MODELLING AND INFERENCE FOR DYNAMIC TRAFFIC NETWORKS

A thesis presented for the degree of
Doctor of Philosophy
in
Statistics
at Massey University, Palmerston North,
New Zealand.

Ahmad Mahmoodjanlou

2020

Contents

List of Tables

List of Figures

Abstract

Nowadays, traffic congestion is a significant problem in the world. With the noticeable rise in vehicle usage in recent years and therefore congestion, there has been a wealth of study into possible ways that this congestion can be eased and the flow of traffic on the road improved. Controlling traffic congestion relies on good mathematical models of traffic systems. Creating accurate and reliable traffic control systems is one of the crucial steps for active congestion control. These traffic systems generally use algorithms that depend on mathematical models of traffic.

Day-to-day dynamic assignment models play a critical role in transport management and planning. These models can be either deterministic or stochastic and can be used to describe the day-to-day evolution of traffic flow across the network. This doctoral research is dedicated to understanding the difference between deterministic models and stochastic models. Deterministic models have been studied well, but the properties of stochastic models are less well understood. We investigate how predictions of the long term properties of the system differ between deterministic models and stochastic models.

We find that in contrast to systems with a unique equilibrium where the deterministic model can be a good approximation for the mean of the stochastic model, for a system with multiple equilibria the situation is more complicated. In such a case even when deterministic and stochastic models appear to have comparable properties over a significant time frame, they may still behave very differently in the long-run.

Markov models are popular for stochastic day-to-day assignment. Properties of such models are difficult to analyse theoretically, so there has been an interest in approximations which are more mathematically tractable. However, it is difficult to tell when approximation will work well, both in a stationary state and during transient periods following a network disruption.

The coefficient of reactivity introduced by Hazelton (2002) measures the degree to which a system reacts to a disruption. We propose that it can be used as a guide to when approximation models will work well. We study this issue through a raft of numerical experiments. We find that the value of the coefficient of reactivity is useful in predicting the accuracy of approximation models. However, the detailed interpretation of the coefficient of reactivity depends to a modest degree on properties of the network such as its size and number of routes.

The experiments discussed in the previous paragraph are restricted to Markov assignment models with short-range memory. The reason is that Hazelton's coefficient of reactivity does not properly account for historical variation in flows for longer memory, nor can it be applied to systems undergoing disruptions lasting longer than one day. We therefore seek to generalize the coefficient of reactivity in two different directions. First, we propose a new definition that does account for variation in historical flows. However, we find that both theoretical evaluation and simulation-based computation are extremely difficult. Nevertheless, we are able to prove an asymptotic equivalence to Hazelton's original definition, which suggests that this original definition may be used more widely than previously thought. Second, we extend the original definition of the coefficient of reactivity to allow for disruptions to the system of arbitrary duration. We illustrate these generalizations using various numerical examples.

Acknowledgements

I would like to express my heartfelt gratitude to my supervisor, Professor Martin Hazelton. It has been an honour for me to work as a Ph.D student with Martin. He is one of the smartest people I know. I appreciate all his contributions of time, ideas, and funding to make my Ph.D. This thesis could not have been completed without his caring guidance, and I feel extremely fortunate to have had the opportunity to work with him. I am indebted to Martin and this association with him will forever be etched in my memory. I could not have imagined having a better advisor and mentor for my Ph.D study.

I express my sincere thanks to my co-supervisor Dr. Katharina Parry. Besides her insightful guidance on my research, her support and understanding during my hardest times helped me go through this tough but rewarding journey. Thank you, Rina, for supporting me during this project.

I also thank my friends (too many to list here but you know who you are!) for providing support and friendship that I needed.

I would like to thank my family back in Iran for their love and support. Last but certainly not least, special thanks go to my dearest wife, Ghazaleh. Thank you so much for your caring, loving and understanding, which always make me feel cherished and protected. Thank you for always cheering me up and staying with me. Without your supports and love, I would not have reached what I have achieved so far.

Publication arising from this thesis

Ahmad Mahmoodjanlou, Martin L Hazelton, and Katharina Parry. Apples versus oranges? comparing deterministic and stochastic day-to-day traffic assignment models. Transportmetrica B: Transport Dynamics, 7(1):1426-1443, 2019.

Chapter 1

Introduction

The increasing population coupled with relative economic prosperity has produced a demand for travel that overloads existing traffic and transit facilities, causing the travelling public to experience unacceptable discomfort and delay. Increasing the number of lanes on roads, increasing the number of buses on routes, etc. can be an obvious solution. Due to the high economic cost of such expansion, it can be implemented only in a very limited fashion.

Other available options to avoid congestion are to increase the system capacity at the congested point or to manage the demand by shifting it in place or time. However, it has been shown in Sheffi (1985) that a naive increase in system capacity can result in a worse situation than existed before the intervention. Another naive method of traffic management is by providing travellers with information on traffic conditions. If too many travellers have very recent information about traffic conditions then the previously underused routes become overloaded. This happens because travellers attempt to minimise their individual route costs. Horowitz (1984) identified this type of instability in simple two-route networks. Ben-Akiva et al. (1991) stated that providing public information about traffic congestion can lead to unpredictable results because it may aggravate travellers coordination problems. The experimental results in Ziegelmeyer et al. (2008) showed that providing more information to drivers about past congestion levels does not significantly decrease congestion levels. Therefore, the reaction of travellers should be taken into account when trying to ease congestion.

Mathematical models of traffic flows are incredibly important for the design and modifications of the transport system. Traffic assignment models play a critical role for traffic management and urban planning. These models translate travel demand through the network into flows and travel times on the individual routes and links in the system, and hence can be used to forecast travel patterns under hypothetical scenarios (e.g. He and Liu 2012; Patriksson 2015). These might correspond to long-term network improvements or adverse events like the failure of a bridge, or to shorter-term network control measures such as the imposition of tolls or refinement of traffic signalling schemes.

The oldest types of traffic assignment models were based on the idea of network equilibrium (e.g. Wardrop; Beckmann et al. 1956; Sheffi 1985; Florian and Hearn 1995; Patriksson 2015). This approach completely ignores the variable nature of the traffic flows, which is inherent to any transport network system.

In fact, the traffic flows on roads have considerable variation across the days. Day-to-day traffic assignment models examine the evolution of traffic flows towards equilibrium. These models determine travellers' route choice behaviour on a given day based on their past experience through a learning and adjusting mechanism. Day-to-day dynamic assignment models are the focus of this thesis. These models can be either deterministic or stochastic in nature. Deterministic models are most widely used in practice by transport planners and traffic engineers. However, when observing traffic systems it can be observed that travel behaviour can appear random with flows varying from one day to the next. It is clear that the question of variability in traffic flows cannot be fully addressed within the framework of conventional deterministic models as it essentially predicts a single outcome given a set of circumstances. By contrast, a wider range of modelling options are now becoming available with stochastic models which represent the traffic flow as a random variable, and such models predict a set of outcomes weighted by their probabilities (e.g. Daganzo and Sheffi 1977; Cascetta 1989). While stochastic models have advantages in describing observed variability, deterministic models can provide a description of "average" behaviour, and are typically much easier to analyse mathematically.

Comparison of deterministic and stochastic traffic assignment models has received surprisingly limited attention in the research literature. Davis and Nihan (1993) studied the properties of stochastic models as travel demand and network capacity become large. In such cases a version of the Law of Large Numbers can be applied to give a deterministic dynamical model that approximates the (mean) behaviour of the stochastic model. The linkage between dynamic Markov models, dynamic deterministic and static equilibrium models has been examined by Cantarella and Cascetta (1995). When Watling and Hazelton (2003) investigated the relationship, they found that as demand increases in the network system, a Markov dynamic model converges asymptotically towards the SUE. Smith et al. (2014) established long-term behaviour of day-to-day traffic assignment models that illustrate, compare and contrast the attributes of deterministic models and stochastic models.

The overall aim of this research project is to gain a better understanding of when two types of traffic assignment models deliver the same conclusions and when they present different stories. In particular, it is important to know if and when deterministic models provide a good approximation to the mean behaviour of stochastic models. Some of those works we have cited partly answer this question. However, (i) they rely on various assumptions (e.g. unique fixed point equilibrium for deterministic models); and (ii) the degree of agreement between a stochastic and deterministic model can depend on what properties you are looking at. Sometimes two models may look quite comparable when describing the short-term behaviour of the system, and yet provide very different representations of long-term behaviour.

Within this framework, Chapter 2 reviews the foundational ideas and methods in transportation research that are used in this thesis.

Previous work has usually focused on comparison of stochastic and deterministic models when the latter has a unique fixed point. Chapter 3 examines what happens in multi-equilibria systems. It places an emphasis on comparing the long-run behaviour of the stochastic and deterministic models. This has important implications for network control. Because the comparison of the stochastic and deterministic model allows us to understand the long-run future of system after some intervention, it is important to know whether the two models will be similar or they will differ in terms of some qualitative outcomes. In this chapter we prove new theoretical results, demonstrating that deterministic and stochastic models can give hugely different estimates of time to reach a desired state.

Typically deterministic models provide a good description of mean behaviour of stochastic systems in which the variation isn't too extreme, even in response to unusual events. However, due to the random fluctuations of traffic flow from day-to-day, stochastic models have the potential to provide a deeper and more complete understanding of the dynamics of the traffic system. Since full day-to-day stochastic models are difficult to analyse mathematically, approximation methods have been developed which are mathematically tractable, computationally more efficient and (hopefully) valid when the travel demand is large. When a deterministic model provides a good approximation to the mean behaviour, we can hope to incorporate this in an approximation of the stochastic model based on a Gaussian auto regressive process. Hence Chapter 4 begins with the review of two different stochastic approximation methods expressed by Davis and Nihan (1993) and Watling and Hazelton (2018) and one approximation method for deterministic model introduced by Cantarella and Cascetta (1995).

Most of the existing work into the quality of approximation methods has focussed on systems in a stationary state. However, it is also important to understand the transient behaviour of a network, for example following some disruption. As a result of this, Chapter 4 presents the definition of coefficient of reactivity, introduced by Hazelton (2002), that measures the stability of a system following a disruption. We hope that the coefficient of reactivity can provide guidance as to when the system is sufficiently stable for such an approximation to work well. In the following, this chapter examines this issue through a string on numerical experiments of a variety of networks. The results indicate that the value of coefficient of reactivity is effective in predicting the accuracy of the approximation models. However, it must be noted that detailed interpretation of the coefficient of reactivity depends on factors like the size of the network and the number of available routes.

In Chapter 5 we first show how the original definition of the coefficient of reactivity works when the length of memory of travellers is more than one which has not been covered in Chapter 4. As stated in Chapter 4, the original definition of the coefficient of reactivity assumes the flow for the days before the disruptions are in a fixed equilibrium state. There are two issues associated with this definition. First, it does not cover properly the variation of the historical flows. Additionally, it does not apply to systems that consider the duration of disruptions longer than one day. Accordingly, we seek to generalize the original definition of the coefficient of reactivity in two directions.

Firstly, a new definition is provided in which the historical flows can vary arbitrarily. In the following, we describe an algorithm to compute the coefficient of reactivity using the new definition. However, it is shown that the implementation of this definition is difficult both analytically and computationally. In spite of this, we prove that when the population size is large enough the new definition and the original one are equivalent. This suggests that the original definition of the coefficient of reactivity still can be used. Secondly, we examine the impact of the extension of disruptions to the system through an extended definition of the coefficient of reactivity.

Finally, Chapter 6 summarises the findings of the dissertation and provides some suggested directions for future research.

Chapter 2

Review of Methods and Techniques

2.1 Chapter overview

This chapter lays the groundwork for succeeding chapters with an overview of some definitions and notations about traffic networks. This chapter also includes an introduction to Markov chains, a type of stochastic model used to describe how traffic flows evolve over time.

2.2 Representation of the traffic network

A transportation system can be defined as a set of elements (road, vehicles, ...) and the interaction between them which are necessary for the movement of passengers or goods. A traffic network model provides a simplified view of the real transportation system. The physical structure of traffic network is represented by a conceptual directed graph, that is, a set of points called nodes, and a set of directed links connecting the nodes. Each node can represent an origin and/or destination of traffic flow, or simply a road intersection. Each link represents a segment of road, connecting a pair of nodes. The network represents the possible routes through a sequence of links from an origin to a destination. An origin node and a destination node is called an OD pair and will be active if there exists demand for travel between the pair.

Mathematically speaking, let $\mathcal{L}$ denote the set of links, $\mathcal{N}$ be the set of nodes and $\mathcal{J}$ be the set of all OD pairs. Let N_j be the traffic demand between OD pair $j \in \mathcal{J}$. We denote by g the total number of routes and by n the total number of links. The link-path relationship is characterised by the incidence matrix $A = (a_{lr})$ with $a_{lr} = 1$ if link l is part of route r, and $a_{lr} = 0$ otherwise for $1 \leq l \leq n$ and $1 \leq r \leq g$. Each route serves a given origin-destination (OD) pair. We write $\mathcal{R}_j$ for the set of routes connecting OD pair $j \in \mathcal{J}$, and define h to be the total number of OD pairs.

Let $\boldsymbol{y}$ denote the vector of link flows and $\boldsymbol{x}$ denote the vector of route flows. We have the relation between link flows and route flows through $A\boldsymbol{x} = \boldsymbol{y}$. Let x_r denote the volume of traffic (number of travellers) on route r. Each link $l \in \mathcal{L}$ in the network has associated with it a cost function c_l. The function $c_l(y)$ gives the cost of traversing the link as a function of the current link flow ((Sheffi 1985),(Cascetta 2009)). Usually due to congestion, the travel cost is an increasing function of flow. Therefore, a performance function rather than a constant travel cost measure should be associated with each of the links of the network.We use a commonly used link cost function which is given by the Bureau of Public Roads (BPR) (Branston 1976). For each link, the link performance function follows the (BPR) function

$$c_l(y_l) = c_l^0 \left[1 + \alpha \left(\frac{y_l}{b_l}\right)^{\varphi}\right]. \tag{2.1}$$

Here c_l^0 represents the free flow cost, y_l denotes the link flow, b_l is commonly referred to a capacity of a link but for the models in this thesis the flow on each link does allow to be larger. α and φ are the calibration parameters that determine the rate at which travel costs increase with traffic volume. The route cost is just the sum of costs on the links forming that route. Let $A_{[r]}$ be the r-th column of incidence matrix, then the cost traversing on route r, is given by

$$C_r = \sum_{l=1}^{n} A_{[r]}^T c_l(\mathbf{y}), \quad r = 1, \ldots, g \tag{2.2}$$

where T represents the transpose operator.

An example of a network with different routes connecting different OD pairs is depicted in Figure 2.1

- Link set $\mathcal{L}$=$\{(O_1, A), (A, D_1), (O_1, B), (B, D_1), (O_2, A), (A, B), (B, D_2)\}$
- Set of OD pairs $\mathcal{J}$=$\{(O_1, D_1), (O_2, D_2)\}$
- Node set $\mathcal{N}$=$\{O_1, D_1, O_2, D_2, A, B\}$
- Route set $\mathcal{R}_1$=$\{(O_1, A, D_1), (O_1, A, B, D_1), (O_1, B, D_1)\}$ and $\mathcal{R}_2$=$\{(O_2, A, B, D_2)\}$
- Link-path incidence matrix

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

In addition we can define another binary matrix Γ called the OD-pair-route matrix with $\Gamma_{jr} = 1$ if route r connects OD pair j and 0 otherwise. If a network has only one OD pair, Γ becomes a vector of 1's.

- OD-pair-route incidence matrix

$$\Gamma = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

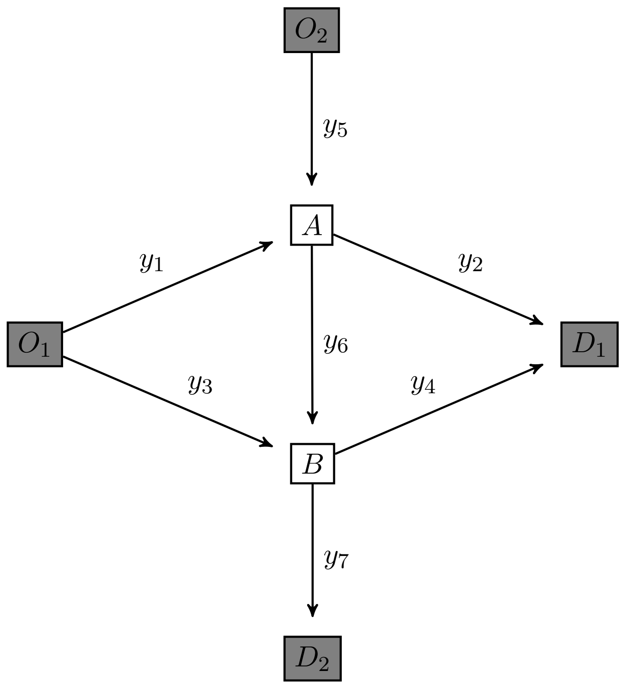


Figure 2.1: Illustrative network, with two OD pairs, six nodes and seven links.

2.3 Traffic assignment models

Mathematical models and computer tools can be used in traffic planning processes to identify problems, generate and evaluate possible solutions, and develop plans for design and modifications of transport systems. Transportation planning is a process that involves the analysis of the current pattern of travel and its forecast, to determine indicators of the use of transport infrastructure to make decisions in the future. Traditionally, this process is based on the classic transport planning model, which is presented as a four-step sequence that includes trip generation, trip distribution, mode choice, and traffic assignment. Through these steps we can describe how many users are going to travel, where are they going to go, what travel mode is used for each trip, and what route will they take when going from one zone to another(Sheffi and Daganzo 1978), (Patriksson 2015).

The first three sub-models are related to the forecast of travel demand, while the sub-model traffic assignment refers to the process of allocating a given set of trips to the specified transportation system. These models translate travel demand through the network into flows and travel times on the individual routes and links in the system, and hence can be used to forecast travel patterns under hypothetical scenarios (Daganzo and Sheffi 1977). These might correspond to long-term network improvements or adverse events like the failure of a bridge, or to shorter-term network control measures such as the imposition of tolls or refinement of traffic signalling schemes. Assignment model outputs describe the state of the system, or rather the mean states and its variations, therefore these models play a pivotal role in comprehensive transportation system models (Cascetta 2009).

The fundamental aim of traffic assignment is to determine in what way the users are distributed among the possible routes associated with a given origin-destination pair. For instance, for the network in Figure 2.1 we may know that the (mean) travel demand from O_1 to D_1 is 10 travellers. Each traveller must choose their travel route from O_1 to D_1 from the possibilities listed earlier. Assume that four travellers have chosen route 1, three of them route 2 and the rest of travellers selected the third option for their trip. The selection of travel routes for all travellers then defines the traffic assignment over the network. This assignment is not generally unique since a pattern of link flows does not necessarily define a unique set of route flows. In Equation 2.3 we demonstrate the linear relationship between route flows and link flows for the network displayed in Figure 2.1.

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_1 \\ x_3 \\ x_2 + x_3 \\ x_4 \\ x_2 + x_4 \\ x_4 \end{pmatrix}. \tag{2.3}$$

Equation 2.3 shows that assignment model is concerned with modelling how travellers choose routes, but this manifests itself in a pattern of link flows.

2.3.1 Equilibrium Models

The bulk of the early work on traffic assignment models was based on the assumption that the traffic system stays in the same state over time (Wardrop; Beckmann et al. (1956); Dafermos and Sparrow (1969); Bar-Gera (2002); Florian et al. (2009)). Interactions between congestion and travel decisions is modelled as a process of reaching an equilibrium. In the following, two kinds of equilibrium models will be reviewed. There are deterministic user equilibrium (DUE) and stochastic user equilibrium (SUE). These models are by far the most well studied traffic equilibrium models.

Deterministic user-equilibrium:

This method is based on Wardrop's first principle stating that no driver can unilaterally reduce travel costs by shifting to another route (Wardrop). From this point of view, no further route switching is feasible, and the traffic system is in a state of equilibrium. If C_j^* denotes the minimum cost of travel between OD pair j, C_r be the cost of route r connecting OD pair j where r is an element of $\mathcal{R}_j$, the set of routes for OD pair j, then the condition for Wardrop equilibrium can be stated as :

$$C_r > C_j^* \quad \text{implies} \quad x_r = 0, \quad \text{and} \quad x_r > 0 \quad \text{implies} \quad C_r = C_j^*. \tag{2.4}$$

An assignment of travellers to routes which satisfies condition 2.4 is called deterministic user equilibrium (DUE). In DUE, users are assumed to have full knowledge of the traffic conditions and to seek the minimum travel time or cost. Travel time (i.e. cost) depends on level of congestion, which in turn depends on route choices. Every traveller wants to follow a cheap route, but if all choose the same one then it will be congested and so there is a benefit in switching to an alternative. This can be thought of as a kind of game between travellers. Indeed, DUE is a form of Nash equilibrium (named after the famous mathematician/economist John Nash)(Charnes and Cooper 1958).

Due to the fact that no explicit assumptions are made concerning either the dynamics of route choice or the probabilistic principles governing traffic generation, this can be an advantage and a weakness of DUE. On the one hand, employing DUE simplifies the analytic task because, when computing the resulting equilibrium, researchers do not have to worry about how traffic population reaches equilibrium. On the other hand, DUE makes no assumptions about dynamics of the system and this means that there is no guidance on the stability of the equilibrium reached. Also, the lack of assumptions concerning the probabilistic properties of equilibrium link flows, coupled with the fact that actual link counts only very roughly approximate their supposed equilibrium values makes it difficult to ascertain whether or not a traffic system is actually in equilibrium (Cascetta 1989). One of the shortcoming of DUE is that it assumes perfect knowledge and identical behaviour for all travellers. If it is not the case then we expect to see usage of sub-optimal routes, which leads us to discuss stochastic user equilibrium (SUE).

Stochastic user-equilibrium:

SUE is an extension of the DUE model, first introduced by Daganzo and Sheffi (1977). SUE is defined as the state where no traveller can improve his or her *perceived* travel cost by unilaterally changing routes. It should be pointed out that the assumption that all the travellers have complete and accurate information about the entire network before their trips is unrealistic even if the travellers have long-term experience with the network, due to the daily variations of travel times and the diversity from travellers' sense of time. This might provide users with knowledge of the system, but (a) not everyone will have it, and (b) it will lead to changes in flow patterns as travellers react to varying congestion, and so an equilibrium may not eventuate and be maintained.

In contrast to DUE, it is assumed that travellers information is not perfect and they think the route they have chosen is the lower cost route, thereby, some use of higher cost routes would be expected. In particular, users are assumed to make errors in their route choice.

This perceived travel cost can be represented by $\hat{C}_r = C_r + \varepsilon_r$ where, C_r is the actual travel cost and ε_r denotes the random perception error term. SUE conditions are achieved when the probability a traveller uses a given route is equal to the probability that the given route is perceived as the cheapest, i.e.

$$p_r = \boldsymbol{p}\{C_r + \varepsilon_r \leq C_s + \varepsilon_s, \quad \text{for all} \quad s\} \tag{2.5}$$

where p_r denotes the probability a traveller chooses route r from the set of routes between OD pair j. For the case where travel demand between OD pair j is a constant N_j, the flow on route r would be $x_r = p_r N$.

It needs to be noted that SUE is a deterministic model, because the route choice probabilities generated are calculated as proportions of travellers taking the various routes. The resulting flows are represented as continuous variables. The word "stochastic" in the name of the SUE model emphasises the model's assumption of stochastic perception errors.

DUE is a special case of SUE if we assume users' perception of travel time is not subject to errors, i.e. $\varepsilon_j = 0 \Rightarrow \hat{C}_r = C_r$. From Equation 2.5 it is clear that the route choice probability will depend on the probability distributions governing the perception error terms. The two models most commonly used in transportation planning assume that the error terms follow Gumbel or Normal distributions. As a result they produce logit or probit models respectively, that are described in Section 2.5.

We demonstrate the difference between DUE and SUE with a simple example.

Example 2.3.1. Consider the following network:

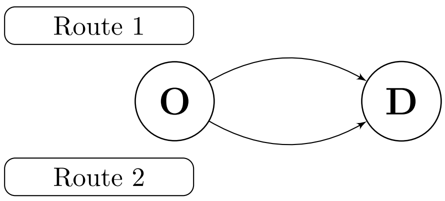


Figure 2.2: An example with two-link network.

Let x_1 and x_2 represent the traffic flow on these routes. Assume that the route choice proportion for route 1 is defined as

$$p_1 = \frac{\exp\left[-C_1(x_1)\right]}{\exp\left[-C_1(x_1)\right] + \exp\left[-C_2(x_2)\right]},$$

where the cost function for each route is $C_1(x_1) = 2 + 3x_1$, $C_2(x_2) = 1 + 5x_2$ and the total travel demand is equal to 2 ($x_1 + x_2 = 2$).

When computing the DUE solution we use the fact that for each O-D pair, the travel cost on all used paths is equal, so that no traveller has a cheaper alternative route available. Hence,

$$C_1(x_1) = C_2(x_2) \Rightarrow 2 + 3x_1 = 1 + 5x_2$$

since $x_2 = 2 - x_1$, this yields :

$$2 + 3(2 - x_2) = 1 + 5x_2.$$

This equation can be solved numerically to obtain

$$x_2 = \frac{7}{8} = 0.875 \quad , \quad x_1 = \frac{9}{8} = 1.125.$$

The model DUE flow on the first route is 1.125 with the flow on route 2 being 0.875.

When calculating the SUE solution the route choice model is defined as:

$$p_1 = \frac{\exp\left[-C_1(x_1)\right]}{\exp\left[-C_1(x_1)\right] + \exp\left[-C_2(x_2)\right]} = \frac{1}{1 + \exp\left[C_1(x_1) - C_2(x_2)\right]}$$

Therefore in our example

$$p_1 = \frac{1}{1 + \exp\left[1 + 3x_1 - 5x_2\right]} = \frac{1}{1 + \exp\left[8x_1 - 9\right]}$$

Solving algebrically:

$$\frac{x_1}{2} = \frac{1}{1 + \exp\left[8x_1 - 9\right]}$$

$$x_1 \simeq 1.10 \Rightarrow x_2 \simeq 0.90.$$

The solution of this equation can be found, numerically, to be $x_1 \simeq 1.10$ meaning that $x_2 \simeq 0.90$. Since $C_1(x_1) \neq C_2(x_2)$ at equilibrium, it then implies that the network is in SUE not DUE.

As we have shown in this example the two equilibrium models are not the same. Even though all the users intend to minimise their perceived travel costs, the perceived travel costs on all the used paths are not equal. Instead, each route is only personally perceived by the users on it to be the shortest one among all the alternatives.

Uniqueness of stochastic user-equilibrium

Uniqueness of the equilibrium in traffic assignment is guaranteed when the condition of monotonicity of the link cost function is established (e.g. Smith (1979); Daganzo (1983);Patriksson (2015)). The failure of this condition does not necessarily mean that there are multiple solutions. However, it is not difficult to find examples where the solution is not unique. To illustrate this issue we consider the same structure as Example 2.3.1 for our network with different cost functions and a different route choice probability function.

Example 2.3.2. Following Smith et al. (2014), travel costs and route choice probability in this model for each route are to be as follows:

$$C_1(\boldsymbol{x}) = 8 - \frac{8x_1}{N}$$

and

$$C_2(\boldsymbol{x}) = 2 + \frac{4x_2}{N}$$

where $x_1 + x_2 = N$ and $p_1 = \frac{1}{1 + \exp\left[-\beta(C_2(x_2) - C_1(x_1))\right]}$. The cost function for each route depends only on the number of travellers using that route. We will look at this example in much more detail in Chapter 3.

In this probability model β is a cost sensitivity parameter. The cost sensitivity influences how sensitive travellers are to cost differences. As β increases, the sensitivity of travellers to the cost increases.

It is important to note that in the first cost function by increasing the number of travellers the cost goes down. From a traffic management point of view, this cost function can be related to the use of public transport. This could relate to lower ticket prices or increased frequency of service for higher demands. This makes sense because it is intended to encourage people to use public transport. In contrast, the cost goes up when more travellers use the second route because the route becomes congested. Consider the case where $N = 10$, for simplicity of exposition. Solving the SUE equation 2.6 numerically, for some fixed value of β, different results for x_1 can be achieved.

$$\frac{x_1}{x_1 + x_2} = \frac{1}{1 + \exp\left[-\beta(C_2(x_2) - C_1(x_1))\right]} \quad \text{where} \quad 0 \leq x_1, x_2 \leq 10. \tag{2.6}$$

As can be seen here the cost functions are not monotonic increasing, nevertheless, for $0 \leq \beta \leq 1$ we have a unique solution for SUE which is $x_1 = 5$. On the other hand, when $\beta > 1$ we start to get multiple solutions for SUE. For instance, when $\beta = 2.1$ then the associated route flows are given by the solutions $x_1 = 0.170$, $x_1 = 9.830$ and $x_1 = 5$. The plot 2.3 displays the numerical and analytical solutions of our example.

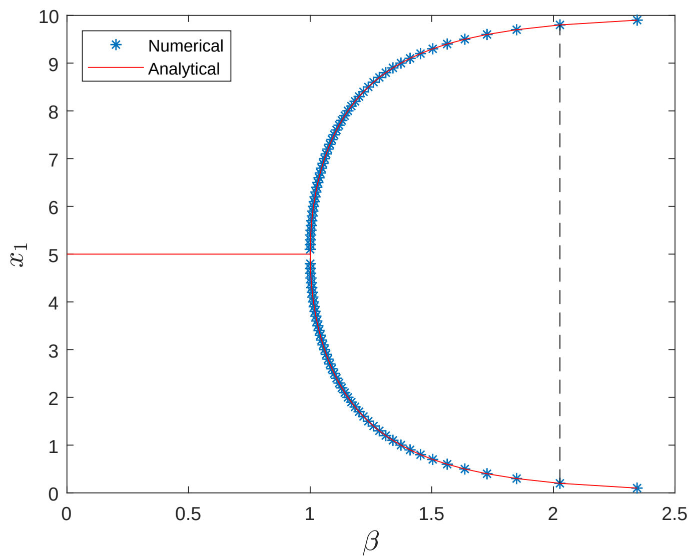


Figure 2.3: Solutions for SUE for a two route network, expressed in terms of the flow x_1 on route 1.The solutions are plotted against β, the cost sensitivity parameter in the SUE assignment model.

Partial Users' Equilibrium

One possible model which can be used to consider travellers' adaptive route choice behaviours under disruptions is the partial user equilibrium (PUE). In PUE user behaviour is characterized as partially adaptive and assumes that only those who are affected are likely to reconsider their original route decisions. This concept of a PUE was introduced in Watling et al. (2004). Here, the underlying assumption is that travellers' on the affected routes aim to find a new UE, while the flows on unaffected routes remained fixed. The resulting flow pattern is deterministic, and can be computed as the solution to a system of linear equations.

2.3.2 Deterministic versus Stochastic

As stated earlier SUE is a deterministic model but the assumptions on traveller's behaviour are stochastic. The classification of the model depends on whether the assignment is a fixed pattern, or a random vector. A deterministic model assumes that starting from the same initial conditions leads to the output of the model being the same. Whereas, as there is at least one random variable in stochastic traffic models, the same initial conditions for starting a simulation of these models may give different results.

Traditionally most traffic network models assume that traffic statistics (for example, travel cost or traffic flows) are deterministic, largely because deterministic models are theoretically and computationally tractable and can produce useful predictions for the purpose of the study. From a transport planning point of view, a deterministic model will be useful for making accurate predictions in the future when haphazard variation is not an important aspect of the network system behaviour.

Nevertheless, in reality traffic networks are subject to random fluctuations. For example, traffic accidents, bad weather, or the fluctuations in the number of total travellers could be the result of random fluctuations. Deterministic models are insufficient to describe these fluctuations on the traffic network. From this point of view, by performing Monte Carlo simulation Cantarella and Cascetta (1995) have shown a similar pattern in link flows for deterministic and stochastic processes with more variation for stochastic ones. Watling (2002) states that stochastic models are more useful when we want to have an idea of the variability in traffic flows as well as the degree of uncertainty in the estimated average flow.

Sheffi and Powell (1981) compared the stochastic and deterministic traffic assignment over congested network. The evolution of traffic flows as a discrete-time deterministic process for the first time appeared in Horowitz (1984). Sheffi (1985) proposed the well-known deterministic and stochastic user equilibrium models by combining the equilibrium approach with the within-day constant demand. The evolution of traffic flows as a discrete-time stochastic process was formulated in Cascetta (1989). Davis and Nihan (1993) explored the relationship between deterministic and stochastic models. They show that, subject to certain technical conditions, stochastic models can be approximated by deterministic models when the quantity of vehicles utilising of the system is permitted to be infinitely large. Watling and Cantarella (2013) proposed several stochastic process models to examine the way in which these models may be used to represent various source of variation in traffic networks. A general deterministic and stochastic modelling approach was presented in Cantarella and Watling (2016) by considering travellers' habits in route choice behaviour.

2.3.3 Static versus Dynamic

Another fundamental distinction between models used for traffic assignment models is that they can be static or dynamic in nature. They are classified according to whether or not they consider a time variable. Static models provide a description of one steady state of the transport system, which is usually representative of the average operating conditions of the system. Static traffic assignment models are used to describe the state of a transportation system as an aggregation of road users individual decisions. DUE and SUE are examples of static models. Static assignment also implies that the origin-destination demand is constant over time (Patriksson 2015).

The limitations of static traffic assignment models have motivated researchers to develop dynamic traffic assignment (DTA) models to adequately represent traffic reality and drivers' behaviour.

As stated, the term dynamic is utilised if a traffic assignment model calculates costs and flows that vary with time. These models are in general much less tractable compared to the static models, both theoretically and computationally. In dynamic process assignment models the evolution of the system state is explicitly simulated based on the mechanisms underlying path choice and information acquisition, which in turn specify user choices in successive reference periods. In these models, the evolution of the system is simulated over a sequence of similar periods (usually identified as day Cascetta (2009)), and the possible convergence of the system over time to a stable condition. Dynamic processes are based on time discrete dynamic system theory or on stochastic process theory, considering the state of the system which could be described by deterministic or stochastic variables respectively. For a more comprehensive review of DTA models, the reader may refer to Peeta and Ziliaskopoulos (2001) and Szeto and Lo (2006).

Two directions of dynamic traffic models exist. First, day-to-day refers to models where travellers typically do daily trips, and learn over time about the choices that they make, like when to start, and possibly also which routes to utilise for habitual trips.

Day-to-day dynamic problems are concerned with how the travel decisions of travellers change over days and how their route choice or departure time on a particular day depend on their experience obtained in previous days, see Watling and Hazelton (2003) and Cascetta and Cantarella (1993). Since our focus in this thesis is on day-to-day dynamic traffic assignment models, we will discuss them in more detail in Chapter 3.

Within-day models focus not so much on habits, but on the build-up of queues and other delay-inducing phenomena in the traffic network. Here, the focus lies on the “loading” of the network, that is, how congestion emerges, and how trip-makers make instantaneous decisions based on that. Models with within-day dynamics allow us to study the dynamics of the systems of transport within the reference time interval; they have the advantage of explicitly considering the effects of congestion, as well as all the strategies of control and information of users implemented in real time. Using these models, the adaptive behavior of choice of the route can then be specified based on the attributes that affect the individual trips, such as schedules, scheduled service, delays and information available; in this way, such behaviour is no longer constant in the interval and depends only on the configuration of the lines, but may vary depending on the time characteristics of the individual trips and offer therefore a representation closer to the real behaviour of users.

The transition from static models to models with within-day dynamics is not immediate; it is necessary to introduce the temporal dimension in all the parts of the models, through a discrete time variable that divides the reference interval into a certain number of sub-intervals and allows them to be analysed. The choice of discrete time for modelling is consistent with the nature of the collective transport services, which are discontinuous in time and therefore would not be correctly represented by continuous variables.

Within-day problems include pure departure time choice problems (Bellei et al. 2006), pure route choice problem (Bogers 2009).Peeta and Ziliaskopoulos (2001) provides a comprehensive review of the within-day dynamic traffic assignment models.

For long term planning purposes, within day models tend to be unnecessarily complicated, computationally expensive and difficult to calibrate. Day-to-day models are much more tractable, but can still hope to represent the critical factors in the evolution on the traffic system.

In the transportation literature, there exist another model called the doubly dynamic traffic assignment model, that can be obtained by combining within-day and day-to-day dynamic traffic assignment models. Literature on doubly dynamic traffic assignment models include Cascetta and Cantarella (1991) and Balijepalli and Watling (2005).

2.4 Route guidance system

Route guidance system refers to a system in which individual vehicles are equipped with devices which communicate information on the best route for that particular vehicle's (user-requested) movement (Watling and Van Vuren 1993). Route guidance system can be classified as a static and dynamic system. Static route guidance system does not respond to traffic conditions actually experienced at that time. In contrast, dynamic route guidance system can provide routing suggestions to users in accordance with current traffic conditions. The information providing mechanism is decided by the current position of users, the availability of information and the destination node.

The route guidance system can impact the travellers' route decision by providing them with useful information regarding the traffic states of the urban regions. Therefore, drivers can follow a series of subregions that has lower cost (in terms of travel time, fuel consumption, etc.), which might lead to a better overall system performance.

The whole philosophy of the day-to-day models is based on the fact that travellers make a decision on the current day based on the memory they have experienced before. These models do not take account of describing what happened during the day. But with route guidance system drivers might ignore what happened in the past and can make their decisions based on the information they have received form the system. The question remains open as to whether day-to-day models can provide robust results for systems in which dynamic route guidance is widely used.

2.5 Route choice models

Route choice models are an essential component of traffic assignment models. Through the route choice model, it can be shown how a traveller chooses one and only one option amongst several mutually exclusive set of alternatives based on expected travel costs. These alternatives can for example be which travel mode to chose or which route to travel.

There are two types of choice, continuous choice and discrete choice (Mannering and Hensher 1987). Since in transportation the options considered normally come from a discrete set, discrete choice models are more commonly used to describe and predict a traveller's choice.

Discrete choice models assume the selection of one alternative from the choice set, based on random utility theory. Random utility theory assumes that a traveller has a set of available choice options and attaches to each of them an unobserved characteristic known as a utility.

A utility function measures the degree of satisfaction that travellers derive from their choice. All travellers evaluate the trade-off between alternatives and choose the alternative that benefits themselves the most. Since utilities are observed only through samples, they must be represented in general by a random variable. This random variable can be written as the sum of measurable utilities and error terms in linear form as:

$$\mathbf{U}(\mathscr{X}) = \mathbf{V}(\mathscr{X}) + \boldsymbol{\varepsilon} \tag{2.7}$$

where $\mathscr{X}$ is a vector of attributes of the alternatives (set of options which can be the possible routes or travel modes like bus, subway, private car or car shared for the daily trip to the workplace), ε is a random variable that is used to account for the unobserved attributes and $\mathbf{U}$ is a perceived utility. The systematic utility $\mathbf{V}$ can be computed as a linear function of the attributes of the alternatives i.e. $\mathbf{V}(\mathscr{X}) = \boldsymbol{\eta}\mathscr{X}$ where $\boldsymbol{\eta}$ is a vector of parameters to be estimated from observations. We will typically define $\mathbf{V}$ (the measured utility) in terms of route costs.

Let $U_{ir}(\mathscr{X})$ be the utility of an individual i who selects alternative r. The probability that i selects r is:

$$\begin{aligned} p_{ir} &= \mathbf{p}(U_{ir}(\mathscr{X}) \geq U_{is}(\mathscr{X}), \quad \forall s \neq r) \\ &= \mathbf{p}(V_{ir} + \varepsilon_{ir} \geq V_{is} + \varepsilon_{is}) \\ &= \mathbf{p}(\varepsilon_{is} - \varepsilon_{ir} < V_{ir} - V_{is}) \\ &= \int I(\varepsilon_{is} - \varepsilon_{ir} < V_{ir} - V_{is}) f(\boldsymbol{\varepsilon}_i) d\boldsymbol{\varepsilon}_i \end{aligned} \tag{2.8}$$

where $\boldsymbol{\varepsilon}_i$ is a vector of random factors with the joint density function of $f(\boldsymbol{\varepsilon}_i)$ and $I(.)$ is the indicator function, which is 1 when the event in the parentheses occurs, and 0 otherwise (Train (2009)).

In route choice rather than an individual's choice the focus is in the aggregated choice. Representative utility is usually specified to be linear in parameters. Assume that the utility associated with a route is exclusively defined based on its perceived travel cost $\hat{C}_r$:

$$U_r = -\theta \hat{C}_r \xrightarrow{\hat{C}_r = C_r + \varepsilon_r} U_r = -\theta C_r - \theta \varepsilon_r. \tag{2.9}$$

The error term in equation 2.7 follows a certain probability distribution. Different choice models are derived from this equation based on different specification of the random term. The most common choices are Gumbel or Normal random errors, leading to logit and probit models respectively. These are described in more detail below.

2.5.1 Logit model

Different types of logit family models have been proposed in the transportation literature. The multinomial logit model (MNL) is considered the most widely used route choice model to compute a route choice probability.McFadden et al. (1973) has shown that when the random error terms in 2.7 are independently and identically Gumbel distributed with zero mean and variance $\pi^2/6\theta^2$, then the function giving the route choice probability has an explicit form given by the multinomial logit formula where the probability of choosing route $r \in \mathcal{R}_j$ can be calculated as

$$p_r = \frac{\exp\{-\theta C_r\}}{\sum_{s \in \mathcal{R}_j} \exp\{-\theta C_s\}}. \tag{2.10}$$

The logit parameter θ determines travellers' sensitivity to differences in utility. This parameter can also be thought of as a spread parameter, with small θ corresponding to a large variance and travel demand will be widely spread over the existing routes. As θ increases, the variability among drivers decreases and flows will be concentrated on routes with lower travel cost.

Because the multinomial logit model possesses a closed-form expression, these types of models are analytically convenient for capturing individual travellers' travel decision.

A special property of the MNL is the independence of irrelevant alternatives (IIA) property which implies that the ratio of the probabilities of any two alternatives is independent of other alternatives. Mathematically speaking, regardless of other alternatives

$$\frac{p_r}{p_s} = \frac{\exp\{-\theta C_r\}}{\exp\{-\theta C_s\}} = \exp\{-\theta(C_r - C_s)\}. \tag{2.11}$$

In a typical network, alternative routes between the same OD pair may have common links, thereby complex correlations among alternatives can be captured. If MNL is applied, links that are common to multiple routes can get overloaded, resulting in serious prediction errors. To show this, consider a transportation network with two options to make a trip for a traveller, personal car and a transit bus. The probability of choosing each of them is 1/2. Originally all buses are red, but then half of them are repainted blue and all three modes have the same utility. Then the predicted probabilities from a logit model will give 1/3 for each alternative. This is unrealistic since we would expect 1/2 as a probability of taking car and 1/4 for selecting each type of coloured bus.

Nevertheless, this model is applied widely for its simplicity and many models that address the correlation problem have been developed. For adding a correction term to account for correlations between routes, C-Logit model which extends the MNL model maintaining the simplicity of computation while improving the prediction accuracy for cases where correlations exist between alternative routes, has been proposed (Cascetta et al. (1996)). The Path Size Logit model Ben-Akiva and Bierlaire (1999) can be considered as an improvement of the C-Logit that incorporates behaviour theory in the utility adjusting process.

2.5.2 Multinomial probit model

In order to address the IIA property as a major drawback of the MNL model, Daganzo (1979) suggested another type of route choice model called multinomial probit model (MNP). This model is based on the assumption that the random error term of each utility follows the multivariate normal distribution with zero mean and a finite variance-covariance matrix ($\mathbf{\Sigma}$). With the multinomial probit model, the choice probability is expressed by:

$$p_r = \int_{U_1=-\infty}^{U_r} \cdots \int_{U_r=-\infty}^{\infty} \cdots \int_{U_J=-\infty}^{U_r} MVN(\mathbf{U}-\mathbf{V}) dU_1 dU_2 \cdots dU_J \tag{2.12}$$

where

$$MVN(\mathbf{U}-\mathbf{V}) = (\frac{\pi}{2})^{\frac{J}{2}} \frac{1}{\sqrt{|\mathbf{\Sigma}|}} \exp\big[-\frac{1}{2}(\mathbf{U}-\mathbf{V})^T \mathbf{\Sigma}^{-1} (\mathbf{U}-\mathbf{V})\big], \tag{2.13}$$

and J is the number of options.

The probit model is more realistic, because it allows the error terms to be correlated to each other, and the covariance matrix is not restricted to be diagonal. Therefore, the model can capture complex correlations between alternative routes. However, the probit model becomes very complicated once there are more than a handful of routes to choose from. This is due to the fact that integral in Equation 2.12 does not have a closed form and cannot be calculated exactly since multi-dimensional integrals of the multivariate normal functions are computationally expensive ((Sheffi 1985), (Cascetta 2009)).

Here, through a simple example we illustrate how this approach defines the covariance matrix of the (dis)utilities.

Example 2.5.1. Consider the network depicted in Figure 2.4 and as shown here we assume the link cost functions are static (do not depend on flows) and therefore, they are fixed. Here, $\varepsilon_1, \ldots, \varepsilon_4$ are independently distributed normal variables with zero mean and variance σ^2. By assuming that route 1 consists of link 1 and link 3, links 2 and 3 together

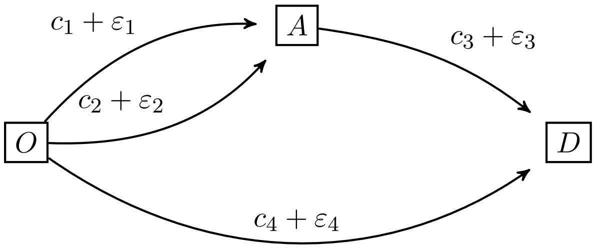


Figure 2.4: Illustrative network, with a single OD pair, four links and three routes.

correspond to the second route and route 3 is the same as link 4, we have

$$\hat{C}_1 = C_1 + \varepsilon_1 + C_3 + \varepsilon_3$$
$$\hat{C}_2 = C_2 + \varepsilon_2 + C_3 + \varepsilon_3$$
$$\hat{C}_3 = C_4 + \varepsilon_4.$$

Consequently,

$$Var(\hat{C}_1) = Var(\varepsilon_1 + \varepsilon_3) = 2\sigma^2$$
$$Var(\hat{C}_2) = Var(\varepsilon_2 + \varepsilon_3) = 2\sigma^2$$
$$Var(\hat{C}_3) = Var(\varepsilon_4) = \sigma^2$$

also $Cov(\hat{C}_1, \hat{C}_3) = Cov(\hat{C}_2, \hat{C}_3) = 0$ and $Cov(\hat{C}_1, \hat{C}_2) = Var(\varepsilon_3) = \sigma^2$. As a result, the variance-covariance matrix ($\mathbf{\Sigma}$), defining the dependence between route costs, is given by

$$\mathbf{\Sigma} = \begin{bmatrix} 2\sigma^2 & \sigma^2 & 0 \\ \sigma^2 & 2\sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{bmatrix}.$$

Observe that the overlap between routes 1 and 2 results in a non-zero covariance between the costs of those routes.

2.5.3 Computation of SUE

Route choice probabilities play a fundamental role in the definition of SUE. Several algorithms have been proposed for finding a solution for the SUE problem. For instance, Cascetta et al. (1997) combined labeling and K-shortest path methods in solving the SUE problem. Dial (2001) proposed an algorithm that applies only for logit route choice models. An algorithm that can be applied for probit-based stochastic user equilibrium was presented by Maher (1992). The reader may refer to Bekhor and Toledo (2005) for a more comprehensive review.

In this thesis, we apply the method of successive averages (MSA) algorithm to solve traditional SUE problems, proposed by Sheffi and Powell (1982). This algorithm has been used in both static and dynamic network equilibrium problems in transportation modeling. Because of its simplicity, this algorithm has been extensively used and it can be applied for both logit and probit cases.

The only drawback of the approach is that it is very slow to converge, because the step length of the search process is not optimised (Huang and Gu 1994). MSA takes the flow on a link as a linear combination of the previous flow on the previous iteration and an auxiliary flow ($\mathbf{z}$) from the present iteration which is obtained by multiplying the route choice probability with travel demand on that route. The method is based on a predetermined series of step sizes such as α_n for overcoming the problem of allocating too much traffic to congested links. With the proper choice of the step size at each iteration, the method converges to the equilibrium solution in static traffic assignment (Sheffi 1985). $\{\alpha_n\}$ is a predetermined step size sequence that guarantees the convergence of the method and satisfies the three conditions, $\lim_{n\to\infty}\alpha_n = 0$, $\sum(\alpha_n) = \infty$ and $\sum(\alpha_n)^2 < \infty$. In following we describe how the conventional MSA works.

