THE APPLICATION OF CHAOS THEORY TO FORECAST URBAN TRAFFIC CONDITIONS

Abraham Tetteh Narh BSc (Hons.) MSc.

A thesis submitted for the degree of Doctor of Philosophy (PhD) at Newcastle University

SCHOOL OF CIVIL ENGINEERING & GEOSCIENCES

Transport Operations Research Group

December 2015

Abstract

This thesis explores the application of Chaos Theory to forecast urban traffic conditions. The research takes advantage of a highly resolved temporal and spatial data available from the Split Cycle Optimisation Technique (SCOOT) system, in order to overcome the limitations of previous studies to investigate applying Chaos Theory in traffic management.

This thesis reports on the development of a chaos-based algorithm and presents results from its application to a SCOOT controlled region in the city of Leicester, UK. A Phase Space Reconstruction method is used to analyse non-linear data from the SCOOT system, and establishes that a 20 second resolved data is suitable for understanding the dynamics of the traffic system. The research develops the Lyapunov exponent as a chaos-based parameter to forecast link occupancy using a multiple regression model based on the temporal and spatial relationships across the links in the network. The model generates a unique forecast function for each link for every hour of the day. The study demonstrates that Lyapunov exponents can be used to predict the occupancy profile of links in the network to a reasonably high level of accuracy (R-values generally greater than 0.6). Evidence also suggests that the predictions from the Lyapunov exponents (rather than occupancy) make it possible to report on the impending conditions over a wider part of the network so that imminent congested conditions can be foreseen in advance and mitigation measures implemented.

Thus, the thesis concludes that incorporating chaos-based algorithms in this way can enable urban traffic control systems to be one-step ahead of traffic congestion, rather than one-step behind. This would improve the management of traffic on a more strategic level rather than purely within smaller network regions thus playing an important role in improving journey times and air quality and making a vital contribution to mitigating climate change.

Acknowledgements

I am grateful to Almighty God for grace and keeping and good health through this journey.

It is worthy to acknowledge also all those who through diverse ways have contributed to the successful completion of this research work.

I am grateful for the support and encouragement of my supervisory team, through whose guidance and contribution throughout this research have made this thesis possible. My sincere appreciation goes to Dr. Neil Thorpe who supported my application and offered advice in order to secure a scholarship for this research, and for the guidance and suggestions throughout the course of my studies, and for his contribution to finalising this thesis. Many thanks goes to Prof. Margaret Bell for bringing on board her experience and knowledge and teaching at tutorial sessions, and for the helpful feedback on my work and correcting of this thesis, and for the kindness to my family. Special thanks goes to Dr. Graeme Hill for his excellent guidance of the mathematical aspects of this research, and for checking the programming work and analysis.

I would like to thank the Engineering and Physical Sciences Research Council (EPSRC) and Newcastle University for granting me the opportunity to pursue this doctoral degree in Civil Engineering. I also thank all the staff and colleagues at the Transport Operations Research Group and my postgraduate research colleagues.

Special thanks go to my friends and well-wishers for their prayers and emotional support. This list is not complete without special mention of Miss Melissa Ware for her many assistance and Graham Patterson for the computing support that was indispensable throughout these years.

Finally, I thank my wife, Afeafa, and children, Nigel, Jeremy, and Rachel, for the ceaseless prayers, love, patience and unflinching support. My parents deserve special acknowledgement because they gave me formal education, and their vision has led to this cherished dream and I will be eternally grateful.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AASHTO	American Association of State Highway and Transportation Officials
ACL	Autocorrelation Lag
ANPR	Automatic Number Plate Recognition
ARCH	Autoregressive Conditional Heteroscedastic
ARIMA	AutoRegressive Integrated Moving Average
ARMA	AutoRegressive Moving Average
ASTRID	Automatic SCOOT Traffic Information Database
CO_2	Carbon Dioxide
CO	Carbon Monoxide
CONG	SCOOT Congestion
CGIF	ConGestion Importance Factor
CCL	Cross-Correlation Lag
DCE	Delay Coordinate Embedding
DETR	Department for Environment, Transport and the Regions
DfT	Department for Transport
DoT	Department of Transportation
EPSRC	Engineering and Physical Sciences Research Council
FH	Forecast Horizon
GDP	Gross Domestic Product
HCM	Highway Capacity Manual
IHT	Institution of Highways and Transportation
ITE	Institute of Transportation Engineers
ITS	Intelligent Transport Systems
KARIMA	Kohonen optimised maps combined with ARIMA
LINSIG	LINcoln SIGnal

LE	Lyapunov Exponent
LOS	Level of Service
MAC	Media Access Control
MESSAGE	Mobile Environmental Sensing System Across Grid Environments
MIDAS	Motorway Incident Detection and Automatic Signalling
MOTION	Method for the Optimization of Traffic signals In On-line controlled Networks
MOVA	Microprocessor Optimised Vehicle Actuation
NEMA	National Electrical Manufacturers Association
NN	Neutral Networks
NO_2	Nitrogen Dioxide
NO	Nitrogen Oxide
ONS	Office of National Statistics
OSCADY	Optimised Signal CApacity and DelaY
PRODYN	PROgramming DYNamic
RHODES	Real-Time Hierarchical Optimized Distributed and Effective System
ROMANSE	Road MANagement System for Europe
SCATS	Sydney Co-ordinated Adaptive Traffic System
SCOOT	Split Cycle Offset Optimisation Technique
SDIC	Sensitive Dependence on Initial Conditions
TRANSYT	TRAffic Network StudY Tool
TRB	Transportation Research Board
TRL	Transport Research Laboratory
TRRL	Transport and Road Research Laboratory
UTC	Urban Traffic Control
UTM	Urban Traffic Management

UTMC	Urban Traffic Management and Control
UTOPIA	Urban Traffic OPtimisation by Integrated Automation
VA	Vehicle Actuation
WHO	World Health Organisation

Chapter 1. Introduction

1.1 Introduction

In most countries, especially those with well-developed or expanding economies, traffic congestion is widely recognised as a major problem given that it makes a significant contribution to global warming, environmental pollution and the depletion of fossil fuels (Manley and Cheng, 2010, O'Flaherty, 2005b, Papageorgiou *et al*., 2003). Evidence that the UK has one of the most congested networks in Europe was presented by Goodwin (2004) (see Figure 1.1). Traffic congestion causes substantial delays, especially in urban areas, which adversely affects journey times and the convenience of road travel in general. Congestion has a significant impact on local air quality with road transport being responsible for over 90% of domestic transport emissions, and the declaration of air quality management areas, in UK urban areas (DfT, 2011, Hitchcock *et al*., 2014). The estimated cost of congestion is £13-£20 billion per year to the UK economy (Centre for Economics and Business Research, 2014, Goodwin, 2004, The Smith Group, 1999). In the USA, vehicle miles has grown by nearly 500% since 1940, which currently costs approximately £72 billion per year (US Census Bureau, 2005).

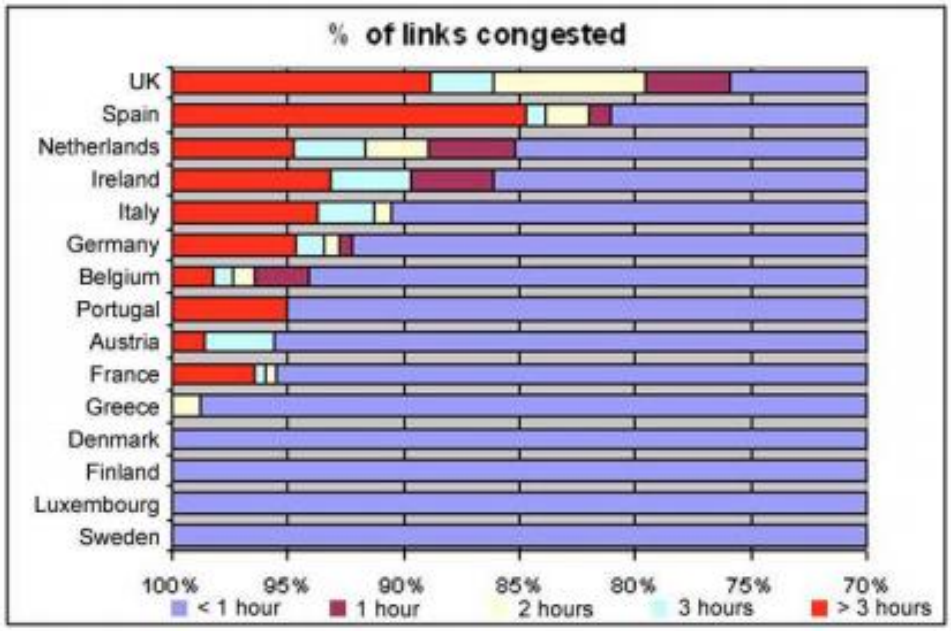


Figure 1.1: Comparison of Congested Links in European Countries
(Source:(Goodwin, 2004))

One of the root causes of this problem is the steady growth since the 1950s in vehicular traffic (with relatively less infrastructural provisions) that has had far-reaching consequences on both the economy and the environment. The average growth in road traffic in the UK is approximately 4% *per annum*, from 46.5 billion vehicle-kilometres in

1950 to 500 billion vehicle-kilometres in 2014 (DfT, 2015b, Goodwin, 2004). By comparison, road length only increased by about 0.5% *per annum* from 297,000 kilometres to 396,000 kilometres, even though road capacity has some additional increment due to localised junction or network improvements and carriageway widening (DfT, 2015a, Goodwin, 2004). Therefore, whilst the amount of road traffic has grown rapidly, road capacity has not expanded sufficiently to accommodate the additional demand. Building new roads to increase urban road capacity on a large enough scale is unlikely as this often tends to be unpopular, economically unaffordable and environmentally unsustainable (Papageorgiou *et al*., 2003).

Almost 50,000 deaths are brought forward annually in the UK as a result of air pollution from congested traffic (ONS, 2014). Further, approximately 3.7 million people die *per annum* across the world due to the poor quality of outdoor air (WHO, 2015). At a time when such externalities are now being a major public concern, increasing network capacity is inadvisable (Bigazzi and Figliozzi, 2013). Instead, networks should be managed to maintain traffic flows at levels that cause minimal risk to health, and alleviating congestion is a significant step towards meeting this goal. Consequently, the UK Government's policy is directed towards investing in carbon efficient transport systems and exploring alternatives including traffic management and by encouraging non-car modes of transport as a means of reducing congestion resulting from car-borne journeys (DETR, 1998, Douglas *et al*., 2011, Eddington, 2006, Hickman and Banister, 2007, Hickman *et al*., 2010).

1.2 UK Transport Policy (1998 – date)

The Kyoto Protocol committed the European Union and national governments to reduce carbon dioxide (CO_2) emissions by 60% in all sectors including transport by 2030 (Hickman and Banister, 2007). As reported, CO_2 is a leading source of emissions in the UK and a major contributor to the estimated 14% of all greenhouse emissions attributed to the transport sector (an estimated 12.9% is due to the road sector) (Stern, 2006). Therefore, a reduction of CO_2 emissions in the transport sector is key to meeting the strategic targets of the Kyoto Protocol. Since its adoption in 1997, UK transport policy has targeted strategies at minimising congestion and its adverse environmental impacts. In 1998, a Government white paper set out a new integrated transport policy to promote the appropriate legislation and transport policies that would ensure continued economic growth, alleviate congestion, and improve journey time reliability, as well as tackling climate change (DETR, 1998). In 2000, the Government published a 10 year investment

plan in surface transport, for the period up to 2010, incorporating strategies to tackle congestion and pollution by improving rail and road transport (Chatterjee and Gordon, 2006). In 2005, the UK Government commissioned the Stern Review and Eddington Transport Study to guide its policies on tackling climate change. The findings of the reports highlighted the adverse environmental impact of transport activities making the need for an urgent solution vital (Eddington, 2006, Stern, 2006). In agreement, Hickman *et al*. (2010) suggested that the current progress towards achieving the targets of the Kyoto Protocol is slow, and therefore argued for urgent action.

A major source of greenhouse emissions is the motor car, which is only second to freight transport contributing significantly to air pollution and CO_2 emissions. Recent advances in technology suggests that there is a vital role for fuel-efficient engines and alternative fuels towards significantly reducing emissions. However, there is concern about the feasibility of deploying these solutions globally given the limited time frame available (Moriarty and Honnery, 2008). Therefore, it is imperative to direct action at minimising car ownership and levels of use. The challenge is the strong correlation between transport and economic growth, and Douglas *et al*. (2011) comment that this relationship poses a problem and likely to offset the gains made through technology (Chapman, 2007). Therefore, it will be necessary to promote these technology-based solutions in parallel with strategies to minimise increasing car ownership and levels of use in order to boost efforts to achieve the environmental targets (Begg and Gray, 2004). One of the principal objectives whilst satisfying economic growth and tackling the negative impacts of car use, is to ensure that the road network operates safely and efficiently with minimal congestion.

According to the Eddington Study, approximately 89% of delay caused by congestion occurs in urban areas. It predicts that the situation is likely to grow worse because urban areas are highly productive centres that attract and generate significant numbers of journeys. If there are no interventions, then by 2015 an estimated 13% of vehicles could be subject to stop-start travel conditions. However, if interventions yielded only a 5% reduction in travel time across all business travel on the roads, this would result in about £2.5 billion of cost savings, which is equivalent to about 0.2% of UK Gross Domestic Product (Eddington, 2006). The monetary and environmental benefits of a free flowing network are immense, and make it necessary to implement the right interventions to alleviate congested urban networks and their catchments in order to enhance network

performance in terms of capacity, delays, reliability and sustain the productivity of the country.

Further, the DfT estimates a future growth for over 24 million UK commuters per day using the transport network, which could further stress the limited available capacity (DfT, 2004). In order to operate the network efficiently, current strategies for managing the network need an overhaul through appropriate policy interventions to promote new technologies, supported by behavioural and lifestyle changes, as well as significant investment in other alternative modes such as public transport and 'zero carbon' walking and cycling (Chapman, 2007, Tight *et al*., 2005). Ultimately, the practice of traffic management needs transforming by enhancing the ability of current systems to anticipate short-term network conditions in advance in order to mitigate appropriately and maintain under-saturated conditions. The focus of this research is to investigate a means of identifying the onset of congestion in real-time to enable proactive traffic management interventions (feed forward control) in order to prevent congestion occurring downstream at a later time. This will be more beneficial than the more traditional approach of operator intervention switching signal plans based on engineering judgement of traffic conditions over many months or years. In practice, these reactive remedial measures referred to as "feedback control'' are implemented at the time congestion occurs, which may be late and ineffective (see section 2.5.6).

1.3 Context of the Research

Current demand responsive signal control systems model, process and interpret real-time traffic data (such as volume, density and link occupancy) to carry out their basic operations in order to determine the type of control strategy to implement (Jianming *et al*., 2003). The bias in their approach, that also limits their functionality however, is that they are only able to implement solutions that tackle adverse network conditions as or after they begin to occur, rather than being pre-emptive of congestion events in order to implement mitigating strategies beforehand and therefore prevent the congestion occurrences. It is argued that, in order for UTC systems to be more effective, they must develop this capability to anticipate congestion events and deal with them in advance. Appropriate mitigation plans, when implemented ahead of congestion events, will improve urban traffic management and reduce congestion and associated emissions, but this can only be possible when the future conditions (in the order of minutes up to one hour) can be forecast spatially and temporally into the future with sufficient accuracy.

Given this background, it is important that modern UTC systems are equipped to forecast network events much more accurately, at least in the short term, as an enhanced benefit of ITS applications to urban traffic management (Han and Song, 2003). It is no surprise, therefore, that the development of traffic condition forecast models has become an area of growing interest among researchers and practitioners alike who are exploring ways to understand the dynamics of traffic flow in order to be able to simulate and forecast real time traffic regimes more accurately (Huang and Sadek, 2009, Kesting and Treiber, 2008, Qi and Ishak, 2014, Vlahogianni *et al.*, 2005). Traffic condition models are able to predict conditions within the coming hour using data from roadway sensors. The ability to predict traffic conditions in real-time makes these models much more attractive than traditional methods that can only make predictions based on (for example) annual growth factors. These are not only more susceptible to inaccurate projections due to their long forecast horizons but also of little value to day-to-day operation of traffic networks. The benefits of a short-term prediction horizon should not be underestimated. Cheslow *et al.* (1992) suggested that "*the ability to make and continuously update predictions of traffic flows and link times for several minutes into the future using real-time data is a major requirement for providing dynamic traffic control*." Indeed, the mere predictions of the flows are less beneficial without the ability to inform the future traffic condition.

In recent years, several traffic condition forecast models have been developed including time series, non-parametric, filtering and phase space models. Of these, the phase space model (based on Chaos Theory) seems to be the most promising for analysing the complexities in a perturbed traffic flow due to the several unpredictable and immeasurable quantities within road traffic systems. The majority of the traffic condition models (including time series analysis, neural networks, Kalman filter theory and regression analysis) tend to assume that the uncertainty of traffic flow is primarily due to the stochastic input into the models. However, modern systems theories prove that the uncertainty of traffic flow is not only due to the external stochastic input but also the internal non-linear structure of the system (Jianming *et al.*, 2003). This knowledge suggests that traffic conditions models can be improved by incorporating both the external stochastic process as well as the internal non-linear structure of the system into prediction models. Chaos Theory therefore emerges as a prime candidate for an enhanced study of short-term traffic conditions given its unique ability to analyse the underlying non-linear structure of evolving systems. Consequently, the overall approach to this research work is to use Chaos

Theory to understand the evolution process of the traffic condition (based on the Lyapunov exponents) through an analysis of time series measurements of traffic data variables.

1.4 Research Gap

Existing literature indicates that previous work that investigated Chaos Theory and the applications to road traffic congestion concentrated on simulated systems such as car following models or modelled gross traffic flows between different areas in a network (Kirby and Smith, 1991). By relying on data aggregated together, these studies ignored the scale of interaction among vehicles in the traffic stream. Consequently, the simulations may be artificially deterministic which can lead to the apparent presence of chaotic patterns where in the real world they do not exist. On the other hand, research work that used actual (as opposed to simulated) traffic flow mainly concentrated on simple systems (such as motorway traffic), whilst the traffic data was of low temporal resolution (Nair *et al*., 2001, Lan *et al*., 2003, Shang *et al*., 2007, Krese and Govekar, 2013). Given that the short-term fluctuation in this data is unaccounted for, there is high likelihood to affect the modelling outputs. This brings into perspective the need for further studies that overcome previous limitations. Due to recent developments, there are now new data sources that are able to provide high-resolution (temporal) traffic information (Shih *et al*., 2001). These sources include pervasive low cost sensors known as Motes and Bluetooth devices (Bell *et al*., 2009, Kay and Jackson, 2012) and sophisticated legacy systems such as Split Cycle Offset Optimisation Technique (SCOOT) (DfT, 1999) and Automatic Number Plate Recognition (ANPR) (Hounsell *et al*., 2009). Of particular relevance is the ability of these sources to provide data over a sufficiently high spatial density. This provides a new stimulus for further research into gaining a better understanding of chaotic traffic data to enhance the current automatic management of urban road traffic. Thus, the benefits of integrating diverse data sources within existing urban control systems are that they will serve a dual function of delivering the traffic management and control function as well as providing the data to enable Chaos Theory to report on the situational state of the urban network in real time (Downes *et al*., 2006). This could lead ultimately to a step change in the implementation of algorithms in existing UTC systems that trigger optimum control strategies that are one-step ahead of real-time traffic congestion, rather than being one-step behind. In order to meet this aspiration, the focus of this research is to examine a complex network, for example, a group of interconnected urban junctions, such as a SCOOT region, using highly resolved traffic data. In this sense, the potential economic benefits resulting

from a successful outcome of this research is substantial as the data are available at no extra cost.

1.5 Research Questions

From the research gap (see Section 1.4), the overarching questions to be answered by this research are:

1. What are the main limitations and which gaps are identifiable for improvement from previous studies for developing Chaos Theory as tool for improving urban traffic control?
2. Are traffic variables such as noise, flow and occupancy appropriate to yield results that would enhance the knowledge of applying Chaos Theory to traffic studies?
3. Is the choice of the temporal resolution of traffic data an important factor in using Chaos Theory to analyse traffic data?
4. Can Chaos Theory detect the onset of congestion within signalised urban networks under real world conditions?
5. Which chaos parameter can be developed as a predictive tool for forecasting the occupancy of links in a SCOOT controlled network region?
6. Is there a forecast time horizon that yields reliable predictions from the traffic model?
7. If so, can this information be used to inform real-time traffic management and control strategies that are 'one-step ahead' of current traffic congestion strategies?'

Answering these research questions will bring new insights that will be invaluable in understanding and managing traffic congestion in signalised urban networks. It may be possible to enact appropriate traffic control and management strategies before the onset of traffic conditions. A better management of congestion through this enhanced understanding will have considerable benefits in terms of improving the environment of our towns and cities by reducing the negative impacts of congestion (for example traffic delays, environmental pollution and safety).

1.6 Aims and Objectives

The overall research aim is to use Chaos Theory to identify a (set of) parameter(s) that enables a better understanding of the traffic evolution process, and then develop a short-term traffic prediction model, which uses this parameter to forecast the occupancy levels of a set of interconnected links in a SCOOT network.

The specific objectives of the research are, as follows:

1. To explore mathematical principles that support Chaos Theory, and apply to the relevant variable(s) that enable understanding of the traffic evolution process.
2. To identify a parameter(s) to characterise the data in terms of chaos and to develop this parameter as an indicator of the onset of congestion.
3. To determine whether chaotic behaviour on the road network occurs at the micro-scale (individual junctions) or across the entire network of inter-connected traffic zones, with respect to a chosen urban SCOOT network.
4. To investigate how network managers may use this information in the prediction and management of future traffic patterns in order to implement appropriate remedial measures in advance.

1.7 Contents of the Main Chapters

In order to present the findings of the work undertaken to achieve the stated aims and objectives, and to answer the research questions, this thesis has been organised into four main parts comprising a total of ten separate chapters summarised as follows:

Part I: A Review of Chaos Theory and the Practice of Congestion Management

Chapter 2 examines congestion and refers to the six levels of service (from A through to F) established by the US Highway Capacity Manual (HCM) to explain how congestion occurs in road networks. It provides a brief review of traffic flow theory in order to understand how congestion relates to key traffic flow variables (e.g. flow, headway, speed, density and occupancy). Based on the mathematical relationships linking congestion and traffic flow, it establishes that occupancy is the most reliable indicator of traffic congestion. Further, Chapter 2 provides an overview of urban traffic control strategies used in the management of congestion in urban areas. It identifies that current signal systems for managing traffic in urban areas lack a coordinated approach to detecting the spatial and temporal evolution of congestion across control regions within city networks. It argues that this inability severely inhibits these systems' ability to detect reliably, on a strategic level, the onset of congestion and implement effective preventative action, for example through timely adjustment of traffic signal settings. It concludes that the traffic system exhibits a time-dependent and non-linear property and therefore emerges as a prime candidate for examining the application of Chaos Theory in urban traffic control environments to improve the operational abilities of control systems for better management of congestion and pollution.

Chapter 3 establishes that the road traffic system is sensitive to initial conditions and examines a number of activities that makes it susceptible to chaos. By reviewing previous research work that has applied Chaos Theory in traffic studies, it draws on their limitations to inform the direction for future studies, thus making a justification for this research project. It argues that a lack of high temporal resolution data has limited previous investigations to relatively uncomplicated motorway and inter-urban networks, arguably where the associated problems of congestion creating vehicle emissions have a less severe impact (for example on public health and climate change) given that motorways tend to be through rural areas for a high proportion of their length. However, these limitations identified in earlier research can be overcome due to new data sources (such as Motes and Bluetooth devices) and sophisticated legacy data sources (such as SCOOT and ANPR) that are able to provide highly resolved (temporal) traffic data. Consequently, it concludes that these developments now provide a new stimulus for further research to gain a better understanding of chaotic traffic behaviour for application to managing urban road traffic.

Chapter 4 presents a critical review of Chaos Theory and its general applications. It looks into the technical meaning of chaos in relation to dynamic systems, and examines the key characteristics of chaotic systems by reference to the logistic curve and bifurcation diagram. It also reviews key studies in this regard and identifies that the discovery of the Lorenz Attractor as pivotal to developing Chaos Theory. Chapter 4 reviews Chaos Theory as applied generally to diverse subject disciplines (such as medicine, engineering and finance), using examples from the extant literature, and concludes that chaos and its application to transport deserves further investigation. Finally, Chapter 4 draws a distinction between dynamical systems with known or unknown equations of motion, and introduces the phase space reconstruction theory as relevant to this research because the evolutionary equations for the traffic system are unknown.

Part II: Mathematical Framework of Chaos Theory and Research Methodology

Chapter 5 presents the framework of Chaos Theory, which forms the main methodology for this work. It conducts a review of traffic condition models, identifies their weaknesses for short-term traffic predictions, and explains the suitability of phase space models (based on Chaos Theory) for the proposed research. More importantly, it presents the methodology for the theory of the phase space reconstruction, and the algorithms for estimating relevant parameters (e.g. delay time, embedding dimension, correlation dimension and the Lyapunov exponents). It uses the Lyapunov exponent as the parameter

for characterising the datasets in terms of chaos. Finally, it interprets the Lyapunov exponents with respect to the cyclical changes in traffic behaviour from free flow through to congestion.

Chapter 6 describes the overall methodology for this work. It presents a pseudo-code of the algorithms for the phase space reconstruction of the traffic data, and the calculation of delay time, embedding dimension, correlation dimension, and the Lyapunov exponents, which were written in the Microsoft C# programming language. Also, Chapter 6 explains the cross-correlation technique adopted to identify the temporal and spatial relationship between network links. Further, it describes the use of the Lyapunov exponents for the hourly short-term traffic prediction of the link occupancies, based on a multiple regression dynamic traffic forecast model.

Part III: A Case Study of Welford Road SCOOT Region, Leicester

Chapter 7 presents the results of a preliminary assessment for the algorithms of Chaos Theory using noise and SCOOT flow and occupancy data. The data were used to test-run the models in order to assess whether the necessary characteristics that identify chaotic behaviour can be identified in the road traffic system. The assessment indicates that chaotic behaviour is present in traffic noise data; however, it was difficult to identify any structure in the system's dynamic behaviour when the Lyapunov exponents derived from the noise data were analysed. On the other hand, it was possible to explain the observed patterns from the profiles of the Lyapunov exponents calculated from the SCOOT link flow and occupancy data, in terms of known traffic behaviour. Finally, an investigation of sampling frequencies of 1,10, 20, 30, 60 and 120 seconds for the SCOOT link data established a 20-second interval data as reasonable for the flow and occupancy data to undertake a detailed study of the chaotic properties in the road traffic system.

Next, Chapter 8 presents the results for the analysis of the chaotic nature of traffic travelling in a network of interconnected links in the Welford Road SCOOT region in Leicester. It identifies that chaos is not an isolated event, but has spatial effects across the network. It also illustrates how the Lyapunov exponents may signal the onset of changes in the dynamic behaviour of the system to prompt the emergence of congestion at a later time and different location in the network. It identified that a minimum of approximately 5-10% of links would indicate chaos during the busiest time of the network activities. Finally, it illustrates the use of the Lyapunov exponent for signalling the onset of congestion, in order

to trigger the implementation of preventative measures that are one-step *ahead* rather than lagging *behind* the saturated network conditions they are designed to resolve.

Chapter 9 presents the results of a multiple regression model that predicts the occupancy of the network links using the Lyapunov exponents. This assessment was undertaken using seven monthly cohorts from 1st January to 31st July 1999. Five assessment scenarios were modelled based on the forecast horizon, threshold lags and correlation coefficients between links in order to assess what predicting the link occupancy from either the occupancy data or Lyapunov exponent profile does. The forecast model was able to undertake the short-term traffic predictions of 15 minutes with a high level of reliability. The results therefore suggest that incorporating the forecast algorithms into existing systems would provide additional intelligence to these systems making them pre-emptive to anticipate congestion and therefore implement mitigating strategies in advance.

Part IV: Discussions, Conclusions and Recommendations

Chapter 10 presents the main conclusions of the study and makes recommendations for future research work. It discusses the results of the study in the context of the extant literature, and demonstrates that the computed lag times, embedding dimension, correlation dimensions, and Lyapunov exponents are consistent with the findings of previous studies. It argues from a traffic management perspective that the Lyapunov exponents are capable of reporting on the situational state of the network better than the traffic occupancy. The exponents are also used in a multiple regression model for forecasting the occupancy of the network links.

The thesis concludes by referring to the research questions enumerated in section 1.5 above. It confirms that Chaos Theory is data hungry and that the sampling frequency of the data is crucial to understanding chaotic behaviour, and traffic flow and occupancy data sampled at 20 seconds intervals is suitable for understanding chaotic traffic behaviour. Further, it identifies that the time series profiles of the Lyapunov exponent are useful for identifying the dynamic state of the individual network links. Furthermore, the Lyapunov exponents can be used to forecast link occupancy reliably.

In terms of study limitations, the investigation of traffic noise lacked detail because only data sampled at one-minute intervals were available. The thesis therefore recommends that the chaos properties in traffic noise requires further investigations through the use of data sampled at various sampling levels less than one minute. Further, the study of the SCOOT

network was limited to interpolated data, the work can be improved by the use of actual observed traffic flow and occupancy data at the required resolution. Future studies may build on this work by using micro-simulation models to simulate events in order to verify the theoretical chaos as hypothesised in this research.

1.8 Conclusions

In conclusion, this Chapter has sought to set out the scene for this thesis by looking at the research proposal in relation to the relevant transport policy, and identifying the appropriate aims and objectives and research questions. Next, Chapter 2 begins Part 1 with a literature review of traffic congestion in order to enhance our understanding of the fundamental mathematical concepts of congestion as well as the existing ways for managing congestion in urban areas.

PART I: A REVIEW OF CHAOS THEORY AND THE PRACTICE OF URBAN CONGESTION MANAGEMENT

Chapter 2. The Practice of Traffic Management in Signalised Urban Networks

2.1 Introduction

In order to discuss the relevance for this research work, this chapter seeks to establish that there are deficiencies in current methods for managing traffic congestion in signalised urban networks. The limitations may be due to either weaknesses in the technological systems making them unable to perform to their full potential or an over-utilisation of the system's capabilities, or both, thus hindering the deployment of optimum solutions. For the purposes of identifying the specific issues, Section 2.2 explores congestion in the context of road traffic, and Section 2.3 provides insight into the process of the traffic evolution from free flow through to a number of degrees of vehicular interaction to congested conditions. Section 2.4 describes the relevant traffic flow parameters, as well as their mathematical relationships, and how they may be used to quantify the amount of traffic flow in order to gauge the performance of a given section of road way. Section 2.5 presents a critical review of the practice of signal traffic control in urban areas with respect to the existing technology and identifies their limitations. It concludes that autonomic forms of traffic control and coordination will revolutionise urban traffic management and therefore the next generation technology will benefit from software that utilise Chaos Theory.

2.2 Traffic Congestion

Congestion is defined as *the impedance vehicles impose on each other on the road network in conditions of unstable equilibrium where the traffic volume exceeds the available capacity of the road space* (Collins, 1995, Goodwin, 2004). According to Weisbrod *et al*. (2003) traffic congestion is defined *"as a condition of traffic delay (i.e. when traffic flow is slowed below reasonable speeds) because the number of vehicles trying to use a road exceeds the design capacity of the traffic network to handle it"*. The main characteristics of congestion are slower speeds, longer journey times and increased vehicle queues. Under normal conditions when there is less traffic on the road, drivers are able to choose their own speeds; however, as traffic increases drivers become constrained by interactions with other vehicles and average speeds fall. As more vehicles are added to the network, vehicle speeds tend to be gradually decreased giving rise to a platoon of slowing vehicles, which may stop and start as traffic flow becomes unstable and congestion emerges. In extreme situations, significant queues occur and traffic remains at a standstill for long periods usually as the traffic demand approaches the capacity of the road. These characteristics are further

examined in Section 2.4.4 based on the speed-flow relationship in Figure 2.1. Dhingra and Gull (2011) view congestion as the product of the effects of interactions of vehicles within the traffic flow leading to disturbances such as stop-and-start traffic. Next, Section 2.3 reviews these interactive patterns as provided by the Highway Capacity Manual (HCM) (TRB, 1994).

2.3 Range of Traffic Conditions

The HCM provides six Levels of Service (LOS) for describing the quality of amenity provided by a given roadway, which are Levels A through to F. The characteristics of Level A are free-flow, low volume, high-speed and comfortable operating conditions, whilst Level F indicates forced-flow, stop and start, and uncomfortable operating conditions (O'Flaherty, 2005a).

At Level A, vehicles operate almost unimpeded with a greater ease of manoeuvrability within the traffic stream and the effects of minor incidents or breakdowns are unnoticed because there is sufficient capacity to absorb queues. This enables drivers to experience a high level of physical and psychological comfort (May, 2005). Traffic manoeuvres may be slightly restricted due to a slight depreciation in LOS at Level B; however, vehicles are still able to operate under reasonably free-flow conditions. Level C offers stable speeds to motorists providing for conditions that are still close to free-flow speeds. However, the ability to manoeuvre within the traffic stream can be noticeably restricted at this stage and lane changes require careful and attention from the driver. Queues become noticeable at bottleneck areas of the networks such as junctions, but there is still capacity to absorb minor traffic incidents.

At Level D, there is limited space to absorb disruptions making minor incidents prone to generate significant queues. There is a further restriction in the degree of manoeuvrability within the traffic stream, and speeds begin to decline as flow increases. At this stage, a small increase in flow has the tendency to cause a substantial deterioration in the road density. The lower boundary of Level E describes driving conditions at capacity, where a roadway becomes extremely unstable and there are virtually no useful gaps within the traffic stream. Consequently, vehicles in the traffic stream need to give way to admit those changing lane or entering the traffic stream from the side road. At the upper level E boundary, freedom to manoeuvre within the traffic stream is extremely limited and minor traffic incidents can cause significant queues. Level F describes forced or flow breakdown,

which occurs when the number of vehicles arriving at a given location or junction exceeds the capacity.

By reference to the levels of service, speed deteriorates sharply with increasing flow, therefore, the onset of congestion is somewhere between level D and lower level E boundary as speed declines sharply with increasing flow, where movement becomes severely restricted. In order to demonstrate how the traffic variables (e.g. flow, occupancy, time or space headway, and speed) relate to congestion, Section 2.4 provides a brief review of theory of traffic flow. It discusses the fundamental principles that apply to traffic models and the tools for analysing the operation of the road network.

2.4 Traffic Flow Theory

The purpose of traffic flow theory is to provide mathematical relationships to describe the interactions between vehicles, drivers and the infrastructure. This includes the highway system and the operational elements such as signal and control devices, signage and markings (Dhingra and Gull, 2011).

Traffic engineers use a number of parameters to determine the performance of a given road section or network, which are determined from on-site measurements of information relating to the traffic stream. These are normally randomly sampled as traffic streams are not uniform but vary over both space and time (Wardrop, 1952, May, 2005). These parameters give an indication of the utilisation of road space or nature of vehicular interactions through measurements such as:

1. rate of flow (vehicles per unit time);
2. density (vehicles per unit road length);
3. concentration (determined by density or occupancy);
4. space headway between vehicles (distance per vehicle);
5. time headway between vehicles (time per vehicle);
6. speed (distance per unit time); and
7. travel time over a known length of road.

Sections 2.4.1 through to 2.4.5 describe the procedure for measuring these on-road parameters and the derivation of the mathematical formulae for estimating their values. Even though this research is largely dependent on flow and occupancy data, all of the enumerated parameters (1-7) are discussed because of their link to congestion.

2.4.1 Concentration and Flow

Concentration and flow indicate the amount of vehicular traffic using a given section of a roadway, but the two parameters differ depending on how they are measured. Concentration refers to measurements over space at a point in time, whilst flow measures observations over time passing a point in space.

Concentration (in terms of density) denoted as k (veh/m) is expressed as the number of vehicles per unit space. It is measured from aerial photographs taken from cameras mounted on tall buildings or poles that record vehicular traffic over defined space at a fixed point in time. The concentration of a given length of road (x) containing some arbitrary number of vehicles (n_x) in one lane at a point in time is:

$$k = \frac{n_x}{x} \qquad \textit{Equation 2-1}$$

Flow (volume) denoted as q (veh/s) is the number of vehicles (n_t) per unit time measured from a specific location in one lane of a roadway for a given time period (t). The data source is normally from manual or pneumatic counters (or some automated sources such as magnetic loop detectors). In mathematical terms, the flow is given by the expression:

$$q = \frac{n_t}{t} \qquad \textit{Equation 2-2}$$

2.4.2 Headways

Headways describe the degree of separation between vehicles within a traffic stream. The two types of headways are space and time headways. Space headways express the gap between traffic stream vehicles in terms of distance, whilst time headway measures the separation in time intervals. They are important determinants of safety for vehicles in the traffic stream; for example, there is a high likelihood of a crash when vehicles travel close together, whilst pedestrians have minimal opportunities to cross the traffic stream with short headways. As vehicles travel much closer when there is (or in the build-up to) congestion, the headways also can give an indication of the onset of congestion.

For any two succeeding vehicles in a traffic stream, the space headway (or spacing), s, is the distance between the rear (or front) bumpers of the front and back vehicles at a point in time. The measurements are normally from aerial photographs. For a length of road (x), the sum of all the space headways is:

$$\sum^{n_x} s_i = x \qquad \textit{Equation 2-3}$$

The time headway (or headway), h, is the time interval between the rear (or front) bumpers of the front and succeeding vehicle crossing a given point in time. In mathematical terms, the sum of all the time headway in the period (t) over which flow is measured gives:

$$\sum^{n_t} h_i = t \qquad \textit{Equation 2-4}$$

By substituting Equation 2-3 in Equation 2-1 gives a reciprocal relationship between the concentration and average spacing (ŝ), which is expressed by Equation 2-5.

$$k = \frac{n_x}{x} = \frac{n_x}{\sum_1^{n_x} s_i} = \frac{1}{\hat{\mathrm{S}}} \qquad \textit{Equation 2-5}$$

Similarly, a reciprocal relationship between flow and the average time headway, ĥ, is given by substituting Equation 2-4 in Equation 2-2 as follows:

$$q = \frac{n_t}{t} = \frac{n_t}{\sum_1^{n_t} h_i} = \frac{1}{\hat{\mathrm{h}}} \qquad \textit{Equation 2-6}$$

2.4.3 Mean Speeds

The quality of the journeys of vehicles within a traffic stream are expressed by the time mean speed ($\bar{u}_t$) and space mean speed ($\bar{u}_s$). These are described here for later reference to the fundamental relationship linking density (occupancy), flow and speed. In order to distinguish between the two speeds, suppose that there are *n* vehicles within a traffic stream travelling at *n* different speeds, u_1 , u_2, ... u_n and the respective journey times are t_1, t_2, … t_n, where $\bar{t}$ is the mean journey time.

The time mean speed is the arithmetic mean of the speeds of all vehicles. In mathematical terms:

$$\bar{u}_t = \frac{\sum_1^n u_i}{n} \qquad \textit{Equation 2-7}$$

The space mean speed however calculates the average speed based on the average journey time of all the vehicles on the assumption that all vehicles in the traffic stream travelled equally a certain known distance, d_s. Consequently, the general expression for the space mean speed is given by the expression:

$$\bar{u}_s = \frac{d_s}{\frac{1}{n}\sum_1^n t_i} = \frac{d_s}{\bar{t}} \qquad \textit{Equation 2-8}$$

This formula includes only those vehicles that travelled the entire distance d_s; however, for long road stretches, it is likely some vehicles in the traffic stream will not complete the entire distance especially in congested traffic. Therefore, calculations based on Equation 2-8 may not always be correct, given its explicit assumption that all vehicles travelled an equal distance of road length (Hall, 1997). Therefore, the IHT (1976) provide a modified approach to address the lapse in the assumption. This takes into account the actual total distance travelled divided by the total time, which provides a more realistic method for estimating the space mean speed. This approach specifies an explicit rectangle on the space-time plane and accounts all travels within it irrespective of whether the vehicle partially or completely undertook the journey. Assuming that the distance travelled by each individual vehicle is given by d_i, and substituting t_i (the distance divided by speed) in Equation 2-8 gives:

$$\bar{u}_s = \frac{d_s}{\frac{1}{n}\sum_i \frac{d_i}{u_i}} \qquad \textit{Equation 2-9}$$

Under certain conditions, for example where all vehicles completed the entire journey, the expression of the mean speed is:

$$\bar{u}_s = \frac{d_s}{\frac{1}{n}\sum_i \frac{d_s}{u_i}} \qquad \textit{Equation 2-10}$$

Therefore, it may be possible to ignore the distance effects and estimate the space mean speed as equivalent to the harmonic mean of the individual vehicles within the traffic stream. Equation 2-11 is the harmonic mean speed at a point over time, for the condition where $\bar{u}_s$ is independent of the distance travelled.

$$\bar{u}_s = \frac{1}{\frac{1}{n}\sum_i \frac{1}{u_i}} \qquad \textit{Equation 2-11}$$

Harmonic mean speeds provide good estimates of $\bar{u}_s$ only for a given road section where vehicles are able to maintain relatively constant speeds across the entire road length. The calculation will be inaccurate where a high variability in speeds exists across the entire length of the road resulting in a difference between the harmonic mean of speeds and the conventional mean calculated using the space speed method (Wardrop, 1952, Lighthill and Whitham, 1955). The likelihood is that the measurements at a point will over-represent the fast moving vehicles and under-represent the slower moving ones, leading to higher than

the true average speeds. For this reason, the time mean speed tends to be generally higher than the space mean speed. Therefore, there is a general preference for the space mean speed, compared to the time mean speed, for calculations involving mean speeds.

The use of point measurements to estimate $\bar{u}_s$ is desirable, if only constant speeds are possible across a given roadway section. In essence, there is hardly any statistically significant difference between the two speeds under uniform free flow conditions, although there will be difference under congested conditions (Hall, 1997). According to Wardrop (1952) the relationship between time and space mean speed (where σ_δ is the standard deviation) is given by:

$$\bar{u}_t = \bar{u}_s + \frac{\sigma_\delta^2}{\bar{u}_s} \qquad \textit{Equation 2-12}$$

2.4.4 The Speed-Flow- Density Relationship

Based on the understanding of the traffic stream variables from the above discussions, the parameters are now pulled together to derive the relationship between flow, density and speed, for understanding traffic flow behaviour in a given roadway section.

Suppose that a traffic stream consists of k-vehicles per unit length of road travelling with an average speed of $\bar{u}_s$. Consider a set i of vehicles within the traffic stream of flow q_i, and speed u_i, and concentration k_i. From Equation 2-6, the time headway is $1/q_i$ which implies the distance travelled in this time (space headway) is u_i/q_i. Hence, the concentration (number of vehicles per unit length) according to Equation 2-5 at any instant is given by:

$$k_i = \frac{q_i}{u_i} \qquad \textit{Equation 2-13}$$

Deducing the overall speed based on the fractional shares of the total density, and substituting Equation 2-13 gives:

$$\bar{u}_s = \frac{\sum_1^n k_i u_i}{k} \qquad \textit{Equation 2-14}$$

$$\bar{u}_s = \frac{\sum_1^n q_i}{k} \qquad \textit{Equation 2-15}$$

$$\bar{u}_s = \frac{q}{k} \qquad \textit{Equation 2-16}$$

$$q = k\bar{u}_s$$ *Equation 2-17*

In traffic engineering, Equation 2-17 is known as the 'fundamental law of traffic flow'. Accordingly, traffic flow theory suggests that flow is the product of mean space speed and density, implying that flow is equal to zero when either one or both parameter(s) are zero. However, the maximum flow occurs at a critical combination of these two parameters. Traffic flow however is very low in traffic jams, and traffic density is very high (or maximum at the highest possible jam). Figure 2.1 presents a generalised relationship between speed, flow and density, which helps to gain a better understanding of the association between these parameters.

Generally, free flow speed (S_f) occurs during low traffic flow conditions; however, the maximum flow (V_m) occurs when the density reaches the critical density (D_0) and the optimum speed (S_0) decreases to approximately the half of the maximum speed $(0.5 * S_f)$. Beyond the critical density, the flow actually decreases, until the density reaches the jam density (D_j), where the flow becomes zero and all traffic is stationary. It is worthy to note further that traffic flow is also low under free flow conditions when traffic density is low whilst speeds are high.

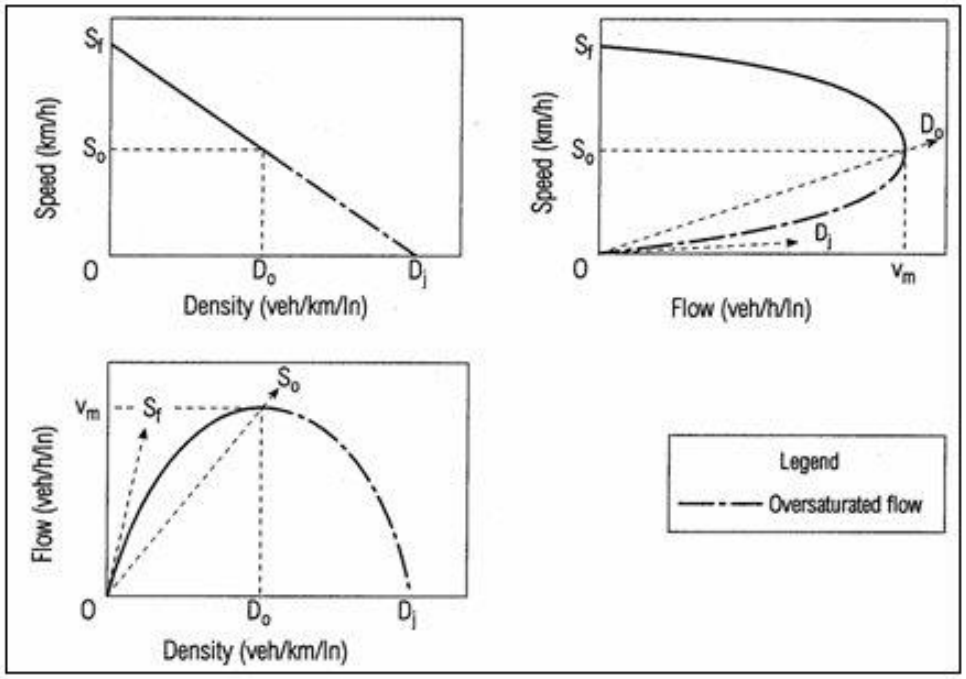


Figure 2.1: Generalised Relationships among Speed, Flow and Density
(US DOT, 2013)

The optimum density therefore defines the critical threshold for traffic congestion to start to emerge on a given roadway. Stable (or uncongested) traffic flow conditions exist when the density is below the critical density, whilst unstable conditions emerge and roadway capacity decreases close to or above this critical density. However, it is important that stable conditions exist to ensure an optimum throughput across a given road way or network of interconnected junctions. The role of network managers therefore is to maintain as much as possible the density below or close to the critical density in order to avoid congestion. This primary objective of traffic management is discussed in detail in Section 2.5. Chaos is characterised by a rapid deterioration in speed and density as flow increases. The relationship of Figure 2.1 to chaos is explained in Section 3.4.2.

It is worth pointing out that the fundamental law of traffic flow in Equation 2-17 is based on the assumption of a constant spacing and uniform speeds within the main traffic stream. Due to the assumption of constant spacing that is unlikely in unstable conditions, and uniform speeds in all sub-streams is invalid because drivers tend to choose their own speeds, the fundamental law may not be applicable for all the ranges of traffic conditions. Hall (1997) pointed out that the estimation of density explicitly from flow may be inaccurate, especially once congestion sets in. Further, the parabolic flow-density curve evaluates two solutions, one on either side of the critical density, for every flow value (except the maximum flow). The indication of a simultaneous congested and free flow conditions at the same flow makes unclear the actual density for a given flow. As a result, there is a preference for the occupancy measure based on the degree of passage over a loop detector on the roadway as a better indicator of the utilisation of a given road space. This is supported by the fact that traffic density is not measureable directly (see 2.4.1) but requires post-processing of measured data from aerial photographs of substantial road lengths, which is not always convenient. On the contrary, occupancy estimated from detector data is always readily available (Heydecker and Addision, 2011).

2.4.5 Occupancy

Occupancy (O) is the proportion of time that a fixed loop detector is occupied by a vehicle (or the proportion of time when the detector can detect the presence of vehicles) (Hazelton, 2004). It is estimated from the aggregated sum of the time that each individual vehicle covers the detector, divided by the observational period (T). In practice, the change in the inductance of a vehicle detector (such as an inductive loop) is polled regularly usually every 0.25 to 0.01 of a second to determine the presence or absence of a vehicle over the loop.

The loop detector outputs results in binary data format of the basic data of vehicles arrival and departure profiles including headways and the duration of every vehicle's occupation of the detector. This information is useful for estimating key statistics such as speed, flow and occupancy (Bell, 2005a). According to Heydecker and Addision (2011), a repeated sampling of the detector provides a time-base estimate of the proportion of the road occupied by vehicles, and consequently an indicator of the degree of utilisation of road space.

For example, suppose an inductive loop detector polled every 0.25 seconds yields the following binary data (where 1 and 0 indicate the loop is occupied and unoccupied respectively) in Figure 2.2.

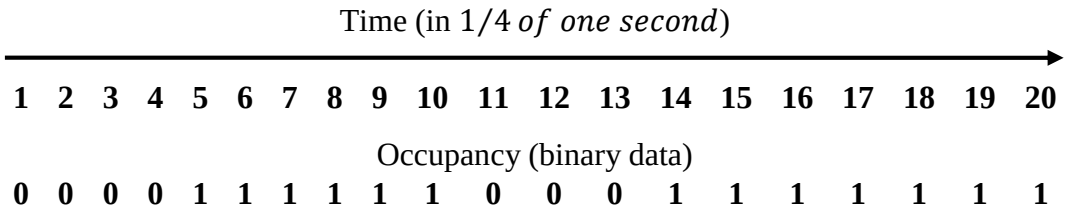


Figure 2.2: Binary Data from an Inductive Loop Detector

The binary information indicates that during the duration of observation of 5 seconds, the occupancy is 65% and the flow is two vehicles. According to the groups of '1' bits (6 and 7), the two vehicle types have different characteristics. The information also indicates that the advantage of occupancy as incorporating both the effects of vehicle length and compositions (car, bus, heavy goods vehicles, etc.), but density ignores these factors. Occupancy is therefore widely used in traffic management systems because it provides a better measurement of demand or the utilisation of road space.

The binary data is also useful for calculating the speed of vehicles across the loop. Assuming that the length of the loop is one metre and the first vehicle has length of 5 metres, then the effective length of the loop is 6 metres. The speed of the vehicle is 4 m/s, given that dwell time of the vehicle is 1.5 seconds (i.e. six of 0.25 seconds according to Figure 2.2). Similarly, the second vehicle's speed will be $6/(7 * 1/4)$ = 3.42 m/s, if both vehicles have the same length. On the other hand, if the second vehicle's length were 6 metres, then the speed of both vehicles will be 4 m/s.

Athol (1965) proved that occupancy is directly proportional to density, and inversely proportional to speed. Assuming an observation over a time T, each vehicle's dwell time on the detector can be determined from the vehicle speed (u_i) and the effective vehicle

length calculated as the length of vehicle (L_i) plus the length of the detector (d). The incorporation of the length of the detector is because the detector activates from the moment the front bumper starts to cross until the rear bumper crosses over the detector.

Accordingly, occupancy (O) is given by:

$$O = \frac{\sum_i (L_i + d)/u_i}{T} \qquad \textit{Equation 2-18}$$

$$O = \frac{1}{T}\sum_i \frac{L_i}{u_i} + \frac{d}{T}\sum_i \frac{1}{u_i} \qquad \textit{Equation 2-19}$$

By multiplying the secondary term in *Equation 2-19* by $N(1/N)$ and substituting *Equation 2-2* and *Equation 2-11*:

$$O = \frac{1}{T}\sum_i \frac{L_i}{u_i} + d.\frac{N}{T}.\frac{1}{N}\sum_i \frac{1}{u_i} \qquad \textit{Equation 2-20}$$

$$O = \frac{1}{T}\sum_i \frac{L_i}{u_i} + d.q.\frac{1}{\bar{u}_s} \qquad \textit{Equation 2-21}$$

$$O = \frac{1}{T}\sum_i \frac{L_i}{u_i} + d.\frac{q}{\bar{u}_s} \qquad \textit{Equation 2-22}$$

Substituting $k = q/u_s$:

$$O = \frac{1}{T}\sum_i \frac{L_i}{u_i} + d.k \qquad \textit{Equation 2-23}$$

$$O = \frac{\sum_i \frac{L_i}{u_i}}{T} + d.k \qquad \textit{Equation 2-24}$$

Introducing a factor of $\frac{1}{N}$ and substituting $T = \sum_i h_i$:

$$O = \frac{\frac{1}{N}\sum_i \frac{L_i}{u_i}}{\frac{1}{N}\sum_i h_i} + d.k \qquad \textit{Equation 2-25}$$

$$O = \frac{\frac{1}{N}\sum_i \frac{L_i}{u_i}}{\hat{h}} + d.k \qquad \textit{Equation 2-26}$$

Assuming a constant vehicle length, L, the equation is simplified as follows:

$$O = \frac{1}{\hat{h}}.L.\frac{1}{N}\sum_i \frac{1}{u_i} + d.\ k \qquad \textit{Equation 2-27}$$

Substituting *Equation 2-6* and *Equation 2-11*:

$$O = L.\frac{q}{\bar{u}_s} + d.k \qquad \textit{Equation 2-28}$$

$$O = L.k + d.k \qquad \textit{Equation 2-29}$$

$$O = (L + d).k \qquad \textit{Equation 2-30}$$

$$O = (L + d).k \qquad \textit{Equation 2-31}$$

$$O = c_k.k \qquad \textit{Equation 2-32}$$

Where: $c_k = L + d$ *Equation 2-33*

On the assumption that, L and d are constant it can be deduced that occupancy and density are constant multiples of each other. Substituting Equation 2-32 into Equation 2-16, an expression of the space mean speed and flow in terms of occupancy is:

$$\bar{u}_s = q.\frac{1}{k} \qquad \textit{Equation 2-34}$$

$$\bar{u}_s = \frac{q}{O}.c_k \qquad \textit{Equation 2-35}$$

$$q = \frac{\bar{u}_s * O}{c_k} \qquad \textit{Equation 2-36}$$

It must be pointed out that the effective vehicle length is sensitive to the traffic composition and conditions, therefore the assumption of uniform length may not be valid under all types of traffic flow conditions (Heydecker and Addision, 2011). Due to the advantage of occupancy over density, it may be reasonable to describe capacity in terms of occupancy multiplied by the speed than as flow expressed in terms of density multiplied by speed.

In concluding the discussion on traffic flow theory, it is important to note that there are a number of traffic stream variables that enable traffic engineers determine the efficiency and quantity of traffic flow on a given roadway including speed, flow and concentration (occupancy and density). The fundamental diagram identifies two main traffic states as stable and unstable (congested) flow conditions, which are regions below and above the critical density respectively. As it is desirable to maintain free flowing conditions across

road networks, a number of traffic control technologies are currently in use to achieve this objective. The *modus operandi* of traffic management is to stabilise traffic flow conditions to below optimal density levels by the use of appropriate technology. Section 2.5 presents a review of the practice of urban traffic control, and explores the potential incorporation of algorithms of Chaos Theory in urban control systems for the analysis of the spatial and temporal evolution of traffic flow in order to implement solutions in advance to tackle imminent congestion problems.

2.5 Urban Traffic Management

Urban Traffic Management is the use of a number of technology-supported applications such as urban traffic control systems and variable message signs to achieve an efficient management of the urban traffic and transportation system (Hounsell *et al*., 2009). The first type of traffic control technology was mainly fixed-time isolated traffic signals introduced primarily to prevent queuing and long delays on minor roads meeting major roads. Originally, signals at junctions did not include pedestrian phases as the safe passage and stoppage of pedestrians was not a priority. They operated with low intelligence compared to current systems; signal plans were fixed for periods (AM, PM and OFF peak) and 'burned-in' to the hardware. They lacked several attributes such as the capability to coordinate the operation of adjacent signals or adjusting signal settings in relation to actual traffic demand, which are embedded in systems today. Currently, traffic signals remain an integral part of Urban Traffic Control (UTC), and incorporates a number of advanced features that enhances their operation. In addition to managing conflicting movements, these systems enable the network to operate with minimal delays and therefore enhance the overall efficiency of the urban road network. Variable message signs have become useful in dealing with the previous weakness in the ability to control traffic over network areas by redistributing traffic over alternate routes to cope with incidents on the road network. Incorporating variable message signs in this way has equipped UTC systems with the ability to respond to the requirements of more complex traffic management and in the provision of traffic information for drivers. They provide mandatory, such as speed limits, or advisory information such as accident, congestion, delays, speed restriction, advisory routes, weather or environmental warnings of adverse road conditions etc. (IHT, 1997b). Local area traffic management has improved significantly by incorporating Vehicle Actuation (VA) and Microprocessor Optimised Vehicle Actuation (MOVA), but demand responsive control systems (e.g. SCOOT, SCATS, UTOPIA and PRODYN) emerges as

the most efficient form of control for area-wide network management. The discussion in subsequent sections focuses on these systems and their role in urban traffic management.

2.5.1 Traffic Signal Control

Traffic signal installations at intersections consist of a set of signal heads connected by cables to a signal controller and loop detectors that are either above or below ground. The signal controller may incorporate a microprocessor device that controls the signal indications via an on-site computer that runs software modules locally. As such, some signals are controlled by loops that put in automated demands handled by the on-board controller but they may also run on 'fixed time' plans. This will put the intersection into a cyclical routine, for example 120 seconds with 40 seconds given to each arm of the junction. These can be programmed to take into account demands on loops using a sort of for-if logic. Alternatively, the signal control is either activated remotely, manually requesting the plan over either mobile data or internet depending on how the intersection has been outfitted (varies on constraints of the signal's location), or in some cases via an automated system. This can be based on journey time thresholds determined for example from ANPR data, or simply an automatic change of signals at different times of the day according to a timetable.

The UK signal aspect is a 4-stage plan (green → amber → red → red and amber) with a fixed-time allocation for amber (3 seconds) and red-and-amber (2 seconds) signal aspects. Except for fixed-time signal plans, the red and green control aspects may be varied depending on the clearance times, stage minimum and maximum, and the overall cycle time. The minimum cycle time is given by:

$$c_m = \frac{0.9 * L}{(0.9 - Y)} \qquad \textit{Equation 2-37}$$

A well-known expression for a good cycle time obtained by Webster (1958) is:

$$c = \frac{(1.5 * L + 5)}{(1 - Y)} \qquad \textit{Equation 2-38}$$

where $L =$ total lost time, $Y = \sum_i q_i / s_i$, where q and s are the expected inflow and saturation flow respectively of the critical link within each stage sequence. Based on the estimated minimum cycle times, the stage duration is given by:

$$g_i = \frac{c_m * q_i}{0.9 * s_i}$$ *Equation 2-39*

Currently, the cycle time and signal plan for the on-street operation of the traffic signal is produced manually from off-line signal optimisation software such as LinSig (LINcoln SIGnals) and OSCADY (Optimised Signal CApacity and DelaY) for isolated junction and TRANSYT (TRAffic Network StudY Tool) for area signal control. However, the latest version of LINSIG has the capability to model junctions in a network. These programs can determine the stage durations as well as generate the optimal stages, stage sequence and transition plans (Bell, 2005a). This is the basic concept underlying the operation of traffic signals; however different types of controls offer enhanced benefits based on modifications to the algorithms. Isolated traffic signals may incorporate vehicle actuation or microprocessor optimised vehicle actuation types of control. Alternatively, a group of signals may be coordinated together by linking the operation of adjacent controllers, which may be operated together through a reactive (demand dependent or responsive) system.

2.5.2 Signal Control at Intersections (Uncoordinated Signals)

Early installations in the 1920s were isolated (uncoordinated) fixed-time traffic signals, where there is only a fixed amount of green time available irrespective of the number of vehicles waiting to be discharged through the junction. By operating different signal plans optimised for significant changes in traffic turning movements during the day, typically AM, PM and OFF peak, efficiency in operation can be introduced from day to day. However, due to not optimising the green splits in real-time, in response to hourly variations in flow and turning movements, excess green could be available when only a few vehicles are awaiting discharge and *vice versa*. This largely renders fixed-time signals unable to tackle the challenges and complexities of urban road congestion. In order to achieve efficiency in operation of these traffic signals, in the 1960s, 'intelligent' stand-alone (isolated) traffic signals such as Vehicle Actuation (VA) and subsequently Microprocessor Optimised Vehicle Actuation (MOVA) techniques were introduced (DfT, 1999).

VA systems operate in response to changes in vehicle demands and modify the start times and duration of phases via communication with upstream loop detectors. In VA systems, a particular phase runs for a minimum time, and subject to the demands on the detector, there is further extension of the green time until a pre-set maximum time. However, VA is prone to extend the green phase inefficiently, especially when there are long queues waiting at red signals, because consideration is not given to the opposing traffic streams. Moreover,

it is difficult to set the duration of maximum greens effectively. These factors can degrade the performance especially if traffic arrival patterns become platooned, which requires coordination of adjacent signals as well as the balance of flows in conflicting traffic streams (Simmonite, 2005). VA systems however can operate better when linked together; for example, a study in Virginia that compared an original uncoordinated VA and final VA junctions gave 30% reduction in journey times (Byungkyu and Chen, 2010).

MOVA generates its own signal timings cycle-by-cycle, which vary continuously with traffic conditions, and therefore are able to respond dynamically to variations in vehicle arrival rates (Meehan, 2003). Simmonite (2005) suggests that MOVA overcomes the problems of VA by implementing a strategy based on delay-and-stops minimisation (under uncongested situations) or capacity-maximisation (under congested conditions). During uncongested situations, MOVA seeks to disperse queue build-up during the red phase by carrying out a delay and stops minimising procedure every half second. If there are predicted benefits, there is extension of the green phase and the calculations repeated; otherwise, the signal changes to the next stage. In the congested mode, MOVA considers the overload of an approach in conjunction with the efficient use of green durations on all approaches, and determines the signal timings and durations that maximise the junction throughput under the prevailing traffic conditions.

Isolated vehicle actuated junctions offer greater flexibility when changing signal plans, compared with coordinated junctions, because the subsequent effects on neighbouring junctions is not considered. Although VA and MOVA are effective controls, they are generally limited to isolated junctions and generally not feasible at the network level, where it is desirable to coordinate the operation of neighbouring junctions, making them of limited use in urban traffic management. However, isolated junctions non-the-less create congestion problems which can spill back and affect signal coordinated regions and should be considered part of a strategic traffic management capability offered by Chaos Theory.

2.5.3 Signal Control Across Networks

Signal coordination is important in urban networks for the complete dispersion of platoons of vehicles between the upstream and downstream intersections, which helps to overcome the limitations of uncoordinated traffic signals. Often, the close proximity of neighbouring junctions in urban areas tends to inhibit traffic moving uninterruptedly across adjacent junctions. Traffic engineers are therefore reliant on signal coordination to establish offsets that enable a platoon (normally formed during the red phase and released during to the

green phase) to pass through a sequence of intersections without stopping. Isolated traffic signals including vehicle actuated systems that have coordinated junctions are often referred to Urban Traffic Control (UTC) systems.

The earliest signal coordination techniques for urban network traffic management were designed in the late 1950s to coordinate the operation of fixed-time signals (but included vehicle actuation systems by the early 1970s) on a route to achieve a *green-wave* along a set of connected links such as a radial network. In networks with across-road traffic linked with adjacent signal controlled junctions (whether chessboard or irregular typical of most UK towns and cities) optimisation of traffic across an area is required. The off-line model TRANSYT developed by the Road Research Laboratory (now the Transport Research Laboratory) is adopted worldwide to create fixed time plans across a region (Robertson, 1969). Currently, LINsig software also offers this capability. Signal junctions are included in the control region provided arrival patterns at stop-lines are platooned. Compared with the original signal plans, TRANSYT gave a reduction of 7.4% to 11.4% in journey times throughout the state of California (Skabardonis, 2001), whilst a trial in Glasgow indicated a 3% reduction in fuel consumption (Robertson, 1983). By optimising a performance index, which measures the aggregated cost of the stops and delays along the route, for example fixed-time signal plans are compiled for different times of the day (and for recurring events on the network) to handle varying traffic demand. The plans are created for periods (typically one to three hours) resulting in three daily fixed time plans namely morning (AM), inter (OFF) and evening (PM) peak hour plans of the day, where flows and routes through the network are similar. Signal coordination ensures minimised risks of significant queues between adjacent intersections, reduction in speed variability on major roads, and minimised stops at intersections, which ultimately minimises vehicular emissions that pollute the environment (Bell, 2005b). However, fixed-time signal coordinated control is unable to respond, in real-time, to within and hour by hour changes in traffic conditions, and is prone to causing lengthy delay of traffic on side roads and a lack of coordination in opposing traffic streams. Crucially, fixed time UTC whether *green-wave* or *area-wide* is not demand responsive, and therefore unable to handle the complex changes in arrival patterns of vehicles at a stop line, or respond to unplanned incidents such as traffic accidents and road works (Hamilton *et al.*, 2012). Due to these drawbacks, their effectiveness in controlling and managing traffic flow is limited to traffic conditions up to about 90% degree of saturation on network links. Such conditions are becoming less likely in urban networks, where the complex nature of the spatial and temporal evolution of

congestion across networks makes fixed-time signals inappropriate. Moreover, their signal plans (compiled for different times of the day and for recurring events on the network) degrade in effectiveness rapidly due to changes in traffic patterns, and require regular updating to be effective (Bell, 1983). The three main limitations of fixed-time area control include:

(i) the need to change from one plan to another can create 9% additional delay, if triggered at peak hour flows (Bretherton, 1979);

(ii) over time signal plans age by 3% per year (Bell, 1984); and

(iii) unless the variation in turning movements and flows remain below about 10%-15% throughout the peak period there is unnecessary delay (when flow transients are lower or higher than this tolerance) (Bell, 1984).

Given these limitations, demand responsive forms of control have become popular in the management of traffic in urban areas. These are equipped to handle the complex traffic flow better than with the traditional UTC systems.

2.5.4 Demand Responsive UTC Systems

In the 1970s, the (then) Transport and Road Research Laboratory (TRRL), in collaboration with suppliers of traffic systems in the UK (e.g. Plessey, Ferranti, GEC), developed the concept of demand responsive UTC systems to overcome the limitations of fixed-time UTC (DfT, 1999). Operating as an on-line traffic model, SCOOT (Split Cycle Offset Optimisation Technique) was launched commercially in 1983 (Bourner, 1984, Hunt *et al*., 1982). Since then, considerable investment in SCOOT has occurred to include bus priority, database facilities and incident detection (Bowen and Bretherton, 1996a, Bowen and Bretherton, 1996b), emission sensitive (Bretherton *et al*., 1998), and gating and optimal queue management (Bretherton *et al*., 2006). Other demand responsive control systems include SCATS (Sydney Co-ordinated Adaptive Traffic System) (Lowrie, 1982), UTOPIA (Urban Traffic Optimisation by Integrated Automation) (Donati *et al.*, 1984), RHODES (Real Time Hierarchical Optimised Distributed and Effective System) (Mirchandani and Head, 2001), PRODYN (Programming Dynamic) (Farges *et al.*, 1990), and MOTION (Method for the Optimization of Traffic signals In On-line controlled Networks) (Busch, 1996). As with SCOOT, these systems make use of detectors that monitor traffic flows continuously across the network in real-time and adjust the green splits, offsets, and cycle times to optimise their operations and reduce delays (Hamilton *et al*., 2013). However, the configuration of detectors and algorithms used to assess demand to modify signals timings are different. Systems typically monitor the network at the signal control region level

(typically 4-12 adjacent junctions), and undertake optimisation procedures to ensure the smooth flow of traffic across a particular section or area of the urban network. These demand responsive systems manage traffic much better than fixed-time systems, with an estimated reduction of approximately 12-20% (on average) in delays in SCOOT, for example, compared to fixed-time signals (DfT, 1999). By comparison, an installation of UTOPIA in Turin was able to achieve significant reduction in journey times of 20% for public transport and 10-15% for other vehicles (Papageorgiou *et al*., 2006); whilst MOTION demonstrated benefits of 30% reduction in traffic delays in Germany (Brilon and Wietholt, 2013). Turner *et al*. (1998) estimated that generally SCATS benefits are about 30% reduction in intersection delay and 19% increase in peak hour speeds, whilst Taylor and Abdel-Rahim (1998) reported up to 32% reduction in travel times in Michigan.

In recent years, an integrated system of UTC with other traffic management systems has made it possible to achieve improved efficiency in the operation of the urban road network. Examples include Urban Traffic Management and Control (UTMC) in the UK (DfT, 2009) and National Transportation Communications in ITS Protocol (NTCIP) in USA (AASHTO, ITE and NEMA, 2009). UTMC was an initiative of the Department for Transport to enable local authorities to obtain maximum benefits from their combined UTC systems and Intelligent Transport System (ITS) applications. The system's design allows different applications used within traffic management systems to communicate and share information with each other, which underpins its ability to provide a more dynamic, intelligent and real-time information system. In order to achieve this, national standards were established to facilitate the unification of initially separate UTC and ITS systems, which also enabled wider, and potentially cheaper, selection of products and services for local authorities. At the heart of the standardisation was the development of a common language to enhance communication among ITS features for an efficient means of sharing information. Consequently, the transport system could utilise other forms of communication such as cable TV to send information without paying the high cost of installations. This offered significant cost savings to operators, which encouraged competition in the mark*et al*ong with a wider choice of systems and suppliers, given the shift away from the costly practise of individual ownership of detectors and communication devices by individual ITS operators (Hamilton *et al*., 2013).

