Technical University of Munich

Chair of Transportation Systems Engineering

Master's Thesis

An Alternative Online Calibration Approach for Dynamic Traffic Assignment Systems

Moeid Qurashi

supervised by

Univ.-Prof.Dr. Constantinos Antoniou

March 7, 2018

Abstract

This thesis provides an alternate online calibration approach for dynamic traffic assignment (DTA) model. This approach has been used so far to estimate the OD demand and can be extended for other supply and demand parameters. The purpose of online calibration is to dynamically update the model parameters using the observed traffic flow data. The proposed approach gains the advantage against other developed approaches in terms of its computational performance and application scale. Online calibration approaches are being mostly restricted due to the problems of non-linearity and dimensionality. This online calibration approach, named as PC-SPSA, combines a stochastic approximation algorithm i.e. Simultaneous Perturbation Stochastic Approximation (SPSA) with a dimension reduction technique i.e. Principal component analysis (PCA). A set of prior estimates are used to calculate their variance in form of PC-directions. Then, these PC-directions are used to evaluate PC-scores of a latest previous estimate. These PC-scores are then calibrated based on the observed traffic flows using SPSA. SPSA has been widely used as an optimization algorithm for model calibration. However, being a random search algorithm, its performance deteriorates as the problem dimension increases. The application of PCA on SPSA provides two major advantages. First, it reduces the number of variables to be estimated significantly. Secondly, it also narrows down the search area of SPSA from a higher dimensional OD flow vector to lower dimensional PC scores, improving SPSA's performance considerably. Case studies of synthetic non-linear problems with different dimensions and a network of Vitoria, Spain are used to test the proposed PC-SPSA approach. The empirical results from these case studies show that PC-SPSA not only performs very well in reducing the error to a very low value, but it also does it rapidly with very few iterations, making it an effective online calibration approach.

Acknowledgements

At first, i would like to thank my supervisor Prof. Dr. Constantinos Antoniou for his admirable support and guidance throughout my thesis. His supervision have not only broaden my skills and knowledge but also changed my attitude of research towards the right direction.

Secondly, i would like to thank Dr. Tao Ma and Msc. Emmanouil Chaniotakis for their help, technical and moral support during the hard parts of my research. There presence have greatly improved the final product of my research shown in this document.

At last, but the most important, i would like to thank and dedicate my thesis to my parents, whom continuous support and praying made me come this far.

Declaration

I hereby declare that this thesis is an outcome of my own efforts and has not been published anywhere else before and not used in any other examination. Also, to mention that the materials and methods used and quoted in this thesis has been properly referenced and acknowledged.

Place and dated:____________________ Signature:__________________

Contents

List of Figures

Chapter 1

Introduction

1.1 Background and motivation

Automobile industry has seen a major revolution in the last few decades, within half a century, it entirely changed the prospect of traveling for mankind. By an estimate from International Organization of Motor Vehicle Manufacturers (OICA, 2015), there is an average of 58 vehicles per 100 persons in Europe and 67 vehicles per 100 persons in North America while the developing regions like Asia and South America are also having a vigorous increase, i.e. 141% in Asia while 60% in South America, within 10 years since 2005. Due to this rapid growth of vehicle mobility, all major cities of the world are facing the problem of congestion. According to a study by INRIX (2014), combined congestion costs for Europe and United States will soar to $293 billion annually by 2030, which will be almost an increase of 50% from 2013. Looking at this trend of vehicles and congestion increase, it is clear that just increasing the supply in form of infrastructure is neither wise nor feasible both physically and economically. Instead, effective management of the existing infrastructure is needed. The diversity in road networks, with complex traffic patterns make traffic models very critical for developing appropriate traffic management strategies. Due to a large number of parameters and inputs, these models tend to be complex. Therefore, calibration and validation of these models is vital.

A traffic model can be separated into two major aspects, demand and supply. Demand side includes travel behaviour modelling and origin-destination demand estimation while supply side includes traffic dynamics i.e. queue formation, dissipation and spill-back based on the network geometry, link capacities etc. Demand in a model, is mainly represented by origin-destination (OD) matrix with each cell representing trips from origin to destination zones. In reality, an OD-matrix is unmeasurable and various methods are used for its estimation. The three most traditional methods include: *'direct OD matrix estimation'* (by doing surveys and applying sampling theory classical estimators (Cochran, 2007)), *'model estimation'*(trip based models, activity based models) and *'estimation using observed traffic flow'*. The approach of estimating OD matrices using observed traffic flow data has been widely adopted, while other approaches can be inefficient due to the issues like biasness, underdeterminacy in surveys etc (Cascetta, 1984). This technique uses prior estimated demand matrix to calibrate and minimize its difference from a true targeted matrix, that if assigned to the network, reproduce the observed traffic flows

(Willumsen, 1978). Traditionally, the models are calibrated over a longer period of time named as *'offline calibration'*. While, using aggregated traffic demand, offline calibrated models are not sensitive for shorter time scale variations in traffic spanning within hours. The concern for managing the short term variation within a facility is increasing with the growing traffic volumes and faster travel speeds. Short term variations that can be due to the effects of weather conditions, public events, holidays etc, increase the risks towards congestions, accidents etc, hence these variations must be incorporated in traffic models for better traffic management and control.

With the evolution of technology, real-time data can be readily collected using different techniques (like loop detectors, bluetooth detection, ANPR plate recognition systems etc). This real-time data has been used for more efficient and dynamic traffic control, through systems like DynaSMART (Mahmassani et al., 1995), DynaMIT (Ben-Akiva et al., 2010) etc. These systems are based on dynamic traffic assignment (DTA) models with dynamic calibration through real-time traffic data. The approach of adjusting the traffic model parameters with real-time traffic data is named as *'online calibration'*. DTA captures the stochastic nature of traffic behaviour effectively, while online calibration improves the model's sensitivity towards the time-scale variability of traffic. In shorter time scale, demand parameters can oscillate the most within a model. Reliable calibration of OD matrices are critical for these systems, as poor quality matrices will eventually affect the system's credibility.

There are two main factors that evaluate an online calibration approach in DTA systems. First is extent of error minimized for model parameters based on the observed vs simulated data through calibration, while second is the time utilized for calibration. As the time span of calibrating a model offline can be from hours to days due to its complexity and scale. But for DTA models, the time for calibration must be reduced within minutes to timely predict, manage and control a facility. Hence, time is considered as a critical constraint for online calibration. The computational performance in terms of time also depends on the scale of the model as with the increase in network dimensions, the model gets more complex and the number of calibration variables also increase, limiting the use of online calibration on larger networks. Recently, improved techniques have been developed, from reducing the dimensions of calibration problem to reducing the complication of the calibration process itself but there is still room for improvement in terms of faster and improved online calibration to increase its scale of application.

1.2 Problem definition and thesis objective

DTA systems consist of estimation and prediction models, first estimating the state of a network based on the surveillance data before prediction. The accuracy of prediction depends on the quality of estimation. For online calibration, the estimation time is also a very important aspect, as much the estimation time prolongs, less will be the time left for prediction and control.

Problem formulation of online calibration is majorly based on the factors that affect its computational time. One of the key factor is non-linearity. As, online calibration based on the observed traffic data, needs mapping of the model parameters into simulated traffic data. The relation between the model parameters and traffic data is highly non-linear, based on either a simulator or a complex set of non-linear equations, increasing the computational time. Another key factor is the dimensionality of the problem, as the network dimensions increases the model gets more complex, increasing the variables to be calibrated and prolonging the calibration time.

Online approaches developed so far, try to address the factors mentioned above, from Extended kalman filter (Antoniou, 2004) addressing the issue of non-linearity to the efforts of applying Principal component analysis (as PC-GLS (Prakash et al., 2017) and PC-EKF (Prakash et al., 2018)) to reduce the dimensions of the problem in order to improve the scale of application. Another approach that can be a potential solution is Simultaneous Perturbation Stochastic Approximation (SPSA) proposed for offline calibration (Balakrishna, 2006). SPSA can perform better in terms of computational time taken as it needs a fixed (twice) number of objective function evaluation per iteration regardless the dimensions or number of variables involved. But, being a random search algorithm, when the model dimensions increase and become more complex the performance of SPSA deteriorate significantly.

The objective of this thesis is to develop an improved online calibration approach for calibration of DTA models. An approach, which can cater the effect due to increase in dimensionality or complexity of a network. The approach is developed based on the use of SPSA, which has been already acknowledged as an offline calibration approach. Application of SPSA is enhanced with the addition of a dimension reduction technique (i.e. Principal Component Analysis) to improve its performance as well as application's scalability towards calibrating DTA models.

1.3 Thesis Outline

The organization for remainder of this document is as follows. Chapter 2 presents a detailed overall literature review about OD demand estimation and calibration. Chapter 3 provides detailed methodologies of two calibration approaches for OD demand calibration. Details for two case studies and their corresponding results are provided in chapter 4, while conclusions and ideas for future research are given in chapter 5.

Chapter 2

Literature Review

This chapter reviews the literature about OD demand estimation and calibration using traffic data. Divided into a total of 5 sections, first section describes the background on the development of OD demand estimation technique, its mathematical model and major types. Section 2 covers the dynamic OD demand estimation in more detail, discussing about the main approaches developed for it. Section 3 describes the Dynamic traffic assignment (DTA) systems, their structure, usage and the concepts about different types of DTA models. While, the next section starts with the concept of calibration in traffic models, then describes offline and online calibration, their literature and comparison. In the end, the chapter concludes with section 5 covering the concept of stochastic approximation, then the comparison of FDSA and SPSA with their usage so far in OD estimation and calibration techniques.

2.1 OD Demand estimation

Traffic demand in a network is in form of an OD matrix. As a network is divided into zones, each cell of an OD matrix represents the number of trips from an origin zone to a destination zone. In accordance with Cascetta (1984), many techniques are developed to estimate an OD matrix. The first set of techniques can be called as "direct sample estimation", where by using sampling theory classical estimators, OD matrix is estimated from different surveys i.e. household surveys, roadside interview etc (Cochran, 2007). Another technique is defined as "model estimation". In this technique, a system of models is used to generate number of trips per mode for a certain time. There are two major types of these systems, first is called as "trip based modelling" while the other as "activity based modelling" ((Ortuzar and Willumsen, 2011)). The third and most widely used technique is defined as "OD matrix estimation by traffic flows". There are two distinctions in this technique. First is to estimate the OD matrix directly from traffic flows, while second is to calibrate the prior estimates of OD matrices by estimating the demand model parameters based on the observed traffic flows (Willumsen, 1978). In a traffic network, mostly the number of observed links are less than the number of unknown OD pairs in an OD matrix, making it an under-determined problem. Using the first approach and relying only on the link flows might lead to a false OD estimation. As a result, second approach of using the prior information and calibrating the OD matrix has been widely followed.

OD Demand Estimation Models

Generally, the approach is to find an estimate of OD matrix using the traffic flow observations. According to Cascetta (1984), the basic mathematical model can be given as:

$$y = Ax \tag{2.1}$$

Here, y is the flow vector representing the observed data, while x is the OD matrix of a network with n pair of centroids, m number of links and A is the assignment matrix of dimension $m \times n$. Equation 2.2 shows a prior OD matrix $\hat{x}$ from a direct or model estimation with ξ as a random vector of mean μ and dispersion matrix V.

$$\hat{x} = x + \xi \tag{2.2}$$

Another error will be because $\hat{A}$ matrix is an assignment based on $\hat{x}$. While, $\hat{x}$ already have an error ξ, so the equation 2.1 will be:

$$y = \hat{A}x + \omega \tag{2.3}$$

where ω is a random vector. As the observed flows considered to be "true" will have some error due to measurement errors and time variations. This biasness of $\hat{y}$ is ignored with respect to ω, and now the model will become as:

$$\hat{y} = \hat{A}x + \eta \tag{2.4}$$

where η is a random vector with mean equal ω and dispersion matrix W. The aim is to find a "true" OD matrix x by combining the information of prior OD matrix and the observed link flows.

There are two major categories of these OD Demand estimation models, static and dynamic models. Static estimation models assume the constant trip desires throughout the estimation period, estimating a demand table based on daily and hourly average traffic counts. Literature on static OD demand estimation can be concluded in a few main approaches, summarized by Zhou (2004). First approach is based on minimum information/ maximum entropy. Van Zuylen and Willumsen (1980) used entropy minimizing principle, while assuming that averaged link counts are Poisson distributed. An OD estimation problem is created which minimizes a log likelihood function subject to equation 2.1. Maher (1983) and Cascetta (1984) proposed, bayesian estimator and a generalized least squares (GLS) estimator respectively, while assuming a multivariate normal distribution for traffic counts. GLS estimator approach is to solve a system of linear stochastic equations of demand modelling errors and flow measurement errors.

2.2 Dynamic OD demand estimation

Dynamic estimation models being the second type in OD demand estimation models incorporate the traffic dynamics and behavioural process more realistically. As an extension of the OD demand estimation problem from equation 2.1, the new dynamic formulation of the estimation is represented as:

$$y_{st_q} = \sum_{t_p} \sum_{r} A_{st_q}^{rt_p} x_{rt_p} \tag{2.5}$$

where y_{st_q} is the flow observed at sensor s in time interval t_q and x_{rt_p} is the OD pair r that departed from its origin during time interval t_p. While $A_{st_q}^{rt_p}$ is the assignment parameter having the proportion of demand from x_{rt_p} observed at sensor s in time interval t_q. One of the main difference from static estimation models is the dependence of the parameter on time intervals t_p and t_q. While, the static estimation models lack the granularity of time representation in dynamic implementations, hence fractions of their assignment matrix (A) cannot have the effects of OD flows related to prior intervals on the link counts of any interval (Ashok, 1996).

2.2.1 Dynamic OD estimation problem

The aim of doing OD estimation is to find the OD demand $\hat{x}_h$ that departs at time h resulting in the simulated link flow $\hat{y}_{h+1}$ to be as close as the observed link flow y_{h+1} as possible. In accordance with Djukic (2014), the dynamic OD estimation problem from equation 2.5 can be redefined as:

$$\hat{x}_h = \underset{x \geq 0}{argmin} f(\sum_{h}^{h+1} A_{h+1}^{h} x_h, y_{h+1}) \tag{2.6}$$

Where f is a function defined to measure the deviation between observed and estimated flows. As discussed before, the technique of OD estimation from observed traffic flows, have two basic approaches. One is to directly estimate the OD matrix from the observed traffic OD flows, but as a system of stochastic equations shown in equation 2.6, the information is not enough to estimate the dynamic flows. Even assuming equation 2.6 as a system of linear equations, the number of observed counts will remain less than the number of OD flows for estimation. Hence, this is an under-determined system which can give infinite number of OD matrices resulting the observed link traffic counts. Here, the second approach become more feasible. Using a prior estimates of the OD matrix to estimate the new OD matrix out of infinite number of potential ones. Considering the second approach of using the previous estimates the new dynamic OD estimation problem can be defined as:

$$\hat{x}_h = \underset{x \geq 0}{argmin} [\alpha f_1(x_h, x_p) + (1 - \alpha) f_2(\sum_{h}^{h+1} A_{h+1}^{h} x_h, y_{h+1})] \tag{2.7}$$

where x_p is the previous estimate used, while α is the weight factor for combining the two sets. The weights act as deciding factors to define importance of each set, specifically to describe the reliability of previous estimates (Djukic, 2014).

In the dynamic OD estimation problem, there are three main components:

- **Input data:** The input data for the estimation problem as shown in figure 2.1. There are two types of data as inputs. First can be any type of traffic data i.e. travel times, counts, densities, OD flows etc representing the true ground information. While second type of data is historical estimates of OD flows. Equation 2.6 also refers x_h and y_{h+1} are the input data variables.

- **Assignment:** Type of the assignment presents the relation between OD flows and observed traffic data. A^h_{h+1} is the mapping between the estimated traffic data $\hat{y}_{h+1}$ and OD flow x_h in 2.5. This relation is mostly represented either analytically (using set of non-linear equations) or by a traffic simulator.

- **Objective function:** The objective function f in equation 2.6 is the function defined to minimize the difference between observed and estimated traffic data. It represents the type of OD estimation approach being used. The type of objective function also defines its application limitations.

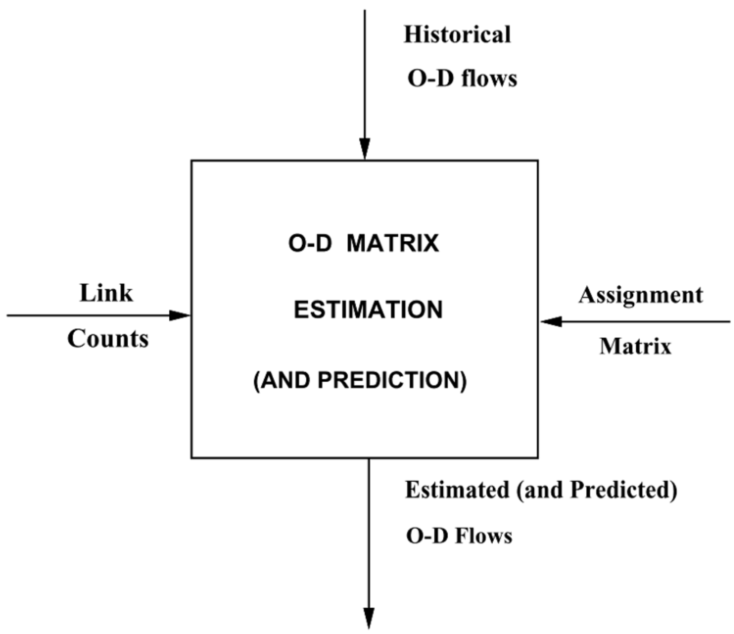

Figure 2.1: Components of OD estimation problem (Ashok, 1996)

In literature, various approaches have been developed to solve the OD estimation problem but the prominent ones can be distinguished as in two: Generalized least squares (GLS) and State space model solving through Kalman filter.

