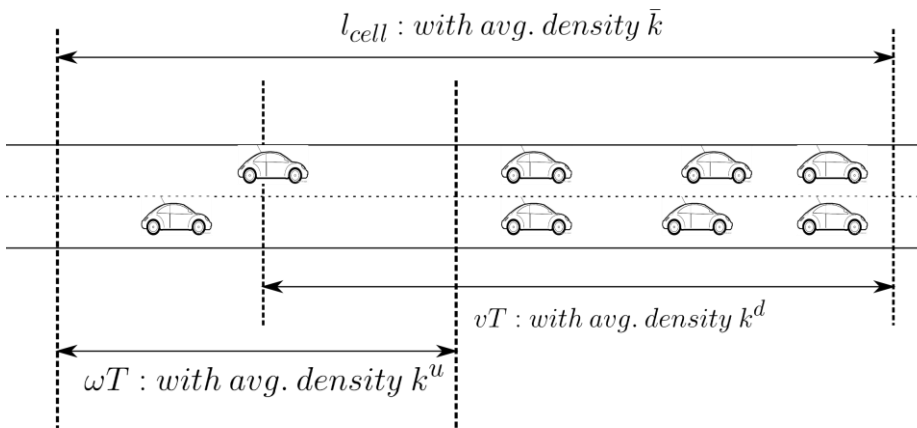
$$Queue\ forming : k^u < \bar{k} < k^d$$

Figure 12 Unevenly vehicle density for queue forming

$$l_{cell} : with\ avg.\ density\ \bar{k}$$

$$vT : with\ avg.\ density\ k^d$$

$$\omega T : with\ avg.\ density\ k^u$$
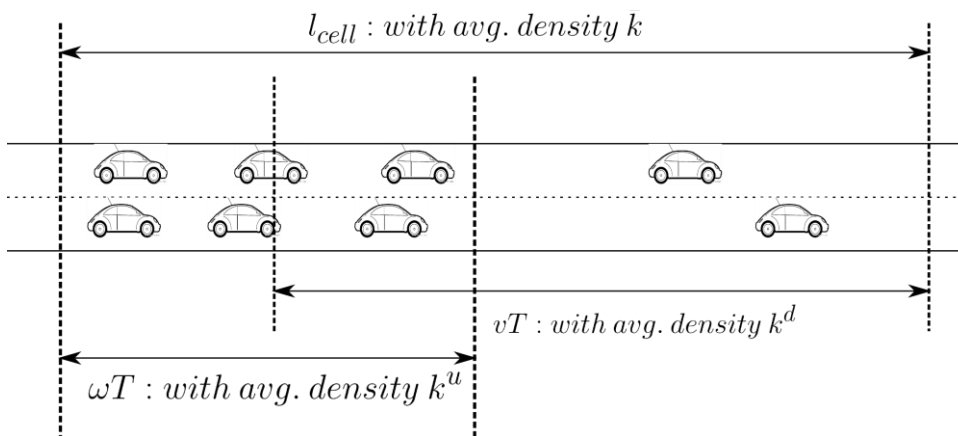
$$Queue\ discharging : k^u > \bar{k} > k^d$$

Figure 13 Unevenly vehicle density in queue discharging

The problem about even distribution assumption is shown by Figure 11, Figure 12, and Figure 13. In order to explain this problem clearly, the cell length used here is slightly longer than the minimum cell length vT. Figure 11 shows an idealized situation. When the modeled traffic flow is stable, the even distribution assumption is reasonable. However, for the situations shown in Figure 12 and

49

Figure 13, the assumption looks unreasonable. Figure 12 shows the unevenly density distribution, which can emerge in a queue forming situation. This situation is frequently observed on both freeway link and urban link, although maybe aroused by different reasons. For example, when a queue is forming at a stop line of an intersection, the vehicle density close to the stop line will change to the jam density at first. As time passing, the queue (jam density) propagates upstream. If we consider the supply traffic volume for this cell, the realistic traffic volume rate is larger than the value calculated using the average density ($\bar{k} > k^u$). Figure 13 shows the unevenly density distribution in a queue discharging situation. When the queue begins to discharge, the discharging phenomenon starts with the downstream part of the queue, and propagates backward to the upstream of the road. During the process, the vehicle located at the downstream end of the cell start up first,

However, in the application of traffic simulation, the most valuable performance of the model is evaluated by the accumulated traffic volume during specific periods. In this example, if we count the volume at any location of the link during long enough time periods, we should get the same results from the two models with different cell lengths, especially when the demand volume is fixed and feasible. This is the main reason for which multiple cell lengths are employed in those cases.

4.2.3 Consideration on interrupted flow

Cell capacity for interrupted flow

Another problem that does not catch much attention is the cell capacity calibration for interrupted flow. A frequent example of interrupted flow is the flow passing a signalized intersection.

In original CTM, when a queue forms before a stop line, the vehicle density of this cell goes into the congested regime. Based on the principle of Godunov's method, the output flow rate, when traffic light turns from red to green, will be the saturation flow. This is to say, the capacity of the critical cell that is just located before the stop line affects the simulation results significantly, especially when the effective green time is relative short, with which it is possible that the cell's density are in congested regime for the entire green time.

A data group based on simulation results can explain this phenomenon. A simplified signal control is modeled. As shown in Figure 14, the signal cycle is 90 seconds, and the displayed green time varies from 5 seconds to 30 seconds. The capacity of the intersection varies with the green displayed green time. In most of CTM applications, the research tend to use a calibrated value of capacity (for a specific green time that is already decided before) in order to ensure the average output flow during modeled effective green time fits the realistic data (here, the microscopic simulation data). However, as seen in Figure 14, the practical cell capacity value varies with green time. This weak point impairs the dependability of CTM when it is used for signal plan optimization. As shown in Figure 14, the cell capacity can be chosen as $0.65\ veh/s/l$, in order to match the maximum observed value of volume rate (as the definition of this parameter), however, this value does not work with the situation when green time is 5 seconds, because the average maximum volume rate of it ($0.6 veh/s/lane$) is 8% smaller than the selected value. In the realistic situation, if facing short green time, the drivers have shorter startup loss time and bigger acceleration; on the contrary, as the green time increasing, the drivers have longer startup loss time and smaller acceleration. This explains why there is a maximum value for the capacity, as the green time changing.
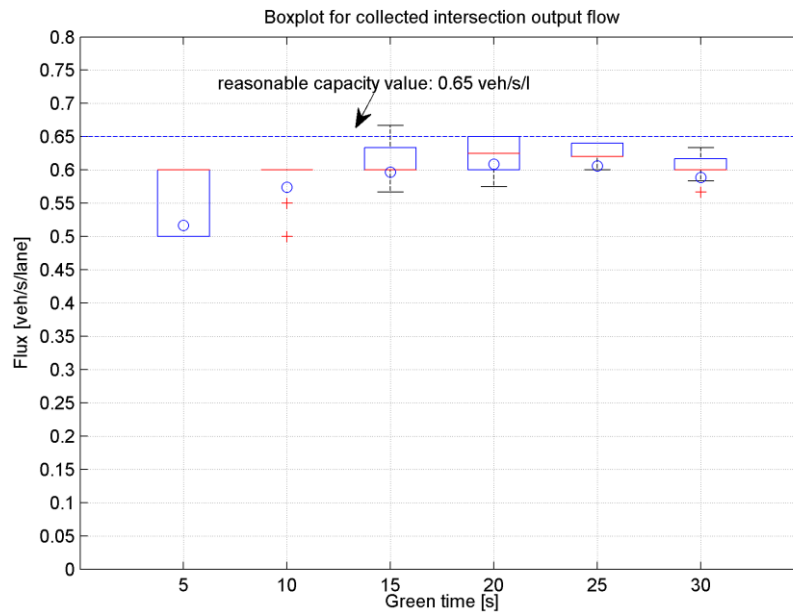
Figure 14 Intersection capacity data obtained from VISSIM model

from a two lane idealized signal control example with default parameters of VISSIM

Temporal fundamental diagram

Generally speaking, CTM is designed for stable flow. One of reasons is that CTM adopts constant FD for each cell. When it is used for modeling freeway with relatively stable flow, the influence is weak and may be ignored. However, when it is used to modeling a highly interrupted flow, such as a flow on urban street that is highly interrupted by signal controls, or considered for hybrid simulation application which needs to maintain the consistency between microscopic level traffic flow and macroscopic level traffic flow, this problem becomes non-negligible. Although the applicability of different FD is still pending, there are at least two situations under which the common FD used in CTM does not match the real.

For vehicles start up situation, the FD used by CTM cannot describe their macroscopic performance. A short example can be used to show this problem. For a startup flow, the initial average speed of the vehicles can be calculated as q_m/k_{jam} in CTM. When the cell capacity is set to $0.6\ veh/s/l$, and the jam density is set to $0.2\ veh/m/l$ which means the average speed of first second of startup is $3\ m/s$, which means the startup acceleration is as high as $6\ m/s^2$. This value exceeds a reasonable value for the average acceleration of vehicle starting up.

For the queue forming process, vehicles tend to move in a low speed that the normal FD predicted, because drivers can see the downstream jam and decelerate earlier. This problem has been noticed by the works of measuring backward wave speed, in which they try to eliminate the data points which come from the queue forming process, but only use the data points that come from the queue discharging process (Lin & Ahanotu, 1995) (Munoz, et al., 2006). In addition the non-concave FD that has a convex tail at the right end of a FD is a possible solution for this problem, but it introduces a mixture of fan wave and shock wave, which destroys its consistence with the continuum flow theory.

In summary, for CTM, a temporal FD, which allows different backward wave speeds for different scenarios, appears more flexible and dependable.

4.2.4 Summary

The cell length problem and evenly distributed density assumption influence the performance of CTM. This influence is stronger when the traffic flow is interrupted deeper. However, if the evaluation interval increases relatively longer, this influence becomes weaker.

A more promising enhancement is on the interface between microscopic flow and CTM (the macroscopic flow), which is a core function of the Flow Transition based Framework (FTF).

4.3 Introduction of the slope-CTM

4.3.1 The theoretical consideration for unevenly vehicle density distribution

LeVeque (1992) summarizes using slope-limiter methods to improve the accuracy of Godunov's method. The basic idea of it is to implement an assumption for density distribution in cells, other than assuming an evenly distributed density within them. The feasibility of the anticipated improvement is based on two basic assumptions:

1) $l_{cell} \geq vT > \omega T$, the discrete interval is longer than the length a wave can travel during one clock interval; otherwise the developed formulations reduce to typical Godunov's method;

2) The new density distribution is more realistic than the evenly distributed density assumption, in other words, the density distribution is possible to be estimated reasonably.

Those two assumptions can be satisfied by the application of traffic flow modeling, in particular for the highly interrupted flow. The first assumption can be satisfied by the practical requirements on CTM application; the second assumption is satisfied by the queue forming and discharging phenomenon when the flow is interrupted, e.g. affected by signal sings, etc. In this thesis, we assume $l_{cell} > vT$.

In this section, we introduce the numerical representation of Slope-CTM, based on the high resolution method (slope-limiter methods) on solving Riemann problems.

Because a piecewise linear fundamental diagram is always adopted in CTM, we discusses the slope-CTM from two perspectives: without nonlinear points and with nonlinear points.

Linear situation

Let $k(x, t)$ be the density of point x at time t, therefore, the traffic volume at point $x_{i+1/2}$ can be formulated as:

$$Q\left(x_{i+\frac{1}{2}}, t_{n,n+1}\right) = \int_{t_n}^{t_{n+1}} f(k(x_{i+\frac{1}{2}}, t)) \, dt$$

Where, $q(x, t) = f\left(k(x, t)\right)$, is the relation between q-k; n represent the index of time intervals.

Considering the situation of shock wave in Figure 15, associated with the triangular fundamental diagram. We assume all the cell average densities ($K_i^n, K_{i-1}^n, K_{i+1}^n$) are at congested regime. For classic Godunov's method, during the time period $t_n \rightarrow t_{n+1}$, we have:

$$k\left(x_{i+\frac{1}{2}}, t\right) = K_{i+1}^n$$

The traffic volume rate at the cell boundary between cell i and cell $i + 1$ is formulated as:

$$q = \omega(K_j - K_{i+1}^n)$$

Cell Transmission Model: discussion and preparation for low resolution level

However, if the realistic density distribution inner cell $i + 1$ can be formulated as:

$$k(x,t) = K_{i+1}^n + \sigma_{i+1}^n(x - x_{i+1})$$

Where, $x \in \left[-\frac{\epsilon}{2}, \frac{\epsilon}{2}\right]$ (in CTM, ϵ is the cell lengh), σ_{i+n}^n is the slope of the density distribution along position. Therefore, for the cell boundary (cell i and cell $i + 1$), the vehicle density can be expressed by:

$$k\left(x_{i+\frac{1}{2}}, t\right) = K_{i+1}^n + \sigma_{i+1}^n\left(\omega t - \frac{\epsilon}{2}\right) \quad t \in [t_n, t_{n+1}]$$

As we assume that $k(x,t)$ is limited under the congested flow regime, we have:

$$f(k) = \omega(k_j - k)$$

Therefore,

$$Q\left(x_{i+\frac{1}{2}}, t_{n,n+1}\right) = \int_0^h \omega \cdot \left\{k_j - \left[K_{i+1}^n + \sigma_{i+1}^n\left(\omega t - \frac{\epsilon}{2}\right)\right]\right\} dt$$

$$\Rightarrow Q\left(x_{i+\frac{1}{2}}, t_{n,n+1}\right) = h\omega k_j - h\omega K_{i+1}^n + \frac{1}{2}h\omega\epsilon\sigma_{i+1}^n - \frac{1}{2}h^2\omega^2\sigma_{i+1}^n$$

It is obvious that, if $\omega h = \epsilon$ (h is the time interval), the equation reduces to:

$$\hat{Q}\left(x_{i+\frac{1}{2}}, t_{n,n+1}\right) = h\omega(k_j - K_{i+1}^n)$$

It equals to the classic Godunov's method. The difference between those two algorithms is:

$$e = Q\left(x_{i+\frac{1}{2}}, t_{n,n+1}\right) - \hat{Q}\left(x_{i+\frac{1}{2}}, t_{n,n+1}\right) = \frac{\omega\sigma_{i+1}^n h}{2}\left(\frac{\omega h}{\epsilon} - 1\right) \qquad \textbf{4-2}$$

Figure 16 shows another linear situation, the forward wave. In this situation, the density profile is assumed to be only within the free flow regime.

Base on similar deduce process, the results of $Q(x_{i+1/2})$ can be formulated by:

$$k\left(x_{i+\frac{1}{2}}, t\right) = K_i^n + \sigma_i^n\left(\frac{\epsilon}{2} - vt\right)$$

$$Q\left(x_{i+\frac{1}{2}}, t_{n,n+1}\right) = \int_0^h v\left[K_i^n + \sigma_i^n\left(\frac{\epsilon}{2} - vt\right)\right] dt$$

$$\Rightarrow Q\left(x_{i+\frac{1}{2}}, t_{n,n+1}\right) = vhK_i^n + \frac{1}{2}\sigma_i^n vh\epsilon - \frac{1}{2}\sigma_i^n v^2 h^2$$

The difference between classic Godunov's method and the stated method is:

$$\hat{Q}\left(x_{i-\frac{1}{2}}, t_{n,n+1}\right) = vhK_{i-1}^n$$

And,

$$e = Q\left(x_{i+\frac{1}{2}}, t_{n,n+1}\right) - \hat{Q}\left(x_{i-\frac{1}{2}}, t_{n,n+1}\right) = \frac{vh\sigma_i^n\epsilon}{2}\left(1 - \frac{vh}{\epsilon}\right) \qquad \textbf{4-3}$$

Equation 4-2 and 4-3 show the difference between original Godunov's method and the slope-limited method. It is obvious that if $v = \epsilon$ ($\omega = \epsilon$), the difference e reduce to zero.
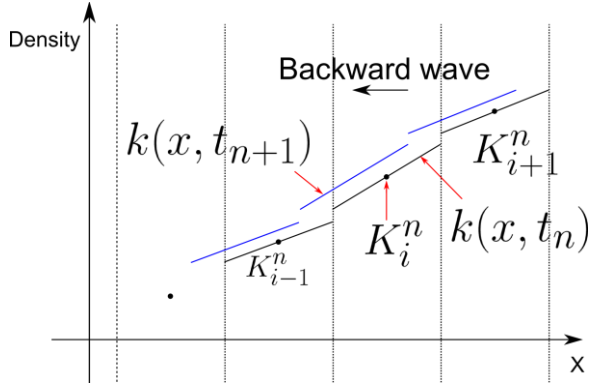
Figure 15 Backward wave under linear situation
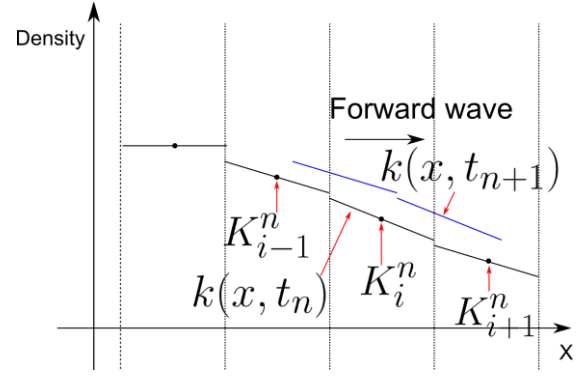
Figure 16 Forward wave under liner situation
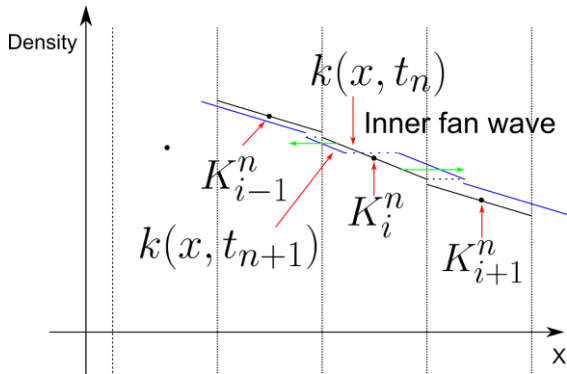
Figure 17 Nonlinear situation (1): inner fan wave
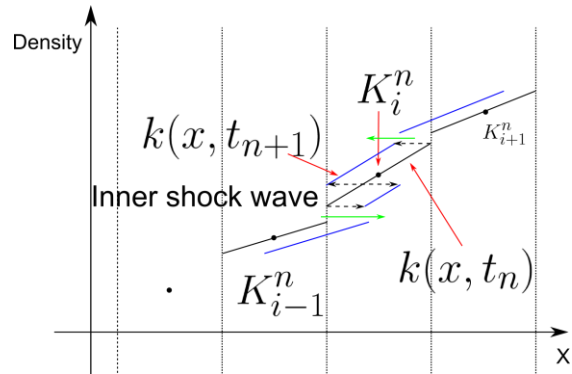
Figure 18 Nonlinear situation (2): inner shock wave

Non-linear situation

The situation of non-linear situation is more complex, because of nonlinear points included in fundamental diagram of traffic flow.

When the vehicle distribution is assumed to be unevenly, it is possible that the redistributed density covers both free-flow regime and congested regime within a single cell. As shown in Figure 17 and Figure 18, when a fan wave or shock wave exists inside a cell, it is possible for the front of the inner wave to reach the cell boundary. Because the slope-limited method, which will be discussed later, is based on the exact solution of conservation equations, and it depends on the linear attributes of conservation law problem, the inner fan wave or shock wave problem cannot be solved directly. therefore, we have to avoid the inner fan/shock wave situations.

One of the possible solutions to avoid the inner fan/shock wave situation is to introduce a variable clock interval. The clock interval (T) should be limited by:

$$T \leq \min\left(\frac{l_{congested}}{\omega}, \frac{l_{free}}{v}\right)$$

Where, $l_{congested}$ is the length of the part which has congested density; l_{free} is the length of the part which has free regime density. It is obviously that $T \leq \frac{l_{cell}}{v}$ can be guaranteed by the limitation above. And also, $l_{cell} = l_{congetted} + l_{free}$.

When the variable time interval limitation is used, the solution that is same to the last section is effective.

Another more effective and concise method which is adopted in this thesis is to avoid both inner fan wave and inner shock wave during the density re-distribution process. In other words, when

54

obtaining the re-distributed vehicle density, the employed algorithm does not allow any cell has a density distribution that covers two different regimes.

4.3.2 Structure of slope CTM

The Slope-CTM is made by add an external step into the original cell density update process, which is called "density re-distribution". The purpose of this step is to estimate the inner cell density distribution according to the current state. As described in the theoretical discussion before, it is obvious that the benefit of Slope-CTM proposed here is more effective when dealing with queue forming and discharging, but cannot appear when a stable flow is modeled. In other words, If there is only stable flow in the network, the Slope-CTM will offer the results same to CTM.

In urban roads, the queue forming and discharging phenomenon happens most likely before stop lines of intersections, especially when the intersection is signalized. In freeway system, a queue is raised by traffic jams, which result from temporal blockages or insufficient capacity.

Figure 19 shows the basic flow chart of Slope-CTM. In this diagram, two new steps are added to the basic CTM update process: slope determination and density re-distribution.
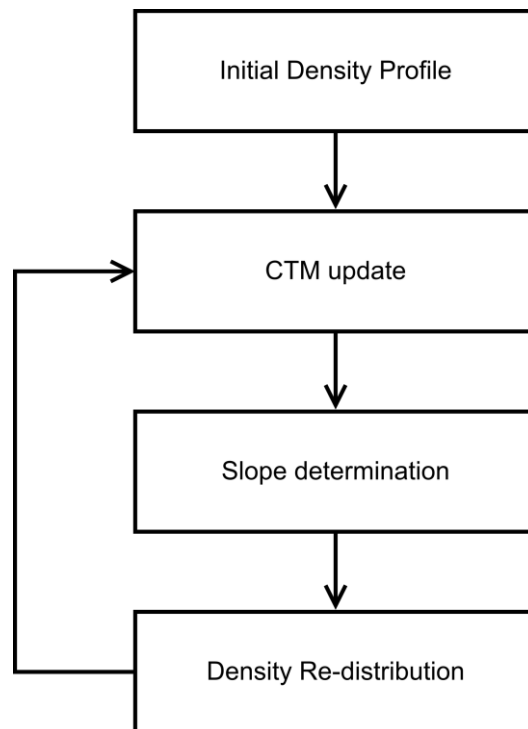
Figure 19 Slope-CTM cell updates progress

The algorithm for slope-CTM is:

Algorithm 1 Slope-CTM

Algorithm: Slope-CTM

Density Initialization: $\{K_i^n | 1 \leq i \leq N\}$
for $t \leftarrow 1$ **to** n **do**
  | *Calculate* $\tilde{k}(x, t_n)$ *based on* $\{K_i^n | 1 \leq i \leq N\}$
  | *Calculate* $\{Q_i^n | 1 \leq i \leq N\}$ *based on* $\tilde{k}(x, t_n)$
  | *Calculate average cell density* $\{K_i^{n+1} | 1 \leq i \leq N\}$

In this algorithm, the density distribution $\tilde{k}(x, t_n)$ is the estimated density, which is based on average density $\{K_j^n | 1 \leq j \leq N\}$, where N is the number of cells. After obtaining $\tilde{k}(x, t_n)$, volumes on the boundary $\{Q_j^n | 1 \leq j \leq N\}$ is calculated, where, Q_j^n is the volume flow out of cell j.

The re-distributed density $\tilde{k}(x, t_n)$ can be obtained by any designed method. Most commonly, a continuous linear function is used.