By combining SCOOT with Intelligent Transport Systems such as variable message systems, for example to redistribute traffic over alternative routes in the event of incidents

occurring on the network, UTC systems have led to significant improvements in urban traffic management. The proposed application of Chaos Theory to enhance congestion management uses data from a SCOOT system; therefore, more details are presented in the next section.

2.5.5 The SCOOT Model

Figure 2.3 shows the flow of information in the SCOOT system. SCOOT makes use of the traffic information from detectors on the upstream end of the approach link, for an online implementation of the traffic model based on the cyclic flow profile. These models use histograms to represent the arrival patterns of traffic, and the cyclic flow profiles indicate the average patterns of traffic flow during one cycle (IHT, 1997a).

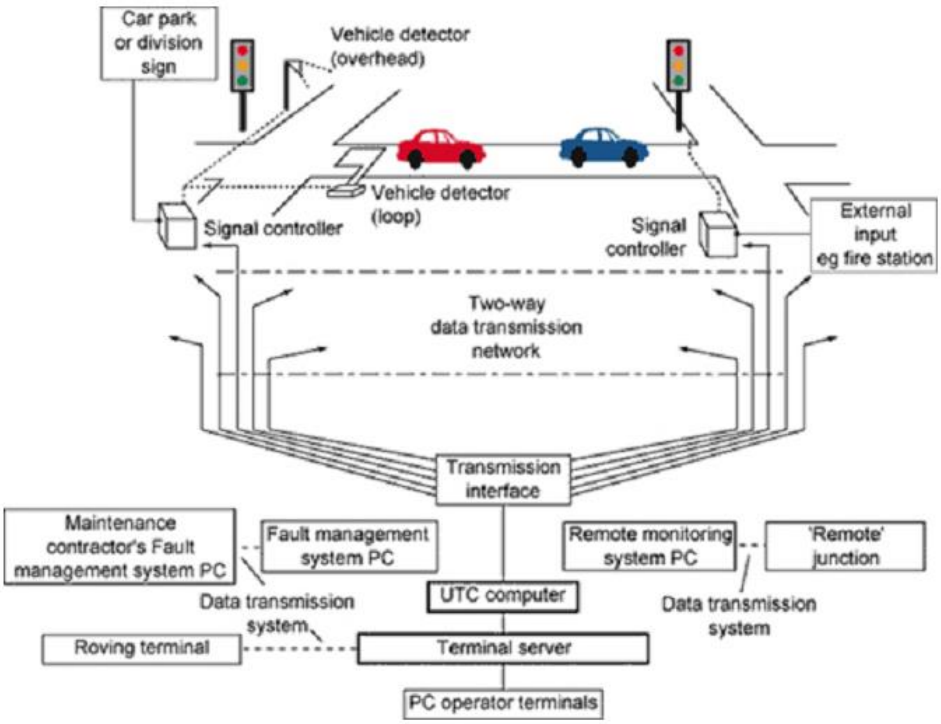


Figure 2.3: The Flow of Information in a SCOOT based Urban Traffic Control System
(Source: (DfT, 1999))

SCOOT uses the cyclic flow profile to undertake the three important optimisations i.e. split, offset and cycle time optimisation to make adjustments to the signal timings in order to improve traffic movement across the network. The split optimisation occurs at every stage change and results in increments/decrements of the green time by between one and four seconds in order to maintain a balance of degree of saturation for each stage at each intersection in the network. In order to maximise the performance index (traditionally a quantification of the delays and stops, but also fuel consumption or pollutant emission), the offset changes by increments/decrements by 4 seconds and occurs once every cycle for

each intersection. The cycle time optimiser operates every five minutes (or two and half minutes if the cycle time is rising rapidly) and increments/decrements the cycle time in steps of 4, 8 and 16 seconds in order to maintain the maximum degree of saturation at close to 90%.

By continuous implementation of the optimisation strategies, SCOOT ensures not only appropriate green waves along specific routes, but also minimises the delays across the network. The main drawback of SCOOT is the incomplete implementation of the 'hill-climbing' offset optimisation procedure making it unlikely to derive and apply the network-wide optimum offsets achievable in the model. In normal 'hill climbing' techniques, for example in TRANSYT, the optimum offset is informed by assessing the effects of a range of positive and negative offsets on improving the performance index. The optimum offset is the offset that makes the most significant improvement to the performance index. In SCOOT, the optimisation is incomplete as only a single step change is made normally to any split, offset or cycle time assessment to inform the implementation of a new set of signal timings. The traffic signal settings are therefore moved towards the optimum offsets suggested by the cyclic flow profile traffic model (Bell, 2005b). Further, the inflow profiles are measured by inductive loops, whilst in TRANSYT, they are calculated in order to achieve an equilibrium. Therefore, a deficiency is identified indicating an inherent weakness that limits exploiting the full capabilities of SCOOT.

In order to manage congestion, SCOOT uses queue data collected via detectors, which are located at the upstream end of the link. Stationary queues on the detectors are an excellent indication for potential exit blocking at the upstream junction. SCOOT Congestion (CONG) is *the proportion of the signal cycle that there is a queue over the detector*. A continuous occupancy of at least 4 seconds is taken to indicate a queue. Therefore, the percentage congestion is given by:

$$\%\,CONG = \frac{No.of\ congested\ intervals * 4 * 100}{cycle\ time\ in\ seconds} \qquad \textit{Equation 2-40}$$

The %CONG may be interpreted as follows: less than 1% "no congestion", 1-25% "some congestion", 25-50% "congestion", and more than 50% "severe congestion" (Siemens Plc, 1999). The %CONG is used in combination with a Congestion Importance Factor (CGIF), which is assigned to every SCOOT link. In normal operation, SCOOT depends on the optimisation procedures to minimise delays in the network. SCOOT runs the lowest cycle

time for the network to reduce delays and at higher cycle times to increase capacity, if needed. However, SCOOT evokes extra logic in order to prevent exit blocking and the potential for gridlocking when the optimisation includes congested links. In basic queue management, SCOOT acts to minimise queues on the links with higher congestion importance through an assessment in which the split optimiser makes use of the %CONG and CGIF to weight the calculation for the merit of the additional green request for the link. SCOOT also includes a congestion offset facility for handling congestion arising from inappropriate offsets, which are major causes of congestion in networks. The facility enables the traffic engineer to specify fixed offset for congested conditions and indicate their relative importance, which allows SCOOT to determine and apply automatically the best solution for the problem (Cowling *et al*., n.d.).

In more sophisticated queue management procedures, SCOOT deploys the gating facility that limits flow into sensitive areas of the network by relocating queues to less problematic areas. SCOOT MC3 (Managing Congestion, Communications and Control) includes a congestion supervisor that continuously monitors congestion in the SCOOT controlled network; identifies problematic links; diagnoses the cause(s) for the problems and recommends probable solutions to enable the user to make use of optimal facilities within SCOOT to manage congestion. Therefore, the facility offers the operator maximum benefits from the management facilities by providing a detailed overview of congestion in the network in order to enable the deployment of optimum solutions (Bretherton *et al*., 2006, Hay *et al*., n.d.).

2.5.6 Limitations and the Future of UTC Systems

In spite of SCOOT's success, the optimisation procedures are only effective within network regions in under-saturated and close to saturated traffic conditions, where queues build-up during the red phase and dissipate during the green phase or can be stored on longer links in the network by appropriate adjustment of offsets, cycle time or splits. In over-saturated traffic conditions however, they are unable to cope with surges and perturbations in the traffic flow or sudden changes in capacity, to which SCOOT's optimiser response can be lethargic. This is because the optimisation procedures are designed to make small incremental changes in timings to avoid inappropriate response to the formation of transient queues. In this way, signal plans evolve and regress to fixed–time *modus operandi* at network saturation. Furthermore, UTC systems are unable to control and coordinate traffic flow between signalised regions or isolated junctions that make up the urban road network at those times in the day when appropriate to do so. Due to the inability to manage traffic

between regions, traffic movement into an already saturated region continues unabated even at those times when congestion begins to develop as detected by the demand responsive control system and existing adverse traffic conditions are made worse. If the ability to control the amount of inflow into network regions existed, then it would be possible to manage these inbound flows upstream, and perhaps more relevant is to better manage traffic flows on the outskirts of a town or city, to ensure that network regions remain in under-saturated conditions at a time in the future.

In practice, all current demand responsive signal control systems analyse and process real-time traffic data (such as volume, density, link occupancy) to carry out their basic operations in order to determine the type of control strategy to implement (Jianming *et al.*, 2003). The bias in their approach, that also limits their functionality, is that they are only able to implement solutions that tackle at a local scale adverse network conditions as, when or after they occur, rather than being pre-emptive of congestion events likely to occur downstream in the network in the future in order to implement mitigating strategies beforehand. Moreover, their reliance on historic (latest) data to determine the delay or control strategy to implement, in real-time traffic management, may be problematic because when real-time flow varies appreciably they become ineffective.

In relation to this, Bell (2005b) argues that it will be reasonable to link the decisions affecting the control of traffic signals to a *feed forward* rather than *feedbac*k form of control. A *feed forward control* makes decisions affecting the control of traffic signals based on the predicted effects on performance taking into consideration the fact that lags are involved in the traffic process; for example, queues take some time to form and dissipate. On the contrary, *feedback control* makes decisions based only on the current situation, which may be liable to achieve a significant loss in efficiency and give rise to instabilities. Given that the traffic evolution process is characterised by natural lags makes a *feed forward* form of control a viable option. In this option, the decisions for traffic control are dependent on the predicted effects of the network performance. This implies that UTC system will be more effective by developing the capability to anticipate congestion events and deal with them in advance. Appropriate mitigation plans, when implemented ahead of congestion events, will improve urban traffic management and reduce congestion and associated emissions. This is achievable only when traffic conditions can be forecast spatially and temporally into the future with sufficient reliability. This will lead to improved journey times and reduce traffic delays, and travel time savings will accrue to economic benefits by taking

appropriate action. As road traffic is a significant contributor of environmental pollutants, a means of reducing emissions will be a significant step forward in tackling climate change and in line with the UK Government's objective of reducing CO_2 emissions by 60% by 2050 over 1990 levels (DfT, 2004), currently superseded by a mandatory target of an 80% reduction by 2050 (Ekins *et al*., 2011).

Given this background, it is important that modern UTC systems are equipped to forecast network congestion events much more accurately, at least in the short term, as an enhanced benefit of ITS applications to urban traffic management (Han and Song, 2003). Consequently, the area of traffic condition forecasting has become of much interest to researchers and practitioners alike. In recent years, several traffic condition forecast models have been developed including time series, non-parametric, filtering and phase state (Huang and Sadek, 2009, Kesting and Treiber, 2008, Qi and Ishak, 2014, Vlahogianni *et al.*, 2005). These models are discussed in detail in Section 3.3. The majority of these methods respond to measured variability in traffic and rely on short-term (cycle-by-cycle) predictions of traffic behaviour. The problem is the limitation that the uncertainty in traffic flow is small and primarily due to the short-term external stochastic behaviour into the models, without considering the uncertainty due to the internal non-linear structure of the system thus limiting the usefulness of these forecasting models (Jianming *et al.*, 2003). The use of Chaos Theory to analyse the uncertainties due to the internal non-linearity in dynamic systems opens up an opportunity to examine its potential for improving the short-term forecast of traffic flow.

2.6 Conclusions

The adverse impacts of congestion are enormous. They include unreliable journey times, unacceptable delays and the misuse of otherwise profitable working time in traffic due to significant loss of roadway capacity. The social, economic and environmental benefits of an uncongested network are immense, which justifies intervention that alleviates congested urban networks and their catchments to enhance network performance in terms of more capacity, less delay, improved reliability as well as sustained productivity.

Demand responsive control systems combined with ITS technologies are currently the most sophisticated for managing traffic in UK urban areas. However, these systems are limited due to for example SCOOT's incomplete 'hill-climbing' procedure, and its ineffectiveness in saturated traffic conditions making it to regress to the fixed-time *modus operandi*. Further, demand responsive systems for managing traffic in urban areas lack a coordinated

approach to detecting the spatial and temporal evolution of congestion across signal control regions within city networks. This severely inhibits these systems' ability to detect reliably, on a strategic level, the onset of congestion and implement effective preventative action, for example through timely adjustment of traffic signal settings to hold queues upstream or reduce traffic volumes.

In order to improve the operational efficiency of these systems therefore, it is important to equip them with enhanced facilities so that they are able to detect the onset of congestion and mitigate imminent flows that lead to adverse congested conditions accordingly to maintain network conditions at close to capacity but not in saturated conditions. For example, they must be equipped with sophisticated traffic assessment tools that will enable them to forecast short-term traffic conditions more effectively. As argued by Bell (2005b), a *feed forward* form of control is reasonable as it will link the decisions affecting the control of traffic signals to predicted effects on performance.

Next, Chapter 3 argues that traffic is a recognised time-dependent and non-linear system and prone to chaos, which creates the opportunity to apply the principles of Chaos Theory to inform the management of urban traffic control systems to reduce congestion and vehicle emissions. It presents a review of the existing literature on the application of Chaos Theory to traffic studies and identifies that highly resolved data is a primary requirement for future studies. It concludes that such data has recently become available due to improved data collection and processing methods, and therefore these developments now provide a new stimulus for further research into gaining a better understanding of chaotic behaviour of traffic for application to managing urban road traffic.

Chapter 3. Chaos Theory and its Potential Role in Traffic Studies

3.1 Introduction

In order to apply Chaos Theory to traffic studies, it is important to establish firstly the chaotic properties that qualify the traffic system as an ideal candidate for such an investigation. This chapter is focussed primarily on this objective, as well as presenting a review of previous research to identify specific issues that need to be addressed and whether Chaos Theory has the properties to overcome the limitations.

This chapter establishes in Section 3.2 that a road traffic network is essentially a non-linear system. Section 3.3 conducts a review of traffic condition models and concludes that Chaos Theory has greater ability to evaluate the non-linearity in traffic flow. Consequently, Section 3.4 identifies those factors that influence the evolution of chaos in the traffic network and the conditions under which chaos can emerge. Section 3.5 reviews previous applications of Chaos Theory to traffic analysis and identifies their limitations, whilst Section 3.6 concludes that these studies have focussed mainly on simple systems such as motorways, or those with low temporal resolution data, or both. Section 3.7 identifies that emerging and existing legacy data sources have the capacity to overcome the data issues of earlier studies, and discusses some computing issues relevant for an efficient implementation of Chaos Theory. Next, Section 3.8 introduces autonomic computing, which is pivotal for the self-management of an efficient transport system, focussing on its four main components namely *self-configuration*, *self-optimisation*, *self-healing* and *self-protection*. Finally, Section 3.9 draws conclusions and makes recommendations for applying Chaos Theory in the future to manage traffic in urban areas in order to reduce problems of congestion and atmospheric pollution, through the incorporation of new chaos-based algorithms within existing UTC systems.

3.2 Non-Linear Characteristics in Traffic Flow

Traffic behaviour gives rise to nonlinear characteristics that makes traffic systems difficult to forecast network-wide events. UTC systems for example only have a snapshot of the traffic state and therefore can model and forecast traffic flows on a link-by-link basis but not strategically network-wide (Cameron and Duncan, 1996, Frazier and Kockelman, 2004, Papageorgiou *et al*., 2003, Mahnke *et al*., 2005, Nair *et al*., 2001, Safanov *et al*., 2002, Hamilton *et al*., 2013). Given the variability in traffic speeds, and vehicles not driving at the speed limit, the amount of traffic flow or occupancy over a loop detector at fixed

intervals over a certain period tend to fluctuate suggesting a non-linear character. Fixed-time traffic signals are unable to cope with these non-linear fluctuations in the traffic system, which then becomes a recipe for chaotic behaviour to develop (Frazier and Kockelman, 2004). Chaos, amongst other reasons, is due to the combined effect of regular and randomly occurring events such as buses stopping, pedestrians crossing, vehicles parking and unparking at the roadside at which point UTMC traffic models are challenged (see Section 3.4.1). This results in traffic signals discharging only a portion of an entire queue length during the green phase in each cycle, which, over time, can lead to severe queuing. At the network level, these activities result in an exponential increase in queues occurring at a rate that is too fast for demand responsive control to react, or the capacity of the junction being exceeded or the ability of the link to store the queues thus causing exit blocking. At this stage, a saturated link cannot accommodate any additional traffic from the upstream link when given the right of way causing further congestion that rapidly develops upstream. As a result, such activities generate chaos on specific links or at junctions in a network on a day-to-day basis and opens up the potential for the technique to identify patterns spatially and temporally in network-wide traffic flows that lead to the congested state.

It must be noted that even though non-linearity is a condition for chaos, chaos is not a property of all non-linear systems; hence the presence of non-linearity does not always guarantee chaos (Gleick, 1988). In traffic studies however, it has been established that non-linearity does generate chaos in the traffic system (Lan *et al.*, 2003, Narh *et al.*, 2014, Shang *et al.*, 2007). The existence of a non-linear structure clearly makes traffic a potential candidate for the investigation of the use of Chaos Theory, which helps to explain the evolution of chaotic systems through a measure of the rate of change of a system known as the Lyapunov exponent. By definition, a positive exponent indicates chaos (instability), a negative exponent indicates a free flowing or congested network (stability) whilst a 'zero' exponent indicates a network in a meta-stable state (Károlyi *et al.*, 2010). In the traffic system, the Lyapunov exponents may be considered as an indicator of the lengthening and shortening of the gap between vehicles in a phase space. Therefore, the Lyapunov exponent may be viewed as an indication of the nature of vehicular interactions by measuring how well the system is able to adapt to perturbations, which reflects available capacity. It is important therefore to understand the evolution of traffic congestion, and its relationship to the Lyapunov exponents, to enable the state of the urban network to be identified. By reference to Figure 2.1, free flow exists when traffic density is low, which is usually during

off-peak periods. As such, asymptotic stable conditions (negative exponents) prevail at off-peak times when mean vehicular speed exceeds the optimal capacity speed, and density is below the safe capacity density. However, traffic density increases as the flow rate increases (as more vehicles join the network) but speed tends to decrease (at an increased rate as flow approaches capacity), thus giving rise to a steady stable condition at synchronised flow where the system exhibits Lyapunov stability at the critical density (or optimum speed). As congestion begins to emerge, the dynamic condition is unstable "chaos" (positive exponents) characterised by greater interaction due to the high flows and lower speeds. Queues develop rapidly at an exponential rate until maximum flow occurs at capacity, where any additional traffic or slight loss of capacity results in congestion (negative exponents). Chaos is characterised by rapid deterioration in speeds and density as flow increases. In extreme conditions, traffic density tends to its maximum when flow rate and vehicular speeds are zero and stationary queues are common. This is further explained in Chapter 8 (Figure 8.10) with a numerical example. The changing conditions within a traffic system reveal a non-linear structure that exposes chaos as a precursor to congestion. By observing the time series profile of the Lyapunov exponents, the turning points can reveal the 'shoulders of congestion' as a signal of impending congestion. A gradient change of the profile from positive to negative reveals a state transition from free flow to unstable flow or chaos, whereas negative to positive change reveals a transition from congestion back to free flow.

3.3 Review of Traffic Condition Models

This section conducts a brief review of the main theoretical approaches in traffic condition modelling in order to identify their uses and limitations for forecasting the traffic system. Based on existing literature, this section concludes that Chaos Theory has the potential to overcome the weaknesses in the reviewed models and to provide a better understanding of the evolution of traffic flow. Five main groups of traffic condition forecast models have emerged in recent years. These are:

Group 1: time series models including autoregressive, integrated and moving average models combining to produce models such as Box-Jenkins, AutoRegressive Moving Average (ARMA), AutoRegressive Integrated Moving Average (ARIMA) and AutoRegressive Conditional Heteroscedastic (ARCH);

Group 2: non-parametric models (e.g. artificial Neural Networks, Regression, Kernel Smoothing and k-Nearest Neighbour);

Group 3: filtering models (e.g. Kalman filter theory, generalised linear models, Bayesian dynamic linear model and least mean square filters);

Group 4: hybrid models (e.g. KARIMA (combining Kohonen maps with ARIMA), ATHENA, Bayesian Neural Networks and Hybrid Fuzzy rule-based models); and

Group 5: phase state models (e.g. Space-Time ARIMA, Bayesian Vector AutoRegressive, and Seemingly Unrelated Regression).

Qi and Ishak (2014) suggested that the majority of models in Group 1, 2 and 3 are constructed to fit data. This can result in the more general issue of 'over-fitting', a problem that reduces a model's ability to capture generalised trends. More specifically, Group 1 time series models (such as ARIMA and Box-Jenkins) assume that the data being analysed are a product of a linear process making them largely inappropriate if the underlying mechanism is non-linear (Zhang *et al.*, 1998). Also ARIMA models concentrate naturally on modelling mean values to the detriment of forecasting extreme values, which makes such models unsuitable for predicting events such as congestion where the associated traffic flows are significantly different from the mean flow (Abdullah *et al.*, 1999, Angeline *et al.*, 1994, Head, 1995). Also in Group 1, non-linear time series models (such as the bi-linear model, the Threshold AutoRegressive (TAR) and ARCH models) are limited as they hypothesise explicit relationships within the data series but with little knowledge of the underlying laws determining the dynamic process. These models are also difficult to formulate due to the many non-linear patterns within traffic datasets, suggesting that a generalised model may not capture adequately all the important features including the evolution of traffic congestion (Zhang *et al.*, 1998). In contrast, non-parametric models (Group 2) make no assumptions about the underlying traffic dynamics. They do not necessarily require training of the data to calibrate the parameters before applying to forecasting, but rely on existing relationship within the data. They have greater capability than Group 1 models to handle non-linear relationships given their ability to capture the underlying relationships without the need for prior assumptions, and to learn from data should the underlying relationships be unclear (Guo *et al.*, 2014, Huang and Sadek, 2009, Zhang *et al.*, 1998). Vlahogianni *et al.* (2005) suggested that these characteristics may indicate an inherent (theoretical) ability of neural networks to model the temporal and spatial variability in traffic flow. However, Khaleghei Ghosheh Balagh *et al.* (2014) argued that Neural Network models are not capable of modelling explicitly the serial correlation in time series data, which is inherent in the build-up of congestion. Also, Neural Network

models suffer from the over-fitting issue and hence the smoothing of flow variability when profiles are changing rapidly which affects the accuracy of congestion forecasts (Qi and Ishak, 2014). If the traffic data does behave as a linear static process, for example as can happen during off-peak periods when traffic is generally free-flowing, linear models such as ARIMA tend to perform better than the nearest neighbour and neural network models (Huang and Sadek, 2009, Zhang *et al.*, 1998).

In order to improve the performance of these standard forecasting approaches, hybrid models (Group 4) have been developed (for example, Guo *et al*. (2014), Huang and Sadek (2009) and Vlahogianni *et al.* (2005)). Although these models do generally perform better than the standard approaches, they are still restricted in their applicability to congested traffic situations due to inherent weaknesses in replicating, in the modelling process, traffic's non-linear behaviour. In order to overcome the various limitations in the models described above (Groups 1-4), phase state models for example based on Chaos Theory have emerged. A phase state model is a mathematical model of a physical system as a set of inputs, outputs and variables related by a first order differential equation, which is an equation of the unknown function and its derivatives and variables. They provide a convenient method to model and analyse systems with multiple inputs and outputs by expressing them as "vectors" (see Section 5.2).

Chaos Theory has the potential for analysing the complexities in perturbing traffic flow, which are due to the various unpredictable and immeasurable quantities within road traffic systems (Jianming *et al.*, 2003). Frazier and Kockelman (2004) identified three major benefits of Chaos Theory for system analysis:

1. It applies to highly nonlinear systems.
2. It naturally accounts for all important system dynamics.
3. It uncovers system information and relationships without having to uncover the laws or equations of the underlying dynamics.

In order to apply Chaos Theory to traffic studies, it is important to understand the non-linear characteristics of traffic flow and establish the conditions under which chaos occurs and propagates through the urban road network.

3.4 Chaos in the Road Traffic Network

The ability of Chaos Theory to analyse uncertainties due to the internal non-linearity of systems makes it appropriate for the study of congestion. The next question is to establish

whether chaos is present in urban road traffic, and, if so, under what circumstances does it occur?

3.4.1 How Does the Evolution of Road Traffic Generate a Chaotic System?

Due to its dynamic nature, traffic fluctuates between free flow and congestion, but any future state of the network cannot be determined accurately. Primarily, this is due to changes in the initial conditions in the evolution process. This is caused by the varied mix of human behaviours, which is partly a function of environmental factors (see section 3.4.2), and partly due to the heterogeneity in the driver population (including personality mix of males and females of a wide age range), which gives rise to varied reaction times consequential of different driving styles (Dendrinos, 1994, Lee *et al.*, 2014). Many other events can affect the initial condition of the traffic system. These include the movements of emergency vehicles (ambulance, police, fire etc.) requiring right of access, the rippling effect of the acceleration/deceleration of successive vehicles and transmission of 'stop-and-start' conditions, and actions of drivers such as stalling the engine, weather, or other environmental conditions. The other factors include temporary diversions, road closures, changes in road layout and road works, and the effects of network incidents (such as accidents, vehicle breakdowns and surges of vehicular or pedestrian volumes).

Safanov *et al.* (2002) attributed the chaotic behaviour in traffic flow to delays in human reactions. Others studies highlighted this chaotic behaviour, which they also identified as posing challenges to forecasting accurately events on the road transport network. Zhu *et al.* (2010) attributed the randomness and uncertainty in traditional transport models to some "mass random" phenomena, but failed to explain the main causes of the underlying dynamics. On the other hand, Xu and Gao (2008) noted the varying physical and human elements as being responsible for the formation and dissipation of a traffic jam, the effect of traffic incidents and the flow evolution. According to Dendrinos (1994), these complexities develop the chaotic state of the road network, which makes accurate long-term predictions difficult.

Road traffic networks can be prone to sudden gridlocking even when only minor events occur. A small change in a single driver's behaviour, for example when brakes are applied, may affect network conditions, especially when the traffic volume is increasing and unstable conditions are beginning to emerge. In other situations, fluctuations in speed or the spacing between vehicles, car parking and unparking or pedestrians jay walking even when traffic density is much lower than the jam density, may result in spontaneous traffic

congestion. Hence, the 'butterfly effect' (see Section 4.3) can be observed due to an infinitesimal change in an initial condition at some location in the network having a significant effect on the long-term state of the system at some other time or place. This has tendency to produce a serious adverse impact on the predictability and hence 'forecast-ability' of the state of the road network (Kirby and Smith, 1991).

Another reason why traffic can be prone to chaos is attributable to the similarity between fluid particles and traffic streams. In practice, the presence of chaotic properties in a deterministic dynamic system (such as fluid particles), can be analysed by the transformation of the time series of the key variables of the system evolution into a phase space. Fluid systems in scientific experiments analysed in this way indicate that particles released in the flow tend to deviate from each other over time, with the distances between pairs of particles growing at an exponential rate, which is normally characterised by the Lyapunov exponents (Károlyi *et al.*, 2010). Given the resemblance between the behaviour of traffic streams and fluid particles, the principles applicable to fluids may also be relevant to streams of vehicles. This similarity can be observed, for example, if one focuses on the waves that ripple through the network as vehicles brake or accelerate in succession (Sipress, 1999). The transmission of 'stop-and-go' movement through a road network, as vehicles change speed, is comparable to the shock waves released when a mass of gas fluid hits a bottleneck and transmits a wave of compressed air back in the opposite direction to the stream of oncoming gas (Sessler, 2007). Therefore, as in fluid systems, the rippling effect of the acceleration/deceleration of successive vehicles and transmission of 'stop-and-go' can potentially generate or exacerbate the chaotic state of the road network.

Furthermore, the occurrence of unstable situations on road networks, as characterised by traffic jams and dynamic changes, is comparable to the change of state of fluid particles from the solid through the liquid to the gaseous state and *vice versa*, which is explained with respect to the levels of service. The degree of manoeuvrability of vehicles (particles) in the road (fluid) system can be described by six levels of service (see Section 2.3) indicating the degrees of interactive patterns from free flow through to congestion (Jou *et al.*, 2013, TRB, 1994). When vehicles operate almost unimpeded with a greater ease of manoeuvrability within the traffic stream, they behave as typical fluid particles that have a high degree of movement (Level A) in the gaseous state. At the phase of "synchronised flow" (Level C), vehicles behave as liquid state fluid particles, where manoeuvrability is noticeably restricted and drivers are forced to drive as part of the overall traffic stream with

restricted ability to change lanes, due to the restricted flow movement (Sipress, 1999). At the extreme (Level F), there is highly restricted manoeuvrability on the road network, which is comparable to particles in the solid state (May, 1990). The evolution of these levels of service is due primarily to forced vehicular interactions on a network with limited capacity, aggravated by the inability to clear the junction during successive back-to-back saturated traffic cycles, and the consequent accumulation of the resulting queues leads to severe congestion. This phenomenon is likely to occur in the transition region between free flow and congestion, within the region of severely restricted movement, before Level F conditions. As chaos indicates that traffic is likely to move out of the free flowing state i.e. congestion is starting to emerge, this can be used as a trigger for remedial action.

In contributing to this discussion, Dendrinos (1994) proposed a utility maximisation model to explain the interactions of motorists in the environment and how this complexity develops the chaotic state of network.

3.4.2 Utility Maximising Model

The basic assumption in this model is that motorists tend to choose speed, acceleration, and spacing (headways) in order to maximise their utility, or the degree of satisfaction enjoyed by a motorist on the road network (Dendrinos, 1994).

In practise, the utility enjoyed by drivers under certain conditions influences their choice of speed and acceleration, which affects the dynamics of the traffic stream. The driving conditions are also functions of other environmental factors, which may include the maximum allowable speed, road surface conditions, weather, pedestrian activities, etc. All these varying mix of conditions faced by each individual motorist, coupled with the heterogeneity in driver population gives rise to varied response/reaction times.

According to the model, the utility of each individual driver is given by the utility function as:

$$U_j(t) = U_j\{\dot{v}_j(t), \ddot{v}_j(t), d_j(t) \dots\dots\dots\dots\dots\} \qquad \textit{Equation 3-1}$$

such that:
$$C_{jk}^{t}\{\dot{v}_j(t), \ddot{v}_j(t), d_j(t), r_j(t),\ s_j, t \dots\dots\dots\dots\dots\} \leq \hat{C}_{jk}(t)$$
$$[\dot{v}_j(t), d_j(t)\,] \geq 0, \quad \ddot{v}_k \neq 0$$

where:

- $\dot{v}_j(t) > 0$ is the operating speed of the j^{th} vehicle at time t $(t \in T_i)$, where T_i is the duration of a time segment i within a larger time interval T, say a day) across a designated cross-section of the road;

- $\ddot{v}_j(t) > 0$ (acceleration) and $\ddot{v}_j(t) < 0$ (deceleration) are operational parameters of the j^{th} vehicle at time t at that point;
- $d_j(t)$ is the vehicle density in the local area at time t (as perceived by the driver);
- $r_j(t)$ is a driver's reaction time; and
- s_j is sensitivity or car response to the driver's handling.

The vehicle specific parameters are constrained also by other factors: $\hat{C}_{jk}(t)$, which includes the constraint imposed by the local maximum allowable speed, road surface conditions, weather, drivers reaction time, sensitivity or car response to the drivers handling.

In the absence of congestion, drivers operate outside of a platooned environment and are unaffected by both the behaviour of neighbouring vehicles, and the distances from other vehicles. However, as vehicular interaction begins to increase in platooned environments, drivers respond to speed, acceleration, and separation of vehicles in front and behind. Consequently, the macroscale density is affected by the choice of speed by the individual drivers. Under these circumstances, motorists respond to the constantly changing influences of other motorists' in attempts to maximise their individual utility. Consequently, there is a constant fundamental feedback as motorists change their driving behaviour in order to maximise utility. Dendrinos *(ibid.)* states that, *"this interaction alone, in effect, may be responsible for non-equilibrium unstable (possibly chaotic) dynamics, depending on the connectances between speed, acceleration/deceleration and immediate density at each individual's utility and perceived constraint functions, and the system wide environment."*

The speed-flow curve in Figure 3.1 shows an upper region of 'stable flow' and lower region of 'unstable flow'. Motorists enjoy their utility on the upper curve but less utility on the lower side is characterised by congestion, increased journey time and significant delays.

As illustrated, speed tends to decrease (as flow increases) after an initial period of little change, but at an increased rate of change of speed as flow approaches capacity. This region of a sharp deterioration in speeds defines the onset of an unstable region (chaos) characterised by low speeds as vehicles are forced closed together into the link (Ortuzar and Willumsen, 2006). Dendrinos (1994) identifies that the utility of the motorist starts to become noticeably restricted due to the increased rate of interactions as flow approaches

capacity, and generates the fundamental feedback interaction that stimulates chaos, which creates conditions for hyper-congestion.

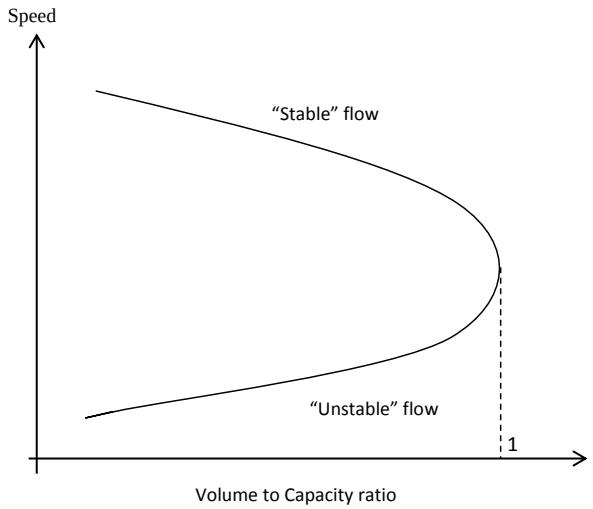


Figure 3.1: The Standard Static Operating Speed versus Volume Relationship
(based on Dendrinos (1994))

Dendrinos (*ibid.*) further alluded to the fact that urban road traffic (measured as traffic volumes or flow rates) exhibits periodicity when assessed over daily, weekly and monthly intervals, with the stronger periodicity observed when the level of aggregation is increased. On the contrary, the behaviour tends to aperiodicity, if analysed in finer time units of sub-one-minute intervals. For this reason, the finely resolved temporal traffic flow information is important for chaos studies enabling the chaotic properties of the system to be analysed. This will be revisited later in Sections 7.2 and 7.3.

Since the 1990s, there have been a number of studies using real-world data to investigate Chaos Theory in traffic analysis; however, as is shown next in Section 3.5, these studies have been limited to ex-urban networks (motorways and trunk roads) and using relatively low-resolution data in which the short-term variability in the traffic flow is often lost.

3.5 Previous Applications of Chaos Theory in Road Traffic Analysis

Several studies have shown that traffic data can display chaotic behaviour close to saturated conditions, making road traffic an ideal candidate for examining the use of Chaos Theory. In the UK, Kirby and Smith (1991) conducted an exploratory study based on the car-following theory to assess the possibility of applying Chaos Theory to traffic. Their results encouraged the need to examine non-linear dynamics in more detail and confirmed that various parameters that arise in transport models, such as car following, do bear characteristic features of chaos. The authors concluded that Chaos Theory is suitable for

understanding the conditions under which instabilities evolve and propagate through road transport systems. The study also highlighted the need for guidance on how to select the forecasting horizon to make predictions realistic and, finally, recommended a study into ways of analysing and summarising data for timely detection of the transition to chaotic behaviour.

More recently, research into Chaos Theory has mainly focussed on motorway networks, as they are relatively simple systems to analyse compared to urban road networks where several complexities exist due to the influence of various interacting elements (such as vehicle types, pedestrians, weather and road layout). Motorways tend to be long stretches of carriageway, often without a dense deployment of detectors to provide sufficient and appropriate traffic data to be able to understand the evolution and propagation of the dynamic states across the network. These studies were based on relatively low temporal resolution data, for example, one-hour intervals, and lacked investigations over a wide spatial area (such as a network of coordinated junctions and adjacent control regions, for example). Nair *et al.* (2001) analysed time series data (consisting of speed, volume and occupancy) from inductive loops installed on the San Antonio freeway system. The data were validated against established relationships between average speed, total volume and average occupancy. However, the actual non-linear analysis of the traffic data focussed on average speed time series aggregated into 5-minute data. By examining the deterministic properties of the data, traffic speed was characterised as chaotic, and techniques based on phase space dynamics were found to be suitable for analysis and prediction. Consequently, the study identified that time series traffic data is deterministic and therefore rendered suitable for modelling by applying phase space techniques.

Lan *et al.* (2003) collected automatic traffic count records sampled at one-minute intervals at 20 selected stations on the Miami Freeway in the USA, for testing and predicting traffic flow dynamics using Chaos Theory. The statistical tests carried out included the Hurst exponent, largest Lyapunov exponent, Correlation dimensions and Kolmogorov entropy. These tests indicated consistently strong evidence of a chaotic rather than a random structure. Furthermore, short-term traffic flow predictions were possible. In a study in Sacramento, California, Frazier and Kockelman (2004) confirmed that a more generalised view could only be drawn when a wider range of traffic flow rates are analysed. The latter study used data from inductive loops collected over a one-month period at 30-second intervals on a section of Interstate 80 highway. Crucially, it concluded that a lack of

extensive data, or sampling at the wrong frequency, fails to capture accurately the important dynamic features in a system. This has severely inhibited progress applying Chaos Theory to traffic congestion management.

Shang *et al.* (2007) identified the existence of a possible geometry (rather than a stochastic character) in traffic speed time series data in Beijing when sampled every two minutes. This tends to suggest that the traffic flow process could be viewed from a new perspective by reference to its chaotic properties. It indicated that traffic time series is deterministic and could be modelled using phase space techniques. Similarly, Zhang and Liu (2007) analysed traffic data from video equipment installed on a freeway passing through a suburb of Tianjin in China using Chaos Theory. The data were analysed for a range of time resolutions between 30 seconds and 6 minutes. A recurrence plot indicated the possibility of chaos, and correlation dimensions could be determined. Analysis showed traffic flow to be chaotic and fractal (irregular or fragmented geometric shapes, each of which is a small part of the whole) under a large range of observation time windows. Correlation dimensions and largest Lyapunov exponents decreased with increasing time windows at one minute scales, which confirms the sensitivity of these parameters. In order to achieve a reliable forecast the study suggested that the sample size should be more than ten times that of the correlation dimension of the short-term traffic flow.

Krese and Govekar (2013) distinguished between the traffic dynamics of a ring road around Ljubljana in Slovenia and several adjoining highways, based on Chaos Theory. The study concluded that the embedding dimensions for the highway were lower than the ring road traffic, whilst the maximum Lyapunov exponents showed higher sensitivity of the highway compared to the ring-road, which is indication that the ring-road was capable of responding better to small traffic perturbations than the highway. This observation was attributed to the higher density of junctions and connections on a ring road, compared to the highway, making it easier for the redirection of traffic to avoid a traffic jam, accident, etc.

In a non-motorway study, Xue and Shi (2008) analysed traffic flow data from Dongjiang Street in Dongguan, China. These data were sampled at five minutes intervals from August to October 2006, which gave more than 30,000 data points for investigating the short-term forecasts based on Chaos Theory. This research deployed the phase space reconstruction method, but used a local polynomial model of the nearest neighbour to forecast the traffic flow. The previous 80-days data was used to provide an historic "picture", whilst the latter 12 days was used to examine the accuracy of the predictions. The results indicated a

maximum relative error of 0.45% with a minimum of 0.04%, whilst 89.5% of the total data had a relative error under 0.10%. The results demonstrated the high precision of local regression models used to make predictions from the reconstructed phase space of the evolution of the traffic flow.

Further, research based on UTC-SCOOT data from the Traffic Management Bureau in Beijing carried out traffic flow forecasts from data sampled at an averaging frequency of 15 minutes (Jianming *et al.*, 2003). The result of one of the roads presented in the paper indicated that the forecast and actual curves fitted well with a mean square error of 7.1% and equalisation coefficient of 95.4%. It is unclear why the study did not analyse the spatial and temporal evolution of congestion across the network, but this could possibly be due to the low temporal resolution data (which conceals the short-term variability upstream-downstream in traffic patterns) and the sparse distribution of available data for the network. Thus, even in the instance where the study used urban area-wide data, it could not investigate the complexities in the urban flow network to investigate the occurrence of chaos at the microscale level or across a network of interconnected junctions. Therefore, it suggests that Chaos Theory is relatively data-hungry and requires highly resolved temporal and spatial data in order to function appropriately to reveal congestion patterns evolving in time and space across a network.

3.6 Limitations of Previous Studies

The above evidence does suggest that the chaos phenomenon can occur within road traffic networks. However, these earlier studies have only focussed on systems with low temporal resolution or simple linear networks such as motorways that are relatively easy to analyse, compared to urban road networks. Consequently, Chaos Theory has limited potential for traffic management in urban networks when the monitoring environment lacks a high-resolution data capture rate. Chaos Theory is data-hungry and dependent upon the availability of large volumes of data from a densely surveyed network with observations over a long time frame and possessing a resolution that preserves the variability in short-term traffic patterns. As indicated by Kirby and Smith (1991), the data must allow the transition to chaotic behaviour to be readily detected in order to enhance our understanding of chaotic traffic flow. This is dependent on data of sufficiently high sampling rate available over a wide spatial area and data quality such that transients, measured variation in congestion is picked up in time and space. This enables “early warning signs” of the state change within the system to be detected giving network managers more time to react.

Suitable data could be for example, sub-one-minute interval data of flow, speed, link occupancy and traffic noise (Frazier and Kockelman, 2004, Lan *et al*., 2003, Nair *et al*., 2001). Despite the problem of data resolution, previous work indicated that chaotic behaviour was present in traffic flow on motorways, inter-urban and urban roads. However, whether chaos is initiated at the microscale on a single link or across a network of interconnected junctions in the urban network remains an unknown. The optimum approach therefore, to enhancing our understanding of road traffic's chaotic behaviour, is to focus research on examining complex networks, for example, a group of interconnected urban junctions, such as a SCOOT region, using highly resolved traffic data.

The next section suggests that there are now new and enhanced data sources that overcome previous limitations with applications of Chaos Theory to urban networks. If chaos occurs network-wide, current signal control systems for managing traffic in urban areas will benefit from being able to take into account the spatial and temporal evolution of congestion across network regions. The literature has revealed evidence to suggest that it is possible to use chaos-based algorithms to design optimal control solutions to address strategic network traffic problems by providing an additional level of intelligence to UTC systems. This will enable these systems to proactively detect congestion and implement strategies to prevent saturated conditions occurring in a network region, over a group of regions or across an entire city region. Chaos Theory will enable assessment of congestion in and between neighbouring network regions through systems that are capable of talking to each other, and therefore enabling the management of congestion strategically across a much wider area and in a more pre-emptive manner, with the overall benefit of more reliable journey times and reduced pollution area-wide.

3.7 New Sources of High-Resolution Traffic Data

The key characteristics of traffic behaviour (e.g. traffic flow and link occupancy) needed to deploy Chaos Theory can vary over a matter of seconds and within tens of metres across a network. Therefore, to detect these variations, the system must capture relevant data at sufficiently high spatial and temporal resolution. Otherwise, the full potential of the Chaos Theory application to urban traffic control will not be realised. These data enable an understanding of the prevailing dynamic state and perturbations in traffic on a link, at a given point in time, and the evolution of the resulting congested state in the near and short-term future. The disadvantage of low-resolution data is the failure to preserve the short-term variations in traffic patterns, making the dynamic property of the system unclear;

therefore, such data do not support the examination of Chaos Theory adequately. Recent innovation in monitoring technologies (such as pervasive sensors known as Motes and Bluetooth devices) along with sophisticated legacy system data sources (such as SCOOT and ANPR) provide highly resolved (temporal) traffic data over a sufficiently high spatial density. This provides a new stimulus for further research into gaining a better understanding of characteristics of chaos in traffic data to enhance the current automatic management of urban road traffic. Thus, the benefits of integrating diverse data sources within existing urban control systems are that they will serve a dual function of delivering the traffic management and control function as well as providing the data to enable Chaos Theory to report on the more strategic situational state of the urban network in real time (Downes *et al.*, 2006). Sections 3.7.1 and 3.7.2 explore these data sources in more detail, in terms of their overall feasibility, and the requirements for high temporal and spatial resolution data.

3.7.1 Pervasive Devices

Motes and Bluetooth are relatively inexpensive devices that are easily and flexibly installed with minimal visual intrusion, enabling mass deployment in urban areas at relatively low capital and on-going maintenance costs. A typical network consists of sensors attached to street furniture (e.g. traffic signal poles, lamp columns and railings). Also, these devices can be programmed to collect traffic and environmental data at varied levels of temporal resolution (including second-by-second).

Motes can be mains or battery-powered, compact and low-cost devices that contain a number of different types of sensors including temperature, humidity, carbon monoxide (CO), nitric oxide (NO), nitrogen dioxide (NO_2) and noise that can be programmed to enable continuous monitoring and simultaneous measurement of ambient conditions including vehicle movement and pollution over a networked area (Akyildiz *et al.*, 2002, Bell and Galatioto, 2013, Bell *et al.*, 2009). Depending on the type of in-built sensor, a mote device can gather traffic information such as traffic flow, link occupancy and speed when it detects vehicles within its zone of surveillance (Shih *et al.*, 2001).

The project MESSAGE (Mobile Environmental Sensing System Across Grid Environments) jointly funded by the UK DfT and the Engineering and Physical Sciences Research Council (EPSRC) developed prototype motes. These were deployed in UK urban networks including Gateshead, Newcastle and Leicester (where the data captured was integrated with real-time traffic data from SCOOT in a database), as well as in London and Palermo in Italy. The project was a collaboration of Imperial College (lead partner) and

the Universities of Newcastle, Cambridge, Southampton and Leeds. The first commercial deployment of the Newcastle MESSAGE mote integrated with SCOOT system was in Medway and Newcastle (see Figure 3.2). Data is sampled at 5Hz, and averaged to one minute, and transmitted from sensor to sensor (up to 5 hops) to a gateway which relays data from the whole array (up to 100 sensors) by GPRS or LAN/Wi-Fi (North *et al.*, 2009).

Bluetooth devices use a low cost transceiver chip to communicate and exchange information with other enabled devices within a global frequency range of 2.45Hz. Figure 3.3 shows a typical Hi-Trac® Bluetooth equipment installed on a traffic signal pole. Each device has an unique identification called a MAC (Media Access Control) address, which facilitates the identification of individual vehicles (with an enabled Bluetooth device) passing through the network.

Figure 3.2: Typical Mote Installed on a Lamp Column

(Source: Bell (2012))

Figure 3.3: Typical Bluetooth Installed on a Traffic Signal Pole

(Source: Lees (2013))

Bluetooth scans an enabled in-vehicle device (such as mobile phone, laptop and in-vehicle electronics/engine management systems). When a vehicle passes through the device's detection zone, a data recorder stores the device address and records the data type to the nearest second (Kay and Jackson, 2012). Through the matching of MAC addresses at successive detection stations, Bluetooth devices provide a large volume of high-resolution traffic data including traffic volume, occupancy (density), speed and travel time (Blogg *et al.*, 2010). Examples of areas where Bluetooth networks are already deployed for traffic

data collection are Queensland (in Australia), Indiana, Pennsylvania and Chicago (in North America), and Altrincham, Manchester, Stockport and Wigan (in England), and Perth and Inverness (in Scotland) (Cragg, 2013, Galas, 2010).

Unlike the Motes and Bluetooth detectors, SCOOT and ANPR systems are capital intensive and not as pervasive (with installations mainly at junctions and intermediate locations such as pedestrian crossings) making them less capable of providing high-density spatial data although they can collect high temporally resolved data.

3.7.2 SCOOT and ANPR

It must be emphasised here that SCOOT and ANPR systems are not new data sources, but the enhanced ability to analyse the data through increased storage capability and computing power makes them potential sources of data for Chaos Theory. Until recently, SCOOT and ANPR datasets were only available in aggregated form at low resolution therefore less suitable for Chaos Theory. Although the raw data is collected in 20 second, 30 second, one minute and two minute intervals, historically measurements have been aggregated typically in time synchronous intervals of 5 and 15 minutes providing average values for each time period. This procedure reduces the size of data files for storage in databases (such as ASTRID) thus minimising the required memory and data writing (or retrieval) which can be both expensive and computationally intensive. Once aggregated, any traffic variation within the time interval is lost; therefore, such low resolution data smoothes the short-term variation in traffic patterns, which ultimately affects the accuracy of short-term forecasts (Zheng and McDonald, 2007). However, during the past few years, the unprocessed data has started to become available to traffic managers enabling processing to take place at second by second intervals. As such, future studies will benefit from datasets whose short-term variation in traffic patterns are known and therefore make possible the opportunity to investigate the underlying non-linear structure available in the detailed measurements. Furthermore, the use of computer technologies such as The Cloud computing system are able to meet the real-time data processing requirements and the computational intensity associated with making these databases more suitable for the Chaos Theory application.

The SCOOT system's layout was described in in Section 2.5.5. As an adaptive system, SCOOT responds automatically to traffic fluctuations by performing three optimisation procedures namely split, cycle and offset and makes slight adjustments to the signal timings depending on traffic conditions within the SCOOT region (Peek Traffic Ltd. *et al.*, 2008). Each optimisation affects a small incremental change in signal timings to avoid

inappropriate response to transient queues that impacts positively on the overall performance of the traffic signal network in the region (DfT, 1999). SCOOT is used in signalised urban networks in more than 250 towns and cities in the UK and overseas to manage traffic in small areas of the urban network (TRL, 2013). The relatively low number of SCOOT and ANPR networks is due to their large capital cost and does imply that Chaos Theory may be applicable only where high-resolution data are available, but it is unlikely that the data are available at an entire citywide level. On the other hand, the spatial resolution of the data itself could be low depending on the distribution of loop detectors within the network.

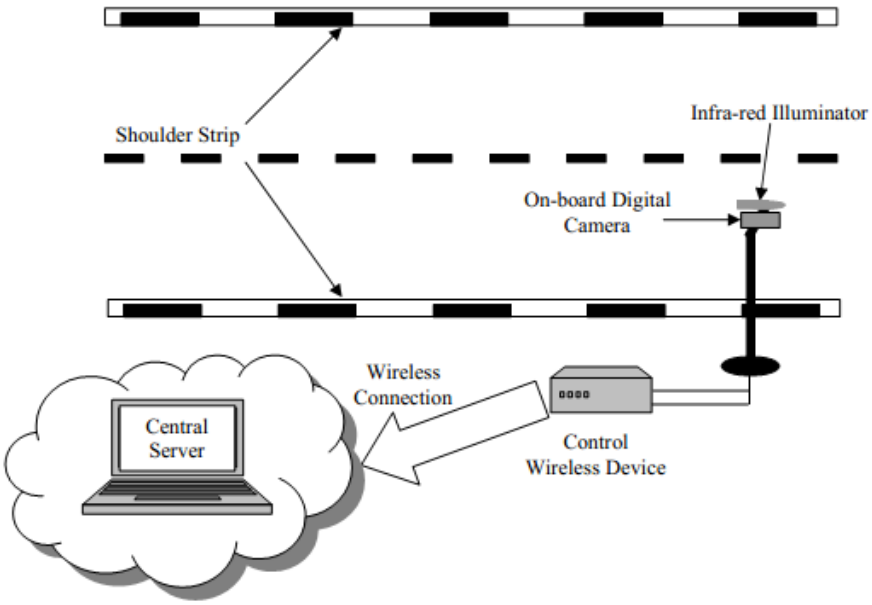


Figure 3.4. The Schematic Diagram for an ANPR System
Yasin *et al*. (2009)

Figure 3.4 shows the system's layout for the ANPR, which provides continuous information from monitoring sites to determine conditions such as flows and journey time on the road network (Hounsell *et al*., 2009). In the UK, a number of regional and national sites (and other local traffic control centres such as the ROMANSE office in Southampton) collate more than six million records of traffic data per day for real-time application (Zheng and McDonald, 2007). It is possible to process ANPR data at any time resolution and therefore the installed sites could provide high-resolution data. As both SCOOT and ANPR are potential data sources of low spatial density, combined with Bluetooth and Motes configured to deliver data at appropriate sampling time and at "critical locations" will be beneficial. This will enable the problems regarding their network coverage to be resolved

for analysing the urban road network using Chaos Theory provided proxy measures of flow and occupancy can be derived from pollution and noise sensors (Galatioto and Bell, 2013). The benefit offered by Chaos Theory is that it can use the wealth of data available as a by-product of existing ITS legacy systems. In an environment of economic restraints, policies of 'more for less' demonstrating additional spin off benefits allows local government to justify ongoing maintenance of existing UTMC systems.

In a real world application, it is possible that data from a single source will be inadequate for the study of an area-wide network, and therefore supplementing with other data sources may be necessary. As SCOOT and ANPR are capital intensive and limited to key locations only, pervasive instruments (such as Mote and Bluetooth) may provide additional information, if required. They may also allow existing areas covered by SCOOT and ANPR to operate through a "linking deployment" of Motes or Bluetooth sensors. The challenge to the implementation of Chaos Theory is whether these different data sources could easily fuse together through an interoperable system.

In extending the principles of Chaos Theory to real-time applications, a traffic system's chaotic property will need to be determined continuously from the discrete measurement of the dynamic traffic stream variables. Consequently, a further requirement will be an efficient and dynamic means of analysing these variables, for example, through stream processing. Stream processing has the capability to execute data continuously over long periods so long as there is a supply of new data for processing. For example, by constantly analysing the traffic data of network links in space and time, it will be possible through an understanding of the spatial correlation between links to use the Lyapunov exponents to forecast the congested links in real-time. This would enable the network manager to implement strategic mitigation plans in advance to avoid or minimise the adverse impacts of congestion. As such, it will be important, therefore, to address any potential memory constraints and computational inefficiencies within a computer system's architecture (Horton and Suh, 2005). In order to discuss other potential problems makes it necessary to introduce the concept of autonomic systems.

3.8 Autonomous Systems

In today's world of pervasive computing, other potential computer issues include integration and interconnectivity problems. There is a vision to connect trillions of computing devices over the internet, which poses a difficulty, but the need to integrate several heterogeneous software environment into corporate-wide computing systems and

beyond even makes matters more complex. With systems becoming more interconnected and diverse, it create challenges making system architects less capable to anticipate and design interactions among components. As these complexities increase, administrators find it more difficult to manage these systems and this has the potential to affect the ability to make timely, decisive responses to fulfil the stream of changing and conflicting demands. In order to address these challenges, in the future, it is likely that the option of *autonomic computing* will be pivotal.

The essence of autonomic computing in urban traffic management is to upgrade current UTMC systems to behave in an autonomous way by equipping them to respond to high-level objectives that will enable them to self-manage, with minimal human interventions, through four key processes namely self-configuration, self-optimisation, self-healing and self-protection. Therefore, ultimately, the role of human administrators in the operation of traffic management systems will be reduced to predominantly relatively less frequent high-level decisions, which the system's capability will enable automatically via numerous, lower-level decisions and actions (Kephart and Chess, 2003). Table 3-1 provides a comparative summary of the current and autonomic computing for the four self-management processes.

Self-configuration refers to the system's ability to readjust itself automatically in support of changing circumstances or to assist in *self-healing*, *self-optimisation* or *self-protection*. This will deal with issues that relate to installation, configuration and integration of large, complex systems that could be time consuming and prone to errors. Therefore, the system can accommodate new features such as software and servers without any disruption to the service (Sterritt, 2005). Autonomous systems will automatically configure themselves based on high-level policies by specifying the desired objective or goal, and not necessarily, the means to achieve it. For example, a new component such as Bluetooth (or mote) will integrate seamlessly into an autonomous system because as the component registers itself and capabilities, the other system components will adapt to its presence and learn how to use this component or modify their own behaviour appropriately.

Self-healing will make computers capable to detect, diagnose and repair localised problems resulting from bugs or failures in the software or hardware. Identifying problems of this nature is usually cumbersome and time consuming, whilst sometimes these problems disappear mysteriously after several unsuccessful attempts at diagnosis. Autonomous systems will enable experts to use their time more profitably because for example the

system will diagnose the problem and match it against known software patches, install the appropriate patch, and retest. Although current UTMC systems are able to detect and report faults, they require human intervention to fix the problem, whereas autonomous systems may alert a human programmer only when necessary.

Concept	Current Computing	Autonomic Computing
Self-configuration	Corporate data centres have multiple vendors and platforms. Installing, configuring, and integrating systems is time consuming and error prone.	Automated configuration of components and systems follows high-level policies. Rest of system adjusts automatically and seamlessly.
Self-optimisation	Systems have hundreds of manually set, non-linear tuning parameters, and number increases with each release.	Components and systems continually seek opportunities to improve their own performance and efficiency.
Self-healing	Problem determination in large, complex systems can take a team of programmers weeks.	Systems automatically detects, diagnoses, and repairs localised software and hardware problems.
Self-protecting	Detection of and recovery from attacks and cascading failures is manual.	System automatically defends against malicious attacks or cascading failures. It uses early warning to anticipate and prevent system wide failures.

Table 3-1: Four Aspects of Self-Management in Autonomous Systems
(Source: Kephart and Chess (2003))

Self-protection means that autonomous systems will automatically protect themselves from malicious attacks and accidental cascading failures not addressed by self-healing measures. They may also anticipate problems based on early reports from sensors and take steps to avoid or mitigate against them. This is necessary because in spite of the existence of firewalls and intrusion-detection tools, current systems still rely on human interventions to decide how to defend protect themselves. However, autonomous systems will undertake this automatically (Kephart and Chess, 2003, Ganek and Corbi, 2003).

Finally, *self-optimisation* means the system will operate with an awareness of what its ideal performance is, and is able to compare the ideal and current performance and initiate policies to improve itself (Sterritt, 2005). This will enable autonomous systems to improve their performance continually by identifying and applying corrective measures to enhance their operational efficiencies. In traffic systems, a number of parameters needs to be set

correctly, in order for the systems to perform optimally; often, there are only few people capable of tuning these parameters appropriately. Further, tuning one parameter can have unexpected impacts on the entire system given that the components of the systems tends to be integrated with each other. However, autonomous systems will have the ability to monitor and learn from experiment or experience and tune their own parameters, and therefore minimises the challenges due to human interventions or the non-optimal tuning of the system. *Self-optimisation* is a prominent feature (to a large extent) in existing traffic systems such as VA, MOVA and demand responsive UTC systems such as SCOOT, SCATS etc., but there is room to enhance this process (see Section 2.5.3 and 2.5.4). This research reveals the potential of Chaos Theory to expose the future dynamical state of the network in advance of real-time traffic conditions, thereby prompting the implementation of appropriate traffic management strategies. Therefore, Chaos Theory can add value to the current means of optimising these systems. Chaos Theory will have a role in the delivery of autonomous systems as a decision support tool in the traffic analysis procedure, which has ability to deliver a self-optimisation property for the self-management of road traffic systems.

3.9 Conclusions

This chapter critically reviewed the current knowledge in the applications of Chaos Theory to traffic studies and establishes that with today's computer processing capabilities, there is much potential for its application to strategically monitor and forecast the onset of congestion.

Therefore, Chaos Theory could play a role delivering a step-change in approach for traffic management, as the debate on how to deliver autonomic properties into traffic control continues to gain momentum among academics and traffic practitioners alike. Chaos is a characteristic property of the road traffic system attributable to the effects of several interacting elements such as cars parking and unparking, pedestrians crossing, buses stopping and the effects of different behaviours consequent to the mix of males and females and their individual driving styles.

In order to apply Chaos Theory to enhance the understanding of the evolution of the road traffic system, it is important to obtain traffic data that is able to capture the dynamics of how the system is evolving. An essential attribute of such data is that it must be able to preserve the short-term variability in traffic patterns, and data (sub-one-minute) must be

available at a high level of spatial density. As such, the acquisition of high-resolution traffic data (including flow, occupancy, speed, travel time and noise) is paramount.

Previous studies have been limited due to either a lack of sufficiently high-resolution data (as the low-resolution data lacks the characteristics of the short-term variation in traffic flow) or confined to relatively simple systems such as motorways and inter-urban links in sparse data environments, or both. However, due to recent developments, new and enhanced data sources such as Motes, Bluetooth, SCOOT and ANPR are now available. These should stimulate further research to examine properly the potential for Chaos Theory to help manage the urban road network. If proven, it is likely that the UTMC industry will seek to configure their data processing to deliver at higher resolution especially if further opportunity emerges for sophistication in decision support resulting in more effective congestion management.

In order to take advantage of highly resolved traffic data now available for progressing research for the application of Chaos Theory to urban networks, the next step is to understand the characteristics of chaotic systems and the mathematical algorithms that support Chaos Theory. Therefore, Chapter 4 is devoted to a critical review of Chaos Theory, whilst Chapter 5 deals with the mathematical algorithms for estimating key chaos parameters such as delay time, embedding dimension, correlation dimension, and Lyapunov exponents. Chapter 4 specifically makes a distinction between chaos and random systems by reference to the logistic map and bifurcation diagram. Finally, it conducts a critical review of a number of pioneering studies in the development of Chaos Theory and concludes by examining some of its general applications.

Chapter 4. A Critical Review of Chaos Theory

4.1 Introduction

Chaos is a phenomenon identified as a property of dynamic systems in the early 19th century, but it was not until the 1960s that Chaos Theory was applied to diverse fields including engineering, medicine, finance, human resources, etc. In spite of the wide applications, investigations for traffic applications has been slow partly due to the lack of highly resolved spatial and temporal traffic data for urban networks. As discussed in previous chapters, road traffic exhibits chaotic properties (see Section 3.4) and limitation of earlier traffic studies is no longer problematic due to the availability of rich new data sources (see Section 3.7). Before exploring chaos for traffic studies in Chapter 5, our general knowledge of this subject area is enhanced through a review of the extant literature and an introduction of the fundamental concepts.

Section 4.2 explains the concept of chaos in relation to dynamic systems, provides a technical definition and describes the characteristics of time-dependent non-linear systems that exhibit chaos. The Logistic curve and Bifurcation diagram are used to illustrate how the analysis of time series data may reveal structured underlying patterns for some systems that appear visually random. Section 4.3 discusses the Lorenz weather prediction model, which is pivotal to the development of Chaos Theory, along with the Lorenz Attractor and other contributory studies. Section 4.4 conducts a general overview of the application of Chaos Theory to many subject disciplines, using examples from the literature, and argues that chaos is an ubiquitous phenomenon and its application to transport deserves further investigation. Finally, section 4.5 distinguishes the methods for analysing chaotic systems, depending on whether the underlying equations of motion are known or unknown, and identifies that the method of the phase space reconstruction would be suitable for the traffic system given that its equations of evolution are unknown.

4.2 Definition and Characteristics of a Chaotic System

In plain language, 'chaos' refers to a messy situation; a state of utter confusion or disorder; a disoriented system that lacks organisation and possesses neither a pattern nor form of structure; or a system characterised by complete randomness. On the contrary, the term suggests otherwise when applied to chaotic systems. In order to understand Chaos Theory properly, it is important to establish the technical meaning of the term 'chaos'. According to Butler (1990) writing with regards to dynamic systems, *"Chaos in this context describes the behaviour of a variable over time which appears to follow no apparent pattern but in*

fact is completely deterministic, that is, each value of the variable over time can be predicted exactly". In more detail, Dreyer and Hickey (1991) state that, "*the term 'chaotic' is generally used to describe non-linear, but deterministic systems whose dynamic behaviour proceeds from stable points through a series of stable cycles to a state where there is no discernible regularity or order*." Chaotic behaviour is almost ubiquitous. It is a property of a number of natural or artificial systems that support our daily lives, but often these systems appear so random that it is impossible to recognise visually any chaotic patterns, although such patterns indeed may exist. Therefore, Chaos Theory techniques enable large arrays of parameters of chaotic systems to be represented in phase space (for example, a multi-dimension plot of a variable in an X-Y-Z plane) so that any possible underlying patterns can be identified to enable forecasting to be undertaken (Uittenbogaard, 2011).

In dynamic system theory, the term is useful in describing motions that are irregular and sensitive to initial conditions. These motions are products of nonlinear dynamic systems, and they exist between the realms of periodic and random behaviour. The latter definition gives a hint that chaos is a prelude to 'randomness' in systems that are sensitive to initial conditions. Such systems are not completely dominated by random behaviour but they have some degree of orderliness that defines their deterministic properties (Wang *et al*., 2005). Therefore, it is important to note that chaotic systems are distinct from random dynamic systems. The difference between the two systems is that chaos is predictable in the short-term, whilst randomness is unpredictable. However, ultimately chaotic systems tend to become completely random at the limits (Hui *et al*., 2005). Stewart (1997) is of the view that chaos is stochastic behaviour in deterministic systems. This statement is essentially correct but seems to lack clarity and could be misconstrued, given that stochastic processes are probabilistic, which implies that only the likelihood of their occurrence can be predicted whilst deterministic systems could be predicted exactly. Andrews (1996) clarifies this indicating that random behaviour could be associated to chaotic systems (which are non-random) only because of their extreme sensitivity to measurements. That is, it is the extreme sensitivity to initial conditions that tends to make chaotic systems appear random, though this is not the case in reality. In order to understand the characteristics of chaotic systems, Sections 4.2.1 and 4.2.2 include a numerical example from Butler (1990), which helps to explain the characteristics of chaos system in detail.

4.2.1 The Logistic Curve

In this section, the Logistics growth map is used to facilitate the understanding of the properties of chaotic systems because of its relatively simple form, but yet possessing characteristic features of dynamic systems. It describes the evolution of time of an arbitrary variable X (for example, population), which is a function of its previous value and a parameter k, which is the growth rate or 'tuning parameter' and the determinant of the steepness of the function.

The mathematical function is given by $X_{t+1} = kX_t(1 - X_t)$ where $0 < X_t < 1$ and $0 < k < 4$, and X_t and X_{t+1} are the previous and current values respectively. As illustrated in Figure 4.1, the parabolic curve of this equation shows an exponential rise in X_{t+1} as X_t is increased below the mean value, but X_{t+1} decreases as X_t is increased above the mean value. Another characteristic feature is that there is increased rate of change in X_{t+1} as the k-value is increased.

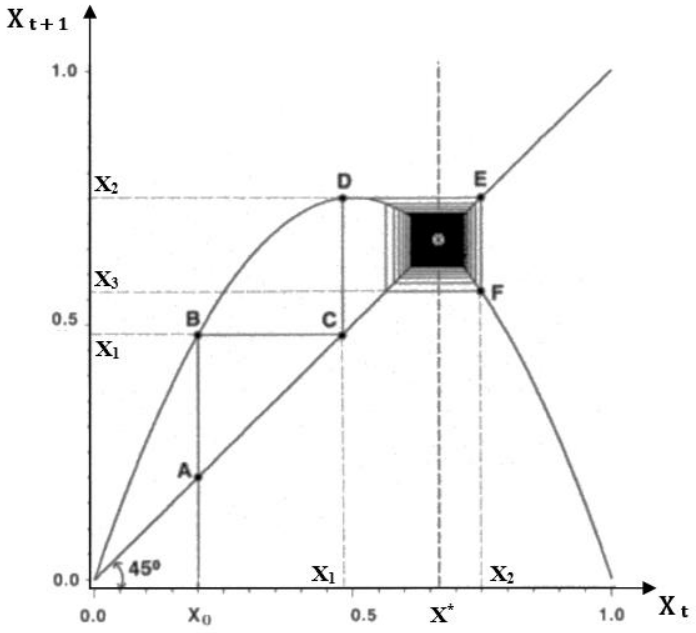


Figure 4.1: A Stable Time Path for a Logistic Growth Curve
(Source: Butler (1990))

Figure 4.1 also illustrates the solution of the logistic equation determined graphically through an iterative process. Let us suppose that we require a solution for the logistic equation $X_{t+1} = 3X_t(1 - X_t)$, where $t = 1, 2, 3 \ldots 50$. All values of X_{t+1} will lie on the parabolic curve, whilst the 45° line shows the path of points when $X_{t+1} = X_t$. Suppose that the initial value $(X_0) = 0.20$ at time $(t) = 0$, then the intercept X_0B on the parabola defines X_1 at the first iteration. At the second iteration, $X_0 = X_1$ (this is determined from

the intercept of X_1BC on the 45° line), point D on the parabola defines the future value (X_2) of the system. Similarly, for the third iteration, let $X_0 = X_2$ and point F on the parabola defines the future value (X_3) of the system. This iterative process is repeated for a number of times until eventually the variable converges to a steady equilibrium X^* at $\frac{2}{3}$. As long as the k-value is less than three, the system is stable and reaches some other steady state values other than the mean (Butler, 1990).

This equilibrium point however becomes unstable as soon as the k-value exceeds three and the system starts to bifurcate (doubling of the solution) exhibiting a two-period cycle where the solution alternates between two values. By increasing the k-value gradually the bifurcations increase further with the periodicity of 2^n (where n = 1, 2, 3…), such that the cycles progress to 4-period cycle (where the system repeats the same sequence after every fifth iteration), then 8-period cycle, etc. The bifurcations become more rapid until the point of accumulation with k-value of approximately 3.57, when the system becomes chaotic, where there can be an infinite number of periodic cycles. This illustrates the characteristic of chaotic systems. The observation reveals that a change in the k-value affects the degree of non-linearity in a deterministic system, which consequently impacts on the output solutions.