The first step is based on the free-flow costs to determine a set of initial link flows and obtain the current flow pattern by employing logit or probit formula. The MSA algorithm then updates the travel costs through the current flow pattern. This method applies $\mathbf{z}-\mathbf{y}$ as the searching direction, with a predetermined step size sequence converging to zero. Then the new current solution can be obtained through $\mathbf{y}^{(n+1)} = \mathbf{y}^{(n)} + \alpha_n(\mathbf{z}^{(n)} - \mathbf{y}^{(n)})$. Finally, if convergence is achieved the algorithm is stopped and the current flows solve the problem. The pseudo-code of MSA algorithm is written as follows:

Algorithm 1 MSA algorithm

Input: network information
Output: SUE solution

Initialisation: Set $y = 0$ in $c_l(y)$
$iter \leftarrow 1$
$gap \leftarrow 1$
$tol \leftarrow 10^{-4}$
while $gap > tol$ **do**
    update $c_l\,(y)$
    Apply logit or probit model to find $\boldsymbol{p}$
    Find an auxiliary flow $\mathbf{z} = N\boldsymbol{p}$
    $gap \leftarrow \sqrt{\frac{1}{iter}(\frac{\mathbf{y}-\mathbf{z}}{N})^2}$
    if $gap > tol$ **then**
        $\mathbf{y} \leftarrow \mathbf{y} + \frac{1}{iter}(\mathbf{z}-\mathbf{y})$
        $iter \leftarrow iter + 1$
    end if
end while
return SUE points

2.6 Stochastic processes

My research will focus on the comparison of deterministic with stochastic day-to-day models. Stochastic day-to-day models use stochastic processes, mathematical abstractions of empirical processes whose development is based on probabilistic laws. In the context of traffic network models, these methods consider the route flows as discrete random variable. We will briefly introduce some basic definitions and properties of stochastic processes to facilitate a better understanding of the concepts for upcoming chapters.

A stochastic process is an infinite collection of random variables typically indexed by time. A continuous-time stochastic process is denoted as $\{X(t), t \in \mathsf{J}\}$, where J is an interval from the set of real numbers such as $[-1, 1]$, $[0, \infty)$, $(-\infty, \infty)$, etc. A discrete-time stochastic process (or a random sequence) is denoted as $\{X(m) = X_m, m \in \mathsf{J}\}$, where J is a countable set such as $\mathbb{N}$ or $\mathbb{Z}$.

Traffic flows can be thought of as a stochastic process. For example, if one observes how vehicles are arriving at a section of road, sometimes several vehicles come together while at other times they pass in a sparse manner.

An important class of random processes are stationary processes. A stochastic process $\{X(t), t \in \mathsf{J}\}$ is stationary if its statistical properties do not change over time. For example, a stationary process at times $X(t)$ and $X(t + \triangle)$ has the same probability distribution:

$$F_{X(t)}(x) = F_{X(t+\triangle)}(x) \quad \text{for all t }, t + \triangle \in \mathsf{J}.$$

Stationary processes can be discrete or continuous depending on whether J is a countable or uncountable set.

2.6.1 Markov chain

We focus here on a special kind of stochastic process called a Markov chain. Consider a discrete-time random process $\{X_m, m = 0, 1, 2, \cdots\}$. The state of the Markov chain at time m is the possible value of X_m. If the X_m's are independent then there is no memory in the system, so each X_m can be considered independently from previous ones, X_{m-1}, X_{m-2} etc.

Nonetheless, the independence assumption is not correct for a large number of real-life processes. For example, imagine X_m denote the stock price of a company at time $m \in \{0, 1, 2, \ldots\}$. Then it is reasonable to assume that X_m is affected by the stock price of the previous day, X_{m-1}, maybe even the stock prices from more distant days leading up to timepoint m. Therefore, we need to develop models where the value of X_m depends on previous values.

In modelling X_m we might want to think about the effects of the states of all previous days. However, that leads to massive complexity, since we need to describe the probability distribution of X_m for every possible history. A Markov chain simplifies this, by assuming that all the information needed to describe the probability model for X_m is provided by the state X_{m-1}. In other words, X_m is conditionally independent of the random variables X_{m-2}, X_{m-3}, ..., X_0 given X_{m-1}.

The evolution of "states" in probabilistic systems can be modeled through Markov chain. Assume we are modelling the people who are in the queue at the bus station. Here, the number of people is a non-negative integer and describe the state of the system. Mathematically speaking, if X_m denotes the number of people in the queue at time m, then $X_m \in S = \{0, 1, 2, \ldots\}$. More precisely, the system is in state i at time m when $x_m = i$. In the queuing example, e.g. $X_3 = 9$ means there are 9 people in the queue at timepoint 3.

The set S is called the state space of the Markov chain. Depending on the particular problem, the states are usually chosen to be 0, 1, 2,..., or 1, 2, 3,....

Consider the stochastic process $\{X_m; m = 0, 1, 2, ...\}$. We say that the process is a Markov chain if it satisfies the Markov property

$$\mathbb{P}(X_{m+1} = j | X_m = i, X_{m-1} = i-1, \ldots, X_0 = i_0) = \mathbb{P}(X_{m+1} = j | X_m = i).$$

$\mathbb{P}(X_{m+1} = j | X_m = i)$ is called the transition probability. The transition probabilities do not depend on time in a time homogenous chain, this means that:

$$p_{ij} = \mathbb{P}(X_{m+1} = j | X_m = i) = \mathbb{P}(X_1 = j | X_0 = i) = \mathbb{P}(X_2 = j | X_1 = i) = \ldots$$

If the process is in state i, then it moves to state j with a probability denoted by p_{ij}. These probabilities can be collected together as a matrix,

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix}.$$

P is called the transition matrix and has the property that

$$\sum_{k=1}^{r} p_{ik} = \sum_{k=1}^{r} \mathbb{P}(X_{m+1} = k | X_m = i) = 1.$$

In a Markov chain model for traffic flows we can potentially define states in terms of the vector of route flows. Consider a simple network depicted in Figure 2.5. Travel demand from O to D is N travellers. The state of the system can be defined purely in terms of flow on route 1. Therefore, the system has $N+1$ different states such as $0, 1, \ldots, N$ where, state 0 means all travellers for example use the car and state N indicates that all travellers have chosen the bus.

n-step transition probabilities

In dynamic traffic models the evolution of the system state can be described through a

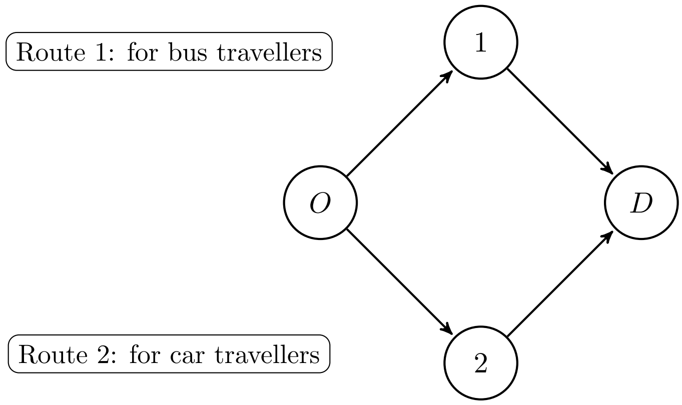


Figure 2.5: A two route network.

transition matrix. Given the transition probability function and a probability distribution for the initial sate, the probabilistic future is then completely characterised. More precisely, given the states of the system on days $0, 1, \ldots, m$ transition probabilities govern the evolution of the system on day $m+1$ given the state of the system on day m. In our work we are not just interested in the transition probability for one step. We need to find the transition probability for multiple steps, therefore it is necessary to explain the n-step transition probabilities.

Consider a Markov chain $\{X_m, m = 0, 1, 2, \ldots\}$, where $X_m \in S$. The probability of going from state i to state j in one step is given as

$$p_{ij} = \mathbb{P}(X_1 = j | X_0 = i).$$

The probability of going from state i to state j in two steps is as follows:

$$\begin{aligned} p_{ij}^{(2)} &= \mathbb{P}(X_2 = j | X_0 = i) = \sum_{k \in S} \mathbb{P}(X_2 = j | X_1 = k, X_0 = i)\mathbb{P}(X_1 = k | X_0 = i) \\ &= \sum_{k \in S} \mathbb{P}(X_2 = j | X_1 = k)\mathbb{P}(X_1 = k | X_0 = i) \qquad \text{(by Markov property)} \\ &= \sum_{k \in S} p_{kj} p_{ik} = \sum_{k \in S} p_{ik} p_{kj}. \end{aligned}$$

This gives us the elements in the ith row and jth column of the matrix P^2. More generally, we can now define the n-step transition probabilities as

$$p_{ij}^{(n)} = \sum_{i, i_1, \ldots, i_{m-1} \in S} p_{ii_1} p_{i_1 i_2} \ldots p_{i_{n-1} j}.$$

Consequently $p_{ij}^{(n)} = \mathbb{P}(X_n = j | X_0 = i)$ and the n-step transition matrix P^n:

$$P^n = \begin{bmatrix} p_{11}^{(n)} & p_{12}^{(n)} & \cdots & p_{1r}^{(n)} \\ p_{21}^{(n)} & p_{22}^{(n)} & \cdots & p_{2r}^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ p_{r1}^{(n)} & p_{r2}^{(n)} & \cdots & p_{rr}^{(n)} \end{bmatrix}.$$

Stationary distribution

In the deterministic approach of traffic assignment models equilibria can be captured through fixed flow patterns. The probabilistic analog of the equilibrium link flow would be the stationary probability distribution of the stochastic process generating the link flows.

A stationary distribution $\boldsymbol{\pi}$ is a row vector giving probabilities for all the possible states. This is a probability distribution such that when the Markov chain reaches the stationary distribution then it remains the same. Mathematically, stationary distributions are obtained by solving $\boldsymbol{\pi} = \boldsymbol{\pi} P$. The term stationary is used because $\boldsymbol{\pi} P^2 = (\boldsymbol{\pi} P)P = \boldsymbol{\pi} P = \boldsymbol{\pi}$ and so $\boldsymbol{\pi} P^n = \boldsymbol{\pi}$ for all $n \geq 0$. A Stationary distribution describes the long-term behaviour of a Markov chain, which can be interpreted as giving, in the long run, the probability of observing any particular pattern of route flows. This distribution could then be used to calculate for instance, the long run average route flows.

When we want to discuss the long-term behaviour of a Markov chain, we would like to know the fraction of time that a Markov chain spends in each state as n becomes large. That is, mathematically we would like to study the distribution:

$$\boldsymbol{\pi}_n = [P(X_n = 0),\ P(X_n = 1),\ P(X_n = 2), \ldots]\text{ as } n \longrightarrow \infty.$$

Let $\boldsymbol{\pi}_0$ denote the probability distribution of Markov chain at time 0 (i.e X_0), the trajectory of the system can be written as $(\boldsymbol{\pi}_0, \boldsymbol{\pi}_0 P, \boldsymbol{\pi}_0 P^2, \ldots, \boldsymbol{\pi}_0 P^n, \ldots)$. If the finite sequence has a limit, then this limit gives such a stationary distribution, i.e.

$$\lim_{n \mapsto \infty} \boldsymbol{\pi}_0 P^n = \boldsymbol{\pi}.$$

The probability distribution $\boldsymbol{\pi} = [\pi_0, \pi_1, \ldots]$ is called the limiting distribution of Markov chain X_n if

$$\pi_j = \lim_{n \mapsto \infty} P(X_n = j | X_0 = i)$$

for all $i, j \in S$ and whenever

$$\sum_{(j \in S)} \pi_j = 1.$$

By the above definition, when a limiting distribution exists, it does not depend on the initial state $(X_0 = i)$ so we can write

$$\pi_j = \lim_{n \mapsto \infty} P(X_n = j) \quad \text{for all} \quad j \in S.$$

A Markov chain will not necessarily converge to a limit, and even if it does, the limiting distribution may not be unique. In order to understand conditions under which convergence to a unique distribution is assured, we must learn about classification of states.

2.6.2 Classification of states

Depending on transition probabilities a Markov chain may visit some states infinitely often and visit other states only a finite number of times. This interaction between the state space and transition probabilities is an important part of a Markov chain. It forms the basis of a classification of states which we now present.

Irreducible states

We say that state j is **accessible** from state i if for some integer n, $p_{ij}^{(n)} > 0$. This is denoted $i \longrightarrow j$, and indicates that starting from state i there is a sequence of transitions that will allow us to eventually reach state j. Two states i and j are said to communicate, written as $i \longleftrightarrow j$, if each state is **accessible** from the other. When all states intercommunicate, then a chain is **irreducible**. In this case, all states communicate with each other. For example, consider a Markov chain with states $\{0, 1, 2, 3\}$ and the transition matrix

$$P = \begin{bmatrix} 1/4 & 0 & 0 & 3/4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

P is irreducible since $0 \to 3 \to 2 \to 1 \to 0$ is a sequence of states that occurs with non-zero probability.

For any chain if there is a self-transition which is $p_{ii} > 0$ for some i, then the chain is aperiodic. We note that this is a sufficient although not necessary condition for aperiodicity. It holds for all plausible traffic models. As an important result, it should be noted that an aperiodic and finite irreducible Markov chain (also known as an ergodic chain) has a unique stationary distribution (Rosenthal 2006).

Recurrent states

A state is said to be **recurrent** if the probability of eventual return is one. That is, for any state i it can be defined as $\mathbb{P}(\exists n \geq 1 : X_n = i | X_0 = i) = 1$.

If the probability of the returning to the state is less than 1, the state is called **transient**. Suppose we have the following transition matrix

$$P = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1/2 & 1/2 \end{bmatrix}.$$

If we start in state 0, we see that the path $0 \to 2 \to 1 \to 0$ must be followed with probability 1. This immediately tells us that the states 0, 1 and 2 are recurrent while states 3 and 4 are transient. State 3 leads to state 2, but state 2 does not lead back to state 3, therefore, state 2 and 3 do not communicate.

The subsequent discussion requires the following definition of first passage time, as it helps us to find the relation between the return probability at time n and the event that a state is recurrent.

Let $f_{ij}^n = \mathbb{P}(X_n = j, X_{n-1} \neq j, \ldots, X_1 \neq j | X_0 = i)$ be the probability of reaching state j for first time in n step starting from state i and the probability that the chain ever visits state j starting from i:

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^n = \mathbb{P}(X \text{ hits } j \text{ after time } 0 | X_0 = i).$$

If $f_{ii} = 1$ then state i is recurrent. State i is transient if $f_{ii} < 1$. The first passage time to state i is the time of the first visit to state i given by $X_0 = i$:

$$T_i = \min\{n \geq 1; X_n = i\}.$$

Absorbing states

Once you enter an absorbing states you never leave them. This concept will be illustrated by an example. For the following transition matrix

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

we determine that state 0 is an absorbing since the probability from going from state 0 to state 0 is one (this means that once the system enters state 0 it does not leave since the probability of moving from state 0 to states 1, 2 and 3 is zero as indicated by the first row) and states 1 and 2 are transient. Similarly, state 3 is an absorbing state.

The difference between recurrent and absorbing states is that, in a recurrent Markov chain, while the probability that you will return to that state at some point after visiting it is 1, you can visit other states afterwards. In contrast, once you visit an absorbing state it is impossible to leave.

A chain with at least one absorbing state and with a possibility to move from every non-absorbing state to some absorbing state, not necessarily in one step, is called an absorbing chain.

2.6.3 Absorbing probabilities

If a chain has only absorbing and transient states, that is, non-absorbing states, then it will visit the transient states for a while before eventually being absorbed.

We want to compute the time it takes until these particular absorbing states are reached given a starting state. We consider finite absorbing chains with a absorbing states and t transient states. We write $\mathcal{A}$ and $\mathcal{T}$ as the subsets of the state space S containing all the absorbing and transient states, respectively.

The transition matrix for such a chain will have the following canonical form (Grinstead and Snell 2012):

$$P = \begin{pmatrix} I & 0 \\ Q & T \end{pmatrix}$$

where I is an $a \times a$ identity matrix, Q is a $t \times a$ matrix specifying the transition probability form transient states to absorbing states and T is a $t \times t$ matrix specifying the transition probability amongst the transient states. Starting from any transient state $i \in \mathcal{T}$ the chain will eventually be absorbed into $j \in \mathcal{A}$, one of the absorbing states with probability f_{ij}. We group these absorption probabilities into a matrix A:

$$A = [f_{ij}]_{i \in \mathcal{T}, j \in \mathcal{A}}.$$

When calculating the absorbing probabilities, we proceed by analysing the possibilities that can arise at the end of the first transition and then using this as the basis of a recursion argument. The Markov property is key for this to work.

$$\begin{aligned} f_{ij} = \mathbb{P}(X_t = j|X_0 = i) &= \sum_{k \in Q} \mathbb{P}(X_t = j, X_1 = k|X_0 = i) \\ &= \sum_{k \in Q} \mathbb{P}(X_t = j|X_1 = k, X_0 = i)\mathbb{P}(X_1 = k|X_0 = i) \\ &= \sum_{k \in Q} \mathbb{P}(X_t = j|X_1 = k, X_0 = i)p_{ik}. \end{aligned}$$

We now consider how to express the first factor for various possible states k. If the starting state is the same as destination state, it is clear that the probability is equal to 1, thus if $k = j \in \mathcal{A}$ then $P(X_t = j|X_1 = j, X_0 = i) = 1$.

In all other cases, if $k \in \mathcal{A}, k \neq j$ then $\mathbb{P}(X_t = j|X_1 = k, X_0 = i) = 0$. This means that the probability of going from one absorbing state to the another absorbing state is zero. Finally, for any transient state it can be shown that, if $k \in \mathcal{T}$, then $\mathbb{P}(X_t = j|X_1 = k, X_0 = i) = \mathbb{P}(X_t = j|X_1 = k) = f_{kj}$.

According to the above statements, we have:

$$\begin{aligned} f_{ij} &= 1.p_{ij} + \sum_{k \in \mathcal{A}, k \neq j} 0.p_{ik} + \sum_{k \in \mathcal{T}} f_{kj} p_{ik} \\ &= p_{ij} + \sum_{k \in \mathcal{T}} p_{ik} f_{kj} \\ f_{ij} &= q_{ij} + \sum_{k \in \mathcal{T}} T_{ik} f_{kj} \\ A &= Q + TA. \end{aligned}$$

The matrix formulation can be represented as below:

$$A = (I - T)^{-1}Q \tag{2.14}$$

where T and Q are the usual components in the canonical form of the transition matrix P. A is an important matrix for a finite Markov chain and is referred to as the fundamental matrix. The column sum of A gives the expected number of transitions prior to absorption for each non-absorbing state.

2.7 Day-to-day Markovian traffic assignment models

As stated before, there are two types of day-to-day dynamic traffic models. In the deterministic version of these models, the model properties remain the same if the starting conditions are unchanged. However, these properties can vary widely even under similar initial conditions for the stochastic models. Taking into consideration the fact that real traffic networks are random by nature has led to increased attention in stochastic assignment models.

Despite the popularity of the deterministic approaches it has been recognised that these models have serious shortcomings. Fixed flow patterns which are meant to represent traffic equilibria cannot account for day-to-day variations. Stochastic day-to-day assignment models taken into account travellers' perception errors and represent day-to-day variability in flows. For the purpose of modeling of the traffic assignment based on this day-to-day variation Cascetta (1989) proposed a stochastic process approach, namely the Markovian traffic assignment.

Markovian assignment represents the evolution of traffic flows over the network under day-to-day variation where the days are indexed by t. Cascetta (1989) first considered a fixed set of feasible routes between all OD pairs, then next assumed OD demand for each OD pair is finite. Cascetta introduced the vector containing the flows of traffic on each route for each OD pair as a state of traffic system and can be denoted by $s(t) = \begin{pmatrix} \boldsymbol{x}_1^{(t)} & \boldsymbol{x}_2^{(t)} & \dots & \boldsymbol{x}_h^{(t)} \end{pmatrix}^T$. Here, $\boldsymbol{x}_i^{(t)}$ denotes the vector of traffic flows on the routes serving OD pair i. He next assumed that traveller's perceived costs on day t is a weighted average of the actual costs occurring the preceding m days, plus a random error term given by

$$\hat{C}_r^{(t)} = \sum_{i=1}^{m} w_i C_r(\boldsymbol{x}^{(t-i)}) + \varepsilon_r, \tag{2.15}$$

where $\hat{C}_r(t)$ can be thought of as a (dis)utility.

In a similar manner to computing stochastic user equilibrium, then this gives route choice probabilities as a function of the weighted average of the recent m days actual costs. Cascetta then assumed that the route choice probability function for all routes in the feasible set is positive and assigned to routes as outcomes of a multinomial random variable.

Example 2.7.1. Consider a network with two non-overlapping routes serving only one OD pair. Suppose that there is a travel demand of N travellers, and they choose their route based on costs remembered over only one day ($t = 1$). Here, links and routes are synonymous in this example (i.e. A is identity matrix). The link cost functions are given as

$$c_l(\boldsymbol{y}) = c_l(y_l) = a_l + \left(\frac{y_l}{b_l}\right)^2 \qquad l = 1, 2. \tag{2.16}$$

Suppose $b_1 = b_2$ in the link cost functions. The average perceived costs by using the last m days' flows can be computed via equation 2.15. For our example we set $m = 1$, so this can be defined as $\hat{C}_r^{(t)} = c_r(x_r^{(t-1)}) + \varepsilon_r$, where ε_r follows a Gumbel distribution. On day t traveller i will then take feasible route r with smallest personal disutility, so the route choice probabilities can be expressed via a logit random utility model:

$$P_r(\boldsymbol{x}^{(t-1)}) = \frac{\exp\{-\theta c_r(x_r^{(t-1)})\}}{\sum_r \exp\{-\theta c_r(x_r^{(t-1)})\}}, \qquad r = 1, 2. \tag{2.17}$$

θ represents the sensitivity that travellers have to differences in route travel costs. Figure 2.6 shows the simulation of traveller route choice via model 2.17. We set $a_1 = 2$ and $a_2 = 1$. The capacity of both routes is directly linked to travel demand with $b = \frac{N}{4}$.

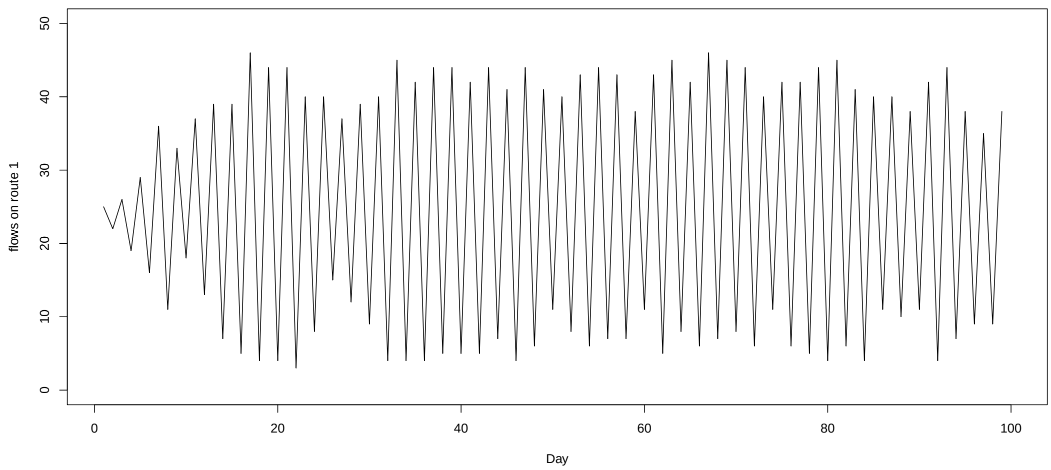


Figure 2.6: Simulation of flows on route 1 for a Markov chain day-to-day model, using a logit route choice model.

Cascetta's (1989) Markov assignment is a finite m-dependent Markov chain. In other words, route choice probabilities depend on a finite number of (m) previous states

$$p_r^{(t)}\left[\boldsymbol{x}^{(t-1)}, \boldsymbol{x}^{(t-2)}, \ldots\right] = p_r^{(t)}\left[\boldsymbol{x}^{(t-1)}, \ldots, \boldsymbol{x}^{(t-m)}\right].$$

Since all states have positive probability of occurring on day t given what happened on days $t-1, \ldots, t-m$ the Markov chain is also ergodic, and so possesses a unique stationary distribution. Finally, since link flows are aggregations of route flows and route costs depends on link flows, the stochastic process generating link flows is also ergodic, and possesses a unique stationary distribution.

The large size of the state space prohibits computation of the stationary distribution or the state to state transition probabilities. In order to obtain estimates of the stationary mean and variance Cascetta (1989) simulated the process and calculated the statistics from the simulated data. This gave him sufficient information to compare the relative accuracy of the stationary distribution's mean as an equilibrium forecast to the probit model SUE.

Cascetta (1989) was able to give a relatively simple proof of the existence and uniqueness of the traffic process's stationary distribution by restricting the model to finite populations. Davis and Nihan (1993) take a different approach, and allow travelling populations to become arbitrary large and have shown the stationary mean in Markovian assignment is well approximated by SUE.

Another Markovian assignment model was introduced by Davis and Nihan (1993). Once a traveller has chosen to make a trip, he or she is faced with a choice of route through the network. They assumed that each acyclic route connecting an origin to a destination has a probability of being chosen by a traveller, and that this probability is a function of the network's past travel costs. Then by assuming that each traveller makes his or her route choice independently of what other travellers are currently doing, the traffic volumes on each route are the outcomes of multinomial random variables. The vector of daily link flows is then given by

$$\boldsymbol{y}^{(t)} = A\boldsymbol{x}^{(t)}. \tag{2.18}$$

Now consider the relation between the cost of travel on the network and travellers' route choice probabilities. They first assumed that the cost of traversing a given path is the sum of the cost of traversing its individual links (2.2). Next it is assumed that travellers do not base their route choice on the most recent link cost, but rather on an anticipated disutility cost denoted by $u(t)$. Through an exponential learning process, which combines the perceived and actual travel cost from the previous day, disutilities are updated each day based upon

$$u_r^{(t)} = \lambda c_r(\boldsymbol{y}^{(t-1)}) + (1-\lambda)u_r^{(t-1)}, \quad r = 1, \ldots, n, \tag{2.19}$$

where λ is a parameter satisfying $0 \leq \lambda \leq 1$. Note that if $\lambda = 1$, perceived costs are the previous day's actual costs, while if $\lambda = 0$, the perceived cost remain constant independent of actual costs. For $0 < \lambda < 1$, the perceived costs will be a weighted average of all the past actual costs, with more recent costs weighted heavily.

Route choice probabilities are computed using random utility theory based on the utilities u_r. The state of the system is defined by $s^t = ((\boldsymbol{u}^{(t)})^T, (\boldsymbol{y}^{(t)})^T)^T$, and s^t is a Markov chain. It's evolution is described by the following steps:

Model 2.7.1. :

Step 1: Given initial perceived costs $\boldsymbol{u}^{(0)}$ and link flows $\boldsymbol{y}^{(0)}$, let $t = 1$,

Step 2: $\boldsymbol{u}^{(t)} = \lambda \boldsymbol{c}(\boldsymbol{y}^{(t-1)}) + (1-\lambda)u_r^{(t-1)}$,

Step 3: $\boldsymbol{x}_j^{(t)}$ are independent multiomial outcomes with parameters $(N_j, \boldsymbol{p}_j(\boldsymbol{u}^{(t)}))$,

Step 4: $\boldsymbol{y}^{(t)} = A\boldsymbol{x}^{(t)}$, Step 5: $t = t + 1$, go to step 2.

Hazelton and Watling (2004) follow similar principles as set out by Cascetta (1989) by assuming that travellers tend to remember their experiences over a finite number of days and develop their perception based on the most recent set of experienced costs. Thus, at the beginning of the day t the travellers update their perceived cost for the routes given by a linear combination of the experienced costs weighted by an appropriate weighted system. As shown in Hazelton and Watling (2004) this can be expressed by

$$\boldsymbol{u}^{(t-1)} = s(\lambda)^{-1} \sum_{j=1}^{m} \lambda^{j-1} \boldsymbol{C}(\boldsymbol{x}^{(t-j)}) \tag{2.20}$$

where $0 < \lambda < 1$, $s(\lambda) = \sum_{j=1}^{m} \lambda^{j-1} = (1 - \lambda^m)/(1 - \lambda)$ is a scaling factor and $\sum_{j=1}^{m} \lambda^{j-1} s(\lambda)^{-1} = 1$. Conditional on the costs that have been experienced in the past, the number of travellers selecting each possible route on day t will follow a multinomial distribution with parameters N_j and $\boldsymbol{p}_j(\boldsymbol{u}^{(t-1)})$.

Recently quite a large number of papers have extensively studied Markovian assignment models (e.g. Cascetta (1989); Davis and Nihan (1993); Cantarella and Cascetta (1995); Hazelton (2002); Watling and Hazelton (2003); Hazelton and Watling (2004); Watling and Cantarella (2013); Parry et al. (2016)).

2.8 Convergence

The concept of convergence is important because we aim to understand the convergence of a stochastic process to a deterministic process when we have a large number of travellers as well as increasing the values of the sensitivity parameter, θ in Equation 2.10 as shown in Figure (2.7).

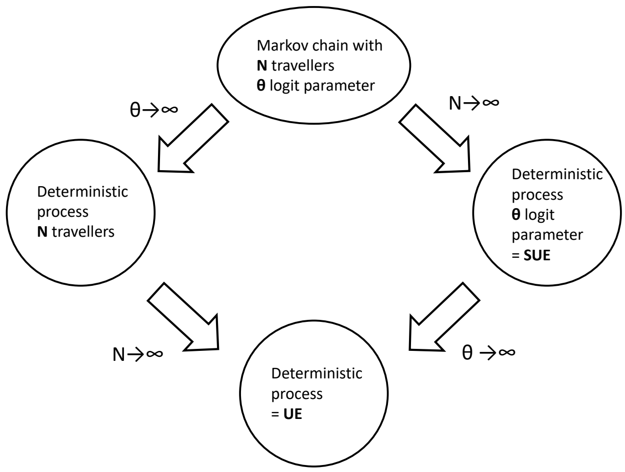


Figure 2.7: Big picture of our goal through the concept of convergence.

In the following, we present some main definitions of convergence (Rosenthal 2006). We consider three types of convergence, as follows:

1. **Convergence in distribution:**

   A sequence of random variables $X_1, X_2, \ldots$ converges in distribution to a random variable X indicated by $X_n \xrightarrow{d} X$ As n tends to infinity, the probability distribution function of X_n tends to that of X. This can be mathematically written as follows

   $$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

   for all points x where $F_X(x)$ is continuous.

2. **Convergence in probability:**

   A sequence of random variables $X_1, X_2, \ldots$ converges in probability to a random variable X denoted by $X_n \xrightarrow{P} X$ if:

   $$\lim_{n\to\infty} P(|X_n - X| \geqslant \varepsilon) = 0$$

   for all $\varepsilon > 0$. The essential thought behind this type of convergence is that with the progress of the sequence, the probability of tangible discrepancy between X and X_n decreases.

3. **Convergence in mean:**

   Let $r \geqslant 1$ be a fixed number. A sequence of random variables $X_1, X_2, \ldots$ converges in the $r-th$ mean or in the L^r norm to a random variable X denoted by $X_n \xrightarrow{L^r} X$ if:

   $$\lim_{n\to\infty} E(|X_n - X|^r) = 0$$

It can be shown that convergence in mean implies convergence in probability, which in turn implies convergence in distribution.

Chapter 3

Comparing the Long Term Behaviour of Deterministic and Stochastic Day-to-Day Traffic Assignment Models

3.1 Chapter overview

As seen in Chapter 2, day-to-day dynamic traffic assignment models can be deterministic or stochastic in nature. Both modelling approaches have their advantages and their disadvantages, and the choice of modelling framework is often determined by the intended use of the model. However, we should certainly hope that results from the two approaches are reasonably comparable. In this chapter we explore this comparability with a particular emphasis on the long-run properties of the model.

Deterministic models are heavily studied, and we understand their properties quite well. For example, Horowitz (1984) suggested a discrete-time day-to-day deterministic model to investigate the stability of stochastic equilibrium for a simple two-link network. A continuous deterministic day-to-day dynamic model to show that travellers switch their route from higher travel cost routes to lower travel cost routes was studied by Smith (1984). By extending the Horwitz's network to a general network Watling (1999) proposed a dynamical adjustment process for analysing the stability of the general asymmetric stochastic equilibrium problem in discrete-time. Watling and Hazelton (2003) reviewed both stochastic and deterministic day-to-day models. In particular, the relationship between day-to-day models with deterministic user equilibrium (DUE) and stochastic user equilibrium (SUE) (as two recognized concepts of economic equilibrium in transport networks) is considered. Other works related to deterministic day-to-day model include Bie and Lo (2010); He et al. (2010); Cantarella et al. (2015); Cantarella and Watling (2016).

For deterministic day-to-day models, convergence to fixed equilibrium points is an essential question. These equilibrium points can be stable or unstable. An equilibrium is stable if we change a flow pattern a bit the natural tendency will be return to that equilibrium. Hence, when the system reaches a stable equilibrium then it should stay there forever. But for an unstable equilibrium, any minor change in the system causes a move from the point. Bie and Lo (2010) show that for a system with multiple equilibria the state space of a system may be partitioned into multiple basins of attraction. Once the system is in such a basin, it will converge towards the equilibrium point within.

Stochastic day-to-day models, however, allow us to incorporate the kind of haphazard variation that is seen in real life. These models have also received considerable attention. To describe the day-to-day fluctuation in transportation network Cascetta (1989) developed a discrete-time stochastic process to analysis of the day-to-day dynamic and the relationship between stochastic process solution and SUE solution has been investigated. An extended stochastic process model was employed by Cascetta and Cantarella (1991) to investigate dynamic flow fluctuations for both day-to-day and with-day dynamic models.Davis and Nihan (1993) have shown that when travel demand becomes large in tandem with network capacity, the stochastic process model can be approximated by the sum of a nonlinear deterministic process and a Guassian multivariate auto-regressive process. Hazelton and Watling (2004) developed an approximation model for calculating the stationary mean and covariance matrix. Some of the papers investigating stochastic day-to-day dynamic models include Akamatsu et al. (1996); Watling and Cantarella (2013); Parry et al. (2016); Watling and Hazelton (2018).

Usually we like to think of deterministic models as describing the behaviour of the system on average, and so approximating the mean behaviour of stochastic models. Cantarella and Cascetta (1995) have shown this is typically the case for "well behaved" systems with a unique equilibrium. But comparison of model types for multi-equilibria systems is not well understood. Smith et al. (2014) began to examine this issue and this chapter builds on their work.

Our focus is on the long-run behaviour of the traffic models under study, and how this compares between deterministic and stochastic models when there are multiple deterministic equilibria. This kind of study has important implications for network control, where we want to impose measures so as to push a system towards some desirable equilibrium state. We might hope that both deterministic and seemingly analogous stochastic models would lead to the same conclusions regarding network control. However, we will show that this is not always the case.

The structure of this chapter is as follows. Section 3.2 describe the deterministic and stochastic day-to-day traffic assignment models which are used in the following. In Section 3.3 we describe the concepts of equilibria and basins of attraction. We address questions about convergence towards equilibria in Section 3.4, and in particular look at the mean time required for stochastic models to achieve some desirable state. It has been a tradition in transportation science to use a simple two route network for preliminary studies (Horowitz 1984), and in that vein sections 3.3 and 3.4 make heavy use of the two route network represented by Smith et al. (2014). We extend the analysis to an example with multiple origin-destination pairs in Section 3.5. A summary of findings from this chapter is provided in Section 3.6.

3.2 Day-to-day dynamic traffic assignment models

To show the variation of the system from day to day and following the notations of chapter 2, let $\boldsymbol{x}^t = (x_1^t, \ldots, x_g^t)$ be the vector of route flows on day t and the sub-vector of $\boldsymbol{x}_j^t$ denote the route flows between OD pair j on day t. The total travel demand $N_j = \sum_{r \in \mathcal{R}_j} x_r^t$ for any given OD pair j is assumed to be constant through time.

In a day-to-day model travel decisions on any given day are (typically) modelled in terms of route flows on a finite number of previous days, and potentially some additional measure of (dis)utility that can in some cases be thought of as drivers predictions of future travel costs (e.g. (Davis and Nihan 1993) and (He and Liu 2012)).

We now look in more detail at the types of deterministic and stochastic assignment models studied by Smith et al. (2014). Our deterministic process model assumes for any given OD pair, the rate at which travellers swap from route r to an alternative route s is determined as a proportion of the product of two factors. The first one is the flow on the higher cost route on day t and the second component considered to be the cost difference $C_r(\boldsymbol{x}^t) - C_s(\boldsymbol{x}^t)$ on any given day $t+1$. Note that if the new route s is cheaper then this cost difference will be positive. We define $\Delta_{r,s}$ as the swap vector from route r to route s, given by

$$\Delta_{rs} = \begin{cases} -1 & \text{in the } r\text{th place} \\ +1 & \text{in the } s\text{th place} \\ 0 & \text{otherwise} \quad . \end{cases}$$

Define k to be a parameter that moderates the rate of route swapping. Then using the notation introduced above, the overall change in the pattern of route flows between days t and $t+1$ is given by

$$\boldsymbol{d}(\boldsymbol{x}^t) = \sum_{j=1}^{h} \sum_{r \in \mathcal{R}_j} \sum_{s \in \mathcal{R}_j} k[C_r(\boldsymbol{x}^t) - C_s(\boldsymbol{x}^t)]_+ \, x_r^t \Delta_{r,s}, \tag{3.1}$$

where $u_+ = \max(0, u)$. The route flows on day $t+1$ are then given by

$$\boldsymbol{x}^{t+1} = \boldsymbol{x}^t + \boldsymbol{d}(\boldsymbol{x}^t) \quad . \tag{3.2}$$

We note that by choosing k to be sufficiently small, we can ensure that elements of $\boldsymbol{x}^{t+1}$ are all non-negative.

Now we look at the stochastic alternative of the deterministic process model described above. When taking a stochastic modelling approach the sequence of route flows $\{\boldsymbol{x}^t : t = 1, 2, \ldots\}$ is modelled as a Markov chain (Cascetta 1989). Using a Markov model, each traveller on the system will independently review their choice to travel each day based upon the costs experienced the day before. Specifically, route choices on day $t+1$ are determined by a vector of route choice probabilities $\boldsymbol{p}^{t+1}$ that in turn depend on costs from day t. Assuming that travellers select their route by following the logit model, p_r^{t+1}, the probability of choosing route r on day $t+1$ between OD pair j is then given by

$$p_r^{t+1} = \frac{1}{\sum_{s \in \mathcal{R}_j} \exp\{\beta[C_r(\boldsymbol{x}^t) - C_s(\boldsymbol{x}^t)]\}} \quad , \tag{3.3}$$

where $\beta > 0$ is a parameter representing travellers' sensitivities to cost differences.

It is assumed that for day $t+1$, travellers make their travel decisions independently, conditional on travels costs from the previous day. Then the route flows on day $t+1$ for each OD pair follow multinomial distributions, conditional on the costs from day t. Specifically, for OD pair j we have

$$\boldsymbol{x}_j^{t+1} \mid \boldsymbol{x}^t \sim \mathrm{Mn}(N_j, \boldsymbol{p}_j^{t+1}) \tag{3.4}$$

where $\mathrm{Mn}(N, \boldsymbol{p})$ denotes a multinomial distribution with N trials and probability vector $\boldsymbol{p}$, and $\boldsymbol{p}_j^{t+1}$ are the entries from $\boldsymbol{p}^{t+1}$ corresponding only to routes serving OD pair j.

For both the deterministic and stochastic model we assume that the temporal dependence is described through a history of only one day. That is, the properties of the system on day $t+1$ are fully regulated by the costs generated from flows on day t. We work with this simple model structure for convenience. Of course, for real world modelling we would want to explore more sophisticated representations of the network dynamics.

3.3 Equilibria and basins of attraction

The traditional goal of many traffic assignment models is to achieve to an equilibrium state which, if attained, would persist indefinitely under certain rational rules of behaviour. However, the concept of equilibrium is not the same for deterministic and stochastic models. In the deterministic case equilibria are fixed points in flow space. For instance, for the deterministic process described in the previous sections, $\boldsymbol{x}^*$ will be an equilibrium if and only if

$$[C_r(\boldsymbol{x}^*) - C_s(\boldsymbol{x}^*)]_+ \, x_r^* = 0 \tag{3.5}$$

for all pairs of routes r and s serving the same OD pair. This happens since,

- if $C_r < C_s$ then $C_r - C_s < 0$ and $x_r \geq 0$ so $(C_r(\boldsymbol{x}) - C_s(\boldsymbol{x}))_+ x_r = 0$

and

- if $C_r > C_s$ then $x_r = 0$ so $(C_r(\boldsymbol{x}) - C_s(\boldsymbol{x}))_+ x_r = 0$.

In fact the fixed point of the system follows the idea of Wardrop's user equilibrium, where more costly routes are not used. In such circumstances the change vector is $\boldsymbol{d}(\boldsymbol{x}^*) = \boldsymbol{0}$ and so the flow pattern does not change from day to day.

An equilibrium state of a stochastic process is defined in terms of a stationary probability distribution, where the system state is not invariant over time but the probability distribution over the possible state is constant over time. This can be define through a stationary distribution which is introduced in Chapter 2. For the Markov day-to-day model with transition matrix P when the probability distribution for the process is updated by $\boldsymbol{\pi}^{t+1} = \boldsymbol{\pi}^t P$, so that $\boldsymbol{\pi}^*$ is an equilibrium distribution if and only if $\boldsymbol{\pi}^* = \boldsymbol{\pi}^* P$.

A point that is worth emphasising here is that the flow pattern does not remain constant in stochastic equilibrium: the flow vectors $\boldsymbol{x}^t$ and $\boldsymbol{x}^{t+1}$ will almost certainly differ. It is the pattern of variation, as characterised by the underlying probability distribution, that is invariant. We illustrate this point in the following example, in which we also describe the system dynamics as the process converges to an equilibrium.

Consider a simple network with a single OD pair connected by two routes, interpreted as corresponding to travel by bus and car respectively. We consider the same route cost functions as used in Section 2.3.2. Specifically, the cost functions are defined by

$$C_1(\boldsymbol{x}) = 8 - \frac{8x_1}{N}$$

and

$$C_2(\boldsymbol{x}) = 2 + \frac{4x_2}{N}.$$

Because the cost functions are not both monotonic increasing, the system has the potential to have multiple equilibria. In fact it has three, as we now demonstrate.

Based on Equation 3.1 we have

$$\boldsymbol{d}(\boldsymbol{x}^t) = k\{[C_1 - C_2]_+ x_1 \Delta_{12} + [C_2 - C_1]_+ x_2 \Delta_{21}\}. \qquad (3.6)$$

Using the fact that $x_2 = N - x_1$, we have $C_2(\boldsymbol{x}) - C_1(\boldsymbol{x}) = \frac{4x_1}{N} - 2$. Following this, it is easy to see $C_2(\boldsymbol{x}) - C_1(\boldsymbol{x}) = 0$ when $x_1 = \frac{N}{2}$. Also, when $x_1 > \frac{N}{2}$ then $\frac{4x_1}{N} > 2$ therefore, $\frac{4x_1}{N} - 2 > 0$. Consequently, $[C_1 - C_2]_+ = 0$. Similarly, when $x_1 < \frac{N}{2}$ we have $C_2(\boldsymbol{x}) - C_1(\boldsymbol{x}) < 0$ and hence $[C_2 - C_1]_+ = 0$.

With this mathematical calculations presented above, the equilibrium will happen when:

1. $x_1 = x_2 = \frac{N}{2} \quad \Rightarrow \quad C_1 = C_2 \quad \Rightarrow \quad \boldsymbol{d} = 0$
2. $x_1 > \frac{N}{2} \quad \Rightarrow \boldsymbol{d} = k\{0 + [C_2 - C_1]_+ x_2 \Delta_{21}\} = k[C_2 - C_1]_+ x_2 \Delta_{21} \quad \text{if} \quad x_2 = 0 \Rightarrow \boldsymbol{d} = 0$
3. $x_1 < \frac{N}{2} \quad \Rightarrow \boldsymbol{d} = k\{[C_1 - C_2]_+ x_1 \Delta_{12} + 0\} = k[C_1 - C_2]_+ x_1 \Delta_{12} \quad \text{if} \quad x_1 = 0 \Rightarrow \boldsymbol{d} = 0$.

As a result we have two stable and one unstable equilibrium. The first stable equilibrium happens at $\boldsymbol{x} = (0, N)^\mathsf{T}$ (universal car usage) with basin of attraction defined by $x_1 \in [0, N/2)$, a second stable equilibrium at $\boldsymbol{x} = (N, 0)^\mathsf{T}$ (universal bus usage) with basin of attraction defined by $x_1 \in [0, N/2)$, and an unstable equilibrium at $\boldsymbol{x} = (N/2, N/2)^\mathsf{T}$. The basins of attraction follow immediately from the signs of the cost differences.

An equilibrium point is accessible if the starting point is within its basin of attraction. The rate of convergence can be determined by the parameter k. Simulating the day-to-day model from (3.1) with demand $N = 50$, Figure 3.1 shows the convergence of the system to three distinct equilibrium points starting from different initial states. As can be seen in the first panel when the initial state is at unstable equilibrium state i.e. $\boldsymbol{x}^0 = (25, 25)^\mathsf{T}$, the system remains in that unstable equilibrium forever. As shown in the second panel starting from an initial state in the basin of attraction of universal car usage, for example, $\boldsymbol{x}^0 = (20, 30)^\mathsf{T}$ and by setting $k = 0.06$, then convergence to universal car usage i.e. $\boldsymbol{x} = (0, 50)^\mathsf{T}$ happens quickly. In the third panel the convergence to universal bus usage i.e. $\boldsymbol{x} = (50, 0)^\mathsf{T}$ occurs after a much large number of days when the initial state is $\boldsymbol{x}^0 = (30, 20)^\mathsf{T}$ and $k = 0.006$.

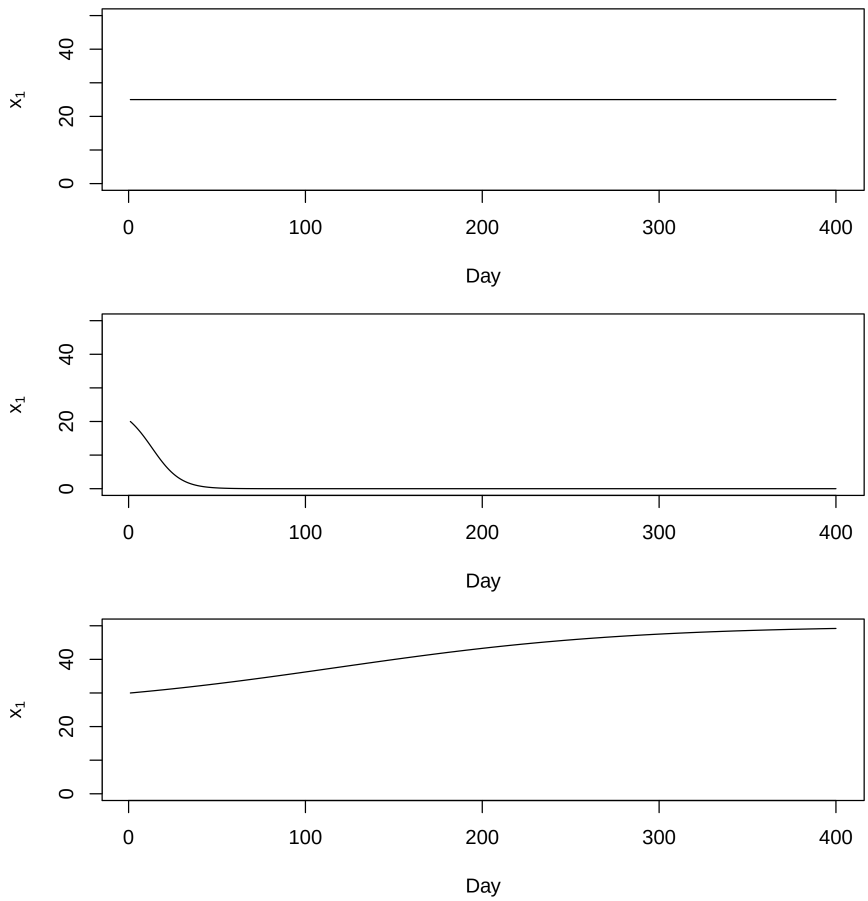


Figure 3.1: Time plots of flows on route 1, x_1, for the two route example. In the top panel the process is initialized at $x_1^0 = 25$, and remains at that unstable equilibrium thereafter. In the middle panel the process is initialized at $x_1^0 = 20$, and converges quickly to the stable equilibrium at $x_1 = 0$ when using a switching parameter of $k = 0.06$. In the bottom panel the process is initialized at $x_1^0 = 30$, when it converges quite slowly to the stable equilibrium at $x_1 = 50$ when using a small switching parameter of $k = 0.006$.

