Traffic Flow Modeling and Forecasting Using Cellular Automata and Neural Networks

A thesis presented in partial fulfillment of the requirements
for the degree of
Master of Science in Computer Science
at Massey University, Palmerston North, New Zealand

Mingzhe Liu

MARCH 2006

Acknowledgement

I would like to take this opportunity to acknowledge those people who have guided and supported me to achieve this qualification.

First of all, I sincerely thank my chief supervisor Dr. Ruili Wang for all his inspiration, support and guidance. His intellectual rigour and honesty kept my research on track and well focused.

My special thanks go to my co-supervisor A/Prof. Ray Kemp for his patience, guidance and suggestions especially over the period of writing the thesis.

I would like to thank many friends that made my life at Massey an enjoyable experience, especially, Yiming, Donn, Frank and Jim.

I also would like to thank many lectures and supporting staff at the IIST for providing the comfortable and well-equipped environment in which to study, especially, Prof. Janina Mazierska, A/Prof. Elizabeth Kemp, Mrs. Christine Allport, Ms. Juliet Newton, Mr. Paul Robertson and Mr. Tim O'Dea.

Finally, special thanks go to my wife Ping Wen and my daughter Xiner for their understanding, encouragement, support and endless love.

Traffic Flow Modelling and Forecasting Using Cellular Automata and Neural Networks

Abstract

In this thesis fine grids are adopted in Cellular Automata (CA) models. The fine-grid models are able to describe traffic flow in detail allowing position, speed, acceleration and deceleration of vehicles simulated in a more realistic way.

For urban straight roads, two types of traffic flow, free and car-following flow, have been simulated. A novel five-stage speed-changing CA model is developed to describe free flow. The 1.5-second headway, based on field data, is used to simulate car-following processes, which corrects the headway of 1 second used in all previous CA models.

Novel and realistic CA models, based on the Normal Acceptable Space (NAS) method, are proposed to systematically simulate driver behaviour and interactions between drivers to enter single-lane Two-Way Stop-Controlled (TWSC) intersections and roundabouts. The NAS method is based on the two following Gaussian distributions. Distribution of space required for all drivers to enter intersections or roundabouts is assumed to follow a Gaussian distribution, which corresponds to heterogeneity of driver behaviour. While distribution of space required for a single driver to enter an intersection or roundabout is assumed to follow another Gaussian distribution, which corresponds to inconsistency of driver behavior.

The effects of passing lanes on single-lane highway traffic are investigated using fine grids CA. Vehicles entering, exiting from and changing lanes on passing lane sections are discussed in detail.

In addition, a Genetic Algorithm-based Neural Network (GANN) method is proposed to predict Short-term Traffic Flow (STF) in urban networks, which is expected to be helpful for traffic control. Prediction accuracy and generalization ability of NN are improved by optimizing the number of neurons in the hidden layer and connection weights of NN using genetic operations such as selection, crossover and mutation.

Publications

Liu, M., Wang, R. and Kemp, R. (2005) Towards a Realistic Microscopic Traffic Simulation at an Unsignalised Intersection. *Proceedings of International Conference on Computational Science and its Applications (ICCSA), Lecture Notes in Computer Science,* Vol. **3481**, pp. 1187-1196, Singapore.

Liu, M., Wang, R., Wu, J. and Kemp, R. (2005) A Genetic-Algorithm-based Neural Network Approach for Short-term Traffic Flow Forecasting. *Proceedings of International Symposium on Neural Networks (ISNN), Lecture Notes in Computer Science,* Vol. **3498**, pp. 965-970, Chongqing, China.

Wang, R. and Liu, M. (2005) A Realistic Cellular Automata Model to Simulate Traffic Flow at Urban Roundabouts. *Proceedings of International Conference on Computational Science (ICCS), Lecture Notes in Computer Science,* Vol. **3515**, pp. 420-427, Atlanta.

Contents

List of Tables

List of Figures

Chapter 1

Introduction and Scope

1.1 Background

Fully understanding traffic flow is essential for traffic engineering. Modelling traffic flow is one way to study traffic flow at a theoretical level. Urban traffic networks have different road features (e.g. straight roads, intersections and roundabouts). Thus, traffic flow in an urban network is more complex than on highways due to the factors of: inhomogeneous driver behaviour, stochastic interactions between drivers, large-scale vehicular movements, various travel demands, complicated road geometries and highly non-linear traffic flow dynamics. Furthermore, traffic flow at unsignalised (i.e. not controlled by traffic signals) intersections and roundabouts are more complicated than that at signalised intersections.

Modelling traffic flow in urban networks requires simulation of vehicular movements on straight roads and at intersections or roundabouts. Many models have been developed to simulate individual road features. For example, gap-acceptance models (e.g., Troutbeck, 1998) have been used exclusively for intersections and roundabouts. However, it is essential to develop a comprehensive framework to describe traffic flow systematically in urban networks.

In recent years, real-time and predictive traffic information has been required by Intelligent Transportation System (ITS), particularly traffic signal control. This kind of information can make traffic resources to be rescheduled and thereby traffic conditions can be greatly improved. For this purpose, Short-term Traffic Flow (STF) forecasting has attracted much attention from different disciplines. Various methods (e.g., regression analysis, Kalman filter, chaotic theory, neural networks) have been used to predict STF. Due to randomness and nonlinearity of traffic flow, STF forecasting to date has not

reached a satisfactory level. There is still great potential for further improvement of the accuracy of prediction.

1.2 Problem Statement

In recent years, Cellular Automata (CA) have been widely accepted in simulating traffic flow for highways and urban networks due to dynamical and discrete characteristics of CA (Toffoli and Margolus, 1987). In addition, CA models (e.g. Nagel and Schreckenberg, 1992) can simulate random driving behaviour (e.g. random deceleration) using different probabilities. In this research, the focus of CA is on vehicle dynamics and driver behaviour in urban networks.

Modelling traffic flow on urban roads is an important part of modelling traffic flow in urban networks, as it is the most common road feature. Previous CA models (e.g., Chopard *et al.*, 1998; Chowdhury and Schadschneider, 1999; Barlovic *et al.*, 2001) used a single-regime acceleration and deceleration to simulate vehicles driving on urban straight roads. Such models are simple but very unrealistic as the speed of a vehicle changes frequently on urban roads because of intersections or roundabouts. Traffic flow on urban straight roads is often interrupted by junctions. Therefore, a detailed and realistic description of vehicle moments (e.g. speed, accelerations and decelerations) is necessary.

Modelling heterogeneity and inconsistency of driver behaviour is important in modelling driver behaviour and interactions between drivers at entrances of unsignalised intersections and roundabouts (Wang and Ruskin 2002, 2006, Akcelik, 2005). Heterogeneous driver behaviour indicates that different drivers behave differently under the same conditions, while inconsistent driver behaviour means that a single driver may behave differently under similar conditions at different times (Ahmed *et al.*, 2002).

Traffic flow at unsignalised intersections and roundabouts has been modelled using two different approaches in recent years: (i) gap-acceptance models (based on statistic analysis), and (ii) CA models. A number of authors have simulated unsignalised intersections and roundabouts using gap-acceptance

models, such as Bonneson and Fitts (1999), Brilon and Wu (1999), Harwood *et al.* (1999), Tian *et al.* (1999), Troutbeck and Kako (1999), Wu (1999), Chodur (2000), Hargring (2000), Tian *et al.* (2000), Tracz and Gondek (2000), Pallatschek *et al.* (2002), Bunker and Troutbeck (2003), Lertworawanich and Elefteriadou (2003) and Tanyel and Yayla (2003). Others have concentrated on CA models, e.g. Essar *et al.* (1997), Ruskin and Wang (2002), Wang and Ruskin (2002, 2003a, 2003b, 2006), Dupuis and Chopard (2003), Fouladvand *et al.* (2003), Campari *et al.* (2004), Liu *et al.* (2005a) and Wang and Liu (2005).

With regard to gap-acceptance models, heterogeneous driver behaviour is divided into just two groups: cautious drivers and risk-loving drivers in a more recent paper (Pollatschek *et al.*, 2002). No consideration has been given to model inconsistent driver behaviour.

With regard to the latest progress in CA models, a Minimal Acceptable sPace (MAP) method for single-lane cross traffic is developed by Wang and Ruskin (2002) and Wang (2003). Multi-stream Minimal Acceptable Space (MMAS) model for multi-lane traffic is also developed by Wang and Ruskin (2006). Heterogeneity and inconsistency of driver behaviour and interactions among drivers in cross traffic at entrances of intersections and roundabouts are simulated by incorporation of four different categories of driver behaviour (i.e. conservative, moderate, urgent and radical), together with reassignment of categories with given probabilities at each time step. Inconsistent driver behaviour can only be roughly described by these four categories in the MAP method.

It is important to realistically simulate both inconsistent and heterogeneous driver behaviour at entrances of unsignalised intersections and roundabouts. Therefore, a need exists to develop new CA models for a better understanding the effects of driver behaviour in urban networks.

1.3 Thesis Objectives

The goals of this research are (i) to develop a novel framework to simulate urban traffic flow realistically, and (ii) to develop a new method for short-term traffic flow forecasting. Specific objectives of this research include the following:

- Build new CA models to simulate traffic flow on urban straight roads, between unsignalised intersections or roundabouts. These CA models are calibrated and validated using field data.
- Develop new methods to simulate driver behaviour and interactions between drivers at entrances of cross traffic, such as unsignalised intersections or roundabouts. Heterogeneous driver behaviour and inconsistent driver behaviour are modelled independently. However, they are related for a single driver. Interactions between drivers should take traffic regulations into account.
- Propose a new method to predict short-term traffic flow on a main road in urban areas. This method should be capable of improving prediction ability. Due to randomness and non-linearity of traffic flow, this method needs to reflect strong generalization ability.

1.4 Main Contributions

This thesis advances the state of the art in modelling traffic flow in urban areas including straight roads and junctions. It enhances the existing models and develops new ones. Another contribution of this thesis is that the proposed Genetic Algorithms-based Neural Networks (GANN) optimization method explores combinations of genetic algorithms and neural networks and extends application fields of genetic algorithms and neural networks. More specifically:

- Fine-grid CA (the length of each cell corresponds to 1 m on a real road) have been adopted in this research. The fine-grid models are able to simulate vehicle dynamics in a more realistic way.
- The five-stage speed-changing CA model is first proposed to describe free flow in urban networks. Moreover, 1.5-second rule based on observations is used to describe car-following processes instead of 1-second rule used in previous CA models.
- Two Gaussian distributions are used for the first time to simulate heterogeneous driver behaviour and inconsistent driver behaviour at entrances of unsignalised intersections and roundabouts.

- Effects of passing lanes on traffic flow on a single-lane highway are investigated using a realistic CA model, which has not been reported in the literature.
- A GAANN method is proposed to improve prediction accuracy and the generalization ability of neural networks and used to forecast STF.

1.5 Thesis Outline

This thesis has eight chapters, including this introductory chapter. Chapter 2 includes a literature review of previously proposed CA traffic flow models for highway traffic and urban networks. Chapter 3 presents new CA models developed in this research for simulating free and car-following traffic flow on urban straight roads. The Normal Acceptable Space (NAS) method is used to model heterogeneous and inconsistent driver behaviour at entrances of unsignalised intersections in Chapter 4. In Chapter 5, driver behaviour at entrances of roundabouts is simulated using the NAS method. In Chapter 6, effects of passing lanes on highway traffic are investigated based on fine-grid CA models. The GANN method is proposed to predict short-term traffic flow in urban networks in Chapter 7. In Chapter 8, study summary and main findings of this thesis are discussed and some suggestions for future improvements are made based on the work to date. In Appendices, interfaces of traffic flow simulation at TWSC intersections and roundabouts and interface of STF prediction are shown. The traditional methods for computing capacity, delay and queue length at urban roundabouts are introduced and the used glossaries of traffic terms in this research are listed.

Chapter 2

Cellular Automata Traffic Flow Models

2.1 Introduction

Cellular automata (CA), first proposed by von Neumann (1966), have attracted scholars from various disciplines (details see Ganguly *et al.*, 2001). Vichniac (1984) defined CA as "dynamical system where space, time and variables are discrete. Cellular automata are capable of non-numerical simulation of physics, chemistry, biology and others and they are useful for faithful parallel processing of lattice models." A cellular automaton has three elements: grids, states and neighbours (Wolfram, 2002). Each cell occupies a regular spatial grid and can have any one of a finite number of states (Wang, 2003). The state of each cell at a given time depends only on the states of itself and its neighbours (Toffoli and Margolus, 1987). All cells in the grid are updated synchronously (Wolfram, 1986) and the state of the grid advances in discrete time steps. The overall structure can be regarded as a parallel processing device (Ganguly and Deutsch, 2004).

The simplest cellular automaton has only 2 states, coloured black and white or through other forms of representation. Besides a finite number of discrete states, continuous values may also be acceptable. The simplest grids are a series of points, like a string of pearls. In a two-dimensional automaton, rectangle, triangular and hexagonal grids may be considered. CA may also be constructed on the n-dimensional integer lattice in arbitrary numbers of dimensions (Wolfram, 2002). The simplest neighbourhoods are the nearest neighbours. As for one-dimentional CA, the nearest neighbours of grid are its left-hand and right-hand neighbours. In two-dimensional CA Von Neumann and Moore neighbourhoods are normally used (see Figure 2.1, Wolfram, 2002).

The grid has either an open or a periodic boundary (Wolfram, 2002). In traffic flow modelling, many papers published to date refer to CA models with periodic boundary conditions, which imply that the road has neither beginning nor end and the number of vehicles is conserved. The use of periodic boundary conditions has been

widely accepted to avoid mathematical complications (Campari and Levi, 2000). For a one-dimensional CA, the right most cell is the neighbour of the left most one and the left most cell is the neighbour of the right most one. Two-dimensional CA grids become a torus under a periodic boundary condition (Nandi *et al.*, 1996).

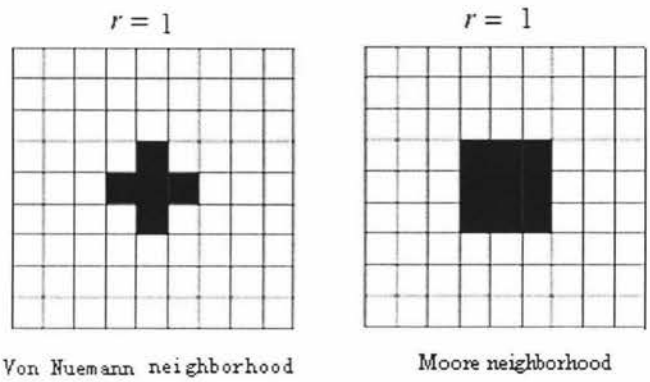


Figure 2.1 Von Neumann and Moore neighborhoods (Wolfram, 2002)

A cellular automaton of traffic flow can be separated into uniform cells on a finite uniform lattice defining the road that vehicles drive on (Wang, 2003). The variables describing each phase of each cell are updated at each time step. The variables may be the speeds of the vehicles or the states of the cells (occupied, empty or any other parameters). The state of a cellular automaton depends on the value of discrete variables at each site (Wolfram, 2002). Each site may have a finite number of discrete variables, but only one value of one variable at any single discrete time step.

A great number of published papers (e.g. Nagel and Schreckenberg, 1992; Nagel, 1996; Chopard *et al.*, 1998; Chowdhury *et al.*, 2000; Knospe *et al.*, 2000; Jiang *et al.*, 2002, 2003; Wang and Ruskin, 2002, 2003a, 2003b; Dupuis and Chopard, 2003; Schadschneider *et al.*, 2005) have shown CA to be powerful tools in modelling traffic flow on highways and in urban networks. This is due to several factors as follows.

- CA models can represent both single- and multi-lane traffic easily, which is particularly important in modelling highway traffic (Nagel, 1996).

- CA can be *computationally advantageous* (Nagel and Schreckenberg, 1992) and are easy to implement in large-scale simulations (Knospe *et al.,* 1999; Schadschneider *et al.,* 2005) as variables in CA only take integer values. Obviously, CA models can be easily computed using parallel computing paradigm.

- Random velocity fluctuations caused by driving behaviour or external conditions (e.g., temporary road close, night visibility or bad weather) can be accurately described due to the stochasticity of CA (Nagel and Schreckenberg, 1992).

- Different driving behaviours (e.g. acceleration/deceleration, turning, lane changing, parking) require appropriate models. CA models are capable of representing vehicle interactions individually. Macroscopic traffic flow variables, such as average flow, average speed and density can be obtained by averaging individual vehicular data (Larraga *et al.*, 2005).

2.2 Nagel and Schreckenberg Model

2.2.1 Description

One of the most widely accepted CA traffic models is Nagel and Schreckenberg's (1992) model or the NaSch model. This is a simple stochastic traffic model for single-lane highways, which can, in general, simulate the basic traffic situations where a car accelerates and decelerates. A road is represented as a line of cells, each being occupied by at most one vehicle. The length of each cell is 7.5 m, which is assumed to represent the real vehicle site occupied on the road (including the length of the vehicle and the space in front of and behind it). Let x_n and v_n denote the current position and velocity of vehicle n, and $d_n = x_{n+1} - x_n$ -1 denotes the distance (also named *gap*) in front of the vehicle n. One time step is one second. Therefore, it is convenient to express the relationship between distance and velocity. For example, $v < d$ is used instead of $v < d/\Delta t$ as $\Delta t = 1$.

The updating rules of N vehicles in parallel on a finite lattice of length L are as

follows:

Step 1: Acceleration

If $v_n < v_{max}$, the velocity, v_n, of the nth car is increased by one, i.e.

$$v_n \rightarrow v_n + 1 \tag{2.1}$$

where $v_{max} = 5$ (this corresponds to a maximum speed of 135 km/h for cars.

Step 2: Deceleration (due to other cars ahead)

If $d_n \leq v_n$, the velocity of the nth car is reduced to d_n - 1, i.e.

$$v_n \rightarrow d_n - 1 \tag{2.2}$$

where $d_n = x_{n+1} - x_n$. This guarantees that collisions are avoided.

Step 3: Randomization

If $v_n > 0$, the velocity of the nth car is decreased randomly by one unit with probability p, i.e.,

$$v_n \rightarrow v_n - 1 \quad \text{with probability } p \tag{2.3}$$

without this stochasticity, natural velocity fluctuations due to the effects of random driving behaviour or external conditions cannot be represented.

Step 4: Vehicle movement

Each car is moved forward according to its updated velocity v_n, i.e.

$$x_n \rightarrow x_n + v_n \tag{2.4}$$

Wagner *et al.* (1997) showed that the NaSch model can qualitatively simulate some known features of traffic flow, e.g. the spontaneous occurrence of congestion, the relationship between traffic flow and traffic density and the back-propagating *stop*-and-*go* wave.

2.2.2 Weaknesses

Although the NaSch model has been widely accepted and adopted as a primary microscopic traffic flow model, there are some fundamental deficiencies in its assumptions and rules. For instance, Wang (2003) suggested that the following should be reconsidered with regard to realistic conditions.

(i) The recommended safe headway is 2 seconds. In the NaSch model, the deceleration rule ($v = gap$) means that the headway of the following car is only 1 second, not 2, so $v = 1/2\ gap$ would be more suitable.

(ii) Steps 2 and 3 are questionable to some extent. According to the assumptions in the NaSch model, 1 unit of velocity is 27 km/h. It is not realistic to decrease speed by a multiple of 27 km/h within one second. In reality, drivers prefer to slow down gradually for safe driving on a freeway.

Other assumptions that also need to be reconsidered are:

- If a car cannot realistically accelerate at 27 km/h, Step 1 is also questionable as acceleration rates of vehicles cannot actually reach 7.5 m/s^2.
- The NaSch model only considers *gaps* in front; the speed and status of the preceding vehicle such as brake (Knospe *et al.*, 2002; Larraga *et al.*, 2005) are not taken into account. This is one of the reasons that the NaSch model cannot demonstrate smooth and steady-state driving.
- Vehicles always try to travel at maximum speed in the NaSch model. However, desired driving speed may not be the maximum speed due to vehicle performance, driving skills. Introduction of a probability to compute desired speed from the maximum speed would improve the model.

2.3 CA Models for Highway Traffic

The NaSch model has mainly been applied to single-lane highway traffic and been extended and generalized in recent years. In this section, newer CA models for single-lane highway traffic are reviewed. Two-lane or multi-lane situations and lane changing rules are reviewed in Section 2.4.

Rickert *et al.* (1996) suggested that the maximum speed in the NaSch model probably differs for different vehicles (cars, trucks). Different maximum speeds, $v_d(i)$, are used to simulate highway traffic flow in their modified NaSch model. However, the maximum travelling speed of a vehicle is related not only to vehicle performance, but also to driver habits.

Fukui and Ishibashi (1996) proposed a high-speed CA model in which speed increase is not gradual, but rapid. That is, if there are m empty cells in front, vehicles can move forward m ($m \leq v_{max}$) cells with probability 1- p, or $m - 1$ cells with probability p at the next time step. However, the speed change in this model is faster than the NaSch model so that their model cannot provide a good agreement between field data and the model data.

Random slowing down was reconsidered by Wagner *et al.* (1997). They introduced a brake probability, P_{brake}, to the NaSch model. The update of speed in Step 3 is as follows:

$$v_n = max\ (0,\ v - 1) \quad \text{with probability } P_{brake} \tag{2.5}$$

$$v_n = v \quad \text{with probability } 1 - P_{brake} \tag{2.6}$$

The influence of the randomization probability, p, on traffic flow was investigated by Barlovic *et al.* (1998). They plotted flow versus density (known as the fundamental diagram) and found a significant dependence between fundamental diagrams and different random probabilities. This led to a new CA model known as the Velocity-Dependent-Randomness model or VDR model. This model can reproduce *meta-stable* state (a state where the nucleus is in an excited state in physics, in traffic it is a state between stable and unstable states) and a *slow-to-start* rule based on two stochastic parameters, p_0 and p, which are described below. The *slow-to-start* rule derives its name from the fact that vehicles experience a small delay before exiting stationary or moving jams (Maerivoet and Moor, 2003).

$$p(v) = p_0 \quad \text{for } v = 0 \tag{2.7}$$

$$p(v) = p \quad \text{for } v > 0 \tag{2.8}$$

Thus, p controls the velocity fluctuations of moving cars while p_0 controls the speeds of cars that did not move in the previous time-step.

The new rules in the VDR model are implemented before the acceleration step. The randomization parameter used in Step 3 of the NaSch model depends on the velocity, $v_n(t)$, of the nth car in the previous time step. Other rules in the NaSch model remain unchanged. If $p_0 > p$, this means that cars in the previous time step have a higher braking probability, p_0, than moving cars.

Recently, Kerner (1998) investigated experimental features and complex spatial-temporal structures of real traffic flow. He proposed a three-phase traffic flow theory, briefly described as follows:

- *Free flow*: in this initial phase, vehicles are not affected by one another. Each car can move at its desired velocity. In this phase, flow increases linearly with the density of cars.
- *Wide moving jams*: wide moving jams normally occur at very high vehicle densities and irrespective of average velocity and flow.
- *Synchronized flow*: in this phase, average velocity is considerably lower than that of free flow. Nevertheless, the flow can be much higher than that of wide moving jams.

Of the three traffic flow phases, *synchronized flow* has attracted the most attention (Helbing *et al.,* 1997; Kerner, 1998, 2002, 2003; Lee *et al.,* 2000; Knospe *et al.*, 2000; Jiang and Wu, 2003, 2005; Lee *et al.*, 2004). Knospe *et al.* (2000) proposed a fine-grid CA model to study synchronized flow with periodic boundary conditions. The length of each cell in their model represents 1.5 m, which greatly decreases unit speed to 5.4 km/h. The model defines a randomization parameter p', which is a function of current velocity and the status of the brake light of the preceding vehicle. This parameter is determined prior to the acceleration step. Knospe *et al.* (2000) claimed that three-phase traffic flow can be reproduced well and good agreement with the empirical data can be found using their model.

Jiang and Wu (2003) challenged the model proposed by Knospe *et al.* (2000) on the basis that the sensitivity of a stopped car is lower than that of a moving car. For this reason, they introduced a parameter t_c, and modified the randomization parameter, p, and other rules are same as Knospe *et al.*'s (2000) model. The reproduction of light synchronized flow is reported in their simulation results. Authors implicitly suggest that light synchronized flow should be related to reflection of drivers to standing cars. In fact, there is little report to support this suggestion in the literature.

Wireless technology allows the speed of the preceding vehicle to be collected more easily. Thus a new extension of the NaSch model was proposed by Larraga *et al.* (2005), which considers both the *gap* and the speed of the vehicle ahead. This model modified the deceleration rule (Step 2) of the NaSch model. Other rules remained unchanged. The new velocity of the vehicle n in the deceleration rule is:

$$v_n = min(v_n , (d_n + (1-\alpha) * v_{n+1})) \tag{2.7}$$

where $d_n = x_{n+1} - x_n - 1$, $0 \leq \alpha \leq 1$. x_{n+1} and v_{n+1} are the position and speed of the preceding vehicle. α is a safety factor and is one of the following values: {0, 0.25, 0.5, 0.75, 1} which are arbitrary assigned by authors. If α takes its maximum value, $\alpha = 1$, the model is equivalent to the NaSch model. At the other end of the scale, when $\alpha = 0$, there is a rear end with 0 gap. With other α values, the results show that if α is smaller, vehicles are easier to crash and thus safety is lower. Although this modification can enhance the speed of vehicles, it is of little use in analyzing traffic flow as it is hard to determine different safety factor α for different drivers.

Xue *et al.* (2005) studied the effects on traffic flow of changing the order of the steps in the NaSch model. Two changing-order models were investigated: (i) Step 1 – 3 – 2 – 4 (Acceleration – Randomization – Deceleration – Movement); (ii) Step 3 – 1 – 2 – 4 (Randomization – Acceleration – Deceleration – Movement). Their simulation results showed that the effects of the CA models were sensitive to the order of the randomization and deceleration steps. However, other changing-order models such as Step 2 – 1 – 3 – 4 (Deceleration – Acceleration – Randomization – Movement), 2 – 3 – 1 – 4 (Deceleration – Randomization – Acceleration – Movement) and 3 – 2 – 1 – 4 (Randomization – Deceleration – Acceleration – Movement) were not discussed and this leads us to further investigate the effects of different orders to traffic flow.

More recently, much emphasis has been put on the study of bottlenecks (Campari and Levi, 2000; Diedrich *et al.*, 2000; Huang, 2002; Jiang *et al.*, 2002; Pedersen and Ruhoff, 2002; Jiang *et al.*, 2003; Jia *et al.*, 2005; Nassab *et al.*, 2005). Bottlenecks are parts of roads where capacity is locally reduced. On- and off-ramps, lane closings, uphill gradients, narrow road sections. are where bottlenecks may occur. However in highway traffic, local capacity decrease is mainly due to on- and off-ramps (Diedrich *et al.*, 2000) and these are reviewed in detail in Chapter 6.

2.4 CA Models for Lane Changing

Much progress has been made in understanding single-lane traffic flow by using above-mentioned models; however, by themselves, such models are not sufficient to simulate realistic traffic flow as two-lane and multi-lane traffic situations exist in the real world. Thus it is necessary to develop two-lane traffic models, which can subsequently be extended to model multi-lane traffic (Nagel *et al.* 1998). In addition to the rules that apply to single-lane models, there is also a rule known as lane changing, which is a vital component of multi-lane traffic flow models.

Lane changing is a common driving behaviour for different needs such as turning, changing speed or avoiding other vehicles (CORSIM, 1996). Generally speaking, lane changing is considered feasible if there are sufficient gaps ahead and behind in the target lane so that the vehicle can change into this lane safely.

Lane changing can be classified into two categories: Mandatory Lane Changing (MLC) and Discretionary Lane Changing (DLC) (CORSIM, 1996). MLC is performed when the driver must leave the current lane, for example when turning, avoiding other vehicles in front and so on, while DLC is performed if the driver perceives traffic conditions to be better in the target lane, such as a good forward perspective, improved velocity and so on. However, DLC is not required and is a stochastic process (Wang, 2003), in that even if all conditions have been satisfied, some drivers still do not change lanes.

Much work on multi-lane models in the literature deals with modelling lane changing using CA models (Nagatani, 1993; Rickert *et al.*, 1996; Yang and Koutsopoulos, 1996; Wagner *et al.*, 1996, 1997; Nagel *et al.*, 1998; Daoudia and Moussa, 2002, 2003; del Rio and Larraga, 2004). Nagatani (1993) considered a CA model for two-lane traffic with periodic boundary conditions. However, his model is deterministic and simply assumes the maximum speed, $v_{max} = 1$ and therefore we can say Nagatani's model just described lane-changing behaviour but did not simulate vehicular movement realistically.

Lane-changing rules can be symmetric (all lanes can be used to overtake) or

asymmetric (only specified lanes can be used to overtake) according to the lanes and the cars (Knospe *et al.*, 2002). Rickert *et al.* (1996) proposed a symmetric rule set where cars change lanes if the following criteria are satisfied:

- Incentive criterion:

$$v_{desire} > gap_{current}, \text{ with } v_{desire} = \min(v + 1, v_{max}) \tag{2.8}$$

- Safety criteria:

$$(1)\ gap_{other} > gap_{current} \tag{2.9}$$

$$(2)\ gap_{back} > v_{max} \tag{2.10}$$

Here v_{desire} and v_{max} are desired and maximum speeds, while $gap_{current}$, gap_{other} and gap_{back} denote the number of empty cells between the object vehicles and the preceding vehicle in the current lane and its two neighbouring cars in the objective lane, respectively.

Daoudia and Moussa (2002) proposed a new version of the symmetric lane-changing rule: at an odd time step, the vehicles in the left lane move ahead, shift to the right lane, or stop according to the CA rules; at an even time step, the vehicles in the right lane move ahead, shift to the left lane, or stop according to the CA rules. However, their model has some deficiencies. Firstly, vehicles from different lanes cannot change lane simultaneously. Secondly, turning light effects are not considered in their model; these are essential to lane changing in the real world. Finally, in order to avoid causing following vehicles in the target lane to decelerate significantly, vehicles usually accelerate when and after changing lanes. This actual driver behaviour has not been considered in their model.

Daoudia and Moussa (2003) extended their previous model (Daoudia and Moussa, 2002) to three-lane traffic with periodic boundary conditions in a car-truck system. An asymmetric lane-changing rule was also investigated in three-lane traffic. Under this rule, cars could change to any one of three lanes, while trucks could only change to the second and third lanes. Their simulation results showed that trucks influenced the system performance less in the asymmetric case in both two-lane and three-lane traffic.

Del Rio and Larraga (2004) used asymmetric lane-changing rules to simulate traffic flow on the highway from Mexico City to Cuernavaca. In their model, the right lane is preferential according to road regulations in Mexico. When a vehicle changes to

the right lane from the left lane, the required time headway ahead and behind in the right lane is assumed to be more than 3 seconds. Although this is a safety criterion, lane changing may not always be successful in peak hour under their model as a 6-second headway in the target lane is quite small.

In lane-changing models (e.g., Nagel *et al.,* 1998; Daoudia and Moussa, 2003; del Rio and Larraga, 2004), a car-truck system is widely used to represent different desired speeds. The common weakness of these models is that vehicles are treated as being of the same length despite having different maximum speeds. Different vehicle lengths need to be considered so that the relationship of flow and density can more accurately reflect actual traffic conditions.

2.5 CA Models for Urban Networks

In urban networks, crossings such as signalised (traffic-light controlled) intersections, unsignalised intersections and roundabouts can be regarded as bottlenecks (FHWA, 2000). The use of CA models to simulate traffic flow at crossings has attracted much attention in recent years. This section reviews different CA models used in signalised intersections. CA models for unsignalised intersections and roundabouts are reviewed in detail in Chapters 4 and 5, respectively.

Chowdhury and Schadschneider (1999) proposed a CA model (the ChSch model) to simulate traffic flow at signalised intersections. Their model modified Step 2 in the NaSch model: slowing down is due to avoiding collision with the preceding vehicle or signals as opposed to the NaSch model. In the case of signal changes, the update rules are given as follows:

Case 1: If the signal is red for the n^{th} vehicle:

$$v_n \rightarrow \min (v_n, d_n - 1, s_n - 1) \tag{2.11}$$

Case 2: If the signal is green for the n^{th} vehicle:

If the signal is going to turn to red in the next time step then

$$v_n \rightarrow \min (v_n, d_n - 1, s_n - 1) \tag{2.12}$$

else $$v_n \rightarrow \min(v_n, d_n - 1) \tag{2.13}$$

v_n is the velocity of the n^{th} vehicle, d_n is the number of free cells to the preceding vehicle and s_n is the number of free cells to the downstream crossing. However, the ChSch

model can also result in a complete stop from a speed of a multiple of 27 km/h; this is too abrupt, in reality a vehicle slows down gradually.