Butler (*ibid.*) further illustrated that the time series plot of X_t versus the number of iterations for the logistic curve shows a fluctuating profile even though the system converges to a stable point. Therefore, in our illustrative example, it may be misconstrued based on the fluctuating times series profile that the system lacks chaotic properties, if the underlying pattern of the system was unknown. This demonstrates that the fluctuating profile of time series data are not always outputs of completely random systems. Andrews (1996) confirms this observation indicating that deterministic systems could generate profiles that may be pseudo-random but this is only due to their sensitivity to initial conditions. This suggests that small changes in the initial conditions may be responsible for disproportionate changes in the future values of the system. The foregoing discussion, in relation to the logistic curve, implies that traffic systems that generate non-linear time series profiles may be products of deterministic rather than stochastic processes. Therefore, it is important to assess the data from these systems critically before drawing a factual conclusion, as to whether they are random or chaotic. In particular, traffic time series data require detailed assessment to determine the type of system responsible for its evolution.

4.2.2 The Bifurcation Diagram

An example of a biological population by May (1976) and May and Oster (1976) identified a progression of states of stable cycle through to a bifurcating hierarchical stable cycle with a period of 2^n before exhibiting chaotic behaviour. Figure 4.2 is a bifurcation diagram that presents a graphical view of the changes in population due to varying growth rates. Similar to the Logistic map, the solutions of these biological and ecological populations also illustrate how the cycles of solutions in chaotic systems develop.

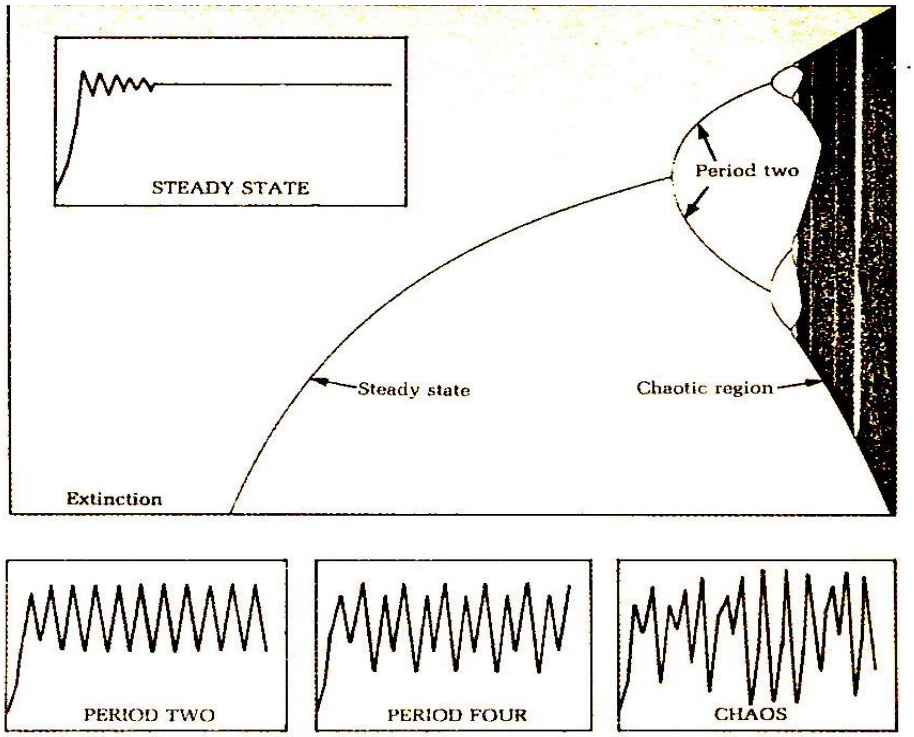


Figure 4.2: The Bifurcation Diagram for the Population Equation
(Source: Gleick (1988))

Figure 4.2 shows that the population is extinct at extremely low k–values (growth rates). At increased (but still low) growth rates, the population always settles to a single number, then further increases in k–values gives rise to bifurcations with a period of 2^n, which increase rapidly at higher rates until chaos appears, where there is no discernible order. That is, the regularity in period-doubling continues until another critical value, beyond which the equation gives solutions without any apparent order, which identifies the region where chaos is said to occur (Kirby and Smith, 1991). According to Gleick (1988), there is still the possibility of stable cycles of odd periodicity i.e. stable cycles of 3, 6, 12…(or 7, 14, 28…) beyond this hazy region, which could eventually plunge the system into a state of renewed chaos. Rae (2006) puts it succinctly that a closer inspection of the white strips reveals a little window of order where bifurcations re-occur prior to a secondary chaos.

From this knowledge emerges the property of *self-similarity* (a characteristic property of chaos), which acknowledges that there was an exact copy of the graph hidden deep inside it. In technical terms, *fractal* describe images that display the attribute of self-similarity; both *self-similarity* and *fractals* are discussed further in Sections 4.3.3 and 5.6. It must be pointed out even though chaos is a property of nonlinear systems the presence of non-linearity does not always ensure chaos (Lorenz, 1993b).

In concluding this section, it is important to acknowledge the work of Xu and Gao (2008) that identified the presence of bifurcations in the nonlinear analysis of traffic flows. The study identified point and period-2 attractors at low densities, period doubling (bifurcations) at high flow, and chaos in unstable flow conditions that induced higher rates of acceleration and deceleration of drivers in the traffic stream. Chaos identifies the onset of congestion, which gives rise to stop-and-start conditions, which is undesirable to motorists. Such conditions also increase fuel consumption, tailpipe emissions and the likelihood of accidents, and decrease the vehicle throughput across the junctions. The study of chaos therefore is important in order to create an awareness of impending unstable conditions and allow the implementation of mitigating strategies to improve the level of service and safety of the road system. The next section reviews important literature in the development of Chaos Theory and in the understanding of chaos in various dynamic systems.

4.3 Pioneering Work of Chaos Theory

The pioneering work in Chaos Theory was in 1961 by Edward Lorenz, an American mathematician and meteorologist, who observed that a small change in the initial setting of a system's conditions could have a significant impact on its long-term state, as in Figure 4.3. This became known popularly as the 'butterfly effect' (Gleick, 1988, Andrews, 1996). The two curves in Figure 4.3 are plots from runs of the same model assuming different values as initial conditions.

In this model, the data in the computer memory was six digits, but in order to save space the printout of the model output only shows three digits. In order to repeat a particular sequence again for a better understanding of the model, Lorenz mistakenly used the conditions (three digits) at some point in the middle of the ***first*** run output as initial conditions for the ***second*** (instead of restarting the entire run). Assuming that a small difference was inconsequential, Lorenz entered this second model's initial conditions as three digits (instead of six digits) based on data from the ***first*** run. Surprisingly, the results

that matched initially began to divert dramatically with the *second* run losing all the resemblance to the *first* run shortly afterwards. Lorenz discovered that the observation was due to a change by a tiny amount of 0.000127 in the initial setting of the weather simulation model, which gave rise to a rapid divergence in the final output of the model.

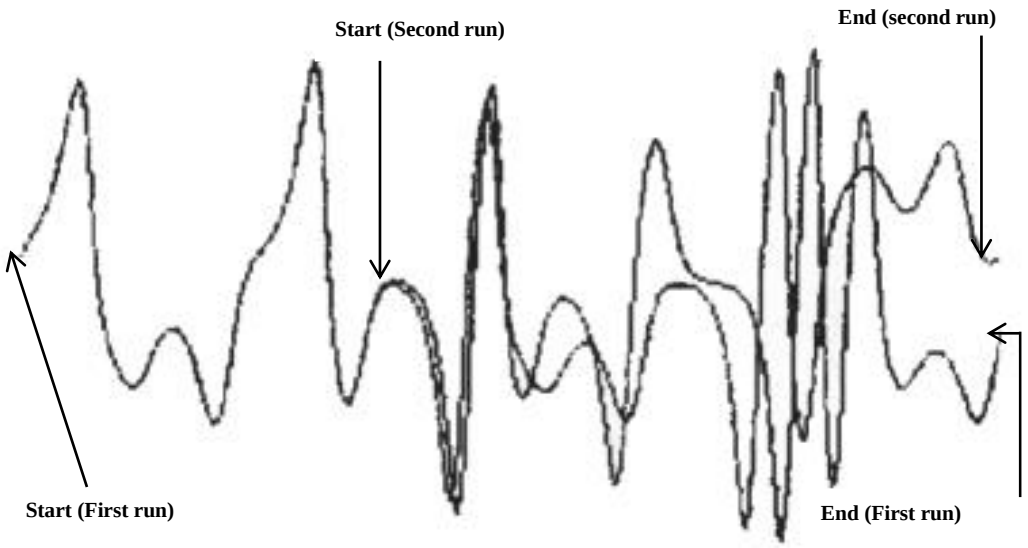


Figure 4.3: Lorenz Experiment Graphs
(Source: Bradley (2010))

Lorenz concluded that small tiny changes in initial conditions could lead to dramatic impacts on the long-term state of the system. Whilst these small changes can be due to real conditions in natural or technical systems, they also may be the outcome of rounding or approximation of values in a model. This important finding by Lorenz became a worthy contribution to the System Theory of the time, and became the main foundation for the development of Chaos Theory when it became a subject discipline many years later.

4.3.1 The Lorenz Attractor

Lorenz's later work in 1963 used the waterwheel to explain the instabilities that occur in the earth's atmosphere, depicting its motion as a "double-spiral well-patterned curve", though the behaviour of the system may appear visually as random. Figure 4.4 is a typical waterwheel that comprises a large wheel mounted under a waterspout, which represents Lorenz' model for investigating the instabilities in the atmosphere.

The wheel has buckets equally spaced on its rim, each mounted on swivels that open upwards with water fed directly above and a small hole at the bottom, so that water that trickles into the top and drips out from the bottom. At initial low flow speeds, the buckets do not fill up enough to overcome the frictional forces and therefore the wheel returns to

its stationary rest position after an initial spin. At some stage when gradually increasing the water flow speed, the heavier buckets empty of water when descending whilst the lighter buckets ascend in the upward motion to be refilled again. The wheel continues to turn at constant flow rates through to average and higher rates of flow. At higher speeds of flow, the waterwheel continues to maintain a clockwise/anticlockwise motion, but suddenly jerks about and revolves in the other direction. The system now enters a chaotic state, where its previously simple rotation becomes awkward and there is no easily recognised pattern or repetition, because the conditions that allows for synchronicity of the bucket filling and emptying no longer operates (Lipa, n.d., Kolář, 1992).

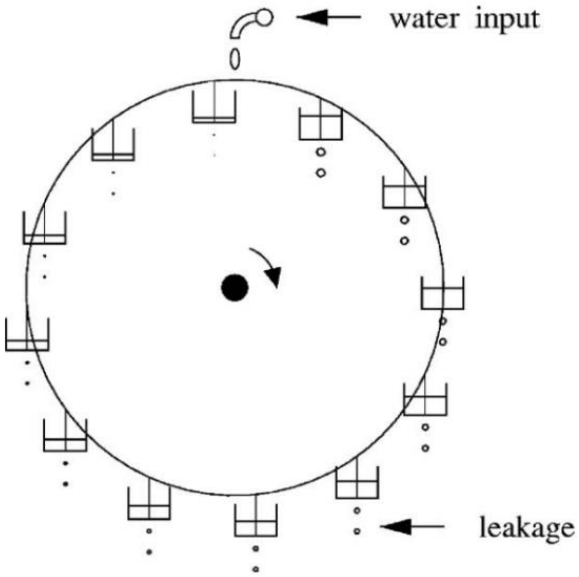


Figure 4.4: Twelve-cup Water Wheel
(Source: Matson (2007))

According to Lorenz, the behaviour depicts the rotation of air columns in the atmosphere and the interpretation of the model is as follows. The water contained in each bucket represents heat in the column of air, and the loss of water signifies heat losses by radiation. The greater force required to push down the heavier bucket corresponds to the force required to push out hot air. When cold air from the atmosphere meets the ground, it is heated and then it rises up back into the atmosphere. The rising hot air replaces the cooler air in the atmosphere inducing a rotation, which causes the denser air to fall to the ground, and consequently repeating the whole cycle.

The Lorenz Attractor was derived from this experiment, and is illustrated in Figure 4.5, which shows a plot of the atmospheric state from time dependent variables in the respective

X, Y and Z planes: convective flow, horizontal temperature and vertical temperature distributions respectively (Yurkon, 1997). The image is neither steady (in which the value of the variable never changes) nor periodic (that systems goes into a loop and repeat themselves indefinitely), but has a definite ordered pattern which neither settles down to a single point nor repeats itself.

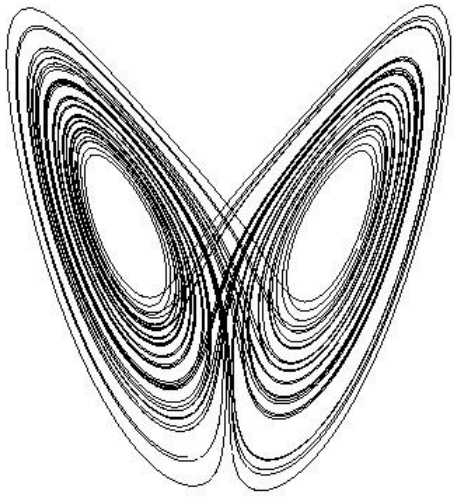


Figure 4.5: The Lorenz Attractor
(Source: Yurkon (1997))

As alluded to by Butler (1990), such occurrences suggests that non-linear time dependent systems that appear to fluctuate may have an underlying pattern, which could be analysed to determine the system's evolution. Until Lorenz's work, only the two states of 'steady' or 'periodic' were recognised. This led to the discovery of the third state of dynamic systems formally became known as "chaos", and the understanding of its mathematical principles became the foundation for Chaos Theory.

4.3.2 Pre-1960s Studies

As far back as the late nineteenth century, scientists had observed an unexplained phenomenon under the prevailing linear systems theory. For example, the failure of linear systems theory to explain certain experiments the results of which led to the logistic maps alerted scientist to possible gaps in the understanding of dynamic systems. However, due to imprecision in measurements or *noise* in the system this was ignored. With the discovery of chaos led to a realisation that the previously unexplained phenomenon attributed to *noise* in the data was actually integral to the studied systems (Abdalah *et al*., 2010). It is not clear why this erroneous view gained credibility for several decades before the Lorenz discovery.

However, it is possible to suggest that the advent of computing power for solving complex mathematical equations has supported immensely Lorenz's work in the 1960s. The computer is indispensable for studying chaotic dynamics, and for applying Chaos Theory in particular, given the need to overcome the challenges of repeated iterations of mathematical formulae that are either cumbersome or impractical by hand due to their computational intensity. The dual importance of the role of computers in the application of Chaos Theory is firstly the relative ease to undertake iterations and secondly the ability to generate graphs and images that describe these systems making it possible to visualise the behaviour of chaotic systems irrespective of the nature and complexities (Dhillon and Ward, 2003).

It is important to appreciate the contribution to this subject of key researchers in the late 19th century. Some authors credit John Clerk Maxwell, who first developed the theoretical unification of electricity and magnetism, as the first person to understand chaos, according to evidence from his work in 1876 that acknowledge the importance of systems that are sensitive to initial conditions (Hunt and Yorke, 1993). Contrary to the view in statistical mechanics that complex behaviour was a product of the interaction of a large number of particles, Maxwell studied the interactions between two gas molecules and noted that sensitivity to initial conditions was pivotal in producing statistical regularity.

One of the early proponents of Chaos Theory was Jules Henri Poincaré (as reported by (Moiola and Chen, 1996)) who, as early as the 1880s, undertook the foundation work for modern day Chaos Theory. Poincaré formerly studied the three-body problem and later the n-body problem (where n is any number more than two orbiting bodies), where he attempted to find the general solution to the motion of two or more orbiting bodies in the solar system. Poincaré discovered periodic orbits that neither increased indefinitely nor approached a fixed point (Sowmya and Sathyanarayana, 2014). He identified the system as having a sensitive dependence on the initial condition, and also developed a topology to explain the general complexity of the evolution over time of the system (Wolfram, 2002). With these findings, Poincaré became the first person to discover a chaotic deterministic system, but this was benchmarked only after Lorenz's ground-breaking research. Lorenz work linked together previous knowledge such as the sensitive dependence on initial conditions, with strange fractals and its erratic dynamics, which all could be visualised together and replicated through a computer simulation model. It is perhaps the relevance of Lorenz's work in this context, and also to weather prediction, and subsequent

applicability to several disciplines of research that gave his work in the 1960s recognition as the third state for dynamic systems and not the mere identification of the phenomenon back in the 1880s.

In 1898, Jacques Hadamard studied the chaotic motion of a free particle sliding on a frictionless surface (Goh *et al*., 2009). The trajectories of the particles were unstable and diverged exponentially from each other, which emerges as positive Lyapunov exponent (Hadamard, 1898, Sowmya and Sathyanarayana, 2014). Interestingly, the use of the Lyapunov exponent is still relevant to present day Chaos Theory, being widely accepted as a key determinant of chaos in dynamic systems. Unlike Poincaré who showed the possibility of strange attractors and sensitive dependence on initial conditions (SDIC), Hadamard presented an explicit model for SDIC.

Poincaré's work inspired a number of theoretical development in the theory of oscillation for application to radio engineering. Dating back to 1927, Van der Pol and Van der Mark found stable periodic oscillations (i.e. limit cycles in current chaos terminology), in electrical circuits employing vacuum tubes (Wolfram, 2002). A limit cycle is a closed trajectory in phase space having the property that at least one other trajectory spirals into it (as time approaches infinity). They reported that while driving the circuits at certain frequencies (close to the natural frequency) an irregular noise was produced before the frequency jumps to next lower level. This was one of the first discoveries about deterministic chaos in radio engineering (Rosser Jr, 2009).

In the early 1940s, Mary L. Cartwright and John E. Littlewood carried out one of the early research to demonstrate the complex nature of the dynamics of the *Duffing* oscillator, which is a model of an electro-magnetic vibrating beam. By varying a crucial control parameter, the model generated double-period cascade of bifurcations in a transition to chaos. Further, Cartwright was the first to study the theory of deterministic chaos, as applied to the *Van der Pol* oscillator. In this work, Cartwright used a model of an electrical circuit with a triode valve whose resistance changes with current, according to the *Van der Pol* equation. The unforced version of the equation is:

$$\frac{d^2x}{dt^2} + a(x^2 - b)\frac{dx}{dt} + x = 0 \qquad \textit{Equation 4-1}$$

Where $a > 0$ and b is a control variable. For $b < 0$ the origin is an attractor, $b = 0$ a bifurcation with a limiting cycle occurs and $b > 0$, a Hopf bifurcation. A Hopf bifurcation

occurs when a periodic solution or limit cycle, surrounding an equilibrium point, arises or goes away as a parameter of the system varies. Cartwright detected chaos in signals from the model that depended sensitively on initial conditions and possessed non-periodic orbits consistent with Lorenz' work (McMurran and Tattersall, 1997). Next, Section 4.3.3 discusses fractals that have also become an important aspect of the study of chaos.

4.3.3 Fractals

Fractal is a property of self-similarity found in many real world systems including blood vessels, the branches of a tree, internal structure of the lungs, graphs of stock market data, the bifurcation diagram and Lorenz Attractor.

Benoit Mandelbrot is often characterised as the father of fractal geometry, but it is important to acknowledge that the ideal of *fractality* goes back to George Cantor (1872) (Peitgen *et al*., 2006). Analysing coffee prices, Mandelbrot found that the results did not fit the normal distribution. However, when viewed from the point of view of scaling, the sequence of changes was independent of the scale applied. For example, Mandelbrot observed a perfect match between the curves for the daily and monthly price changes. Mandelbrot applied this principle to the coastline noting an increase in the number of bays in the coastline as the length of the coastline was magnified (Rae, 2006). A typical example of this concept of scaling was captured graphically in Helge von Koch's mathematical construction, called the Koch curve (Gleick, 1988). Formed from an equilateral triangle, the Koch curve reveals a line of an infinite length surrounding a finite area (a circumscribed circle around the original triangle), which is best described by its *fractal dimension* (see Section 5.6.1). The Koch curve has a *fractal dimension* of about 1.26, which suggests that the Koch curve is rougher than a smooth curve or line that has a dimension of one. Since the dimension is less than two, it has no area and therefore poor at filling up space compared to a square. In chaos mathematics, the *fractal dimension* provides a statistical measure used to judge the complexity of the attractor (Jianming *et al*., 2003).

Although all of these works are of relevance for later studies of chaotic dynamics, it is interesting to note that Lorenz's work was an independent investigation without any link to these previous works. In spite of these, Lorenz's work accelerated research and applications to several fields. Lorenz's work led to formal recognition of Chaos Theory and accelerated research in many disciplines arousing extensive investigations especially among mathematicians and professionals within the scientific research community. The remainder of this chapter provides examples of chaotic systems and applications to

illustrate the capability of Chaos Theory in analysing varying forms of non-linear dynamic systems such as the traffic system.

4.4 Chaos Theory: Examples and Applications

Chaos Theory is a branch of nonlinear analysis used to analyse chaotic behaviour in dynamic systems, which includes complex systems such as traffic flow that appears too difficult to be handled by traditional techniques (Xu and Gao, 2008). These systems include the transport environments such as the interaction of highly variable physical and human elements (such as vehicles, weather and road surface conditions, road layout, topography and pedestrian activities). In these systems, it is the proper representation of the mix of interactions in a model, especially the immeasurable quantities and inconsistent relationships among these elements, that makes forecasting difficult. Chaos Theory is capable of dealing with these complexities, by naturally accounting for all of the factors that are responsible for the system's dynamics. It can extract all the system information and relationships without the need to first determining the laws or equations of the underlying dynamics (Fraser and Swinney, 1986). It is applicable to conditions (motions) that are highly sensitive to initial conditions and leads to aperiodic behaviour in deterministic nonlinear systems, thereby rendering long-term prediction impossible. In such systems, the future (and even the past) state of the system can be determined according to the evolutionary equations, based on Chaos analysis, if the initial value is known (Frazier and Kockelman, 2004). The difficulty is that the initial conditions are difficult to determine explicitly due to the continuous evolution of these systems.

Núñez Yépez *et al*. (1989) noted that the dynamic instability that produces chaotic behaviour can be found in many electrical, optical, mechanical, chemical, hydro-dynamic, and biological systems. Furthermore, chaos can be seen in the motion of a tossed coin (Levy, 1994, Ford, 1983), and experimental systems such as the damped and driven pendulum (Bevivino, 2009). Chaos has been identified in the alternating cycles of economic booms and recessions (Levy, 1994, Becker, 2008), turbulent flow in fluids (Lorenz, 1993a), and the dripping pattern of a water tap (Ambravaneswaran *et al*., 2000, Coullet *et al*., 2005, Dreyer and Hickey, 1991, Núñez Yépez *et al*., 1989).

Levy (1994) cites the toss of a coin or the roll of a dice is an example of a deterministic system, yet it yields unpredictably different outcomes for each event due to chaotic behaviour. He notes that two variables that influence the outcome of each event are: (i) speed of the spin attributed to difficulty in tossing the die or flipping the coin in the same

manner twice and (ii) rate of fall due to the influence of circulating air currents on the downward motion of the coin or die (Rae, 2006). The combined effect of both factors implies that, if it were even possible to toss the coin in exactly the same manner, the change in the circulating air current will affect the outcome (Ford, 1983).

Chaos has been detected in the drops of water that fall from a leaky tap under appropriate conditions (Dreyer and Hickey, 1991, Núñez Yépez *et al*., 1989, Coullet *et al*., 2005, Ambravaneswaran *et al*., 2000). The drop is subject to three main forces (gravity, liquid tension force and friction) at the tip of the tap. The liquid tension force ensures that the drop does not fall immediately but grows until the gravitational force exceeds the liquid tension force. Experiments identified a relationship between successive drops, based on evidence that each drop that exits the tap creates oscillations in the residue at the nozzle that serves as initial condition for the next drop. Further assessment of incremental changes to the rate of flow of fluid through the tap induced bifurcations such that a transitional change of state occurred where the drop behaviour changed from steady to chaotic through a sequence of stable cycles until the flow became continuous.

A number of related studies have also confirmed this. Dreyer and Hickey (1991) measured the time interval (t_N) between successive drops then plotted the time series data in a delay-time coordinate system (t_{N+1} versus t_N) to generate the discrete time map. The return map indicated patterns (clusters) that are characteristic of chaos. Furthermore, Scientists at the University of California Santa Cruz recorded the time interval of the dripping tap, analysed the data and identified an uneven time interval between drips at certain flow velocities. However, the plot of the data indicated that an underlying pattern was present (Rae, 2006). Another study investigated the dripping behaviour of the tap as a function of the flow rate. At low flow of fluid, the tap dripped at constant rate and there was equal interval between successive drops. Subsequently, the drops become subject to bifurcations with multiple drops exiting the tap, which increased with increasing flow until the emergence of chaos when there was no discernible order or pattern (Coullet *et al*., 2005).

Analysis of chaotic behaviour helped to provide insights about complex systems such as the earth's weather, the behaviour of water boiling on a stove, the migration patterns of birds, and the spread of vegetation across the continent. Initial work in this area led to the discovery that complex systems often run through some kind of cycles, although these cycles rarely duplicate and repeat exactly. Furthermore, the graphical profiles of these

systems also revealed settling down to some sort of equilibrium, irrespective of the initial conditions (Uittenbogaard, 2011).

Levy (1994) noted that Chaos Theory is useful in predicting weather conditions based on simulations of measurements from thousands of points dispersed across the globe. However, accurate forecasts are limited currently to only a few days, but can give sufficient information to warn of impending natural disasters such as earthquakes, hurricanes, tornadoes, etc. Due to deterministic properties inherent in the earth's system, the condition of the system at time t, if known, could be used to determine the future state at time $t + 1$. Even though the earth's system may appear to be visually unstable and complex, data measurements indicate that they are not completely random but do possess a certain degree of orderliness, which makes short-term forecasts possible. Consequently, if a simulation model of the system is available (or constructed), then with accurately determined specific starting (initial) conditions it will be possible to provide useful forecasts over several consecutive time-periods. The major challenge is that the initial conditions are difficult to know exactly introducing a component of error, and there is always the influence of small randomness that is difficult to incorporate in the model. The impact of such uncertainties could result in unexpected outcomes (see Figure 4.3), which limits the reliability of weather forecasts to only a few hours (or days). However, the long-term predictions may not be able to be determined for example when and exactly where events that lead to chaotic properties may occur, but may provide some information about the likely path of a potential occurrence.

The stock and financial markets do alternate between cycles of recession and boom, although the depth and duration of each event is difficult to predict. According to Becker (2008) previous studies have linked the difficulty of predicting such markets to the randomness of announcements and impacts of *noise* in both traders' decisions and financial statements that do not affect cash flow but influence the choice of stock options. The analysis of the time series information using Chaos Theory also may be useful to study the volatility of the stock market. Chaos Theory may enable examining the relationships between business cycles and other variables such as demand, interest rates and availability of credits. Becker (*ibid.*) identified chaos effects within business organisation attributing this to constant changes that are both internal and external to these establishments. These changes that affect existing practise may be due to the effect of technology, deregulation, removal of trade barriers and mergers and acquisitions. Therefore, the awareness of these

factors that influence chaos in the environment, if acted upon in a positive way, through for example acquisition of innovative technology or review of existing practise, could effectively increase operational capability and globalisation competitiveness in industries.

Chaos Theory has also been widely applied in the field of medical sciences. Chaos Theory has enabled medical practitioners to analyse the electrical activity of the human brain for the control of epileptic seizures (Kumar and Hedge, 2012, Oestreicher, 2007, Yambe *et al*., 2005). Research has revealed that the cranial activity of a healthy brain is essentially chaotic and therefore any anomaly is indicative of an unhealthy system (Sarbadhikari and Chakrabarty, 2001), whilst cardiac chaos may be a useful physiological marker for the diagnosis of an unhealthy heart (Denton *et al*., 1990). Studies indicate that a brain's chaotic pattern could suddenly change to abnormal regularity during epileptic seizures (Kumar and Hedge, 2012). Scientists use an EEG (electroencephalogram) machine to measure the brain's electrical activity which detects brain disorders when significant changes occur in measurements, where high-amplitude synchronised waves are observed for epileptic seizures (Oestreicher, 2007). In order to prevent the onset of the disease, patients benefit from an implanted device that induces more chaos once a change is detected (Yambe *et al*., 2005). Similarly, an ECG (electrocardiography) machine that measures the electrical currents produced by the heart detects arrhythmia (abnormal irregular rhythms). This enables the activation of chaos control programs to manipulate and normalise the heart's rhythm (Schewe, 2007, Garfinkel *et al*., 1992).

Chaos and fractals are also useful in the creation of computer art. For example, a computer can create a beautiful tree using a simple formula that creates irregular patterns of the back of the tree that do not exactly repeat itself (Rae, 2006). It also is applied to modelling shapes and rendering textures in digital images. Barnsley *et al*. (1988) reported that images of clouds, mists and surf, seascapes and landscapes, and faces are reproducible from the chaos models of the original photographs.

4.5 Types of Systems Deploying Chaos Theory

After giving a detailed introduction of the subject in previous sections, this section identifies that Chaos Theory may be used to quantify chaos in dynamical systems depending on the known or unknown equations of motion of the systems. Whether the equations of motions are known or unknown, a number of parameters (including fractal power spectra, entropy, fractal dimension, Lyapunov exponent, etc.) can be used to determine if chaos exists or not in the system. The research reported in the thesis

investigates the Lyapunov exponent as a parameter for analysing the system. Theoretically, the Lyapunov exponent relates to the expansion and contraction in different directions of the spheres of points of the phase space, which are defined accordingly by the Lyapunov spectrum. The rate of separation can be different for different orientations of the initial separation and maximum in a certain direction. When this largest Lyapunov exponent is positive for at least one direction, the system is chaotic. This method is determined by monitoring the evolution of an n-sphere of initial conditions, which deforms into an n-ellipsoidal shape over time due to the local deforming nature of the system (Wolf *et al*., 1985). In this way, this research relies on the use of the Lyapunov exponent to identify chaos, as well as to understand the time dependent evolution of the dynamical states, in terms of the onset and subsidence of congestion, in the traffic system.

Systems and their Equations of Motion	Parameter Values	Lyapunov Spectrum (bits/s)	Lyapunov Dimension
Hénon:			
$X_{n+1} = 1 - aX_n^2 + Y_n$ $Y_{n+1} = bX_n$	$\begin{cases} a = 1.4 \\ b = 0.3 \end{cases}$	$\lambda_1 = 0.603$ $\lambda_2 = -2.34$	1.26
Rossler-chaos:			
$\dot{X} = -(Y + Z)$ $\dot{Y} = X + aY$ $\dot{Z} = b + Z(X - c)$	$\begin{cases} a = 0.15 \\ b = 0.20 \\ c = 10.0 \end{cases}$	$\lambda_1 = 0.13$ $\lambda_2 = 0.00$ $\lambda_3 = -14.1$	2.01
Lorenz:			
$\dot{X} = \sigma(Y - X)$ $\dot{Y} = X(R - Z) - Y$ $\dot{Z} = XY - bZ$	$\begin{cases} \sigma = 16 \\ R = 45.92 \\ b = 4.0 \end{cases}$	$\lambda_1 = 2.16$ $\lambda_2 = 0.00$ $\lambda_3 = -32.4$	2.07
Rossler-hyperchaos:			
$\dot{X} = -(Y + Z)$ $\dot{Y} = X + aY + W$ $\dot{Z} = b + XZ$ $\dot{W} = cW - dZ$	$\begin{cases} a = 0.25 \\ b = 3.0 \\ c = 0.05 \\ d = 0.5 \end{cases}$	$\lambda_1 = 0.16$ $\lambda_2 = 0.03$ $\lambda_3 = 0.00$ $\lambda_4 = -39.0$	3.005
Mackey-Glass:			
$\dot{X} = \frac{aX(t+s)}{1 + [X(t+s)]^c} - bX(t)$	$\begin{cases} a = 0.2 \\ b = 0.1 \\ c = 10.0 \\ s = 31.8 \end{cases}$	$\lambda_1 = 6.30 \times 10^{-3}$ $\lambda_2 = 2.62 \times 10^{-3}$ $\lvert\lambda_3\rvert < 8.0 \times 10^{-6}$ $\lvert\lambda_4\rvert = -1.39 \times 10^{-2}$	3.64

Table 4-1: Examples of Model Systems and their Lyapunov Spectra and Dimensions
(Source: Wolf *et al*. (1985))

Table 4-1 provides a summary of previous research applicable to dynamical systems whose equations of motion are known. For such systems, the Ordinary Differential Equations (ODE) approach provides a relatively straight forward procedure for computing the largest Lyapunov exponents (Benettin *et al*., 1980, Shimada and Nagashima, 1979). Other forms of these systems were also studied by Froehling *et al*. (1981), Fraser and Swinney (1986) and Rosenstein *et al*. (1993).

In these models, it is difficulty to specify the initial condition along the principal axes because these are unknown explicitly. Therefore, it is accepted practise to specify the initial condition with a value as small as the precision of the variable or the value allowable in the computer, and then with the non-linear ODE of motions allow the boundary condition to emerge from a series of iterations. Given the constantly evolving nature of chaotic systems, it is difficult to monitor the deformation given by the exponents along the principal axes because these directions change continuously, especially as the separation becomes rapid. Therefore, it is not possible to guarantee a condition of small separation, which is a requirement for the convergence of the system, making this approach limited in chaotic systems.

In order to address this problem, a 'fiducial' (the centre of the sphere) trajectory is defined by the ODE based on some initial conditions. In addition, the trajectories of points on the surface of the sphere, which are separated infinitesimally from the fiducial trajectory, are defined by linearised equations of motions. The principal axes of the linearised equations of motion are initially an orthonormal vector frame anchored to the fiducial trajectory. In this way, the 'fiducial' trajectory, working in conjunction with its tangent space within a phase space, ensures the principal axis always exists. Using this approach, any further diverges may be eliminated by renormalisation of the separations by a repeated use of the Gram-Schmidt Reorthonormalisation (GSR) procedure. However, Murison (1995) proved that the difference in the separation calculated with or without the GSR procedure is indiscernible.

In experimental data, the equations of motion are unknown and therefore the ODE approach is not directly applicable. In order to handle such data, Wolf *et al*. (1985) proposed a method using an approximate GSR procedure in a reconstructed phase space (described in detail in section 5.4). Unlike the ODE method, which gives only the largest Lyapunov exponent, this approximate GSR method enabled all possible instantaneous Lyapunov exponents in all directions to be calculated. Similarly, for traffic systems, the equations of motion governing its evolution are unknown, therefore the ODE method is inappropriate. Applying the phase space reconstruction plus the approximate GSR procedures however needs modification to enable not only the positive but also the non-positive values of the Lyapunov exponents to be calculated. In this research, the method applied is based on the finite-size approach proposed by Károlyi *et al*. (2010) discussed in Section 5.7. Analysing the system in this way will make available the entire range of exponents of the system's

evolution that could be processed further in order to understand how chaotic systems transition from one state to another.

By incorporating the conclusions from Murison (1995), the research reported in this thesis makes use of the phase space reconstruction approach and calculates the separations based on the fiducial trajectories only. However, instead of the approximate GSR procedures, the research applies a modified version of the finite-size Lyapunov exponent method proposed by Károlyi *et al*. (2010). In this approach, the final separation of every individual point p_i and the average separation of points within a specified distance d_s from p_i is calculated.

There are few studies in the literature which have reported practical uses of the Lyapunov exponent in traffic systems (see Section 3.5). For example, Shang *et al*. (2005) applied the Lyapunov exponent to simply determine the presence or otherwise of chaos in traffic data whilst Hui *et al*. (2005) used the parameter in combination with a proposed radial base function neural network to predict chaos. This study presents a first attempt at using the Lyapunov exponent from an **unknown** dynamical (traffic) system to understand the evolution of the flow regimes in order to gain an insight into the cyclical nature of the dynamic states resulting from the traffic characteristics. This study is important because chaos gives rise to stop-and-start conditions and breakdown of traffic flow causing traffic delay, which is undesirable to motorists. Such conditions also increase fuel consumption, tailpipe emissions and the likelihood of accidents, and decrease the vehicle throughput across the junctions. Therefore, this study will investigate the use of the Lyapunov exponent to identify chaos in order to create awareness of impending unstable conditions and thus allow mitigating strategies to be implemented in real-time in order to improve the level of service and safety of the road system.

4.6 Conclusions

It is concluded that chaos is ubiquitous and widely ingrained in many natural processes such as the earth's system and those that relate to the human system, such as rhythmic patterns in the function of the human heart, brain and immune systems, and the physiological process such as menstruation through to menopause (Kumar and Hedge, 2012). The above discussion demonstrates that Chaos Theory has wide applications in numerous fields, and the application to UTC system traffic may be feasible but requires in-depth analysis in order to confirm this.

Chaotic and random systems are different because chaos is predictable in the short-term whilst randomness is unpredictable. Chaos is a product of non-linear dynamic systems that are highly sensitive to initial conditions. The literature review also identified that non-linearity is *sine qua non* for the development of chaos, but the presence of non-linearity is not always a guarantee for chaos (Lorenz, 1993b). Non-linearity ultimately gives rise to randomness (but not necessarily chaos); however, when both randomness and chaos are present there is always a progression of states from stable cycles through chaos before randomness.

In this research, it is important to establish the validity of the assumption that road traffic is chaotic (rather than random) system. According to the literature review, a discrete map may indicate chaos (Dreyer and Hickey, 1991), as well as proof of a progression of states from stable cycles through bifurcating hierarchal stable cycles with a period of 2^n (Butler, 1990, Coullet *et al*., 2005, Gleick, 1988). Other tests include the evaluation of embedding dimension, correlation dimension or positive Lyapunov exponents for the system (Jianming *et al*., 2003). These will give a general indication of the possibility of chaos in the traffic system due to the interactions of highly variable physical and human elements (such as vehicles, weather and road surface conditions, road layout, topography and pedestrian activities). Such analysis will require time series information of traffic stream variables (such as flow, occupancy, speed) for the assessments.

Once this initial first step is established with the available datasets, there will be a firm basis to apply Chaos Theory to determine the characteristics that will enable practical application to traffic networks. It is anticipated that Chaos Theory will enable traffic forecasts to be determined more reliably to allow traffic management strategies to be implemented ahead of real-time to maintain under-saturated operating conditions in the network.

This Chapter concludes Part 1, and against this background, Part 2 begins in Chapter 5 that deals with the mathematical framework that supports Chaos Theory in order to gain a firm understanding of its algorithms. It presents a step-by-step procedure for applying Chaos Theory, and explains the key parameters (such as delay time, embedding dimensions, correlation dimensions, and Lyapunov exponents), which are subsequently coded in Microsoft C# program for application in this research.

PART II: MATHEMATICAL FRAMEWORK OF CHAOS THEORY AND RESEARCH METHODOLOGY

Chapter 5. Mathematical Framework of Chaos Theory

5.1 Introduction

This Chapter will focus on the mathematical framework that underpins the study of chaotic systems. A good understanding of the fundamental mathematical concept is crucial to successful application of Chaos Theory in this research work. Broadly, it discusses the Phase Space Reconstruction Method, which deploys the embedding coordinate technique to transform a one-dimensional time series data into a high-dimensional phase space. This transformation is necessary to enable a one-to-one representation of the true, but unknown, dynamic system for subsequent determination of chaotic parameters of the system (see Section (Zhu *et al*., 2010).

In order to present the mathematical framework for this research, Section 5.2 introduces the technique of the phase space reconstruction for the transformation of a time series data into a phase space to enable the study of the system's chaotic properties. The subsequent discussions will focus on the procedures used to determine important parameters of the system. Section 5.3 illustrates the overall process for applying Chaos Theory, whilst Section 5.4 describes the process of the phase space reconstruction. Section 5.5 discusses the delay time (or lag) based on the autocorrelation and mutual information functions, whilst Section 5.6 deals with the algorithms for estimating the correlation and embedding dimensions. Finally, Section 5.7 illustrates how these parameters are important inputs into the reconstructed phase space used to calculate the Lyapunov exponents, which gives an indication of the dynamic states of the system.

5.2 Analytical Approach

The two main analytical approaches in non-linear data analysis are (i) the delay-coordinate embedding technique and (ii) the time-frequency method based on the Hilbert transform (Lai and Ye, 2003). The Hilbert-transform method performs well for nonlinear and/or non-stationary time series, and in the analysis of the distribution of instantaneous physical frequencies of, for example, signal processes. The literature review did not find evidence of the Hilbert transform method applied to traffic studies. However, the delay-coordinate method has been widely used in many traffic studies, and specifically proven useful for analysis where the dynamic system that is responsible for the observed time series is of low dimensional, deterministic or mostly deterministic and the influence of 'noise' is minimal. The research presented in this thesis adopted the delay-coordinate embedding (DCE)

technique for analysing the traffic data because relative to Hilbert transform it has been used more widely in traffic studies and thus provides a good foundation for this work.

Evidence suggests that Jianming *et al*. (2003), Hui *et al*. (2005) and Shang *et al*. (2005) used the DCE technique to reconstruct the phase space from scalar quantities. Packard *et al*. (1980) applied the DEC technique to analyse a time series data from a single coordinate and constructed a phase portrait and determined the dimensionality of an attractor, whilst Eckmann and Ruelle (1985) used this method to examine the dynamics (turbulence) of a fluid system. Froehling *et al*. (1981) noted that the method was sensitive to instrumental *noise*, and therefore required large amount of data points. Frazier and Kockelman (2004) comment on the data issue indicating that the lack of extensive data or sampling at the wrong frequency will not accurately capture important dynamics of the system. For any system, the time series information is the main data requirement for analysis using Chaos Theory. The remainder of this section therefore illustrates how time series information can be applied in this analysis.

5.2.1 Time Series Information

The time series information normally consists of data points sampled at regular and frequent intervals, which may be restricted to either one or several variables at each sample time.

Suppose an observed time series consists of the following k-vectors (where k is the number of variables), which are:

$$\bar{X}_1,\ \bar{X}_2, \bar{X}_3 \ldots\ldots\ldots \bar{X}_k \qquad \textit{Equation 5-1}$$

For the N-observations of each variable, the vector elements will be:

$$\bar{\mathrm{X}}_{\mathrm{i}} = [X_i(1), X_i(2), X_i(3) \ldots \ldots \ldots X_i(N)] \qquad \textit{Equation 5-2}$$

Consequently, for a time series observation of only one variable (one-dimensional scalar quantity), Equation 5-2 becomes:

$$\bar{X} = [X(1), X(2), X(3) \ldots \ldots \ldots X(N)] \qquad \textit{Equation 5-3}$$

In summary, this theory deems that by using the vector elements in Equation 5-3, it is possible to create a phase space (based on the phase space reconstruction theory), which is a representation of the true but unknown system in Equation 5-2.

5.3 Deploying Chaos Theory

The mathematical framework for the research is underpinned by the Phase Space Reconstruction theory (Jianming *et al*., 2003). Figure 5.1 presents the four main stages, whilst a fifth stage makes use of the outputs from the reconstructed phase space for further analysis

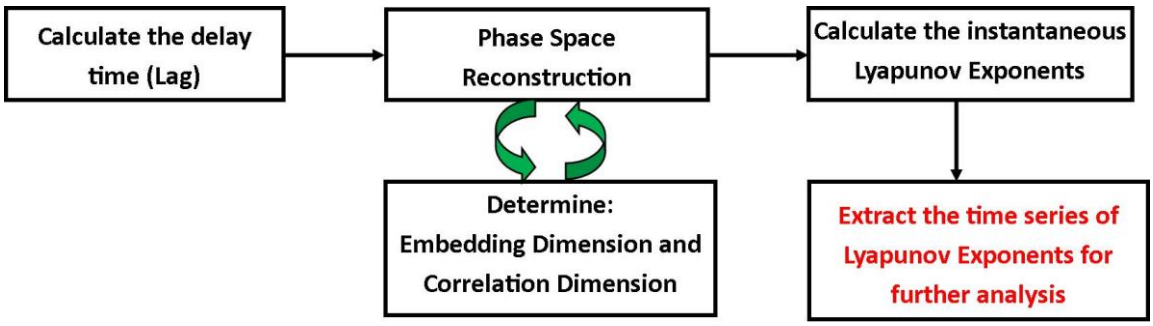


Figure 5.1: Overall Summary of the Chaos Theory Analytical Procedure

The key steps include:

Step 1. Construct a generic model of the phase space of the traffic flow series using the Packard and Takens Phase Space Reconstruction theory.

Step 2. Determine the optimum autocorrelation lag (delay time).

Step 3. Input the optimum lag and reconstruct the phase space for any arbitrary number of embedding dimensions

Step 4. Calculate saturated embedding dimension and correlation dimension.

Step 5. Reconstruct the phase space for the saturated embedding dimension and calculate the instantaneous Lyapunov exponents.

Step 6. Undertake analyses including assessing the variability of Lyapunov exponents across the loops in the network based on the descriptive statistics and the testing of a number of hypotheses to answer the research questions 3, 4, 5, 6 and 7.

5.4 Phase Space Reconstruction Theory

The DCE technique developed by Takens (1981) provides the mathematical foundation of the phase space reconstruction method for the non-linear analysis of dynamic systems. Using this method, Takens (*ibid*.) indicated that it was possible to construct the underlying dynamics of a nonlinear system using the time series of the system's evolution. He proved the existence of a one-to-one correspondence between the actual but unknown dynamic system and the reconstructed dynamic system. That is, Takens (*ibid.*) showed that it was

possible to understand the characteristics of a dynamic system based on only the time series information without the need of any additional information (or evolutionary equations).

Often, the equations governing these systems are generally unknown and their derivation often involves difficult mathematical computations. Therefore, the delay-coordinate technique overcomes the challenges of deriving complex mathematical equations for the system, and instead simply focuses on the time series data to understand the system's evolution. The raw time series may not be suitable to understand the system's dynamic properties, but the transformed data enables key parameters to be determined thus enhancing our knowledge of the system (Lai and Ye, 2003). In traffic studies, this technique can enable understanding of the temporal changes of the stable, unstable and chaotic nature of the network, as well as the evolution of free flowing and congested states.

The main thrust of this technique is that even though the dynamical property of a system may involve several variables, the evolution of an arbitrary component (one variable) of the system is dependent on its other relative components (other variables). For example, speed, flow and density relationships. This influence means that the developing process of any arbitrary component contains relevant information of the other components. Therefore, the theory suggests that one observed variable can reveal detailed information of the whole system (Jianming *et al*., 2003).

5.4.1 Notation

Suppose a scalar observed time series:

$$\mathrm{X(i)} = \mathrm{x(t_i)}, \text{for } t_i = t_0 + \mathrm{i}\Delta\mathrm{t}, \mathrm{i} = 1, 2, \ldots\ldots \mathrm{N} \qquad \textit{Equation 5-4}$$

where $\Delta\mathrm{t}$ is a constant sampling interval; t_0 and N are the initial and final times respectively within the sampling interval; and t_i evaluates a specific time for the time series observation of interest.

According to the delay coordinate embedding method, each scalar measurement $\mathrm{x(t)}$ is a component of a vector $\mathrm{X(t)}$ of m-dimensions and τ is the delay time (lag), as follows:

$$\begin{aligned}
\mathrm{X(1)} &= [\mathrm{x(1)}, \mathrm{x}(1+\tau), \mathrm{x}(1+2\tau), \ldots\ldots\ldots, \mathrm{x}(1+(\mathrm{m}-1)\tau)] \\
\mathrm{X(2)} &= [\mathrm{x(2)}, \mathrm{x}(2+\tau), \mathrm{x}(2+2\tau), \ldots\ldots\ldots, \mathrm{x}(2+(\mathrm{m}-1)\tau)] \\
\mathrm{X(3)} &= [\mathrm{x(3)}, \mathrm{x}(3+\tau), \mathrm{x}(3+2\tau), \ldots\ldots\ldots, \mathrm{x}(3+(\mathrm{m}-1)\tau)] \\
&\ldots\ldots\ldots \\
\mathrm{X(k)} &= [\mathrm{x(k)}, \mathrm{x}(\mathrm{k}+\tau), \mathrm{x}(\mathrm{k}+2\tau), \ldots\ldots\ldots, \mathrm{x}(\mathrm{k}+(\mathrm{m}-1)\tau)]
\end{aligned} \qquad \textit{Equation 5-5}$$

where:

- m is the embedding dimension
- $N = k + (m-1)\tau$ is the length of the time series

Equation 5-5 suggests that the first observation in times series will be a point with coordinates $x(1), x(1+\tau), x(1+2\tau), \ldots\ldots\ldots, x(1+(m-1)\tau)$ in the phase space, the second observation becomes point $x(2), x(2+\tau), x(2+2\tau), \ldots\ldots\ldots, x(2+(m-1)\tau)$ and so on. All the k vectors will form a reconstructed phase space that maintains both the geometric and topological structure as well as the dynamic properties of the original system (Jianming *et al*., 2003). Therefore, the reconstructed phase space preserves the invariants (such as the dimension and Lyapunov exponents) of the true but unknown system, which was responsible for the time series observations of variables such as flow, occupancy and speed. Consequently, it is possible to investigate the dynamics of the original system based on the phase space constructed from the time series data.

The above notation indicates that every reconstructed vector for a data point at an arbitrary position p will require the data point at a relative position $p + (m-1)\tau$ to define its coordinates correctly. If k is the last data point of the time series with a valid reconstructed vector, then the reconstructed phase space will contain $N - k$ fewer vectors than the length of original time series. This means therefore that it is not possible to construct any vector beyond $X(k)$ because the value of the m^{th} co-ordinate in the reconstructed vector $X(k)$ will be equivalent to the value of the N^{th} observation (i.e. the last observation of the original time series). Given the length of the original time series N, any delay coordinate embedding vector beyond $X(k)$ will require components beyond the N^{th} component of the original data series, which would be unavailable and unknown. The shortfall in length of the reconstructed vectors could reduce the number of points in the phase space and consequently limit its intended uses. This is also reason why a large number of observations is required for analysing dynamical systems. It is reasonable therefore to agree with Froehling *et al*. (1981) and Frazier and Kockelman (2004) that large volumes of data of appropriate sampling frequency is key to applying Chaos Theory.

Lai and Ye (2003) identified some of the parameters that can be derived from the reconstructed phase space, as the Kolmogorov entropy, the relative weights of the deterministic and stochastic components embedded in the time series and the fractal dimension (which defines the self-similarity property of the underlying dynamic system).

In addition, it is useful in determining the paths of the unstable periodic orbits that constitute the skeleton of the invariant set responsible for the observed dynamics as well as the degree of sensitivity of initial conditions as characterized by the Lyapunov exponents.

Equation 5-5 reveal the delay time (τ) and embedding dimension (m) as two important parameters whose values need to be determined explicitly in order to fully construct the phase space. Without these, it will be impossible to construct the phase space accurately and therefore the key invariants of the system cannot be determined. In order to demonstrate the uses of these parameters, let us consider a dynamic system with parameters $m = 3$ and $\tau = 1$, where the regular interval sample data of N values are $4,\ 2,\ 6,\ 1,\ 5,\ 3,\ 8, 9, \ldots$ In the reconstructed phase space, the coordinates of the first three points will be: $X(1) = (4,2,6);\ X(2) = (2,6,1);\ X(3) = (6,1,5)$, etc. Roux *et al*. (1983) is of the view that the determination of these values involves a trial and error procedure, which can be cumbersome involving the test of a number of embedding dimensions and delay times in order to select the optimum saturated values. In this illustrated example, a range of τ (e.g. $1, 2, 3 \ldots 10$) will tested to determine the optimum lag time based on an autocorrelation value of 0.4 (see section 5.5.2), which serves as input into Equation 5-5 in order to estimate the embedding dimension. Similarly, the m values (e.g. $1, 2, 3 \ldots 10$) may be tested for a range of distances of separation between points in the phase space in order to determine the saturated value using the function of the embedding dimension (see Section 5.6.2). Therefore, this procedure could involve a large amount of iterations that may be difficult to undertake manually without computer software, and this confirms an initial perception in Section 4.3.2 of the indispensability of the computer in deploying Chaos Theory. Sections 5.5 and 5.6 provide a more complete description of the process of calculating these parameters.

5.5 Delay Time

5.5.1 Purpose of the Delay Time (Lag)

The purpose of the delay time is to determine the length of interval when the individual observations of the time series are neither too independent nor less independent on each other. Therefore, this parameter provides an indication of the delayed time intervals that observations correlate moderately for subsequent analysis. The lag is useful for determining the coefficients of the independent variables (or coordinates) of the points in the reconstructed phase space of the dynamic system. In addition, it is a crucial initial step

prior to determining the embedding dimension that defines the number of independent variables for describing the phase space accurately.

It is important that the calculation of the delay time yields the optimum value as the choice of an incorrect lag could render the reconstructed phase space incapable of its intended uses. In order to be sure that the reconstructed phase space of a time series dataset is of practical use, it is necessary to specify an optimum time interval such that the results are relatively insensitive to the chosen time interval (Froehling *et al*., 1981). According to Lai and Ye (2003) and Jianming *et al*. (2003), the choice of a small value of τ makes the adjacent components too correlated to serve as independent variables. This makes the coordinates appear almost consistent such that the trajectories dwindle into a straight line in the phase space. On the other hand, a large τ is undesirable as they generate neighbouring coordinates that are highly uncorrelated and potentially give rise to phase space distortions – a phenomenon where the trajectory of a simple geometric curve could appear to be complex. The effect of the extreme values is that it makes it impossible to establish a one-to-one correspondence between the reconstructed phase space and the true but unknown system.

The literature review identified two methods for selecting the optimum lag: (i) autocorrelation coefficient (AC) and (ii) mutual information (MI) (Boker, 1996, Jianming *et al*., 2003). There are mixed views on the reliability of both approaches among authors; for example, Shang *et al*. (2005) state that "*currently, there is no agreement among authors on which of the two methods should be used in order to determine the delay time*". Reiss (2001) and Doniosio *et al*. (2003) are critical of the autocorrelation method on the basis that it provides a linear measure and may not always provide an accurate measure of serial dependence if there is significant non-linearity in the dataset. Reiss (*ibid.*) for example indicates a preference for the mutual information method over the autocorrelation function given its ability to measure both linear and non-linear correlations. However, Reiss *(ibid.)* appears to contradict this view in the conclusion to his work, where he assumes a neutral position whilst cautioning that neither mutual information nor autocorrelation coefficient method is suitable for all situations and therefore Reiss *(ibid.)* recommends the use of both functions.

With reference to the mutual information method, Reiss *(ibid.)* states, "*this is preferred over the autocorrelation function, because the mutual information can measure nonlinear correlations whereas the autocorrelation function only takes linear correlations into*

account. However, neither is definitive so it is recommended that both functions be used." However, Jianming *et al*. (2003) indicates a preference for the autocorrelation coefficient method for analysing traffic data, stating that the, *"Mutual Information method determines delay time according to seeking the corresponding delay time with the first minimum of mutual value. But this method cannot guarantee that the mutual information always exists. Besides, the computation of mutual information is very difficult."* This statement may imply that the mutual information method is appropriate when this is a property of the system; otherwise, the autocorrelation method is a better choice. From the foregoing, it will appear that the choice of a method seems to depend on what works well in the particular application.

According to Boker (1996) the mutual information and autocorrelation functions are similar when applied within a single time series because both methods seek to indicate a degree of dependence. However, mutual information has the ability to measure mutual dependence even when the system is non-linear or chaotic, whereas autocorrelation may suggest an independent random process. This means that the autocorrelation method may be disadvantaged because it assumes a linear property as an underlying characteristic of the system when calculating a value for dependency, and therefore the value may be inaccurate if the underlying process is non-linear. Furthermore, autocorrelation requires that the distribution of the original time series and its lagged components be jointly normal but mutual information requires no such assumptions. It is important to note however that Frede and Mazzega (1999) proved that the lag computed from both methods yield equivalent results. Boker (1996) confirms this but indicates that both methods produce the same value of the optimal delay if the system is linear.

In theory, the function of the mutual information provides the optimal lag based on the first minimum of the curve [i.e. $\frac{d}{dt}f(MI) = 0$], whereas the autocorrelation methods calculates this value as the first zero crossing of the autocorrelation function [i.e. $f(AC) = 0$] (Wang *et al*., 2005). For non-linear systems, Boker (*ibid*.) stated that the first minimum of the mutual information function only provides a minimum value for the optimum lag, and therefore the choice of the optimum lag must provide a balance between eliminating autocorrelation and preserving non-linear dependency. However, there is no exact science for determining the optimum lag in this case, which allows for some degree of subjectivity in making this decision. Boker (*ibid*.) further argued that the shortest possible lag is optimum because it minimises the degree of autocorrelation in the time series. However,

Sauer *et al*. (1991) suggests that there is very low probability that the choice of the lag from either methods can affect the phase space construction of the true yet unknown dynamical process of the studied system.

By considering their mathematical formulae, the mutual information method appears to be complex with a crucial step that involves the computation of probabilities whereas the autocorrelation coefficient method is relatively simple mainly requiring the computation of the mean of the entire datasets. Whereas the parameters for using autocorrelation coefficient method are directly available and can be computed directly from the data, the mutual information method is logarithmic and requires probabilities as input that in turn increases the computation process making it cumbersome. Therefore, the autocorrelation method emerges as the convenient and less complicated method for calculating the optimum lag time.

Consequently, the research presented in this thesis used the autocorrelation coefficient method because of its simplicity and evidence of application in previous research. Autocorrelation refers to the correlation of a time series with its own past and future values. This real-world property helps to determine the tendency of the time series to remain in the same state from one observation to the next. Regarding time series traffic data, a positive autocorrelation coefficient suggests for example the likelihood of an unstable condition in the future over a link detector because the link was unstable in the past. The value of the autocorrelation can vary depending on the chosen lag, and therefore it is usual to sample the lags and evaluate the autocorrelation function in order to determine the optimum lag (or ideal lag interval). Wang *et al*. (2005) identified different approaches that evaluate the lag as the autocorrelation function that reaches a certain value, for example $\frac{1}{e}$, $1-\frac{1}{e}$, 0.1 or 0.5, instead of zero; because a value of zero ensures no correlation between the data points giving rise to disconnected trajectories in the phase space. Therefore, in practise, the lag is chosen with a degree of subjectivity. Jianming *et al*. (2003) successfully applied an empirical value of $\frac{1}{e}$ (or 0.4). Further evidence is given by Ratner (2009) in section 9.5.1, which indicates a threshold value of 0.4 where the autocorrelation of the time series is neither strongly nor weakly correlated. Consequently, this research also adopted this figure of 0.4. Initial trials in sections 7.2.1 and 7.3.1 using a Microsoft C# program yielded satisfactory outcomes, and therefore all calculations presented here used this method. Section 5.5.2 is devoted to describing the mathematical algorithms for the autocorrelation coefficient method.

5.5.2 Autocorrelation Coefficient Method

The autocorrelation coefficient method measures the correlation of a time series with its own past and future values. The autocorrelation function for any arbitrary delay time (lag time) is given by:

$$C(\tau) = \frac{\sum_{i=1}^{N-r}[x(i) - \bar{x}][x(i+\tau) - \bar{x}]}{\sum_{i=1}^{N}[x(i) - \bar{x}]^2} \qquad \textit{Equation 5-6}$$

where:

- $\bar{x}$ is the mean of observed data series
- $x(i)$ is the preceding time observation
- $x(i + \tau)$ is the observation at the lagged time (τ).

Equation 5-6 will be revisited in Section 6.6.2 for the coding in the C# program for computing the lag of the spatial and temporal research data. The first step is to calculate the autocorrelation coefficient of a time series x(i) and the lag time value $x(i + \tau)$ under a number of pre-selected time lags. From the calculated autocorrelations, it is then possible to select the optimal time delay from the plot of the correlation coefficients against the delay time (lag). This graph is known as a Correlogram (or autocorrelation plot). Correlograms are commonly used to determine the presence of randomness within data sets, or to establish whether indeed a relationship exists between two or more adjacent observations. If randomness exists, the autocorrelations should be close to zero for one or all time lags; otherwise, one or more autocorrelations will be significantly non-zero (Xiufen *et al*., 2002).

The widely used approach estimates the optimal delay as the duration of a lag such that $C(\tau)$ is equivalent to approximately $\frac{1}{e}$. This optimal delay is the equivalent of the delay or lag duration that corresponds to an autocorrelation of 0.4 (as described above), an indication of when the autocorrelation of data points is moderate (Jianming *et al*., 2003). The next two sections explain the mathematical concepts underlying fractals and dimensions (introduced at Sections 4.3.3 and 5.6.3 respectively) in which the lag plays a vital role.

5.6 Dimensions

Chaos Theory uses a number of dimension to describe the characteristics of a system under investigation. These include the fractal dimension, embedding dimension and correlation dimension, which are discussed in sections 5.6.1, 5.6.2 and 5.6.3.

5.6.1 Fractal Dimensions

The term 'fractal' describes systems with fractional dimensional. Fractals possess a property of self-similarity at all scales by magnifying several pieces of the same objects, each of the magnified pieces is indistinguishable to the whole system. The theoretical fractal dimension of an object equals the set of Euclidean or topological dimensions that ordinarily describe the geometric shapes. For example, a 0-dimensional fractal sets describe points; a 1-dimensional set describes an object by length only. Surfaces with both length and width are 2-dimensional objects, whilst those with length, width and depth are 3-dimensional objects. However, some strange objects in nature have fractions for dimensionality, unlike the simple geometrical objects that have whole number dimensionality (Lorenz, 1993b).

Both fractal and non-fractal dimensions help to determine the number of equations (or unknown variables) necessary to describe the shape of the attractor. Whilst non-fractal dimensions are normally associated with simple objects, fractal dimensions define dimensions of complex systems, or as Mandelbrot (1983) puts it "*objects that are described as oddly shaped without any term to describe them in either the sciences or the arts*". Where a fractal dimension exists, the number of equations is incremented to the next higher integer (Orcutt and Arritt, 1995).

Fractal dimensions are useful in describing natural objects and evaluating trajectories of dynamic systems (see section 4.3.3). The fractal dimension is a statistical index that characterises the complexity of an object according to the ratio of how the detail in the fractal pattern changes with the scale of measurement (Mandelbrot, 1983).

The mathematical derivation by Christos and Ibrahim (1994) is useful in describing the relationship between the fractal dimension and the scale of measurement. Let us consider a simple object (of self-similarity property) of linear size l that consists of a set of points in an ordinary Euclidian space of dimension D, which may also define the embedding dimension. By dividing the length into equal parts by a factor of $1/l$ in each spatial direction, the number of similar smaller objects, required to describe the original object is given by:

Equation 5-7

$$N = \left(\frac{1}{l}\right)^{-D} = l^{D}$$

Figure 5.2 illustrates the number of similar small objects resulting from the combinations of different values of l and D, according to Equation 5-7. For example, if each side of a cube of dimension D equals to 3 is divided into equal parts by a factor of $1, \frac{1}{2}$ and $\frac{1}{3}$ (which are the inverse of l), the resulting number of similar small objects are 1, 8 and 27 respectively.

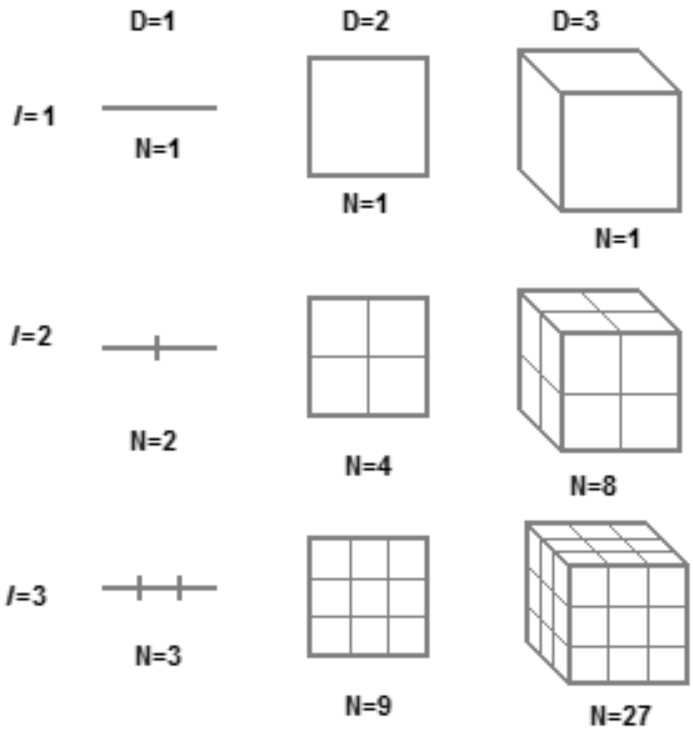


Figure 5.2: Traditional Notations of Geometry for Defining Scaling and Dimension
(Source: Brendan (2007))

By taking the natural log of both sides of Equation 5-7, the fractal dimension of a complex object is given by:

$$D = \frac{ln\ N}{\ln l} = \lim_{l \to 0} \frac{\ln N}{\ln l} \qquad \textit{Equation 5-8}$$

Where in the limit $l = 0$, $D = \ln N$, which is non-integer.

5.6.2 Correlation and Embedding Dimensions

The correlation dimension is a type of fractal dimension used to determine the possible fractal dimension of fractal objects. In time series analysis, the correlation dimension is used to determine the degree of complexity of the chaotic structure of the underlying dynamics of the system (Shirer *et al*., 1997). In this application, the correlation dimension provides an index to judge the type and complexity of an attractor, which also describes the

level of stability and the ability to determine the entire changing process of the system. This index is useful in determining the rough estimate of the embedding dimension, which defines the dimension of the high order phase space capable of containing the evolutionary process of the chaotic attractor, in order that the orbits do not overlap in the phase space (Jianming *et al*., 2003). In other words, this index helps to specify the size of the phase space that would embed the evolution process of the chaotic system without any distortions.

The process of constructing the phase space and subsequent determination of the correlation dimension and embedding dimension is as follows, and the relevant equations will be used in analysing the research data and therefore are revisited in Sections 6.6.3 and 6.6.4.

Consider two points in the reconstructed phase space:

$$\mathrm{X(j)} = [\mathrm{x(j)}, \mathrm{x(j+\tau)}, \mathrm{x(j+2\tau)}, \ldots\ldots\ldots\ldots\ldots\ldots \mathrm{x(j+(m-1)\tau)}]$$
$$\mathrm{X(i)} = [\mathrm{x(i)}, \mathrm{x(i+\tau)}, \mathrm{x(i+2\tau)}, \ldots\ldots\ldots\ldots\ldots\ldots \mathrm{x(i+(m-1)\tau)}] \qquad \textit{Equation 5-9}$$

Let $r_{ij}(m)$ denote the distance between them, so that:

$$r_{ij}(m) = \left\| X_i - X_j \right\| \qquad \textit{Equation 5-10}$$

Given a certain distance r, the ratio of the number of paired phase points whose distance of separation is less than r to all the pairs of phase points is given by the accumulation function:

$$C_2(r,m) = \frac{1}{N(N-1)} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \theta(r - \left\| X_i - X_j \right\|) \qquad \textit{Equation 5-11}$$

where $\theta(z)$ is the Heaviside function, which is given by:

$$\theta(z) = \begin{cases} 0, & z < 0 \\ 1, & z \geq 0 \end{cases} \qquad \textit{Equation 5-12}$$

$C_2(r,m)$ describes the probability that two points in the phase space are separated by a distance less than or equal to r. Hence, this condition of $0 < C_2(r,m) < 1$ must be valid. The ranges of values define the "scaling" region, which is the region sandwiched between the "depopulated" and "saturated" regions of each curve. The pairs of points whose separation is below the lower end r-range is a "depopulated" region, whilst at the upper

end the r-value above which there is no further increase in the correlation function is the "saturated" point.

In order to enable the selection of the appropriate correlation and embedding dimensions, it is important to examine the relationship that exists between $C_2(r,m)$ and r. As the correlation dimension is a form of fractal dimension, it is possible to adopt the approach of the latter in the derivation of $C_2(r,m)$, where r is the scaling factor. Firstly, let us consider an extremely small value ∂ (which represents a certain multiple of r). In the case of a 1-dimension and linear attractor, the number of correlation points will be proportional to r/∂. If the attractor is 2-dimensional, then the number of correlation points is proportional to $(r/\partial)^2$. Consequently, the number of correlation points for a $D_2(m)$-dimensional attractor is directly proportional to $(r/\partial)^{D_2(m)}$. Hence:

$$C_2\,(r,m) \propto \left(\frac{r}{\partial}\right)^{D_2(m)} \qquad \textit{Equation 5-13}$$

$$C_2\,(r,m) = f\left(\frac{r}{\partial}\right)^{D_2(m)} \qquad \textit{Equation 5-14}$$

Let $\partial = 1$ and $f = 1$, then:

$$C_2\,(r,m) = (r)^{D_2(m)} \qquad \textit{Equation 5-15}$$

Taking the natural log of both sides:

$$D_2(m) = \frac{\ln C_2(r,m)}{\ln r} \qquad \textit{Equation 5-16}$$

$D_2(m)$ is the function of the embedding dimension, which is the estimated value of the Correlation Dimension.

