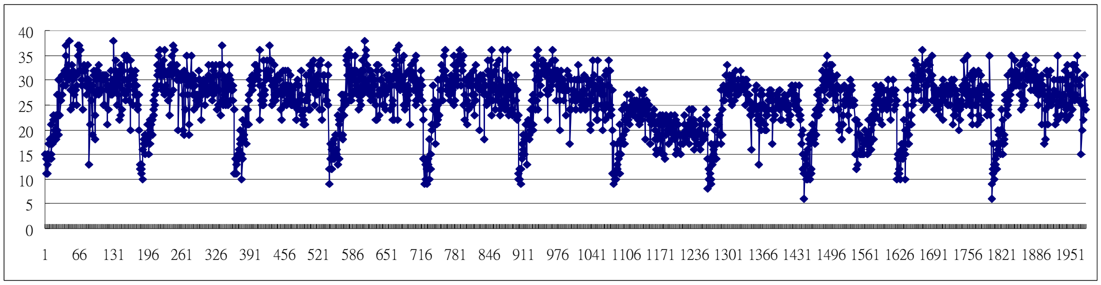
CHAPTER 1. INTRODUCTION

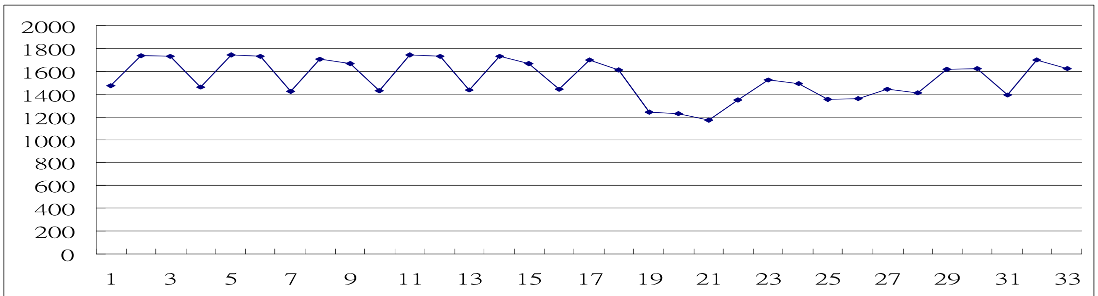
1.1 The Problem and Its Significance

The characteristics of short-term traffic flows, such as one-minute or shorter-time traffic counts, are considered as the key inputs to the development of the intelligent transportation systems (ITS). Many ITS applications have emphasized the importance of automatic collection of short-term traffic data, which is served as the foundation of the real-time traffic control strategies formulated by various traffic management authorities and is applied in traffic incidents simulation and monitoring (Ahmed and Cook, 1982). Only when the short-term traffic information be automatically collected and accurately predicted can an advanced traffic management and information system (ATMIS) operates in an efficient and effective manner (Davis and Nihan, 1991; Smith and Demetsky, 1997).

Essentially, the roadways traffic flows are derived from travel demands for goods and people, thus the traffic patterns are fully reflected by the trip decisions, including origin/destination, departure time, mode and route choices, made by the road users. If we look at the roadway traffic flow dynamics by examining the hourly flow time series at any location, the trace plot of the flow time series is typically not reproducible from day to day. If we repeat the observation another day under what seem to be identical conditions, the new hourly flow plot may be nearly, but not exactly, the same as the previous days. The hourly flow plots will generally have wiggles in different days and most previous researches have treated such non-reproducible traffic flow dynamics as stochastic (see, for instance, Newell, 1982; Daganzo, 1997). However, should the same traffic time series have been counted in a short time interval (e.g., one-minute or even shorter) rather than one hour, we would visualize more conspicuous fluctuations and larger wiggles as illustrated in Figure 1-1.



(a) Minute-flow Dynamics



(b) Hourly -flow Dynamics

Figure 1-1 Comparison of hourly- and minute-flow patterns for ten morning rush hours (Station 50, US I-35 Freeway, Minneapolis)

It is interesting to notice from Figure 1-1 that the short-term traffic flow patterns never repeat themselves, but the patterns from day to day look very alike, suggesting that there must intrinsically exist some deterministic rules governing the nature of such short-term flow patterns. The intrinsic rules could be a combination of several effects. For instance, the majority of trip makers might get to work by 9 am and take off at 5 pm in the workdays. They might depart from homes or work places at approximately the same times, using the same modes, and/or choosing the same routes everyday. These regularities could lead to similar flow patterns from day to day. Moreover, the roadway capacity at any location has limited the maximum number of vehicle throughputs such that the observed flow rates are always bounded within two extreme values: zero and capacity flows. However, the flow dynamics can never become eventually fixed points or stable equilibria at these two extreme points due to the heterogeneity of vehicle and driver behaviors. All of these factors might have governed the short-term flow dynamics to be nearly but not exactly the same as the previous days. In other words, the short-term traffic flow time series trajectories have "self-similarity" but never repeat themselves, which is essentially characterized as a chaotic system.