4.3.3 Slope determination

Theoretically speaking, the slope can happen when the densities of adjacent cells are different. If $K_{i-1}^n > K_i^n > K_{i+1}^n$, cell i tends to have a decreasing density as shown in Figure 20 ; if $K_{i-1}^n < K_i^n < K_{i+1}^n$, cell i tends to have an increasing density as shown in Figure 21 .
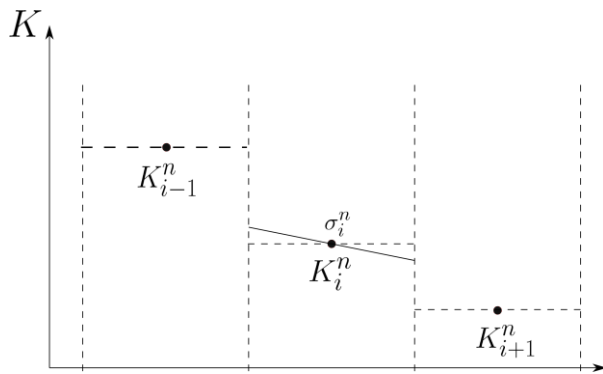
Figure 20 Decreasing density distribution
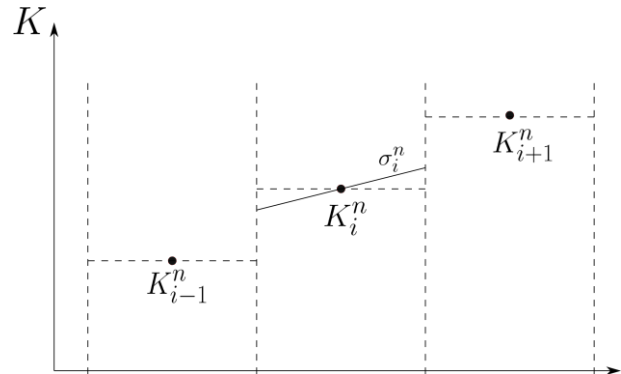
Figure 21 Increasing density distribution

However, when it applies to traffic flow, it has to be assumed that the vehicle density can be estimated. In addition, the fundamental diagram used for traffic flow has its own attributes, which makes some of the slopes do not have helpful influence on the results.

We describe the principles applied to Slope-CTM as following:

For CTM, whether a cell (e.g., C_i) has internal density slope is determined by the average density profile of its upstream and downstream cell ($\{K_{i-1}, K_i, K_{i+1}\}$). In order to keep the conciseness of CTM, a simplified method is used to judge it. C_i has inside density slope if and only if all of the following conditions are matched:

$$\begin{cases} K_{i-1} \geq K_{i-1,c} & (a) \\ K_i \geq K_{i,c} & (b) \\ K_{i+1} < K_{i+1,c} & (c) \\ K_{i-1} > K_i & (d) \end{cases}$$

4-4

Condition (a) and (b) mean the upstream and current cell have to be under the congested regime; condition (c) means the downstream cell has to be under the free regime; condition (d) means the average density of the upstream cell must be larger than that of the current cell.

Therefore, the slope only happens for the queue dissipation process as shown in Figure 22.

l_{cell} : with avg. density $\bar{k}$

vT : with avg. density k^d

ωT : with avg. density k^u $k_u > \bar{k}$
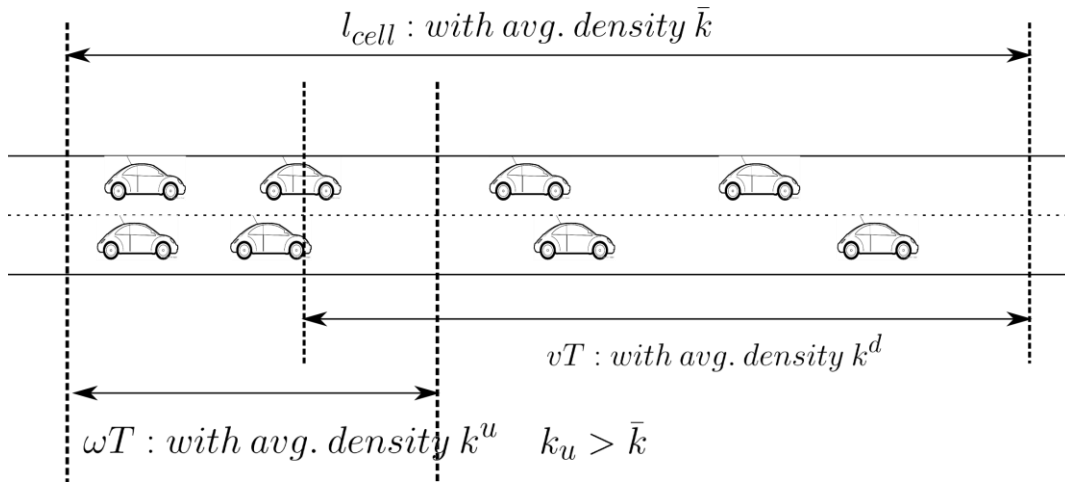
Figure 22 The realistic situation:

upstream density is significantly higher than average density, k^u leads to lower income volume than $\bar{k}$

The slope condition described in equation 4-4 can be shown in Figure 23.
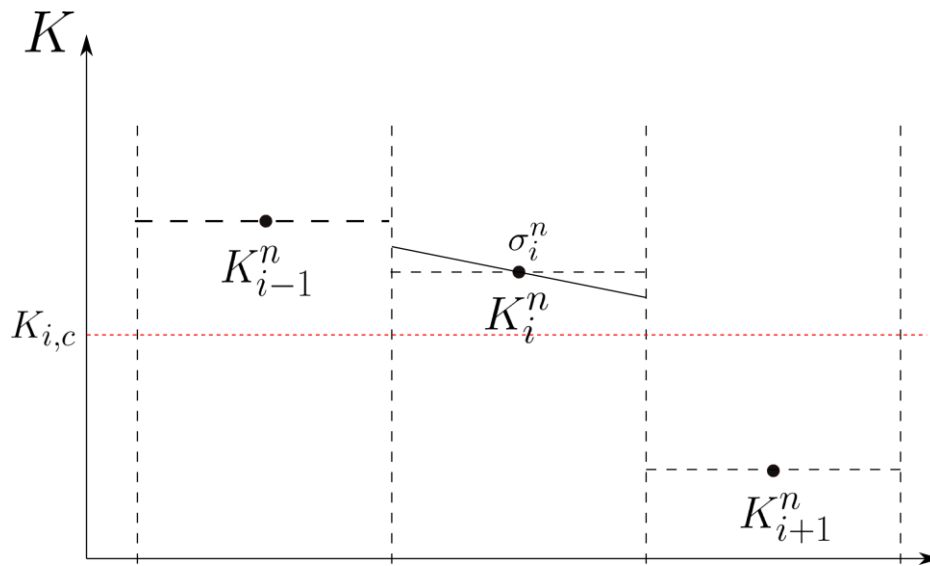
Figure 23 Recognized scenario for density redistribution for cell i

4.3.4 Density re-distribution

The density re-distribution process generates the estimated inside vehicle density for specific cell. In order to ensure the method calculates the results converge to the exact solution, the density re-distribution process should be Total Variance Demining method (TVD method).

$$TV\left(\tilde{k}(x, t_n)\right) \leq TV\big(K(t_n)\big) \qquad \textbf{4-5}$$

We cite the following theorem from (LeVeque, 1992):

"If the condition 4-5 is satisfied in the step of calculating $\tilde{k}(x,t_n)$ of algorithm of slope limitation method, then the method is TVD for scalar conservation laws."

Based on this consideration, the density re-distribution should satisfy condition 4-5, which means the re-distribution process is Total Variance Demining.

The minmod slope is a possible method to be adopted in this application. The minmod slope can be described in equation

$$\sigma_i = \frac{1}{\epsilon} minmod(K_{i+1} - K_i, K_i - K_{i-1}) \tag{4-6}$$

Where, ϵ is the cell lengh, and

$$minmod(a,b) = \begin{cases} a & if \ |a| < |b| \ and \ ab > 0 \\ b & if \ |b| < |a| \ and \ ab > 0 \\ 0 & if \ ab \leq 0 \end{cases} \tag{4-7}$$

Under situation 4-4, equation 4-6 reduces to:

$$\sigma_i = \frac{1}{\epsilon} \max\{(K_{i+1} - K_i), (K_i - K_{i-1})\} \tag{4-8}$$

When a slope is considered by the CTM update process, two parameters have to be calculated: equivalent upstream density k^u and equivalent downstream density k^d.

$$k^u = K_i^n - \frac{\sigma_i^n}{2}(\epsilon - \omega) \tag{4-9}$$

$$k^d \geq K_i^c \tag{4-10}$$

The expression of k_d is simplified because all the densities $k \geq K_i^c$ have a same demand curve based on the basic setting of CTM.

Another method for determining the slope is based on the regime limitation. In this method, the slope can be described as:

$$\sigma_i^n = \max\left\{\frac{2}{\epsilon}(K_i^n - K_{i-1}^n), \frac{2}{\epsilon}(K_{i,c}^n - K_i^n)\right\} \tag{4-11}$$

Based on this slope, k_u and k_d still be calculated from equation 4-9 and 4-10.

One point worth noting is that during this process, the demand volume from the current cell does not change, because its regime does not change.

4.4 Introduction of the Temporal Fundamental Diagram (TFD)

As discussed before, the slope-CTM does not improve the output volume of startup flow, because of the regime limitation. However, a more nature and effective perspective can be used to solve this problem: the temporal fundamental diagram (TFD). Generally speaking, TFD is a group of FD that designed for specific scenarios. TFD also comes from the consideration on interrupted flow.

4.4.1 FD for startup flow

One of the clear disadvantages of CTM is to implement FD of stable flow for flow that is not in the stable state yet. For example, when a queue starts to dissipate, CTM offers a volume rate equals cell capacity, which means the average speed of the flow is about 3m/s (10.8km/h) with the acceleration of $6m/s^2$. This value is too high for a realistic vehicle. The reason for this problem is the realistic vehicles need time to start up and accelerate.

The basic idea of FD for startup flow is to calculate a more reasonable demand curve for the very beginning of flow starts, which considers the factor of driver's reaction time and acceleration process. As time passing, the supply curve evolves to the normal one gradually.

The TFD for startup flow can be described as following (Figure 24): when the volume starts from static state, a period of startup lost time t_{loss} has to be considered. During this time period, the demand volume stays at zero. After t_{loss}, vehicles are under acceleration process. During this period, the demand volume of the cell can be estimated as following:

$$D_i^n = min\left(\frac{K_i^n at_a}{2}, \frac{\omega at_a K_{jam}^n}{at_a + 2\omega}, Q_{i,m}^n\right) \qquad \textbf{4-12}$$

Where, K_i is the vehicle of the cell; a is the average vehicle start acceleration; t_a is the past acceleration time. t_s is the passed start time; therefore, $t_a = t_s - t_{loss}$. Based on this definition, the average speed for time t_s is $v_t = \frac{1}{2}at_a$. As t_s increases, v_t increases consequently. When $v_t \geq v_{free}$, the demand curve evolves into the stable one.
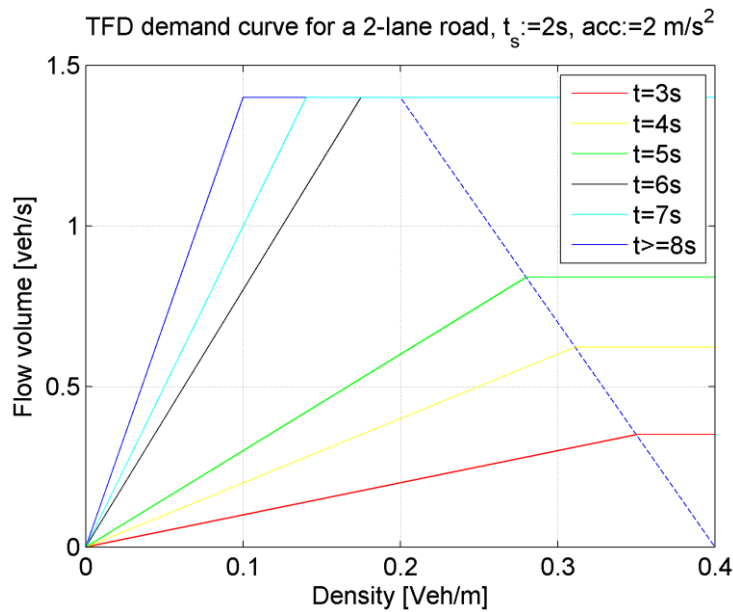
Figure 24 Variable demand curve for vehicle acceleration

In the application of CTM, the startup flow is easy to be recognized.

We describe the reorganization rules from two perspectives: intersection related and non-intersection related.

For a model including stop lines, the startup flow always happens in the cells just before signal stop lines. These situations can be recognized easily.

$$\begin{cases} K_{i+1}^n = 0 \\ q_{i,i+1}^n > 0 \\ q_{i,i+1}^{n-1} = 0 \\ K_i^{n-1} \neq 0 \end{cases} \qquad \textbf{4-13}$$

If conditions described in equation 4-13 are satisfied, we record the past start time t_s=1 for the current cell (cell i). It means the signal light changes from red to green and vehicles start to move. Since this time spot, the start time increases as the simulation time, until the demand curve fits the normal demand curve. for the cells which are not just before stop lines, their recognition is based on its downstream cells (only if its downstream cell is involved in the startup flow, in other words, assigned a start time)and together with the following conditions:

$$\begin{cases} q_{i,i+1}^n > 0 \\ q_{i,i+1}^{n-1} = 0 \end{cases} \qquad \textbf{4-14}$$

For a model not including stop lines, the following, equation 4-15 is suitable for recognizing all cells with startup flow.

$$\begin{cases} K_i^n = K_{i,jam} \\ q_{i,i+1}^n > 0 \\ q_{i,i+1}^{n-1} = 0 \end{cases} \qquad \textbf{4-15}$$

The difference between condition 4-13 and 4-15 is that the first one describes the situation of intersection, and the second one describes the situation of normal link. If an intersection is modeled, the startup flow can be recognized, even the current cell is not with jam density (the situation under which the queue length is very short). If it is a normal link, the algorithm can only recognize a cell with jam density; otherwise it is illogical to recognize it as stop.

4.4.2 FD for queue forming

Another situation under which the simplified FD cannot be described very well is when the queue is forming.

During this period, the density at the queue forming area with a congested density should have lower density than what the normal FD described. It means that the backward wave should have lower speed than queue dissipation process.

Figure 25 shows an example of the temporal FD. The blue curve shows the normal supply curve, at the same time, the red one shown a supply curve for congested situation.

When applying the FD for congested supply curve, the situation under which the flow is performing queue forming has to be recognized. Therefore, the following conditions have to be satisfied:

$$\begin{cases} K_{i-1}^n > K_{i-1,c} \\ K_i^n > K_{i,c} \\ K_{i+1}^n > K_{i+1,c} \\ K_{i-1}^n < K_i^n < K_{i+1}^n \end{cases} \qquad \textbf{4-16}$$

For application, two backward wave speed values are predefined, backward wave speed for normal flow (ω_n) and backward wave speed for queue forming (ω_{con}). When condition 4-16 is satisfied, ω_{con} is used for the current cell, otherwise, ω_n is used.
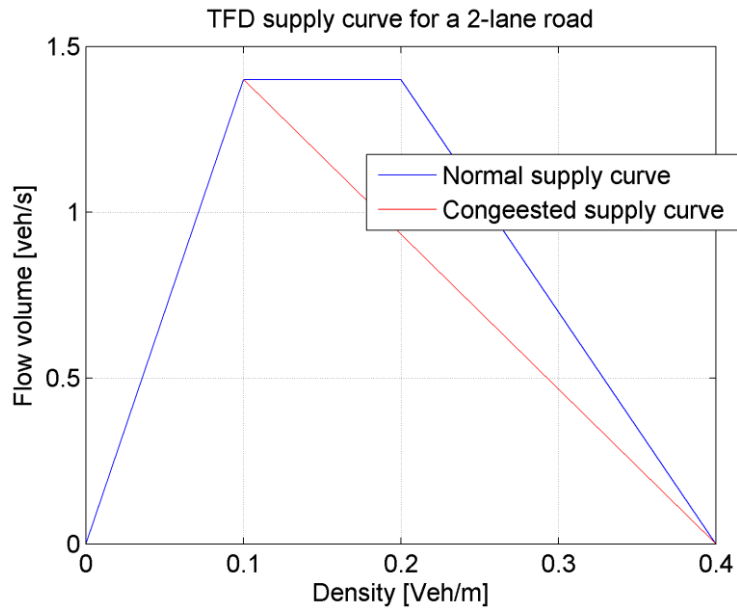
Figure 25 Temporal supply curve

4.5 Consideration on model performance of CTM

For the data fusion based integration framework, which will be discussed in the next chapter, the problem of what factors influence the final performance of a CTM model is important. In this section, several related points are discussed.

4.5.1 Parameters of CTM and IFCTM

This chapter describes several enhancements on CTM. However, because the main goal of this thesis is to discuss the method of model integration, both the enhanced and original CTM will be adopted for HyTran by different integration framework (FTF and DFF), the performance of the macroscopic model is mostly concerned.

Table 3 Parameters comparison: CTM and IFCTM

	CTM	IFCTM
Free flow speed	fixed	variable
Backward wave speed	fixed	variable
Jam density	fixed	fixed
cell capacity	fixed	variable
Other parameters	-	Parameters for density distribution Startup loss time Startup acceleration

Table 3 shows the basic situation of parameters of CTM and IFCTM. One point worth noticing is that the IFCTM is a method that improving the feasibility of CTM under highly interrupted flow, but not designed for DFF. There are two reasons: first, IFCTM has more parameters to be calibrated, it needs more computational work for calibration process; second, the improvements that can be obtained from enhancement on CTM also can be obtained by using DFF when the evaluation scenario is

practical. Therefore, IFCTM is chosen for FTF in order to improve the performance of flow transition at model boundaries; and CTM is used for DFF due to its concise parameters structure.

Four basic parameters of CTM includes: free flow speed, backward wave speed, jam density, and cell capacity. They influence the final performance of a CTM model by their own ways.

The free flow speed determines how fast a vehicle can travel across a link under the free flow condition. For a given density within specific range, this value can determine the traffic volume. However, this range is relatively narrow for a typical application. In addition, at the free flow regime, traffic volumes are more significantly determined by the input volume. In addition, from the perspective of realistic data about volume-density plot, at the low density range, free flow speed is fair stable.

The backward wave speed can be shown clearly when a queue is forming. And it also affects how fast a formed queue discharges together with jam density. In addition, it affects the link density at specific queue condition. From the perspective of realistic data, backward wave speed is highly unstable under different situation.

The cell capacity determines the performance of an intersection model, especially when the demand is high enough to give rise to a saturated flow. This is because if the vehicle density in the cell that is just before the stop line is also in the congested regime, this density leads to a corresponding volume rate that equals the cell capacity. One point worth noting is that a cell capacity does not determine the capacity of a link consequently, the link capacity is a comprehensive performance influenced by diverse factors.

Jam density determines how much vehicle can be stored in a link at the fully congested condition, which affects the simulated link density.

4.5.2 Signal expression considering pedestrian crossing

In the application of CTM for urban intersection evaluation, the signal plan has to be modeled by CTM. As summarized in section 2.4.4, the signal plan which includes flashing green, flashing red, amble, and red/amble, can be simplified to green and red signal signs by a concise principle.

In terms of traffic engineering, a more effective method is to model an effective green time. However, the relation between effective green time and display green time varies as the change of green ratio, which makes it not able to be calculated by a simplified method.

Another important situation which affects the effective green time is the conflicted pedestrian crossing, which cannot be described by the CTM. There are two methods can be used to express the influence of pedestrian crossing. The first one is to use the limited cell capacity for the critical cell; the second one is to use modified green time, in other words, the effective green time is reduced by the conflicting crossing pedestrians.

In this thesis, we do not try to find a more detailed way to explain the exact modeling method for different signal signs and signal programs. But, we will solve this problem from another perspective, the hybrid simulation. Generally speaking, microscopic models are highly developed for modeling complex signal plans, the effective of diverse signal signs and pedestrian crossing are well described. The basic idea of this thesis is to utilize microscopic model's ability of modeling detailed intersection infrastructure in the hybrid modeling.

4.5.3 The influence of FIFO (First In First Out) principle

Based on the update principles of the diverging cells that are typically used to model entrance links of intersections, CTM follows a FIFO rule. It means that for each diverging cell, the sending volume of it has constant turning percentages. If one of its downstream cells is blocked, the diverging cell will be completely blocked (as shown in Figure 26, red curve and red cell mean they are blocked). This phenomenon is possible to cause problems.
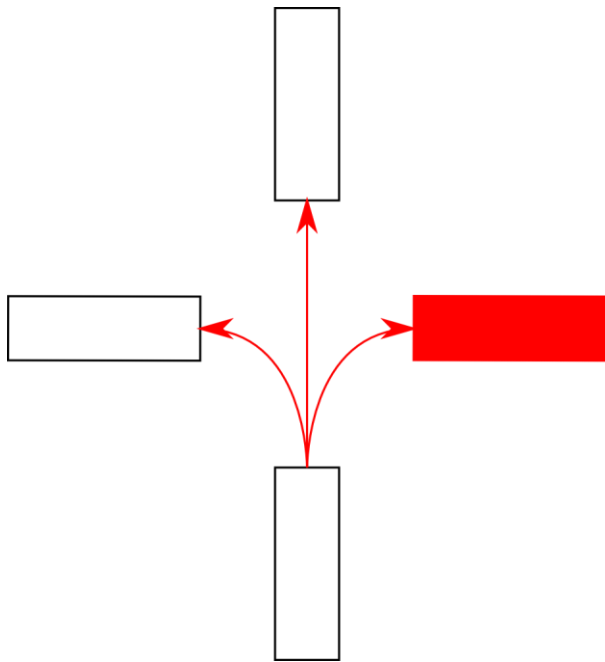
Figure 26 the influence of FIFO
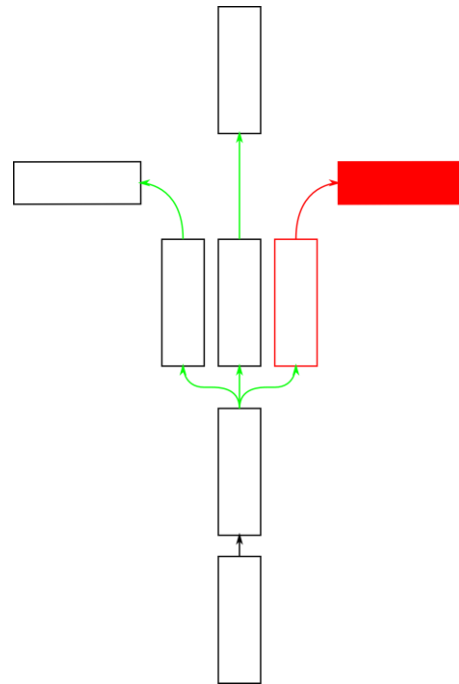
Figure 27 Parallel cells for intersection entrance

One commonly used method for relieving this problem is to use parallel cells after the diverging cell (as shown in Figure 27). The essential idea of this method is to enlarge the load ability for each direction. Under this situation, even one turning direction is blocked (the right turning in the figure), the blocked flow does not influence the upstream diverging cell instantly. Alternately, vehicles can wait in the corresponding parallel cell (the red rectangle with no fill color). Before this cell is fully occupied, the diverging cell can keep sending vehicles out of it. This solution explains the realistic phenomenon: if one downstream direction is blocked long enough, the entire upstream diverging link will be blocked by vehicles that tend to run into the blocked direction but cannot change to their target lane (because the lane has been fully occupied by queuing vehicles).

In addition, how long the parallel cells are set will influence the link capacity largely; especially when this link leads to a signal controlled intersection. Therefore, the topological design has to be carefully done, and analyzed with respect to specific modeled objects.

4.5.4 Lane change related capacity drop

CTM describes neither the lane-changing behavior nor the lane-changing related capacity drop. This is a main reason for which CTM has weaker predictability than a microscopic model. When CTM is used for modeling a segment with frequent lane-change behavior, the resulted cell capacity has to be determined during the calibration process. Especially, when the changing ratio differs, new calibration has to be made. However, microscopic models are capable of modeling lane-changing

behavior (although the analytical ability of those models still can be improved), therefore, their influence on general performance is better predicated.

4.6 Summary

This chapter discusses CTM, including its basic attributes and enhancement, together with the summarization in related sections in literature review, the author intends to make a solid preparation for implementing CTM with the proposed hybrid simulation framework, HyTran.