The G-P Algorithm (GPA), a method by *Grassberger and Procaccis* (Xue and Shi, 2008, Wang *et al*., 2005, Zhang and Liu, 2007), is widely used to select the desired correlation dimension and embedding dimension. The method is quite popular due to its simplicity and the additional benefits of the intermediate steps for calculating both correlation dimension and entropy. It is extremely sensitive to the number of data points, reconstruction delay time, embedding dimension and highly reliant on large, noise-free datasets. Given the large number of data points of the order of thousands required, the Lyapunov exponent may be preferred to the correlation dimension as an indicator of chaotic

behaviour because the scale of the data gives a more likely representation of the underlying process within the system (Rosenstein *et al*., 1993).

In practise, the first step to determine the correlation dimension is to select an arbitrary set of m-values and r-values, and the computation process is dependent on a combination of these values. For each m –value, the process involves computing $C_2(r, m)$ for all r-values within the pre-defined range. The next step is to generate a G-P curve of $C_2(r, m)$ against r for all $m -$ values independently i.e. a log-log profile of the accumulation function versus the pre-defined distances (r-values) for any arbitrary number of m-values. The *rationale* is to construct a phase portrait for a number of pre-selected m-values, starting from some arbitrary minimum and proceeding through a stepwise increment by one, until the phase portrait fails when an extra dimension is added. In other words, this process requires increasing the embedding dimension gradually to reach the saturated embedding dimension, where any further increase in embedding dimension does not change the value of the correlation dimension. The saturated embedding dimension is the preceding dimension before the phase portrait failure. According to Zhu *et al*. (2010), the slope of the saturated curve (i.e. the curve whose m-value is selected as the embedding dimension, according to Equation 5-16 is the correlation dimension of the attractor.

It is common practise to examine the sequential progression of the log-log m-curves. The slope of the linear (straight-line) part will change gradually before stabilising for the remainder of the curves. The stabilised slope is the desired correlation dimension and the embedding dimension m at which this occurs is the saturated embedding dimension. This process normally involves a visual inspection of the curves to select the saturated embedding (and subsequently the correlation dimension), which could be problematic if the change in gradient is not obvious.

Probably it is in response to this that Shang *et al*. (2005) proposed the least squares fit of a straight line plotted over the "scaling region" to select the saturated embedding and correlation dimensions. This revised method plots the local slopes (correlation exponents) against the corresponding embedding dimension values. The saturated value of the correlation exponent gives the expected optimum, whilst the nearest integer above the embedding dimension defining this saturated value indicates the minimum number of parameters (coordinates) required to model the dynamics of the attractor. In this method, if the correlation exponent increases indefinitely with increases in the embedding dimension, it is an indication that the system has a stochastic behaviour.

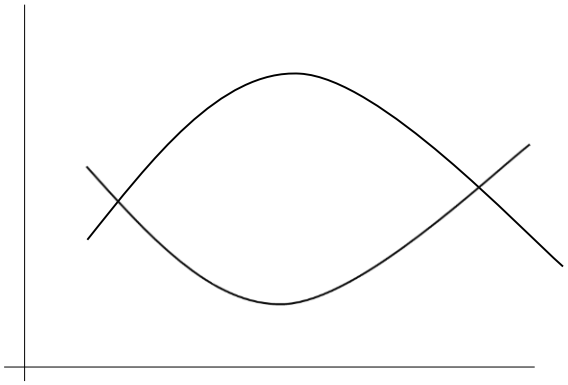
5.6.3 Relationship between Correlation and Embedding Dimensions

According to the theory of Takens (1981), the saturated embedding dimension (m^*) should satisfy $m^* \geq 2D_2{}^* + 1$, where $D_2{}^*$ is the saturated correlation dimension. This is to ensure that the embedding dimension is sufficiently large in order to preserve the property of chaotic attractors, in which their orbits do not intersect or overlap with each other. If m^* is too small the attractor cannot spread out completely; when m^* is too large, there is the likelihood of phase space distortions.

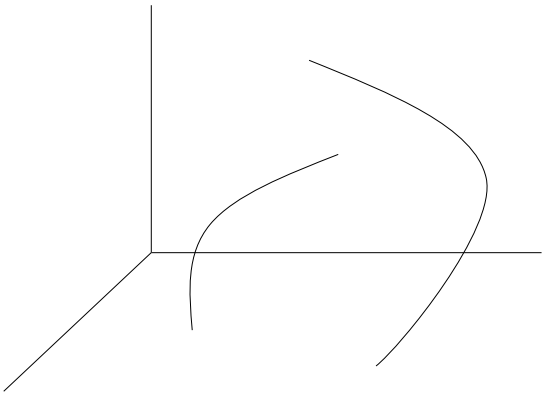
Rosenstein *et al*. (1993) proved that the orbits intersect or overlap when the attractor has an embedding dimension that is less than that needed to describe the dynamics of the system. This is because chaotic systems could become effectively stochastic due to inadequate capacity to accommodate the true dynamics of the system. Takens (1981) criteria for selecting the appropriate embedding dimension seems to be a very robust one, given that Rosenstein *et al*. (1993) reported of an algorithm for calculating chaotic properties that performed well, even for m^*-values below the Takens' criteria. The work of Echman in 1985 further strengthens this position, and provides proof that a range of $D_2{}^* + 1 \leq m^* \leq 2D_2{}^* + 1$ is acceptable (Jianming *et al*., 2003, Zhu *et al*., 2010, Fenghua *et al*., 2005). However, many authors who made use of this criteria also referenced Takens' work. Based on this evidence, Takens' criteria is applied to this study.

Lai and Ye (2003) derived a mathematical relationship to prove the validity of Takens' theorem. If one considers the set of intersection of two smooth surfaces of dimensions d_1 and d_2 in an M-dimensional space, the dimension of the intersection set $d_I = d_1 + d_2 - M$. If $d_I \geq 0$, a self-intersecting (or generic) intersection occurs. Figure 5.3(a) shows the intersection of two one-dimensional surfaces d_1 and d_2 in a two-dimensional plane $M = 2$ will be a point, and the calculated dimensionality evaluates as zero (i.e. $d_I = 0$). Similarly, if $M = 3$, $d_1 = 2$ and $d_2 = 1$, then $d_I = 0$ and an intersection of the curves will occur (see Figure 5.3(c)).

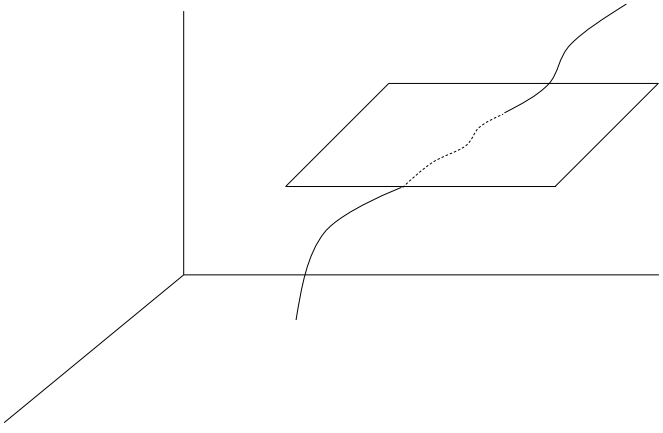
Finally, Figure 5.3(b) indicates that there will be no intersection of the curves because $M = 3$, $d_1 = d_2 = 1$, and therefore $d_I = -1 < 0$. Therefore, this phase space is considered large enough to accommodate the trajectory of the system's evolution.



(a) $d_1 = d_2 = 1$ and $M = 2$ ($d_I = 0$ therefore generic intersection)



(b) $d_1 = d_2 = 1$ and $M = 3$ ($d_I = -1$, which is < 0 therefore non-generic intersection)



(c) $d_1 = 1$, $d_2 = 2$ and $M = 3$ ($d_I = 0$ therefore generic intersection)

Figure 5.3: Illustration of Generic and Non-Generic Intersections of Simple Geometric Sets
(Source: Lai and Ye (2003))

The illustration in Figure 5.3 indicates that a one-to-one correspondence is only possible when a self-intersection does not occur i.e. $d_I < 0$. In mathematical terms, if $d_1 = d_2 = d$ and $m = M$, then $m > 2d$, which does imply that $m \geq 2d + 1$. By this illustration, Lai and Ye *(ibid.)* proved that it is important to select the appropriate embedding dimension otherwise it may be impossible to capture the invariant sets of the true dynamic system.

5.7 Lyapunov Exponent

In this research, the Lyapunov exponent (λ) has been identified as a suitable chaos parameter that can be used to characterise the traffic data in terms chaos, and subsequently developed as a suitable tool for the short-term forecasting of traffic states or conditions. The method deployed in this research is motivated by the finite-size Lyapunov exponent method proposed by Károlyi *et al*. (2010). However, the analysis focused on the rate of separation of the fiducial trajectory based on the separation measured from the centres of the spread of points in the phase space (see Section 4.5).

The Lyapunov exponent is one important product of the reconstructed phase space. This parameter is useful for determining the presence of chaos in a nonlinear dynamic traffic system. According to Rosenstein *et al*. (1993), the Lyapunov exponent provides a useful classification for chaotic systems through the measurement of the degree of sensitivity to initial conditions based on the rate of separation of infinitesimally close trajectories in a dynamic system. Over the observed interval, if it yields at least one positive λ-value, then the system can be classified to exhibit chaotic behaviour (Lai and Ye, 2003).

Figure 5.4 shows the orbits of two points X_0 (reference orbit) and $X_0 + \delta X_0$ (test orbit) and their initial separation in phase space is δX_0.

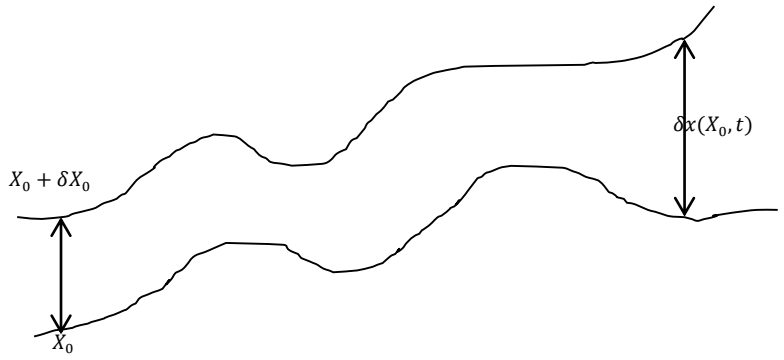


Figure 5.4: Representation of Two Orbits in Phase Space
(Source: Elert (2007))

Each point generates its independent orbits in the space using some time-dependent parameter equations or system of equations. Consequently, the separation between the two points is also time-dependent. Given that the system's sensitivity also affects the rate of propagation, the future state also depends on the initial condition. After a certain period t, the separation between the reference and test orbits will be $\delta x(X_0, t)$.

In mathematical terms, the rate of separation of two trajectories that diverge in phase space with initial separation δX_0 (if the divergence can fit within the linearized approximation) is given by:

$$|\delta X(t)| \approx e^{\lambda t}|\delta X_0| \qquad \textit{Equation 5-17}$$

In stable systems, the rate of separation diminishes asymptotically, and their orbits are characterised by attracting (fixed) points or periodic orbits. In unstable systems, the rate of separation diverges at an (initial) exponential rate behaving erratically before eventually settling down. By re-arranging Equation 5-17, the mean exponential rate of divergence of two initial close points is:

$$\lambda = \lim_{\substack{t\to\infty \\ |\Delta X_0|\to 0}} \frac{1}{t} ln \left|\frac{dx(X_0, t)}{dX_0}\right| \qquad \textit{Equation 5-18}$$

The work of Károlyi *et al*. (2010) is supported by Equation 5-18 given by Elert (2007). This equation is deployed in this research to primarily understand the evolution of the traffic flow in the system. The uniquenss of this method is the ability to generate the time series profile of the instantaneous Lyapunov exponent. This enables the propagation of the system over time to be understood. Rosenstein *et al*. (1993) criticizes the approach of Wolf *et al*. (1985) for failing to take advantage of all the available data because it focuses on one fiducial trajectory. However, Rosenstein *et al*. (1993) took advantage of the work of Sato *et al*. (1987) to provde an improved method for estimating the largest Lyapunov exponent. Existing literature reveals that Shang *et al*. (2005) applied the method proposed by Rosenstein *et al*. (1993) to calculate the largest Lyapunov exponent of traffic time series data comprising of speed, flow and occupancy. Rosenstein's approach is emerging as particularly attractive because it requires less data and can simultaneously yield the correlation dimension. In addition, it is useful in determining the longest forecast length,

which is the reciprocal of the largest Lyapunov exponent. Notwithstanding this evidence, the algorithm is not considered ideal for this research, because it does not provide the instantaneous exponents that this work seeks to investigate in detail for the potential application of Chaos Theory to forecast urban traffic conditions.

In order to apply Equation 5-6 to this research, the formula was coded in C# program (see Section 6.6.5). Existing literature indicates that there are three kinds of orbits, which can be distinguished depending on whether the magnitude of λ is negative, zero or positive. If $\lambda < 0$, the orbits exhibits asymptotic stability, where greater negativity indicates greater stability of the system. Hence, $\lambda = -\infty$ is indicative of super-stable systems that reach equilibrium as soon as possible. The nature of dissipative (or non-conservative) behaviour in dynamic systems is identified by a negative λ, which indicates that the orbits are attracting (fixed) points of stable periodic orbits. Other systems, which are characterised by $\lambda = 0$, exhibit Lyapunov stability, which is a steady state, where the system eventually settles to a neutral fixed point. Chaotic behaviour is present if $\lambda > 0$, where the orbits diverge exponentially with time irrespective of the initial close proximity of the points within the phase space (Elert, 2007).

Traffic behaviour can relate to these states in many respects. During free-flowing conditions, it is possible to maintain the separation between two vehicles for as long as possible. That is, the initial and final separations remain constant throughout the observation period. Therefore, traffic networks exhibit stable conditions because it is free flowing, making free flow conditions likely of more negativity and nudge towards super-stable conditions during the extremely quieter overnight periods. Stable conditions may also be present during congestion. Erratic behaviour is more likely in the lead up to congestion when there is a high interaction between vehicles, which is likely to be characterised by frequent changes in the separation between the vehicles. This condition is termed chaos in dynamic systems behaviour and identified by a positive Lyapunov exponent.

5.8 Conclusions

This chapter has set out the mathematical framework for Chaos Theory based on the Phase Space Reconstruction method, which is suitable for analysing non-linear time-dependent chaotic systems. It discussed the procedure for estimating parameter values (such as Lag

(delay time), embedding dimension and correlation dimensions, and subsequently the use of these parameters as inputs for the reconstruction of the phase space. Section 7.2.1 illustrates the computation of the delay time from the autocorrelation method, whilst Section 7.2.2 deals with the embedding and correlation dimensions. Finally, this chapter explored the use of the Lyapunov exponent to determine the dynamic state of a non-linear system. The discussion in this chapter underpins the research presented in the remaining chapters of this thesis. The next chapter discusses the methodology for this research work in entirety. It sets out how the contents of this chapter fits within the overall process, the development of computer programming codes to accelerate the delivery of this work, and the types of analyses undertaken to achieve the research aims and objectives.

Chapter 6. Methodology for the Research Study

6.1 Introduction

In order to answer the research questions, it is important to identify the appropriate research techniques as well as the systematic procedures that are required to implement these techniques. It is important to understand for example, why the choice of a particular dataset will be more suitable among the options available, and why the particular analysis techniques for the data were chosen, and in addition, what are the appropriate models to use and how to interpret effectively their outputs.

This chapter therefore follows on from the introduction to Chaos Theory in the preceding chapters, and describes how the literature review has enabled the formulation of this research methodology. The data used in the research through to the systematic development of the programming codes are described, and the types of analysis used to explore how Chaos Theory might facilitate the management of traffic congestion in signalised urban networks are elaborated upon.

Section 6.2 provides a brief description of the study network in the context of the local area. Section 6.3 describes the types of traffic data and their sources for this work. It summarises the traffic data in the SCOOT database and discusses why the network analyses were based on the 1999 data. Further, it describes the process of repair, interpolation and synchronisation to account for missing data. Section 6.4 summarises the type of descriptive statistics used to explore the data. Section 6.5 describes the modelling steps for deploying Chaos Theory focussing on the process of calculating the lags, embedding dimension, correlation dimension, and the Lyapunov exponents, and how the preliminary model was tested using noise and subsequently flow and occupancy data (from a single link) and subsequently the extension to a network of interconnected links.

In order to make the modelling process easy to understand, Section 6.6 describes the chaos algorithms and presents a pseudo code of the Microsoft C# programs, which are more fully covered in Appendix A. Section 6.7 describes the analysis of the Lyapunov exponents, based on the curve fitting using the LOWESS approach (see Section 6.7.1) and subsequently the derivation of relationships using cross-correlation techniques. Section 6.8 describes the spatial and temporal analysis of the network by using the Lyapunov exponents

(derived from flow data) to develop a multiple regression model for predicting the occupancy of the network links.

6.2 Study Area

The study area is located in Leicester, a city and unitary authority area in the East Midlands of England, and the county town of Leicestershire. The city lies on the River Soar and at the edge of the National Forest. Currently the largest unitary authority in the East midlands, the population of the Leicester Unitary Authority was 330,000 according to the 2011 census. The city is located at the intersection of the north/south Midland Main Line and east/west Birmingham/Leicester/Cambridge CrossCountry railway lines, to the south of the A46 and east of the M1/M69 motorways intersection, and west of the confluence of the A6 and A47 trunk roads.

The map in Figure 6.1 shows the location of the study area, which is the Welford Road SCOOT Region (Region KA), in relation to the local road network in Leicester City, which is to the south of University of Leicester. This network is approximately two kilometres in length and includes 10 signalised junctions incorporating 40 network links.

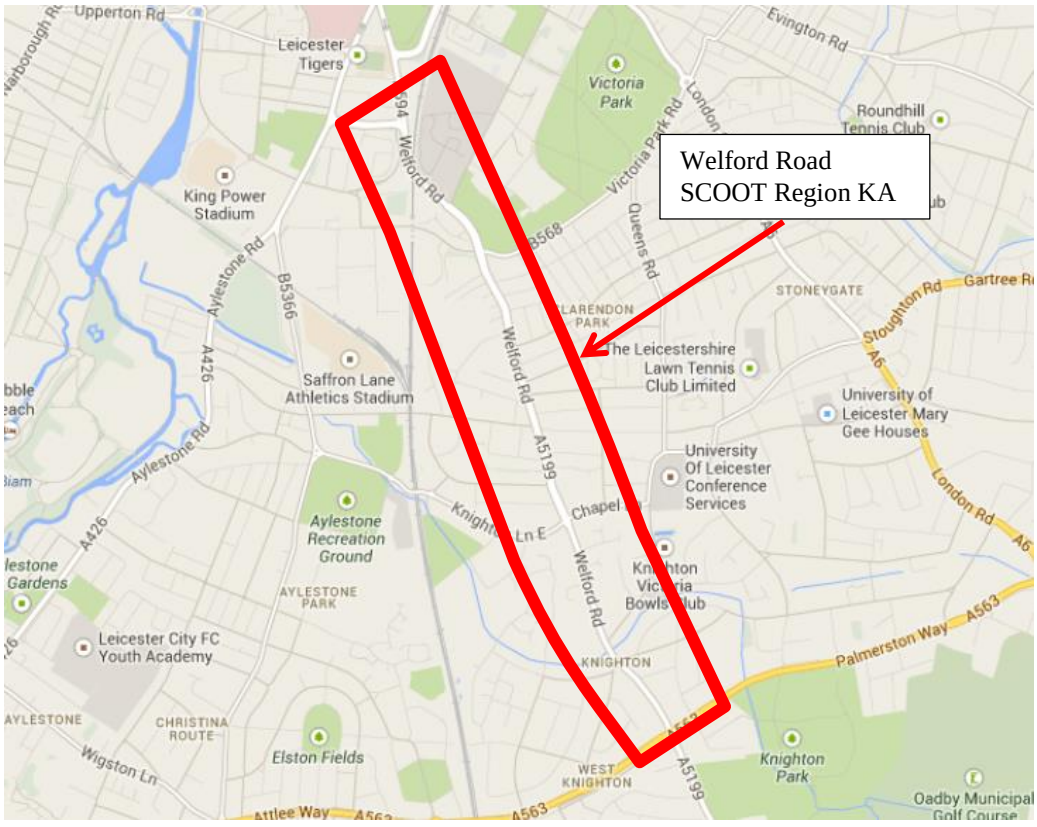


Figure 6.1: Network Layout – Welford Road SCOOT Region, Leicester

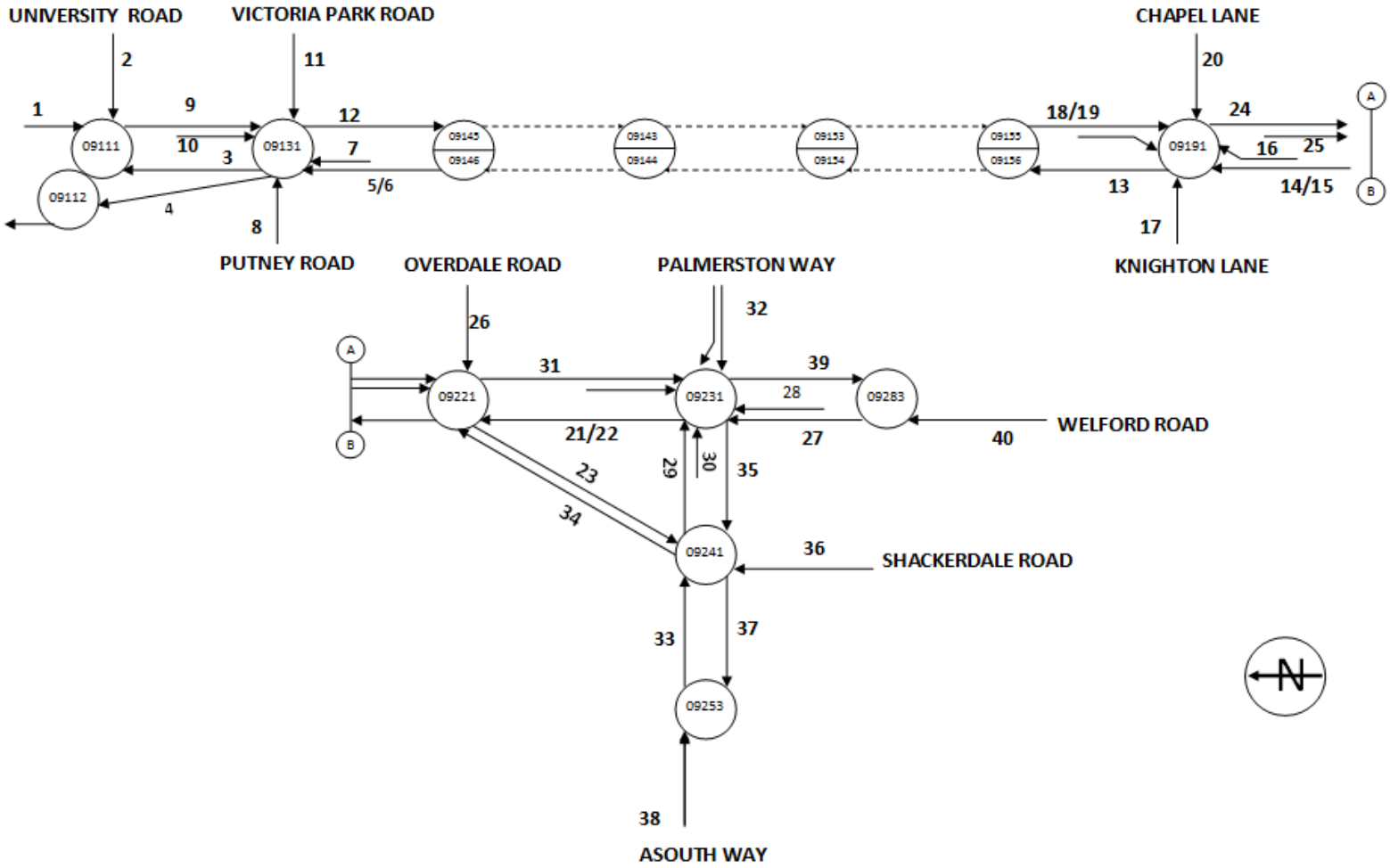


Figure 6.2: Schematic Diagram of the Welford Road SCOOT Region ("Region KA"), Leicester[1]

[1] Based on Schematic Drawing for SCOOT Region KA – Welford Road (dated 8th April, 2004) with permission of Leicester City Council

Figure 6.2 shows the schematic layout of the SCOOT Region KA. The traffic data available from the SCOOT Region KA network is analysed in detail in Chapter 8 and Chapter 9. Table 6-1 provides details of the actual SCOOT link identifications, which were re-labelled as 1, 2, 3…40 (link IDs) for ease of reference in this research.

Link IDs	SCOOT IDs	Link IDs	SCOOT IDs	Link IDs	SCOOT IDs	Link IDs	SCOOT IDs
1	09111A1	11	09131N1	21	09221R1	31	09231N1
2	09111B1	12	09145B1	22	09221R2	32	09231P1
3	09111C1	13	09156I1	23	09221S1	33	09241A1
4	09112R1	14	09191A1	24	09221T1	34	09241B1
5	09131I1	15	09191A2	25	09221U1	35	09241C1
6	09131I2	16	09191B1	26	09221V1	36	09241D1
7	09131J1	17	09191C1	27	09231I1	37	09253V1
8	09131K1	18	09191D1	28	09231J1	38	09253W1
9	09131L1	19	09191D2	29	09231K1	39	09283A1
10	09131M1	20	09191E1	30	09231L1	40	09283B1

Table 6-1: Look-up Table for Relabelled Link Numbers

6.3 Data Types & Acquisition

Traffic analysis requires measured observations of key variables of vehicular traffic over both space and time, and depending on the data type, may provide an indication of the level of traffic in the traffic stream (May, 2005, Wardrop, 1952). The variables that describe traffic in a stream include flow, density, concentration, spacing (or space headway), speed and travel time were described in detail in Section 2.4. In this research, the data sources were limited mainly to flow and occupancy from SCOOT inductive loop detectors, but noise data from motes were used for the preliminary testing of the models. Flow and occupancy data were available from SCOOT inductive loop detectors.

Flow (denoted as q (veh/s)) is the number of vehicles(n_t) passing a point in one lane for a period (t). In Mathematical terms, flow is given by $q = \frac{n_t}{t}$. Occupancy is the proportion of time that a loop is occupied by a vehicle (or detects the presence of a vehicle) (Hazelton, 2004). This is computed as the sum of the time intervals that a detector is covered by a vehicle divided by the observational period (T). In SCOOT, a unit time interval is 0.25 seconds given that the sampling rate is SCOOT detectors is 4Hz.

6.3.1 Data Sources

Data 1 is a minute-by-minute dataset from 20 mote sensor locations in Medway (in Kent) for April, May, June and July 2011. Data 2, 3, 4 and 5 comprise 1, 10, 20, 30, 60 and 120 seconds resolution flow and occupancy data from the Leicester SCOOT system. The preliminary research used Data 1, 2 and 3 only. Data 2 and 3 is data for a single loop detector used for the preliminary testing of the procedures of Chaos Theory (together with Data 1) before applying Data 4 and 5 to the network of spatially and temporally connected loops in the SCOOT network in Leicester (see Figure 6.2).

It is important to note that Leicester is one of UK instrumented cities, where the Institute of Transport Studies (ITS) at Leeds University has been accumulating and archiving traffic and environment data in support of its research activities. The Leicester data sets cover the period from 1987 for SCOOT messages and since 1997 for pollutant levels and meteorological conditions until 2011 (Bennett *et al*., n.d.).

The Leicester SCOOT database provided flow and occupancy data compiled for the study area for 1998-2011. For the SCOOT system data, it is possible to identify missing data sections. The missing data was typically attributable to incidences such as faulty loops, maintenance issues or system failures. The most conspicuous being in 2010, where all data were unavailable. However, data were available for all months in 1998, 2001, 2002, 2003 and 2004; 11 months of data was available for 1999, whilst for the remaining years the number of month available varied but there was at least one full month of data.

The Leicester SCOOT database provided flow and occupancy data compiled for the study area for 1998-2011. For the SCOOT system data, it is possible to identify missing data sections. The missing data was typically attributable to incidences such as faulty loops, maintenance issues or system failures. The most conspicuous being in 2010, where all data were unavailable. However, data were available for all months in 1998, 2001, 2002, 2003 and 2004; 11 months of data was available for 1999, whilst for the remaining years the number of month available varied but there was at least one full month of data.

In order to select the appropriate data (measured by the amount of missing data in the observations) from the database, an assessment of the monthly capture rate of the recorded data was carried out to determine the data quality. For the purpose of this exercise, the capture rate of the on-road SCOOT detectors was defined as the ratio of the number of

recorded to the number of expected data points. The summary of this assessment is presented in Table 6-2 and indicates that the capture rate ranged between 1% (June 2007) and 93% (September 2011). The normality test at 5% significance level of the ratios gave a significant p-value of 0.03 indicating that the distribution is not normal. The statistics are median 59%, whilst the first and third quartiles are 40% and 73% respectively.

	Number of Recorded Data Points (%)												
	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2011
Jan	51	81	59	40	41	58	47	41	28	-	78	43	-
Feb	29	89	-	78	49	66	84	-	28	-	-	47	-
Mar	69	81	-	60	65	52	48	26	62	38	-	37	-
Apr	71	79	-	68	59	37	61	58	-	32	-	21	-
May	57	77	67	84	74	39	34	-	21	28	-	12	-
June	65	81	48	81	86	46	66	-	65	23	28	32	3
July	62	81	54	69	75	50	78	-	45	1	32	6	73
Aug	68	46	45	74	78	73	42	71	35	17	40	-	-
Sept	69	73	89	63	83	93	65	40	43	-	51	18	93
Oct	76	35	85	49	70	74	74	30	44	61	47	-	-
Nov	75	50	63	55	79	43	27	64	17	62	62	32	-
Dec	76	-	48	40	48	40	36	55	-	59	59	-	-

Table 6-2: Monthly Breakdown of the Capture Rate of SCOOT Flow and Occupancy Data (Data 2, 3, 4 and 5) (1998 – 2011)

In this work, the selected baseline for quality of the data is the upper 75th percentile; therefore, the sampled months used in this work are of at least 75% capture rate. This figure was based on assessing the effect of missing data on the total daily flow using SCOOT data for link 1, sampled at 20-seconds intervals for the entire day on 1st January 1999. In this exercise, a random number generator function in Excel software was used to remove missing data from the complete data in increasing steps of 1.25%. For each missing datum, the ratio of the standard deviation to the mean was computed (see Figure 6.3). The curve shows a gradual change in ratio from the lower end, which rises sharply from when the missing data are about 25%, and peaks at about 60% and falls away afterwards. The results suggest a consistent change in ratio when the proportion of missing data in the sample is

less than 25%. Therefore, applying data of a capture rate of at least 75% was satisfactory for this research work.

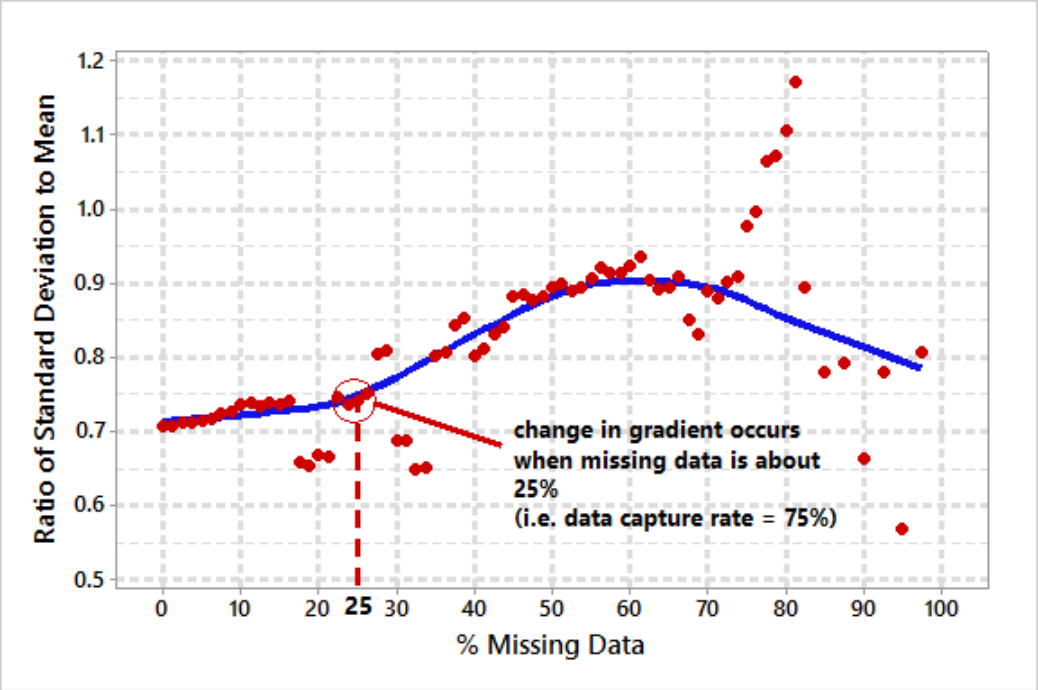


Figure 6.3: Graph of the Ratio of Standard Deviation and Mean against Missing Data

The assessment of the capture rate helped to establish a reasonable (statistical) basis for identifying suitable datasets for the research work by focussing mainly on datasets of high capture rate and therefore (relatively) minimal amounts of missing information. The year 1999 was found to have the highest number of qualifying months (against the prescribed criteria of more than 75%) with continuous data from January to July. Therefore, the relevant months in the 1999 datasets were analysed in this research. A summary description of the data sources studied in this research is given in Table 6-3.

Data ID	Data Type	Year	Data Source	Stage in Research Data was Applied
Data 1	Noise	2011	Motes in Medway	Preliminary
Data 2	Flow	1999	Leicester SCOOT System	Preliminary
Data 3	Occupancy	1999	Leicester SCOOT System	Main
Data 4	Flow	1999	Leicester SCOOT System	Main
Data 5	Occupancy	1999	Leicester SCOOT System	Main

Table 6-3: Summary Description of the Data Type and Sources

6.3.2 Data Repair/Interpolation

In order to enhance the data quality for the analysis, it was necessary to fix the missing information. However, this must not compromise the integrity of the data structure itself (which would consequently be of no use for chaos analysis), which is very likely if a high degree of repair was required. As demonstrated above, the choice of the 75th percentile was appropriate. Each data point recorded by the SCOOT system will also include a "period" variable that will state the time between subsequent measurements (which determines the current measurement frequency as the inverse of the period) of that particular message. By analysing the time stamp of subsequent messages it was possible to identify missing data points within the overall data set.

There are two options to address the problem of missing data points:

1. *Ignore the missing data*: This is normally only possible when any subsequent analysis does not rely on a complete data set with which to work. Typically, this would be analysis requiring aggregate statistics. For example, the average occupancy over a one-hour period will not need a complete data set and can provide a result with acceptable level of accuracy even with very patchy data, assuming no systematic bias is embedded in the missing data.
2. *Interpolation*: This is the most frequently used process when the subsequent analysis is reliant on complete data and therefore estimates of the missing data will be required. This research required complete datasets making the data interpolation option essential.

For the initial run of this work, interpolation of the data was undertaken using a simple linear interpolation between the two measured data points on either side of the missing data. The justification for this simple interpolation is two-fold:

1. The missing data is typically random and infrequent; therefore, a simple linear interpolation is expected not to drastically affect the results.
2. The amount of data that will need to be checked and then subsequently rewritten in a repaired format is extensive. Therefore, a more sophisticated interpolation (for example, LOWESS (see Section 6.7.1)) will increase the time for the data repair quite substantially.

6.3.3 Synchronisation

Although SCOOT data is collected at a known frequency (as logged by the "Period" variable in the messages reported by the SCOOT system), this frequency is not necessarily homogeneous across the entire SCOOT network, or indeed within the same SCOOT link at different time periods. Typically, the SCOOT data is logged either at a frequency that is determined by the current state of the traffic system typically cycle-by-cycle data consolidation or at a fixed pre-set time usually five minutes. This is not a problem as most SCOOT analysis takes place at the aggregate level (five minutes, for example) so the non-synchronised data (both in terms of phase and absolute time) does not affect the analysis. However, for the chaotic analytical techniques it is necessary for the data to be both synchronised in phase and absolute time. The sampling interval for the detectors must be homogenous across the network.

Therefore, the data stream has also been manipulated so that each SCOOT cycle is completely synchronised. In order to accomplish this the data have been processed in a similar fashion to that used to repair missing data (see Section 6.3.2). A measurement time period was chosen such that the interpolation measurement cycle was a divisor of both the phase and the absolute time of the real world data. The algorithm then interpolates between each measured pair, creating and inserting new data points and time periods so that there exists a continuous periodic data set for all of the SCOOT data that is being currently analysed. The time period chosen in general is typically the longest suitable time-period, which may be different for each dataset investigated.

Again, as with repairing the data, the algorithm used is quite simplistic but has the advantage that large quantities of data may be processed between different time periods and different SCOOT links can be easily compared.

6.4 Analysing the Raw Data

The descriptive statistics (i.e. mean, median and standard deviation) were derived and found useful to gain understanding of the datasets and to compare parameters across the independent links. This initial assessment of the traffic data included generating visualisations of the data such as boxplots, time series plots and histograms that helped to visually identify possible extreme values in the datasets. Furthermore, this initial analysis

assisted in identifying the normality of the datasets and thus to determine whether a parametric or non-parametric test was suitable for the statistical scrutiny of the datasets.

The computed delay (lag) times of the datasets were used to investigate whether or not relationships exist between other metrics and whether the lag times of the respective datasets also revealed possible trends (see Section 8.3.2. The importance of this step was to determine whether it was possible to evaluate the lag times directly from the calculated metrics of the datasets or whether there was evidence of autocorrelation in which case alternative approaches were needed as described in Section 5.5.2. The next section sets out the procedure in using Chaos Theory to analyse road traffic data.

6.5 Modelling the Process

The flow chart of the methodological process included in Figure 6.4 gives an overview of the relationship among the key tasks undertaken to achieve the research aims and objectives. Figure 6.4 also demonstrates how the preliminary investigations carried out feeds into the main research.

6.5.1 Preliminary Research

The first step was to translate the mathematical equation supporting the key stages into algorithms for coding in Microsoft C# program language. All the program codes are included in Appendix A (Program 1, Program 2 and Program 3 are codes for the autocorrelation lag, embedding and correlation dimension, and Lyapunov exponent respectively). The purpose of the programs are to provide a quick and efficient way to undertake the iterations of the mathematical formulae of Chaos Theory.

Once completed, the model was checked, calibrated and validated at the preliminary research stage using noise data from a network of motes in Medway, Kent (Data 1) and a secondary data source of SCOOT flow and occupancy from a single link (Data 2 and 3 respectively). This process was important for establishing a reasonable range of values to derive variables such as the optimum lag time, embedding dimension, maximum separations in phase space, etc.

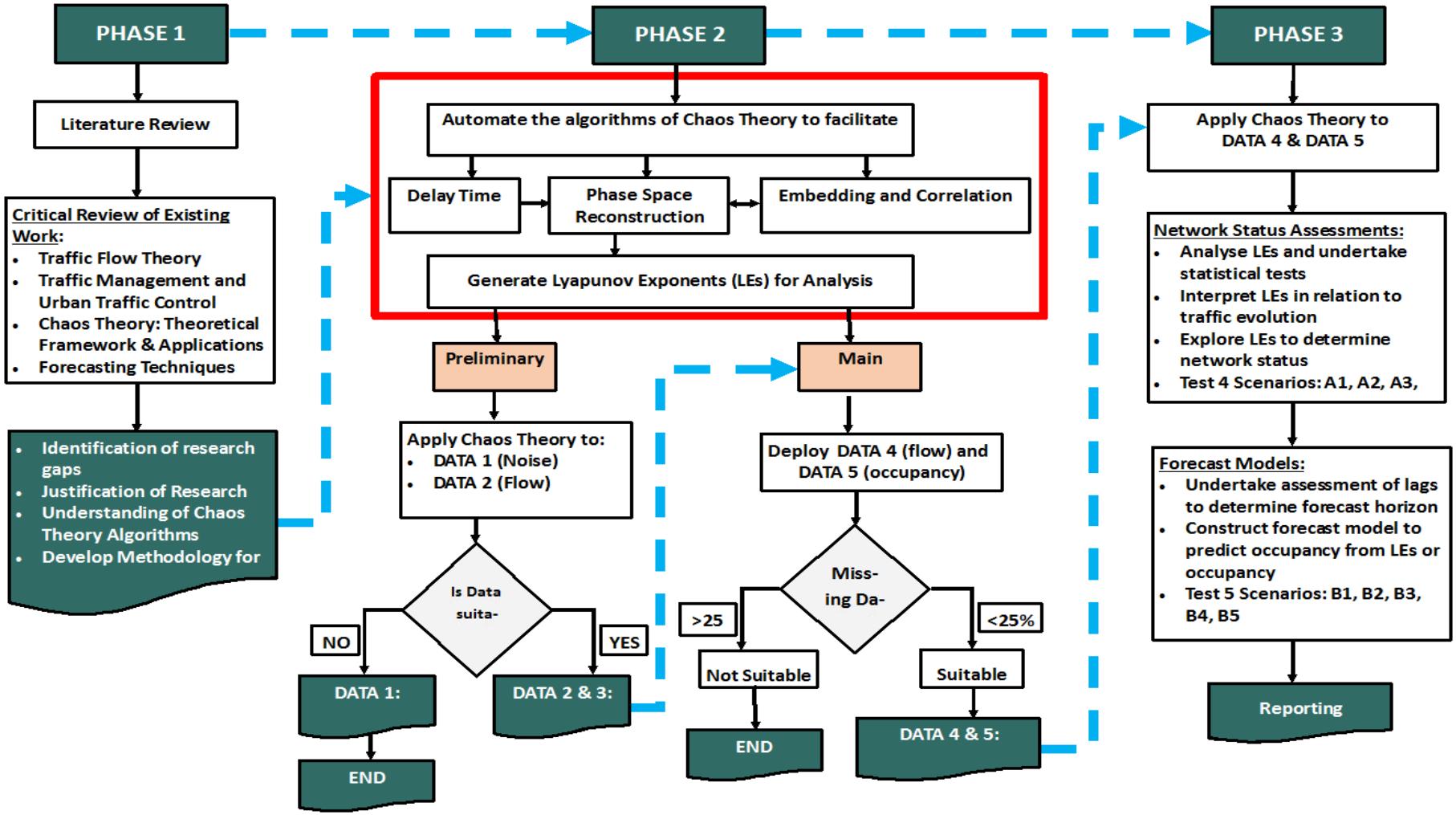


Figure 6.4: Flow Chart of the Methodological Process

In order to achieve this, the initial runs of the model tested several ranges for these parameters in the form of a jagged array. The process identified that traffic noise, for example, required an embedding dimension range of 5-10 (see Section 7.2.2), whilst SCOOT variables were of a lower dimension of 2-3 (see Section 7.3.2).

Moreover, a range of lags of durations between 0-1000 steps (appropriately 0-333 minutes) was reasonable for the model runs when computations of autocorrelation coefficients are required. This suggests that the optimal lag for a time series of SCOOT data sampled at 20 seconds interval was more likely to fall within this range. This was the conclusion from an initial assessment of 10 lag ranges between 0-5000 (see Program 1 in Appendix A).

The profiles of the Lyapunov exponents extracted from the outputs of the models were useful to identify the dynamic states of the signalised urban road network. The results were able to distinguish between free-flow, steady and chaotic traffic conditions in the urban network. The model tested various sampling frequency of 1, 10, 20, 30, 60 and 120 seconds of flow and occupancy SCOOT data (Data 2 and 3 respectively). The profile of the Lyapunov exponent showed no visible structure at the lower and upper ends of the range, whilst a 20-second sampling frequency showed an underlying structure and proved a suitable range for assessing chaos (see Section 7.3.3). These conclusions shaped the next step of expanding the modelling work to cover the SCOOT network of interconnected signalised links within the study area in the main research.

6.5.2 Main Research

In the main research, the computer programs were used to investigate an entire network of SCOOT links in order to identify whether chaos is an isolated phenomenon or an occurrence across a network of flows on links spatially and temporally connected in the urban environment. The area studied was described in Section 6.2. The construction of the network model was fundamental to the understanding of chaos and the extent of chaotic behaviour in the SCOOT network. Data 4 and 5 were used at this stage, which included seven monthly cohorts of flow and occupancy data chosen based on the 75^{th} percentile capture rate that were analysed. The cohorts are January, February, March, April, May, June and July 1999. The results of the network models are discussed in Chapter 8 and Chapter 9.

6.6 Coding of the Modelling Stages

6.6.1 General

The purpose of the program is to provide a quick and efficient way to undertake the iterations of the mathematical formulae of Chaos Theory. All the program codes are included in Appendix A (Program 1, Program 2 and Program 3 are codes for the autocorrelation lag, embedding and correlation dimension, and Lyapunov exponent respectively) but some extracts are included in the text here for ease of reference. Each of the programs are capable of dealing with an arbitrary number of files, undertaking specific calculations and storing the outputs to a designated location for further (or detailed) analytical processing.

Once the program starts to run, a *folder browser* prompts the user to select the data folder. The program selects in a chronological order according to the filenames all the files in this location. Then, all each file is written as a text file for subsequent calculations. Next, the user must specify the start and end date as well as the time intervals of the time series of flow and occupancy data. The program reads and splits all the lines of data within a prescribed interval with each line occupying a dedicated row with all data in the cell separated by a comma. This process continues until the selection of all the data within the range is completed or the next line is empty. In the case of the SCOOT data, for example, the order of the data is: Link ID, $date, time, refence\ number,\ occupancy\ and\ flow$ respectively. The flow chart in Figure 6.5 illustrates this process.

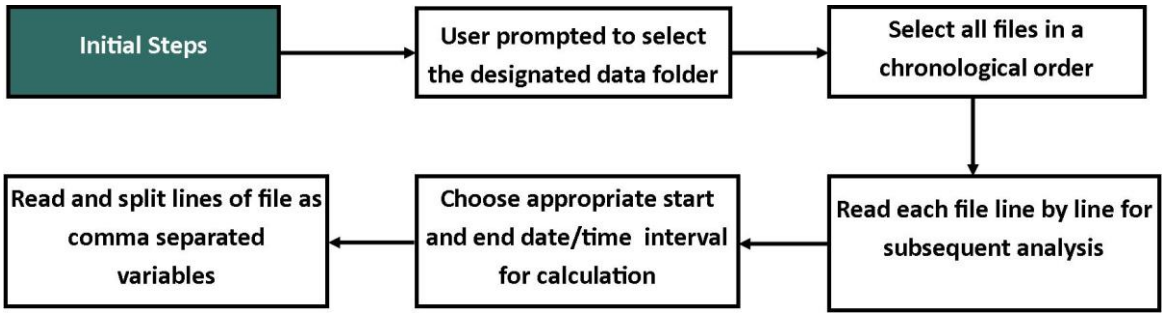


Figure 6.5: Initial Steps Algorithm for Opening and Reading Files in a Designated Folder

6.6.2 Time Delay or Autocorrelation Lag (ACL)

Given that the optimum lag is unknown, the results of an initial assessment using a number of arbitrary ranges of delay time is undertaken in order to identify a suitable range for further analysis. The initial run of the tested a range of 10 sets of delay times to determine

the optimum estimate of the delay time corresponding to an autocorrelation of 0.4, but the analyses concluded that the acceptable data range of 0-1000 for flow and occupancy (and 0-500 for noise) data for this research (as indicated previously in section 6.5.1). The delay times were stored in a jagged array. The algorithm is illustrated in Figure 6.6. On completion of the model runs, the outputs were analysed based on *correlograms*, which are graphical plots of autocorrelation versus the lagged time. According to the available literature, the optimum delay time corresponds to the solution for the autocorrelation at 0.4.

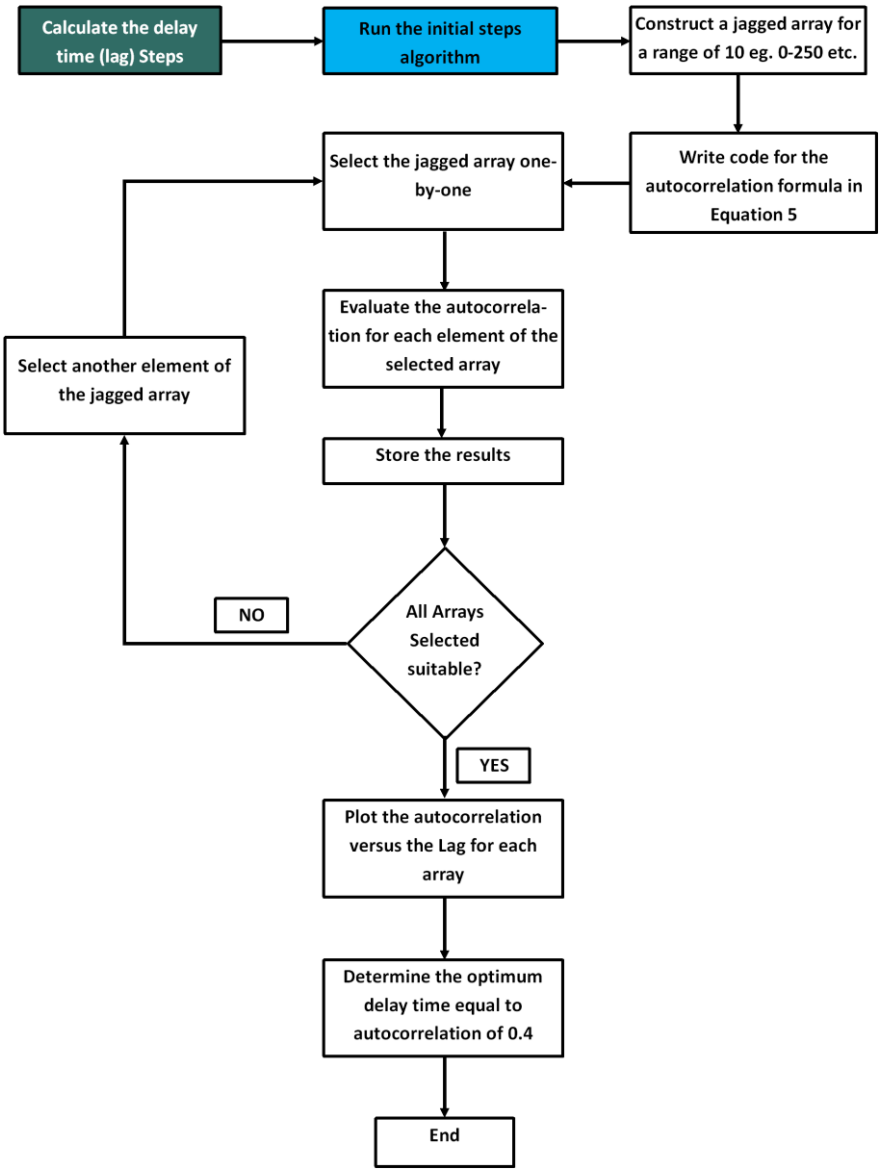


Figure 6.6: Generic Algorithm for Calculating the Delay Time

6.6.3 Phase Space Reconstruction

By using the optimum lagged time, the program code for Stage 2 translates a one-dimensional time series into any arbitrary n-dimensional phase space. Given that the actual dimension is unknown, there is a systematic trial of the dimension number to evaluate the correlation and embedding dimensions.

Figure 6.7 illustrates the generic code for constructing the phase space, which shows that an observation $X(1)$ translated into a 5-dimensional phase space will have its 1st, 2nd, 3rd, 4th and 5th components as $x(1)$, $x(1+\tau)$, $x(1+2\tau)$, $x(1+3\tau)$ and $x(1+4\tau)$ respectively, where τ is the length of the lag (calculated in stage 2). The components are variables i, j, k, l and m respectively.

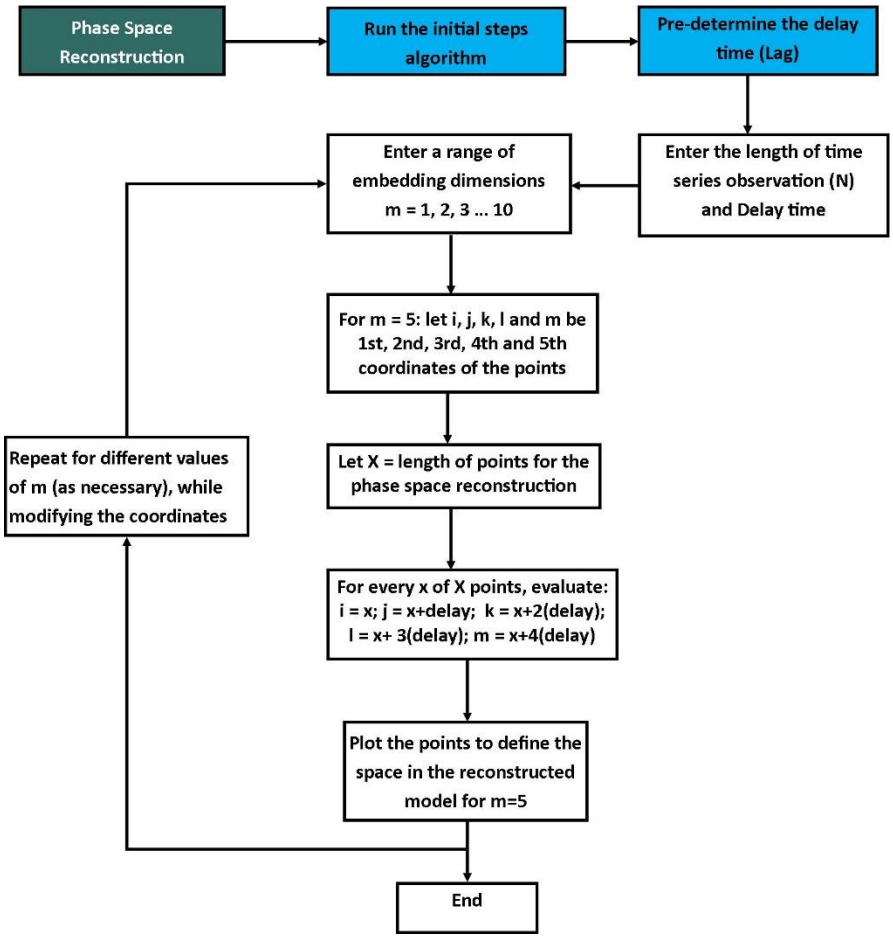


Figure 6.7: Generic Algorithm for the Phase Space Reconstruction Process

The observation recorded at the positions corresponding to i, j, k, l and m in the data series are then selected for the construction of the phase space. Given that all of the individual components making up the coordinates will correspond to a specific numeric observation

in the dataset, the transformation identifies each observation as an exact point in the phase space. The next step was to use the reconstructed phase space to determine the embedding and correlation dimensions and subsequently the instantaneous Lyapunov exponents.

6.6.4 Embedding Dimension

Figure 6.8 is flow chart of the generic algorithm used for calculating the embedding and correlation dimensions.

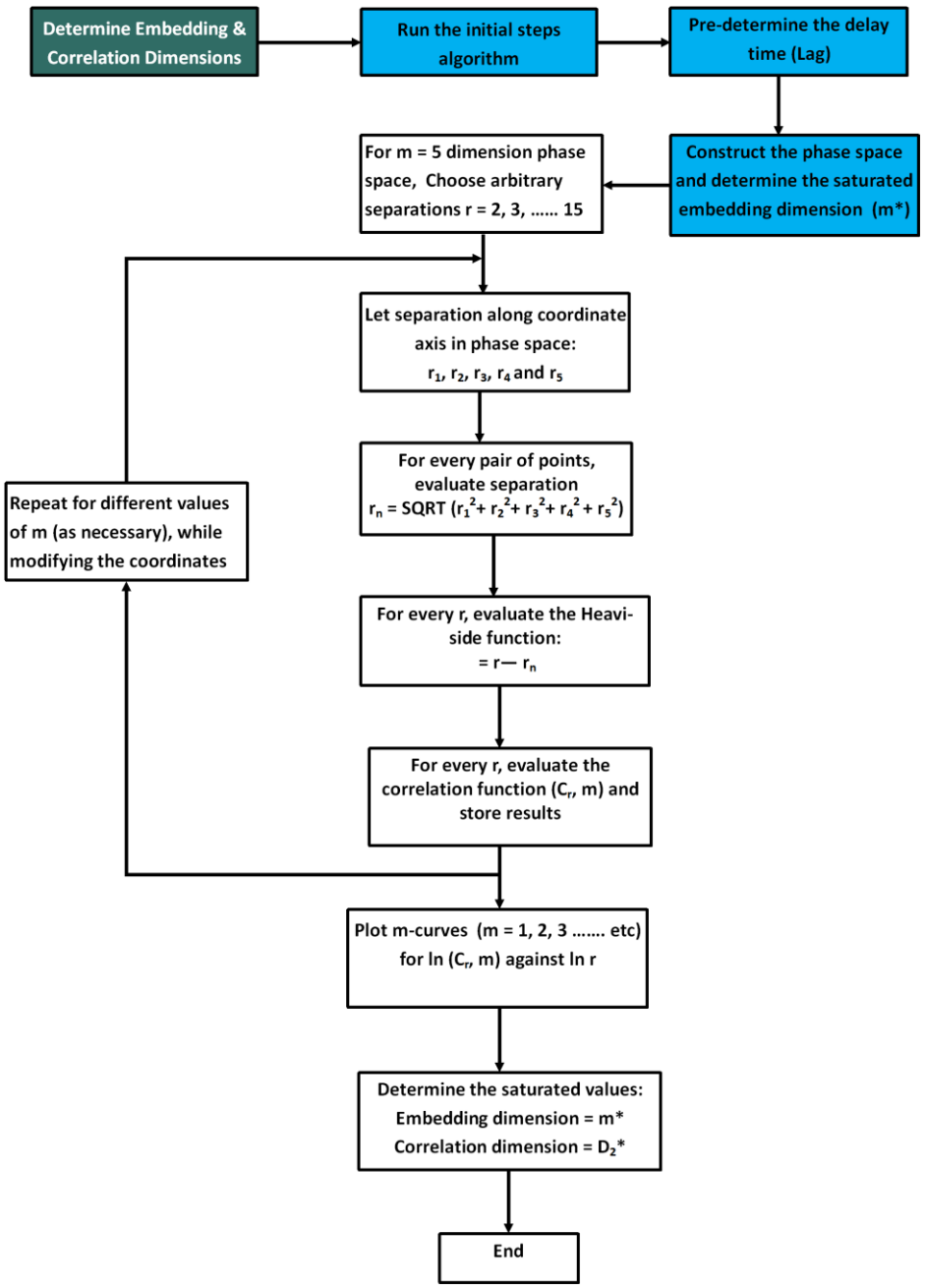


Figure 6.8: Generic Algorithm for Determining the Embedding and Correlation Dimensions

Equation 5-10 provided the algorithm to code the program for the calculation of the embedding and correlation dimensions using pre-selected values of the embedding dimension (m) and set distances (r). The program investigated a combination of m and r values of between 1 and 10, and 2 and 15 respectively. In the ensuing code for an assumed 5-dimensional phase space for example, the distance between two points along each axis is r_1, r_2, r_3, r_4 and r_5 respectively. The variable r_n represents the actual separation, which is input into evaluating the Heaviside function (Equation 5-11).

The outcome of the correlation function $C_2(r, m)$ for the all m-values is stored, and finally the saturated embedding and correlation dimensions calculated from the m-curves given by the plot of ln $C_2(r, m)$ against ln r using the Grasberger-Procaccia (G-P) algorithm method (see section 5.6.2). This will be used in the analysis and will be revisited in Sections 7.2.2 and 7.3.2.

6.6.5 Lyapunov Exponent

The program code for the instantaneous Lyapunov Exponent is consistent with Equation 5-18. In the generic code illustrated in Figure 6.9, the separation between a point and neighbouring points in phase space is calculated for every point in the phase space.

Next, the average separation with points less than a certain maximum distance, and the average time interval, is calculated and stored. An initial separation between points is assumed for the phase space, and the logarithmic term of Equation 5-18 calculated as the ratio of the mean separation and the initial separation, which is divided by the average time to give the instantaneous Lyapunov exponent. This will be used in the analysis and will be revisited in sections 7.2.3 and 7.3.3.

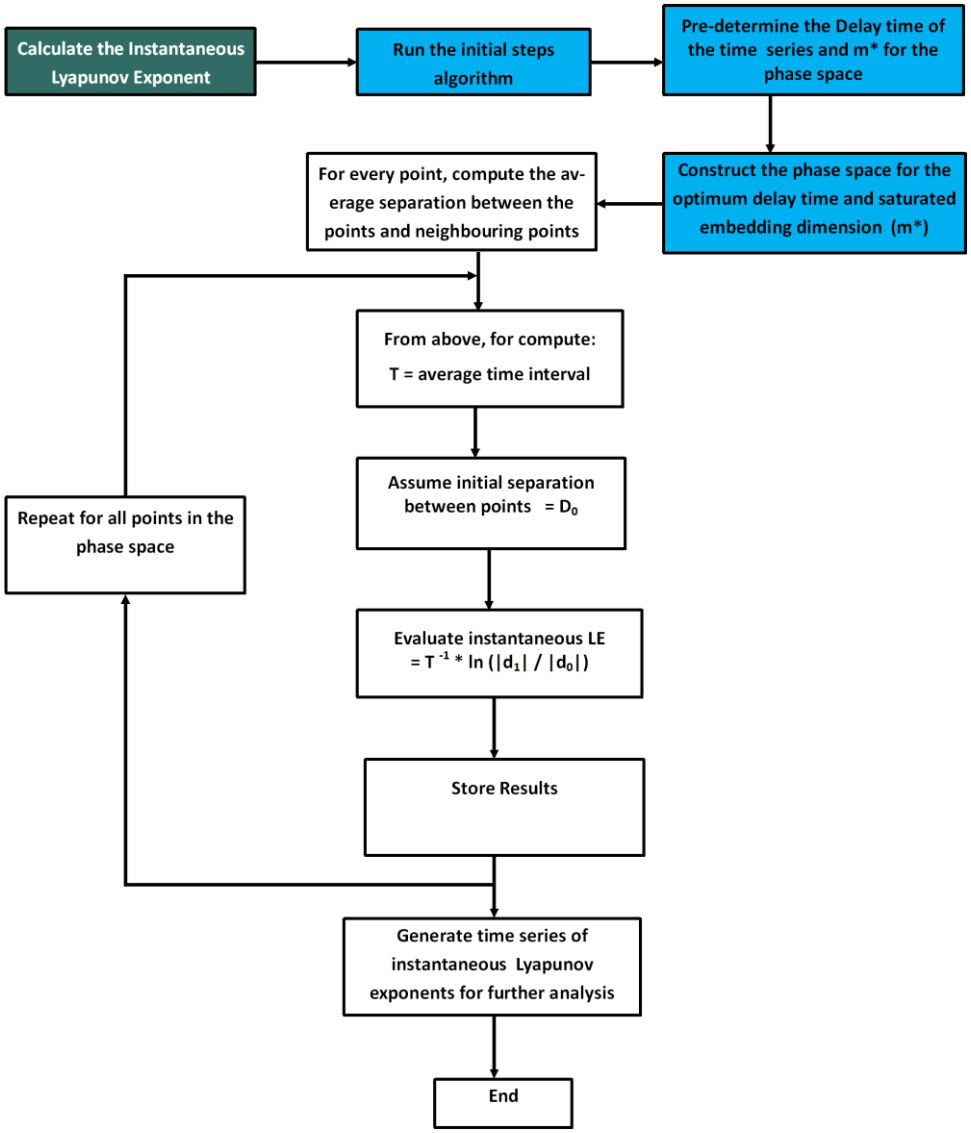


Figure 6.9: Generic Algorithm for Calculating the Instantaneous Lyapunov Exponent

6.7 Network Status Analysis

The instantaneous Lyapunov exponents output for all links in the network for each 20-second interval for a 24-hour duration from midnight were calculated resulting in a matrix of dimension 4320 x 32 for each day.

The chaotic state of the network was identified, with reference to the sign of the instantaneous Lyapunov exponents, by identifying the number of links indicating a 'positive', 'zero' and 'negative' Lyapunov exponent at each sampling interval. The profiles of the positive Lyapunov exponents were subsequently analysed with the assumption that they were the main indicator of chaos (see Section 8.6). For each cohort, the daily positive Lyapunov exponents were summarised at the weekly and same day of the week ("same

weekday") levels across the entire month. The weekly analysis included profiles in the same week from Monday to Sunday, whilst the "same weekday" comprised, for example, all-Mondays, all-Tuesdays etc. A fitted curve over the summarised data helped to understand the propagation of chaos in the network and identified the critical number of links exhibiting chaos characteristics that signal the onset of congestion (see Section 8.6.3). Finally, the data were also analysed to assess the degree of correlation among the pairs of the daily positive Lyapunov exponents.

6.7.1 LOWESS (LOc WEighted Scatterplot Smoothing)

A fitted curve using the LOWESS mathematical procedure over the occurrence of the positive Lyapunov exponents helped to determine the distributions of the positive Lyapunov datasets whilst minimising the influence of the noisy components of the system. LOWESS is a non-parametric regression where, unlike a 'standard' linear regression analysis, there is no assumption that the form of the relationship is a straight line. It depicts the 'local' relationships between two variables over parts of their range, which may generally tend to differ from the 'global' relationship determined from the entire datasets. The Minitab LOWESS fitted curves over the data points in the sample explores the relationship between two variables without the need for a particular function and the only specification required being the degree of the polynomial and smoothing parameter value. In this research, the LOWESS regression procedure was used to explore differences and similarities for datasets over a day, from day to day and weeks.

6.7.2 Correlation Coefficient

Equation 6-1 is the mathematical formula for calculating the correlation coefficient between any two variables (x and y). The coefficients provides a measure of the degree and type of relationship between two variables, and the squared value gives a ratio of the explained to the total variation in the data. The correlation coefficient (r) of +1, 0 and -1 respectively means there is 'perfect correlation', 'no correlation' and 'perfect negative correlation'.

$$r = \frac{\sum_i (x_i - \bar{x})\,(y_i - \bar{y})}{\left[\left[\sqrt{\sum_i (x_i - \bar{x})^2}\right]\left[\sqrt{\sum_i (y_i - \bar{y})^2}\right]\right]} \qquad \textit{Equation 6-1}$$

where:

- x_i and y_i are observations at time instant i of variables x and y respectively
- $\bar{x}$ and $\bar{y}$ are the mean values of variables x and y respectively.

By using the correlation coefficients, the weekly and same weekday profiles were analysed according to the following four scenarios, which were formulated to understand the patterns of the frequency of occurrence of chaos across the entire network of interconnected links:

1. ***Scenario A1***: All the paired combinations of the "same weekday" for Monday, Tuesday, Wednesday, Thursday, Friday, Saturday and Sunday across the month. For example, the correlation coefficient for Monday 1st February against Monday 8th, 15th, and 22nd February. Similarly, the coefficient for Tuesday 2nd February against Tuesday 9th, 16th and 23rd February etc.
2. ***Scenario A2***: All the paired combinations of each day versus the average profile for its corresponding same weekday. For example, the correlation coefficient of the profile for Monday 1st February 1999 against the average Monday profile for the month of February. Similarly, the coefficient of the profile for Tuesday 2nd February 1999 against the average Tuesday profile for the month of February etc.
3. ***Scenario A3***: All the paired combinations of each day versus its corresponding average weekly profile. For example, the correlation of Monday 1st and Tuesday 2nd February each against the average profile for 1st–7th February. Similarly, the correlation of Monday 8th and Tuesday 9th February against the average profile for 8th – 15th February.
4. ***Scenario A4***: All the paired combinations of the average profile of the seven "same weekday" analysed. For example, the correlation of the average profile for Mondays against the average profile of Tuesday, Wednesday, Thursday, Friday, Saturday and Sunday etc.

6.7.3 Determining the Longest Lag

The Cross-Correlation function is simply the linear correlation function expressed as a function of the time lag. The Cross-Correlation helps to find the value of the shift that maximises the correlation coefficient. The cross correlation r at delay d is defined as:

$$r = \frac{\sum_i (x_i - \bar{x})\,(y_{i-d} - \bar{y})}{\left[\sqrt{\sum_i (x_i - \bar{x})^2}\right]\left[\sqrt{\sum_i (y_{i-d} - \bar{y})^2}\right]} \qquad \textit{Equation 6-2}$$

where:

- x_i and y_i are observations at time instant i of variables x and y respectively
- $\bar{x}$ and $\bar{y}$ are the mean values of variables x and y respectively.

The Cross-Correlation helps to find the value of the shift that maximises the correlation coefficient. For any range of d values, the lag is given by the d-value with the highest correlation coefficient. This value is the cross-correlation lag (CCL). Normally, the function evaluates the lag as negative because it measures the relationship between time series of variable x (first variable) and the past time series of variable y (second variable).

The following steps, also coded in R in Appendix B (see program 1), were used to determine the combination of variables that gave the longest lag:

1. Calculate the cross-correlation for all the possible pairs of the hourly flow across all detectors in the network.
2. Select the highest lag associated with the highest correlation coefficient for every hour for each link.
3. Calculate the average lag for every link for every hour of the day.
4. Repeat step 1 through to 3 for flow-flow, occupancy-occupancy, Lyapunov exponent- Lyapunov exponent, occupancy-Lyapunov exponents and Lyapunov exponents-occupancy.
5. Calculate the percentage increase of lag for flow-flow, occupancy-occupancy, occupancy-Lyapunov exponent and Lyapunov exponent-occupancy over the Lyapunov exponent-Lyapunov exponent.