Now, for stochastic day-to-day models with logit route choice model from (3.3), all routes have non-zero probability of selection on day t regardless of the flows on day $t-1$. As a consequence all states are aperiodic, and so the Markov process is ergodic and possesses a unique stationary distribution $\boldsymbol{\pi}^*$. It is useful to note that the system cannot be in equilibrium if we started with any initial flow pattern $\boldsymbol{x}^0$. This happens because the initial probability distribution $\boldsymbol{\pi}^0$ will be a vector with a one in the position corresponding to the initial state, and zero for all other elements while $\boldsymbol{\pi}^*$ is a vector with non-zero elements. Nonetheless, when t becomes large enough regardless of the initial state, the state probabilities will be close to $\boldsymbol{\pi}^*$.

Considering our bus-car example above with multiple deterministic equilibria, we can say that travellers behaviour with the stochastic model is somewhat similar to the deterministic model. To illustrate this, Figure 3.2 shows the simulation results for flows on route 1 over 1000 days assuming that the total travel demand is $N = 50$. Starting with the same initial state $\boldsymbol{x}^0 = (20, 30)^\mathsf{T}$, it has two different scenarios. The first panel shows the flows quickly move to the deterministic equilibrium near $\boldsymbol{x} = (0, 50)^\mathsf{T}$ while the second panel shows the flows at $\boldsymbol{x} = (50, 0)^\mathsf{T}$. Due to the stochastic nature of the model we have fluctuations in the neighbourhood of the deterministic equilibrium states.

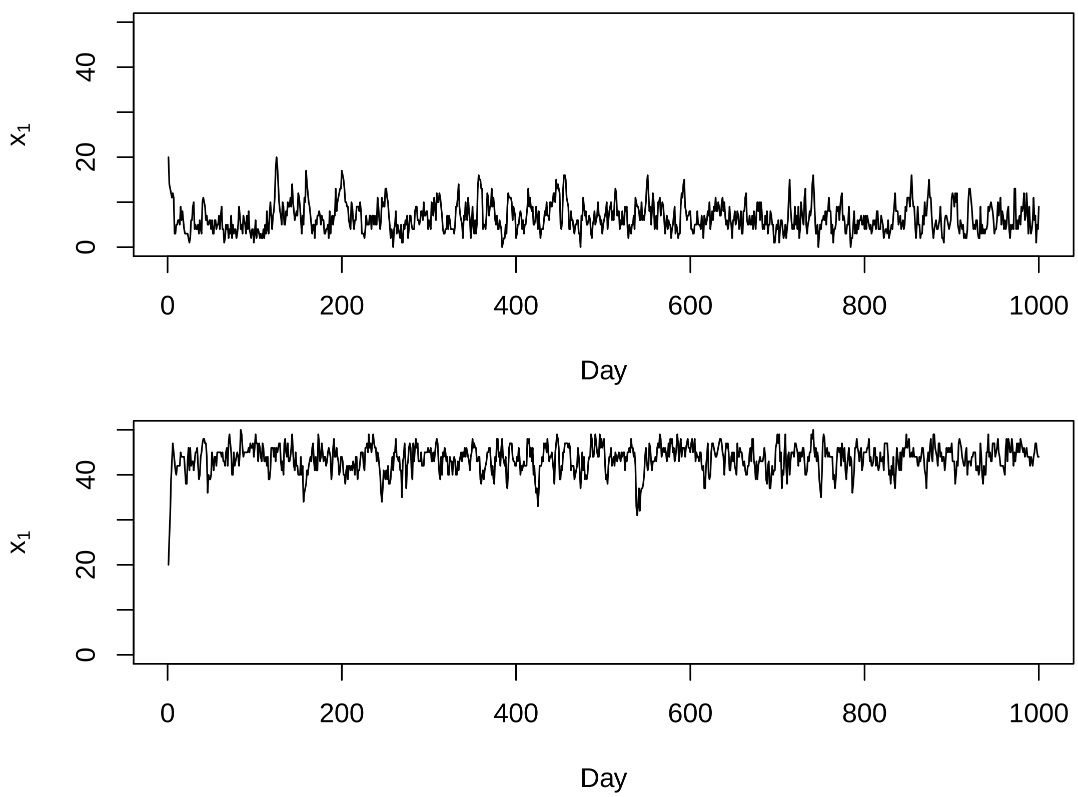


Figure 3.2: Time plots of route 1 flows, x_1, for simulations of the stochastic day-to-day traffic model. For the simulations in both panels the initial state is $\boldsymbol{x}_1^0 = 20$ and the logit parameter is set to $\beta = 1.32$.

As can be seen in Figure 3.3 since the equilibrium distribution in this case is bimodal, there is no sense in which the system can 'converge' to either full bus usage or full car usage. This bimodal distribution is not the same as initial distribution with probability one for the state $x_1 = 20$ and zero for the other states. It can be shown that the relative discrepancy between the flow probability vector at day $t = 100$ and the stationary distribution is

$||\boldsymbol{\pi}^t - \boldsymbol{\pi}^*||/||\boldsymbol{\pi}^*|| = 0.165$. One in fact needs to wait until day $t = 8467$ for this relative difference to shrink below 1%.

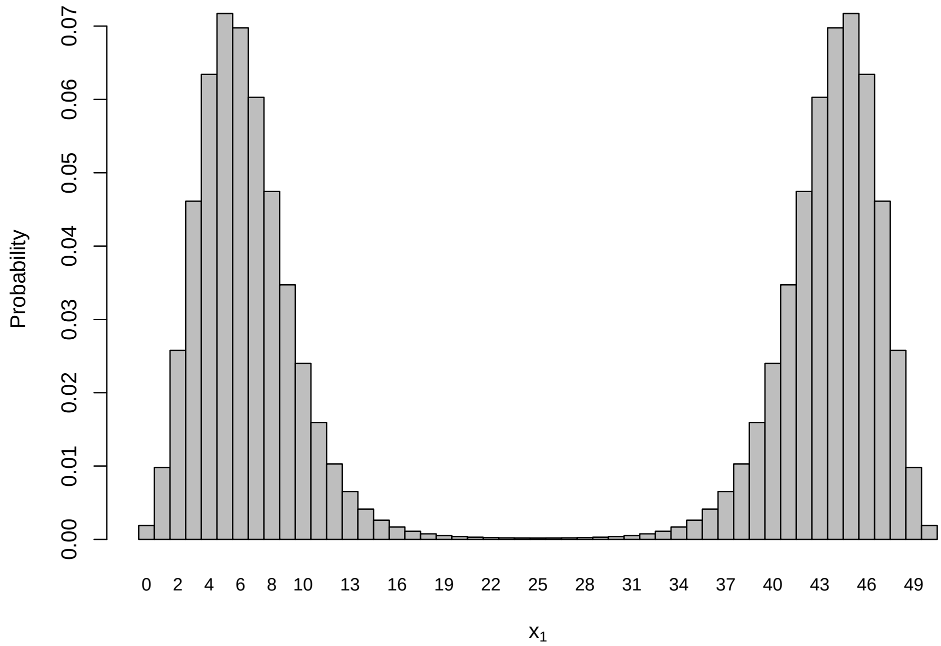


Figure 3.3: Stationary distribution for flows on route 1 for the stochastic day-to-day model with logit parameter $\beta = 1.32$.

As another way to looking at this, it can be said that the initial state is forgotten as the stochastic process converges to its stationary distribution. Now, let us assume that the starting state of the process is $\boldsymbol{x}^0 = (0, 50)^\mathsf{T}$, then the system will persist to be close to that deterministic equilibrium for a period of time. Nonetheless, there is a chance (even though it is tiny) that many travellers switch to the more expensive route. The system may then flip to the neighbourhood of the other deterministic equilibrium. When this flip-flopping between the extremes has occurred a number of times, any memory of the initial state will be negligible and the process will have converged. This is illustrated in the long simulation run depicted in Figure 3.4.

The relationship between the deterministic and stochastic models is complicated. In some ways they appear very similar. For example, the bimodal stationary distribution for the stochastic model (which places most of the probability weight near the extreme flow patterns) reflects the existence of the two stable equilibria in the deterministic model. Furthermore, if you look at the behaviour of the stochastic model over the first 10000 days in the long-run simulation, it looks very much like a noisy form of the deterministic model, converging to the stable equilibrium at $\boldsymbol{x} = (0, 50)^\mathsf{T}$. However, the sudden jumps between equilibria in the stochastic model is unlike any deterministic model behaviour. Furthermore, the deterministic model does not provide a good approximation to the stationary mean of the stochastic model, which lies at $\boldsymbol{\mu} = (25, 25)^\mathsf{T}$. This last point reflects the fact that the mean can be a poor summary of a bimodal distribution. This observations motivates us to explore further instances where the mean behaviour of a stochastic model differs a lot

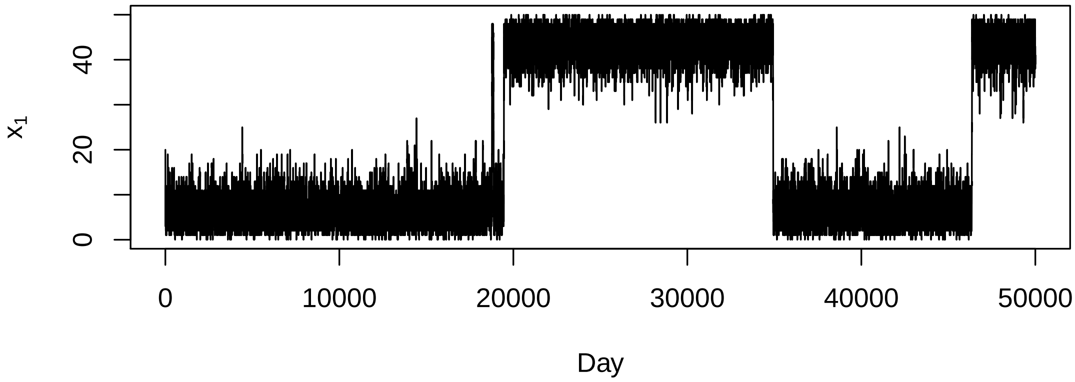


Figure 3.4: Time plot of route 1 flows, x_1, for a long simulation of the stochastic day-to-day traffic model, with the logit parameter set to $\beta = 1.32$. The pattern of route flows is characterized by long periods close to one of the stable deterministic equilibria, interspersed with occasional flips between these.

from what might appear a similar deterministic model.

3.4 Mean hitting time and its implications for network control

A deterministic traffic model can be obtained by looking at some limiting version of a stochastic traffic assignment model. This limiting process might involve the travel demand becoming large (in tandem with network capacity), or travellers becoming increasingly sensitive to cost differences so that there is less and less randomness in route choice. It is of interest to see what properties of the stochastic model are exhibited by the deterministic one under various limiting processes.

Network control is concerned with guiding a traffic system to move to some desirable state. To achieve the desired equilibrium in a deterministic traffic model when the initial state is not in the basin of attraction of that equilibrium point, we can make a temporary change to the network to ensure that the flow patterns enters the basin of attraction of the desired equilibrium, and then let the system dynamics run their course. For a specific illustration, think about our two-route example. Imagine that the system is in its unstable equilibrium, where the number of cars users and bus users are perfectly matched. As a network controller, we may be interested in trying to get all travellers to use the bus. For the deterministic process model described earlier, this can be achieved by inducing a majority of travellers to use the bus by introducing cheaper bus fares, or imposing tolls on private cars for a time, when convergence to total bus usage will follow automatically over the next few days. Even if the initial incentives to change to bus are removed, the system will nevertheless converge to universal bus usage ($\boldsymbol{x} = (N, 0)^{\mathsf{T}}$) eventually.

The analysis of network control is more complex for stochastic models because the behaviour of the system is less certain. Can we be sure that the system will every reach total bus usage? And even if it does, will this be achieved in a relatively small number of days, or will it take a long time?

To begin to answer the first question, as shown in Smith et al. (2014), for models such as the truncated linear probability model

$$p_r^{t+1} = \varphi(C_2^t - C_1^t) = \begin{cases} 1 & \frac{1}{2} + \frac{\beta}{4}(C_2^t - C_1^t) > 1 \\ 0 & \frac{1}{2} + \frac{\beta}{4}(C_2^t - C_1^t) < 0 \\ \frac{1}{2} + \frac{\beta}{4}(C_2^t - C_1^t) & \text{otherwise,} \end{cases}$$

the regularity of Markov chains are not guaranteed because we can get zero values for some conditional probabilities p_{ij}. In such a case, the stationary distribution may not be unique as the state space can be divided into transient and recurrent states. However, for the logit model, all transition probabilities can get close but never quite reach zero (i.e. $p_{ij} > 0$ for all i, j) and the Markov chain is regular. As a result all states are non-null recurrent, and consequently, if given sufficient time, the system will eventually reach any pre-specified flow pattern. For example, in our two-route network if all travellers use the car at the initial point (that is, $\boldsymbol{x}^0 = (0, N)^\mathsf{T}$) then through one of the 'flips' that we saw in Figure 3.4 the system will reach the state where all travellers are use the bus. It is to be expected that the system flips back to car usage at some point in the future, but when the goal is to encourage people to use the bus then it is possible to keep all travellers as bus users by banning car usage. This can be done for instance by closing the undesirable route. The second issue concerns the amount of time the system has to spend to get to the desired state. A network intervention that requires 10 years for the system to finally reach its target equilibrium might be of limited use, for example. Even if we are to consider less extreme cases, the time for any given initiative to take effect should be an element in a comparison of alternative network control schemes. In our deterministic model based on a swapping rate parameter k and the size of the route cost differences there is a fixed time for a system to reach to the desirable state.

In a typical Markov chain, we frequently observe that states are left and re-entered again and again. It is known that $p_{ij}^{(n)}$ describes the probability that a system will leave state i and be in state j after n transitions. But this does not give us any information about whether the system entered state j at any time before the nth transition. When we are specifically interested in the probability that a system leaves state i and enters state j for first time after n steps, we are looking at first passage probability.

For a stochastic model, the time to hit a target flow pattern (or get sufficiently close to it) for the first time is a random variable. In general, the mean of this random variable will be important for network control. Following standard terminology for stochastic processes, we will refer to this quantity as a *mean hitting time*. We might give serious consideration to some intervention for which the mean hitting time is $m = 20$ days, but would discount an alternative for which the corresponding mean time is $m = 10,000$ days. We might hope to relate mean hitting time to the kinds of trace plots displayed in Figure 3.2. In particular, if we can impose some control measure that temporarily moves the system close to universal bus usage in our example, say $\boldsymbol{x}^0 = (45, 5)^\mathsf{T}$, then we can expect the system to visit the target state $\boldsymbol{x}^0 = (50, 0)^\mathsf{T}$ reasonably quickly. Conversely, the time required for the system to reach universal car usage ($\boldsymbol{x} = (0, 50)^\mathsf{T}$) from that starting point will be considerably greater.

To analyse this issue further, let us define the random variable T_{ij} to be the time (in days) taken for the system to reach state χ_j for the first time, having started in state χ_i. That is,

$$T_{ij} = \min\{t \in \mathbb{N} \colon \boldsymbol{x}^t = \chi_j \mid \boldsymbol{x}^0 = \chi_i\}. \tag{3.7}$$

For general pairs of states χ_i, χ_j the quantity T_{ij} is typically referred to as the first passage time. When χ_j has some special significance (like for us, where χ_j is the target state for network control) the term *first hitting time* is often preferred. We will look at the mean of this random variable, and define $m_{ij} = \mathsf{E}[T_{ij}]$. m_{ij} is the expected number of time-steps for reaching state χ_j for the first time given that the chain was initially in state χ_i. In cases where the choice of target state is clear, we may drop the second subscript and simply refer to m_i as the mean hitting time given that the initial state is $\boldsymbol{x}^0 = \chi_i$.

Now we want to rewrite m_{ij} in terms of matrices. Mean hitting times can be computed using the result

$$m_{ij} = \frac{z_{jj} - z_{ij}}{\pi_j^*} \quad \text{for } i \neq j. \tag{3.8}$$

In this formula z_{ij} is the i,jth element of the fundamental matrix $Z = (I - P + \boldsymbol{e}\boldsymbol{\pi}^*)^{-1}$, where I is the appropriately sized identity matrix, P is the Markov transition matrix and $\boldsymbol{\pi}^*$ the row vector stationary distribution introduced earlier, and $\boldsymbol{e}$ is a column vector of ones. See for example Hunter (2007).

Table 3.1 displays the values for mean hitting time to reach to universal bus usage in our bus-car example by assuming the OD demand is $N = 10$. The results are provided for the different values of the logit parameter β and starting from different initial states. When i people travel by bus in the initial flow pattern then the initial state is $\boldsymbol{x}^0 = \chi_i = (i, N - i)$. Suppose that $\beta = 0$ then the people do not care about cost differences at all, therefore, the probability of choosing a route for each traveller is the same as the probability of tossing a coin. Since we have 10 independent travellers and for each, there is 1/2 chance to take the bus thus the probability of universal bus usage would be $1/2^{10}$ and consequently the mean hitting time would be $m_i = 1024$ irrespective of initial state.

	Initial route 1 (bus) usage, x_1^0					
	0	2	4	6	8	9
$\beta = 0.1$	981	981	981	980	980	979
$\beta = 0.5$	377	376	375	373	367	362
$\beta = 1.0$	65.3	63.8	59.9	52.7	42.7	36.6
$\beta = 2.0$	1.12×10^4	1.12×10^4	9.63×10^3	1.59×10^3	30.7	6.28
$\beta = 3.0$	1.77×10^8	1.77×10^8	1.69×10^8	7.17×10^6	1.16×10^3	19.9
$\beta = 4.0$	4.01×10^{12}	4.01×10^{12}	3.97×10^{12}	3.93×10^{10}	5.63×10^4	108

Table 3.1: Mean hitting time for the state of universal bus usage when i travellers use the car initially. The system has $N = 10$ travellers and uses logit parameter β.

When $\beta = 0.1$, travellers are less sensitive to cost differences. As a result, the values of mean hitting time lie between $(979 - 981)$ for different starting states as shown in Table 3.1. Obviously, it does not show noticeable changes compared to the mean hitting time for $\beta = 0$. As β increases, travellers take the cost differences into more consideration. Therefore, it is expected that when there is a small number of initial bus usage (i.e. x_1^0) the mean hitting time becomes larger for increased values of β. If the car is the choice of all travellers, which implies that $x_1 = 0$, then the system will be in the deterministic basin of attraction for the car use equilibrium and remain for an extended period; cf. Figure 3.4. Considering the results for column $x_1^0 = 9$, the mean hitting time declines when β goes up and drops to 6.28 for $\beta = 2$. At first glance, smaller values are expected for mean hitting time when β becomes more than 2. Instead, the mean hitting time increases to 108 for $\beta = 4$ as shown in the table.

To understand why, note that

$$m_9 = \sum_{t=1}^{\infty} t\,\mathbb{P}(T_{9,10} = t)$$

where $T_{9,10}$ is the first passage time from state χ_9 to χ_{10}. Now the probability mass for the distribution of $T_{9,10}$ is concentrated very heavily on small outcomes. For example, $\mathbb{P}(T_{9,10} = 1) = \mathbb{P}(x_1^1 = 10|\, x_1^0 = 9) = 0.984$. However, this first passage time distribution has a very long tail, so that even though $\mathbb{P}(T_{9,10} = t)$ will be small for large values of t, the products $t\mathbb{P}(T_{9,10} = t)$ will add appreciably to the value of the sum for even huge values of t. In simple terms, there is a very high chance for the system to reach universal bus usage based on the initial states, however, the system might suddenly shift to universal car usage with a very tiny possibility. As β is large, the system spends a very long time to exit from that equilibrium point and will be waiting for one of the flip events depicted in Figure 3.4.

We can generalize this result to show that there is no limit on the size of the mean hitting time μ_i through the following lemma.

Lemma: *Consider our two route system with $N = 4$ travellers. Given any $\Lambda > 0$ and initial state $\boldsymbol{x}^0 = \chi_i = (i, N - i)$ for $i < N$, there exists β such that*

1. $\sum_{t=1}^{\infty} \mathbb{P}(x_1^t = N, x_1^{t-1} \neq N, x_1^{t-2} \neq N, \ldots, x_1^1 \neq N | x_1^0 = i) = 1.$
2. $\mu_i > \Lambda.$

.

The first part of this result indicates that the system will eventually reach universal bus usage with probability 1, no matter what the initial state. The second part shows that the mean time required to do so can be arbitrarily large. While we focus on the case $N = 4$ for simplicity, the result should be extendible to any value of N.

Proof of part (1): We know our transition matrix elements, are

$$\binom{n}{x} \frac{e^{-\beta(n-x)(C_2-C_1)}}{\left(1 + e^{-\beta(C_2-C_1)}\right)^n}$$

when we used the logit model for the probability of bus choice. For any value of $\beta > 0$, these probabilities are non-zero, and hence all states intercommunicate and are aperiodic. It follows that the Markov chain $\{x_1^t : t = 1, 2, \ldots\}$ is finite regular and therefore all states are non-null recurrent. Part 1 of the lemma follows immediately.

Turning to part 2, we will work with an asymptotic expansion of the transition matrix in terms of $\gamma = e^{-\beta}$. Note that $\gamma \to 0$ and $\beta \to \infty$. For $N = 4$ the exact form of the transition matrix for the system is

$$P = \begin{bmatrix} \frac{e^{8\beta}}{(1+e^{2\beta})^4} & \frac{4e^{6\beta}}{(1+e^{2\beta})^4} & \frac{6e^{4\beta}}{(1+e^{2\beta})^4} & \frac{4e^{2\beta}}{(1+e^{2\beta})^4} & \frac{1}{(1+e^{2\beta})^4} \\ \frac{e^{4\beta}}{(1+e^{\beta})^4} & \frac{4e^{3\beta}}{(1+e^{\beta})^4} & \frac{6e^{2\beta}}{(1+e^{\beta})^4} & \frac{4e^{\beta}}{(1+e^{\beta})^4} & \frac{1}{(1+e^{\beta})^4} \\ \frac{1}{16} & \frac{4}{16} & \frac{6}{16} & \frac{4}{16} & \frac{1}{16} \\ \frac{e^{-4\beta}}{(1+e^{-\beta})^4} & \frac{4e^{-3\beta}}{(1+e^{-\beta})^4} & \frac{6e^{-2\beta}}{(1+e^{-\beta})^4} & \frac{4e^{-\beta}}{(1+e^{-\beta})^4} & \frac{1}{(1+e^{-\beta})^4} \\ \frac{e^{-8\beta}}{(1+e^{-2\beta})^4} & \frac{4e^{-6\beta}}{(1+e^{-2\beta})^4} & \frac{6e^{-4\beta}}{(1+e^{-2\beta})^4} & \frac{4e^{-2\beta}}{(1+e^{-2\beta})^4} & \frac{1}{(1+e^{-2\beta})^4} \end{bmatrix}.$$

Routine expansions then give

$$P = \begin{bmatrix} 1-4\gamma^2 & 4\gamma^2 & 0 & 0 & 0 \\ 1-4\gamma+10\gamma^2 & 4\gamma-16\gamma^2 & 6\gamma^2 & 0 & 0 \\ \frac{1}{16} & \frac{4}{16} & \frac{6}{16} & \frac{4}{16} & \frac{1}{16} \\ 0 & 0 & 6\gamma^2 & 4\gamma-16\gamma^2 & 1-4\gamma+10\gamma^2 \\ 0 & 0 & 0 & 4\gamma^2 & 1-4\gamma^2 \end{bmatrix} + O(\gamma^3)$$

where it is understood that the big O order notation applies elementwise to entries of the matrix.

The stationary distribution $\boldsymbol{\pi}^*$ (a row vector) satisfies $\boldsymbol{\pi}^* = P\boldsymbol{\pi}^*$. Tracking the requisite terms in the asymptotic expansions, it follows that

$$\boldsymbol{\pi}^* = \left(\tfrac{1}{2}, 2\gamma^2, \tfrac{192}{5}\gamma^4, 2\gamma^2, \tfrac{1}{2}\right)(\boldsymbol{e} + O(\gamma)) .$$

It follows that

$$I - P + W = \begin{bmatrix} \frac{1}{2} & -2\gamma^2 & \frac{192}{5}\gamma^4 & 2\gamma^2 & \frac{1}{2} \\ -\frac{1}{2} & 1 & -6\gamma^2 & 2\gamma^2 & \frac{1}{2} \\ \frac{7}{16} & -\frac{1}{4} & \frac{10}{16} & -\frac{1}{4} & \frac{7}{16} \\ \frac{1}{2} & 2\gamma^2 & -6\gamma^2 & 1 & -\frac{1}{2} \\ \frac{1}{2} & 2\gamma^2 & \frac{192}{5}\gamma^4 & -2\gamma^2 & \frac{1}{2} \end{bmatrix} (I + O(\gamma))$$

where $W = \boldsymbol{e}\boldsymbol{\pi}^*$ and so

$$Z = (I - P + W)^{-1} = \begin{bmatrix} \frac{1}{48\gamma^4} & \frac{1}{12\gamma^2} & -\frac{1536\gamma^4}{3} & -\frac{1}{\gamma^2} & -\frac{1}{48\gamma^4} \\ \frac{1}{48\gamma^4} & \frac{1}{12\gamma^2} & \frac{48\gamma^2}{5} & -\frac{1}{\gamma^2} & -\frac{1}{48\gamma^4} \\ -\frac{7}{10} & \frac{2}{5} & \frac{8}{5} & \frac{2}{5} & -\frac{7}{10} \\ -\frac{1}{48\gamma^4} & -\frac{1}{12\gamma^2} & \frac{48\gamma^2}{5} & \frac{1}{\gamma^2} & \frac{1}{48\gamma^4} \\ -\frac{1}{48\gamma^4} & -\frac{1}{12\gamma^2} & -\frac{1536\gamma^4}{3} & \frac{1}{\gamma^2} & \frac{1}{48\gamma^4} \end{bmatrix} (I + O(\gamma)) .$$

Collect together mean first passage times between each pair of states into the matrix

$M = (m_{ij})$. Then using equation (3.8), we find

$$M = \begin{bmatrix} 0 & \frac{1}{4\gamma^2} & \frac{1}{24\gamma^4} & \frac{1}{12\gamma^4} & \frac{1}{12\gamma^4} \\ \frac{1}{4\gamma^2} & 0 & \frac{1}{24\gamma^4} & \frac{1}{12\gamma^4} & \frac{1}{12\gamma^4} \\ \frac{1}{24\gamma^4} & \frac{1}{24\gamma^4} & 0 & \frac{1}{24\gamma^4} & \frac{1}{24\gamma^4} \\ \frac{1}{12\gamma^4} & \frac{1}{12\gamma^4} & \frac{1}{24\gamma^4} & 0 & \frac{1}{4\gamma^2} \\ \frac{1}{12\gamma^4} & \frac{1}{12\gamma^4} & \frac{1}{24\gamma^4} & \frac{1}{4\gamma^2} & 0 \end{bmatrix} (I + O(\gamma)) \ .$$

It follows immediately that all off the mean first passage times between different states can be made arbitrarily large by making $\gamma = e^{-\beta}$ sufficiently small, completing the proof of part 2 of the Lemma.

Notice that the final column of M gives the asymptotic mean hitting times for state $x_1 = 4$ (i.e. universal bus usage) as $\beta \to \infty$. These theoretical results mirror the empirical findings from the large ($N = 10$) system supplied in Table 3.1. In particular, the mean hitting times are very large when starting in states $x_1 = 0$ or $x_1 = 1$ when β is large. The mean hitting time when starting in state $x_1 = 3$ also grows with β, but at an asymptotically much slower rate.

Finally, we observe that the theoretical result from the Lemma could be extended to a system with any number of travellers. This can be achieved by combining states in the large system, to produce a Markov process with 5 states when the proof can then follow the same lines as above. The idea behind of the combining states is, assume we are given an $N+1$-state Markov chain with transition probability P. Let $A = \{A_1, A_2, \cdots, A_t\}$ be a partition of the set of states. We form a new process as follows. The outcome of the $j-$th iteration in the new process is the set A_k that contains the outcome of the $j-$th step in the original chain. We define the branch probabilities as follows: At the zero level we assign $P(X_0 \in A_i)$. At the first level we assign $P(X_1 \in A_j | X_0 \in A_i)$. In general at the $n-$th level we assign branch probability $P(X_n \in A_t | X_{n-1} \in A_s \cap \ldots \cap X_1 \in A_j \cap X_0 \in A_i)$.

The above procedure could be used to reduce a process with a very large number of states to a process with a smaller number of states. In our two-route example we are considering the odd number of states (even number of travellers) then we combined the states between the first state and the middle state to a new state and also combined the states between middle state and the last state together. An illustration of how this works for $N = 6$ is provided in the Appendix A. We obtained the transition matrix in general should be as follows.

$$P = \begin{pmatrix} 1 - N\gamma^2 & N\gamma^2 & 0 & 0 & 0 \\ 1 - N\gamma^{\frac{2N-4}{N}} & N\gamma^{\frac{2N-4}{N}} & \ldots & 0 & 0 \\ \frac{1}{2^N} & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & N\gamma^{\frac{2N-4}{N}} & 1 - N\gamma^{\frac{2N-4}{N}} \\ 0 & 0 & 0 & N\gamma^2 & 1 - N\gamma^2 \end{pmatrix}.$$

The stationary vector for combining states is

$\boldsymbol{\pi} = \left(\frac{1}{2} - \frac{N}{2}\gamma^2, \frac{N}{2}\gamma^2, \gamma^\ell, \frac{N}{2}\gamma^2, \frac{1}{2} - \frac{N}{2}\gamma^2\right)$ where $\ell \geq 3$.

If we named our new states as $\{A_0, B, A_{N/2}, D, A_N\}$ then to prove that the lemma applies to a model with N travellers, we need to show $\mu_{BA_0} = \mu_{DA_N}$ is equal to some negative power of γ. By doing the same that we showed for $N = 4$ the mean first passage time from state B to state A_0 and from state D to state A_N is given by $\mu_{BA_0} = \mu_{DA_N} = \frac{1}{N\gamma^2}$. It is obvious that when γ tends to zero then $\mu_{BA_0} = \mu_{DA_N}$ tends to infinity and this completes the proof of the second part and this demonstrates how the results can extend to arbitrary values of N.

3.5 Numerical results for a network with different OD pairs

There is a history in transportation research of using toy networks comprising two parallel routes for preliminary investigations (e.g. Horowitz (1984)). Much can be learned from such simple examples, as we saw in the previous sections. However, transport networks gain a further dimension of complexity and interest when there are links carrying traffic from two or more routes. Working with such systems, we can examine the effects of interaction between traffic from different OD pairs.

With those comments in mind, we consider in this Section the 7 link network depicted in Figure 3.5. This network has two OD pairs, $1 \to 5$ and $3 \to 5$, with correspondsing (fixed) demands N_1 and N_2. The link path incidence matrix is

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} .$$

The vector $\mathbf{c}$ of link cost functions has ith element defined by

$$c_i(y_i) = a_i + b_i y_i^{\phi_i} \ . \tag{3.9}$$

This example could be considered as an extension of the two link network taking into account three different transport mode choices. For the first OD pair, travellers can take a fixed cost walk (link 1) and then use the demand-responsive bus service (link 4), or they can take a fixed cost walk (link 3) and then travel by car (link 6). These options correspond to routes 1 and 2; that is, columns 1 and 2 of A. Assuming that route 3 corresponds to bus usage and car usage will be taken route 4, then we have similar travel options for the second OD pair. Notice that the bus (link 4) carries travellers from both OD pairs. The parameters used are given in Table 3.2.

We first examine the long-term behaviour of the system for the deterministic day-to-day model described by equations (3.1) and (3.2). We assume that both OD pairs have the same OD demand of 50 travellers (i.e.$N_1 = N_2 = 50$). Three different equilibria exist. We

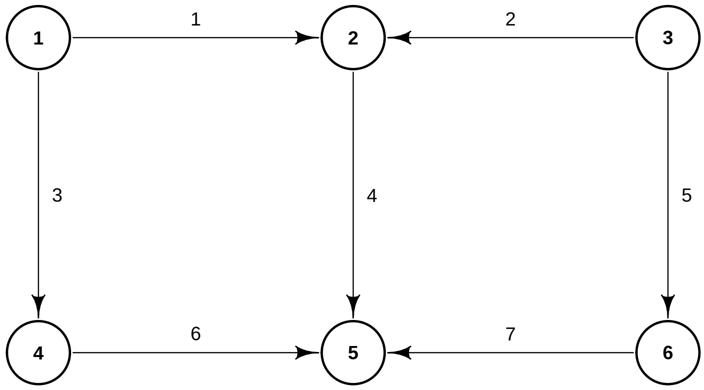


Figure 3.5: In this 7 link network, the origin destination (OD) pairs are $1 \to 5$ and $3 \to 5$. Each OD pair is serviced by two routes.

Link (i)	a_i	b_i	ϕ_i
1	1	0	0
2	1	0	0
3	1	0	0
4	8	$-8/(N_1 + N_2)$	1
5	1	0	0
6	2	$4/N_1$	1
7	2	$4/N_2$	1

Table 3.2: Parameters of link cost functions for the 7 link network.

have two stable and one unstable equilibria. A stable equilibrium point corresponding to universal bus usage occurs at route flow pattern $(50, 0, 50, 0)^\top$. When the route flow pattern is at $(0, 50, 0, 50)^\top$, the system reaches another stable equilibrium point related to the car usage equilibrium. $(25, 25, 25, 25)^\top$ is an unstable equilibria where all routes have the same number of travellers. Figure 3.6 represents the basins of attraction for each of the equilibria with swapping parameter $k = 0.1$. Due to the symmetric nature of the network, the state of the system is characterised by bus usage for both OD pairs.

The darker upper triangle is the domain of attraction for universal bus usage. So, for any initial flow vector in this area the system will quickly go to the stable equilibria corresponding the universal bus usage in response to a temporary intervention. Similarly, the system will absorb into another stable equilibrium point due to any temporary change to the network when the initial flow is in the light shaded lower triangle which is the basin of attraction for universal car usage. The mid-grey diagonal line (from top left to bottom right) is the basin of attraction for the central equilibrium, but the width has been exaggerated for display purposes. Any intervention for the unstable equilibria the system will lead to one the two stable equilibrium points.

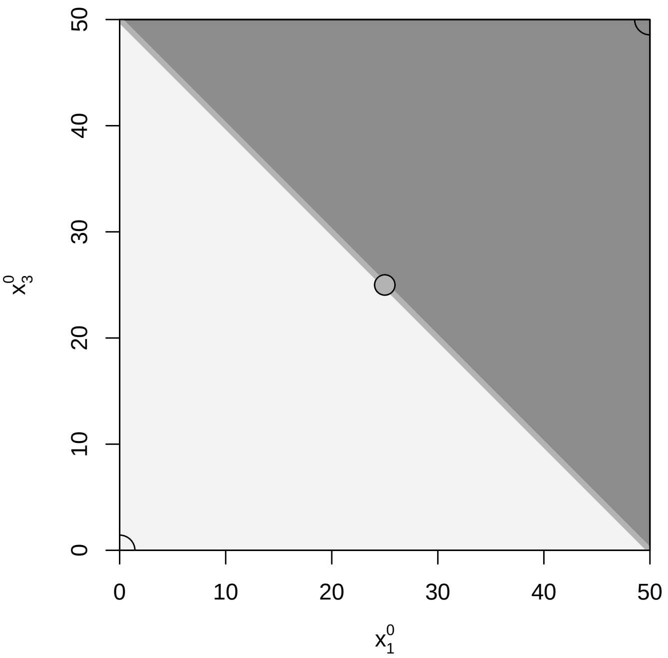


Figure 3.6: Equilibria (marked by circles or segments of circles) and corresponding basins of attraction for the deterministic day-to-day model on the 7 link network. The basin of attraction for the central equilibrium has been exaggerated for display purposes: in reality it is a line of zero width.

In order to illustrate the application of our stochastic day-to-day model to the network with respect to the value of mean hitting time, we set the logit parameter to be β = 3.0. Calculating the mean hitting time for every possible initial flow pattern is impossible because the system has $51 \times 51 = 2601$ different states and through equation (3.8) we need to find the inverse of a matrix with the dimension of 2601×2601. To avoid this, we obtain the estimates based on long-run simulation. The square area of the heat map plot in Figure 3.7 defined by $\{x_1^0 > 25, x_3^0 > 25\}$ shows the initial states with low mean hitting time and darker shading indicates the larger values for mean hitting time. The square shape of the high smaller hitting times in 3.7 contrast the triangular shape of the domain of attraction. It seems we need high initial bus usage on both routes in order to hit universal bus usage relatively quickly.

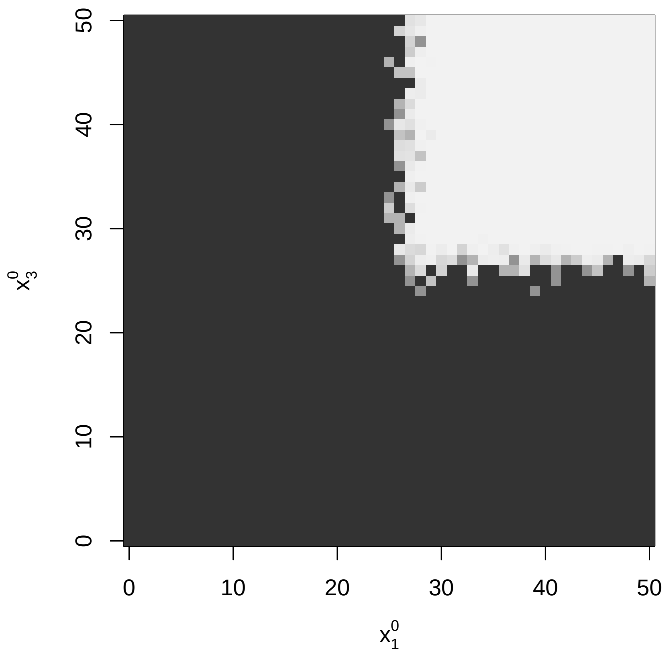


Figure 3.7: Estimated mean hitting times for universal bus usage using the stochastic day-to-day model on the 7 link network. Darker shading indicates larger values. The means are truncated above at $m = 100,000$ days, explaining the lack of differentiation in the dark areas of the plot.

It would be interesting to take a closer look to the situations in our example when for one OD pair the number of bus travellers is close 50 and at the same time, this number is close to 0 for the other OD pair. To explore the system behaviour for such cases, Figure 3.8 displays the trace plots of flows on routes 1 and 3 by day using the deterministic model with swapping parameter values $k = 0.05$ and $k = 0.5$ when the initial states is $\boldsymbol{x}^0 = (48, 2, 5, 45)^\mathsf{T}$. Since $x_1^0 + x_3^0 > 50$, according to the Figure 3.6 this initial flow vector is within the domain of attraction for universal bus usage, therefore, as expected the system converges to universal bus usage equilibrium point (i.e.$(50, 0, 50, 0)^\mathsf{T}$). But the convergence behavior for these two is slightly different and for a system with a larger value of k a few flip-flopping behaviours are observed in the initial stages. In addition, for the more reactive system (second plot), it took fewer days to reach the equilibrium point. For a hyper-reactive system with $k = 0.8$ more drastic fluctuations happen and the system never reaches equilibrium. In this case the flow patterns fall well outside the feasible set (we get negative flows and flows well beyond the demand), and so the theory for basins of attraction does not apply.

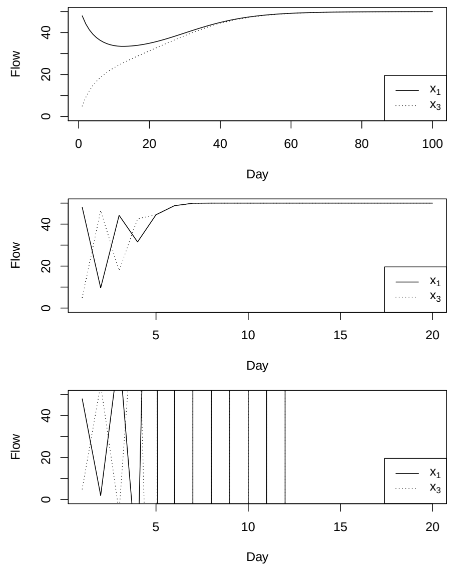


Figure 3.8: Trace plots of flows on routes 1 and 3 (bus usage for each of the OD pairs) for the deterministic model. The upper plot was obtained using swap parameter $k = 0.05$; the middle plot using $k = 0.5$ and the lower plot using $k = 0.8$.

Now, we move the initial flow vector towards the bus usage equilibrium point by considering $\boldsymbol{x}^0 = (45, 5, 20, 30)^\mathsf{T}$. Still $x_1^0 + x_3^0 > 50$ and the starting point is inside of the domain of attraction for universal bus usage. As shown in Figure 3.9, applying the same swapping parameter values for the deterministic model, the patterns of flow are pretty similar to what we have seen in the previous case, except when travellers are very sensitive to the cost differences (i.e. $k = 0.8$). As can be seen in the third plot albeit the flow routes go above the total than the total demand for a few days but going further, convergence to the desired equilibrium state is attained.

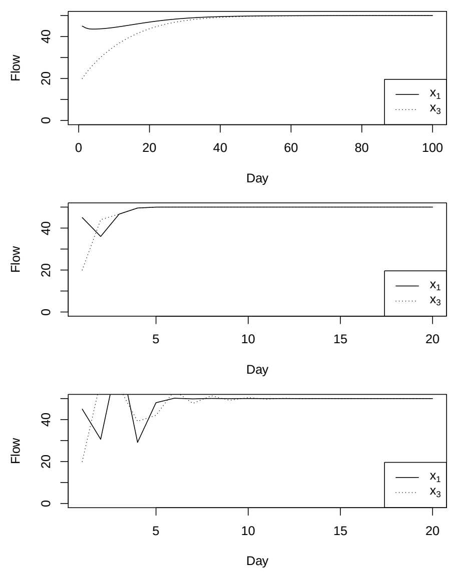


Figure 3.9: Trace plots of flows on routes 1 and 3 (bus usage for each of the OD pairs) for the deterministic model. The upper plot was obtained using swap parameter $k = 0.05$; the middle plot using $k = 0.5$ and the lower plot using $k = 0.8$.

For comparison purposes, Figure 3.10 shows the results from stochastic model with initial state $\boldsymbol{x}^0 = (48, 2, 5, 45)^\mathsf{T}$ and with logit parameters $\beta = 3$ and $\beta = 1$ respectively. As can be seen, a larger value for β widens the variation in route flows, because higher sensitivity to cost differences results in drastic changes of travellers' daily route choice behaviour. Accordingly, as shown in the first plot when all travellers use the car for first OD pair nobody use the car in the second OD pair. Strictly speaking, the system oscillates between two non-equilibrium states $(0, 50, 50, 0)^\mathsf{T}$ and $(50, 0, 0, 50)^\mathsf{T}$.

Considering the lower sensitivity parameter $\beta = 1$, the second plot displays behaviour rather different from anything that we have seen earlier in this chapter. The random fluctuations are not centred around the stable equilibria. In other words, we cannot produce anything through the deterministic model that is comparable to this situation.

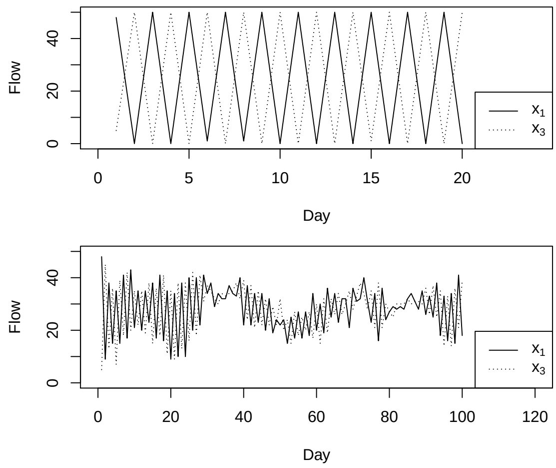


Figure 3.10: Trace plots of flows on routes 1 and 3 (bus usage for each of the OD pairs) for the stochastic model. The upper plot was obtained using logit parameter $\beta = 3$; the lower plot using $\beta = 1$.

3.6 Summary

Stochastic models for dynamic traffic assignment have the capacity to account for haphazard variation in a way that deterministic models do not. In the transportation literature some works (e.g. Davis and Nihan (1993), Cantarella and Cascetta (1995), Watling and Cantarella (2013)) addressed the relationships between deterministic processes, together with corresponding equilibrium states, and stochastic probability distributions. They have shown for a system with single deterministic equilibrium, provided the OD demand and capacity values are large enough, the deterministic equilibrium route flows tends to be close to the mean values obtained through a stochastic process model.

In this chapter, we investigated the relationships between deterministic and stochastic models when the network has more than one equilibrium point. In such a case, the equilibrium distribution for stochastic models would be expected to be multimodal, with peaks approximated by the stable equilibria. Therefore, the mean flow is not a useful summary of the system. Furthermore, we found that the transition times between various states are also affected. As shown in Section 3.3 and 3.4, with deterministic models, once a state in the domain of attraction is reached it moves towards the desired equilibrium. However, for stochastic models, shifts between states are not necessarily in a certain direction and in some cases this means that it can take a very long time to reach a desired equilibrium state. This is another failure of the mean as a useful summary of the system.

It should be noted that the deterministic model that we have chosen is not the closest analogue of the stochastic model under study. We could have defined our deterministic model following Cantarella's (1995) model where, $\boldsymbol{x}_j^t = N\boldsymbol{p}_j^t$, which seems like a closer match to our Markov model. However, the trouble with using such a model is that the equilibria are SUE flow patterns which are interior points in the flow space. Therefore, the system will never reach extreme states like universal bus usage, rendering impossible the kind of comparison that we draw in Section 3.4.

Our achievements suggest that, to compare the results of the deterministic and stochastic models for a multi-equilibria system, substantial care would be required. Apart from the mean, results of the deterministic model in many cases correspond to the most probable stochastic model outcome. However, in the bottom panel of Figure 3.10 we saw behaviour from the stochastic model that deterministic models cannot imitate.

Chapter 4

Assessing the Stability of Approximation Methods for Traffic Assignment

4.1 Chapter overview

Stochastic models are attractive for representing traffic dynamics over a network because of the unpredictable fluctuations in traffic flow from day-to-day. Nonetheless, it is easier to work with deterministic models, which seek to describe the behaviour of the system "on average". In particular, there is now a large body of theory on deterministic models.

A potentially attractive hybrid approach is to approximate discrete-flow stochastic models by a Gaussian random process, in which the dynamics of the mean follow one of these well understood deterministic models. This is known to work well for stable traffic networks, but Watling and Hazelton (2018) have recently shown that such approximations can break down badly for highly unstable systems. The question remains, how can one tell if a system is sufficiently stable for the Gaussian approximation to be serviceable?

In this chapter we show that this question can be examined through a measure called the coefficient of reactivity introduced by Hazelton (2002). The coefficient of reactivity measures stability of a system following a disruption. This is illustrated through a range of numerical experiments, including a variety of different road networks.

This chapter is organized as follows. Section 4.2 reviews the literature on approximation methods and introduces the different stochastic approximations that will be studied. In Section 4.3, the definition of coefficient of reactivity will be given. This is followed by an illustrative example along with some simulation results to show how this quantity works for models with memory length set to one day and having one day disruption. Section 4.4 presents simulation results based on different networks and shows the relationship between the value of coefficient of reactivity and the quality of the approximation methods. Major finding are recapped in Section 4.5.