IFCTM is aiming at improve the description ability of CTM. Based on those enhancements, we pay more attention to the unstable traffic flow, and try to describe their behavior that is different to a stable flow. However, a more desired purpose of doing this is not for using CTM independently, but using CTM with the hybrid framework. Specifically, the IFCTM has better performance when it is used for boundary interfaces flow transfer both from macroscopic to microscopic and microscopic to macroscopic. This performance will be discussed in the part of model evaluation.

This chapter also discusses the model attributes of CTM. We want to clarify the influence of diverse factors to the performance of CTM. This is a preparation for implementing CTM with the data fusion based framework (DFF).

5   HyTran: the integration framework

As stated in section 3.3, for HyTran, the discrete version of LWR model, the cell transmission model, is selected as the macroscopic level of HyTran. And, the VISSIM, whose fundamental core model is psycho-physical car-following model, is selected as the microscopic level of HyTran. The general introduction about mechanism of VISSM has been presented in section 3.3. And the attributes about CTM are discussed in section 2.4 and chapter 4.

From this chapter, we focus on the consistency problem that is the essential problem included in any MRM research. The rest part of this chapter is arranged as below: section 5.1 discusses the consistency problems that have to be faced by HyTran; section 3.2 proposes basic definition of FTF and DFF; section 5.3 discusses FTF in details; section 5.3.4 discusses the basic algorithms used in DFF, and section 5.5 discusses DFF in details; at last, section 5.6 make a short summary for HyTran.

5.1   Consistency problems of discrete continuum flow model and car-following based microscopic models

HyTran tends to focus on traffic flow modeling, but not network modeling; therefore HyTran takes path flows as input. However, there is no theatrical limitation of not including traffic assignment with the simulation process. The pre-trip route choice is expressed by assigning turning percentages for each intersection, which can be effectually implemented both in CTM and VISSIM. For network representation, we follow the a principle: macroscopic level models all the links which need to be considered in the modeling operation, and microscopic level models the sub-networks which need more detailed description. This principle will be discussed in detail in later sections.

From the perspective of flow consistency, there are two points needing consideration: boundary consistency and performance consistency.

The boundary consistency refers to the total volume of traffic flow does not change because of the flow transition. This requirement is mandatory for any integration framework, HyTran is no exception. Boundary consistency is maintained by approach for FTF, which will be discussed in section 5.3.

Another more complex but not fully investigated problem is the performance consistency. Magne (Magne, et al., 2000) makes a useful step about this topic. Magne's research summaries that, in order to make a macroscopic model and a microscopic model to be compatible, one solution is to make the microscopic model to have the same macroscopic characteristics as the macroscopic one. From the numerical perspective, this research proposes five principles for model consistency. They are:

(1). $q = kv$;

(2). $q|_{k=k_m} = q_m$, where k_m is the density corresponding to the maximum traffic volume rate;

(3). $\left.\dfrac{\partial q}{\partial k}\right|_{k=0} = v_l$, where, v_l is the speed limit

(4). $q|_{k=0} = 0$;

(5). $q|_{k=k_{jam}} = 0$, where, k_{jam} is the jam density;

Condition (1) and (4) is suitable for all continuous flow from the physical perspective, regardless of on which resolution level they are. Condition (3) describes the state of a free flow, in which vehicles do

not affected by others and can travel at their desired speed. If considering a strictly homogeneous traffic composition, this condition is also satisfied by CTM and microscopic model, although this condition is better satisfied by CTM, and car-following based microscopic model can also describes a mixture of vehicles with different desired speeds. Except for condition (1), (3), and (4), more serious problems emerge when considering the rest two principles ((2) and (5)) with fundamental diagrams of CTM (Figure 28) and the corresponding q-k relationship of car-following based microscopic model (Figure 29).
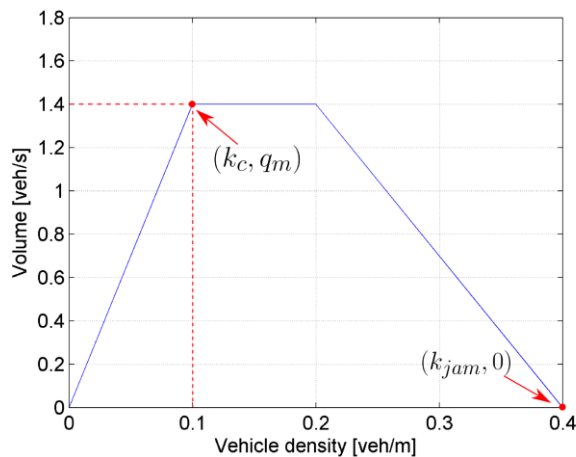
Figure 28 Typical trapezoid fundamental diagram (piecewise linear q-k relation for 2-lane road)
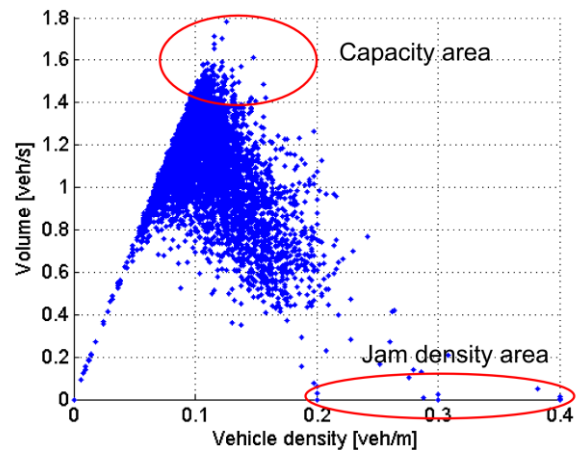
Figure 29 q-k distribution obtained from VISSIM simulation, 2-lane road, obtained from a bottleneck recovery case, data averaged at 10s intervals

As described by the condition (2), a maximum flow rate (q_m) exits, and it corresponds to a specific density k_m. This condition is true for CTM, the maximum flow rate corresponds to cell capacity, and k_m corresponds to the critical density. But, for microscopic model, this condition becomes ambiguous. It is difficult to determine the density value according to which the volume rate gets its maximum value; also the maximum volume is blurry. This phenomenon is shown by the "capacity area" in Figure 29. As described by the condition (5), a unique density point, the jam density, should exist, on which the corresponding traffic volume equals zero. For macroscopic model, this point exists and well defined. However, for microscopic model, this point is hardly to be determined, because when the vehicle density is high, the simulation results show a large variance (Jam density area in Figure 29). Vehicles may stop with various front-to-front distances or moving with very low speed without complete stop, which lead to the "various jam densities". A point worth noting is that the data set used for Figure 29 is only a specific example. If the setting for data obtaining method is changed, the shape of the point cloud changes. However, the basic pattern that is described in the last paragraphs still performs.

The consistency in traffic performance for two models is also a discussed by other research. Burghout (2004) explains it as "for those facilities that can be simulated sufficiently well by both models, they need to produce (nearly) the same results on the mesoscopic level of aggregation". The basic idea of this explanation is that if a simple network, for example, a single link, is modeled by a microscopic model and a macroscopic model independently, two models should offer similar results under same simulation scenarios. This viewpoint broadens the limitation proposed by Magne by allowing reasonable discrepancy between two models. From this point of view, it can be considered that condition (2) and (5) are also satisfied by CTM and VISSIM to some extent.

In summary, it is difficult to find a theoretical perfect match between CTM and VISSIM, and from the practical perspective, this is also not necessary. A logical strategy is to ensure the volume consistency, and make weak requirement on theoretical consistency about sharing same macroscopic characteristics. This is the key task of FTF.

In the current thesis, we tend to go one step further at the direction of performance consistency. The motivation comes from a simple question: although CTM has weakness in modeling complex behaviors, it is possible to calibrate it according to a well-developed microscopic model and obtain a consistent macroscopic level simulation results? If the answer is yes, we name it as "advanced performance consistency". To discuss the advanced performance consistency is the foundation and main goal DFF.

5.2   Definition of modeling framework

The integration framework designed in HyTran consists of two parallel components: the Flow Transition based Framework (FTF, shown in Figure 30) and the Data Fusion based Framework (DFF, shown in Figure 31).

In HyTran, two levels of model are included: macroscopic level and microscopic level. The main component of the macroscopic level is a macroscopic traffic flow model; the main component of the microscopic level is a microscopic traffic flow model; the macroscopic level involved the complete road network which is simulated, from the perspective of function, this level offers an evaluation on performance of the entire network; from topological perspective, it works as a "skeleton" of the model structure. The microscopic level consists of several independent sub-networks modeled by microscopic traffic flow model. The sub-networks can cover various extents; they can be as small as a single road intersection, or as big as an arterial including several intersections. From the perspective of function, microscopic level enhances the analysis ability for those simulated interesting points; from the topological perspective, they are some extra models attached to the skeleton at the corresponding locations. Whether to implement a microscopic model is determined by the requirements of the specific analytical cases. For example, if a new designed traffic signal plan needs to be evaluated in this project, it is obvious that the intersection with the signal plan needing evaluation should be modeled by microscopic model, in order to offer a more accurate model response to the signal plan.

If some part of the network are modeled both by microscopic level and macroscopic level, a problem has to be answered is how does the user treat the results of them, which has been proposed as performance consistency problem in the last section. A traditional solution is to trust the high resolution model fully. One of the reasons is that a flow transition based approach does not have an overlapped macroscopic level model if some part is modeled by microscopic level. This solution is included in HyTran as FTF. In addition, HyTran tends to realize the "advanced performance consistency" by introducing a data fusion idea.

The framework described in Figure 30 is the traditional solution, the Flow Transition based Framework (FTF). Under this framework, when a simulation starts, both the macroscopic level and microscopic level begin to simulate. The traffic flow transition algorithms work when flows move across the boundaries from macroscopic network to microscopic network, and inversely. And it maintains the consistency of traffic volume during the transition. The performance of microscopic model is used as the model output.

The framework described in Figure 31 is a new concept proposed by HyTran, the Data Fusion based Framework (DFF). Under this framework, when a simulation starts, the calibration plan generator calls the microscopic model and makes probing run for multiple times with a simulation period as short as possible. Afterwards, the collected data from probing runs is used to calibrate the corresponding part of macroscopic model. When the calibration process succeeds, the macroscopic model will be used to simulate the entire network for any period. During the calibration process, in order to show the representative attributes of the target sub-network, the probing runs of microscopic model have to be carefully designed and planned.

The essence of DFF is to extract useful information from short time microscopic model running (the probing runs), and then, to integrate this information into macroscopic. That information possibly comes from some infrastructures that cannot be enough described by CTM, and also possibly comes from the roughly calibrated parameter values. From this point of view, the target of DFF is to improve the description ability of the macroscopic level. The task of data fusion is achieved by calibration algorithms, and the influence of those extracted information is expressed by using suitable values for parameters of CTM.
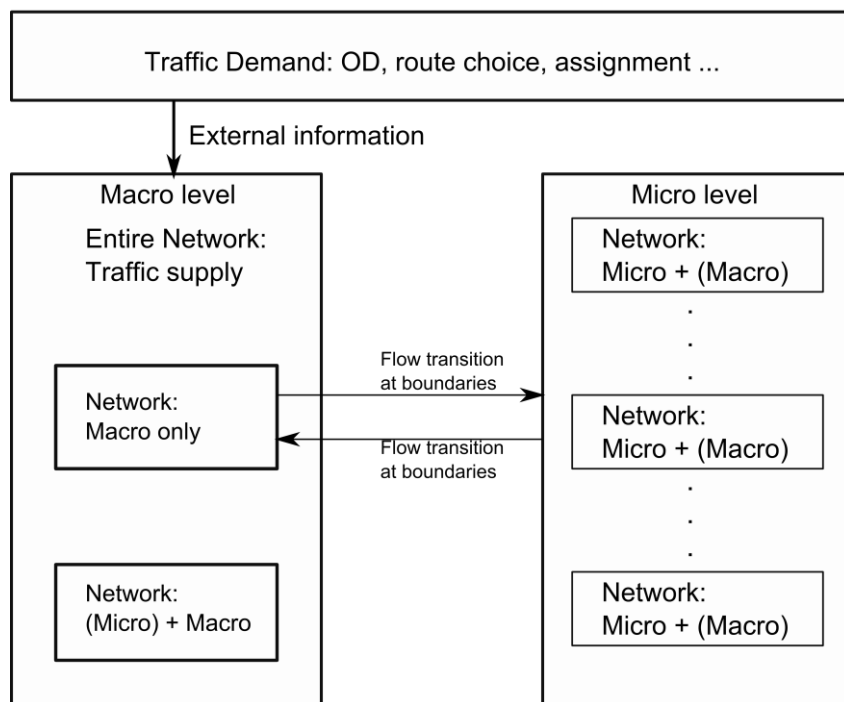
Figure 30 System Architecture of HyTran: Flow Transition based framework (FTF)

Data Fusion Based Framework (DFF)
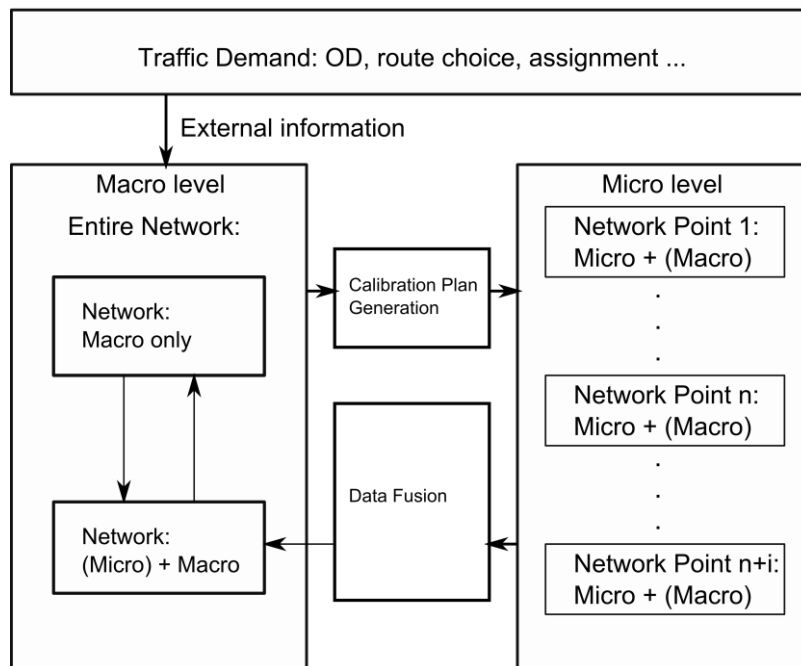
Figure 31 System Architecture of HyTran: Data Fusion based framework (DFF)

5.3 Flow Transition based Framework (FTF)

The abstract function of flow representation transition, as it expresses literally, is to make a transformation between flow representation (described by vehicle density, traffic volume, and so on) and individual representation (described by modeling each individual vehicle). Another way of describing this function is to maintain the consistency of traffic flow at the boundaries between microscopic sub-network and macroscopic sub-network. Traffic flow volumes across the boundary should be described correctly regardless of the volume rate is determined by which side of the boundary. If the volume rate is determined by the demand side (upstream side), vehicles should be generated in the downstream side appropriately; at the meanwhile, if the volume rate is determined by the supply side (downstream side), traffic flow exiting the upstream side has to be controlled appropriately.

Generally, this problem is discussed in terms of two directions: macroscopic to microscopic and microscopic to macroscopic.

For the flow representation transition from macroscopic level to microscopic level, the implemented algorithm has to generate vehicles in the microscopic link, according to the flow demand from its upstream macroscopic link and the current traffic states of the downstream microscopic link. The system has to decide on which time a new vehicle should be generated on the microscopic link, which can be described by the headway (time or space) between the newly generated one and its preceding one; in addition, the parameters of the generated vehicles, including position (lane selection), initial speed, and so on, also has to be determined.

For the flow representation transition from microscopic level to macroscopic level, the implemented algorithm has to calculate the traffic volume that come across the boundary; in addition, it is also necessary to control the vehicles that are spatially close to the boundary in the upstream microscopic

link, in order to show the influences of the downstream macroscopic link when its supply cannot satisfy the demand of its upstream link, such as a queue is propagating upstream.

In this section, both the two problems stated above are discussed. In addition, an acceptable method for lane determination is discussed.

5.3.1 Boundary interface from macroscopic to microscopic

Traditionally, the boundary interface calculates suitable time headway for vehicle generation according to the desired traffic flow rate. This thesis adopts a relatively new method for vehicle generation which is suitable for short macroscopic time interval. The rest of this section discusses the potential problems of the time headway based method first, and then introduces the method used in this thesis.

About the time headway based method

As summarized in the chapter of literature review, to transfer the desired traffic volume rate ([veh/s]) into time headway ([s]) is a feasible method. Those two variables have a direct relationship. However, there are two weak points existing in this method.

First, the time headway based method is suitable for the case in which the time interval of macroscopic model is much greater than the cycle of microscopic model. Under this situation, the time interval of microscopic model can be considered as continues value. Therefore, the desired time headway can be followed strictly, and errors can be ignored. However, when this assumption is not satisfied, the calculation error may accumulate. For example, if the link capacity is set to $0.65\ veh/s$, the corresponding time headway, or can be addressed vehicle generation interval, is $1.5s$. When the CTM model uses a cycle of $1s$ and the microscopic model uses a cycle of $0.2s$, the used generation interval can only be $1.6s$ (with the errors of 7% for each step). The system error will accumulate during simulation period. Another problem is when the macroscopic cycle is very short (such as 1 second), although the desired headway changes for each cycle, it does not take effect in the microscopic model during such short period. For example, vehicle inputs in VISSIM follows a randomly generation process, the desired input volume is realized only when the simulation period is long enough.

Second, the time headway based method is suited for one lane situation. Under this situation, the latest generated vehicle must be the preceding vehicle of the new generated vehicle; in addition, it must also be the adjacent and leading vehicle of the new generated one. Therefore, the determination of generation and the initial speed of the new generated vehicle are both determined by the status of the last one. However, if the boundary link has more than one lane, this phenomenon is possible to be not true. One example is: assuming a 2-lane link is modeled and the current desired headway is t_h; when a vehicle is generated in lane 1 at time t; at time $t + \Delta t$, if $\Delta t < t_h$, according to the time headway based principle, no vehicle is allowed to be generated. However, at this time, lane 2 is empty; therefore, a new vehicle can be generated on lane 2 regardless of how long is the headway in lane 1. In reality, the same situation happens. This is why the observed traffic volume rate for a short time period can be much larger than the long time one. Despite the possible enhancement for headway based method for multiple-lane situation (such as generate vehicles on different lanes separately), this simplified example shows that the generation interval is not the key point, but the accumulated generation volume is.

One simplified alternative of time headway based method is to set the input volume directly. This method does not generate individual vehicle directly, but only change the input volume setting of the

microscopic model. From practical perspective, this method has one attractive advantage: the operation process will be relatively concise, because most of the complex tasks, including initial speeds, new vehicle generation (to initialize a new object in microscopic program), are handled by the original software package. However, two risks cannot be ignored: first, the volume consistency is not guaranteed. Second, the volume errors may accumulate during the simulation process. Because the two risks are significantly affected by the vehicle generation mechanism used by the current microscopic software package, it is difficult to evaluate the risks without knowing its source code.

The method can be summarized as the following algorithm:

Algorithm 2 Vehicle generation

Algorithm: Vehicle generation
while current simulation second (Macro) $\leq$ simulation period (Macro) **do**
    for *each relative link* **do**
        Read desired supply volume from CTM
        Obtain handle to Vehicle Input
        Calcualte input value
        Set the input volume

Dynamic demand/supply equilibrium

HyTran proposes a method addressed as "dynamic demand/supply equilibrium" for vehicle generation. Without focusing on calculation of vehicle generation interval, this method chose the perspective of balancing demand and supply. The basic idea of this method is to obtain the demand volume from upstream macroscopic link, and then, try to generate vehicles in the downstream microscopic link as much as possible. After the current time interval finishes, the realistic number of generated vehicle is fed back to macroscopic model as the result value. With this definition, the demand value is possible to be a decimal value. Considering the adopted macroscopic model cycle, the decimal part is not ignorable compared with the integer part. For example, assuming a 2 lane link with cycle of 1 second, the calculated demand value is between 0 to 1.35 veh, which makes the decimal part (0.35) cannot be ignored. The adopted method uses the following equation to calculate the demand volume:

$$D(t) = \lceil \min\{vk^t t, q_m t\} \rceil \qquad \textbf{5-1}$$

The algorithm of dynamic demand/supply equilibrium is described as below.

Algorithm 3 Vehicle generation

Algorithm: Vehicle Generation

Model initialization

while current simulation second (Macro)$\leq$ simulation period (Macro) **do**

    Compute the desired passing value $D(t)$

    for each micro cycle **do**

        if $D(t) < 1$ **then**

            break

        else

            Determine empty space for each lane $d(t)$

            Calculate the minimum required space headway d_{min}

            for each lane **do**

                if $d(t) \geq d_{min}$ **and** $D(t) \geq 1$ **then**

                    Add a new vehicle

                    Record vehicle information

                    Subtract 1 from $D(t)$

The minimum space headway is the necessary required empty space for generating a new vehicle, in other words, the system generates a new vehicle only if the current space headway is longer than the minimum space headway. This value is determined by two factors: speed of preceding vehicle (v_p), and average vehicle occupied length (L_{veh}).

$$d_{min} = v_p \cdot t_r + L_{veh} \qquad \textbf{5-2}$$

Where, t_r is a parameter named as related time. The value of this parameter should be determined by considering the minimum acceptable time headway for vehicle generation, in order to get the generation rate as high as possible.

The function of minimum space headway is to prevent vehicles from being added into the link too frequently, which is probable to make the vehicle to decelerate instantly after generation. This unexpected deceleration phenomenon reduces capacity of the interface significantly.

The initial speed of the new generated vehicle is determined by a concise method: the speed of new generated vehicle is assigned same speed with its leading one.

5.3.2 A method for Mandatory Lane Change (MLC)

The lane change behavior has to be discussed with the attributes of the used microscopic model. Therefore, this problem is discussed according to VISSIM.

In CTM, lane information is not described, unless independent parallel cells are used (but this approach cannot be used for arbitrary situation, and has to be determined according to related signal program). However, when the macroscopic-microscopic boundary is located on a link with multiple lanes, it has to be determined on which lane the new generated vehicle should be placed, considering the desired turning percentages. Otherwise, more lane change actions are probable to happen on the microscopic link than in reality. This phenomenon reduces the capacity of the link significantly. In realistic road, drivers make lane choice earlier enough even it does not looks really necessary. This behavior is still not well simulated by microscopic models. In microscopic modeling, the user should make the link for lane change longer than the anticipated queue length, in order to

ensure that the vehicles have enough opportunities to change to their ideal lanes. However, for a hybrid application, modeling process should maintain the link consistency, which means the length of a specific link is strictly determined by reality of modeled road.

Lane change related capacity drop emerges when demand for lane change is stronger in the model running than in reality. Therefore, the designed method aims for reducing the demand for lane change on a logical level. In addition, the designed approach should try to minimize its influences on traffic. Three methods can be used to solve the potential problem. First, the system makes the microscopic link before turning points long enough, therefore, simulated vehicles have more time and space to search possible acceptable lane changing gap. Second, the system generates vehicles on its desired lane, which is decided according to the turning percentage at the upcoming intersection. Third, the system moves vehicles which are not on its desired lane to its desired lane.

The first strategy is probable to violate the requirement of link consistency between microscopic level and macroscopic level. If the entrance link for an intersection is extended, the load capacity of this link will be larger than the realistic situation. The second strategy is ideal for solving this problem, but considering the implement, vehicles should be assigned with route information when it is generated, this is not natural for VISSIM. In addition, this method mixes the vehicle generation and lane choose together, which makes it difficult to be handle in practical programming. The third one avoids the problems mentioned above, and will be adopted in this thesis. This method is named as "Mandatory Lane Change (MLC)".

MLC takes effects only under specific condition. If the demand volume is low enough, the vehicles are able to make lane change based on the original strategy involved in the microscopic model software package. Under this situation, no extra mandatory lane change process is needed. On the contrary, if the demand volume is too high to be satisfied by the current infrastructure, the extra mandatory lane change process is neither needed. For example, if one of the three downstream legs of a link has too much relative flow rate than the other two, vehicles are possible to have problems with change to the desired lane, because the desired lane is already occupied. Under this situation, it is not useful to move the vehicles to its desired lane in mandatory. Except for the two situations stated above, the system should only carry out mandatory lane change when the demand volume and demand for lane change are at an adequate level.