Barlovic *et al.* (2001) extended the ChSch model to a global urban network consisting of N x N intersections. They tried to maximize network flow by finding optimal model parameters. Three types of signal control strategies were used: synchronized traffic lights, a random offset strategy and a green wave strategy. Their results indicated that the green-wave strategy improved network flow more than the other two strategies, but that cycle time affected network flow greatly. In their model, there were no dominant roads, which meant that the green and red phases were of the same length for each direction. In reality, major and minor roads are differentiated and green and red phases for each direction may have different cycles. Furthermore, turning manoeuvres were not considered.

Brockfeld *et al.* (2001) examined Barlovic *et al.*'s (2001) model and pointed out that the optimal time periods should be determined by the geometric characteristics of the network, i.e. the distance between intersections. Green wave and random switching strategies were used to improve network throughput.

Barlovic *et al.* (2004) further explored the impact of different traffic-light control strategies on traffic flow. Three adaptive strategies were presented: queue length, waiting time and a neural network. Although their adaptive strategies showed quite good results, their network topology was relatively simple (one-way to one-way intersections). Thus this model was unable to reflect the real world traffic as traffic lights are seldom installed in the above simple road topology.

2.6 CA Models in Traffic Simulation Software

The effectiveness of traffic models depends on their ability to accurately reproduce actual traffic conditions. Calibration and validation (Benekohal, 1989) are normally used to measure simulation results. The Highway Capacity Manual (HCM, 2000) defines calibration as the process of adjusting the model parameters in order to obtain a close match to the actual traffic measurements while validation is the process of determining whether the selected model is appropriate for the given conditions and task.

The CA models discussed above have been successfully utilized for traffic simulation in many pieces of software. TRANSIMS (TRansportation ANalysis SIMulation System), developed by Los Alamos National Laboratory USA, is one such packages. Details of the system, including lane changing, complex turns and intersection configurations, are fully represented and each driver is given a destination and a preferred path (TRANSIMS, 2000). Large-scale simulation networks including route choice (Raney *et al.*, 2002) have been carried out in Switzerland using a multi-agent-based TRANSIMS. In their simulation, each vehicle is regarded as an agent and agents' movement is updated according to the NaSch model.

OLSIM (On-Line SIMulation) is another CA-based real-time traffic information system, developed by Duisburg-Essen University, Germany. It can provide users with the current traffic states such as congested of free flow and a 30 and 60 minute prediction of urban networks flow (Schreckenberg *et al.*, 2003). In order to simulate traffic flow realistically, some improvements have been made to the OLSIM (Table 2.1), which has been used in simulating a traffic network in Germany. Table 2.1 shows the differences between the CA models used in TRANSIMS and OLSIM.

Table 2.1 Comparison of TRANSIMS and OLSIM

	TRANSIMS	OLSIM
Cell size	7.5 m	1.5 m
Velocity unit	27 km/h	5.4 km/h
Vehicle types	One	Two (cars and trucks)
Lane changing	Symmetric	Asymmetric or symmetric
VDR	No	Yes
Anticipation	No	Yes
Brake lights	No	Yes

2.7 Summary

In this chapter, CA models for highway traffic and urban networks have been comprehensively reviewed. The advantages of using CA models are widely acknowledged and these models constitute fundamental building blocks for this research.

Some weaknesses in previous CA models have been pointed out. Based on this review, new CA models in this thesis are to be developed to simulate realistic traffic flow for highways and urban networks.

The design of lane changing rules is a primary requirement for multi-lane traffic modelling. Many criteria for lane changing have been defined in terms of speed and space in previous researches. In lane changing models, the car-truck system is widely used to represent different desired speeds. The common weakness of these models is that vehicles are assumed to be of the same length despite having different maximum speeds and this is undesirable because the relationship between flow and density has not been actually reflected. This topic is further discussed in Chapter 6.

Currently, CA models have been built into traffic simulation software such as TRANSIMS and OLSIM. Both TRANSIMS and OLSIM have been used for large-scale traffic simulation.

Chapter 3

Urban Straight Roads

3.1 Introduction

Modelling traffic flow in urban networks requires both cross traffic simulation and straight road simulation. A number of models (e.g. Chopard *et al.*, 1998; Wang and Ruskin, 2002, 2003a, 2003b, 2006; Dupius and Chopard, 2003; Campri and Levi, 2004) have been developed to simulate traffic flow at intersections and roundabouts. However, these models focus on driver behaviour and interactions in cross traffic situations, and assume that vehicular arrival rates follows a Poisson distribution, without directly modelling urban straight road traffic in detail. This assumption makes the above-mentioned models focus on cross traffic flow modelling. However, it is necessary to describe vehicular motion on straight roads when simulating an urban network.

Biham *et al.* (1992) developed a CA model for city traffic (known as the BML model) in which each of the sites of a square lattice represents an intersection of an east–west and north–south streets. Initially, vehicles are randomly distributed among the streets. The state of east-bound vehicles are updated in parallel, at every odd discrete time step, whereas those of the north-bound vehicles are updated in parallel at every even time step. The update rule is the same as that of the NaSch model (see Section 2.2) with $v_{max} = 1$ and $p = 0$, i.e. a vehicle moves forward one cell if and only if the cell in front is empty; otherwise, the vehicle does not move at that time step. A phase transition from free flow to congested states can be reproduced in the BML model (Horiguchi and Sakakibara, 1998). As Chowdhury *et al.* (2000) noted, the dynamics of the BML model are fully deterministic and randomness arises only from the initial random conditions. Therefore, stochastic traffic flow is not reproduced in the BML model. One reason for this refocus is that the BML model is the first discretemodel to simulate vehicular traffic flow on urban straight roads using CA approach.

The BML model has been generalized and extended in order to more realistically simulate city traffic conditions. Freund and Pőschel (1995) extended

one-way traffic to two-way traffic on all streets. Each east-west (north-south) street is assumed to consist of two lanes, one of which allows east-bound (north-bound) traffic while the other has west-bound (south-bound) traffic. However, the update rules in their model are also deterministic as the BML model.

Tőrők and Kertèsz (1996) considered the phenomenon of 'green-wave' synchronization, which involves the synchronization of traffic lights in order to allow continuous traffic flow and which is often found along major streets. In this situation, vehicles move as a platoon (one vehicle follows the other with no empty cells between all vehicles) and greatly improve traffic volumes in major streets.

Chowdhury and Schadschneider (1999) combined the BML and NaSch models to give a new model (the ChSch model). This model assumes that vehicles must stop at entrance of intersections. Thus, the conditions causing vehicles slowing down include: collision avoidance with preceding vehicle and being still at entrance of intersections. The ChSch model can be applied to traffic-light-controlled intersections rather than unsignalised intersections. This is because that priority among conflicted flow has not been considered in the ChSch model.

In general, the BML model and its extensions are incapable of realistically simulating traffic flow for one main reason: the coarse grid leads to sharp speed changes. In other words, in these models the length of each cell is assumed to be 7.5 m, the maximum speed to be 5 cells and a time step to be 2 seconds.

Different acceleration and deceleration rates, used in other kinds of models (e.g. car-following models) for highway traffic and urban networks, are summarised in Table 3.1. The maximum acceleration rate in Table 3.1 is approximately 2.9 m/s^2 ($= \mu + 3*$□) (Suzuki *et al.*, 2003). It can be seen that acceleration rates in Table 3.1 normally range from 1 m/s^2 to 2 m/s^2. In previous CA models (e.g. ChSch model), the normal acceleration rate is 3.75 m/s^2 (= 7.5/2 m/s^2). The maximum deceleration rate for collision avoidance in Table 3.1 is 6.4 m/s^2. The deceleration rates normally range from 1 m/s^2 to 3 m/s^2, except for emergencies. In previous CA models, the maximum deceleration rate

Table 3.1 Acceleration rates and deceleration rates in the literature

Types	Values (m/s^2)	References
Acceleration rates	1.1 for speeds < 12.19 km/h 0.37 for speeds > 12.19km/h	Treiterer (1975)
	2.0 (maximum)	Gipps (1981)
	Varies with street conditions	Barbosa *et al.* (2000)
	Normal distribution (μ = 2.0, □= 0.3) for cars Normal distribution (μ = 1.5, □= 0.2) for trucks	Suzuki *et al.* (2003)
	1.6 for cars, 1.2 for city buses 1.3 for trams, 1.0-1.2 for trucks	Luttinen (2004)
	2.5 for cars, 1.2 for bus/coaches 1.8 for light goods vehicles 1.4 for straight Trucks 1.1 for truck with trailers	Oketch *et al.* (2004)
	1.5 (normal) for cars 2.0 (maximum) for cars 1.2 (normal) for trucks 1.6 (maximum) for trucks	Bonsall *et al.* (2005)
Deceleration rates	3.0 for avoidance collision	Gipps (1981)
	Varies with street conditions	Barbosa *et al.* (2000)
	Normal distribution (μ = 2.0, □= 0.3) for cars Normal distribution (μ = 1.5, □= 0.2), for trucks	Suzuki *et al.* (2003)
	3.05 (normal) deceleration 6.4 for emergent conditions	Bham and Benekohal (2004)
	1.9 for cars, 1.3 for city buses 1.3 for trams, 1.2 for trucks	Luttinen (2004)
	4.5 for cars, 3.7 for bus/coaches 3.9 for light goods vehicles 3.7 for straight Trucks 3.2 for truck with trailers	Oketch *et al.* (2004)
	2.0 (normal) for cars 5.0 (maximum) for cars 1.5 (normal) for trucks 2.5 (maximum) for trucks	Bonsall *et al.* (2005)

is 18.75 m/s^2 (from maximum speed unit 5 to 0) and the normal deceleration rate is 3.75 m/s^2 (= 7.5/2 m/s^2). Therefore, acceleration rates and deceleration rates used in previous CA models are too extreme to be realistic. In this thesis, realistic acceleration and deceleration rates are used in simulating straight roads traffic.

In this chapter, realistic fine grid CA models are proposed to simulate traffic flow on urban straight roads. The length of each cell corresponds to 1 m, which allows realistic speed changes (see details in Section 3.3.1), while one time step corresponds to 1 second in real time.

3.2 Methodology

Generally speaking, traffic in urban networks consists of free flow and car-following flow. In a car-following flow, a vehicle has to follow the vehicle in front. In this case, speed changes of the following vehicle depend on the distance to the vehicle in front. In free flow, a vehicle can be driven at its desired speed. In this case, speed changes of the object vehicle depend on the driver's preference, because the driver has enough time steps to stop the vehicle.

In this section, free flow and car-following flow on urban straight roads are modelled with multi-value CA models (the values are the speeds of the vehicles). If the headway between the leading vehicle and the following vehicle is less than 3 seconds, it is assumed that the following vehicle is in the car-following flow phase. Otherwise, it would be in the free flow phase. This assumption is comparable to that suggested by Aycin and Benekohal (1998), who used a headway value of approximately 2.6 seconds to define free flow versus car-following flow.

3.2.1 Free flow

Previous CA models normally used a constant acceleration rate to model traffic flow in urban networks and on highways (Nagel and Schreckenberg, 1992; Simon and Nagel, 1998; Chowdhury *et al.*, 2000; Schadschneider *et al.*, 2005). However, the assumption of driving with one acceleration rate is not realistic. For example, drivers with manual gears will not use the same gear for different speeds and different gears

have different acceleration capacities. Even automatic gears have similar effects.

In this Chapter, a five-stage model for speed changes is developed to describe free flow vehicles on a straight road. The five stages are: (i) high acceleration, (ii) low acceleration, (iii) steady state, (iv) low deceleration, and (v) high deceleration stages. These five stages are illustrated in Figure 3.1 and apply to a vehicle driving between unsignalised intersections or roundabouts.

In the high acceleration stage (stage A), speed increases quickly due to a high acceleration rate. In practice, drivers normally use the lowest gear. During this stage, the speed of a vehicle increases up to a certain level (about is 40 km/h based on field observation). In stage B, speed increases with relatively moderate acceleration rate, which gradually leads to a desired speed in stage C. Clearly, the acceleration rate is much lower in stage B than in stage A, but the vehicle drives at higher speeds.

In steady stage (stage C), speed is assumed either to be unchanged or to randomly fluctuate within a certain range i.e. $1 m/s^2$ (Transportation Engineering Handbook, 1992). Speed in this stage must neither exceed the legal limit (50 km/h in New Zealand) nor be too low. The duration of this stage depends on the length of the road.

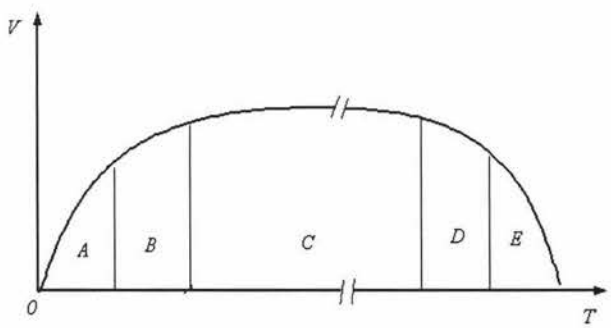


Figure 3.1 Five-stage model of speed changes with travel time. V and T denote current speed and travel time respectively. A, B, C, D and E are the five stages.

Based on field observation, a vehicle slows down with two different deceleration stages depending on how close to the downstream junction it is: a low and a high deceleration stages. At the low deceleration stage (stage D), a vehicle will

decelerate gradually with a moderate deceleration rate, followed by a high deceleration stage (stage E), in which speed decreases rapidly until the vehicle is stationary.

The update rules of the nth vehicle depends on its speed, $v_n(t)$, and the distance between its current position and the downstream junction, $d_n(t)$, at time step t:

1. Speed adjustments

$$\text{A: } v_n(t) \rightarrow v_n(t) + a_1 \tag{3.1}$$

$$\text{B: } v_n(t) \rightarrow v_n(t) + a_2 \tag{3.2}$$

$$\text{C: } v_n(t) \rightarrow v_n(t) + a_3 \tag{3.3}$$

$$a_3 = \begin{cases} 1 & with\ probability\ p_1 \\ -1 & with\ probability\ p_2 \\ 0 & with\ probability\ p_3 \end{cases}$$

$$\text{D: } v_n(t) \rightarrow v_n(t) + b_1 \tag{3.4}$$

$$\text{E: } v_n(t) \rightarrow v_n(t) + b_2 \tag{3.5}$$

where

a_1 and a_2 are the acceleration rates of the high and low acceleration stages

p_1 is the probability of acceleration

p_2 is the probability of deceleration

p_3 is the probability of unchanged speed

b_1 and b_2 are the deceleration rates of the low and high deceleration stages

As vehicles drive at individual desired speeds within a certain range with random fluctuation, in stage C three probabilities (p_1, p_2 and p_3) are used to reflect the fluctuation process. For simplicity, all probabilities (p_1, p_2 and p_3) are simply assumed to be equal, namely, $p_1 = p_2 = p_3 = 1/3$. For details of the values of a_1, a_2, a_3, b_1 and b_2, see Section 3.3.2.

2. Vehicle movement

$$x_n(t+1) \rightarrow x_n(t) + v_n(t+1) \tag{3.6}$$

Each vehicle moves to a new position according to its current position and updated speed obtained from the above steps.

3.2.2 Car-following flow

A novel two-phase, car-following CA model has been developed, based on the five-stage free flow model. Car-following models (Brackstone and MacDonald, 1999; Ahn *et al.*, 2004; Zhang and Kim, 2005), not based on CA, have emphasised the interaction between the leading and following vehicles. In other words, the speed of the following vehicle changes in relation to the gap between it and the leading vehicle. Vehicles in this situation are defined as being in the *close tailing* phase.

However, most drivers favour driving at their own preferred speed rather than close-tailing other vehicles and thus change their speed depending on their speeding stage rather than on the gap. This situation is similar to a vehicle in free flow. Vehicles in this case are therefore defined as being in the *semi-free-flow* phase.

The update rules of the nth vehicle depend on its position $x_n(t)$, speed $v_n(t)$, $v_{n+1}(t)$ and gap $g_n(t)$, $g_{n+1}(t)$ (empty cells between its current position and that of the preceding vehicle) at time step t:

1. Close tailing

$$\text{If } g_n(t) \leq 1.5v_n(t), \text{ then: } v_n(t+1) \rightarrow 2/3g'_n(t+1) \tag{3.7}$$

$$g'_n(t+1) = g_n(t) + \min\{v_{n+1}(t), g_{n+1}(t)\}$$

This rule is based on the 1.5-second rule which is obtained from the averaged 1.5-second car-following process in the real data. This averaged 1.5-second headway means that some headways are less than 1.5 second due to smaller gap synchronization and others not due to larger gap. $g'_n(t+1)$ denotes the anticipated travelling distance in the next time step. In other words, the vehicle can only drive up to 2/3 of the total anticipated distance between it and the preceding vehicle in one time step.

2. Semi-free flow

If $3v_n(t) \geq g_n(t) > 1.5v_n(t)$, then, although the following vehicle is still in car-following, its behaviour will be similar to that of free flow.

3. Vehicle movement

$$x_n(t+1) \rightarrow x_n(t) + v_n(t+1) \tag{3.8}$$

Each vehicle moves to a new position according to its current position and updated speed obtained from the above steps.

It should be noted that free flow and car-following flow do not occur in isolation from each other. In reality and in the proposed models, one is often converted to the other and vice versa. Correspondingly, vehicle speeds are updated under the different flow phases.

3.3 Calibration and Validation

3.3.1 Model formulation

In previous CA models, the length of each cell is 7.5 m. The disadvantages of this length have already been discussed (see Section 2.3). In this research, the length of each cell corresponds to 1 m on a real road, thus providing better resolution in modelling actual traffic flow than other earlier models. This cell length is the shortest that has ever been used so far. Each time step is 1 second and therefore one unit of speed is equal to 3.6 km/h. One unit of acceleration is $1m/s^2$ and this corresponds to a 'comfortable acceleration' rate (Transportation Engineering, 1992).

In urban networks, a lower speed is required than on open roads due to speed constraints. The legal speed limit in urban networks is usually 50 km/h in New Zealand; however, some people drive at round 58 km/h, which is just below the apprehension limit (61 km/h in New Zealand). Therefore it is assumed that the maximum speed of each vehicle is between 46.8 km/h – 57.6km/h. This means that a vehicle can move through 13 -16 cells in 1 second when driving at the maximum speed.

Normally, previous models have implicitly assumed that the headway (=distance/speed) is 1 second (e.g. Nagel and Schereckenberg, 1992, 1996; Campri and Levi, 2000; Knospe *et al.*, 2002; Larraga, 2005). However, a safe headway of 2 seconds is recommended to all drivers and a minimum headway of 1.44 seconds is recommended by the New Zealand Land Transport Safety Authority (New Zealand Road Code, 2003). In practice, the headways that drivers use are normally shorter than 2 seconds (Neubert *et al.*, 1999) and longer than 1 second. In this research, over 10 hours of traffic data was

recorded between 16 August 2004 and 27 August 2004. An average car-following headway of approximately 1.5 seconds was observed in local urban networks and this 1.5-second rule has been built into the proposed models. Table 3.2 shows a comparison between the parameters of the proposed models and those of other CA models.

Table 3.2 Comparison of parameters in the proposed models versus those of earlier models. Ave A or D $(m/s^2)^{\#}$ - Average acceleration or deceleration, Ave headway (s)* - Average headways

Parameters (unit)	Proposed models	Earlier CA models
Length of each cell (m)	1	7.5
Time step (s)	1	1
Speed unit (km/h)	3.6	27
Ave A or D $(m/s^2)^{\#}$	1	7.5
Ave headways (s)*	1.5	1

3.3.2 Calibration

Most CA models treat all vehicles as being of the same length. In reality, different vehicle types occupy different numbers of cells in the model. In order to simulate traffic flow realistically, vehicle data (type and length) were recorded during 10 hour of morning peak hour traffic and are used in the models (see Table 3.3).

Table 3.3 Vehicle types, number of cells occupied and composition of traffic

Vehicle Types	Cells	Percentage of traffic (%)
Motorcycles (M)	3	2
Personal Vehicles (P)	5	78
Vans or minibuses (V)	7	11
Buses (B)	10	6
Other large vehicles (O)	13	3

The free flow model has been calibrated using real data recorded during off-peak hours, when there was less interaction between vehicles. These field observations were made in good weather in order to avoid driver behaviour fluctuations due to different weather conditions. The data was collected on five roads in Palmerston North city, New Zealand.

Over 20 data sets are used in this research. Five speeding stages were observed. On the steady stage (stage C), driver behaviour is basically uniform. Therefore the main concern in this chapter is driver behaviour relating to acceleration and deceleration between junctions. In order to compare driver behaviour, relating to acceleration and deceleration on different roads, the different speed-time relationships were collected into one graph. Figures 3.2 and 3.3 show speed changes of vehicles on urban roads in different acceleration stages. Figures 3.4 and 3.5 show speed changes of vehicles on the same roads in different deceleration stages.

The results clearly demonstrate that the dual-regime of acceleration and deceleration matches the real behaviour of vehicle drivers. As mentioned above, 1 unit of speed represents 1 m/s (3.6 km/h) in the real world and 1 unit of acceleration is equivalent to 1 m/s^2 (3.6 km/h^2). The parameters in expressions (3.1) ~ (3.5) have been calibrated and are shown below:

$$a_1 = \begin{cases} 1 & probability = 0.36 \\ 2 & probability = 0.64 \end{cases} \tag{3.9}$$

$$a_2 = \begin{cases} 0 & probability = 0.5 \\ 1 & probability = 0.5 \end{cases} \tag{3.10}$$

$$a_3 = \begin{cases} 1 & probability = 0.33 \\ -1 & probability = 0.33 \\ 0 & probability = 0.33 \end{cases} \tag{3.11}$$

$$b_1 = \begin{cases} 0 & probability = 0.5 \\ 1 & probability = 0.5 \end{cases} \tag{3.12}$$

$$b_2 = \begin{cases} 1 & probability = 0.22 \\ 2 & probability = 0.78 \end{cases} \tag{3.13}$$

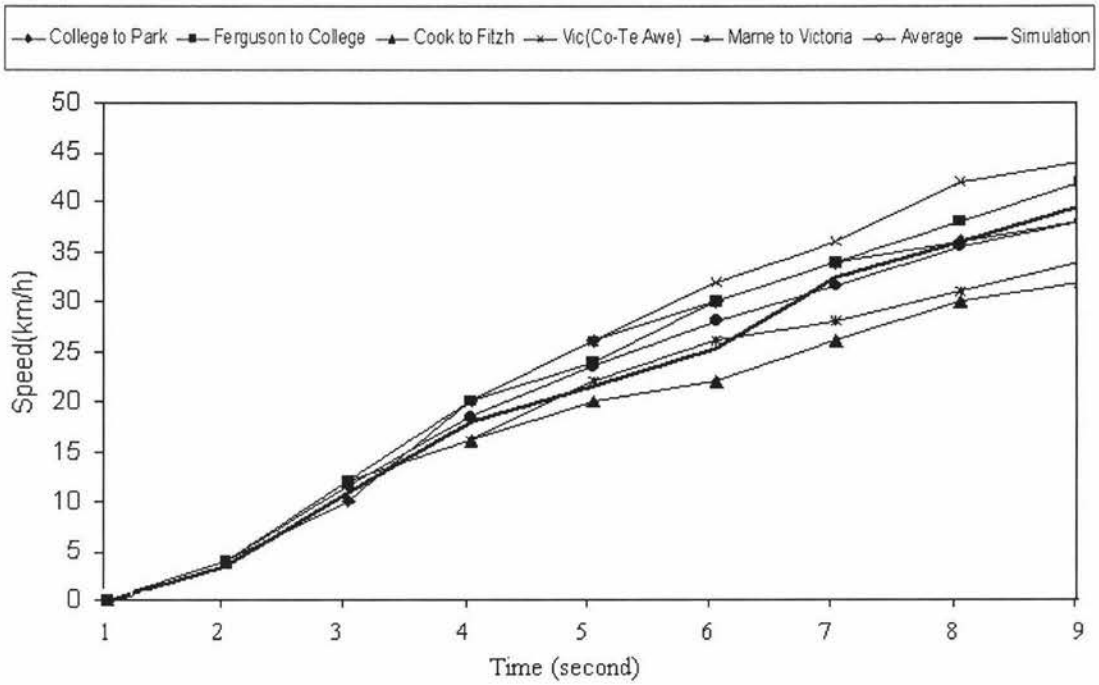


Figure 3.2 Acceleration stage A in actual traffic and the proposed model

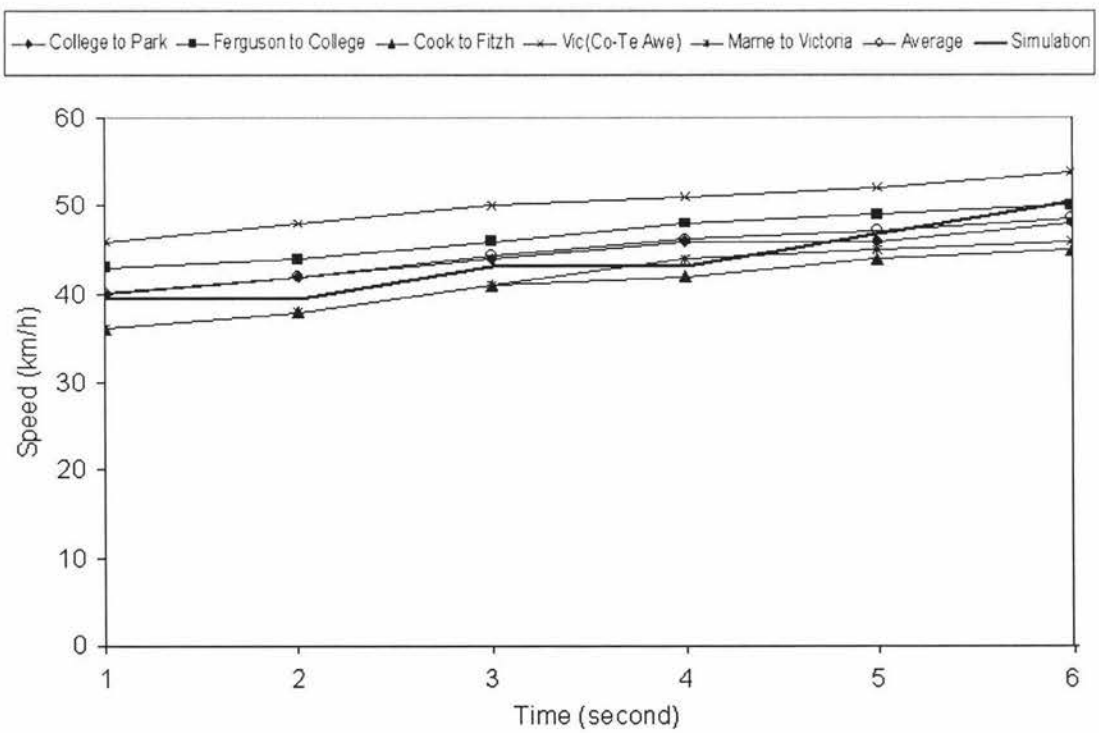


Figure 3.3 Acceleration stage B in actual traffic and the proposed model

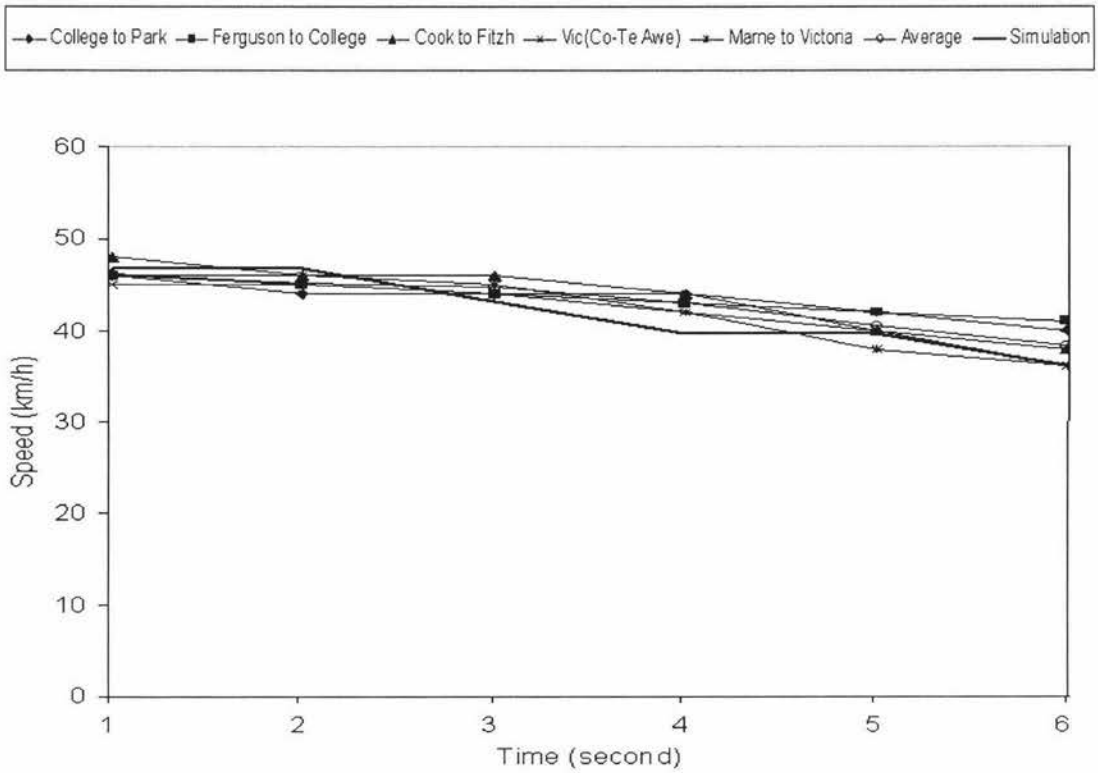


Figure 3.4 Deceleration stage D in actual traffic and the proposed model

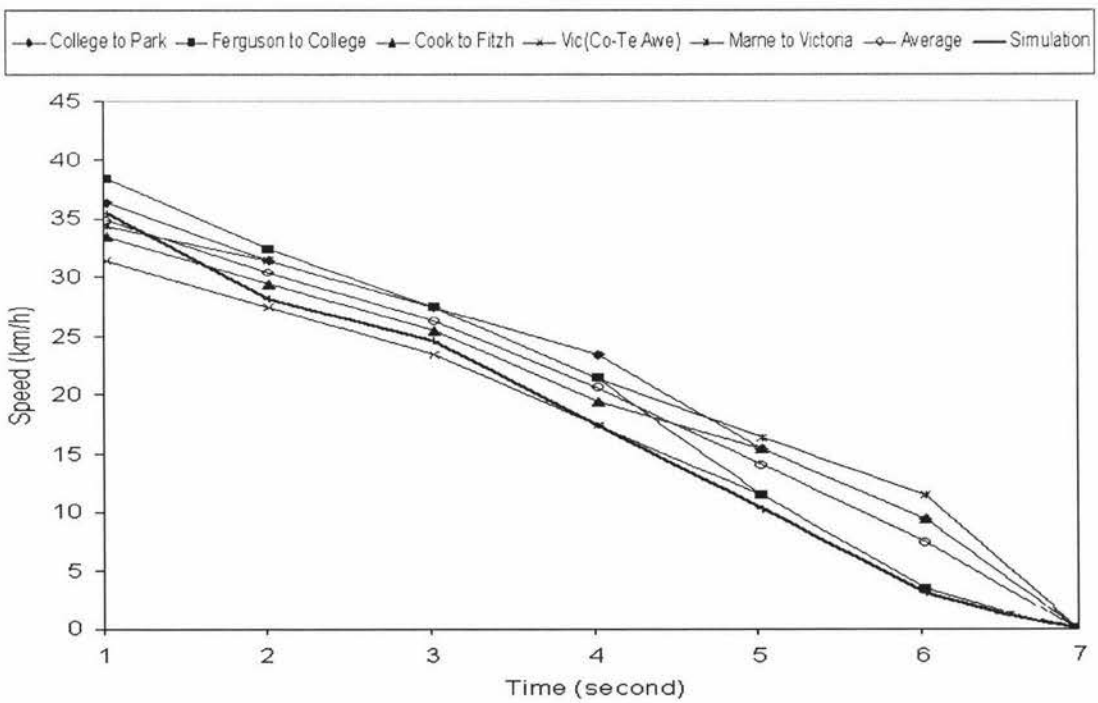


Figure 3.5 Deceleration stage E in actual traffic and the proposed model

Error tests are used to quantitatively measure the ability of fitness from simulation, compared to the field data. The mean relative error (MRE) test is one such test, which is expressed as:

$$MRE = \frac{1}{n}\sum_{i=1}^{n}\frac{|x_i - x_s|}{x_i} \tag{3.14}$$

where x_i is the actual observation value and x_s is the simulated value.

Both the simulation results and the MRE results (see Figure 3.6) show that the proposed models can closely reproduce the field data with regards to the speeds of individual vehicles in different acceleration and deceleration stages.