4.2 Approximation methods

As described in Chapter 2, Markovian assignment models represent the day-to-day evolution of traffic flows on a transport network. Such stochastic day-to-day models can represent the system both in a stationary state, but also in transitional states (e.g. following an intervention). We will be interested in modelling systems in both cases. Usually, stochastic models are analysed with the help of Monte Carlo techniques to generate a pseudo-random observation of the whole process over a given period of time. But such methods are not free from drawbacks, such as confirming the stationarity of the process, and are not easily amenable to mathematical analysis and reproducibility (as results vary from time to time). In addition, they require a lot of computer time for networks and demand levels of reasonable practical significance. Facing the fact that the implementation and interpretation of stochastic simulation of Markovian assignment models is not that simple, this has motivated research to develop fast and efficient methods of working out properties of the equilibrium probability distribution (e.g. Cantarella and Cascetta 1995; Hazelton and Watling 2004; Balijepalli and Watling 2005). However, it does not help us model the transient behaviour of the system.

Cantarella and Cascetta (1995) noted that, provided the OD demand and capacity values are large enough, the deterministic equilibrium flows become closer to the mean values obtained through a stochastic process model. Later on Hazelton and Watling (2004) developed a practical approach for estimating the equilibrium covariance matrix for a particular class of Markovian assignment models characterised by a traveller learning mechanism based on linear learning filters with weights that are exponentially decreasing. They attempted to compute the equilibrium covariance matrix in a day-to-day context (assuming static within day cost flow functions) and proved that the equilibrium probability distributions can be approximated in a fraction of time as compared to solving Davis and Nihan's (1993) fixed point equations or the method of simulating the route choices using Monte Carlo techniques.

Theoretically investigating the behaviour in the limiting case when the population becomes large is an alternative to Monte Carlo simulation techniques. By considering this, Davis and Nihan (1993) demonstrated that Markovian assignment models could be approximated by Gaussian multivariate autoregressive processes as travel demand becomes large. They estimated the mean of a Markov chain assignment in a day-to-day context and demonstrated that it converges to the stochastic user equilibrium assignment in stationary conditions, hence motivating further research in approximating the properties of equilibrium probability distributions and transient behaviour. This significant result established the foundation for further efforts in unifying the deterministic and stochastic approaches in traffic modelling. Watling and Hazelton (2018) provided a more straightforward way of producing "large demand" approximations, but it requires more assumptions than Davis and Nihan (1993). The practical cosequence is seen in how well the two methods work during transient periods.

Behavioural effects after any network disruption could have significant impact on day-to-day travel demand. However, there have been a few works considering the issue of transience in a day-to-day dynamic context. Zhu et al. (2010) studied the day-to-day dynamics of the system, with flow oscillation observed after the 2007 I-35 Bridge collapse in Minneapolis, Minnesota. The oscillations lasted for approximately six weeks before travellers behaviour returned to normal. A study by Watling et al. (2012) analysed the impact of planned disruptions in the city of York on traveller behaviour. Using data from number plate surveys, they successfully determined considerable impact in routing patterns. The effectiveness of day-to-day traffic models to describe the transient behaviour after network disruption (the 2007 I-35 Bridge collapse in Minneapolis, Minnesota) has been verified in He and Liu (2012).

In the next section we describe the technical details of two stochastic day-to-day approximation methods introduced by Davis and Nihan (1993) and Watling and Hazelton (2018) and also one deterministic day-to-day method introduced by Cantarella and Cascetta (1995).

4.2.1 Modelling framework

In this chapter we focus on Markov day-to-day models with memory length $m = 1$. Route choices on day $t+1$ depend only on costs from the previous day. In a small change to previous notation, for this subsection we now let X_r be the flow on route r. We introduce ξ to be a demand multiplier, and define $x_r = X/\xi$ to be a standardized version of flow. Focussing on the standardized flows will allow us to examine cases in which the demand becomes large in a coherent manner. We first describe the exact Markov model, which other models will seek to approximate.

Our model sets the disutility $\boldsymbol{u}^t = \boldsymbol{C}(\boldsymbol{x}^{t-1})$ where $\boldsymbol{x}^{t-1}$ is the standardized flow vector from the previous day. Note that setting costs in terms of standardized flows means that we can work with a fixed set of cost function parameters. Now the route choice probabilities are defined via a logit model, $p_r = \frac{\exp\{-\theta u_r^t\}}{\sum_s \exp\{-\theta u_s^t\}}$, and the flows on day t are determined by multinomial distributions. That is,

$$\boldsymbol{X}_j^t | \boldsymbol{u}^t = \mathrm{Mn}(\xi, \boldsymbol{p}_j(\boldsymbol{u}^t)). \tag{4.1}$$

The process $\boldsymbol{X}^t$ is a Markov chain with discrete state space. The state of the system at time t is defined by $\boldsymbol{s}^t = \begin{pmatrix} \boldsymbol{u}^t \\ \boldsymbol{x}^t \end{pmatrix}$.

We first develop the approximation methods (i.e. Davis and Nihan's (1993) and Watling and Hazelton's (2018)) for systems with a single OD pair. For our setting with one day memory, the approximations can be written as

$$\begin{pmatrix} \boldsymbol{u}^t \\ \boldsymbol{x}^t \end{pmatrix} = \begin{pmatrix} \bar{\boldsymbol{u}}^t \\ \bar{\boldsymbol{x}}^t \end{pmatrix} + e^t \tag{4.2}$$

where $\bar{\boldsymbol{u}}^t = \boldsymbol{C}(\bar{\boldsymbol{x}}^{t-1})$ and $\bar{\boldsymbol{x}}^t = \xi \boldsymbol{p}(\bar{\boldsymbol{u}}^t)$ are dynamic deterministic quantities. Because the conditional mean is represented as a dynamic process, this approximation does not just cover the process in its stationary state, it also includes its behaviour during transient periods. We linearize these mean functions, using Jacobians evaluated at the current state. So, for both approximation methods the mean vector $\boldsymbol{\mu}^t = \mathsf{E}[\boldsymbol{s}^t]$ evolves through

$$\boldsymbol{\mu}^t = \boldsymbol{s}^* + M^t(\boldsymbol{\mu}^{t-1} - \boldsymbol{s}^*)$$

and the covariance matrix which evolves in time via

$$\Sigma^t = M^t \Sigma^{t-1} {M^t}^T + V^t. \tag{4.3}$$

The difference between the two methods is in how M^t is evaluated. For Davis and Nihan (1993), this matrix evolves with time but for Watling and Hazelton (2018) it is fixed. Here, the definition of M^t for Davis and Nihan (1993) approximation method is given by

$$M^t = \begin{pmatrix} 0 & \frac{\partial \boldsymbol{C}}{(\partial \boldsymbol{x})} \\ 0 & \frac{\partial \boldsymbol{C}}{(\partial \boldsymbol{x})} \frac{\partial \boldsymbol{p}}{(\partial \boldsymbol{u})} \end{pmatrix},$$

where, $\frac{\partial \boldsymbol{C}}{(\partial \boldsymbol{x})}$ and $\frac{\partial \boldsymbol{p}}{(\partial \boldsymbol{u})}$ are Jacobian matrices evaluated at $\boldsymbol{s}^{t-1}$. Hence making the matrix M^t changes with time. Also, V^t is a covariance matrix given by

$$V^t = \begin{pmatrix} 0 & 0 \\ 0 & \operatorname{diag}(\boldsymbol{x}^t) - \boldsymbol{x}^t {\boldsymbol{x}^t}^T \end{pmatrix}.$$

For Watling and Hazelton (2018), M^t is given by

$$M^t = \begin{pmatrix} 0 & \frac{\partial \boldsymbol{C}}{(\partial \boldsymbol{x}^*)} \\ 0 & \frac{\partial \boldsymbol{C}}{(\partial \boldsymbol{x}^*)} \frac{\partial \boldsymbol{p}}{(\partial \boldsymbol{u}^*)} \end{pmatrix},$$

where the Jacobians are evaluated at SUE. The covariance matrix V^t can then given by

$$V^t = \begin{pmatrix} 0 & 0 \\ 0 & \operatorname{diag}(\boldsymbol{x}^*) - \boldsymbol{x}^* {\boldsymbol{x}^*}^T \end{pmatrix}.$$

It should be noted that the normality of the error term e^t in 4.2 follows by applying the Central Limit Theorem to the (conditional) multinomial distribution as ξ becomes large.

As Davis and Nihan (1993) and Watling and Hazelton (2018) showed, these results can also be extended to networks with multiple OD pairs. Following the notation of Watling and Hazelton (2018), we define Γ to be the route-OD incidence matrix, and $\xi \boldsymbol{\mathcal{W}}$ to be the vector of OD demands. Then the matrix M^t generalizes to equation 4.4 where B is the Jacobian matrix $\partial \boldsymbol{C}/\partial \boldsymbol{x}$ and D the Jacobian matrix $\partial \boldsymbol{p}/\partial \boldsymbol{u}$ evaluated at s^t in the case of Davis and Nihan's (1993) approximation, and evaluated as SUE in the case of Watling and Hazelton's (2018) approximation. Note that M^t remains unchanged through time in the latter case,

$$M^t = \begin{pmatrix} I \\ \mathrm{diag}(\Gamma\boldsymbol{\mathcal{W}})D \end{pmatrix} \begin{pmatrix} 0 & B \end{pmatrix}. \tag{4.4}$$

In the case of multiple OD pairs, the covariance matrix V has a block diagonal structure, with each block corresponding to a single OD pair. The matrix V_j in the jth block is given by

$$V_j = \begin{pmatrix} 0 & 0 \\ 0 & \mathcal{W}_j^{-1}\mathrm{diag}(\boldsymbol{x}_j^*) - (\boldsymbol{x}_j^*)(\boldsymbol{x}_j^*)^T \end{pmatrix}$$

for Davis and Nihan's (1993) method, and by

$$V_j^t = \begin{pmatrix} 0 & 0 \\ 0 & \mathcal{W}_j^{-1}\mathrm{diag}(\boldsymbol{x}_j^{t-1}) - (\boldsymbol{x}_j^{t-1})(\boldsymbol{x}_j^{t-1})^T \end{pmatrix}.$$

The other approximation method that we use is a deterministic approximation method introduced by Cantarella and Cascetta (1995). This method works by computing the mean flow vector on day t conditional on the network state on day $t-1$. Mathematically speaking,

$$\boldsymbol{x}_j^t = \mathsf{E}[\boldsymbol{x}_j^t | \boldsymbol{x}_j^{t-1}] = N\boldsymbol{p}_j^t. \tag{4.5}$$

Since the travellers' route choice behaviour depends on the anticipated costs it is modeled through $\boldsymbol{p}_j^t = \boldsymbol{p}_j(\boldsymbol{u}_j^t)$.

This approximation method is fully deterministic because it assumes that the route flow vector is equal to its expected value. Therefore, this model seeks only to approximate the mean flow for day t. For future reference, we note that the idea can be expanded trivially when the probability of choosing a route on day t depends on a weighted mean cost from the m.

4.3 The coefficient of reactivity

As stated before, the day-to-day evolution of traffic flows, where the decision made by travellers' at any given time point governed by travellers' experiences during the finite past, can be captured through Markovian assignment models. The structure of these models provides a framework for investigating travellers day-to-day behaviour adjustment under day-to-day variation. Hazelton (2002) suggests that it would be potentially useful to define a new tool that can measure the magnitude of day-to-day fluctuations in flow following a disruption to the network. From this point of view, for a Markovian assignment model with memory length m, Hazelton (2002) established the following measure called coefficient of reactivity

$$\omega = \sup_{\boldsymbol{x}\in\mathcal{X}} \frac{\|\psi_m(\boldsymbol{x}) - \boldsymbol{x}^*\|}{\|\boldsymbol{x} - \boldsymbol{x}^*\|}, \tag{4.6}$$

where $\mathcal{X}$ gives the set of feasible flows, $\boldsymbol{x}^*$ denote the stationary mean for route flows vector and $\psi_m(\boldsymbol{x})$ is the conditional expectation of $\boldsymbol{x}^{(t)}$ as a function of the route flow on day $t-1$ given by

$$\psi_m(\boldsymbol{x}) = E[\boldsymbol{x}^{(t)}|\boldsymbol{x}^{(t-1)} = \boldsymbol{x}, \boldsymbol{x}^{(t-2)} = \boldsymbol{x}^*, \ldots, \boldsymbol{x}^{(t-m+1)} = \boldsymbol{x}^*]. \tag{4.7}$$

Suppose that due to a disruption in the system, the flow pattern at day t is $\boldsymbol{x}$. Then what can we expect from the flow pattern at day $t+1$? Of course, because the model is stochastic, there will be a range of possible outcomes on day $t+1$ in response to the disruption at day t. To simplify things, we focus on the mean flow pattern on day $t+1$, viewed as a function of $\boldsymbol{x}$. That's what $\psi_m(\boldsymbol{x})$ represents.

The coefficient of reactivity therefore compares the "oddness" of the expected flow pattern on day $t+1$ (i.e. the distance of that expected flow pattern from the long term mean $\boldsymbol{x}^*$) with the "oddness" of the disrupted flow pattern on day t. If the system is unresponsive then we expect the flow pattern on day $t+1$ to largely revert back to its usual state ($\boldsymbol{x}^*$), but if the system is very reactive we expect $\psi_m(\boldsymbol{x})$ to be very different to $\boldsymbol{x}^*$.

The coefficient of reactivity, with regard to route choice behaviour of travellers, measures the amount of average traveller response to variation in flow patterns in the recent past. In other words, this quantity represents how much the system is affected on average by unusual flow patterns on previous days. The coefficient of reactivity does not directly measure the time it takes for the system to settle back to normal, but we do expect a system with a high coefficient of reactivity to take a relatively long time to do so.

We can see from equation 4.7, the marginal influence of today's flow pattern on the flow pattern expected for tomorrow is given by $\dfrac{\|\psi_m(\boldsymbol{x}) - \boldsymbol{x}^*\|}{\|\boldsymbol{x} - \boldsymbol{x}^*\|}$. This influence is measured in terms of the position relative to the stationary mean. The largest value this fraction can take over the feasible set represents the value of coefficient of reactivity. When the expected flow pattern tomorrow compared to today's flow pattern is closer to the stationary mean, the output of the corresponding fraction will be less than one. On the contrary, if the expected flow pattern tomorrow compared to today's flow pattern is further from the stationary mean, a value greater than one would be expected for this fraction.

It is computationally difficult to find ω through the definition (4.1). To make it easy to calculate, Hazelton (2002) explored a surrogate definition for ω as follows.

Definition (*Asymptotic coefficient of reactivity*). For a Markovian assignment model with memory length m, the asymptotic coefficient of reactivity, ω_0 is defined to be

$$\omega_0 = \lim_{\epsilon \to 0} \sup_{\boldsymbol{x} \in \mathcal{X}_\epsilon} \frac{\|\psi_m(\boldsymbol{x}) - \hat{\boldsymbol{x}}\|}{\|\boldsymbol{x} - \hat{\boldsymbol{x}}\|}, \tag{4.8}$$

where $\mathcal{X}_\epsilon = \{\boldsymbol{x} : \boldsymbol{x} \in \mathcal{X}, \|\boldsymbol{x} - \hat{\boldsymbol{x}}\| < \epsilon\}$. In this definition $\hat{\boldsymbol{x}}$ is the stochastic user equilibrium route flow vector. The following theorem enables us to compute asymptotic coefficient of reactivity through straightforward calculations.

Theorem 1:(Hazelton, 2002) *Assume that ψ and all its first-order partial derivatives exist in some neighborhood of $\hat{\boldsymbol{x}}$, and let* $\mathbf{J}$ *denote the Jacobian matrix of ψ evaluated at $\hat{\boldsymbol{x}}$. Let* $\mathbf{M} = I_n - n^{-1}\boldsymbol{1}_n\boldsymbol{1}_n'$ *where n is the length of the vector $\boldsymbol{x}$, I_n is the n-dimensional identity matrix and $\boldsymbol{1}_n$ is a vector of n ones. Then ω_0^2 is the largest eigenvalue of the matrix* $\mathbf{MJ'J}$.

We illustrate some properties of the coefficient of reactivity using a simple example and show the calculation of the asymptotic coefficient of reactivity.

Example 4.3.1. We consider the two route network described in Example 2.7.1 where the link cost function is $c_l(\boldsymbol{y}) = c_l(y_l) = a_l + \left(\frac{y_l}{b}\right)^2$ for $l = 1, 2$, and the route choice probability is determined by a logit random utility model. As noted in chapter 2 we can find the SUE for a two-route network using the following equation:

$$\frac{\hat{x}_1}{N} = \frac{\exp\{-\theta c_1(\hat{x}_1)\}}{\exp\{-\theta c_1(\hat{x}_1)\} + \exp\{-\theta c_2(\hat{x}_2)\}}.$$

The above equation can be written as

$$\hat{x}_1 = \frac{N}{1 + \exp\left\{-\theta\left[a_1 - a_2 + \left(\frac{\hat{x}_1}{b}\right)^2 - \left(\frac{\hat{x}_2}{b}\right)^2\right]\right\}}.$$

Using the fact that $\hat{x}_1 + \hat{x}_2 = N$, $\hat{x}_2$ can be calculated by $\hat{x}_2 = N - \hat{x}_1$. Since we have a simple two route network the number of travellers on day t follows the binomial distribution, therefore $E(\boldsymbol{x}^{(t)}|\boldsymbol{x}^{(t-1)}) = N\boldsymbol{p}^t$, where $\boldsymbol{p}^t$ is a vector of probabilities. As a consequence $\psi_1(x_1^{t-1}, x_2^{t-1})$ can then given by

$$\frac{N}{1 + \exp\left\{-\theta\left[a_2 - a_1 + \left(\frac{x_2^{t-1}}{b}\right)^2 - \left(\frac{x_1^{t-1}}{b}\right)^2\right]\right\}}.$$

Similarly, $\psi_2(x_1^{t-1}, x_2^{t-1})$ can be expressed by

$$\frac{N}{1 + \exp\left\{-\theta\left[a_1 - a_2 + \left(\frac{x_1^{t-1}}{b}\right)^2 - \left(\frac{x_2^{t-1}}{b}\right)^2\right]\right\}}.$$

In order to simplify the notation define $\exp\left\{-\theta\left[a_2 - a_1 + \left(\frac{x_2^t}{b}\right)^2 - \left(\frac{x_1^t}{b}\right)^2\right]\right\} = \varrho_1$ and $\varrho_2 = \exp\left\{-\theta\left[a_1 - a_2 + \left(\frac{x_1^{t-1}}{b}\right)^2 - \left(\frac{x_2^{t-1}}{b}\right)^2\right]\right\}$.

The Jacobian matrix of ψ evaluated at SUE is given as

$$\boldsymbol{J} = \begin{bmatrix} \frac{\partial\psi_1(x_1^{t-1}, x_2^{t-1})}{\partial(x_1^{t-1})} & \frac{\partial\psi_1(x_1^{t-1}, x_2^{t-1})}{\partial(x_2^{t-1})} \\ \frac{\partial\psi_2(x_1^{t-1}, x_2^{t-1})}{\partial(x_1^{t-1})} & \frac{\partial\psi_2(x_1^{t-1}, x_2^{t-1})}{\partial(x_2^{t-1})} \end{bmatrix}$$

by elementary calculus it can then be written as

$$\boldsymbol{J} = \begin{bmatrix} \frac{-N\theta\left(\frac{2x_1^{t-1}}{b^2}\right)\varrho_1}{(1+\varrho_1)^2} & \frac{N\theta\left(\frac{2x_2^{t-1}}{b^2}\right)\varrho_1}{(1+\varrho_1)^2} \\ \frac{N\theta\left(\frac{2x_1^{t-1}}{b^2}\right)\varrho_2}{(1+\varrho_2)^2} & \frac{-N\theta\left(\frac{2x_2^{t-1}}{b^2}\right)\varrho_2}{(1+\varrho_2)^2} \end{bmatrix}.$$

It can be seen that $\frac{\varrho_1}{(1+\varrho_1)^2} = \frac{\varrho_2}{(1+\varrho_2)^2}$ and $\hat{x}_1\hat{x}_2 = \frac{N^2\varrho_1}{(1+\varrho_1)^2}$. Following that, the Jacobain matrix can be summarised as

$$\boldsymbol{J} = 2\theta\hat{x}_1\hat{x}_2 b^{-2}N^{-1}\begin{bmatrix} -\hat{x}_1 & \hat{x}_2 \\ \hat{x}_1 & -\hat{x}_2 \end{bmatrix}.$$

Now, since

$$I_n - n^{-1}\mathbf{1}_n\mathbf{1}_n' = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \end{bmatrix}$$

the matrix $\mathbf{M}$ as defined in Theorem 1 is $\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$. Define $\gamma = 2\theta\hat{x}_1\hat{x}_2 b^{-2}N^{-1}$. The eigenvalues of $\mathbf{MJ'J}$ solve the equation $\left|\mathbf{MJ'J} - \vartheta I\right| = 0$ which implies that $\vartheta(\vartheta - \gamma^2 N\hat{x}_1 - \gamma^2 N\hat{x}_2) = 0$. After solving this equation the largest eigenvalue is $\vartheta_1 = N^2\gamma^2$. Hence, the asymptotic coefficient of reactivity is given by

$$\omega_0 = \gamma N = 2\theta\left(\frac{\hat{x}_1}{b}\right)\left(\frac{\hat{x}_2}{b}\right). \tag{4.9}$$

It can be seen that for networks with parallel routes, when the functions giving the route choice probabilities have an explicit form given by logit formula, the logit dispersion parameter θ and coefficient of reactivity are proportional to one another.

The question that can be raised is that, why can't the logit parameter be used to measure the reactivity of the system? In response to this question, it is worth mentioning that firstly, unlike θ, the coefficient of reactivity is unitless. If θ doubled but costs rescaled to be halved, then the coefficient of reactivity remains unchanged. Secondly, the coefficient of reactivity is not restricted to logit model and can be applied for any route choice model such as probit, etc.

It should be pointed out in most cases we cannot get an analytic expression for the coefficient of reactivity, but instead must compute the coefficient numerically (including getting a numerical approximation to the Jacobian matrix, for example).

4.3.1 Simulation Results For Asymptotic Coefficient of Reactivity

We continue to work with the network from Example 4.3.1 in order to investigate the relation between the value of the coefficient of reactivity and the impact of network disruption on the day-to-day flow evolution through the simulation of flows. To do so we maintain constant parameters. We set route 1 to be twice as long as route 2 with $a_1 = 2$ and $a_2 = 1$. The capacity of both routes is directly linked to travel demand with $b = \frac{N}{4}$. We compute the SUE values by using the method of successive averages. To show there is no limitation for choosing the route choice probability model, for the simulation purposes, two different models (logit and probit) for selecting the route for each traveller are considered.

Figures 4.1, 4.2 and 4.3 illustrate how the coefficient of reactivity works based on one day memory and one day disruption when traveller route choice probability is expressed via a logit model. Following on from Example 4.3.1, in this case we can adjust the coefficient of reactivity by changing the logit parameter which we do by choosing $\theta = 0.002$, $\theta = 0.112$ and $\theta = 0.13$ respectively for simulation of traffic flow on route 1 over a sequence of 100 days. We would like to understand when the cost on a particular route increases significantly, how the system would react. To see this, we allow the system to run normally for a while and then we assume that an unusual event (road closure) happens on day 41 for route 1 and the road reopens on day 42. This unusual event has been modelled by multiplying a very large value to the cost function for route 1 on that day.

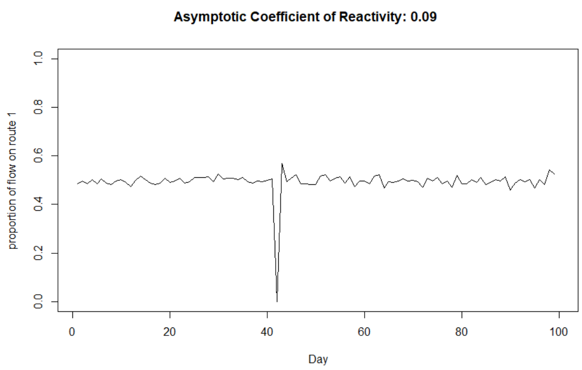


Figure 4.1: Route 1 flows simulated over 100 days as a fraction of total demand with coefficient of reactivity 0.09.

Figure 4.1 shows the unforeseen occurrence only had an affect on travellers route choice behaviour for that day and the system quickly returns to normal the next day. This happens because the small value of coefficient of reactivity shows the system is not reactive.

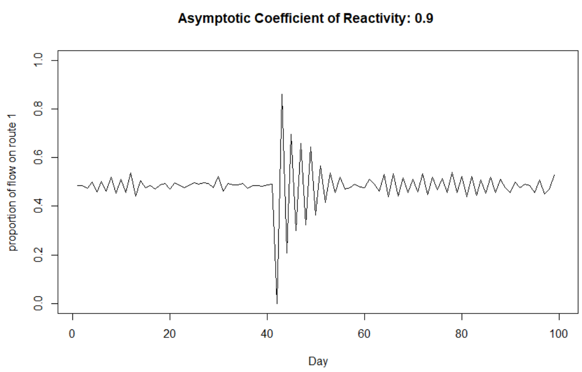


Figure 4.2: Route 1 flows simulated over 100 days as a fraction of total demand with coefficient of reactivity 0.9.

The next two figures, display the system with higher values of coefficient of reactivity. As shown in Figure 4.2 it takes around 20 days for the system to resume to usual flow patterns after the shock of a road closure. It is clear that the variation of flows between days 40 and 60 would result from drastic changes of travellers' daily route choices. Figure 4.3 depicts travellers' responses to the unusual flow observed on day 41 when the coefficient of reactivity is greater than 1. This leads to more drastic changes of travellers behaviour, with more switching between route 1 and route 2. Based on the definition of ω_0 this is not an unexpected behaviour, since the value more than one for the fraction indicates that the expected flow pattern for next day is far from the asymptotic stationary mean. This implies that oscillation of traffic flow remains for long period and wouldn't stabilise when ω_0 is larger than one.

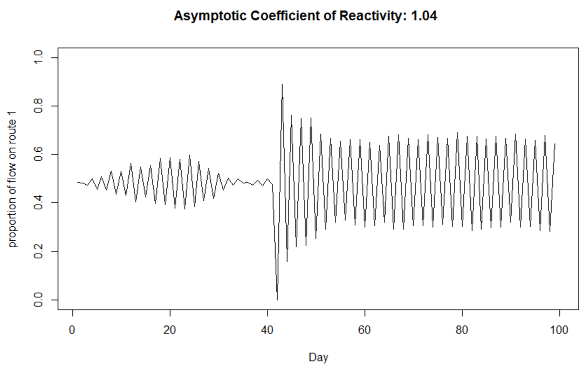


Figure 4.3: Route 1 flows simulated over 100 days as a fraction of total demand with coefficient of reactivity 1.04.

We now explore the use of the coefficient of reactivity when a probit route choice model is employed. All the link cost parameters remain unchanged. Note that coefficient of reactivity does not have a closed form, so all the Jacobians need to be computed numerically. By adjusting the tuning parameter in the probit model (describing the variance of the link costs) we set the coefficient of reactivity to the same values (approximately) as for the logit model example. The results are pretty similar to those from the logit model. As can be seen for the plot with a very low value of coefficient reactivity of 0.09 (Figure 4.4) the system does not react at all for this disruption and the system immediately returns to normal flow behaviour.

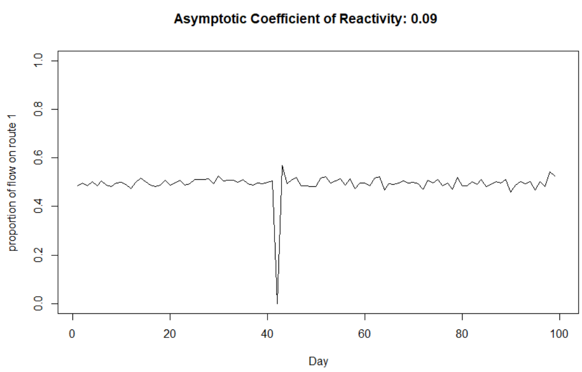


Figure 4.4: Route 1 flows simulated over 100 days as a fraction of total demand based on probit model with coefficient of reactivity 0.09.

For the second plot, Figure 4.5, with a bit high coefficient of reactivity of 0.9 there is reaction for the day after disruption and the proportion of travellers using route 1 goes up to around 0.9. Also it takes around 20 days until the system returns to normal flow behaviour. Finally, for Figure 4.6 with the value of coefficient of reactivity greater than one i.e.(1.04) the system become more reactive. There is a severe change for the days after disruption and the system never returns back to normal within the period of 100 days.

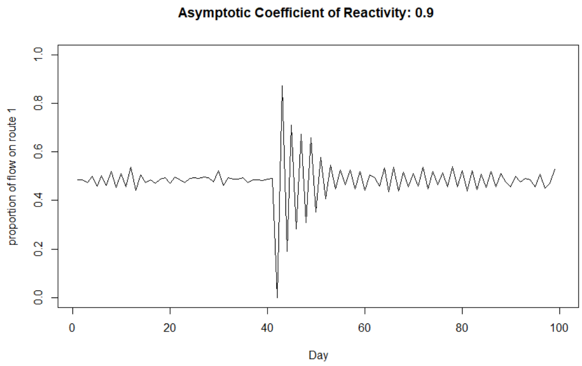


Figure 4.5: Route 1 flows simulated over 100 days as a fraction of total demand based on probit model with coefficient of reactivity 0.9.

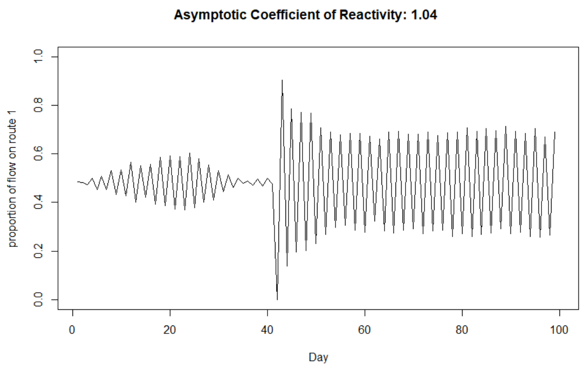


Figure 4.6: Route 1 flows simulated over 100 days as a fraction of total demand based on probit model with coefficient of reactivity 1.04.

This example illustrates the coefficient of reactivity can be applied to any kind of day-to-day Markov model of the type introduced earlier.

4.4 Simulation results

The general purpose here is to examine how well approximation methods work, and how that relates to coefficient of reactivity. For a variety of networks we will perform multiple simulations from the true stochastic model. We will use this to compute estimates of the true dynamics of the mean flow, and will also compute prediction intervals for flows. We will then compute corresponding quantities using approximation models (but not prediction intervals for deterministic approximation). Coefficient of reactivity will be computed for each experiment. Results will be computed for various values of coefficient of reactivity, adjusted by controlling the logit parameter.

Results will be presented through plots for selected networks. We will plot the mean of Davis and Nihan (1993) approximation method, the mean of Cantarella and Cascetta (1995) deterministic model, the true mean of Watling and Hazelton (2018) linear approximation and the true mean flows resulting from the simulation. Also following Watling and Hazelton (2018) we used 'mean $\pm$ 1.96 standard deviation' to find the 95% prediction intervals for stochastic approximation methods. All results from the true model are based on 1000 simulations of the process.

For each network and choice of coefficient of reactivity, different plots with three different initial flows are provided. In each Figure, for the top plot the starting point is at SUE, the initial point for the middle one is fairly close to SUE, and quite far from SUE for the plot in the bottom. We have used R version 3.4.4 to obtain the numerical results running under Windows 10 on a computer with 8 GB memory.

All the results are based on a Markov process with memory length 1 and choices based solely on the flow from the previous day (i.e. with $\lambda = 1$ from equation 2.19). The reason for these choices are (i) Hazelton's (2002) coefficient of reactivity is not designed work with models including a separate utility term; and (ii) because flows from day $t-2$ and earlier are set at SUE, the usefulness of the coefficient of reactivity for models with memory length 2 or more is uncertian. We will return to this issue in Chapter 5, where we look at the extensions of Hazelton's coefficient of reactivity.

For all experiments the gray dotted central lines is the mean of Davis and Nihan (1993) approximation method. The green dotted lines is the Cantarella and Cascetta (1995) deterministic model. The jagged light purple lines show the realized time plots of traffic flows from 10 simulations of the model. The unbroken lines depict the true mean flow (red line) and black lines show the true mean of Watling and Hazelton (2018) linear approximation. The outer dotted lines correspond to 95% prediction intervals as prediction for flows (the same as mean, gray one for Davis and Nihan (1993) approximation model, black one for Watling and Hazelton (2018) approximation model and the red one for flows generated by simulation).

Experiment 4.4.1. The first network is just a continuation of a particular version of Example 4.3.1 where a simple two route network serves a single OD pair with demand, $N = 100$ travellers. Since links and routes are identical in this case, we can express the link costs in terms of route flows. The link cost functions are

$$C_1(\boldsymbol{x}) = 2 + x_1^2$$

and

$$C_2(\boldsymbol{x}) = 1 + x_2^2.$$

It is assumed that users' route choice probability follows the logit model, and is given as,

$$p_r(\boldsymbol{x}) = \frac{\exp\{-\theta C_r(x_r)\}}{\sum_r \exp\{-\theta C_r(x_r)\}}, \quad r = 1, 2.$$

Since the flow on route 2 simply complement flow on route 1, just the evolution of flow on route 1 is displayed. It can be seen that for a low value of coefficient of reactivity (Figures 4.7 and 4.8), both approximation methods work really well for different initial flow patterns. Notice the differences in behaviour of the system over the initial transient periods for the cases where the processes do not start at SUE. For the bottom plots the initial flow is $\boldsymbol{x}^0 = (35, 65)^T$, which is away from $\boldsymbol{x}^* = (49.94, 50.06)^T$. More changes in mean behaviour can be observed in the first few days for the plot with higher value of coefficient of reactivity (Figure 4.8).

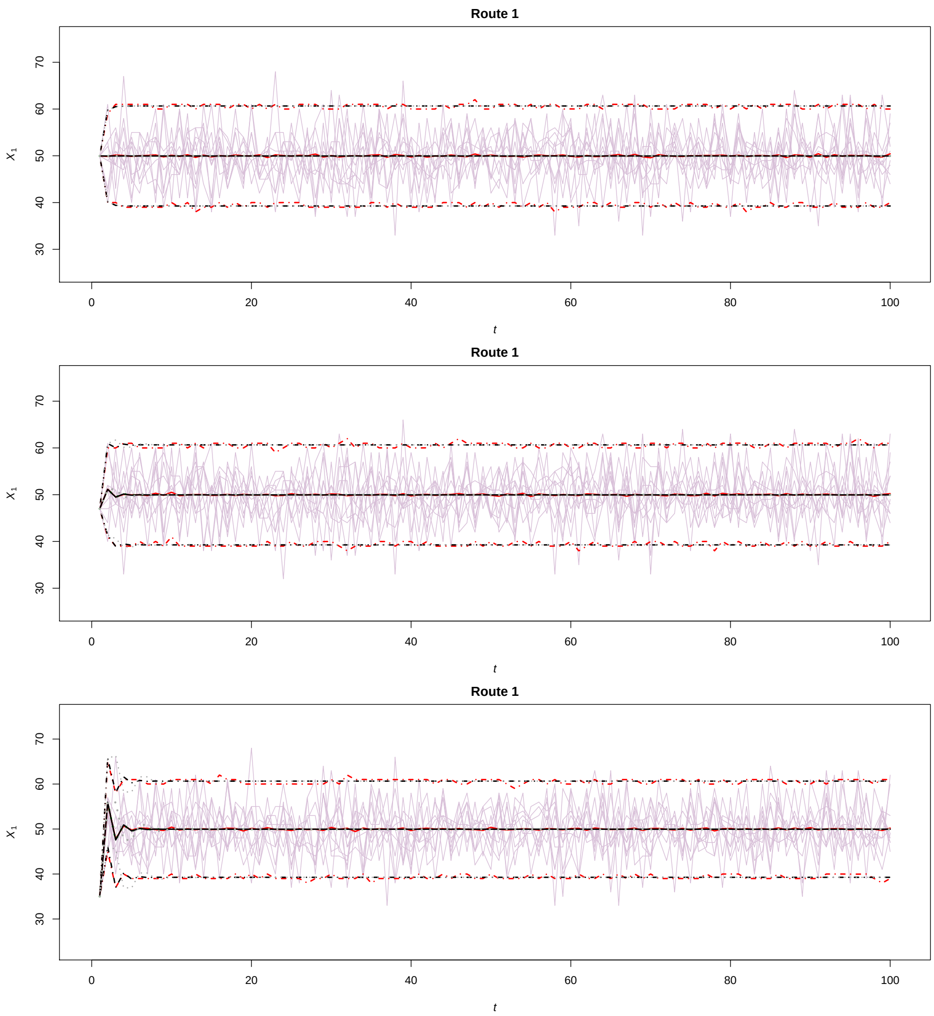


Figure 4.7: Flows on route1 in simple two-route example with coefficient of reactivity 0.48.

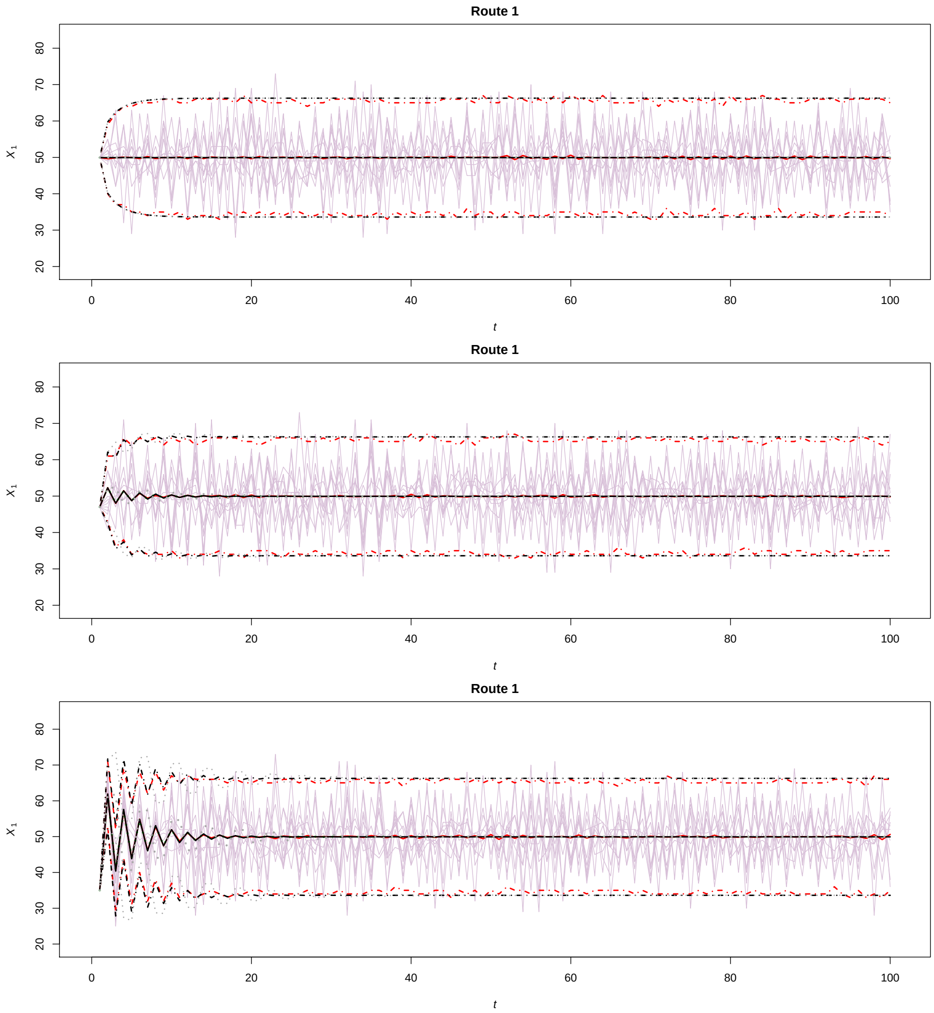


Figure 4.8: Flows on route1 in simple two-route example with coefficient of reactivity 0.8.

When the value of the coefficient of reactivity is increased to 1, we see the approximation of standard errors begin to break down. See the top plot in Figure 4.9, where the approximate prediction intervals (based on those standard errors) diverge. The mean approximations remain stable. However, changing the initial flow fairly close to SUE (the middle plot) to $\boldsymbol{x}^0 = (47, 53)^T$, leads to unstable mean flow patterns for both stochastic approximation methods but the deterministic approximation remains in steady state. As the bottom plot shows, when we start from a flow pattern which is far from SUE, all methods fail completely.

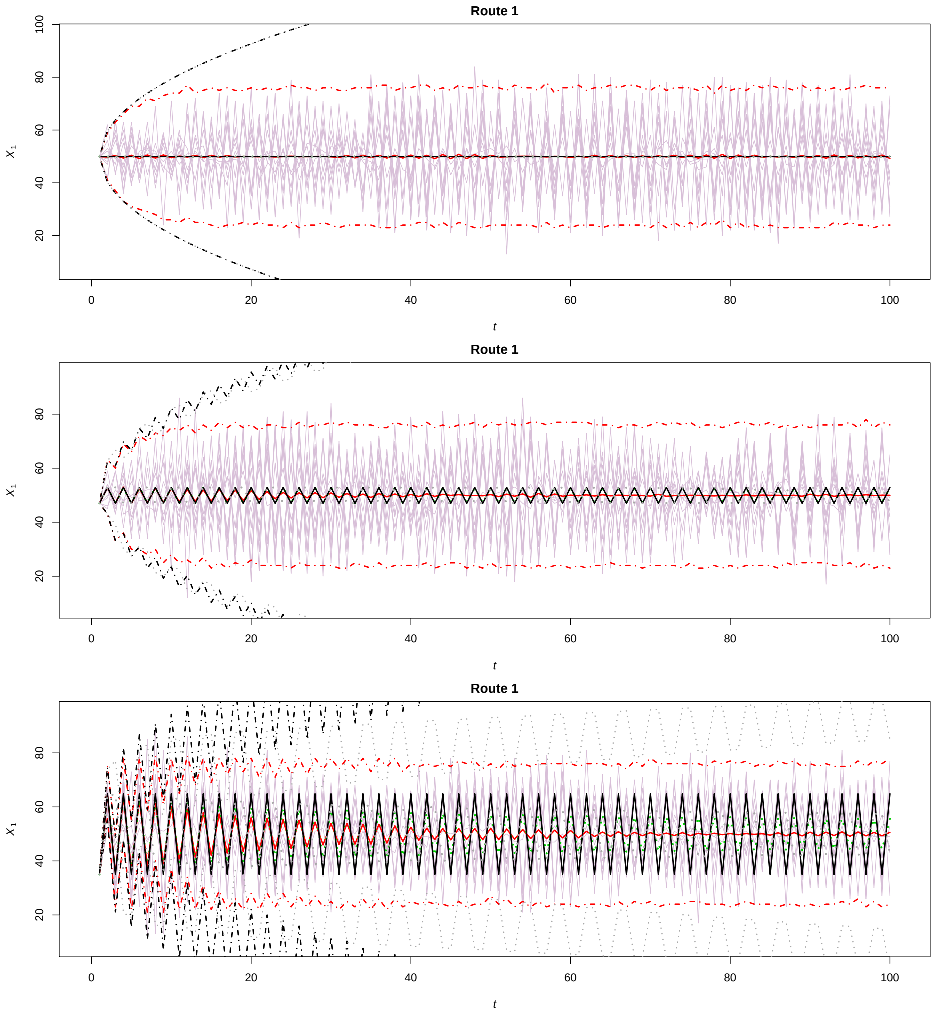


Figure 4.9: Flows on route1 in simple two-route example with coefficient of reactivity 1.

For a higher value of coefficient of reactivity 1.10 (Figure 4.10) the failure in 95% confidence interval in top plot happens sooner compared to Figures 4.9. For the middle plot, along with 95% confidence interval the mean of both stochastic approximation methods and the deterministic model start to break down. For all approximation methods, the size of fluctuations increases when the period of days for simulation rise. Similarly, with initial flow pattern $\boldsymbol{x}^0 = (35, 65)^T$ we see the failure of all methods with more bouncing from the beginning.

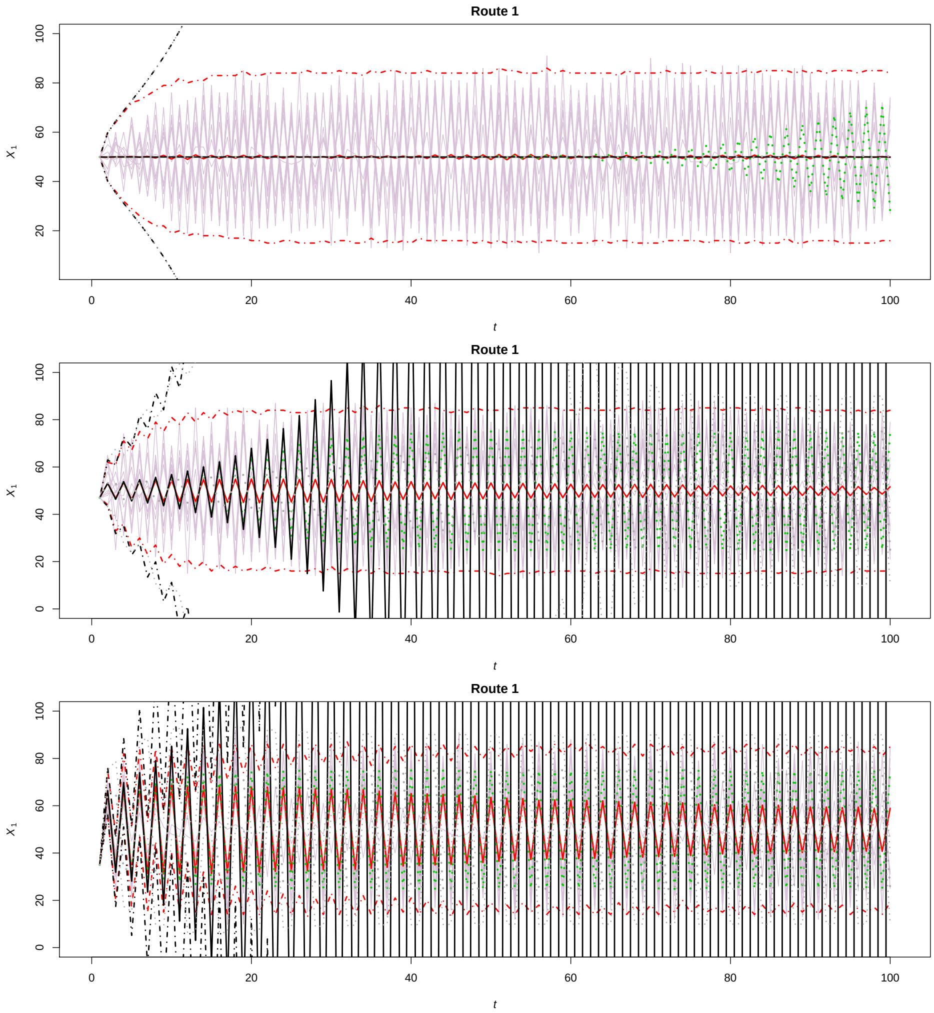


Figure 4.10: Flows on route1 in simple two-route example with coefficient of reactivity 1.10.

Experiment 4.4.2. The second illustrative example is a network with a single OD pair, a demand of $N = 40$ travellers, and three parallel routes with different cost functions. The cost functions are as follows:

$$c_1(\boldsymbol{x}) = 2 + 8x_1,$$

$$c_2(\boldsymbol{x}) = 3 + 10x_2^2,$$

$$c_3(\boldsymbol{x}) = 6 + 25x_3^2.$$

Again logit model for route choice probabilities has been applied. All simulations has been generated over a period of 100 days. The SUE flow pattern with two decimal places is $\boldsymbol{x}^* = (15.15, 16.61, 8.24)^T$. We can see similar behaviour to 4.4.1 when the coefficient of reactivity is low (Figures 4.11 and 4.12). In the second and third panels of Figure 4.12 as the initial flow pattern is “unusual” (not typical of stationary distribution), the first 5-10 days are all about approximating the system during a transient phase.