2.2.2 Generalized least squares (GLS)

Generalized least squares (GLS) method is defined as minimizing the sum of least squares of the residuals. Cascetta et al. (1993) gave an approach of optimizing the two objective functions f_1 and f_2 from the dynamic OD estimation problem (equation 2.7), such that the sum of squares of the residuals for both functions is minimized.

As the first part minimizes the difference between the estimated OD of an interval x_h and a prior OD matrix x_p from a prior interval p. The frequency based GLS estimator is given as:

$$f_1(x_h, x_p) = (x_h - x_p)^T V^{-1}(x_h - x_p) \tag{2.8}$$

where V is the variance-covariance matrix of the sampling errors vector that is affecting the prior estimate x_p. While the second objective function f_2 based on basic equation 2.4 is given as:

$$f_2(x_h, y_{h+1}) = \delta^T W^{-1}\delta \tag{2.9}$$

$$\delta = Ax_h - y \tag{2.10}$$

where δ is given as the difference between observed traffic flow y and estimated traffic flow $\hat{y}$. $\hat{y}$ comes by assignment of OD demand x_h through the assignment matrix A, while W is the dispersion matrix of the assignment and measurement errors as referred in equation 2.4. Cascetta et al. (1993) proposed two possible approaches to estimate the OD flows. One is referred as simultaneous estimators, which estimates a whole OD demand for all intervals n by utilizing available counts of all n intervals together. The mathematical model is:

$$(\hat{x}_h....\hat{x}_{h+n}) = \underset{[x_h.....x_{h+n}]\geq 0}{argmin} \sum_{1}^{n} \Big[(x_h - x_p)^T V^{-1}(x_h - x_p) + \delta^T W^{-1}\delta\Big] \tag{2.11}$$

While, the second approach is sequential estimator. It estimates OD matrix step by step for each interval using counts of current and previous intervals in f_2 as well as the previous interval's OD estimates in f_1. As h is the current interval, then x_p can be any of the previous estimates $[\hat{x}_{h-1}, \hat{x}_{h-2,....}]$. The mathematical model is:

$$\hat{x}_h = \underset{x_h\geq 0}{argmin}\Big[(x_h - x_p)^T V_h^{-1}(x_h - x_p) + \delta_h^T W_h^{-1}\delta_h\Big] \tag{2.12}$$

where V_h and W_h are considered to be diagonal matrices, implying that they do not have any covarience between them. As in equation 2.7, weights can be given to the respective objective function based on the degree of confidence on the data being used like prior estimates etc. Comparing both approaches of simultaneous and sequential GLS, sequential have a more computational advantage due to its step wise estimation as well as it can also use the estimated OD flows of an interval into next intervals as a prior estimate. Application of a sequential GLS on real time dynamic OD demand estimation is actually a state space model.

2.2.3 State Space model

A state space model has two basic equations, measurement equation and transition equation. Okutani and Stephanedes (1984) made the first effort to create a state space formulation for dynamic OD estimation problem. A form that can be solved using kalman filter. Measurement equation is stated as the same in equation 2.4:

$$y_h = A_h x_h + v_h \tag{2.13}$$

While, the transition equation is:

$$x_h = F_{h-1}^{h} x_{h-1} + w_h \tag{2.14}$$

where v_h and w_h are normally distribution independent matrices of random errors. A_h is the assignment matrix and F_{h-1}^{h} is a transition factor which relates spatial and temporal OD relationship between time intervals $(h-1)$ and h, capturing the effect of a state vector from interval $(h-1)$ on the state vector from interval h. Ashok (1996) redefined the approach where the state vectors are the difference between OD flow and average prior OD flow. Instead of estimating the actual OD matrix, it estimates the deviation from prior OD martix. Using the basic approach stated in equation 2.7 and the state space model equations 2.13 and 2.14, a new model is derived as:

$$\hat{x}_h = \hat{x}_{h-1} + k_h(y_h - A_h \hat{x}_{h-1})$$

where k_h is called the Kalman gain matrix. k_h is actually a time-varying weighting matrix estimated on the criteria of minimizing the sum of squared errors. The overall model approach can be defined in the following equations:

$$\hat{x}_{0|0} = \bar{x}_0$$

$$\hat{x}_{h|h-1} = F_{h-1}\hat{x}_{h-1|h-1} \tag{2.15}$$

$$\hat{x}_{h|h} = \hat{x}_{h|h-1} + k_h(y_h - A_h \hat{x}_{h|h-1}) \tag{2.16}$$

where $\bar{x_0}$ can be a prior estimate by direct or model estimation providing the base OD matrix and $h = 1, 2, 3, ...$ are the time intervals. This approach by Ashok and Ben-Akiva (2000) contains better practical adaptability from its predecessor. As, the state augmentation is a way of improving the OD flow estimates by exploiting the prior information about OD departure intervals. OD deviations from past intervals are added to the state vector, and re-estimated periodically as future intervals are being processed. This way of state augmentation can provide a bargain between computational effective sequential estimator (eq. 2.12) and a more efficient simultaneous estimator (eq. 2.11).

Djukic (2014) provided a comprehensive summary of the efforts made in the past three decades for solving the dynamic OD estimation problem in figure 2.2. This summary contains all the information about the three major components of dynamic OD estimation problem mentioned before, i.e. the types of input observed data, mapping approach and the objective function utilized.

	Input data			Mapping		Objective function		
Author	ODF	RF	LC	A-DTA	S-DTA	GLS	SS	ME
Bell (1991)	+		+	+		+		
Cascetta et al. (1993)	+		+	+	+			
Chang & Wu (1994)	+		+		+		+	
Chang & Tao (1996)	+	+	+	+			+	
Wu (1997)	+		+		+			+
Van Der Zijpp & De Romph (1997)	+	+	+	+			+	
Tavana (2001)	+		+		+	+		
Sherali & Park (2001)	+		+		+	+		
Dixon & Rilett (2002)	+		+	+		+		
Ashok & Ben-Akiva (2002)	+	+	+	+			+	
Mishalani et al. (2002)	+		+		+		+	
Tsekeris (2003)	+		+		+			+
Lindveld et al. (2003)	+		+		+	+		
Bierlaire & Crittin (2004)	+		+		+	+		
Kwon & Varaiya (2005)	+	+	+		+		+	
Zhou & Mahmassani (2006)	+		+		+	+		
Antoniou et al. (2006)	+	+	+	+			+	
Balakrishna (2006)	+		+		+	+		
Zhou & Mahmassani (2007)	+	+	+		+		+	
Barcelo (2010)	+	+	+	+			+	
Cipriani et al. (2011)	+		+		+	+		
Frederix et al. (2011)	+		+		+	+		

Input data: ODF, OD flow data; RF, route flow data; LC, link condition data; Mapping of OD flows to inout data: A-DTA, analytical based dynamic traffic assignment; S-DTA, simulation based dynamic traffic assignment; Objective functions: GLS, generalized least squares; SS, state-space; EM, entropy maximization.

Figure 2.2: Summarize literature review about the efforts for dynamic OD estimation-(Djukic, 2014)

2.3 Dynamic Traffic assignment systems

Dynamic traffic assignment (DTA) systems are an important part of the modern traffic management systems. Also referred as traffic estimation and predication systems, figure 2.3 provide the generic structure of a DTA system with two major processes, state estimate and predication-based information generation. The key components of a DTA framework are the demand and supply simulator. Traffic demand represented by OD matrices, oscillates the most in shorter scale of time, hence their estimation and prediction is crucial for a DTA systems. Supply simulator is usually a detailed high level simulator like DynaMIT (Ben-Akiva et al., 2010) and DYNASMART (Mahmassani et al., 1995), representing the traffic dynamics and supply parameters (Antoniou, 2004). Different types of DTA models used in DTA systems are discussed below.

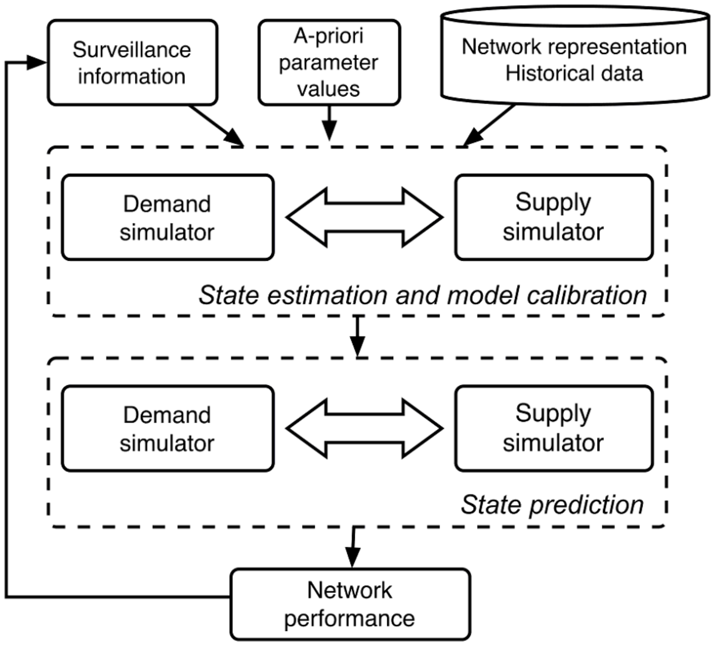

Figure 2.3: Structure of a DTA System-(Antoniou, 2004)

Dynamic traffic assignment models

The complexity of DTA systems require, detailed simulation based DTA models. DTA models replicate the complex interaction of demand and supply model parameters to simulate traffic scenarios, as shown in figure 2.1. Demand models include the estimation/prediction of OD matrices and modelling travel behaviours. While, the supply models include the traffic dynamics due to lane changing, acceleration, merging/weaving, incidents etc. (Ben-Akiva et al., 2002).

Balakrishna (2006) have summarized the efforts for the evolution of DTA models. Starting with the technique classified as "quasi-dynamic" assignment, with the aim of introducing dynamic considerations by repeating the static method application to sub-intervals of a period. Time of an event is divided into intervals and solved by static user equilibrium to get a dynamic scenario. But the reliance on static assignment had limitations which led the research to other techniques. Further research

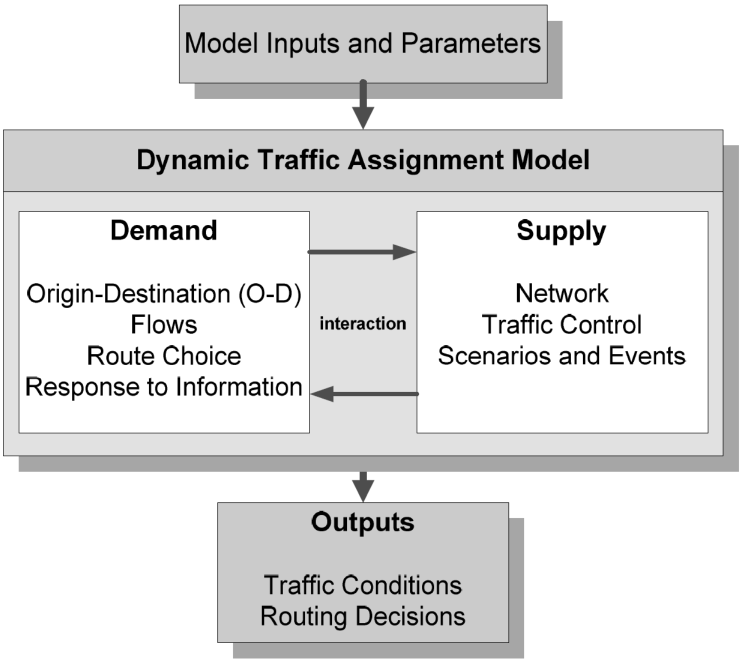

Figure 2.4: Generic interactions in a DTA model-(TRB, 2011)

can be categorized into two major classes of DTA, Analytical and simulation based dynamic traffic assignment. In analytical DTA approaches, assignment problem is based on a set of equations with constraints, to approximate the DTA problem for a specific objective like user equilibrium (Wardrop, 1952). These equations evaluate the observed counts by creating a list of contributions by each OD pair at the locations. The solutions are mostly based on traditional optimization algorithm solving for the unknown variables. Peeta and Ziliaskopoulos (2001) can provide more detail review on the analytical based modelling concepts.

The simulation based DTA models are classified based on the level of detail achieved for individual drivers and driving behaviour. Major three classes are *Microscopic*, *Marcoscopic* and *Mesoscopic* models. *Microscopic* models include vehicular interactions i.e. car following, lane changing, merging and yielding. Individual drivers with their behavior and decisions are simulated in a high level of detail, capturing the traffic dynamics. While, *Marcoscopic* models consider demand as a homogeneous flow and use physical concept like fluid dynamics to propagate through the network without capturing the stochasticity of driver interactions and behavior patterns. The run time of *macroscopic* models is a lot faster than *microscopic*, coming on an expense of the lack of individual driver behaviour modelling. *Mesoscopic* models blend the *microscopic* approach of modelling behaviour patterns with *macroscopic* models approach of traffic dynamics. They are significantly faster than microscopic models and yet capture the behavioural patterns, making them most feasible ones for real-time estimation of traffic conditions (Balakrishna, 2006).

2.4 Calibration

Calibrating a model, is a process of adjusting the model parameters so that the model can be a close representation of the real traffic system (Olstam and Tapani, 2011). As discussed before, DTA models are complex, involving a large number of parameters from both supply and demand models. Calibration of DTA models is an important aspect, considering the level of variability or stochasticity involved in traffic systems. The concept of demand calibration is the same as of OD demand estimation with a prior estimate, discussed in previous sections. Based on the application, calibration of a DTA model can be classified into two main approaches: offline calibration and online calibration.

The objectives of offline and online calibration are identical i.e. to estimate the DTA model parameters, which give similar outputs as the observed traffic data. Both combine all the available data for estimation, as figure 2.2 shows the inputs, outputs and the interaction of both calibration processes. The main difference between them is the time-based representation by the model's output after calibration. Offline calibration establishing a historical database, enables the model to simulate average traffic conditions observed over multiple days. This helps in the planning aspects like developing and evaluating traffic management strategies, but fails to replicate the short term time varying conditions within a day. To achieve this level of performance, the offline parameters must be fine tuned in real-time. The abundance of historical data from offline calibration can be integrated within a real-time online process as the offline calibrated parameters. With the approach of online calibration, DTA models are sensitive enough to capture the shorter time variations in traffic due to an incident, weather conditions etc (Antoniou et al., 2009).

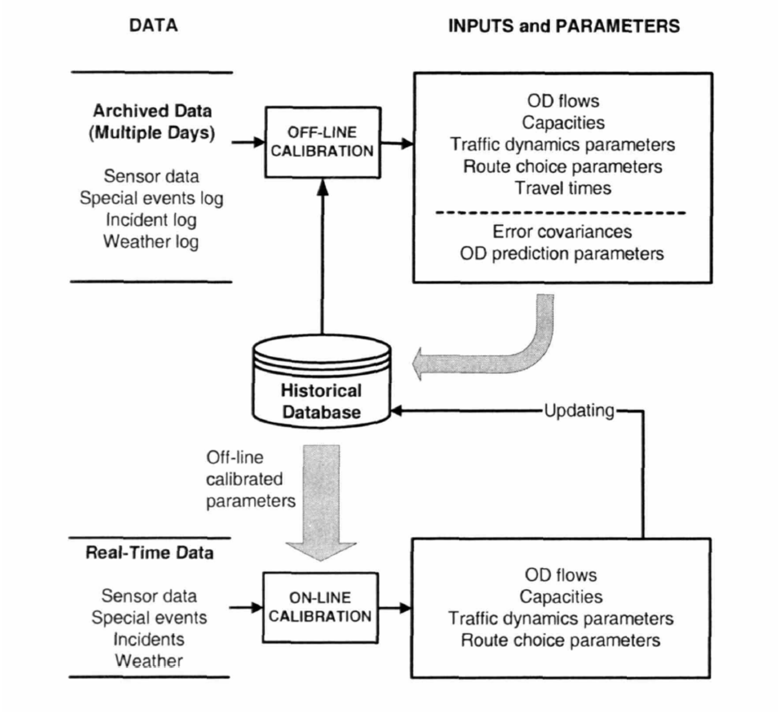

Figure 2.5: Offline and Online Interaction-(Balakrishna, 2006)

Calibration of a DTA model as a generic optimization problem is formulated as:

$$\underset{\beta,x}{Minimize}\ z(y, y', x, x^p, \beta, \beta^p) \tag{2.17}$$

Where,

$$y' = f(x, \beta, G)$$

where β and x are OD flows and model parameters with β^p and x^p as their prior values. y is the observed traffic measurements with y' their simulated counterparts. y' is based on x, β and G parameters, with G representing the road network and $f()$ denotes the DTA model. While, the optimization is based on the goodness of fit z, capturing the error between simulated and observed values (Balakrishna, 2006).