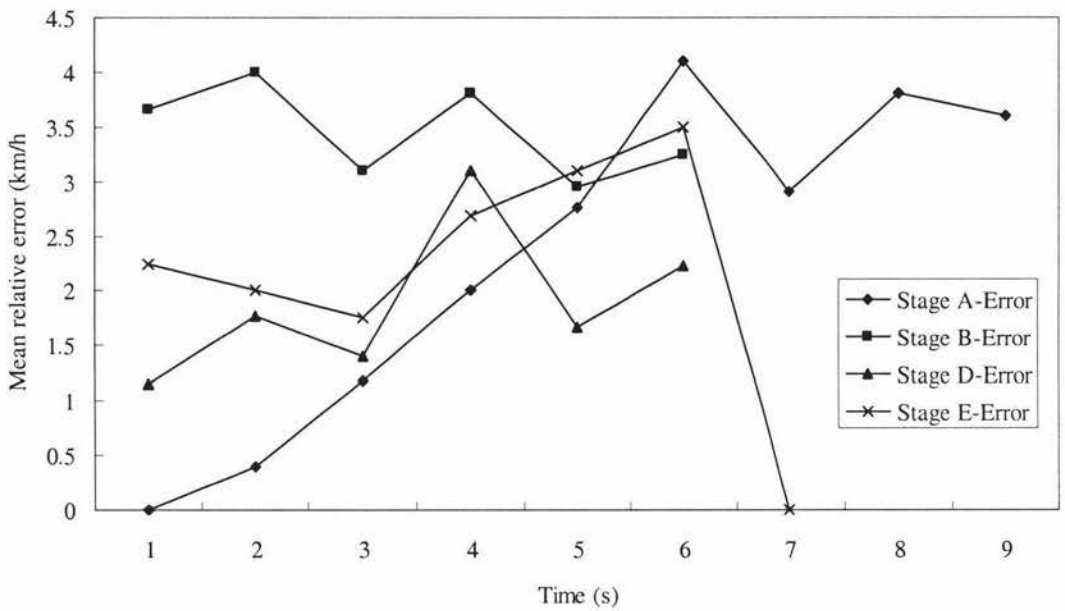


Figure 3.6 Mean relative errors in the different acceleration/deceleration stages

3.3.3 Validation

Benekohal (1989) indicated that validation of models should be performed at both the microscopic and macroscopic levels. For microscopic validation, the speed or position of individual vehicles should be compared to the field data. For macroscopic validation, average speed, average density and flow should be compared to corresponding values from the field data. In this research, due to the lack of average speed and flow data within a certain time intervals, only microscopic validation is carried out. Fig. 3.7 shows observed single-vehicle speeds from three different drivers

(driver 1, driver 2 and driver 3) and the results of simulation using the proposed method with calibrated values for acceleration and deceleration rates. The field data in this study was recorded using a digital video camera between Cook Street and Fitzherbert Avenue in Palmerston North, New Zealand. The multi-regime of acceleration and deceleration means that the model maintains a good alliance with the real data, especially in the initial acceleration and final deceleration stages.

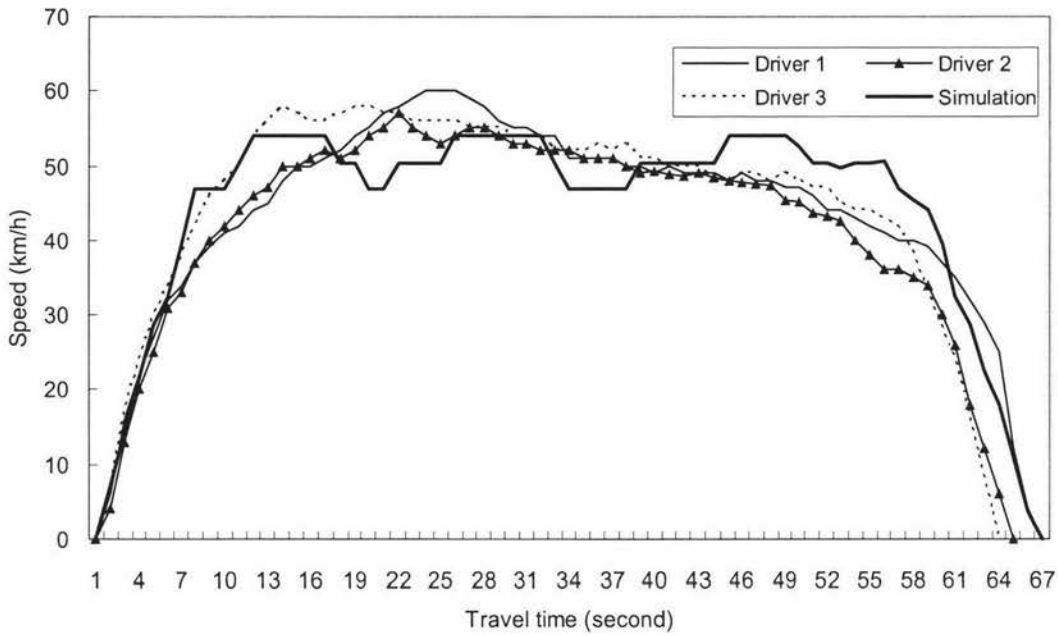


Figure 3.7 Observed single-vehicle speeds and simulation results

3.4 Summary

Modelling traffic flow in urban networks requires both straight road simulation and cross traffic simulation. With regards to modelling traffic flow on urban straight roads, two types of traffic flow, free and car-following flow, have been considered in this chapter. In free flow, a vehicle can move at its desired speed. In this case, the speed changes depend on the driver's preference, because there are enough time steps to stop the vehicle. In this chapter, a novel five-stage (two acceleration stages, one steady state and two deceleration stages) CA model is developed to model free flow. The results clearly demonstrate that the dual-regime of acceleration and deceleration matches the real behaviour of vehicle drivers. This approach can simulate traffic flow more realistically.

In the car-following flow, a vehicle has to follow the vehicle in front. In this case, the speed changes of the object vehicle depend on the distance to the preceding vehicle. Two stages of speed change (close tailing and semi-free flow) are presented to model the car-following flow. The average headway of car following was observed to be 1.5 seconds in local urban networks. We call this phenomenon as 1.5-second rule. This 1.5-second rule has been implemented in modelling car-following processes, which corrects 1-second headway used in previous CA models.

Fine-grid CA are used in our models. The length of each cell corresponds to 1 m on a real road. Correspondingly, one speed unit is 3.6 km/h. In other previous urban models, such as the improved BML models, the length of each cell corresponds to 7.5 m and one speed unit corresponds to 13.5 km/h. Therefore, our models can describe traffic flow more precisely than those CA models in which the cell length is longer than 1 m.

The proposed models are used to simulate traffic flow on single urban straight roads. These models can easily be extended to simulate traffic flow on multiple straight roads.

Chapter 4
Unsignalised Intersections

4.1 Introduction

Unsignalised intersections are an important component of urban networks. There are two main types of unsignalised intersections: Two-Way Stop-Controlled (TWSC) intersections and All-Way Stop-Controlled (AWSC) intersections (HCM, 2000). AWSC intersections are typical in North America. TWSC intersections are mostly used in the UK, Ireland, Australia and New Zealand. In New Zealand, there are Two-Way Give-Way-Controlled (TWGWC) intersections, which are similar to TWSC intersections. The focus of this chapter is single-lane TWSC/TWGWC intersections, which are suitable for the New Zealand environment.

The study of traffic flow at TWSC intersections is mainly focused on performance measurements, such as capacity (the number of vehicles passing an entrance during a specified time period, usually expressed as vehicles per hour), delay (the additional travel time experienced by a driver, passenger, or pedestrian) and queue lengths.

Interactions between drivers at TWSC/TWGWC intersections have attracted attention from many disciplines, such as physics, mathematics and computer science. Various interaction models have been developed, for example, gap-acceptance models (Tian *et al.*, 1999; Troutbeck and Kako, 1999; Bunker and Troutbeck, 2003) and cellular automata (CA) models (Chopard *et al.*, 1998; Ruskin and Wang, 2002; Dupius and Chopard, 2003; Wang, 2003). The limitations of gap-acceptance models have been analysed and detailed by Wang (2003). In addition, Wang (2003) suggests that different road features can be regarded as a unified system which means different road geometries can be modelled using the same method. Thus, in this chapter, the focus is on using a CA model to simulate traffic flow at a single-lane TWSC

intersection.

In this chapter, a novel CA model, based on the Normal Acceptable Space (NAS) method, is proposed to simulate driver behaviour and interactions between drivers at entrances of cross traffic, which is derived from the MAP (Minimal Acceptable sPace) method proposed by Wang and Ruskin (2002, 2006). The NAS method can facilitate understanding of the realistic interactions between drivers and provide a unified framework to model traffic flow at TWSC intersections and roundabouts.

4.2 Background

Traffic flow at TWSC/TWGWC intersections has to obey both priority and stop or give-way rules. Figure 4.1 illustrates a hierarchy of streams. Westbound and eastbound streets are major streams, while southbound and northbound are minor streams. Figure 4.1 (a) shows a classification of four levels of priority, in terms of the twelve arrows produced for traffic driving on the right. This ranking system is presented by the Transportation Research Board (HCM, 2000). However, this method cannot be directly applied to driving on the left side, particularly when special traffic rules are dominating the traffic flow. For instance, in New Zealand, turning-left vehicles give way to turning-right vehicles travelling in the opposite direction. Figure 4.1 (b) shows the streams ranking for left-side driving, according to the New Zealand Road Code (2003). The lower rankings mean higher priorities.

The priority rules at TWSC/TWGWC intersections, based on the New Zealand Road Code (2003), can be summarized in the following ways:

- All entering vehicles give way to all vehicles on the intersection
- A left-turning (LT) vehicle from a major-stream gives way to the oncoming, straight-ahead vehicle (ST) or a right-turning vehicle (RT) from another major-stream.
- A vehicle from a minor street gives way to all vehicles on the major road
- A LT vehicle from a minor street gives way to the oncoming ST or RT

vehicles from another minor street.

A stop or give way rule means that a vehicle from a minor street must stop or give way to other vehicles, according to the priorities, before entering the intersection, even if there is no vehicle on the major street.

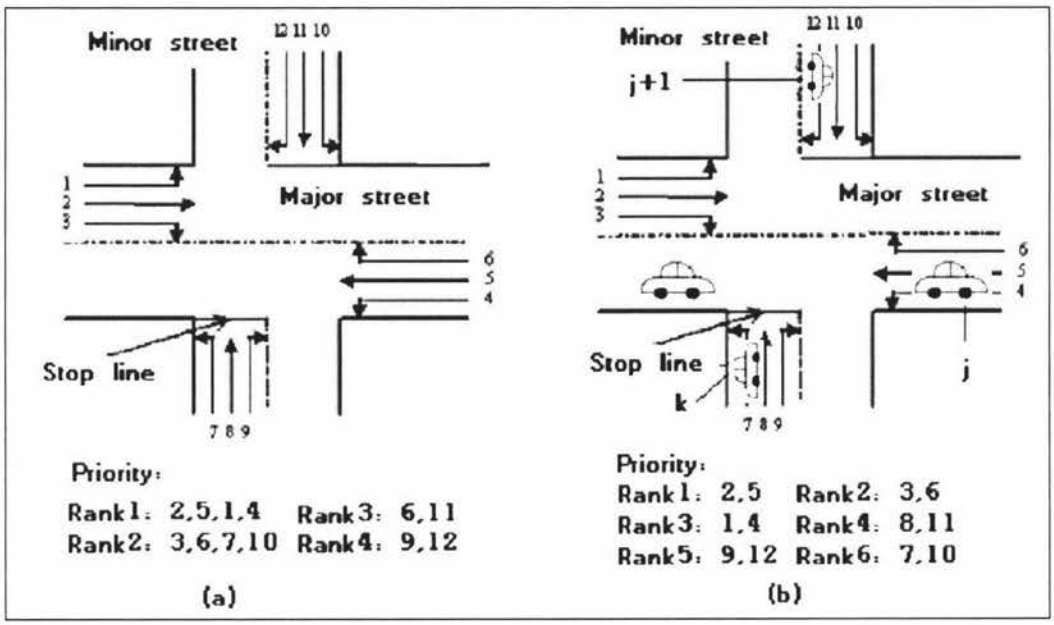


Figure 4.1 Illustration of roads and an unsignalised intersection. (a) Sourced from the Highway Capacity Manual (2000). (b) Indicates the ranking of streams at TWSC intersections in New Zealand, according to the Road Code (2003)

Chopard *et al.* (1998) indicate that the BML model (see Section 3.1) did not consider the street exits. That is to say, vehicles can only move forward, not turn. Therefore, Chopard *et al.* (1998) propose a CA ring model (hereafter it is referred to as the CDL model) to simulate traffic flow at crossings. The CA ring model allows several one-dimensional CAs to be interconnected, as shown in Figure 4.2. The car inside the rotary has priority to move forward over any entering car. The intersection may have any number of branches and therefore numerical implementation is relatively simple. The rotary acts as a connection, which connects all branches to form a system.

The CDL model has had relative success in exploring some features of intersections without traffic lights (Chopard *et al.*, 1998; Dupuis and Chopard, 2003; Wang, 2003). However, one problem with their model is that all cars, from different branches, have equal priority to move in the rotary (Wang, 2003). In reality however, on a crossing without traffic lights, traffic flow is generally governed by yield rules, priority regulations or the 'give-way' rule. Therefore, vehicles from major roads have priority over vehicles from minor roads.

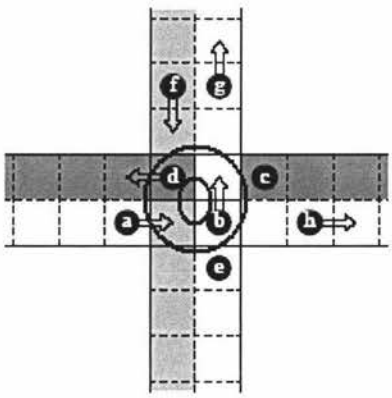


Figure 4.2 is a schematic diagram of a CA ring with a traffic configuration (Chopard *et al,*. 1998) together with a slight modification (two rings are added). The four central cells represent the rotary which is travelled counter clockwise. The dots labelled a, b, c, d, e, f, g and h denote cars which move forward and this is indicated by the arrows. A car without an arrow (e. g. dots c and e) is not allowed to move at the next time step

A common deficiency of the above-mentioned models, including gap acceptance and CA (e.g. Tian *et al.*, 1999; Troutbeck and Kako, 1999; Bunker and Troutbeck, 2003; Chopard *et al.*, 1998; Dupius and Chopard, 2003), is that until fairly recently driver behaviour was not considered as an input parameter. Akcelik (2005) indicates that driver behaviour is one of the main influences on capacity and performance characteristics. Therefore, driver behaviour is of paramount importance in analysing traffic flow at unsignalised intersections.

As Ahmed *et al.* (2002) suggest, driver behaviour is heterogeneous and inconsistent at cross traffic. The heterogeneous driver behaviour implies that different drivers may behave differently under the same conditions, while inconsistent driver behaviour means that a single driver may behave differently under similar conditions at each time step.

Ruskin and Wang (2002) propose a CA-based Minimal Acceptable sPace (MAP) method to simulate interactions between drivers at single-lane TWSC intersections. This method is able to simulate heterogeneous driver behaviour and inconsistent driver behaviour. In their model, heterogeneity and inconsistency of driver behaviour and interactions in *cross traffic* at entrances of intersections and roundabouts are simulated by incorporation of four different categories of driver behaviour (i.e. conservative, moderate, urgent and radical), together with reassignment of categories with given probabilities at each time step. If the entry criteria are met, vehicles can enter intersections, otherwise drivers have to wait and inconsistent driver behaviour is updated at the next time step.

Figure 4.3 illustrates interaction at an intersection entrance for a ST vehicle from a minor street (Wang and Ruskin, 2002b). A rational driver needs to observe 7 marked cells before s/he can drive onto the intersection (Figure 4.3 a). In contrast, a conservative driver needs to observe 9 marked cells (Figure 4.3 b). The dots marked with 0, a and b have the following meanings:

- '0' means that the cell is vacant
- 'a' means that the cell is either vacant or occupied by a LT vehicle
- 'b' means that the cell is not occupied by a RT vehicle

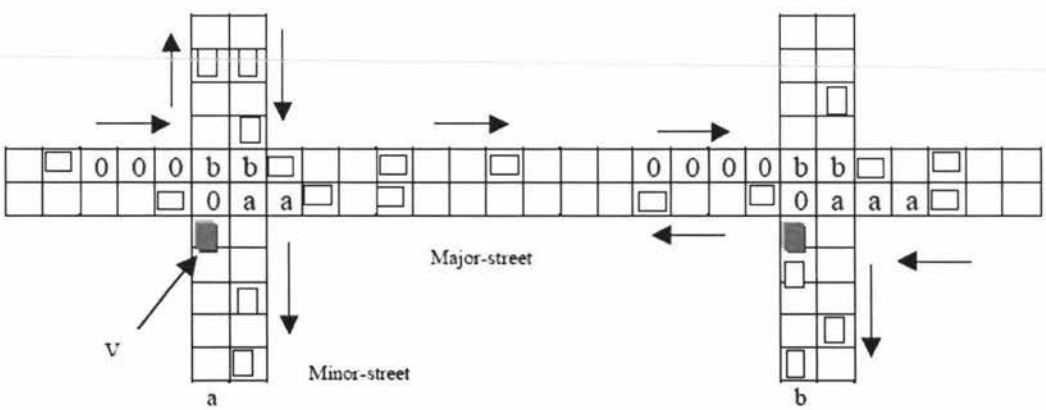


Figure 4.3 A ST vehicle from a minor street: (a) rational, (b) conservative (Wang, 2003)

The speed of vehicles in the MAP method is set to be one of values: 0, 1 or 2 (v_{max} = 2), which corresponds to a speed of 0, 25 km/h and 50 km/h. The length of each cell is approximately 7 m. This approach is the first model to reveal the impact

of driver behaviour on traffic flow at unsignalised intersections. The MAP method has been further used to simulate traffic flow at two-lane TWSC intersections (Wang and Ruskin, 2003a), single-lane roundabouts (Wang and Ruskin, 2002) and two-lane roundabouts (Wang and Ruskin, 2003b), respectively. Different lane partition for LT, ST and RT in two-lane intersections and roundabouts leads to different critical traffic flow according to the simulations in Wang and Ruskin (2002a and b). However, inconsistent driver behaviour, in the MAP method, can only be roughly described by these four categories in the MAP method.

4.3 Methodology

Modelling traffic flow at TWSC intersections includes vehicles moving onto straight roads, driver behaviour modelling at entrances, vehicles moving onto intersections, exiting from intersections and moving onto straight roads again. This process can be described in Figure 4.4. Vehicles moving onto straight roads have been modelled in Chapter 3. Vehicles moving onto and exiting from intersections are simply assumed to accelerate. Thus, in this section, driver behaviour assignment and interactions rules for entry intersections are the main points of the discussion.

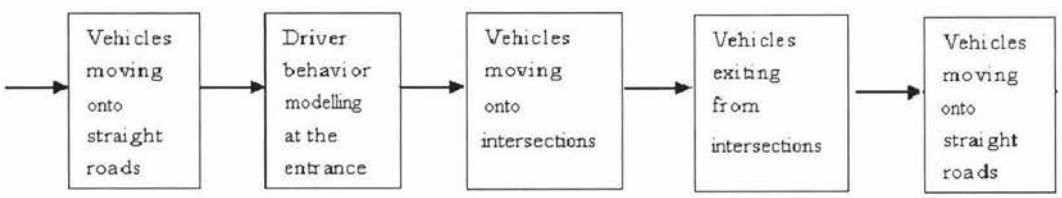


Figure 4.4 Illustration of the modelling process of traffic flow at TWSC intersections

4.3.1 Normal acceptable space (NAS) method

A novel and realistic CA model, based on the Normal Acceptable Space (NAS) method, is proposed to simulate heterogeneous driver behaviour and inconsistent driver behaviour at entrances of cross traffic, respectively. The NAS method is theoretically based on statistics and practically based on the fine grid (the length of each cell corresponds to 1 m in a real road). Statistical theory is used to describe driver behaviour; while the fine grid guarantees population distribution have a usual shape. The latter is extremely important in the NAS method. If the size of CA

cells is large, the population distribution of driver behaviour can not be described by Normal Distribution. If the grid is too small (< 1 m), sample accuracy is enhanced but it provides little help to analyse driver behaviour instead increases the computational complexity.

In this chapter, heterogeneous driver behaviour and inconsistent driver behaviour are modelled within a statistical framework and then the relationship between heterogeneous driver behaviour and inconsistent driver behaviour is discussed.

According to the central limit theorem (Hays, 1994), distribution of space required for all drivers at entrances of cross traffic is assumed to follow a Gaussian distribution. This population distribution is referred to as GDH (Gaussian distribution for heterogeneous driver behaviour). There are two pre-conditions required to judge whether a distribution can be roughly viewed as a Gaussian distribution in the central limit theorem: (1) independent sample size N is very large and (2) the population has a finite variance and a finite mean.

With regard to case (1), it is clear that the number of drivers is large, which meets the requirement of using very large samples in statistics. Although there is still no consensus of standard, on how large the sample size must be in order to apply the central limit theorem, a sample size of 30 or more (Hays, 1994) is usually regarded as sufficient. For a population of drivers, the condition of independent sample size $N \geqslant 30$ is quite easily satisfied.

With regard to case (2), each driver has his/her own special space criteria which are normally based on age, gender, driving skill and so on. However, in general, these space criteria fluctuate along a certain range and do not have very unusual values. That is to say, a finite variance and a finite mean exists. Therefore, according to case (1) and (2), population distribution (heterogeneous driver behaviour) can be approximately considered as a Gaussian distribution. The density function of GDH can be written as follows:

$$f(x) = \frac{1}{\sigma_H \sqrt{2\pi}} e^{\frac{(x-\mu_H)^2}{2\sigma_H^2}} \tag{4.1}$$

where x is random variable, μ_H is the mean and $\square_H$ is the deviation of GDH. Thus, the value y_n, defined as the required space to enter intersections for a driver n, can be determined by Eqn. 4.2. In Gaussian distribution, 99.73% confidence interval is given by $[\mu_H - 3\square_H, \mu_H + 3\square_H]$ (Bhattacharyya and Johnson, 1977). In Eqn. 4.2, a confidence interval of $[\mu_H - 3\square_H, \mu_H + 3\square_H]$ is simply extended to be 100%. Meanwhile, μ_H is assumed to fall in [0.4, 0.6]. Therefore, heterogeneous driver behaviour for driver n can be described as follows:

$$y_n = \begin{cases} \mu_H - 3\sigma_H & if \quad 0 \le p_n \le 0.0212 \\ \mu_H - 2\sigma_H & if \quad 0.0212 < p_n \le 0.1586 \\ \mu_H - \sigma_H & if \quad 0.1586 < p_n < 0.4 \\ \mu_H & if \quad 0.4 \le p_n \le 0.6 \\ \mu_H + \sigma_H & if \quad 0.6 < p_n \le 0.8414 \\ \mu_H + 2\sigma_H & if \quad 0.8614 < p_n \le 0.9788 \\ \mu_H + 3\sigma_H & if \quad 0.9788 < p_n \le 1 \end{cases} \tag{4.2}$$

With regard to a single driver n at entrances of TWSC intersections or roundabouts, the value $d(n)$ can be viewed as the required space to enter (the number of required unoccupied cells in CA models). Therefore, $d(n)$ can be one of a set of data sets in the data records, namely, $d(n) \in \{d_1(n), d_2(n), d_3(n), \ldots, d_l(n)\}$, $l = 1, 2, 3, \ldots, L$. L is the number of possible values which is required by a driver to enter the junction. The value $d(n)$ may consist of two independent parts:

$$d(n) = d + \square_n \tag{4.3}$$

That is, $d(n)$ is a sum of a constant part d plus a random and independent *error* component $\square_n$. Moreover, the *error* portion can itself be thought of as a sum of m components (Hays, 1994):

$$\square_n = g(e_1 + e_2 + e_3 + \cdots + e_m) \tag{4.4}$$

Here, e_i is a random variable that can only take on one of following two values:

$e_i = 1$, when factor i is valid

$e_i = 0$, when factor i is not valid

Thus, the error $\square_n$ can be roughly assumed to follow a Gaussian distribution, i.e., $\square_n \sim N(0, \square_n^2)$, if the inconsistent driver behaviour of driver n is sampled many

times. In reality, the basic value d stands for the driver's habits under normal driving circumstance. $\square_n$ is influenced by many stochastic factors such as temporary road construction, a pedestrian crossing, bad weather, night visibility, peak hour. However, these stochastic factors normally impact slightly on constant part d. $d(n)$ fluctuates along the basic value d. Therefore, $d(n)$ is also assumed to approximately follow a Gaussian distribution , i.e., $d(n) \sim N(d, \square_n^2)$. The possibility density function of $d(n)$ can be written as follows:

$$f_n(x) = \frac{1}{\sigma_n \sqrt{2\pi}} e^{\frac{(x_n - d_n)^2}{2\sigma_n^2}} \tag{4.5}$$

In general, a driver may accept a value which is less than $d(n)$, due to a long waiting time or other urgent conditions, while a driver may also accept a value which is larger than $d(n)$, due to bad weather, night visibility or other factors. Let x_{min} represent the number of minimum acceptable cells and x_{max} stand for the number of maximum acceptable cells for a driver to interact with other drivers. If $x > x_{max}$, a vehicle can pass the intersection without delay and there is no interaction. Values less than x_{min} are rejected, due to safety factors and values larger than x_{max} do not need to be considered, due to no interaction involved (free flow). Therefore, the resulting model can be viewed as a *truncated* Gaussian distribution (Barr and Sherrill, 1999), where the left and right parts have been cut off (Johnson, 2002). Thus, the distribution of space required for a single driver can be assumed to follow a *truncated* Gaussian distribution, which is referred to as a Gaussian distribution for inconsistent driver behaviour (GDI). It can be mathematically written in Eqn. 4.6. Therefore, the value z_n of inconsistent driver behaviour for the nth driver can be expressed by Eqn. 4.7. According to Eqn. 4.6 and 4.7, the value of z_n can be determined by Eqn. 4.8. Therefore, inconsistent driver behaviour for driver n can be described as follows:

$$f_n(x) = \frac{1}{\sigma_n \sqrt{2\pi}} e^{\frac{(x_n - \mu_n)^2}{2\sigma_n^2}} \qquad x_{\min} \le x \le x_{\max} \tag{4.6}$$

$$z_n = \begin{cases} \mu_n - 3\sigma_n & if \quad 0 \le p'_n \le 0.0212 \\ \mu_n - 2\sigma_n & if \quad 0.0212 < p'_n \le 0.1586 \\ \mu_n - \sigma_n & if \quad 0.1586 < p'_n < 0.4 \\ \mu_n & if \quad 0.4 \le p'_n \le 0.6 \\ \mu_n + \sigma_n & if \quad 0.6 < p'_n \le 0.8414 \\ \mu_n + 2\sigma_n & if \quad 0.8614 < p'_n \le 0.9788 \\ \mu_n + 3\sigma_n & if \quad 0.9788 < p'_n \le 1 \end{cases} \tag{4.7}$$

$$z_n = \begin{cases} \max(z_n, x_{\min}) & if \quad z_n < x_{\min} \\ \min(z_n, x_{\max}) & if \quad z_n > x_{\max} \end{cases} \tag{4.8}$$

The relationship between GDH and GDI is that for the *n*th driver the value of y_n in Eqn. 4.2 is equal to the value of μ_n in Eqn. 4.7. In this way, the heterogeneous driver behaviour and inconsistent driver behaviour can be systematically simulated by the NAS method.

In this chapter, Gaussian distribution is used to describe required space bias from normal or average values for driver population and a single driver. This assumption is comparable to other Gaussian distribution examples noted by Carlson and Thorne (1997) such as heights of boys in the same ages, weights of food packages, total sales or production, daily sales. *Moreover, according to centre limit theory (Carlson and Thorne, 1997), distributions of space requirement for driver population and a single driver can also be assumed to be Gaussian distributions.* That is, Gaussian distribution normally models spatial features of samples. In contrast, Poisson distribution and exponential distribution are usually related to temporal features of samples.

The Poisson distribution is used to compute the probability of x events in a unit of time (Carlson and Throne, 1997). For instance, arrival rates of vehicles in a one-hour period are normally assumed to follow a Poisson distribution, which has been adopted by Ruskin and Wang (2002), Fouladvand *et al.* (2003), Wang and Ruskin (2003a, 2003b and 2006). The exponential distribution is used to model a number of waiting line problems (Carlson and Throne, 1997), including the time between cars arriving at toll booths, customers arriving at a bank teller and telephone calls to a business. The negative exponential distribution has been widely used to

measure time headway that the next vehicle will arrive at intersections (Troutbeck, 1998; Tian *et al.*, 1999).

4.3.2 Interaction rules for entering intersections

At the entrances of intersections, interactions between drivers rely on priority rules. These rules can be represented by a classification of six levels of stream ranking in Figure 4.1(b). Lower ranking means higher priority. When a vehicle is preparing to enter an intersection, the driver must check if there are vehicles with lower rankings in conflicting streams. For instance, if a vehicle is in front of the stop line in lane 7, the driver must check if vehicles are coming from lanes 5 and 12 in Figure 4.1(b).

The following notations are used in the NAS method (Figure 4.5):

l_k = the length of vehicle k;

$p_{k,j}(t)$ = the number of free cells between vehicle k and j at time step t;

$s_{k,n}(t)$ = gap between vehicle k and n at time step t;

$v_k(t)$ = speed of vehicle k at time step t.

$\square_k$ = the deviation from GDI of vehicle k;

y_k = the value of heterogeneous driver behaviour of vehicle k;

$z_k(t)$ = the value of inconsistent driver behaviour of vehicle k;

The following entry rules are performed by driver k and other drivers at entrances in parallel.

1. Assigning expected space to enter TWSC intersections for all drivers according to Eqn. 4.2 randomly.
2. Counting $p_{k,j}(t)$, $j = 1, 2, \ldots, q$. q is the number of vehicles with higher priorities than vehicle k. Vehicle k must give way to other vehicles with higher priorities in conflicting streams. Space between vehicle k and other vehicles in conflicting streams is counted.
3. Comparing $z_k(t)$ and $p_{k,j}(t)$. If $z_k(t) \leq \min \{ p_{k,j}(t), j = 1, \ldots, q\}$ and $l_k(t) \leq s_{k,n}(t)$, vehicle k can enter the unsignalised intersection, otherwise vehicle k has to wait. If the required space of vehicle k is not larger than the

space for other vehicles in conflicting streams, vehicle k can move into the intersection.

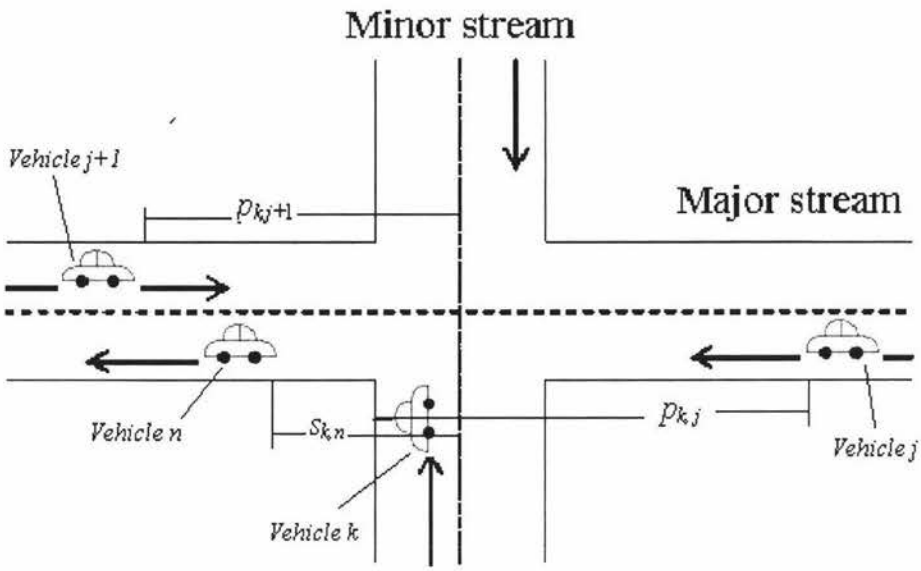


Figure 4.5 Schematic diagram of entry rules at single-lane TWSC intersections

4. Updating inconsistent driver behaviour. If vehicle k is waiting for entry, the update rule at each time step is as follows:

$$z_k(t) = y_k + \square_k \tag{4.9}$$

□is one of seven values {-3, -2, -1, 0, 1, 2, 3} and is determined by probability p_k in Eqn. 4.7. p_k is a random number within[0,1]. Different p_k corresponds to different □value.

4.3.3 Comparing NAS method with other methods

There are many differences between the NAS method and other CA methods (e.g. MAP method, model of Chopard *et al.*) and gap acceptance. The model of Chopard *et al.* (1998) is hereafter referred to as the CDL model.