One of the requirements of MLC is to minimize the unexpected influence of mandatory lane change on current flow state. Therefore, the system searches for exchangeable vehicle pairs on specific area. The exchangeable vehicle pair (Figure 32) refers to two vehicles that are running on different lanes of a link, but their target lanes happen to be another's current running lane. If two vehicles satisfy this condition, the system can exchange the route information of them. After this operation, the lane changing demand of those two vehicles does not exist anymore. During this process, MLC only exchange the route information of the exchangeable pair, and does not change any other states of them. Because no speed, acceleration, and location of vehicles are changed during the process, the solution does not bring extra influence to the current traffic flow statue. The basic process is shown below.

Algorithm 4 MLC

Algorithm: Mandatory lane change
while current simulation second ≤ simulation period **do**
    for each relative link **do**
        Read vehicle information on the current link
        Search for exchangeable pairs
        for each exchangeable pair **do**
            Change the route choice of the two vehicles
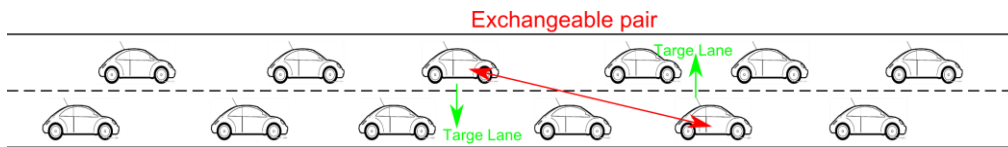
Figure 32 MLC: searching for exchangeable pair, only exchange their route information, minimize its influence on current traffic state

5.3.3 Boundary interface from microscopic to macroscopic

The transition from microscopic representation to macroscopic representation is based on controlling the speed of the vehicles which are about to leave the microscopic link soon, and recording the vehicle amount getting across the boundary line. The boundary interface from microscopic to macroscopic has to have at least two functions: under free flow situation, transfer the passing volume in upstream microscopic model to the downstream macroscopic model; under the congested flow situation, to block the vehicles in upstream microscopic model from leaving, which can be used to modeling the phenomenon of queue propagation.

The main idea is to control the running speed of the vehicles that are leaving the upstream microscopic link, and the desired speed is calculated according to the vehicle density and FD of the CTM (5-3).

$$v_{desired} = \frac{Q(K_i^n)}{K_i^n}$$ **5-3**

A possible problem in the interface is the suddenly speed change of the controlled vehicles, as discussed in (Burghout, 2004) (however, this thesis uses an idea of virtual vehicles). The current thesis avoids this problem by using a relative small macroscopic model time interval (1s). With this setting, the vehicle density of the downstream cell in CTM does not change largely between two time steps, therefore the relative desired speed neither changes largely. Although the determination of macroscopic time interval does not focus on this problem originally (it is determined by considering the expression of traffic signal), it does bring benefits for boundary interface from microscopic to macroscopic.

Another possible problem is the speed control does not generate a zero flow. In order to maintain a blocked flow when it is necessary, the minimum vehicle speed is defined. When the desired speed is lower than the minimum speed, the speed of the controlled vehicle will be set to zero. In realistic application, if the minimum vehicle speed is not applied, before downstream CTM link goes into jam density, extra vehicles pass the boundary unexpectedly.

Algorithm 5 Speed control

Algorithm: Speed control

while current simulation second (Macro) $\leq$ simulation period (Macro) **do**

    for *each micro cycle* **do**

        for *each related link* **do**

            Determine controlled vehicle

            Determine desired speed

            for *each controlled vehicle* **do**

                if desired speed $\leq v_{limit}$ **then**

                    Change speed to desired speed

                else

                    Change speed to zero

5.3.4 System synchronization

Under FTF, the system has to run both the macroscopic model and the microscopic model, therefore it has to be decided that how the two models are synchronized.

The solution to the synchronization is to make a basic assumption that the macroscopic model has a simulation cycle that is an integral multiple of the simulation cycle of the microscopic model. It means that if the macroscopic cycle is T_M and the microscopic cycle is T_m, the following relation should be true:

$$T_M = n \cdot T_m \qquad \textbf{5-4}$$

Where, n is an integer.

Based on this assumption, the algorithms stated in the current section can work without synchronization problem. For the boundaries from macroscopic model to microscopic model, the vehicle generation algorithm is executed during every macroscopic cycle, and the desired demand volume is updated for each T_M. MLC is executed for each microscopic cycle. For the boundaries from microscopic model to macroscopic model, the vehicle speed is controlled for each microscopic cycle.

5.4 Foundation of Data Fusion based Framework (DFF)

One of the main functions of DFF is to calibrate the macroscopic level by data obtained from microscopic model (probing runs). Based on this requirement, HyTran has to find appropriate algorithms for the data fusion process.

As summarized in chapter 2, various algorithms have been applied to calibration of traffic simulation model. From the literature, it is difficult to tell which one is superior to another, especially when diverse perspectives of the algorithms have to be considered. This chapter presents two well-known algorithms, Simultaneous Perturbation Stochastic Approximation (SPSA) and Bootstrap Particle Filter (BPF). Before the two algorithms are presented, this chapter starts with the discussion of the attributes of CTM calibration at first.

5.4.1 Calibration of CTM for HyTran

Section 2.5.2 reviews the preceding research on CTM calibration. Generally speaking, the calibration strategies come as two categories.

The first category is based on the physical definitions of CTM parameters. As introduced in pervious related chapters, CTM calculates traffic volumes on discrete points according to attributes of the used fundamental diagram, which is defined by four basic values: free flow speed, backward wave speed, jam density, and cell capacity. Because those four values correspond to physical values directly, they can be calibrated by dada analysis on volume-density dataset. For example, backward wave speed can be estimated by applying the method of least-square to data points located in congested regime. Cell capacity can be estimated by using the maximum observed traffic volume rate.

The second category is based on optimization of model performance. These methods do not focus on obtain the parameters separately and directly according to their definitions, but on adjusting the values of parameters in order to obtain better model performance, which has to be measured by predefined indicators. Compared to the first category, this strategy is more effective and fit the mechanism of system simulation better. The main reason is that as a model of the real world, the parameters do not mimic the real world exactly, especially for the models with relative low resolution. Therefore, although the very exact values of parameters are found according to their definitions, it does not guarantee the developed model will offer a better simulation results compared with realistic world or the field data, due to the simplification and idealization during model building process. From another perspective, users have to make several compromises or compensation during the parameter value selection, to ensure a better comprehensive attributes of the model, which should be the main target of model calibration.

Another problem needing consideration is about the data which can be obtained for a calibration process. It is not difficult to obtain traffic volume data from realistic world and simulation model. However, it does not guarantee that any dataset will include information abundant enough to involve all the necessary information for CTM calibration. For example, when running a microscopic model, it is seldom to obtain a vehicle density data that equals the jam density. This is because in car-following based microscopic model, when simulating a queue forming process, cars tend to move in the pattern of "stop and go" rather than stop instantly with maximum vehicle density. Therefore, the observed vehicle density will be always lower than the real jam density (by definition), unless the queue lasts long enough. Considering this problem, the performance based calibration strategy still possesses advantages.

Compared with model calibration by realistic data, to fuse information from microscopic model has its own characteristics. The most important one is that the system can read traffic volume data at any point at any time intervals. In contrast to realistic world, these data source in simulation model is not limited by physical devices' install locations and working conditions, such as detect loops. This is a key point for choosing the performance indicator.

Based on the statement in the last paragraph, the second strategy, optimization of model performance, is applied in this thesis. In addition, from the perspective of traffic engineering, the traffic volumes, synchronized with time axis, are always the most important performance indicator of a traffic network. Therefore, groups of traffic volume data at predefined points are used as the indicators of model performance. The advantage of choosing this indicator is to have a uniform expression of model performance. In addition, because the traffic volume data are obtained at short

time intervals and for the entire simulation time (starting with an empty network), this means the vehicle density between two control points is already considered. For other high-level indicators, such as delay and travel time, the difference more depends on the effectiveness of the algorithm used for estimating them.

As summarized in chapter 4, the possible calibrated parameters of CTM are free flow speed, backward wave speed, jam density, and cell capacity. It is very difficult to tell which parameters are not necessary to be selected as calibrated one.

In addition, the signal control is also open for calibration, since there is no solid method for effective green time estimation. However, the assignment rate is not constant but depends on several factors, including vehicle behaviors, signal cycle, green time, and red time, and so on. Therefore it is also reasonable to consider calibrating several key signal changing time point, when a signal controlled intersection is modeled. However, it is clear that, in a microscopic simulation projects signal time is commonly considered as model input but not parameter, which is different to what is done here.

5.4.2 Simultaneous Perturbation Stochastic Approximation (SPSA)

The Simultaneous Perturbation Stochastic Approximation (SPSA) is a promising optimization algorithm that is proposed by James C. Spall (Spall, 1998). The main purpose of this algorithm is to handle the optimization problems that are difficult or impossible to obtain an analytical solution. In addition, this algorithm works effectively for multivariate problems. Furthermore, the code for SPSA is very concise that makes its implementation easy. In summary, based on the description and reported application cases of SPSA, it is logical to suppose SPSA to be suitable for the application of traffic flow model calibration. Some research (e.g. (Lee & Ozbay, 2009)) on this topic have been executed.

The core idea of SPSA is to estimate the approximations of gradient of the objective function by and only by objective function measurements. Based on the gradient approximation, a new solution is proposed. In each iteration, a new proposed solution is formulated as:

$$\hat{\theta}_{k+1} = \hat{\theta}_k - a_k \hat{g}_k(\hat{\theta}_k)$$

Where, $\hat{g}_k$ is the function for estimating the gradient $g(\theta) = \partial L(\theta)/\partial \theta$. $L(\theta)$ is the loss function. $g(\theta)$ is estimated by:

$$\begin{bmatrix} \vdots \\ \hat{g}_{ki}(\hat{\theta}_k) \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ \dfrac{y(\hat{\theta}_k + c_k \Delta_{ki}) - y(\hat{\theta}_k - c_k \Delta_{ki})}{2 c_k \Delta_{ki}} \\ \vdots \end{bmatrix}$$

Where, $\Delta_k = \{\Delta_{k1}, \Delta_{k2}, \dots\}$ is the set for user-specified random perturbation vector.

The basic algorithm of SPSA can be seen below:

Algorithm 6 SPSA

Algorithm: SPSA

parameter initialization

for $k \leftarrow 1$ **to** n **do**

    $Update \; a_k$

    $Updata \; c_k$

    $Get \; random \; perturbation \; vector$

    $Calculate \; positive \; perturbed \; vector \; \theta_+$

    $Calculate \; negative \; perturbed \; vector \; \theta_-$

    $Calculate \; loss \; function \; L(\theta_-) \; and \; L(\theta_+)$

    $Calculate \; estimated \; gradient \; \hat{g}$

    $Update \; parameter \; vector \; \hat{\theta}$

The sequence a_k and c_k are calculated by equation 5-5 and 5-6.

$$a_k = \frac{a}{(k+1+A)^\alpha} \qquad \textbf{5-5}$$

$$c_k = \frac{c}{(k+1)^\gamma} \qquad \textbf{5-6}$$

The random perturbation vector is generated according to a Bernoulli ± 1 distribution (Δ_{ki}, where, k is the iteration number, i is the no. of vector components of $\hat{\theta}$). The value of α and γ affects the performance of the algorithm significantly. Based on the introduction in (Spall, 1998), $\alpha = 0.602$ and $\gamma = 0.101$ are used in HyTran.

The perturbed parameter vectors are calculated by equation 5-7 and 5-8.

$$\theta_- = \theta - c_k \Delta_k \qquad \textbf{5-7}$$

$$\theta_+ = \theta + c_k \Delta_k \qquad \textbf{5-8}$$

For the application of CTM calibration, the loss function is defined as the Mean Absolute Error (MAE):

$$L(\theta) = \frac{1}{n} \sum_{j=1}^{n} \frac{\left| V_{Mj}(\theta) - V_{mj} \right|}{V_{mj}}$$

More typically, the Root Mean Square Normalized Error (RMSNE) can be used:

$$L(\theta) = \sqrt{\frac{1}{n} \sum_{j=1}^{n} \left(\frac{V_{Mj}(\theta) - V_{mj}}{V_{mj}} \right)^2} \qquad \textbf{5-9}$$

Where, V_{Mj} is the volume measured at control point j in CTM for specific time period; V_{mj} is the volume counted at control point j in microscopic model; n is the total number of control point.

The updated parameter vector can be calculated by:

$$\theta_{k+1} = \theta_k - a_k \cdot \hat{g} \qquad \textbf{5-10}$$

5.4.3 Bootstrap Particle filter

Particle Filter (PF) is a combination of Bayesian estimation and Monte Carlo method, therefore, it is also known as Sequential Monte Carlo method (SMC). The original purpose of developing this method is to obtain a more applicable filter algorithm than Kalman Filter (KF) for non-linear and non-Gaussian problems. When traditional Kalman Filter is employed for solving non-linear problems, a

linearization process is unavoidable. For example, the Extended KF adopts Taylor series expansion to obtain a linear approximation. In contrast to the EKF, Unscented Kalman Filter (Julier & Uhlmann, 1997) offers a better approximation than EKF by describing a distribution by sigma points, which makes UKF do not need to calculate a linear approximation of the measurement equation. However, UKF highly depends on the Gaussian attributes of estimated posterior possibility distribution. PF is a more general method having similar orientation with UKF. PF does not using a group of parameters (which means the distribution function should be predefined) to describe the estimated distribution, but using Monte Carlo method to simulate a "real" distribution of them. Although PF needs more samples than sigma points used in UKF, PF handles the non-Gaussian and non-linear problems more effectively.

HyTran adopts a typical Bootstrap Particle Filter. This decision is made according to two considerations: first, to run CTM needs relative low computation resource, therefore, to sample more points does not lead to serious computational problems, this attributers makes the application of Monte Carlo method possible; second, PF has more widely feasibility for different noise distribution; the distribution of noises are not well known in this application, in other words, if other assumption on the distribution of noises are required, PF will still be suitable (both for the observation noise and state update noise). One possible problem with using BPF is hard to determine the necessary particle numbers in for the Monte Carlo approach, and, this problem is not fully investigated from theoretical perspective.

In order to fit the basic structure of Bayesian estimation, the macroscopic level is formulated as the following form:

Let $x(t)$ is the parameters of CTM; $p(t)$ is other related parameters, including effective green time, etc.

$$system\ state: \theta(t) = [x(t) \quad p(t)]^T$$
$$observations: V(t)$$
$$state\ update: \theta(t) = \mathcal{F}\big(\theta(t-1)\big) + N_s^t$$
$$observation\ equation: V(t) = \mathcal{M}\big(\theta(t)\big) + N_o^t$$

Where, the system state is a vector consists of parameters needing adjusting. Observation is a vector for collected traffic volume data from microscopic level.

The system update, described by $\mathcal{F}(\theta)$, is the predictable change of parameters. A possible solution is to model $\mathcal{F}$ as an autoregressive process, however, in the current thesis, we model it as a constant transfer.

The observation equation, $\mathcal{M}(\theta)$, is the developed CTM model itself. The attributes of this transformation is not analytical. The nonlinearity of this observation has to be considered when performing state estimation.

The algorithm of BPF can be described as below:

Algorithm 7 Bootstrap Particle Filter

Algorithm: Bootsrap Particle Filter

Particles initialization
Weights initialization
for *Each iteration* **do**
  Update system parameter
  for Each particle **do**
    Update particle
    Calculate linkihood function
    Update particle weight
  Normalize particle weights
  if need resamling **then**
    Resampling

The particle set is $\theta(t) = \{\theta_i(t) : 1 \leq i \leq N_p\}$. The initial particle set $\theta(0)$ is generated according to the predefined prior parameter distribution. For the initial particle group, each parameter is assumed to follow a uniform distribution on its possible range. The initial weights of particles are set to be uniform.

$$W_i(0) = \frac{1}{N_p} \qquad \text{5-11}$$

Where, N_p is the number of particles.

The main process of BPF is to update the weights. The update principle of particle weights follows the basic equation:

$$W_i(t) = W_i(t-1) \cdot \frac{Pr\big(V(t)|\theta_i(t)\big) \cdot Pr(\theta(t)|\theta_i(t-1))}{q\big(\theta(t)|\theta_i(t), V(t)\big)} \qquad \text{5-12}$$

Where, $q(\theta(t)|\theta_i(t), V(t)$ is the importance sampling distribution, which is the key component of MC. In bootstrap particle filter, we use the transition prior as the importance proposal, which means $Pr\big(\theta(t)|\theta_i(t-1)\big) = q\big(\theta(t)|\theta_i(t), V(t)\big)$. Therefore, equation 5-13 reduces to:

$$W_i(t) = W_i(t-1) \cdot Pr\big(V(t)|\theta_i(t)\big) \qquad \text{5-13}$$

Equation 5-13 is highly determined by $Pr\big(V(t)|\theta_i(t)\big)$, the likelihood function. For each update, the weights should be normalized by equation 5-14.

$$\mathcal{W}_i(t) = \frac{W_i(t)}{\sum_{i=1}^{N_p} W_i(t)} \qquad \text{5-14}$$

The resampling happens when the following variable ($N_{eff}(t)$) is smaller than a predefined thresh ($N_{eff} \leq N_{thresh}$).

$$N_{eff}(t) = \frac{N_p}{E_q\{W^2(\theta_t)\}} \approx \frac{1}{\sum_{i=1}^{N_p} W_i^2(t)} \qquad \text{5-15}$$

This function of this strategy is to resampling the particles based on the current estimate, if some existing particles have too low weights to be considered any more. Theoretically, this method does not change the estimation results, but is possible to save computational resource when the parameters are stable.

Based on PF estimation, the posterior distribution of the system state can be estimated as:

$$\widehat{Pr}(\theta(t)|V(t)) \approx \sum_{i=1}^{i=N_p} \mathcal{W}_i(t) \cdot \delta\big(\theta(t) - \theta_i(t)\big)$$

5-16

Where, $\delta(\theta)$ is a typical Dirac delta function. And the posterior estimation $(\hat{\theta})$ is:

$$\hat{\theta}(t) = \sum_{i=1}^{N_p} \mathcal{W}_i(t) \cdot \theta_i(t)$$

5-17

5.5   Data Fusion based Framework (DFF)

As described in the above section, HyTran uses online calibration-based interface to fuse the information from microscopic level into macroscopic level. The main purpose of this interface is to introduce the model behavior obtained from microscopic level into the macroscopic level, by this, the macroscopic level will be able to offer a consistency performance as the microscopic model. This is called "advanced consistency maintenance (ACM)" The direct consequence of ACM is the microscopic level can be paused after that a specific time period, and only the macroscopic level runs until a system signal requires a restart of the microscopic level.

Compared with FTF, the DFF) has several advantages. In other words, DFF solves several problems which are impeding the application of Hybrid simulate model. They are summarized as 3 points in section 5.2 as:

- Short microscopic model run;

- Results stability (multiple runs);

- Application easiness;

However, this strategy also brings some problems need to be solved.

The most important fact is that a single arbitrary run of microscopic model is probable not be able to offer a data group that is informative enough for CTM model tuning. For example, if the demand is strongly feasible for each cycle, the cell capacity will never be shown from the obtained data. Therefore carefully designed data obtaining methods and related calibration process are necessary for effective application.

The rest part of this chapter introduces the general steps of DFF. Furthermore, those steps are explained in detail one by one. The last subsections introduce practical considerations during application.

5.5.1   General description

The main purpose of DFF is to extract information from short time microscopic model runs, and to inject them into macroscopic model. The target of this process is to improve the accuracy or / and analysis ability of macroscopic level.

To tune parameters of macroscopic level can improve the accuracy of macroscopic level, especially the ability of prediction. Although a normal application of macroscopic modeling calibrate the parameters carefully, when the simulated scenario changes, the feasibility of the calibrated model drops largely. In other words, those applications use unique parameter values and unique setting for all modeled scenarios. This situation reduces analytical accuracy by using improper parameter values

and ignoring influences of scenario changes. DFF calibrates parameter values for the current simulated scenario according to the corresponding microscopic model. If there is big difference between microscopic results and corresponding macroscopic results, DFF adjusts macroscopic parameters in order to reproduce results that are closer to microscopic results, which can be considered as the best simulation performance we can get. Besides the scenario changes, CTM also lack of description ability for some objects, e.g., pedestrian crossing, lane-changing related capacity drop, variable cell capacity, etc. The influences of those factors have to be described by modifying model parameters (or inputs). Essentially, it is an indirect solution but not a direct one as microscopic model does.

In order to realize the two stated targets, the following process is used (Figure 33).

Those steps included in this process need clarification:

- Model initialization: the system initialize models on two levels, the tasks including model loading, set reading, etc.;

- Microscopic model probing run: the system runs the microscopic model for several times with different random seeds for predefined simulation period. During this process, data for ACM is recorded. In addition, the system judge the stability of results in order to determining whether finish the data collection process;

- Micro level result selection: the system selects the suitable data group, which is considered most representative, as the reference data of ACM;

- Macro level model calibration: the system calls predefined calibration algorithms, tunes selected parameters of the macroscopic model, in order to obtain better model performance. This is the first step of ACM.

- Macro level model fast validation: because of the attributes of performance based calibration, several groups of values of parameters are possible to produce simulation results that are similar good (by coincidence). However, some of them may not suitable for other data group. In order to filter those coincidental groups, the system has to operate a validation process. This is the second step of ACM.

- Macroscopic model run: based on the calibrated parameters, the system runs the macroscopic model for desired analytical period; the outputs from macroscopic level are used as the final results of the analytical process.

As described about, the Advanced Consistency Maintenance (ACM) is realized by two sequential steps: macro level model calibration and fast validation. Those two steps agree with the basic routine of model calibration. However, there is at least one main advantage can be utilized relating to this specific case. The simulation based data source is more convenient to be accessed. Because the data source is a microscopic model, a user can set data collection point at any desired location, and the data can be obtained at any desired time intervals. The limitations that exist when realistic data source is used do not exist with this case, such as damaged detectors, fixed locations, etc.

Last but not least, before implementing the DFF, it is assumed that the CTM network is properly built, which means this macroscopic model is comparable with the corresponding microscopic model. However, this thesis does not focus on the methodology of modeling urban or freeway network with CTM (these problems are highly investigated by existing research). However, the points worth noting will be mentioned when talking about examples.
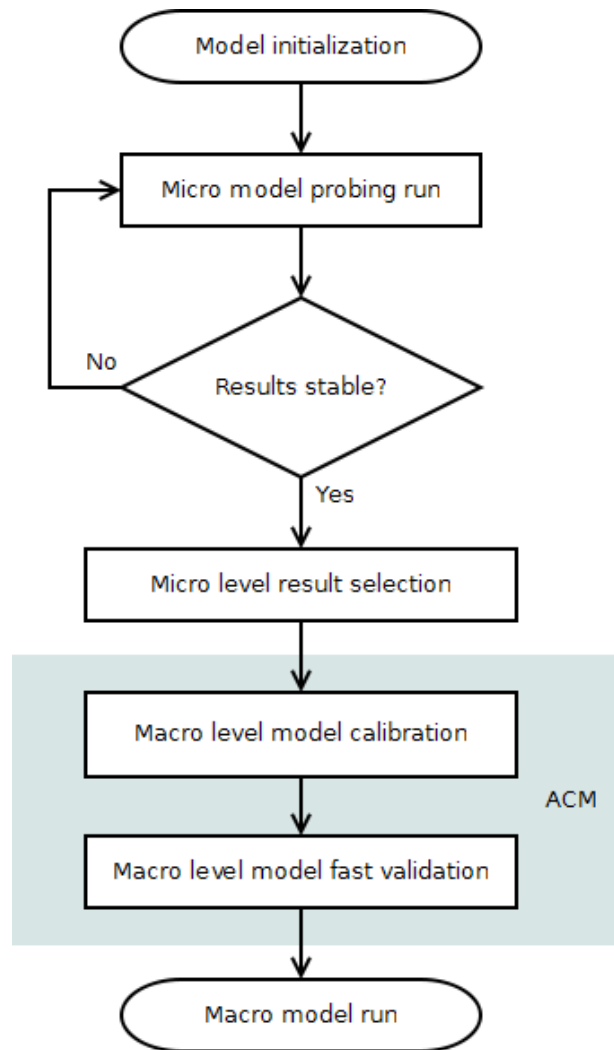
Figure 33 DFF general process

5.5.2 Model performance indicators for model calibration and fast validation

HyTran uses traffic volume data from several control points at specific time intervals as indicators of model performance. The discrepancy of traffic volumes between corresponding data collectors (in microscopic model and macroscopic model) is used to measure the difference between them. In addition, to minimize the measured discrepancy is the main tasks of the calibrating operation.