2.4.1 Offline Calibration

As described before, offline calibration addresses the problem of calibrating a set of initial DTA model parameters using time-dependent traffic data aggregated upon multiple days. In accordance with Antoniou et al. (2009), the offline calibration problem is developed as a special case of equation 2.17 as:

$$(\beta^*, x^*) = \underset{\beta,x}{argmin}\ z = \sum_{h=1}^{n}[z_1(y_h, y'_h) + z_2(x_h, x_h^p)] + z_3(\beta_h, \beta_h^p) \tag{2.18}$$

subject to:

$$y' = f(x_1, ..., x_h; \beta_1....\beta_h; G_1...G_h) \tag{2.19}$$

$$l_h^x \leq x_h \leq \mu_h^x \tag{2.20}$$

$$l_h^\beta \leq \beta_h \leq \mu_h^\beta \tag{2.21}$$

Where β^* and x^* are the calibrated model and demand parameters and $z_1()$, $z_2()$, $z_3()$ are the function evaluating the difference depending on the goodness of fit measures, as equation 2.8 gives the estimation relation based on prior estimate x_p by generalized least squares. While, the above constraints apply on all intervals $h \in \{1, ..., n\}$. Due to the scale and complexity of the offline calibration problem, focus has been more on solving a sequence of smaller sub-problems. Using a part of available data, each sub-problem estimates a subset of DTA parameters assuming all other to be exogenous. One of the prominent solution applied for offline calibration is given by Balakrishna (2006), using Simultaneous Perturbation Stochastic Approximation (SPSA) (Spall, 1998a). As, SPSA approximates the gradient by perturbing all calibration parameters simultaneously, it requires only a fixed number of function evaluations (twice) regardless of the problem dimensions making it very computational effective against large scale problems. Details on SPSA and its effectiveness is covered in section 2.4.3 and chapter 3. Antoniou et al. (2009) gives in detail the literature review about the solution approaches for offline calibration problem. Major efforts towards the demand side parameters are by Balakrishna et al. (2005), who sequentially estimated the OD flows and route choice parameters using multiple days count, while efforts of Ashok and Ben-Akiva (2000) and Hazelton (2000) entirely focused on estimating OD flows.

2.4.2 Online Calibration

Online calibration is a real-time process concerning about a given day's system performance. Prior estimates, which could be the offline calibrated parameters are fine tuned, to calibrate the models for shorter time variations due to an incident or weather variation etc. Online calibration is to estimate the demand and supply parameters at a time interval h using observed data from time intervals $\{1, ..., h\}$ and prior estimated parameters. While the idea of calibration is same as, referred in equation 2.17, but online calibration is done sequentially using the information available at the given interval. The online calibration problem formulation is very much similar to that of a state space model described in section 2.2.3 proposed by Ashok and Ben-Akiva (2000). Similar to equation 2.14, which gives a transition equation using a typical autoregressive process to relate the current interval state to the states from previous interval. Equation 2.22 and 2.23 give the direct measurement of the current state vectors x_h and β_h from x_h^p and β_h^p, which are the prior values with l_h and q_h as random error vectors.

$$x_h^p = x_h + l_h \tag{2.22}$$

$$\beta_h^p = \beta_h + q_h \tag{2.23}$$

While, the indirect measurement equation is given in equation 2.24 similar to equation 2.13, with n to be the no. of intervals required for the longest trip inside the network.

$$y_h = f(x_{h-n}, ..., x_{h-1}, x_h; \beta_{h-n}, ..., \beta_{h-1}, \beta_h; G_{h-n}, ..., G_{h-1}, G_h;) + v_h \tag{2.24}$$

Most of the methodologies researched for solving the online OD calibration problem have been based on the kalman filtering technique (Kalman, 1960). Initially, Ashok and Ben-Akiva (2000) made an effort to define the OD estimation problem as a state space model and estimated the deviations from a prior OD matrix using Kalman filtering algorithm. Then, extension of Kalman Filter to non-linear systems with Extended Kalman Filters (EKF) is proposed (Chui and Chen, 1999), implying the non-linear relationship of measurement equation towards linearization. While, in simulation based systems, EKF required a large number of function evaluations for linearization using numerical derivatives, making it computationally expensive. Then an improved EKF in form of Limiting Extended Kalman Filter (LimEKF) is proposed, with a more effective computational performance. LimEKF, instead of computing Kalman gain matrix in real-time, replaces it with a constant gain matrix called limiting Kalman gain matrix (Chui and Chen, 1999). While, this constant gain matrix can be computed and updated offline or a method of moving average could be applied on previous estimates of gain matrices. Antoniou et al. (2005) defined an online calibration problem for speed-density relationship to test the three solution approaches i.e. EKF, Iterated EKF and Unscented Kalman Filter (UKF) and later Antoniou et al. (2007a) applied them to jointly estimate demand and supply parameters of dynamic traffic assignment systems. Zhou and Mahmassani (2007) also developed a similar Kalman Filter based procedure with its transition equation as polynomial trend filter capturing historical demand deviations.

In recent years, efforts have been towards increasing the application scalability of online calibration. As larger networks have larger OD matrix dimensions and simulation or mapping periods, the computational run-time is extensively increased. Djukic et al. (2012) used the concept of principal component analysis (PCA) on OD matrix to reduce its dimensions. Using, principal components instead of original OD matrix in the dynamic OD estimation using Extended Kalman Filter. While, Prakash et al. (2017) proposed the usage of PCA with generalized least square problem to significantly improve the online calibration's scale. Afterwards Prakash et al. (2018) proposed a generic framework for the usage of a dimensional reduction technique with any set of parameter of either supply or demand using wide range of field measurements. Prakash et al. (2018) used a case study of Singapore highway network to evaluate the performance of PCA towards estimation and prediction using a constrained EKF approach. These recent efforts of incorporating the application of PCA are very important in terms of increasing the application scalability of online calibration. The application of PCA on OD demand estimation is discussed in detail in section 3.2.

2.5 Stochastic Approximation

Stochastic approximation (SA) methods are iterative type of optimization algorithms. These type of optimization algorithms minimize a specified error for a given problem when its objective function has no known analytical form, but can be only be estimated based on noisy observations. The approach is to iteratively find an order of parameter estimates which converge the objective function towards zero. Stochastic approximation has been another important aspect in literature of OD calibration. OD estimation was first converted into a generic optimization problem by Cascetta et al. (1993). Then, the dynamic OD estimation problem is formulated as a state space by Okutani and Stephanedes (1984) and Ashok (1996), which is also a direct optimization formulation.

The first major application of stochastic approximation was to linearize the non-linear relationship of measurement equation in the application of Extended Kalman Filter (Chui and Chen, 1999). EKF required a large number of function evaluations for linearization using numerical derivatives, making it computationally expensive. While, Antoniou et al. (2007b) provided another approach of approximation of measurement equation by simultaneous perturbation to reduce the computational efforts. Instead of taking the numerical derivatives with a large number of function evaluations, simultaneous approximation needs only two evaluations of objective function per iteration. This idea is similar to the comparsion of Finite Difference Stochastic Approximation (FDSA) versus Simultaneous Perturbation Stochastic Approximation (SPSA), given by Spall (1998a) in figure 2.6. Both,

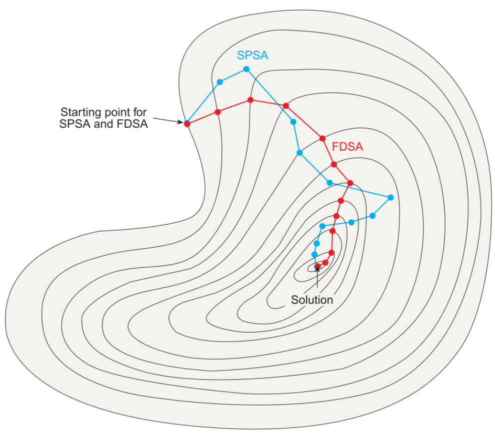

Figure 2.6: FDSA vs SPSA-(Spall, 1998a)

FDSA and SPSA are gradient free stochastic algorithms for the problems which does not allow the approximation of its true gradient.

FDSA is the classical gradient free stochastic approximation algorithm (Kiefer and Wolfowitz, 1952). Equation 2.25 describes the gradient approximation for FDSA, as:

$$g'_k(\theta_k) = \begin{bmatrix} \frac{y(\theta_k + c_k\xi_1) - y(\theta_k - c_k\xi_1)}{2c_k} \\ . \\ . \\ . \\ \frac{y(\theta_k + c_k\xi_p) - y(\theta_k - c_k\xi_p)}{2c_k} \end{bmatrix} \tag{2.25}$$

where, ξ_i is a vector of dimension p with 1 in the ith place and 0 elsewhere. While, p equal to the θ_k vector dimensions. Gradients are evaluated with two evaluations of the objective function y for each i from $[1, 2, ..., p]$ within each iteration. So, the gradients are evaluated with perturbing one element from the θ vector at a time (Spall, 2003). Equation 2.26 defines the gradient approximation for SPSA.

$$g'_k(\theta_k) = \frac{y(\theta_k + c_k\Delta_k) - y(\theta_k - c_k\Delta_k)}{2c_k\Delta_k} \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ . \\ . \\ \Delta_h \end{bmatrix} \tag{2.26}$$

where, instead of ξ_i, Δ_k is used. Δ_k is a random vector based on ± 1 bernoulli distribution. The remaining notations are described in section 3.1. With the usage of Δ_k are the elements of the vector θ are perturbed simultaneously and the gradient is evaluated with only two evaluations of the objective function y.

The number of iterations for SPSA are always more than FDSA. But as, SPSA needs two evaluations of objective functions for each iteration, the number of gradient evaluations for FDSA increases p-folds than SPSA. Hence, SPSA needs very less computational effort than FDSA, making its run-time alot more faster.

SPSA is first proposed by Balakrishna (2006) and Balakrishna et al. (2007) as a solution approach for offline calibration of DTA models, due to its advantage of fixed number of function evaluations (two) irrespective of the dimensions of the problem. Later, Antoniou et al. (2007b) utilized this approach to reduces the computational efforts in Extended Kalman Filter. Recently, Lu et al. (2015) and Antoniou et al. (2015) made the efforts to improve its application efficiency for DTA model calibration by introducing the idea of Weighted Simultaneous Perturbation Stochastic Approximation (W-SPSA). W-SPSA tries to incorporate the effect of non-homogeneous structural correlation with the parameters of the traffic simulation model by the determination of a weight matrix approximating the actual correlation patterns between the model parameters.

Chapter 3

Methodology

This chapter is divided in three major sections. First section describes about the application of an optimization algorithm, Simultaneous perturbation stochastic approximation (SPSA), for OD demand calibration. The section starts with the description of major steps involved in SPSA. Then, comes the modifications proposed in generic SPSA to use it for OD Demand calibration. Afterwards, its complete algorithm is explained with the corresponding flow chart. Next section covers the development of PC-SPSA. It is an application of principal component analysis (PCA) with SPSA. PCA being a dimension reduction technique, improves the computational performance of SPSA. The section starts with the description about PCA and its implementation on OD demand. Then, the steps of converting SPSA to PC-SPSA are discussed, while the section ends with the description of the complete algorithm of PC-SPSA and its corresponding flowchart. Third and the last section concludes with a comparison between SPSA and PC-SPSA through analyzing the major differences in both algorithms.

3.1 Simultaneous Perturbation Stochastic Approximation (SPSA)

SPSA proposed by Spall (1998a), comes from the family of stochastic approximation (SA) methods, which are iterative optimization algorithms. SPSA provides a major advantage over other SA algorithms because it needs only two evaluations for the given objective function to calculate its gradient for minimization (for detail see section 2.4.3). There are some basic parameters that define SPSA's performance, these parameters are defined manually with the guidelines given by Spall (1998b), the parameters are:

- θ The decision variable
- c, γ To specify c_k (where k is the iteration number)
- a, A, α To specify a_k
- c_k, a_k Gain sequence
- Δ Random vector based on bernoulli distribution symmetrically distributed over zero with discrete value of either -1 or 1, with size equal θ.

The performance of SPSA for a specified problem is based on the definition of gain sequence c_k and a_k and their reduction pattern. c_k defines the magnitude of perturbation in the decision variable θ. While, a_k defines the magnitude of minimization for θ in each iteration. c_k and a_k are being specified as:

$$c_k = \frac{c}{(k)^{\gamma}} \qquad a_k = \frac{a}{(k+A)^{\alpha}}$$

So that:

$$a_k > 0, c_k > 0, a_k \to 0, c_k \to 0, \sum_{k=0}^{\infty} a_k = \infty, \sum_{k=0}^{\infty} a_k^2/c_k^2 < \infty$$

Where parameters c and a are defining the magnitude of gain sequence c_k and a_k and γ, α and A define the pattern of reduction in a_k and c_k with the increase in number of iterations as shown in figures 3.1 and 3.2.

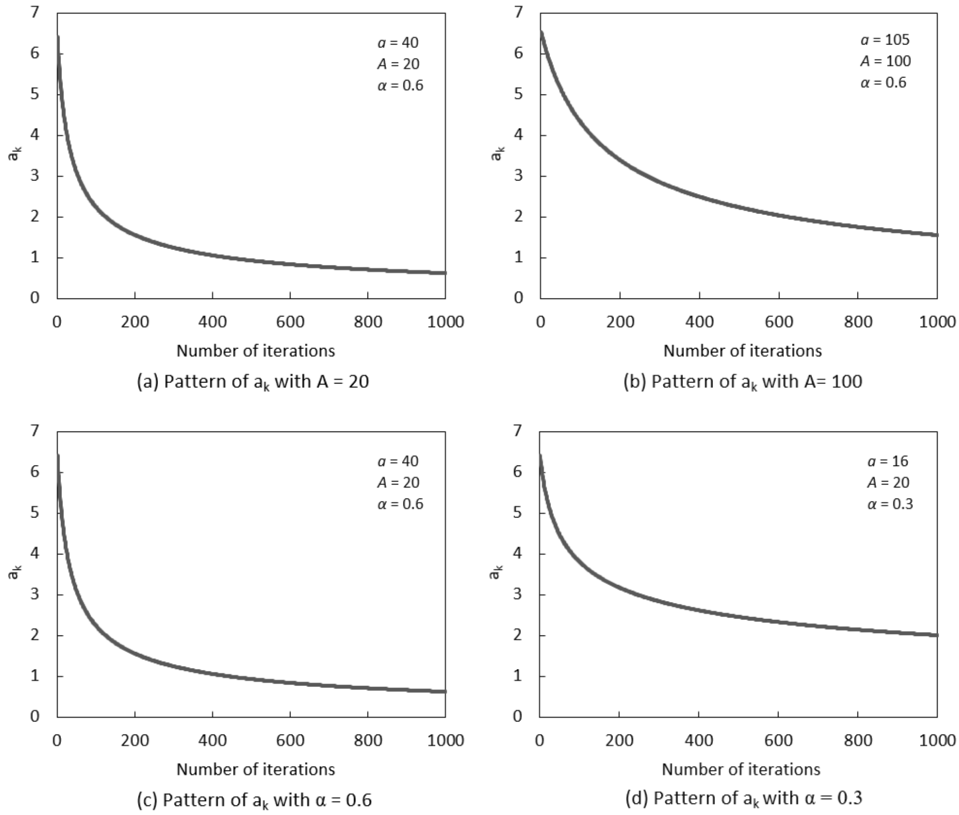

Figure 3.1: Reduction pattern of a_k (Source: own)

The comparison of figure 3.1 (a) and (b) shows the difference in the reduction pattern of a_k due to change in A. The value of A defines the starting number of iterations in which the minimization will be alot more than the upcoming iterations. As figure 3.1(a) shows that the value of a_k is reduced significantly in the first 20 iterations as A is 20, while in figure 3.1(b), the value of A is 100 and the reduction pattern smoothened out in the first 100 iterations. It can be proclaimed that A defines the point of inflection in the reduction pattern of a_k.

Figure 3.1 (c) and (d) shows the comparison of a_k patterns due to change in parameter α. This comparison shows that with the larger value of α the range of reduction increases from its starting point, as the relation of α and a_k is exponential. The values of a are also changed in all the graphs with the aim of showing the change in pattern due to A and α with similar starting points.

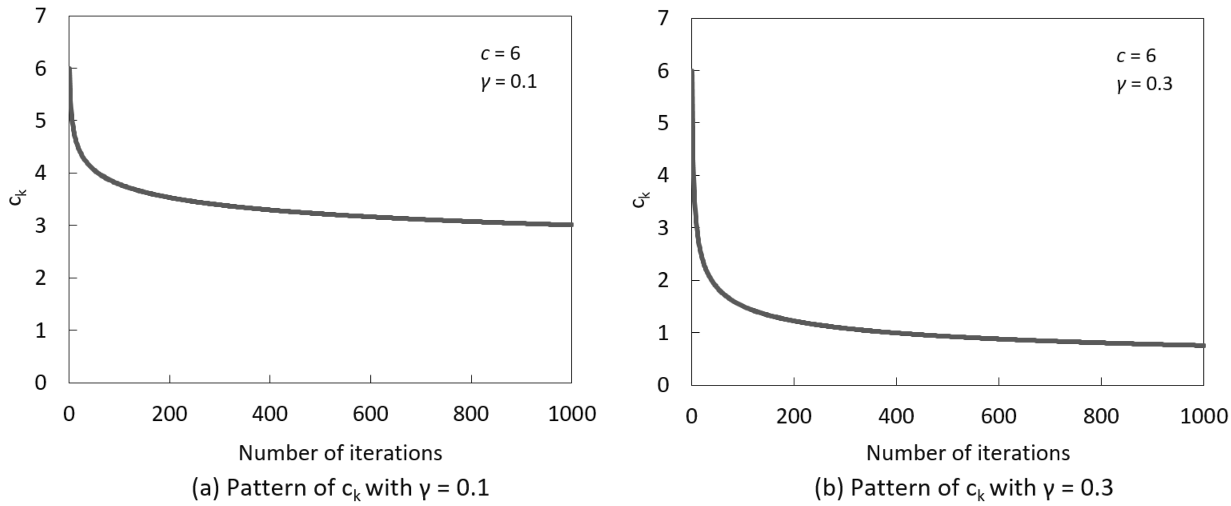

Figure 3.2: Reduction pattern of c_k (Source: own)

Figure 3.2 shows the comparison of the reduction patterns in c_k due to two different values of γ. The comparison is similar to that of α on a_k as γ has the similar relation with c_k. Larger values of γ tend to increase the range as well as the amount of reduction with the increasing number of iterations.