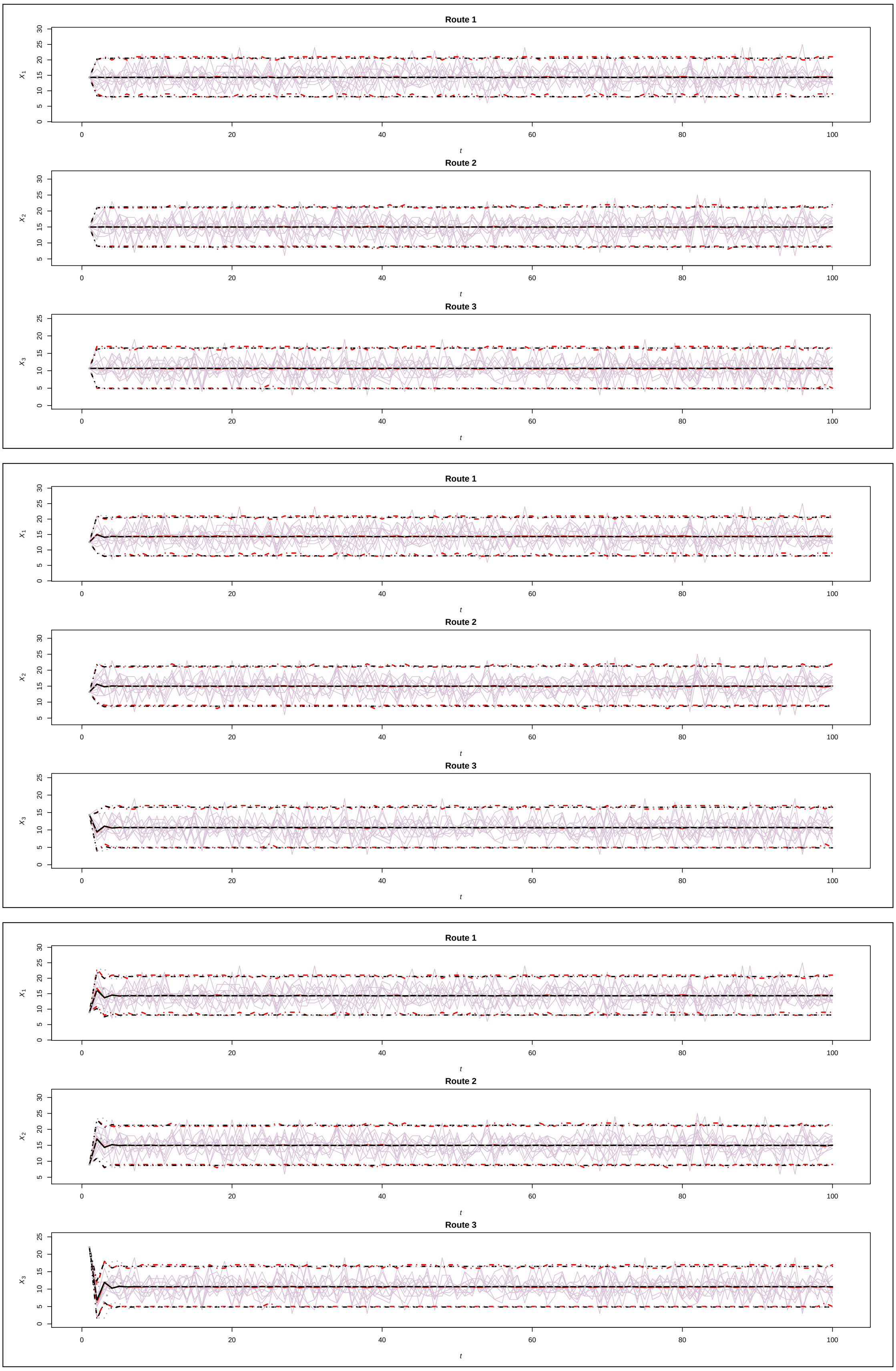


Figure 4.11: Route flows in three-route network example with coefficient of reactivity 0.34.

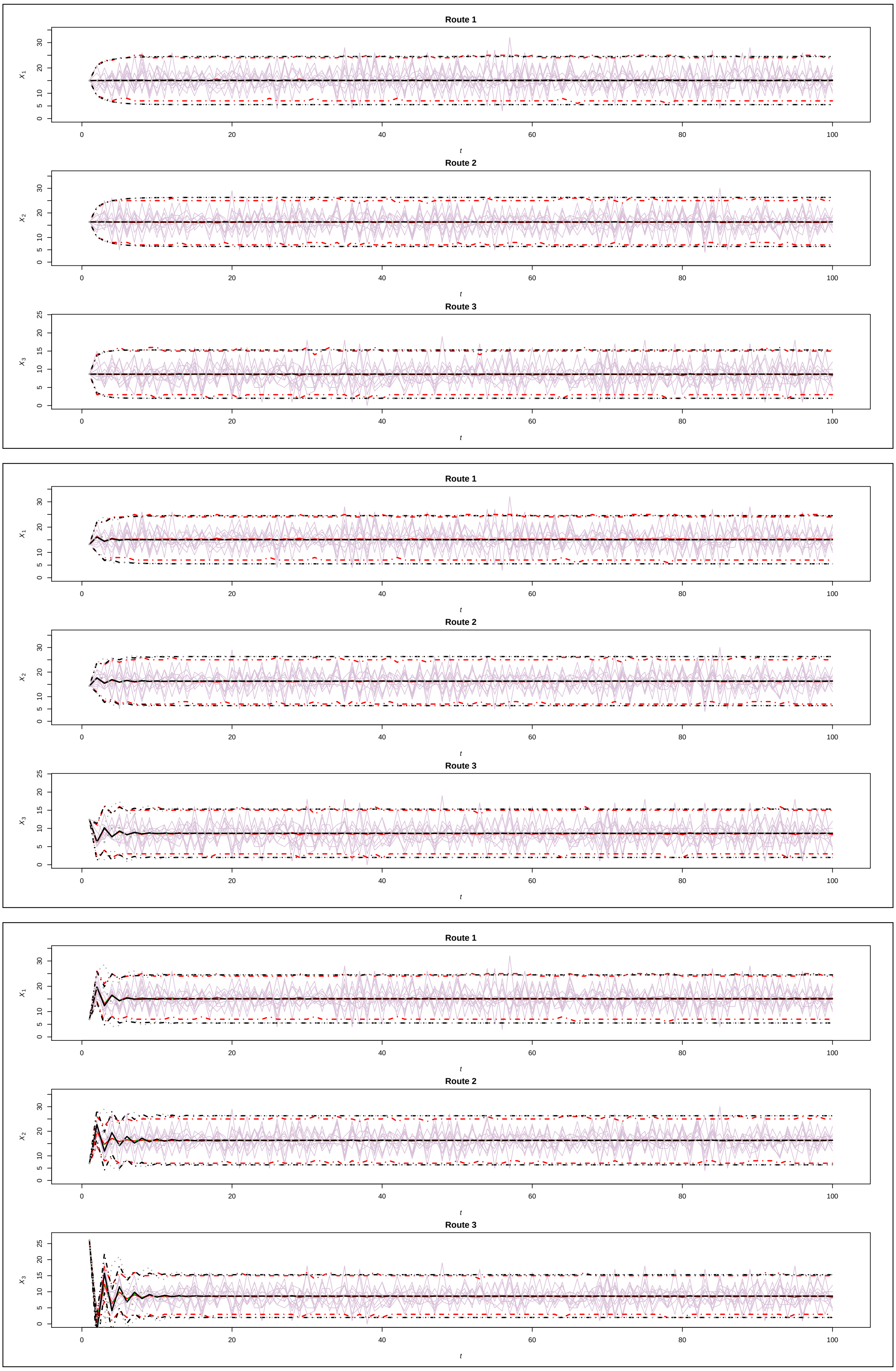


Figure 4.12: Route flows in three-route network example with coefficient of reactivity 0.8.

Figure 4.13 shows that the approximations become worse when the value of coefficient of reactivity increases. This can be seen in all panels as the limits of the prediction intervals diverge for the first two routes even with starting point at SUE.

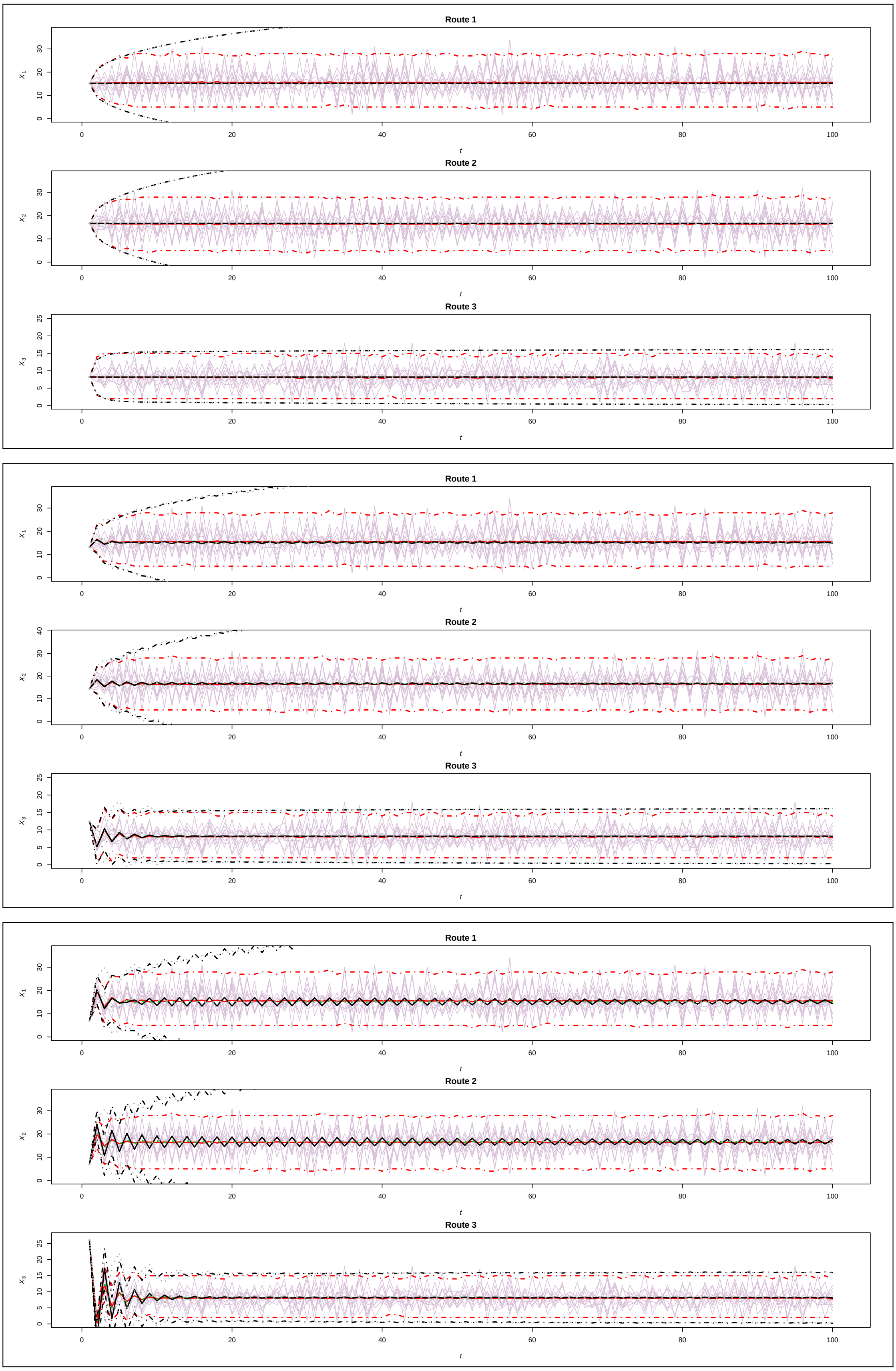


Figure 4.13: Route flows in three-route network example with coefficient of reactivity 0.99.

The transient phase can be seen in route 3 for the bottom panel around 10 days. Also, for this route all approximation methods stabilised at the stationary states, whereas for the other routes in that panel Watling and Hazelton's (2018) mean approximation fails to properly settle down, other methods are fine.

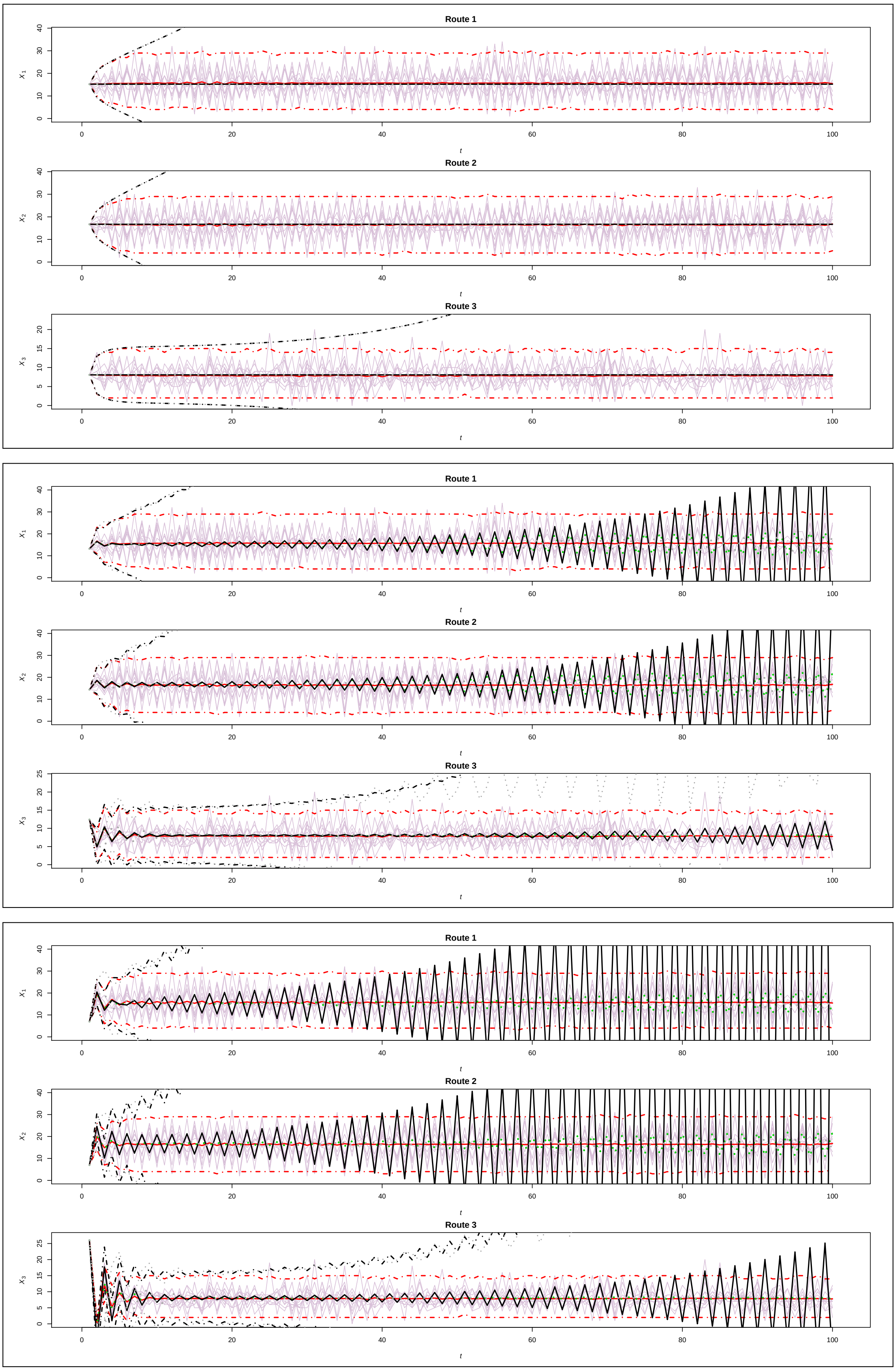


Figure 4.14: Route flows in three-route network example with coefficient of reactivity 1.04.

Finally, for more reactive system with coefficient of reactivity 1.04 (Figure 4.14), the 95%

prediction interval breaks down after a few days for both approximation methods and it does not matter what the initial flow is. Looking at route 3 when the starting point is not at SUE demonstrates the boundary of prediction interval for both approximation methods is divergent but in contrast to the linear approximation method, the approximate mean process for nonlinear method along with deterministic model remain steady over running days. For the first two routes, all approximation methods break down with more drastic changes for the approximate mean process of the linear approximation model. The earlier fluctuations happened when the initial flow is $\boldsymbol{x}^0 = (7, 7, 26)^T$ which is far away from SUE flow pattern.

Experiment 4.4.3. This example uses a network described by Han et al. (2018) with six links, five nodes and one OD pair from origin 1 to destination 5. The total OD demand is 2000. Three routes are available between the OD pair. Route 1 composed of links 1 and 4, route 2 composed of links 2,5 and 6 and route 3 only has link 3. For this experiment we consider the BPR link cost function from 2.1. The parameters used are given in Table 4.1.

Link (l)	c_l^0	α_l	b_l	φ_l
1	2	0.15	500	4
2	1	0.15	800	4
3	5	0.15	800	4
4	2	0.15	800	4
5	1	0.15	500	4
6	1	0.15	500	4

Table 4.1: Parameters of link cost functions for test network.

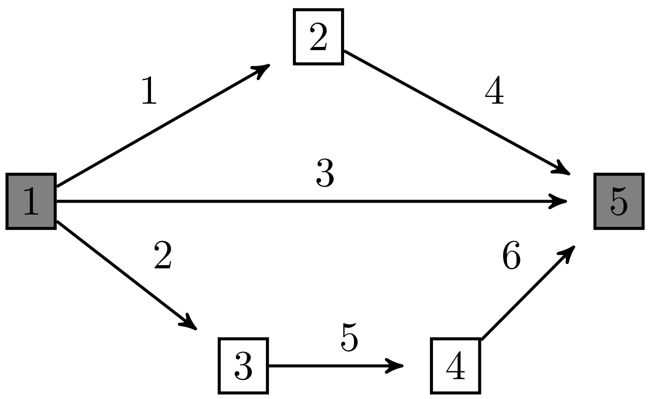


Figure 4.15: One OD pair with six links network.

As illustrated in the top panel of Figures 4.16 and 4.17 for this network all processes have settled down from starting point to their stationary distribution when the value of coefficient of reactivity is less than 1 and the initial flow is at SUE. The next two panels in Figures 4.16 and 4.17 show different transient behaviours of travellers. Strong variation is observed when the initial flow patters are far from SUE. It takes a longer duration to approximate the system within a transient phase for the higher coefficient of reactivity. However, it should be noted that the approximation methods do an excellent job of approximating the stochastic behaviour, even when it is far from SUE.

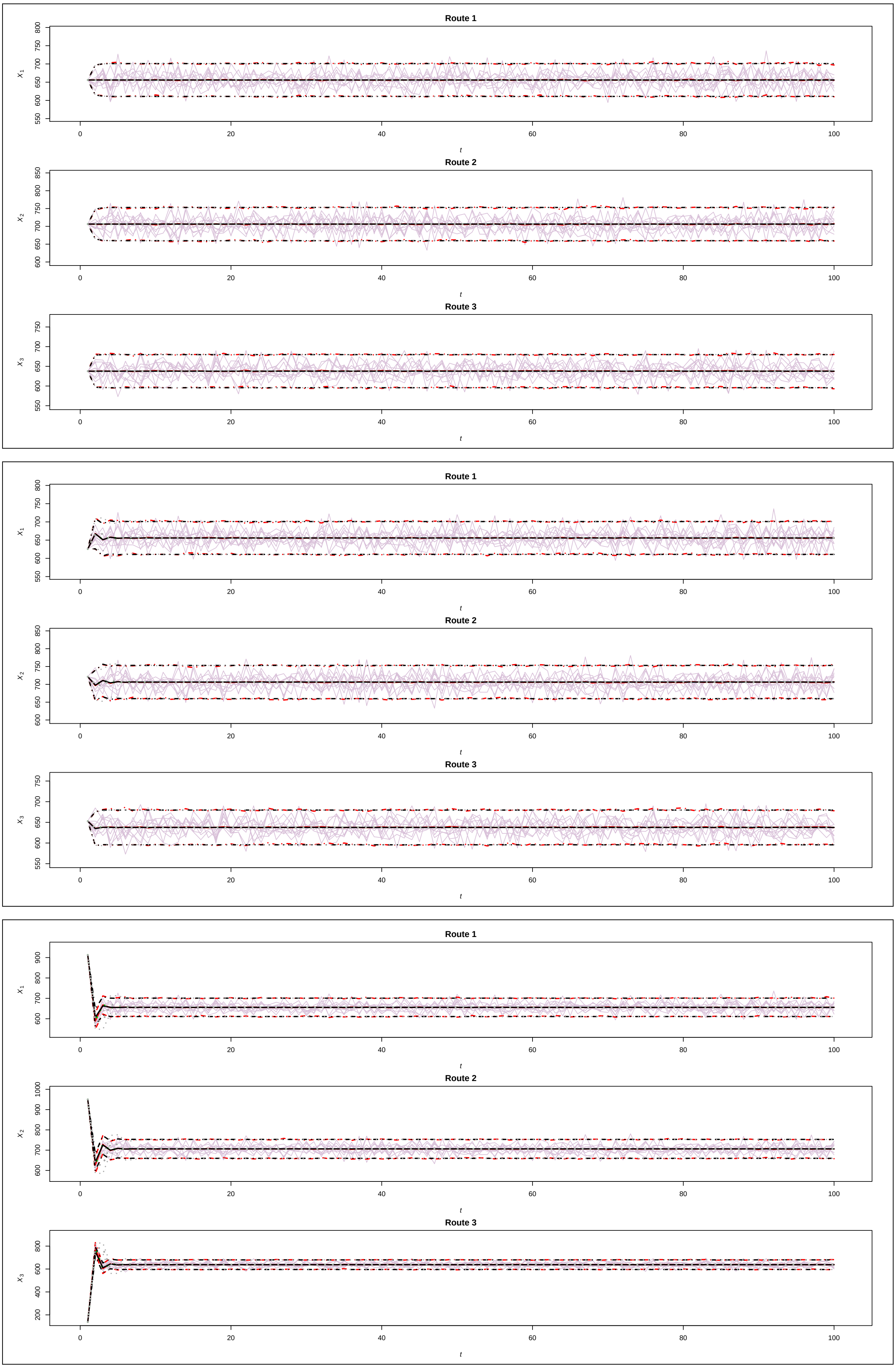


Figure 4.16: Route flows in six-link network example with coefficient of reactivity 0.46.

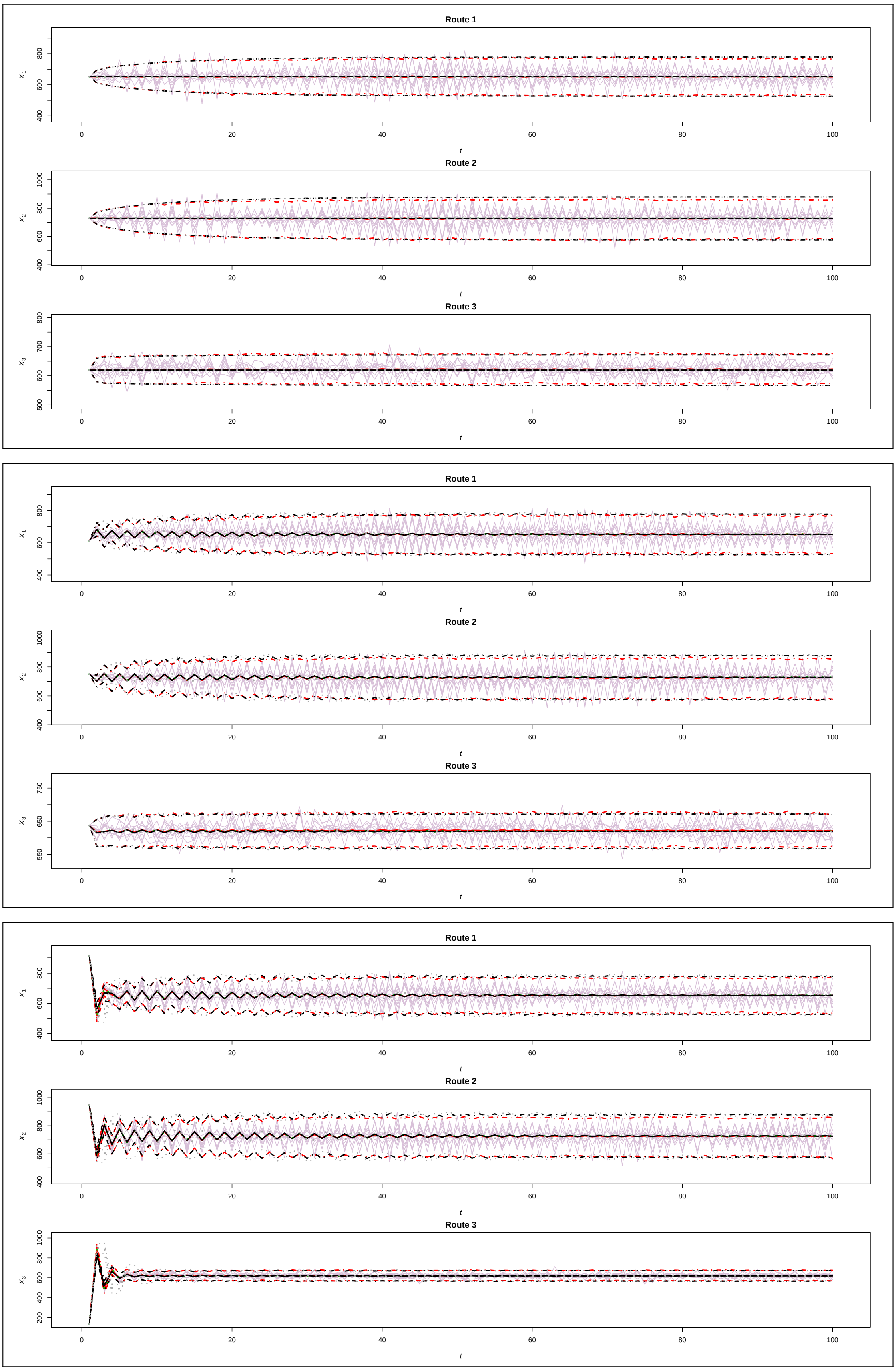


Figure 4.17: Route flows in six-link network example with coefficient of reactivity 0.96.

Here, even when the coefficient of reactivity is close to 1 ($\omega_0 = 0.96$) it seems the mean flows of all methods work well.

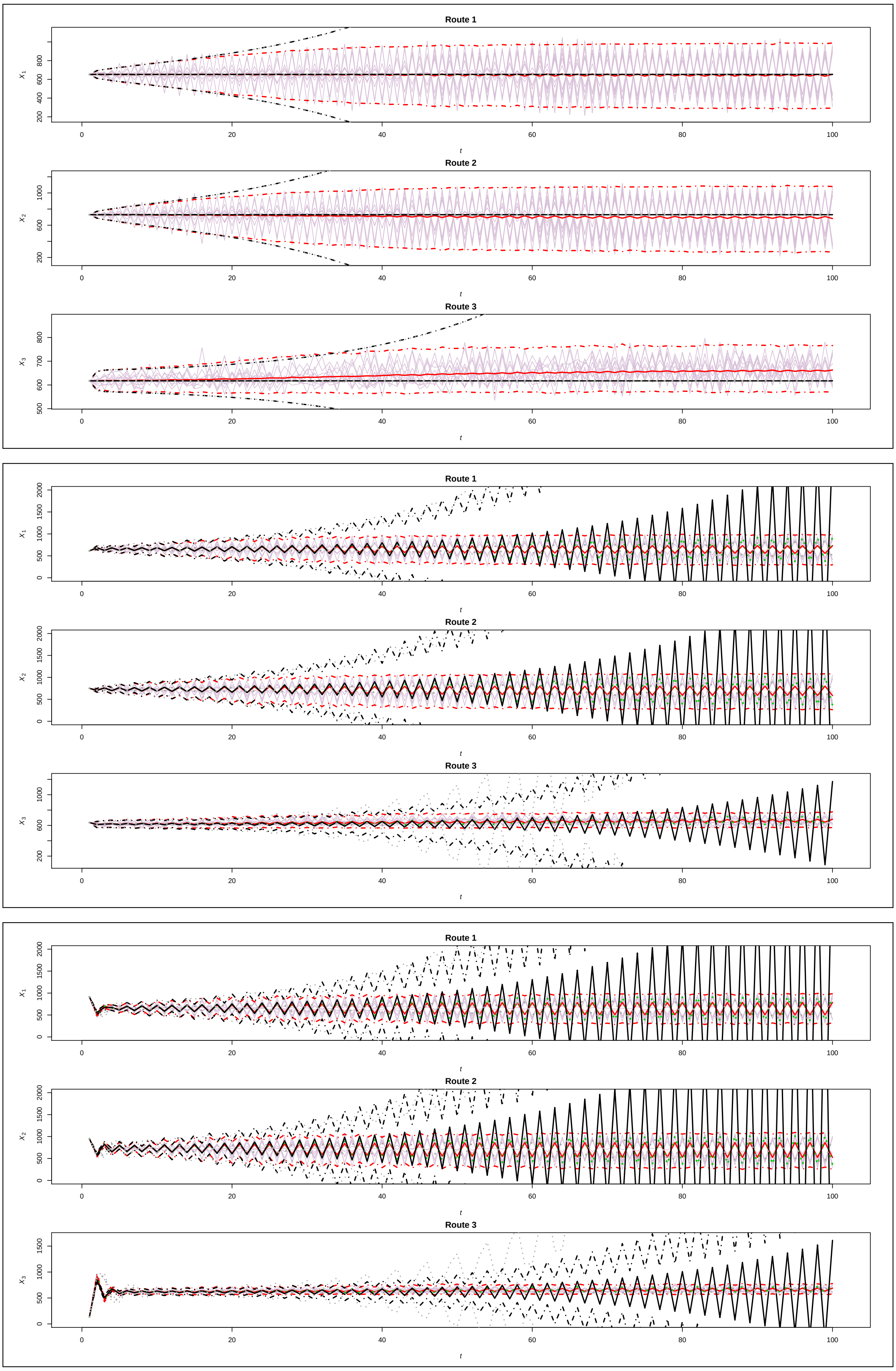


Figure 4.18: Route flows in six-link network example with coefficient of reactivity 1.05.

We expect to see the mean flow become unstable, as seen in Figure 4.18, when the value of the coefficient of reactivity is greater than 1. In such a case, when the initial flow is at SUE the 95% prediction intervals are out of the range while the mean flows of all methods work well. When the starting point is near the SUE, mild oscillations rapidly evolve into very strong ones for the mean of Watling and Hazelton's (2018) method. The deterministic and Davis and Nihan (1993) approximation methods (shown in the two lower panels of Figure 4.18) appear to be less affected by this phenomena. The further away from SUE the starting point, the more pronounced the onset of the extreme fluctuations in predicted flows.

Our experiments so far have involved systems with a single OD pair. We observed that 1 is something of a threshold value of the coefficient of reactivity. For larger values, the approximation methods begin to break down at least in 95% prediction interval for both approximation methods and for mean flows for Watling and Hazelton's (2018) method.

From the following experiments, we investigate the connection between aforementioned methods and the value of coefficient of reactivity for networks with multiple OD pairs.

Experiment 4.4.4. The example network shown in Figure 4.19 has two different OD pairs, seven links with six nodes. It is taken from Hazelton and Watling (2004). In this network nodes 1 and 3 are the origins and node 5 is a destination. Four different routes connecting these OD pairs consist of nodes $1-4-5$, $1-2-5$, $3-2-5$ and $3-6-5$ as routes 1 to 4 respectively. The link cost functions used for our network is

$$c_l(y_l) = 5 + 2.5\left(\frac{y_l}{N}\right)^2$$

where N is the (common) demand for each OD pair. For our simulation experiment travel demand for both OD pairs is assumed to be $N = 50$.

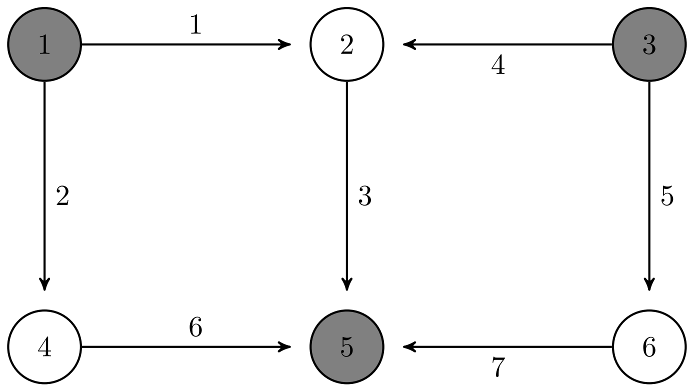


Figure 4.19: Illustrative network, with four routes and seven links.

Similar to the three preceding experiments, all approximation methods work well when the value of coefficient of reactivity is low as shown in Figure (4.20). Note that the approximations are effective during both transient and stationary period.

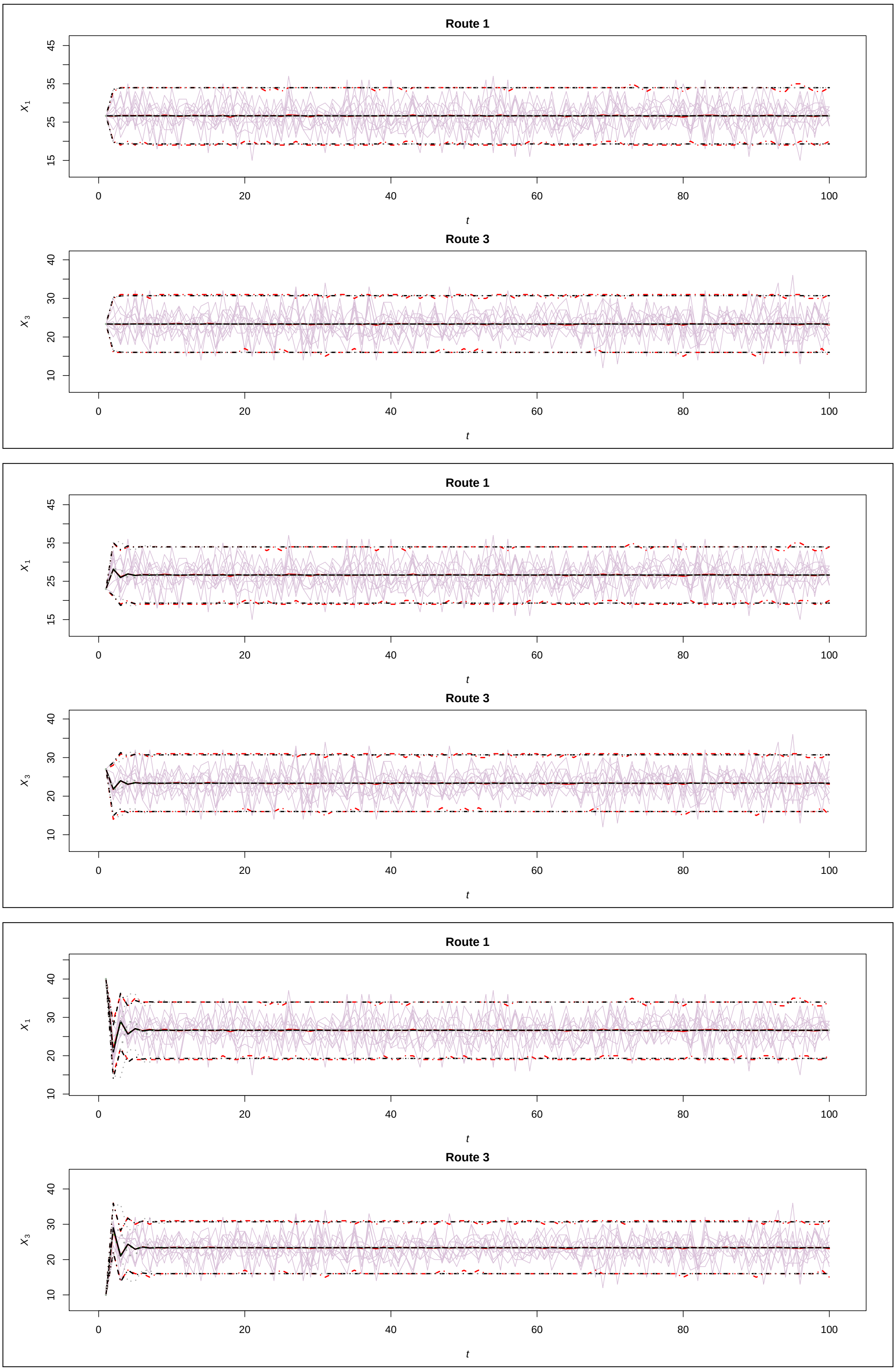


Figure 4.20: Seven links network example with two different OD pairs with coefficient of reactivity 0.42.

As the value of coefficient of reactivity gets close to 1 ($\omega_0 = 0.97$) Figure 4.21 demonstrates the 95% prediction intervals for both stochastic approximation methods are too wide even when the initial flow pattern is at SUE. That is because the flows are less variable than those approximations indicate. For the starting point close to SUE, besides the 95% prediction interval, the mean flows for both stochastic approximation methods remain unstable until about 80 days, but Cantarella and Cascetta's (1995) deterministic method converges to stationary state around day 60.

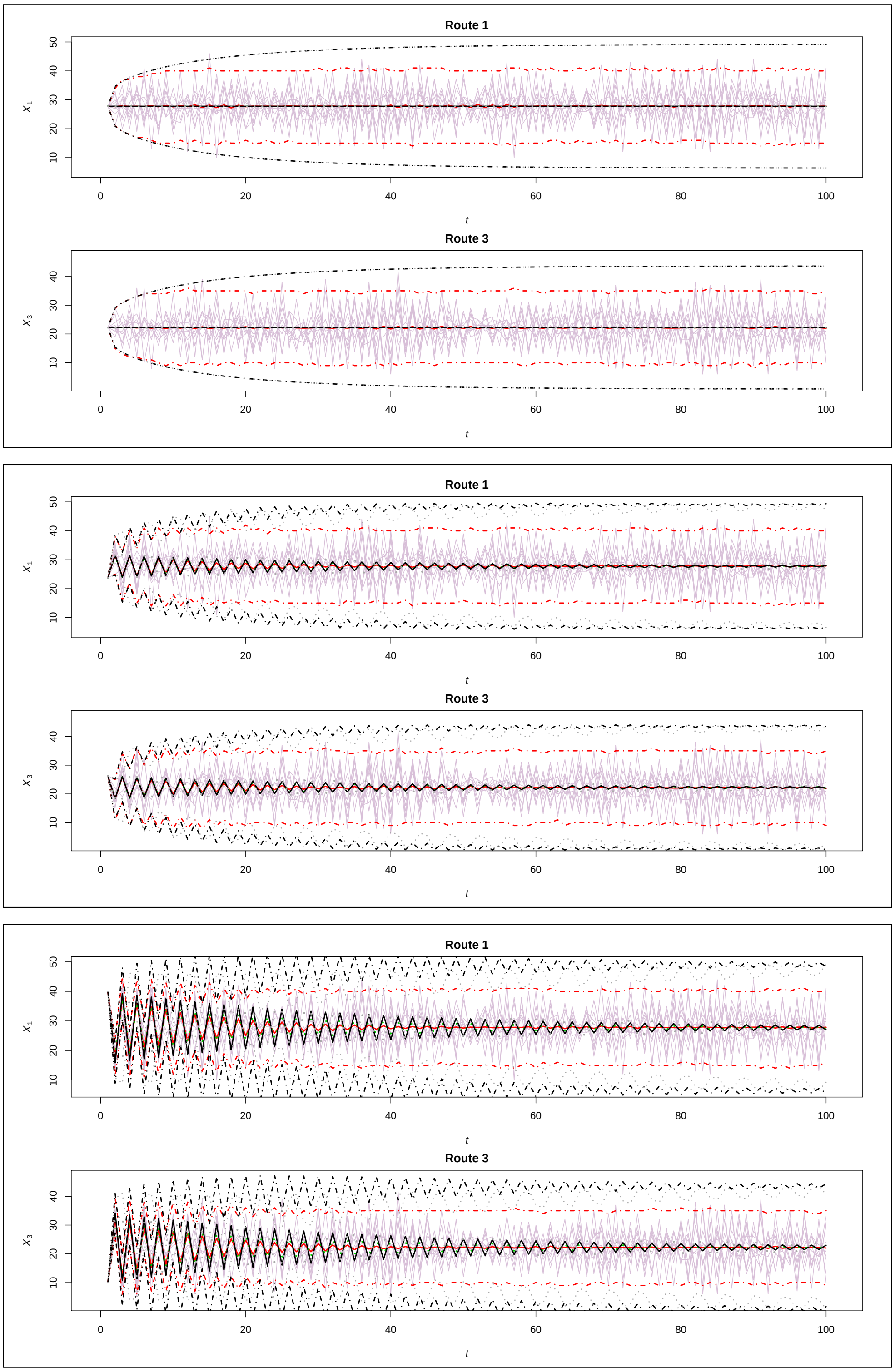


Figure 4.21: Seven links network example with two different OD pairs with coefficient of reactivity 0.97.

The situation is a bit different when the initial flow is far away from SUE. As displayed in the bottom panel of Figure 4.21 the mean flows drastically change over the first 40 days for all approximation methods and it takes a longer period to reach to the steady state.

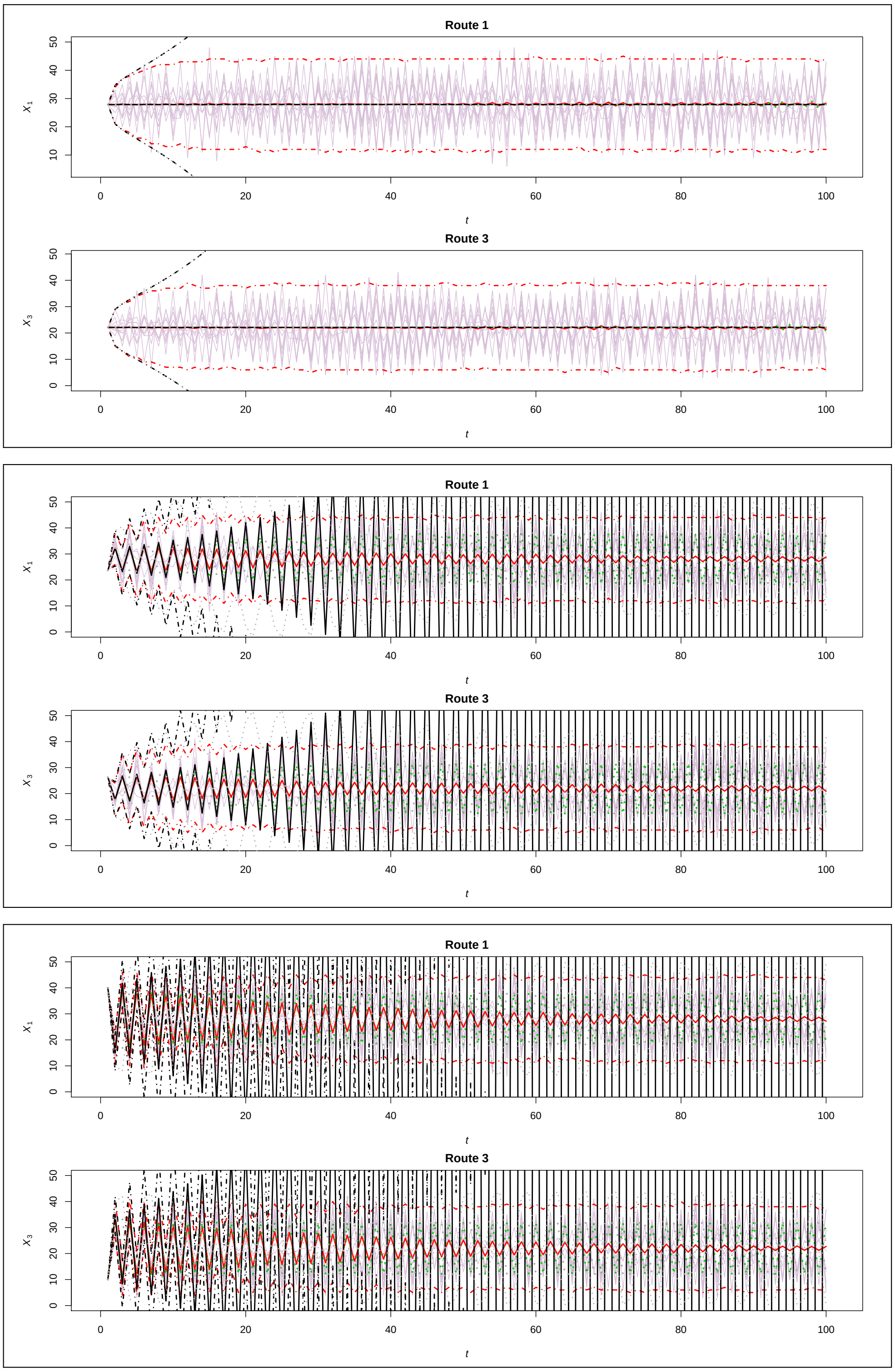


Figure 4.22: Seven links network example with two different OD pairs with coefficient of reactivity 1.06.

When the value of the coefficient of reactivity hits 1 Figure 4.22 shows the bounds of the 95% prediction interval diverge quickly even for the initial flow at SUE. As we expected, the failure of all approximation methods happens from the beginning and as time goes by the size of oscillations increase. As reflected in the third panel of Figure 4.22 more drastic changes in the mean flows of all approximation methods can be seen when the initial flow is distant from SUE.

Experiment 4.4.5. The network depicted in Figure 4.23 consists of six nodes, seven links and four different OD pairs. Node and link numbers are shown in this figure. The travel demand of O-D pairs (1,3), (1,4), (2,3) and (2,4) are 50, 10, 10, and 60, respectively. The network has six different routes which are denoted by link number sequence (1, 2-5-6, 2-5-7, 4-5-6, 4-5-7, and 3). The link cost function is of BPR type from 2.1, where $\alpha = 0.15$ and $\varphi = 4$. The link free flow cost c_l^0, respectively, is 10, 4, 12, 4, 5, 5 and 4. Also the link capacity b_l is assumed to be 50, 40, 40, 25, 60, 25 and 25.

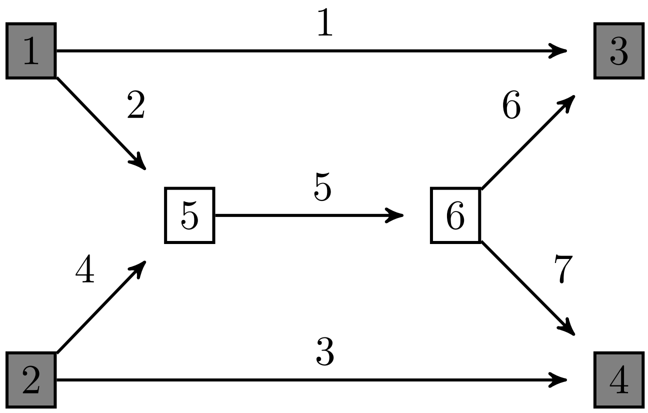


Figure 4.23: Four OD pairs with six routes network.

Two different routes have been selected for presentation of results: route 2, which is a combination of three different links and route 6. Again as illustrated in Figure 4.24 all methods work really well for a low value of coefficient of reactivity $\omega_0 = 0.53$. Of course when the starting state is far from SUE a first few days display the transient behaviour of the system. Similarly, when the value of coefficient of reactivity becomes very close to 1 ($\omega_0 = 0.98$) Figure 4.25 shows all processes have settled down to their stationary distribution from starting point when the initial flow is at SUE. Even when the initial flow is far from SUE it can be seen that after 20 days all methods stabilised.

Contrary to the experiments we have seen so far, the failure of approximation methods did not happen for this network when the value of coefficient of reactivity is bigger than 1 ($\omega_0 = 1.06$). This is demonstrated in Figure 4.26. As we expected to see the unstable mean flows when the coefficient of reactivity is greater than 1, we decided to increase the value of coefficient of reactivity. In such a case, Figure 4.27 shows when the value of coefficient of reactivity reaches to 1.10, this leads to a break down for the lower limit of the prediction interval at least for route 6 regardless of the initial flow pattern.