In their latest paper (Pollatschek *et al.*, 2002) for the gap-acceptance model, heterogeneous driver behaviour is simply divided into two groups: cautious drivers and risk-loving drivers. Clearly, the assumption to categorize driver behaviour into

two groups is also coarse. Inconsistent driver behaviour has not been considered in their gap-acceptance model. Therefore, the NAS method is mainly compared with the MAP method and the CDL model, particularly in input parameters. Table 4.1 lists input parameters (except for turning rate, arrival rate) used in the NAS, MAP and CDL models.

Table 4.1 Input parameters in NAS, MAP and CDL

Input parameters	NAS	MAP	CDL
State of each cell	2	3	2
Length of each cell (m)	1	≈ 7	5
Speed unit (km/h) for legal limit 50km/h	3.6	25	50
Maximum speed (cells)	≈ 14	2	3
Vehicle types	5	2	1
Time step (s)	1	1	0.36
Normal acc/dec (m/s^2)	1	≈ 7	≈ 14
Heterogeneous driver behaviour	Gaussian distribution	4 groups	N/A
Inconsistent driver behaviour	Gaussian distribution	4 groups	N/A

The state of each cell in the MAP has three meanings: unoccupied (0), occupied (1) or occupied (2). The state of occupied (1) means that the cell is occupied and the current speed of the car is 1 (corresponding to 25 km/h). The state of occupied (2) implies that the cell is occupied and the current speed of the car is 2 (corresponding to 50 km/h). Each cell in the NAS and CDL has two states: empty (0) or occupied (1). However, a data structure is used to describe the properties of vehicles in the NAS compared to the MAP and CDL. The data structure of vehicles can be written as follows:

```
Typedef struct {
        int ength;
        int position;
        int destination;   /* Exit in roundabouts */
        int speed;
        int maxspeed; /* 16 cells, corresponding to 57.6 km/h */
        int gap;   /* distance to the preceding vehicle */
        int indicator;   /* 1=LT, 2=RT, 0 = other conditions */
```

```
        int types;   /* 5 types, for more details see Section 3.3.2 */
        int acceleration rate;
        int deceleration rate;
} vehicle;
```

There are five types of vehicles calibrated and used in the NAS method (see Section 3.3.2 for more details). Two types of vehicles are used in the MAP method. For small vehicles, the length is designated by one cell ($\approx$ 7 m) and two cells ($\approx$ 14 m) for large vehicles. There is only one type of vehicle used in the CDL model and the length of each vehicle is 5 m.

In the MAP method (Ruskin and Wang, 2002), heterogeneity and inconsistency of driver behaviour and interactions in *cross traffic* at entrances of intersections and roundabouts are simulated by incorporation of four different categories of driver behaviour (i.e. conservative, moderate, urgent and radical), together with reassignment of categories with given probabilities at each time step. Heterogeneous driver behaviour and inconsistent driver behaviour have not been modelled in the CDL model.

Compared to the MAP and CDL methods and gap acceptance models, the NAS method uses an analogous but more realistic and flexible methodology to simulate driver behaviour at entrances of cross traffic using two different Gaussian distributions.

4. 4 Simulation Results

In the simulations, the length of each cell corresponds to 1 m in a real road, which provides a better resolution modelling actual traffic flow than other models (e.g. 5 m in the CDL model and 7.5 m in model of Simon and Nagel (1998)). A unit of speed is therefore equal to 3.6 km/h and each time step is 1 second. Since 1 unit of acceleration is 1 m/s^2, this also corresponds to a 'comfortable acceleration' (Transportation Engineering, 1992).

4.4.1 Calibration

Field studies were performed to observe vehicle interactions at the entrance of the single-lane TWSC intersection. Traffic data was collected for three days, using a video camera recorder for one hour per day; during the morning peak hour period (8:00-9:00). Normally, there is quite high traffic flow during peak hours at this intersection.

Four-hour traffic data-sets, arrival rates, turning rates and vehicle types were extracted from the recorder. For convenience in computing different space requirements, the video data was transferred into a Microsoft Media Maker format and divided into 162 media clips. Therefore, traffic flow can be re-examined in a computer with playback and slow motion facilities. The observed distribution can be approximately considered as Gaussian distribution. Heterogeneous driver behaviour follows a GDH with the mean (μ_H) = 20 and deviation ($\square_H$) = 2, while inconsistent driver behaviour follows a GDI with deviation = 1 (mean is obtained from the GDH). The number of minimum acceptable empty cells x_{min} is approximately 14, while the number of maximum acceptable empty cells in interactions between drivers x_{max} is approximately 26.

In order to examine the feasibility of the NAS method, the capacity of a minor stream of a TWSC intersection is compared with the MAP method (Ruskin and Wang, 2002) under the same values of traffic flow parameters, such as arrival rates and turning rates. The following assumptions are made in the simulations for comparing the capacity of a minor stream:

- Only RT (Right-turning) and LT (Left-turning) vehicles are in the minor stream in a T-intersection.
- Only ST (Straight ahead) vehicles are in the major stream.
- The total number of vehicles per hour in major streams are assumed to be 1440 vph, which consists of the near-lane stream (vehicles coming from the right) and the far-lane stream (vehicles coming from the left).
- The total number of vehicles per hour in a minor stream is assumed to be 1080 vph.

Table 4.2 shows the capacity of a minor stream in the NAS method and in the MAP method. Both NAS and MAP methods have the same trend. For instance, the capacity of the minor stream decreases, when the number of LT vehicles decreases. There is an average increase of over 3% compared to the MAP method. This can be explained as follows: In the MAP method, heterogeneous driver behaviour is randomly divided into four categories: *conservative* (4 empty cell spaces needed to enter), *moderate* (3 empty cell spaces needed to enter), *urgent* (2 empty cell spaces needed to enter) *and radical* (1 empty cell needed to enter). Inconsistent driver behaviour is assumed to be one of four categories with a specified probability at each time step. This assumption is questionable. For instance, if a driver's behaviour, in a current time step, is radical (1 empty cell space needed to enter and the length of each cell approximately corresponds to 7 m in the MAP model), at the next time step it may be *conservative* (4 empty cells needed to enter). For the same driver, such great disparity in space requirement is unacceptable. However, the capacity in the NAS method is basically in agreement with the MAP method. Simulation results show that the NAS method is feasible.

Table 4.2 A comparison of the capacity of a minor stream with the NAS and MAP methods. TRR is turning rate ratio and FRR is flow rate ratio

TRR (= LT rate : RT rate)	Capacity (vph)									
	FRR (= Flow rate of near lane : Flow rate of far lane)									
	1440:0		1080:360		720:720		360:1080		0:1440	
1:0	196	212	397	413	585	600	755	801	900	960
0.75:0.25	193	200	363	401	483	502	527	556	415	452
0.5:0.5	190	195	331	340	413	422	408	420	286	310
0.25:0.75	183	194	308	304	361	365	337	351	217	230
0:1	177	180	288	290	321	335	286	314	180	188

By contrast with the MAP method, the NAS method uses two different Gaussian distributions to model heterogeneous driver behaviour and inconsistent driver behaviour at cross traffic, respectively. Thus, inconsistent driver behaviour can be simulated within a reasonable range. Different driver behaviour can be calibrated by adjusting the mean and deviation of the corresponding Gaussian distribution.

Therefore, the NAS method is more suitable for simulating inconsistent driver behaviour and heterogeneous driver behaviour than other CA or gap-acceptance models.

4.4.2 Validation

The NAS method has been validated using field data provided in Table 4.3. Input parameters, such as turning rate, volume, Peak Hour Factor (= the ratio of the hourly volume to the peak 15-minute flow rate) (HCM, 2000), and the percentage of vehicle types is the same as in the literature (WWW1). The length of normal vehicles, such as cars, is set as five cells and heavy vehicles, such as buses or trucks, are set as 10 cells in the NAS method. In the literature, westbound and eastbound are minor streams, which are different from Section 4.2, where they are major streams.

The method is then validated by comparing it with the results of capacity, delay and queue length, illustrated in Table 4.4. The 95% queue-length is calculated in terms of the queue equation in HCM (2000). The NAS method appears to agree well with the field data at the validation stage.

4.4.3 Application of the method

For the purpose of a realistic microscopic simulation, vehicle arrival rates, turning rates, vehicle types, driver behaviour, special interaction (limited priority mechanism), well-designed division of speed. have been considered and built into the MAP method. By calibrating the method, a realistic simulation is implemented.

The limited priority mechanism, proposed by Troutbeck and Kako (1999), is based on the gap-acceptance model. The definition of 'limited priority mechanism' is that drivers in the major stream at a merging area may incur a delay, which allows drivers in the minor stream to accept smaller gaps. This mechanism has been used in modelling traffic flows at roundabouts and freeway merging by Troutbeck and Kako (1999) and Bunker and Troutbeck, 2003), respectively.

Table 4.3 Field data used in the NAS method. VOL = Volume, PHF = Peak-Hour Factor, HFR = Hourly Flow Rate, PHV = Percent Heavy Vehicles, LT = Left Turn, ST = Straight ahead, RT = Right Turn

	Major Street						Minor Street					
	Northbound			Southbound			Westbound			Eastbound		
	LT	ST	RT	LT	ST	RT	LT	ST	RT	LT	ST	RT
VOL	8	184	5	165	273	10	10	35	116	1	67	11
PHF	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
HFR	8	193	5	173	287	10	10	36	122	1	70	11
PHV	3	0	0	3	0	0	3	3	3	3	3	3

Table 4.4 Comparison results of the NAS method and field data in the literature

Variables	Westbound		Eastbound	
	Field data	NAS	Field data	NAS
Capacity	485	506	272	301
V/C	0.35	0.33	0.3	0.27
95% Queue Length	1.53	1.4	1.23	1.0
Delay	16.3	13.25	23.8	18.8

The rate of limited priority (Bunker and Troutbeck, 2003) of vehicles from minor streams is arbitrarily set as 1%; that is, one out of every 100 minor-stream vehicles would enter the intersection, in terms of a limited priority rule.

Theoretically, if all arriving vehicles pass the intersection without queuing, the flow rate λ= 0.1, 0.2, 0.3... equivalents are 360vph (vehicle per hour), 720vph, 1080vph... respectively. A vehicle departure is assumed to follow a Poisson distribution. In the simulations, the departure rate of the major-stream vehicles vary between 0.1 and 0.2, while the departure rate of the minor-stream vehicles is fixed to be 0.1. Vehicles move forward in terms of the proposed model in Chapter 3. The turning rates of LT, ST, and RT are 0.2, 0.6, and 0.2, on all approaches.

Capacity and performance (delay and queue length) of a TWSC intersection have been investigated and detailed under different arrival rates (traffic volume) and

turning rates (turning proportion) by Wang (2003). Here, capacity is studied, based on different vehicle configuration and arrival rates (= westbound: eastbound). Experiments were carried out on ten occasions. Each time, the simulations were run, 3600 time steps (equivalent to 1 hour) for a street-length of 100 cells occurred on all approaches. Every simulation is repeated ten times and the resulting capacity is obtained by averaging the capacity of ten experiments.

Table 4.5 Capacities of a minor stream for various vehicle components, arrival rate and the limited priority

	With limited priority				Without limited priority			
Vehicles type (%)	Arrival rate on major streams (westbound : eastbound)				Arrival rate on major streams (westbound : eastbound)			
M:P:V:B:O	0.1:0.1	0.1:0.2	0.2:0.1	0.2:0.2	0.1:0.1	0.1:0.2	0.2:0.1	0.2:0.2
0:100:0:0:0	358	336	322	286	356	331	317	276
0:80:11:9:0	350	328	310	267	348	323	304	255
2:78:11:6:3	341	320	301	243	335	311	287	234

Table 4.5 shows that vehicle components, arrival rate and the limited priority rule can affect capacity. It can be seen that the capacity of the minor-stream generally decreases when the arrival rate of major-stream vehicles increases. This agrees with the results in the MAP method (Wang, 2003). The capacity of the minor-stream also decreases, when the percentage of private cars decreases and the percentage of other vehicles increases. However, this effect differs, as vehicle types vary. It can also be seen that the capacity has a lower increase with limited priority than that with no limited priority, under the same conditions, especially for congested traffic.

4.5 Summary

In this chapter, a novel and realistic CA model, based on the Normal Acceptable Space (NAS) method, is proposed to simulate heterogeneity and inconsistency of driver behaviour at entrances of TWSC intersections.

The NAS method is based theoretically on statistics and practically on a fine grid (length of each cell corresponds to 1 m on a real road). The fine grid is extremely important to assume driver behaviour follows a distribution. If the size of CA cells is large, the population distribution of driver behavior is hardly to be described by a normal distribution. If the grid is too small (< 1 m), the sample accuracy can be enhanced but it provides little more help to analyze driver behavior instead of increasing much more computational complexity.

The NAS method uses GDH to describe heterogeneity of driver behaviour and GDI to describe inconsistent driver behaviour. This dual-regime distribution provides a more realistic and flexible methodology, compared to that of other CA models and gap acceptance models.

A stream ranking in Figure 4.1(b) was obtained from the New Zealand Road Code (2003) and used to simulate interactions between drivers at entrances. Inconsistent driver behaviour is updated, according to combinations of mean and deviation of GDI with different probabilities.

The queue-length and delay of each stream, at each time step, can be directly obtained from the CA model. Also, the relationship between the performance measurements of intersections and parameters of traffic flow are easily obtained from the simulation.

Although the capacity of a minor stream in the CA model has an average of over 3% increase compared to the MAP method, simulation results agree well with the MAP method. Therefore, the NAS method is feasible to model driver behaviour at cross traffic.

Different driver behaviour can be easily calibrated by adjusting the mean and deviation of the corresponding Gaussian distributions. The proposed CA model has been validated using field data. Input parameters, such as turning rates, volume and percentage of vehicle types are the same as the field data. Simulation results of the CA model show a strong agreement with field data.

Chapter 5

Urban Roundabouts

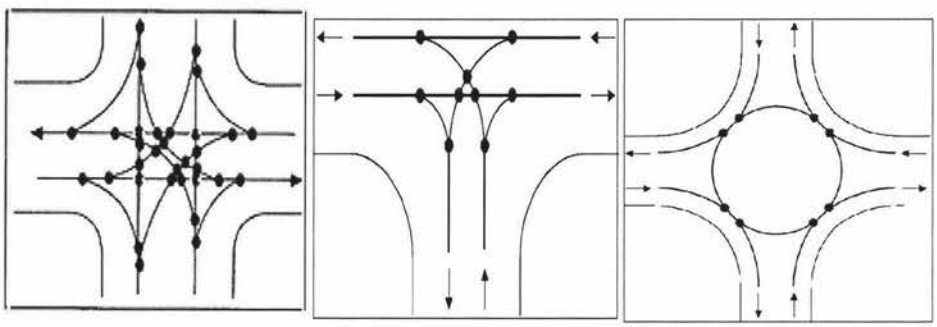
5.1 Introduction

Roundabouts have been considered as an alternative traffic facility that can improve safety and operational efficiency, compared to intersection controls. Some are controlled by traffic lights and others are governed by yield-at-entry. The definition of yield-at-entry (HCM, 2000) is that vehicles from the secondary roads give way to the vehicles on the circulatory road.

The main advantages for using roundabouts to improve traffic safety are as follows:

- Conflict points are decreased from 32 (single-lane intersections) to 8 (single-lane roundabouts) (see Figure 5.1). A study conducted by the Insurance Institute of Highway Safety (IIHS) has shown that total crashes decreased by 39% when using roundabouts instead of intersections (Russell and Rys, 2004).
- The speed is normally reduced to 20 ~ 30 km/h (Transportation Engineering Handbook, 1992).
- All traffic inside the roundabout comes from one direction.
- It is easy to determine passage priorities when there are more than four arms, compared to other intersection control methods.
- It makes U-turns easier, compared to T-intersections and Y-intersections.

However, roundabouts are not always convenient for pedestrians, cyclists and over-sized vehicles (Russell and Rys, 2004).



(a) (b) (c)

Figure 5.1 Number of conflicts points in: (a) 32 for single-lane intersections, (b) 9 for T-intersections and (c) 8 for single-lane roundabouts (Cape Cod Commission , 2003)

5.2 Background

Both empirical and theoretical methods (Akcelik, 2003) are proposed to measure roundabout capacity and performance such as delay and queue length. With regard to both methods, gap-acceptance criteria (Flannery and Datta, 1997; Troutbeck, 1998) are commonly used. Gap-acceptance models are, however, generally unrealistic in assuming that drivers are consistent and homogenous (Wang, 2003). A consistent driver would be expected to behave in the same way in all similar situations, while in a homogenous population, all drivers have the same critical gap (the minimum time interval, between two major-stream vehicles, required by one minor-stream vehicle to pass through) and are expected to behave uniformly. The limitations of gap-acceptance models have been analysed and detailed by Wang and Ruskin (2002).

More recent studies in gap-acceptance models have been made by Pollatschek *et al.* (2002) and Bunker and Troutbeck (2003). Pollatschek *et al.* (2002) proposed a microscopic decision model for driver gap-acceptance behaviour, when waiting at a roundabout entrance. Heterogeneous driver behaviour is divided into two groups: cautious drivers and risk-loving drivers. One parameter α is used to represent the propensity to take a risk. In their model, a small α is typical of a risky driver while a large α typifies a cautious driver. The entry capacity is assumed to be the function of the parameter α and the aggregated gap. Their simulation results show that the smaller the parameter α, the larger is the entry capacity at roundabouts. However, the assumption to

categorise driver behaviour into two groups is also coarse and the percentage of heterogeneous driver behaviour is not quantified. Inconsistent driver behaviour has not been considered in their model.

Recently, a limited priority gap-acceptance model (Bunker and Troutbeck, 2003) has been proposed to estimate minor stream delays at uncongested freeway merging. In the limited priority model, drivers in the major stream at a merging area may incur delay, which allows minor stream drivers to accept smaller gaps. The headway distributions are assumed to follow Cowan's M3 (Cowan, 1975). However, limited priority seems to be unnecessary for uncongested conditions. Their model is probably best suited for congested highway merging or roundabouts.

The employment of the CA to model traffic flow at roundabouts has attracted attention in the last few years (Fouladvand *et al.*, 2003; Wang and Ruskin, 2002, 2003; Campari *et al.*, 2004, Wang and Liu, 2005), due to their dynamic and discrete characteristics (Toffoli and Margolus, 1987) and its connection with the stochastic nature of traffic flow (Nagel and Schreckenberg, 1992).

Fouladvand *et al.* (2003) propose a stochastic CA interaction model. With this model, a waiting vehicle can enter the roundabout, only if there are no vehicles on the roundabout in its left side quadrant. Obviously, if a vehicle is on the roundabout in its left side quadrant and will exit from its left exit, the waiting vehicle may certainly enter the roundabout. This condition has not been mentioned in their paper. Each time step in the model is equivalent to 2 seconds. Clearly, this model is not able to describe vehicular dynamics in detail, such as acceleration or deceleration. A time step in micro-simulation is recommended between 0.1 and 1 second (Brackstone and McDonald, 2003).

A simpler entry rule is also presented by Campari *et al.* (2004), that is, if the cell located in front of the entrance is not occupied by a vehicle, a waiting vehicle is randomly generated and the cell is then occupied by a vehicle. However, the Yield-at-Entry rule is not observed by and the speed of the following vehicle on the circulatory lane is not considered. Therefore, their model is unlikely to be a safe model.

A common deficiency of the models put forward by Fouladvand *et al.* (2003) and Campari *et al.* (2004) is that the effects of driver behaviour is not considered. Akcelik (2005) indicates that driver behaviour is directly related to capacity and performance measurement. Therefore driver behaviour is of paramount importance in analysing traffic flow in urban networks.

The MAP method (Wang and Ruskin, 2002) has also been applied to simulate interactions between drivers at single-lane roundabouts. Assignment of driver behaviour at the entrance to roundabouts is the same as that of single-lane TWSC intersections. The weakness of the MAP method has been analysed in section 4.2. Furthermore, a two-lane urban roundabout simulation has been carried out using two CA-rings, with the same number of cells in each ring (Wang and Ruskin 2003b). As the speed is lower in the inner-lane than that in the outer-lane, the same number of cells in each ring is acceptable. Different lane-allocation patterns are well analysed by Wang and Ruskin (2003b, 2006). In this chapter, the NAS method and several realistic features described in Chapter 4 are also used to simulate traffic flow at a single-lane roundabout, operating under the yield-at-entry rule. Capacity and performance (delay, queue length) are also studied.

5.3 Methodology

Modelling traffic flow at a roundabout includes vehicles moving onto straight roads, driver behaviour modelling at the roundabout entrance, vehicles moving onto the roundabout, exiting from the roundabout and moving onto a straight road again. These processes are illustrated in Figure 5.2. Vehicles moving on a straight road between intersections have been modelled in Chapter 3. In this chapter, interaction rules at entrances, moving onto a roundabout and exiting from a roundabout are analysed in detail.

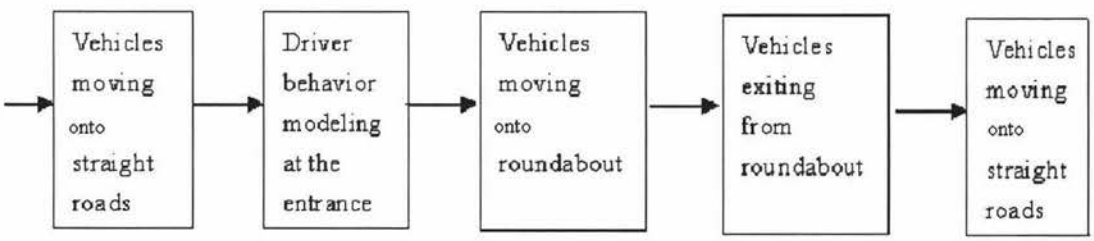


Figure 5.2 Illustration of modeling process of traffic flow at roundabouts

5.3.1 Interaction at roundabout entrances

Under the Yield-at-Entry rule, vehicles from secondary roads must give way to vehicles on a roundabout. Therefore, there is only one conflicting stream at a roundabout's entrance, compared with the TWSC intersections, where drivers need to simultaneously observe more than one conflicting stream before entering.

Drivers need to judge whether the current available space on the roundabout is sufficient for them to drive onto it and they need to ensure that their vehicles will not cause following ones to slow down significantly. The judgement process involves the comparison of current available space and expected space. If the current space is larger than the expected space, the vehicle can enter the roundabout. Otherwise, it has to wait. Factors that influence a driver's decision include age, gender, driving skills, the weather, the performance of their vehicle, motivation of travel, road conditions. This decision-making process may vary for each individual driver. These factors make driver behaviour heterogeneous and inconsistent. Similar to the method used in the TWSC intersections, the distribution of driver behaviour can be represented by GDH and GDI which have been used in Chapter 4.

For the purpose of simplicity, a four-armed single-lane roundabout is used to describe the interaction process in Figure 5.3. Figure 5.3 illustrates the location of vehicles and the topology of the road and the roundabout. Vehicles on the roundabout are numbered in the circulatory lane, namely, vehicle $n + 1$ precedes vehicle n; vehicle n precedes vehicle $n - 1$, and so on. Vehicle $n + 1$ has passed the point P and vehicle n is approaching the point P. Vehicle k is waiting to enter the point P at the entrance.

When a vehicle is approaching an entrance to a roundabout, the driver needs to observe other vehicles to the right of the entrance and their indicators (left-side driving). For instance, in Figure 5.3, even if the indicator of the left-turn (LT) vehicle n is on, the driver still needs to count gap to the vehicle $n - 1$ for the purpose of safety.

The following notations are used for subsequent descriptions:

l_k = the length of vehicle k;

$s_{k,n-1}(t)$ = the number of free cells between vehicle k (point P) and the front bumper of vehicle n-1 at time t.

$s_{k,n}(t)$ = the number of free cells between vehicle k (point P) and the front bumper of vehicle n at time t.

$s_{k,n+1}(t)$ = the number of free cells between vehicle k (point P) and the rear bumper of vehicle n+1 at time t.

$v_n(t)$ = speed of vehicle n at time step t.

σ_k = the deviation from GDI of vehicle k;

y_k = expected space for heterogeneous driver behaviour of vehicle k;

$z_k(t)$ = expected space for inconsistent driver behaviour of vehicle k;

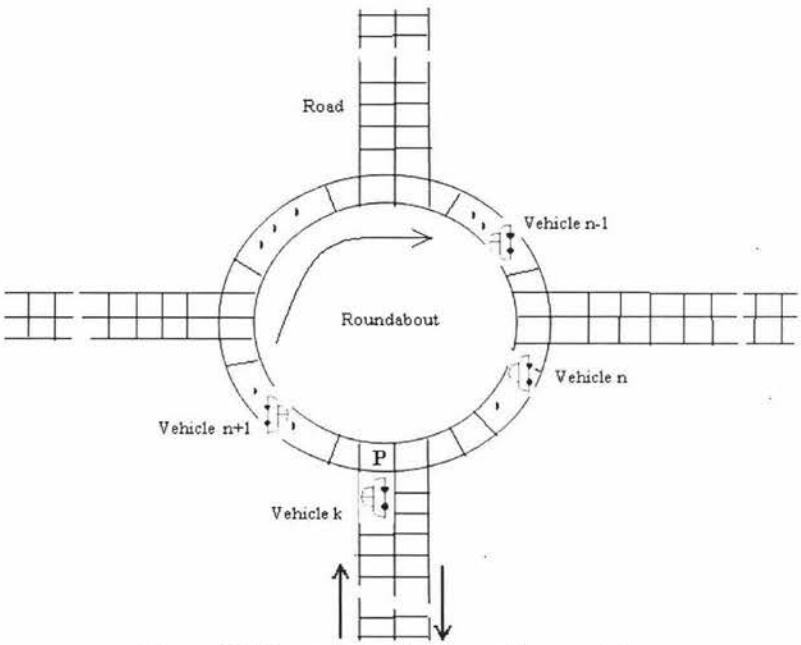


Figure 5.3 A topology of roads and a roundabout

The following entry rules are performed for vehicle k and other vehicles at a roundabout entrance in parallel:

1. Assigning y_k (expected space of heterogeneous driver behaviour) for vehicle k according to Eqn. 4.2 in Chapter 4.
2. Obtaining expected space of inconsistent driver behaviour for vehicle k according to the following Equation. :

$$z_k(t) = y_k + \lambda\sigma_k$$

λ has one of seven values {-3, -2, -1, 0, 1, 2, 3}, which is determined by probability p in Eqn. 4.7 in Section 4.3.1.

3. Counting $s_{k,n+1}(t)$, $s_{k,n}(t)$ and $s_{k,n-1}(t)$.

   The vehicle k can enter the roundabout if:

   - $z_k(t) \leq s_{k,n}(t)$ and $l_k \leq s_{k,n+1}(t)$, namely, current space is larger than expected space and number of free cells in front is larger than length of vehicle k, or;
   - $z_k(t) > s_{k,n}(t)$, $z_k(t) \leq s_{k,n-1}(t)$ and $l_k \leq s_{k,n+1}(t)$, but the indicator of left-turn of vehicle n is on, or;
   - $l_k \leq s_{k,n+1}(t)$ and the indicators of left-turn of vehicle n and $n-1$ are on.

   Otherwise vehicle k cannot enter the roundabout.

4. Updating expected space of inconsistent driver behaviour. If vehicle k is waiting for entry, the update rule at the next time step is as follows:

$$z_k(t) = y_k + \lambda\sigma_k$$

   λ has one of seven values {-3, -2, -1, 0, 1, 2, 3} and is determined by probability p in Eqn. 4.7 in Section 4.3.1.

5.3.2 Update rules on roundabouts

The state of cells on roundabouts has different meanings in the MAP and NAS methods. Each cell on the roundabout has two states: empty (0) and occupied (>0) in the MAP method (Wang and Ruskin, 2002). If the cell is occupied, the value of the state represents the distance to the destination exit. The distance will decrease by one if the vehicle moves forward one cell. If the number is equal to 1, it indicates that the vehicle will leave the roundabout at the next time step. In the NAS method, each cell also has two states, but this only indicates whether the cell is empty or occupied. A detailed data structure is used to describe input parameters of vehicles including length, types, speed, destination exit which are given in Section 4.3.3.

The update rules for vehicles on the roundabout are as follows. Let $s_n(t)$ denote the number of unoccupied cells in front of vehicle n at time t, $v_n(t)$ denotes the current speed of vehicle n, and $e_n(t)$ denotes the number of cells between current position and the exit of vehicle n. The update rules of vehicle n on the roundabout can thus be summarized as follows:

- If $v_n(t) < \min\{s_n(t), e_n(t)\}$, then vehicle n keeps current speed(does not accelerate on the roundabout) (5.1)

- Otherwise, $v_n(t) = \min\{s_n(t), e_n(t)\} - 1$ (5.2)

The assumption that vehicles do not accelerate on a roundabout (Eqn. 5.1) is based on the following considerations: firstly, for a normal roundabout (Ashley, 1994), the inscribed circle diameter is about 30 m and the limited speed is about 30 mph. It will take about 7 seconds for a vehicle to do a U-turn. So the time taken on roundabouts is no more than 4 seconds averagely. Therefore, accelerating around a roundabout provides little help to shorten travel time. Secondly, a roundabout's geometry limits the opportunity to speed up since the circulatory road is a circle, not a straight line road. Finally, a vehicle normally approaches a destination exit at low speed. Eqn. 5.2 indicates that vehicles decelerate to avoid collision or they leave a roundabout at the next time step.

5.3.3 Exiting from roundabouts

Drivers clearly have their own destinations in mind, so that the destination exits of each vehicle can be assigned before entering. Characterizing a given exit for each vehicle before entry is clearly more realistic than assuming that such a decision is made once entry is affected. The approach used here is introduced by Wang (2003). In other words, the sum of probabilities to all exits is equal to 1 for each entrance. Each exit is randomly assigned an integer number. This number is equal to the number of the cells that a vehicle needs to pass in order to arrive at its destination exit.

5.4 Calibration and Validation

The fine grids are also used here. That is to say, the length of each cell corresponds to 1 m in a real road. A unit of speed is therefore equal to 3.6 km/h and each time step corresponds to 1 second.

5.4.1 Calibration

Field data was collected to observe interactions between drivers at a single-lane roundabout entrance. Traffic data has been recorded for two days, using a video camera recorder for one hour per day during the morning peak period (8:00-9:00). Normally there is quite high traffic flow during the peak hours at this roundabout.

Two-hour traffic data-sets, arrival rates, turning rates and vehicle types were extracted from the recorder. To make it easier to count available space, the video data was transferred into a Microsoft Media Maker format and divided into 121 media clips. Therefore, the traffic flow can be re-examined in a computer with playback and slow motion facilities. The observed distribution can be approximately considered as Gaussian distribution. Heterogeneous driver behaviour follows a GDH with mean = 14 and standard deviation = 2, while inconsistent driver behaviour is assumed to follow a GDI with standard deviation = 1 (mean is obtained from GDH). The number of minimum acceptable empty cells x_{min} is about 8, while the number of maximum acceptable empty cells in interaction between drivers x_{max} is about 20.

Preliminary work to calibrate the model involved checking the method itself, which included checking assumptions and rules. Vehicle motions and interactions between drivers have been visualized (see Appendix B). The first step was to directly observe the graphics interface to check whether there were any obvious errors, such as collisions and unreasonable entry. The data checking was then made in detail. Update rules were tested by recording the speed and position of each vehicle and the state of each cell in each time step.

Details of the entry process were checked by observing interactions between vehicles when vehicles entered the roundabout, according to different driving behaviour. The total number of vehicles entering the roundabout was checked and the number of vehicles on the roundabout and passing through the roundabout was also counted.

In this research, volume, delay and queue length could easily be recorded at each time step, using the proposed CA method. Thus, simulation results can be directly compared with the empirical computation results.

5.4.2 Validation

In method validation, capacity and performance measurement from a method are compared with observed values. In this effort, consistency of input parameters (turning rates, arrival rates, headway distribution, and vehicle types) and real-world observations

must be ensured. Therefore, it is important to collect relevant data at the same time. For roundabouts, relevant data includes turning rates, arrival rates, vehicle types, headway distribution, available space, and delay and queue length. For each approach and this data needs to be recorded at the same time. Unfortunately, it is difficult to get whole roundabout data at present, due to limited manpower and video facilities. Therefore, the validation of the NAS method is just based on comparison with other models (e.g. aaSIDRA, UK Linear Regression, HCM 2000, NAASRA 1986, MAP method) and field data provided on the Internet (WWW2).

The NAS method has been compared with previous models (aaSIDRA, UK Linear Regression, HCM 2000, NAASRA 1986, the MAP). The roundabout capacity models (Akcelik, 2001) mainly analyse the relationship between the entry capacity and circulating flow rate. In order to compare the NAS method with previous models, three different groups of space considerations are used, which are called NAS 1, NAS 2 and NAS 3, respectively. Table 5.1 shows assumed different distributions of driver behaviour in the NAS method. The standard deviations in GDI for all drivers are assumed to have the same value, i.e. $\sigma_I = 1$.