3.1.1 Major steps with SPSA

SPSA being an iterative process, contains four major steps within an iteration. These steps are:

1. Perturbation:

First step is to perturb the decision variable by adding and subtracting the gain sequence c_k times a random vector resulting in two variables θ^+ and θ^-. The random vector Δ randomly increases half of the vector variables by c_k and reduces the other half to create θ^+. While for θ^-, the sign is changed for c_k, so the vector variables that were increased before are reduced while increasing the other half.

$$\theta^{\pm} = \theta \pm c_k \Delta \tag{3.1}$$

2. Objective function evaluation:

In this step, the objective function is evaluated twice based on the resulting θ^+ and θ^- from perturbation. In case of OD estimation, due to incomparable relation of OD flows are traffic counts, this objective function $f()$ can be specified based on a goodness of fit, used to specify the difference between the observed and the simulated traffic data.

$$y^{\pm} = f(\theta^{\pm}) \tag{3.2}$$

where, $y^{\pm}$ is the goodness of fit value.

3. Gradient Approximation:

Next step is to approximate the gradient by taking the difference between the two evaluated outputs from the objective function y^+ and y^- and dividing it by the perturbation magnitude $c_k \times \Delta$.

$$g' = \frac{y^+ - y^-}{2c_k} \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ . \\ . \\ \Delta_h \end{bmatrix} \tag{3.3}$$

4. Minimization:

After the gradient is approximated, it is used with the gain sequence a_k to minimize the decision variable. θ_k is the minimized decision variable estimated at iteration k.

$$\theta_{k+1} = \theta_k - a_k g'_k(\theta_k) \tag{3.4}$$

3.1.2 Algorithm

Implementation of SPSA on a OD matrix calibration problem is done with a few modifications from a basic SPSA algorithm. The stepwise summary for SPSA is described below with figure 3.3 showing the flow chart being followed in its implementation.

1. Initially, θ^k is set to be the h-dimensional starting OD matrix from the previous estimate. The iteration counter i.e. k is set zero and the SPSA parameters c, γ, A, α (except a) are defined, with respect to the problem characteristics and guidance mentioned in Spall (1998b).

2. Setting the no. of gradient replications for obtaining an average gradient estimate at each iteration for θ^k.

3. Incrementing the iteration counter by one and calculation of the gain sequence c_k. a_k is also estimated at this step in each iteration, except for the first iteration.

4. Generation of the h-dimensional random vector Δ_h by Monte Carlo. Δ_h is generated by a probability distribution similar to ± 1 Bernoulli distribution with equal probability.

5. Perturbation of θ^k by equation 3.1 with specific constraints (including lower and upper bound constraints) to result in θ^+ and θ^-. Then, evaluation of objective function with θ^+ and θ^- to obtain y^+ and y^-.

6. Gradient approximation in form of a h-dimensional vector g' based on the difference between y^+ and y^- and step size c_k.

7. Repetition of step 4 till step 6 until the required number of gradients are approximated (specified in step 2) and then calculation of an averaged gradient.

8. Calculation of the SPSA parameter a which specifies the magnitude of step size a_k for minimization based on the evaluated gradient g' so that:

$$a_k = \frac{t \times c_k}{g'} \qquad 0 < t < 1 \tag{3.5a}$$

$$a = a_k(k + A)^{\alpha} \tag{3.5b}$$

where t is a parameter to specify the value of a_k with respect to c_k. This step is done only in the first iteration to specify the starting step size of a_k against the value of parameter g' and c_k.

9. Minimization of θ^k with specific constraints (including lower and upper bound constraints) by the step size a_k upon the evaluated gradient g' by equation 3.4.

10. Termination of the algorithm, depending on the specified criteria i.e. either some specified number of iterations based on a time constraint or stabilized convergence values of θ and its function evaluation y upon several iterations.

3.1.3 Modifications from generic SPSA:

SPSA, being a stochastic random search algorithm is used to find approximations for any stochastic large scale problem. To improve its application on OD estimation problem, a few modifications have been proposed. These modifications listed below are either in form of constraints or feedback upon the major steps in the algorithm, each with its application procedure and description.

- **Non-negative:** Constraints on the decision variable θ_k to stay non-negative during perturbation and minimization. As, the OD matrix cell values cannot be negative, so instead of converting the value as a negative one, its previous state is retained.

- **Relative segmented change:** Another constraint during both perturbation and minimization is that the change is applied relatively based on defined intervals. To describe this constraint mathematically, c_k and a_k will become as:

  $$a_k = \frac{a_k \times i \times n}{\mu} \qquad c_k = \frac{c_k \times i \times n}{\mu}$$

  where i is defined interval size, μ the mean of non-zero values of the vector θ and n is the interval number for which the gain sequence is calculated. n can be $[1, 2, ..., j]$, with $j \times \mu$ the maximum value that can occur in vector θ.

  This method of relative change is important because the OD-estimation problem is under-determined and the variables to be estimated are far more than the observed data. This means that there can be many solutions fulfilling the required results and reliance on the previous estimate is very important in terms of finding a solution having similar demand patterns as the previous one.

- **Gradient feedback:** Also defined above in step 8 of algorithm. The parameter a is defined based on the mean evaluated gradient so that $a_k = t \times c_k$ with $0 < t < 1$.

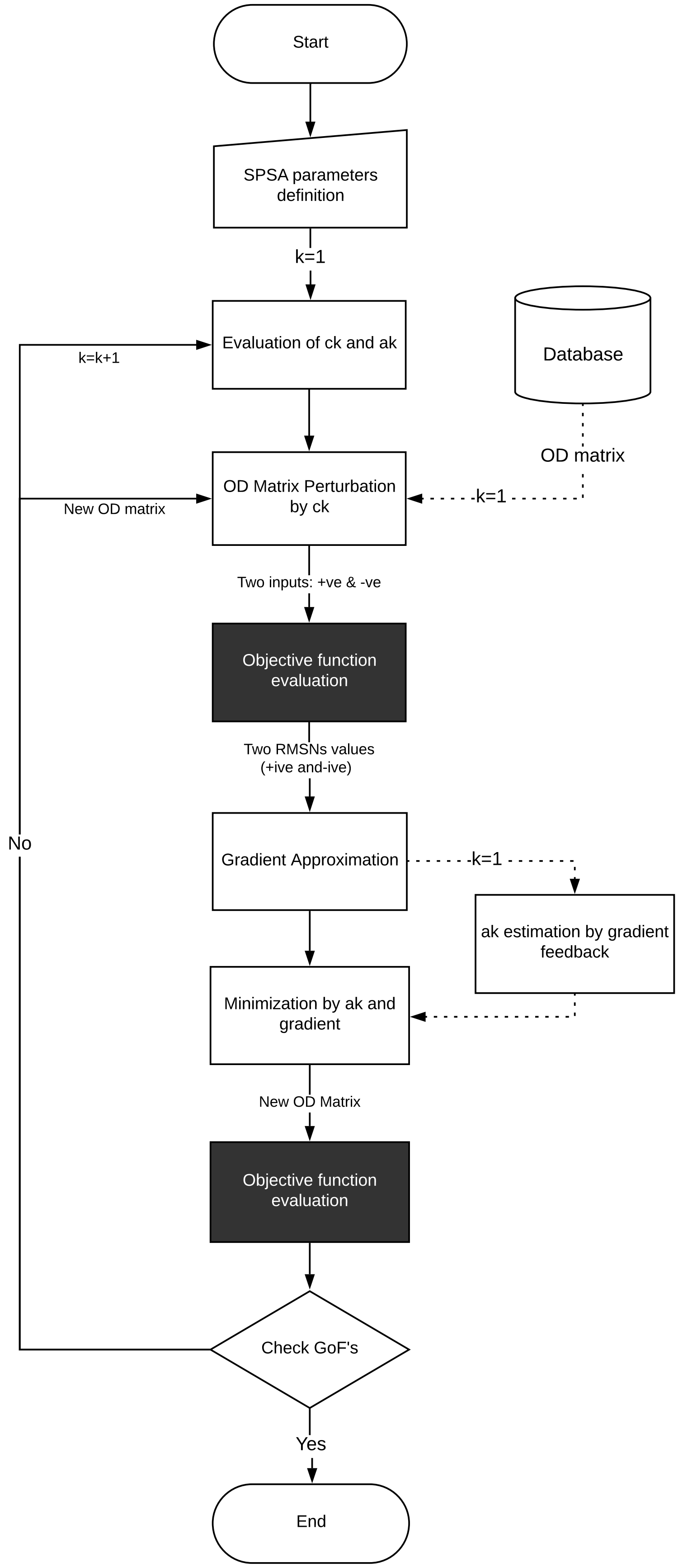

Figure 3.3: SPSA algorithm flow chart

3.2 Principal Component Analysis with SPSA (PC - SPSA)

SPSA being proven to be very effective in stochastic large scale problems, still have limitations against its usage for OD demand estimation. These limitations are due to the factor that OD estimation by counts must rely on previously estimated OD demand. SPSA on the other hand, cannot incorporate any factors from the previous estimate except, its starting variable state to be the preivous estimate. This aspect limits the performance of SPSA for this problem. This section is to define another approach which incorporates a dimension reduction technique (PCA) with SPSA. PCA provides SPSA, the search directions from the variance of previous estimates and reducing the number of estimation variables by limiting its search area within the variance of previous estimates.

3.2.1 Principal component analysis (PCA)

PCA (Jolliffe, 2002) is a technique used to describe the variation in a multivariate dataset in form of a set of uncorrelated variables by extracting those components from data set that can describe most of the variance, reducing a high dimensional data into its lower dimensional PC components. For the application of PCA, a data set is needed have each row or column to be multiple measurements of a single entity.

PCA is actually an eigenvalue decomposition of the covariance matrix $X^T X$, resulting in two matrices V and $\sum$. The columns of V are the eigen vectors or PC vectors representing the direction of variance. While, $\sum$ is a digonal matrix containing the eigen values or PC score corresponding to their respective eigen or PC vectors. The first PC component represents the largest sample variance while the second PC is representing the second largest sample variance subject to being orthogonal to the first one, while continuing till a few of the first PCs represent most of the data variance.

3.2.2 PCA on OD demand

In accordance with Djukic et al. (2012) and Prakash et al. (2017), the application of PCA can significantly reduce the dimensions of OD estimation problem, by capturing the systematic variations of OD flows in lower dimensions. To apply PCA, a data set of multiple estimates for OD flows are required, which can be the previous estimates from daily calibration using an offline or online calibration approach. These estimates are used to create a data matrix X with dimensions of $\{n_k \times n_x\}$. Each row of data matrix n_k represent the number of data points, while each column n_x is the corresponding estimated OD matrix of that point in form of a vector. To evaluate the PCs, single value decomposition is performed on a data matrix resulting in to three matrices U, $\sum$ and V with the following relation:

$$X = U \sum V^T \tag{3.6}$$

where $\sum$ is a rectangular-diagonal matrix with dimensions of $n_k \times n_x$ containing positive PC scores, while U and V are left and right singular vectors of dimensions $\{n_k \times n_k\}$ and $\{n_x \times n_x\}$ respectively. It must be noted that the V matrix containing

PC vectors, only capture the structural relationships between the OD matrix cells (Prakash et al., 2017). As first few of the PCs can capture most of the variance, columns of matrices V and $\sum$ can be reduced from n_x to n_d. The new V matrix will be:

$$V = [v_1 \quad v_2 \quad v_3 \quad ... \quad v_{n_d}]$$

Where v_1, v_2 till v_{n_d} are the PC vectors and a new $\{n_d \times 1\}$ PC score vector can be generated for an OD flow vector x as:

$$z = V^T x \tag{3.7}$$

while, this OD flow x can be reconstructed again as:

$$x = Vz \tag{3.8}$$

The above equations 3.7 and 3.8 are representing the main idea of applying PCA on OD flows, with V containing the variance patterns of previous estimates, dimensions of an OD matrix x can be reduced in form of a new vector z containing its PC scores.

3.2.3 SPSA to PC-SPSA

With the application of PCA, the number of estimation variables are reduced from number of cells in an OD matrix to a few PC scores corresponding to their PC-vectors which contain most of the variance over previous estimates. To transform the approach from SPSA to PC-SPSA, following changes are used:

- Addition of a data matrix consisting of previous estimates and creation of PC directions vector V from this data matrix.
- The decision variable θ is changed to be PC scores vector z instead of the OD matrix for the steps of perturbation and minimization.
- Instead of using the approach of relative segmented change in perturbation and minimization, both steps are applying the change in terms of percentage increase or decrease through multiplication. The definition of gain sequences c_k and a_k must be different, as they are percentage change in the decision variable, as shown by the following relations:

$$\text{Perturbation:} \qquad \theta^{\pm} = \theta \pm \theta \times c_k \Delta \quad . \tag{3.9}$$

$$\text{Minimization:} \qquad \theta^{\pm} = \theta - \theta \times a_k g' \quad . \tag{3.10}$$

- Conversion of PC scores to OD matrix after perturbation and minimization to evaluate the objective function through a DTA model or simulator.

3.2.4 Algorithm

After the amendments to incorporate PCA in SPSA, the new stepwise summary for PC-SPSA is described below. Figure 3.4 provides the flow chart for the steps being followed in the algorithm.

1. Initialization with iteration counter k set as zero and definition of SPSA parameters c, γ, A, α (except a) with respect to the problem characteristics and guidance given by Spall (1998b).

2. Estimation of PC directions V from the data set of previous estimates. Then, evaluating the z score from the latest estimated previous OD matrix and setting it as θ (a d-dimensional vector).

3. Setting the no. of gradient replications for obtaining an average gradient estimate at each iteration for θ^k.

4. Incrementing the iteration counter by one and calculation of the gain sequence c_k. a_k is also estimated at this step in each iteration, except for the first iteration.

5. Generation of the d-dimensional random vector Δ_d by Monte Carlo. Δ_d is generated by a probability distribution similar to ± 1 Bernoulli distribution with equal probability.

6. Perturbation of θ^k by equation 3.9 with specific constraints (like non-negativity) to result in θ^+ and θ^-.

7. Conversion of θ^+ and θ^- in to their respective OD matrices and then, evaluation of objective function by these OD matrices to obtain y^+ and y^- values.

8. Gradient approximation in form a d-dimensional vector g' based on the difference between y^+ and y^- and step size c_k.

9. Repetition of step 5 till step 8 until the required number of gradients are approximated (specified in step 2) and then calculation of an averaged gradient.

10. Calculation of the SPSA parameter a which specifies the magnitude of step size a_k for minimization based on the evaluated gradient g' so that:

$$a_k = t \times c_k \qquad 0 < t < 1 \tag{3.11a}$$

$$a = a_k(k + A)^\alpha \tag{3.11b}$$

    where t is a parameter to specify the value of a_k with respect to c_k. This step is done only in the first iteration to specify the starting step size of a_k against the value of parameter g' and c_k.

11. Minimization of θ^k with specific constraints (including lower and upper bound constraints) by the step size a_k upon the evaluated average gradient g' by equation 3.4.

12. Termination of the algorithm, depending on the specified criteria i.e. either some specified number of iterations based on a time constraint or stabilized convergence of θ and its function evaluation y upon several iterations.

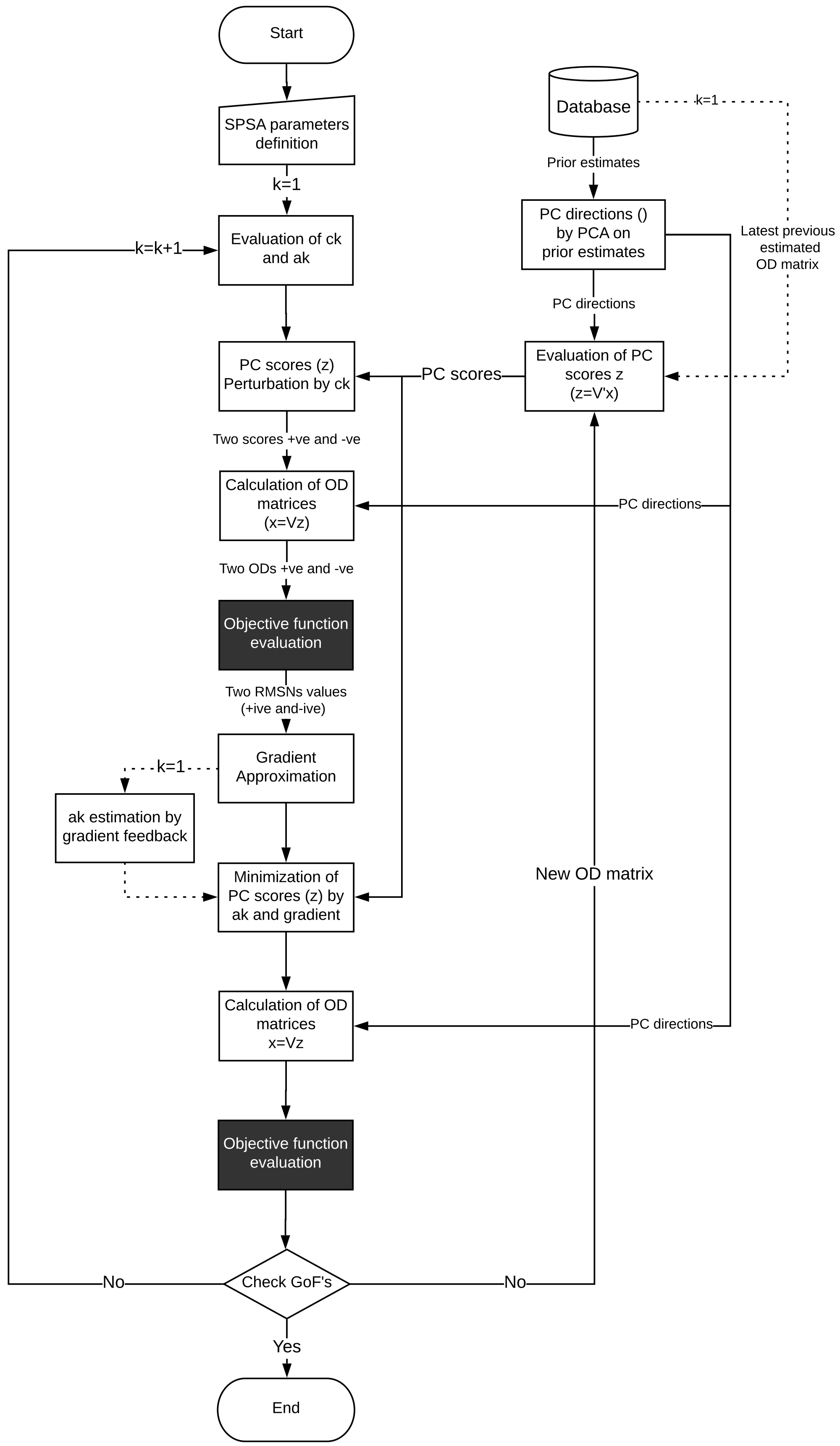

Figure 3.4: PC-SPSA algorithm flow chart

3.3 Conclusion

SPSA and PC-SPSA both use the same basic methodology of perturbing a vector to evaluate its gradient and then minimizing the vector upon it. The major differences between both algorithms are in the steps of perturbation and minimization where SPSA directly changes vector's cell values randomly, either positively or negatively. While, PC-SPSA perturbs and minimizes the PC scores evaluated based on PC directions. These directions are calculated through previous estimates.