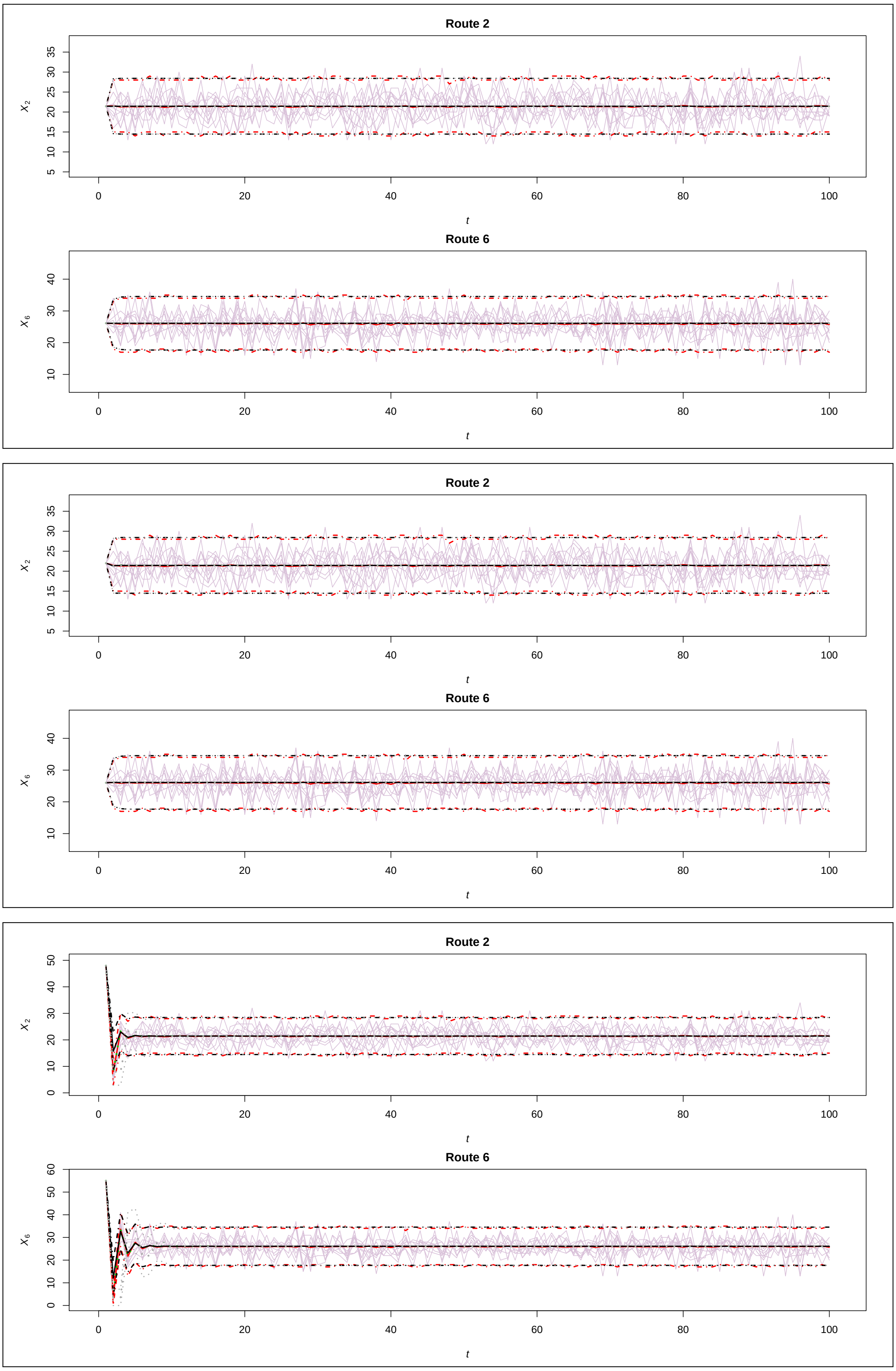


Figure 4.24: Route flows in four OD pair and six route network example with coefficient of reactivity 0.53.

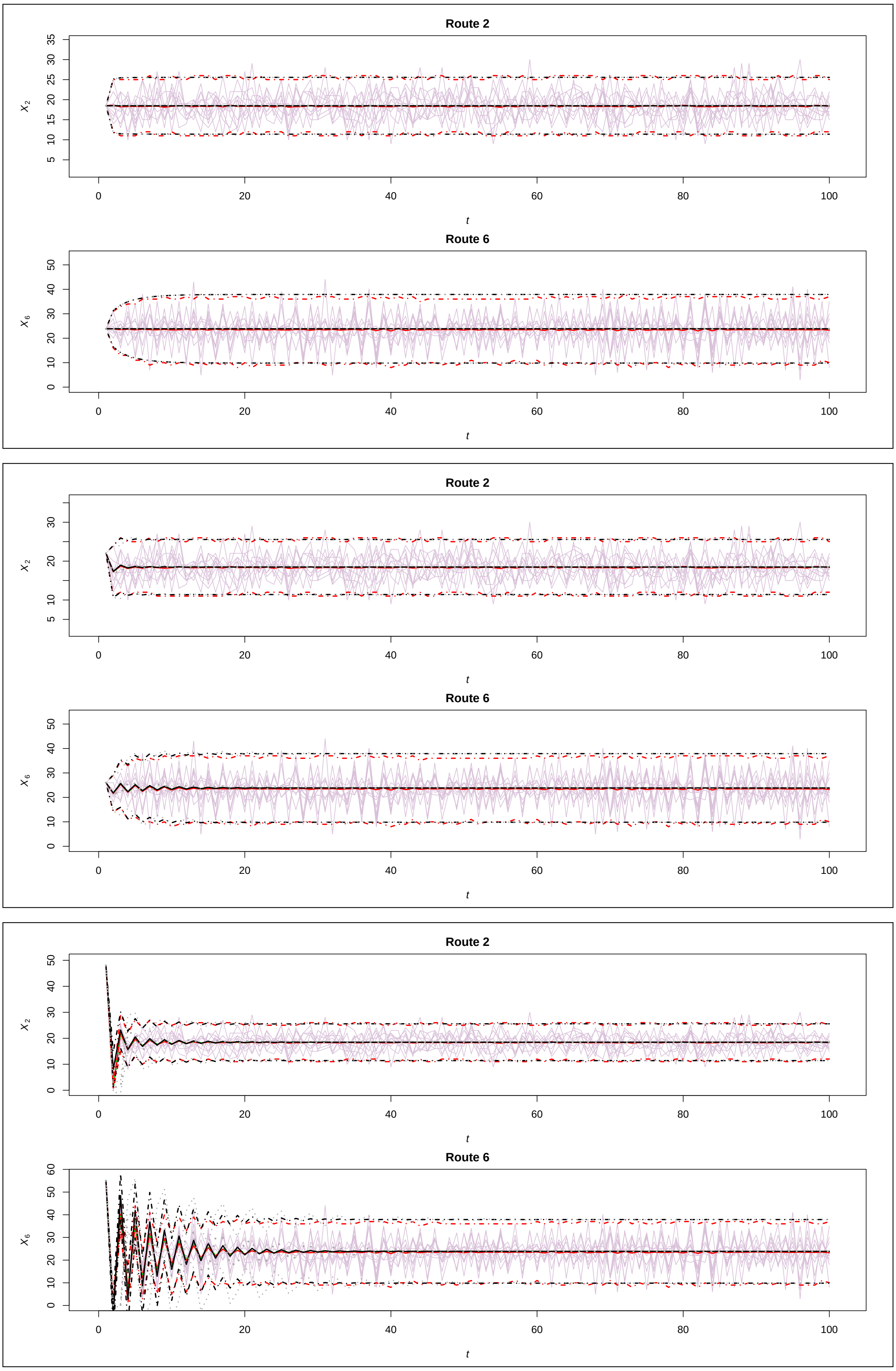


Figure 4.25: Route flows in four OD pair and six route network example with coefficient of reactivity 0.98.

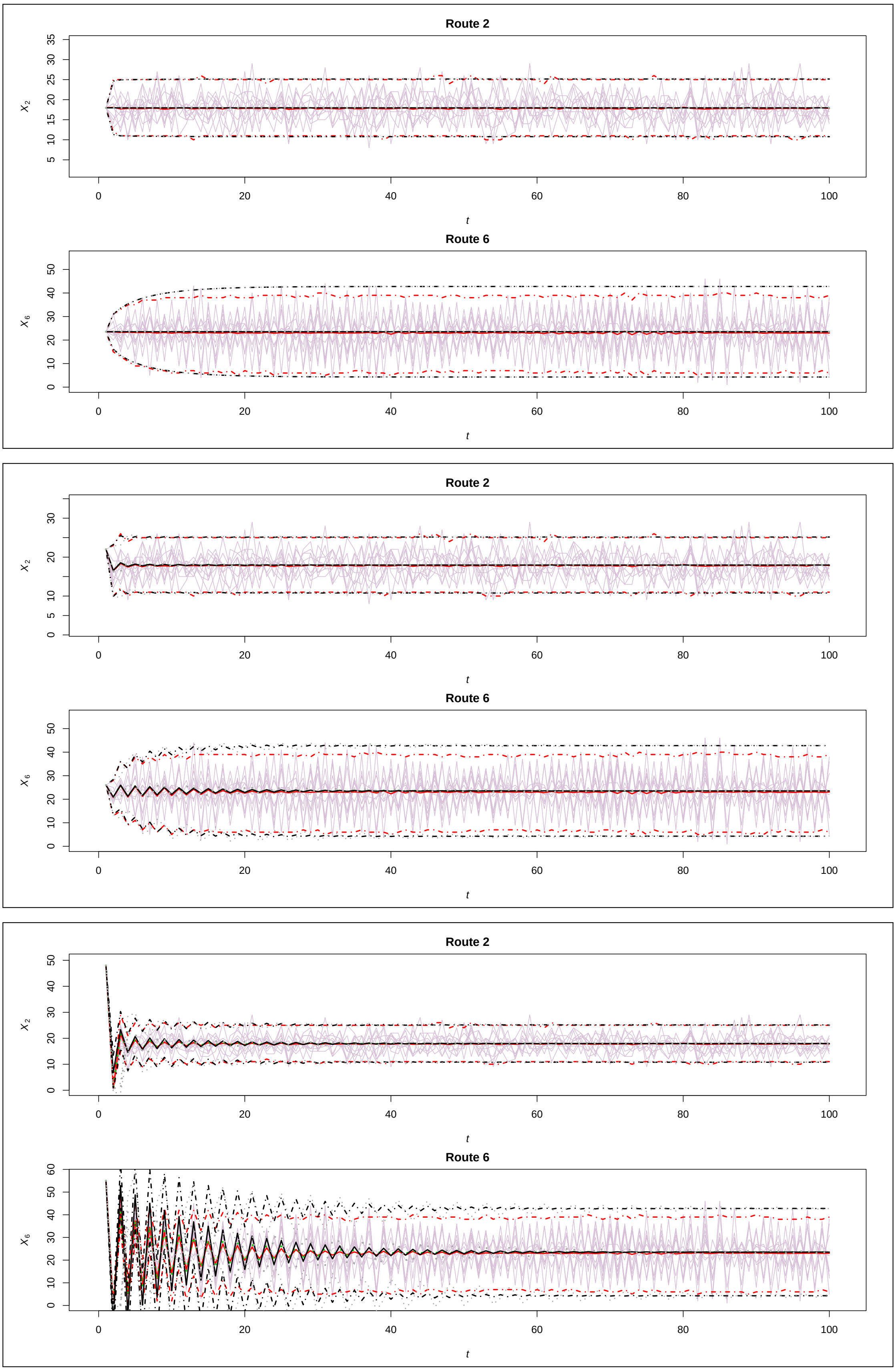


Figure 4.26: Route flows in four OD pair and six route network example with coefficient of reactivity 1.06.

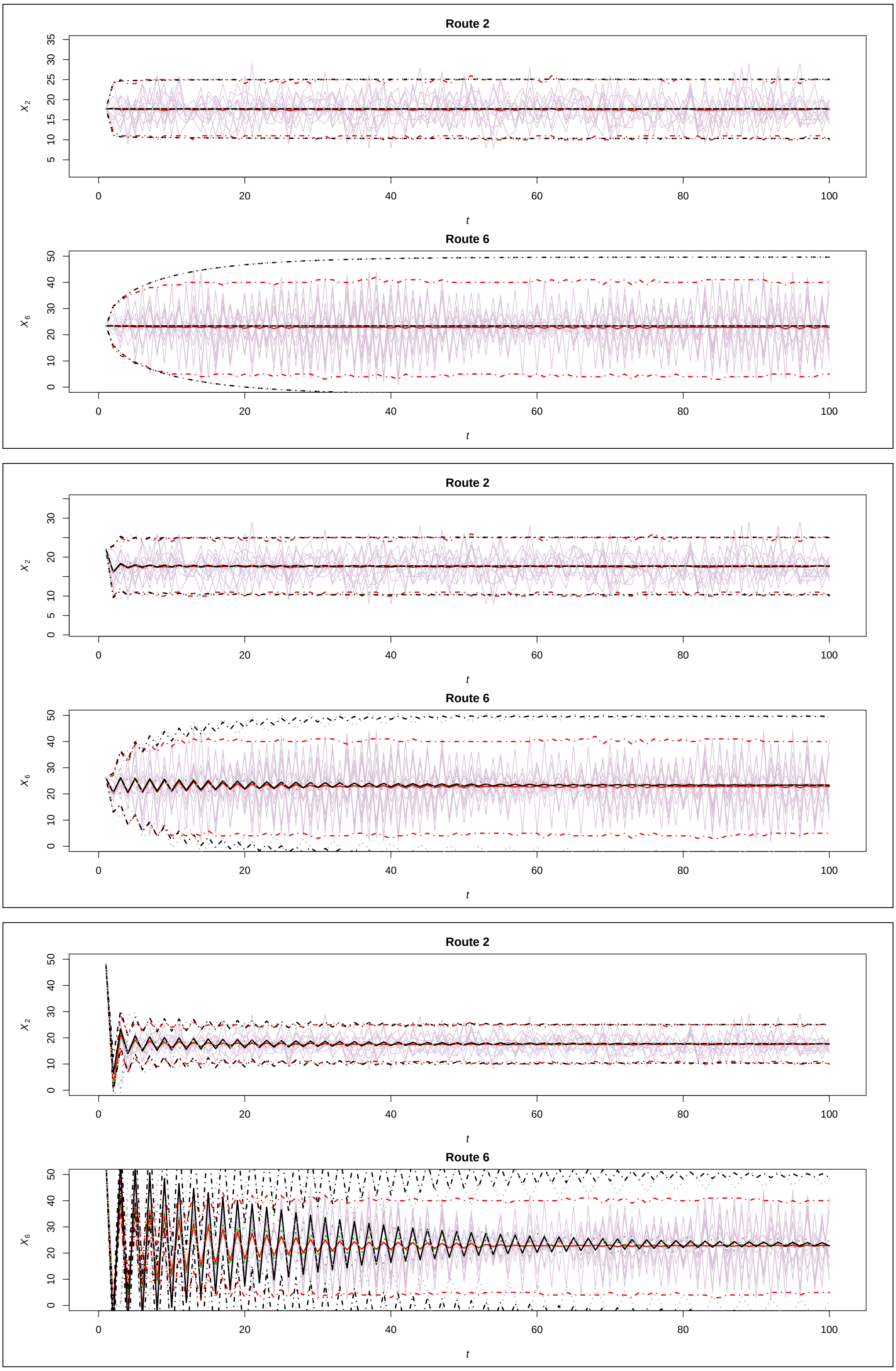


Figure 4.27: Route flows in four OD pair and six route network example with coefficient of reactivity 1.10.

Experiment 4.4.6. We consider the aggregated network of the city Sioux Falls, South Dakota shown in Figure 4.28, with 24 nodes and 76 links. Their link cost function follows the BPR form (2.1) with parameters are taken from “https://github.com/bstabler/Transportation Networks Table”. For this experiment four OD pairs from origin 4 to destination 20, from 6 to 24, from 1 to 19 and from 2 to 23 has been considered. The feasible routes are given in Table 4.2.

Feasible routes for OD (4,20)	Feasible routes for OD (6,24)
route 1: 4 → 5 → 6 → 8 → 7 → 18 → 20	route 5: 6 → 8 → 7 → 18 → 20 → 21 → 24
route 2: 4 → 11 → 14 → 15 → 19 → 20	route 6: 6 → 5 → 4 → 3 → 12 → 13 → 24
route 3: 4 → 5 → 9 → 10 → 16 → 18 → 20	route 7: 6 → 5 → 4 → 11 → 14 → 23 → 24
route 4: 4 → 5 → 9 → 10 → 15 → 19 → 20	route 8: 6 → 2 → 1 → 3 → 12 → 13 → 24
	route 9: 6 → 8 → 16 → 17 → 19 → 15 → 22 → 21 → 24
Feasible routes for OD (1,19)	Feasible routes for OD (2,23)
route 10: 1 → 2 → 6 → 8 → 16 → 17 → 19	route 14: 2 → 1 → 3 → 12 → 13 → 24 → 23
route 11: 1 → 3 → 4 → 5 → 9 → 10 → 15 → 19	route 15: 2 → 1 → 3 → 4 → 11 → 14 → 23
route 12: 1 → 2 → 6 → 5 → 9 → 10 → 17 → 19	route 16: 2 → 1 → 3 → 4 → 5 → 9 → 10 → 15 → 22 → 23
route 13: 1 → 3 → 12 → 11 → 10 → 16 → 18 → 20 → 19	route 17: 2 → 6 → 5 → 9 → 8 → 16 → 17 → 10 → 15 → 14 → 23

Table 4.2: Feasible routes in Sioux falls network

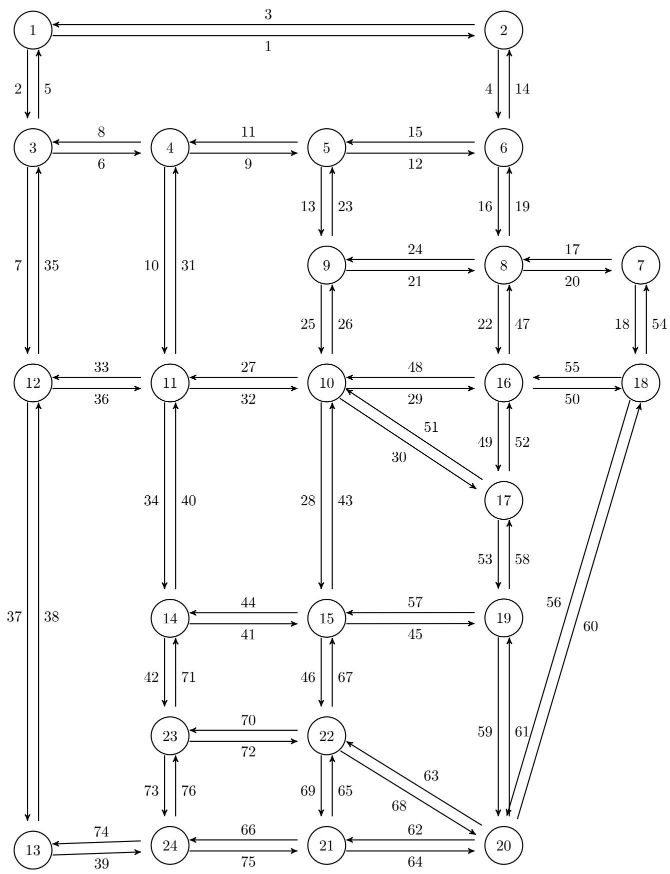


Figure 4.28: Sioux falls network

One route from each OD pair has been chosen and for all plots in this experiment, simulation has been done over 100 and 500 days.

Like the other experiments for small value of coefficient of reactivity ($\omega_0 = 0.68$) and also for the value closest to 1 (Figures 4.29 and 4.30), the approximations have worked well as the process has settled down to its stationary distribution. This can be seen even when the initial state of the system is far away from SUE after around 20 days which is the transient period.

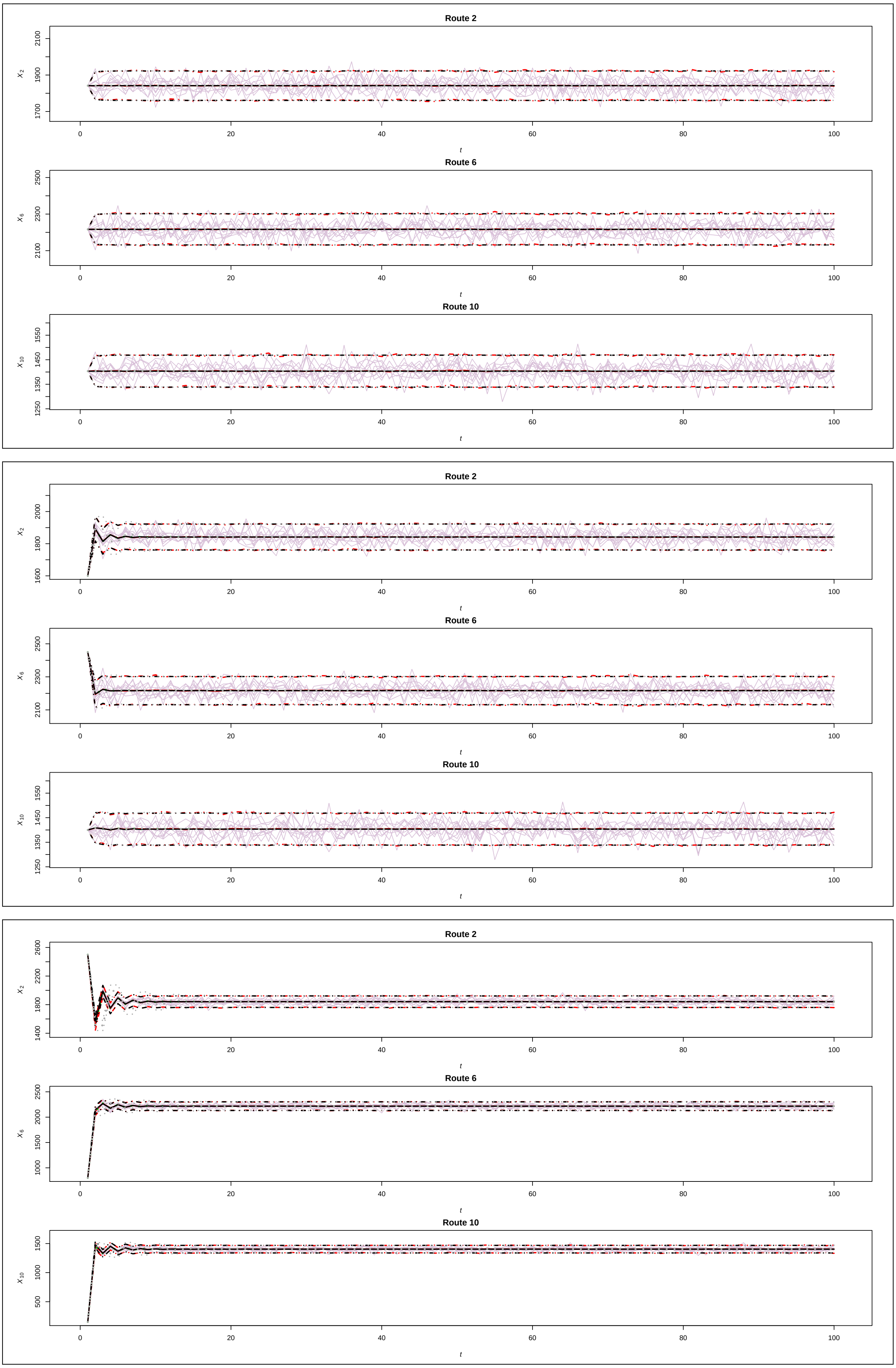


Figure 4.29: Sioux falls network with 17 feasible routes and four OD pairs with coefficient of reactivity 0.68.

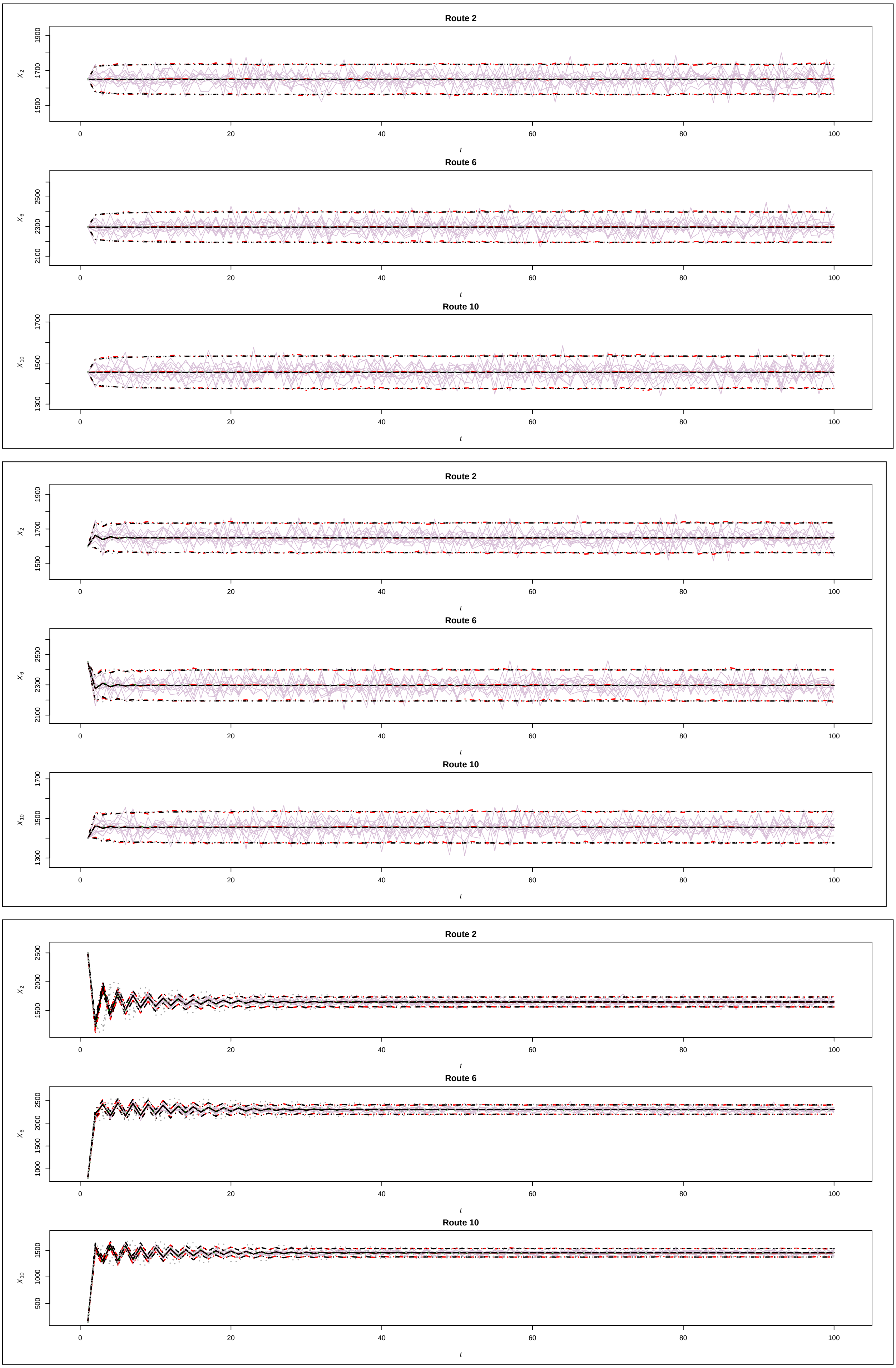


Figure 4.30: Sioux falls network with 17 feasible routes and four OD pairs with coefficient of reactivity 0.96.

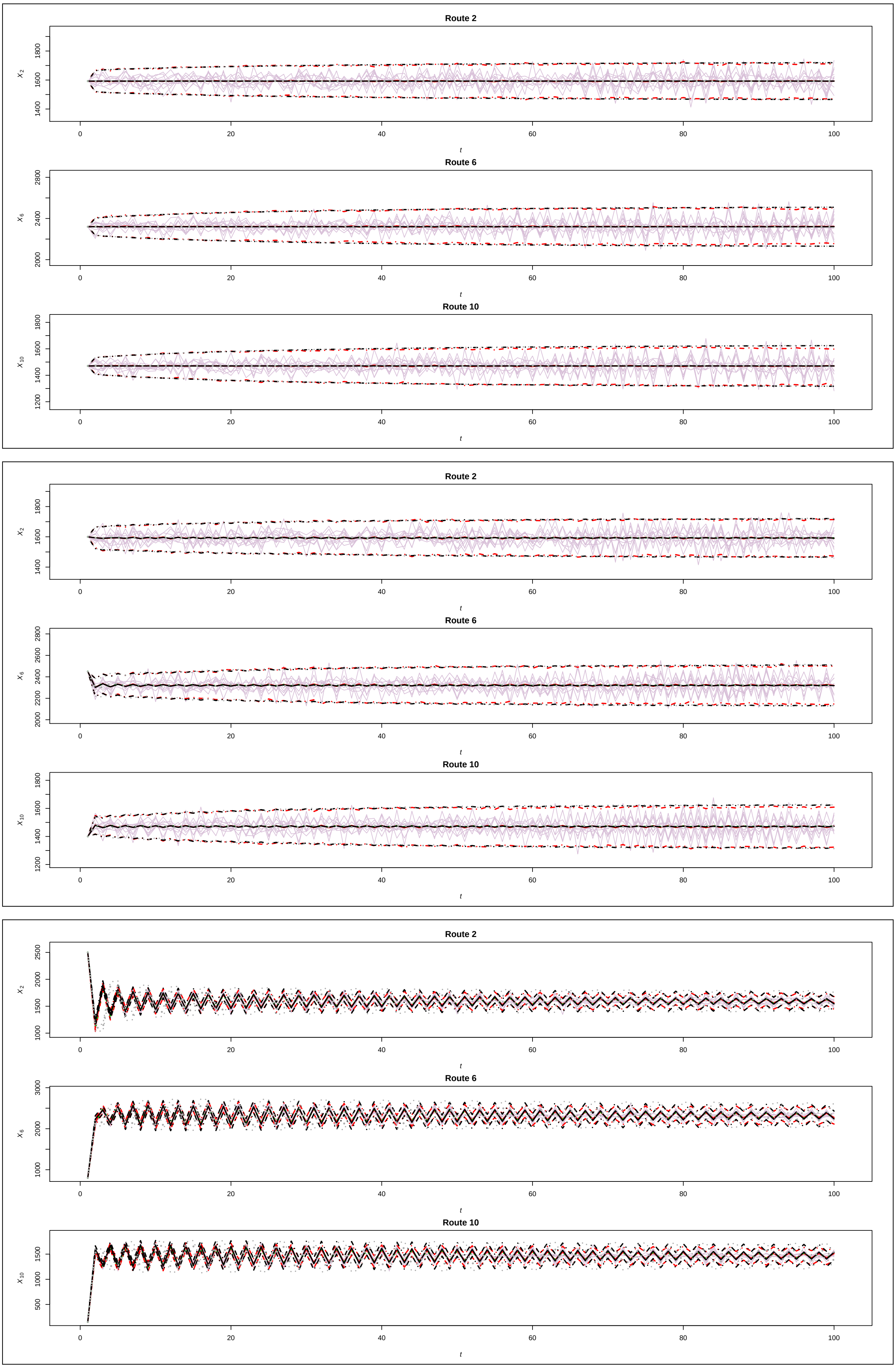


Figure 4.31: Sioux falls network with 17 feasible routes and four OD pairs with coefficient of reactivity 1.06.

Clearly the approximation methods work less well during the transient period when we raise the value of coefficient of reactivity to 1.06. This is reflected in the bottom plot in Figure 4.31 as we see fluctuations in mean flows. However, if we run the simulation over 500 days we observe that there is no oscillation for mean flows and no failure for prediction intervals (Figure 4.32).

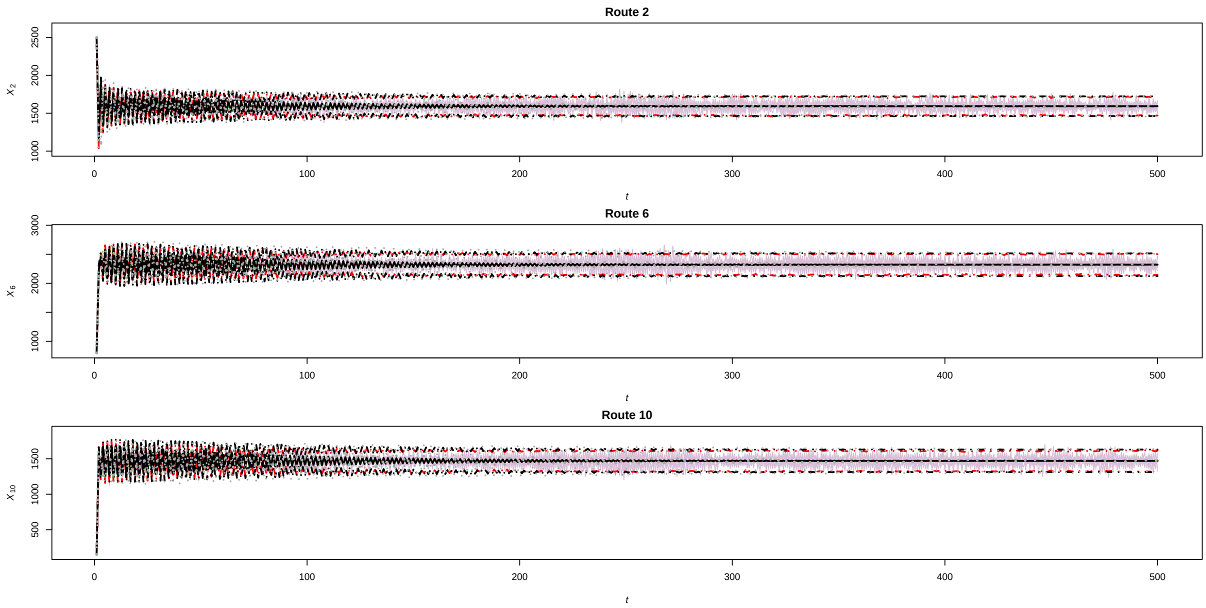


Figure 4.32: Sioux falls network with 17 feasible routes and four OD pairs with coefficient of reactivity 1.06 over 500 days.

Considering a higher value for the coefficient of reactivity ($\omega_0 = 1.15$), Figure 4.33 displays the limit of the prediction interval going out of range before 40 days when the initial state is at SUE. As can be seen, route 16 is a bit more stable and starts to fail after day 50. Furthermore, for all routes the mean approximation methods remain stable.

Looking at the middle panel of Figure 4.33, for which the initial flow pattern is fairly close to SUE, we see that the prediction intervals break down quickly. The mean for Watling and Hazelton's (2018) approximation also starts to fail from around day 60. Also it is clear to see as time goes by the size of the oscillations increase. The mean behavoiur of flows for Watling and Hazelton's (2018) approximation method is less stable compared to Davis and Nihan's (1993) ones. This is evident at least in route 16 where the nonlinear method is still stable at the stationary mean on day 100.

When the value of the coefficient of reactivity reaches 1.15 and the starting state is extremely far away from SUE, the failure of all approximation methods happens very fast as shown in bottom panel of Figure 4.33.

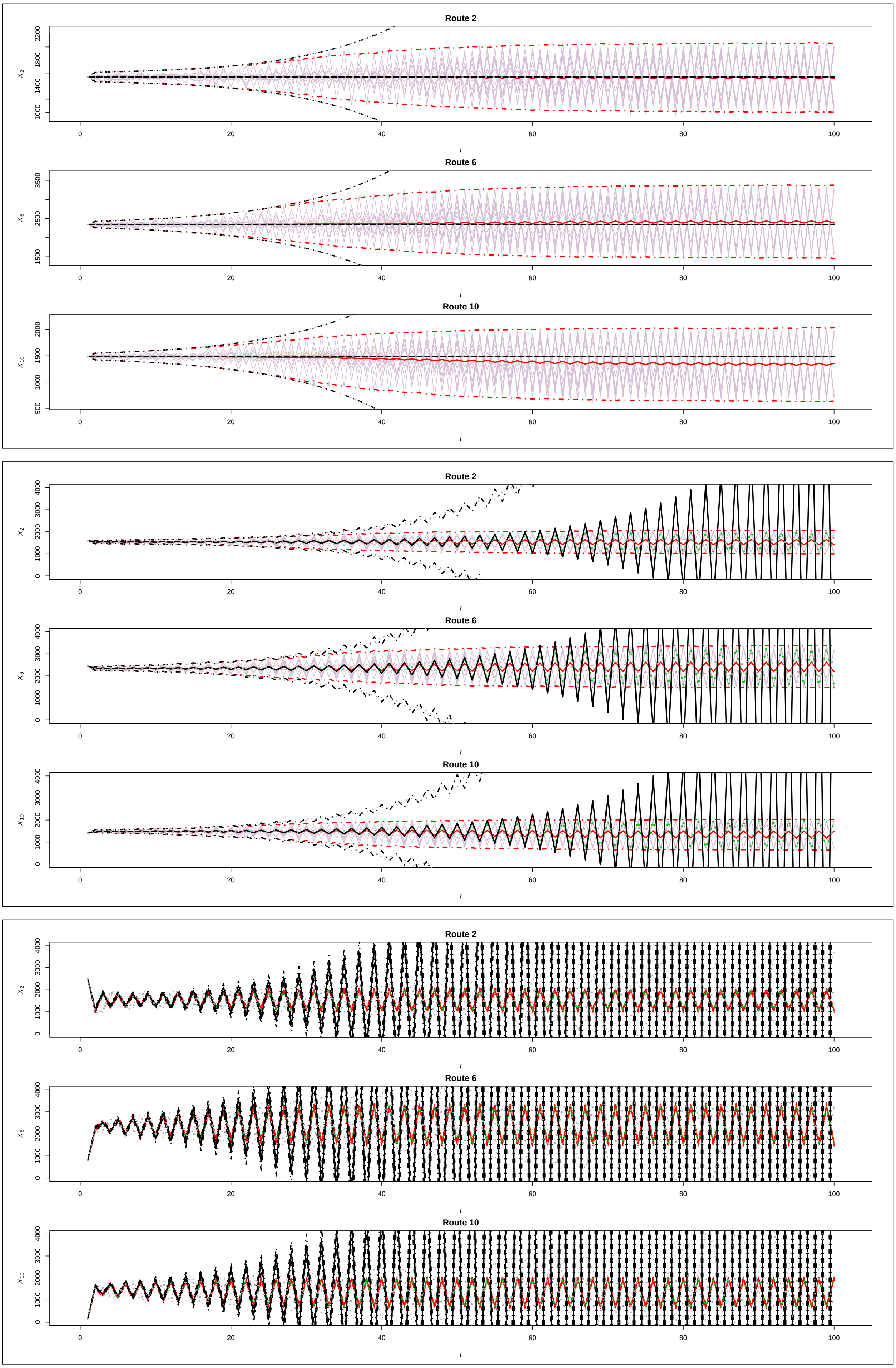


Figure 4.33: Sioux falls network with 17 feasible routes and four OD pairs with coefficient of reactivity 1.15.

4.5 Summary

The coefficient of reactivity value, introduced by Hazelton (2002) and described in Section 4.3, offers a research tool for investigating the relation between stochastic and deterministic traffic assignment models. In other words, the coefficient of reactivity can be used as a tool for assessing when approximations are likely to hold. As shown through the experiments we found that for all networks when the value of the coefficient of reactivity goes up, then approximation methods get worse.

Another conclusion can be drawn from the results for all networks: the value of 1 is a kind of threshold for approximation methods starting to break down (at least the prediction interval) but is not perfect. There is evidence in the examples provided that approximation methods work better for the networks with a greater number of alternative routes serving one OD pair. The reason might be for networks with more routes and more OD pairs there is an option to spread the flows further across the network.

In our experiments we examined the performance of approximation methods during transient periods by initializing the flows away from SUE. We saw that the further away the initial flow pattern is from SUE the more fluctuations happen for a higher value of the coefficient of reactivity. As expected, Davis and Nihan's (1993) approximation works better than Watling and Hazelton's (2018) method for more reactive systems. Because Davis and Nihan's approach involves constant recalculation of the model Jacobians, the quality of the approximation should remain reasonable for flow patterns well away from SUE. In contrast, Watling and Hazelton's (2018) method uses a fixed Jacobian computed at SUE. As those authors showed, the approximations should remain valid for initial flows within a neighbourhood of SUE, but will break down when the initial flow pattern is far from SUE. However, Davis and Nihan's method requires constant recalculation of a very large matrix which makes this method really slow when compared with Watling and Hazelton's (2018) approximation method. For instance, for the longer simulation experiment (Figure 4.32), Watling and Hazelton's (2018) approximation method required 0.18 CPU seconds to run while Davis and Nihan's1993 approximation took 18.55 CPU seconds.

Chapter 5

Extending the Coefficient of Reactivity

5.1 Chapter overview

Travellers on their daily trips make their decisions to reach their destinations based on their experience in terms of time or cost. They select a route or travel mode which optimises their travel cost (time). Any network disruption may have an impact on traveller route choice behaviour. A diverse range of events, from minor accidents on the road to catastrophic events such as natural disasters (earthquake, flood, etc.) or failure of civil infrastructure (bridge collapse, sinkhole, etc.), may cause disruptions to the network. Some of those events have long-term impacts. On the other hand events such as traffic accidents cause short-term disruptions. After any network disruption, travellers may change their normal route, switch to alternative travel modes, change their destination or change their schedule depending on the type of disruption.

Let us assume that an unforeseen occurrence happens to the network on a particular day. For example, temporary road closures or an accident disrupt the normal flow of traffic. We would observe that on such days the cost of using an affected route increases, given that it takes a really long time to pass through that route. Travellers are assumed to make their route choices based on their prior knowledge.

Understanding route choice behaviour is crucial to predict day-to-day traffic flow evolution after an unexpected network disruption. The focus of this chapter is about understanding how travellers adapt to change in a network.

In order to explore the impact of disruption on the transportation network Zhu et al. (2010) investigated the effects of the I-35W bridge collapse in Minneapolis, Minnesota. They found substantial impacts on travel behaviour. Owing to travellers learning and adjusting travel decisions, it took approximately six weeks for the system to return to an (approximate) equilibrium flow behaviour. Zhu and Levinson (2012) provided a review of different types of disruptions and studies on their impacts on travel decisions. Lu et al. (2011) evaluated travellers' route selection changes by introducing the en-route real-time information on the occurrence of an incident. Wang et al. (2013) developed a dynamical framework to integrate individual perception and decision schemes under risk with group learning. An empirical study by Watling et al. (2012) analysed two planned network disruptions in the city of York, one a bridge closure and the other a capacity reduction for maintenance works. They developed a four step model and they reported how well a traffic equilibrium model would predict real network impacts affecting road capacity.

He and Liu (2012) proposed a deterministic day-to-day traffic assignment model to capture the characteristics of traffic flow evolution after an unexpected network disruption by using field data collected after the I-35W Bridge collapse in Minneapolis, Minnesota. These works provide insight into how long people take to learn about the impacts of a major disruption and how they adjust their routing decisions in the long term.

Assume for a moment that travellers make their route choice decision entirely based on one previous day. Now, if the route was expensive the previous day because of a disruption, then they will most likely switch their route. But if travellers select their route based on more than one day's memory, they might ignore it and use the same route again. In terms of route choice behaviour, it is interesting to investigate how travellers' will react to various types of one-off event, and examine the scale of response by travellers in response to a disruption.

In chapter 4 we reviewed a measure called the coefficient of reactivity suggested by Hazelton (2002) to summarise the reaction of a traffic system to a disruption when the flow pattern is changed to some unusual state (e.g. because of a major accident) for one day. Here, we introduce a refined definition of the coefficient of reactivity, which takes better account of variation in historical flows. Extension of the definition of the coefficient of reactivity to allow for assessment of the impact of longer disruptions to the system is another novelty of this chapter.

This chapter is organized as follows. Section 5.2 presents an m-memory Markov model that will be used in this chapter, and applies that to find the value of the original coefficient of reactivity reviewed in Chapter 4. We illustrate how the coefficient of reactivity relates to the flow properties of this type of model through some simulations. In Section 5.3 we present the new definition of the coefficient of reactivity where we allow the system to have more variation for flows for days before the disruption, contrary to the original definition where the flows are set to be at a fixed mean state. The approach to calculating the newly defined coefficient of reactivity is illustrated using the same example as in Section 4.3. In the following, we propose an algorithm for computing the new coefficient of reactivity. Then through a theorem, the relationship between the original and new definition of the coefficient of reactivity will be demonstrated. In Section 5.4 we replace the original definition of the coefficient of reactivity by another definition where we extend the length of an unusual event. Simulation results for this generalized coefficient of reactivity are provided in Section 5.5 We present the findings in Section 5.6.

5.2 Models with memory length m

As stated in Chapter 2, the day-to-day evolution of a traffic system can be modelled as a discrete-time Markov process. In such models, the route choice probability at day t are functions of the route costs observed in the finite past. Here, we assume m is the length of memory, which is the number of days the system looks back for day t (i.e. $t-1, t-2, \ldots, t-m$). It is assumed that travellers update their perceived cost for the route through a (dis)utility function which is a linear filter of past costs given by $\boldsymbol{u}^{(t-1)} = \sum_{j=1}^{m} \delta_j \mathbf{C}(\boldsymbol{x}^{(t-j)})$. Typically trvellers tend to remember the most recent experience very well compared to the old one. Hence, $\delta_1, \ldots, \delta_m$ are usually a decreasing sequence summing to unity. Travellers will take feasible routes with minimum personal disutility then the vector of route flows on day t will follow a multinomial distribution,

$$\boldsymbol{x}_j^{(t)} | \boldsymbol{u}^{(t-1)} \sim \text{Multinomial}(N_j, \boldsymbol{p}_j(\boldsymbol{u}^{(t-1)})),$$

where $p_r(\boldsymbol{u}^{(t-1)}) = \frac{\exp\left\{-\theta u_r^{(t-1)}\right\}}{\sum_r \exp\left\{-\theta u_r^{(t-1)}\right\}}, \quad r = 1, 2.$

Since the probability distribution of the state on day t is fully determined by the previously realised values of the states $\boldsymbol{x}^{(t-j)} \quad j = 1, 2, \ldots, m$, $\boldsymbol{x}^{(t)}$ is a m-dependent Markov process. This means the state $S^t = (\boldsymbol{x}^t, \boldsymbol{x}^{t-1}, \ldots, \boldsymbol{x}^{t-m+1})$ is a Markov process.

In Chapter 4 we have shown that the original coefficient of reactivity can measure the impact of a short-term disruption (such as a road closure) when travellers make their decisions based on experiences on the previous day. In essence, we demonstrated how the coefficient of reactivity works when travellers' have one day memory and one day disruption. However, Hazelton's (2002) definition works when we have longer memory with one day disruption in the network. Here, we demonstrate how the value of the coefficient of reactivity changes purely according to the weight given to the most recent memory when $m > 1$.