Figure 5.4 shows the capacity of a roundabout, obtained from the NAS method and other models. It can be seen that the proposed method agreed well with most methodologies. In particular, entry capacity in the NAS 1 is very similar to the Map method, when circulating flow is less than 900 pcu per hour. When circulating flow is larger than 1200 pcu/h, entry capacity in NAS 3 keeps agreement with the MAP method. The expected space in the MAP method is 21 m. The pcu (Passenger Car Units) is usually used to convert heavy vehicles to passenger car equivalence (HCM, 2000). Normally, car = 1 pcu, heavy vehicle = 2 pcu.

Table 5.1 Different driver behaviour distribution

	μ_H in GDH	σ_H in GDH	σ_I in GDI	x_{min}	x_{max}
NAS 1	10	2	1	6	16
NAS 2	14	2	1	8	18
NAS 3	18	2	1	12	24

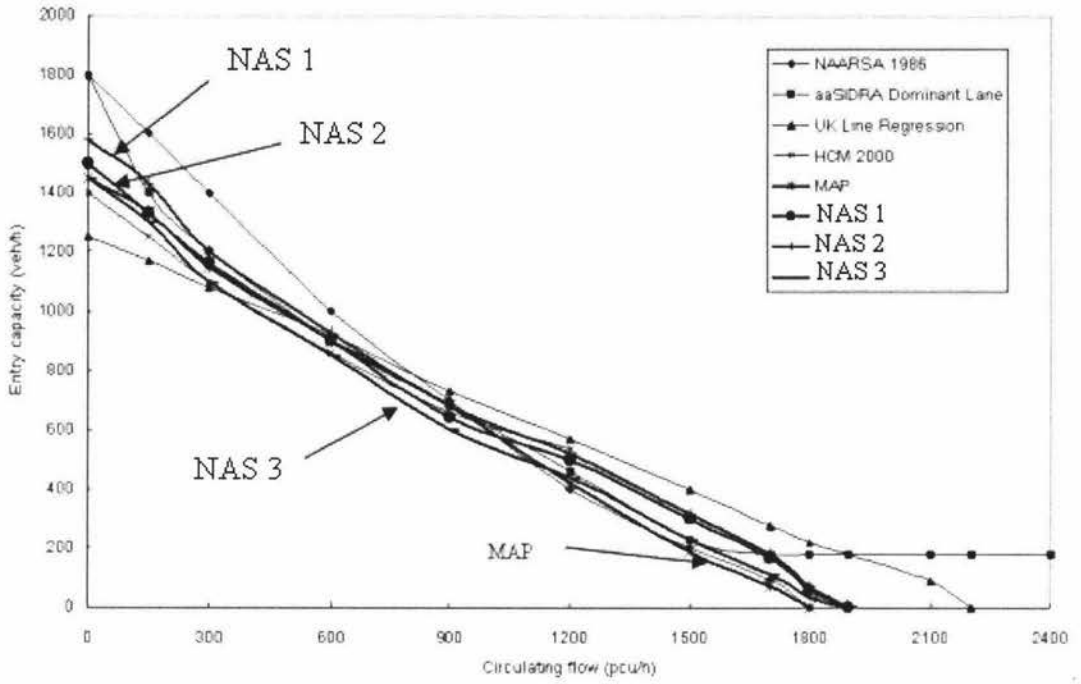


Figure 5.4 Comparison of entry capacities estimated by the NAS method and other models (the aaSIDRA, TRL (UK) Linear Regression, HCM 2000, NAASRA 1986, MAP)

The effects of a roundabout size, vehicular arrival rates, and vehicular turning rates on throughput (the number of vehicles that pass through a roundabout in a given time) have been investigated in this chapter. Our simulation results show that (i) throughput is less sensitive to roundabout size if similar topology and other parameters held constant. This is because that when the topology and other parameters keep unchanged, the inflow of vehicles from the entrances holds almost constant. Therefore, the outflow from the roundabout does not depend on roundabout size; (ii) throughput increases with arrival rate until the arrival rate reaches a critical value on one or more roads. When the arrival rate is larger than the critical value, saturation occurs on one or more roads; (iii) throughput increases (decreases) as left-turning rate (right-turning rate) increases, as vehicles on average need to travel shorter (longer) distances on the roundabout. These findings agree with the results obtained by Wang and Ruskin (2002) and Akcelik (2005). Furthermore, capacity, delay, and queue lengths are also discussed in detail. In this simulation, the NAS method is validated using more realistic input parameters and results (volume, delay and queue length) are compared with the empirical

method obtained from Eqns. D.1, D.2 and D.3 in Appendix D.

In this chapter, each cell in the simulation equals to 1 m in the real world and a unit velocity is assumed as 3.6 km/h. In this regard, each time step is 1 second. Generally!, the legal maximum velocity in urban networks is taken as 50 km/h. Normal speed on a roundabout is about 20 – 30 km/h. Thus, the number of corresponding cells ranges from 6 to 9 cells.

The NAS method is applied to a case study. Experiments were implemented for 36000 time steps (equivalent to 10 hours) for a street-length of 1000 cells on all approaches. The average values of capacity, delay and queue length are used to validate the NAS method. The distribution of driver behaviour in the NAS method is assumed to follow three different Gaussian distribution, which are listed in Table 5.1. All driver behaviour ranges within $[x_{min}, x_{max}]$, where x_{min}, x_{max} are different in NAS 1, 2 and 3. The standard deviation of each truncated Gaussian distribution is assumed to be 2 cells.

To carry out a realistic simulation, many input parameters are required, such as vehicles components, occupied cells, turning rates, arrival rates. The data source are from the webpage (WWW2) are used to verify the proposed method, where vehicles types and components are given. Table 5.2 shows the data retrieved on the Internet, including left turning (LT) rate, straight ahead (ST) rate, right turning (RT) rate, volume, capacity, delay and 95% queue length. Arms 1, 2, 3 and 4 are four roads connecting with the roundabout. Table 5.3 shows the numerical results of capacity, delay and queue length under different driver behaviour distributions. Figure 5.5 shows illustration of capacity, delay and queue length under different driver behaviour distributions. To make it clear, we assume that driver behaviour in NAS 1, 2 and 3 represents aggressive, normal and conservative, respectively.

It can be seen that aggressive drivers in NAS 1 can accept smaller required space so capacity has an increase, delay is reduced and queue length is shortened compared with other types of drivers in NAS 2 and 3. With regard to normal drivers, values of delay and queue length are slightly larger that that in NAS 1. It is found that the curves of capacity, delay and queue length in the empirical computational (EC) method lie in be-

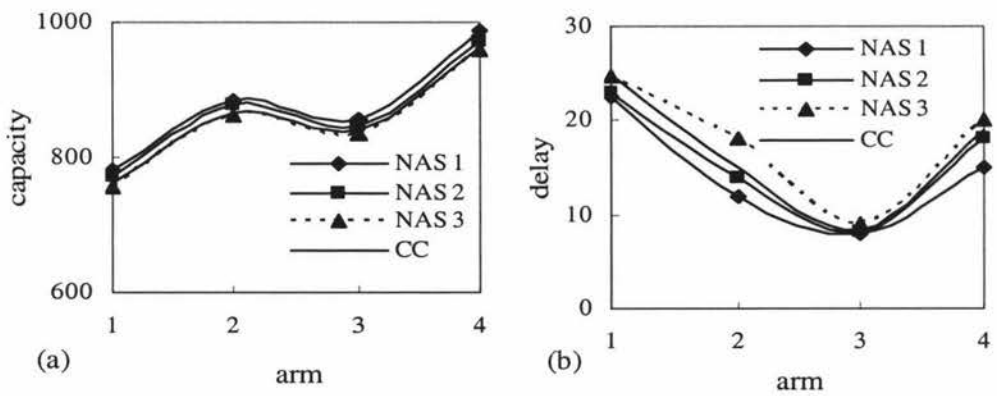
tween two curves of NAS 2 and 3. This implies that our method can simulate same results as that in the empirical computational method by just simply adjusting GDH and GDI. However, entry capacity has a few increases and delay and queue length is slightly decreased in the NAS method, compared to the empirical computational (EC) methods. The rate of volume (v) and capacity (c) is an important parameter to measure roundabout saturation and it is believed that when $v/c > 0.85$, saturation occurs (Akcelik, 2003). In figures 5.5 (a) and (b), it can be seen that there are the minimal values of v/c (0.489 for NAS 1, 0.49 for NAS 2, 0.498 for NAS 3 and 0.498 for empirical method) in Arm 3 which corresponds to the smallest delay and shortest queue length in Arm 3 for both NAS and empirical methods. This means that when arrival rate on a road is low, delay and queue length decrease correspondingly, which is consistent with the daily experience.

Table 5.2 Data retrieved on the Internet

Road	Vehicles Turning			Volume	Empirical method		
	LT	ST	RT		Capacity	Delay	95% queue length
Arm 1	118	377	150	645	762	25	10
Arm 2	88	454	100	642	865	15	6.86
Arm 3	107	258	54	419	840	8.4	2.85
Arm 4	133	586	78	797	963	18.9	9.8

Table 5.3 Capacity, delay and 95% queue length with different driver behaviour distribution. [C] = capacity; [D] = delay; [Q] = 95% queue length

	NAS 1			NAS 2			NAS 3		
Road	[C]	[D]	[Q]	[C]	[D]	[Q]	[C]	[D]	[Q]
Arm 1	782	22.4	9.2	735	23	9.4	770	24.8	9.7
Arm 2	883	12	6.5	880	14	6.63	868	18	8.2
Arm 3	856	8	2.5	848	8.2	2.74	842	9	2.81
Arm 4	988	15	8.6	971	18	9.6	966	20	10



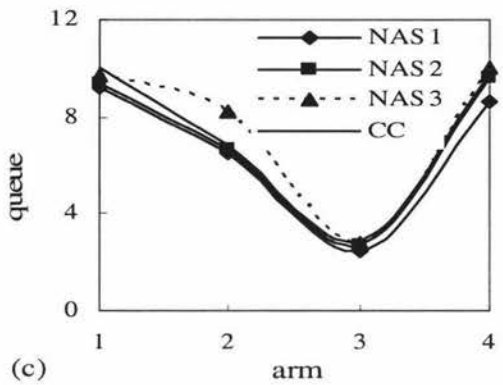


Figure 5.5 Comparison of capacity, delay and 95% queue length between NAS method and empirical computational (EC) method. Driver behavior in NAS 1, 2 and 3 see Table 5.1.

5.5 Summary

The NAS method has been developed to simulate driver behaviour and interactions between drivers at entrances of single-lane roundabouts. Modelling driver behaviour, entering rules at entrances, updating on roundabouts and exiting from roundabouts have been analysed in detail.

The relationship between entry capacity and circulating flow is obtained from the NAS method. Comparison with other methods demonstrates that the NAS method agrees well with most methodologies, particularly the MAP method.

Theoretical analysis has shown that throughput is less sensitive to roundabout size if same topology and other parameters held constant. This is because that when the topology and other parameters keep unchanged, the inflow of vehicles from the entrances is almost constant. Therefore, the outflow from the roundabout does not depend on roundabout size.

Throughput increases with arrival rate until the arrival rate reaches a critical value on one or more roads. When the arrival rate is larger than the critical value, saturation occurs on one or more roads. Moreover, throughput increases (decreases) as left-turning rate (right-turning rate) increases, as vehicles on average need to travel shorter (longer) distances on the roundabout. These findings agree with the results obtained by Wang and Ruskin (2002) and Akcelik (2005).

The NAS method has been further validated by data source from the Internet. Turning rates, arrival rates and vehicle types are directly inputted into the NAS method. Capacity, delay and queue length are computed in the simulations. Our simulation results agree with that in the empirical computational method. However, entry capacity has a few increases and delay and queue length is slightly decreased in the NAS method, compared to the empirical computational (EC) methods.

Chapter 6

Passing Lanes

6.1 Introduction

In recent decades a variety of mathematical models (e.g. car-following models [Gazis *et al.*, 1961], cellular automata models [Nagel and Schreckenberg, 1992], gas-kinetic models [Helbing et al., 2001] and hydrodynamic models [Lighthill and Whitham, 1955]) for modelling traffic flow on highways have been suggested. These models help to understand the nature of traffic flow better. Among them, CA traffic flow models have become quite popular, due to the advantages introduced in Chapter 2. However, the single-lane CA models are not sufficient to simulate realistic traffic, mainly for one reason: Two and multi lanes exist in the real world (Wang, 2003). In order to model more realistic traffic flow, multi-lane CA models (e.g. Wagner *et al.*, 1997; Knospe *et al.*, 1999 and Daoudia and Moussa, 2003) have been developed. Unfortunately, multi-lane highways are not always available in many countries such as New Zealand, due to terrain constraints or construction costs. As a result, passing lanes have been constructed as an alternative traffic facility, for the improvement of traffic volumes.

Passing lanes provide an overtaking opportunity for single-lane highway traffic. When vehicles approach passing lane sections, they are released from platoons and are driven at desired speeds on passing lane sections (Koorey and Tate, 1999). This process can be viewed as an off-ramp process and leads to an increase of the mean travel speed. When exiting passing lane sections, vehicles are merged into the main road. This process is similar to an on-ramp process. Figure 6.1 shows a schematic diagram of a single-lane section and a passing lane section. The road is divided into three sections: section A (upstream of the passing lane), section B (passing lane section), section C (downstream of the passing lane).

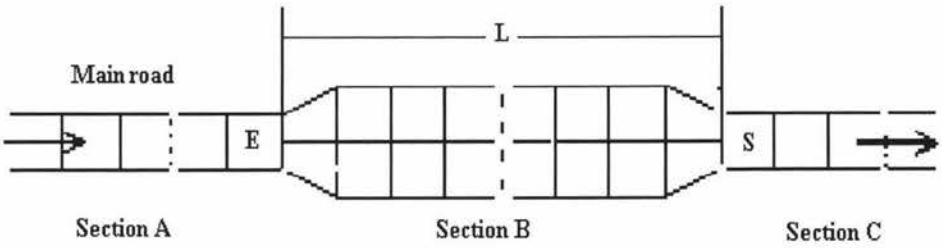


Figure 6.1 Schematic diagram of a single-lane highway with a passing lane

The effects of passing lanes depend on the length of the passing lane sections. Khan *et al.* (1991) found that longer passing lanes can reduce platoons more effectively than shorter lanes. However, Koorey and Tate (1999) indicate that a 1–2 km passing lane is more effective than a 3 km passing lane, if net benefits are considered.

AustRoads (1993) also gives recommended lengths for passing-lanes. For a 100 km/h design speed highway these are:

- Minimum: 600 m
- Recommended: 800 m
- Normal maximum: 1200 m

Although different lengths of passing lanes have been suggested by many aforementioned researchers and practitioners, there is still a potential to study of effects of passing lanes using microscopic traffic models such as CA. This is because that a gradual increase in number of vehicles makes roads busy. In this case, it is necessary to re-examine current traffic facilities to judge whether they can meet current and future transportation demand.

The NaSch model, proposed by Nagel and Schreckenberg (1992) has been widely used to simulate on- and off-ramps issues (Diedrich *et al.*, 2000). However, the study of the off- and on-ramps process, using CA models, has been little reported in the literature. Passing lane sections can be seen as such a similar process, which normally includes entering sections, driving onto sections and leaving from sections. These three processes can be analogous to off-ramps, two-lane traffic and on-ramps, respectively.

In this chapter, the modified NaSch model is used to study the effects of passing lanes on traffic flow. A fine grid CA (length of each cell corresponds to 1 m) and the 1.5-second rule are adopted for this research (details see Section 6.3.1). The main concerns are the influence of different lengths of passing lane sections on traffic flow and the effects of different lane-changing rules on driver behaviour. The effects of single-lane highways with and without passing lanes have also been discussed.

6.2 Background

In recent years, much emphasis has been put on the study of bottlenecks, sections of roads where the capacity is locally reduced. In highway traffic, on- and off-ramps are the main bottlenecks (Diedrich *et al.*, 2000). The impacts of on- and off-ramps in highway traffic have been widely investigated by many researchers (Lee *et al.*, 1999; Campari and Levi, 2000; Jiang *et al.*, 2002; Huang, 2002; Kerner, 2002; Nassab *et al.*, 2005; Jia *et al.*, 2005).

Campari and Levi (2000) simulated on-ramps on a two-lane highway without periodic boundary conditions. In their study there are no off-ramps, except for the highway's end. Also, in their model, on-ramps are regarded as an entrance cell. At each time step, the system checks whether the cell is empty or not. If the cell is unoccupied and the generated random number is less than the given threshold, a car is generated and occupies this cell with a speed at 2 cells per second. The length of each cell corresponds to 5 m in the real world and the maximum speed is assumed to be 9 cells per second (162 km/h). Therefore, an on-ramp process is simplified by authors. With regard to two-lane traffic, if the lane-changing criteria are satisfied, vehicles change to the destination lane. This lane-changing mechanism in their model may also lead to "ping-pong lane changing" (Nagatani, 1993; Knospe *et al.*, 1999). In reality, the ping-pong phenomenon hardly exists in actual traffic system. Moreover, discretionary lane changing (see Section 2.5) has not been considered in their model. Therefore, in order to model realistic lane changing, "ping-pong" effects should be avoided and random lane changing should be reproduced.

Diedrich *et al.* (2000) studied single-lane highway traffic with on- and off-ramps with periodic boundary conditions. Their focus is on the on-ramps because Diedrich *et al.* (2000) believe that off-ramps play a secondary role within traffic dynamics. Two

different types of on-ramps are introduced. One type corresponds to inconsiderate driver behaviour and another type implies both inconsiderate and cautious driver behaviour. The authors showed that there is no qualitative difference between these two types. However, both types lead to a local perturbation that divides the system into free flow and congested flow.

Stochastic noise (allowing vehicles not to change lanes even when the environmental criteria are met) is also considered by Huang (2002) when studying the effects of on-ramp and off-ramp on a multi-lane traffic flow. Pedersen and Ruhoff (2002) present a novel idea in designing on-ramps. They suggest placing *"shadow cars"* on entry ramps, which would enable drivers to take ramp cars into account.

Jiang *et al.* (2002) investigated the influence of the main road on the on-ramp to a single-lane highway. The simulation results show that main roads can impact greatly on on-ramps and cause big delay on on-ramps. Further effects of two-lane roads with on-ramps and randomization rules (Jiang *et al.*, 2003) have also been studied.

Usually there is an accelerating lane on an on-ramp, which merges with the main road. The impacts of the length of this accelerating lane on traffic flow and the impacts of traffic regulations on traffic behaviour have been investigated by Jia *et al.* (2005). Their simulation results show the introduction of an accelerating lane can improve the capacity of the on-ramp system.

Nassab *et al.* (2005) used the NaSch model and the VDR model (Section 2.4) to study the effects of spacing Δx ($\Delta x = 70$ cells, 1 cell = 7.5 m) between on- and off-ramps on a same periodic road. Two different types of spacing: $\Delta x > 0$ (off-ramps then on-ramps) and $\Delta x < 0$ (on-ramps then off-ramps) are considered. The basic phenomena of traffic flow, such as plateau formation, reduction of capacity and phase transformation from free flow to congested flow, have been simulated using the NaSch model. The effects of both types of spacing are qualitatively similar. The phenomena of large moving jams and stop-and-go waves have been reproduced using the VDR model. However, the effects of different spacing Δx and the different length of ramps on traffic flow have not been discussed in their paper. And so we can say effects of different spacing on traffic flow have not been investigated yet.

Preliminary investigations (e.g., Campari and Levi, 2000; Jiang *et al.*, 2002; Huang, 2002; Kerner, 2002; Nassab *et al.*, 2005) have shown that on-ramps have attracted more attention and they can act as a local blockage, which can then cause local capacity to reduce. The effects of off-ramps on traffic flow are overlooked. The road geometry, where vehicles from section A enter section B (Figure 6.1), can be analogous to an off-ramp. The focus in this chapter is on the effects of passing lanes on single-lane highway traffic.

6.3 Methodology

Passing lane sections can be regarded as a hybrid system, integrating off- and on-ramps and two-lane traffic. For two-lane traffic, lane changing plays an essential role. So in this subsection, first the modified NaSch model is proposed and then traffic flow at off- and on-ramps is modelled. Finally, symmetric lane-changing rules are discussed. Symmetric rules imply that both cars and trucks can use the main lane and the passing lane.

6.3.1 Model

The road is divided into cells which can be either empty or occupied by part of a vehicle with speed $v = 0, 1, \ldots, v_{max}$. The vehicles move from left to right with periodic boundary conditions. Let $x_n(t)$ and $v_n(t)$ denote the position and velocity of vehicle n at time step t; $d_n(t)$ be the gap (the number of free cells in front) of vehicle n at time step t. Vehicle n-1 precedes to vehicle n. The updated rules of the model can be described as follows:

S1: Acceleration

$$v_n\,(t+1) \rightarrow \min\{v_{\max},\, v_n(t)+1\} \tag{6.1}$$

S2: Deceleration (avoiding collision)

$$v_n\,(t+1) \rightarrow d_n(t) * 2 / 3 \tag{6.2}$$

S3: Randomization

$$v_n(t+1) \rightarrow \max\,\{0,\, v_n(t)-1\} \text{ with probability } p_b, \tag{6.3}$$

S4: Vehicle movement

$$x_n(t+1) \rightarrow x_n(t) + v_n(t+1) \tag{6.4}$$

Rules S1, S3 and S4 are the same as the NaSch model. Rule S2 is the main modification to the NaSch model. Rule S2 is based on the 1.5-second rule. The headway of 1.5 second for car-following processes has been observed in urban roads as well as on highways. In other words, a vehicle can only drive up to 2/3 of the total gap between the vehicle and the preceding one. In order to count gaps accurately, the fine grid CA is used, i.e. the length of each cell corresponds to 1 m in reality, compared to 7.5 m in the NaSch model. Each time, step is 1 second. A unit of velocity is therefore equal to 3.6 km/h, compared to 27 km/h in the NaSch model. Hence, 1 unit of acceleration rate corresponds to 1 m/s^2, which is a 'comfortable acceleration' (Traffic Engineering Handbook, 1992).

6.3.2 Off- and on-ramps

The following notations are used in this subsection:

L length of the passing lane section

B_l main lane (left lane) in the passing lane section

B_r passing lane (right lane) in the passing lane

S the first cell of section C

E the last cell of section A

A_{first} the leading vehicle on section A

B_{lfirst} the leading vehicle on lane B_l

B_{rfirst} the leading vehicle on lane B_r

B_{llast} the last vehicle on lane B_l

B_{rlast} the last vehicle on lane B_r

C_{last} the last vehicle on section C

X_{Bllast} position of the last vehicle on lane B_l

X_{Brlast} position of the last vehicle on lane B_r

Figure 6.1 shows a sketch of a single-lane highway with a passing lane. To clearly describe the problem, the road upstream of the passing lane section, the passing lane section and the road downstream of the passing lane section are denoted as sections A, B and C, respectively. The two branches of section B are denoted as B_l and B_r, respectively.

Fine-grid CA models have been used to simulate realistic traffic flow, while two

actual considerations are still needed in the following simulations: (1) vehicles enter section B from section A, which corresponds to an off-ramp. Drivers have to decide which lane to choose. (2) vehicles enter section C from section B, which corresponds to an on-ramp.

For case 1, the driver needs to decide which lane in which to drive. When the car A_{first} follows a car, which is a farther distance from it, i.e., the empty cells in front of A_{first} are assumed to be $d = \max\{d_{lAfirst}, d_{rAfirst}\}$, where $d_{lAfirst} = X_{Bllast} - X_{Afirst} - 1$ and $d_{rAfirst} = X_{Brlast} - X_{Afirst} - 1$, the relationship of position of car A_{first} and its two leaders in section B can be described as follows.

- If $X'_{Afirst} > X_E$ and $d_{lAfirst} > d_{rAfirst}$, then car A_{first} enters B_l (main lane);
- If $X'_{Afirst} > X_E$ and $d_{lAfirst} < d_{rAfirst}$, then car A_{first} enters B_r (passing lane);
- If $X'_{Afirst} > X_E$ and $d_{lAfirst} = d_{rAfirst}$, then car A_{first} enters B_l if the driver does not want to overtake or enters B_r if the driver wants to overtake. In this case, it is assumed that a driver has the same probability to determine overtaking or not. So if the random number in [0, 1] is less than 0.5, the driver is expected to drive on the left lane with no overtaking. Otherwise, the vehicle is driving on the right (passing) lane. Here X_E is the position of cell E, X'_{Afirst} and is the new position of the car A_{first}.

For case 2, it is similar as an on-ramp process. Cars B_{lfirst} and B_{rfirst} are checked first at cell S at each time step. If only one of them can arrive at S, then this car enters section C. If both of them pass at cell S in one time step, the time t_{Bl} and t_{Br} needed to arrive at cell S for cars B_{lfirst} and B_{rfirst} is calculated without randomization in Expression 6.2 and 6.3 (Jiang *et al.*, 2002):

$$t_{Bl} = \frac{x_s - x_{Bllead}}{\min(v_{\max}; x_{Clast} - x_{Bllead} - 1; v_{Bllead} + 1)} \tag{6.2}$$

$$t_{Br} = \frac{x_s - x_{Brlead}}{\min(v_{\max}; x_{Clast} - x_{Brlead} - 1; v_{Brlead} + 1)} \tag{6.3}$$

The following time relationship decides which car can enter section C first.

- If $t_{Bl} \neq t_{Br}$, then the car which needs less time to arrive at cell S can first enter section C.
- If $t_{Bl} = t_{Br}$, then the car which is nearer to cell S has a priority to enter section C.
- If $t_{Bl} = t_{Br}$ and the distances to S are the same for both B_{lfirst} and B_{rfirst}, according to

left-side driving regulations, the vehicle on the right lane can enter section *C* first.

6.3.3 Lane changing behaviour

The following notations are used in this subsection:

$d_{n,\,left}$ number of empty cells between n^{th} vehicle and its front neighbour car on the main (left) lane at time t

$d_{n,right}$ number of empty cells between n^{th} vehicle and its front neighbour car on the passing (right) lane at time t

$d_{n,\,back}$ number of empty cells between n^{th} and its back neighbours on the destination lane

v_n velocity of n^{th} vehicle at time t

v_{ob} velocity of the following vehicle on the destination lane at time t

p_1 probability of changing lane from the main lane to the passing lane

p_2 probability of changing lane from the passing lane to the main lane

Generally speaking, lane changing is considered feasible if there are sufficient gaps ahead and behind in the target lane, so that the vehicle can safely change to the target lane without causing other vehicles behind, in the target lane, to slow down significantly. However, lane changing is not required. Intrinsically, lane changing is a stochastic process, so even when all conditions have been satisfied, some drivers still do not change lanes (Wang, 2003).

When a car is approaching cell *S*, the driver must decide to change lanes or not. This decision-making process, on a single lane, has been analysed in case 1 of Section 6.3.1. Here a car on the two-lane traffic is discussed. According to these two regulations, two sets of lane-changing rules are given as follows.

For rule 1, from main lane to passing lane, if conditions

$$d_{n,\,left} < v_n,\ d_{n,\,right} > v_n,\ d_{n,\,back} > v_{ob} \text{ and random } () < p_1 \tag{6.4}$$

are met, the n^{th} car directly changes to the passing lane from the main road. Otherwise, it keeps on the main lane and speed is reduced to $d_{n,\,left} * 2/3$ (1.5-second rule, see Section 3.2.2). Condition $d_{n,\,right} > v_n$ means the driver can move onto the passing lane and overtake; and condition rand $() < p_1$ means the driver would like to overtake.

For rule 2, right lane to left lane, if conditions

$$d_{n,\,left} > v_n,\ d_{n,\,back} > v_{ob} \text{ and random}() < p_2 \tag{6.5}$$

are fulfilled the n^{th} car can change to the main lane. Otherwise, it keeps in the passing lane. The condition $d_{n,\,right} < v_n$ is not a necessary condition because most cars will change to the main lane after they have finished overtaking. Condition $d_{n,left} > v_n$ means the driver can move onto the main lane and give way to overtaking cars; and condition random () < p_2 means the driver would like to keep in the main lane. The above-mentioned lane-changing rules are similar to Wagner *et al.* (1997), however, the 1.5-second rule is considered in current lane-changing rules.

6.4 Simulations Results

In this chapter, an inhomogeneous vehicle system is considered. Two different types of vehicles are adopted here, the fast (such as cars) and the slow (such as trucks). v^f_{max} and v^s_{max} corresponds to the maximum velocity unit of cars and trucks, respectively. R_f and R_s denotes the ratio of cars and trucks to all vehicles. In the simulations, the values of v^f_{max} and v^s_{max} are assumed to be 30 cells (108 km/h) and 25 cells (90 km/h) in normal conditions, respectively.

In the simulations, the length of each cell corresponds to 1 m compared to 7.5 m in the NaSch model. The time step is 1 second. R_s / R_f = 0.1 / 0.9 is adopted, which has been used in two-lane traffic (Nagel *et al.*, 1998; Helbing and Schreckenberg, 1999; Daoudia and Moussa, 2003). The length of normal vehicles, such as cars, is set to 5 cells and heavy vehicles, such as buses or trucks, are set to 10 cells. The first 10,000 time steps are discarded to let the system reach a stationary state. The flow and speeds are obtained by counting the vehicles that pass a virtual detector at cell *S* in the next 10,000 time steps. In order to reduce the influence of scattering of fast and slow vehicles on traffic flow, each simulation has been run for 50 times and the averages of values (e.g., velocity, flux) in 50 trials are used in this chapter.

Each lane of the highway can be divided into cells which are empty or occupied by vehicles. Vehicles move in the direction of increasing cell number. Section *A* is divided into 3,000 cells and section *C* into 2,000 cells. Although the length of section *B*

is recommended from 600 to 1,200 m for a 100 km/h design speed highway (Austroads 1993), in this chapter the length of the passing lane is extended to 2,000 m, so that there is a better understanding about how different passing section lengths influence traffic behaviour.

In this research, the effects of two aspects are investigated. One is how different lengths of passing lanes affect flow, density and speeds. The other is how lane-changing probabilities affect flow, density and speeds on the passing lane. In this case, two typical values of $p_1 = 0.5$ and $p_2 = 0.5$ are discussed.

6.4.1 Effects of length of passing lanes

At first, a fundamental graph of flow versus density is plotted showing the proposed model and the NaSch model, both without passing lanes (Figure 6.2). In this chapter, the maximum speed of vehicles in the NaSch model is taken as 108 km/h since the legal speed is limited to 100 km/h in New Zealand. It can be seen in Figure 6.2 that there is an increase of the maximum flow in the homogeneous system, compared to the NaSch model. This can be explained as follows: In rule S2 (Expression 6.2), if the following vehicle can keep the same speed as the leading vehicle, rear-ending can be avoided. In this case, platoon formation (Evans, 1997) is observable and jam flow in the NaSch model is transformed into synchronized flow. In addition, since overtaking is not allowed in a single-lane system, the slow vehicles clearly dominate the average flow at low densities. This implies that lower speed leads to lower flow under the same density.

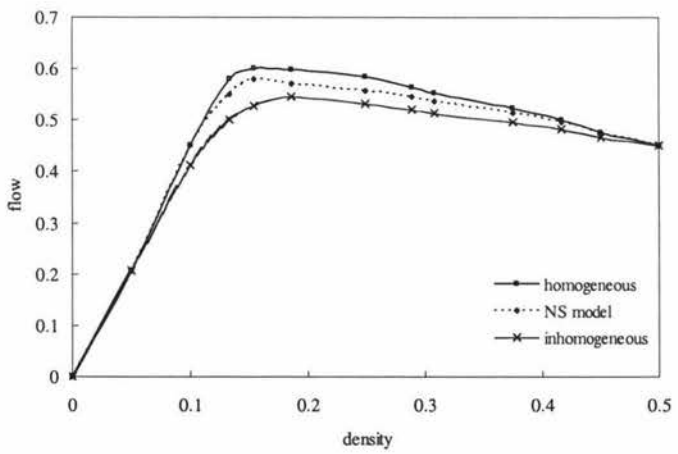


Figure 6.2 Fundamental diagram of single-lane highway traffic with $p = 0.5$

With regard to homogeneous traffic, the maximum flow in our model and NaSch model in Fig. 6.2 is about 2160 and 1952 vehicle per hour (vph) respectively at density $\rho^{flow}{}_{max} \approx 1.6$. While for heterogeneous traffic, the maximum flow in Fig. 6.2 is about 1872 vph at density $\rho^{flow}{}_{max} \approx 2.0$. It can be seen that the maximum flow decreases when vehicle composition changes to heterogeneous from homogeneous flow. On the other hand, the density in which the maximum flow is reached moves forward when vehicle composition changes to heterogeneous from homogeneous flow. When density $\rho > \rho^{flow}{}_{max}$, traffic flow is in congested states. The congested flow can be divided into synchronized flow and wide moving jams according to the three-phase traffic theory proposed by Kerner (1998). The synchronized flow is characterized by high flux without any clear flow-density relation (Kerner, 2002; Lee *et al.*, 2004). Thus, different explanations to the formation of synchronized flow are proposed such as density profiles, road bottlenecks (Kerner, 1998), vehicle composition (Treiber and Helbing, 1999), changes of driver behaviour (Krauβ, 1998) and anticipation effects of several vehicles ahead (Schreckenberg and Wolf, 1998; Lenz *et al.*, 1999). With the increase of density, the regions of synchronized flow decrease while the queuing regions increase. And when the density is larger than 0.5, flows in homogeneous and heterogeneous traffic are almost equal. This implies that the effects of slow vehicles are not clearer in high density than in low density.