OD demand calibration is an underdetermined non-linear problem, so the calibration must rely on previous estimates. A generic SPSA can change the variance pattern significantly making itself very hard to converge towards a solution. The modifications used in SPSA's algorithm, can somewhat retain the variance patterns of the decision variable from its starting state. But in PC-SPSA, with the incorporation of PC vectors or directions capturing the variance of previous estimates, the reliance on previous estimates is very strong. These PC vectors provide SPSA the search directions, with perturbing only the PC scores.

So in summary, application of principal component analysis (PCA) with SPSA can provide two advantages over traditional SPSA. First, it can store the variance from historical OD demand over time in form of PC vectors or directions, which provide SPSA a search direction instead of a random search while the second advantage is that it reduces the dimensions of an OD matrix considerably, reducing the number of variables for calibration. Both advantages can make the calibration alot more rapid, as the number of iterations will decrease significantly.

Chapter 4

Case Studies

This chapter consists of three sections. First section starts with the description of major parameters involved in developing the case studies. Section 2 discusses a case study based on a synthetic non-linear function, including a description of the synthetic function and a comparison for the performance of SPSA and PC-SPSA for three different scenarios. Later, results are shown to analysis the robustness of PC-SPSA against different dimensions and demand patterns. Section 3 describes another case study based on the network of Vitoria, Spain. This section starts with the explanation of the major characteristics involved in the case study and the demand scenarios created to test both, SPSA and PC-SPSA. Later results are shown for the created demand scenarios with corresponding arguments.

4.1 Introduction

Definition of an appropriate testbed is very important to test a calibration algorithm. For the evaluation of the both approaches SPSA and PC-SPSA, two major case studies are defined including different demand scenarios. Both approaches need to have evaluation of an objective function. Figure 4.1 provides the basic structure for evaluation of this objective function. With the inputs of an OD matrix and observed traffic flows, this function is responsible to map the OD matrix in to traffic flows and then evaluate the difference between the observed and simulated traffic flows based on an evaluation criteria. Case studies have been developed based on the approach of converting OD matrix in to traffic flows.

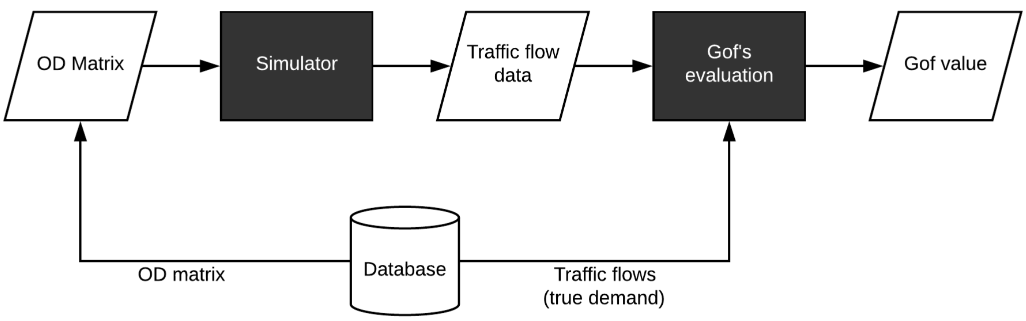

Figure 4.1: Basic structure of an objective function evaluation

In literature, mapping of OD matrix into counts has been a critical aspect. As the relation is very complex and non-linear, the conversion comes by the interaction of a set of complex supply and demand models using a simulator (details in section 2.3.1). Simulating through a DTA model requires time, specially with larger network dimensions.Hence, two types of case studies have been developed. One case study is a simulator based DTA model, using a simulator for the mapping of OD flows into counts. While, the other case study uses a synthetic non-linear function for the conversion.

For the definition of different demand scenarios, an OD demand is assumed as a ground truth and its simulated traffic flows as the observed traffic flows. Then, different target OD matrices are created with the purpose of testing the performance of SPSA and PC-SPSA to calibrate and converge them, towards the true demand.

4.1.1 Evaluation criteria

Evaluation criteria is the goodness of fit measure specified to evaluate the difference between the simulated and observed traffic data. As OD estimation problem is nonlinear, the simulated counts cannot be directly related back to OD flows. Hence, the minimization is done based on this specified goodness of fit criteria. In this thesis, for the evaluation of the objective function shown in figure 4.1, Normalized Root Mean Square (RMSN) error is used as the goodness of fit criteria. RMSN have been chosen previously in many research's (for example, in Ashok and Ben-Akiva (2000), Prakash et al. (2018) etc.). Equation 4.1 gives the mathematical relation for the calculation of RMSNs between simulated and observed values.

$$RMSN = \frac{\sqrt{n \sum_{i=1}^{n} (\hat{y}_i - y_i)^2}}{\sum_{i=1}^{n} y_i} \tag{4.1}$$

where y are the observed traffic flows and $\hat{y}$ are their simulated counterparts. n presents the total number of values and i is from the set $[1, 2, ..., n]$.

4.1.2 Estimation of PC directions

For the application of PC-SPSA, a data matrix of 25 previous estimates is used. Each of the previous estimate is generated using equation 4.2.

$$x_p = x \times (1 - q_p R \Delta) \tag{4.2}$$

where, R is normally distributed random vector with mean 0 and values between 0 and 1. Δ is a ± 1Bernoulli distribution random vector to randomize reduction or increase for each value of x_p. While, q_p is a randomization coefficient to specify the scale of randomization.

After the generation of the data matrix, PC-directions are calculated by principal component analysis (detail in section 3.2.2). Then, these PC-directions are reduced till, the remaining PC-components contain 95% variance of the data matrix. This reduction is done based on the values of $\sum$ vector. Figure 4.2 provide a graph for the cumulative percentage variance explained over the increasing number of principal components (PC). This graph is from the data matrix for the two case study of Vitoria. It shows that the first 80 PCs explain more than 95% of variance from the data matrix.

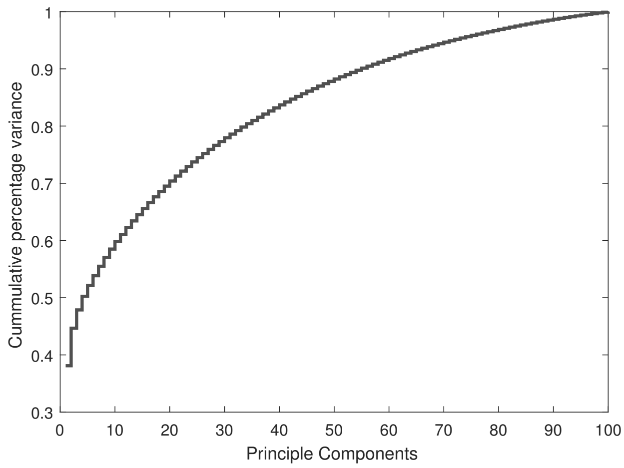

Figure 4.2: RMSN values from SPSA and PC-SPSA for the synthetic function

4.2 Synthetic function

A synthetic function is developed as a case study to skip the use of a simulator to map OD matrix into counts. This case study have been used to test the algorithm during their development because their evaluation time is very less than a simulator based problem. This synthetic problem is using two randomly generated weight matrixs W and W_s as the assignment matrices to map the OD matrix x into counts y. The dimensions of weight matrices are $[x \times y]$. The reason for using two weight matrices is to have a non-linear function, replicating the non-linear OD estimation problem. Equation 4.3 shows the mathematical relation of this synthetic non-linear function, with both x and y as vectors.

$$y = Wx + W_s x^2 \tag{4.3}$$

4.2.1 Demand scenarios

To evaluate the performance of both approaches different demand scenarios have been created with the basic formula shown in equation 4.4

$$x_h = x_{h-1} \times r + x_{h-1} \times q\Delta \tag{4.4}$$

where h is the current interval and x_h is the new target vector, required to achieve after calibration. r is the reduction coefficient, while q is a randomization coefficient and Δ is a random vector similar to the one used in the steps of perturbation and minimization (Chapter 3). d is be used to specify dimensions of x in form of number of zones, so that the dimensions of x as a vector will be $[d^2 \times 1]$.

At first, three demand scenarios are created with different values of r and q, while the dimensions d are set at 60. These scenarios are used to evaluate the performance comparison of PC-SPSA with SPSA. Later, results are generated with PC-SPSA for a range of each coefficient of r, q and d separately.

4.2.2 Results

Results of synthetic function case study are divided into two groups. First group of results compare the performance of SPSA and PC-SPSA for a problem of reasonable dimensions, having three different demand scenarios. While, the second group of results evaluate the robustness of PC-SPSA for different problems defined by a range of dimensions d, reductions r and randomizations q to create the target vector x_h based on the equation 4.1.

Comparison of SPSA and PC-SPSA

To compare the performance of SPSA and PC-SPSA for the defined synthetic function, dimensions d for x are set to be 60, so x will be a vector of $[3600 \times 1]$. The number of sensor counts y are chosen to be 720. The dimensions d of x are based on a comparison to the second case study for the network of Vitoria, which is a reasonable real-life network of a moderate size having the dimensions of 57 zones or $[3249 \times 1]$ OD vector (details in section 4.3.1).

With d set to 60, three scenarios are created from the equation 4.1 with different values of r and q. These three scenarios are:

Scenario 1: $x_h = 0.70x_{h-1} + 0.15x_{h-1}\Delta$

Scenario 2: $x_h = 0.80x_{h-1} + 0.20x_{h-1}\Delta$

Scenario 3: $x_h = 0.70x_{h-1} + 0.25x_{h-1}\Delta$

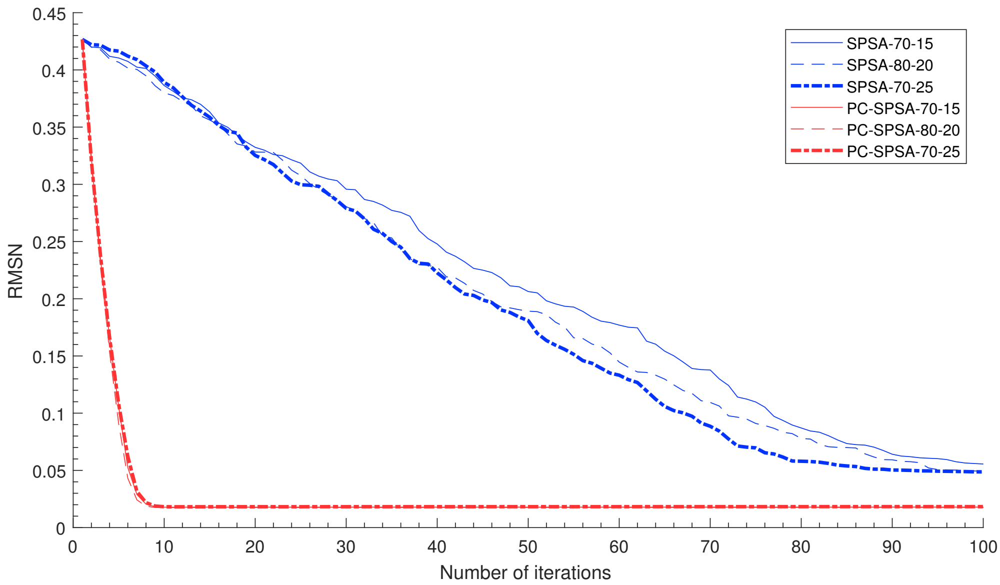

Figure 4.3: RMSN values from SPSA and PC-SPSA for the synthetic function

Figure 4.3 shows the results of calibration done by SPSA and PC-SPSA for all three scenarios. The performance of PC-SPSA in all three scenarios is far better than SPSA. PC-SPSA reduces the RMSN error to a minimal value(i.e. under 3%) within 10 iteration. While, SPSA takes more than 90 iterations to reduce this RMSN error around 5%. These results depict that through the application of PCA, SPSA is able to calibrate OD demand x very rapidly in comparison with the generic

SPSA. Figure 4.4 provides further results of calibrated counts from both approaches against the observed counts for third scenario. As the final RMSN value from the calibration of PC-SPSA is half than the values from SPSA, figure 4.4 shows that the data points from the graph of PC-SPSA are more closer to the 45° plot line than the data points of SPSA.

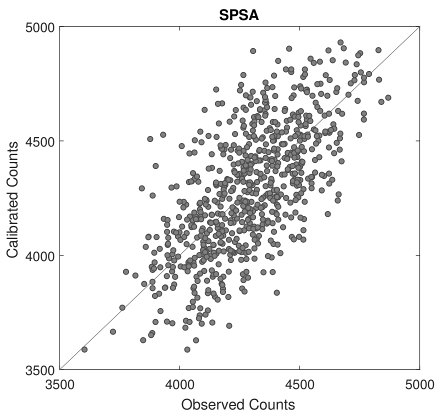

Figure 4.4: Comparison of calibrated and observed counts for scenario 3

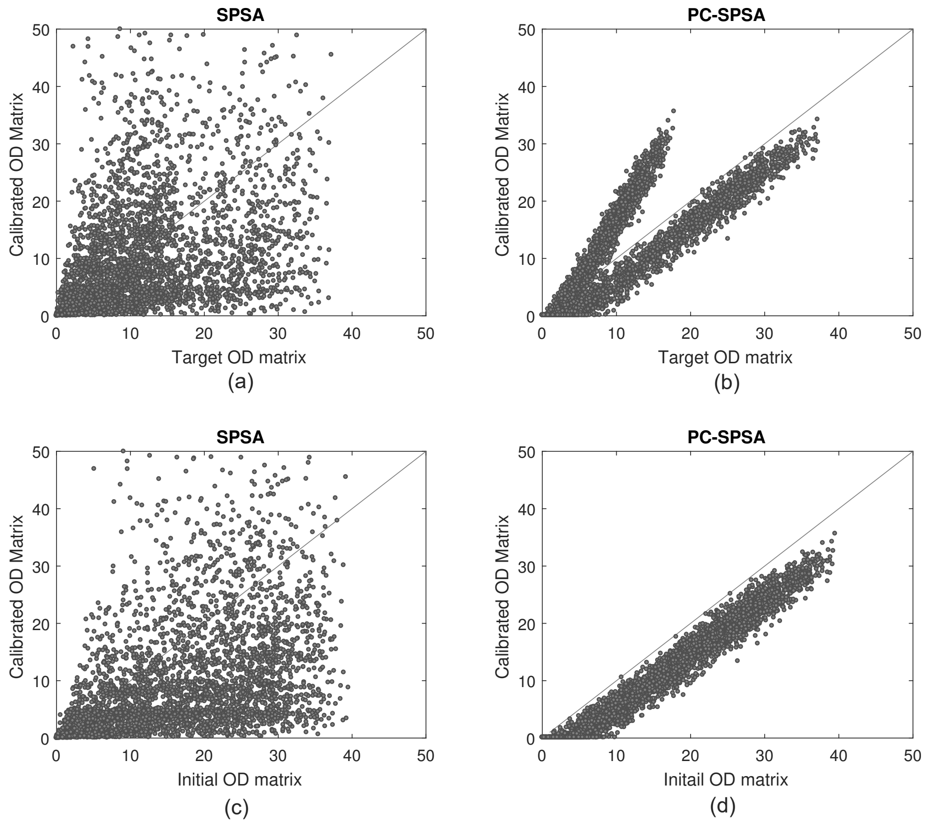

Figure 4.5: Comparison of calibrated OD matrices of SPSA and PC-SPSA with inital and targeted OD matrices for scenario 3

Figure 4.5, graphically shows the reason for the improved performance of PC-SPSA against SPSA. It shows the comparison of the calibrated OD demands $\hat{x}$

against the initial and targeted OD demands for both approaches. As figure 4.5(c) shows that the application of SPSA changes the variance patterns of calibrated OD matrix drastically from their initial values, scattering all the data points of this graph. The reason for having such changes is that SPSA is a random search algorithm and it is perturbing the OD matrix values directly without any search directions, changing its variance in each iteration. While, figure 4.5(d) shows that due to the application of PCA, the variance patterns of the calibrated OD matrix is kept within the variance from the previous estimates. In other words, PC directions evaluated from the previous estimates provide SPSA the directions to search, limiting the changes due to its perturbation within the variance patterns from prior estimates. Figure 4.5(a) and (b) depict the OD demand values calibrated by SPSA and PC-SPSA with reference to the targeted OD demand. Both graphs show similar patterns as in the graphs referred with the initial OD demand. Figure 4.5(a) confirms that calibration by SPSA tends to converge towards a false solution.