According to Sprott (2003), chaotic systems have several important characteristics: (1) they are aperiodic, namely trajectories or orbits never repeat; (2) they exhibit sensitive dependence on initial conditions (SDIC), hence they are unpredictable in the long run; (3) they are governed by one or more control parameters, a small change in which can cause the chaos to appear or disappear; and (4) their governing equations are nonlinear. In addition, the geometry with non-integer dimensionality, named "fractal," also plays an essential role in chaotic systems (Mandelbrot, 2000). Due to fractals, chaotic systems often exhibit self-similarity -- a portion of the system, once suitably magnified, will look like the system itself.

Chaos is part of the broader field of dynamical systems, which can be stochastic or deterministic. A stochastic system will change with time according to some random processes, including uncorrelated noise (or called white noise or pure noise process), Brownian noise (or called random walk process) and other correlated noises such as pink, black and blue noises which are related to different stochastic processes). A deterministic system, by contrast, will evolve under some deterministic rules (or mathematical equations) in such a way that the present state is uniquely determined by the past states. Such deterministic chaos can only occur when the governing rules or equations are nonlinear (Sprott, 2003). In the case illustrated in Figure 1-1, stochastic models might successfully elucidate the fluctuation of hourly flow patterns, but it could fail to capture the conspicuous fluctuations or drastic wiggles of minute-flow dynamics. As explained by the deterministic rules which intrinsically govern the nature of such short-term flow patterns, we presume that chaotic models are likely to work better than the stochastic models in capturing how a short-term traffic flow dynamics will evolve as time progresses.

1.2 Motivations and Purposes of the Research

Chaos theory has been widely applied to various areas, including nature science (ecological, biological, etc.), engineering (hydraulic, mechanical, chemical, etc.) and social science (financial, economic, etc.) systems, to explain and predict the nonlinear complex systems. In the area of traffic flow, some studies also utilize chaos theory to describe the traffic flow phenomena (for instance, Disbro and Frame, 1989). Addison and Low (1996, 1998) find that some specific parameter values of GM car-following models exist in chaos dynamics. Iokibe, et al. (1995), Dendrinos (1994), Frison and Abarbanel (1997) and Zhang and Jarrett (1998) prove that the short-term traffic flow has nonlinear chaotic phenomena. Lan and Lin (2003) apply Lyapunov exponent and cellular automata (CA) procedure to examine the urban traffic flow time series. In the short-term prediction of traffic flows, local fuzzy reconstruction method (Iokibe, et al., 1995) and phase space local approximation method (Farmer and Sidorowich, 1987) are utilized for the urban street as well as freeway traffic flow time series and both methods obtain satisfactory prediction accuracy (Lan and Chen, 1998; Lan and Lin, 2001).

Previous focus was primarily on reproducing traffic conditions through simulation and the traffic states with different densities (Pottmeier, et al. 2002; Schadschneider, 2000; Wolf, 1999). It has been argued that traffic flow behavior with low density (non-congested state) is like a random walk, whereas its behavior in the high-density (congested) state is characterized as a chaotic motion. Without further empirical evidence, this argument may not stand. Whether there exist chaotic or stochastic phenomena for the short-term traffic dynamics is still not generally clear to us. Only demonstrated with enough evidences can we utilize chaos theory to elucidate the traffic flow phenomena and to apply it to traffic prediction for various purposes, incident detection for non-recurrent congestion management, or chaos control for recurrent congestion management.

Chaos is one type of behaviors exhibited in nonlinear dynamical systems. Some sudden and drastic changes in nonlinear systems may give rise to the complex behavior called chaos. It is used to describe the time behavior of a system when that behavior is aperiodic (it never exactly repeats) and is apparently "noisy" like random (Hilborn, 1994). A chaotic system, which may appear to be random but is in effect generated by a deterministic model, cannot be examined by standard statistical techniques. Various diagnostic tools for distinguishing deterministic chaos from random systems have been attempted. However, it is argued that testing for chaos is more artistic than scientific and that no recipe will guarantee success for every case (Sprott, 2003). Therefore, seeking for more efficient and effective ways to test if chaotic phenomena exhibit in the nature of short-term traffic flow time series is a challenging issue and that motivates this research.