It must be noted that the cross-correlation lags are different from the autocorrelation lags. The study identified cross-correlation lags in the order of minutes and the autocorrelation lag in hours. The algorithm above is based on the cross-correlation analysis, which gives shorter lags. However, the sampling interval of one hour is more related to the standard practise in traffic studies, where it is normal to study traffic in hourly segments to identify for example peak hour or inter-peak etc.

6.8 Short-term Link Predictions

Sections 9.4 and 9.5 are devoted to the spatial analysis and short-term prediction of the network links. In order to facilitate the process, an R program was developed to undertake the tasks outlined in Sections 6.7.3 and below, as well as for additional procedures for the short-term link forecasts. The algorithms below suggests that the dynamic model has the ability to determine a unique function for forecasting each link for each hour in the day (for

all days in each cohort) depending on the qualifying links meeting a specified criteria for lags and correlation coefficient.

A translated version of the program is described below for ease of understanding of the key steps, whilst the full R program is included in Appendix B (see Program 2). The steps are:

1. To create files to store the monthly data of the network links in columns and grouped under flow, occupancy and Lyapunov exponent respectively.
2. To select the file of data for a specific month for modelling each link.
3. To identify the *model lag*, which specifies the range of the data for testing of the regression of the model.
4. To create a day to carry out the modelling exercise.
5. To create two empty matrices one for the predicted and the other for the actual data to compare for the links for the chosen day. The matrices are labelled '*full_pred*' and '*full_test*' respectively.
6. To choose the occupancy data for every link one-by-one, and for each selected link process the data hour-by-hour.
7. To create two blank matrices labelled '*mlag*' and '*mcor*' in which to store the lag and maximum correlations respectively, from a cross-correlation assessment.
8. To perform cross-correlation of the chosen link's hourly occupancy data and the Lyapunov exponent data of all other links.
9. Repeat step 8 by varying the two datasets depending on what is being predicted and the data used. The scenarios analysed are presented in section 6.8.1.
10. To select the maximum correlation coefficient and corresponding cross-correlation lag (CCL) and insert into matrices '*mcor*' and '*mlag*' receptively.
11. To repeat the exercise for all 24 hours in the chosen day for all 40 links.
12. To select from the data generated for each link for every hour, the links with correlation coefficient and lag values meeting a specified criteria.
13. To repeat step (11) above for all scenarios adjusting the specified criteria (see Section 6.8.1) appropriately.
14. To construct a regression model for each link for every hour using the appropriately lagged data of Lyapunov exponents of the links that passed the criteria to forecast the link occupancy.

15. To test the constructed models for a number of scenarios, which are described in detail in Section 6.8.1. The scenarios vary in terms of the assigned critical values of correlation coefficients and lags, and the selected links for predicting the lane occupancies.
16. To modify the *model lag* appropriately depending on the forecast horizon in order to choose the test data.
17. To insert the test data for each link for every hour into a new matrix called '*full_test*', which stores all the daily predictions for the links at the required resolution (i.e. sampling frequency of the data).
18. To insert the lagged test data in the regression model in order to predict the link occupancy.
19. To insert the predicted data for each link for every hour into a new matrix *'full_pred'*, which stores all the daily predictions for the links at the required resolution.
20. To calculate the coefficient of regression for the linear relationship between the '*full_pred*' versus '*full_test*'.

6.8.1 Modelling Scenarios

In order to test the model, the Lyapunov exponents from a preceding time segment (based on the lag) were input into the model, and the model used to predict the occupancy for a succeeding time segment. The scenarios were developed to assess the suitability of predicting the link occupancy from the Lyapunov exponents rather than the occupancy data. The five scenarios tested are:

1. ***Scenario B1*** predicts the hourly occupancy profile **24 hours ahead** (for example, Lyapunov exponent (LE) data from 0700h-0800h is used to construct the model, then the next day 0700h-0800h LE data is selected to forecast the 0700h-0800h occupancy). The "specified criteria" for the "selected links" is maximum correlation coefficients and lags greater than 0.4 and 100 seconds respectively.
2. ***Scenario B2*** predicts the hourly occupancy profile **15 minutes into the next hour** (for example, LE data from 0700-0800 is used to construct the model, then the 0715-0815 LE data is selected to forecast the 0800-0815 occupancy). The

"specified criteria" for the "selected links" is maximum correlation coefficients and lags greater than 0.4 and 100 seconds respectively.

3. ***Scenario B3*** is same as ***Scenario B2***, except that occupancy data in used instead of LE data.
4. ***Scenario B4*** is same as ***Scenario B2***, except that the threshold criteria for the "selected links" for the maximum correlation coefficient is greater than 0.7(only the stronger coefficients).
5. ***Scenario B5*** is same as ***Scenario B2***, except that the threshold criteria for the "selected links" for the maximum correlation coefficient is greater than 0.3 (inclusive of the weaker coefficients).

6.9 Spatial Analysis

The cross-correlation analysis helps to understand the relationship between the causes of events and corresponding spatial effects on the network. By using the cross-correlation function, it was possible to determine the spatial impacts of events at one location by an indication of the strength of the relationship with other locations of influence and the duration interval over which the effects occur. This analysis is presented in section 9.6. In brief, each link is cross-correlated with all other links in the network. That is, if there are links 1, 2, 3 … 40, link 1 will be cross-correlated with all the other 39 links, then link 2 with the other 39 links, and so on.

The process involved in assessing the spatial impacts of how activities at one location relate to another location in the network is as follows:

1. Calculate the cross-correlation for the hourly Lyapunov exponents across all detectors in the network.
2. Summarise the analysis for each hour giving two 32 x 32 matrices for (i) maximum correlation coefficient and (ii) corresponding lag times. There are therefore 24-paired matrices for the correlation coefficient and the corresponding lagged times for each day.
3. Focus only on the correlation coefficients that are significant by using a threshold of significance for the correlation coefficient set as a minimum of 0.4 and extract the qualifying links meeting this criteria together with the corresponding lagged times, for those "upstream links" (links gaining right of way) and their "feeder links" (preceding links losing right of way).

4. Identify the most critical link in the network, which is the links among those gaining right of way that has the highest number of correlations to those losing right of way. The critical link has wider spatial impact, whilst the range of lagged time gives an indication of the window of opportunity to fix network problems ahead of real-time.

6.10 Conclusions

The methodology for undertaking the research requires analysis of an extensive amount of traffic data and involves several iterations of the steps enumerated in this chapter (see Sections 6.6, 6.7 and 6.8). Therefore, it is important to process each individual step by coding in either Microsoft C# or R program in order to process the analyses quicker, minimise the time demands of the user and systematically iterate through the steps as the number of iterations required is not feasible to achieve manually. The necessity of the programming aspects of the work confirm that the computer is an indispensable tool for analyses deploying Chaos Theory endorsing the literature review (see Section 4.3.2)

Given the general lack of normality in the distribution of the raw data and output of the models, non-parametric tests (including Mood's Median, Mann-Whitney and Kruskal-Wallis) were used in the analysis when exploring relationships and to determine the statistical significance of the results. These were mainly undertaken using Microsoft Excel and Minitab, which were also used for displaying descriptive statistics, regression analysis, time series plots as well as graphical displays such as scatterplots, bar charts, boxplots and histograms.

Therefore, this chapter has set out the *rationale* for the proposed methodology for this research work. Consequently, in subsequent chapters, this methodology is applied to flow and occupancy data for a SCOOT controlled region in Leicester (Data 4 and 5). Prior to the study of this network, data including noise from motes (in Medway) (Data 1) and flow and occupancy (of a single link located in Leicester) (Data 2 and 3) were tested in order to ascertain the performance of the models, determine the requirements for selecting the research data, and to enhance understanding of the key parameters of Chaos Theory. Therefore, Chapter 7, which begins Part III, presents the results of the models as applied in the preliminary study to noise, flow and occupancy data (Data 1, 2 and 3 respectively).

PART III: A CASE STUDY OF WELFORD ROAD SCOOT REGION, CITY OF LEICESTER

Chapter 7. Preliminary Assessments of Noise, Flow and Occupancy Data

7.1 Introduction

The purpose of the preliminary assessment is to answer research question numbers two and three. More specifically, it is to identify the datasets with appropriate characteristics (see Section 3.6) such that the underlying structure, which will allow analysis using Chaos Theory exists. Also, the temporal resolution of the traffic data was investigated given that this is an important factor in determining the suitability of the dataset for analysis using Chaos Theory. Both the data characteristics and temporal resolution are important to ensure that the right attributes (such as the sampling rate) exists in the traffic data in order to progress the analysis. This Chapter also develops understanding of the evolution of chaos across a single loop detector, before applying this knowledge to a network of detectors within a SCOOT controlled region in Chapter 8 and Chapter 9. This assessment will also enable prior testing of the models, described in Section 6.6 in order, to demonstrate their suitability for analysing the study network for the main research work.

These preliminary assessments make use of three types of datasets to run the models[2]. These are (i) noise (Data 1) (ii) flow (Data 2) and (iii) occupancy (Data 3). The noise data was from a network of motes in Medway, whilst the flow and occupancy data were from a SCOOT controlled region in Leicester City. The sampling frequency of traffic noise was 60 seconds, but interpolating the raw SCOOT data gave 1, 10, 20, 30, 60, and 120 seconds interval flow and occupancy data. The data was available in Comma Separated Variable (*.csv*) format, in four/five column entries with each row containing a unique data string, the mote identification number (which was the same for all entries within a particular file), date and time stamp, and noise level measurement. This chapter discusses the results of the analysis of the three main parameters (delay time, embedding dimension and Lyapunov exponents) derived from the data and associated with Chaos Theory applied to the aforementioned data types. Section 7.2 discusses these parameters based on traffic noise, and the parameters from flow and occupancy are discussed in section 7.3. As referenced

[2] The raw data (noise, flow and occupancy) are not included in the appendix due to its large volume, but available in electronic format, if required

in Section 3.4, Sections 7.2.3 and 7.3.3 specifically looks into issues regarding the resolution of data when analysing chaotic systems.

7.2 Noise Data (Data 1)

Data 1 was available from 20 mote locations in the Medway urban area, and sampled at one-minute interval for the months of April, May, June, and July 2011.

7.2.1 Delay Time (Autocorrelation Lag) for Noise

The autocorrelation lag was computed independently for each mote location using data for weekly, monthly, and then for a continuous 2-month and 3-month period, in order to assess the sensitivity of the lag value to the total length of the data stream. The ranges in minutes of lags tested were 0-100, 0-500, 0-1000, 0-1500, 0-2000, 0-2500, 0-3000, 0-3500, 0-4000, 0-4500 and 0-5000. For the longer (upper end) range of lags, the plot of autocorrelation coefficient versus delay time showed a polynomial relationship with multiple solutions at an autocorrelation coefficient of 0.4. On the other hand, there were no solutions at the shorter (lower end) lags. All ranges of lags were tested and the single range of 0-500 was found to be most appropriate based on the standard criteria of an autocorrelation coefficient equal to 0.4.

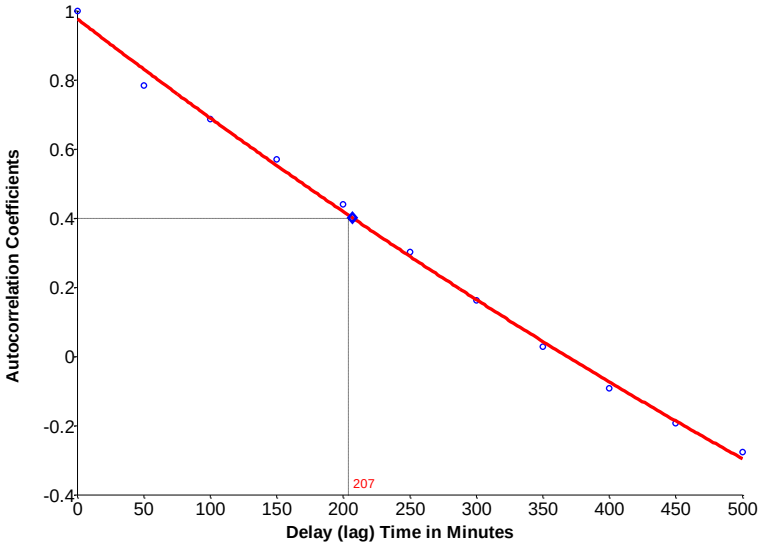


Figure 7.1: Typical Autocorrelation Plots for Mote 106 (May 2011)(based on Data 1)

Figure 7.1 is an example of a correlogram showing a plot of the autocorrelation coefficient versus delay time (or lag) for mote 106 in the month of May 2011. Table 7-1 shows all the results of the optimal autocorrelation lags for the tested scenarios based on the autocorrelation plots for the weekly and monthly scenarios.

	April 2011					May 2011					June 2011					July 2011				
Mote ID	Week 1	Week 2	Week 3	Week 4	All	Week 1	Week 2	Week 3	Week 4	All	Week 1	Week 2	Week 3	Week 4	All	Week 1	Week 2	Week 3	Week 4	All
21						278	212	215	215	234	227	205	210	240	234	220	215	212	215	234
36						165	160	165	177	175	175	160	167	162	175	165	175	165	175	175
39						195	155	160	175	175	188	165	155	157	175	155	157	149	150	175
43						156	150	150	165	165	155	150	155	150	165	155	160	150	155	165
46						150	140	149	155	154	148	147	145	138	154	141	145	140	145	154
50						152	140	150	155	156	149	137	147	143	156	155	150	145	155	156
53						170	165	170	200	175	172	165	170	175	175	170	170	168	166	175
58						115	105	93	115	113	120	110	117	103	113	97	87	87	90	113
59						143	138	120	150	148	133	131	125	117	148	125	124	118	125	148
65						125	125	145	155	151	128	148	147	133	151	150	158	131	130	151
87						200	190	190	200	214	202	*	200	186	214	*	*	*	*	214
102						184	177	180	170	183	181	175	179	175	183	175	180	170	180	183
105		250	220	208	265	202	200	200	204	203	202	205	205	*	202	204	225	205	*	182
106		212	227	213	230	210	207	207	206	205	208	210	209	*	207	204	225	210	*	212
107		228	215	190	225	187	165	180	185	185	183	172	174	170	185	180	185	218	185	185
108		220	215	170	210	195	187	187	195	189	178	168	175	*	179	191	200	181	*	190
109		250	230	195	235	203	207	207	208	210	206	204	200	*	201	205	209	180	*	200
112						214	182	195	226	212	210	205	210	195	212	212	200	185	190	212
115						163	150	165	187	166	155	150	155	162	166	163	160	145	145	166
116						162	150	155	160	163	155	151	153	146	163	155	155	150	155	163

Table 7-1: Summary of Delay Times (based on Data 1)

For reasons explained in Section 5.5.2, there is a deliberate attempt in these calculations to ensure that the length of the lag is neither too short nor too long in order that when used in subsequent processes the ability of the phase space to represent the actual, but unknown, system is not compromised. Accordingly, a solution for an autocorrelation of 0.4 yields a relatively moderate length of lag for the phase space reconstruction.

From Figure 7.1, when autocorrelation reaches the value 0.4, the lag is 207. A closer inspection of the graph through sensitivity analysis indicates that when autocorrelation reaches the value 0.7, 0.6, 0.5, 0.3, 0.1 and zero, the lag is approximately 100, 133, 167, 232, 299 and 334 respectively. Therefore, the results reveal that the 1st, 2nd and 3rd scenario tests (with stronger correlations) gave relatively shorter lags compared to an autocorrelation of 0.4, whilst the 4th, 5th and 6th scenarios (with weaker correlations) gave relatively longer lags. Given that the longer and shorter lags adversely affect the evolution of the trajectories in the phase space, as discussed in Section 5.5.1, the results confirm the choice of an autocorrelation of 0.4 as ideal.

There was no consistency in terms of the magnitude of the delay times, across the monthly data stream or combination of months for that matter. The delay times for the various scenarios were evaluated to establish whether the optimal values were sensitive to the duration of the time series data stream. As the distribution of the optimum delay times was not normal, the Kruskal-Wallis statistical test was used to compare the optimal delay times derived from the model. A comparison of the delay times based on the weekly or monthly total gave a p-value of 0.000 indicating statistically significant differences in the medians across the motes location. Similarly, when comparing the delay times using a different duration of data streams both weekly and monthly, the test gave a statistically significant p-value of less than 0.05. Based on data sampled at one-minute intervals, this result implies that the delay time is sensitive to the total duration of the data stream.

These findings are important for the subsequent stages in the application of Chaos Theory, because the delay time (or lag) serves as input into the construction of the traffic phase space. As the lag values vary from one location to another, it is incorrect to generalise the lag durations across different spatial areas. Therefore, every location's lag needs calculating independently for the construction of a unique phase space for each mote location.

The duration of the lag is sensitive to the total length of the data, and gives different outcomes for the weekly and monthly datasets. That is, the same time series information when assessed based on different lengths of time series data stream (for example hourly, daily, weekly, monthly or yearly) will yield different outcomes. Therefore, it is important to maintain a consistent data duration across all the locations, but each of these locations will generate its own delay time (or lag) that must be used for its phase space construction. As explained in section 5.4.1, the duration of the data stream required to construct the phase space varies depending on the dimensionality of the phase space. Therefore, a large volume of data (several thousand data points) is required for construction of the phase space, specifically because the actual embedding dimension is unknown and it is necessary to test multiple possibilities in order to determine this saturated value that would adequately embed the chaotic attractor. This assessment used a dataset of one-month duration for each location to ensure a robust assessment of the dimensionality of the phase space.

7.2.2 Embedding and Correlation Dimension for Noise

In order to determine the embedding and correlation dimensions, a number of phase spaces were constructed with assumed embedding dimensions of 1, 2, 3… 10. The lag values derived from the monthly data streams were input. Figure 7.2 is a typical *m*-curve derived for noise (Data 1) based on mote 105 for the month of April 2011. The existence of these log-log curves confirms the existence of chaotic character in the noise data as discussed in Section 5.4.2 (Jianming *et al*., 2003).

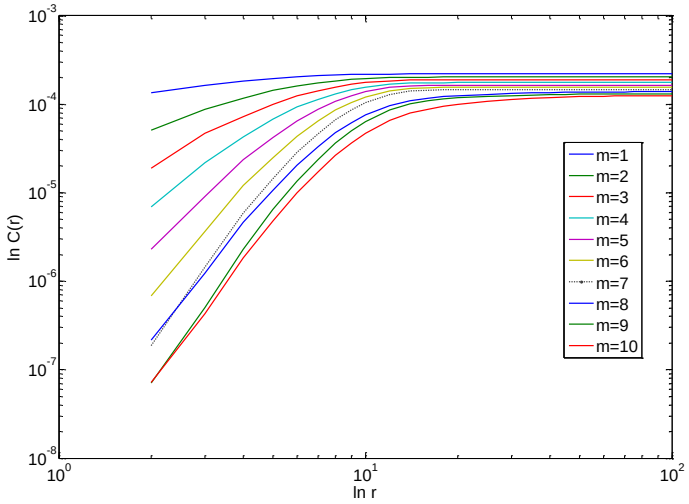


Figure 7.2: m-Curves Showing Relationship between Ln C(r, m) and Ln r (based on Data 1)

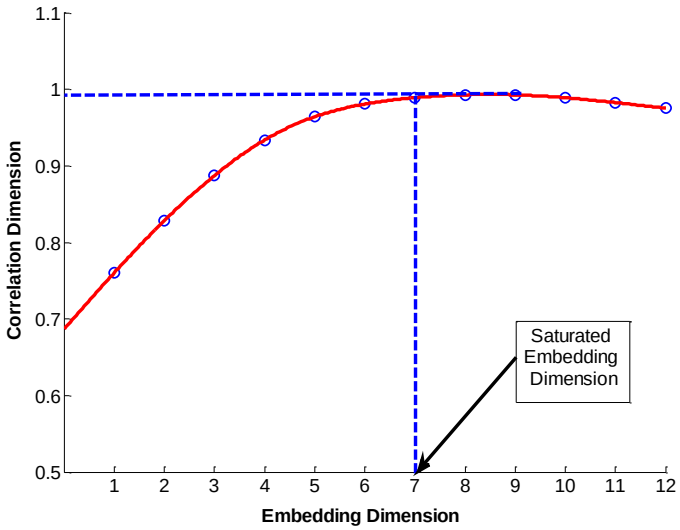


Figure 7.3: Relationship between Embedding and Correlation Dimensions (based on Data 1)

Figure 7.3 shows the relationship between the embedding dimension and the correlation dimension, which indicates that the saturated embedding dimension is 7 and the corresponding correlation dimension is 0.99. Table 7-2 presents a summary of the correlation and embedding dimensions for motes 105 – 109 for the inclusive months April – July 2011. Therefore, the results of the analysis confirm that chaos does exist in traffic noise (Data 1) because embedding and correlation dimensions are unique properties of chaotic systems.

Table 7-2 shows that the embedding and correlation dimensions for every location were generally similar irrespective of month but can be different for each mote. The estimated p-values are 0.494, 0.440, 0.615, 1.000 and 1.000 respectively for motes 105, 106, 107, 108 and 109. The p-values indicate that the results of the embedding dimensions for each mote across the months are not statistically significantly different at 95% confidence interval.

	Embedding Dimension				Correlation Dimension			
ID	April	May	June	July	April	May	June	July
105	7	6	6	7	0.99	0.98	0.98	0.99
106	6	5	5	9	1.0	0.99	0.99	0.99
107	11	9	9	9	0.97	0.95	0.95	0.95
108	6	6	6	6	0.99	0.98	0.98	0.98
109	11	11	11	11	0.97	0.96	0.96	0.97

Table 7-2: Summary of Embedding and Correlation Dimension for the Inclusive Months April-July 2011 (based on Data 1)

There is however, a difference in results from one mote location to another and the test indicated a statistically significant p-value of 0.007. Similarly, the results of the correlation dimensions were not statistically different for every mote, but different when compared across the different spatial locations. The consistency of values of the embedding and correlation dimensions from month to month at every location is an indication that there is a structure, which is inherent in the noise data. The results suggests that these parameters are unique properties, which are unaffected by the length of the times series data string so long as the associated delay time is also used to construct the phase space.

In summary, the delay time is sensitive to the length of the data series string. However, the embedding and correlation dimensions are constant determinants of chaotic systems if the phase space and lagged times are derived from the same volume of data. Given this observation, it is important to use a consistent set of data to determine the optimal delay, embedding dimension and its corresponding correlation dimension and then subsequently for all other calculations (Lees, 2013).

7.2.3 Lyapunov Exponents Profile for Noise

By using the saturated embedding dimension and the lag time, the instantaneous Lyapunov exponents[3] were calculated from the reconstructed phase space of the time series string of noise data to assess the underlying structure of the system's evolution. Figure 7.4 presents a typical time series profiles of the Lyapunov exponents based on the traffic noise data.

For dynamic systems, a positive Lyapunov exponent indicates chaos; a negative value indicates a stable state whilst zero is the meta-stable state (see Section 3.2). The entire observation of Lyapunov exponents was either wholly positive or negative (for example Figure 7.4 shows only positive exponents), which therefore did not enable an understanding of the characteristics of the cyclical nature of the dynamic state of the network. The daily profiles were of a fluctuating character, but generally rising over the course of the day so that the start and end observations were completely at variance. Ideally, it would be expected that the curve does not rise indefinitely but descend during the evening period when there is relatively less chaos at this time. Consequently, this structure is not suitable

[3] The Lyapunov exponent profiles for noise, flow and occupancy are not included in the appendix due to its large volume, but available in electronic format, if required

for explaining any patterns in traffic flow and therefore does not support the aspirations of this research work.

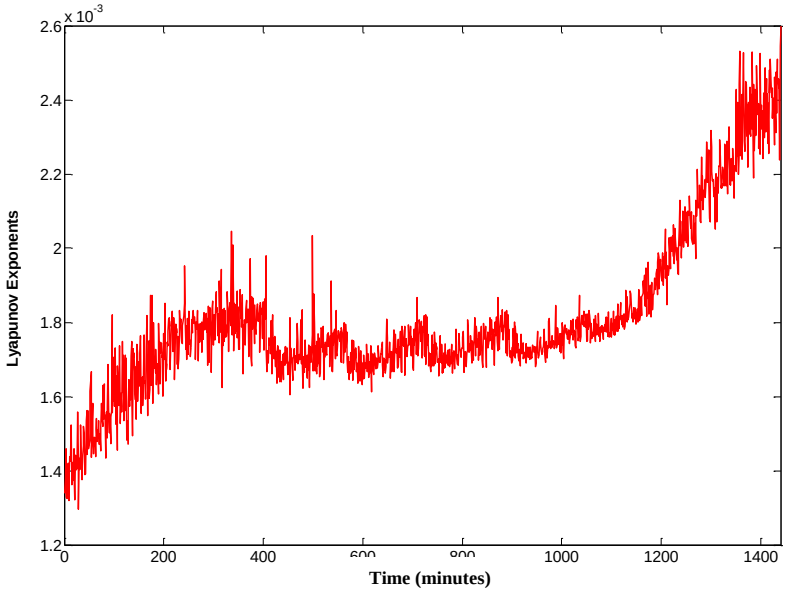


Figure 7.4: Typical (Daily) Time Series of Lyapunov Exponents (based on Data 1)

Given the evidence of the existence of delay duration (or lag time) along with embedding and correlation dimensions, which were indicative of some underlying chaotic behaviour, a more structured pattern is expected. It has been reported in the existing literature that the sampling of the traffic data at the incorrect frequency may not capture the system's chaotic properties with sufficient statistical significance (Frazier and Kockelman, 2004). Therefore, it could be possible that noise data averaged over 60 seconds is inappropriate for identifying chaos. However, with time resolved data for example averaged out over 20 seconds, or even lower frequencies down to the sampling rate, may reveal chaotic behaviour. However, the raw noise data (Data 1) is not available, only the 60 second averaged data is available. Therefore, it was not possible to interrogate the data extensively or undertake further assessments to establish the suitability of such data for analysing dynamic systems.

In SCOOT for example, due to the cycle times varying in response to changes in traffic demand at different times of the day, there is additional data that makes it possible to harmonise the data at a finer resolution. Therefore, representations of the SCOOT data were generated at various sampling intervals, which was not possible with the noise data because of the nature of how the noise data are generated.

Therefore, given that the Lyapunov exponent profile (Figure 7.4) suggests that the traffic condition at the location of this mote is in a state of chaos throughout the 24-hour period, which is very unlikely, there was no (reasonable) underlying structure to justify further investigations using the traffic noise dataset (Data 1) for the main research work. This problem may be due to the traffic noise data (Data 1) for computing the separation between points in the phase space being logarithmic (given by $L = 16\log(speed) + 9\log(flow)$). Once transformed into the phase space, the distances between the points, as used in the Lyapunov exponent function, no longer become meaningful. This makes the instantaneous Lyapunov exponents (from a logarithmic expression) based on reconstructed phase space invalid and therefore potentially explains why it is may be difficult to identify a suitable structure in the time series profile.

7.3 SCOOT Flow and Occupancy Data (Data 2 and 3)

Repeating the same procedure, the SCOOT flow and occupancy data for link N03221I1 (Data 2 and 3 respectively) were analysed for the months of May, June, and July 2002. The data were reconstructed at frequencies of 1, 10, 20, 30, 60 and 120 seconds intervals in order to determine the range of sampling frequencies that affect the generation of the underlying structure for the chaotic dynamics. This was found to be necessary because the 60 second frequency noise data (Data 1) proved to be unsuitable in the preliminary stage of this research. Therefore, for each sampling interval, the delay times and correlation and embedding dimensions, which indicate that the time series data is a product of dynamic system that generates chaos, were calculated.

Similar to the noise data (Data 1), there was no structured profile found for the Lyapunov exponent for data sampled at 1, 10, 30, 60 and 120 second intervals for Data 2 and 3. However, the Lyapunov exponent profiles for the 20 second interval data yielded structured patterns and therefore were deemed useful in achieving the outcomes of this research work by enabling an understanding of the propagation of the dynamic states in signalised urban networks by applying Chaos Theory. The results for the data sampled at 20 second intervals is presented in Sections 7.3.1, 7.3.2 and 7.3.3 respectively with respect to calculating the delay time, embedding and correlation dimensions, and the Lyapunov exponents.

7.3.1 Delay Time (Autocorrelation Lag) for Flow and Occupancy

The optimal delay times based on the SCOOT flow data for link N03221I1 (Data 2) for April, May, June and July were 303, 280, 243 and 253 respectively. Similarly, the corresponding values for occupancy (Data 3) are 290, 323, 240 and 273 respectively.

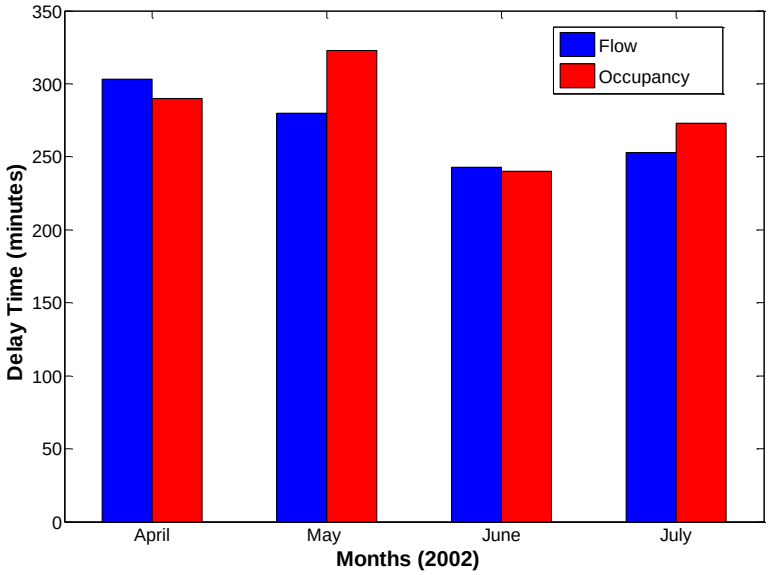


Figure 7.5: Delay Times for Flow and Occupancy for Inclusive Months April-July 2002 (based on Data 2 and 3)

Figure 7.5 presents a graphical display of the delay times for ease of comparison. There was no observed trend in the observations, but a comparison of the two groups using the Mood's median test gave a significant p-value of 1.000 at 95% significance level, which indicates that there was no statistical difference in the observations of the flow and occupancy (Data 2 and 3) across all months.

7.3.2 Embedding and Correlation Dimensions for Flow and Occupancy

Both the flow and occupancy data for link N03221I1 (Data 2 and 3 respectively) consistently indicated an embedding dimension of 3 across all months, whilst the correlation dimensions ranged from 0.14–0.16 and between 0.28–0.26 for flow and occupancy, respectively, as shown in Figure 7.6. For flow (Data 2), the correlation dimensions are 0.15, 0.16, 0.15 and 0.14, and occupancy (Data 3) are 0.28, 0.24, 0.28 and 0.26, for April, May, June and July respectively.

Figure 7.6 shows that the correlation dimension for occupancy is generally higher than that for flow. The Mood's median test indicated that the correlation dimension of the two groups was statistically different with a significant p-value of 0.005. However, the

correlation dimension was not statistically different between months for all the observations for each group. The Kruskal-Wallis test (which was appropriate because more than two months were sampled) indicated p-values of 0.440 and 0.440 for flow and occupancy. This result was similar to the case of the traffic noise data (Data 1), and confirmed that the embedding and correlation dimensions are unique characteristics of an evolving dynamic system.

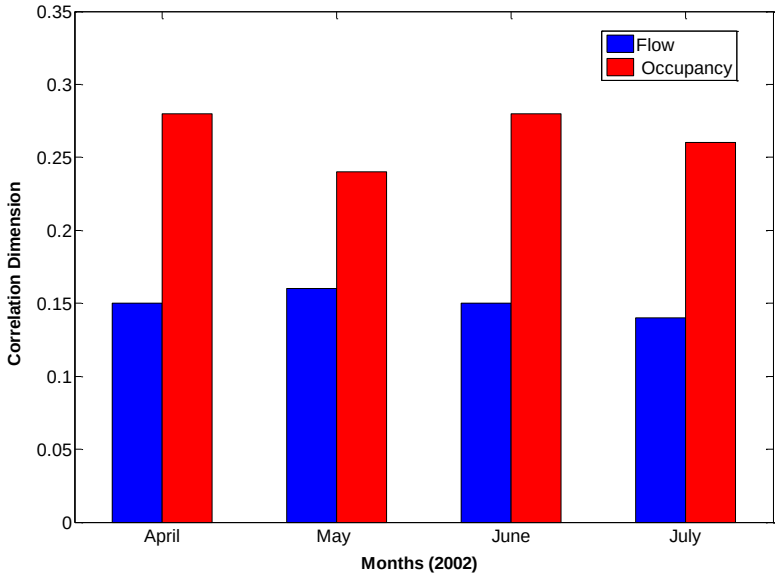


Figure 7.6: Correlation Dimensions for Flow and Occupancy for Inclusive Months April-July 2002 (based on Data 2 and 3)

In order to ensure confidence in the suitability of the phase space, the computed values of the parameters were further tested against Takens' criterion, which is given by $m^* \geq 2{D_2}^* + 1$, where m^* and ${D_2}^*$ are the saturated embedding and correlation dimension respectively. By using the highest value of ${D_2}^*$, and m^* equal to 3, it is proved that for flow (Data 2): $3 \geq (2 * (0.16) + 1 = 1.32)$ and occupancy (Data 3): $3 \geq (2 * (0.28) + 1 = 1.56)$. In both cases, the results indicate that the embedding dimension is more than 10 times greater than the correlation dimension, as suggested by Shang *et al*. (2007). Therefore, the valid m^* and ${D_2}^*$ values indicate that the phase space is sufficiently large to embed the trajectory of the attractor from both flow and occupancy (Data 2 and 3) time series. The results further indicate that flow and occupancy (Data 2 and 3) time series have low dimensional properties for the traffic flow system, and therefore subsequent analysis of the profile of Lyapunov exponents presented in the next section will confirm the suitability of the Phase Space Reconstruction theory for low dimensional time series data.

7.3.3 Lyapunov Exponent Profile for Flow and Occupancy

The independently reconstructed phase space for flow and occupancy (Data 2 and 3) enabled the instantaneous Lyapunov exponents to be calculated. In each case, the calculations were based on monthly traffic flow and occupancy data, and the phase space that relied on inputs of the estimated delay time and embedding dimension. Unlike the case of the traffic noise data (Data 1), the profiles fluctuate about the zero reference line and show a structured pattern that enables the interpretation of the traffic flow in terms of its dynamic behaviour.

Figure 7.7 shows graphically typical hourly profiles of the Lyapunov exponent of the flow and occupancy variables (Data 2 and 3) for 0800h-0900h on 10th May 2002. These profiles were plotted on the same graph for ease of comparison; however, the plots are on different scales due to the difference in the order of their magnitude values. The primary and secondary axes are Lyapunov exponents related to occupancy (Data 2) and flow (Data 3) respectively.

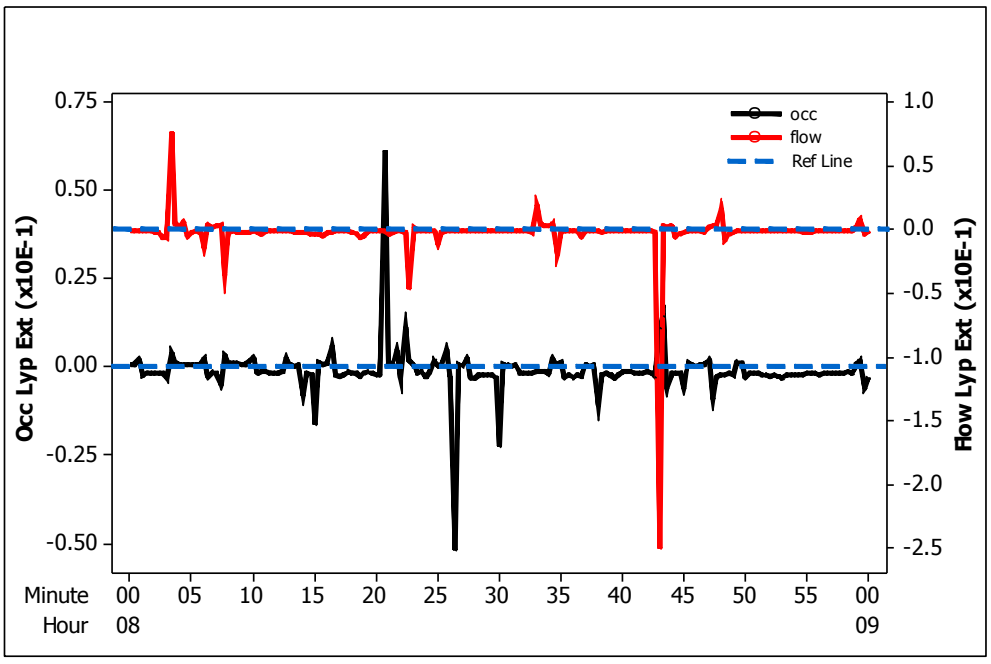


Figure 7.7: One-Hour Graphical Profile of Lyapunov Exponent Based on SCOOT Data (0800h-0900h) (based on Data 2 and 3)

The points above the zero reference line indicate an unstable state (chaos) and the points below represent the stable state whilst the zero line shows the transition between these two states. The profiles generally provide a similar message from their respective Lyapunov

exponents. For example, both profiles indicate a steady state behaviour during the period 08.28-08.34, 08.36-08.42 and 08.50-08.59; and chaos between 8.03-8.04 and 8.59-9.00. Flow (Data 3) indicated chaos occurred at the following times: 08.07, 08.33 and 08.48, and it is relevant to note that occupancy (Data 2) also indicated chaos almost one minute after each occurrence at times 08:08, 08:34 and 08:49. This is perhaps an indication that the flow and occupancy may detect chaos at slightly different times, an issue that is investigated further in Chapter 9; however, the similarity may also be indicative of the fact that these variables are directly proportional at low values (see Equation 2-32) (Athol, 1965).

Further analysis comparing the Lyapunov exponents profiles from flow (Data 2) and occupancy (Data 3) showed statistical similarity. The tests performed for different days consistently showed a statistical similarity in the hourly Lyapunov exponents of flow and occupancy variables. The Kruskal-Wallis test consistently gave p-values of greater than 5%, which indicates that the Lyapunov exponents from the two sets of data were not statistically significantly different. Given that flow and occupancy and *vice-versa* infer the same results, both variables were considered suitable in this research for understanding the evolution of the dynamic state of traffic in a signalised network.

The observations in these profiles are evidence of the capability of Chaos Theory to be useful in understanding the dynamic behaviour of traffic on links in a signalised urban network. An identification of chaos in traffic flow or occupancy data is very important for identifying the sources of data for this research work. Chaotic traffic causes "stop-and-start" traffic that forces acceleration and deceleration to occur often, which is uncomfortable and undesirable by motorists and can lead to accidents. The awareness of chaos is therefore important for transport planners and engineers, and traffic operators to initiate plans to maintain stable flow by improving the level of service and consequently the safety of the network (Xu and Gao, 2008).

7.4 Conclusions

In this chapter, times series data for noise, flow and occupancy were examined to identify the types of datasets suitable for an in-depth study of the potential use of Chaos Theory in traffic studies. The analyses are based on the algorithms of the mathematical framework outlined in Chapter 5, and described in Chapter 6. Firstly, the main parameters to calculate were delay time, embedding and correlation dimensions, which were used for

reconstructing the traffic phase space. Secondly, the phase space was used to calculate the instantaneous Lyapunov exponent. It was important in this process to ensure that the same dataset was used consistently to estimate the delay time, embedding and correlation dimensions, and the Lyapunov exponents.

The analysis revealed that noise data (Data 1) sampled at a rate of 1 second but only available as averages over a 60 second interval were unable to generate the patterns of Lyapunov exponents necessary for analysis using Chaos Theory. It was likely that the noise data averaged over 60 seconds were inappropriate; however, the raw noise data (Data 1) was not available so the data could not be interrogated at other (higher) resolutions to establish the suitability of such data. Given the aspirations of this research to understand the evolution of traffic based on Chaos Theory, the lack of unsuitable patterns from the noise generated Lyapunov exponents did not make Data 1 a viable option for further investigations in this research. Therefore, Data 1 was not used in the main research analysis in Chapter 8 and Chapter 9.

Further, the analyses indicated that the 20 second interval data Lyapunov exponent profiles for flow (Data 2) and occupancy (Data 3) could explain the underlying dynamics responsible for the evolution of the system. However, the profiles for flow and occupancy (Data 2 and 3) at 1, 10, 30, 60 and 120 second interval data were unsuitable. Therefore, the results confirm that reasonably highly resolved traffic data are required in order to explore applying Chaos Theory to traffic networks. Given that the SCOOT flow and occupancy data (Data 2 and 3) were shown to help improve our understanding of the road condition on a single lane, it offers potential for extension to the analysis of the urban road network. In conclusion therefore, flow and occupancy data (Data 2 and 3) of 20 second intervals are suitable types of traffic data for investigating chaotic behaviour in urban traffic systems. The suitability of both variables demonstrates that the data source is limited to not only SCOOT but also any system. For example MIDAS that the provides only flow data, but not occupancy.

This type and characteristics in terms of resolution of traffic data will be the basis for further analysis in this research. Chapter 8 and Chapter 9 extend the knowledge of dynamic behaviour from this preliminary assessment to a study area comprising of a SCOOT network of 40 interconnected links in Leicester City. Chapter 8 makes use of this

understanding of dynamic systems to identify the status of a signalised urban network, whilst Chapter 9 explores the potential for the short-term forecasts of traffic lane-occupancy.

Chapter 8. Network Status Analysis

8.1 Introduction

This chapter extends the application of the algorithms of Chaos Theory to a network of SCOOT links, in order to assess whether or not chaos as a phenomenon occurs network-wide across connected links, with reference to the control region in Figure 6.2. The fundamental understanding of chaos, in the context of the evolution of traffic congestion at a single loop (section 7.3.3), identified that traffic flow and occupancy data sampled at 20 second intervals were a suitable temporal resolution for a detailed study of chaotic traffic on links in a network. Given that Chaos Theory is data-hungry, requiring large volumes of data, the assessment makes use of continuously sampled data of seven monthly cohorts that passed the 75% capture rate qualifying criteria (see Section 6.3.1). It is envisaged that the properties of the dataset applied to a signalised urban network will provide deeper insights into the chaotic behaviour of traffic that will ultimately be of benefit in the management of traffic in urban areas.

In order to present the assessments undertaken, the structure of this chapter is as follows. Section 8.2 analyses the raw flow and occupancy data (Data 4 and 5 respectively) for each link in order to extract the descriptive statistics, and assesses the similarity and differences between links across the SCOOT region. Then, the subsequent sections present the results of applying the methodology of Chaos Theory to the study network. Section 8.3 presents and discusses the computation and analysis of the lag times based on the autocorrelation coefficient method.

Section 8.4 makes use of the lag times to construct the phase space of the traffic flow on the links in the network for calculating the embedding and correlation dimension. Section 8.5 presents the results of the calculated Lyapunov exponents from the reconstructed phase spaces, and interprets the exponents in relation to identifying the temporal traffic flow condition of the network links. Section 8.6 further analyses the exponents, for every cohort, with respect to the frequency of the links indicating chaos across the network links at every sampled interval. Finally, Section 8.7 summarises the results of this chapter and presents the main conclusions.

8.2 Appreciation of SCOOT Flow and Occupancy Data

A summary of the basic statistics of the analysed raw SCOOT (flow/occupancy) data[4] for each link for every cohort is included in Appendix C (see Tables 1C-7C). This includes the median, mean and standard deviations. A normality test indicated that generally p-values are consistently less than 0.05. Therefore, at a 5% significance level, there was insufficient evidence to suggest that the observation in the datasets (deviation from the normal curve) was likely to be due to chance. It is concluded that the datasets are not normally distributed, which is confirmed by the probability plots in Figure 1C through to Figure 40C (in Appendix C). Further assessments indicated that the distribution of the descriptive statistics for mean, median, variance and range of the network links were also not normal. Given the distribution of the data, the Kruskal-Wallis, Mann-Whitney and Mood's Median tests were the main methods applied to identify the statistical significance level for the assessments presented in the remainder of the thesis.

A comparison of the groups of the flow medians month-by-month for each link pair gave a non-significant p-value of 0.255, which suggests that the median values are not statistically significantly different across the seven cohorts at the 95% confidence interval. Similarly, the distribution of the link occupancies was not normal with p-values consistently less than 0.005. The test of the monthly groups of the medians of the link occupancies indicated a p-value of 0.187, which tends to suggest that the median values of the link occupancies are not significantly different from month to month.

The inferences from the flow and occupancy data (Data 4 and 5) indicate similar characteristics in the two datasets, likely due to the influence of the underlying proportional relationship between the two variables at low values. These results suggest that the median flows or occupancies behave in a similar way across the links. Figure 8.1 present a box plot of the median of the medians for the cohorts, for flow and occupancy, which indicate a statistically significant difference between the two variables with p-value of 0.0014 according to the Mann-Whitney test.

[4] The raw data (flow and occupancy) are not included in the appendix due to its large volume, but available in electronic format, if required

The available datasets provide a simultaneous measurement of the link flow and occupancy variables throughout the observed period. The next step was to use one of these parameters initially to run the algorithms of Chaos Theory in order to identify the chaos-parameter (Lyapunov exponent) to be used to develop a basic understanding of the evolution of the system. Then, subsequently, the Lyapunov exponent profiles (derived from the first parameter (i.e. flow) were used to predict the other parameter (occupancy) in the next stage.

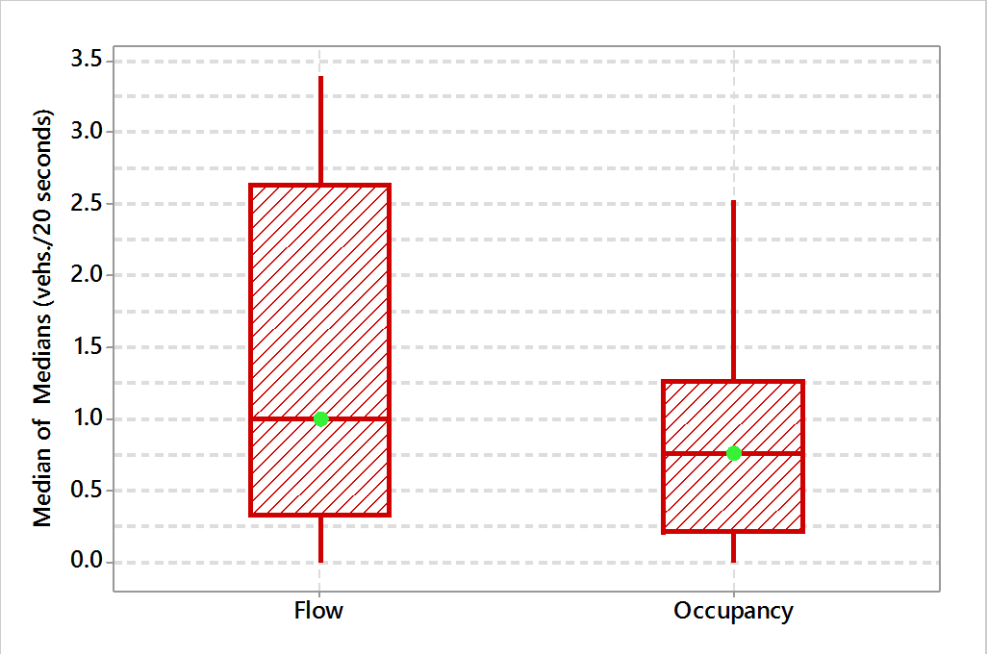


Figure 8.1: Box Plot of the Average Median of Flow and Occupancy

Since the existing literature suggested that occupancy is a better indicator of congestion (see section 2.4.4), the decision was made to predict ultimately the occupancy variable. The *rationale* is that flow does not necessarily determine traffic flow regimes unlike occupancy, which not only indicates when flow is low or high, but also whether it is free flowing, unstable or congested.

It must be noted however that not all detector systems measure both variables (e.g. MIDAS measures only flow but not occupancy). Therefore, the remainder of this chapter mainly presents the analysis in relation to the traffic flow data, whilst Chapter 9 relates to the predictions of the link occupancy based on the Lyapunov exponents (derived from flow).

8.3 Delay Time (Autocorrelation Lag)

In this section, the optimum lags (delay times) are determined from the time series profiles of the individual links. The optimum lags are then modelled with the metrics of the link flow data to determine a polynomial function for the forecasting of the lags.

8.3.1 Autocorrelations and Correlograms

The autocorrelation lags from the correlograms are included in Table 1D (see Appendix D), which also includes the profile of the lag times of each link, computed independently for each of the seven cohorts (from January through to July inclusive) of traffic flow data sampled at 20 second intervals. Figure 8.2 presents a scatterplot of the individual links lags categorised by month, as well as the monthly average, which indicates an average lag range between 140 and 270 minutes.

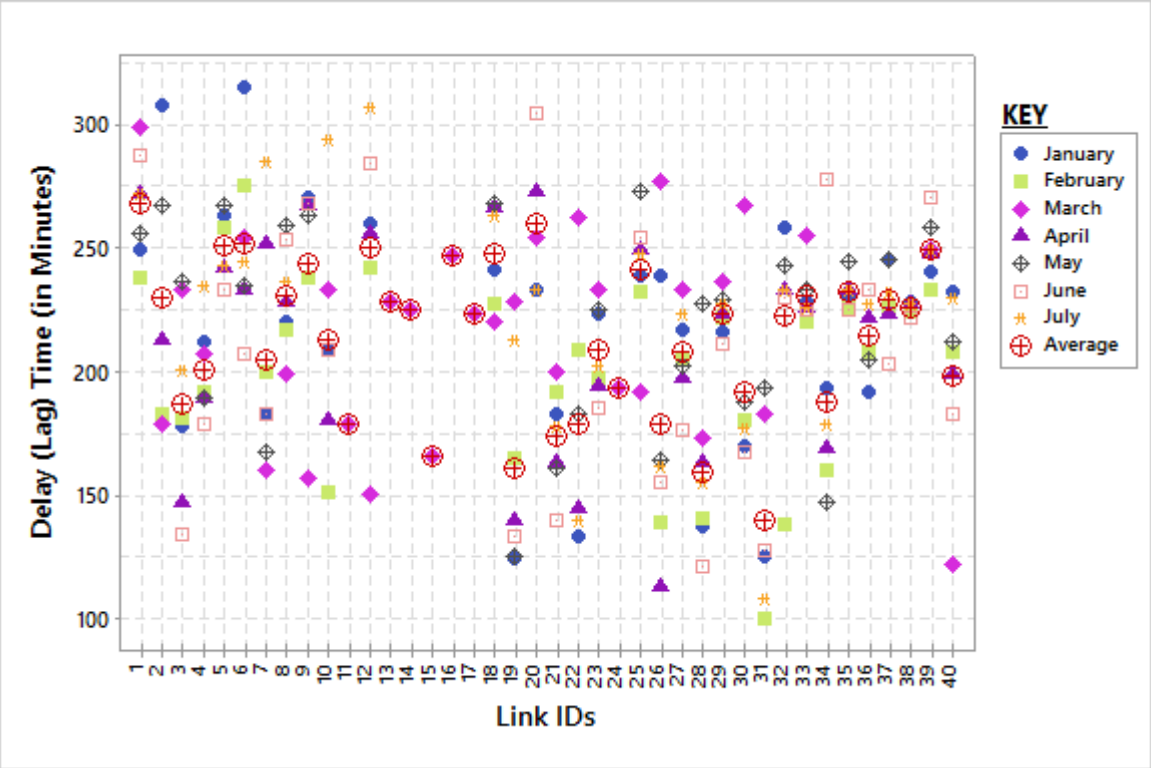


Figure 8.2: Scatterplot of the Lags of the Respective Links Categorised by Month

Further, Figure 8.3 presents a histogram of the distribution of the links. Generally, the lag times for the links were normally distributed except for links 9, 30, 32 and 33 (see Table 2D in Appendix D). It must be noted however that only the data for March (Cohort 3) was available for links 11, 13, 14, 15, 16 17 and 24 (mainly connected to SCOOT node 09191) because of data problems (such as large amounts of missing data).

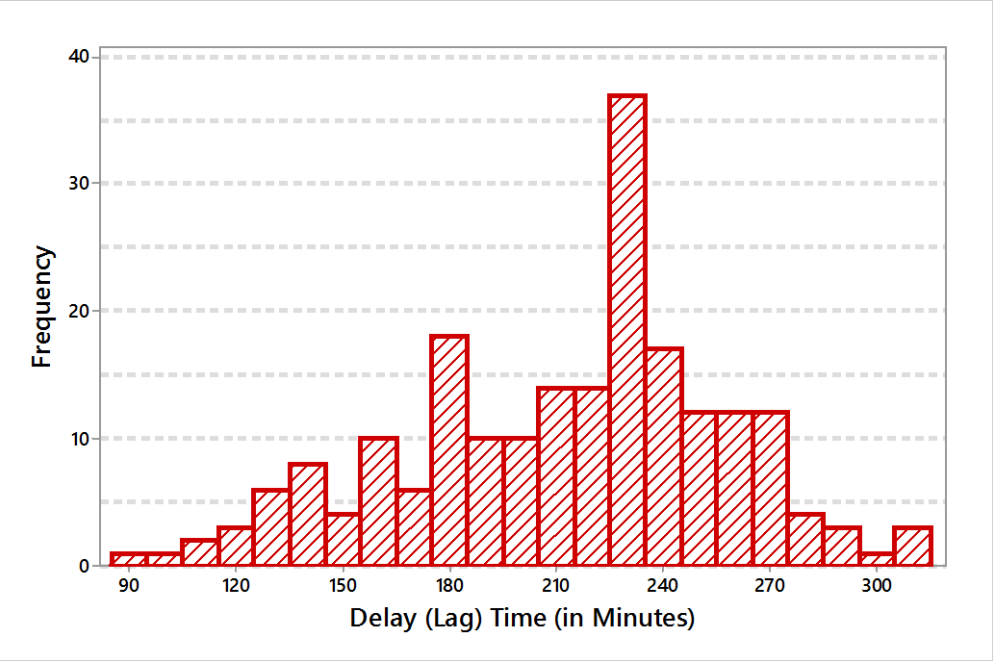


Figure 8.3: Distribution of Link Lags for All Cohorts of SCOOT Data 4[5]

The Kruskal-Wallis test showed that globally there was a statistical difference between the lag durations of the links in the network. Table 8-1 presents the results for the Kruskal-Wallis test based on the groups of adjacent links along with the statistical significance level of the lags. In the southbound direction, the test gave a p-value of 0.000 at a 5% significance level. The equivalent result in the northbound direction was 0.001. In order to identify clusters of links in which the lag times were shown to be statistically significantly similar in the network, groups of adjacent links were identified for analysis. In spite of the conclusions of independently statistically significantly differences for the northbound and southbound links, the analysis showed that there were groups of links where the lags were not statistically significantly different.

Table 8-1 shows that there are link groups (mainly of two to four links without any in-between intermediary links) where the difference between the lags are not statistically significant. These groups experience traffic platooning effects as there is minimal traffic flow variability across the links within these groups, and therefore the lags are not statistically significantly different. This suggests that an understanding of the spatial

[5] Lags for links 9, 30, 32 and 33 were based on median lag for the cohorts

distribution of lags may enable the identification of regions of well-coordinated platooning traffic in the network.

Group	Node Number	Links Numbers	p-value[6]
1	Northbound links	Northbound links	0.001
2	Southbound links	Southbound links	0.000
3	09111/09131	1, 9 ,12	**0.570**
4	09111/09131/09191	1, 9, 12, 18, 19	0.005
5	09191	18, 19	0.007
6	09131	3, 7	**0.482**
7	09111/09131	3, 5, 6, 7	0.013
8	09111/09131	3, 5, 6	0.008
9	09112/09131	4, 5, 6	0.007
10	09231/09283	31, 39	0.003
11	09191/09231/09283	18, 31, 39	0.003
12	09283/09231	27, 40	**0.848**
13	09221/09283/09231	21, 27, 40	**0.045**
14	09221/09283/09231	22, 27, 40	**0.488**
15	09221/09283/09231	21, 22, 27, 40	**0.132**
16	09231/09241/09253	29, 33, 38	**0.645**
17	09231/09241/09253	32, 35, 37	**0.843**
18	09221/09241/09253	34, 33, 38	**0.076**
19	09221/09241	34, 36	**0.078**
20	09253/09241	23, 37	**0.054**
21	09221/09241	23, 26	**0.180**
22	09221/09241/09253	26, 23, 37	**0.092**

Table 8-1: Output of the Significance Testing of Lags of the Groups of Adjacent Links (p-values)

On the other hand, mainly the groups that incorporated links connected to node 09191 (signal at Chapel Lane/Knighton Lane East) showed a statistically significant difference in lags. Due to the location of the node, at an approximate midpoint of the network and the distance between consecutive junctions, it is likely that the platoons may have become dispersed by the time the traffic arrives at this location (in both the northbound and southbound directions). Figure 8.4 shows a number of examples of the areas of the network where the results suggest the likelihood of traffic platooning through groups of links.

[6] The p-values in bold relate to adjacent links where traffic platooning occurs in the network

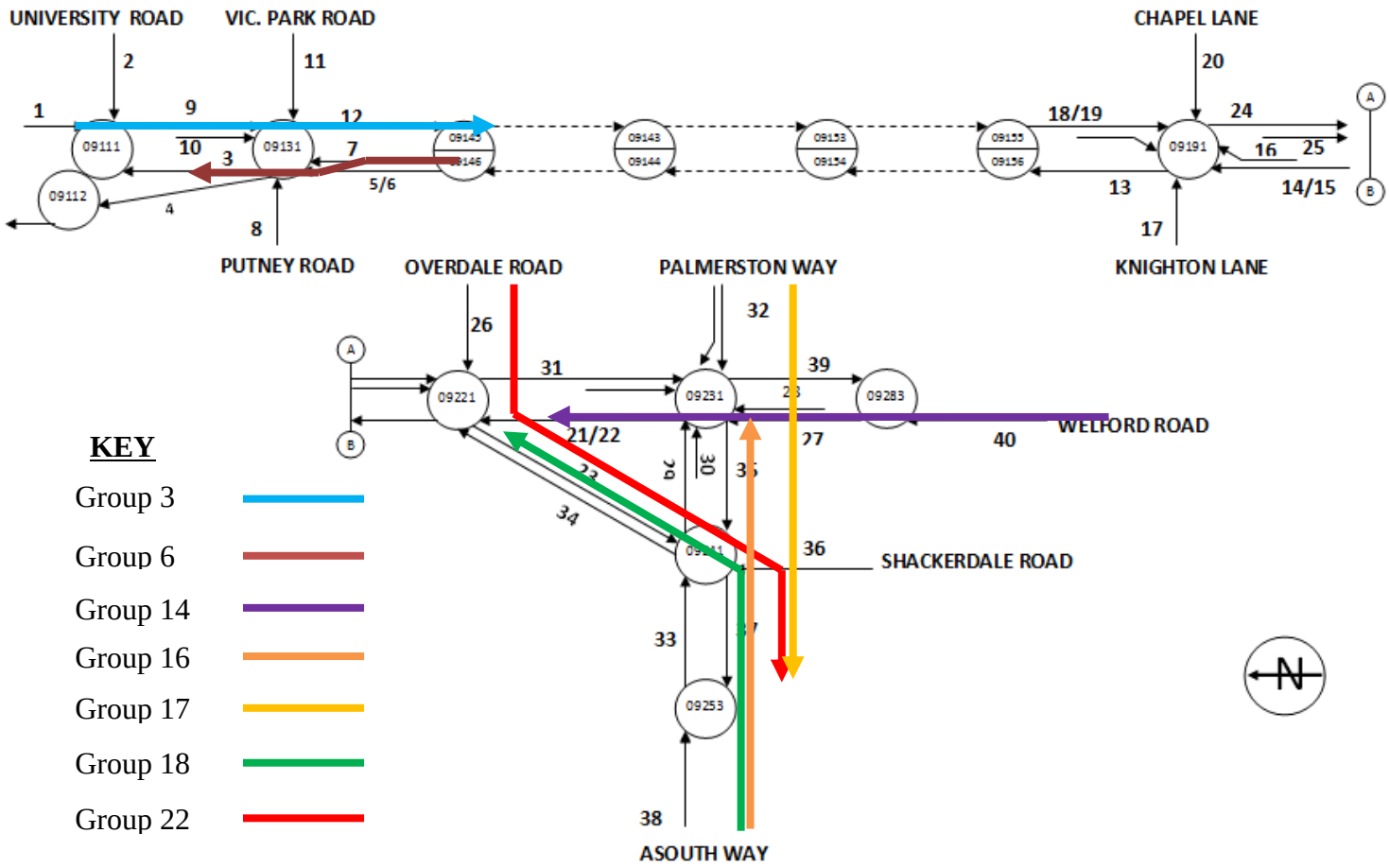


Figure 8.4: Examples of Groups of Links (Regions) of Platooning Traffic in the SCOOT Network

8.3.2 Modelling of the Autocorrelation Lags and the Metrics of the Data

In order to determine the relationship between the lags and the statistics (mean, median and standard deviation) of the data, a regression model was fitted to the scatter plot of the lags and the statistics of the datasets. Table 8-2 presents a summary of the coefficient of determination (R^2) for the non-linear model for the cohorts.

	R^2 Coefficients (%) for the model of Lag versus:		
Cohorts	Mean	Median	Standard Deviation of the mean
Individual Cohorts			
1: January	0.17	0.16	0.17
2: February	0.51	0.35	0.52
3: March	**0.02**	**0.03**	**0.03**
4: April	0.41	0.34	0.40
5: May	0.42	0.31	0.41
6: June	0.51	0.49	0.50
7: July	0.28	0.23	0.29
Average	0.33	0.27	0.33
Combined Cohorts			
All Cohorts	0.26	0.22	0.26

Table 8-2: Summary of R^2 Coefficient (between Lags and Statistics (Data 4))

The results for March are particularly low, compared with the other months. However, the average values of the independent cohorts compare well with the results of the model from the combination of all the cohorts data. According to the models, on average, the mean, median and standard deviation independently explain about 32.9%, 27.3% and 33.1% respectively. This result suggests that the strength of correlation of the lag time with mean, median and standard deviation is 57.3%, 52.2% and 57.5% respectively.

The model for the combined datasets for all the cohorts, in Figure 8.5, Figure 8.6 and Figure 8.7, show that the respective generalised models for the lag based on mean, median and standard deviation (s.d) from all the cohorts are included in Table 8-3.

Lags	Generalised Models
Lag_{mean}	$134.2 + 119.6 \times mean - 54.8 \times (mean)^2 + 8.7 \times (mean)^3$
Lag_{median}	$153.5 + 84.6 \times median - 32 \times (median)^2 + 4 \times (median)^3$
$Lag_{s.d}$	$124.9 + 166.8 \times s.d - 105 \times (s.d)^2 - 24.5 \times (s.d)^3$

Table 8-3: Generalised Models for the Lags Based on the Descriptive Statistics (Data 4)

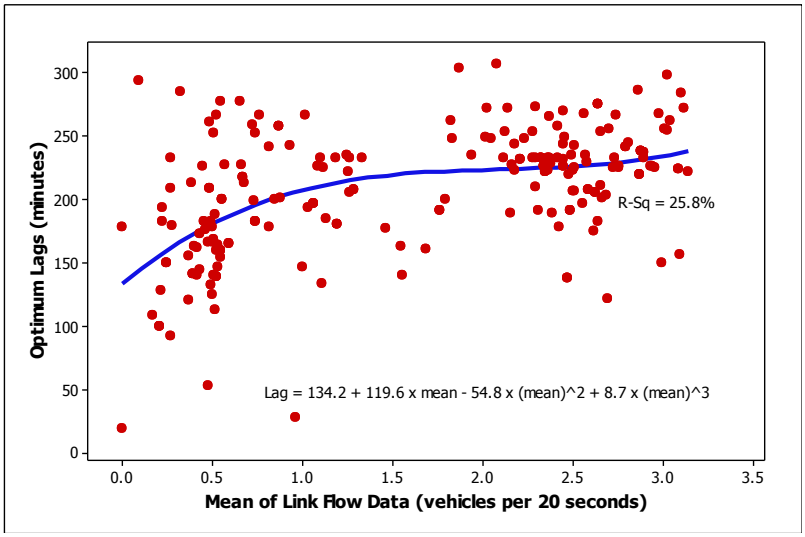


Figure 8.5: Graph of the Link Optimum Lags versus Mean (Data 4)

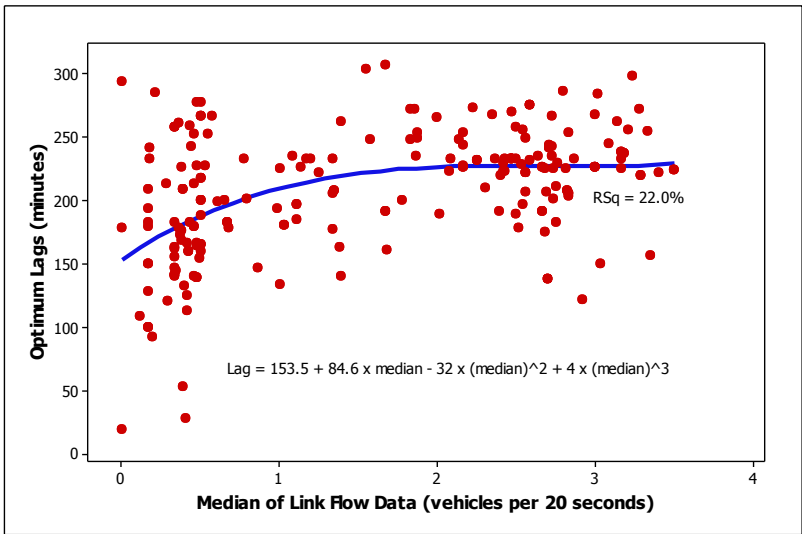


Figure 8.6: Graph of the Link Optimum Lags versus Median (Data 4)

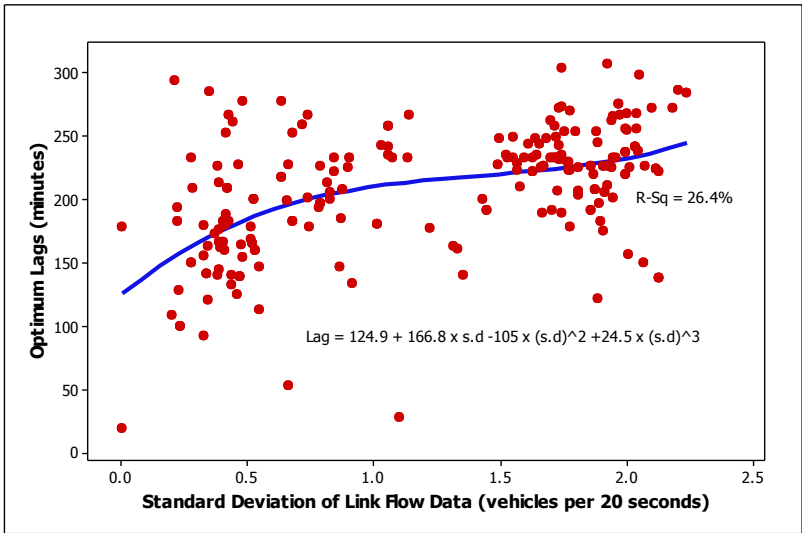


Figure 8.7: Graph of the Link Optimum Lags versus Standard Deviation (Data 4)

In order to test the model, the lags were predicted by using the statistics for each cohort as inputs for the above equations. The R and p values for the assessments are included in Appendix D (Tables 3D-9D), whilst Table 8-4 presents a summary of the correlation coefficients between the observed and predicted values (where significant p-values are in bold).

	Mean		Median		Standard Deviation	
Cohorts	R-value	P-value	R-value	P-value	R-value	P-value
1: January	0.694	0.000	0.593	0.000	0.700	0.000
2: February	0.778	0.000	0.677	0.000	0.783	0.000
3: March	0.734	**0.067**	0.178	**0.364**	0.022	**0.913**
4: April	0.643	0.000	0.560	0.001	0.597	0.000
5: May	0.629	0.000	0.548	0.002	0.630	0.000
6: June	0.513	0.004	0.460	0.011	0.524	0.003
7: July	0.414	0.023	0.375	0.041	0.493	0.006
Cohorts 1-7	-0.217	0.001	0.236	0.000	0.509	0.000

Table 8-4: Summary of R Coefficient (between Lags and Metrics (Data 4))

With the exception of Cohort 3 (March), the results indicate statistically significant relationships. The observation in Cohort 3 is expected because the fitted model for Cohort 3 gave extremely low R^2 values of approximately 3% on average (see Table 8-2). Even though the R^2 values for the fitted model for Cohort 1 (January) are comparatively low as well (of approximately 17% on average (see Table 8-2)), the model of the combined data from all cohorts indicates a statistically significant relationship between the observed and predicted values for this cohort. This is perhaps due to the advantage of the large sample of data points (cohorts 1-7 combined), which has addressed the weakness of the models caused by the small sample of data captured in some months. The case of Cohort 3 (March) to some extent is an anomaly. As alluded to previously in Section 8.3.1, Cohort 3 was based on fewer links because the lags were derived only for links 11, 13, 14, 15, 16 and 17.