Robustness of PC-SPSA

The above results from the comparison of SPSA and PC-SPSA make it evident that PC-SPSA performs far better than SPSA and its rate of convergence is also very high. But the performance of PC-SPSA must be tested against three major factors that might limit its performance. These three major factor are the one used for creating different scenarios i.e. the coefficients of dimension d, reduction r and randomization q. Table 4.1 shows the initial and calibrated RMSN values with the application of PC-SPSA for different dimensions d, randomizations q and a constant 70% reduction. These calibrated RSMN values are only after the first 10 iterations.

Dimensions		Randomization						
		0%	5%	10%	15%	20%	25%	30%
20	Initial	1.022	1.021	0.975	0.917	0.910	0.773	0.718
	Calibrated	0.026	0.115	0.072	0.072	0.085	0.098	0.122
30	Initial	1.011	1.008	0.974	0.988	0.838	0.655	0.713
	Calibrated	0.022	0.041	0.070	0.064	0.066	0.063	0.069
40	Initial	1.008	1.011	0.972	0.933	0.847	0.721	0.731
	Calibrated	0.056	0.070	0.024	0.049	0.068	0.049	0.060
50	Initial	1.007	0.999	0.963	0.902	0.858	0.858	0.662
	Calibrated	0.041	0.048	0.101	0.100	0.038	0.054	0.047
60	Initial	1.006	0.998	0.960	0.930	0.891	0.827	0.769
	Calibrated	0.011	0.018	0.061	0.031	0.040	0.037	0.043
70	Initial	1.006	0.999	0.976	0.938	0.901	0.788	0.741
	Calibrated	0.050	0.020	0.022	0.023	0.049	0.030	0.035
80	Initial	1.006	0.994	0.966	0.933	0.886	0.771	0.746
	Calibrated	0.055	0.051	0.025	0.043	0.023	0.026	0.031
90	Initial	1.005	1.000	0.976	0.916	0.851	0.788	0.709
	Calibrated	0.023	0.050	0.015	0.056	0.021	0.026	0.025

Table 4.1: PC-SPSA calibrated RMSN values for different dimensions and randomization coefficients (after 10 iterations)

The above results demonstrate the performance of PC-SPSA against different dimensional and randomization scenarios. It is clearly evident that within the first 10 iterations it converges rapidly within 10% error. To further test and depict the robustness of PC-SPSA against each of these three coefficients separately, results are evaluated by fixing two coefficients and setting a range of the third coefficient. The group of results for each coefficient is shown below separately.

Dimensions

Dimensionality is one of the major factor that limits the application of online calibration approaches on larger networks. PC-SPSA with the application of PCA reduces the dimensions of the OD demand variables into a few PC scores. Figures 4.6, 4.7 and 4.8 provide the graphs for the calibration of x for the synthetic function case study. These graphs show the results for a range of dimensions of x i.e. from 20 till 90 zones, with different but fixed coefficients of reduction and randomization.

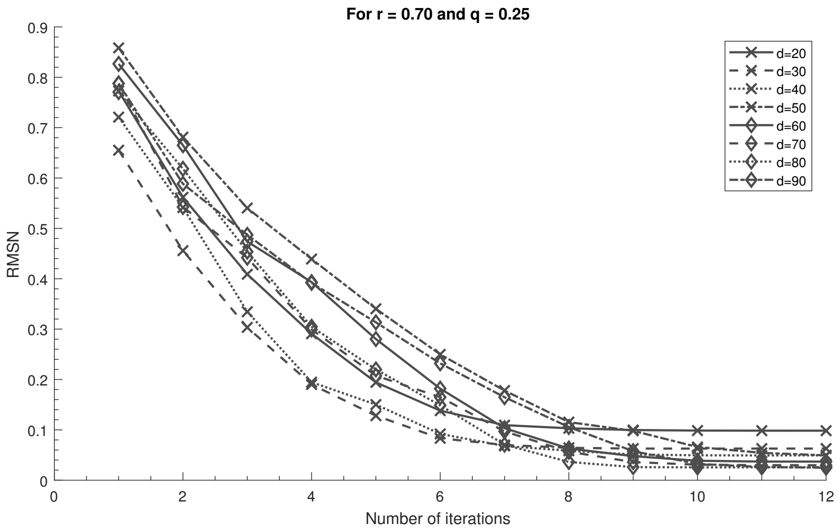

Figure 4.6: RMSN comparison for PC-SPSA against different dimensions of x with 70% reduction and 25% randomization

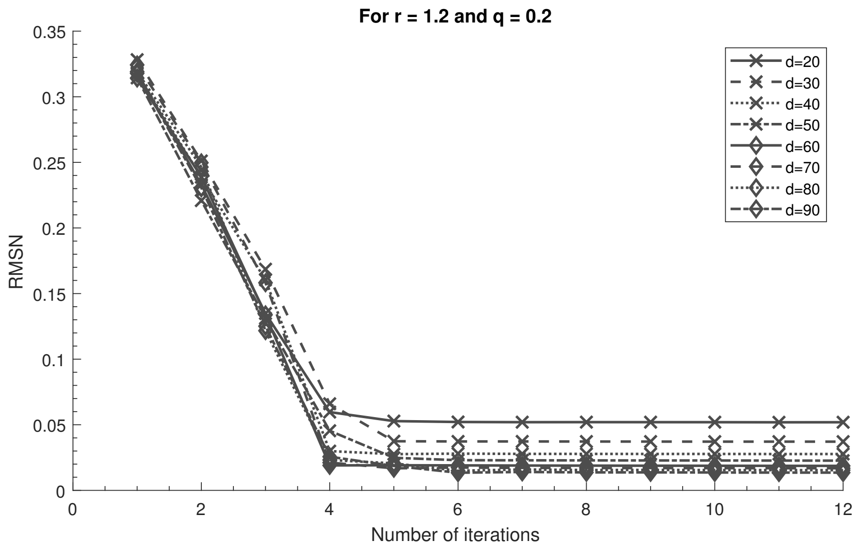

Figure 4.7: RMSN comparison for PC-SPSA against different dimensions of x with 20% increase and 20% randomization

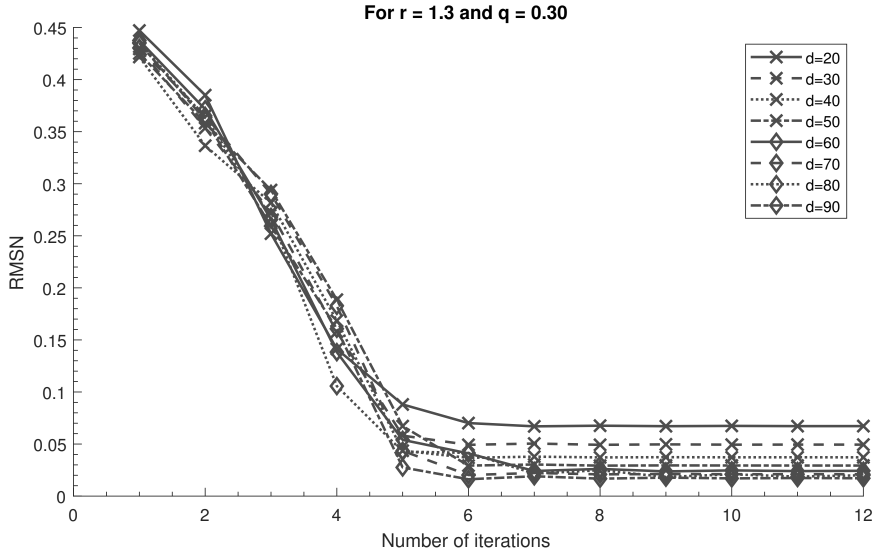

Figure 4.8: RMSN comparison for PC-SPSA against different dimensions of x with 30% increase and 30% randomization

Figures 4.6, 4.7 and 4.8 show the comparison of the performance of PC-SPSA for calibration upon different dimensions of OD demand x. It can be seen that there is no obvious change in the rate of convergance for PC-SPSA due to the increase of dimensions for all three scenarios. The reason is that, PC-SPSA is not perturbing and minimizing the cell values of OD matrix directly. Instead, it perturbs and minimize the PC-scores values. The number of PC-scores does not increase significantly with the increase in dimensions, instead they increase with the increase in the variance of previous estimates as shown in section 4.1.2. While, for the results of this case study, the variance between the previous estimates is kept constant i.e. 30% randomization coefficient q_p for equation 4.2. Hence, the number of PC-score of each scenarios are similar regardless the dimensions of the problem.

Randomization

Another major factor that can effect the performance of PC-SPSA is the amount of change in the variance pattern of the target OD demand from the latest previous estimates used to calculate the PC-scores. To test the robustness of PC-SPSA for different randomization coefficients q, three scenarios are used fixing different values of dimensions d and reduction r while plotting the calibration result for a range of q coefficients. Figures 4.9, 4.10 and 4.11 provide the results for each indicated scenario.

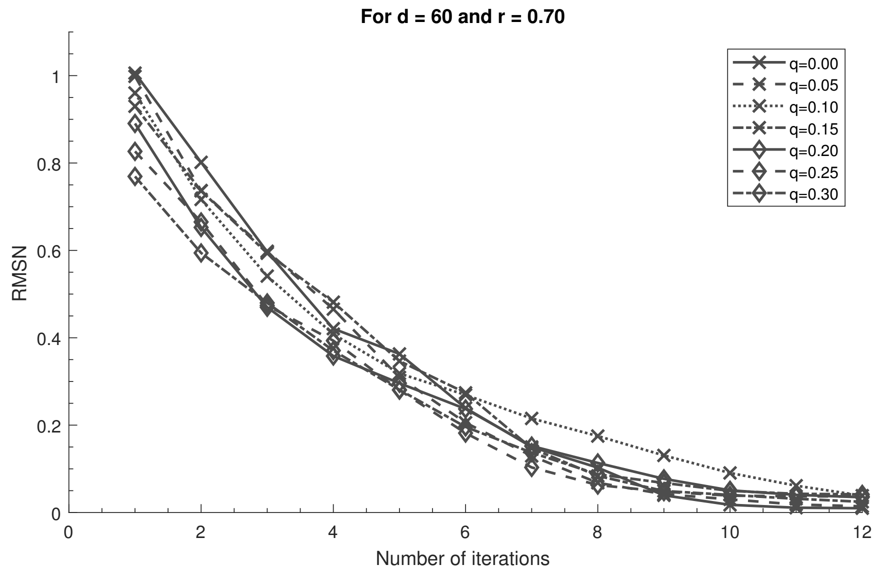

Figure 4.9: RMSN comparison for PC-SPSA against different randomization in x with dimension as 60 and 30% reduction

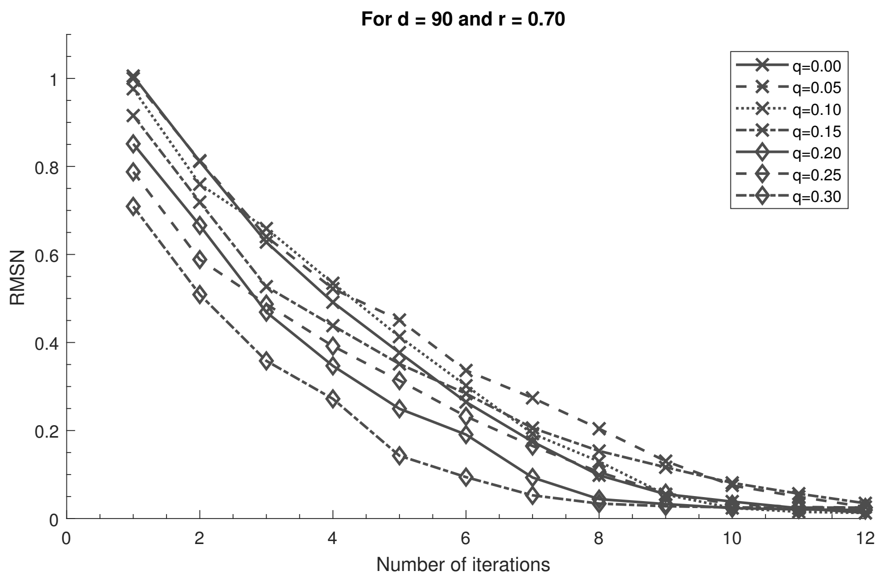

Figure 4.10: RMSN comparison for PC-SPSA against different randomization in x with dimension as 90 and 30% reduction

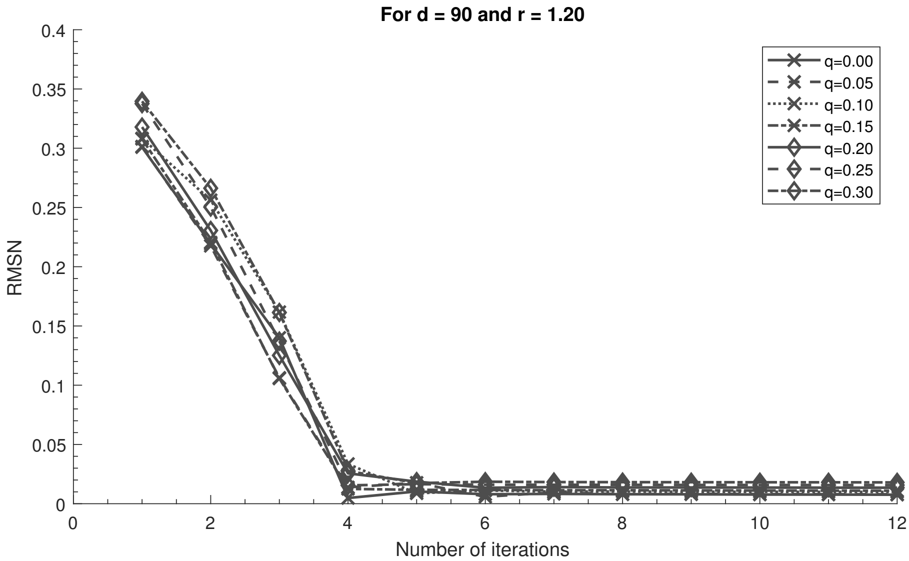

Figure 4.11: RMSN comparison for PC-SPSA against different randomization in x with dimensions as 90 and 20% increase

The above three figures 4.9, 4.10 and 4.11 provide the results for calibration of PC-SPSA against different randomization coefficients. These results have shown consistency due to the factor that the amount of randomization for a target OD matrix against the latest previous estimate does not significantly effect the performance of PC-SPSA, until the variance pattern of the target OD matrix is within or closer to the variance captured in all the previous estimates through PC directions. For PC-SPSA to be robust against this coefficient, the quality or relevance of previous estimates is very important. As the randomization coefficient q_p for generating the previous estimates is set to 30% and the range of randomization coefficients q is from 0 to 30%. Hence the results show consistency for the values of q. A future research can be done for assessing the robustness of PC-SPSA for variance pattern of a target matrix different than that of the previous estimates.

Reductions

Another important factor is the over all change of target demand from the previous estimates. These changes might be a reduced demand due to change in weather conditions (i.e. snow) or an increased demand due to a holiday or a festival. The robustness of PC-SPSA is tested for a range of reduction coefficients r with two scenarios having fixed values of dimensions d and randomization q. Figures 4.12 and 4.13 show the plotted results for the two developed scenarios.

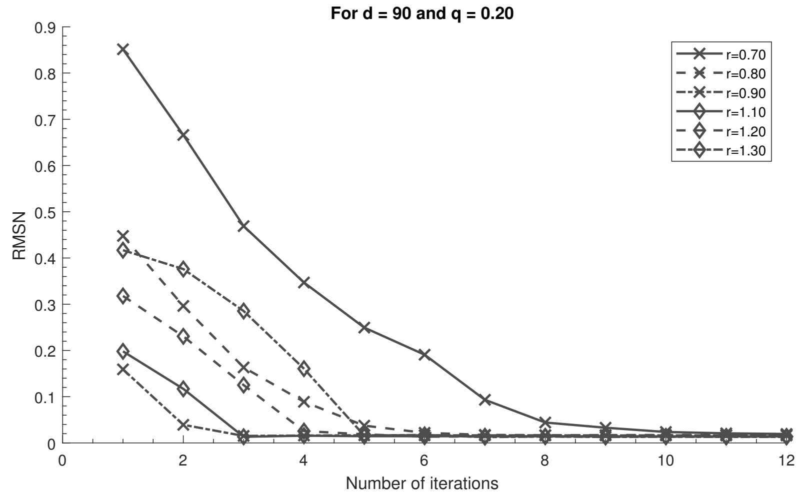

Figure 4.12: RMSN comparison for PC-SPSA against different reductions in x with dimensions as 90 and 20% randomization

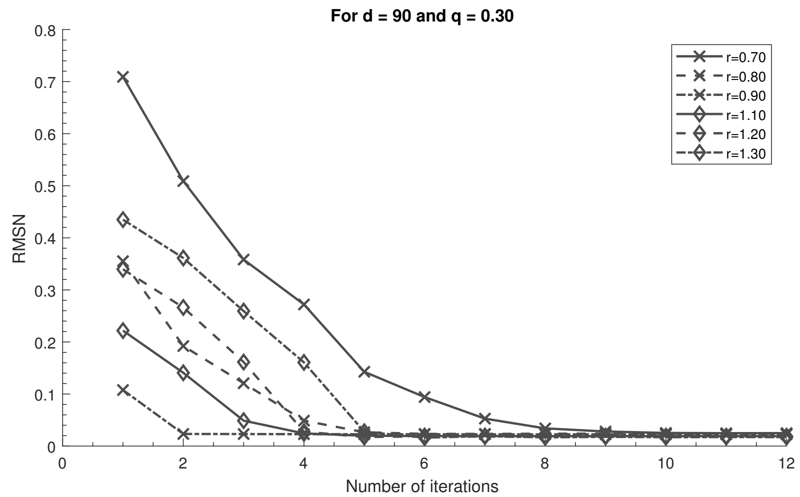

Figure 4.13: RMSN comparison for PC-SPSA against different reductions in x with dimensions as 90 and 30% randomization

The results of PC-SPSA against different scale of changes r in the target demand shown in figures 4.12 and 4.13 depict that the rate of convergence for PC-SPSA is similar overall. This consistency is because of the properties of SPSA. As, SPSA perturbs and minimizes the variable to be estimated, the amount of change in the targeted variable from a starting variable only increase the number of iterations for convergence, while the rate of convergence is the same.