Many studies (Iokibe, 1995; Farmer and Sidorowich, 1987; Sakawa, et al. 1998) have shown that the chaos approach by means of state space reconstruction has the advantages of prediction. However, different prediction reasoning concepts can come up with various prediction models that might perform quite differently. Therefore, seeking for appropriate chaos prediction reasoning to predict short-term traffic flow dynamics is another challenging topic worthy of exploration. And this also motivates this research.

The first purpose of this research is to develop a parsimony procedure to test if chaotic phenomena exist in a traffic flow dynamics. A comprehensive comparison of geometric plots and statistics between the observed freeway traffic flow time series data and their surrogates are carried out. The most crucial indexes are selected to develop a parsimony testing procedure. Some known time series data generators are examined to validate the proposed procedure, which is then applied to testing for chaos of traffic flows at different locations.

The second purpose of this research is to attempt different prediction reasoning concepts to develop temporal confined (TC), spatiotemporal confined (STC) and spatial confined (SC) prediction models for short-term traffic forecast. The proposed TC model utilizes temporal similarity for flow trajectories in the prediction reasoning. The STC model incorporates both spatial and temporal similarities of the trajectories into the prediction reasoning. The SC model employs the spatial similarity to perform the prediction reasoning. The one-minute flow time series data from the United States I-35 Freeway are used to examine the prediction performance of these three models.

1.3 Methodologies and Framework of the Research

It is easy to incorrectly think a random system as chaos or a chaotic system as random by only judging their time series one-dimensional (trace) plots because they are very much alike. Therefore, we must introduce other powerful indexes that could make more distinction between stochasticity and chaoticity. The known geometric plots in chaos and time series literatures include return maps (plots of each local maximum versus the previous maximum), phase-space plots (slopes of the trajectories), Poincare maps (or Poincare movies), iterated function systems (IFS) clumpiness maps, autocorrelation function plots, probability distributions, and power spectra. The known statistics include the largest Lyapunov exponent, Kolmogorov entropy, Hurst exponent, relative complexity, capacity dimension, embedding dimension, correlation dimension, and delay time. Chapter three will highlight the definitions and properties of these promising geometric plots and statistics. This research will attempt batteries of these geometric plots and statistics to test for chaos by using the field observed one-minute flow time series data and their surrogates. The most crucial ones will be selected to develop a parsimony testing procedure.

Building the chaotic prediction model from a time series mainly involves two steps: reconstruction of state space from the time series data and development of reasoning for state-space prediction. As aforementioned, it is almost impossible to distinguish, by visualization methods, between a chaotic system and a stochastic system because both have very similar irregularity. However, if we reconstruct these time series in higher dimensional state spaces, we might be able to see the difference. Therefore, many studies (Iokibe, 1995; Farmer and Sidorowich, 1987; Sakawa et al., 1998) have shown that the chaos approach by means of state space reconstruction has the advantages of prediction.

The order of any traffic flow time series, in general, can be described by *N* state variables (*N* stands for the number of variables affecting a system) and represented as trajectories in the *N*-dimensional state space. For a single observed variable, its trajectory can be reconstructed in an *N*-dimensional space using delay time. Let $\zeta(t)$ denote the observed time series data and *N* and τ represent the embedding dimension and delay time, then the vector $Z(t)=\{\zeta(t),\zeta(t-\tau),...,\zeta(t-(N-1)\tau)\}$ indicates one point of the *N*-dimensional reconstructed state space. As such, a trajectory can be drawn in the *N*-dimensional space by changing *t* with τ fixed. When the embedding dimension *N* is sufficiently large, we can say that the trajectory is embedded in the reconstructed state space. Taken (1981) proved that by retaining the phase structure in the original *O*-dimensional state space, namely the reconstructed trajectory is embedded as $N \geqq 2O+1$, if the observed time series is chaotic then the trajectories of the time series will follow a certain deterministic regularity. Thus, if the deterministic regularity can be estimated, then the data in the near future (before the deterministic causality is lost) can be predicted.