Further, the models were evaluated using the Mean Square Error (MSE) criterion (and confirmed by the Mean Absolute Error (MAE)) to identify the best model for predicting the lag of a time series based on the statistics derived from the time series. The MSE is the average of the square of the errors in prediction, which is expressed mathematically as:

$$MSE = \frac{\sum {e_i}^2}{n}$$

Where:

- n = number of data points which are predicted
- e_i = error in the forecast, which is the magnitude of the difference between the predicted and expected values.

The Mean Absolute Error (MAE) criterion was used in a second evaluation to validate the results of the MSE approach. The MAE is the average of the absolute values of the differences between the predicted and observed values. The expression for the MAE is:

$$MAE = \frac{\sum |e_i|}{n}$$

As a standard convention in the analysis of forecasting models, the model with the lowest MAE or MSE is ranked as the highest performing model. Due to primarily evaluating the average total errors between the observed and predicted values, the model with minimal errors generates the lowest vales in MAE or MSE.

Table 8-5 presents the results for the comparative evaluation for the forecast models. That is, both approaches show that the model of the lag predicted from the standard deviation emerged as the best model, outperforming the models based on the mean and median. As evidenced in Table 8-5, the two methods both ranked the standard deviation as the best (in bold text), mean as second and median as third best.

Lags predicted from:	**MAE**	**MSE**
Mean	44	44,525
Median	57	96,219
Standard Deviation	**33**	**1,911**
Conclusion	Model of lag and standard deviation is more appropriate for predictions.	

Table 8-5: Model Evaluation Results using MAE and MSE Criteria (Data 4)

The outcomes suggest that the lag times depend more on the variation within the data rather than the mean or median flow level. As the lag time indicates the time interval a particular dataset correlates within its own past and future values, it was expected *a priori* that the lag

would be influenced by the variation within the data. Therefore, this result offers a tentative endorsement of this assertion.

In summary, the models indicate relationships between the lags and the statistics of the traffic flow data. However, the general trends of the models suggest that the lag time increases as the value of the statistic (mean, median and standard deviation) increases. Both models of the lag against mean and standard deviation follow a positive gradient and therefore the lag increases with increasing values of these two statistics. In the lag/median model however, the lag saturates at about 237 minutes from a median value of about 2.0 upwards. The results imply that any changes in the median values beyond about 2.0 do not affect the lag in the model. This is essentially a limitation of this model that makes the model weak in explaining errors beyond this threshold value. This ultimately affects MAE/MSE resulting in large errors and making this model less suitable.

8.4 Correlation and Embedding Dimensions

As the system's dynamics are initially unknown, a number of embedding dimensions, ranging between 1 and 10 were tested for the network links in order to determine the saturated embedding dimensions. The assessments showed that the embedding dimensions ranged between 2 and 5, and the range of correlation dimension is approximately 0.22-0.35. This confirms earlier results in the preliminary assessments that showed that the embedding dimension of flow, and occupancy, is equal to 3 (see section 7.3.2).

Figure 7.2 showed a typical G-M log-log curve and the relationship between the embedding dimension and the correlation dimension for Link 2, which indicate that the embedding and correlation dimensions are equal to 2 and 0.347 respectively. The evaluation of embedding and correlation dimension is an indication of the existence of a chaotic character in the time series data. According to Takens (1981), the choice of parameters must satisfy the relationship given by $m \geq D_2 + 1$, where m and D_2 are the embedding and correlation dimension respectively. The theorem ensures that the embedding dimension is sufficiently large so that the orbits of the attractors do not intersect in the phase space, and it is often used to test the suitability of the choice of parameters (see section 5.6.3). The calculated values of m and D_2 were at all times consistent with the prescribed criteria.

8.5 Model Outputs of Lyapunov Exponents

This section presents the results of the model output for the Lyapunov exponents. It describes the statistical inference and interprets the Lyapunov exponents in the context of dynamical systems. Finally, it relates the exponents to the individual links, and discusses their implication for traffic management based on the understanding of the link behaviours.

8.5.1 Description of the Results

The respective lag and embedding dimension for every link, derived from the time series of the traffic flow information, were used to construct the unique phase spaces independently for each link in order to calculate its instantaneous Lyapunov exponents[7]. The assessment gave 20 second Lyapunov exponent profile of the links in the network from the 1st January to 31st July 1999, separated into monthly cohorts (1 through to 7). Table 8-6 presents a summary of the minimum and maximum Lyapunov exponents for each cohort. Overall, the exponents range between -2.34 x 10^{-2} and 1.12 x 10^{-2}.

Cohort	Minimum	Maximum
1: January	-2.34 E-02	1.09 E-02
2: February	-2.34 E-02	1.09 E-02
3: March	-2.24 E-02	8.23 E-03
4: April	-2.12 E-02	8.90 E-03
5: May	-2.31 E-02	1.01 E-02
6: June	-2.32 E-02	1.10 E-02
7: July	-2.32 E-02	1.12 E-02

Table 8-6: Summary of Lyapunov Exponents for Seven Monthly Cohorts

The normality test showed that the profiles of the Lyapunov exponents were not normally distributed. The link median values were averaged over the seven cohorts as shown in Figure 8.8 for visualisation of the differences that exist in the Lyapunov exponents from link to link.

The Kruskal-Wallis test gave a p-value of 0.001, when the average median profiles were compared, which indicates statistical differences in profiles. Further, this result was confirmed by a comparison of the monthly groups of the median of the Lyapunov exponent

[7] The Lyapunov exponent profile not included in the appendix due to its large volume, but available in electronic format, if required

profiles, which gave a significant p-value of 0.000. The statistical differences highlight the uniqueness in the evolution of the traffic system from one location to the other.

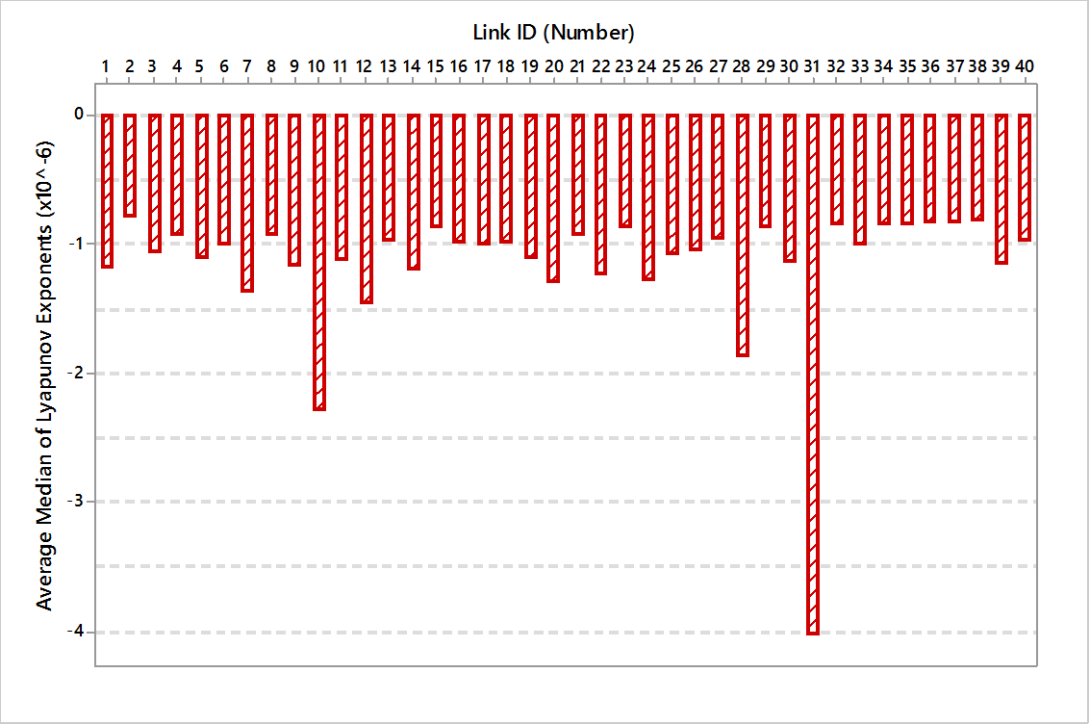


Figure 8.8: Average Median Profiles of Lyapunov Exponents (based on Data 4) for Seven Monthly Cohorts

The Lyapunov exponents were examined further across the network in order to gain a better understanding of the dynamics of the traffic system. In order to achieve this, the median Lyapunov exponents across all links in the network were extracted at every 20 seconds interval on 5th July 1999. This provided a total of 4320 median points for every day in each of the seven cohorts.

Figure 8.9 is an example histogram, which is not-normal and bi-modal, typical of the daily median Lyapunov exponents, whilst Figure 8.10 shows in detail the distribution for link 29.

The next step of the analysis was, for each cohort, to examine the median Lyapunov exponents of all the days with respect to the degree of correlation with due consideration of the p-values. The results of the assessments are included in Tables 3D-9D (in Appendix D), whilst Table 8-7 presents a breakdown for each month of the number of correlations in each individual category.

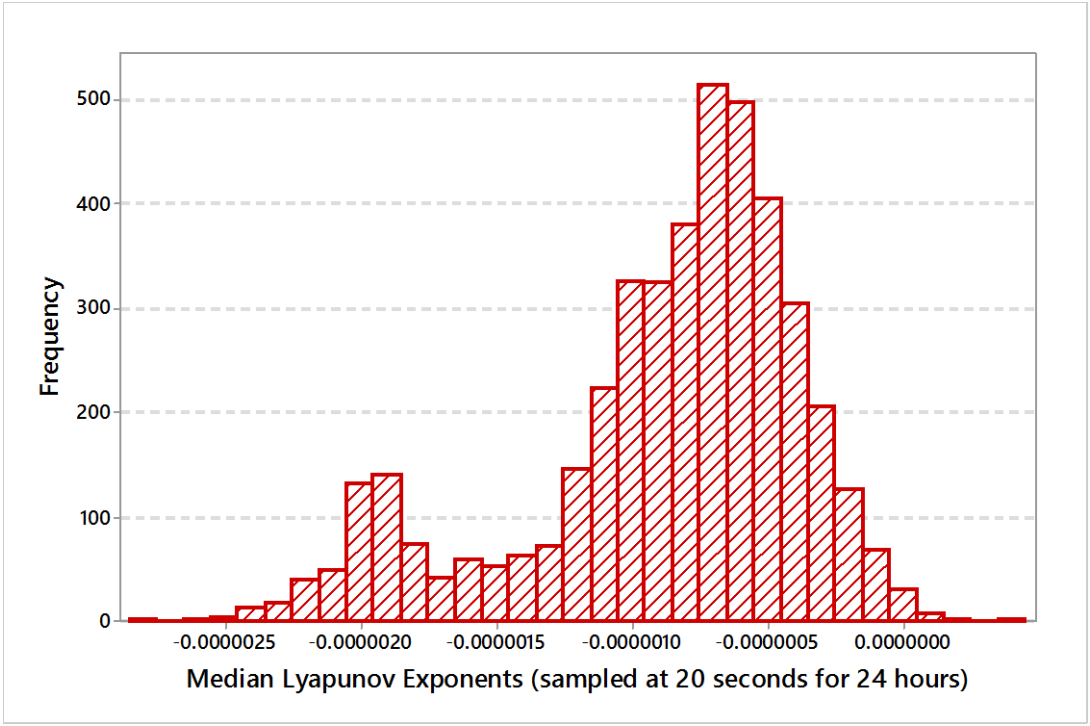


Figure 8.9: Typical Histogram of the Daily Median Lyapunov Exponents across all Links

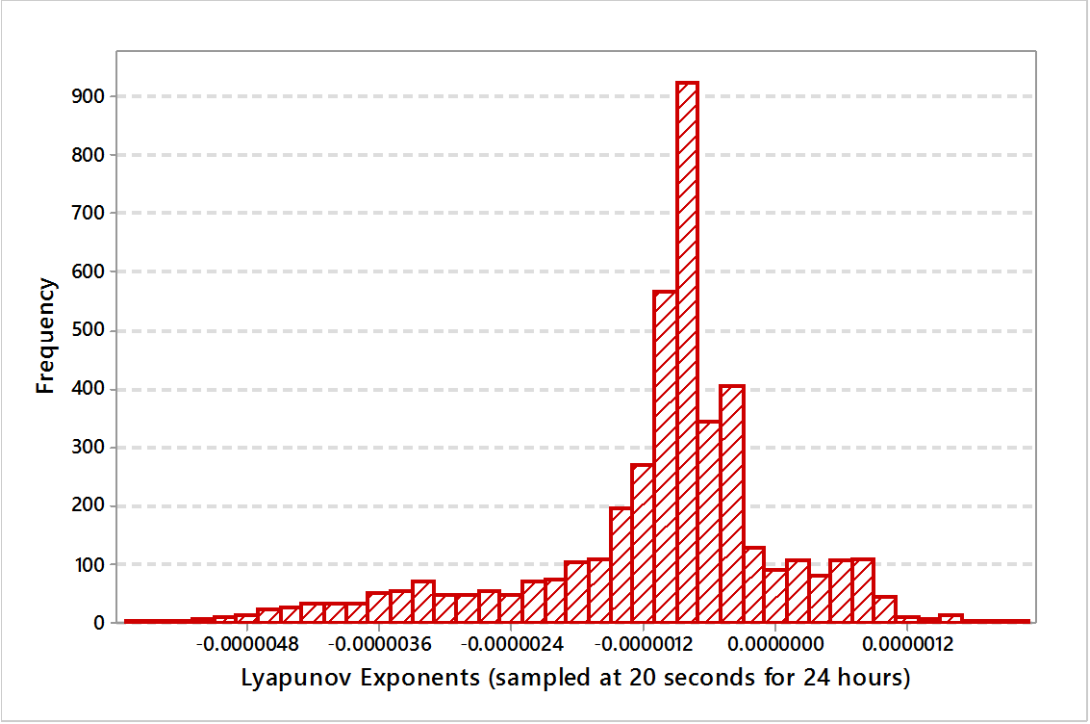


Figure 8.10: Typical Histogram of Lyapunov Exponents for Link 29

The number of days assessed for each month are 28, 27, 29, 28, 29, 29 and 29 respectively for cohorts 1 through to 7. Given that the daily Lyapunov exponents were correlated with all other days gave a matrix table of R-values of 28 x 28, 27 x 27, 29 x 29, 28 x 28, 29 x

29, 29 x 29 and 29 x 29 respectively. Therefore, the strength of the association of the exponent values of two links was measured by counting the number of links falling between certain ranges of values.

Degree of Correlation	Cohorts							
	1	2	3	4	5	6	7	Total
Very Strong ($\|R\| \geq 0.70$)	253	270	194	174	253	222	242	1608
Strong ($0.6 < \|R\| \leq 0.69$)	64	56	111	101	80	129	91	632
Moderate Strong ($0.4 < \|R\| \leq 0.59$)	12	3	25	23	27	10	28	128
Weak ($0.2 < \|R\| \leq 0.39$)	9	12	17	24	22	32	13	129
No Correlation ($0 < \|R\| \leq 0.19$)	40	10	59	84	24	13	32	262
Total ($0 < \|R\| \leq 1$)	378	351	406	406	406	406	406	2759

Table 8-7: Summary of Correlation Assessment for Median Lyapunov Exponents

Discounting the main diagonal (to eliminate the correlation of a day with itself) and one-half of the matrix (to eliminate double counting because of reversal in the pairing for example Day 1-Day 2 and Day 2-Day 1 etc.) gave the respective totals for the cohorts in Table 8-7. The results showed that, in spite of the variability in the dynamics of the evolution of the traffic system from link to link (according to the Kruskal-Wallis test), the dynamics of the individual days do correlate reasonably well. This conclusion was informed by the high proportion of p-values less than 5% given that, approximately 96% (2648 divided by 2759) of the correlation exercises exhibited statistically significant relationships.

8.5.2 Interpreting the Lyapunov Exponent Profiles

In dynamic systems, the interpretations of instantaneous Lyapunov exponents are as follows. Positive, negative and zero exponents indicate unstable, stable and meta-stable states respectively, as explained in Section 0. The unstable states mainly occur in the daytime when congestion is more likely. Generally, the daily plots show a fluctuating profile peaking between 0600h-0900h and 1500h-1800h, and sometimes around midday. The identified peaks are indications of a change of state of the links, which are precursors to the links becoming congested and traffic behaviour transitions to chaos occur prior to congestion (Lyapunov exponents falling) following free flow (Lyapunov exponents increasing). Appendix F (Figure 1F through to Figure 10F) includes a number of daily

profiles for links 21, 23, 29, 33, 34, 35, 36, 38, 39 and 40 to illustrate the uniqueness of the dynamic behaviour for the individual links in the network.

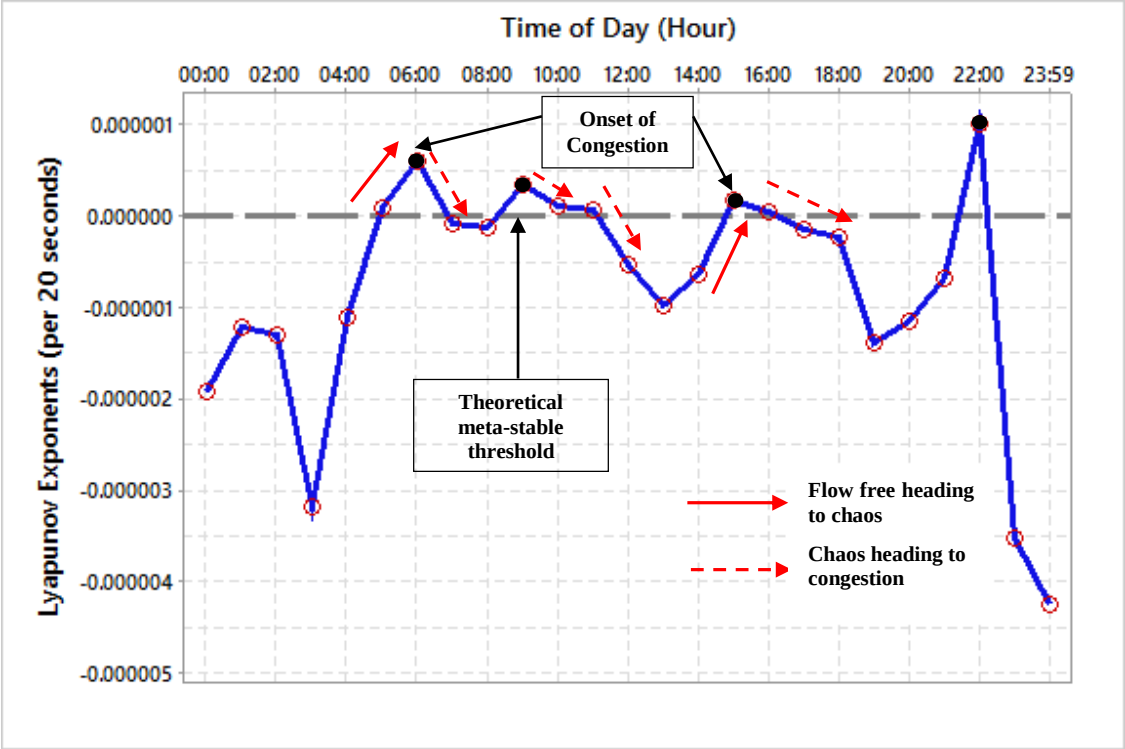


Figure 8.11: Typical 24-Hour Profile of Lyapunov Exponents (5th July 1999) for Link 29

Figure 8.11 shows an example of a 24-hour profile of instantaneous Lyapunov exponents computed for a link 29 on 5th July 1999. Ideally, the region below the Lyapunov exponent of zero (theoretical) reference line indicates that the network is extremely stable (i.e. when the link is free flowing due to low traffic flows or when in a congested state), whilst above the zero threshold indicates unstable states of the link. Therefore, the area below the meta-stable reference line determines the times when the network is less sensitive to perturbations in traffic flow, greater negativity and insensitivity. This is most likely to occur during the overnight period, which is not associated with the onset of prolonged congestion, but can occur also in the following two situations. The first is in free flow conditions when the network control system can respond to the randomness in the traffic arrivals at the stop line because there is plenty of spare capacity. The second is in the congested state when traffic is generally stationary and very slow moving, in which state the traffic is calmed. In both of these states, the network is considered to be stable and insensitive, as there are virtually no interactions between vehicles, and the Lyapunov exponent is negative, zero or a small positive value. The link starts to become sensitive as

the degree of interaction between vehicles increases with high positive Lyapunov exponents indicating extreme sensitivity.

The Lyapunov exponent is sensitive to states of rapid change. On short links, queues build up quickly and is reflected in the occupancy measured by SCOOT. However, well-coordinated traffic signal system ensures this link clears at each cycle therefore smaller positive values of the Lyapunov exponents are acceptable. This situation occurs frequently overnight when signal timings are not coordinated, but instead operate in the demand dependent mode and during low flow periods of the day on a dedicated right turn link with short reservoir length. For these reasons, there may be instances when the theoretical threshold value of zero is impractical (see Figure 2F in Appendix F). Strict adherence to the zero criterion to an unlikely case dominated by negative Lyapunov exponents suggest the link is hardly in a chaotic state during the daytime period, which tends to be associated with the onset of prolonged congestion. Therefore, the theoretical meta-stable datum may need adjusting to reflect the lesser likelihood of chaos over night. That is, the low values of exponents overnight were used as a benchmark to determine the *practical* meta-stable points where dynamic behaviour transitions between the stable and unstable regions. In future applications of chaos in real networks, it will be necessary to carry out further work to establish whether this threshold value for the Lyapunov exponent will be generic for a network or needed to be fine-tuned for each link. This process could be implemented as a component of an autonomic or a *feed forward* system in the future. This will be discussed further later in the thesis (in Section 10.2.6).

By observing the time series profile of the Lyapunov exponents, the state change can be detected at the turning points of the curve, which determines the onset of congestion as a signal of impending congestion. As shown in Figure 8.11, the onset of congestion is identified as the turning point from a positive to negative gradient, which is a state transition from free flow to chaos, whereas a negative to positive change reveals the instability associated with the onset of congestion or free flow. This means that a positive-negative transition does indicate that the link is operating at or close to the optimum degree of saturation (usually 90-95% saturation). Figure 8.11 identifies two instances of the onset of congestion in the morning around 0600h and 0900h but the link starts to become congested from about 0700 and 1100h respectively, whilst the evening peak is around 1500h (and

becomes congested from 1600h). However, the sensitivity or vulnerability of the system during the evening peak is not as pronounced as the morning peak, indicating that relatively it is easier to stabilise perturbations during the evening peaks.

At the network level, it is important to assess the chaos situation across all the individual links, which make up the network, in order to determine the total dynamic state across the network. Such an assessment is a key step in determining the traffic conditions in the network because the instantaneous state of any arbitrary link may not necessarily define the state of the entire network. Consequently, the raw Lyapunov exponents were analysed with respect to all other links in the network in order to determine the dynamical state of the entire SCOOT network.

It was found, as shown in Figure 8.8 and Figure 8.10, that negative exponents dominate the profiles, which is revealed also in Figure 8.9 by the median profiles. This was considered reasonable because it suggests that generally, traffic flow is stable, and the number of instances of chaos is relatively less compared with stable conditions as expected. In order to analyse chaos across the entire network, a method is proposed in this research based on analysing the frequency of positive Lyapunov exponents for all the links in the network. The outcome from this method, together with an understanding of the link Lyapunov exponents profile will help to determine the critical times that should trigger the onset of congestion. Section 8.6 seeks to demonstrate how this methodology is used to develop the Lyapunov exponent as an indicator of the onset of congestion.

8.6 Exploring the Lyapunov Exponents to Determine the Network Status

In order to understand the global state of the network, all daily profiles of the Lyapunov exponents for all the links need to be considered together. Therefore, the next step is to count the number of positive Lyapunov exponents at each interval across all the links in the network. In this way, the locations as well as how chaos occurs and propagates through the network can be investigated. The number of links indicating negative, zero or positive Lyapunov exponents were determined every 20 seconds for all the surveyed days. For example, the results at 0900 hours on the 21st July 1999 indicated that there are 21, 6 and 0 links indicating positive, negative and neutral values of Lyapunov exponents respectively. The profiles of the positive exponent were then used in the subsequent analysis to understand the patterns of the frequency of occurrence of chaos across the entire network

of interconnected links. A summary of the full results of the analyses, which are used in analysing the scenarios described in Section 6.7.2, are sampled at one hour intervals and included in Appendix G (see Table 1G through to Table 49G). These are described briefly below for each of the cohorts. The discussion for the cohorts here are to establish whether chaos is a microscale event or has spatial effects in a network of interconnected links based on the analysis of positive Lyapunov exponents.

8.6.1 Analysis of Results – Scenario A1

Section 8.6.1 presents the results for Scenario A1, with Scenario A2 and A3 in Section 8.6.2, and Scenario 4 in Section 8.6.3. An assessment of all scenarios are necessary in order to establish how to address issues of missing Lyapunov exponent data especially when using this data for short-term traffic forecast of link occupancy.

Table 8-8 presents the summary of the results for Scenario A1, based on the correlation coefficients of frequency (20 seconds) of positive exponents between link pairs on the same days (Scenario A1). Table 8-8 includes the most and least correlated profiles, based on the average correlation coefficients, for cohorts 1 through to 7, for each 'same day' analysed. Table 8-8 shows that the most and least correlated profiles for Cohort 1 are Friday (0.68) and Monday (0.36) respectively, and for Cohort 2 are Saturday (0.75) and Wednesday (0.38) respectively.

Further, the corresponding figures for Cohort 3 are Saturday (0.60) and Friday (0.34) respectively, and for Cohort 4 are Sunday (0.61) and Monday (0.22) respectively. Furthermore, the most and least correlated profiles for Cohort 5 are Friday (0.64) and Monday (0.12) respectively. Therefore, Cohort 5 repeats the observation in Cohort 1. Similar to the observation in Cohort 4, Table 8-8 shows the corresponding results for Cohort 6 are Sunday (0.71) and Monday (0.32) respectively. Finally, Table 8-8 shows that the highest and least correlated profiles for Cohort 7 are Friday (0.61) and Tuesday (0.21) respectively.

Cohorts	Day of Week	Minimum	Maximum	Average
1: January	*Monday*	*0.01*	*0.70*	*0.36*
	Tuesday	0.08	0.70	0.39
	Wednesday	0.29	0.57	0.44
	Thursday	0.48	0.73	0.60
	Friday	*0.37*	*0.91*	*0.68*
	Saturday	0.26	0.79	0.52
	Sunday	0.49	0.79	0.63
2: February	Monday	0.68	0.72	0.70
	Tuesday	0.15	0.70	0.42
	Wednesday	*0.04*	*0.69*	*0.38*
	Thursday	0.48	0.75	0.61
	Friday	0.47	0.75	0.61
	Saturday	*0.72*	*0.77*	*0.75*
	Sunday	0.70	0.73	0.72
3: March	Monday	0.41	0.68	0.56
	Tuesday	0.38	0.65	0.51
	Wednesday	0.52	0.61	0.57
	Thursday	0.16	0.67	0.39
	Friday	*0.18*	*0.71*	*0.34*
	Saturday	*0.44*	*0.76*	*0.60*
	Sunday	0.38	0.71	0.54
4: April	*Monday*	*0.60*	*0.47*	*0.22*
	Tuesday	0.39	0.62	0.53
	Wednesday	0.21	0.69	0.43
	Thursday	0.33	0.69	0.53
	Friday	0.46	0.75	0.59
	Saturday	-0.19	0.79	0.29
	Sunday	*0.55*	*0.72*	*0.61*
5: May	*Monday*	*-0.14*	*0.40*	*0.12*
	Tuesday	0.12	0.83	0.51
	Wednesday	0.59	0.63	0.61
	Thursday	0.61	0.65	0.63
	Friday	*0.54*	*0.72*	*0.64*
	Saturday	0.35	0.73	0.58
	Sunday	-0.16	0.72	0.18
6: June	*Monday*	*0.12*	*0.62*	*0.32*
	Tuesday	0.13	0.70	0.46
	Wednesday	0.18	0.65	0.41
	Thursday	0.63	0.70	0.66
	Friday	0.63	0.71	0.68
	Saturday	0.44	0.74	0.55
	Sunday	*0.70*	*0.72*	*0.71*
7: July	Monday	0.40	0.65	0.56
	Tuesday	*-0.18*	*0.58*	*0.21*
	Wednesday	0.28	0.58	0.44
	Thursday	0.22	0.56	0.38
	Friday	*0.57*	*0.64*	*0.61*
	Saturday	0.46	0.60	0.53
	Sunday	0.60	0.61	0.60

Table 8-8: Summary of Correlation Coefficients (R) for Scenario A1 (All Cohorts)

Table 8-8 further indicate the minimum range of the correlation coefficients are 0.01 and 0.49 (Cohort 1), 0.04 and 0.72 (Cohort 2), 0.16 and 0.52 (Cohort 3), -0.19 and 0.60 (Cohort 4), -0.10 and 0.61 (Cohort 5), 0.12 and 0.70 (Cohort 6) and -0.18 and 0.60 (Cohort 7). The corresponding ranges of the maximum correlations are 0.57 and 0.91 (Cohort 1), 0.69 and 0.77 (Cohort 2), 0.61 and 0.76 (Cohort 3), 0.47 and 0.79 (Cohort 4), 0.40 and 0.73 (Cohort 5), 0.62 and 0.74 (Cohort 6) and 0.56 and 0.65 (Cohort 7).

The results indicate high variations across the minimum correlations compared with the maximum correlations. The results further indicate that the weekdays had the maximum correlations in all the cohorts, whilst the weekends dominated the minimum correlations. Table 8-9 gathers the highest and lowest of the average correlation coefficients, and the corresponding ranges of maximums.

Cohorts	Highest	Lowest	Minimum Range
1: January	0.68	0.36	0.34
2: February	0.75	0.38	0.08
3: March	0.60	0.34	0.15
4: April	0.61	0.22	0.32
5: May	0.64	0.12	0.43
6: June	0.71	0.32	0.12
7: July	0.61	0.21	0.06

Table 8-9: Overall Summary of Correlation Coefficient (R) for All Cohorts (Scenario A1)

Clearly, the summary statistics illustrate that patterns in daily Lyapunov exponents across the network were prevalent, which is neither limited to any particular month nor weekday but occurs on different days including a Saturday and on a Sunday depending on the month. This result confirms the need to continuously monitor the evolution of congestion in real-time so that when the thresholds are reached in particular link pairs the onset of congestion is sent as an alert to an operator who can respond appropriately.

8.6.2 Analysis of Results – Scenarios A2 and A3

This section describes the results of the analysis in hierarchical order from Cohorts 1 through to Cohort 7 for scenarios A2 and A3 (see Section 6.7.2), which are as follows:

1. Scenario A2: each day profile versus the respective average same-day profile.
2. Scenario A3: each day profile versus the respective average weekly profile.

		Scenario A2		Scenario A3	
Cohorts	Day of Month	Correlation Coefficients			
		Highest	Lowest	Highest	Lowest
1: January	3rd Wednesday			0.89	
	16th Saturday	0.93			
	18th Monday		0.26		0.04
	Average	0.79		0.68	
2: February	3rd Wednesday		0.11		0.13
	6th Saturday	0.91			
	11th Thursday			0.89	
	Average	0.82		0.78	
3: March	11th Thursday		0.45		
	19th Friday				0.46
	27th Saturday	0.91			
	29th Monday			0.92	
	Average	0.78		0.76	
4: April	7th Wednesday			0.90	
	11th Sunday	0.94			
	24th Saturday		-0.13		-0.16
	Average	0.74		0.70	
5: May	3rd Monday		0.02		-0.09
	7th Friday	0.90		0.86	
	Average	0.74		0.71	
6: June	8th Tuesday				0.32
	13th Sunday	0.90			
	28th Monday		0.33		
	29th Tuesday			0.97	
	Average	0.80		0.77	
7: July	26th Monday		0.04		-0.07
	27th Tuesday	0.88		0.85	
	Average	0.77		0.71	

Table 8-10: Summary of Correlation Coefficients (R) for Scenarios A2 and A3 (All Cohorts)

Figure 1H through to Figure 6H (in Appendix H) present the output of the bar chart analysis of the cohorts, whilst the key features of the graphical profiles are summarised in Table 8-10 that shows the highest and lowest correlation of the daily profiles with the average same-day and weekly profiles together with the corresponding day. The results of the cohorts consistently indicate a strong relationship between the daily and average same day profile, suggesting that the average same-day profiles may be useful surrogate data in the model for the forecast of occupancy. This will be explored further in Chapter 9.

Figure 8.12 presents the profile of Cohort 1, which also illustrates the characteristics of the profiles for Scenario A2 compared with Scenario A3. Figure 8.12 indicates that Day 16

(Saturday) has the highest correlation of 0.93 for Scenario A2, whilst Day 18 (Monday) has the lowest correlation of 0.26. Similarly, Day 15 (Friday) has the highest correlation of 0.89, whilst Day 18 (Monday) has the lowest correlation of 0.04 for Scenario A3. Figure 8.12 shows that generally, the relationships in Scenario A2 are stronger than Scenario A3, and the results indicate that the average correlation coefficients for scenarios A2 and A3 are 0.79 and 0.68 respectively.

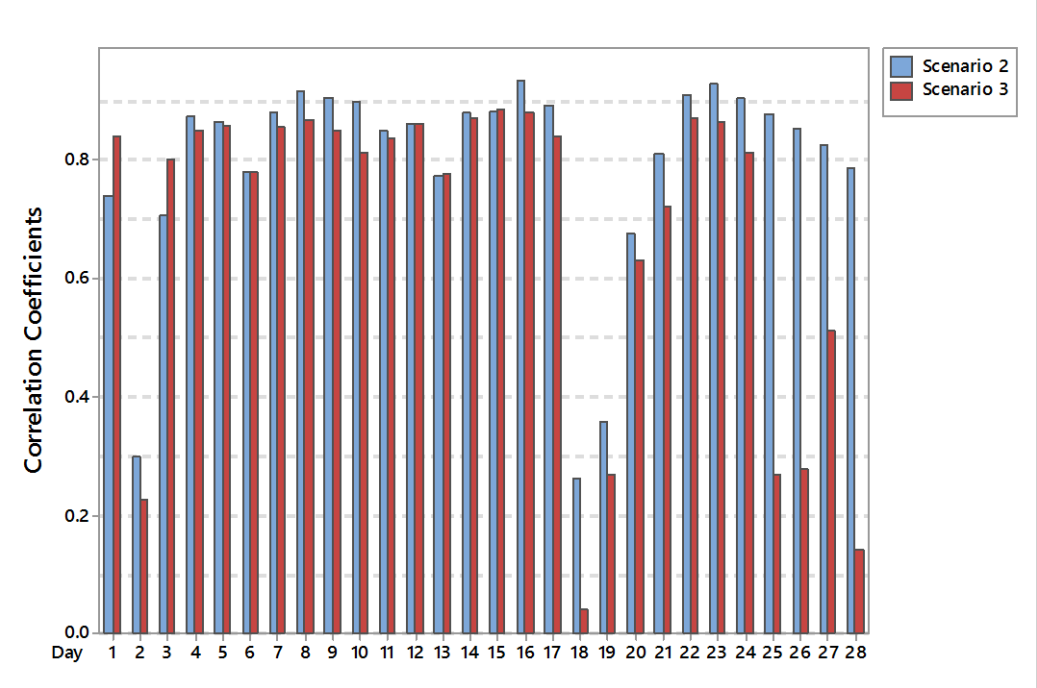


Figure 8.12: Profiles for Scenario A2 compared with Scenario A3 (January 1999)

The inferences that can be made for the other cohorts in terms of the range of correlation coefficients are presented in Table 8-11.

Cohorts	Scenario A2		Scenario A3	
	Minimum	Maximum	Minimum	Maximum
2: February	0.11	0.91	0.13	0.83
3: March	0.45	0.91	0.46	0.92
4: April	-0.13	0.94	-0.16	0.90
5: May	0.02	0.90	-0.09	0.86
6: June	0.33	0.90	0.31	0.97
7: July	0.04	0.88	-0.07	0.85

Table 8-11: Overall Summary of the Range of Correlation Coefficients (R) for Scenarios A2 and A3 (All Cohorts)

The Mann-Whitney test gave a significant p-value of 0.0098 to confirm that the correlations of the daily profiles with the average same-day correlations were statistically significantly higher than the average weekly in Cohort 1. The results therefore confirm, as might be expected, that the average same day of the week profiles tend to correlate with the respective days. Cohorts 2, 5 and 7 also gave the same conclusions since the Mann-Whitney test outputs reveal statistically significant values of 0.0019, 0.01 and 0.01 respectively.

Cohorts	Scenario A2			Scenario A3		
	Highest	Lowest	Minimum Range	Highest	Lowest	Minimum Range
1: January	0.93	0.26	0.67	0.89	0.04	0.84
2: February	0.91	0.11	0.80	0.89	0.13	0.76
3: March	0.91	0.45	0.46	0.92	0.46	0.46
4: April	0.94	-0.13	1.07	0.90	-0.16	1.06
5: May	0.90	0.02	0.88	0.86	-0.09	0.95
6: June	0.90	0.33	0.57	0.97	0.31	0.66
7: July	0.88	0.04	0.84	0.85	-0.07	0.92

Table 8-12: Overall Summary of Correlation Coefficient (R) for All Cohorts (Scenarios A2 and A3)

On the other hand, the statistics for Cohorts 3, 4 and 6 indicated that there was no statistically significant difference between the correlations of the daily profiles with the average same-day and weekly profiles. According to the Mann-Whitney test, the p-values are 0.23, 0.27 and 0.07 respectively, all greater than the critical value of 0.05 for 95% confidence interval.

Overall, the results show that the daily profiles are either statistically similarly correlated with the average same-day and weekly profile or correlated more strongly with the average same-day profiles than the average weekly profiles. It is important to note there is no instance that the results indicated the average weekly profiles correlated more strongly than the average same-day profiles. Consequently, the average same-day profiles were used for further analysis.

8.6.3 Analysis of Results – Scenario A4

This section presents the analysis of the same-day of the week profiles for the seven cohorts, which falls under Scenario A4 (see Section 6.7.2). The graphical profiles for the seven

cohorts are included in Appendix I (see Figure 1I through to Figure 7I), which consistently indicate at least about 10% of links in chaos during the daytime period (0600h–1600h) when congestion may be expected. Figure 8.13 shows the profile for Cohort 7, which is included here to help to explain how the proportions of links that experience chaos vary throughout the day. It is must be noted that unlike the commonly used flow profile Figure 8.13 is a graph illustrating the onset of, rather than, the presence of congestion.

The curves in Figure 8.13 start with an upward slope for the frequency of occurrence of congestion onset building up from midnight and peak between 0500h and 0700h during the weekdays (but slightly delayed peaks over the weekend). It shows two distinct peaks that make them unique. The weekday morning and evening onset of congestion occur around 0600h and 1500h respectively. The highest average frequency is 30% that occurs in the morning and evening onset of peaks on Friday and Monday respectively. The weekend peaks are around 0800h with a frequency of *circa* 29%, whilst the not too pronounced evening peak occurs around 1700h on Saturday only with a frequency of *circa* 21%. A further, less frequently, occurring peak occurs about 2030h.

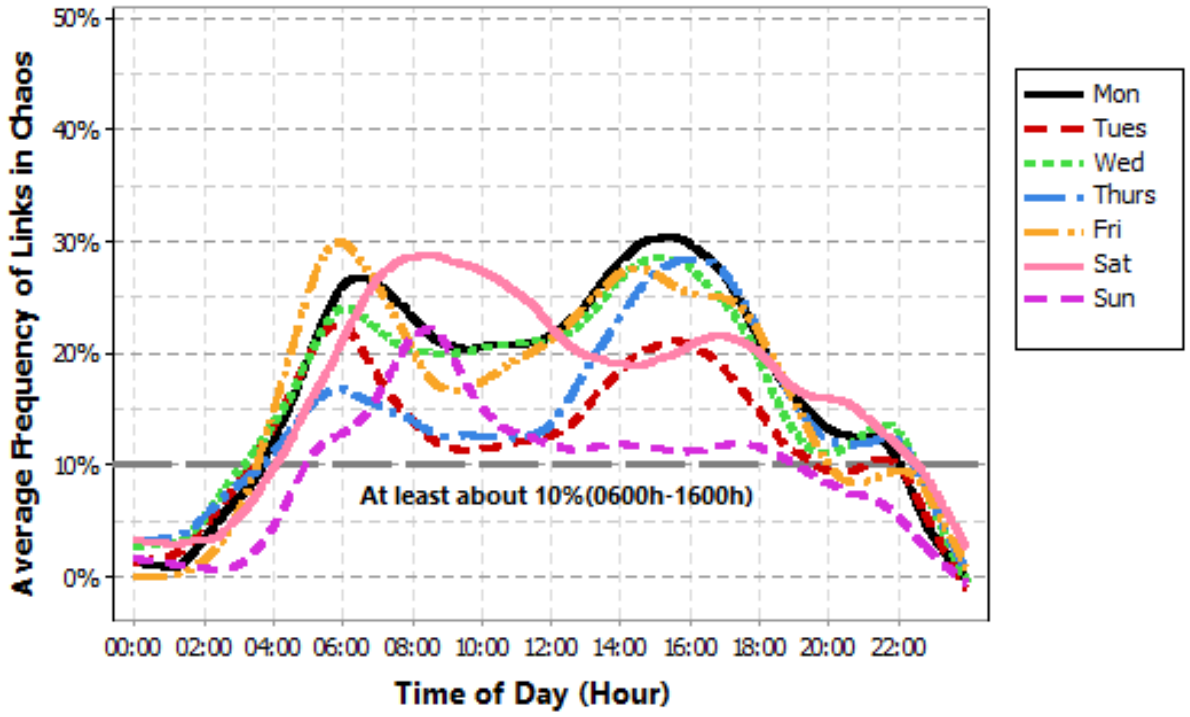


Figure 8.13: Graph of the Average 'Same Day' Profiles (July 1999)

The curves reveal that there is a delay of about two hours between the morning weekend and weekday peaks. That is, the weekday peak is approximately two hours ahead of the

weekend peak. This is consistent with commuter traffic during the week and shopping at the weekend. The absence of an evening peak on Sundays reflects early closing. The late evening peak on Saturdays is consistent with leisure and entertainment activities.

Across the cohorts analysed, the highest frequency of links exhibiting chaos at the peak time ranged between approximately 28-32%, which occurred on different days. The profiles generally indicated that the shapes of the weekend profiles (especially between 0600h-1100h) were very similar, as were the weekday profiles (see Figure 1I through to Figure 7I in Appendix I), and therefore gave distinct clusters. This observation, whilst illustrating the degree of consistency of patterns on particular days throughout the year month to month, reflects the different characteristic traffic patterns in terms of volumes over the weekends. Therefore, there is a natural expectation for stronger correlations between the pairs of weekday-weekday or weekend-weekend than the weekday-weekend.

Table 8-13 presents a summary of the results of the correlation assessments for the average same-day profiles for Cohort 4. Generally, the results show that the correlations between the pairs of weekday-weekday or weekend-weekend are stronger than the combination of the weekday-weekend. Cohort 2 is the only exception, where the average coefficients of the weekday-weekend indicate stronger relationship than the weekend-weekend, albeit less than the weekday-weekday. This might be because February is a shorter month and possibly contains a school half-term.

	Correlations of all daily pairs			Average correlations		
Cohorts	Minimum	Maximum	Average	weekday-weekday	weekend-weekend	weekday-weekend
1: January	0.58	0.87	0.77	0.78	0.87	0.76
2: February	0.51	0.89	0.73	0.79	0.66	0.68
3: March	0.59	0.87	0.79	0.83	0.77	0.76
4: April	0.60	0.87	0.76	0.77	0.85	0.73
5: May	0.66	0.85	0.77	0.82	0.78	0.77
6: June	0.58	0.89	0.77	0.84	0.82	0.69
7: July	0.46	0.85	0.68	0.78	0.78	0.58
Average	0.57	0.87	0.75	0.80	0.79	0.71

Table 8-13: Summary of Correlation Coefficients for Scenario A4 (All Cohorts)

The results indicate that the average monthly minimum correlation between the daily profiles is 0.57 ranging between 0.46 and 0.66. Similarly, the maximum is 0.87 ranging between 0.85 and 0.89, whilst the average is 0.75 ranging between 0.68 and 0.77. Further, the average monthly correlation of the pairing of weekday-weekday ranges between 0.77 and 0.84 and on average 0.80. Similarly, the weekend-weekend ranges between 0.66 and 0.87 and is on average 0.79, whilst the weekday-weekend is on average 0.71 ranging between 0.68 and 0.77.

The Kruskal-Wallis test gave consistent p-values to indicate there is no statistical significant difference between the seven cohorts in terms of the minimum, maximum and average correlations in Table 8-13. A similar assessment for the weekday-weekday, weekend-weekend and weekday-weekend options also revealed the same conclusions. The p-values ranged between 0.435 and 0.499. This result endorses the observations from Scenario A3.

8.6.4 Summary and Inferences

By undertaking the same analysis for the seven cohorts, the main inferences from the analyses of the positive Lyapunov exponents presented above are:

1. generally, the correlations of the daily profiles against the average same-day profiles show stronger relationships than with the average weekly profiles. Otherwise, there is no statistical difference in the results although the average coefficients of the former are greater than the latter;
2. there is a stronger relationship between the pairs of weekday-weekday or weekend-weekend than the pairs of the weekday-weekend profiles (0330h-2200h);
3. there is at least about 10% links showing chaos during daytime period; and
4. the highest frequency of links showing chaos at the busiest time is *circa* 30%.

The findings in 1 and 2 (above) are expected because the patterns of traffic flows tend to be similar for the same-day of the week, and therefore there is a stronger relationship between the week's daily profiles and its respective average same-day profile across the month. Given the variability in traffic flow from day-to-day, the daily profiles as expected were found not to correlate strongly with the average weekly profile. It is likely however, that some of the weekday daily profiles will correlate well with the average weekly profile due to the smoothing effects from the averaging of the profiles over the week. The

relationship between the daily and average same day profile implies that the average same-day profile may be useful in the short-term forecast model, as replacement data in the event of missing data, whether an entire or segment of the day.

Furthermore, the curves show distinct patterns in the profiles for the weekdays and weekends. For example, the highest proportion of links indicating chaos on the curves tend to concentrate around 0530h-0700h for the average weekday profiles, whilst the delayed peaks of the weekends occur one or two hours later. The curves have distinct shapes for the two categories. Such patterns make the relationships within the weekday-weekday and weekend-weekend groups stronger with high correlation coefficients compared with the weekday-weekend combinations.

The results indicate that there is more than one link indicating chaos during the busy daytime period when chaos is expected to be imminent. The minimum frequency of links experiencing chaos is *circa* 10%, whilst the highest is *circa* 30%. The observation therefore suggests that at the time when the network is in a chaotic state, there will be a number of links indicating a positive value of Lyapunov exponent. The results demonstrate that chaos has spatial effects network-wide, which might be a result of natural lags in traffic flow and of events (such as the sudden application of brakes, high flow volumes, accidents) in one network location giving rise to secondary effects in other locations.

8.7 Conclusions

The analyses have revealed that the optimum lag time of the traffic flow time series of a link can be predicted from a multiple regression model of the lag against mean or standard deviation, although the standard deviation model performed better. The lag times ranged between 140 and 270 minutes for the links in the network studied, which shows distinct lag times across the entire network. However, there are identifiable groups of adjacent links in which the lag times are not statistically significantly different. This is expected because platooning of traffic exists across these groups of links through SCOOT's offset optimisation and therefore similar traffic patterns were observed. The results imply that the lag durations monitor the extent of traffic platooning in the network such that if adjacent links indicate that the lags are not statistically significantly different then this is likely a result of platooning traffic on these links.

It is important to note that it was not possible to use the raw Lyapunov exponents of the individual links to generate a single parameter for analysing the entire network-wide system. The results show a high frequency of the negative values in the time series of the Lyapunov exponents, which is useful to determine the stable and unstable episodes on the link detector but not network-wide. From the traffic management point of view, it is important to provide an analysis tool for understanding the network-wide traffic behaviour necessary in order to keep the network operating in under-saturated conditions.

By focusing on the frequency of occurrences of positive Lyapunov exponents offers the potential to target those individual network links occurrence of chaos to address the network-wide congestion problems. This frequency of occurrence analysis revealed that the correlations between same-day (Monday-Monday, Tuesday-Tuesday, etc.) profiles did not yield any specific pattern, nor any indication of a statistically consistently strong relationship. This suggests substantial inconsistency of traffic from day to day in a week, or same weekday in different weeks. Therefore, there was further analysis of correlation of the daily and respective average same-day profiles (for example, Monday versus average of all Mondays, or Tuesday versus average of all Tuesdays etc.), and daily and respective average weekly profiles (for example, Monday, Tuesday etc. in week 1 versus average of all days (Mondays to Sunday inclusive) in week 1). Based on the understanding that the daily correlations with the average same-day profiles were stronger than the weekly profiles, the average same-day profiles were examined to determine the progression of chaos in the network of SCOOT links. According to the analysis, there is at least about 5-10% links indicating chaos during the daytime period and a maximum of 30% during the onset of congestion at the busiest time of the day.

The lag interval between an event occurring at one location and its effects at other locations whether upstream or downstream at a later time will be investigated in detail in Chapter 9. Subsequently, by analysing the causes and effects of the temporal and spatial evolution of traffic behaviour using the Lyapunov exponents, this research work demonstrates two important possible applications to signalised urban networks. These are discussed in Chapter 9 making use of the Lyapunov exponents, which characterises the different traffic flow regimes (free-flow, unstable and congested), to investigate the forecast of link

occupancy and, subsequently how this may be beneficial for managing traffic in signalised urban networks.

Chapter 9. Exploring the Temporal and Spatial Relationships of Lyapunov Exponents for Short Term Forecasting

9.1 Introduction

The previous chapter identified the Lyapunov exponent as a suitable parameter for determining the dynamic status of the traffic in a signalised urban road network. The Lyapunov exponents were subject to further analysis to demonstrate successfully potential for practical use in managing traffic in actual urban networks. Consequently, this chapter presents the results of a real-world investigation into the application of the Lyapunov exponents for predicting the short-term traffic in a SCOOT region in an urban road network. It also explores the use of the Lyapunov exponent for anticipating where and when congestion will occur enabling traffic management strategies to be implemented in advance to deal with impending traffic conditions. This will be achieved by an assessment of the temporal and spatial correlations between links in the network.

In line with the stated objectives, Section 9.3 introduces the cross-correlation analysis that underpins the assessments in this chapter and explains how an understanding of the lags may be helpful for enhancing our knowledge of the evolution of traffic flow. Section 9.4 describes the forecasting model developed to deliver the short-term traffic predictions, enumerates the scenarios assessed based on the model, and presents the descriptive statistics of the results of the modelled scenarios. Section 9.5 discusses the link-by-link results of the short-term predictions of link occupancies, and Section 9.6 presents the network analysis to identify the critical link based on the Lyapunov exponents and demonstrates how this understanding could enhance traffic management practise in urban areas.

9.2 Review of Spatial-Temporal Time Series Models

A review of traffic condition models was discussed at Section 3.3. This section provides a further review of spatial-temporal regression models in order to justify the forecast technique based on nonparametric regression model adopted in the thesis. By recognising that lags are naturally embedded in the traffic evolution process, the ideal model must have the capability to forecast the traffic flow at a given location using the time series information of the neighbouring locations in the network. Essentially, this model must be

able to incorporate not only the temporal but also the spatial relationships of the predicted (dependent) link with the other (independent) links in the network.

Models that describe the spatial-temporal evolution of a single or multiple variable relationship in space and time has received significant attention in the last thirty years due to the advancement in computer technology and existence of large databases. Equation 9-1 provides generalisation of the spatial-temporal model proposed by Cliff and Ord (1975a) incorporating autoregressive, moving average and regression techniques (and various combination of these) to model the behaviour of economic series.

$$Y = \rho W y + X\beta + \varepsilon + \gamma A\varepsilon \qquad \textit{Equation 9-1}$$

In this model, W and A represent the matrices of known 'structural weights', X is a matrix of known (non-stochastic) regressor or independent variables, ρ, γ and β are unknown parameters, and ε is a random disturbance term with zero mean. The first, second and fourth terms of the equation are the autoregressive, regressive and spatial moving average component of the model. Therefore, if $\gamma = 0$ and $\beta = 0$, the model is purely spatially autoregressive; if $\rho = 0$ and $\gamma = 0$, the model is a conventional spatial regressive model of Y upon a set of independent variables, X; and if $\beta = 0$ and $\rho = 0$, the model reduces to a spatial moving average.

Following this work, various techniques have evolved in response to different inferential needs and data types in order to examine the relationship between two variables in space and time. These include methodologies for space-time analysis such as Space-Time AutoRegressive Integrated Moving Average (STARIMA), Bayesian Vector AutoRegressive (BVAR), Dynamic Space-time (DS-t) and Seemingly Unrelated Regression (SUR) (Cliff and Ord, 1975b). Both the STARIMA and BVAR models apply when investigating the spatio-temporal evolution of a single variable. The distinction is that STARIMA performs better for large spatial and temporal dimensions, whilst BVAR is better suited for studies that involve a small number of locations. The DS-t and SUR models describe relationships between multiple variables with a spatio-temporal reference.

It is important to note that in STARIMA, each time series is modelled simultaneously as a linear combination of past observations and disturbances at neighbouring sites.

Fundamentally, in some respects, the spatial-temporal models are an extension of the univariate modelling techniques. For example, STARIMA models depend on the basic univariate ARIMA principle, where the more recent exerts more influence than the distant past. Such models are unsuitable in this research for two reasons. Firstly, STARIMA models a variable (flow) with only information about the same variable (flow) without considering other information (for example, occupancy and speed) that may be relevant also. However, in the research reported in this thesis a variable was modelled with the values of other variables.

In spatial analysis which focusses on time series based on multiple variables, the spatio-temporal evolution of traffic conditions, or functional relationship between variables, is not as important as the relationship between measurement of different variables (such as flow and density) observed in space and time. Kamarianakis and Prastacos (2003) proposed a generalised first order serial and spatial autoregressive distributed lag model in vector form in macroeconomic modelling to examine the relationship between time series variables such as Gross Domestic Product (GDP), employment, labour force participation, productivity etc. representing data collected for different regions. The model provided generalised relationships between variables in space and time based on aggregated data of the entire time series. Therefore, such models are limited due to their inability to take advantage of disaggregate time series data to provide robust models that focus on, for example, specified time periods.

For the reasons already discussed, the above models that are primarily based on autoregression are unsuitable for this research. However, the lagged distribution of the time series observation was found to be necessary in any model considered for this work due to the time lag between causes and effects at different spatial locations in the network. Further, the model must be capable of making predictions based on disaggregate data to construct specific models for specific duration of time series observations.

In deciding the choice of model for this work, attention was paid to the knowledge that regression models can treat time series data in a disaggregate form. This suggests that it is possible to construct a spatial regression model for different locations over different time periods and thus the geographical relations can be modelled implicitly in the covariance of the system of equations (Anselin, 1988, Zellner, 1962).

Regression models, unlike the models discussed above, have the inherent ability to model a variable based on information of another variable, rather than on its own information. Linear regression is used to model the value of a dependent variable based on its linear relationship to one or more predictors. Nonlinear regression is appropriate when the relationship between the dependent and independent variables is not intrinsically linear. Regression models may also be either parametric or non-parametric. A finite number of parameters, which when defined remains constant irrespective of the forecast time segment, represent a parametric model. On the contrary, the nonparametric model is not represented by a finite number of parameters but instead has the ability to compute a whole function instead of a set of parameters. Therefore, when applied, a specific unique model is generated for every forecast scenario.

A class of nonparametric model is the modelling, which can learn from local conditions and therefore provides a more accurate model estimator than other models such as neural networks that are global in nature (Atkeson *et al*., 1997). Local models are real-time models that generate the functions from an infinite stream of training data. These models include the local constant regression such as the k-nearest neighbours (Smith *et al*., 2000), kernel estimator (Elfaouzi, 1996) and local linear regression (Fan and Gijbels, 1996). Rice *et al* (2001) in a study of travel time applied the time varying coefficient linear model, which is closely related to the local linear regression model, where the weighting function was determined by the difference in departure time rather than the distance of the covariates space.

A potential problem with regression models applied to time series data is that the data tends to be generally autocorrelated. As such, if there are common elements in the autocorrelation structure of the dependent and independent variables, then the equation may show strong relation when in fact the independent variables have no real explanatory power of the dependent variable. Therefore, the concept of "dynamic regression" models for time series data has been applied in this research. Dynamic regression (distributed lag regression) models enhance times series modelling by taking into account the possible time lagged relations between the input and output variables (Pankratz, 2012).

Equation 9-2 gives a purely regressive model of Equation 9-1 for a dependent variable Y_t and a set of independent variables X_1, X_2 … X_k observed at a given time interval t.

Similarly, Equation 9-3 presents the lag distributed regression model for the dependent and independent variables.

$$Y_t = \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \cdots \beta_k X_{k,t} + e_t \qquad \textit{Equation 9-2}$$

$$Y_{t-s} = \beta_0 + \beta_1 X_{1,t-1} + \beta_2 X_{2,t-1} + \cdots \beta_k X_{k,t-1} + e_{t-1} \qquad \textit{Equation 9-3}$$

where:

- Y_t is the linear function of the k predictor variables $(X_{1,t}, X_{2,t}, \ldots X_{k,t})$
- Y_{t-s} is the lag distribute model for the k predictor variables $(X_{1,t-1}, X_{2,t-1}, \ldots X_{k,t-1})$
- β_0 is the population intercept (a constant term)
- $\beta_1, \beta_2 \ldots \beta_k$ are the coefficients of the independent variables
- e_t and e_{t-1} are usually assumed to be an uncorrelated error term

By reflecting on the knowledge of regression models in the background literature review, it was concluded that the distributed lag model in Equation 9-2 would be most suitable for this research. Prior to applying Equation 9-2, the model undertakes a cross-correlation assessment exercise to evaluate the lag between the dependent and each of the independent variables. Then, a specified criteria is imposed to select a particular set of independent variables to develop the model function, which was generated uniquely for every hour of the day for every predicted link.

9.3 Exploring the Lyapunov Exponents for Predicting Link Occupancy

This section develops a forecast function to enable a prediction of the network link occupancies using the Lyapunov exponents. Due to natural lags embedded in the traffic data, the traffic behaviour at a given point at time $t(s)$ tends to be related to the data at another location at time $t(s-\tau)$, where τ is lagged (delay) time. The lag interval between two locations gives an indication of how long it takes for the traffic flow at an upstream location to influence the traffic flow at a downstream location. In order to forecast traffic accurately therefore, the natural lag within the data is required to provide a good reference for the selection of the forecast horizon. A sound understanding of lags in the datasets are important – the longer the lags, the longer the time horizon for the prediction of a traffic condition at a particular network location enabling timely selection and implementation of remedial measures to prevent congestion. Further, as there are three variables involved in

this research, namely traffic flow and occupancy variables, and the computed Lyapunov exponents, it is important to determine which of these variables achieve the longest time horizon for the forecast.

Consequently, this section aims to understand the existing lag within traffic flow and occupancy data variables as well as the computed Lyapunov exponents. This analysis is carried out based on the cross-correlation function given by Equation 6-2 (see section 6.7.3), which helps to determine the shift in time (lag) that maximises the correlation coefficient between any pairs of network links.

In these assessments, the tested lags were within the range of zero to 1000 seconds, in steps of 100 seconds. It must be noted that these lags from the cross-correlation method are different from the delay time (calculated from the autocorrelation method) for the phase space reconstruction. The autocorrelation lags correlate the data series with its own past or future values at various lag values and uses the correlogram plots of the resulting coefficients to select the optimum lag as equivalent to the solution at a coefficient of 0.4. Therefore, this method ensures that in the process of the phase space reconstruction the length of lag when the data series is neither too correlated nor uncorrelated is used, whilst the cross-correlation method gives lags that maximises the correlation between the time series of two links.

9.3.1 Determining the Cross-correlation Lags

Cross-correlation assessment was undertaken for all possible pairs of links in order to examine the relationship between the maximum correlations and corresponding optimum lags (see Program 1 in Appendix B). The four scenarios assessed were (i) Flow - Flow (ii) Occupancy – Occupancy (iii) Lyapunov exponents – Lyapunov exponents (iv) Occupancy - Lyapunov exponents and (v) Lyapunov exponents - Occupancy, for January 1999. These scenarios describe how the datasets were correlated; for example, scenario (i) implies flow data from one location is correlated with flow data from another location, whilst scenario (iv) implies occupancy data from one location is correlated with Lyapunov exponents profile from another location. The results are based on the averaging of the lags for all active loops over a 24-day period in January 1999 (see Table 1J through to Table 6J in Appendix J). Figure 9.1 shows a boxplot of the comparative distributions of the scenarios based on the average daily lags, where higher negative values indicate that the lags are

longer and *vice-versa*. The results indicate that the lags for the 'Lyapunov exponent to Lyapunov exponent' were longer than 'flow to flow' and 'occupancy to occupancy'. The lags for flow occur earlier than occupancy, confirming the observation in section 7.3.3, where the flow detects chaos slightly earlier than occupancy.

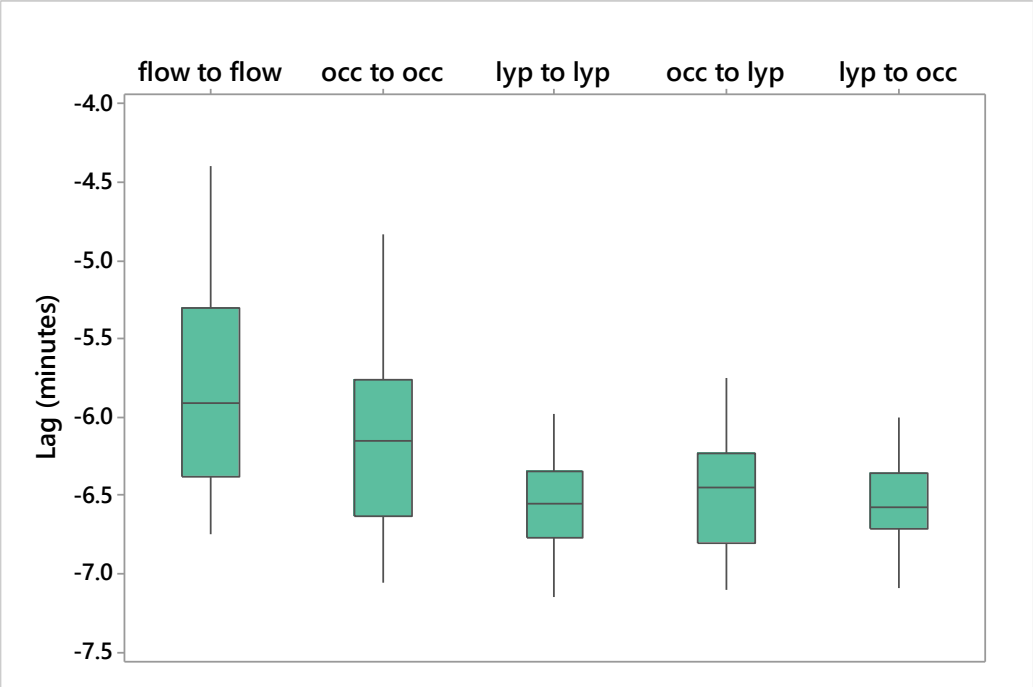


Figure 9.1: Box Plot of Lags for Flow, Occupancy and Lyapunov Exponents

From the traffic management perspective, this outcome is an indication that the exponents provide a longer forecast duration than flow and occupancy and therefore gives traffic managers more time to react. On average, the lags of the exponent were between 5.97 and 7.14 minutes (average of 6.54 minutes), which is 10.9% and 7.2% more than flow and occupancy respectively, but only 1.7% more than 'occupancy-exponents'. This is confirmed by the Mann-Whitney statistical test of equality at 95% confidence interval. The lags for the 'exponents-exponents' option ranged between 5.97 and 7.14 minutes (average of 6.54 minutes), and were not statistically significantly different compared with the 'occupancy-exponent' option. The 'exponents-occupancy' option however ranged between 6 and 7.1 minutes (average of 6.56), which was 0.3% better than the 'exponent-exponent' option. The results suggests that, among all the scenarios considered, the 'exponents-occupancy' option will forecast the occupancy longer. Further, this implies that by choosing a forecast horizon of between 6 and 7.10 minutes, the forecast data would

be correlated strongly with actual data irrespective of predicting the occupancy from itself (occupancy) or the exponents.

The result is reasonable justification for predicting the occupancy from the exponents. Firstly, it is an objective of the research to identify a parameter based on Chaos Theory suitable for predicting a traffic flow parameter. Since the computation of the exponents used traffic flow data, it is reasonable to use the exponents to predict the occupancy that has no direct connection to the calculation of the exponents. Given that the lags of the 'occupancy-exponents' are relatively long, and they offer benefits for traffic management, this is the optimum choice for further investigations of the link-to-link interactions within the study network. It indicated that the Lyapunov exponents could enable selection of a reasonably lagged yet well-correlated link pairs in the prediction model and therefore produce a better predictor of the occupancy of a link.

9.4 Construction of Multiple Regression Models

Figure 9.2 presents a generalised form of the modelling procedure, which relies on Lyapunov exponents (from the phase space constructed using traffic flow data). The predictions of the link occupancies were determined from a multiple regression model, which generates a unique forecast model for each link for every hour of the day (see Program 2 in Appendix B). Suppose there are 40 links numbered *1, 2, 3,…40* and the occupancy and Lyapunov exponent profiles are $O_1, O_2, O_3, \ldots O_{40}$ and $LE_1, LE_2, LE_3, \ldots LE_{40}$ respectively. The first step is to carry out a cross-correlation for every link's occupancy against all the Lyapunov exponents of all other network links. For example, the occupancy profile O_1 (link (1)) is cross-correlated with $LE_2, LE_3 \ldots LE_{40}$. The next step is to determine the links whose Lyapunov exponent profile to include in the regression model (the "selected links" are those links where the maximum correlation coefficient and associated lags meets a "specified criteria" in the model).

In this example illustrated in Figure 9.2, there are four selected links to predict the occupancy of an arbitrary link at time T+15. The cross-correlation lags of the links are 15, 10, 20 and 17. The forecast model is constructed using one-hour occupancy data between T and T-60. Taking into consideration the respective lags of the selected links, the range of the Lyapunov data are as follows: between T-15 and T-75, T-10 and T-70, T-20 and T-80, and T-17 and T-77 for the first, second, third and fourth link respectively. Once

completed, the model predicts the next 15-minute occupancy by using the respective links data ranging between T and T-60, T+5 and T-55, T-5 and T-65, and T-2 and T-62 to forecast the occupancy for the T+15 and T-45.

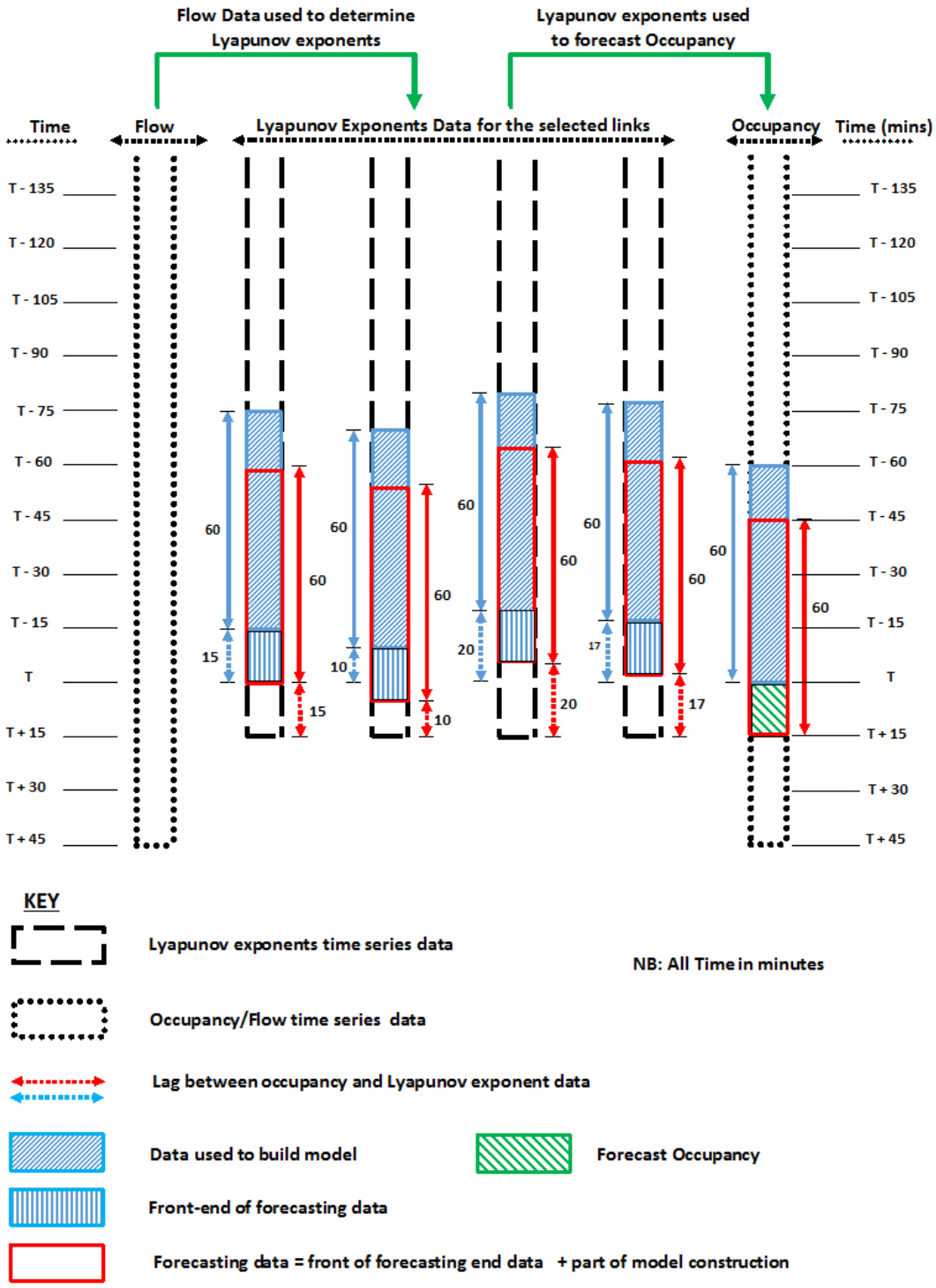


Figure 9.2: Diagram of the Generalised Form of the Forecast Model

By way of illustration, Table 9-1 presents the 24-hour profile for the qualifying links (the respective correlation coefficients and corresponding lags) for constructing the regression model for the Link 2 forecast on the 19th May 1999. The "selected links" were found to be consistently different, even for the same link at different hours of the same day or the same hours on different days reflecting the variations in traffic flows under different varying conditions of the day. This implies that when applied to traffic management remedial measures will also need to be monitored and continuously updated.