The simulation results here are for two-lane symmetric lane-changing behaviour. That is to say, fast as well as slow vehicles may use both lanes. The length of the passing lane is assumed to be 1,000 cells, corresponding to 1,000 m of real space. In figure 6.3, the effects of the passing lanes on traffic flow is compared with the fundamental diagram of a single-lane road. It can be seen that the introduction of passing lanes can improve the average flow, both in homogeneous and inhomogeneous systems. The flow in a homogeneous system with passing lanes increases by about 15% and 8%, compared to a homogeneous system without passing lanes and an inhomogeneous system with passing lanes, respectively.

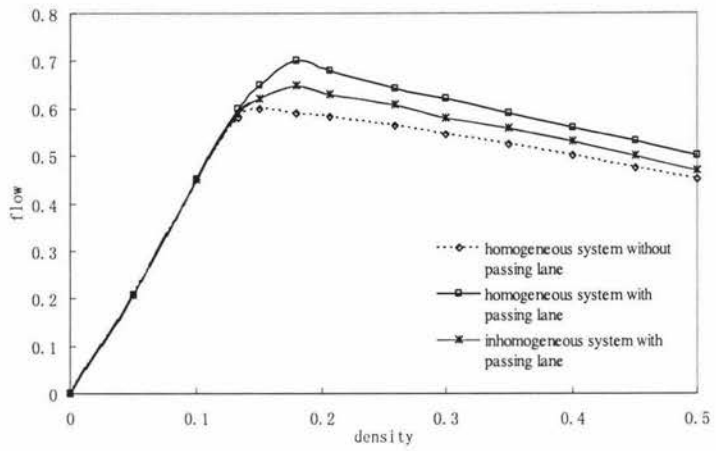


Figure 6.3 Effects of passing lanes on homogeneous and inhomogeneous systems with $p = 0.5$

The relationship between flow and the length of the passing lane is shown in Figure 6.4. The length of the passing section ranges from 600 cells to 2,000 cells. Flow and speed at cell S have been recorded from the time step 40,000 to 43,600 (one hour). This process is repeated 20 times and the results are obtained by average. It can be seen that flow slightly increases and then decreases; then considerably increases and reaches a maximum at about 1,500 cells. With the increase of the length of the passing lane, more drivers would prefer to drive in the passing section, leading to an increase of traffic volume. According to the simulation, when the length of the passing lane falls in 1,300 ~ 1,700 cells, traffic volume basically approaches saturation point. Therefore, a passing lane which is too long is not clear to improving road throughput, but will instead just increase the construction costs.

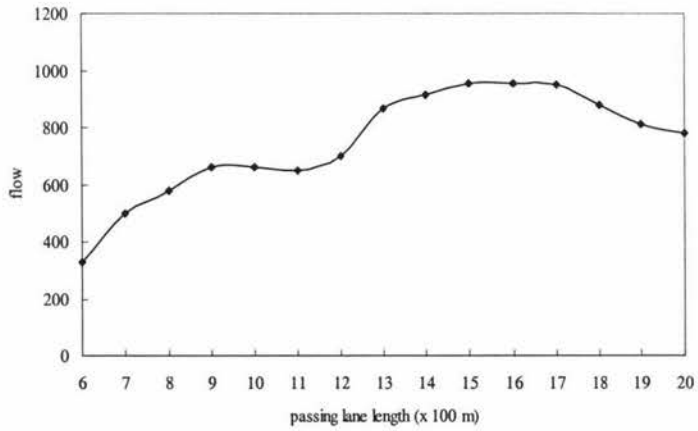


Figure 6.4 Flow variations with different lengths of passing lane sections

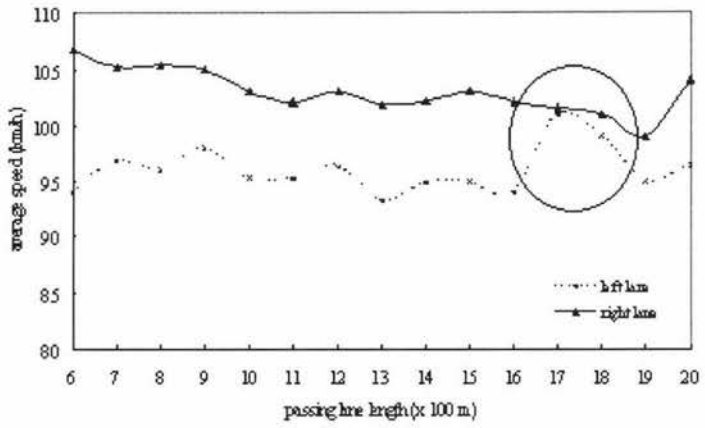


Figure 6.5 Average speed variations with different length of passing lane section

Figure 6.5 shows the relationship between average speed and different lengths of passing lanes. One can see that average speed fluctuates moderately through the whole length on the right lane. On the left lane, average speed fluctuates stronger than on the right lane due to the existence of different types of vehicles. The circle part in Figure 6.5 shows abrupt speed increase and decrease. This can be explained as follows. At the near end of the passing lane, many cars have finished overtaking and are driving at speed of about 100 km/h on the left lane, which causes an increase of average speed on the left lane. And when approaching the merging section, these cars have to slow down to give way to cars on the right lane according to the left-hand traffic rules in New Zealand.

6.4.2 Effects of lane changing

Lane-changing rules have been discussed in Section 6.3.3. Here the effects of lane changing are analysed. The probability of lane changing $p_1 = p_2 = 0.5$ are arbitrarily chosen. At each time step, the number of vehicles in a main lane and passing lane can be counted through a virtual detector. Thus, lane usage is obtained. It can be seen that lane usage basically reaches a near-optimal state with a range from 0.4 to 0.6 in Figure 6.6. This can be explained by the same probability of lane changing imposing little on lane usage. The initial lane choice (discussed in case 1 of Section 6.3.2) plays an important role in lane usage, under the same probability of lane changing.

In this research the probability of lane changing p_1 and p_2 are just assumed to be 0.5 due to no much field data. That is to say, even criteria of lane changing are satisfied,

only a half of drivers would like to change lanes. In reality, the probability of lane changing p_1 and p_2 are various. Clearly, different probabilities can lead to different speed, density, flow and land usage in main and passing lanes. Further study will focus on this issue.

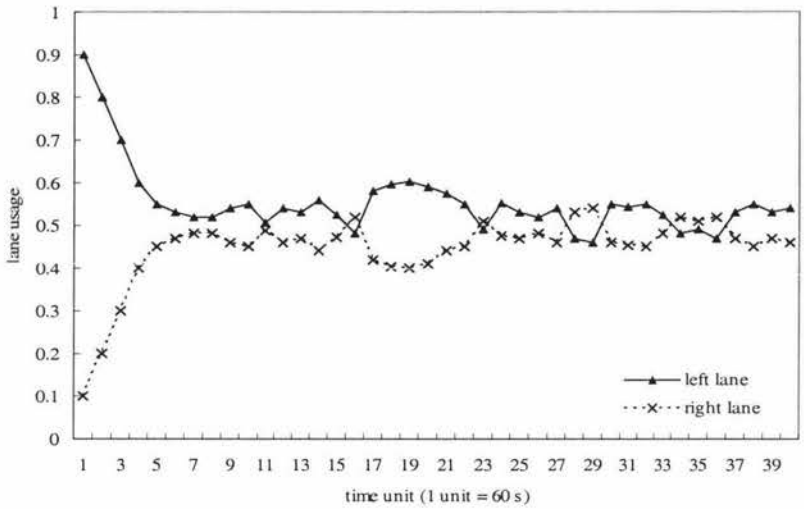


Figure 6.6 Lane usage variations over time

6.5 Summary

In this chapter, the effects of passing lanes on highway traffic have been investigated. Two realistic features have been taken into account. Firstly, the 1.5-second rule is used for the headway (=distance /velocity) in modelling car-following processes; (ii) a road is divided into fine grid CA, that is, the length of each cell corresponds to 1 m, which is the shortest cell length used in CA models for the best of our knowledge.

Vehicles entering and exiting from passing lane sections (analogous to off-ramps and on-ramps processes) have been analysed. An inhomogeneous (car-truck) system and symmetric lane-changing behaviour are applied in this chapter.

The fundamental diagram in Figure 6.2 indicates that trucks can significantly influence flow, on roads without passing lanes. When a passing lane is introduced, the flow increases by almost 15%. The flux reaches the maximum value when the length of passing lane is 1,500 m in the simulations and the flux and the speed becomes slightly fluctuated while the length is over 1,300 m. Lane usage is analysed under the same

probability of lane changing $p_1 = p_2 = 0.5$. Two-lane traffic can basically reach a near-optimal state with a range of lane usage from 0.4 to 0.6.

In future work, the effects of passing lane sections will further be investigated from the point of view of statistical physics, that is, phase transitions among free flow, synchronized flow and wide moving jams will be studied in a passing lane section.

Chapter 7

Short-term Traffic Flow Forecasting

7.1 Introduction

The demand for predictive information on traffic conditions is becoming more and more essential for Intelligent Transportation Systems (ITS), such as route guidance systems, signal control and Variable Message Signs (VMS) (Sun et al., 2003). This type of advanced information system can effectively redistribute traffic resources in real time (Lingras and Sharma, 1999). As such, an accurate and fast prediction of short-term traffic conditions has become a main concern to researchers.

The prediction of traffic conditions prediction commonly focuses on three macroscopic traffic variables: flow, speed and occupancy (Vlahogianni *et al.*, 2004). However, flow prediction dominates this field (Vlahogianni *et al.*, 2004). Short-term Traffic Flow (STF) forecasting provides short-term flow estimation in the next few time steps, usually in the range from five minutes to an hour according to different application requirements.

Clearly, prediction accuracy is strongly related to the choice of length of prediction time (also called horizon), length of interval (also called step) and different algorithms. The forecasting horizon denotes the extent of time into the future. The forecasting step represents the frequency of predictions in the forecasting horizon (Vlahogianni *et al.*, 2004). For example, an algorithm predicts traffic flow 15 minutes ahead in 5-minute intervals. The forecasting horizon is 15 minutes and step is 5 minutes. It should be noted that both larger horizon and shorter step can lead to inaccurate forecast. Ishak and Al-Deek (2002) concluded that the prediction accuracy degrades as the forecasting horizon increases. They also indicate that decrease of forecasting accuracy has also contributed to the strong variability of traffic flow when observed in short time intervals.

The forecasting horizon and step are typically different hen modelling highways and urban networks. As Kirby *et al.* (1997) noted, traffic in highways can travel great distances in a specific period, for example 30 minutes, compared with traffic in urban areas. In this case, the forecasting horizon and step can take larger values than in urban networks. However, with regard to real-time adaptive traffic control systems in urban networks, the predictive interval should be significantly decreased. Head (1995), for example, suggested that forecasts in 30 or 40 seconds into the future are able to reschedule the signal time logic. Ledoux (1997) predicted traffic flow at 1-minute intervals in a signalised intersection.

Apart from traffic flow forecasting, occupancy (proportion of total length of vehicles on a road in total length of the road) and speed forecasting have also been discussed by many researchers. Dougherty and Cobbett (1997) developed three different methods for predicting traffic flow, occupancy and speed in highways. They found that their methods were satisfactory for predicting traffic flow and occupancy, but not satisfactory for speed. The failure in predicting speed, as the authors argued, was due to the low speed vehicles in low flow conditions. Innamaa (2000) used both flow and speed to predict flow. The findings, when forecasting both variables with one method, gave worse results than two separate methods for flow and speed. So we can say that different data resources should select different methods. Lin *et al.* (2002) suggested that occupancy is a better indicator of traffic conditions. Their studies imply that an ANN approach may not be versatile when different data source such as flow, speed and occupancy are used as input. Different data characters should choose suitable algorithms.

This chapter concentrates on short-term traffic flow forecasting in urban networks, which attempts to provide accurate flow prediction for traffic signal control. Signal timing cycle lengths usually range from 45 to 120 seconds (New York City Department of Transportation, 2002). Traffic signal timing is based on traffic volume and traffic patterns such as speed of vehicles, distance between traffic signals, and number of non-signalised intersections along the roadway system (Housatonic Valley Council of Elected Officials, 2005). Table 7.1 shows cycle lengths under different traffic conditions. In Table 7.1 one can see that flow prediction of 60-second interval into the future is appropriate for traffic signal control. If the time interval is larger than cycle

length, the system cannot obtain real-time input. Too short time interval may lead to computational complexity, even disorder signal schedule. Thus, traffic flow at 1-minute intervals is predicted in this research.

Table 7.1 Timer cycle lengths under different traffic conditions (Housatonic Valley Council of Elected Officials, 2005)

Distance between traffic signals	Low volumes	High volumes
0.25 mile (≈ 402 meters)	70 seconds or less	90 – 120 seconds
0.5 mile (≈ 804 meters)	60 – 120 seconds	

This research concerns flow prediction in urban roads. Flow prediction is a very important parameter in traffic control. In this chapter, a new hybrid algorithm GAANN (Genetic Algorithm-base Artificial Neural Networks) is proposed to predict traffic flow. In order to improve prediction accuracy and generalization ability of Artificial Neural Networks (ANN), the global searching strategies of Genetic Algorithms (GA) are used to optimize ANN architecture (e.g., number of neurons in the hidden layer) as well as network parameters (e.g., connection weights). In fact, when we apply ANN and GA to numerical prediction, we do not only set up optimal ANN architecture and appropriate network parameters but also seek optimal combination between ANN architecture and network parameters. Thus, the difference between the GAANN and other methods (e.g., Ferreira *et al.*, 2004; Yang *et al.*, 2004; Kim *et al.*, 2004; Niska *et al.*, 2004; Fish *et al.*, 2004 and Tian and Noore, 2005) is that other methods optimize either ANN architecture or network parameters, not both. More detailed review on other methods, see section 7.2.3. Our results of using GAANN demonstrate that the proposed approach is suitable for traffic flow prediction.

7.2 Background

The study on STF (Short-term Traffic Flow) forecasting has attracted extensive attention from various disciplines, both inside and outside of traffic and transportation areas. A variety of methodologies and techniques have been adopted to study STF forecasting in the recent decades, such as the Auto-Regressive Integrated Moving Average (ARIMA) (Smith and Demetsky, 1997; William, 1999), Kalman filter (Okutani and Stephanides, 1984), regression analysis (Hobeika and Chang, 1994), chaotic theory

(Hu *et al.*, 2003), Markov chain method (Yu *et al.*, 2003), ANN (Dougherty and Cobbett, 1997).

Most of the above methodologies (e.g., ARIMA, Kalman filter, Markov chain method) are based on statistical methods and concentrate on centralizing metrics such as mean mode and normally ignore extreme values; and thus, cannot accurately approximate an abrupt change in traffic flow prediction (Vlahogianni *et al.*, 2004). In practice, these extreme values often severely affect traffic conditions. They can cause large-scale flow fluctuations, even congestions. The chapter focus on using a GAANN method, particularly a multi-layered feedforward ANN, to forecast traffic flow in urban networks.

7.2.1 Artificial neural networks

The use of ANN is expected to be valuable in modelling optimization problems due to the ANN capacity to approximate a given numeral at any desired accuracy (Hornik, 1989) and in solving regression problems (White, 1990). It has been proven that a 3-layered ANN has an ability to approximate any given non-linear function (Hornik, 1991). When a ANN is to be applied, the network structure (number of neurons in the input, hidden and output layers) and learning parameters (e.g., learning rate, momentum factor) should be determined. However, there is no strict mathematical theory to guide people how to select the network structure and how to tune learning parameters (Gallant and Aitken, 2003). Basically, they are usually determined empirically.

With regard to a three-layered feedforward ANN, the selection of the network topology normally depends on the selection of the number of input neurons, output neurons and hidden neurons (Haykin, 1999). The selection of input and output neurons are often based on experience and practical application. Thus, the selection of the hidden neurons becomes a crucial factor in evaluating performance of ANN.

If the number of neurons in the hidden layer is too small, it is difficult for the ANN to converge at a desired error. On the other hand, if the number is too large, the ANN learning process will be very long (Srinivasan, 1994). There are no clear

mathematical formulations to determine the optimum number of hidden neurons which can obtain an optimal solution in the literature as the number of hidden neurons is data-dependent in different applications (Rui and Ei-Keib, 1996).

The learning parameters in ANN mainly include the learning rate and momentum factor. The learning rate controls learning speed and effectiveness the network converges to a near-optimal or optimal solution. A larger learning rate will lead to a more rapid convergence for smooth functions, while a smaller learning rate is more suitable for steep and narrow curves (Haykin, 1999). So, the optimal value of the learning rate is not fixed, but also data-dependent. Indeed, many interesting things can be done by making the different learning rates for different stages of the network. Values for the learning rate may range from 0 to 1 (Haykin, 1999).

Except for the learning rate, another variable used in weight adjustment is the momentum factor, which can prevent the learning process from terminating in a narrow local minimum on the error surface to a certain degree (Haykin, 1999), that is, the momentum factor can improve the learning performance in some situations, by helping to smooth out unusual conditions in the training set. However, generally the optimal value of the momentum factor is also data-dependent and perhaps large than 1, which depends on conditions.

7.2.2 Genetic algorithms

More recently, utilizing GA (Genetic Algorithms) to optimize ANN has become a popular approach to solve optimization and classification issues. The brief review sees Subsection 7.2.3. Before this, a brief introduction of GA is recalled. GA have proven to be powerful techniques in solving linear and non-linear problems due to their ability to globally search all state space (maximum or minimum points) through genetic operations such as crossover, mutation and reproduction (Yao, 1999).

Genetic algorithms are derived from evolutionary theory (Holland, 1975). They are based on the genetic process of biological organisms. Over many generations, natural populations evolve according to the principles of natural selection, i.e. *survival of the fittest*, first clearly stated by Charles Darwin (1909) in *The Origin of Species*. By

following this process, genetic algorithms are able to map solutions to real problems, if chromosomes have suitable codes.

Before a genetic algorithm can be carried out, a suitable encoding (or representation) for the problem must be devised. A fitness function is also required, which assigns a value of merit to each encoded solution. During the run, parents must be selected for reproduction, and recombined to generate offspring (see Figure 7.1).

It is assumed that a potential solution to a problem may be represented as a set of parameters. These parameters (known as genes) are joined together to form a string of values (chromosomes). In genetic terminology, the set of parameters represented by a particular chromosome is referred to as an individual. The fitness of an individual depends on its chromosome and is evaluated by the fitness function (Holland, 1975). The individuals, during the reproductive phase, are selected from the population and recombined, producing offspring, which comprise the next generation. Parents are randomly selected from the population using a scheme, which favors fitter individuals. Having selected two parents, their chromosomes are recombined, typically using mechanisms of crossover and mutation (Holland, 1975). Mutation is usually applied to some chromosomes to guarantee population diversity.

The main deficiencies of GA optimizing ANN are a high computational cost and considerable search space (Miller *et al.*, 1989). Many researchers such as Back *et al.* (1997) and Eiben *et al.* (1999) have suggested some efficient computation methods to address this problem. In addition, with the rapid increase in CPU clock speeds, the high computational requirement has not caused application obstacles. This means that more attention have been paid on efficiency of algorithms rather than complexion.

7.2.3 Genetic algorithms optimizing artificial neural networks

The hybrid methods of GA optimizing ANN are not novel. Different evolutionary strategies have been used in the various fields (e.g., Ferreira et al., 2004; Kim *et al.*, 2004; Niska *et al.*, 2004; Fish *et al.*, 2004; Tian and Noore, 2005). Ferreira *et al.* (2004) applied GA to optimize the number of time lags, the number of hidden neurons and the training algorithm in a time-lagged network structure. An approach of

using GA to optimize wavelet networks has been developed and used in traffic flow forecasting (Yang *et al.*, 2004). Kim *et al.* (2004) adopted GA to determine network parameters such as the number of the hidden neurons, the momentum factor and the learning rate. GAs have also been used for selecting the inputs and designing the high-level architecture of a Multi-Layered Perceptron (MLP) method for air pollution time series prediction in Helsinki (Niska *et al.*, 2004). Fish *et al.* (2004) indicated that a simpler ANN architecture led to better generalization in the case of the multilevel choice and then used a GA-based neural network to model individual consumer choices. Tian and Noore (2005) used GA to optimize the number of the delayed input neurons and the number of the hidden neurons. This optimized network was developed to model software cumulative failure time prediction.

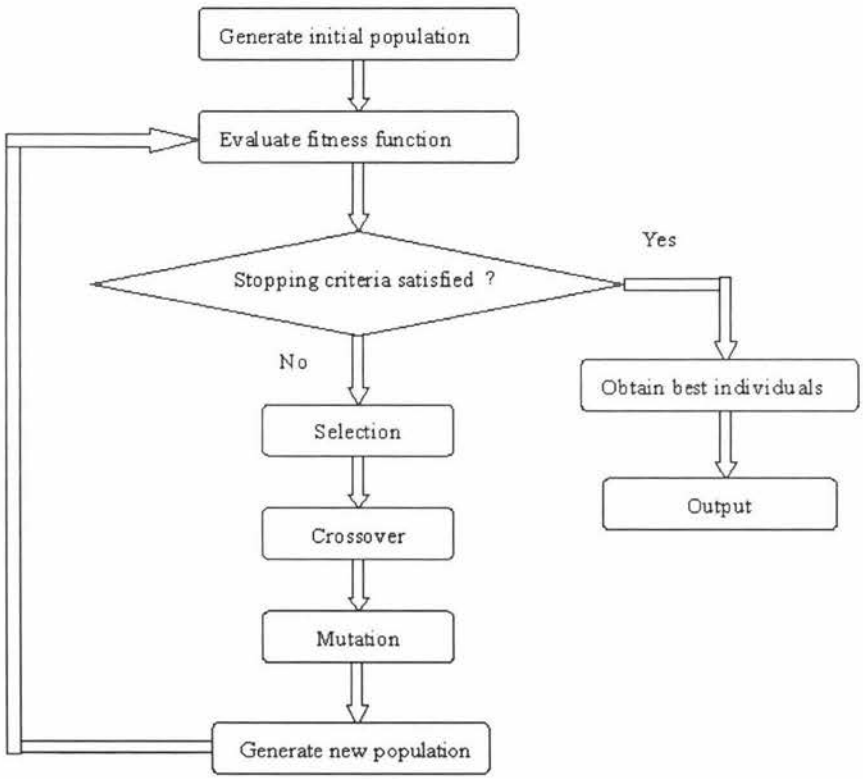


Figure 7.1 Illustration of standard genetic algorithms

Evolving ANN architecture and network parameters using GA has been extensively studied by a great number of researchers. However, these methods optimize either ANN architecture or network parameters only. The method of evolving ANN architecture (e.g., number of neurons in the hidden layer) as well as network parameters (e.g., connection weights) using GA has not been reported in the literature. To

effectively apply ANN and GA to numerical prediction, it is not sufficient to just set up an optimal ANN architecture or appropriate network parameters. Seeking optimal combination between ANN architecture and network parameters is necessary. As connection weights are of paramount importance in that they determine the accuracy and generalization ability of the methods, thus, in this chapter GA is employed to optimize connection weights and the number of neurons in the hidden layer and apply GAANN to predict short-term traffic flow.

7.3 Methodology

7.3.1 Chromosome representation and encoding

Number of neurons in the hidden layer and connection weights need to be encoded, which are related to ANN architecture as different ANN architecture may have different number of neurons in the hidden layer. Figure 7.2 is a three-layered neural network with m neurons in the input layer, h neurons in the hidden layer, and n neurons in the output layer.

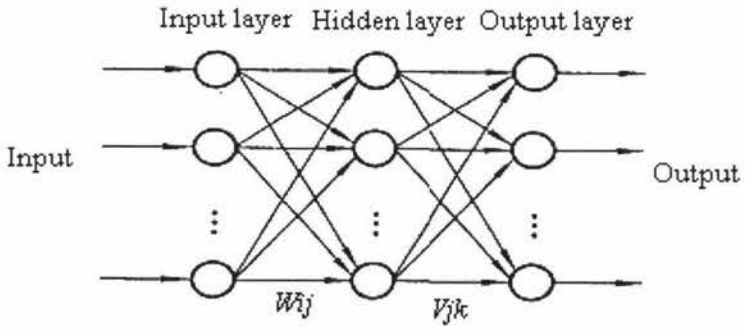


Figure 7.2 Schematic diagram of a three-layered neural network with m neurons in the input layer, h neurons in the hidden layer, and n neurons in the output layer

This denotes:

w_{ij}: connection weight from ith neuron in the input layer to jth neuron in the hidden layer. $i = 1, 2, \ldots, m; j = 1, 2, \ldots, h$.

v_{jk}: connection weight from jth neuron in the hidden layer to kth neuron in the output layer. $j = 1, 2, \ldots, h; k = 1, 2, \ldots, n$.

γ_j: connection degree of jth neuron in the hidden layer.

θ_j: threshold of jth neuron in the hidden layer. $j = 1, 2, \ldots, h$.

β_k: threshold of kth neuron in the output layer. $k = 1, 2, \ldots, n$.

The number of neurons in the hidden layer is dynamically determined according to genetic operations (reproduction, crossover and mutation). Accordingly, the weight matrix also changes dynamically. If the neuron in the hidden layer has connection with input and output layers, it is called a connected neuron and denote $\gamma_j = 1$; otherwise, $\gamma_j = 0$. Thus, a binary value (0 or 1) is used to encode the number of neurons in the hidden layer. As connection weights are real values, so real values are directly used to encode connection weights.

The whole encoding string of each chromosome L consist of two parts: connection code G and weight code H, namely, $L = G + H$, where G is the length of connection code of the number of hidden neurons and H is the length of real-valued code of connection weights.

For the purpose of simplicity, it is assumed that there is only one neuron in the output layer. Thus, the network architecture of $m * h * 1$ expounds chromosome representation and encoding.

The length of G is the number of neurons in the hidden layer. The length of H can be calculated as follows, $H = m * h + h + h * 1 + 1$ (m is the number of input neurons, h is the number of hidden neurons, output neuron is taken to be 1).

The whole string corresponds to a chromosome, which consists of some gene sections. According to network structure illustrated in Figure 7.2, gene sections can be tabulated as follows:

String G: $\gamma_1 \gamma_2 \ldots \gamma_h$

Weights matrix from input layer to hidden layer:

$$\begin{bmatrix} w_{11} & w_{12} & w_{13} & \cdots & \cdots & w_{1h} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ w_{m1} & w_{m2} & w_{m3} & \cdots & \cdots & w_{mh} \end{bmatrix}$$

Weights matrix can be written as the following gene string: $w_{11}\ w_{12}\ w_{13} \ldots w_{1h}$ $w_{21}\ w_{22} \ldots w_{2h} \ldots \ldots w_{m1}\ w_{m2} \ldots w_{mh}$

Weights from hidden layer to output layer: $v_1\ v_2 \ldots v_h$

The overall encoding string can be given by: $\gamma_1\ \gamma_2 \ldots \gamma_h\ w_{11}\ w_{12}\ w_{13} \ldots w_{1h}\ w_{21}\ w_{22} \ldots w_{2h} \ldots \ldots w_{m1}\ w_{m2} \ldots w_{mh}\ v_1\ v_2 \ldots v_h$

7.3.2 Optimization method

Generally, partition of the sample set is related to the proposed approach and it often includes a training set and a test set. In this research, cross-validation (Haykin, 1999) is adopted to further partition the training set into two disjoint subsets: estimation subset and validation subset. Estimation subset is used to select the method. In this chapter estimation subset is used to determine the basic state space of connection weights. Validation subset is used to test or validate the method. As many researchers (e.g. Haykin, 1999) have suggested, cross-validation can avoid over-fitting to a certain extent. Thus the sample set is divided into three parts: estimation subset φ_1, validation subset φ_2 and test subset φ_3.

Figure 7.3 shows the flow chart of a GA-based ANN (GAANN) prediction algorithm. The GAANN algorithm is introduced as follows:

Step 1: In order to determine a basic state space of connection weights matrix, the traditional BP (back-propagation) algorithm is first implemented in a 3-layered feedforward neural network. Connection weights are initialized within [-1, 1] for estimation subset φ_1 and validation subset φ_2 and adjusted until the desired tolerance of errors ε_1 and ε_2 is satisfied. The maximum and minimum of weights are denoted as u_{max} and u_{min}, respectively. The value of weights are taken within $[u_{min} - \delta_1,\ u_{max} + \delta_2]$, where δ_1, δ_2 are adjustment parameters.

$$\min\ E_i = \frac{1}{2}\sum^{\phi_i} [\, y_k(t) - \hat{y}_k(t)]^2 < \varepsilon_i \tag{7.1}$$

where i = 1, 2, 3 which corresponds to three data subsets. $\hat{y}_k(t)$, $y_k(t)$ are the desired output and real data, respectively.

Step 2: Encode gene string as described in Section 7.3.1.

Step 3: Generate randomly a population of chromosomes.

Step 4: Test if an individual is satisfied with $E_1 < \varepsilon_1$, $E_2 < \varepsilon_2$, and $E_3 < \varepsilon_3$. If yes, go to Step 10, otherwise go to next step.

Step 5: Calculate fitness individually according to Equation 7.2 below.

$$F = 1/(1 + \min E) \tag{7.2}$$

Step 6: Copy the highest fitness individual directly to a new offspring and select other individuals by means of the method of spinning the roulette wheel (Messai *et al.*, 2002).

Step 7: Use basic crossover and mutation operations to the connection code, namely, if a hidden neuron is deleted (added) according to mutation operation, the corresponding connection code is encoded 0 (1). The crossover and mutation operators of weights are encoded as follows:

- Crossover operation with probability p_c

$$X_i^{t+1} = c_i \cdot X_i^t + (1 - c_i) \cdot X_{i+1}{}^t \tag{7.3}$$

$$X_{i+1}{}^{t+1} = (1 - c_i) \cdot X_i^t + c_i \cdot X_{i+1}{}^t$$

where X_i^t, $X_{i+1}{}^t$ are a pair of individuals before crossover, X_i^{t+1}, $X_{i+1}{}^{t+1}$ are a pair of individuals after crossover; and c_i is taken as random value within [0, 1].

- Mutation operation with probability p_m

$$X_t^{i+1} = X_t^i + c_i \tag{7.4}$$

where X_i^t is individual before mutation, X_i^{t+1} is individual after mutation; and c_i is taken as random value within ($u_{\min}$ - δ_1 - X_i^t , $u_{\max}$ + δ_2+ X_i^t).

Step 8: Generate the new population and replace the current population.

Step 9: The above procedures (Step 4 ~ 8) are repeated until convergence conditions ($E_1 < \varepsilon_1$, $E_2 < \varepsilon_2$ and min $E_3 < \varepsilon_3$) are satisfied.

Step 10: Decode the highest fitness individual; obtain a corresponding number of the hidden neurons and connection weights.

Step 11: Input training and test subsets into the network. If the current errors (E_1, E_2 or E_3) can further decrease, connection weights are updated. Otherwise, keep the current network architecture and connection weights and output prediction results.