Figure 1-2 illustrates the concept of reconstruction of state space by embedding the one-dimensional time series data into a three-dimensional state space. After plotting the time series data from the latest observations in the three-dimensional space, we can then investigate the historical observations neighboring to the latest observation. If the observed time series data (Figure 1-2(a)) is characterized with deterministic chaos, by Taken's theorem, after the state space reconstruction (Figure 1-2(b)), we can look for the "similarity" of trajectories by differencing the variables of the historical vectors $Z_m(t)$ and the present vector $Z_p(t)$. If the difference between $Z_p(t) - Z_p(t-v)$ and $Z_m(t) - Z_m(t-v)$ is smaller than a specific threshold value, then we can view that the present vector has "similar pattern" as the historical ones, where *v* denotes the time steps ahead of the present time *t*. By using the changes in the "historical similar trajectories" from $Z_m(t)$ to $Z_m(t+s)$ we can perform the prediction reasoning for the latest observed vector from $Z_p(t)$ to $Z_p(t+s)$, where *s* denotes the time steps after the present time *t*.

Of course, seeking for appropriate prediction reasoning for chaotic traffic flow dynamics forecast will play the most critical role in developing the prediction models. Various rationales for the "prediction reasoning" will come up with different models leading to diverse predictive performances. This research will develop three prediction models based on various rationales. The first model, based on temporal confined (TC) concept, only employs the "temporal similarity" of flow trajectories to perform the prediction reasoning. The second model, based on spatiotemporal confined (STC) concept, incorporates both "spatial and temporal similarities" into the prediction reasoning. The third model, based on spatial confined (SC) concept, only considers the "spatial similarity" to carry out the reasoning.

The scopes of this research are only limited in two important fundamental issues:

testing and predicting the chaotic short-term traffic dynamics. Other chaotic related studies or applications, such as diagnosis for chaotic traffic patterns due to traffic incidents, examinations for the transients from non-chaotic to chaotic states, non-recurrent and recurrent traffic congestion management by chaos control, are not covered in this research. The framework of this study is shown in Figure 1-3.