4.3 Simulator based case study

The first case study setup presented the results of SPSA and PC-SPSA using a synthetic non-linear function for mapping OD flows into traffic counts. While, for DTA systems, this mapping is done by simulating the OD flows through a DTA model. This conversion comes by the interaction of a set of complex supply and demand models, which make the OD estimation a complex non-linear problem.

This case study is using a DTA simulator Aimsun (TSS (2014)), to map the OD matrix into traffic flows. The main agenda considered while choosing a traffic simulator is the numerical stability of its results i.e. small perturbation in OD demand should result in small variations in simulated results, consistently (Antoniou et al. (2016)). Secondly, being a mesoscopic model, Aimsun give the benefit of modelling detailed driving behaviour patterns with a far better efficiency of time against microscopic models (as described in section 2.3). This Aimsun based platform with its network and demand is inherited from the research case study of Antoniou et al. (2016).

4.3.1 Vitoria Network

The network used in the simulator based case study is from Vitoria, a city located in Spain. Figure 4.14 shows the geometry of the network consisting a total of 5,799 links, which are about 600 km in length and have 2884 nodes. The network is divided in to 57 zones each having a centroid connected to its specific links for production and attraction of vehicles, in and out of the network.

Figure 4.14: Vitoria network with loop detector locations

This size of the network used, can be considered a reasonable real-life network with its congestion levels and route choices, as in an urban area. The demand is represents an afternoon peak hour. Having a total of five intervals of 15-min each, four intervals consist of the actual one-hour demand with fifth as a pre-warmup period. To detect the traffic flow data, a total of 389 detectors are installed with locations shown in figure 4.14 providing traffic counts, density, speeds and occupancies.

4.3.2 Demand scenarios

The type of traffic demand changes occurring are very important for the performance of a calibration algorithm. Two demand scenarios are created to depict two different regular demand fluctuations. These scenarios depict different changes in daily traffic demand. As mentioned in the first case study, the targeted demand is created by defining two major factors. First factor is for a reduction or increase. While, the second factor is to randomize using a random vector. Considering different values of both factors, two demand scenarios are created as:

• **Scenario 1:** This scenario represents a sudden low demand, reduced 70% from the prior estimate. As in equation 4.5, the target demand is 70% reduced and then randomly perturbed with a factor of 15% using ±1 Bernoulli distribution. So, an OD pair cell can have a value between 55% to 85% of the prior demand estimate.

$$x_h = x_{h-1} * [0.7 + 0.15 \times \Delta] \tag{4.5}$$

where x_h is the new targeted demand of interval h, x_{h-1} is the prior estimate of the previous interval $(h-1)$ and Δ is the ±1 Bernoulli distribution based random vector of dimensions equal to x_{h-1}.

• **Scenario 2:** This scenario represents a target demand of medium reduction but more randomization, reduced only 80% from its prior estimate. As in equation 4.6, the target demand is 80% reduced and then randomly perturbed with a factor of 20% using ±1 Bernoulli distribution. So, an OD pair cell can have a value between 60% to 100% of the prior demand estimate. The purpose of setting this scenario is to test the performance of both algorithms with an increased randomization.

$$x_h = x_{h-1} * [0.8 + 0.20 \times \Delta] \tag{4.6}$$

Both the regular demand scenarios are defined with a reduction and not a single scenario is created with a increase in traffic demand from its prior estimate. The reason for this assumption is the effect of congestion in to the network. If, a scenario is created with an increase in demand, perturbation of the OD matrices might take it higher than the targeting demand which might take the network in to a state of congestion. This state will then act as a point of no return, so the factor of proper congestion mitigation is important which will be described in section 4.6.2.

4.3.3 Traffic Assignment

The network model of vitoria through AIMSUN, provides the basic demand-supply models interaction with some major assumptions. One of the major assumption is about keeping the route choice parameters constant. Apart from the OD Demand, route choice parameter also come under demand models. It is the type of traffic assignment approach used for deriving traffic counts from OD matrix. These parameters are also very critical, as it dictates the level of stochasticity, driver perception and congestion mitigation in the model. The traffic assignment is done with dynamic user equilibrium (DUE) with the method of successive averaging MSA. At first, assignment paths are being generated by the aforementioned approach, then the parameters of assignment are fixed for 100% vehicles to follow these paths. This assumption is done to make the model simple.

The major reason for using DUE assignment is congestion. Congestion mitigation is a very important aspect. As the fundamental diagrams of traffic dynamics are followed under uncongested conditions. When a roadway is near or at capacity, the fundamental diagrams are not followed. The simulators are built on the basics of these fundamental diagrams, due to congestion the results of the simulator might not stay reliable. DUE assignment evaluates the assignment paths with multiple iterations to find the best paths for all vehicles with the minimal travel time incorporating the level of existing volume/capacity ratios, thus reducing the chances of congestion to its minimal.

4.3.4 Iterations

After the definition of assignment paths using DUE assignment approach. The simulation is run through the stochastic scenario with assignment paths to be fixed from DUE assignment. Stochastic scenario incorporates the stochastic nature of drivers behaviour, randomized vehicle input into the network etc. The stochastic patterns are based on a specific random seed. While, each stochastic scenario run provide different results due to a different random seed. A minimum number of simulation runs are needed to make the results statistically significant. Aimsun (2013) provides the guidelines of the number of replications to be 10. The reason specified behind using 10 number of replications is that, every 10 number of generated random seeds are normally distributed, so the mean of 10 replication outputs, provide the appropriate results.

4.3.5 Estimation of PC directions

For the application of PC-SPSA, a data matrix of 25 previous estimates is used. Each of the previous estimate is generated using equation 4.2, with q_p set to 30%. Then, PC-directions are calculated using principal component analysis on the data matrix. Afterwards, these PC-directions are reduced to 80, which contained 95% variance of the data matrix as shown in figure 4.2. So, with the application of PCA, the number of estimating variables are reduced from 3249 variables (i.e. $[57 \times 57]$ OD matrix) to just 80 PC-scores.

4.3.6 Results

The results of this case study is based on two demand scenarios defined before. For each scenario, a comparison of the RMSN values is given for both SPSA and PC-SPSA. Then, the comparison of the calibrated OD matrix and their simulated traffic counts is provided with the corresponding true demand and its observed or true counts, on a 45° plot. Results of each scenario are shown and discussed below:

Scenario 1

This scenario contains a target demand, set by 70% reduction and 15% randomization from the latest previous estimate. The demand contains four 15-min intervals and the calibration is done for each interval separately. Three types of results are provided in figures 4.15, 4.16 and 4.17, each discussed below:

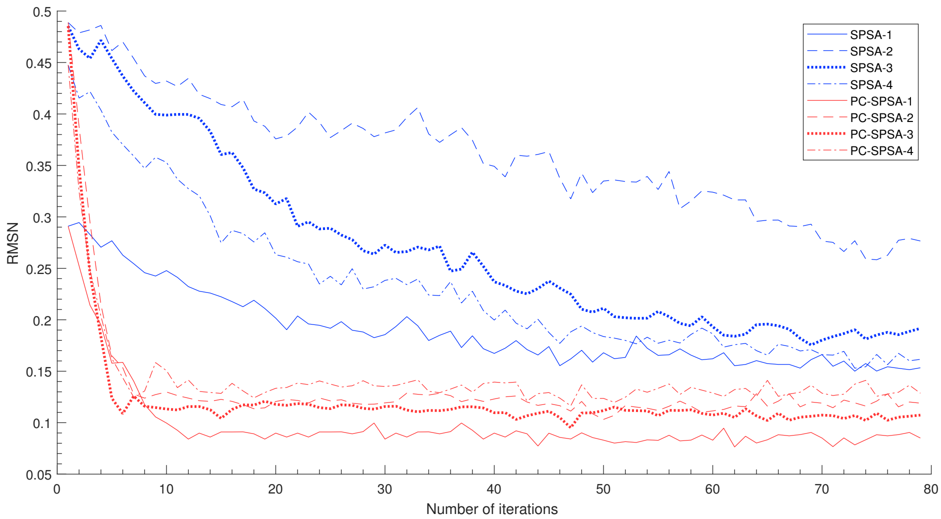

Figure 4.15: Comparison of RMSN values for SPSA and PC-SPSA

Figure 4.15 shows the RMSN values by the calibration of SPSA and PC-SPSA over a run of 80 iterations. The performance of PC-SPSA is far better than SPSA and it reduces the RMSN error value less than 15% for all four demand intervals within first 10 iterations. While, SPSA is able to reduce the RMSN error value around 20% for three intervals while for interval number 2 the error remains more than 30%. This results validate the performance of PC-SPSA in synthetic function case study, where PC-SPSA has shown similar convergence. For SPSA, results for the synthetic case study are better than this case study because the first case study is based on just a non-linear function, while this case study incorporate the complex non-linearity of a moderate size DTA model network. Figure 4.16 gives further results for the calibrated counts from both approaches against the observed counts. This graph being created for the counts of second demand interval, depicts that as the RMSN value from the calibration of PC-SPSA is far reduced than that of SPSA, the data points representing calibrating counts from PC-SPSA are also more closer to the 45° plot line than the counts by SPSA.

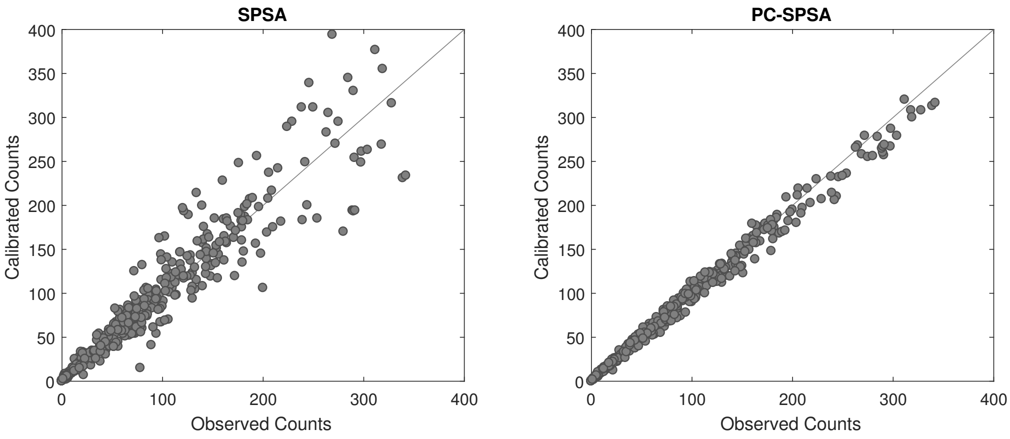

Figure 4.16: Comparison of calibrated and observed counts of SPSA and PC-SPSA for scenario 1

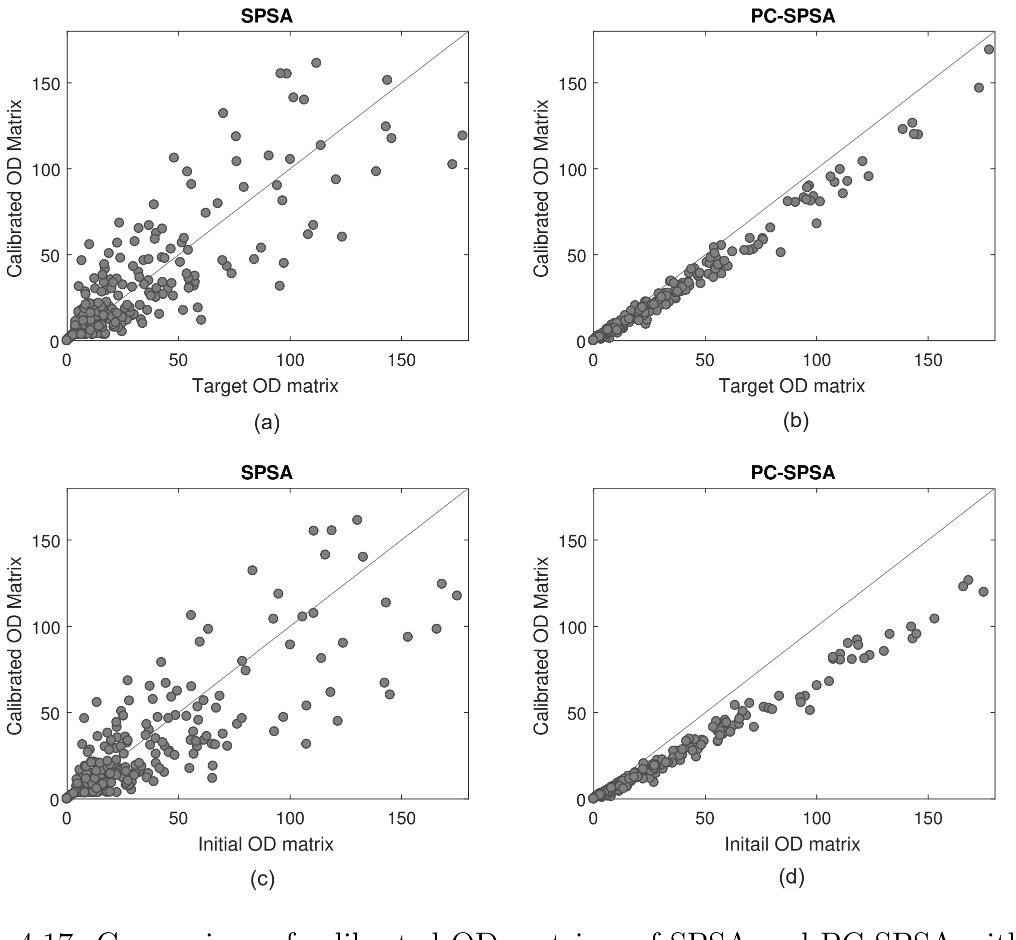

Figure 4.17: Comparison of calibrated OD matrices of SPSA and PC-SPSA with inital and targeted OD matrices for scenario 1

Figure 4.17 provides the comparison of the calibrated OD matrix for the second demand interval from both SPSA and PC-SPSA. These results show similar patterns from the synthetic case study, as the calibrated OD matrix from SPSA has very scattered plot of data points shown in figure 4.17(a) and (c) .While, for the calibrated OD matrix of PC-SPSA the plots shown in figure 4.17(b) and (d) are much closer to the 45° plot line, as its patterns is changed within the search area provided by PC-directions.

Scenario 2

This scenario contains a target demand with more randomization and less reduction from the previous one. As the coefficients are set by to have 80% reduction and 20% randomization from the latest previous estimate. The number of iterations done for this scenario are 50 and only the graphs for the values of RMSN and comparison of the calibrated OD matrices are provided for this case study.

Assessing the performance of SPSA and PC-SPSA from the figure 4.18, it can be

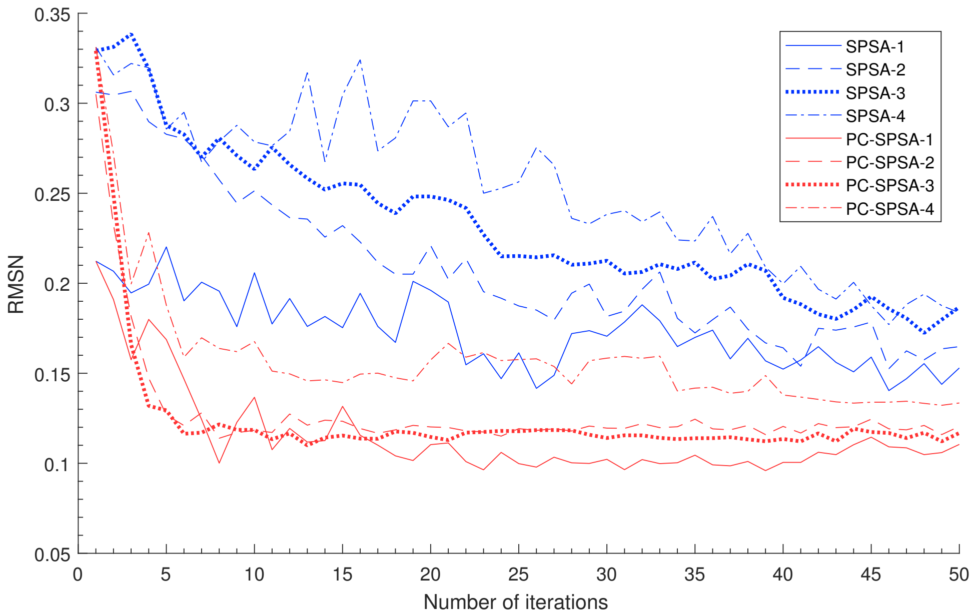

Figure 4.18: Comparison of RMSN values for SPSA and PC-SPSA

seen that the convergence patterns for both approaches are similar to the previous scenario. As this scenario included only a 20% reduction for creating the target OD matrix, the starting RMSN values are less than the starting RMSN values of the previous estimate. While, due to the increased randomization the pattern of convergence for both approaches seems to be more irregular comparing to the previous scenario.

Figure 4.19 shows the results for the comparison of calibrated OD matrices from both approaches with the initial and target OD matrix. These results are also a replication of the results shown in the previous scenario, except for the comparison of the calibrated OD matrix of PC-SPSA with the target matrix shown in figure 4.19(b). Due to the increase in randomization for setting the target OD matrix, the calibrated OD matrix of PC-SPSA seems to show some scatteredness along the 45° plot line.