Hour of Day	Feeder Links (cc > 0.4)		
	IDs	**Correlation Coefficients**	**Forecast Horizon**
01:00	6, 10	0.40, 0.43	-3.3, -16.7
03:00	5, 32	0.43, 0.40	-6.0, 2.0
04:00	2, 3, 40	0.57, 0.45, 0.44	-2.7, -2.0, -6.0
05:00	1, 21, 27	0.43, 0.42, 0.40	-7.3, -16.7, -5.7
06:00	1, 25, 32	0.41, 0.47, 0.41	-9.3, -4.0, -9.7
07:00	1, 7, 12, 30	0.46, 0.51, 0.42, 0.44	-8.3, -2.0, -15.7,-4.7
08:00	4, 12, 25, 33	0.49, 0.45, 0.42, 0.49	-12.3, -5.7, -5.3, -15.0
11:00	10, 31, 39	0.41, 0.48, 0.52	-16.7, -9.3, -10.7
12:00	3, 4, 5, 8, 19, 23	0.46, 0.49, 0.60, 0.41, 0.44, 0.55	-7.0, -10.3, -11.0, -15.7, -6.7, -11.0
13:00	27	0.40	6.0
15:00	35	0.46	-2.3
16:00	8, 40	0.55, 0.42, 0.42	-5.7, -4.0
19:00	9	0.45, 0.45	-9.0, -6.7
20:00	6, 28	0.41, 0.46	-3.7, -9.7
23:00	22, 25	0.40, 0.41	-3.3, -7.7

Table 9-1: Summary of Model Information for Link 2 (15th May 1999)

In order to predict the hourly occupancy of a link, a multiple regression forecast function is determined for each hour using the profiles of hourly Lyapunov exponents (or occupancy) of the selected links. The generalised form of the regression formula for the hourly prediction is given by:

$$O_i = a + \sum_{x=1}^{n} b_x * L_x \qquad \textit{Equation 9-4}$$

where:

- O_i is the occupancy for Link (i), which is the predicted occupancy per link per hour

- n is the number of variables per hour explaining the variance in the predicted occupancy (O_i) per link per hour
- a is the constant or intercept of the regression line estimated per link per hour
- b_x is the parameter for L_x estimated per link per hour
- L_x is the Lyapunov exponents profile per hour of link i (LE_i)
- L_n is the n^{th} (Last) independent Lyapunov exponents explaining the variance in O_i per link per hour

For example, the regression model for 0815h-0915h will be based on the hourly Lyapunov exponents data from links 4, 12, 25 and 33, which are lagged at approximately 12, 6, 5 and 15 minutes respectively, relative to the 0800h-0900h occupancy data from link 2. The data for the independent variables (links 4, 12, 25 and 33) in the model will therefore correspond to data from 0748h-0848h, 0754h-0854h, 0755h-0855h and 0745h-0845h respectively.

Section 5.5.2 notes that the shortfall in duration between the actual time series and the reconstructed phase space could limit its application. Figure 9.2 illustrates this gap as the differential in length between the exponents and the flow/occupancy data. In order to fill in this missing data to enable real time application, this work suggests given the evidence of a strong correlation between the same-day and the respective average same day profiles (see section 8.6.2) that the average same day profile of the recent weeks is used as surrogate data in the model. Alternatively, a localised regression model would be useful to generate this data.

9.4.1 Modelling Scenarios

In order to test the model, the Lyapunov exponents from a preceding time segment (based on the lag) were input into the model, and the model used to predict the occupancy for succeeding time segments. Five scenarios were considered:

1. Scenario B1 predicts the hourly occupancy profile **24 hours ahead** (for example, LE data from 0700h-0800h is used to construct the model, then the next day 0700h-0800h LE data is selected to forecast the 0700h-0800h occupancy). The "specified criteria" for the "selected links" is maximum correlation coefficients and lags greater than 0.4 and 100 seconds respectively.

2. Scenario B2 predicts the hourly occupancy profile **15 minutes into the next hour** (for example, LE data from 0700-0800 is used to construct the model, then the 0715-0815 LE data is selected to forecast the 0800-0815 occupancy). The "specified criteria" for the "selected links" is maximum correlation coefficients and lags greater than 0.4 and 100 seconds respectively.
3. Scenario B3 is same as Scenario B2, except that occupancy data in used instead of LE data.
4. Scenario B4 is same as Scenario B2, except that the threshold criteria for the "selected links" for the maximum correlation coefficient is greater than 0.7.
5. Scenario B5 is same as Scenario B2, except that the threshold criteria for the "selected links" for the maximum correlation coefficient is greater than 0.3.

9.4.2 Scatter Plot Analysis

Figure 9.3 is a typical scatter plot of the predicted versus actual occupancy for link 1 (Scenario 5) on the 28th July 1999. Figure 9.3 includes a total number of 4320 predicted points at every 20 second interval throughout the day. It indicates a "very strong" correlation coefficient between the predicted and actual occupancy of approximately 93%.

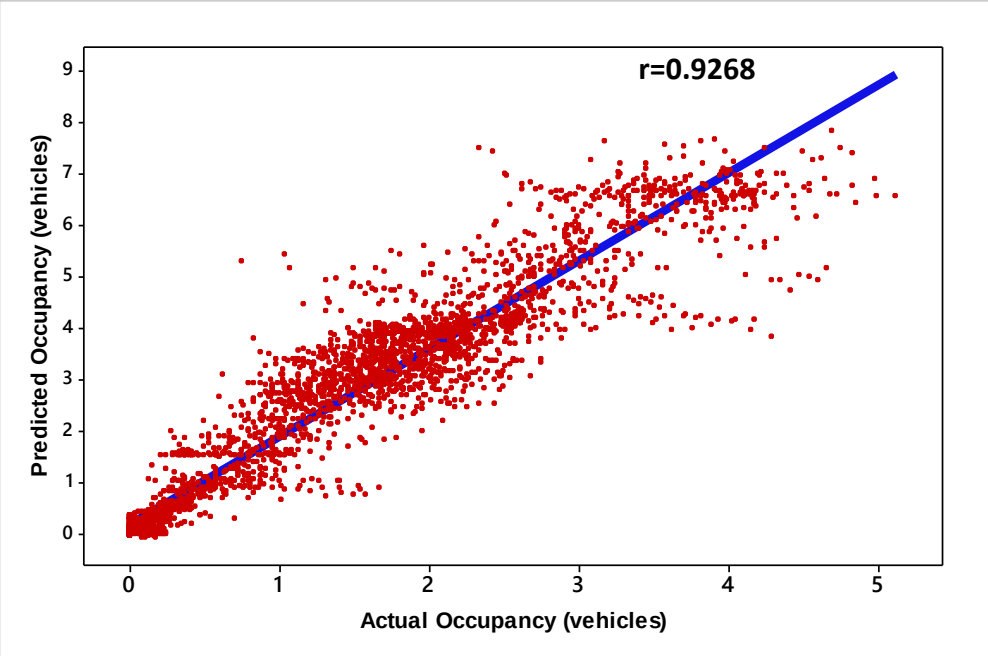


Figure 9.3: Scatterplot of the Predicted versus Actual Occupancy of Link 1 (typical)

9.4.3 Descriptive Statistics of Correlation Coefficients

Figure 9.4 presents a boxplot of the median values for the correlation coefficients of the predicted versus actual occupancy based on the daily hourly predictions in Scenario B1.

The number of days that were forecast in January, February, March, April, May, June and July are 25, 26, 29, 28, 29, 28 and 28 respectively.

The respective minimum and maximum correlation coefficients sampled from the daily hourly correlation coefficients are included in Table 9-2, together with the median for this data range for the entire months.

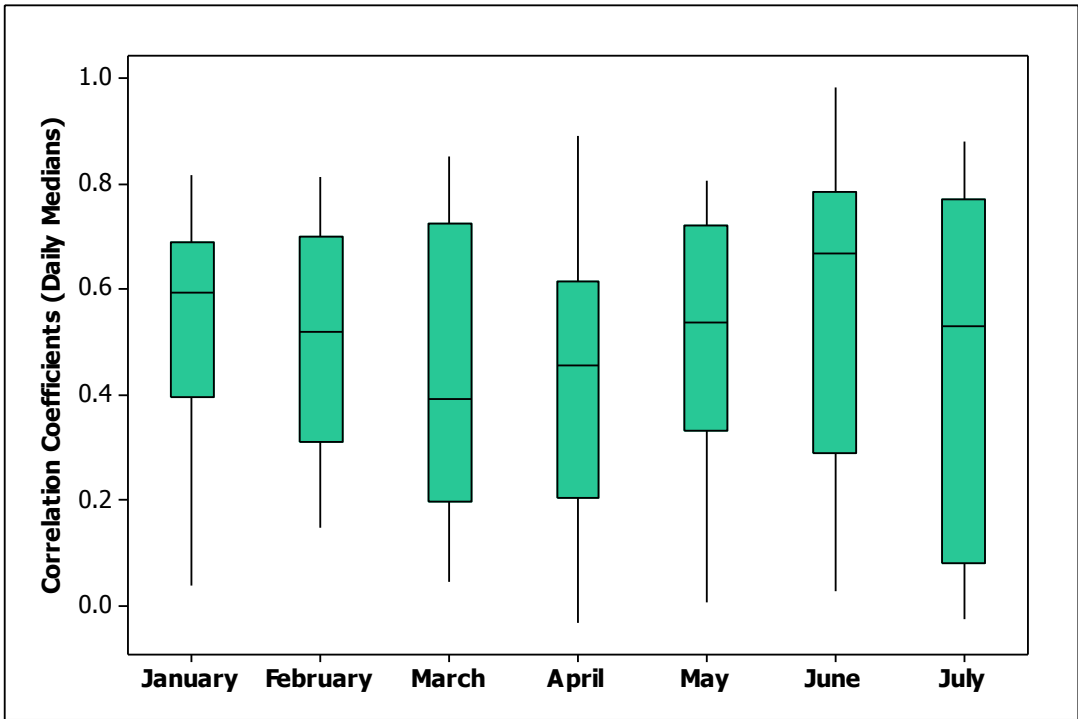


Figure 9.4: Boxplot of the Median Correlation Coefficients of the Hourly Actual versus Forecast Occupancy (Scenario B1)

Cohort	Minimum	Maximum	Median
1: January	0.0381	0.8163	0.5961
2: February	0.1495	0.8143	0.5198
3: March	0.048	0.8538	0.3935
4: April	-0.0306	0.8923	0.4581
5: May	0.0087	0.806	0.5382
6: June	0.0274	0.9829	0.6678
7: July	-0.0238	0.8797	0.5324

Table 9-2: Summary of Median Correlation Coefficients of the Hourly Actual versus Forecast Occupancy (Scenario B1)

Similar to Figure 9.4, the following four figures presents the boxplots for the median values for the correlation coefficients of the predicted versus actual occupancy based on the daily

hourly predictions for other scenarios analysed, whilst a summary of the minimum, maximum and median correlation coefficients for their cohorts are included in Table 9-3.

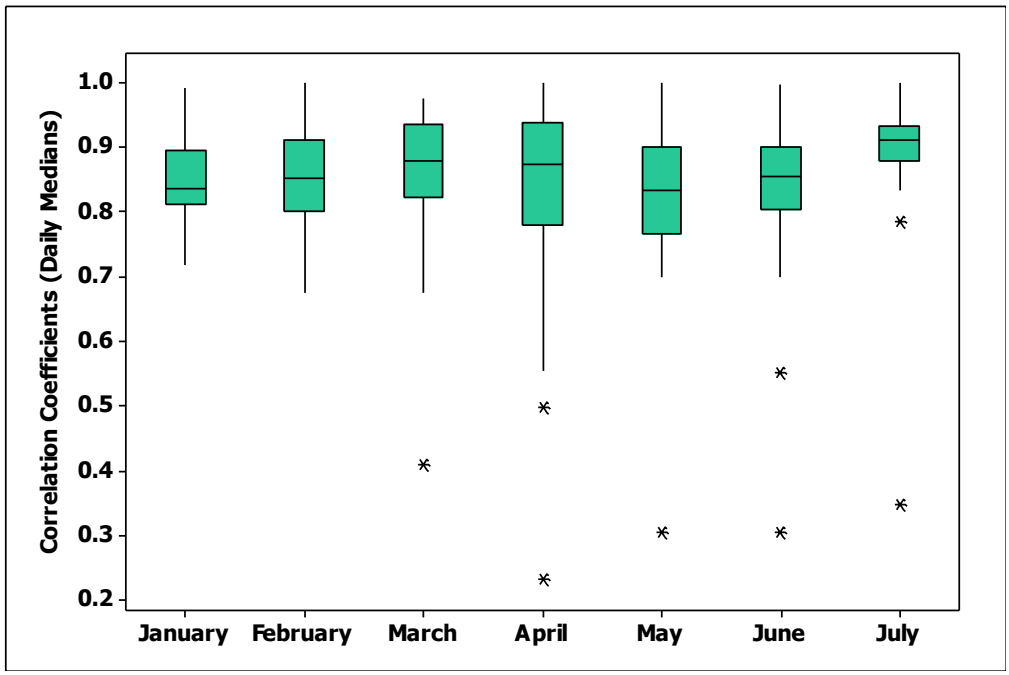


Figure 9.5: Boxplot of the Median Correlation Coefficients of the Hourly Actual versus Forecasts Occupancy (Scenario B2)

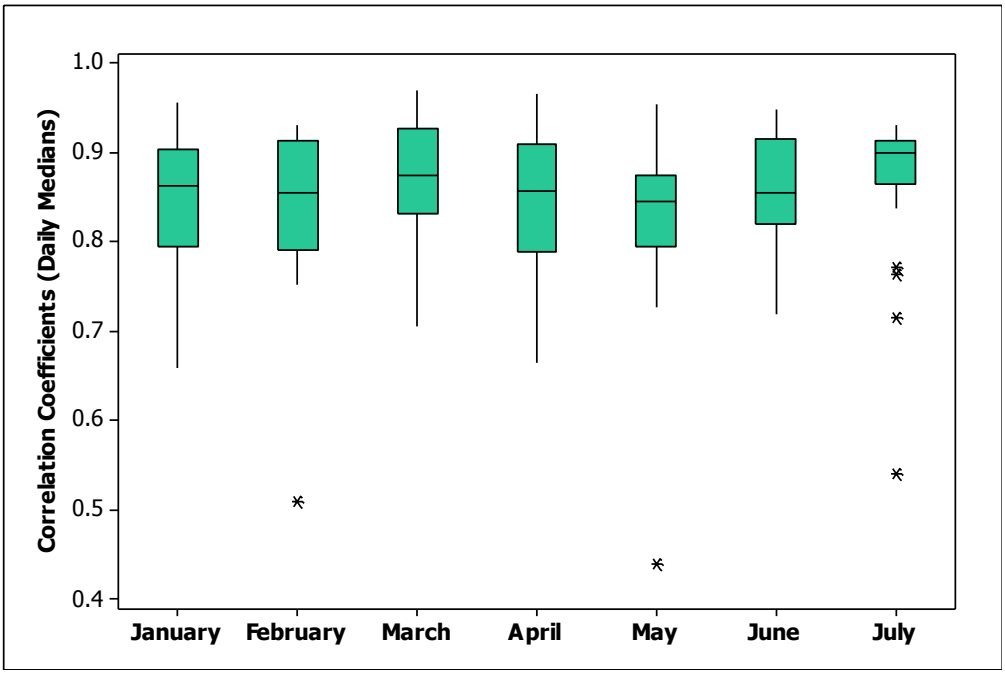


Figure 9.6: Boxplot of the Median Correlation Coefficients of the Hourly Actual versus Forecasts Occupancy (Scenario B3)

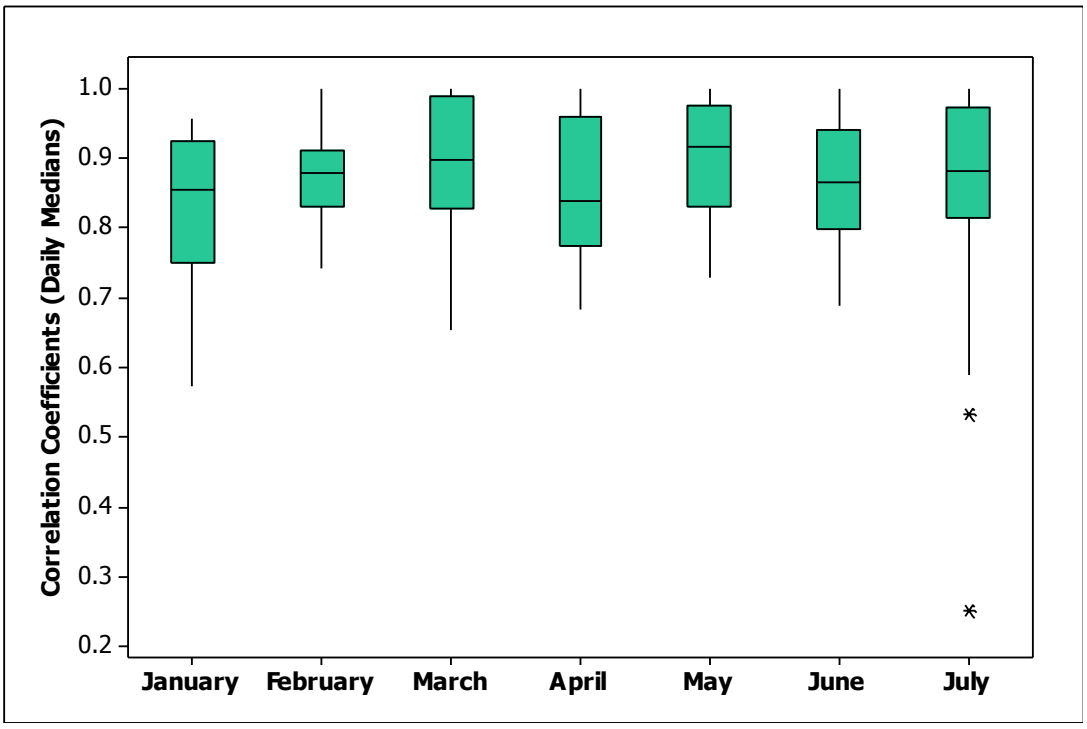


Figure 9.7: Boxplot of the Median Correlation Coefficients of the Hourly Actual versus Forecast Occupancy (Scenario B4)

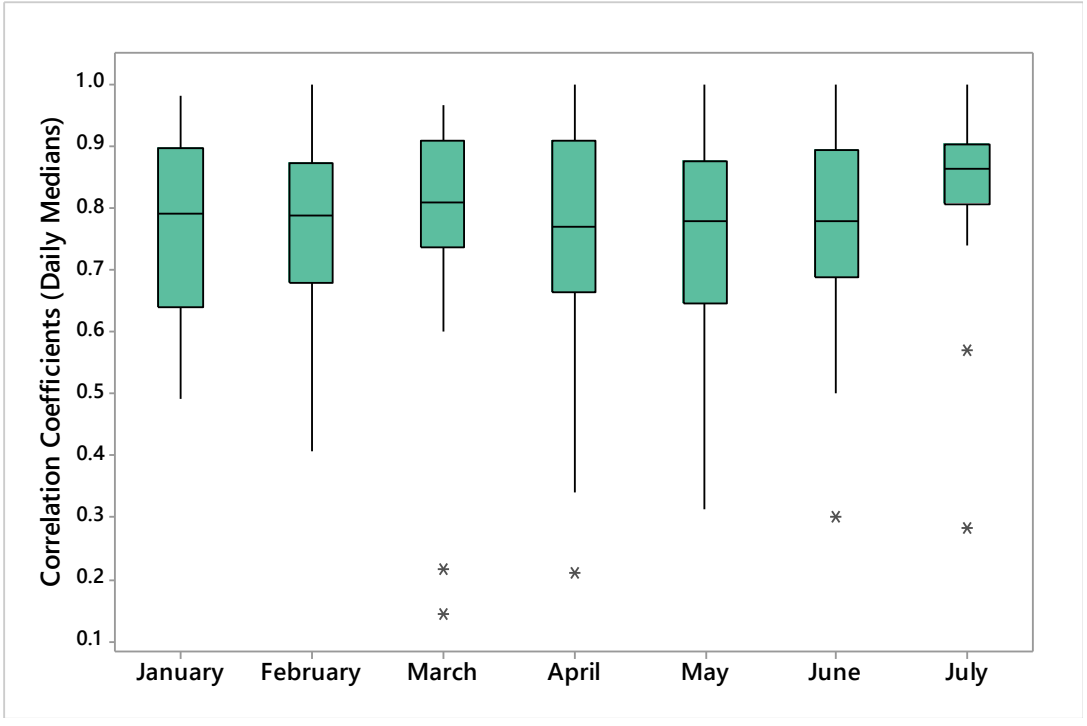


Figure 9.8: Boxplot of the Median Correlation Coefficients of the Hourly Actual versus Forecast Occupancy (Scenario B5)

	Cohorts						
	1	2	3	4	5	6	7
Scenario	**Minimum Correlation Coefficients**						
Scenario B2	0.7189	0.6754	0.4110	0.2318	0.3064	0.3064	0.3487
Scenario B3	0.6589	0.5095	0.706	0.6643	0.4399	0.7196	0.5401
Scenario B4	0.5726	0.7418	0.6529	0.6844	0.7277	0.6874	0.2570
Scenario B5	0.4920	0.4081	0.1426	0.2106	0.3127	0.3013	0.2830
	Maximum Correlation Coefficients						
Scenario B2	0.9936	0.9997	0.9759	0.9998	0.9999	0.9989	0.9993
Scenario B3	0.956	0.9301	0.970	0.9659	0.954	0.948	0.9307
Scenario B4	0.9568	0.9998	0.9997	0.9999	0.9999	0.9999	0.9999
Scenario B5	0.9817	0.9996	0.9660	0.9997	0.9998	0.9995	0.9992
	Median Correlation Coefficients						
Scenario B2	0.8366	0.8527	0.8794	0.8745	0.8329	0.8561	0.9122
Scenario B3	0.8630	0.8549	0.8748	0.8579	0.8448	0.8555	0.9005
Scenario B4	0.8537	0.8777	0.8971	0.8398	0.9172	0.8641	0.8808
Scenario B5	0.7907	0.7880	0.8095	0.7699	0.7789	0.7774	0.8635

Table 9-3: Summary of Median Correlation Coefficients of the Hourly Actual versus Forecast Occupancy (Scenarios B2, B3, B4 and B5)

Figure 9.5 presents a boxplot of the median correlation coefficients in Scenario B2. The number of days analysed in the respective months from January to July are 27 (January), 27 (February), 30 (March), 30 (April), 29 (May), 29 (June) and 29 (July).

Figure 9.6 presents a boxplot of the median correlation coefficients in Scenario B3. The number of days that were predicted are 26 (January), 26 (February), 30 (March), 27 (April), 26 (May), 27 (June) and 28 (July). Similarly, Figure 9.7 presents the results for Scenario B4. The number of days for which predictions were undertaken are 10 (January), 27 (February), 30 (March), 29 (April), 26 (May), 29 (June) and 29 (July). Further, Figure 9.8 presents a boxplot of the medians of the hourly predictions in Scenario B5. The number of days that were predicted are 10 (January), 27 (February), 30 (March), 29 (April), 26 (May), 29 (June) and 29 (July).

In general, the correlation coefficients are higher in Scenario B2, compared with Scenario B1. By comparing the groups of medians for the monthly cohorts, the Kruskal-Wallis test of equality gave a statistically significant p-value of 0.785 for Scenario B1, which indicates that the medians from the cohorts can be considered similar. The results of the analysis for scenarios B3, B4 and B5 also indicate the same conclusion with p-values of 0.218, 0.481 and 0.481 respectively. On the contrary, Scenario B2 gave a statistical significant p-value

of 0.022 indicating the medians were not the same, which was expected due to the effect of outliers in the outcomes of all but two cohorts.

A Kruskal-Wallis test of equality for the medians within each month for Scenario B1 indicated statistical equality with p-values ranging 0.462 and 0.465. The conclusion of statistical similarity is valid for all the other scenarios. The p-values ranged from 0.463 to 0.465 (Scenario B2), 0.462 to 0.465 (Scenario B3), 0.437 to 0 0.465 (Scenarios B4 and B5). Further, all the scenarios were tested together for similarity and the Kruskal-Wallis test consistently gave p-values of 0.486 for all the scenarios (except Scenario B2 with p-value of 0.487), which also showed that the median of the daily correlation coefficients were statistically significantly similar. The tests therefore confirm that the hourly predictions compared to the forecasts are statistically not different from each other.

9.5 Analysing the Outputs of the Link Predictions

Appendix K includes a summary of the correlation coefficients for the analysed five scenarios described in section 9.3.1, which are presented as follows: Cohort 1 (Table 1K through to Table 7K), Cohort 2 (Table 8K through to Table 14K), Cohort 3 (Table 15K through to Table 21K), Cohort 4 (Table 22K through to Table 28K), Cohort 5 (Table 29 through to Table 35). Before discussing the results, it is important to describe how the correlation coefficients were categorised and applied in these assessments.

9.5.1 Categorizing the Correlation Coefficients

In order to analyse the predictions further, the correlation coefficient between the predicted forecast and actual occupancy were analysed for five categories defined as follows: "very strong" ($0.70 \leq |r| \leq 1.0$), "strong" ($0.60 \leq |r| \leq 0.69$), "moderate" ($0.40 \leq |r| \leq 0.59$), "weak" ($0.20 \leq |r| \leq 0.39$) and "no" ($-0.19 \leq |r| \leq 0.19$) (Ratner, 2009). There was no observed pattern in the distribution of correlation coefficients for the all links over the day. The correlation coefficients ranged between -1.0 and 1.0, but generally in favour of the moderate to high correlations. Consequently, at 95% confidence level, the link-by-link predictions were not significantly different within a particular day (p-values > 5%); however, there was statistically significant difference when the link-by-link predictions were compared for different days (p-values < 5%). Similarly, the monthly predictions were statistically significantly different from each other.

Table 9-4 presents a summary of the cumulative monthly predictions for the analysed cohorts for scenarios B1, B2 and B3. Scenario B2 shows that there was generally an increase in the number of "strong" correlations, compared with Scenario B1. This equates to about 38% more suggesting the short-term 15 minutes forecasts yielding better results than Scenario B1 where the next day forecast is 24 hours ahead.

	Cohorts							
	1	2	3	4	5	6	7	Average
	Jan	Feb	Mar	April	May	June	July	
Scenario B1 (model build from LE, r=0.4, lag > 100 seconds, FH[8] = 24-hours)								
No. of Days	25	26	29	28	29	28	28	28
No. of Forecast Links	817	861	965	936	963	929	914	912
	Number of Correlations:							
Very Strong	284	317	300	274	337	393	321	318
Strong	233	202	230	232	257	178	158	213
Moderate	66	51	67	68	82	50	46	61
Weak	42	61	70	76	81	41	48	60
No	192	230	298	286	206	267	341	260
Scenario B2 (model build from LE r = 0.4 and lag >100 seconds, FH = 15 minutes)								
No. of Days	27	27	30	30	29	29	29	29
No. of Forecast Links	979	969	1064	1092	1036	1028	990	1023
	Number of Correlations:							
Very Strong	702	676	779	774	732	739	780	740
Strong	155	162	163	184	160	152	96	153
Moderate	38	45	42	47	48	44	51	45
Weak	24	33	34	40	39	28	28	32
No	60	53	46	47	57	65	35	52
Scenario B3 (same as Scenario 2, but model build from occupancy data)								
No. of Days	27	27	30	30	29	29	29	29
No. of Forecast Links	902	878	1003	886	882	912	917	911
	Number of Correlations:							
Very Strong	667	648	759	667	635	699	709	683
Strong	148	165	159	155	169	155	121	153
Moderate	39	20	43	34	31	20	34	32
Weak	27	21	19	23	19	15	27	22
No	21	24	23	7	28	23	26	22

Table 9-4: Summary of the Frequency of the Categorised Monthly Output of the Daily Average Hourly Forecasts for the Network (Scenarios B1, B2 & B3)

[8] FH is the Forecast Horizon

9.5.2 Analysis of Model Outputs: Scenario B1

In order to translate the results in Table 9-4 into an understanding of the performance of the model in making daily hourly forecasts of congestion events in the network, Figure 9.9, Figure 9.10 and Figure 9.11 present the average daily proportion of the link forecasts made for scenarios B1, B2 and B3. Based on the five categories identified, Figure 9.9 indicates that the relative performance of the predictions in Scenario 1 are 29-42% ('Very Strong'), 17-29% ('Strong'), 5-9% ('Moderate'), 4-8% ('Weak') and 21-37% ('No').

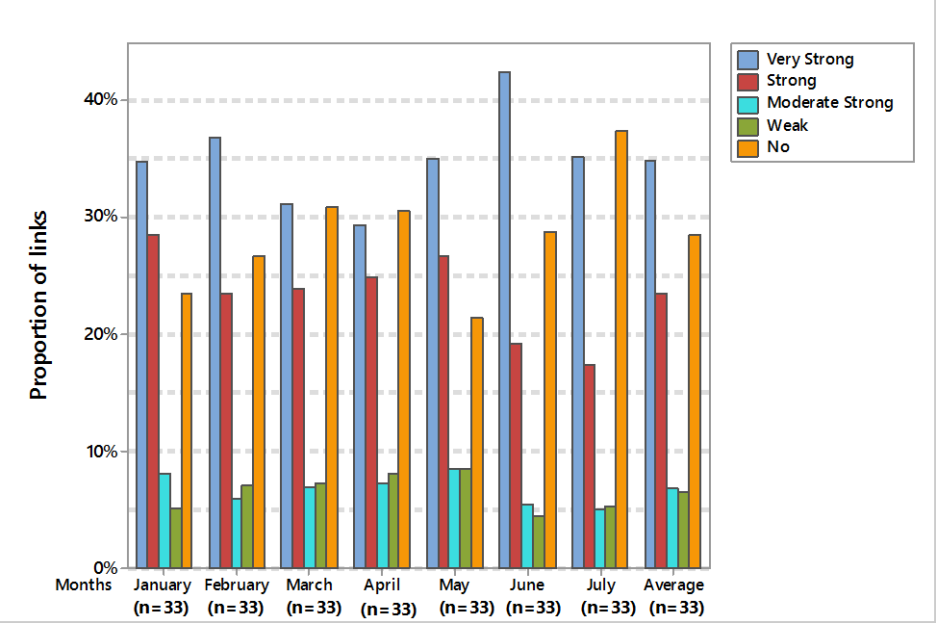


Figure 9.9: Profiles of Average Hourly Forecasts 24 Hours Ahead (Scenario B1)

The results indicate that the proportion of the 'No' category alone in Scenario B1 is substantially more than the combined category of 'moderate', 'Poor' and 'No' correlations in Scenario B2. The results therefore highlight the difficulty of forecasting the long-term (24 hours ahead) link occupancy of traffic on links in the network, given that the number of predictions with better than 'poor' correlations are up to 65%, but the proportion of 'Weak' and 'No' correlations were as high as up to 35%. The comparative figures for Scenario 2 are 92% and 8% respectively.

This result is due to the highly temporal nature of the network characteristics due to a number of factors that disturb the stability of the network's dynamic system. Such factors may occur frequently or infrequently, but importantly do not repeat themselves exactly from time to time over the course of the hour (or day). They result from variations in traffic volumes, traffic behaviour, rat-running, pedestrians' jay walking and disruption, planned

road works or unplanned minor/major network incidents (see Section 3.4.1). For this reason, there is a likelihood of a higher degree of variability in the sensitivity of the system (Lyapunov exponent) from day to day, compared with short time horizons of 15 minutes.

Due to the Lyapunov exponent being prone to this variability, the predictions of the link occupancies for one hour 24 hours in advance were subject to a high proportion of link correlations being weak or non-existent correlation (35% 'weak or 'no' correlation). The model therefore suggests that it is difficult to predict accurately the short-term link occupancy one day in advance. However, the results can be useful for strategic management of the network, though unsuitable for managing short-term variability.

9.5.3 Analysis of Model Outputs: Scenario B2

In Scenario B2 that tested the predictions for short-time forecasts of 15 minutes, the results in Figure 9.10 showed that the proportion of weak and non-existent correlations were about 6%-9% (compared to 35%). In this instance, the system's sensitivity is less susceptible to high variability in the short term, and therefore the predictions were found to be highly correlated to the actual link occupancies.

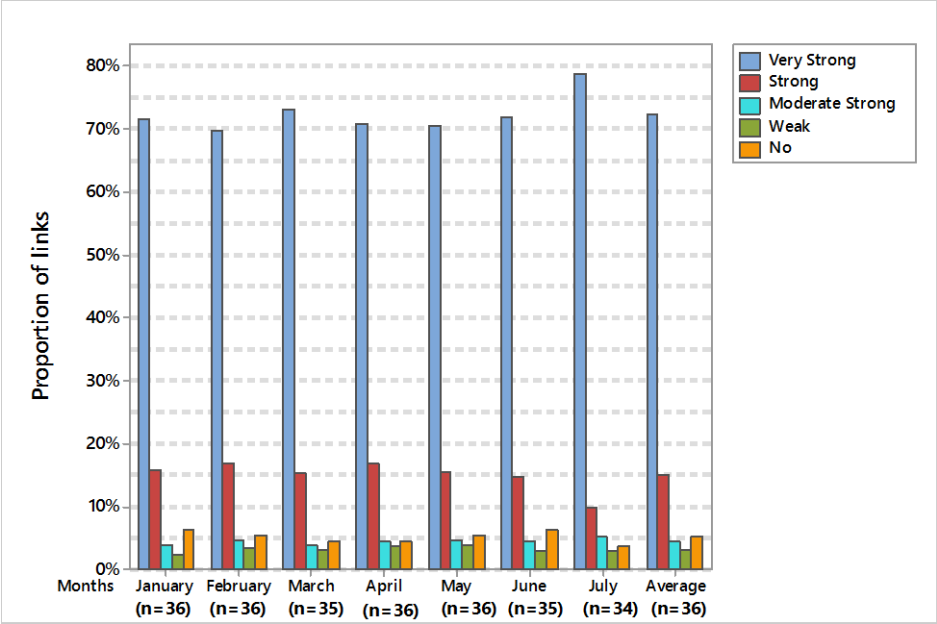


Figure 9.10: Profiles of Average Hourly Forecasts 15 minutes Ahead (Scenario B2)

On average, about 91% of the links obtained moderate correlations or better. It is important to note that a lag criteria (forecast horizon) of 15 minutes was used in selecting links for the forecast model. Therefore, it is in line with the expectation that short-term 15 minute (close to the average lag) forecasts were correlated more strongly than the long-term 24-

hour forecasts. Further work may be needed to assess the relationship between the chosen lag and the duration of forecasts.

Reflecting on the results for Scenario B2 in Table 9-4 based on the total monthly correlations gave the proportions for the five categories in Table 9-5. On average, the results in Table 9-5 are similar to Figure 9.10 (based on the results for the daily averages). Overall, there was a likelihood of a "moderate" to "very strong" correlation of 92%, with "weak" correlation" of 3% and "no" correlation of 5%.

	Cohorts							
	1	2	3	4	5	6	7	Average
	Jan	Feb	Mar	April	May	June	July	
Very Strong	72%	70%	73%	71%	71%	72%	79%	72%
Strong	16%	17%	15%	17%	15%	15%	10%	15%
Moderate	4%	5%	4%	4%	5%	4%	5%	4%
Weak	2%	3%	3%	4%	4%	3%	3%	3%
No	6%	5%	4%	4%	6%	6%	4%	5%

Table 9-5: Summary of the Proportions of the Categorised Correlation Coefficients within the Five Categories (Scenarios B2)

A comparison of Figure 9.9 and Figure 9.10 illustrates the deficiencies in the model results of the two scenarios. The results indicate that the degree of correlation in Scenario B2 is higher than Scenario B1. In line with expectations, there were better correlations for the short term compared to the long term duration interval and on average 90% (an increase of about 30% compared to Scenario B1) of the links indicated moderate to high correlations with actual observations. On average, only 3-5% of the links showed predictions that were uncorrelated with the actual observations.

9.5.4 Analysis of Model Outputs: Scenario B3

In order to demonstrate the reliability of the short-term forecasts, Scenario B3 repeats the analysis from Scenario B2, but instead the link occupancies were predicted from occupancy data (rather than the Lyapunov exponents). Figure 9.11 indicates as 3-6% of the links were weakly correlated or had no correlation. The distribution of the correlation coefficients were 72%-77% ('very strong'), 13%-19% ('strong'), 2%-4% ('moderate'), 2%-3% ('weak') and 1%-3% ('no'). This was confirmed by the Mann-Whitney test, which revealed that the number of predictions in Scenario B3 was statistically significantly greater than Scenario B2 for the 'very strong' and 'strong' categories, and *vice-versa* for the

'moderate', 'poor' and 'no' categories. On a month-by-month comparison the actual difference in the categories of Scenarios B2 and B3 is less than two links. The highest difference of approximately 5% was recorded in June, which equates to about two additional 'very strong' forecasts in terms of actual counts. Therefore, although the results are statistically significantly different, the differences between the two scenarios in terms of actual numbers are small, and therefore the outputs from the two forecasts agree well.

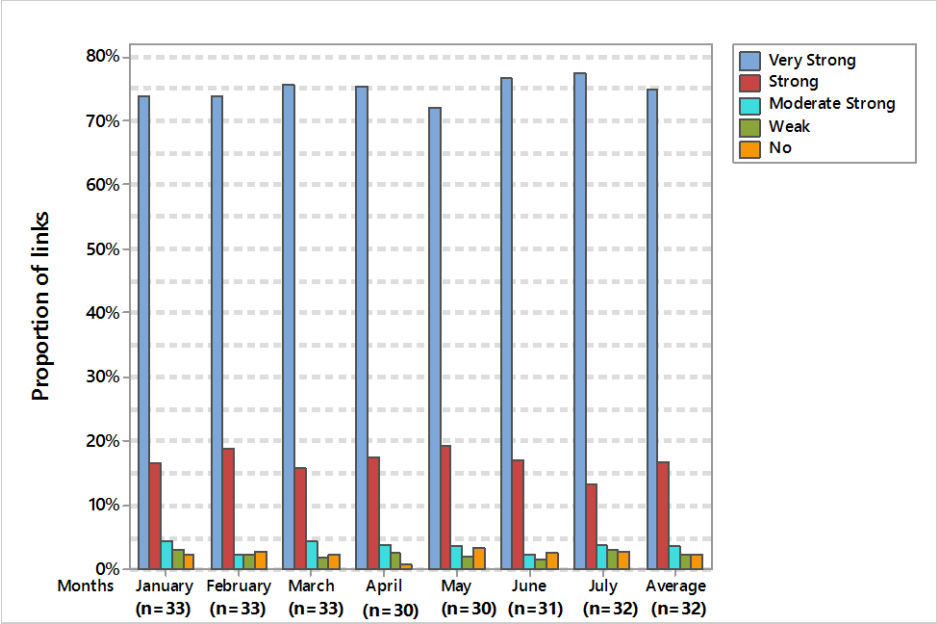


Figure 9.11: Profiles of Average Hourly Forecasts 15 Minutes Ahead (Scenario B3)

However, it is important to note advantages of using the Lyapunov exponents instead of the occupancy data that emerge from the data. Spatially, and over all the seven cohorts analysed, there were a total of 7158 links forecasted the onset of congestion for Scenario B2 whilst there were 6380 in Scenario B3. This difference is attributable to the high failure rate of the occupancy data, against the adopted criteria based on the degree of correlation coefficient and the lag time, for choosing the selected links. This means that the forecasts from the Lyapunov exponents (rather than occupancy) make it possible to report on the impending conditions over a wider part of the network and with longer time horizon.

Overall, the results for Scenario B3 confirm that the short-term forecasts of link occupancy based on Lyapunov exponent (Scenario B2) with a reasonable lag time are better. Therefore, the short-term predictive capability of the Lyapunov exponents could be exploited for traffic management in urban areas. This has the potential to overcome existing limitations of current demand responsive systems that only address network

problems ***as*** or ***after*** they occur. The results suggested that incorporating the forecasting algorithms into existing systems would provide additional intelligence to these systems making them pre-emptive to anticipate congestion and therefore implement mitigating strategies in advance.

9.5.5 Analysis of Model Outputs: Scenarios B4 and B5 (Sensitivity Testing)

The results discussed above indicated that the predictions of the link occupancies from the Lyapunov exponents are a better approach, compared with occupancy itself, to forecast congestion. The next step is to assess the robustness of the model to parameters that govern the results. For example, the threshold correlation coefficient value of 0.4 is the assumed optimum however, the model performance may be improved by varying this value. Table 9-6 presents a summary of the total number of forecast for each month for the analysed cohorts for Scenarios B4 and B5 (sensitivity tests), where the assumed correlation coefficients are 0.7 and 0.3 respectively. The former includes only 'very strong' correlations, whilst the latter incorporates 'weak', 'moderate', 'strong' and 'very strong' correlations.

	Cohorts							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**Average**
	Jan	**Feb**	**Mar**	**April**	**May**	**June**	**July**	
Scenario B4 (same as Scenario 2, but r = 0.7)								
No. of Days	27	27	30	30	29	29	29	29
No. of Forecast Links	198	606	740	701	716	706	577	606
	Number of Correlations:							
Very Strong	142	476	609	546	583	549	449	479
Strong	36	81	100	103	99	101	78	85
Moderate	9	10	10	16	15	16	17	13
Weak	6	13	10	16	17	16	31	16
No	5	26	11	20	2	24	2	13
Scenario B5 (same as Scenario 2, but r = 0.3)								
No. of Days	27	27	30	30	29	29	29	29
No. of Forecast Links	403	992	1090	1073	1015	1040	998	944
	Number of Correlations:							
Very Strong	255	607	709	654	634	676	715	607
Strong	74	165	163	183	194	174	108	152
Moderate	23	55	49	76	46	52	36	48
Weak	19	73	64	58	51	51	46	52
No	32	92	105	102	90	87	93	86

Table 9-6: Summary of the Frequency of the Categorised Monthly Output of the Daily Average Hourly Forecasts for the Network (Scenarios B4 and B5)

The results indicate that compared to Scenario B2, the numbers of the predicted links decreases significantly in Scenario B4 when the threshold correlation coefficient increased to 0.7 as expected (see Section 9.5.3). The Mann-Whitney test indicated at 95% significance level that the number of forecast links in Scenario B2 is statistically significantly greater than Scenario B4, the result indicating a significant p-value of 0.0005.

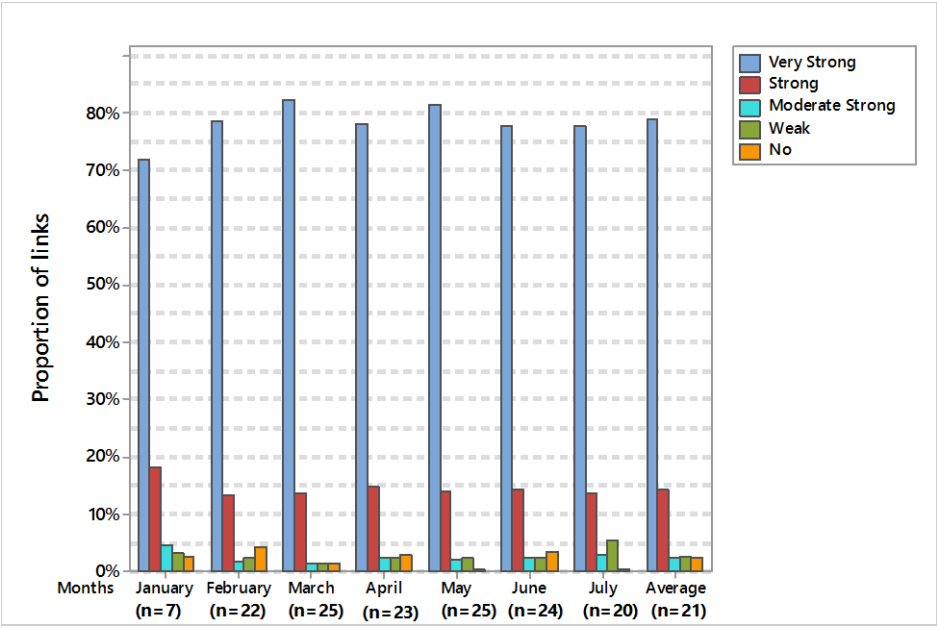


Figure 9.12: Profiles of Average Hourly Forecasts 15 Minutes Ahead (Scenario B4)

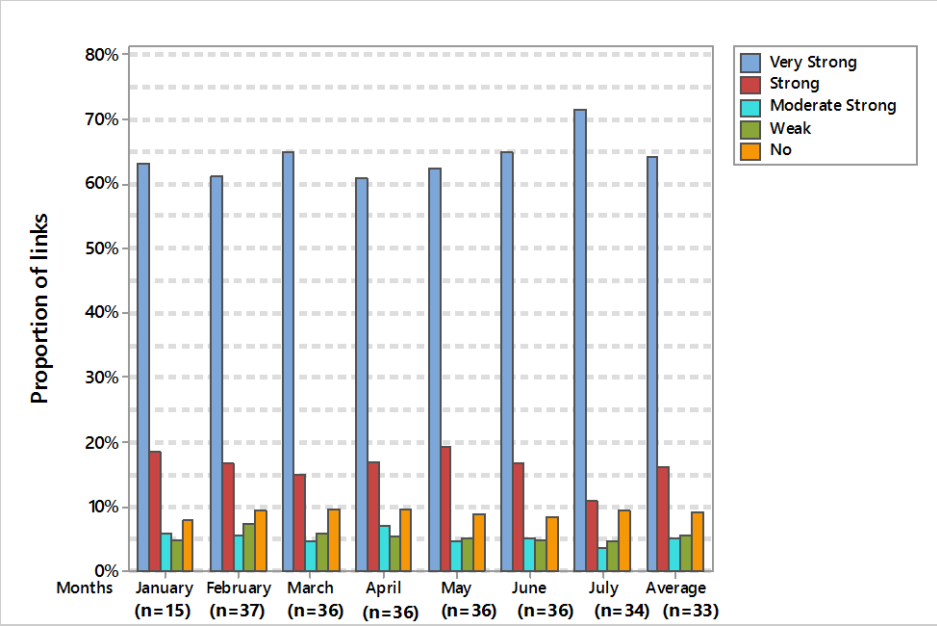


Figure 9.13: Profiles of Average Hourly Forecasts 15 Minutes Ahead (Scenario B5)

This means that there is a high failure rate of the Lyapunov exponents profile when screened against the correlation coefficient of 0.7, mainly because of the filtering out of the

data at the onset, which limits its ability to predict. As was mentioned previously (in Section 9.5.4), a major disadvantage of this result is that, by considering only those link pairs with 'very strong' correlations, it inhibits the ability of the exponents to report on the situational dynamic state over a wide spatial area of the urban network.

Turning now to the results of Scenario B5 with correlation coefficient threshold of 0.3 presented in Table 9-6, the Mann-Whitney test indicated that the number of predictions in scenarios B2 and B5 were statistically similar given a p-value of 0.3964. The results therefore indicate that reducing the threshold correlation coefficient to 0.3 (in Scenario B5) does not give any significance to the results from the forecast. This suggests the 'weak' correlation do not make any significant impact in this respect. In terms of actual counts, the total number of average monthly predictions in Scenario B2 was 1023, which is more than the 944 (in Scenario B5).

Figure 9.12 and Figure 9.13 present the average hourly profiles for Scenario B4 (with threshold of 0.7) and Scenario B5 (with threshold of 0.3) respectively. The results indicate that the average hourly predictions for Scenario B4 are 95% above "weak" with only 5% below "weak" correlations. The corresponding figures for Scenario B5 are 85% and 15% respectively. This suggests that reducing the criteria for selecting qualifying links does not necessarily improve the forecasting function. It is important to note that Scenario B2 (with respective figures of 92% and 8%) provides a higher quality of predictions than Scenario B4, even though there is statistically no difference between the two scenarios in terms of the actual number of forecasts. Finally, in terms of actual counts and reporting on the spatial network impacts, the results indicates Scenario B2 is more suitable.

9.6 Exploring the Lyapunov Exponent for Traffic Management

Apart from the forecasts, there was an investigation of the uses of the Lyapunov exponents for understanding the spatial interactions between the network links in order to explore how the Lyapunov exponents may be beneficial for managing the road network traffic conditions. In order to improve traffic flow, it is common for traffic operators and managers to monitor the network using Close Circuit Television (CCTV) surveillance and apply remedial strategies in the vicinity of crucial bottlenecks to maintain traffic conditions at below saturated levels. This section explores the use of the Lyapunov exponents for identifying critical links in a network in advance of them experiencing congestion by

investigating the spatial and temporal correlation across the network of the time series Lyapunov exponents of all the links. The methodology was set out in Section 6.9.

Given that the Lyapunov exponents are an indication of the nature of vehicular interaction, the spatial correlation of its value for all the links helps to understand how the dynamic behaviour of a link at time t (s) is influenced by other links at a time in the past t $(s-\tau)$, where τ is the lagged time interval. Therefore, through this investigation, the 'critical' link is identified as a link whose future dynamic behaviour is influenced by significant dynamic behaviour in the past of traffic 'feeder' links (which may or not be physically connected to the critical link). This method is similar to the assessment carried out in Section 9.3.1 and is based on the cross-correlation of the Lyapunov exponent – Lyapunov exponent link interaction for every hour of the day at a sampling interval of 20 seconds. For every link, the cross-correlation is for the link's profile against all the feeder links for every hour of the day. The feeder links with correlation coefficients greater than a threshold of 0.4 and the associated lag times were selected for further analysis. Then, for every link, the frequency of feeder links above the correlation threshold of 0.4 was computed. For each hour, the link with the highest number of feeder links was identified as the critical link for the period, and the associated lag time provide a window of opportunity for the mitigation of imminent network problems.

Table 9-7 presents a summary of the result for the 0800h-0900h peak for 24 days in January 1999. Table 9-7 suggests that, the critical links and their associated feeder links vary from day-to-day for the 0800h-0900h period, and also from hour-to-hour within the same day. The variation in what are the critical links and their feeder links is widespread. It is pertinent to note that this also confirmed the similar observation in Section 9.4 in the selection of links for forecasting. Therefore, the results do confirm that the dynamic behaviour of the traffic system does not repeat itself exactly due to initial conditions changing continuously. Due to the variability in the traffic conditions on those links which are critical at a particular point in time, and which are the feeder links, makes long-term forecasts based on (for example) annual growth factors unreliable. This suggests that it may be impossible for traffic behaviour on a specific critical link to be linked consistently with a particular set of other links due to the inconsistent patterns in the evolution of traffic

flow. However, this is not necessarily a problem, because the dynamic process is re-assessed continually as time progresses and 'learns' this building intelligence over time.

Day	Critical Link IDs	Feeder Links (correlation Coefficient >= 0.4)			
		Frequency	Links IDs	Correlation Coefficients	Lags (minutes) [9]
1	37	3	10, 21, 38	0.52, 0.42, 0.45	-1.3, -4.7, -5
4	18	2	20,29	0.52, 0.41	-9.7, -7
5	10	3	19, 21, 32	0.41, 0.42, 0.43	-3.3, -15.67, -9
6	31	4	23, 26, 30, 36	0.42, 0.41, 0.44, 0.49	-15.7, -1, -0.7, -5.3
7	8	3	4, 12, 23	0.44, 0.45, 0.46	-2.7, -13.7, -4
9	33	3	3, 5, 29	0.50, 0.42, 0.46	-1, -4, -3.7
10	40	4	20, 21, 28, 30	0.41, 0.42, 0.41, 0.42	-9.7, -2.3, -4.3, -6.7
11	38	2	2, 21	0.46, 0.63	-13, -4.3
12	30	3	3, 21, 28	0.40, 0.40, 0.42	-14, -5, -16.7
13	19	3	25, 29, 34	0.41, 0.43, 0.44	-11.3, -6.7, -15
14	9	3	1, 10, 31	0.43, 0.56, 0.44	-12.3, -1.3, -16.7
15	23	4	3, 18, 33, 36	0.91, 0.45, 0.57, 0.77	-12, -8.7, -12.3, -0.3
16	1	4	5, 6, 12, 26	0.84, 0.44, 0.77, 0.48	-8.3, -12.3, -12, -16.7
17	29	2	2, 21	0.41, 0.56	-7, -16.3
19	10	11	1, 8, 19, 21, 22, 23,29, 30, 31, 35, 40	0.61, 0.79, 0.54, 0.67, 0.68, 0.60, 0.49, 0.67, 0.72, 0.77, 0.79	-14.3, -14.3, -14.3, -12.3, -14.3, -10, -14.3, -14.3, -14.3, -14.3, -14.3
20	39	6	7, 19, 33, 36, 38, 39	0.54, 0.44, 0.56, 0.56, 0.52, 0.57	-10.7, -12.3, -6.7,-12.7,-12.7, -12.3
21	33	4	4, 5, 32, 37	0.48, 0.44, 0.94, 0.54	-1.7, -8.3, -14, -16
22	33	6	1, 4, 20, 38, 39, 40	0.74, 0.49, 0.52, 0.77, 0.50, 0.49	-9, -1.7, -0.3, -1.3, -2, -9
23	30	6	5, 7, 19, 22, 28, 39	0.43, 0.43, 0.43, 0.40, 0.42, 0.45	-6, -13.7, -10.3, -8.3, -5.3, -5.7
24	1	4	3, 21, 29, 36	0.43, 0.40, 0.40, 0.66	-3.3, -16.7, -15, -10.7
25	7	4	19, 23, 26, 39	0.43, 0.40, 0.40, 0.41	-2.7, -6.7, -2, -0.3
26	9	3	4, 21, 27	0.52, 0.45, 0.49	-2.7, -16, -16.3
27	30	6	3, 23, 33, 34, 35, 38	0.49, 0.43, 0.42, 0.48, 0.45, 0.43	-0.3, -3.7, -6.7, -3.7, -15.7, -0.3
28	23	5	3, 19, 21, 26, 31	0.48, 0.45, 0.60, 0.45, 0.46	-6, -13.7, -4.7, -11.7, -15.7

Table 9-7: Summary of Cross-Correlation Analysis for January 1999

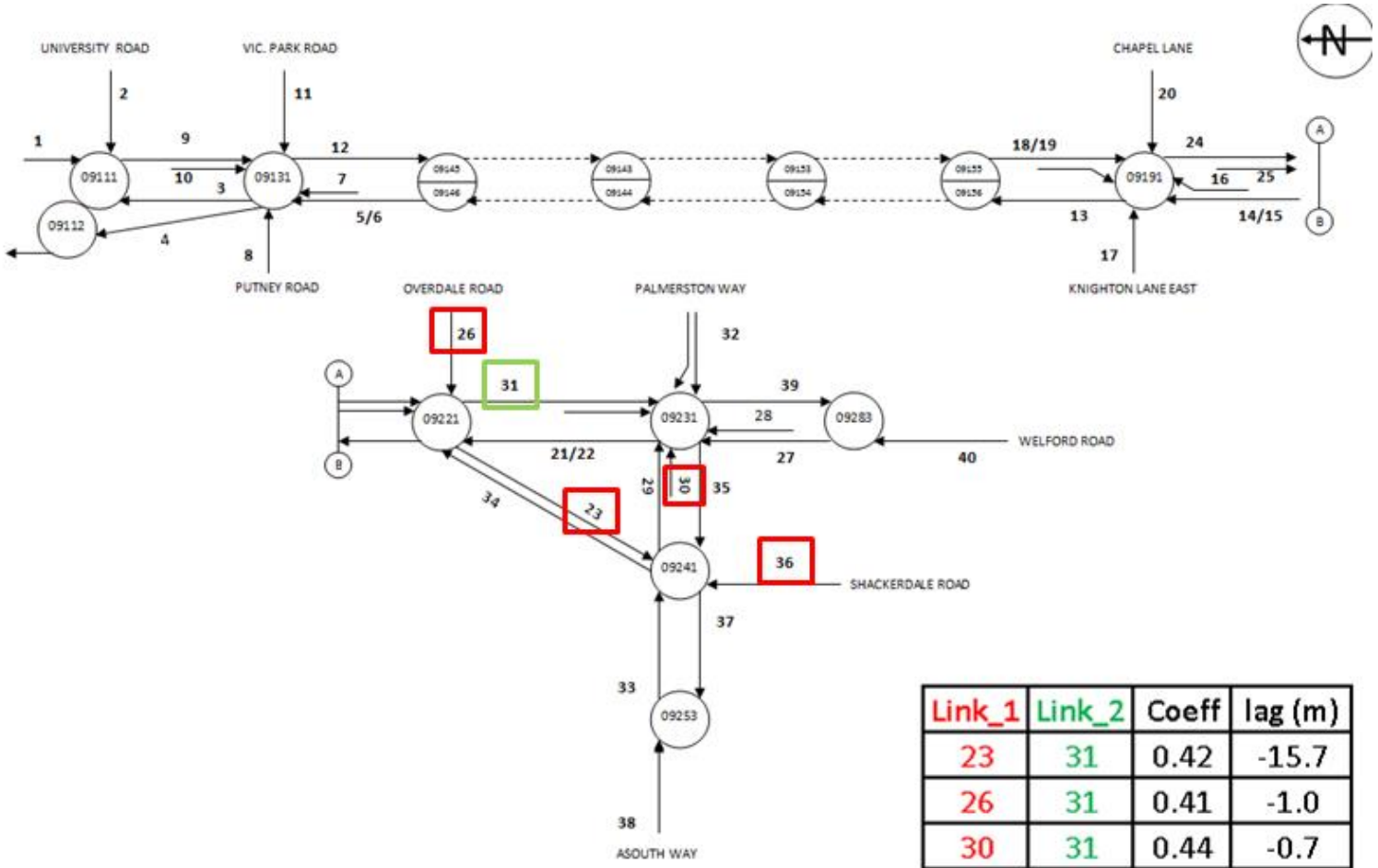
[9] Negative lag indicate that the computation of the correlation coefficient is between the first variable and the past time series of the second variable.

The analysis also shows different lags associated with the critical links throughout the 24 days. The computed lags indicate the time interval that the dynamic behaviour of a 'feeder' link maximises its influence on a critical link and governs the response time available for operators and managers to implement appropriate alleviation measures.

Therefore, by selecting the relevant lag of those feeder links associated with the critical link provides a window of opportunity to implement remedial measures in order to address impending network problems in advance. The longer the lags the longer the time available to the traffic manager to take action on how to keep the network operating below saturated levels. However, the longer the lag times, the greater the uncertainty in the accuracy of the forecast. Overall, the results indicated an average lag of 8-10 minutes as optimum, which was entirely consistent with the average lag computed in Section 9.3.1, which was a maximum of approximately 8 minutes. Given that the select forecast horizon of 15 minutes in Section 9.5 was robust, the results will be even better with a lower forecast horizon.

The link layout plan (Figure 9.14) illustrates the interconnectivity between the critical and feeder links for the assessment for 0800h-0900h on the 6th of January 1999. Figure 9.14 shows that the critical link "31" is correlated with four upstream links (23, 26, 30 and 36). The correlation coefficients for the links are 0.42, 0.41, 0.44 and 0.49, and lags -15.7, -1.0, 0.7 and 0.53 minutes respectively. Link "23" has the highest lag with link "31", and therefore gives a window of opportunity of up to 15.7 minutes in advance to mitigate conditions at links "23", "26", "30" and "36".

In this example, whilst the highest lag is 15.7 minutes, the other lags are relatively small and therefore the choice of any of these low values do not allow as much time to implement remedial measures, even though an action designed at link 23 has potential to resolve the situation at links 26, 30 and 36 in a timely manner. Therefore, the traffic manager may find it appropriate to make decision based on the highest lag; however, the averaging of the lags may be reasonable assumption if the lags are not significantly different. The challenge is to get the right balance between a significantly high enough correlation and enough time to implement effective remedial action. The choice of the lag may therefore need further investigations to determine the optimum value to traffic management decisions.



Link_1	Link_2	Coeff	lag (m)
23	31	0.42	-15.7
26	31	0.41	-1.0
30	31	0.44	-0.7
36	31	0.49	-5.3

Figure 9.14: Link Interactions in SCOOT Region KA Welford Road (0800h-0900h)

9.7 Conclusions

In previous chapters, it was identified that Chaos Theory can be used to identify the onset of congestion in UTC network regions. By analysing individual links, Chaos Theory provides strong indication that the Lyapunov exponents can be used as a measure to expose the shoulders of congestion in a SCOOT controlled region in Leicester. Further, chaos is not an isolated incidence at an individual junction but has spatial effects across the independent links on the entire network of interconnected junctions. It was estimated that during the busy daytime period, more than 10% of the links will indicate positive Lyapunov exponents before the effects of chaos are fully-developed within the network region. Consequently, this chapter examined Lyapunov exponents in combination with adjacent network links to improve short-term traffic forecasts through optimising neighbouring links together in order to harmonise their operations.

It is recognised that there are a number of processes in the urban network that tends to impact on the accuracy of forecasts of traffic volumes or occupancy across a loop detector or a section of the road network. These include changes in a single driver's behaviour, effects of the movement of emergency vehicles, changes in weather conditions, road closures, diversions, rat-running, unusual conditions such as major sporting events. Therefore, in order to make accurate predictions, it is necessary to consider the relationship of a link with both downstream and upstream links across the entire network. This is important as the dynamic behaviour of traffic flow on a link is a result of the temporal variability in interactions between the interconnected links. Therefore, regression models that seek to undertake predictions of a link based on only specific links are likely to compromise on the dynamic evolution of the road network and consequently affect the forecasting accuracy. Consequently, Chapter 9 demonstrated that a dynamic regression model that selects links for the model construction based on a cross-correlation assessment of all links in the network can forecast the occupancy of the network links with high certainty. Table 9-4 shows that out of the 7158 links tested, the breakdown of the number of forecasts for "very strong", "strong", "medium", "weak" and "no" correlations were 5182, 1072, 315, 226 and 363 respectively. Therefore, 6569 links representing approximately 92% of links indicated at least a moderate correlation between the actual and forecast occupancy.

From a traffic management perspective, the inference from a cross-correlation exercise of the Lyapunov exponents of links in a SCOOT region suggests that Chaos Theory is suitable

for managing traffic in urban network. Section 9.3 indicated that the lags between the Lyapunov exponents of all links can enable identifying a time window of approximately 8-10 minutes (on average) as a trigger for implementing strategies ahead of actual time of occurrence of congestion in order to ameliorate imminent saturated traffic conditions. Currently, network managers optimise a set of parameters and conditions, or continuously learn and apply corrections, to control parameters at the time of implementing strategies to manage transport networks and systems. Therefore, the incorporation of chaos-based algorithms into current systems would be beneficial towards autonomic applications for traffic systems, which require that systems are able to self-manage their operations through various capabilities with minimal human support, thus minimising the role of network managers (see Section 3.8). It is important to note therefore that Chaos Theory could play a role in this step change requirement for traffic management, as the debate for autonomic traffic control continues to gain momentum among academics and traffic practitioners alike.

Chapter 9 concluded chaos-based algorithms will enable enacting appropriate traffic control and management strategies that are "one step ahead" rather than "one step behind" of prevailing traffic conditions. This would radically improve the management of traffic on a strategic level rather than purely locally within smaller network regions. Thus playing an important role in improving journey times and air quality and making a vital contribution to mitigating climate change.

PART IV: DISCUSSIONS, CONCLUSIONS AND RECOMMENDATIONS

Chapter 10. Discussions, Conclusions and Recommendations

10.1 Introduction

This chapter is devoted to discussing the findings of the analysis in the context of the existing literature, in order to identify how this work relates to the knowledge established in previous investigations of the application of Chaos Theory in traffic studies. It highlights principally the contribution this research work makes to knowledge and understanding of whether road traffic exhibit properties of a chaotic system. Available literature shows that the study of Chaos Theory, and its application in traffic studies, is still at the fundamental stage. The majority of previous works have mainly tested traffic data such as flow (Lan *et al*., 2003, Zhang and Liu, 2007, Frazier and Kockelman, 2004, Xue and Shi, 2008, Jianming *et al*., 2003, Krese and Govekar, 2013), occupancy (Nair *et al*., 2001) and speed (Shang *et al*., 2007) in order to identify the presence or absence of chaos in these datasets. Frazier and Kockelman (2004), Hui *et al*. (2005), Jianming *et al*. (2003), Xue and Shi (2008) and Zhu *et al*. (2010) applied Chaos Theory in combination with standard techniques for forecasting purposes. However, only a few studies have explored Chaos Theory as a forecasting technique of traffic conditions in urban networks. Other than confirming that datasets of road traffic variables exhibit chaotic properties, the majority of studies have not examined traffic data in detail in order to explore the benefits of the practical application of Chaos Theory in traffic management such as forecasting in advance the onset of congestion. This study, therefore, looked at traffic variables such as flow, occupancy and noise, and examined how these variables help to answer the research questions enumerated in Section 1.5 through novel analysis. This research demonstrates that traffic flowing in urban signalised networks has chaotic properties, and concludes on how Chaos Theory may be useful for traffic management in urban areas.

In order to present the discussion, conclusion and recommendations for future work, this concluding chapter of the thesis is set out as follows. Section 10.2 presents the discussions of the results under six main themes: (i) data requirements (ii) lag times (iii) dimensions (iv) Lyapunov exponents (v) forecast horizon and (vi) spatial-temporal analysis to forecast the onset of unstable traffic predictions. It argues that whilst chaos may be an inherent property in time dependent non-linear systems, in order to derive benefits other than the mere detection of chaos, the data must be sampled at a sufficiently high temporal resolution. Furthermore, it argues that the Lyapunov exponents are very suitable for understanding the spatial relationship between network links and incorporating this knowledge within

existing traffic control systems will enable initiating control measures that are one-step ahead of congestion. Section 10.3 presents a summary of the limitations of the research work and Section 10.5 the overall conclusions of this study, and demonstrates the importance of the research, and how the results of the study help to answer the research questions enumerated in Section 1.5. Finally, some recommendations for future studies are discussed in Section 10.6.

10.2 Discussions of the Results

The results demonstrated the potential role for Chaos Theory in traffic management. The Lyapunov exponent is the principal parameter developed for monitoring and understanding the evolution of traffic states prevalent in networks. The lag (delay time) and embedding dimensions, which serve as input to the process of reconstructing the phase space need to be pre-determined before the profile of Lyapunov exponents can be derived. The discussion in this section focusses on these main parameters in addition to data issues, the forecast horizon and the analysis of the temporal and spatial correlations of the datasets.

10.2.1 Data Requirements

The literature review identified in previous research that significant variability exists in the sampling interval of data used for detecting chaos. For example, for speed five-minutes (Nair *et al*., 2001) and two and half minutes (Shang *et al*., 2007); for flow data one-minute interval (Lan *et al*., 2003), 15 minutes (Jianming *et al*., 2003), 30 seconds (Frazier and Kockelman, 2004) and ranging between 30 seconds and six minutes (Zhang and Liu, 2007). The research results carried out in this thesis endorsed these earlier observations because chaos was detectable at varied degrees of sampling. However, understanding the development and propagation of chaos in the traffic system was found not to be possible with data of either too high or low resolution (e.g. 1, 30, 60 and 120 seconds). Therefore, for practical application of Chaos Theory rather than simply determining the chaos status of the data requires highly resolved temporal traffic information.

On the suggestion of Frazier and Kockelman (2004), this research analysed three types of traffic related data (flow, occupancy and noise) at a wide range of sampling rates demonstrating that the urban road traffic system is a chaotic system. The results of the research shows that identifying chaos from a given dataset may be independent of the sampling interval, but the choice of sampling interval is crucial to the understanding of the propagation of chaos in the system.

