

PC-SPSA: Employing dimensionality reduction to limit SPSA search noise in DTA model calibration

Moeid Qurashi*, Tao Ma, Emmanouil Chaniotakis, and Constantinos Antoniou

Abstract—Calibration and validation have long been a significant topic in traffic model development. In fact, when moving to Dynamic Traffic Assignment (DTA) models, the need to dynamically update the demand and supply component creates a considerable burden on the existing calibration algorithms, often rendering them impractical. These calibration approaches are mostly restricted either due to non-linearity or increasing problem dimensionality. Simultaneous Perturbation Stochastic Approximation (SPSA) has been proposed for DTA model calibration, with encouraging results, for more than a decade. However, it often fails to converge reasonably with the increase in problem size and complexity. In this research, we combine SPSA with Principal Components Analysis (PCA), to form a new algorithm we call PC-SPSA. PCA limits the search area of SPSA within the structural relationships captured from historical estimates in lower dimensions, reducing the problem size and complexity. We formulate the algorithm, demonstrate its operation and explore its performance using an urban network of Vitoria, Spain. Practical issues that emerge from the scale of different variables and bounding their values are also analyzed through a sensitivity analysis using a non-linear synthetic function.

Index Terms—Model calibration, principal component analysis (PCA), simultaneous perturbation stochastic approximation (SPSA).

I. INTRODUCTION

E VOLUTION of the automobile industry entirely changed the prospect of traveling for mankind. Serving as the most widely used mode of transportation, it comes at a price of congestion, costing us socially, economically and environmentally. Practitioners worldwide try to tackle this problem through real-time traffic management and operations based on Dynamic Traffic Assignment (DTA) systems. However, the efficacy of this practice is based on the level of accuracy for traffic state estimation and prediction, resulting from DTA models. Hence, calibration of DTA models is a crucial aspect towards the performance of these systems.

Dynamic Traffic Assignment (DTA) models are based on the interaction of demand parameters [including Origin-Destination (OD) flows, travel behaviors] with network characteristics (link capacities, speed-density relationships), to simulate road traffic conditions. Calibration of DTA models involves adjustment of these supply and demand parameters using observed traffic measurements. These adjustments are mostly done both offline (for long-term traffic planning)

[12] and online (for real-time traffic management) [3]. The calibration approaches developed so far include two types of formulations: assignment matrix based and non-assignment matrix based. Assignment matrices provide linear mappings between OD flows and traffic data, allowing their updating to be done analytically [32] without the use of a DTA model. Assignment matrix based formulations limit the use of traffic measurements to link counts primarily (limited extensions have been proposed, e.g. [2] considered ways to add Automated Vehicle Identification information), which might fail to address the Dynamic OD Estimation (DODE) problem for congested networks, due to the dynamics of density-flow relationships (i.e. similar count values possible for both congested and uncongested traffic regimes) [22].

Having said that, several researchers have enhanced assignment-based approaches with non-linear features to overcome these limitations. Frederix et al. [22] address the DODE problem for congested networks by suggesting that instead of just using a linear response function, a second term should also be included (the authors used a first order Taylor expansion). Thus, the correlation between OD flows and those link counts, which are indirectly affected e.g. due to congestion (i.e. spill-backs) is also captured. Toledo and Kolehckina [33] also provide valuable insight on the performance of different approaches. By using the enhanced assignment matrix approximation, the authors show that the analytical approaches, based on relative gradient and quasi Newton methods, can outperform meta-heuristic algorithms, using traffic counts as the information.

On the other hand, non-assignment matrix-based approaches can use any source of traffic information, due to a more generic (often non-analytical) formulation [3]. A widely adopted, non-assignment matrix-based approach for calibration of DTA models is Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm [28]. Besides being simple and computationally efficient, SPSA provides the advantage of using any traffic measurement due to a generic problem formulation. Being stochastic in nature, SPSA is sensitive to the definition of its initial parameters (gain sequence), scale of the estimating variables, the objective function and its corresponding gradient. Due to these characteristics, its performance often reduces significantly with the increase in DTA model size and complexity.

The motivation behind this research is to increase the robustness of SPSA for complex, high-dimensional and non-linear problems. DTA model calibration is typically a highly non-linear, undetermined large size problem. Due to its scale and complexity, SPSA is typically applied for offline calibration, and requires a significant number of iterations (more than a

M. Qurashi, T. Ma, E. Chaniotakis and C. Antoniou are with the Department of Civil, Geo and Environmental Engineering, Technical University of Munich, Munich, Bavaria, 80333 Germany e-mail: (see <http://tse.bgu.tum.de>).

Manuscript received December 31, 2018; revised version received March 25, 2019; final version submitted May 4, 2019

*corresponding author: Moeid.Qurashi@tum.de

hundred) to reach an appropriate solution. The role of SPSA's predefined gain sequence and its problem characteristics is very significant in determining its performance regarding error reduction and computational time. Increasing dimensions and complexity of the calibrated DTA model significantly affect SPSA's calibration accuracy and temporal efficiency.