To illustrate, we consider the same two route network as 2.7.1 with N travellers. Similarly the route choice probabilities are given by the logit model. The utilities when travelers make their choice are based on the experience of the last two days, and are given by

$$u_1^{t-1} = \delta_1 \left(a_1 + \left(\frac{x_1^{t-1}}{b} \right)^2 \right) + (1 - \delta_1) \left(a_1 + \left(\frac{x_1^{t-2}}{b} \right)^2 \right)$$

and

$$u_2^{t-1} = \delta \left(a_2 + \left(\frac{x_2^{t-1}}{b} \right)^2 \right) + (1 - \delta) \left(a_2 + \left(\frac{x_2^{t-2}}{b} \right)^2 \right).$$

Therefore, the probability of choosing route 1 i.e. $p_1(\boldsymbol{x}^{(t-1)}, \boldsymbol{x}^{(t-2)})$ can be written as

$$p_1(\boldsymbol{x}^{(t-1)}, \boldsymbol{x}^{(t-2)}) = \frac{\exp\{-\theta[u_1^{t-1}]\}}{\exp\{-\theta[u_1^{t-1}]\} + \exp\{-\theta[u_2^{t-1}]\}}.$$

Here we set $\boldsymbol{x}^{(t-1)} = \boldsymbol{x}$ and $\boldsymbol{x}^{(t-2)} = \hat{\boldsymbol{x}}$ following the definition of $\psi(\boldsymbol{x})$ from equation

4.7. As a consequence $\psi_1(\boldsymbol{x})$ is then given by

$$\psi_1(\boldsymbol{x}) = \frac{N}{1+\exp\left\{-\theta\left[\delta(a_2+(\frac{x_2}{b})^2)+(1-\delta)(a_2+\left(\frac{\hat{x_2}}{b}\right)^2)-\delta(a_1+(\frac{x_1}{b})^2)-(1-\delta)(a_1+\left(\frac{\hat{x_1}}{b}\right)^2)\right]\right\}}.$$

Similarly $\psi_2(\boldsymbol{x})$ can be expressed by

$$\psi_2(\boldsymbol{x}) = \frac{N}{1+\exp\left\{-\theta\left[\delta(a_1+(\frac{x_1}{b})^2)+(1-\delta)(a_1+\left(\frac{\hat{x_1}}{b}\right)^2)-\delta(a_2+(\frac{x_2}{b})^2)-(1-\delta)(a_2+\left(\frac{\hat{x_2}}{b}\right)^2)\right]\right\}}.$$

To simplify the notation, we define

$$\exp\{-\theta[\delta(a_2+\left(\frac{x_2}{b}\right)^2)+(1-\delta)(a_2+\left(\frac{\hat{x_2}}{b}\right)^2)-\delta(a_1+\left(\frac{x_1}{b}\right)^2)-(1-\delta)(a_1+\left(\frac{\hat{x_1}}{b}\right)^2)]\} = \nu_1$$

and

$$\exp\{-\theta[\delta(a_1+\left(\frac{x_1}{b}\right)^2)+(1-\delta)(a_1+\left(\frac{\hat{x_1}}{b}\right)^2)-\delta(a_2+\left(\frac{x_2}{b}\right)^2)-(1-\delta)(a_2+\left(\frac{\hat{x_2}}{b}\right)^2)]\} = \nu_2.$$

The Jacobian matrix of ψ with respect to $\hat{\boldsymbol{x}}$ is then given by

$$\boldsymbol{J} = \begin{bmatrix} \frac{-N\theta\delta\left(\frac{2\hat{x_1}}{b^2}\right)\nu_1}{(1+\nu_1)^2} & \frac{N\theta\delta\left(\frac{2\hat{x_2}}{b^2}\right)\nu_1}{(1+\nu_1)^2} \\ \frac{N\theta\delta\left(\frac{2\hat{x_1}}{b^2}\right)\nu_2}{(1+\nu_2)^2} & \frac{-N\theta\delta\left(\frac{2\hat{x_2}}{b^2}\right)\nu_2}{(1+\nu_2)^2} \end{bmatrix}.$$

Since $\frac{\nu_1}{(1+\nu 1)^2} = \frac{\nu_2}{(1+\nu 2)^2}$ and $\hat{x_1}\hat{x_2} = \frac{N^2\nu_1}{(1+\nu 1)^2}$, The Jacobian matrix can be written as

$$\boldsymbol{J} = 2\theta\delta\hat{x}_1\hat{x}_2 b^{-2}N^{-1}\begin{bmatrix} -\hat{x}_1 & \hat{x}_2 \\ \hat{x}_1 & -\hat{x}_2 \end{bmatrix}.$$

Similar to the steps we took in Section 4.3.1, it can be concluded

$$\omega_0 = \gamma N = 2\theta\delta\left(\frac{\hat{x_1}}{b}\right)\left(\frac{\hat{x_2}}{b}\right). \tag{5.1}$$

This example also shows that the coefficient of reactivity depends on the weight attributed (δ) to memory from just one recent day in the cases where $m > 1$.

It is straightforward to show that the asymptotic coefficient of reactivity ω_0 is exactly the same for any memory length m so long as the coefficient of $C(x^{t-1})$ remains unchanged as δ_1. The values of the other coefficients are irrelevant. That is because the more historical flow patterns are set to SUE.

Now we present some simulation results when travellers consider the memory of previous days to select their route for any day. This will help us gain a better understanding of the relation between the value of the coefficient of reactivity and travellers behaviour after a disruption. We continue to use the two-route network from above. The next three plots illustrate how the coefficient of reactivity works based on three days of memory and one-day disruption. To do the simulation we assign traffic to each route for a sequence of 50 days. We assume the network disruption happens on day 41 for route 1. This disruption was modelled by setting the flow for route 2 on day 41 at 0.

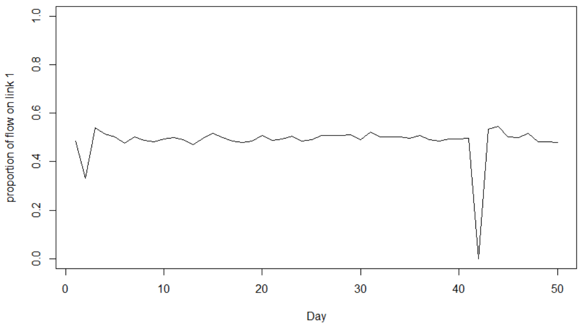


Figure 5.1: Route 1 flows simulated over 50 days as a fraction of total demand with $\omega_0 = 0.1$.

Through the definition of the coefficient of reactivity, we can describe what happens directly after a network disruption. As can be seen in Figure 5.1 with a very low value of coefficient of reactivity 0.1 the system does not react at all to the disruption. In this case, travellers' are not going to change their route the day after the disruption. Therefore, the system immediately returns to normal flow behaviour.

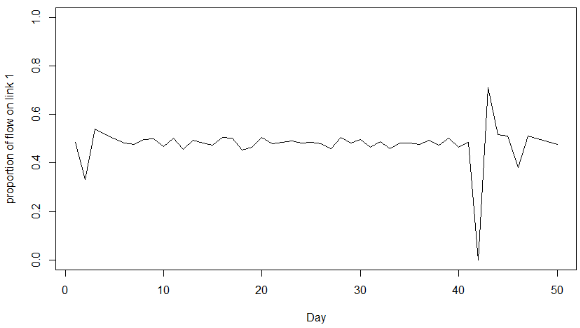


Figure 5.2: Route 1 flows simulated over 50 days as a fraction of total demand with $\omega_0 = 0.5$.

Figure 5.2 shows a flow pattern when the value of the coefficient of reactivity is 0.5. We can see there is a bit of a reaction on the day after the disruption since the proportion of flows on route 1 on day 42 is around 0.7. Also, there is a small change to normal behaviour, but the system quickly tracks back to the normal flow pattern as it was before disruption.

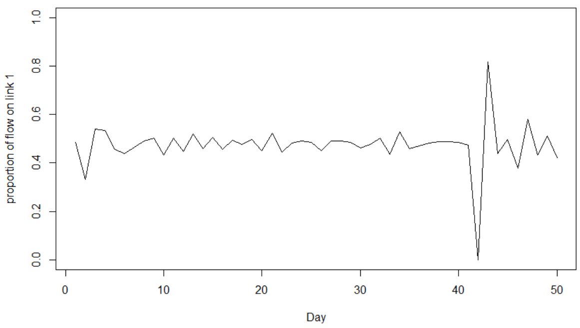


Figure 5.3: Route 1 flows simulated over 50 days as a fraction of total demand with $\omega_0 = 0.9$.

Increasing the value of the coefficient of reactivity leads to the system becoming more reactive. Figure 5.3 depicts for a higher value of the coefficient of reactivity 0.9, more travellers' using route 2 for the next day after a disruption.

It should be noted that when travellers have a memory length greater than one day, then they will also remember the state of the system on earlier days. Hazelton (2002) dealt with this by setting all flows before the disruption to SUE. However, that ignores the important historical variation, and hence motivate us to define a new definition.

5.3 New version of coefficient of reactivity

As stated before, in Hazelton's (2002) original definition of the coefficient of reactivity, the flows from day $t-2$ and before are set to the stationary mean. This means that we are considering a system that had no variation up to day $t-2$, which is of course unrealistic. Here, we attempt to relax that assumption by allowing for variation (according to the stationary distribution) for flows on day $t-2$ and before. We specify a new definition as follows:

$$\psi_m^{new}(\boldsymbol{x}) = E[\boldsymbol{x}^{(t)}|\boldsymbol{x}^{(t-1)} = \boldsymbol{x}], \tag{5.2}$$

then the value of the coefficient of reactivity is

$$\sup_{\boldsymbol{x}} \frac{||\psi_m^{new}(\boldsymbol{x}) - \hat{\boldsymbol{x}}||}{||\boldsymbol{x} - \hat{\boldsymbol{x}}||}$$

where $\hat{\boldsymbol{x}}$ is the mean of stationary distribution. Here, we consider variation in $\boldsymbol{x}^{(t-2)}, \ldots, \boldsymbol{x}^{(t-m)}$ when computing the expectation in contrast to Hazelton's definition where these historical flows all set to be at the stationary mean. It can be seen that for a one day memory the new definition of ψ and $\psi_m(\boldsymbol{x})$ (4.7 and 5.2) are exactly the same. When m is greater than

1 the new definition can be expanded as follows:

$$\begin{aligned}\psi_m^{new}(\boldsymbol{x}) &= \mathsf{E}[\boldsymbol{x}^{(t)}|\boldsymbol{x}^{(t-1)} = \boldsymbol{x}] \\ &= \mathsf{E}[\mathsf{E}[\boldsymbol{x}^{(t)}|\boldsymbol{x}^{(t-1)} = \boldsymbol{x}, \boldsymbol{x}^{(t-2)}, \ldots, \boldsymbol{x}^{(t-m)}] \\ &= \underbrace{\sum \cdots \sum}_{m-2\, times} \mathsf{E}[\boldsymbol{x}^{(t)}|\boldsymbol{x}^{(t-1)} = \boldsymbol{x}, \boldsymbol{x}^{(t-2)}, \ldots, \boldsymbol{x}^{(t-m)}] \\ &\times f(\boldsymbol{x}^{(t-2)}, \ldots, \boldsymbol{x}^{(t-m)})\end{aligned} \tag{5.3}$$

where $f(\boldsymbol{x}^{(t-2)}, \ldots, \boldsymbol{x}^{(t-m)})$ is the joint distribution of the route flow vectors from day $t-m$ to $t-2$. This joint distribution is describing how the flow patterns on day $(t-2)$ back to day $(t-m)$ vary under the stationary distribution for the model. This new version of coefficient of reactivity averages over all possible sets of historical flow patterns.

5.3.1 An illustrative example for the new definition of coefficient of reactivity

For this part we reconsider our two route network example (2.7.1) to obtain the coefficient of reactivity through the new definition of $\psi_m^{new}(\boldsymbol{x})$. In order to apply the new definition of ψ, we look at a case with a single traveller. Specifically, this is a situation in with the coefficient of reactivity can be computed exactly. We set $m = 2$ for the memory length of travellers. The probability that the travellers takes route r is given by

$$P_r(\boldsymbol{x}^{(t-1)}, \boldsymbol{x}^{(t-2)}) = \frac{\exp\{-\theta[\delta c_r x_r^{(t-1)} + (1-\delta) c_r x_r^{(t-2)}]\}}{\sum_r \exp\{-\theta[\delta c_r x_r^{(t-1)} + (1-\delta) c_r x_r^{(t-2)}]\}}, \quad r = 1, 2.$$

We can rewrite the definition of $\psi_m^{new}(\boldsymbol{x})$ as follow:

$$\psi_m^{new}(\boldsymbol{x}) = \sum_z \mathsf{E}[\boldsymbol{x}^{(t)}|\boldsymbol{x}^{(t-1)} = \boldsymbol{x}, \boldsymbol{x}^{(t-2)} = \boldsymbol{z}] \times f(\boldsymbol{x}^{(t-2)} = \boldsymbol{z})$$

since we have one traveller it can then be written as

$$\begin{aligned}\psi_m^{new}(\boldsymbol{x}) &= \mathsf{E}\left[\boldsymbol{x}^{(t)}|\boldsymbol{x}^{(t-1)} = \boldsymbol{x}, \boldsymbol{x}^{(t-2)} = \begin{bmatrix}0\\1\end{bmatrix}\right] \times f\left(\boldsymbol{x}^{(t-2)} = \begin{bmatrix}0\\1\end{bmatrix}\right) \\ &+ \mathsf{E}\left[\boldsymbol{x}^{(t)}|\boldsymbol{x}^{(t-1)} = \boldsymbol{x}, \boldsymbol{x}^{(t-2)} = \begin{bmatrix}1\\0\end{bmatrix}\right] \times f\left(\boldsymbol{x}^{(t-2)} = \begin{bmatrix}1\\0\end{bmatrix}\right).\end{aligned}$$

In order to compute the stationary distribution, it is necessary to find the transition probabilities matrix. In accordance with two days memory of the system the state on day t is given by $S^t = \begin{bmatrix}\boldsymbol{x}^t \\ \boldsymbol{x}^{t-1}\end{bmatrix}$. Since we have one traveller and two days memory, therefore, four possibilities for the state on day t are then given as $\{[1 \quad 0 \quad 1 \quad 0]^T, [1 \quad 0 \quad 0 \quad 1]^T, [0 \quad 1 \quad 1 \quad 0]^T, [0 \quad 1 \quad 0 \quad 1]^T\}$.

Because the system has only two routes we can describe the state of the system based on flow on route 1 on days t and $t-1$. Hence, the first element of this matrix is defined as

$$\mathbb{P}(s^t = \begin{bmatrix}1\\0\end{bmatrix}|s^{t-1} = \begin{bmatrix}1\\0\end{bmatrix}),$$

which describes the probability that a travellers start off at state $\{(1 \quad 0 \quad 1 \quad 0)^T\}$ and finishes at the same state, can be calculated as:

$$\begin{aligned}
&\mathbb{P}\big(s^t = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right] | s^{t-1} = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]\big)\\
=&\mathbb{P}\big(x^t = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right], x^{t-1} = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right] | x^{t-1} = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right], x^{t-2} = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]\big)\\
=&\mathbb{P}\big(x^t = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right] | x^{t-1} = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right], x^{t-2} = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]\big).
\end{aligned}$$

Applying the route choice probability, we have

$$\begin{aligned}
&\mathbb{P}\big(s^t = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right] | s^{t-1} = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]\big)\\
=&\frac{\exp\{-\theta[\delta c_1(1) + (1-\delta)c_1(0)]\}}{\exp\{-\theta[\delta c_1(1) + (1-\delta)c_1(1)]\} + \exp\{-\theta[\delta c_2(0) + (1-\delta)c_2(0)]\}}\\
=&\frac{\exp\{-\theta[\delta(a_1 + (\frac{1}{b})^2) + (1-\delta)(a_1 + (\frac{1}{b})^2)]\}}{\exp\{-\theta[\delta(a_1 + (\frac{1}{b})^2) + (1-\delta)(a_1 + (\frac{1}{b})^2)]\} + \exp\{-\theta[\delta(a_2 + (\frac{0}{b})^2) + (1-\delta)(a_2 + (\frac{0}{b})^2)]\}}\\
=&\frac{\exp\{-\theta[(a_1 + (\frac{1}{b})^2)]\}}{\exp\{-\theta[(a_1 + (\frac{1}{b})^2)] + \exp(-\theta[a_2]\}}.
\end{aligned} \tag{5.4}$$

It is easy to see that $\mathbb{P}\big(s^t = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right] | s^{t-1} = \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]\big) = 0$. This happens since as shown in equation 5.5 the flow on day $t-1$ do not match together.

$$\begin{aligned}
&\mathbb{P}\big(s^t = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right] | s^{t-1} = \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]\big)\\
=&\mathbb{P}\big(x^t = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right], x^{t-1} = \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right] | x^{t-1} = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right], x^{t-2} = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]\big)\\
=&0.
\end{aligned} \tag{5.5}$$

In order to simplify the calculations, by replacing $\rho_1 = \exp\{-\theta[a_1 + (\frac{1}{b})^2]\}$ and $\rho_2 = \exp\{-\theta[a_2]\}$ in equation 5.4 the first element of our transition matrix is then $\frac{\rho_1}{\rho_1 + \rho_2}$. Following the same calculations as 5.4 and 5.5 we can find the remaining elements of the transition matrix. By defining

$$\begin{aligned}
\rho_3 &= \exp\{-\theta[a_1 + \delta(\frac{1}{b})^2]\}\\
\rho_4 &= \exp\{-\theta[a_2 + (1-\delta)(\frac{1}{b})^2]\}\\
\rho_5 &= \exp\{-\theta[a_1 + (1-\delta)(\frac{1}{b})^2]\}\\
\rho_6 &= \exp\{-\theta[a_2 + \delta(\frac{1}{b})^2]\}\\
\rho_7 &= \exp\{-\theta[a_1]\}\\
\rho_8 &= \exp\{-\theta[a_2 + (\frac{1}{b})^2]\}
\end{aligned}$$

then the transition matrix can be shown as

$$\begin{pmatrix} \frac{\rho_1}{\rho_1+\rho_2} & 0 & \frac{\rho_2}{\rho_1+\rho_2} & 0 \\ \frac{\rho_3}{\rho_3+\rho_4} & 0 & \frac{\rho_4}{\rho_3+\rho_4} & 0 \\ 0 & \frac{\rho_5}{\rho_5+\rho_6} & 0 & \frac{\rho_6}{\rho_5+\rho_6} \\ 0 & \frac{\rho_7}{\rho_7+\rho_8} & 0 & \frac{\rho_8}{\rho_7+\rho_8} \end{pmatrix}.$$

Now it is necessary to find the fixed row vector (stationary distribution), to do so it is needed to solve the following equations:

$$\begin{aligned} \frac{\rho_1}{\rho_1+\rho_2}\pi_1 + \frac{\rho_3}{\rho_3+\rho_4}\pi_2 &= \pi_1 \\ \frac{\rho_5}{\rho_5+\rho_6}\pi_3 + \frac{\rho_7}{\rho_7+\rho_8}\pi_4 &= \pi_2 \\ \frac{\rho_2}{\rho_1+\rho_2}\pi_1 + \frac{\rho_4}{\rho_3+\rho_4}\pi_2 &= \pi_3 \\ \frac{\rho_6}{\rho_5+\rho_6}\pi_3 + \frac{\rho_8}{\rho_7+\rho_8}\pi_4 &= \pi_4 \\ \pi_1+\pi_2+\pi_3+\pi_4 &= 1. \end{aligned}$$

Solving the above equations we obtain $\pi_2 = \pi_3$ and

$$\begin{aligned} \pi_1 &= \frac{\rho_3\rho_7(\rho_5+\rho_6)(\rho_1+\rho_2)}{\rho_3\rho_7(\rho_5+\rho_6)(\rho_1+\rho_2)+2\rho_2\rho_7(\rho_3+\rho_4)(\rho_5+\rho_6)+\rho_2\rho_6(\rho_3+\rho_4)(\rho_7+\rho_8)} \\ \pi_2 &= \frac{\rho_2\rho_7(\rho_5+\rho_6)(\rho_3+\rho_4)}{\rho_3\rho_7(\rho_5+\rho_6)(\rho_1+\rho_2)+2\rho_2\rho_7(\rho_3+\rho_4)(\rho_5+\rho_6)+\rho_2\rho_6(\rho_3+\rho_4)(\rho_7+\rho_8)} \\ \pi_4 &= \frac{\rho_2\rho_6(\rho_7+\rho_8)(\rho_3+\rho_4)}{\rho_3\rho_7(\rho_5+\rho_6)(\rho_1+\rho_2)+2\rho_2\rho_7(\rho_3+\rho_4)(\rho_5+\rho_6)+\rho_2\rho_6(\rho_3+\rho_4)(\rho_7+\rho_8)}. \end{aligned}$$

According to the new definition of ψ,

$$\begin{aligned} \psi(\begin{bmatrix}1\\0\end{bmatrix}) = \mathsf{E}\left[\boldsymbol{x}^{(t)}|\boldsymbol{x}^{(t-1)} = \begin{bmatrix}1\\0\end{bmatrix}, \boldsymbol{x}^{(t-2)} = \begin{bmatrix}0\\1\end{bmatrix}\right] \times f(\boldsymbol{x}^{(t-2)} = \begin{bmatrix}0\\1\end{bmatrix}) \\ + \mathsf{E}\left[\boldsymbol{x}^{(t)}|\boldsymbol{x}^{(t-1)} = \begin{bmatrix}1\\0\end{bmatrix}, \boldsymbol{x}^{(t-2)} = \begin{bmatrix}1\\0\end{bmatrix}\right] \times f(\boldsymbol{x}^{(t-2)} = \begin{bmatrix}1\\0\end{bmatrix}) \end{aligned}$$

where $f(\boldsymbol{x}^{(t-2)} = \begin{bmatrix}1\\0\end{bmatrix}) = \pi_1 + \pi_3$ and $f(\boldsymbol{x}^{(t-2)} = \begin{bmatrix}0\\1\end{bmatrix}) = \pi_2 + \pi_4$. Since we have a simple two route network the number of travellers on day t follow the Binomial distribution, therefore $\mathsf{E}[\boldsymbol{x}^{(t)}] = (P(x_1 = 1), P(x_2 = 1))^T$. Applying the stationary results and route choice probability, we obtain

$$\begin{aligned} \psi(\begin{bmatrix}1\\0\end{bmatrix}) &= \frac{\rho_3}{\rho_3+\rho_4}(\pi_2+\pi_4) + \frac{\rho_1}{\rho_1+\rho_2}(\pi_1+\pi_3) \\ \psi(\begin{bmatrix}0\\1\end{bmatrix}) &= \frac{\rho_7}{\rho_7+\rho_8}(\pi_2+\pi_4) + \frac{\rho_5}{\rho_5+\rho_6}(\pi_1+\pi_3). \end{aligned}$$

The maximum value of $\frac{\|\psi_m^{new}(\boldsymbol{x}) - \hat{\boldsymbol{x}}\|}{\|\boldsymbol{x} - \hat{\boldsymbol{x}}\|}$ represents the coefficient of reactivity. By considering the same Example 2.7.1 with the same values for the parameters a_1 and a_2 and setting the other parameters as $\delta = 0.6$, $\theta = 0.5$ and $b = 2$, the transition matrix is given by

$$P = \begin{bmatrix} 0.3486451 & 0.0000000 & 0.6513549 & 0.0000000 \\ 0.3716838 & 0.0000000 & 0.6283162 & 0.0000000 \\ 0.0000000 & 0.3834335 & 0.0000000 & 0.6165665 \\ 0.0000000 & 0.4073334 & 0.0000000 & 0.5926666 \end{bmatrix},$$

and the stationary distribution is then

$$\boldsymbol{\pi} = \begin{bmatrix} 0.140 & 0.245 & 0.245 & 0.371 \end{bmatrix}.$$

The mean of the stationary distribution is $\hat{\boldsymbol{x}} = \begin{bmatrix} 0.385 \\ 0.615 \end{bmatrix}$. As a result we obtain the value of the coefficient of reactivity by taking the maximum value from $\frac{\|\psi(\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]) - 0.385\|}{\|1 - 0.385\|}$ and $\frac{\|\psi(\left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]) - 0.615\|}{\|1 - 0.615\|}$ in 3 decimal places as 0.565.

The new version of the coefficient of reactivity is more attractive from a theoretical perspective since it takes proper account of historical variation of flows. However, there is no simple way for computing it, since it depends on the joint distribution of flow patterns between times $t-2$ and $t-m$. In principle, it can be approximated through simulation. In the following subsection we introduce an algorithm to compute the value of the coefficient of reactivity based on our new definition.

5.3.2 An algorithm for computing the new coefficient of reactivity

As stated before, the new coefficient of reactivity requires calculating expected values of the flows at day t given the value of flow on day $t-1$ is $\boldsymbol{x}$, irrespective of what happen previously. The point of this section is to describe how the new coefficient of reactivity can be computed using simulation. We will approximate the true value of ψ by averages taking over simulations. In particular, we need to create simulation for each "history" of flows, then impose the disrupted flow pattern $\boldsymbol{x}$ on day t, and then work out the mean response. We simulated history of flows over 150 days. To do so, we fixed the flows on days $1, 2, \ldots, m$ at SUE, then we simulated the flows for the next 150 consecutive days using a multinomial distribution and employing the logit formula for route choice probability. We repeated this procedure 100 times to create different histories and we kept the last $m-1$ days of each histories.

In order to find ψ it is needed to force the system to take all possibilities of flows at day t which can be thought of as a disruption in the network. It should be pointed out that due to the complexity, the following algorithm only works for networks with a maximum of two routes per OD pair. Moving forward, for each combination of flow history and setting of flow on day t, 10 different instances (route flows) have been simulated for day $t+1$. For each set flow pattern x on day t (the day of disruption), the mean flow on day $t+1$ was computed by averaging over all histories and all instances. By replacing these values as ψ in $\dfrac{\|\psi_m(\boldsymbol{x}) - \hat{\boldsymbol{x}}\|}{\|\boldsymbol{x} - \hat{\boldsymbol{x}}\|}$, the maximum value of these outputs can be considered as the value of the coefficient of reactivity.

Pseudo code for calculating the value of the coefficient of reactivity based on the new definition is presented in Algorithms 2 and 3. In these algorithms we let Nrun denote the history length, Nff the number of feasible flows for day $t+1$, Nh the number of histories, and Ni the number of instances.

Algorithm 2 History of Flows

Input: network information, θ, $\boldsymbol{\delta}$, m
Output: Create travellers' history of flows over 150 days

function HISTORY.FLOW(network information, θ, $\boldsymbol{\delta}$, m)
    Apply MSA algorithm 1 to find SUE route flows i.e. $\hat{\boldsymbol{x}}$
    Select initial states by replacing SUE for first m days
    Compute utility $u_r^{t-1} = \delta_1 C_r(\boldsymbol{x}_r^{t-1}) + \delta_2 C_r(\boldsymbol{x}_r^{t-2}) + \ldots + \delta_m C_r(\boldsymbol{x}_r^{t-m})$
    for $t = 1$ to Nrun **do**
        $p_r(\boldsymbol{x}^{(t-1)}) = \frac{\exp\{-\theta u_r^{(t-1)}\}}{\sum_r \exp\{-\theta u_r^{(t-1)}\}}$,
        $\boldsymbol{x}_j^{(t)} | \boldsymbol{u}^{(t-1)} \sim \text{Multinomial}(N_j, \boldsymbol{p}_j^{(t-1)})$
        Update utility for last m days
        $t \leftarrow t + 1$
    end for
    return Route flows for last m days
end function

The R code for our new definition of the coefficient of reactivity is provided in Appendix B.

Algorithm 3 coefficient of reactivity

Input: network information, θ, $\boldsymbol{\delta}$, m
Output: value of coefficient of reactivity

function COEFFICIENT.REACTIVITY(network information, θ, $\boldsymbol{\delta}$, m)
  for $k \leftarrow 1$ to Nh **do**
    Recall history.flow function
    store last m days $\boldsymbol{x}^{t-(m-1)}, \boldsymbol{x}^{t-(m-2)}, \ldots, \boldsymbol{x}^{t-1}$
    $k \leftarrow k + 1$
  end for
  for $\mathtt{h} \leftarrow \mathtt{1}$ to Nff for each route **do**
    Set $\boldsymbol{x}^t = \boldsymbol{x}[\mathtt{h}]$
    for $i \leftarrow 1$ to Nh **do**
      $\boldsymbol{x}new \leftarrow \boldsymbol{x}^{t-(m-1)}, \boldsymbol{x}^{t-(m-2)}, \ldots, \boldsymbol{x}^{t-1}, \boldsymbol{x}^t$
      Compute utility $u_r^{t-1} = \delta_1 C_r(\boldsymbol{x}_r^{t-1}) + \delta_2 C_r(\boldsymbol{x}_r^{t-2}) + \ldots + \delta_m C_r(\boldsymbol{x}_r^{t-m})$
      for $l \leftarrow 1$ to Ni **do**
        $p_r(\boldsymbol{x}^{(t-1)}) = \frac{\exp\{-\theta u_r^{(t-1)}\}}{\sum_r \exp\{-\theta u_r^{(t-1)}\}}$,
        $\boldsymbol{x}_j^{(t)} | \boldsymbol{u}^{(t-1)} \sim \text{Multinomial}(N_j, \boldsymbol{p}_j^{(t-1)}) \quad j = 1, \ldots, h$
        Update utility for last m days
        $l \leftarrow l + 1$
        Return instance $x^t[l]$
      end for
      $\bar{\boldsymbol{x}} = \frac{1}{Ni} \sum_l \boldsymbol{x}^{(t)}[l]$
      $i \leftarrow i + 1$
    end for
    Average of all histories i.e. $\psi(\hat{\boldsymbol{x}}[\mathtt{h}]) = \frac{1}{\mathtt{Nh}} \sum_{\mathtt{i}} \bar{\boldsymbol{x}}[\mathtt{i}]$
    calculate $\dfrac{\|\psi_m(\boldsymbol{x}) - \hat{\boldsymbol{x}}\|}{\|\boldsymbol{x} - \hat{\boldsymbol{x}}\|}$
    $\mathtt{h} \leftarrow \mathtt{h} + \mathtt{1}$
  end for
  return $\omega^{new} = \max_{\mathtt{h}} \dfrac{\|\psi(\boldsymbol{x}[\mathtt{h}]) - \hat{\boldsymbol{x}}\|}{\|\boldsymbol{x}[\mathtt{h}] - \hat{\boldsymbol{x}}\|}$
end function

We continue to use the two route network from the previous section, but now we set the travel demand $N = 3$ and the memory length to $m = 2$. We consider day 31 to be the time at which the intervention is applied. For the purposes of illustration we consider just $Nh = 2$ histories and $Ni = 2$ instances per history. The simulations required to estimate the coefficient of reactivity are displayed in Figures 5.4-5.7. Each plot corresponds to a particular choice of x^t (indexed by h in the algorithm above). In each Figure there are two panels, corresponding to each of the two histories. Each panel depicts the flow on route 1. The flows in the top panel of each Figure are identical for the first 30 days, since they are all based on the same history. Likewise, the flows on the bottom panels are all the same for the first 30 days, since they all correspond to the second history. The flow pattern on day 31 reflects the value of x^t for the Figure in question. For example, in Figure 5.4 we consider the case where $x_1^t = 0$, so the flow on route 1 is set to zero in both panels on day 31. Finally, we plot two instances for day 32 in each panel, distinguished by plotting colour. The value of the estimate of $\psi(x)$ for $x^t = x$ is obtained by averaging the instances on day 32 over both panels in each Figure.

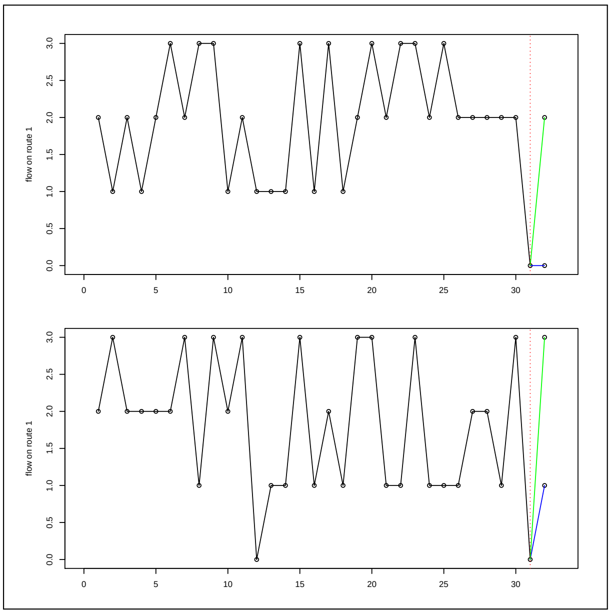


Figure 5.4: Simulated flows on route 1 corresponding to $\psi(\begin{bmatrix} 0 \\ 3 \end{bmatrix})$.

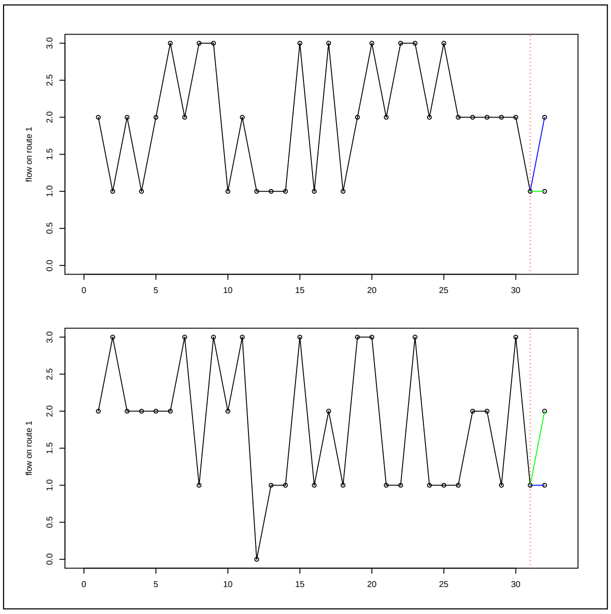


Figure 5.5: Simulated flows on route 1 corresponding to $\psi(\begin{bmatrix} 1 \\ 2 \end{bmatrix})$.

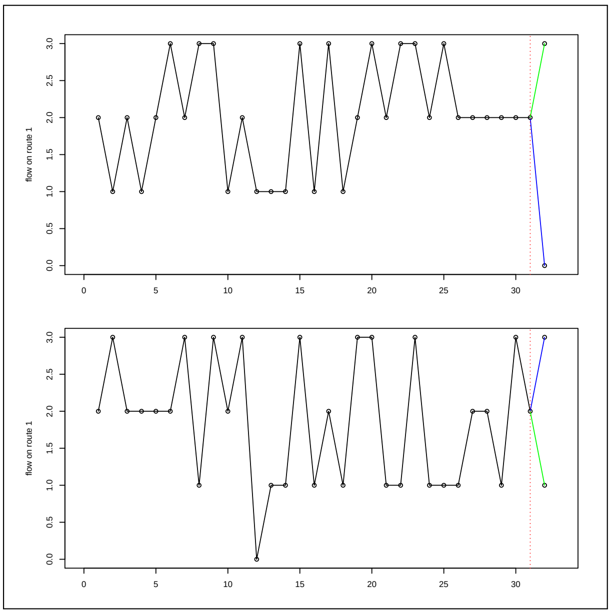


Figure 5.6: Simulated flows on route 1 corresponding to $\psi(\begin{bmatrix} 2 \\ 1 \end{bmatrix})$.

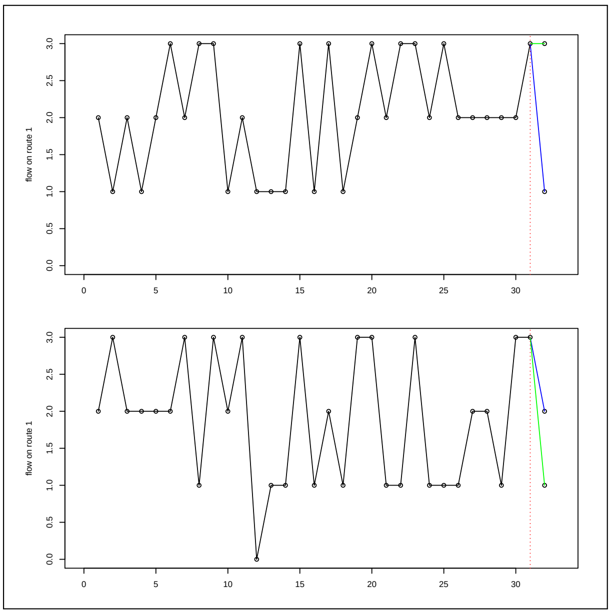


Figure 5.7: Simulated flows on route 1 corresponding to $\psi(\begin{bmatrix} 3 \\ 0 \end{bmatrix})$.

5.3.3 Simulation result for new definition of coefficient of reactivity

Considering the same route cost function as 2.16, we suppose that only the previous three days memory of 100 travellers impact on their decision to select their route for today. We assume that a traveller gives 50% weight to what happened yesterday, 30% weight to what happened on the day before yesterday and 20% for three days ago. These values correspond to our $\boldsymbol{\delta}$ parameter and then $u_r^{(t-1)} = 0.5c_r(x^{(t-1)}) + 0.3c_r(x^{(t-2)}) + 0.2c_r(x^{(t-3)})$. The evolution of flows are simulated over 100 days. The results are shown below.

To illustrate the behaviour of the model, Figures 5.8-5.10 display the proportion of flows on route 1. The asymptotic coefficient of reactivity which depends on the definition of the model is computed. We have computed the new coefficient of reactivity using the simulation of histories and instances. Considering the same parameters gives us different values for the asymptotic and the new definition of the coefficient of reactivity. As can be seen in Figure 5.8 disruption in the network does not affect the system for the day after a disruption and the behaviour of the system is similar to flow behaviour before the unusual event happens to the network. This happens as the value of the coefficient of reactivity is small from both definition.

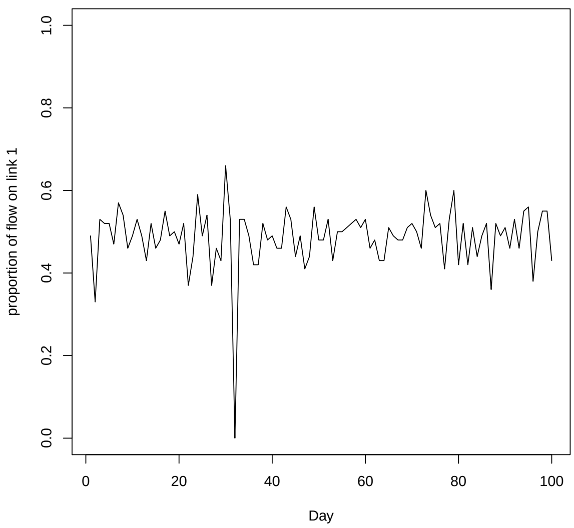


Figure 5.8: Route 1 flows simulated over 100 days as a fraction of total demand based on logit model with $\omega_0 = 0.16$ and $\omega_0^{new} = 0.32$.

We know that for the network with a higher value of the coefficient of reactivity, there will always be fluctuation. The value of the asymptotic coefficient of reactivity corresponds to the plot depicted in Figure 5.9 is 0.8 and we see a reaction for the day after the disruption and the proportion of travellers using route 1 is around 0.8. However, the value of the coefficient of reactivity computed through the new definition is 1.03. As the value hits 1 we might expect to see more variations in flow pattern after the disruption day. But as described before the value of the coefficient of reactivity from the new definition is based on different histories and instances. Each time we run the simulation we obtain a different value for the coefficient of reactivity. So if we repeated the simulation several times, then there is a possibility to get a value less than one for the coefficient of reactivity.

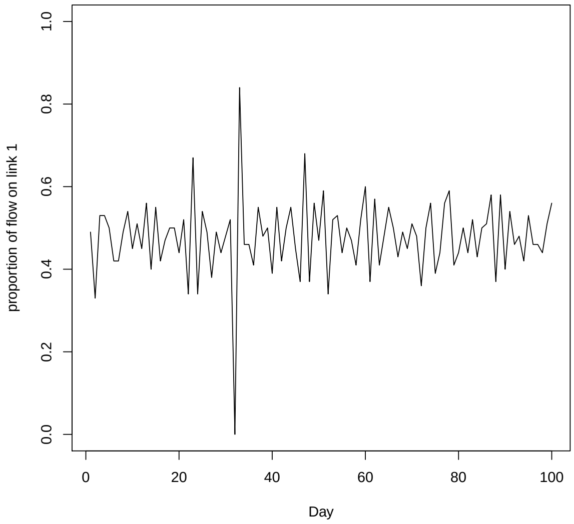


Figure 5.9: Route 1 flows simulated over 100 days as a fraction of total demand based on logit model with $\omega_0 = 0.8$ and $\omega_0^{new} = 1.03$.

Figure 5.10 exhibits the same behaviour as Figure 5.9, the difference is that here the system is more reactive because the value of the coefficient of reactivity is higher either for the old and the new ones. As a result, it will have a greater impact on travelers' behaviour. As shown in Figure 5.10 the proportion of travellers' using route 1 in response to the unusual event on the network goes around 0.9 for the days after the disruption. Here the value of the coefficient of reactivity for both definitions are higher than 1.

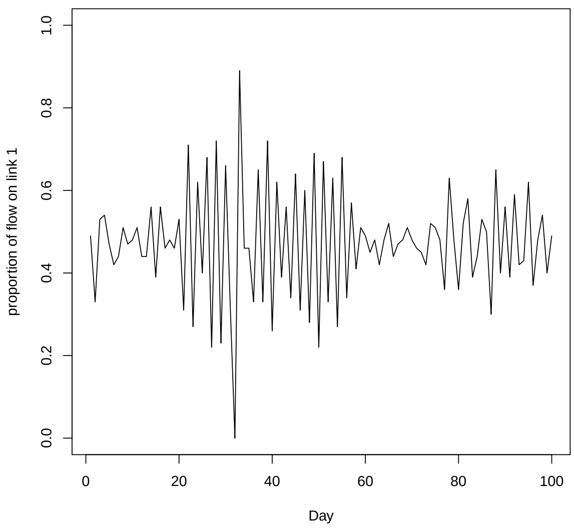


Figure 5.10: Route 1 flows simulated over 100 days as a fraction of total demand based on logit model with $\omega_0 = 1.19$ and $\omega_0^{new} = 1.92$.

Through these simulations we have not seen big differences between the original and the new definition of the coefficient of reactivity. In the following we theoretically demonstrate the results of these two definition will be identical for large travel demand.

5.3.4 Theoretical comparison

We saw previously that direct calculation of the new coefficient of reactivity is not feasible except in tiny examples. The simulation-based method of computation can be applied to larger problems, but is still feasible only for systems with a small number of routes and low demand.

Let $\boldsymbol{N}$ denote the vector of demands for the network. We will allow the travelling population to become arbitrary large, so it is adequate to specify the demands by $\boldsymbol{N}=\zeta\boldsymbol{N}_0$ where ζ is a scalar and $\boldsymbol{N}_0$ denoting the fixed vector of demand. We are interested to show that the coefficient of reactivity with both $\psi_m(\boldsymbol{x})$ and $\psi_m^{new}(\boldsymbol{x})$ when $\zeta \to \infty$ are the same. In this limiting regime, costs are based on the normalized flows $\boldsymbol{x}/\zeta$.

Theorem 2: *Assume the system has a unique stochastic user equilibrium then as* $\zeta \to \infty$ *we have* $\omega_0 \to \omega_0^{new}$.

According to the definition of coefficient of reactivity it is sufficient to show that $\psi_m \to \psi_m^{new}$. Before proving the theorem we need the following well known lemma, a proof is provided by Polansky (2011), for example.

Lemma: *Let* X_n *be a sequence of random variables, such that* $X_n \xrightarrow{\mathbb{P}} X$. *If* $g : \boldsymbol{R} \mapsto \boldsymbol{R}$ *is a continuous mapping,then* $g(X_n) \xrightarrow{\mathbb{P}} g(X)$.

In order to prove our theorem, let us define

$$g_{\boldsymbol{x}}(\boldsymbol{x}^{(t-2)}, \ldots, \boldsymbol{x}^{(t-m)}) = \mathsf{E}[\boldsymbol{x}^{(t)}|\boldsymbol{x}^{(t-1)} = \boldsymbol{x}, \boldsymbol{x}^{(t-2)}, \ldots, \boldsymbol{x}^{(t-m)}].$$

As stated in Hazelton (2002) when the flows from day $t-2$ to day $t-m$ are equal to the stationary mean, then $g_{\boldsymbol{x}}(\boldsymbol{x}^{(t-2)}, \ldots, \boldsymbol{x}^{(t-m)}) = \psi(\boldsymbol{x})$. From Hazelton and Watling (2004), $\boldsymbol{x}^{(t)} = \boldsymbol{x}^* + O(\zeta^{-1/2})$ for all t, when process follows stationary distribution. Hence $g(\boldsymbol{x}^{(t-2)}, \ldots, \boldsymbol{x}^{(t-m)}) = g(\boldsymbol{x}^*, \ldots, \boldsymbol{x}^*) + O(\zeta^{-1/2})$ by the lemma. Therefore,

$$\psi^{new}(\boldsymbol{x}) = \mathsf{E}[g(\boldsymbol{x}^{(t-2)}, \ldots, \boldsymbol{x}^{(t-m)})] = g(\boldsymbol{x}^*, \ldots, \boldsymbol{x}^*) + O(\zeta^{-1/2}) = \psi(\boldsymbol{x}) + O(\zeta^{-1/2}).$$

Theorem result then follows.

The new coefficient of reactivity is very difficult to compute. However, we have seen that the original coefficient of reactivity provides a good approximation to the new definition when travel demand is a least moderately large. This suggests we continue to work with original definition in practice.

5.4 A generalization of the original coefficient of reactivity

The original definition of coefficient of reactivity is in terms of disruption in the system for one day, for example road closure. However, we might have disruption for a longer duration. Therefore, the aim of this section is to extend the Hazelton (2002) definition where the length of the disruption can be arbitrarily long. Here, we define the disruption length using a parameter r. Let us suppose the system not only has been disturbed for one day, but for r days. The definition of $\psi_m(\boldsymbol{x})$ in equation 4.7 can be replaced by

$$\psi_{m,\mathrm{r}}(\boldsymbol{x}) = \mathsf{E}[\boldsymbol{x}^{(t)}|\boldsymbol{x}^{(t-1)} = \boldsymbol{x}, \ldots, \boldsymbol{x}^{(t-\mathrm{r})} = \boldsymbol{x}, \boldsymbol{x}^{(t-\mathrm{r}-1)} = \boldsymbol{x}^*, \ldots, \boldsymbol{x}^{(t-m)} = \boldsymbol{x}^*], \tag{5.6}$$

where r is related to the unusual behaviour of the system. The generalized coefficient of reactivity is given by

$$\omega^{\mathrm{r}} = \sup_{\boldsymbol{x}\in\mathcal{X}} \frac{\|\psi_{m,\mathrm{r}}(\boldsymbol{x}) - \boldsymbol{x}^*\|}{\|\boldsymbol{x} - \boldsymbol{x}^*\|}. \tag{5.7}$$

Also, it should be noted that the asymptotic version will be described in the corresponding way:

$$\omega_0^{\mathrm{r}} = \lim_{\epsilon\to 0} \sup_{\boldsymbol{x}\in\mathcal{X}_\epsilon} \frac{\|\psi_{m,\mathrm{r}}(\boldsymbol{x}) - \hat{\boldsymbol{x}}\|}{\|\boldsymbol{x} - \hat{\boldsymbol{x}}\|}.$$

Obviously $\psi_m(\boldsymbol{x})$ is a special case of $\psi_{m,\mathrm{r}}(\boldsymbol{x})$ when $\mathrm{r} = 1$. But in fact there is more connection between the original and generalized definition of coefficient of reactivity.

Theorem 3: *For the m-memory models described in Section 5.2, for $r \geq m$ $\psi_{m,r}(\boldsymbol{x}) = \psi_{1,1}(\boldsymbol{x})$.*

Proof: Let us define a set of models indexed by m; say $\{M(m) : m = 1, 2, 3, \ldots\}$. All models have the same network, demand and route choice probability model, with the probability of choosing route r on day t given by a logit model

$$\frac{\exp(u_r^{(t-1)})}{\sum_s \exp(u_s^{(t-1)})}.$$

However, for model $M(m)$ the utility is given by $\boldsymbol{u}^{t-1} = \delta_1\boldsymbol{C}(\boldsymbol{x}^{(t-1)}) + \delta_2\boldsymbol{C}(\boldsymbol{x}^{(t-2)}) + \ldots + \delta_m\boldsymbol{C}(\boldsymbol{x}^{(t-m)})$. Now we want to show that $\psi_{m,\mathrm{r}}(\boldsymbol{x})$ for model $M(m)$ with $\mathrm{r} \geq m$ is the same as $\psi_{1,1}(\boldsymbol{x})$ for model $M(1)$.

The inequality $\mathrm{r} \geq m$ indicates that flows on day $t-1$ to day $t-m$ are set to be at $\boldsymbol{x}$. So, the definition of $\psi_{m,\mathrm{r}}(\boldsymbol{x})$ from 5.6 is then given by

$$\psi_{m,\mathrm{r}}(\boldsymbol{x}) = \mathsf{E}[\boldsymbol{x}^{(t)}|\boldsymbol{x}^{(t-1)} = \boldsymbol{x}, \ldots, \boldsymbol{x}^{(t-m)} = \boldsymbol{x}].$$

Since $\psi_{m,\mathrm{r}}(\boldsymbol{x})$ depends on the utility, therefore, we have $\boldsymbol{u}^{(t-1)} = \delta_1\boldsymbol{C}(\boldsymbol{x}) + \delta_2\boldsymbol{C}(\boldsymbol{x}) + \ldots + \delta_m\boldsymbol{C}(\boldsymbol{x})$. Consequently, $\boldsymbol{u}^{(t-1)} = \boldsymbol{C}(\boldsymbol{x})(\delta_1 + \delta_2 + \ldots + \delta_m)$. Since $\delta_1 + \delta_2 + \ldots + \delta_m = 1$ as a result, $\boldsymbol{u}^{(t-1)} = \boldsymbol{C}(\boldsymbol{x})$. This is equal to the utility employed when computing $\mathsf{E}[\boldsymbol{x}^{(t)}|\boldsymbol{x}^{(t-1)} = \boldsymbol{x}]$ with model M(1). Hence $\psi_{m,\mathrm{r}}(\boldsymbol{x}) = \psi_{1,1}(\boldsymbol{x})$ for all $\boldsymbol{x}$, completing the proof.