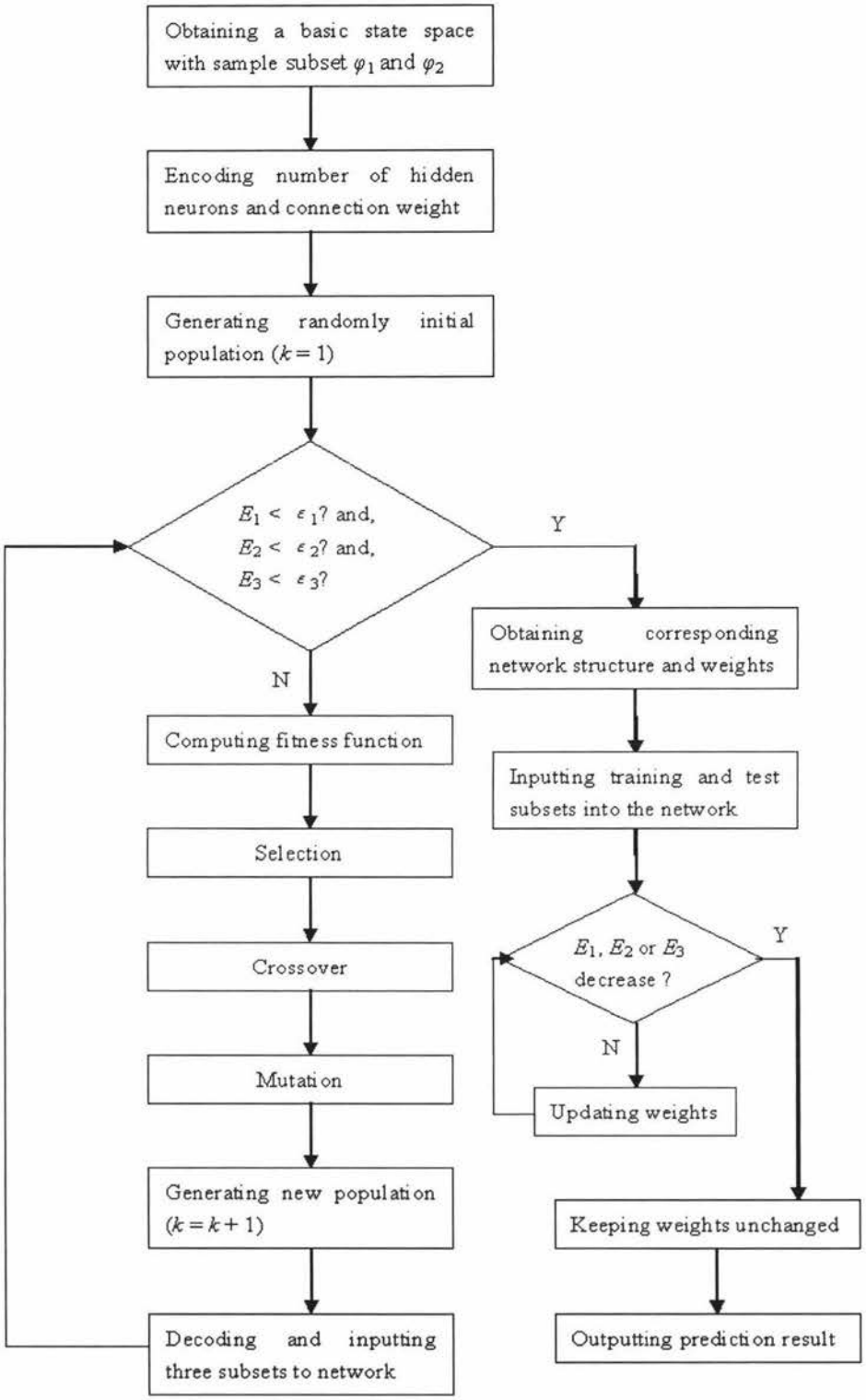


Figure 7.3 Flow chart of GA-based ANN prediction algorithm

7.4 Implementation and Results

7.4.1 Data collection and pre-processing

In order to evaluate the performance of the proposed method, field observation was performed to count headways (time between passage of the front ends of two successive vehicles) and traffic flow in a double-lane urban road. The research recorded 5 days traffic flow data between 16 August 2004 and 20 August 2004. The data set was sampled by 1-minute intervals in the morning peak period (7:50 – 8:50) each day. Each subset thus has 60 observations and 5 data sets are named as *Dataset* 1, *Dataset* 2, *Dataset* 3, *Dataset* 4 and *Dataset* 5, respectively.

As mentioned in Section 7.3.2, each data set is divided into three parts: estimation subset φ_1, validation subset φ_2 and test subset φ_3. The percentage of each subset is different from different papers. For instance, φ_1: φ_2: φ_3 = 0.6:0.1:0.3 (Dia, 2001); φ_1: φ_2: φ_3 = 0.78:0.125:0.095 (Xu and Wong, 2003). In this research, φ_1: φ_2: φ_3 = 0.6:0.2:0.2 is arbitrary set. Table 7.2 shows the classification of observations in the three data sets. Different partition of datasets how to influence prediction and generalization ability has been little reported in the literature. Probably, it is also data-dependent. These issues are expected to clarify in the near future works.

Table 7.2 Data sets partition used for method development

Data set	Percentage of data set (%)	Number of observations
Estimation subset	60	36
Validation subset	20	12
Test subset	20	12

Before the neural network starts learning, the data must be pre-processed and scaled. It has been widely acknowledged that data quality is a key issue with data analysis. Statistical results show that 80% of data mining efforts are spent on data quality (Goharian, 2003).

In order to increase the accuracy of the method, pre-processing should be implemented in advance. There are many methods to pre-process data (Goharian, 2003), such as data cleaning (cleaning incomplete, noisy, or inconsistent data), data integration (integrating different data sources), data normalization (scaling data value in a range,

such as [0, 1]), data reduction (reducing huge dataset to smaller representation). This research adopted normalization to pre-process traffic flow data.

When using data normalization, the maximum value of traffic flow in the sample sets should be clarified. Although the maximum observation value can be found, the value is not ensured to be the maximum in theory. According to field observation, the minimum headway is about 1 second, that is to say, the maximum traffic flow is 60 vehicles in a sample unit (1 minute) in a single lane. So the maximum theoretical traffic flow is 120 vehicles in two lanes. The normalization equation is given as follows:

$$y = x / x_{\max} \tag{7.5}$$

where y is the scaled value fed into the network, x is the actual value before scaling and x_{max} is the maximum value theoretically. Here $x_{max} = 120$.

After the stopping criteria are satisfied with three sample subsets (estimation, validation and test subsets), the network output is scaled back to its actual value via the following equation:

$$x = y * x_{\max} \tag{7.6}$$

7.4.2 Implementation and evaluation

Traffic flow can be normally viewed as a time series, namely, the problem of prediction can be formulated as the estimation $d(l + t)$, where t is a small time interval, given a set of historical data series say, $d(1)$, $d(2)$, ..., $d(l)$ and where $d(i)$ represents the value at the i^{th} time step, $1 \leq i \leq l$. In STF forecasting, the value of t is recognized as less than 30 minutes (Lingras and Sharma, 1999), even equals to 20 seconds into the future for predicting vehicle velocity (Dia, 2001). Since a shorter time interval and acceptable prediction accuracy are very necessary for traffic information systems, in this chapter the time interval t is set as 1 minute. This can validate the performance of the method more microscopically.

It was proposed to optimize the number of neurons in the hidden layer and connection weights using genetic algorithms. However the optimal number of neurons in

the input layer was not known. So the number of neurons in the input layer as 3, 4, 5, 6 and 7 was taken to evaluate the performance of the proposed method.

A three-layered BP neural network is first employed to estimate basic state space of connection weights. For instance, when the number of neurons in the input layer is taken as 4, the minimum and maximum values of weights are obtained, -1.21 and 0.96, respectively. Let δ_1 = -0.09 and δ_2 = 0.04, therefore, the range of weights is assumed to be within [-1.3, 1.0]. The activation function adopted here from input to hidden layer is Sigmoid, while from hidden to output layer is Purelin (y = x) function. For the proposed hybrid neural network, the following system parameters in Table 7.3 are applied to training sample and prediction.

Table 7.3 Required method parameters

Name of Variables	Value	Name of Variables	Value
Learning rate	0.9	Crossover Probability p_c	0.80
Momentum	0.7	Mutation Probability p_m	0.05
MSE of subsets φ_1 and φ_2	0.003	MSE of subset φ_3	0.004
Epochs	1000	Populations	30

Neural network ensemble is a learning paradigm where a number of neural networks are trained for the same task (Sullish and Krogh, 1996), which shows that the generalization ability of a neural network system can be significantly improved through training many neural networks and then combining their results (Zhou and Chen, 2002). This kind of technique has been widely used in different fields such as face recognition (Mao, 1998) and medical diagnosis (Li *et al.*, 2005) and so on. In general, a neural network ensemble is constructed in two steps, i.e., training a number of neural networks and then combining the component predictions. The most popular approaches for generating individuals of a neural network are Boosting (Schapire, 1990) and Bagging (Breiman, 1996). As for classification tasks, plurality voting or majority voting (Hansan and Soloman, 1990) are normally adopted. As for regression problems, simple averaging (Penene and Cooper, 1993), limited weighted averaging (Merz and M. J. Pazzani, 1999) or selective ensemble (Zhou *et al.*, 2002) are normally used. In this chapter, 50 times computation is carried out for each dataset (totally 5 datasets). The simple averaging is used as the ensemble strategy.

Table 7.4 below shows comparison of errors for different neurons in the input layer in this study. The error is represented by MSE (mean squared error). The MSEs of three subsets and average MSE are obtained from different methods (GAANN and MLP). These results demonstrate the better predictive performance of GAANN when compared to the MLP.

When the number of neurons in the input layer equals to 4, the predictive accuracy in Table 7.4 is higher than that of other numbers. Thus, the same network structure is used to train other four sample sets, *Dataset* 2, *Dataset* 3, *Dataset* 4, and *Dataset* 5. In the network structure, the number of neurons in the input layer is 4 and the corresponding number of neurons in the hidden layer is 8. Table 7.5 lists the results for five sample sets. These results clearly show that the GAANN method is capable of forecasting traffic flow for 1 minute into the future. The best average MSE (0.0015) is obtained for the sample set *Dataset* 3.

Figure 7.4 shows the learning results and the actual value in sample set *Dataset* 1. It demonstrates that the GAANN method has superior approximation ability. Figure 7.5 shows the prediction of the next 60 minutes from the GAANN and MLP. These results show that the GAANN method has a better prediction performance when compared to the actual data and the MLP. However, in peak values the GAANN method is not satisfied. It suggests that the GAANN method should be further improved the prediction accuracy. Moreover, the prediction horizon should be extended to 2, 5, or 15 minutes in order to further validate prediction capability in the proposed method. Currently, as no much field data is available, the effects of different prediction horizons has not made in this research.

7.5 Summary

In this chapter a Genetic Algorithm-based Neural Networks (GAANN) approach is proposed for Short-term Traffic Flow (STF) forecasting in urban networks. Prediction of STF is indispensable for traffic control. In order to improve prediction accuracy and generalisation ability of Neural Network (ANN), network structure is optimized and connection weights are adjusted through the implementation of genetic operations (selection, crossover and mutation). The advantage of the GAAN is that it

does not only optimize the ANN architecture and network parameters but also develops an optimal combination between them which is different from other previous methods presented in the literature.

Data normalization is adopted for improving data quality. The different numbers of neurons in the input layer and different sample sets are used to test GAANN method. The results show that the predictive performance of the proposed method is better than that of the traditional BP neural network. Furthermore, the 4 x 8 x 1 architecture shows better prediction accuracy and generalization ability compared to other network architectures.

However, the GAANN method does not behave perfectly in peak values. The consideration of further improvements in the method performance should include the following factors, such as, different prediction horizon (2, 5, or 15 minutes), crossover, mutation operators and classification of data sets.

Table 7.4 Errors for different neurons in the input layer

	Number of input neurons	Number of Hidden neurons	MSE of esti-mation subset	MSE of valid-ation subset	MSE of test subset	Average MSE of three subset
GAANN	3	5	0.00167	0.0015	0.0023	0.00182
MLP	3	5	0.0018	0.0025	0.0033	0.00353
GAANN	4	8	0.0019	0.0012	0.002	0.0017
MLP	4	8	0.0016	0.0026	0.0027	0.0023
GAANN	5	9	0.0019	0.0017	0.0023	0.002
MLP	5	9	0.0021	0.0032	0.0027	0.00267
GAANN	6	11	0.002	0.0018	0.002	0.0019
MLP	6	11	0.0016	0.0021	0.003	0.0022
GAANN	7	12	0.0017	0.0018	0.0034	0.0023
MLP	7	12	0.0022	0.0025	0.0036	0.0031

Table 7.5 Errors for five sample sets

Sample set	Network type	MSE of estimation subset	MSE of validation subset	MSE of test subset	Average MSE of three subset
Dataset 1	GAANN	0.0019	0.0012	0.002	0.0017
	MLP	0.0016	0.0026	0.0027	0.0023
Dataset 2	GAANN	0.0018	0.0012	0.0023	0.00176
	MLP	0.0022	0.002	0.0032	0.0025
Dataset 3	GAANN	0.0015	0.0012	0.0019	0.0015
	MLP	0.0014	0.0025	0.0023	0.00206
Dataset 4	GAANN	0.0017	0.0013	0.0022	0.0017
	MLP	0.0020	0.0019	0.0031	0.0023
Dataset 5	GAANN	0.0021	0.0026	0.0033	0.00267
	MLP	0.0026	0.0025	0.003	0.0027

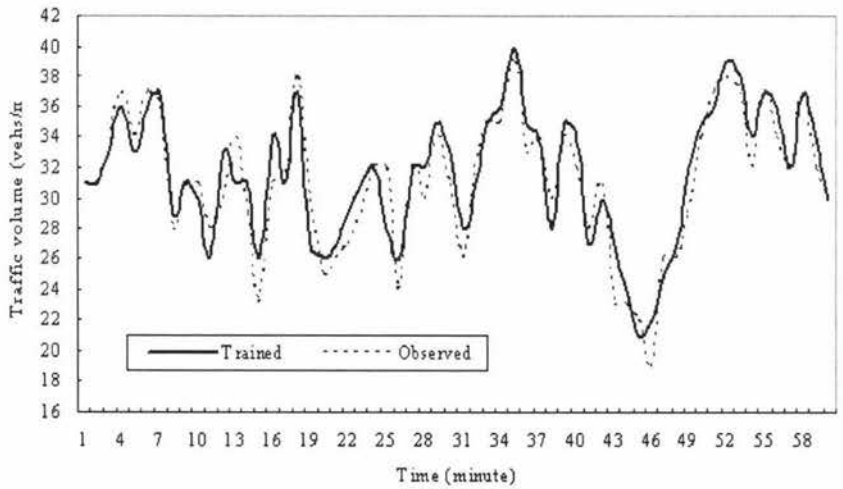


Figure 7.4 Observed traffic flow and GAANN trained results

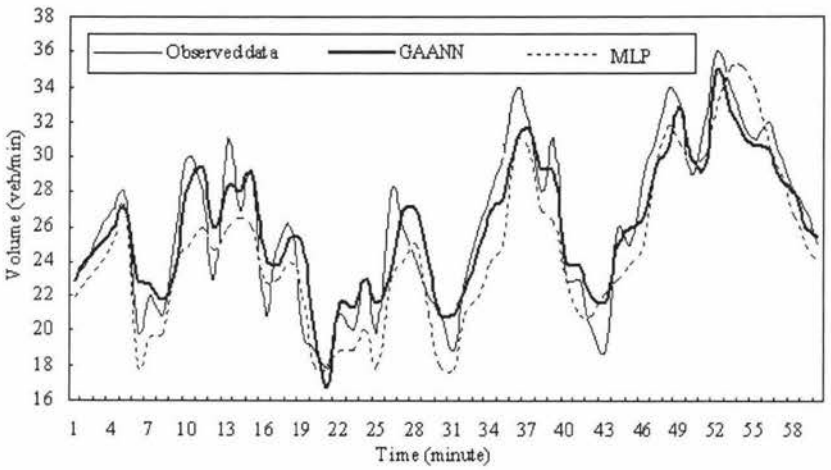


Figure 7.5 Traffic flow of observations, GAANN and MLP

Chapter 8

Summary and Future Work

8.1 Study Summary

The main intention of this thesis has been to give an insight into traffic flow modelling using Cellular Automata (CA) to gain an understanding of the effects of driver behaviour on urban unsignalised intersections and roundabouts. New methodologies have been developed to provide enhanced modelling features towards realistic traffic simulation. The effects of passing lanes for highway traffic flow have been investigated. Short-term Traffic Flow (STF) forecasting in urban networks using Neural Networks (NN) has been explored.

In Chapter 1 an introduction to this thesis is given. Chapter 2 critically reviews CA traffic models for highways and urban networks. Two kinds of traffic flow: free flow and car-following flow were modelled on urban straight roads in Chapter 3. For free flow, five stages (dual-regime acceleration, one steady state and dual-regime deceleration) are developed. Based on the five-stage speed changes, a novel and realistic CA model is proposed to simulate free flow. For car-following flow, 1.5-second rule is used to describe the car-following processes, which is based on the field observation. Simulation results indicate that the dual-regimes of acceleration and deceleration make CA models keep in good agreement with the real world.

Two types of road features have been modelled in Chapters 4 and 5, single-lane Two-Way Stop-Controlled (TWSC) intersections and single-lane roundabouts. A TWSC intersection is controlled by priority and stop/give-way rules. A hierarchy of streams (priority-based) at TWSC intersections for left-side driving has been highlighted in Section 4.2. Interaction rules at roundabouts are known as yield-at-entry (by which a vehicle entering gives way to one already on the roundabout).

With regard to modelling driver behaviour at single-lane TWSC intersection and roundabout entrances, a novel and realistic CA model is developed, which is based on the

Normal Acceptable Space (NAS) method. As for each driver, driver behaviour has her/his own special space criteria. If these are met, the vehicle can then move onto intersections or roundabouts. If not, the vehicle has to wait. Space requirement of a driver at an entrance is updated at each time step. Therefore, the NAS method can dynamically describe space considerations required in detail. There are three advantages using the NAS method to simulate heterogeneous and inconsistent driver behaviour and interactions between drivers for different traffic conditions and for a variety of urban road features as follows.

Firstly, the NAS method simulates traffic flow more realistically and precisely comparing with previous gap- acceptance models and MAP method. For gap-acceptance models, the more recent paper (Pollatschek *et al.*, 2002) just considered driver behaviour as two groups: cautious drivers and risk-loving drivers. And, no consideration has been given to model heterogeneous and inconsistent driver behaviour. The MAP (Wang, 2003) method has been used to simulate heterogeneous and inconsistent driver behaviour in urban networks, however, driver behaviour has only been categorised into four groups. The NAS method uses an analogue of the MAP but with two Gaussian distributions to describe heterogeneous and inconsistent driver behaviour rather than four arbitrary groups. This assumption is theoretically based on statistical distributions and practically based on fine-grid CA and field observations. .

Secondly, the NAS method is generic and can integrate different road geometries into a unified system. It can provide a unified framework to simulate traffic flow at intersections and roundabouts, and to simulate traffic flow in the large-scale urban networks as well as to facilitate designing simulation software.

Finally, the NAS method observes realistic traffic rules at unsignalised intersections and roundabouts. That is to say, the stream ranking at TWSC intersections (Section 4.2) and yield-at-entry at roundabouts (Section 5.1) can be closely integrated into the NAS method.

The capacity and performance (delay and queue lengths) of intersections and roundabouts have been studied. The proposed method has been calibrated and validated using field data and data source from the Internet. Simulation results show that the NAS method provides a satisfactory description of the empirical findings.

In Chapter 6 the effects of passing lanes have been investigated using the modified Nagel and Schreckenberg (1992)'s model. Passing lane sections can be regarded as hybrid systems with off- and on-ramps and two-lane traffic. In off-ramps (entering passing lane sections), vehicles' road choices have been discussed according to position of the preceding vehicle on the passing lane. In on-ramps (emerging into the main lane), the decision-making process that vehicle has a priority to first exit from passing lane section is analyzed. The effects of different length of passing lane on flux, density and speed are studied. The flux reaches the maximum value when the length of passing lane is 1,500 m in the simulations. In fact, the flux and speed fluctuate slightly while the length is over 1,300 m. It is suggested that a passing lane which is too long is not conducive to improving road capacity, but will instead just increase the construction costs.

The aforementioned CA models adopt fine grid CA (the length of each cell in CA corresponds to 1 m in reality) which provides a better resolution than other previous CA models (e.g. NaSch model, VDR model, BML model, CDL model and MAP method). Since 1 unit of acceleration in CA corresponds to 1 m/s^2, this equals to a "comfortable acceleration" suggested in Traffic Engineering Handbook (1992). Therefore, the proposed CA models can simulate traffic flow more realistically. The calibration and validation of the proposed CA models against the field data reveal a good agreement with the empirics on description of both macroscopic and microscopic levels.

In Chapter 7, Short-term Traffic Flow (STF) forecasting is investigated on an arterial road in urban networks. Prediction of STF is expected to provide real-time traffic information for traffic signal control, Advanced Traveller Information Systems (ATIS) and vehicular route choice. A Genetic Algorithms-based Neural Networks (GANN) method is proposed. The numbers of neurons in the hidden layer and connection weights are optimized by genetic operations (selection, crossover and mutation). Different number of neurons in the input layer and different sample sets are used to test the proposed method. Numeric results show the GANN method can effectively improve prediction accuracy and generalization ability of neural networks.

The major contributions in this research can be summarized as follows:

- Fine-grid CA models (the length of each cell corresponds to 1 m on a real road) have been adopted in this research. The fine-grid models are able to

simulate vehicle dynamics in a more realistic way.

- The five-stage speed-changing CA model is first proposed to describe free flow in urban networks. Moreover, 1.5-second rule based on observations is used to describe car-following processes instead of 1-second rule used in previous CA models.
- Two Gaussian distributions are used for the first time to simulate heterogeneous driver behaviour and inconsistent driver behaviour at entrances of TWSC intersections and roundabouts.
- Effects of passing lanes on traffic flow on a single-lane highway are investigated using a realistic CA model, which has not been reported in the literature.
- GANN method is first proposed to improve prediction accuracy and generalization ability of NN and used to forecast STF.

8.2 Future Work

Several areas identified for further research are summarized below:

- There is a clear need to gather more real data to validate current models and give these a basis and reinforcement for further development. We have recorded some videos of traffic flow. Clearly, collaboration with the relevant local government bodies would be useful in this regard.
- The modelling methodologies developed in this research should be further extended to simulate traffic flow at two-lane TWSC intersections and roundabouts.
- The preliminary findings of this research can be applied in modelling traffic flow at signalised intersections and traffic signal control.
- Using genetic algorithms to optimize neural network architecture should consider how to avoid eliminated individuals reappearing in the next generation. This can address more efficient crossover and mutation operators. In addition, how to obtain an optimal partition of data sets is also an interesting issue.

References

Adams, W.F. (1936) Road traffic considered as a random series. Institution of Civil Engineers, 9, London.

Ahmed, A., Khanal, R. and Rust, R. (2002) Evaluation of Freeway Diversion Route Plans Using Microscopic Simulation and Real-Time Data: The Case of Idaho's Treasure Valley Corridor. Transportation Research Board Annual Meeting in Washington D.C.

Ahn, S., Cassidy, M.J. and Laval, J. (2004) Verification of a simplified car-following theory. *Transportation Research Part B: Methodological*, Vol. **38**, Issue 5, pp. 431-440.

Akcelik, R. and Besley, M. (2001) Acceleration and deceleration models. The 23rd Conference of Australia Institutes of transport Research (CAITR 2001), Monash University, Melbourne, Australia, 10-12 December.

Akçelik, R. (2003) A Roundabout Case Study Comparing Capacity Estimates from Alternative Analytical Models. The 2nd Urban Street Symposium, California, 28-30 July.

Akcelik, R. (2005) Roundabout Model Calibration Issues and a Case Study. The TRB National Roundabout Conference, Vail, Colorado, USA, 22-25 May.

AustRoads. (1993) *Rural Road Design* Sydney.

Aycin, M.F., Benekohal, R.F. (1998) Linear Acceleration Car-Following Model Development and Validation. *Transportation Research Record*, Vol. **1644**, TRB, National Research Council, Washington, DC, pp. 10–19.

Bäck, T., Fogel, D. and Michalewicz, Z. (1997) *Handbook of Evolutionary Computation*, Institute of Physics Publishing Ltd., Bristol and Oxford University Press, New York.

Bando, M., Hasebe, k., Nakayama, A., Shibata, A. and Sugiyama, Y. (1995) Dynamical model of traffic congestion and numerical simulation. *Physical Review E,* Vol. **51**, pp. 1035-1042.

Barbosa, H.M., Tight, M.R. and May, A.D. (2000) A model of speed profiles for traffic calmed roads. *Transportation Research Part A*, Vol. **34**, pp. 103-123.

Barlovic, R., Santen, L., Schadschneider, A. and Schreckenberg, M. (1998) Metastable states in cellular automata for traffic flow, *The European Physical Journal B*, Vol. **5**, 793.

Barlovic, R., Brockfeld, E., Schreckenberg, M., Schadschneider, A. (2001) Optimal traffic states in a cellular automaton model for city traffic. *Traffic and Granular Flow*, 15 – 17 October, Nagoya University, Japan.

Barlovic, R., Huisinga, T., Schadschneider, A. and Schreckenberg, M. (2004) Adaptive traffic light control in the ChSch Model. *Traffic and Granular Flow'*03, (Springer 2004).

Barr, D., Sherrill, E. (1999) Mean and Variance of Truncated Normal Distributions. *The American Statistician,* Vol. **53**, pp. 357-361.

Benekohal, R. F. (1989) Procedure for Validation of Microscopic Traffic Flow Simulation Models. *Transportation Research Record* **1320**, TRB. National Research Council, Washington, D.C., pp. 190-202.

Benjaafar, S. and Dooley, K.J. (1997) Cellular Automata for Traffic Flow Modelling. Univ. of Minnesota, Minneapolis.

Benyoussef, A., Boccara, N. (2001) Traffic Flow in a 1D Cellular Automaton Model with Open Boundaries, *Chinese Journal of Physics,* Vol. **39**, No. 5.

Berg, P. (2001) Optimal-velocity models of motorway traffic. Doctoral Thesis, University of Bristol, UK.

Bhattacharyya, G.K., Johnson, R.A. (1977) Statistical Concepts and Methods. John Wiley & Sons, Inc. Printed in the United States of America.

Bham, G.H. and Benekohal, R.F. (2004) A high fidelity traffic simulation model based on cellular automata and car-following concepts. *Transportation Research Part C: Emerging Technologies*, Vol. **12**, Issue 1, pp. 1-32.

Biham, O., Middleton, A.A. and Levine, D. (1992) *Physical Review A*, Vol. 46, R6124.

Bonsall, P., Liu, R. and Young, W. (2005) Modelling safety-related driving behaviour - impact of parameter values. *Transportation Research Part A: Policy and Practice,* Vol. **39**, Issue 5, pp. 425-444.

Brackstone, M. and McDonald, M. (1999) Car-following: a historical review. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. **2**, Issue 4, pp. 181-196.

Breckfeld, E., Barlovic, R., Schadschneider, A. And Schreckenberg, M. (2001) Optimizing traffic lights in a cellular automaton model for city traffic. *Physical Review E*, Vol. **64**, Issue 5, p.056132.

Bunker, J. and Troutbeck, R.J. (2003) Prediction of minor stream delays at a limited priority freeway merge. *Transportation Research part B*, Vol. **37**, pp. 719-735.

Campari, E.G. and Levi, G. (2000) A cellular automata model for highway traffic. *The European Physical Journal B*. Vol. **17**, pp. 159-166.

Campari, E.G., Levi, G. and Maniezzo, V. (2004) Cellular automata and roundabout traffic simulation. Proceedings of the Sixth International Conference on Cellular Automata for Research and Industry, 25 – 27 October, Amsterdam, Netherland.

Cape Cod Commission. (2003) Guidelines for Transportation Impact Assessment. Technical Bulletin 96-003.

Chopard, B., Dupuis, A. and Luthi, P. (1998) Traffic *and Granular Flow'97*, World Scientific. pp. 153-168.

Chowdhury, D. and Schadschneider, A. (1999) Self-organization of traffic jams in cities: Effects of stochastic dynamics and signal periods. *Physical Review E*, Vol. **59**, R1311–R1314.

Chowdhury, D., Santen, L. and Schadschneider, A. (2000) Statistical Physics of Vehicular Traffic and Some Related Systems. *Physics Reports*, Vol. **329**, pp. 199.

CORSIM User Manual, Version. 1.01, The Federal Highway Administration, US Dept. of Transportation.

Cowan, R.J. (1975) Useful headway models. *Transportation Research* **9/6**, pp. 371–375.

Daoudia, A.K. and Moussa, N. (2002) A New Version of the Nagatani Model for Traffic on a Two-lane Roadway. *Chinese Journal of Physics,* Vol. **40**, No. 5, pp. 484-489.

Daoudia, A.K. and Moussa, N. (2003) Numerical Simulations of a Three-Lane Traffic Model Using Cellular Automata. *Chinese Journal of Physic,* Vol. **41**, No. 6, pp. 671-681.

Del Rio, J.A. and Larraga, M.E. (2004) Transient situations in traffic flow: Modelling the Mexico City Cuernavaca Highway. Proceedings of Second Mexican Meeting on Mathematical and Experimental Physics, El Colegio Nacional, Mexico City, Mexico.

Dia, H. (2001) An object-oriented neural network approach to short-term traffic forecasting. *European Journal of Operational Research*, Vol. **131**, pp. 253-261.

Diedrich, G., Santen, L., Schadschneider, A. and Zittartz, J. (2000) Effects of on- and off-ramps in cellular automata models for traffic flow. *International Journal of Modern Physics C,* Vol. **11**, pp. 335.

Dougherty, M.S and Cobbett, M.R. (1997) Short-term inter-urban traffic forecasts using neural networks. *International Journal of Forecasting,* Vol. **13**, Issue 1, pp. 21-31.

Dupuis, A., Chopard, B. (2003) Cellular Automata Simulation of Traffic: A Model for the City of Geneva. *Networks and Economics*, Vol. **3**, pp. 9-21.

Eiben, A.E., Hinterding, R. and Michalewicz, Z. (1999) Parameter control in evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, Vol. **3**(2), pp. 124–141.

Evans, M.R. (1997) *Physica A*, Vol. **30**, pp. 5669.

Ferreira, T.A.E, Vasconcelos, G.C and Adeodato, P.J.L. (2004) A Hybrid Intelligent System Approach for Improving the Prediction of Real World Time Series. *Proceedings*

of Congress on Evolutionary Computation, Vol. **1**, 19-23 June, pp.736 – 743.

Fish, K.E., Johnson, J.D., Dorsey, R.E. and Blodgett, J.G. (2004) Using an artificial neural network trained with a genetic algorithm to model brand share. *Journal of business Research*, Vol. **57**, pp. 79-85.

Flannery, A., Datta, T. (1997) Operational performance measures of American roundabouts. *Transportation research Record*, Vol. **1572**, pp. 68 -75.

Fouladvand, M.E., Sadjadi, Z. and Shaebani, M.R. (2003) Characteristics of Vehicular Traffic Flow at a Roundabout. Preprints. cond-mat/0309560.

Freund, J. and Poschel, T.P. (1995) *Physica A*, Vol. **219**, pp. 95.

Fukui, M. and Ishibashi, Y. (1996) Traffic Flow in 1D Cellular Automaton Model Including Cars Moving with High Speed. *Journal of Physical Society of Japan*. Vol. **65**, pp. 1868.

Gallant, P.J. and Aitken, J.M. (2003) Genetic Algorithm Design of Complexity-controlled Time-series Predictors. *Proceedings of the IEEE XIII Workshop on Neural Networks for Signal Processing*. 17-19 September, pp. 769 – 778.

Ganguly, N., Sikdar, B.K., Deutsch, A., Canright, G. and Chaudhuri, P.P. (2001) A Survey on Cellular Automata. www.cs.unibo.it/bison/publications/CAsurvey.pdf

Ganguly, N. and Deutsch, A. (2004) A Cellular Automaton Model for an Immune Derived Search Algorithm. *Lecture Notes in Computer Science,* Vol. **3305**, pp. 61-70.

Gartner, N.H., Messer, C.J. and Rathi, A.K. (1997) the Revised Monograph on Traffic Flow Theory.

Gipps, P.G. (1981) A behavioural car-following model for computer simulation. *Transportation Research 15B*, pp. 105–111.

Goharian, N. (2003) http://ir.iit.edu/~nazli/cs422/CS422-Slides/DM-Preprocessing.pdf Accessed on 4 March 2005.

Haykin, S. (1999) *Neural Networks: A Comprehensive Foundation*. Printice Hall.

Hays, W.L. (1994) Statistics (the 5th Edition). University of Texas at Austin, Harcourt Brace College Publishers.

Helbing, D. (1995) Improved fluid-dynamic model for vehicular traffic. *Physical Review E*, Vol. **51**, pp. 3164-3169.

Helbing, D. (1997) Fundamentals of traffic flow. *Physical Review E,* Vol. **55**, pp. 3735-3738.

Helbing, D. And Schreckenberg, M. (1999) Cellular automata simulating experimental properties of traffic flow. *Physical Review E*, Vol. **59**, R2505–R2508.

Hobeika, A.G. and Chang K.K. (1994) Traffic Flow Prediction Systems Based on Upstream Traffic. *Proceedings of vehicle Navigation and Information System Conference.*

Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor.

Hornik, K. (1989) Multilayer feedforward networks are universal approximates. *Neural Networks*, Vol. **2**, pp. 359-366.

Hornik K. (1991) Approximation capability of multiplayer feedforward networks. *Neural Networks*, Vol. **4**, pp.251-257.

Horiguchi, T. and Sakakibara, T. (1998) *Physica A*, Vol. **252**, pp. 388.

Housatonic Valley Council of Elected Officials. (2005) http://www.hvceo.org/transport/signal_coordination.php Accessed on 28th December 2004

Hu, J., Zong, C., Song, J., Zhang, Z. and Ren, J. (2003) An Applicable Short-term Traffic Flow Forecasting Method Based on Chaotic Theory. *Proceedings of the IEEE International Conference on Intelligent Transportation Systems.* Vol. **1**, pp. 608-613.

Huang, D.W. (2002) Effects of Ramps in the Nagel-Schreckenberg Traffic Model. *International Journal of Modern Physics C*, Vol. **13**, Issue 6, pp. 739-749.