To use traffic volume as indicator is based on three considerations:

First, traffic volume data is easy to be obtained from both macroscopic model and microscopic model. In contrast to traffic volume, other indicators, which are widely used for traffic evaluation, such as delay, queue length, travel time, etc., are hard to be estimated effectively, especially when considering the different definitions of them in macroscopic and microscopic model. For example, the delay data obtained from microscopic model is probable to be different from the same data

obtained from macroscopic model, even if the traffic states in two models are same. This is because the delay cannot be measured directly, but has to be estimated based on the basic flow states, and different model has its own estimation method. However, traffic volume data is always obtained according to its definition in any model. Therefore, the system does not have to consider the difference between the indicator definitions of two models.

Second, from the perspective of traffic engineering, the traffic volume at special time intervals are widely accepted indicators for analysis, both for simulation models and field data. In the realistic world, to measure delay or stops are more difficult (the results highly depend on the used algorithms for estimation). On the contrary, loop detectors that are widely equipped worldwide on roads offer reliable traffic volume data effectively at flexible time intervals.

Third, when collected at suitable time intervals, the traffic volume data is also representative for vehicle density situation between two control points, if the measurement starts with the empty network. The average vehicle density for road segments (between two control points) is base of calculation of further indicators, such as average delay.

For HyTran, to select proper locations of control points is an important task. The selection should ensure that the attributes of simulated network can be fully shown, on the other side, to have fewer control points is possible to reduce the computational complexity of calibration process. In summary, a logical and effective way is to decide the control points' locations in terms of specific cases, but not by an abstract description.

5.5.3 Calibrated parameter preparation

Three questions have to be answered calibrated parameters:

1. What parameters are chosen to be calibrated?

2. How to decide their possible ranges?

3. Which part of the network share the same values of parameters?

These two points also have to be determined according to specific application.

Calibrated parameters

Generally speaking, the factors that satisfy two conditions will be considered for calibration: (1) factors affecting the model performance; (2) factors that are not clear and fixed. Based on these criteria, the calibrated parameters used in DFF come as two categories:

- Basic parameters of CTM

- Control parameters

The group of basic parameters of CTM includes free flow speed, backward wave speed, cell capacity, and jam density. Furthermore, for different part of the network, the values of parameters may vary. The question of how to cluster the cells sharing same parameter values has to be answered according to specific case. The model structure is the main factor for deciding this.

The group of control parameters mainly refers to the effective green time that is used in macroscopic model. Generally speaking, a normalized effective green time calculation method is probable to be not suitable for different cases. Because of the existence of the microscopic model, it is a dependable reference for obtaining better description of effective green time.

Value range

Except for determination of selection of calibrated parameters, to determine the acceptable range of each parameter value is important for effectiveness of calibration process.

The principle of determine parameter range can be described as following: first, to determine an average value. The average value can be obtained from experiences (e.g. the cell capacity, free flow speed), or normalized calculation (e.g., $\omega = v/2$); second, to determine the upper bound and lower bound, according to the average value. Generally, the upper bound and lower bound could be symmetrical to the average value, and determined by experiences.

Network partition

Besides the selection of calibrated parameters, the selection of calibrated parts of the network is also important for the final performance of calibration process. In order to express the attributes of the simulated road network, different parts of the network are probable suited by different values of parameters. For example, the capacity of a specific leg of an intersection is largely depended on how strong is the demand for lane change happens in the entrance link. This is to say, different legs are probable have different parameter value to each other. The determination on selection of calibrated links depends on the practical condition of the modeled network, including link channeling, signal program, or even the demand intensity, therefore, it has to be determined with specific cases. However, one extreme solution is to make each cell has their own values of parameters, it is not feasible because it will increase the dimension of calibration to be hundreds-order. This situation will bring impassable obstacle for the selected algorithm for calibration.

5.5.4 Advanced Consistency Maintenance (ACM)

ACM consists of two main components: calibration and fast validation. For implementing ACM, necessary data set has to be obtained from the corresponding microscopic model.

Requirements for data

The calibration and fast validation process have their respective requirements on the data set. And the requirements represent the different attributes of those two processes.

As discussed in section 4, the dynamic behavior of CTM is determined by its capacity. When the demand is strictly feasible, CTM tends to converge to a stable state in which the cell densities have a converged solution; when the demand is not feasible, a blockage happens on specific point, queue starts to form until the demand recover to feasible. In other words, if the demand is always feasible, the simulated traffic volumes of CTM do not show significant difference even different parameter values are adopted. However, the different attributes of a network emerges when demands that is not always feasible are loaded on the simulated network. Therefore, to find out the capacity of each road segment becomes the key task of calibrating a CTM network. In a word, the capacity performance of each part of a model plays an important role in accuracy of a model.

Generally speaking, the segment capacity obtained by CTM is likely higher than the realistic condition (some exceptions may exist when modeling intersections), because the model does not express the spontaneous capacity drop. The flows in a CTM network moves in a more ideal state than the realistic network or a microscopic model. However, to describe those phenomena happens to be the advantage of microscopic model. In order to obtain a more accurate CTM, DFF tends to investigate the capacity of intersection or specific segment, such as a merging area, by performing a probing microscopic model run.

However, to obtain accuracy segment capacities is not an easy task for microscopic model, even the microscopic model is sufficiently calibrated. Because whether the simulation run will show a

segment's real capacity are affected by various factors, such as the input volume, the variety of input volume, and difference of turning percentages. Considering those practical difficulties, DFF set the demand volume to be 10% to 20% higher than the normal demand according to the basic demand. The higher demand volume will lead to more crowded traffic state than the basic demand, which can be considered closer to the saturation state than the basic demand.

The fast validation process needs to check if the calibrated parameters can lead to suitable performance for the basic demand. Therefore, when obtaining the data set for validation, the basic demand is used.

The target of the probing runs is to obtain the reference data for macroscopic level calibration. In addition, the simulation period for each probing run should be as short as possible if it guarantees to show the basic attributes of the simulated objects. The detailed introduction about probing runs, especially about the run times, will be discussed in the next subsection.

Calibration and validation

When the required data has been collected from the probing runs, calibration process can be started. The calibration work is based on one of the algorithms between SPSA and BPF. When SPSA is adopted, the calibration process is an optimization problem. When BPF is adopted, the calibration process is a state estimation problem. The detailed operations about those two algorithms have been discussed in section 5.5.6.

The calibration process is probable to generate several groups of parameter values that have similar effect according to the calibration data. Those groups of parameters values are considered as the candidate group. The process of fast validation chooses the best parameters from the candidate groups. Generally, the fast validation needs a different data group which has not been used for capacity calibration; it is the validation data that has been discussed.

5.5.5 Data acquisition

In order to obtain the reference data for calibration and validation, several problems have been clearly considered.

Simulation period for Probing runs

For obtaining calibration data, the main purpose of probing microscopic model runs is to investigate the performance of the specific model when the demand is higher than the normal level, ideally, the capacity of them emerges. Therefore, the running period is not necessary to be as long as the real evaluation period, because if the demand is high enough, the capacity volumes emerge in short period. For an intersection, the probing period should be longer than at least one signal cycle. In order to avoid the biased value arouse by random attributes of microscopic model, a probing period that is longer than 3 times of signal cycle seems logical.

For obtaining the validation data, the main purpose of probing microscopic model runs is to get sample of model performance under normal demand level. Because of the demand level used for validation data is lower than the probing runs for calibration data, and this proving run will only be done for one time (explained in the next paragraph), a longer running period can be adopted. 10-15min is a recommended selection.

Number of runs

Different to CTM, microscopic model takes the randomness of modeled system into consideration. The benefit of this strategy is to model the realistic behavior more naturally; however, it brings

instability to the simulation results. Therefore, when using microscopic simulation results as reference data for CTM calibration (and validation), multiple runs are necessary for obtaining unbiased data.

(FHWA, 2004) describes a guideline for determining necessary model running times. This method makes an independent samples t-test on collected data by assuming the data following a normal distribution. This t-test shows whether there is a significant difference between the obtained mean value and true mean according to the desired confidence level (CI). The test statistic can be expressed as equation 5-18.

$$T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \sim t(n-1) \qquad\qquad \textbf{5-18}$$

Where,

s: the standard deviation of the obtained model results;

n: the number of data;

μ_0: the true mean;

$\bar{X}$: the sample mean;

α: the confidence level;

This test can be used to test whether $\bar{X}$ is significantly different to μ_0, according to predefined confidential interval and confidential level. Based on this setting, the minimum required repetitions can be calculated by equation 5-19 .

$$N = \left(2 * t_{(1-\alpha/2)}(N-1) \cdot \frac{s}{CI_{1-\alpha\%}} \right)^2 \qquad\qquad \textbf{5-19}$$

Where,

$t_{(1-\alpha/2)}(N-1)$: the student's t inverse cumulative distribution function, with the degree of freedom of $N-1$ (N data points) for the corresponding probability in $(1-\alpha/2)$;

$CI = X - \mu_0$: the confidential interval;

There are two ways for applying this conclusion. The first one is to determine CI at first, and calculate N by results of n runs. And then, to check if $n \geq N$. If it is true, it means the current simulation repetition is acceptable; otherwise, add one more repetition and do the comparison again until the repetitions become acceptable. Another more concise way is to set the desired value of CI/s instead of CI. Using this way, desired CI depends on s, which means the minimum required repetition can be decided without simulation results.

A table (Table 8.) is illustrated in (FHWA, 2004) for several typical values. For example, given $CI/s = 1.5$ and $1-\alpha = 0.9$, the minimum required repetitions are 9.

Data selection

After determining the minimum required repetitions, HyTran has to determine a specific data group used as reference of calibration process. A typical evaluation by microscopic model tends to use the mean value of results as a final result; however, this solution is not suitable for HyTran. One reason is that if the mean value is selected, no single run can produce the same result, which means the different items are not comparable with each other.

In HyTran, an alternative method is used: the result that is most similar to the mean value is selected as the reference data. The similarity is measured by a root-mean-square normalized deviation (RMSND).

$$D = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{m_i - m_i^{ref}}{m_i^{ref}}\right)^2}$$

5-20

m_i is the $i\,th$ measured value, m_i^{ref} is the corresponding reference value.

The scenarios that do not include intersections use traffic volume data on control for similarity evaluation; the scenarios that include intersections can also use turning percentages for similarity evaluation. In addition, for these scenarios, the predefined turning percentages obtained from external assignment model are used as the first reference values, and traffic volumes on control points are used as the second reference values. This decision is based on the variance of simulated turning percentages for short time period in microscopic model.

One concise method to choose the reference data of validation is to use the same random seed as the selected capacity data.

5.5.6 Application consideration

When applying DFF, several points are worth noting:

- Turning percentage consistency: although turning percentages can be set for each intersection in microscopic model, the stochastic characteristic of microscopic model leads to a random performance of the simulated turning percentages. Those turning percentages approximate to the predefined value as time passing. However, for a time period as short as 5min, the simulated turning percentages vary largely. In order to make the volumes in intersections comparable, the temporal turning percentages that are calculated by the reference data are used in CTM for the validation process if necessary.

- Calibration period: one goal of DFF is to minimize the simulation time of microscopic model; however, the minimum proper time period is difficult to be determined. One possible principle is that the calibration period should cover several signal cycle, and definitely not shorter than one cycle. Another principle is that the selected time period should be long enough for showing the basic attributes of the modeled network.

- Warming period: in a typical application of microscopic model, the warming period is necessary in order to avoid the biased evaluation results that are affected by abnormal network state at the beginning part of simulation, and the data obtained during this time period should not be used for calibration or validation. However, because the CTM also starts with an empty network, the warming period is not necessary for HyTran calibration. In other words, the reference data are obtained from the very beginning of the simulation to the desired time point, therefore they are directly comparable.

There are several points needing clarification when using SPSA:

- The parameters of SPSA: The parameters, including a, c and A, affect the stability of SPSA. Because different dimensions of the parameter vector (θ) has different scales, c is determined by measure the possible range of value for each specific dimension. And a is determined according to observing the change interval in each iteration. A is set to 100.

- Starting point and allowance range: the starting point of parameters is the point consisting of the common values of those parameters. The allowance ranges are determined according to maximum and minimum value of specific parameter.

There are several points needing clarification when using BPF:

- Initialization of particles: a uniform distribution can be used to initialize the particles, which means the original particle set can be generated by the equation 5-21, where, θ_L and θ_U is the low and up limit of current parameter, δ ($0 \leq \delta \leq 1$) is a uniformly distributed random decimal.

$$\theta = \theta_L + (\theta_U - \theta_L) \cdot \delta \qquad \textbf{5-21}$$

- Covariance matrix for observation noise: When the flow data on control points are obtained for N random seeds (as described in section 5.5.5), N groups of data are obtained. This data is used to estimate the covariance matrix of the observation noise by definition (equation 5-22).

$$Cov(X_i, X_j) = E\big[(X_i - \mu_i)(X_j - \mu_j)\big] \qquad \textbf{5-22}$$

5.6 Comparison of FTF and DFF

Compared with FTF, the main difference of DFF is to employ a calibration method as an additional information interface between those two levels, in order to realize the advanced performance consistency (by ACM). The possible benefits of this strategy and the weak points that cannot be ignored come from three points:

First, in FTF, microscopic model has to keep running for the entire analysis period. On the contrary, in DFF, the microscopic model only runs for a short period to show the attributes of the simulated objects. This method can save computational resource, especially when the evaluation period is long and total traffic volume large. In microscopic level, all the vehicles that run into the simulated network will consume memory and to simulate their behaviors always consume the computational resource. As the amount of simulated vehicles increases, the simulation running speed drops significantly. When vehicles are accumulated inside a network, the simulation speed after a relative long period, such as 30 minutes, this phenomenon is probable to emerge.

Second, the macroscopic level results are more stable and are probable to be more representative. There is a difference of running process between macroscopic models and microscopic models. Because the macroscopic model focuses on offering stable and representative results for the simulated system, stochastic behaviors are not described. This feature makes the user of

macroscopic models only need to run the simulation once. In contrast to macroscopic one, microscopic models consider the stochastic behaviors of traffic participants, therefore the results changes with different random seeds even if the basic input does not change. This feature makes the simulations with microscopic models need multiple runs, otherwise the result maybe the very biased one. Therefore, in FTF, an unavoidable problem worth arguing is about the random seed selection. When a simulation runs with a selected random seed, it is difficult to ensure the results obtained with the selected random seed represent the realistic performance of the simulated objects. It is possible that the selected random seed leads to very biased results. However, to carry out multiple runs with different random seeds in FTF will reduce the advantage of hybrid solution in terms of saving computational resource. Compared with FTF, DFF avoids this problem by introducing a data process based on multiple random seeds.

Third, although the method for flow transition has been discussed widely, there are still several problems are pending and seems difficult to be solved. In addition, more tough problems come from the technological perspective. One problem is about the lane and position selection when generating vehicles on a microscopic link. From the perspective of technology, to add extra vehicles into a running network will bring extra computational and operational time. VISSIM, for instance, has a COM interface function can be used to add extra vehicles into an existing running network, but when the amount of added vehicles is large, the speed of simulation will be reduced significantly. A possible reason for this phenomenon is that the inserting point of vehicles is not determined carefully enough, therefore some of the vehicles changes speed instantly, which needs much more computational resource than a freely running vehicle. One possible solution for solving this problem is to use a more sophisticated algorithm to determine the speed, lane, and location of the newly inserted vehicles. But, there is a dilemma that this method will also introduce more operations about reading data and calculation, which also reduce the speed of running significantly. In contrast to FTF, DFF does not depend on the flow transition, and during the microscopic model run, the outside interruption is limited as few as possible. In other words, the computational resource is mostly devoted to the calculation that generates useful results, but not handling the flow transition, which does not bring any useful information directly. From this point of view, DFF is more suitable to be implemented with the developed simulation tools.

Although DFF has those benefits and FTF has those disadvantages, FTF has its own advantages and DFF also has several (possible) weak points.

The most valuable advantage of TFF is that the keep running microscopic model can bring the maximum analysis ability. A premise of DFF is that the possible phenomenon has to be able to be described by the basic parameters of macroscopic model. In other words, when some items are very difficult to be described by macroscopic model (at least by CTM that is adopted now), FTF is still the first choice. For example, CTM is not designed for evaluating a scenario with adaptive signal control for bus priority. However, this object can be modeled by microscopic model very well. Because of the limitation of the model structure of CTM, if new extension on CTM for bus priority is not developed, using microscopic model is still the best choice.

The second point that may be argued about is the effectiveness of the calibration based data fusion. DFF is based on a basic premise that if the parameters are well tuned, CTM is able to simulate the model sufficiently enough. However, in a practical example, it is possible that this premise is false. Fortunately, the existing applications of CTM cover diverse cases and scenarios; this is a strong evidence for the premise.

The third possible weak point of DFF is about simulate time. The original idea of applying hybrid traffic simulation is to reduce the needed simulation time. However, in DFF, at least one component will need extra simulation time: the calibration algorithm. If the calibration process needs too much time, and the total running time is possible to be larger than FTF.

In HyTran, FTF and DFF are both implemented, because it is hard to say which one is superior to another till now. In addition, to implement those two strategies does not face any unsolvable conflicts.

5.7   System architecture

In order to implement and test the basic methods and algorithms involved in HyTran, we prototype HyTran with a series of code packages. The main code is programed in Matlab, which ensures the code to be flexible. In addition, the most appealing benefit of coding in Matlab is to utilize the highly optimized mathematical functions integrated in Matlab. This benefit accelerates the prototyping process significantly.

The general structure of the experiment environment is described in Figure 34. CTM and VISSIM work as two elementary components of HyTran. FTF and DFF are both based on them. CTM network covers the entire simulated object; VISSIM model covers some parts of it, but no more areas than what are covered by CTM. DFF and FTF work as two parallel approaches. DFF includes 3 modules: probing run, ACM, and Macro run. The probing run module run the VISSIM model as probing runs for obtaining reference data; the ACM module operates CTM calibration and validation, based on the reference data offered by those probing runs; the CTM, whose several parts are calibrated, is used for final evaluation run and get the output results. For the FTF, a module VIS-MM, which is coded in Matlab based on VISSIM COM interface, is used to implement the flow transition method.
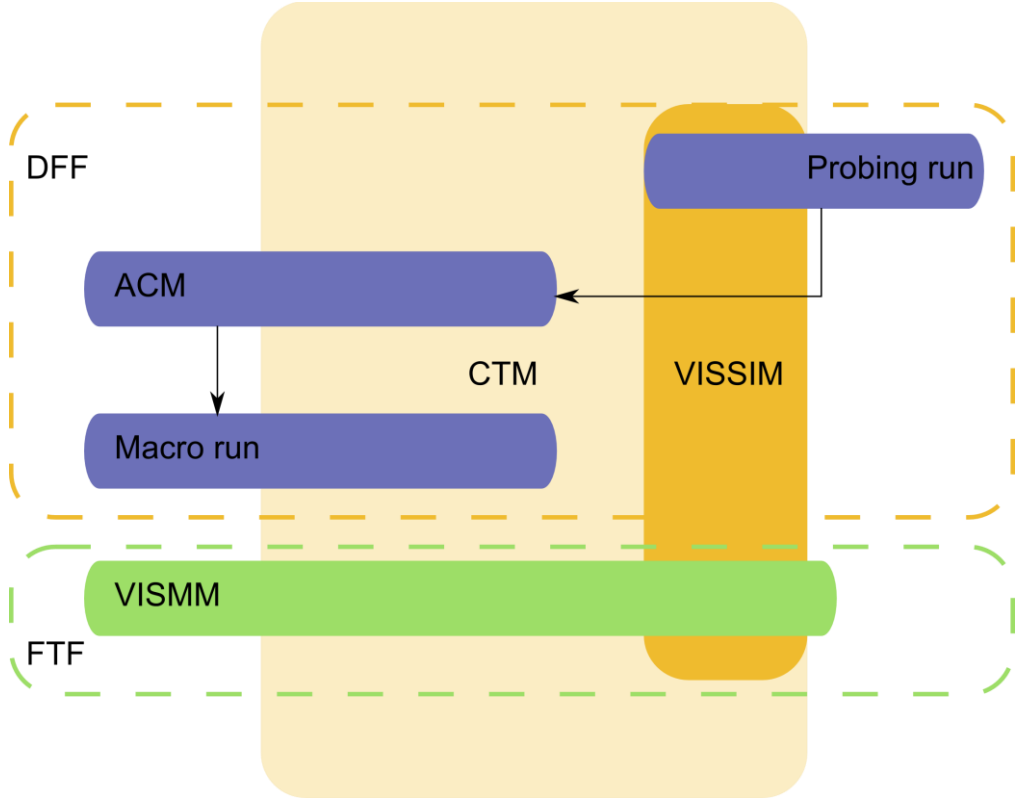
Figure 34 System structure

All those code are implemented on a PC with Intel Core Duo E8600 @ 3.33GHz processor and 4GB memory. Matlab R2010b on MS Windows is the main coding environment.

An introduction to the developed software environment is presented in Appendix A.

6   Evaluation of HyTran

In this chapter, we investigate the feasibility, effectiveness, and limitation of HyTran by several cases. First, the two main components of HyTran, FTF and DFF are tested separately in terms of their own functions. Afterward, an example of an urban intersection from a realistic road network in Graz (Austria) is used to compare the performance of FTF and DFF.

6.1   Determination of basic parameter values of CTM

The tests operated in this chapter need a set of basic values of CTM parameters.

Considering one of the targets of this thesis, to investigate the difference of model performance between microscopic model (VISSIM) and macroscopic model (CTM), the basic set of parameter values is measured from the data obtained from VISSIM simulation runs. In other words, these values correspond to the default setting in VISSIM model.

The free flow speed is determined as the value of common free flow speed in a urban area, i.e. $v = 50 \, km/h \approx 14 m/s$. This value is also reasonable for the default value of urban traffic composition in VISSIM simulation, although VISSIM considers a stochastic distribution of desired speed for each simulated cars.

The cell capacity is determined by considering the observed maximum traffic flow rate in multiple VISSIM simulations. A typical value is set to $q_m = 0.7 \, veh/s/lane$.

The jam density is determined by the observed maximum vehicle density in simulation examples with queues, i.e. $k_{jam} = 0.15 \, veh/m/lane$.

The determination of backward wave speed is based on the flow-density data sets obtained from VISSIM simulation. The data set is filtered in order to pick out the data points which are under the situation of queue dissipation but not queue forming. We operate a constrained least-squares fit on the filtered data set. The final used value is set to $\omega = 7m/s = v/2$.

In addition, the simulation cycle is set to 1 second in this paper. Correspondingly, the cell length is ensured to be longer than 14 meters. Correspondingly, the used VISSIM model has the simulation cycle of 0.2 seconds.

6.2   Basic evaluation of FTF

6.2.1   Boundary interface with original CTM and IFCTM

Under FTF, the boundary performance of flow transformation affects the effectiveness of the hybrid approach significantly. Generally, we hope the adopted interface has impact on the traffic flows as weak as possible. This section evaluates the dynamic performance the boundary flow of FTF, both from macroscopic to microscopic and microscopic to macroscopic.

Figure 35 shows the basic testing cases. The first case (Figure 35 (a)) tests the performance of boundary transition from a macroscopic flow to a microscopic flow (M2m). A 280 meters long test link (2-lane) is equally divided into 2 segments; each of them is 140 meters long. The upstream segment is modeled with CTM (IFCTM), and the downstream segment is modeled with VISSIM. The second case (Figure 35 (b)) tests the performance of boundary transition from a microscopic flow to a macroscopic flow (m2M). Similar tested object is used as the first case; however, the main

difference is the upstream segment is modeled with VISSIM and the downstream segment is modeled with CTM (IFCTM).