The chaos status of traffic noise is unknown, according to the literature reviewed, and therefore this research represents a first attempt at discovering this. This work found that traffic noise is chaotic, but because the data was available only at a 60-seconds sampling frequency a detailed assessment was not possible. However, for the reconstructed flow and occupancy data, chaos was observed at 10, 30, 60 and 120 seconds. However, the profiles of the Lyapunov exponents did not produce any underlying structure to enable examining this data for practical uses except at 20 second resolution. The highly resolved 10 second data did not prove to be suitable consistent with the view of Kirby and Smith (1991) that stated that the data at the appropriate sampling frequency was essential if the transition to chaos was to be readily and correctly detected. This research also showed that the flow and occupancy data revealed consistency in results as expected because these variables are directly proportional at low values (see Equation 2-32) (Athol, 1965). Therefore, the outcomes of the analysis of the two datasets collected at the same location and sampling interval would reveal similar patterns.

There is currently no specific guidelines on how to select an appropriate sampling interval. However, this work suggested that the best approach is to carry out an initial assessment based on varying sample frequency to identify the optimum sampling frequency for analysing the chaotic behaviour in non-linear systems. Based on the analysis conducted, typically sampling at reasonably high frequency of 20 seconds is needed to harness the properties of traffic flow data to detect the transition to chaos. Therefore, this research shows that a highly sampled data can overcome the limitations of previous studies. However, the study cautions that after a certain optimum level of temporal resolution, there is no improvement (or even a deterioration) in the performance of the algorithm for more finely resolved times series. There is therefore a need for caution in the choice of the temporal resolution of the data that will be suitable for revealing chaotic properties in the systems.

10.2.2 Delay Time (Autocorrelation Lags)

The autocorrelation lag indicates the time interval that the time series data points correlates with itself. The autocorrelation method gave lag values ranging between 140 and 270 minutes for the links within the network studied (see Section 8.3), which shows a global disparity in the lag times across the entire network. The large number of independent lags computed for the 40 links in the network presents a reasonably wide range of autocorrelation lags for comparison with previous work. However, due to the different sampling frequency and the length of time series data, statistical comparison was

inappropriate. Similar to this research, Jianming *et al*. (2003) applied the autocorrelation method to time series of traffic flow data sampled at 15 minute intervals over a duration of one-month, which gave an estimated lag time of 210 minutes for the time series information. An alternative approach based on the mutual information method applied by Frazier and Kockelman (2004) gave a lag time of 175 minutes but for traffic flow data sampled at 5 minute intervals for a 5 week time series sample. In spite of the unequal sampling interval of data, these results based on the autocorrelation coefficient or mutual information method respectively are within the range of the lags computed in this research. The results may also suggest some credibility in the assertion of Frede and Mazzega (1999) about the equivalence of the results from the autocorrelation and mutual information methods. However, this may be possible only when both methods are used to test the same datasets because a combination of factors affect the lags including the sampling interval, length of the datasets and type of data. The study of Krese and Govekar (2013) confirm these factors by revealing lag times of between 300 and 360 minutes in traffic flow data sampled hourly for three years on a ring-road and motorway (in Ljubljana, Slovenia). The higher lag values from this latter study are likely to be due to inherent properties of the data with hourly sampling frequency and the long duration of the data collection.

10.2.3 Dimensions

The embedding dimension for flow and occupancy ranged between 2 to 5, and correlation dimension between 0.22 to 0.35. These figures satisfy the Takens' criterion, which is given by $M \geq 2D + 1$, where M and D are the embedding and correlation dimension respectively. However, the figures from the research are low compared with the existing literature. Krese and Govekar (2013) estimated an embedding dimension of 6 (motorway) and 8 (ring-road) in Ljubljana, Slovenia. This is the only literature that actually interprets the variables in terms of traffic flows. The authors suggested that 6 and 8 variables can describe the dynamics of the motorway and ring-road respectively. The differences being due to the fact that the ring-road, compared with the motorway, has relatively a higher number of junctions offering an ease of connectivity to the spatial network. In principle, urban networks are more complicated in terms of design than ring-roads and motorways, and traffic movement may be severely restricted to a few links in the network, depending on the traffic management in operation. Interpreting the results in line with the thoughts of Krese and Govekar (2013) suggests that the relatively low embedding dimensions emerging from this research is due to much less opportunity to divert away from the study area in the event of network incidents. On the contrary, the work of Jianming *et al*. (2003)

yielded embedding dimension of 8 and correlation dimension of 2.7 ± 0.2 for traffic flow sampled at 15 minutes intervals from UTC/SCOOT system in Beijing, China, which is comparatively higher than this research. Similarly, Xue and Shi (2008) estimated an embedding dimension of 8 and correlation dimension of 3.4 ± 0.2 from traffic flow data sampled at five minute intervals over a three-month period in Dongguan, China. The previous two studies and this research used data from an urban area, unlike the ring-road and motorway data used by Krese and Govekar (2013). However, the result of this research appear in conflict with that of the others. The relatively low dimensions from this research were therefore more likely due, among other factors, to the high temporal resolution of the traffic flow data compared with the previous studies.

10.2.4 Lyapunov Exponents

In relation to traffic studies, the Lyapunov exponents may be considered as an indicator of the nature of the lengthening or shortening of the gaps between vehicles in the road traffic network. This research is unique in its approach to using the instantaneous Lyapunov exponents to explain the changing traffic conditions on an urban link in a SCOOT network. The research adopts the established criteria that negative, zero and positive values of the exponent respectively are indicative of the traffic status as stable, meta-stable and chaos. It is useful to draw attention to the fact that when interpreting the Lyapunov exponents the idealised meta-stable state is assigned the value of zero. However, in order to apply Chaos Theory in practice, it may be necessary to adjust the theoretical reference line to match the duration of the observed traffic condition.

Overall, the values of the exponents ranged between $-0.234*10^{-1}$ and $+0.0578*10^{-1}$, which are generally higher than the values of $0.0353 \pm 6*10^{-4}$ (ring-road) and $0.139 \pm 6*10^{-4}$ (highway) reported by Krese and Govekar (2013) for the interurban networks. The results thus indicate that the traffic in the urban network studied in this research is more sensitive to small perturbations than a ring road and highway, which are networks with more capacity and typically can more easily dissipate congestion and minor network incidents. Previous research mainly used the maximum Lyapunov exponent as the indicator of chaos, without close examination of the implications in the context of the evolution of the traffic system.

This research concluded that there is a maximum of about 30% links (or a minimum of 10%) exhibiting chaos at the busiest times of the day. This information is crucial for urban traffic management; the network manager can take advantage of this knowledge to inform the design of remedial measures to implement in advance of impending congestion. The

research thus provided a potential tool for network management suggesting that by monitoring the progression of the states of the Lyapunov exponent profiles of the individual links, it will be possible to identify the critical links in the network for timely implementation of remedial measures to inform the network operation to keep traffic flow below saturated levels. Furthermore, the research also proved the Lyapunov exponents to be an independent variable within a spatial-temporal multiple regression model to forecast successfully the link occupancy.

10.2.5 Forecast Horizons from the Cross-correlation Lags

In this study, the method for calculating the maximal prediction length T_m was dependent on the lag time that corresponds with the highest cross-correlation coefficient. The analysis carried out indicated that a forecast horizon of about 8 minutes was good reference; and the modelling results based on the 15 minute forecast horizon gave satisfactory results. However, existing literature has indicated that this value was the reciprocal of the highest Lyapunov exponent λ_1, $T_m = {}^{1}/_{\lambda_1}$ (Kant'z method). Shang *et al*. (2005) suggested that a 20 second raw data feed of flow, speed and occupancy aggregated into 2 minute data gave good predictions within a window of 9-12 minutes, even though the reciprocal formula gave a maximum window of 8 minutes. On the other hand, the study by Zhang and Liu (2007) proved that the maximal prediction lengths increase proportionally as the time window of the aggregated data. From their analysis, time windows of 30, 45, 60, 120, 180, 240, 300 and 360 seconds for the aggregated data gave corresponding maximal predicted lengths of 13.5, 15.0, 15.5, 17.5, 17.5, 18.5, 19.0 and 23.0 minutes respectively. Extrapolating this profile indicates that the maximal prediction window for a 20-second interval data is approximately 12.5 minutes. The results therefore suggest that the reference forecast horizon estimated from the method used in this research is consistent with the Kantz', and the actual forecast horizon was consistent with Shang *et al*. (2005) and Zhang and Liu (2007). Kantz' method is generally limited to time series with low noise and therefore unsuitable for data with a strong noise component, and the process of determining the maximal Lyapunov exponent is quite laborious. Therefore, this research proposed a simplified method based on the cross-correlation assessments as an alternative approach for calculating the forecast horizon for the forecast of the onset of congestion.

10.2.6 Spatial-Temporal Analysis and Traffic Predictions

In order to forecast the link occupancies, this research developed a novel multiple regression model that depends on the spatial and temporal relationship of link occupancy and Lyapunov exponent (calculated from flow). The model, described in detail in Chapter

9, involved a dynamic process requiring filtering to identify the influencing links to include in the prediction model at every stage depending on the correlation coefficient and the lag time intervals. With an average of 90% of the link predictions resulting in medium to high correlation with the actual observations, the research shows that the model can validly predict the link occupancy from the exponents for the short-term predictions of up to 15 minutes. Given the reliability of the predictions, the methodology adopted for the selection of this forecast horizon may well provide answers to Kirby and Smith (1991) who highlighted the need for guidance on how to select the forecasting horizon to make predictions realistic. The degree of fit between the actual and predicted observations typically by the model is in line with the conclusions of Jianming *et al*. (2003) who gave an equalisation coefficient of 95.44% in their model. However, it must be pointed out here that the two models have different data sampling frequency (15 minutes and 20 seconds for Jianming *et al.* and this work respectively), whilst in the former, the result was for an isolated location and did not data analyse spatially over the network.

In this research, it may be argued that the predictions of the link occupancy from the link occupancy data gave slightly higher degrees of correlation coefficient compared to the Lyapunov exponents data. Considering this, the question arises "Is there any benefit in the use of the Lyapunov exponents?" Even though in terms of the degree of correlation the model forecasting the link occupancy from the occupancy data slightly out-performs the Lyapunov exponent's model, the latter has a higher number of valid predictions spatially than the occupancy model. Over all the seven cohorts, the number of valid predictions were 7158 and 6380 links for the model based on the Lyapunov exponents and occupancy respectively. Consequently, the Lyapunov exponents (rather than the occupancy) have greater ability to report on the situational state of the network over a wider area of the network. Therefore, from the temporal-spatial analysis perspective, this observation suggests that the Lyapunov exponents-based forecast model will provide a better understanding of the future traffic flow considering the network as a whole.

The model's ability to predict to such a high level of accuracy means that incorporating these algorithms into existing demand responsive systems such as SCOOT has the potential to enhance the current practice of urban traffic management. The prediction of the link occupancies based on the temporal and spatial analysis can provide a rich source of information for the operation of SCOOT systems. Current operations rely on historic traffic information used to undertake optimisation procedures off-line to modify engineering

judgements or override SCOOT splits, offsets and cycle times in response to recurrent congestion events to maintain the network operation at the highest level of service. Even though currently the most effective form of control, this approach has two main limitations. Firstly, the optimised solution may not be suitable given the difficulty in judging the consistency of the current with previous congestion event through CCTV surveillance and there can be significant disparity between the historic and the actual flows. Secondly, operation is reactive rather than proactive. Operators respond to what is happening in the network without the ability to determine congested conditions in advance. Chaos Theory opens up the potential to operate with more sophistication with potential to avoid saturated traffic conditions. The incorporation of these algorithms means optimisation procedures can be derived to influence the build-up of flows downstream into the future and therefore avoid saturated conditions. This approach would address the current operational deficiencies due to the disparity in the historic and observed flows in a reactive system.

Further, this research provides a novel approach to determine continuously from the current predicted flows the ***future*** network traffic conditions in terms of being saturated or unsaturated. Since the future flows can be predicted with sufficient accuracy to calculate the Lyapunov exponents, the dynamic condition of the traffic in the network can be forecast ahead of real-time to enable network managers to initiate measures to avoid such adverse network conditions. Therefore, this approach maintains the use of the existing legacy control systems but provides a functional support by implementing additional algorithms based on Lyapunov exponents to improve effectively the outputs of the optimisation of the splits, offsets and cycle times. Given the computational ability of current computer systems (the use of The Cloud and parallel processing) it is envisaged that faster than real time modelling will enable a bespoke signal control intervention to be derived.

The algorithms therefore present intelligent traffic (transport) systems an opportunity to enhance their potential through an ability to forecast the onset of congestion based on a model of short-term traffic flows with acceptable levels of accuracy for the planning of traffic management strategies to be derived to minimise the impacts of adverse traffic conditions. The importance of this is summarised by Cheslow *et al.* (1992) that "*the ability to make and continuously update predictions of traffic flows and link times for several minutes into the future using real-time data is a major requirement for providing dynamic traffic control.*" The model developed in this thesis fits into the requirement for providing dynamic control considering that it possesses the ability to make accurate predictions, and

the algorithms allow for continuous updating of the predictions so that the network's future state is known and thus to alert managers to implement mitigation measures to avoid the adverse impacts. Consequently, the results offer a potential to render the decision making a *feed forward* rather than a *feedbac*k form of control.

10.3 Policy Implications for Government and Local Authorities

The forecast model developed has the ability to predict to a high level of accuracy, and therefore provides Local Authorities (LAs) a tool to enhance the current practice of urban traffic management by incorporating Chaos Theory into existing demand responsive systems such as SCOOT. The prediction of the link occupancies based on the temporal and spatial analysis can provide a rich source of information for the operation of SCOOT systems. Current operators in LAs rely on historic traffic information to undertake optimisation procedures off-line to modify engineering judgements or override SCOOT splits, offsets and cycle times in response to recurrent congestion events to maintain the network operation at the highest level of service. Even though currently the most effective form of control, this approach has two main limitations. Firstly, the optimised solution may not be suitable given the difficulty in judging the consistency of the current with previous congestion event through CCTV surveillance and there can be significant disparity between the historic and the actual flows. Secondly, operation is reactive rather than proactive. Operators respond to what is happening in the network without the ability to determine congested conditions in advance. Chaos Theory opens up the potential to operate with more sophistication with potential to avoid saturated traffic conditions. The incorporation of these algorithms means optimisation procedures can be derived to influence the build-up of flows downstream into the future and therefore avoid saturated conditions. LAs would benefit from this approach to address the current operational deficiencies due to the disparity in the historic and observed flows in a reactive system.

Further, by making use of rich data sources from existing legacy systems, this research provides Local Authorities a novel approach to determine continuously from the current predicted flows the ***future*** network traffic conditions in terms of being saturated or unsaturated. Since the future flows can be predicted with sufficient accuracy to calculate the Lyapunov exponents, the dynamic condition of the traffic in the network can be forecast ahead of real-time to enable network managers to initiate measures to avoid such adverse network conditions. Therefore, this approach enables LAs maintain the use of the existing legacy control systems but provides a functional support by implementing additional

algorithms based on Lyapunov exponents to improve effectively the outputs of the optimisation of the splits, offsets and cycle times. Given the computational ability of current computer systems (the use of The Cloud and parallel processing), it is envisaged that LAs could feasibly implement faster than real time modelling to enable a bespoke signal control intervention to be derived.

The algorithms therefore present LAs an opportunity to enhance their potential of existing intelligent traffic (transport) systems through an ability to forecast the onset of congestion based on a model of short-term traffic flows with acceptable levels of accuracy for the planning of traffic management strategies to be derived to minimise the impacts of adverse traffic conditions. The importance of this is summarised by Cheslow *et al*. (1992) that "*the ability to make and continuously update predictions of traffic flows and link times for several minutes into the future using real-time data is a major requirement for providing dynamic traffic control.*" The model developed in this thesis fits into the requirement for providing dynamic control considering that it possesses the ability to make accurate predictions, and the algorithms allow for continuous updating of the predictions so that the network's future state is known and thus to alert managers to implement mitigation measures to avoid the adverse impacts. Consequently, the results offer LAs a potential to render the decision making a *feed forward* rather than a *feedbac*k form of control.

Furthermore, the research offers an additional reason to justify the maintenance of existing UTMC systems because LAs will find beneficial the wealth of data available as by-product these systems useful for implementing Chaos Theory. However, in real world application, LAs may require highly resolved temporal data from different spatial locations to enable the tempo-spatial modelling using Chaos Theory. In the event that insufficient spatial data is available, LAs may find it appropriate to investigate the potential use of pervasive sensors such as Bluetooth and motes operating through a "linking deployment" with existing legacy systems such as SCOOT and ANPR.

10.4 Limitations of the Study

The use of only one-minute interval noise data due to the unavailability of highly resolved traffic noise data at various sampling intervals did not allow for a detailed assessment of the potential of traffic noise in this research. The raw noise data was not available so the data could be not interrogated at other (higher) resolutions to establish the suitability of such data. Consequently, the study failed to establish the suitability of noise data in understanding the evolution of the traffic in a signal controlled system.

Further, the study has been limited to Welford Road SCOOT region in Leicester. Time constraints did not allow replicating the study in other SCOOT regions and non-urban networks. Notwithstanding, the scope of the research proposal has been executed fully. However, the potential to enrich further the conclusions by the assessment of other scenarios have not been possible.

Finally, it would be helpful to apply the approach developed in the study to examine in more detail the results of previous work from similar networks, but this was not possible because of the disparity in the type and scale of the networks. This would have given a richer understanding of the differences found in the results. This study offers a comprehensive analysis of the traffic evolution of 40 links in a network, whilst the majority of studies reported in existing literature have focussed on only a single link location that are not necessarily in urban areas. Therefore, there was no other work to compare the spatial aspect of the output from this research.

10.5 Conclusions

The introductory chapter highlighted a number of objectives targeted at addressing specific questions in order to enhance our knowledge about applying Chaos Theory in traffic studies. Having fulfilled these objectives, and presented the results and discussions, this concluding section reflects upon the findings of the study in order to establish how the research undertaken answers the research questions enumerated in section 1.5, which are:

1. *What are the main limitations and which gaps are identifiable for improvement from previous studies for developing Chaos Theory as tool for improving urban traffic control?*
2. *Are traffic variables such as noise, flow and occupancy appropriate to yield results that would enhance the knowledge of applying Chaos Theory to traffic studies?*
3. *Is the choice of the temporal resolution of traffic data an important factor in using Chaos Theory to analyse traffic data?*
4. *Can Chaos Theory detect the onset of congestion within signalised urban networks under real world conditions?*
5. *Which chaos parameter can be developed as a predictive tool for forecasting the occupancy of links in a SCOOT controlled network region?*
6. *What is the forecast time horizon that yields reliable predictions from the traffic model?*

7. *If so, can this information be used to inform real-time traffic management and control strategies that are 'one-step ahead' of current traffic congestion strategies?'*

The conclusion refers to the research questions in order in the following paragraphs.

The extant literature indicated that previous attempts at applying Chaos Theory in traffic studies mainly focussed on identifying or proving the presence of chaotic properties in traffic data, but the studies conducted were found to be mainly for inter-urban and motorway environments and were not exhaustive in exploring the practical uses of Chaos Theory for urban traffic control. The traffic data used in previous studies was of low temporal resolution, and coupled with low spatial density and relatively short observation period did not enable the investigation of valid practical uses for Chaos Theory. The research confirmed that Chaos Theory is data-hungry and dependent on high temporal and spatial resolution data. The research conducted in this thesis has taken advantage of the appropriate data with these requisite qualities to overcome the weaknesses of earlier studies, and to address where possible their recommendations thus to gain further insight into the chaotic properties embedded in traffic time series data. Specific issues addressed relate to: (i) how to determine the forecast horizon to ensure reliable predictions; (ii) how to analyse the data for a timely detection of the transition to chaos and (iii) the verification of the importance of the sampling frequency in understanding properties of significance in the system.

The previous studies mainly focused on time series data of traffic flow, occupancy and speed. This study showed that traffic noise, flow and occupancy data are variables that indicate that road traffic has properties that enhances the potential for applying Chaos Theory in traffic studies. However, the traffic noise data available for use in this research lacked a detailed assessment due to the constraint of the data availability at only 60 second sampling intervals. Although the study identified traffic noise is chaotic, the profile of the Lyapunov exponents did not indicate any underlying structure to enable interpretation of the exponents in relation to known traffic flow behaviour. This was possibly due to the unsuitable use of the logarithmic noise data in the exponent formula, which is also logarithmic. Perhaps, a highly resolved temporal data could give a suitable profile, but the data limitations did not enable this investigation. On the other hand, the results of this research indicated that flow and occupancy at 20 second sampling frequency generated the appropriate profiles to investigate and gain understanding of the evolution of the traffic in

a signal control system. Flow and occupancy data sampled at 10, 30, 60 and 120 seconds did not; whilst they demonstrated the presence of chaos were insensitive to the onset of congestion.

The research highlighted the importance of the choice of the temporal resolution in applying Chaos Theory that the appropriate resolution must be identified by a systematic analysis of the data to be studied. The results showed that the resolution needed to be reasonably high and for the flow and occupancy data in this research, 20 seconds was most suitable for understanding the system's evolution. Frazier and Kockelman (2004) raised this concern about sampling drawing attention to the fact that at the wrong frequency important system dynamics were unable to be captured. Whilst this research has tested different sampling frequencies to reach this conclusion, a specific criterion to determine the most suitable frequency was not established. Future work may find it prudent to test a number of sampling frequencies and data types to endorse the approach in this research or recommend an alternative. For flow and occupancy, this research showed that data sampled at 20 seconds frequency is suitable.

A novelty of this research is the ability to use Chaos Theory to determine the onset of traffic congestion. Extant literature has not been able to connect Chaos Theory to congestion in traffic networks. However, this work shows that the profiles of the Lyapunov exponents can give an indication of the traffic flow regimes in the road network. A change in the progression of the curve from positive to negative gradient (both in the unstable region) indicated a state change from free-flow to chaos (unstable). On the other hand, a change from negative to positive gradient (both in the stable region) is a state change from chaos (unstable) to free flowing/congestion.

The study found the Lyapunov exponents were a useful parameter to characterise the datasets in terms of chaos. Further, the study identified the Lyapunov exponent as a parameter suitable for developing as a predictive tool for forecasting the link occupancy. Consequently, the study proposes a dynamic multiple regression model using the Lyapunov exponents as the independent variable for the forecast of the link occupancy based on the understanding of the temporal and spatial correlation of the flows on the links across the network.

The developed model suggested that the link occupancy can be predicted with an acceptable degree of accuracy up to 15 minutes in advance based on the Lyapunov exponents. The

results of this model compared with a forecast horizon of one hour indicated that the moderate to strong predictions are 92% and 65% respectively. The comparative figure from an alternative model of the link occupancy from the occupancy data itself is 96%. The actual numbers of links predicted 15 minutes ahead were fewer in the occupancy model than the Lyapunov exponents' model. The problem of the occupancy data was the high failure rate against the adopted criteria based on the degree of correlation and the lag time for choosing the selected links to include in the model. This means that the predictions from the Lyapunov exponents (rather than occupancy) make it possible to report on the impending conditions over a wider part of the network. That is, the Lyapunov exponents provide a better indication of the overall state of the network.

The prediction model has the potential to overcome existing limitations of current demand responsive systems that only address network problems ***as*** or ***after*** they occur. The results suggest that incorporating the forecasting algorithms into existing systems would provide additional intelligence to these systems making them pre-emptive to anticipate congestion and therefore implement mitigating strategies in advance.

10.6 Recommendations for Further Studies

The results of the urban network analysed were compared with the existing literature of studies from the both urban and non-urban networks. The comparisons, based on different resolution data, of the different networks gave reasonable conclusions. In order to obtain better understanding of the chaos parameters (lag, embedding and correlation dimensions and Lyapunov exponents), however, it will be necessary for a future comprehensive study to focus on the comparative appraisals of the different types of networks such as ring-road, highway etc. using traffic data of the same resolution and characteristics.

This thesis has analysed data and developed an understanding of free flow, congestion and transition from/to free-flow to/from congestion going through the states associated with chaos. This was hypothesised using engineering judgement based on traffic flow data from traffic management systems such as SCOOT. Future studies may build on this understanding by using micro-simulation models to simulate events that will actually result in chaos. By developing experimentally controlled events of chaos, the theoretical chaos hypothesised in this research can be verified. By repeating these events using different seeds and changes in input variables and analysing outputs together with the Lyapunov exponents, future studies can confirm that the interpretations based on engineering judgement and real-time traffic are in fact true. This would enhance further our

understanding of the Lyapunov exponents, their sensitivity and their relationships to different traffic flow regimes, by mapping of the exponents to specific conditions particularly in determining the onset of congestion.

Finally, this thesis identified that Chaos Theory has the potential for application to traffic management, by predicting the temporal and spatial evolution of traffic in the network building up to create events using the cross-correlation of Lyapunov exponents on links within the network. Future studies can explore the comparative modelling of the observed and predicted flows to assess the capability of Chaos Theory to forecast the concentrations of environmental pollutants. If this proves to be feasible and yields reasonable outcomes, then it would justify a further use of the Lyapunov exponents in traffic studies.

References

AASHTO, ITE & NEMA. 2009. *National Transportation Communications for ITS Protocol: The NTCIP Guide* [Online]. Available: www.ntcip.org/library/documents/pdf/9001v0406r.pdf [Accessed 9th June 2015].

ABDALAH, S. F., AL-NAIMEE, K. A., MEUCCI, R., AL MUSLET, N. & ARECCHI, F. T. 2010. Experimental Evidence of Slow Spiking Rate in a Semiconductor Laser by Electro-optical Feedback: Generation and Control. *Applied Physics Research,* 2 (2).

ABDULLAH, B., PORWAL, H. & RECKER, W. 1999. Short-term Freeway Traffic Flow Predictioon Using Genetically-Optimised Time-Delay-Based Neural Networks. *78th Transport Research Board Annual Meeting.* Washington, DC: Transport Research Board.

AKYILDIZ, I. F., SU, W., SANKARASUBRAMANIAM, Y. & CAYIRCI, E. 2002. Wireless Sensor Networks: A Survey. *Computer Networks,* 38 (4)**,** pp. 393-422.

AMBRAVANESWARAN, B., PHILIPS, S. D. & BASARAN, O. A. 2000. Theoritical Analysis of a Dripping Faucet. *Physical Review Letters,* 85 (25)**,** pp. 5332-5.

ANDREWS, M. 1996. *Introduction to Chaos Theory* [Online]. Google. Available: http://www.gweep.net/~rocko/sufficiency/node10.html [Accessed 12th December 2012].

ANGELINE, P. J., SAUNDERS, G. M. & POLLACK, J. B. 1994. An Evolutionary Algorithm that Constructs Recurrent Neutral Networks. *IEEE Transactions on Neutral Networks* 5(1)**,** pp. 54-65.

ANSELIN, L. 1988. Lagrange Multiplier Test Diagnostics for Spatial Dependence and Spatial Hterogeneity. *Geographical Analysis,* 20 (1)**,** pp. 1-17.

ATHOL, P. 1965. Interdependence of Certain Operational Characteristics Within a Moving Traffic Stream. *Highway Research Record,* 72**,** pp. 58-87.

ATKESON, C. G., MOORE, A. W. & SCHAAL, S. 1997. Locally Weighted learning for Control. *Artificial Intelligence Review***,** pp. 75-113.

BARNSLEY, M. F., JACQUIN, A., MALASSENET, F., REUTER, L. & SLOAN, A. D. 1988. Harnessing Chaos for Image Synthesis. *Computer Graphics* 22**,** pp. 131-40.

BECKER, H. 2008. The Chaotic Blue Ocean. *Review of Business Research,* 8 (6)**,** pp. 125-131.

BEGG, D. & GRAY, D. 2004. Transport Policy and Vehicle Emission Objectives in the UK: Is the Marriage Between Transport and Environment Policy Over? *Environmental Science & Policy,* 7 (3)**,** pp. 155-163.

BELL, M. C. 1983. A Survey of the Methods Used to Define and Change Signal-Plans in the UK. *Transport Operations Research Group,* Research Report 50.

BELL, M. C. 1984. Ageing of Fixed-time Signal Plans. *16th Universities Transport Studies Group Conference* Loughborough, United Kingdom: Universities Transport Studies Group.

BELL, M. C. 2012. Newcastle University Integrated Database Platform (NUIDAP). *Transport Operations Research Group*.

BELL, M. C. & GALATIOTO, F. 2013. Novel Wireless Pervasive Sensor Network to Improve the Understanding of Noise in Street Canyons. *Applied Acoustics,* 74 (1)**,** pp. 169-180.

BELL, M. C., GALATIOTO, F., HILL, G. & NAMDEO, A. 2009. Modelling Environmental Impacts of Traffic Using a New Generation of Pervasive Sensors. *16th World Congress for ITS Systems and Services.* Stckholm, Sweden: Intelligent Transport Systems.

BELL, M. G. H. 2005a. Signal Control at Intersections. *In:* O'FLAHERTY, C. A. (ed.) *Transport Planning and Traffic Engineering* England Elsevier Butterworth-Heinemann.

BELL, M. G. H. 2005b. Signal Control in Networks. *In:* O'FLAHERTY, C. A. (ed.) *Transport Planning and Traffic Engineering* England Elsevier Butterworth-Heinemann.

BENETTIN, G., GALGANI, L., GIORGILLI, A. & STRELCYN, J.-M. 1980. Lyapunov Characteristic Exponents for Smooth Dynamical Systems and for Hamiltonian Systems: A Method for Computing all of them. Part 1: Theory. *Meccanica,* 15**,** pp. 9-20.

BENNETT, S., WATSON, A. & BELL, M. n.d. *Realising an Integrated Database for the Instrumented City* [Online]. Institute of Transport Studies, University of Leeds. Available: http://www.engineering.leeds.ac.uk/civil/research/cae/papers/UTSG.pdf [Accessed 3rd September 2015].

BEVIVINO, J. 2009. The Path From the Simple Pendulum to Chaos. *Dynamics at the Horsetooth* 1**,** pp. 1-24.

BIGAZZI, A. Y. & FIGLIOZZI, M. A. 2013. Marginal Costs of Freeway Traffic Congestion with On-Road Pollution Exposure Externality. *Transportation Research Part A: Policy and Practice,* 57 (0) pp. 12-24.

BLOGG, M., SEMLER, C., HINGORANI, M. & TROUTBECK, R. 2010. Travel Time and Origin-Destination Data Collection using Bluetooth MAC Address Readers. *Australasian Transport Research Forum 2010.* Canberra, Australia: Australasian Transport Research.

BOKER, S. M. 1996. *Linear and Non-linear Dynamical Systems Data Analytic Techniques and an Application to Developmental Data.* Doctor of Philosophy, University of Virginia.

BOURNER, M. 1984. SCOOT Experience in Kent. *Institution of Electrical Engineers Colloqium* Digest No. 1984 (58).

BOWEN, G. T. & BRETHERTON, R. 1996a. Latest developments in SCOOT-Version 3.1. *Eighth International Conference on Road Traffic Monitoring and Control.* London: IET.

BOWEN, G. T. & BRETHERTON, R. D. 1996b. Latest Developments in SCOOT-Version 3.1. *Eighth International Conference on Road Traffic Monitoring and Control.* London: IET.

BRADLEY, L. 2010. *The Butterfly Effect* [Online]. Available: http://www.bibliotecapleyades.net/ciencia/ciencia_climatechange25.htm [Accessed 10th June 2015].

BRENDAN, R. 2007. Aid to Defining Fractal Dimensions. *In:* FRACTALDIMENSIONEXAMPLE.PNG (ed.). On-line.

BRETHERTON, D., BODGER, M. & COWLING, J. 2006. SCOOT - Managing Congestion, Control and Communications. *Traffic Engineering & Control,* 47 (3)**,** pp. 88-92.

BRETHERTON, D., WOOD, K. & RAHA, N. 1998. Traffic Monitoring and Congestion Management in the SCOOT Urban Traffic Control System. *Transportation Research Record: Journal of the Transportation Research Board,* 1634**,** pp. 118-122.

BRETHERTON, R. D. 1979. Five Methods of Changing Fixed-time Traffic Signal Plans. *Transport and Road Research* LR 879.

BRILON, W. & WIETHOLT, T. 2013. Experiences with Adaptive Signal Control in Germany. *Transportation Research Record: Journal of the Transportation Research Board***,** pp. 9-16.

BUSCH, F. 1996. Traffic Telematics in Urban and Regional Environments. *Inter-Traffic Conference.* Amsterdam: PTRC Education and Research Services Limited.

BUTLER, A. 1990. A Methodological Approach to Chaos: Are Economists Missing the Point? *Federal Reserve Bank of St. Louis,* 72 (13)**,** pp. 36-48.

BYUNGKYU, B. & CHEN, Y. 2010. *Quantifying the Benefits of Coordinated Actuated Traffic Signal Systems: A Case Study* [Online]. Available: http://www.virginiadot.org/vtrc/main/online_reports/pdf/11-cr2.pdf [Accessed 8th June 2015].

CAMERON, G. D. B. & DUNCAN, G. I. D. 1996. PARAMICS—Parallel Microscopic Simulation of Road Traffic. *The Journal of Supercomputing,* 10 (1)**,** pp. 25-53.

CENTRE FOR ECONOMICS AND BUSINESS RESEARCH 2014. The Future Economic and Environmental Cost of Gridlock in 2030. *An Assessment of the Direct and Indirect Economic and Environmental Costs of Idling in Road Traffic Congestion to Households in the UK, France, Germany and the USA.* London: Cebr Consultancy.

CHAPMAN, L. 2007. Transport and Climate Change: A Review. *Journal of Transport Geography,* 15 (3)**,** pp. 354-367.

CHATTERJEE, K. & GORDON, A. 2006. Planning for an Unpredictable Future: Transport in Great Britain in 2030. *Transport Policy,* 13 (3)**,** pp. 254-264.

CHESLOW, M., HATCHER, S. G. & PATEL, V. M. 1992. An Initial Evaluation of Alternative Intelligent Vehicle Highway Systems Architectures. *MTR 92W0000063.* Bedford, MS: MITRE Corporation.

CHRISTOS, F. & IBRAHIM, K. 1994. Beyond Uniformity and Independence: Analysis of R-trees Using the Concept of Fractal Dimension. *Proceedings of the thirteenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems.* Minneapolis, Minnesota, USA: ACM.

CLIFF, A. & ORD, J. 1975a. Model Building and the Analysis of Spatial Pattern in Human Geography. *Journal of the Royal Statistical Society. Series B (Methodological)***,** pp. 297-348.

CLIFF, A. & ORD, J. K. 1975b. Space-Time Modelling with an Application to Regional Forecasting. *Transactions of the Institute of British Geographers***,** pp. 119-128.

COLLINS, M. S. 1995. The Evaluation of Urban Congestion Management Schemes. *Urban Congestion Management: Colloquium Organized by Professional Group C12 (Transport Electronics and Control).* London: The Institution of Electrical Engineers.

COULLET, P., MAHADEVAN, L. & RIERA, C. S. 2005. Hydrodynamical Models for the Chaotic Dripping Faucet. *Journal of Fluid Mech.,* 526**,** pp. 1-17.

COWLING, J., HAY, G. & BRETHERTON, D. n.d. *SCOOT Advise Leaflet 2: Congestion Management in SCOOT* [Online]. Available: http://www.scoot-utc.com/documents/2_Congestion.pdf [Accessed 7th June 2015].

CRAGG, S. 2013. Bluetooth Detection - Cheap but Challenging. On-line: Transport Scotland.

DENDRINOS, D. S. 1994. Traffic-Flow Dynamics: A Search for Chaos. *Pergamon: Chaos, Solitons & Fractals,* 4 (4)**,** pp. 605-617.

DENTON, T. A., DIAMOND, G. A., HELFANT, R. H., KHAN, S. & KARAGUEUZIAN, H. 1990. Fascinating Rhythm: A Primer on Chaos Theory and its Application to Cardiology. *American Heart Journal,* 120 (6)**,** pp. 1419-1440.

DETR 1998. A New Deal for Transport: Better for Everyone United Kingdom: DfT Publications.

DFT 1999. The "SCOOT" Urban Traffic Control System. *Traffic Advisory Leaflet, 7 (99).* United Kingdom: DfT Publications.

DFT 2004. The Future of Transport - A Network for 2030. *Presented to Parliament by the Secretary of State for Transport by Command of Her Majesty.* United Kingdom: DfT Publications.

DFT. 2009. *UTMC Initiative* [Online]. Available: www.utmc.uk.com/background/02.php [Accessed 14th September 2012].

DFT 2011. Transport Energy and Environment Statistics 2011. *Statistical Release.* United Kingdom: DfT Publications.

DFT 2015a. Road Length in Great Britain 2014 *Statistical Release.* United Kingdom: DfT Publications.

DFT 2015b. Road Traffic Estimates: Great Britain 2014 *Statistical Release.* United Kingdom: DfT Publications.

DHILLON, G. S. & WARD, J. 2003. Chaos Theory as a Framework for Studying Information Systems. *Advanced Topics in Information Resources Management,* 2**,** pp. 320-338.

DHINGRA, S. L. & GULL, I. 2011. Traffic Flow Theory Historical Research Perspectives. *Transport Research Circular E-C 149: 75 Years of the Fundamental Diagram of Traffic Flow* June 2011**,** pp. 45-62.

DONATI, F., MAURO, V., RONCOLINI, G. & VALLAURI, M. A Hierarchical Decentralisated Traffic Light Control System - the First Realisation. 9th International FAC World Congress, 1984 Budapest. Federation of Automotive Control.

DONIOSIO, A., MENEZES, R. & MENDES, D. A. 2003. *Mutual Information: A Dependence Measure for Non-linear Time Series* [Online]. On-line: Research Division of the Federal Bank of St. Louis. [Accessed 30th August 2012].

DOUGLAS, M. J., WATKINS, S. J., GORMAN, D. R. & HIGGINS, M. 2011. Are Cars the New Tobacco? *Journal of Public Health,* 33 (2)**,** pp. 160-169.

DOWNES, I., RAD, L. B. & AGHAJAN, H. 2006. Development of a Mote for Wireless Image Sensor Networks. *Proc. of COGnitive Systems with Interactive Sensors (COGIS).* Paris, France.

DREYER, K. & HICKEY, F. R. 1991. The Route to Chaos in a Dripping Water Faucet. *American Journal of Physics,* 59 (7)**,** pp. 619-627.

ECKMANN, J.-P. & RUELLE, D. 1985. Ergodic Theory of Chaos and Strange Attractors. *Reviews of Modern Physics* 57(3) Part 1**,** pp. 617-56.

EDDINGTON, R. 2006. The Eddington Transport Study. *The Case for Action: Sir Rod Eddington's Advice to Government* United Kingdom: DfT Publications.

EKINS, P., ANANDARAJAH, G. & STRACHAN, N. 2011. Towards a Low-Carbon Economy : Scenarios and Policies for the UK. *Climate Policy,* 11 (2)**,** pp. 865-882.

ELERT, G. 2007. Measuring Chaos. *The Chaos Hypertext Book* On-line.

ELFAOUZI, N.-E. Nonparametric Traffic Flow Prediction Using Kernel Estimator. Internaional Symposium on Transportation and Traffic Theory, 1996 Lyon, France. Pergamon Press, Oxford, Royaume-Uni Etats-Unis, pp. 41-54.

FAN, J. & GIJBELS, I. 1996. *Local Polynomial Modelling and its Applications: Monographs on Statistics and Applied Probability 66,* CRC Press.

FARGES, J. L., KHOUDOUR, I. & LESORT, J. PRODYN: On Site Evaluation. 3rd International Conference on Road Traffic Control 1990 Stevenage, England. IEE Publications, pp. 62-66.

FENGHUA, W., ZHIJIAN, J., ZISHU, Z. & XUSHENG, W. Modeling the DC Electric Arc Furnace Based on Chaos Theory and Neural Network. Power Engineering Society General Meeting, 2005, 12-16 June 2005 San Francisco, CA, USA. IEEE Publications, pp. 2503-2508.

FORD, J. 1983. How Random is a Coin Toss. *Physics Today,* 36 (4)**,** pp. 40-47.

FRASER, A. M. & SWINNEY, H. L. 1986. Independent Coordinates for Strange Attractors from Mutual Information. *Phys,* Rev. A (33)**,** pp. 1134-1140.

FRAZIER, C. & KOCKELMAN, K. M. 2004. Chaos Theory and Transportation Systems: An Instructive Example. *83rd Annual Meeting of the Transportation Research Board, January 2004.* Washington D. C.: TRB Publications.

FREDE, V. & MAZZEGA, P. 1999. Detectability of Deterministic Non-linear Processes in Earth Rotation Time-Series—I. Embedding. *Geophysical Journal International,* 137 (2)**,** pp. 551-564.

FROEHLING, H., CRUTCHFIELD, J. P., FARMER, D., PACKARD, N. H. & SHAW, R. 1981. On Determining the Dimension of Chaotic Flows. *Physica 3D***,** pp. 605-617.

GALAS, J. 2010. *Traffic Data Collection with Bluetooth* [Online]. On-line: Ilinois Department of Transport Available: http://www.conferences.uiuc.edu/traffic/2010PDF/File_10_Galas.pdf [Accessed 13th August 2013].

GALATIOTO, F. & BELL, M. C. 2013. Exploring the Processes Governing Roadside Pollutant Concentrations in Urban Street Canyon. *Environmental Science and Pollution Research,* 20 (7)**,** pp. 4750-4765.

GANEK, A. G. & CORBI, T. A. 2003. The Dawning of the Autonomic Computing Era. *IBM Systems Journal,* 42(1)**,** pp. 5-18.

GARFINKEL, A., SPANO, M. L., DITTO, W. L. & WEISS, J. N. 1992. Controlling Cardiac Chaos. *Science,* New Series 257 (5074)**,** pp. 1230-1235.

GLEICK, J. 1988. *Chaos - Making a New Science* London, Great Britain William Heinemann Ltd.

GOH, S. M., NOORANI, M. S. M. & HASHIM, I. 2009. Efficacy of Variational Iteration Method for Chaotic Genesio System – Classical and Multistage Approach. *Chaos, Solitons & Fractals,* 40 (5)**,** pp. 2152-2159.

GOODWIN, P. 2004. The Economic Costs of Road Traffic Congestion. *A Discussion Paper.* United Kingdom: Rail Freight Group.

GUO, J., HUANG, W. & WILLIAMS, B. M. 2014. Adaptive Kalman Filter Approach For Stochastic Short-term Traffic Flow Rate Prediction and Uncertainty Quantification. *Transportation Research Part C: Emerging Technologies,* 43, Part 1**,** pp. 50-64.

HADAMARD, J. 1898. Les Surfaces `a Courbures Oppos´ees et Leurs Lignes G´eod´esique. *Journal de Mathématiques Pures et Appliquées* 4**,** pp. 27-73.

HALL, F. L. 1997. Traffic Stream Characteristics. *Traffic Flow Theory: A State of the Art Report - Revised Monograph on Traffic Flow Theory* Tennessee: Transport Research Board.

HAMILTON, A., WATERSON, B., CHERRET, T., A., R. & SNELL, I. 2012. Urban Traffic Control Evolution. *44th Universities Transport Studies Group Conference.* Aberdeen, United Kingdom: Universities Transport Studies Group.

HAMILTON, A., WATERSON, B., CHERRETT, T., ROBINSON, A. & SNELL, I. 2013. The Evolution of Urban Taffic Control: Changing Policy and Technology. *Transportation Planning and Technology,* 36 (1)**,** pp. 24-43.

HAN, C. & SONG, S. A Review of Some Main Models for Traffic Flow Forecasting. IEEE International Conference on Intelligent Transport Systems, 12-15 October 2003 Shanghai, China. Institute of Electrical and Electronic Engineers, pp. 216-219.

HAY, G., COWLING, J. & BRETHERTON, D. n.d. *SCOOT Advise Leaflet 1: The "SCOOT" Urban Traffic Control System* [Online]. Available: http://www.scoot-utc.com/documents/1_SCOOT-UTC.pdf [Accessed 7th June 2015].

HAZELTON, M. L. 2004. Estimating Vehicle Speeds from Traffic Counts and Occupancy Data. *Journal of Data Science,* 2**,** pp. 231-244.

HEAD, I. 1995. Event-based Short-term Traffic Prediction Model. *Transport Research Board,* 1510**,** pp. 45-52.

HEYDECKER, B. G. & ADDISION, J. D. 2011. Measuring Traffic Flow Using Real-Time Data. *Transport Research Circular E-C 149: 75 Years of the Fundamental Diagram of Traffic Flow* June 2011.

HICKMAN, R., ASHIRU, O. & BANISTER, D. 2010. Transport and Climate Change: Simulating the Options for Carbon Reduction in London. *Transport Policy,* 17(2)**,** pp. 110-125.

HICKMAN, R. & BANISTER, D. 2007. Looking Over the Horizon: Transport and Reduced CO< sub> 2</sub> Emissions in the UK by 2030. *Transport Policy,* 14 (5)**,** pp. 377-387.

HITCHCOCK, G., CONLAN, B., KAY, D., BRANNIGA, C. & NEWMAN, D. 2014. Air Quality and Road Transport. *Impacts and Solutions.* London: RAC Foundation.

HORTON, M. & SUH, J. 2005. A Vision for Wireless Sensor Networks. *Microwave Symposium Digest, 2005 IEEE MTT-S International.*

HOUNSELL, N. B., SHRESTHA, B. P., PIAO, J. & MCDONALD, M. 2009. Review of Urban Traffic Management and the Impacts of New Vehicle Technologies. *Intelligent Transport Systems, Institution of Engineering and Technology,* 3 (4)**,** pp. 419-428.

HUANG, S. & SADEK, A. W. 2009. A Novel Forecasting Approach Inspired by Human Memory: The Example of Short-term Traffic Volume Forecasting. *Transportation Research Part C: Emerging Technologies,* 17 (5)**,** pp. 510-525.

HUI, F., JIANMIN, X. & LUNHUI, X. 2005. Traffic Chaos and its Prediction Based on a Nonlinear Car-following Model. *Journal of Control Theory and Applications,* 3**,** pp. 302-307.

HUNT, B. R. & YORKE, J. A. 1993. Maxwell on Chaos. *Nonlinear Science Today,* 3(1)**,** pp. 1-4.

HUNT, P. B., ROBERTSON, D. I., BRETHERTON, R. D. & ROYLE, M. C. 1982. The SCOOT On-line Traffic Signal Optimisation Technique. *Traffic Engineering and Control***,** pp. 190-192.

IHT 1976. *Transportation and Traffic Engineering Handbook,* USA, Prentice-Hall.

IHT 1997a. Coordinated Traffic Signals. *Transport in the Urban Environment* London, United Kingdom: The Institution of Highways & Transportation.

IHT 1997b. Technology for Network Management. *Transport in the Urban Environment* London, United Kingdom: The Institution of Highways and Transport.

JIANMING, H., CHUNGUANG, Z., JINGYAN, S., ZUO, Z. & JIANGTAO, R. An Applicable Short-term Traffic Flow Forecasting Method Based on Chaotic Theory. IEEE International Conference on Intelligent Transport Systems, 2003 Shanghai, China. Institute of Electrical and Electronic Engineers pp. 608-613.

JOU, R.-C., KOU, C.-C. & CHEN, Y.-W. 2013. Drivers' Perception of LOSs at Signalised Intersections. *Transportation Research Part A: Policy and Practice,* 54 (0)**,** pp. 141-154.

KAMARIANAKIS, Y. & PRASTACOS, P. 2003. Spatial Time Series Modeling: A Review of the Proposed Methodologies. *The Regional Economics Applications Laboratory*.

KÁROLYI, G., PATTANTYÚS-ÁBRAHÁM, M. & JÓZSA, J. 2010. Finite-Size Lyapunov Exponent: A New Tool for Lake Dynamics. *Engineering and Computational Mechanics,* 163 (EM4)**,** pp. 251-259.

KAY, A. & JACKSON, P. 2012. An Appraisal of Emerging Bluetooth Traffic Survey Technology. *Scottish Transport and Applications Research Conference.* On-line.

KEPHART, J. O. & CHESS, D. M. 2003. The Vision of Autonomic Computing. *IEEE Computer Society,* 36(1)**,** pp. 41-50.

KESTING, A. & TREIBER, M. 2008. How Reaction Time, Update Time, and Adaptation Time Influence the Stability of Traffic Flow. *Computer-Aided Civil and Infrastructure Engineering,* 23 (2)**,** pp. 125-137.

KHALEGHEI GHOSHEH BALAGH, A., NADERKHANI, F. & MAKIS, V. 2014. Highway Accident Modeling and Forecasting in Winter. *Transportation Research Part A: Policy and Practice,* 59 (0)**,** pp. 384-396.

KIRBY, H. R. & SMITH, M. J. 1991. Can Chaos Theories Have Transport Applications? *23rd Universities Transport Studies Group Conference.* Nottingham, England: Universities Transport Studies Group.

KOLÁŘ, M. 1992. Theory for the Experimental Observation of Chaos in a Rotating Waterwheel. *Physical Review A,* 45 (2)**,** pp. 626-637.

KRESE, B. & GOVEKAR, E. 2013. Analysis of Traffic Dynamics on a Ring Road-Based Transportation Network by Means of 0–1 Test for Chaos and Lyapunov Spectrum. *Transportation Research Part C: Emerging Technologies,* 36 (0)**,** pp. 27-34.

KUMAR, A. & HEDGE, B. M. 2012. Chaos Theory: Impact On and Applications in Medicine. *Nitte University Journal of Science,* 2 (4)**,** pp. 93-99.

LAI, Y.-C. & YE, N. 2003. Recent Developments in Chaotic Time Series Analysis. *International Journal of Bifurcation and Chaos,* 13 (6)**,** pp. 1383-1422.

LAN, L. W., LIN, F.-Y. & KUO, A. Y. 2003. Testing and Prediction of Traffic Flow Dynamics with Chaos. *Journal of the Eastern Asia Society for Transport Studies,* 5 (Oct. 2003)**,** pp. 1975-1990.

LEE, J. F. J., KWOK, P. K. & WILLIAMS, J. 2014. Heterogeneity among Motorists in Traffic-Congested Areas in Southern California. *Transportation Research Part A: Policy and Practice,* 70 (0)**,** pp. 281-293.

LEES, A. 2013. Image of Hi-Trac® Blue TDC Systems Limited.

LEVY, D. 1994. Chaos Theory and Strategy: Theory, Application and Managerial Implications. *Strategic Management Journal,* 15**,** pp. 167-178.

LIGHTHILL, M. J. & WHITHAM, G. B. 1955. On Kinematics Waves: II. A Theory of Traffic Flow on Long Crowded Roads. *Proceedings of the Royal Society,* A229**,** pp. 317-415.

LIPA, C. n.d. *Chaos and Fractal - Introduction to Chaos* [Online]. Available: http://www.math.cornell.edu/~lipa/mec/lesson1.html [Accessed 15th January 2013].

LORENZ, E. N. 1993a. Deterministic Non-Periodic Flow. *Journal of the Atmospheric Sciences,* 20**,** pp. 130-41.

LORENZ, E. N. 1993b. *The Essence of Chaos* London, United Kingdom, UCL Press Ltd.

LOWRIE, P. R. The Sydney Co-ordinated Adaptive Traffic System: Principles, Methodology, Algorithms. IEE International Conference on Road Traffic Signalling, 1982 London, United Kingdom. Institution of Electrical Engineers.

MAHNKE, R., KAUPUŽS, J. & LUBASHEVSKY, I. 2005. Probabilistic Description of Traffic Flow. *Physics Reports,* 408 (1-2)**,** pp. 1-130.

MANDELBROT, B. B. 1983. *The Fractal Geometry of Nature* New York, USA, W. H. Freeman and Company

MANLEY, E. & CHENG, T. Understanding Road Congestion as an Emergent Property of Traffic Networks. International Multiconference on Complexity, Informatics and Cybernetics, 2010 Orlando, USA. pp. 109-114.

MATSON, L. E. 2007. The Malkus–Lorenz Water Wheel Revisited. *American Journal of Physics,* 75 (12)**,** pp. 1114-1122.

MAY, A. D. 1990. *Traffic Flow Fundamentals,* Englewood Cliffs, NJ, Prentice-Hall.

MAY, A. D. 2005. Introduction to Traffic Flow Theory. *In:* O'FLAHERTY, C. A. (ed.) *Transport Planning and Traffic Engineering.* England: Elsevier Butterworth-Heinemann.

MAY, R. M. 1976. Simple Mathematical Models With Very Complicated Dynamics. *Nature* 26 (1)**,** pp. 459-467.

MAY, R. M. & OSTER, G. F. 1976. Bifurcations and Dynamic Complexity in Simple Ecological Models. *The American Naturalist,* 110 (974)**,** pp. 573-599.

MCMURRAN, S. L. & TATTERSALL, J. J. 1997. Cartwright and Littlewood on Van der Pol's Equation. *Contemporary Mathematics,* 208**,** pp. 265-276.

MEEHAN, D. 2003. We Know It's Good, But Just How Good Is It? *Traffic Engineering and Control,* 44 (8)**,** pp. 295-298.

MIRCHANDANI, P. & HEAD, L. 2001. A Real-time Traffic Signal Control System: Architecture, Algorithms, and Analysis. *Transportation Research Part C: Emerging Technologies,* 9 (6)**,** pp. 415-432.

MOIOLA, J. L. & CHEN, G. 1996. *Hopf Bifurcation Analysis - A frequency Domain Approach,* MA, USA, World Scientific.

MORIARTY, P. & HONNERY, D. 2008. Low-Mobility: The Future of Transport. *Futures,* 40 (10)**,** pp. 865-872.

MURISON, M. A. 1995. Notes on How to Numerically Calculate the Maximum Lyapunov Exponent Massachusetts: U.S. Naval Observatory.

NAIR, A. S., LIU, J. C., RILETT, L. & GUPTA, S. 2001. Non-Linear Analysis of Traffic Flow. *IEEE Intelligent Transportation Systems Conference* Oakland (CA), USA: Institution of Electric and Electronic Engineers.

NARH, A. T., THORPE, N., BELL, M. C. & HILL, G. A. 2014. Using Chaos Theory to Examine the Dynamical States of Road Traffic in Signalised Urban Networks. *46th Universities Transport Studies Group* Newcastle, United Kingdom.: Universities Transport Studies Group.

NORTH, R., COHEN, J., WILKINS, S., RICHARDS, M., HOOSE, N., POLAK, J., BELL, M., BLYTHE, P., SHARIF, B., NEASHAM, J., SURESH, V., GALATIOTO, F. & HILL, G. 2009. Feild Deployments of the MESSAGE System for Environmental Monitoring. *Traffic Engineering and Control,* 50 (11)**,** pp. 484-488.

NÚÑEZ YÉPEZ, H. N., SALAS BRITO, A. L., VARGAS, C. A. & VICENTE, L. A. 1989. Chaos in a Dripping Water Faucet. *Eur. J. Phys.,* 10**,** pp. 99-105.

O'FLAHERTY, C. A. 2005a. Road Capacity and Design-Standard Approaches to Road Design. *In:* O'FLAHERTY, C. A. (ed.) *Transport Planning and Traffic Engineering.* London: Butterworth-Heineman.

O'FLAHERTY, C. A. 2005b. Traffic Planning Strategies. *In:* O'FLAHERTY, C. A. (ed.) *Transport Planning and Traffic Engineering.* Oxford: Butterworth-Heinemann.

OESTREICHER, C. 2007. A History of Chaos Theory. *Dialogues Clin Neurosci,* 9 (3)**,** pp. 279-89.

ONS 2014. Deaths Registered in England and Wales. *Statistical Release, Series DR (2013).* United Kingdom: ONS.

ORCUTT, K. F. & ARRITT, R. W. Comparative fractal dimensions for daytime and nocturnal surface layer turbulence. 11th Symp Boundayr Layer & Turb, 1995. Charlotte: NC Amer Meterol Soc.

ORTUZAR, J. D. D. & WILLUMSEN, L. G. 2006. Assignment. *Modelling Transport.* Third ed. London: John Wiley & Sons Ltd.

PACKARD, N. H., CRUTCHFIELD, J. P., FARMER, J. D. & SHAW, R. S. 1980. Geometry From a Time Series. *Physical Review Letters,* 45 (9) pp. 712-716.

PANKRATZ, A. 2012. *Forecasting with Dynamic Regression Models*, John Wiley & Sons.

PAPAGEORGIOU, M., BEN-AKIVA, M., BOTTOM, J., BOVY, P. H., HOOGENDOORN, S., HOUNSELL, N. B., KOTSIALOS, A. & MCDONALD, M. 2006. ITS and traffic management. *Handbooks in Operations Research and Management Science,* 11**,** pp. 743-754.

PAPAGEORGIOU, M., DIAKAKI, C., DINOPOULOU, V., KOTSIALOS, A. & YIBING, W. 2003. Review of Road Traffic Control Strategies. *Proceedings of Insitution of Electrical and Electronic Engineers,* 91 (12)**,** pp. 2043-2067.

PEEK TRAFFIC LTD., SIEMENS & TRL. 2008. *ASTRID - Automatic SCOOT Traffic Information Database* [Online]. Available: http://www.scoot-utc.com/ASTRID.php?menu=Technical [Accessed 30th January 2013].

PEITGEN, H.-O., JÜRGENS, H. & SAUPE, D. 2006. *Chaos and Fractals: New Frontiers of Science,* New York, Springer Science & Business Media.

QI, Y. & ISHAK, S. 2014. A Hidden Markov Model for Short Term Prediction of Traffic Conditions on Freeways. *Transportation Research Part C: Emerging Technologies,* 43, Part 1 (0)**,** pp. 95-111.

RAE, G. 2006. *Chaos Theory: A Brief Introduction* [Online]. Available: https://doublebubbleun1.godaddysites.com/fractals.html [Accessed 10th June 2015].

RATNER, B. 2009. The Correlation Coefficient: Its Values Range Between +1/-1, or Do They? *Journal of Targeting, Measurement and Analysis for Marketing,* 17**,** pp. 139-142.

REISS, J. D. 2001. *The Analysis of Chaotic Time Series.* Doctor of Philosophy in Physics, Georgia Institute of Technology.

ROBERTSON, D. I. 1969. TRANSYT: A TRAffic Network StudY Tool. *Transport and Road Research,* LR 253.

ROBERTSON, D. I. Coordinating Traffic Signals to Reduce Fuel Consumption. Conference of the Royal Society of London. Series A, Mathematical and Physical Sciences, 1983 London. The Royal Society, pp. 1-19.

ROSENSTEIN, M. T., COLLINS, J. J. & DE LUCA, C. J. 1993. A Practical Method for Calculating Largest Lyapunov Exponent from Small Data Sets. *Physica D,* 65**,** pp. 117-134.

ROSSER JR, J. B. 2009. Chaos Theory Before Lorenz. *Nonlinear Dynamics, Psychology, and Life Sciences,* 13(3)**,** pp. 257.

ROUX, J. C., SIMOYI, R. H. & SWINNEY, H. L. 1983. Observation of a Strange Attractor. *Physica D: Nonlinear Phenomena,* 8 (1-2)**,** pp. 257-266.

SAFANOV, L. A., TOMER, E., STRYGIN, V. V., ASHKENAZY, Y. & HAVLIN, S. 2002. Delayed-Induced Chaos with Multi-Fractal Attractor in a Traffic Flow Model. *Europhysics Letters,* 57**,** pp. 151-57.

SARBADHIKARI, S. & CHAKRABARTY, K. 2001. Chaos in the Brain: A Short Review Aluding to Epilepsy, Depression, Exercise and Lateralization. *Medical Engineering & Physics,* 23 (7)**,** pp. 447-457.

SATO, S., SANO, M. & SAWADA, Y. 1987. Practical methods of measuring the generalized dimension and the largest Lyapunov exponent in high dimensional chaotic systems. *Progress of Theoretical Physics,* 77**,** 1-5.

SAUER, T., YORKE, J. & CASDAGLI, M. 1991. Embedology. *Journal of Statistical Physics,* 65 (3,4)**,** pp. 95-116.

SCHEWE, P. 2007. *Controlling Cardiac Chaos* [Online]. American Institite of Physics Available: http://www.aip.org/pnu/2007/split/840-1.html [Accessed 16th January 2013].

SESSLER, P. 2007. *Traffic on the Roads and Links to the Choas Theory* [Online]. Helium. Available: http://www.helium.com/items/717680-traffic-on-the-roads-and-links-to-the-chaos-theory [Accessed 16th October 2012].

SHANG, P., LI, X. & KAMAE, S. 2005. Chaotic Analysis of Traffic Time Series. *Science Direct: Chaos, Solitons and Fractals,* 25**,** pp. 121-28.

SHANG, P., WAN, M. & KAMA, S. 2007. Fractal Nature of Highway Traffic Data. *Computers and Mathematics with Applications,* 54**,** pp. 107-116.

SHIH, E., CHO, S.-H., ICKES, N., MIN, R., SINHA, A., WANG, A. & CHANDRAKASAN, A. Physical Layer Driven Protocol and Algorithm Design for Energy-Efficient Wireless Sensor Networks. 7th Annual International Conference on Mobile Computing and Networking, 2001 Rome, Italy. Association of Computing Machinery, pp. 272-287.

SHIMADA, I. & NAGASHIMA, T. 1979. A Numerical Approach to Ergodic Problem of Dissipative Dynamical Systems. *Progress of Theoretical Physics,* 61 (6)**,** pp. 1605-1616.

SHIRER, H. N., FOSMIRE, C. J., WELLS, R. & SUCIU, L. 1997. Estimating the Correlation Dimension of Atmospheric Time Series. *Journal of the Atmospheric Sciences,* 54 (1)**,** pp. 211-230.

SIEMENS PLC 1999. *SCOOT User Guide,* UK, Siemens Plc.

SIMMONITE, B. 2005. An Introduction to Traffic Signals United Kingdom: JCT Consultancy Ltd.

SIPRESS, A. 1999. Lab Studying Science Behind Traffic Patterns. *Washington Post* August 5, 1999.

SKABARDONIS, A. 2001. ITS Benefits: The Case of Traffic Signal Control Systems. *Transportation Research Board 80th Annual Meeting.* Washington DC: Transport Research Board.

SMITH, B. L., WILLIAMS, B. M. & OSWALD, R. K. Parametric and Nonparametric Traffic Volume Forecasting. TRB 79th Annual Meeting, 2000 Washington DC. Transportation Research Board.

SOWMYA, S. & SATHYANARAYANA, S. V. Symmetric Key Image Encryption Scheme with Key Sequences Derived from Random Sequence of Cyclic Elliptic Curve Points Over GF

(p). International Conference on Contemporary Computing and Informatics (IC3I), 2014 Mysore. IEEE, pp. 1345-1350.

STERN, N. 2006. Review on the Economics of Climate Change. United Kingdom: HM Treasury.

STERRITT, R. 2005. Autonomic Computing. *Innovations in Systems and Software Engineering,* 1(1)**,** pp. 79-88.

STEWART, I. 1997. *Does God Play Dice? The Mathematics of Chaos* London, Penguin.

TAKENS, F. 1981. Detecting Strange Attractors in Turbulence. *Dynamic Systems and Turbulence: Lecture Notes in Mathematics,* 898**,** pp. 366-81.

TAYLOR, W. C. & ABDEL-RAHIM, A. S. 1998. Final Report on Analysis of Corridor Delay Under SCATS Control (Orchard Lake Road Corridor). On-line: University of Michigan, Intelligent Transportation Systems Laboratory, Technology Planning and Evaluation Group.

THE SMITH GROUP 1999. Clearing the Way - Practical Solutions for Urban Road User Charging, London. United Kingdom: The Smith Group.

TIGHT, M. R., BRISTOW, A. L., PRIDMORE, A. & MAY, A. D. 2005. What is a Sustainable Level of CO2 Emissions from Transport Activity in the UK in 2050? *Transport Policy,* 12 (3)**,** pp. 235-244.

TRB 1994. Highway Capacity Manual. *Special Report 209.* Washington DC, USA: TRB.

TRL 2013. SCOOT. *Split Cycle Offset Optimisation Technique.* United Kingdom: TRL.

TURNER, S., STOCKTON, W. R., JAMES, S., ROTHER, T. & WALTON, C. M. 1998. *ITS Benefits: Review of Evaluation Methods and Reported Benefits* [Online]. Texas Transportation Institute, Texas A & M University System. Available: http://ntl.bts.gov/lib/10000/10500/10502/1790-1.pdf [Accessed 8th June 2015].

UITTENBOGAARD, A. 2011. *Chaos Theory for Beginners* [Online]. Available: http://www.abarim-publications.com/ChaosTheoryIntroduction.html [Accessed 12th December 2012].

US CENSUS BUREAU. 2005. *Trends in Vehicle Miles Travelled* [Online]. On-line. Available: www.pewclimate.org/global-warming-basics/facts_and_figures/us_emissions/vmt.cfm [Accessed 9th September 2012].

US DOT 2013. Control and Management Concepts for Freeways. *Traffic Systems Control Handbook.* On-line.

VLAHOGIANNI, E. I., KARLAFTIS, M. G. & GOLIAS, J. C. 2005. Optimized and Meta-optimized Neural Networks for Short-term Traffic Flow Prediction: A Genetic Approach. *Transportation Research Part C: Emerging Technologies,* 13 (3)**,** pp. 211-234.

WANG, J., SHI, Q. & LU, H. The Study of Short Term Traffic Flow Forecasting Based on Theory of Chaos. Intelligent Vehicles Symposium, 6th-8th June 2005 Nevada, USA. Institute of Electrical and Electronic Engineers

WARDROP, J. G. 1952. Some Theoretical Aspects of Road Traffic Research. *Proceedings of the Institution of Civil Engineers,* Part II (1)**,** pp. 325-362.

WEBSTER, F. V. 1958. Traffic Signal Settings. *Road Research Technical Paper 39.* London: Road Research Laboratory.

WEISBROD, G., VARY, D. & TREYZ, G. 2003. Measuring Economic Costs of Urban Taffic Congestion to Business. *Transportation Research Record: Journal of the Transportation Research Board***,** pp. 98-106.

WHO 2015. Ambient (Outdoor) Air Quality and Health. *Facts and Figures (updated March, 2014).* On-Line: World Health Organisation.

WOLF, A., SWIFT, J. B., SWINNEY, H. L. & VASTANO, J. A. 1985. Determining Lyapunov Exponents from a Time Series. *Physica* 16D**,** pp. 285-317.

WOLFRAM, S. 2002. Some Historical Notes - Notes for Chapter 7: Mechanism in Program and Nature. *Section: Chaos Theory and Randomness from Initial Conditions* Wolfram Media.

XIUFEN, Y., MENG, M. Q. H., LIU, P. X. & GUOBIN, L. Statistical Analysis and Prediction of Round Trip Delay for Internet-Based Teleoperation. IEEE/RSJ International Conference on Intelligent Robots and Systems, 2002 Lausanne, Switzerland Insitute of Electrical and Electronic Engineers, pp. 2999-3004.

XU, M. & GAO, Z. 2008. Nonlinear Analysis of Road Traffic Flows in Discrete Dynamical System. *Journal of Computational and Nonlinear Dynamics,* 3 (2)**,** pp. 0212061-0212066.

XUE, J. & SHI, Z. 2008. Short-Time Traffic Flow Prediction Based on Chaos Time Series Theory. *Journal of Transportation Systems Engineering and Information Technology,* 8 (5)**,** pp. 68-72.

YAMBE, T., ASANO, E., MAUYAMA, S., SHIRAISHI, Y., SHIBATA, M., SEKINE, K., WATANABE, M., YAMAGUCHI, T., KUWAYAMA, T., KONNO, S. & NITTA, S. 2005. Chaos Analysis of Electro Encephalography and Control of Seizure Attack of Epilepsy Patients. *Biomedicine & Pharmacotherapy,* 59, Supplement 1(0)**,** pp. S236-S238.

YASIN, A. M., KARIM, M. R. & ABDULLAH, A. S. 2009. Travel Time Measurement in Real-Time Using Automatic Number Plate Recognition for Malaysian Environment. *Journal of the Eastern Asian Society for Transportation Studies* 8**,** pp. 1738-1751.

YURKON, G. T. 1997. *Introduction to Chaos and It's Real-World Applications* [Online]. Available: http://www.csuohio.edu/sciences/dept/physics/physicsweb/kaufman/yurkon/chaos.html [Accessed 28th February 2013].

ZELLNER, A. 1962. An Efficient Method of Etimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American statistical Association,* 57 (298)**,** pp. 348-368.

ZHANG, G., PATUWO, B. E. & HU, M. Y. 1998. Forecasting with Artificial Neural Networks: The State of Art. *International Journal of Forecasting,* 14**,** pp. 35-62.

ZHANG, J. & LIU, J. Non-linear Characteristics of Short-term Traffic Flow and Their Influences to Forecasting. IEEE International Conference on Automation and Logistics, 2007 Jinan, China. Institute of Electrical and Electronic Engineers, pp. 847-851.

ZHENG, P. & MCDONALD, M. 2007. Optimal Traffic Data Archive Scheme. *Intelligent Transport Systems, IET,* 1 (20)**,** pp. 144-149.

ZHU, C., XU, X. & YAN, C. The Research of Method of Short-Term Traffic Flow Forecast Based on GA-BP Neural Network and Chaos Theory. 2nd International Conference on Information Science and Engineering (ICISE), 4-6 Dec. 2010 Hangzhou, China. Institute of Electrical and Electronic Engineers, pp. 1617-1620.

Appendices