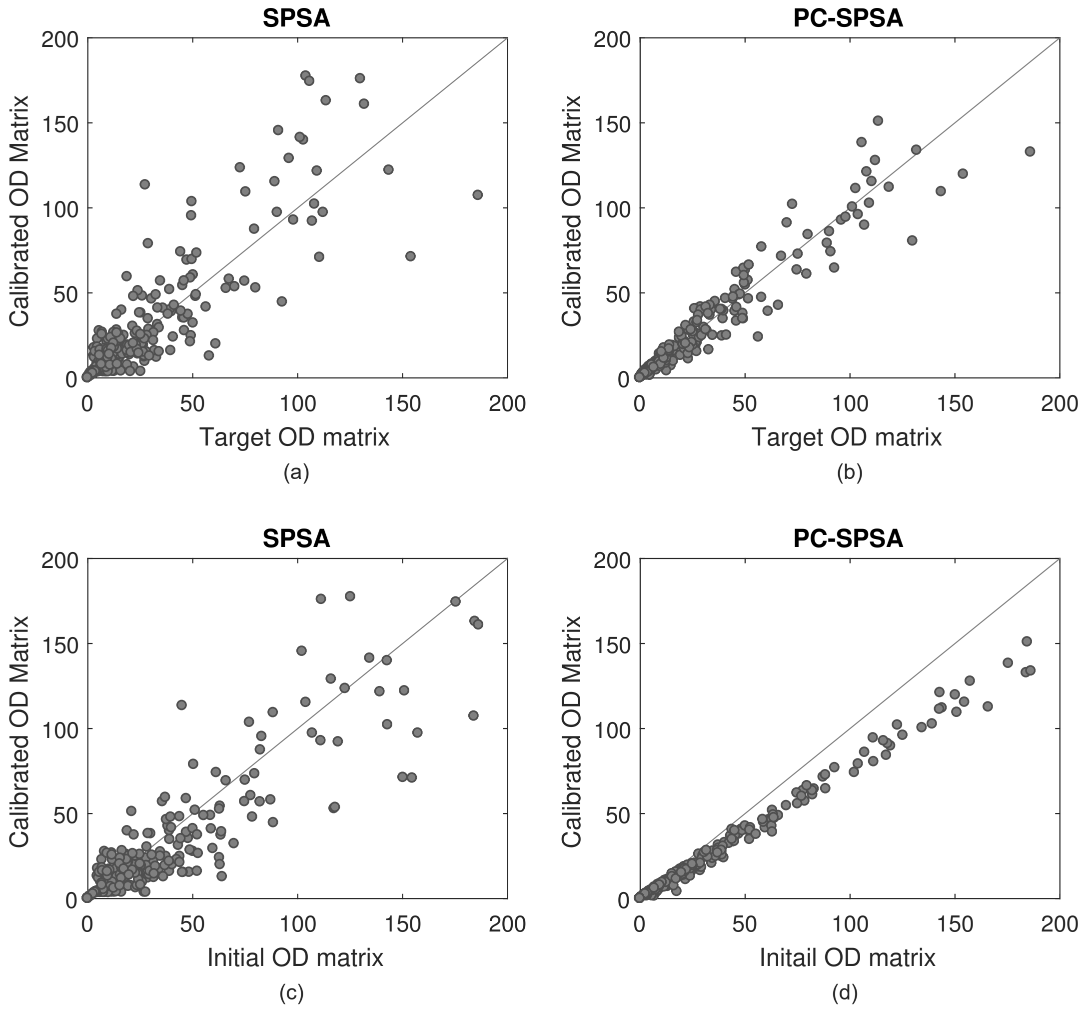

Figure 4.19: Comparison of calibrated OD matrices of SPSA and PC-SPSA with inital and targeted OD matrices for scenario 2

Chapter 5

Conclusion

5.1 Summary

An online calibration approach for dynamic traffic assignment has been proposed, named as PC-SPSA. This approach addresses the issues of non-linearity and dimensionality in a more effective manner than generic SPSA and has the potential to improve the scale of application for online calibration on larger networks due to its computational performance.

The purpose of an online calibration approach is to dynamically update the model parameters using the observed traffic data, so that the model outputs are closer to the observed ones. This calibration needs to be time efficient, as the estimated parameters are used as an input in DTA systems for further prediction and control. Hence, the scope of utilization for an online calibration approach is limited by its computational time. For an online calibration approach, dimensions of a model and its complexity, limit its application on the model. As the dimensions increase the number of estimating parameters are also increased, while as it gets more complex the time required for simulating the model during calibration increases. These two parameters play the major role in defining the performance of an online calibration approach.

Many of the classical methods of calibration have been mostly based on the state space model. For most of these developed approaches the calibration problem is formulated by direct and indirect measurement equations, with direct measurement equation relying the calibration on prior estimates and the indirect measurement equations relying on the surveillance data. From most of the previously developed approaches, it can be concluded that calibration from an online calibration approach must rely on the previous estimate, to make it time efficient. Recent efforts have been focusing on incorporating the previous estimates as well as reducing the dimension of the online calibration problem with the application of Principal Component Analysis (PCA), resulting into some fruitful conclusions.

The approach developed in this thesis has used the similar concept of applying Principal Component Analysis (PCA) for reducing the dimensions of the problem. But instead of using a state space model, the solution is developed using PCA with a stochastic approximation algorithm, Simultaneous Perturbation Stochastic Approximation (SPSA), named as PC-SPSA. SPSA is a random search algorithm with an advantage of having a fixed number of objective function evaluations (two) for approximating its gradient. To develop the approach, PCA is applied on a set of previous estimates of OD demand matrices. Which evaluates the structural spatial relationship within the OD matrix over all the previous estimates in the form of PC-directions. Then, an OD matrix is calibrated by converting it into the PC-scores using these directions. PCA can capture most of the structural variance from the previous estimates into a few principal components, hence the number of PC-scores against the dimensions of an OD matrix are reduced significantly. These PC-scores are then used as the estimation variables for SPSA. Due to the properties of SPSA, PC-SPSA also uses only two evaluations of the objective function within each iteration and as the number of estimation variables are reduced significantly this approach converges very rapidly.

Two case studies have been developed to test the performance of PC-SPSA in comparison of SPSA. Results from both case studies have been very promising for PC-SPSA. Especially, the second case study which is based on a moderately sized network of Vitoria, provides the validation that PC-SPSA converges the solution within very few iterations. Further analysis is also done to evaluate the factors that can potentially effect the performance of PC-SPSA. But within all the results provided, PC-SPSA have proven itself to be a robust online calibration algorithm, especially against the factor of dimensionality.

The effectiveness of PC-SPSA hides under few aspects that arise due to the combination of PCA with SPSA. These aspects can be concluded as:

- The application of PCA significantly reduces the dimension of estimation variables (e.g. in second case study the dimension are reduced by a factor of 40 even keeping 95% cumulative variance)

- As the dimension reduction is dependent on the cumulative variance, so the scale of reduction is dependent on the amount of variance present in the previous estimate. In other words, the scale of reduction can reduce with the increase of number of previous estimate or if they variate a lot from each other.

- The reduction in the performance of PC-SPSA is not directly proportional to the increase in dimension, either it relates with the amount of variation present in previous estimates. But mostly for a traffic network with the increase in dimensions the variance can increase significantly.

- As, PC-scores are used to calibrate the OD matrix, the search area of SPSA is limited within the variance patterns captured in from the previous estimates. The means that the performance of PC-SPSA depends on the quality of previous estimates. The more the variance pattern of a targeted demand variable is away from the patterns of the prior estimates, the harder it is for PC-SPSA to converge towards a better solution. This implicates that the performs of PC-SPSA can be reduced due to an implusive change in the targeted demand.

In consideration with the above mentioned aspects, it can be concluded that PC-SPSA have the potential to be an effective online calibration algorithm. The performance of PC-SPSA depends on the relevance of previous estimates with the targeted estimate and the scale of variance present within previous estimates. While the case study of Vitoria has shown promising results, PC-SPSA can be utilized as an online calibration algorithm and with appropriate computational infrastructure and utilization of the concept like parallel computing, the time of taken for each iteration can be further reduced significantly.

5.2 Future works

5.2.1 PC-SPSA

Further potential research areas under PC-SPSA can be summarized as:

- First potential research can be to implement PC-SPSA for demand scenarios with changes due to special events (i.e. football match, carnival etc.). These demand scenarios can have an impulsive demand changes for demand being attracted to a few number of zones changing its structural patterns. Research can be done to test and evaluate PC-SPSA's performance and find the potential factors that might be limiting its performance.

- Another potential research can be to find the methods to incorporate the historical estimates for the specific events causing impulsive demand changes to evaluate the PC-directions which can perform better against these impulsive change. Or proposed a methodology of providing a database of different PC-directions based on a different regular and impulsive historical estimates and utilize them alternatively to find the most calibrated solution.

- Comparison of PC-SPSA with the alternative online calibration approaches that have been proposed to evaluate and compare their performance under different scenarios.

- Research on improving the methodology followed for PC-SPSA e.g. applying the relative segmented perturbation and minimization in PC-SPSA which might improve its converging performance.

- Extension of PC-SPSA into the entire spectrum of model variables, including the supply side parameter, where PCA can be more suitable to extract the relationship between the model variables, thus allowing for the development of more detailed traffic simulation models with better computational efficiency

5.2.2 SPSA

An area of modification can be researched for the application of SPSA for OD demand calibration. This modification can be that once the decision variable and its function evaluation stabilizes but the required criteria of goodness of fit is not achieved. The gain sequence of c_k and a_k can be kicked to larger values and a number of outputs can be obtained. Then, the previous algorithm can be restarted with the lowest output from this swarm search. This method can help the search when its stuck on a local minimum, with the swarm search skipping the local minimum and move it to a global optimum. But, the consequence of this approach can be that it can change the patterns in the OD matrix drastically, changing it from what should have been achieved based in the historical estimates.

Bibliography

Aimsun (2013). Aimsun Traffic Simulation Model Development Manual. Technical Report October, Goverment of South Australia.

Antoniou, C. (2004). *On-line calibration for dynamic traffic assignment*. PhD thesis.

Antoniou, C., Azevedo, C. L., Lu, L., Pereira, F., and Ben-Akiva, M. (2015). W-SPSA in Practice: Approximation of Weight Matrices and Calibration of Traffic Simulation Models. *Transportation Research Procedia*, 7(April):233–253.

Antoniou, C., Balakrishna, R., Koutsopoulos, H. N., and Ben-Akiva, M. (2009). Off-line and on-line calibration of dynamic traffic assignment systems. In *IFAC Proceedings Volumes (IFAC-PapersOnline)*, volume 42, pages 104–111. IFAC.

Antoniou, C., Barceló, J., Breen, M., Bullejos, M., Casas, J., Cipriani, E., Ciuffo, B., Djukic, T., Hoogendoorn, S., Marzano, V., Montero, L., Nigro, M., Perarnau, J., Punzo, V., Toledo, T., and van Lint, H. (2016). Towards a generic benchmarking platform for origin-destination flows estimation/updating algorithms: Design, demonstration and validation. *Transportation Research Part C: Emerging Technologies*, 66:79–98.

Antoniou, C., Ben-Akiva, M., and Koutsopoulos, H. (2005). On-line calibration of traffic prediction models. *Transportation Research Record: Journal of the Transportation Research Board*, (1934):235–245.

Antoniou, C., Ben-Akiva, M. M., and Koutsopoulos, H. N. (2007a). Nonlinear Kalman filtering algorithms for on-line calibration of dynamic traffic assignment models. *Ieee Transactions on Intelligent Transportation Systems*, 8(4):661–670.

Antoniou, C., Koutsopoulos, H. N., and Yannis, G. (2007b). An Efficient Non-Linear Kalman Filtering Algorithm Using Simultaneous Perturbation and Applications in Traffic Estimation and Prediction. *Proceedings of the 10th International IEEE Conference on Intelligent Transportation Systems (ITSC 2007)*, pages 217–222.

Ashok, K. (1996). Estimation and Prediction of Time-Dependent Origin-Destination Flows. page 155.

Ashok, K. and Ben-Akiva, M. (2000). Alternative approaches for real-time estimation and prediction of time-dependent origin-destination flows. *Transportation Science*, 34(1993):21–36.

Balakrishna, R. (2006). *Off-line calibration of dynamic traffic assignment models*. PhD thesis.

Balakrishna, R., Ben-Akiva, M. E., and Koutsopoulos, H. N. (2007). Offline Calibration of Dynamic Traffic Assignment: Simultaneous Demand-and-Supply Estimation. *Transportation Research Record: Journal of the Transportation Research Board*, 2003:50–58.

Balakrishna, R., Koutsopoulos, H. N., and Ben-Akiva, M. E. (2005). Calibration and Validation of Dynamic Traffic Assignment Systems. In *Proceedings of the 16th International Symposium on Transportation and Traffic Theory: Flow, Dynamics and Human Interaction.*, pages 407–426.

Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H. N., and Mishalani, R. (2002). *Real Time Simulation of Traffic Demand-Supply Interactions within DynaMIT*, pages 19–36. Springer US, Boston, MA.

Ben-Akiva, M., Koutsopoulos, H. N., Antoniou, C., and Balakrishna, R. (2010). *Traffic simulation with DynaMIT*, volume 145, pages 363–398. Springer New York, New York, NY.

Cascetta, E. (1984). Estimation of trip matrices from traffic counts and survey data: A generalized least squares estimator. *Transportation Research Part B*, 18(4-5):289–299.

Cascetta, E., Inaudi, D., and Marquis, G. (1993). Dynamic Estimators of Origin Destination Matrices Using Traffic Counts. *Transportation Science*, 27(4):363–373.

Chui, C. K. and Chen, G. (1999). *Kalman filtering: With real-time applications.*

Cochran, W. G. (2007). Sampling Techniques.

Djukic, T. (2014). *Dynamic OD demand estimation and prediction for dynamic traffic management*. PhD thesis.

Djukic, T., Flotterod, G., Van Lint, H. J., and Hoogendoorn, S. (2012). Efficient real time OD matrix estimation based on Principal Component Analysis. *Intelligent Transportation Systems (ITSC), 2012 15th International IEEE*, (October):115–121.

Hazelton, M. L. (2000). Estimation of origin-destination matrices from link flows on uncongested networks. *Transportation Research Part B: Methodological*, 34(7):549–566.

INRIX (2014). Annual Cost of Gridlock in Europe and the US will Increase 50 Percent on Average to $293 Billion by 2030 — INRIX.

Jolliffe, I. T. (2002). Principal Component Analysis, Second Edition. *Springer Series in Statistics*, 98:487.

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82(1):35.

Kiefer, J. and Wolfowitz, J. (1952). Stochastic Estimation of the Maximum of a Regression Function. *The Annals of Mathematical Statistics*, 23(3):462–466.

Lu, L., Xu, Y., Antoniou, C., and Ben-Akiva, M. (2015). An enhanced SPSA algorithm for the calibration of Dynamic Traffic Assignment models. *Transportation Research Part C: Emerging Technologies*, 51:149–166.

Maher, M. J. (1983). Inferences on trip matrices from observations on link volumes: A Bayesian statistical approach. *Transportation Research Part B*, 17(6):435–447.

Mahmassani, H. S., H., T. H., and Jayakrishnan, R. (1995). Dynamic traffic assignment and simulation for advanced network informatics (DynaSMART). In *Urban traffic networks: Dynamic flow modeling and control*. Springer.

OICA (2015). Vehicles in Use.

Okutani, I. and Stephanedes, Y. J. (1984). Dynamic prediction of traffic volume through Kalman filtering theory. *Transportation Research Part B*, 18(1):1–11.

Olstam, J. and Tapani, A. (2011). A review of guidelines for applying traffic simulation to level-of-service analysis. In *Procedia - Social and Behavioral Sciences*, volume 16, pages 771–780.

Ortuzar, J. d. D. and Willumsen, L. G. (2011). *Modelling Transport*.

Peeta, S. and Ziliaskopoulos, A. K. (2001). Foundations of Dynamic Traffic Assignment: The Past, the Present and the Future. *Networks and Spatial Economics*, 1(3):233–265.

Prakash, A. A., Seshadri, R., Antoniou, C., Pereira, F., and Ben-Akiva, M. E. (2018). Improving Scalability of Generic Online Calibration for Real-Time Dynamic Traffic Assignment Systems. *Transportation Research Record: TRB*, (in press).

Prakash, A. A., Seshadri, R., Antoniou, C., Pereira, F. C., and Ben-Akiva, M. E. (2017). Reducing the Dimension of Online Calibration in Dynamic Traffic Assignment Systems. *Transportation Research Record: TRB*, 2667(March 2017):96–107.

Spall, J. (2003). Stochastic Approximation and the Finite-Difference Method. In *Introduction to Stochastic Search and Optimization: Estimation, Simulation and Control*, number Chapter 6, pages 150–175.

Spall, J. C. (1998a). An Overview of the Simultaneous Perturbation Method for Efficient Optimization. *Johns Hopkins Apl Technical Digest Volume: 19, Issue: 4*, 19(4):482–492.

Spall, J. C. (1998b). Implementation of the simultaneous perturbation algorithm for stochastic optimization.

TRB (2011). Dynamic Traffic Assignment. A. Primer. Technical Report June.

TSS, T. S. S. (2014). *Aimsun 8 Users Manual*. Transport Simulation Systems, Barcelona, Spain.

Van Zuylen, H. J. and Willumsen, L. G. (1980). The most likely trip matrix estimated from traffic counts. *Transportation Research Part B: Methodological*, 14(3):281–293.

Wardrop, J. G. (1952). Some Theoretical Aspects of Road Traffic Research.

Willumsen, L. G. (1978). Estimation of an O-D matrix from traffic counts - a review. *Online*, 64:49–67.

Zhou, X. (2004). Dynamic origin-destination demand estimation and prediction for off-line and on-line dynamic traffic assignment operation.

Zhou, X. and Mahmassani, H. S. (2007). A structural state space model for real-time traffic origin-destination demand estimation and prediction in a day-to-day learning framework. *Transportation Research Part B: Methodological*, 41(8):823–840.