For both two cases, data collectors are located close to the boundary and both at the macroscopic side and microscopic side. In addition, a signal light is placed at the downstream end of the simulated link, which can produce interruptions to a stable traffic flow.

The main purpose of this setting is to find out how well the traffic flows are transformed without delay and errors. Two indicators are used to show the performance: accumulated traffic volume at the boundary points, and average density of the entire links.

For each test, the entire simulation lasts 600 seconds. The signal light is set to red from 120s – 360s (lasting for 4 minutes). During this period, traffic flows are blocked; therefore, this setting gives rise to a queue forming and dissipating process. Therefore, the examples show whether the queue propagate across the boundaries. In addition, three different demand values are used for test, including 1000 veh/h, 1500 veh/h, and 2000 veh/h.

In order to evaluate the influence of the boundary interface, a corresponding microscopic link (280 meters long) is used as a reference (Figure 35 (c)). In the reference microscopic model, the test set one data collector at the location (140 meters away from the upstream end) where the boundary is.

In the two tests, CTM and IFCTM have 14 meters long cell. The basic values of parameters determined in section 6.1 are used. And a matched setting in VISSIM is used. The Root Mean Square Normalized Error (RMSNE, defined by equation 5-9) is used to evaluate the results. The data of accumulated traffic volumes are sampled at the time intervals of 60s. For the average link density, average values for each 300s (5 minutes) is used.

Figure 36, Figure 37, Figure 38, and Figure 39 show the summarized simulation results. The accumulated volume curves are included in Appendix B.

(a) M2m

(b) M2M
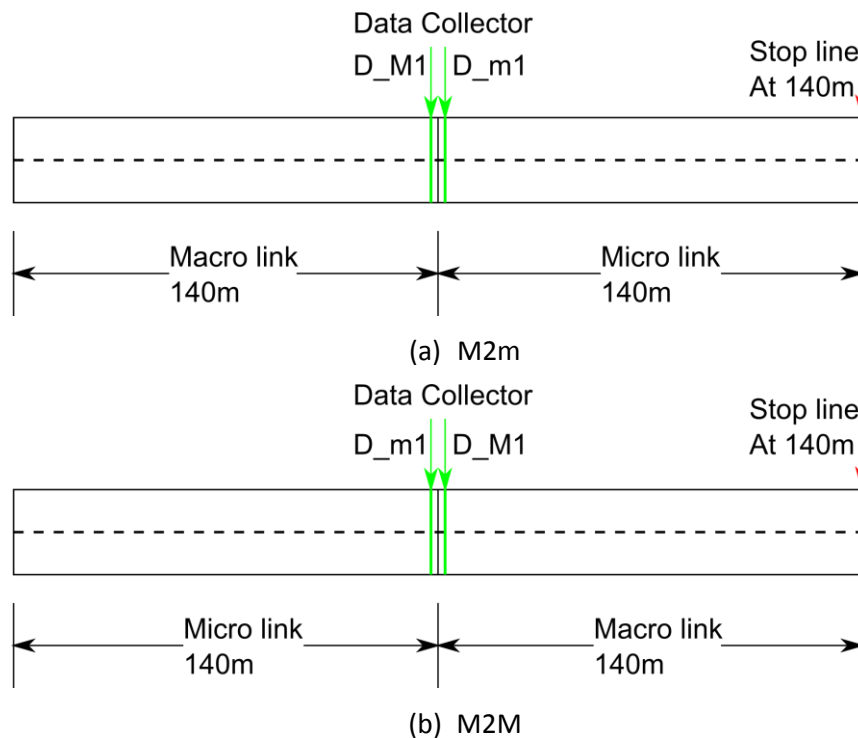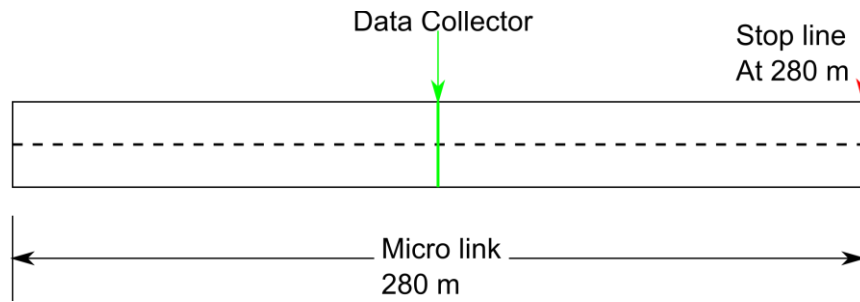
(c) Microscopic reference

Figure 35 Test cases for boundary interface

Based on the simulation results, several points can be discussed as following:

- The accumulated volume curves show that the discrepancies between macroscopic side volume and microscopic side volume are ignorable. This is a logical consequence caused by mechanism of the flow transition algorithms. This means the developed flow transition algorithms maintain the volume consistency at boundaries successfully. Therefore, in the volume comparison (in Figure 36, Figure 37, Figure 38, and Figure 39), only one data (the upstream side data) is used to be compared with the reference data.

- The queue forming and dissipation propagates smoothly from one more to another. However, there are discrepancies between the hybrid model and reference model in terms of accumulated volume and link density. Generally speaking, the discrepancy increases with the demand raising.

- In m2M, flows with CTM start up earlier than the reference flow; the application of IFCTM alleviates this phenomenon. In M2m, flows with CTM have a larger volume rate at the beginning of the flow starts; the application of IFCTM also alleviates this problem.

- Generally speaking, IFCTM has better boundary performance than CTM. However, the difference varies with different demand level. For example, in the volume comparison shown in Figure 38, when the demand volume is 1000 veh/h, the difference of simulation results are fair small (from 4.3% to 4.1%), however, when the demand volume is set to 2000 veh/h, the same item shows higher difference (from 6.6% to 2.0%).

	m2M	M2m
volume	Figure 36 Volume error comparison under different demand intensity, m2M	Figure 37 Volume error comparison under different demand intensity, M2m
density	Figure 38 Density discrepancy under different demand intensity, m2M	Figure 39 Density discrepancy under different intensity, M2m

6.2.2 Influence of Mandatory lane change (MLC) for FTF

As described in section 5.3.2, the method for mandatory lane change (MLC) aims at reducing the unexpected capacity drop aroused by excessive demand for lane change. In this chapter, a test for with and without the algorithm of MLC is carried out, in order to compare the capacity changes under those two situations.

The performance of MLC is tested by the following example. Figure 40 shows the basic setting. A short diverging segment is modeled as a 2 lane link. During running on this link, vehicles have to change their lanes if it is necessary, in order to run into their desired downstream branches. In this test, different settings of turning percentages are used (1:1, 2:2, and 3:1). The different levels of demand volumes are used: 2500 veh/h, 3500 veh/h, 4500 veh/h. The VISSIM model uses its default setting. The observed traffic volumes for 15 min are used as the indicator.
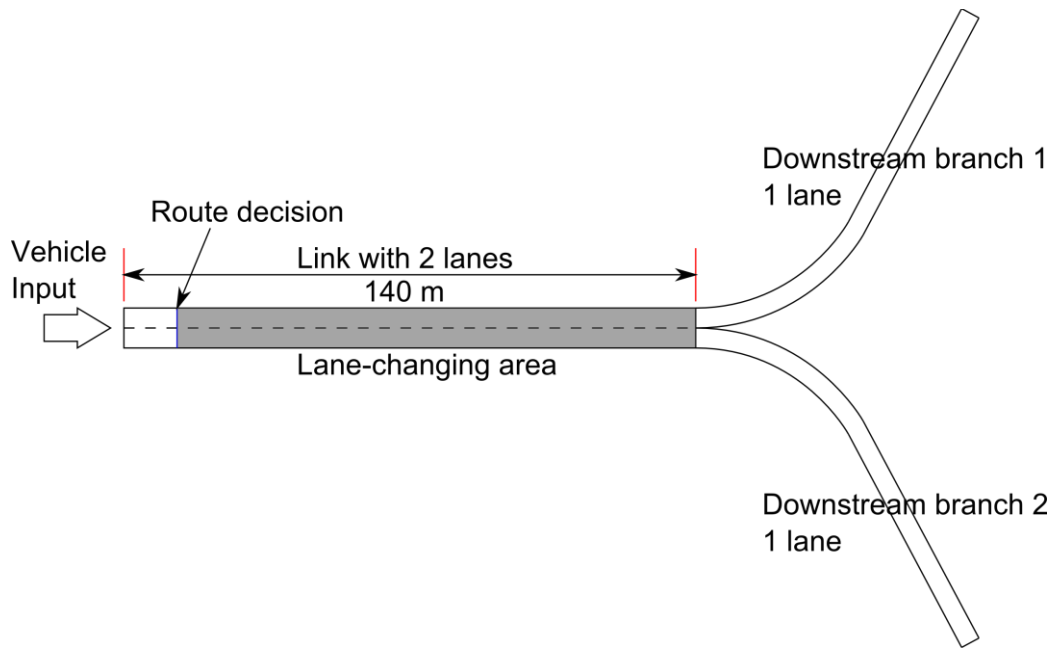
Figure 40 Example for short weaving segment

The simulation results are shown in Table 4. In this table, the simulated traffic volumes for 900s on two branches are listed. It is obvious that the MLC increases the passing volumes.

Table 4 Simulated traffic volume for different scenarios

Demand [veh/h]	Turning ratio	With MLC?	Downstream branch 1 volume [veh for 900s]	Downstream branch 2 volume [veh for 900s]	Total volume [veh for 900s]
2500	1:1	Y	276	337	613
		N	278	334	612
	2:1	Y	386	226	612
		N	384	225	609
	3:1	Y	444	167	611
		N	432	164	596
3500	1:1	Y	389	459	848
		N	335	404	739
	2:1	Y	500	303	803
		N	457	271	728
	3:1	Y	508	191	699
		N	484	178	662
4500	1:1	Y	435	511	946
		N	336	403	739
	2:1	Y	513	308	821
		N	470	277	747
	3:1	Y	502	186	688
		N	501	186	687

Figure 41 shows the rate of increase of the total traffic volumes when MLC is effective, compared with the same simulation but without MLC. As shown in this figure, when the demand is 4500 veh/h and turning ratio is 1:1, MLC has most significant influence (28% increase).

Because MLC is based on exchangeable pairs, it does not increase the link capacity into an unreasonable level. For example, when demand level is 4500 veh/h, and turning ratio is 3:1, MLC shows similar total passing volume as the model without MLC. This phenomenon can be explained as: when the turning ratio is higher, fewer exchangeable pairs are found, which means the efficiency of MLC is weakened under this situation.
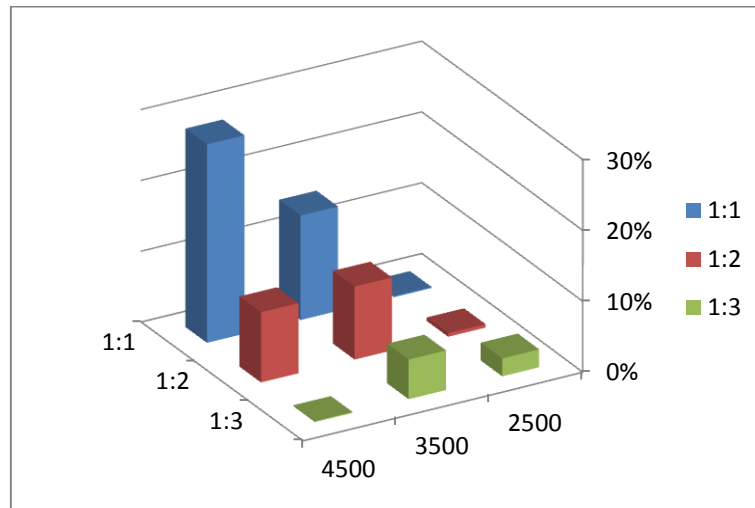

Figure 41 Increase of total passing volume when MLC is effective

6.3 Basic performance of calibration algorithm

This section discusses the effectiveness of the two calibration algorithms discussed in section 5.3.4; furthermore, the performance of them are compared according to the studied example. This test focuses on two points: 1) how well do the calibration algorithms work to obtain more suitable and representative values for those calibrated parameters; 2) how well is the CTM able to reproduce traffic flow state that has been described by microscopic mode. The first point is a general purpose for a common calibration process. The second one is special for HyTran because HyTran tends to decide under which situation the DFF works and under which situation FTF has to be used. However, the second point will also be discussed in the next case study in section 6.4.

The basic tested example can be described as following (Figure 42): a two-lane road (link 1) leads to two branches (link 2 and link 3); each of them has one lane. Vehicles run into the simplified network from the upstream end of link 1. After assigned with their route information (to link 1 or to link 2), vehicles run into their desired branches. Therefore, there will be lane-changing behaviors happening in link 1. The effects of lane-changing cannot be predicted by CTM, but can be well described by VISSIM. This example studies DFF's ability of adjusting the CTM model, in order to maintain the advanced consistency with microscopic model.

Three data collectors (DC_1, DC_2, and DC_3) are set on the links. DC_1 locates at 20m away from the input point; DC_2 locates on the connector from link 1 to link 2, and DC_3 locates on the connector from link 1 to link 3. This network is modeled both using VISSIM and CTM independently.

For testing the effectiveness of the calibration process, several factors have been considered:

- Change of desired speed of input vehicles;

- Change of traffic composition;

- Different demand of lane change behavior;

- Different input volumes;

For each scenario, 9 different random seeds are used. The data group whose total input volume has smallest discrepancy with the average value is selected as the reference. When several data groups have similar total input volume, the one whose simulated turning percentages match the predefined value best is selected.

The calibration period is 300s (5 minutes). Input volumes vary among 3 levels: high demand (3500 veh/h), middle demand (2500 veh/h), and low demand (1500 veh/h); three different traffic compositions are used, including: C1 (the default traffic composition in VISSIM) : 98% car (50 km desired speed), 2% HGV (50 km/h desired speed); C2: 60% car (40km/h desired speed), and 40% HGV(30km/h desired speed); C3: 100% car (30 km/h desired speed). Two different turning ratios are used: 1:1 and 4:1.
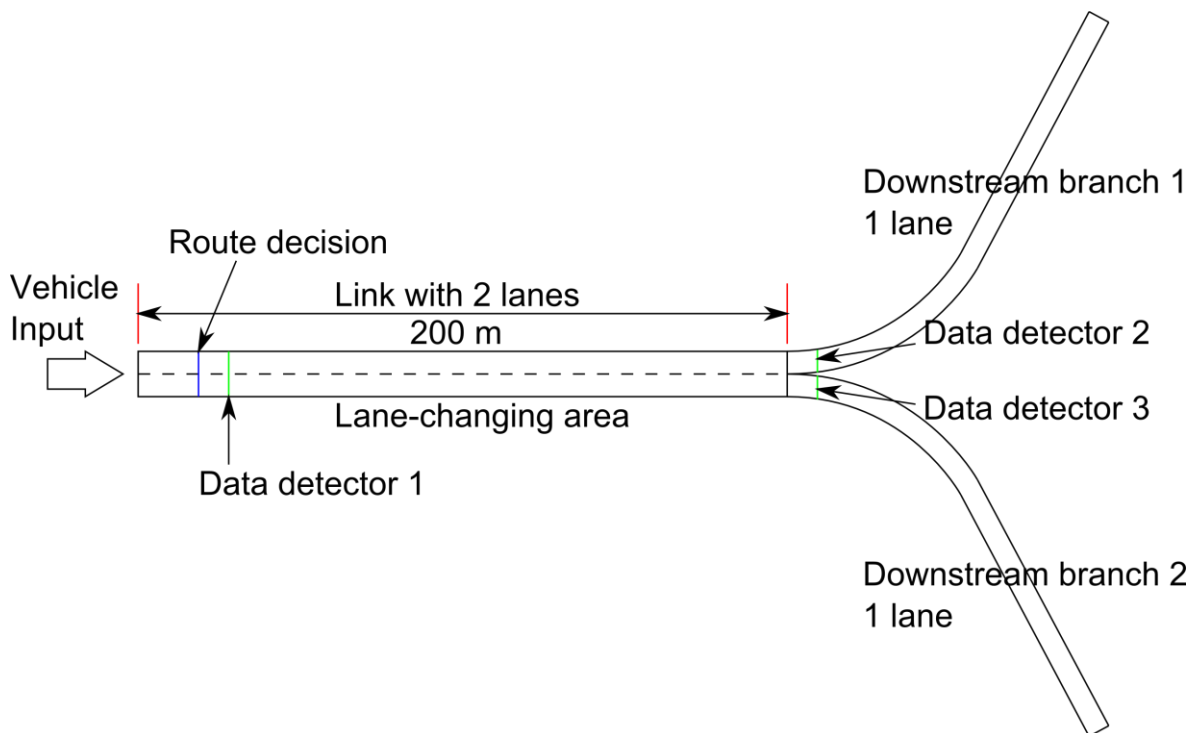
Figure 42 T-intersection profile with no traffic signal control

The total traffic volume data of 301-1800s is used for validation. However, the purpose of the validation process operated here is different to that discussed in DFF. The purpose of the validation process here is to prove the effectiveness of calibration process.

The test plan is shown in Table 5; this plan includes 18 tests.

Table 5 Test plan

No.	Demand Intensity	Traffic Composition	turning ratio
1	2500	3	1:1
2	2500	2	1:1

No.	Demand Intensity	Traffic Composition	turning ratio
3	2500	1	1:1
4	1500	3	1:1
5	1500	2	1:1
6	1500	1	1:1
7	3500	3	1:1
8	3500	2	1:1
9	3500	1	1:1
10	2500	3	1:4
11	2500	2	1:4
12	2500	1	1:4
13	1500	3	1:4
14	1500	2	1:4
15	1500	1	1:4
16	3500	3	1:4
17	3500	2	1:4
18	3500	1	1:4

In the calibration tests, link 1, link 2, and link 3 have independent parameter values; therefore, the calibrated vector is expressed by:

$$\theta = [v_i \quad \omega_i \quad q_{m,i} \quad k_{j,i}]_{i=1,2,3}^T$$

For both calibration and validation, three data points are used, which are shown in Figure 42 as Data detector 1, data detector 2, and data detector 3.

RMSNE (equation 5-9) values on those three data points are used for evaluation. Table 6 shows the calibration performance of SPSA and Table 7 shows the calibration performance of BPF.

The main task of this evaluation is to prove that calibrated parameters will have better performance during validation. Therefore, the validation process performed here is not a typical validation process for picking out parameters as described with DFF. Generally, all the two algorithms are able to improve the model performance during calibration process. The key problem here is whether an improvement during calibration leads an improvement during validation process.

In order to investigate this problem, we perform a correlation analysis on simulation results. The improvement of calibration is defined as following:

$$I_{cal} = \left[RMSNE_{cal}^i - RMSNE_{original}^i\right]_{i=1,\cdots,18}^T \tag{6-1}$$

Correspondingly, the improvement of validation is defined as following:

$$I_{val} = \left[RMSNE_{val}^i - RMSNE_{original}^i\right]_{i=1,\cdots,18}^T \tag{6-2}$$

The correlation analysis should check if I_{cal} and I_{val} have a strong positive correlation or not, and the results are shown in Table 8. The results show that both SPSA and BPF offer a strong positive correlation between calibration results and validation results, which proofs the effect of the calibration process.

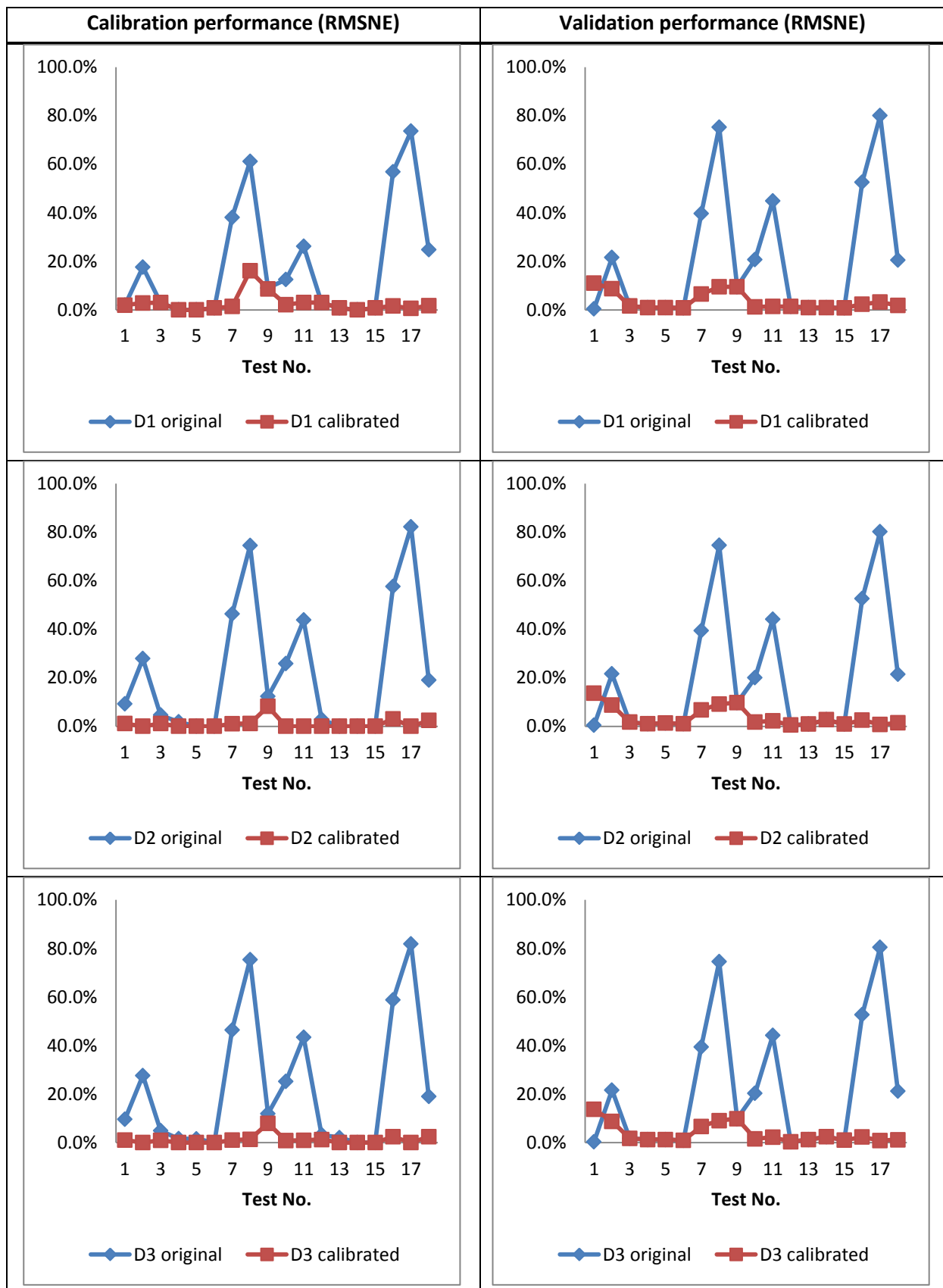**Table 6 Performance of SPSA, RMSNE value for 3 data points of 18 tests**

Calibration performance (RMSNE)	Validation performance (RMSNE)