As shown through the definition of $\psi_{m,\mathrm{r}}$ (5.6) we are able to extend the length of disruption. Here, with an example, we demonstrate the relationship between the value of the coefficient of reactivity by applying the definition 5.6 and the travellers' route choice behaviour. We examine this issue when the disruption happens to the network for more than one day both in a mathematical way and through the simulations. To do this, we consider our simple two route network 2.7.1 with assumptions from Example 4.3.1, but here we assume that travellers make their decision for today based on the cost they experienced in three previous days.

In such a case, the utilities are given as

$$u_1^{t-1} = \delta_1 \left(a_1 + \left(\frac{x_1^{t-1}}{b} \right)^2 \right) + \delta_2 \left(a_1 + \left(\frac{x_1^{t-2}}{b} \right)^2 \right) + (1 - \delta_1 - \delta_2) \left(a_1 + \left(\frac{x_1^{t-3}}{b} \right)^2 \right)$$

and

$$u_2^{t-1} = \delta_1 \left(a_2 + \left(\frac{x_2^{t-1}}{b} \right)^2 \right) + \delta_2 \left(a_2 + \left(\frac{x_2^{t-2}}{b} \right)^2 \right) + (1 - \delta_1 - \delta_2) \left(a_2 + \left(\frac{x_2^{t-3}}{b} \right)^2 \right)$$

where δ_1 and δ_2 are the weight attributes assigned to each day. Taking similar steps as in Example 4.3.1, it is easy to show that for $\psi_{3,1}(\boldsymbol{x}) = E[\boldsymbol{x}^{(t)} | \boldsymbol{x}^{(t-1)} = \boldsymbol{x}, \boldsymbol{x}^{(t-2)} = \boldsymbol{x}^*, \boldsymbol{x}^{(t-3)} = \boldsymbol{x}^*]$, then the asymptotic generalized coefficient of reactivity is given by

$$\omega_0^1 = 2\theta\delta_1 \left(\frac{\hat{x_1}}{b} \right) \left(\frac{\hat{x_2}}{b} \right). \tag{5.8}$$

In the same way, when we have disruptions for two days then ψ can be written as $\psi_{3,2}(\boldsymbol{x}) = E[\boldsymbol{x}^{(t)} | \boldsymbol{x}^{(t-1)} = \boldsymbol{x}, \boldsymbol{x}^{(t-2)} = \boldsymbol{x}, \boldsymbol{x}^{(t-3)} = \boldsymbol{x}^*]$. As a consequence,

$$\omega_0^2 = 2\theta(\delta_1 + \delta_2) \left(\frac{\hat{x_1}}{b} \right) \left(\frac{\hat{x_2}}{b} \right). \tag{5.9}$$

Finally, when the duration of disruption days is the same as the length of travellers memory then, $\psi_{3,3}(\boldsymbol{x}) = E[\boldsymbol{x}^{(t)} | \boldsymbol{x}^{(t-1)} = \boldsymbol{x}, \boldsymbol{x}^{(t-2)} = \boldsymbol{x}, \boldsymbol{x}^{(t-3)} = \boldsymbol{x}]$ and

$$\omega_0^3 = 2\theta \left(\frac{\hat{x_1}}{b} \right) \left(\frac{\hat{x_2}}{b} \right), \tag{5.10}$$

which is exactly equal to the value of coefficient of reactivity from the original definition based on a model with $m = 1$ day memory.

5.5 Simulation results for the generalized coefficient of reactivity

So far, our simulation results illustrated how the coefficient of reactivity works when travellers' have one day memory and one day disruption and also longer memory with one day disruption in the network. Here, we demonstrate a few simulations when we have disruption for more than one day. The route flow evolution patterns based on three days memory and two days disruption are shown in Figures 5.11, 5.12 and 5.13. For these plots we assume that the unusual events happen on day 40 and 41 for route 1. The most important feature that can be captured through these plots is for more than one day disruption coefficient of reactivity works exactly in the same way of one day disruption. The result for a system with very low value of coefficient of reactivity 0.1 is shown in Figure 5.11.

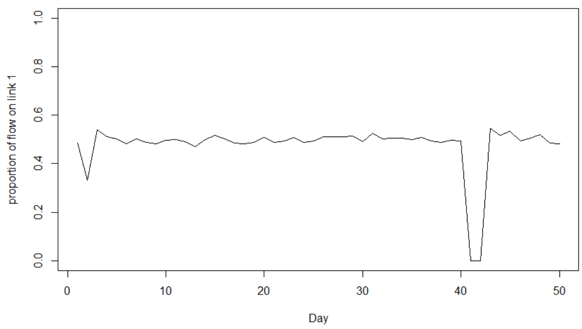


Figure 5.11: Route 1 flows simulated over 50 days as a fraction of total demand with $\omega_0^2 = 0.1$.

We can see that travellers' are not going to switch to an alternative route. As shown by Figure 5.12, when the value of coefficient of reactivity is 0.5 there is a little reaction on the day immediately following the disruption. The proportion of travellers' using route 1 is around 0.7. Figure 5.13 shows a very big immediate impact on travellers behaviour with higher value of coefficient of reactivity 0.9.

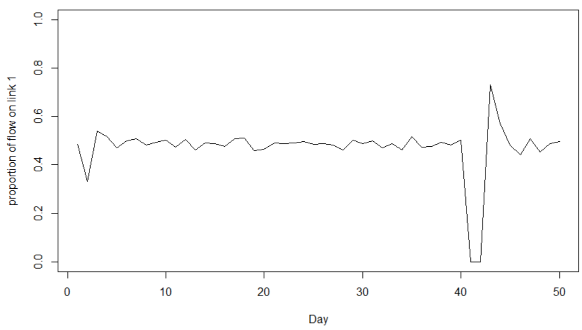


Figure 5.12: Route 1 flows simulated over 50 days as a fraction of total demand with $\omega_0^2 = 0.5$.

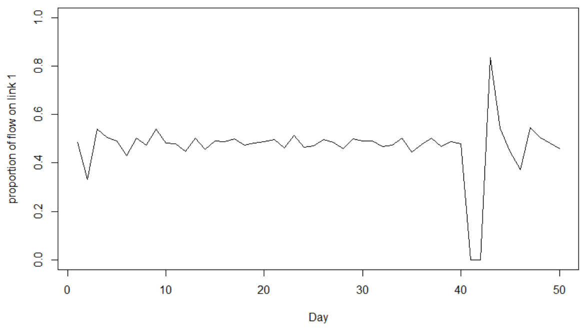


Figure 5.13: Route 1 flows simulated over 50 days as a fraction of total demand with $\omega_0^2 = 0.9$.

Now, to show why it is important to consider the generalized definition of the coefficient of reactivity, we examine the travellers' reaction to the disruption in our two route network through some simulations by looking at two different models. For the first model we set the length of memory $m = 1$ and for the second one we assume $m = 3$. The route choice probability is determined by logit model. We assign $\boldsymbol{\delta} = (0.4, 0.3, 0.3)$ to the utility function for the second model. To find the value of the coefficient of reactivity and for the simulation results we set $\theta = 0.06$ for the first model and $\theta = 0.15$ for the second one. Table 5.1 shows the values of the coefficient of reactivity for the different scenarios. As can be seen the value

m	r	ω_0^{r}
1	1	0.48
1	2	0.48
1	3	0.48
3	1	0.48
3	2	0.83
3	3	1.19

Table 5.1: The value of generalized coefficient of reactivity corresponding to the different values of m and r.

of coefficient of reactivity ω_0^1 for both models is the same. Therefore we would expect to see similar behaviour (based on the simulation) after one day's disruption. This is exhibited in Figure 5.14

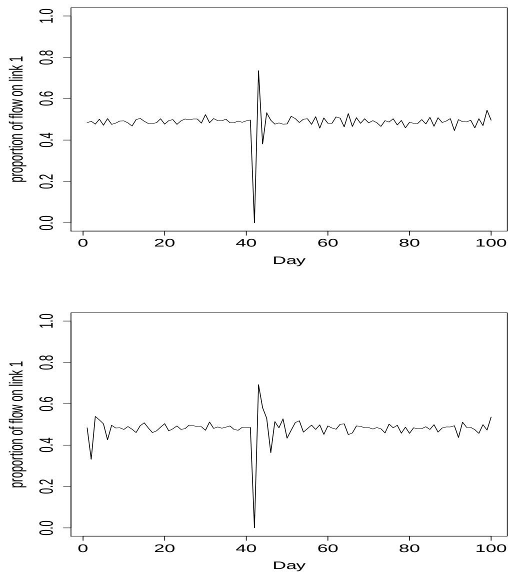


Figure 5.14: Simulation of memory length $m = 1$ (top) and memory length $m = 3$ (bottom) Markov models when the system undergoes a disruption of length 1 day.

However, the situation is not the same when the duration of the disruption is more than one day. Figures 5.15 and 5.16 show different behaviour for the day after the disruption days. In particular, as depicted in Figure 5.16, for disruption length r = 3, model 1 shows a mild reaction for the day after disruptions and the proportion of travellers using route 1 is around 0.6 and the model stabilizes quite quickly. But for the second model the system becomes more reactive that is, it can be seen the proportion of travellers goes around 0.9 when the memory length is $m = 3$.

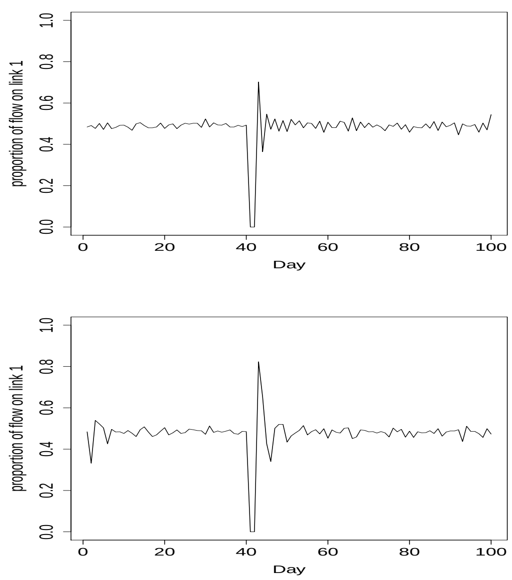
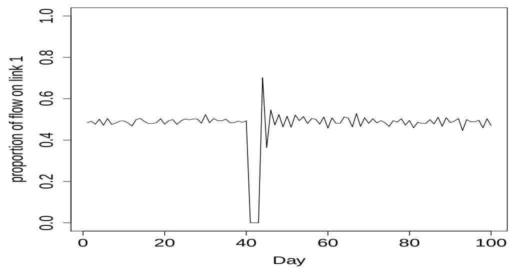


Figure 5.15: Simulation of memory length $m = 1$ (top) and memory length $m = 3$ (bottom) Markov models when the system undergoes a disruption of length 2 days.



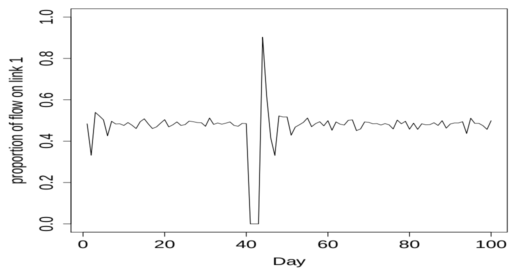


Figure 5.16: Simulation of memory length $m = 1$ (top) and memory length $m = 3$ (bottom) Markov models when the system undergoes a disruption of length 3 days.

The original definition of the coefficient of reactivity cannot distinguish the difference between the two models in terms of reactivity when the length of disruption increases. The original definition of coefficient of reactivity indicates that both models are equally reactive. Specifically, we see that $\omega_0^1 = 0.48$ for both models. However, our simulation shows that the second model is far more reactive to long disruptions. This is correctly reflected in the generalized coefficient of reactivity for r = 3, where we have $\omega_0^3 = 0.48$ for the first model but $\omega_0^3 = 1.19$ for the second model. This demonstrates the usefulness of the generalized coefficient of reactivity.

5.6 Conclusion

Disruptions to transportation networks cause travel delays and decreased transportation efficiency. A good understanding of behavioral reactions to such incidents is crucial for traffic management and planning. Depending on the impact and effect of disruptions travellers may change their attitudes towards travel and transport. In the case of short-term disruption, travellers may change their travel mode or choose alternative travel routes for a few days and return back to their routine afterward. Long term disruptions may have effects that last for a long time, perhaps leading to more permanent changes for traveller choice behaviour.

In order to analyse traveller behaviour and day-to-day flow variation after network disruption, the coefficient of reactivity has been introduced by Hazelton (2002). We found that coefficient of reactivity (ω_0) plays an important role in the stability of traffic evolution when travellers make their decisions based on one day memory. As we have shown for the lowest value of coefficient of reactivity the system quickly returns back to normal flow pattern, but for a higher value of coefficient of reactivity ($\omega_0 > 1$) we have a more drastic change in traveller behaviour for the days after disruption.

Following that, we have generalized the idea from one day to multiple days of disruption. We found that the coefficient of reactivity still shows a very big immediate impact on traveller behaviour and tells us the size of this. More precisely, the coefficient of reactivity measures how extreme the initial reaction of travellers is for the day after disruption. While the original definition only looks at short-term disruptions, the generalized coefficient of reactivity can deal with longer durations of disruption. Hence, the generalized coefficient of reactivity can help us to understand behaviour of the system better than original definition. One important feature we can observe from the figures is that as the value of ω_0^r increases the route flow on affected route changes more significantly and 1 as a value of coefficient of reactivity still is the critical cut-off.

All the examples that we have studied in Chapters 4 and 5 have related to systems with a unique deterministic equilibrium. We examined multi-equilibria systems in some detail in Chapter 3. However, further study using the coefficient of reactivity is not practically feasible. This is because the only way we can calculate the value of the coefficient of reactivity is through the SUE. As stated in Chapter 2 for a system with multiple equilibria the SUE solution is not unique, therefore, we have not provided an example for the network with multiple equilibria that requires the calculation of the value of the coefficient of reactivity.

Overall, the coefficient of reactivity is useful in describing the stability of a system after disruption. However, because it is a single number summary, one would expect there to be cases where it fails to fully characterise the stability. This can happen with systems with long memory, depending on weights applied in the utility. Looking at results for different values of r can help. Of course, we then no longer have a single summary of the stability of the system, but rather a collection of values indexed by the disruption length. Whether these can be combined usefully into a single value remains an open question.

Chapter 6

Conclusion

6.1 Summary of Thesis

This chapter summarises the major finding and contributions in this thesis. Also, possible future extensions are highlighted.

In the real world, traffic flows vary from one day to the next. Therefore, day-to-day traffic modelling has been suggested where the goal is different from the static traffic assignment model, which distributes traffic flow in a network such that a predefined goal, e.g., user equilibrium, is achieved. Day-to-day modelling approach is useful for long term transportation planning purposes and comes in two varieties, deterministic and stochastic models. As stated in Chapter 1 the main objective proposed for this study was investigating the relationship between deterministic and stochastic day-to-day traffic assignment models. We wanted to investigate how the long term transitional properties of the system differ between deterministic and stochastic models, and more particularly, when deterministic models can be a good approximation for the average of stochastic models.

Our first step was to review the basic concepts, notations and techniques from transportation and stochastic processes in Chapter 2. Those methods and techniques have been used in the following chapters and helped us to achieve our desired goals. Chapter 3 examined the travellers route choice behaviour for a system with multiple equilibria. In transportation literature such as Davis and Nihan (1993), Cantarella and Cascetta (1995) and Watling and Cantarella (2013), the focus was on the network with single deterministic equilibrium. All these works show, in a network with a single equilibrium, when travel demand and link capacities tend to infinity in tandem, then stochastic day-to-day models will converge to corresponding deterministic models.

We found that the situation for systems with multiple equilibria is not the same and is more complicated. By examining two different properties of a number of systems we discovered that the mean of stochastic models is not a helpful summary for those kinds of systems. On the one hand, as shown in Section 3.3 for our simple two route network with multiple equilibria the stationary distribution of the stochastic model was bimodal where the most probability weight was placed near the extreme flows. This was because of the existence of two stable equilibria in the deterministic model.

On the other hand, we also examined the transition time between various states using the mean hitting time concept. We observed for the deterministic model in our two route network if we push a system a little bit then the system quickly converged towards the desired equilibrium (e.g. every one use the bus). But for the stochastic model as shown in Table 3.1 the first passage time distribution has a very long tail. Therefore, the time plot in Figure 3.4 illustrated the long-run simulation where the route flow was at one equilibrium for a long period, then through a flip went to another equilibrium. That was the failure of considering deterministic models as an approximation of mean behaviour of stochastic models.

We also applied the idea for a network with multiple OD pairs. We have seen in Figure 3.10 a flow behaviour from stochastic model that deterministic model is not able to create anything that can be comparable. Our findings in this chapter indicated that rather than the mean, the result from the deterministic model can be a good approximation to show the average behaviour of the stochastic model in many cases. However, it was highlighted that comparing the deterministic and stochastic models for a system with multiple equilibria needs considerable care.

Investigating the properties of the networks with multiple equilibria has got implications for network control. While in a traditional deterministic model with certainty we can move the system towards the desired equilibrium, in a stochastic model it is not necessarily the case and it is more complicated. We can attempt to reach a desired equilibrium by making some changes and one of these changes can be congestion pricing.

One of the limitations of day-to-day modelling is that it won't necessarily reflect detailed differences between control strategies. So for example, subtle changes to congestion charges through the course of a day can only be represented in an approximate way. Also day-to-day modelling it is ill suited to time-critical problems, like swift changes to traffic patterns due to critical hazards. A practical traffic management tool to protect urban road networks from over-saturation in, for example, disaster response is gating control (Keyvan-Ekbatani et al. 2012). The mutual impact of gating control and traffic assignment models has been shown in Bu et al. (2019). They demonstrated the effectiveness of a gating control strategy for traffic operations in emergency management to obtain improved traffic assignments through a nonlinear programming optimization model.

From a traffic management perspective, after any network disruption (short term or long term) it is really important to predict the evolution of day-to-day traffic flow. Essentially, they want to know how travellers will react to the disruption and how traffic flow evolves from a disequilibrium state to a new equilibrium state. At first, in Chapter 4, we investigated the reaction of travellers for the day after a network disruption through a measure called coefficient of reactivity presented by Hazelton (2002). Since this definition only works for short term disruptions, the whole idea for this chapter was based on the assumption that we have one day disruption in the system and travellers make their decision based on one day memory.

Considering that the main focus of this research has been to explore the relationship between stochastic and deterministic day-to-day models, in the second step of Chapter 4, we examined this issue with regard to the value of coefficient of reactivity. This chapter contained the two stochastic (Davis and Nihan 1993), (Watling and Hazelton 2018) and one deterministic (Cantarella and Cascetta 1995) approximation methods derived for the simple traffic generating model introduced in Chapter 2. Comparison based on stability of traffic flows has been done through simulations for different networks from a simple two route network with single OD pair to the networks with multiple OD pairs and many different routes.

To see the transient dynamics of the system, for our experiments in this chapter three different initial flow patterns were consider; starting point at SUE, fairly close to SUE and extremly away from SUE. Our finding indicated that for all networks when the value of the coefficient of reactivity is small enough (i.e. less than one) irrespective of the initial state, all approximation methods settled down at their stationary distribution. However, the first few days showed the transient behaviour of the system when the initial state is not at SUE.

When the value of the coefficient of reactivity hits 1, then for small networks the failure of stochastic approximation methods was seen in the prediction intervals even for simulation over a short period of time. But for larger networks, the situation was a bit different. As we have seen in experiments 4.4.5 and 4.4.6 the approximation methods work well when the value of coefficient of reactivity is a bit higher than 1. At least those experiments demonstrated increasing the value of the coefficient of reactivity leads to an increase in the period of the transient behaviour of the system. Overall, the fine detail of how well the approximations work alongside the value of the coefficient of reactivity will depend on starting values, details of the network (size of the network, the link cost function, etc.) and so on.

As stated in Chapter 4 the original definition of coefficient of reactivity only works when the disruption happens in the system only for one day. Extending this definition was the main idea for Chapter 5 and has been done in two directions. First, in contrast to Hazelton's 2002 definition where the flows on day $t-2$ and before are set to the stationary mean, in our new definition we allowed the system to have variation for those days, which seems more realistic. However, we have shown that working with this definition is not easy due to the fact that it depends on the joint distribution of flow patterns between days $t-2$ and $t-m$. Hence, we approximated this definition through simulation. We also have seen that the implementation of the algorithm presented in section 5.3.2 is difficult. However, we proved that when travel demand is large then the original coefficient of reactivity is a good approximation for the new definition.

Consequently, since the calculation of the value of coefficient of reactivity is easier with the original definition, the second direction has been motivated. We modified the original definition by increasing the length of the disruption to r days. This revised version can be applied to any model and when the length of the disruption is the same as travellers' memory length then it is exactly equivalent to the new definition of coefficient of reactivity. We saw that while the original definition of the coefficient of reactivity can fail to distinguish between models with very different behaviour after a lengthy disruption, the generalized definition is able to recognize these differences.

It should be pointed out that we have spent a lot of time looking at the differences between the deterministic and stochastic models but we have not made any statement about which are more realistic. The reason for that is models depend on parameters, and the performance in terms of approximation of the real system depends on how well the model behaves when we estimate those parameters. More complex models have the potential for greater realism, but the inclusion of extra parameters makes them harder to properly calibrate. Instead this thesis has focussed on comparing the properties of models. Improving our understanding of those differences will hopefully help others to recognize the likely strengths and weaknesses of these models when applied in practice, and should provide a foundation for future research.

6.2 Future research

As we have discussed in Chapter 3 networks with multiple equilibria have not received much attention in the transportation literature. We investigated some interesting properties of stochastic day-to-day dynamic models. Hence, it would be attractive to study the day-to-day dynamic in stochastic models with multiple deterministic equilibria. For instance, developing a new day-to-day model and novel theoretical approaches can be useful.

We studied the relationship between the value of the coefficient of reactivity and the approximation methods in Chapter 4. However, we only considered the memory length and the duration of the disruption to be 1. Further research can explore the relationship when the length of both of them can be arbitrarily large.

In Chapter 5 we saw the generalized coefficient of reactivity can be useful in distinguishing between systems after prolonged disruptions. However, consider Figures 6.1 and 6.2, in which we plot simulated flows for two models when there is just a one day disruption. The coefficient ω_0^1, is the same for both models, but it is clear from the plots that the first system is much more reactive than the second. In principle we could learn more from this example by computing the generalized coefficients of reactivity for longer disruptions, since these would confirm differences in reactivity for the two models. However, we have then begun to lose sight of Hazelton's (2002) original aim, which was to provide a single number to describe the reactivity of a system. A possible research direction is to see whether we can combine values of the generalized coefficient of reactivity over different lengths of disruption to get some kind of aggregate value.

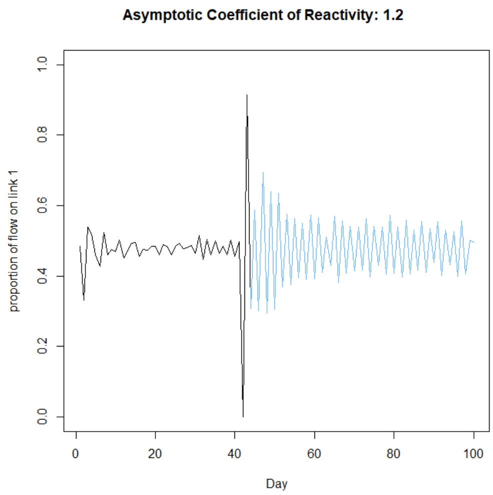


Figure 6.1: Route 1 flows simulated over 100 days as a fraction of total demand with $\theta = 0.3$ and $\delta = (0.5, 0.3, 0.2)$.

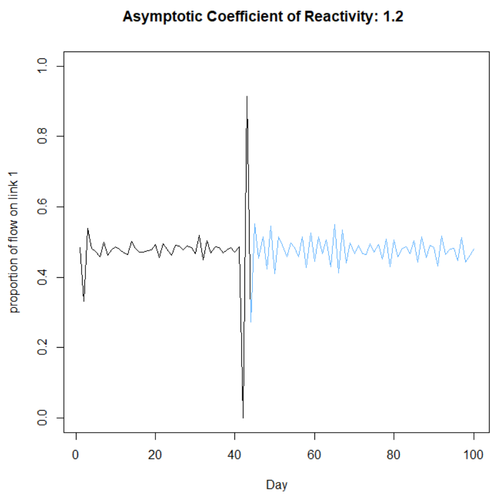


Figure 6.2: Route 1 flows simulated over 100 days as a fraction of total demand with $\theta = 0.25$ and $\delta = (0.6, 0.3, 0.1)$.

Appendix A

Combining states

Let us assume $N = 6$, therefore the number of states is 7 (i.e. 0, 1, 2, 3, 4, 5, 6). We demonstrate how we can combine the states 1 and 2 to a new state called $1'$, and states 4 and 5 to a new state called $4'$. Here, the elements of the new transition matrix can be calculated as follows.

$$
\begin{aligned}
&p_{01'} = \mathbb{P}(X_t = 1' | X_{t-1} = 0) = \\
&\mathbb{P}(X_t = 1 \text{ or } 2 | X_{t-1} = 0) = \\
&\frac{\mathbb{P}(X_t = 1 \text{ or } 2, X_{t-1} = 0)}{\mathbb{P}(X_{t-1} = 0)} = \frac{\mathbb{P}(X_t = 1, X_{t-1} = 0) + \mathbb{P}(X_t = 2, X_{t-1} = 0)}{\mathbb{P}(X_{t-1} = 0)} = \\
&\frac{\mathbb{P}(X_t = 1 | X_{t-1} = 0) \times \mathbb{P}(X_{t-1} = 0) + \mathbb{P}(X_t = 2 | X_{t-1} = 0) \times \mathbb{P}(X_{t-1} = 0)}{\mathbb{P}(X_{t-1} = 0)}
\end{aligned}
$$

$\implies \boxed{p_{01'} = p_{01} + p_{02}}$.

Similarly, $\boxed{p_{04'} = p_{04} + p_{05}}$. Transition probability when the user start off at state 0 and finishes at state 0, p_{00} remains unchanged. Similar situation happens for p_{03} and p_{06}. These values are corresponding to the first row of our new transition matrix. Now the first element of the second row is obtained by

$$
\begin{aligned}
&P_{1'0} = \mathbb{P}(X_t = 0 | X_{t-1} = 1') = \\
&\mathbb{P}(X_t = 0 | X_{t-1} = 1 \text{ or } 2) = \\
&\frac{\mathbb{P}(X_t = 0 | X_{t-1} = 1) \times \mathbb{P}(X_{t-1} = 1) + \mathbb{P}(X_t = 0 | X_{t-1} = 2) \times \mathbb{P}(X_{t-1} = 2)}{\mathbb{P}(X_{t-1} = 1) + \mathbb{P}(X_{t-1} = 2)}
\end{aligned}
$$

$\implies \boxed{P_{1'0} = \dfrac{P_{10}\pi_1 + P_{20}\pi_2}{\pi_1 + \pi_2}}$.

The rest of our new matrix elements can be obtained through similar steps that have been shown. In conclusion, our new transition matrix is given by

$$\begin{pmatrix} p_{00} & p_{01}+p_{02} & p_{03} & p_{04}+p_{05} & p_{06} \\ \frac{P_{10}\pi_1+P_{20}\pi_2}{\pi_1+\pi_2} & \frac{(P_{11}+P_{12})\pi_1+(P_{21}+P_{22})\pi_2}{\pi_1+\pi_2} & \frac{P_{13}\pi_1+P_{23}\pi_2}{\pi_1+\pi_2} & \frac{(P_{14}+P_{15})\pi_1+(P_{24}+P_{25})\pi_2}{\pi_1+\pi_2} & \frac{P_{16}\pi_1+P_{26}\pi_2}{\pi_1+\pi_2} \\ p_{30} & p_{31}+p_{32} & p_{33} & p_{34}+p_{35} & p_{36} \\ \frac{P_{40}\pi_4+P_{50}\pi_5}{\pi_4+\pi_5} & \frac{(P_{41}+P_{42})\pi_4+(P_{51}+P_{52})\pi_5}{\pi_4+\pi_5} & \frac{P_{43}\pi_4+P_{53}\pi_5}{\pi_4+\pi_5} & \frac{(P_{44}+P_{45})\pi_4+(P_{54}+P_{55})\pi_5}{\pi_4+\pi_5} & \frac{P_{46}\pi_4+P_{56}\pi_5}{\pi_4+\pi_5} \\ p_{60} & p_{61}+p_{62} & p_{63} & p_{64}+p_{65} & p_{66} \end{pmatrix}.$$

Appendix B

R code for new definition of the coefficient of reactivity

```
link.costs <- function(y,a,b,pow){
  costs <- a + (y/b)^pow
  costs
}

logit=function(rcost,ODpair,theta){
  ll=length(rcost)
  for(i in 1:length(rcost)){
    rset=which(ODpair==ODpair[i])
    ll[i]=exp(-theta*rcost[i])/sum(exp(-theta*rcost[rset]))}
  ll
}

#initialstates
SUE = function(A,ODdemand,ODpair,a,b,pow,theta,tol=1e-5,x1=NA){
  n.ODpair = length(ODdemand)
  n.routes = c(table(ODpair))
  q.rep = rep(ODdemand,n.routes)
  if (is.na(x1)){
    Lcost = link.costs(rep(0,nrow(A)),a,b,pow)
    Rcost = t(A)%*%Lcost
    p=logit(Rcost,ODpair,theta)
    x1=q.rep*p
  }

  x = x1
  eps = 1
```

```
  m = 1
  while(eps > tol){
    y = A%*%x
    Lcost = link.costs(y,a,b,pow)
    Rcost = t(A)%*%Lcost
    p=logit(Rcost,ODpair,theta)
    eps = sqrt(mean((x/q.rep-p)^2))
    if (eps > tol){
      x = x + (q.rep*p-x)/m
      #    X = rbind(X,x)
      m = m + 1
    }
  }
  # X
  x
}
```

```
flow.history <-function(A,a,b,ODpair,theta,ODdemand,numofruns,mem,
                        delta,x.ini=NA){
  n.links <- nrow(A)
  n.routes <- ncol(A)
  x=matrix(0,nrow=n.routes,ncol=numofruns)
  if (is.na(x.ini[1])){
    s.u.e <- SUE(A,ODdemand,ODpair,a,b,pow,theta)
    for (i in 1:mem){
      x[,i] <- s.u.e
    }
  }
  lc <- matrix(0,nrow=n.links,ncol=mem)
  rc <- matrix(0,nrow=n.routes,ncol=mem)
  for (i in 1:mem){
    lc[,i] <- link.costs(y=A%*%x[,i],a,b,pow)
    rc[,i] <- delta[mem-i+1]*t(A)%*%lc[,i]
  }
  u <- rowSums(rc)


  for(i in (mem+1):numofruns){
    ptemp=logit(rcost=u,ODpair,theta)
    for( od in 1:length(ODdemand)){
      x[ODpair== od,i]=rmultinom(1,ODdemand[od],ptemp[ODpair== od])
    }
```

```
    for (j in 1:mem){
      lc[,j] <- link.costs(y=A%*%x[,i-mem+j],a,b,pow)
      rc[,j] <- delta[mem-j+1]*t(A)%*%lc[,j]
    }
    u <- rowSums(rc)
  }
  x
}


coefficient.reactivity <- function(A,a,b,pow,delta,ODpair,theta,ODdemand,numofruns,mem,
                                   n.instance,n.history){
  n.links <- nrow(A)
  n.routes <- ncol(A)
  history.store <- array(matrix(0,nrow=n.routes,ncol=mem-1),
                         dim=c(n.routes,mem-1,n.history))
  for( k in 1:n.history ) {
    x.hist <- flow.history(A,a,b,ODpair,theta,ODdemand,numofruns,mem,delta)
                                                  [,-c(1:(numofruns-mem+1))]
    history.store[,,k]=x.hist
  }

  n.ODpair <- length(ODdemand)
  n.routes.per.OD.pair <- c(table(ODpair))
  if (max(n.routes.per.OD.pair)>2) stop("Only works for 2 routes per OD pair")

  s.u.e <- SUE(A,ODdemand,ODpair,a,b,pow,theta)

  my.list <- list()
  for (i in 1:n.ODpair){
    if(n.routes.per.OD.pair[i]==2) my.list[[i]] <- 0:ODdemand[i]
    if(n.routes.per.OD.pair[i]==1) my.list[[i]] <- ODdemand[i]
  }

  feasible.route.1.flows <- expand.grid(my.list)


  n.feasible.flows <- nrow(feasible.route.1.flows)
  Av.hist.flow <- numeric(0)
  try <- numeric(n.feasible.flows)
  history_flow = vector("list", length=n.feasible.flows)


  for (h in 1:n.feasible.flows){                    # create all possibilities for x
```

```
    Av.route.flows <- numeric(0)

    x <- numeric(0)
    for (j in 1:n.ODpair){
      if(n.routes.per.OD.pair[j]==2) x <- c(x,feasible.route.1.flows[h,j],
                                        ODdemand[j]-feasible.route.1.flows[h,j])
      if(n.routes.per.OD.pair[j]==1) x <- c(x,feasible.route.1.flows[h,j])

    }

      for( i in 1 : n.history ){
      runs<- numeric(0)
      xnew=matrix(0,n.routes,mem)
      xnew=cbind(history.store[,,i],x)
      lc <- matrix(0,nrow=n.links,ncol=mem)
      rc <- matrix(0,nrow=n.routes,ncol=mem)
        for (k in 1:mem){
        lc[,k] <- link.costs(y=A%*%xnew[,k],a,b,pow)
        rc[,k] <- delta[mem-k+1]*t(A)%*%lc[,k]
        }
     u <- rowSums(rc)
     xsim <- rep(0,n.routes)
       for(l in 1:n.instance){
       ptemp=logit(rcost=u,ODpair,theta)
       for( od in 1:length(ODdemand)){
        xsim[ODpair==od]=rmultinom(1,ODdemand[od],ptemp[ODpair== od])
       }
      runs <- cbind(runs,xsim)

     }

     Av.route.flows <- cbind(Av.route.flows,rowMeans(runs))
  }

    history_flow[[h]] = Av.route.flows
  Av.hist.flow <- rowMeans(Av.route.flows)
  try[h] <- sqrt(sum((Av.hist.flow-s.u.e)^2))/sqrt(sum((x-s.u.e)^2))

  }
 coef.reac <- max(try)
 res =list("coef.reac"=coef.reac, "history_flow"=history_flow)
 return(res)

}
```

Appendix C

DREC 16 Form

DRC 16

MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Ahmad Mahmoodjanlou
Name/title of Primary Supervisor:	Prof. Martin Hazelton
Name of Research Output and full reference:	
Mahmoodjanlou, A., Hazelton, M. L., & Parry, K. (2019). Apples versus oranges? Comparing deterministic and stochastic day-to-day traffic assignment models. Transportmetrica B: Transport Dynamics, 7(1), 1426-1443.	
In which Chapter is the Manuscript /Published work:	Chapter 3
Please indicate:	
• The percentage of the manuscript/Published Work that was contributed by the candidate:	60%
and	
• Describe the contribution that the candidate has made to the Manuscript/Published Work:	
The candidate performed computational work and produced figures; develop theoretical results including proof of theorem; and co-wrote the manuscript with the lead supervisor.	
For manuscripts intended for publication please indicate target journal:	
Candidate's Signature:	Ahmad Mahmoodjanlou Digitally signed by Ahmad Mahmoodjanlou Date: 2020.03.27 11:46:41 +13'00'
Date:	27/03/2020
Primary Supervisor's Signature:	Martin Hazelton Digitally signed by Martin Hazelton DN: cn=Martin Hazelton, c=NZ, o=University of Otago, ou=Department of Mathematics and Statistics, email=martin.hazelton@otago.ac.nz Date: 2020.03.27 12:59:03 +13'00'
Date:	27/03/2020

(This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis)

Bibliography

Takashi Akamatsu et al. Cyclic flows, markov process and stochastic traffic assignment. *Transportation Research Part B: Methodological*, 30:369–386, 1996.

Chandra Balijepalli and David P Watling. Doubly dynamic equilibrium distribution approximation model for dynamic traffic assignment. 2005.

Hillel Bar-Gera. Origin-based algorithm for the traffic assignment problem. *Transportation Science*, 36(4):398–417, 2002.

Martin Beckmann, Charles B McGuire, and Christopher B Winsten. Studies in the economics of transportation. Technical report, 1956.

Shlomo Bekhor and Tomer Toledo. Investigating path-based solution algorithms to the stochastic user equilibrium problem. *Transportation Research Part B: Methodological*, 39(3):279–295, 2005.

Giuseppe Bellei, Guido Gentile, Lorenzo Meschini, and Natale Papola. A demand model with departure time choice for within-day dynamic traffic assignment. *European Journal of Operational Research*, 175(3):1557–1576, 2006.

Moshe Ben-Akiva and Michel Bierlaire. Discrete choice methods and their applications to short term travel decisions. In *Handbook of Transportation Science*, pages 5–33. Springer, 1999.

Moshe Ben-Akiva, Andre De Palma, and Kaysi Isam. Dynamic network models and driver information systems. *Transportation Research Part A: General*, 25(5):251–266, 1991.

Jing Bie and Hong K Lo. Stability and attraction domains of traffic equilibria in a day-to-day dynamical system formulation. *Transportation Research Part B: Methodological*, 44(1):90–107, 2010.

Enide Astrid Irene Bogers. Traffic information and learning in day-to-day route choice.(phd). *Technische Universiteit Delft*, 2009.

David Branston. Link capacity functions: A review. *Transportation Research*, 10(4):223–236, 1976.

Lei Bu, Feng Wang, Xuesong Zhou, and Chuanzhong Yin. Managed gating control strategy for emergency evacuation. *Transportmetrica A: transport science*, 15(2):963–992, 2019.

Giulio E Cantarella and Ennio Cascetta. Dynamic processes and equilibrium in transportation networks: towards a unifying theory. *Transportation Science*, 29(4):305–329, 1995.

Giulio E Cantarella and David P Watling. Modelling road traffic assignment as a day-to-day dynamic, deterministic process: a unified approach to discrete-and continuous-time models. *EURO Journal on Transportation and Logistics*, 5(1):69–98, 2016.

Giulio E Cantarella, Pietro Velonà, and David P Watling. Day-to-day dynamics & equilibrium stability in a two-mode transport system with responsive bus operator strategies. *Networks and Spatial Economics*, 15(3):485–506, 2015.

Ennio Cascetta. A stochastic process approach to the analysis of temporal dynamics in transportation networks. *Transportation Research Part B: Methodological*, 23(1):1–17, 1989.

Ennio Cascetta. *Transportation systems analysis: models and applications*, volume 29. Springer Science & Business Media, 2009.

Ennio Cascetta and Giulio Erberto Cantarella. A day-to-day and within-day dynamic stochastic assignment model. *Transportation Research Part A: General*, 25(5):277–291, 1991.

Ennio Cascetta and Giulio Erberto Cantarella. Modelling dynamics in transportation networks: state of the art and future developments. *Simulation practice and theory*, 1(2):65–91, 1993.

Ennio Cascetta, Agostino Nuzzolo, Francesco Russo, and Antonino Vitetta. A modified logit route choice model overcoming path overlapping problems. specification and some calibration results for interurban networks. In *Transportation and traffic theory. Proceedings of the 13th international symposium on transportation and traffic theory, Lyon, France, 24-26 July 1996*, 1996.

Ennio Cascetta, Francesco Russo, and Antonino Vitetta. Stochastic user equilibrium assignment with explicit path enumeration: comparison of models and algorithms. *IFAC Proceedings Volumes*, 30(8):1031–1037, 1997.

A Charnes and WW Cooper. Extremal principles for simulating traffic flow in a network. *Proceedings of the National Academy of Sciences of the United States of America*, 44(2):201, 1958.

Stella C Dafermos and Frederick T Sparrow. The traffic assignment problem for a general network. *Journal of Research of the National Bureau of Standards B*, 73(2):91–118, 1969.

Carlos Daganzo. *Multinomial probit*. Elsevier, 1979.

Carlos F Daganzo. Stochastic network equilibrium with multiple vehicle types and asymmetric, indefinite link cost jacobians. *Transportation Science*, 17(3):282–300, 1983.

Carlos F Daganzo and Yosef Sheffi. On stochastic models of traffic assignment. *Transportation Science*, 11(3):253–274, 1977.

Gary A Davis and Nancy L Nihan. Large population approximations of a general stochastic traffic assignment model. *Operations Research*, 41(1):169–178, 1993.

Robert B Dial. Equilibrium logit traffic assignment: elementary theory and algorithm. In *80th Annual Meeting of the Transportation Research Board, Washington, DC*, 2001.

Michael Florian and Donald Hearn. Network equilibrium models and algorithms. *Handbooks in Operations Research and Management Science*, 8:485–550, 1995.

Michael Florian, Isabelle Constantin, and Dan Florian. A new look at projected gradient method for equilibrium assignment. *Transportation Research Record: Journal of the Transportation Research Board*, (2090):10–16, 2009.

Charles Miller Grinstead and James Laurie Snell. *Introduction to probability*. American Mathematical Soc., 2012.

Linghui Han, Huijun Sun, David ZW Wang, and Chengjuan Zhu. A stochastic process traffic assignment model considering stochastic traffic demand. *Transportmetrica B: Transport Dynamics*, 6(3):169–189, 2018.

Martin L Hazelton. Day-to-day variation in markovian traffic assignment models. *Transportation Research Part B: Methodological*, 36(7):637–648, 2002.

Martin L Hazelton and David P Watling. Computation of equilibrium distributions of markov traffic-assignment models. *Transportation Science*, 38(3):331–342, 2004.

Xiaozheng He and Henry X Liu. Modeling the day-to-day traffic evolution process after an unexpected network disruption. *Transportation Research Part B: Methodological*, 46(1):50–71, 2012.

Xiaozheng He, Xiaolei Guo, and Henry X Liu. A link-based day-to-day traffic assignment model. *Transportation Research Part B: Methodological*, 44(4):597–608, 2010.

Joel L Horowitz. The stability of stochastic equilibrium in a two-link transportation network. *Transportation Research Part B: Methodological*, 18(1):13–28, 1984.

Hai-Jun Huang and Chang-Yao Gu. Stochastic user equilibrium assignment model and its parameter calibrating. *IFAC Proceedings Volumes*, 27(12):953–958, 1994.

Jeffrey J Hunter. Simple procedures for finding mean first passage times in markov chains. *Asia-Pacific Journal of Operational Research*, 24(06):813–829, 2007.

Mehdi Keyvan-Ekbatani, Anastasios Kouvelas, Ioannis Papamichail, and Markos Papageorgiou. Exploiting the fundamental diagram of urban networks for feedback-based gating. *Transportation Research Part B: Methodological*, 46(10):1393–1403, 2012.

Xuan Lu, Song Gao, and Eran Ben-Elia. Information impacts on route choice and learning behavior in a congested network: experimental approach. *Transportation Research Record*, 2243(1):89–98, 2011.

MJ Maher. Sam-a stochastic assignment model. In *Mathematics in Transport Planning and Control (Institute of Mathematics & Its Applications Conference Series 38)*, 1992.

Fred Mannering and David A Hensher. Discrete/continuous econometric models and their application to transport analysis. *Transport Reviews*, 7(3):227–244, 1987.

Daniel McFadden et al. Conditional logit analysis of qualitative choice behavior. 1973.

Katharina Parry, David P Watling, and Martin L Hazelton. A new class of doubly stochastic day-to-day dynamic traffic assignment models. *EURO Journal on Transportation and Logistics*, 5(1):5–23, 2016.

Michael Patriksson. *The traffic assignment problem: models and methods*. Courier Dover Publications, 2015.

Srinivas Peeta and Athanasios K Ziliaskopoulos. Foundations of dynamic traffic assignment: The past, the present and the future. *Networks and Spatial Economics*, 1(3):233–265, 2001.

Alan M Polansky. *Introduction to statistical limit theory*. Chapman and Hall/CRC, 2011.

Jeffrey S Rosenthal. *A first look at rigorous probability theory*. World Scientific Publishing Co Inc, 2006.

Yosef Sheffi. *Urban transportation networks*, volume 6. Prentice-Hall, Englewood Cliffs, NJ, 1985.

Yosef Sheffi and Carlos Daganzo. A unified approach to transportation forecasting: Hypernetworks and supply-demand equilibrium with disaggregate demand models. In *57th TRB Meeting, Washington, DC*, 1978.

Yosef Sheffi and Warren Powell. A comparison of stochastic and deterministic traffic assignment over congested networks. *Transportation Research Part B: Methodological*, 15(1):53–64, 1981.

Yosef Sheffi and Warren B Powell. An algorithm for the equilibrium assignment problem with random link times. *Networks*, 12(2):191–207, 1982.

Michael J Smith. The existence, uniqueness and stability of traffic equilibria. *Transportation Research Part B: Methodological*, 13(4):295–304, 1979.

Michael J Smith. The stability of a dynamic model of traffic assignment—an application of a method of lyapunov. *Transportation Science*, 18(3):245–252, 1984.

Mike Smith, Martin L Hazelton, Hong K Lo, Giulio E Cantarella, and David P Watling. The long term behaviour of day-to-day traffic assignment models. *Transportmetrica A: Transport Science*, 10(7):647–660, 2014.

Wai Yuen Szeto and Hong K Lo. Dynamic traffic assignment: properties and extensions. *Transportmetrica*, 2(1):31–52, 2006.

Kenneth E Train. *Discrete choice methods with simulation*. Cambridge university press, 2009.

Guangchao Wang, Shoufeng Ma, and Ning Jia. A combined framework for modeling the evolution of traveler route choice under risk. *Transportation Research Part C: Emerging Technologies*, 35: 156–179, 2013.

John Glen Wardrop. Some theoritical aspects of road traffic research.

David Watling. Stability of the stochastic equilibrium assignment problem: a dynamical systems approach. *Transportation Research Part B: Methodological*, 33(4):281–312, 1999.

David Watling. A second order stochastic network equilibrium model, i: Theoretical foundation. *Transportation Science*, 36(2):149–166, 2002.

David Watling and Martin L Hazelton. The dynamics and equilibria of day-to-day assignment models. *Networks and Spatial Economics*, 3(3):349–370, 2003.

David Watling and Tom Van Vuren. The modelling of dynamic route guidance systems. *Transportation Research Part C: Emerging Technologies*, 1(2):159–182, 1993.

David Watling, Agachai Sumalee, Richard Connors, and Chandra Balijepalli. *Advancing methods for evaluating network reliability*. The Institute, 2004.

David Watling, David Milne, and Stephen Clark. Network impacts of a road capacity reduction: Empirical analysis and model predictions. *Transportation Research Part A: Policy and Practice*, 46(1):167–189, 2012.

David P Watling and Giulio E Cantarella. Modelling sources of variation in transportation systems: theoretical foundations of day-to-day dynamic models. *Transportmetrica B: Transport Dynamics*, 1(1):3–32, 2013.

David P Watling and Martin L Hazelton. Asymptotic approximations of transient behaviour for day-to-day traffic models. *Transportation Research Part B: Methodological*, 118:90–105, 2018.

Shanjiang Zhu and David M Levinson. Disruptions to transportation networks: a review. In *Network reliability in practice*, pages 5–20. Springer, 2012.

Shanjiang Zhu, David Levinson, Henry X Liu, and Kathleen Harder. The traffic and behavioral effects of the i-35w mississippi river bridge collapse. *Transportation Research Part A: Policy and Practice*, 44(10):771–784, 2010.

Anthony Ziegelmeyer, Frédéric Koessler, Kene Boun My, and Laurent Denant-Boèmont. Road traffic congestion and public information: an experimental investigation. *Journal of Transport Economics and Policy (JTEP)*, 42(1):43–82, 2008.