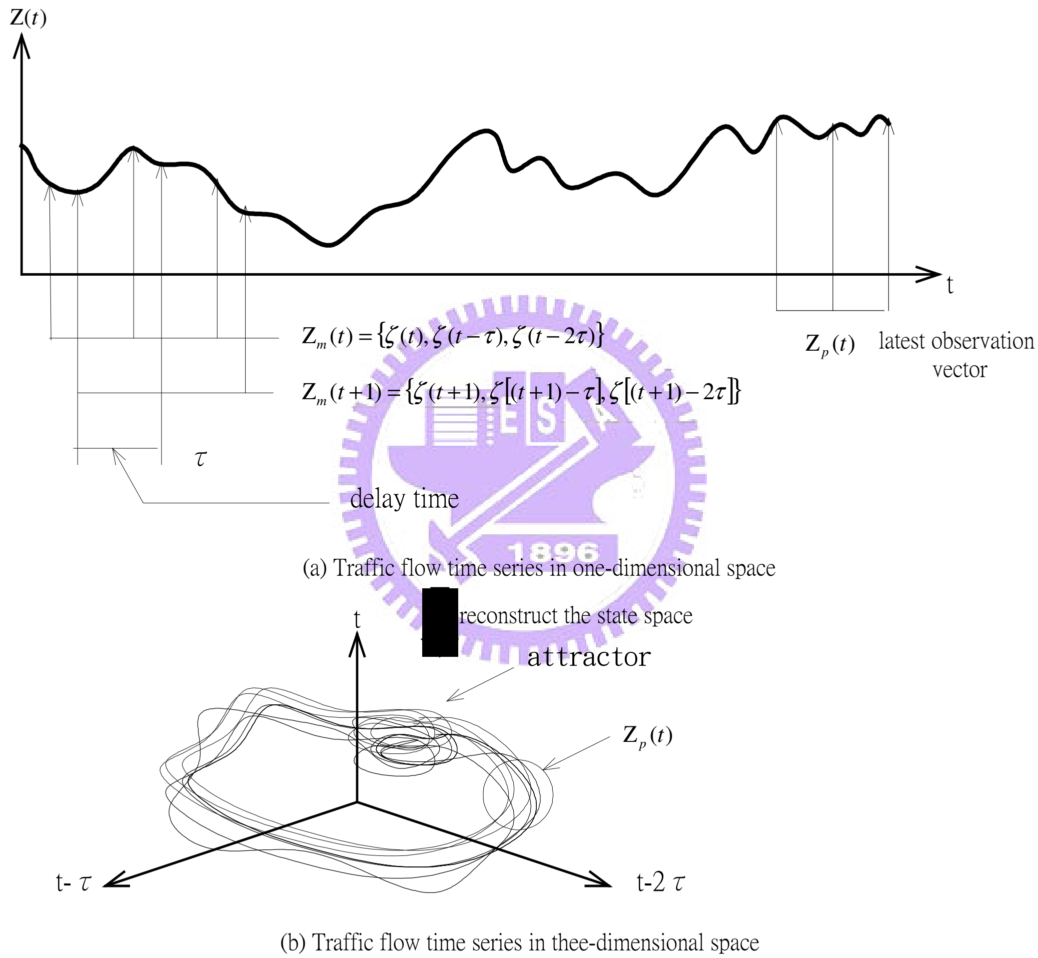


Figure 1-2 State space reconstruction

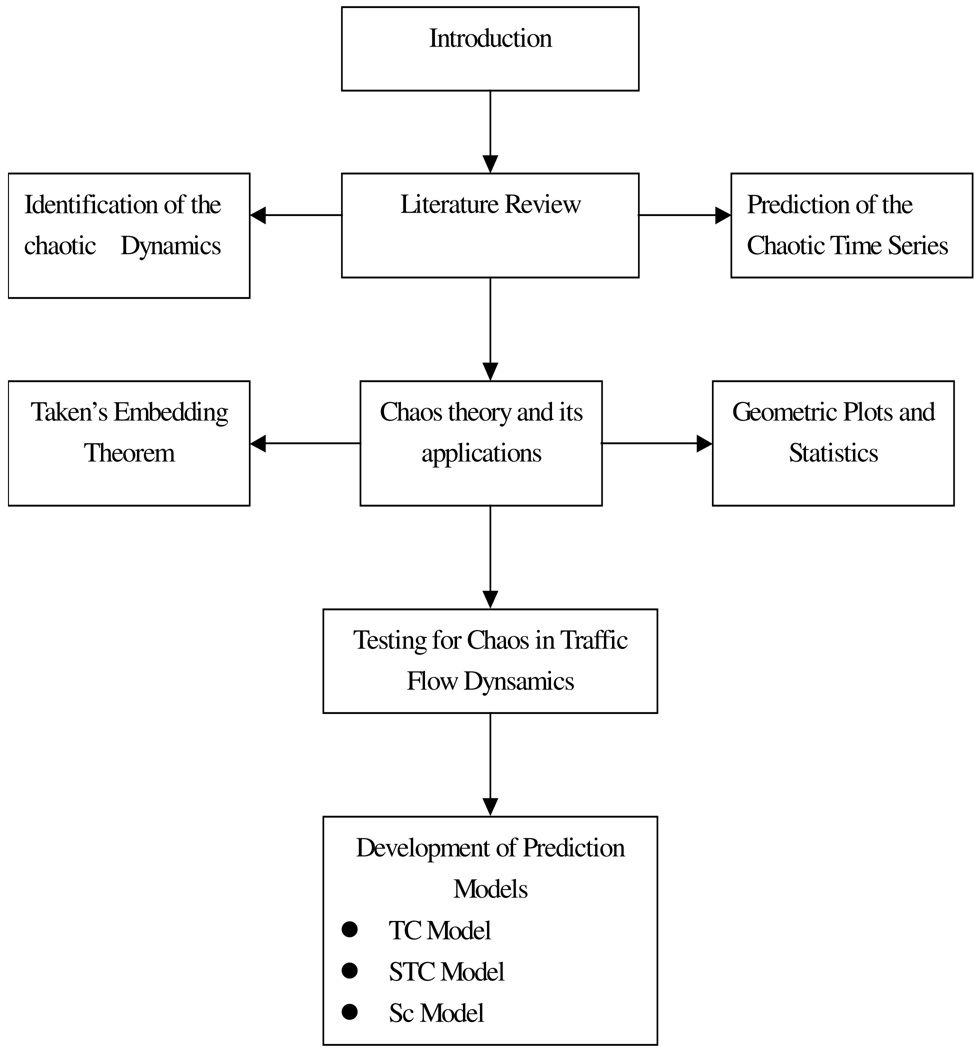


Figure 1-3 Framework of this research

1.4 Organization of the Research

This research is composed of six chapters. Chapter 1 describes the problem, motivation, purposes, and methodology of this research. Chapter 2 describes reviews the previous works on the identification of the chaotic traffic flow dynamics, prediction of the chaotic time series and some comments. Chapter 3 describes the methodology and framework for this research, including some properties of chaos, geometric plots, statistics and Taken's embedding theorem. Chapter 4 proposes a parsimony procedure to test if chaotic phenomena exist in a traffic flow dynamics.

Chapter 5 develops three chaotic prediction models, including the temporal confined (TC), the spatiotemporal confined (STC) and the spatial confined (SC) models. Chapter 6 summarizes the general conclusions and possible extensions.