Table 7 Performance of BPF, RMSNE value for 3 data points of 18 tests
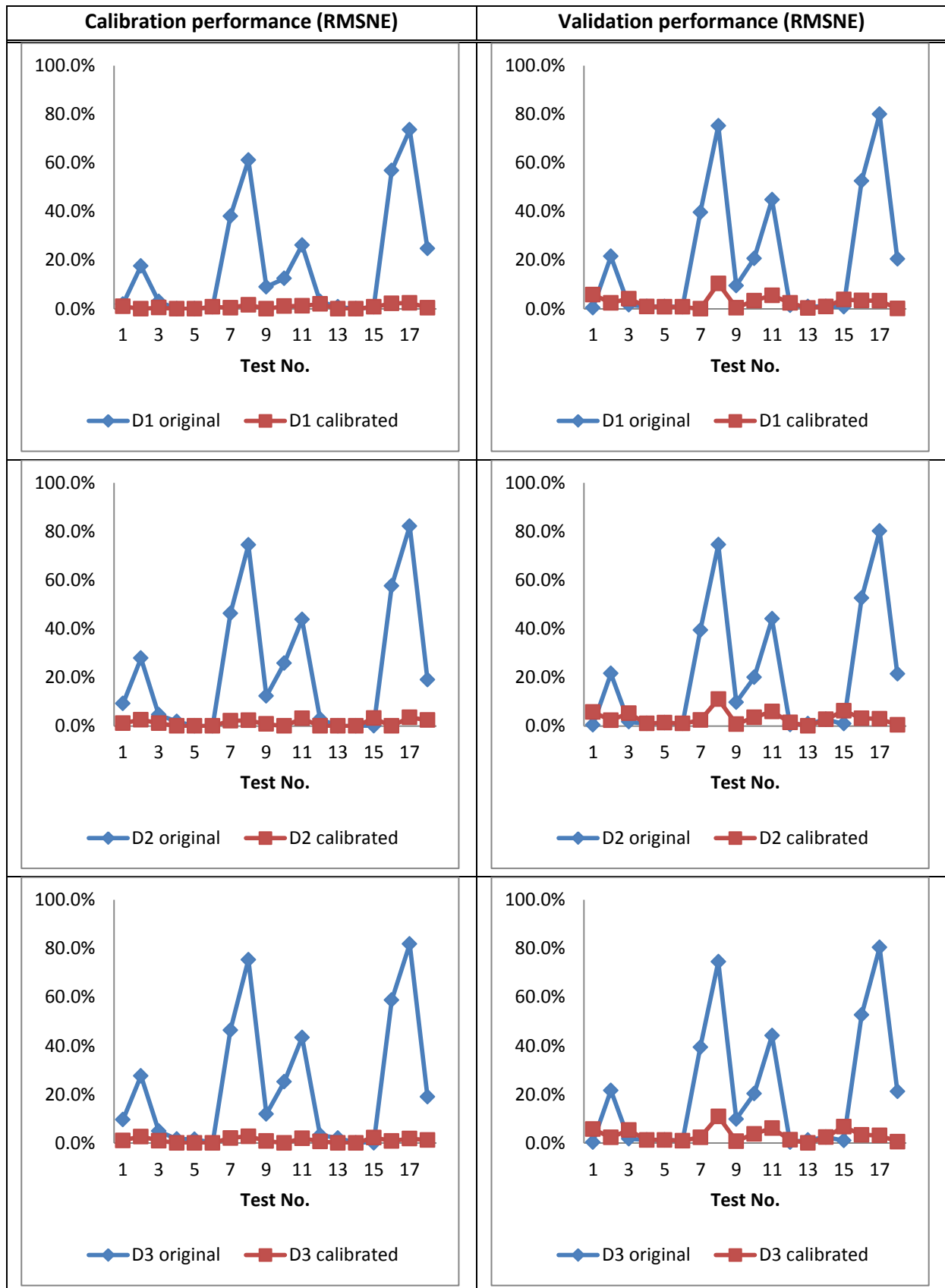
Table 8 Correlation analysis of BPF and SPSA

	Correlation coefficient of I_{cal} and I_{val}	P-value
SPSA	0.9792	1.0000
BPF	0.9915	1.0000

6.4 Comparison of FTF and DFF: an intersection case study

6.4.1 Basic environment

To model urban networks with signalized intersections is a main application of hybrid traffic modeling. Within a specific network, the importance of intersections varies. Therefore, some of them should be modeled by high resolution model, and others can be modeled by low resolution model. By using this strategy, HyTran is able to save computational resource when simulating large networks, and also reduce the simulation running time greatly.

In this example, a part of the road network in Graz (Austria) is selected as a studied case (Figure 43). Glacis Street is a main road in the city of Graz, lays in the city center in a north-south / south-north direction. As shown in Figure 43 (a), we build a macroscopic model for an area (marked by the gray lines) that includes four intersections (I-1, I-2, I-3, and I-4).

Figure 45 shows the basic layout of macroscopic model. The rectangles represent links, and lines with arrows represent the connection relations and flow directions. Macroscopic level is modeled with CTM. The cell length is set to 20m. Furthermore, when using FTF (and FTF + MLC), the cells close to the model boundaries are modeled with IFCTM.

The main purpose of this study is to compare the performance of FTF and DFF on intersection I-1. The road layout of intersection I-1 is shown in Figure 43(b). The original signal program of intersection I-1 is shown in Figure 44. The intersection layout and signal program of other three intersections (I-2, I-3, and I-4) are not listed in this thesis, but coded in CTM according to the realistic setting.

Assuming that the intersection between Glacis Street and Elisabeth Street is the critical intersection (I-1), I-1 is also modeled with VISSIM.

I-1 is a 4-way intersection. The western approach of I-1 is a one-way road, which has only exit link. For the north approach, a left-turn lane with independent signal phase is modeled. In addition, because the south approach does not have left turn flow, there is no permitted left-turn in this example. Pedestrian is also considered in VISSIM model, which influences the right turn flows.

In order to implement FTF and DFT with this network, a VISSIM model for intersection I-1 is built (Figure 45) based on necessary basic information. The building process of this model follows the basic model developing workflow, which makes it a dependable model for application. In addition, 10 data control points are placed in this model (P1…P10). P1, P5, and P9 are placed close to the upstream end of north, east, and south approaches. The other data points are placed close to the intersection stop line for collecting traffic volumes on each direction. The data collections collect traffic volume data during simulation runs.

The demand volumes are provided by VISUM model for city of Graz. The VISUM model also provides traffic assignment information in the form of turning percentages for each intersection approach.

The CTM (IFCTM) part has a cell length of 20 meters; the simulation cycle is set to 1 second. The VISSIM model has a simulation cycle of 0.2 second.
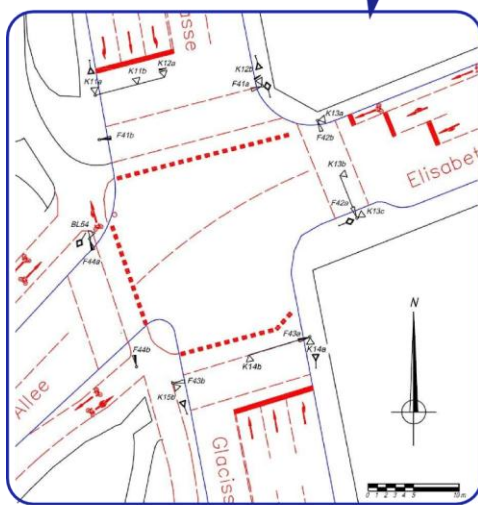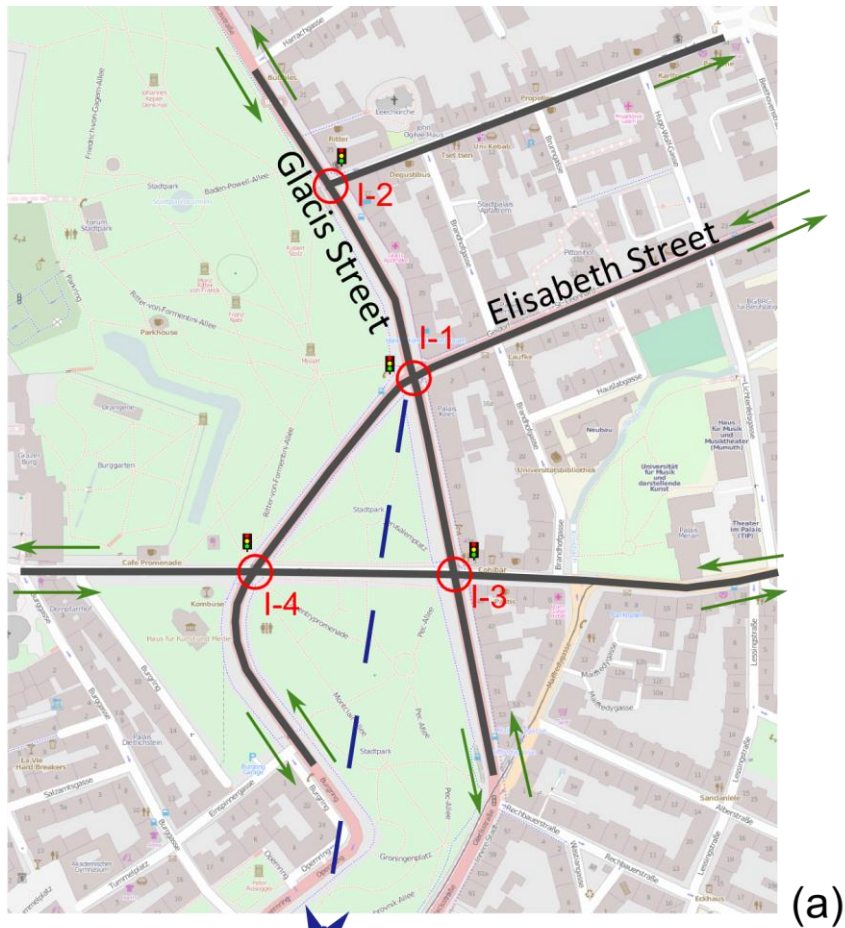
(a) the modeled network; (b) the critical layout of intersection I-1

Figure 43 Modeled Area in Graz (Austria);
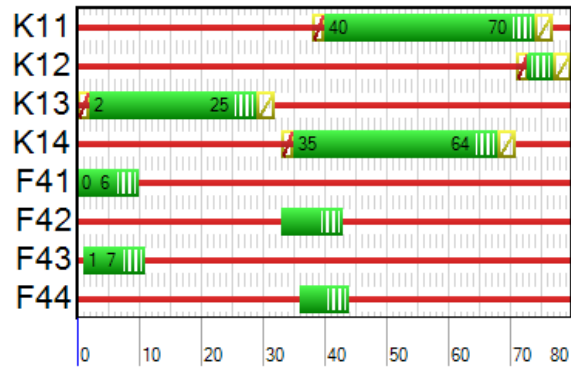
Figure 44 Signal program of I-1
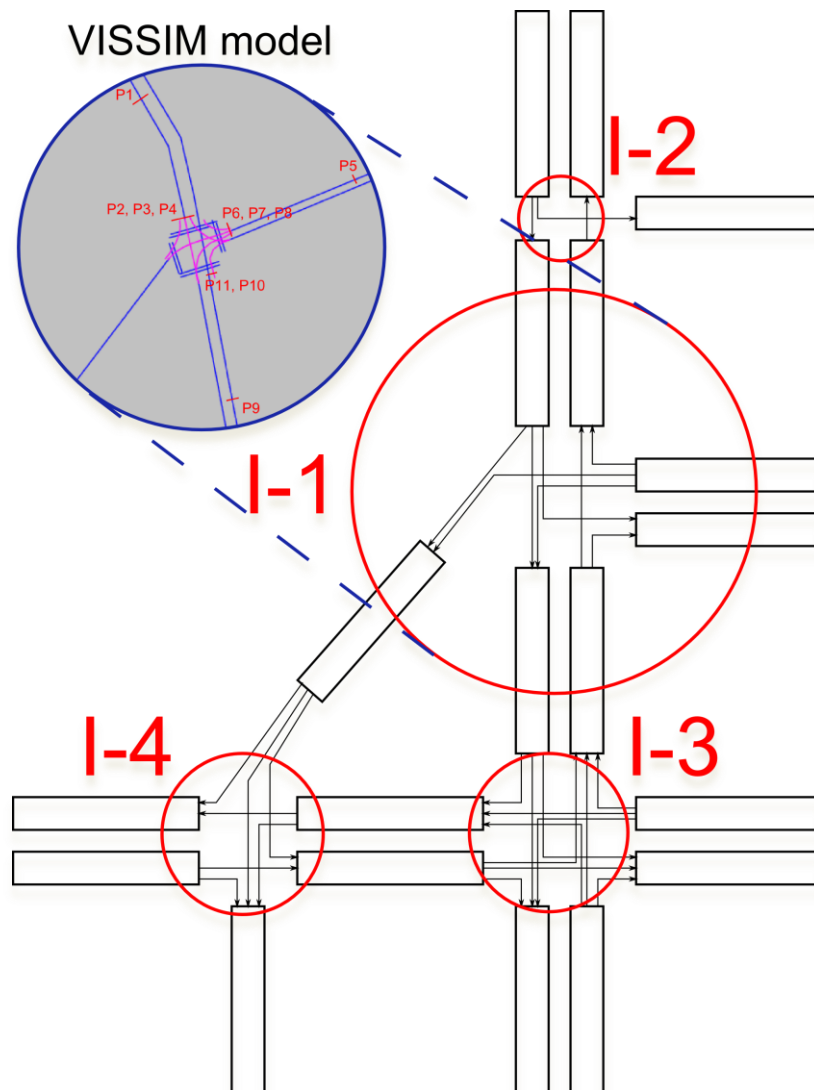
Figure 45 Hybrid model layout

6.4.2 Results comparison and discussion

The target of this comparison is to show the difference between FTF and DFF in terms of simulation results.

Traffic volumes at the time intervals of 5 minutes on predefined control points are used as the performance indicator. Calibration process uses 300s (5 min) traffic volumes as reference, and the performance evaluation uses 1800s (30 min) traffic volume with constant demand volume. The detailed simulation results, which are described by 5 min traffic volumes on each control point, are included in Appendix C.

The total volumes on the three studied approaches are shown in Figure 46. An obvious difference can be seen with the north approach. Because this link is under saturated state, FTF (w/o MCL) and FTF (w/ MLC) show very different performance. The results obtained from DFF (both with BPF and SPSA) shows volumes close to the results obtained from FTF (W/ MLC).
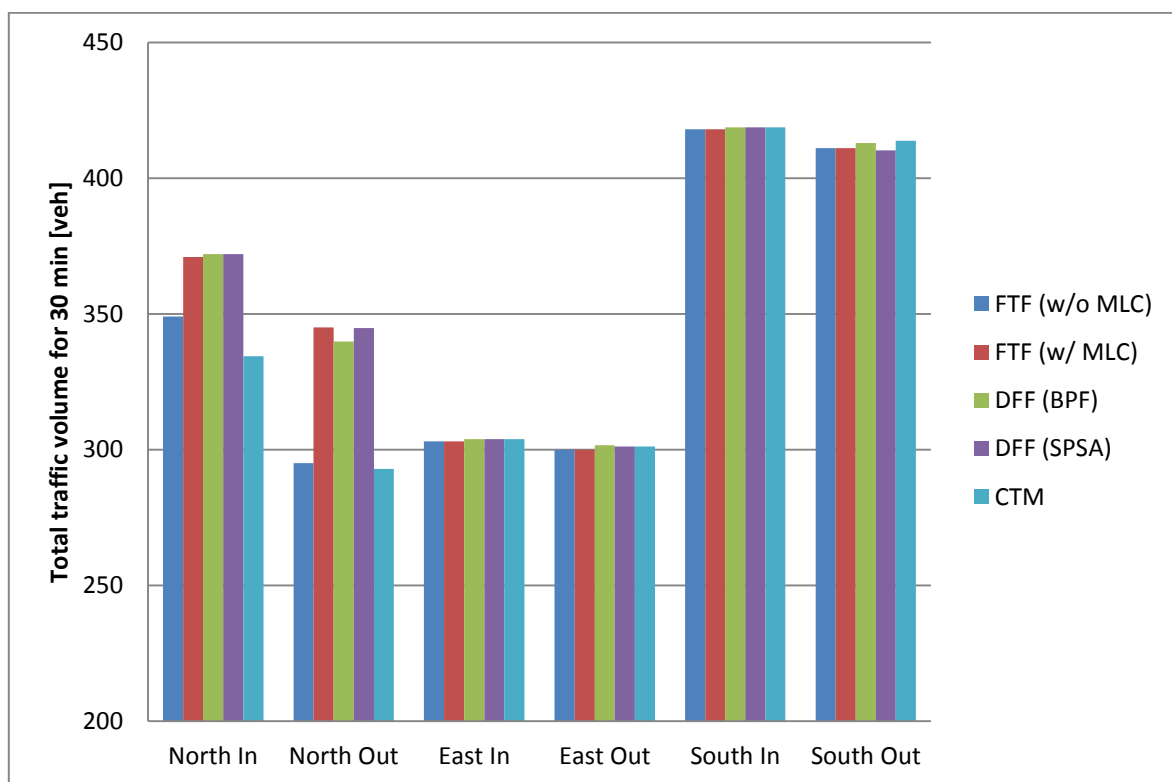
Figure 46 total volume for each

Figure 47 shows the average link density of the three approaches. The density values are calculated for each 5 minutes, and this figure shows the average values for the simulation period (the 30 min). The significant difference still emerges with the north approach. DFF (both BPF and SPSA) has lower average link density, and more closer to the results obtained from FTF (w/ MLC).

Figure 47 Link densities

Several points can be summarized from the simulation results:

- When the simulated flow has relatively low degree of saturation, all the 5 methods show similar performance (for east approach and south approac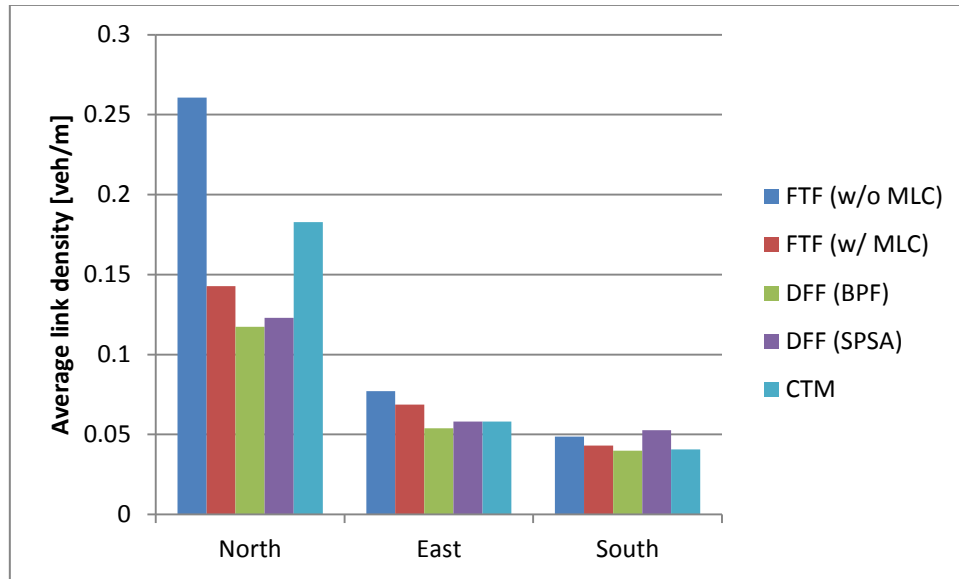h); the difference increases as the degree of saturation grows. This phenomenon means that the degree of saturation is one of the main factors that need considerations when determining whether a microscopic level is necessary.

- DFF (BPF) and DFF (SPSA) have similar results, and the results are significantly different to the results obtained from original CTM (for the north approach link). This phenomenon means those two algorithms do adjust the dynamic behavior of the calibrated model. The similarity of the results from BPF and SPSA can be considered as a confirmation of their effectiveness.

- Results from DFF are closer to results from FTF (w/ MLC) than to results from FTF (w/o MLC). This can be explained as the attributes obtained from probing runs match better the performance of FTF (w/o MLC). This phenomenon fits the basic setting. Within the probing runs, the lane change related capacity drop are considered and corresponding measures are implanted to reduce the demand for lane change.

Another problem worth noting is the time consuming for the simulation runs with different methods. In the current example, the FTF (w/o MLC) run cost about 10 minutes, and FTF(w/ MLC) run cost about 4 minutes; however, only one random seed is used for them.. For DFF case, the probing runs (30 random seeds are used) cost about 4 minutes, and the rest of the steps, including the calibration and the model run, cost less than 1 minute. However, this comparison of time consuming between FTF and DFF may be not fair. The main reason is that the efficiency of FTF is significantly determined by the programming efficiency. For example, to call COM interface may be cost too much time for which the entire speed of FTF is slow. In addition, the simulation speed of microscopic level (VISSIM) is significantly affected by saturation of simulated link. If one link is highly saturated, the simulation

speed drops largely (this is the main reason for which FTF (w/o MLC) and FTF (w/ MLC) has very different time cost in the current example). However, the degree of saturation does not have impacts on the simulation speed of the macroscopic level (CTM).

7 Conclusions

7.1 Summary and findings

The aim of this thesis is to investigate the approach for the combination of macroscopic and microscopic traffic flow simulation models. An integration framework, HyTran, has been proposed and evaluated. HyTran is applicable to simulation based traffic analysis, especially when the requirements for analytical ability vary from one part of the simulated road network to another.

This framework includes two independent approaches for model integration: the flow representation transition based framework (FTF) and data fusion based framework (DFF). For FTF, three main works have been done: first, IFCTM is proposed for improving the flow transition performance; second, the flow transformation algorithms are developed; third, a Mandatory Lane Change algorithm (MLC) is developed for solving the problem of excessive lane-changing behavior. For DFF, two basic calibration algorithms are selected and tested. In addition, the DFF process is proposed and discussed.

FTF derives from a traditional approach for model integration. The simulation results show that the flow transformation algorithms proposed in this thesis maintains the consistency of traffic flow at model boundaries effectively. Besides flow transformation, two points are investigated for improving the performance of FTF. First, the IFCTM, which considers an unevenly distributed vehicle density (slope CTM) and the Temporal Fundamental Diagram (TFD), improves the performance of CTM when used for FTF. Based on the performance of boundary interface that is presented, IFCTM shows a more realistic behavior about the propagation of queue forming and dissipation across the boundary interface. Second, the method for Mandatory lane change (MLC). Because the microscopic link lengths are probable to be limited by modeled objects, the lane change behavior on this link influence the performance of the microscopic modeled part significantly. The implement of MLC shows a strong influence on link capacity and provide a method to lane changing related capacity drop if it is necessary.

DFF is a new proposed approach for model integration. This approach introduces online-calibration into hybrid simulation. DFF extracts useful information from short time microscopic model running (the probing runs), and then, to integrate this information into macroscopic level by performing a modified parameter calibration. The feasibility of DFF is based on several assumptions, which have been proved by the examples in evaluation. At least, to some specific extent (in the tested examples), DFF works successfully. Two calibration algorithms are investigated for DFF: the simultaneous perturbation stochastic approximation (SPSA) and Bootstrap Particle filter (BPF). The performance of those two algorithms is comparable. Although the testing results show BPF has slightly better performance than SPSA, it needs more iteration of simulation.

Although DFF is a new method, we do not draw a conclusion that DFF is more recommended. The selection between FTF and DFF should be based on specific application. FTF is more feasible under diverse scenarios, in which the microscopic model behavior is too complex to be "projected" into macroscopic model; however, DFF saves huge amount of computational resource, and offers more stable simulation results. From this perspective, DFF uses macroscopic model as a filter to rule out the obviously biased results obtained from microscopic simulation results.

7.2 Contributions

The main contribution of this thesis can be summarized as "has investigated the approach of MRM in the domain of traffic flow modeling". This contribution consists of three parts:

- An approach for the combination of macroscopic and microscopic traffic flow simulation models, that is named "HyTran", is proposed.

  - The Flow Transition based Framework (FTF). Algorithms for flow transformation are developed. In addition, two components are developed for improving the performance of flow transition: method for Mandatory Lane Change (MLC) and IFCTM.

  - The Data Fusion based Framework (DFF). The new approach for model integration, DFF (Data Fusion based Framework), is investigated. In addition, the difference, advantages, and disadvantages are explored, compared with FTF. From the perspective of MRM, DFF is a new approach for consistency maintenance that achieves the advanced consistency.

- The possibility of combine a macroscopic and a microscopic traffic flow is investigated, based on examples of CTM (Cell Transmission Model, a macroscopic traffic flow model) and VISSIM (a simulation system whose kernel is a car-following based microscopic traffic flow model). This thesis discusses the detail stuff of integration according to the attributes of CTM and VISSIM, which has not been reported before.