Jia, B., Jiang, R. and Wu, Q. (2005) The effects of accelerating lane in the on-ramp system. *Physica A,* Vol. **345,** Issues 1-2, pp. 218-226.

Jiang, R., Wu, Q.S. and Wang, B.H. (2002) Cellular automata model simulating traffic interactions between on-ramp and main road, *Physical Review E*, Vol. **66**, 036104.

Jiang, R., Jia, B. and Wu, Q.S. (2003) The stochastic randomization effect in the on-ramp system: single-lane main road and two-lane main road situations. *Journal of Physics A: Mathematical and General,* Vol. **36**, pp. 11713-11723.

Jiang, R. and Wu, Q.S. (2003) Cellular automata models for synchronized traffic flow, *Journal of Physics A: Mathematical and General*, Vol. **36**, pp. 381-390.

Kerner, B.S. and Rehborn, H. (1997) Experimental Properties of Phase Transitions in Traffic Flow. *Physical Review Letter*, Vol. **79**, pp. 4030–4033.

Kerner, B.S. (1998) Experimental Features of Self-Organization in Traffic Flow. *Physical Review Letter,* Vol. **81**, pp. 3797.

Kerner, B.S. (2000) Experimental features of the emergence of moving jams in free traffic flow. *Journal of Physics A: Mathematical and General*, Vol. **33**, pp. L221 – L228.

Kerner, B.S. (2002) Empirical macroscopic features of spatial-temporal traffic patterns at highway bottlenecks. *Physical Review E*, Vol. **65**, pp. 46138.

Khan, A.M. (1991) *Cost-effectiveness of Passing lanes: Safety, Level of Service, and Cost Factors*, Research and Development Branch, Ministry of Transportation, Ontario.

Kim, G., Yoon, J., An, S., Cho, H. and Kang, K. (2004) Neural network model incorporating a genetic algorithm in estimating construction costs. *Building and Environment*, Vol. **39**, pp. 1333-1340.

Kimber, R.M. (1980) The Traffic Capacity of Roundabouts. TRRL Laboratory Report 942. Transportation and Road research Laboratory, Crowthorne, Berkshire, UK.

Klar, A. and Wegener, R. (2000) Vehicular Traffic: From Microscopic to Macroscopic Description. *Transport Theory and Statistical Physics*, Vol. **29**(3-5), pp. 479-493.

Knospe, W., Santen, L., Schadschneider, A. and Schreckenberg, M. (1999) Disorder effects in cellular automata for two-lane traffic. *Physica A,* Vol. **265**, pp. 614.

Knospe, W., Santen, L., Schadschneider, A. and Schreckenberg, M. (2000) Towards a realistic microscopic description of highway traffic. *Journal of Physics A,* Vol. **33**, pp. 477 - 485.

Knospe, W., Santen, L., Schadschneider, A. and Schreckenberg, M. (2002) A realistic two-lane traffic model for highway traffic. *Journal of Physics A,* Vol. **35**, pp. 3369-3388.

Knospe, W., Santen, L., Schadschneider, A. and Schreckenberg, M. (2002a) Single-vehicle data of highway traffic: microscopic description of traffic phases *Physical Review E*, Vol. **65**, 056133.

Knospe, W., Santen, L., Schadschneider, A. and Schreckenberg, M. (2002b) Human behavior as origin of traffic phases. *Physical Review E*, Vol. **65**, 015101.

Koorey, G. and Tate, F. (1999) Road infrastructure assessment model - Incorporation of passing-lane sections. *IPENZ Transactions,* 1999, Vol. 26, No. 1, pp. 13-19.

Lárraga, M.E., del Río, R.A. and Alvarez-Icaza, L. (2005) Cellular automata for one-lane traffic flow modelling. *Transportation Research Part C: Emerging Technologies,* Vol. **13**, Issue 1, pp. 63-74.

Lee, H.K., Lee, H.W. and Kim, D. (2004) Steady state solutions of hydrodynamic traffic models. *Physical Review E*, Vol. **69**, 016118.

Lee, H.Y., Lee, H.W. and Kim, D. (1999) Dynamic states of a continuum traffic equation with on-ramp. *Physical Review E*, Vol. **59**, pp. 5101-5111.

Lee, H.Y., Lee, H.W. and Kim, D. (2000) Phase diagram of congested traffic flow: An empirical study. *Physical Review E*, Vol. **62**, pp. 4737-4741.

Lighthill, M.H. and 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, pp. 229.

Lingras, P. and Sharma S. (1999) Short-Term Traffic Volume Forecasts: Existing And FutureResearch, *Proceedings of the Annual Conference of Canadian Society of Civil Engineers*, Regina, Saskatchewan, IV, pp. 429-438.

Liu, M., Wang, R. and Kemp, R. (2005a) Towards a Realistic Microscopic Traffic Simulation at an Unsignalised Intersection. *Lecture Notes in Computer Science,* Vol. **3481**, pp. 1187-1196.

Liu, M., Wang, R., Wu, J. and Kemp, R. (2005b) A Genetic-Algorithm-Based Neural Network Approach for Short-term Traffic Flow Forecasting. *Lecture Notes in Computer Science,* Vol. **3498**, pp. 965-970.

Luttinen, R.T. (2004) Capacity and Level of Service at Finnish Unsignalized Intersections. Finnish Road Administration, Finnra Reports 1/2004, ISBN 951-803-180-0, ISSN 1457-9871.

Maerivoet, S. and De Moor, B. (2003) Development of an improved traffic cellular automaton model for traffic flows on a highway road network. *Proceedings of the 10th World Congress and Exhibition of Intelligent Transport Systems and Services (ITSS03)*, Madrid, Spain, pp. 16.

Messai, N., Thomas, P., Lefebvre, D. and Moudni, A.E. (2002) A Neural Network Approach for Freeway Traffic Flow Prediction. *Proceedings of the IEEE International Conference on Control Applications*, 18-20 September 2002, Glasgow, Scotland, U.K.

Miller, G., Todd, P., Hedge, S. (1989). Designing neural networks using genetic algorithms. In: Schaffer, J. (Ed.), the Third International Conference on Genetic Algorithms and Their Applications, CA, San Mateo.

Nagatani, T. (1993) Self-organization and phase transition in traffic-flow model of a two-lane roadway. *Journal of Physics A,* Vol. **26**, pp. 781.

Nagel, K. and Schreckenberg, M. (1992) A cellular automaton model for freeway traffic. *Journal of Physics,* I (France) Vol. **2**, pp. 2221-2229.

Nagel, K. (1996) Particle Hopping vs. Fluid-Dynamical Models For the Traffic Flow, Traffic and Granular Flow'95, pp. 41-56, World Scientific.

Nagel, K., Wolf, D. E., Wagner, P. And Simon, P. (1998) Two-lane traffic rules for cellular automata: a systematic approach. *Physical Review E*, Vol. **58**, pp. 1425-1437.

Nassab, K., Schreckenberg, M., Ouaskit, S. And Boulmakoul, A. (2005) Impacts of different types of ramps on the traffic flow. *Physica A*, Vol. **352**, Issue 2-4, pp. 601-611.

Neubert, L., Santen, L., Schadschneider, A. and Schreckenberg, M. (1999) Single-vehicle data of highway traffic: A statistical analysis. *Physics Review E*, Vol. **60**, pp. 6480.

Neumann, J.V. (1966) The Theory of Self-Reproducing Automata. Burks (ed), Univ. of Illinois Press, Urbana and London.

New York City Department of Transportation. http://www.nyc.gov/html/dot /html/about /faqs_signals.html, Accessed on 12 January, 2006.

Newell, G.F. (1961) Operation Research. Vol. **9**, pp. 209.

Niska, H., Hiltunen, T., Karppinen, A., Ruuskanen, J. and Kolehmainen, M. (2004) Evolving the neural network model for forecasting air pollution time series. *Engineering Application of Artificial Intelligence*, Vol. **17**, pp. 159-167.

Oketch, T., Delsey, M. and Robertson, D. (2004) Evaluation of Performance of Modern Roundabouts Using Paramics Micro-simulation Model. Transportation Association of Canada (TAC) Annual Conference, Quebec City, 19-22 September.

Okutani, I. and Stephanides, Y.I. (1984) Dynamic Prediction of Traffic Volume through Kalman theory. *Transportation research Part B*, Vol. **18**, No. 1, pp. 1-11.

Payne, H.J. (1971) Models of freeway traffic and control. Mathematical Models of Public Systems. Simulation Council Proceedings Series, Vol. **1**, No. 1, pp. 51-61.

Pedersen, M.M. and Ruhoff, P.T. (2002) Entry ramps in the Nagel-Schreckenberg model. *Physical Review E*, Vol. **65**, 056705.

Pollatschek, M.A., Polus, A. and Livneh, M. (2002) A decision model for gap acceptance and capacity at intersections. *Transportation Research Part B,* Vol. **36**, pp. 649-663.

Rakha, H. and Van Aerde, M.W. (1997) Comparison of simulation modules of TRANSYT and INTEGRATION models. *Transportation Research Record*, Vol. **1566**, pp. 1-7.

Richards, P.I. (1956) Shockwaves on the Highway. *Operations Research*, Vol. **4**, pp. 42-51.

Rickert, M., Nagel, K., Schreckenberg, M. and Latour, A. (1996) Two lane traffic simulations using cellular automata. *Physica A*, Vol. **231**, pp. 534.

Ruskin, H.J. and Wang, R. (2002) Modelling Traffic Flow at an Urban Unsignalised Intersection, Proceedings of the International Conference on Computational Science, 21-24 April, Amsterdam, *Lecture Notes in Computer Science*, Vol. **2329**, pp. 381-390.

Schadschneider, A. (1999) The Nagel-Schreckenberg model revisited. *The European Physical Journal B*, Vol. **10**, pp. 573-584.

Schadschneider, A. (2002) Traffic flow: a statistical physics point of view, *Physica A*: Statistical Mechanics and its Applications, Vol. **313**, Issues 1-2, pp. 153-187.

Schadschneider, A., Knospe, W., Santen, L. and Schreckenberg, M. (2005) Optimization of highway networks and traffic forecasting. *Physica A*, Vol. **346**, pp. 165-173.

Simon, P.M., Nagel, K. (1998) Simplified cellular automata model for city traffic. *Physical Review E*, Vol. **58**, pp. 1286.

Smith, B.I. and Demetsky, M.J. (1997) Traffic Flow forecasting: Comparison of Modeling Approaches. *Journal of Transportation Engineering*, Vol. **123**, pp. 261-265.

Srinivasan, B., Prasad, U.R. and Rao, N.J. Back propagation through adjoints for the identification of nonlinear dynamic systems using recurrent neural models, *IEEE Transactions on Neural Networks.* Vol. **5**(2), pp. 213-228.

Sun, H., Liu, H., Xiao, H., Ran, B. (2003) Short term traffic forecasting using local linear regression model. TRB annual meeting, Washington D.C.

Suzuki, H. and Nakatsuji, T. (2003) Effect of adaptive cruise control (ACC) on traffic throughput: numerical example on actual freeway corridor. *JSAE Review*, Vol. **24**, pp. 403–410.

Takehito, S., Yasushi, H. and Tsuyoshi, H. (2000) *Physica A,* Vol. **276**, pp. 316-337.

The Official New Zealand Road Code (2003) Land Transport Safety Authority.

Tian, L. and Noore, A. (2005) Evolutionary neural network modeling for software cumulative failure time prediction. *Reliability Engineering and System Safety*, Vol. **87**, pp. 45-51.

Tian, Z., Vandehey, M., Robinson, B.W., Kittelson, W., Kyte, M., Troutbeck, R., Brilon, W. and Wu, N. (1999) Implementing the maximum likelihood methodology to measure a driver's critical gap. *Transportation Research Part A*, Vol. **33**, Issues 3, pp. 187-197.

Toffoli, T., Margolus, N. (1987) *Cellular Automata Machines--A New Environment for Modelling* http://pm1.bu.edu/~tt/cambook, MIT Press.

Torok, J. and Kertesz, J. (1996) *Physica A*, Vol. **231**, pp. 515.

TRANSIMS (2002) User Manual, Version 3.0.1, Los Alamos National Laboratory.

Transportation Research Board. (2000) Highway Capacity Manual. National Research Council, Washington, D.C., U.S.A.

Traffic Engineering Handbook. (1992) Institute of Transportation Engineers, U.S.A.

Treiterer, J. (1975) Investigation of Traffic Dynamics by Aerial Photogrammetry Techniques. Final Report EES278, Transportation Research Center, Department of Civil Engineering, Ohio State University.

Troutbeck, R.J. (1986) Average delay at an unsignalized intersection with two major streams each having a dichotomized headway distribution. *Transportation Science.*

Troutbeck, R.J. (1998) Background for HCM section on analysis of performance of roundabouts. *Transportation research Record,* Vol. **1646**, pp. 54 - 62.

Troutbeck, R.J. and Kako, S. (1999) Limited priority merge at unsignalized intersections. *Transportation Research Part A,* Vol. **33**, pp. 291-304.

Vichniac, G.Y. (1984) Simulation physics with cellular automata. *Physica* **D**, 10, 96-116.

Vlahogianni, E.I, Golias, J.C. and Karlaftis, M.G. (2004) Short-term Traffic Forecasting: Overview of Objectives and Methods. *Transport Review,* Vol. **24**, No. 5, pp. 533-557.

Wagner, P., Nagel, K. and Wolf, D. (1997) Realistic multi-lane traffic rules for cellular automata. *Physica A*, Vol. **234**, pp. 687-698.

Wang, R. (2003) Modelling Unsignalised traffic Flow with Reference to Urban and Interurban Networks. Doctorate Thesis. Dublin City University.

Wang, R. and Liu, M. (2005) A Realistic Cellular Automata Model to Simulate Traffic Flow at Urban Roundabouts. *Lecture Notes in Computer Science,* Vol. **3515**, pp. 420-427.

Wang, R., and Ruskin, H.J. (2002) Modeling Traffic Flow at a Single-lane Urban Roundabout, *Computer Physics Communications*, Vol. **147**, pp. 570-576.

Wang, R. and Ruskin, H.J. (2003a) Modelling Traffic Flow at a Multilane Intersection. *Proceedings of the International Conference on Computational Science and its Applications*, 18 – 21 May, Montreal, Lecture Notes in Computer Science (LNCS).

Wang, R. and Ruskin, H.J. (2003b) Modelling Traffic Flow at a two-lane Roundabout. *Proceedings of International Conference on Computer Science, Software Engineering, Information Technology, e-Business and Applications*, 5 – 7 June, Rio de Janeiro, Brazil.

Wang, R. and Ruskin, H.J. (2006) Modelling traffic flow at multilane urban roundabouts. *International Journal of Modern Physics C*, Vol. **17**, No. 3.

White, H. (1990) Connectionist non-parametric regression: multilayer feedforward networks can learn arbitrary mapping. *Neural Networks*, Vol. **3**, pp. 535-549.

William, B.M. (1999) Modeling and Forecasting Vehicular Traffic Flow as a Seasonal Stochastic Time Series process. Doctoral dissertation, Department of Civil Engineering, University of Virginia, Charlottesville.

von Neumann, J. (1966) *The Theory of Self-reproducing Automata*, A. Burks, ed., Univ. of Illinois Press, Urbana, IL.

Wolfram, S. (1986) Theory and Applications of Cellular Automata. World Scientific, Singapore. ISBN 971-50-124-4.

Wolfram, S. (2002) A New Kind of Science. Wolfram Media, Inc., ISBN 1-57955-008-8.

WWW1: http://www.co.st-ohns.fl.us/BCC/gmsvcs/Planning/DRI/Transportation.Appendix/HCS.2-4-03.pdf. Accessed on 10 November 2004.

WWW2: http://www.rpi.edu/dept/cits/files/ops.ppt. Accessed on 20 October 2004.

Xu, Y. and Wong, K.W. (2003) Effect of decay function on the generalization ability of twdrls algorithms. IEEE International Conference on Neural Networks and Signal Processing. 14-17 December Nanjing, China.

Xue, Y., Dong, L., Li, L. and Dai, S. (2005) Effects of changing orders in the update rules on traffic flow. *Physical review E*, Vol. **71**, 026123.

Yang, L., Jia, L. and Wang, H. (2004) Wavelet Network with Genetic Algorithm and its Applications for Traffic Flow Forecasting. *Proceedings of the 5th World Congress on Intelligent Control and Automation*, 5-19 June, Hangzhou, P.R.China.

Yang, Q. and Koutsopoulos, H.N. (1996) A Microscopic Traffic Simulator for Evaluation of Dynamic Traffic Management Systems. *Transportation Research Part C*, Vol. **4** (3), pp. 113-129.

Yao, X. (1999). Evolving artificial neural networks. *Proceedings of the IEEE Transactions on Neural Networks,* Vol. **87**, pp. 1423–1447.

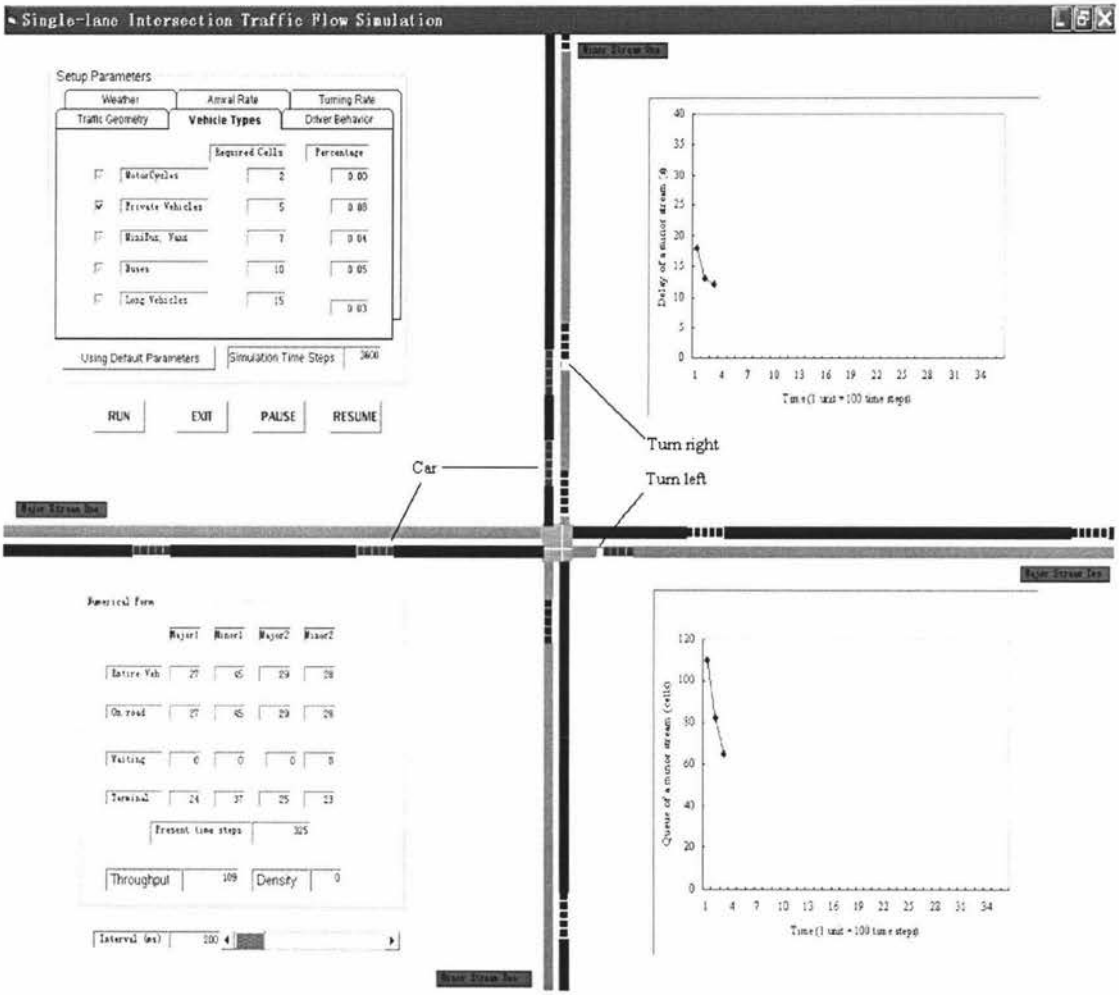
Yu, G., Hu, J., Zhang, C., Zhuang,L. and Song, J. (2003) Short-term Traffic Flow Forecasting Based on Markov Chain Model. *Proceedings of the IEEE Intelligent Vehicles Symposium*, 9-11 June, pp. 208 - 212.

Yukawa, S., Kikuchi, M., and Tadaki, S. (1994) Dynamical Phase Transition in One–Dimensional Traffic Flow Model with Blockage. *Journal of Physical Society of Japan*, Vol, **63**, No. 10, pp. 3609-3618.

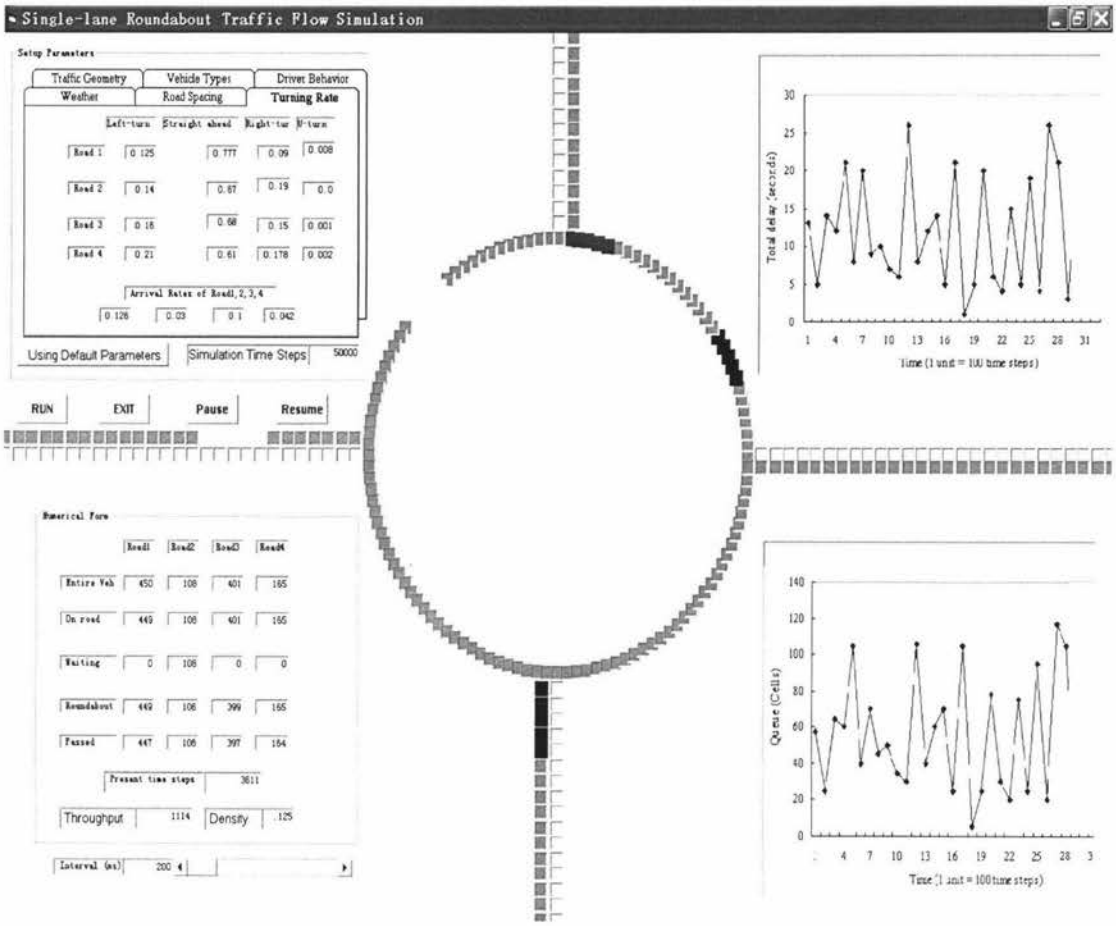
Zhang, H.M. and Kim, T. (2005) A car-following theory for multiphase vehicular traffic flow. *Transportation Research Part B: Methodological,* Vol. **39**, Issue 5, pp. 385-399.

Appendixes

Appendix A: Interface of traffic flow simulation at a single-lane TWSC intersection



Appendix B: Interface of traffic flow simulation at a single-lane urban roundabout



Appendix C: Interface of a Genetic-Algorithm-based Neural Networks (GANN) method for short-term traffic flow forecasting

Appendix D: Traditional methods for computing capacity, delay and queue length

Capacity is a main determinant of performance measures (delay, queue-length etc.). The relationship between performance measures and capacity is often expressed in terms of degree of saturation (volume – capacity ratio). The typical desirable maximum degree of saturation is 0.85. The higher degree of saturation results in more unstable operation, increased delay and queuing.

Two kinds of approaches, analytical method and empirical method, have been used to evaluate capacity. For analytical method, capacity is looked as a function of critical gap, follow-up time and proportion of bunched vehicles, while capacity is a function of different geometric features for empirical method (Akcelik, 2001). However, capacity estimate illustrated in Figure D.1 has been widely accepted and used to measure single-lane roundabout performance.

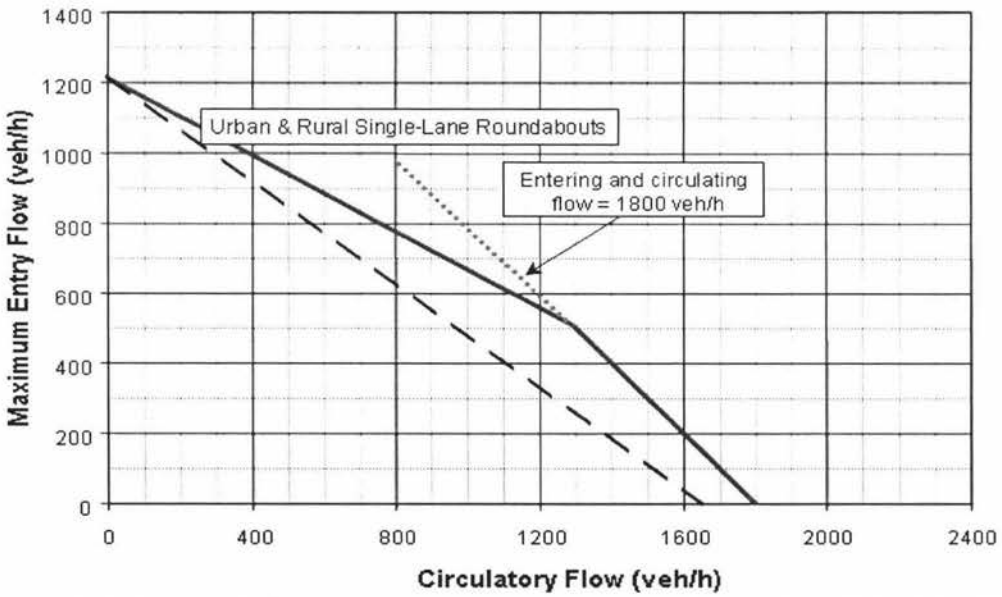


Figure D.1 Relationship between entry capacity and circulatory flow (Akcelik, 2001)

Delay is usually divided into control delay and geometry delay. The control delay includes initial deceleration delay, queue move-up time, stopped delay, and final acceleration delay while geometry delay is caused by geometry features. The control delay is typically used and computed in the following equation (Akcelik, 2001).

$$D = \frac{3600}{c} + 900\ T \times \left[\frac{v}{c} - 1 + \sqrt{\left(\frac{v}{c} - 1 \right)^2 + \frac{\left(\frac{3600}{c} \right)\left(\frac{v}{c} \right)}{450\ T}} \right] \quad \text{(D.1)}$$

where

D = control delay, s/veh

v = entry flow, veh/h

c = capacity, veh/h

T = time period = 0.25 h

Queue length is often estimated by average queue (50^{th} percentile) and maximum queue (95^{th} percentile). Queue length is computed in the following equations D.2 and D.3, respectively (Akcelik, 2001).

$$L = v * d / 3600 \tag{D.2}$$

where:

L = average queue length, veh

v = entry flow, veh/h

d = average delay, s/veh

$$Q_{95} = 900\ T \times \left[\frac{v}{c} - 1 + \sqrt{\left(1 - \frac{v}{c}\right)^2 + \frac{\left(\frac{3600}{c}\right)\left(\frac{v}{c}\right)}{150\ T}} \right] \times \frac{c}{3600} \tag{D.3}$$

where:

Q_{95} = queue length, veh

v = entry flow, veh/h

c = capacity, veh/h

T = time period = 0.25 h

Appendix E: Glossary of Traffic Terms

Aggregate Delay

The summation of delays for multiple lane groups usually aggregated for an approach, an intersection, or an arterial route.

All-Way Stop-Controlled (AWSC)

An intersection with stop signs at all approaches. The driver's decision to proceed is based on the rules of the road (e.g. the driver on the right has the right-of-way) and also on the traffic conditions of the other approaches.

Arrival Rate

The mean of the statistical distribution of vehicles arriving at a point or uniform segment of a lane or roadway.

Bottleneck

A road element on which demand exceeds capacity.

Calibration

The process of comparing model parameters with real-world data to ensure that the model realistically represents the traffic environment. The objective is to minimize the discrepancy between model results and measurements or observations.

Capacity

The maximum sustainable flow rate at which vehicles or persons reasonably can be expected to traverse a point or uniform segment of a lane or roadway during a specified time period under given roadway, geometric, traffic, environmental, and control conditions; usually expressed as vehicles per hour, passenger cars per hour, or persons per hour.

Circulating Flow

The volume of traffic on the principal roadway of a roundabout at a given time.

Critical Gap

The minimum time, in seconds, between successive major-stream vehicles, in which a minor-street vehicle can make a manoeuvre.

Delay

The additional travel time experienced by a driver, passenger, or pedestrian.

Density

The number of vehicles on a roadway segment averaged over space, usually expressed as vehicles per kilometre or vehicles per kilometre per lane. Also see Pedestrian density.

Downstream

The direction of traffic flow.

Flow Rate

The equivalent hourly rate at which vehicles, bicycles, or persons pass a point on a lane, roadway, or other traffic way. It is computed as the number of vehicles, bicycles, or persons passing the point, divided by the time interval (usually less than I h) in which they pass, and, expressed as vehicles, bicycles, or persons per hour.

Flow Ratio

The ratio of the actual flow rate to the saturation flow rate for a lane group at an intersection.

Free Flow

A flow of traffic unaffected by upstream or downstream conditions.

Gap Acceptance

The process by which a minor-street vehicle accepts an available gap to manoeuvre.

Headway

(1) The time, in seconds, between two successive vehicles as they pass a point on the roadway, measured from the same common feature of both vehicles (for example, the front axle or the front bumper).

(2) The time, usually expressed in minutes, between the passing of the front ends of successive transit units (vehicles or trains) moving along the same lane or track (or other guide way) in the same direction.

Level of Service (LOS)

A qualitative measure describing operational conditions within a traffic stream, based on service measures such as speed and travel time, freedom to manoeuvre, traffic interruptions, comfort, and convenience.

Major Street

The street not controlled by stop signs at a two-way stop-controlled intersection.

Minor Street

The street controlled by stop signs at a two-way stop-controlled intersection; also referred to as a side street.

Passenger Car Equivalent (PCE)

The number of passenger cars displaced by a single heavy vehicle of a particular type under specified roadway, traffic, and control conditions.

Passing Lane

A lane added to improve passing opportunities in one direction of travel on a conventional two-lane highway.

Peak-Hour Factor (PHF)

The hourly volume during the maximum-volume hour of the day divided by the peak 15-min flow rate within the peak hour; a measure of traffic demand fluctuation within the peak hour.

Platoon

A group of vehicles or pedestrians travelling together as a group, either voluntarily or involuntarily because of signal control, geometrics, or other factors.

Queue

A line of vehicles, bicycles, or persons waiting to be served by the system in which the flow rate from the front of the queue determines the average speed within the queue. Slowly moving vehicles or people joining the rear of the queue are usually considered part of the queue. The internal queue dynamics can involve starts and stops. A faster-moving line of vehicles is often referred to as a moving queue or a platoon.

Roundabout

An unsignalised intersection with a circulatory roadway around a central island with all entering vehicles yielding to the circulating traffic.

Two-Way Stop-Controlled (TWSC)

The type of traffic control at an intersection where drivers on the minor street or a driver turning left from the major street wait for a gap in the major-street traffic to complete a manoeuvre.

Upstream

The direction from which traffic is flowing.

Validation

Determining whether the selected model is appropriate for the given conditions and for the given task; it compares model prediction with measurements or observations.

Volume

The number of persons or vehicles passing a point on a lane, roadway, or other traffic-way during some time interval, often 1 h, expressed in vehicles, bicycles, or persons per hour.

The Glossaries of traffic terms are taken from the Highway Capacity Manual (2000), Transportation Research Board, National Research Council, Washington, D.C., USA