7.3 Further research

Within the research, there are also problems that are not fully solved, which can be considered as the limitations of current work, and also can be considered as the direction for further research.

- The feasibility of DFF depends on the performance of calibration algorithms. In the current work, the calibration algorithms are proposed and investigated. However, there is still a potential to improve the efficiency of the data fusion process.

- The feasibility of DFF also depends on the dynamic behavior of macroscopic model. There is a possible situation that no matter how much effort is devoted into macroscopic level calibration, the performance of macroscopic cannot be as good as expected. In order to investigate this problem, more detailed analysis on the dynamic behavior of macroscopic model is needed. In addition, further research should focus on the practical examples, not only on the abstract examples.

- The attributes of CTM and VISSIM are used when it is necessary. Therefore, HyTran may not suitable for other models. It has to be considered when implementing HyTran with other models.

- The efficiency of code for FTF is limited by the COM Interface of VISSIM. In the further work, two solutions may solve this problem: first, to implement HyTran with an open source microscopic simulation tool. This solution probably needs a large amount of coding work. Second, to improve the flow transition algorithms from the perspective of software engineering.

Bibliography

Alecsandru, C. D., 2006. *A stochastic mesoscopic cell-transmission model for operational analysis of large-sacle transportation networks, doctoral thesis.* s.l.:Louisiana State University.

Almasri, E., 2006. *A New Offset Optimization Method for Signalized Urban Road Networks (doctoral thesis).* s.l.:University Hannover.

Antoniou, C., Ben-Akiva, M. & Koutsopoulos, H., 2007. Nonlinear Kalman Filtering Algorithms for On-Line Calibration of Dynamic Traffic Assignment Models. *Intelligent Transportation Systems, IEEE Transactions on ,* 8(4), pp. 661 - 670 .

Barcelo, J., 2010. *Foundamentals of Traffic Simulation.* 1st ed. s.l.:Springer.

Bourrel, E. & Henn, V., 2002. *Mixing micro and macro representations of traffic flow: a first theoretical step.* Bari, Italy, Proceeding of the 9th meeting of the Euro Working Group on Transportation.

Bourrel, E. & Lesort, J.-B., 2003. Mixing Microscopic and Macroscopic Representations of Traffic Flow: a Hybrid Model Based on the LWR theory. *Transportation Research Record,* Issue 1852, pp. 193-200.

Burghout, W., 2004. *Hybrid microscopic-mesoscopic traffic simulation.* Stockholm, Sweden: Royal Institute of Technology.

Burghout, W. & Wahlstedt, J., 2007. Hybrid Traffic Simulation with Adaptive Signal Control. *Transportation Research Record,* pp. 191-197.

Daganzo, C. F., 1994. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological,* 28(4), pp. 269-287.

Daganzo, C. F., 1995. The cell transmission model, part II: Network traffic. *Transportation Research Part B: Methodological,* 29 (2), pp. 79-93.

Daganzo, C. F., 1999. *The lagged cell-transmission model.* Jerusalem, Israel, 14th International Symposium on Transportation and Traffic Theory.

Davis, P. K. & Bigelow, J. H., 1998. *Experiments in multiresolution modeling (MRM),* Washington, D.C.: RAND (U.S.).

Dervisoglu, G., Gomes, G., Kwon, J. & Varaiya, P., 2009. *Automatic Calibration of the Fundamental Diagram and Empirical Observations on Capacity.* s.l., TRB 88th Annual Meeting.

Fellendorf, M. & Vortisch, P., 2010. Microscopic Traffic Flow Simulator VISSIM. In: *Fundamentals of Traffic Simulation.* s.l.:Springer, pp. 63-92.

FGSV, 2001. *Handbuch für die Bemessung von Straßenverkehrsanlagen.* Köln : FGSV Verlag.

FHWA, 2004. *Traffic analysis toolbox, volume II: guidelines for applying traffic microsimulation,* s.l.: FHWA, US Department of Transportation.

Flötteröd, G., Bierlaire, M. & Nagel, K., 2011. Bayesian demand calibration for dynamic traffic simulations. *Transportation Science,* 45(4), pp. 541-561.

Flötteröd, G. & Nagel, K., 2005. *Some practical extensions to the cell transmission model.* s.l., Intelligent Transportation Systems, 2005. Proceedings. 2005 IEEE.

Flötteröd, G. & Rohde, J., 2011. Operational macroscopic modeling of complex urban road intersections. *Transportation Research Part B: Methodological,* 45(6), p. 903–922.

Gomes, G. et al., 2008. Behavior of the cell transmission model and effectiveness of ramp metering. *Transportation Research Part C: Emerging Technologies,* Volume 16, pp. 485 - 513.

Helbing, D., Farkas, I. & Vicsek, T., 2000. Simulating dynamical features of escape panic. *Nature,* 09, 407(6803), pp. 487-490.

Helbing, D. & Molnar, P., 1995. Social Force Model for Pedestrian Dynamics. *PHYSICAL REVIEW E,* May, Volume 51, pp. 4282-4286.

Ishak, S. A. C. S. D., 2006. Improvement and Evaluation of the Cell-Transmission Model for Operational Analysis of Traffic Networks: A Freeway Case Study. *Transportation Research Record,* Volume 1965, pp. 171-182.

Jayakrishnan, R., Sahraoui, A.-E.-K. & Jun-Seok, O., 2007. Calibration and Path Dynamics Issues in Microscopic Simulation for Advanced Traffic Management and Information Systems. *Transportation Research Record: Journal of the Transportation Research Board,* 1771(2001), pp. 9-17.

Julier, S. J. & Uhlmann, J. K., 1997. *A New Extension of the Kalman Filter to Nonlinear Systems.* s.l., Proc. of AeroSense: The 11th Int. Symp. on Aerospace/Defence Sensing, Simulation and Controls..

Kates, R. & Poschinger, A., 2000. *Investigation of a stochastic disaggregation model.* Torino, World Congress on Intelligent Transportation Systems.

Leclercq, L., 2007. Hybrid approaches to the solutions of the "Lighthill-Whitham-Richards" model. *Transportation Research Part B: Methodological,* Volume 41, pp. 701-709.

Lee, J.-B., 2008. *Calibration of Traffic Simulation Models Using Simultaneous Perturbation Stochastic Approximation (SPSA) Method extended through Bayesian Sampling Methodology.* New Brunswick: Rutgers, The State University of New Jersey (Doctoral thesis).

Lee, J.-B. & Ozbay, K., 2009. New Calibration Methodology for Microscopic Traffic Simulation Using Enhanced Simultaneous Perturbation Stochastic Approximation Approach. *Transportation Research Record,* Issue 2124, pp. 233-240.

Lee, S., 1996. *A cell transmission based assignment-simulation model for integrated freeway/surface street systems, master thesis.* s.l.:Ohio State University.

LeVeque, R. J., 1992. *Numerical methods for conservation laws.* 2nd ed. s.l.:Birkhaeuser.

Lighthill, M. J. & Whitham, G. B., 1955. On Kinematic Waves. II. A Theory of Traffic Flow on Long Crowded Roads. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences,* 229(1178), pp. 317-345.

Lin, W. & Ahanotu, D., 1995. *Validting the basic cell transmission model on a single freeway link,* s.l.: Uniersity of California, Berkeley.

Magne, L., Rabut, S. & Gabard, J.-F., 2000. *Towards a hybrid macro-micro traffic flow simulation model.* Salt Lake City, USA, s.n.

Ma, J., Dong, H. & Zhang, H. M., 2007. Calibration of Microsimulation with Heuristic Optimization Methods. *Transportation Research Record,* Issue 1999, pp. 208-217.

Molina, G., Bayarri, J. M. & Berger, O. J., 2005. Statistical Inverse Analysis for a Network Microsimulator. *Technometrics,* Volume 47, pp. 388-398.

Munoz, L. et al., 2004. *Methodological calibration of the cell transmission model.* s.l., American Control Conference, 2004. Proceedings of the 2004.

Munoz, L., Sun, X., Horowitz, R. & Alvarez, L., 2006. Piecewise-Linearized Cell Transmission Model and Parameter Calibration Methodology. *Transportation Research Record,* Issue 1965, pp. 183-191.

Muralidharan, A., Dervisoglu, G. & Horowitz, R., 2011. Probabilistic Graphical Models of Fundamental Diagram Parameters for Simulations of Freeway Traffic. *Transportation Research Record,* Issue 2249, pp. 78-85.

Nagel, K. & Nelson, P., 2005. A Critical Comparison of the Kinematic-Wave Model with Observational Data. In: H. S. Mahmassani, ed. *Transportation and Traffic Theory. Flow, Dynamics and Human Interaction. 16th International Symposium on Transportation and Traffic Theory.* s.l.:Elsevier, pp. 145-163.

Nagel, K. & Schreckenberg, M., 1992. A cellular automaton model for freeway traffic. *J. Phys. I France ,* 12, Volume 2, pp. 2221 - 2229.

Newell, G. F., 1961. Nonlinear Effects in the Dynamics of Car Following. *Operations Research,* 9(2), pp. 209-229.

Pohlmann, T., 2010. *A new method for online control of urban traffic signal systems (doctoral thesis).* s.l.:Technischen Universität Carolo-Wilhelmina zu Braunschweig .

Poschinger, A., Kates, R. & Juergen, M., 2000. *The flow of data in coupled microscopic and macroscopic traffic simulation models.* Torino, Wordl Congress on Intelligent Transportation System.

PTV, 2011. *VISSIM 5.40-01 User Manual,* Karlsruhe: Planung Transport Verkehr AG.

Retzko, H.-G. & Boltze, M., 1987. Timing of inter-green periods at signalized intersections: The German method. *ITE Journal,* 9, 57(9), pp. 23-26.

Richards, P. I., 1956. Shock Waves on the Highway. *Operations Research ,* Volume 4, pp. 42-51.

Rothery, R. W., 1992. *Revised Monograph on Traffic Flow Theory: Car following models,* s.l.: FHWA.

Sewall, J., Wilkie, D. & Lin, M. C., 2011. Interactive Hybrid Simulation of Large-Scale Traffic. *ACM Transaction on Graphics (Proceedings of SIGGRAPH Asia),* Volume 30.

Spall, J. C., 1998. An Overview of the Simultaneous Perturbation Method for Efficient Optimization. *Johns Hopkins APL Technical Digest,* 19(4), pp. 482-492.

Spall, J. C., 1998. Implementation of the simultaneous perturbation algorithm for stochastic optimization. *IEEE Transcations on aerospace and electronic systems,* Volume 34, pp. 817-823.

Sumalee, A., Zhong, R., Pan, T. & Szeto, W., 2011. Stochastic cell transmission model (SCTM): A stochastic dynamic traffic model for traffic state surveillance and assignment. *Transportation Research Part B: Methodological,* Volume 45, pp. 507-533.

Szeto, W. Y., 2008. Enhanced Lagged Cell-Transmission Model for Dynamic Traffic Assignment. *Transportation Research Record,* Volume 2085, pp. 76-85.

Szeto, W. Y., 2008. Enhanced Lagged Cell-Transmission Model for Dynamic Traffic Assignment. *Transportation Research Record,* 2085(1), pp. 76-85.

Toledo, T. et al., 2003. Calibration and Validation of Microscopic Traffic Simulation Tools: Stockholm Case Study. *Transportation Research Record,* Volume 1831, pp. 65-75.

Williams, J. C., 1992. *Revised Monograph on Traffic Flow Theory: Macroscopic Flow Models,* s.l.: FHWA.

Zhang, M. H., Kuhne, R. & Michalopoulos, P., 1992. *Revised Monograph on Traffic Flow Theory: continuum flow models,* s.l.: FHWA.

Appendix A: Introduction to software environment

To implement HyTran, three components are necessary: the component for CTM, the component for FTF, and the component for DFF.

HyTran has three independent software modules for implementing these components, they are:

- CTMMAT: the module for CTM and IFCTM. This module describes basic data structure for a CTM (IFCTM) model, and the related functions.

- VIS-MM: the code package for implementing FTF, including flow transition algorithms and MLC. This module utilizes COM Interface provided with PTV VISSIM.

- CTMCAL: the code package for implementing DFF, including code for SPSA and BPF.

CTMMAT

The following class diagram shows CTMMAT. CTMMAT has two classes: CTMMAT and CTMMAT_IFCTM. From the perspective of programing, CTMMAT_IFCTM is a class that derives from CTMMAT, inherits all the attributes and operations from CTMMAT, and has several extra attributes and operations for IFCTM.

```
           CTMMAT
+dim: Integer
+demandSplitMat: Matrix
+supplySplitMat: Matrix
+demandList: Matrix
+supplyList: Matrix
+timestep: integer
+signalList: Matrix
+stateVec: Vector
+numofreal: Integer
+numofsource: Integer
+numofsink: Integer
+inputList: Vector
+leftturnList: Matrix
+exterCon: Vector
+chkPoint: vector
+curstep: Integer
+singleStep()
+demMatFun()
+supMatFun()
+traMatFun()
+datBacFun()
```

```
       CTMMAT_IFCTM
+relatedCellList: Vector
+startupTimeList: Vector
+tfdFun()
+slopeFun()
+singeStepIF()
```

The general process of CTM and IFCTM is shown by activity diagrams below.
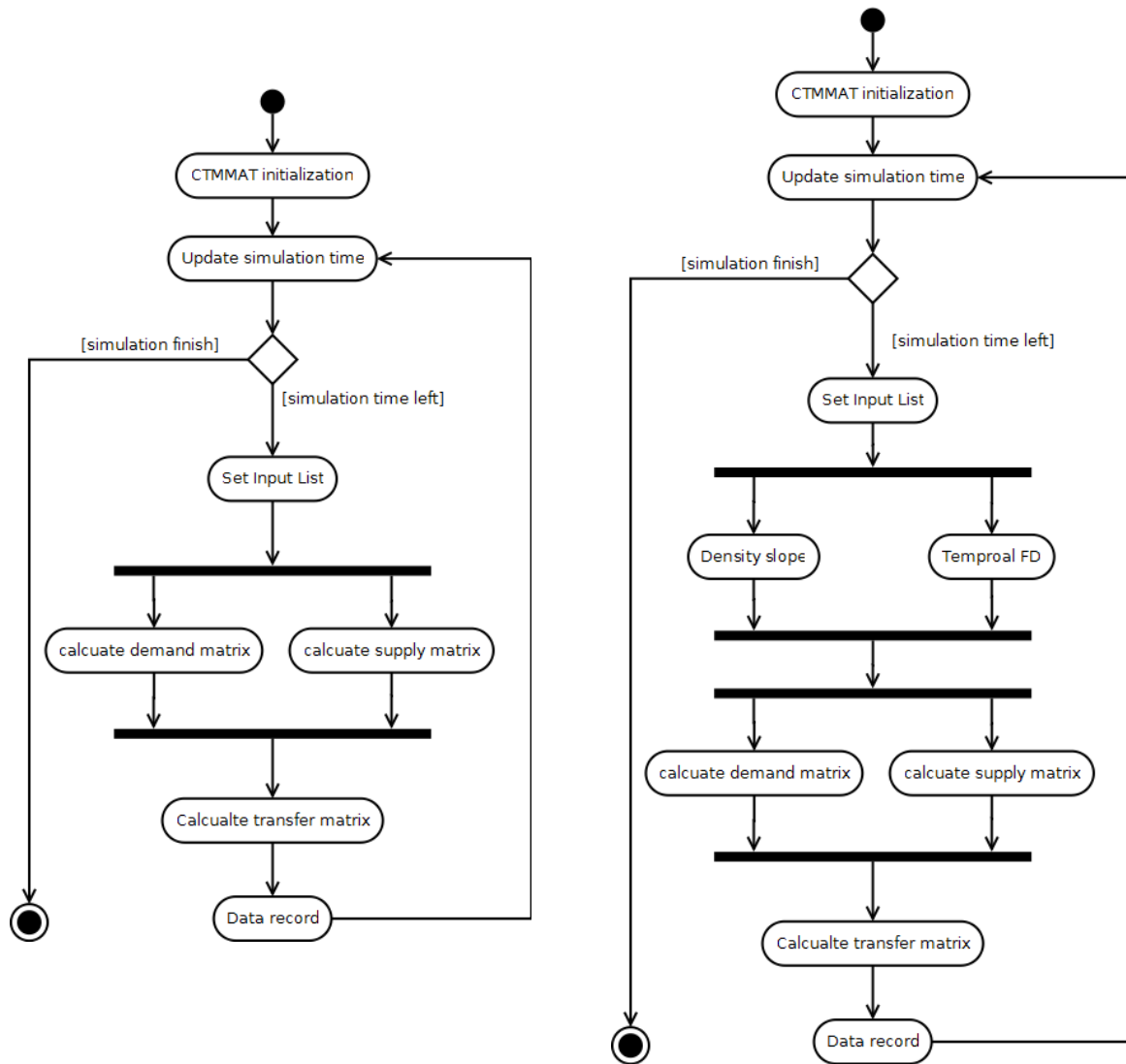
Activity diagram for CTM   **Activity diagram for IFCTM**

In CTMMAT, the basic elements are matrices. In addition, two technologies are used to speed up the calculation: 1) the calculation should be vectorized; 2) functions coded in C can be compiled with MEX ("Matlab executable"), and called in Matlab code as binary MEX-files.

VIS-MM

VIS-MM is a module for implementing FTF. The activity diagram of VIS-MM is shown following. During the process of CTM simulation, VIS-MM is called at each time step.
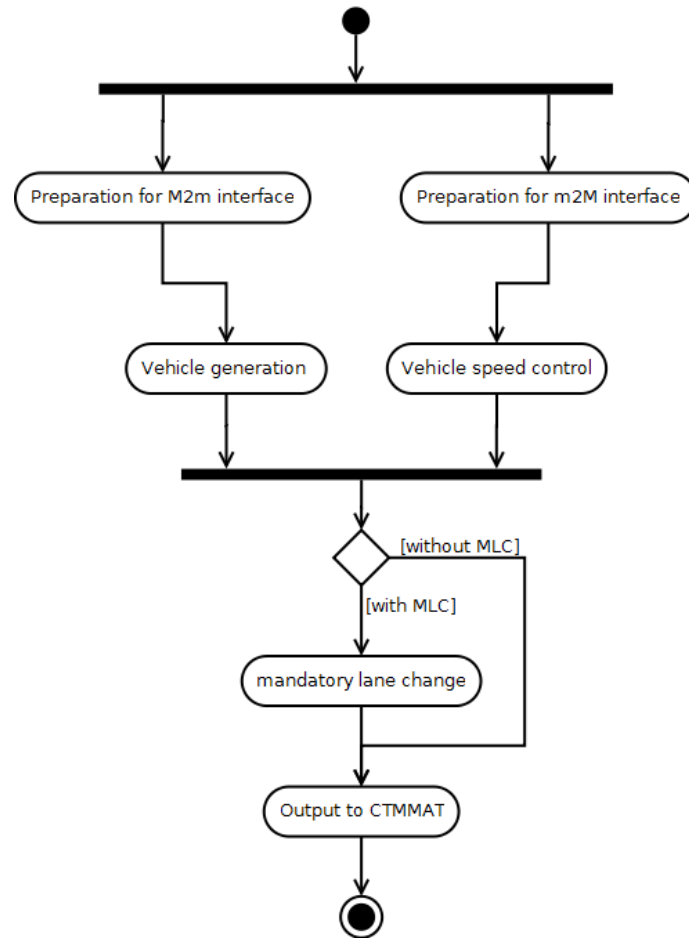
Activity diagram for VIS-MM

VIS-MM implements all the communications with VISSIM though COM interface provided by VISSIM.

CTMCAL

CTMCAL includes two functions for SPSA and BPF.

The implementation of SPSA is based on the algorithm described in 5.4.2, coded as a Matlab function.

The implementation of BPF is based on ReBEL (Recursive Bayesian Estimation Library), an open source matlab toolkit for performing Bayesian estimation.

Appendix B: Results of basic FTF evaluation

The following figures show the accumulative volume curves in section 6.2.1. These figures show how well the volume curves math with each other.

1000 veh/h, m2M, original CTM

1000 veh/h, m2M, IFCTM

1500 veh/h, m2M, original CTM

1500 veh/h, m2M, IFCTM

2500 veh/h, m2M, original CTM

2500 veh/h, m2M, IFCTM

118

1000 veh/h, M2m, original CTM

1000 veh/h, M2m, IFCTM

1500 veh/h, M2m, original CTM

1500 veh/h, M2m, IFCTM

2500 veh/h, M2m, original CTM

2500 veh/h, M2m, IFCTM

Appendix C: Results of the comparison of FTF and DFF

The simulation results of section 6.4:

Traffic volumes from FTF

Time [min]	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
0-5	60	32	9	12	48	20	21	0	63	12	47
5-10	58	20	16	13	51	27	23	1	71	13	55
10-15	63	14	1	14	51	20	34	0	73	13	55
15-20	42	18	15	12	51	18	27	0	67	14	63
20-25	55	33	14	10	51	19	34	2	73	15	57
25-30	71	41	8	13	51	27	25	2	71	21	46
Total	349	158	63	74	303	131	164	5	418	88	323

Traffic volumes from FTF +MLC

Time [min]	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
0-5	60	32	11	11	48	20	21	0	63	12	49
5-10	58	27	13	15	51	28	26	1	71	12	54
10-15	67	24	10	14	51	20	30	0	73	12	57
15-20	65	42	21	15	51	16	28	1	67	14	62
20-25	63	26	4	12	51	25	29	1	73	15	58
25-30	58	36	18	14	51	20	32	2	71	22	44
Total	371	187	77	81	303	129	166	5	418	87	324

Traffic volumes from DFF with BPF

Time [min]	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
0-5	60.8	27.9	13.8	11.6	48.9	19.0	23.8	0.4	64.2	14.3	47.9
5-10	58.2	25.5	12.6	16.9	51.0	23.9	29.9	0.5	70.8	15.4	51.6
10-15	66.8	32.1	15.8	16.9	51.0	23.1	28.9	0.5	72.6	15.5	52.0
15-20	64.8	32.0	15.8	16.9	51.0	18.8	23.5	0.4	67.3	17.4	58.3
20-25	63.2	27.9	13.7	12.7	51.0	23.9	29.9	0.5	73.1	16.9	56.6
25-30	58.2	20.8	10.3	16.9	51.0	23.9	29.9	0.5	70.8	15.4	51.6
Total	372.1	166.1	82.0	91.8	303.9	132.7	165.9	3.0	418.7	95.0	318.0

Vehicle volumes from DFF with SPSA

Time [min]	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
0-5	60.8	27.9	13.8	11.3	48.8	18.8	23.5	0.4	64.2	14.1	47.3
5-10	58.2	25.2	12.4	15.7	51.0	23.9	29.9	0.5	70.8	14.9	50.0

Appendix C: Results of the comparison of FTF and DFF

10-15	66.8	32.1	15.8	15.7	51.0	23.2	29.0	0.5	72.6	15.9	53.3
15-20	64.8	32.1	15.8	15.7	51.0	18.7	23.4	0.4	67.3	17.4	58.3
20-25	63.2	31.0	15.3	11.8	51.0	23.9	29.9	0.5	73.1	17.0	57.0
25-30	58.2	25.2	12.4	15.7	51.0	23.9	29.9	0.5	70.8	14.9	50.0
Total	372.1	173.4	85.6	85.8	303.8	132.5	165.6	3.0	418.7	94.3	315.8

Traffic volumes from CTM

Time [min]	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
0-5	60.8	27.9	13.8	10.3	48.8	18.8	23.5	0.4	64.1	14.1	47.1
5-10	58.2	25.6	12.6	14.0	51.0	23.9	29.9	0.5	70.8	15.8	53.0
10-15	66.8	28.1	13.9	14.0	51.0	23.2	29.0	0.5	72.6	15.3	51.3
15-20	52.1	22.9	11.3	14.0	51.0	18.7	23.4	0.4	67.3	17.4	58.3
20-25	48.3	22.9	11.3	10.5	51.0	23.9	29.9	0.5	73.1	16.7	56.0
25-30	48.3	17.2	8.5	14.0	51.0	23.9	29.9	0.5	70.8	15.8	53.0
Total	334.4	144.7	71.4	76.8	303.8	132.5	165.6	3.0	418.7	95.2	318.6