The objectives of this research are two-fold: 1) Develop a modified SPSA algorithm, with the application of Principal Component Analysis (PCA) concepts, thus calibrating the model parameters in a reduced dimensional space; 2) Evaluate the performance of PCA based SPSA (which we call PC-SPSA) on a real medium-sized urban network of Vitoria, Spain, against traditional SPSA; and 3) Examine the robustness of PC-SPSA against the increase in problem dimensionality.

This research contributes to the literature by proposing a statistics-based enhancement for SPSA. The new algorithm (PC-SPSA) shows an enhanced application scalability towards larger and more complex problems. PC-SPSA statistically reduces its problem dimensions and non-linearity through recognizing the structural patterns in its historical estimates and calibrates it in a reduced dimensional space using SPSA. The performance of PC-SPSA in calibrating the considered network and solving its DODE problem illustrates that using PCA along SPSA not only increases its application feasibility for calibrating large size complex networks in an offline context, but it can also be a potential online calibration approach for networks of medium size and complexity. As PCA restricts the search area of SPSA within the structural relationships captured from historical estimates, it not only makes PC-SPSA perform much faster, but also addresses the issue of indeterminacy by keeping the solution pattern closer to the patterns observed in historical estimates. Along reducing the estimation variables more than 50-fold, PC-SPSA calibrates a vector of PC-scores corresponding to orthogonal principal components and reduces non-linearity, which is otherwise present if calibrating actual DTA model parameters, such as OD flows.

II. LITERATURE REVIEW

DTA models capture complex interactions between demand and supply parameters. The extensive literature on DTA model calibration consists of approaches in both offline and online contexts; in this paper we focus on the offline context. Offline calibration approaches estimate model parameters to replicate average traffic conditions over multiple days. Earlier works on offline calibration that focused on the DODE problem include Generalized least square based approach (GLS) by Cascetta et al. [17] and state space formulations by Ashok and Ben-Akiva [9], Okutani and Stephanedes [24]. These approaches are categorized as assignment matrix-based, constrained by using only traffic counts for calibration and formulated in an optimization framework using assignment matrix as the source of mapping between OD demand and traffic counts. Later, a more generic non-assignment matrix-based approach is proposed by Balakrishna et al. [12] using Simultaneous Perturbation Stochastic Approximation (SPSA) [31] to estimate both supply and demand parameters. SPSA has a generic

gradient based problem formulation and it can use any type of traffic measurements for calibration [5]. Balakrishna et al. [11] demonstrates the benefits of not using assignment matrix formulation, while solving a DODE problem. We refer the reader to Antoniou et al. [4, 6] for further detailed literature on offline calibration with a focus on SPSA and its variants.

Many efforts have been made to improve the performance and stability of SPSA for DTA model calibration. Balakrishna et al. [12] applied SPSA for DTA model calibration with modifications, such as scaling of perturbation coefficient c_k and use of multiple gradient replications to improve the gradient approximation, resulting in improved stability of SPSA performance. Cipriani et al. [18] proposed forward (asymmetric) differencing for gradient approximation, which reduces the number of function evaluations, but adds more bias in gradient estimation, compared to using a central differencing technique. Cipriani et al. [19] and Cantelmo et al. [14] also present a sensitivity analysis on SPSA by investigating the effects of multiple perturbation coefficient c_k values, as well as varying the number of gradient replications and applying gradient smoothing. Cantelmo et al. [15] focus on the problem of seed-matrix improvement and present a two-step procedure, based on the concept of dividing the problem into smaller problems: the first step focuses on the optimization of a subset of OD variables (those generating the higher flows or contributing to bottlenecks), while the second step operates on all OD pairs.

An adaptive SPSA algorithm [30] based on Hessian estimates is also proposed, which automatically scales the estimation variables along adaptive scaling and shifting of the gain sequence parameters [13]. Another algorithm, referred to as weighted SPSA (W-SPSA), has been proposed by Lu et al. [23] incorporating the network information through a weight matrix defined based on spatial and temporal correlations between the model parameters and traffic measurements. The use of a correlation weight matrix helps to reduce the noise generated by uncorrelated measurements. Antoniou et al. [7] demonstrated the practical implications of W-SPSA by exploring multiple techniques for evaluating an effective weight matrix along the comparison of W-SPSA with SPSA. Another SPSA-based approach named cluster-wise SPSA (c-SPSA) has been proposed by Tympakianaki et al. [34], segmenting OD flows into clusters to approximate the gradient separately for each cluster and reduce the gradient bias. C-SPSA acts as a hybrid of FDSA and SPSA with increased number of objective function evaluations, i.e. number of clusters \times 2 (traditional SPSA). Clustering is done based on two major techniques: magnitude based k-means clustering [34] and clustering based on spatial interactions of OD pairs computed through free-flow travel times proposed in [35]. The latter technique also helps reducing the non-linearity in the estimation problem by combining non-correlated variables in a cluster.

Efforts have also been made to increase the calibration scalability for DTA models by reducing the problem dimensions using principal component analysis (PCA). Djukic et al. [21] proposed the technique of applying PCA to reduce the dimension of OD matrices. PCA converts the problem parameters into a set of uncorrelated PC-scores through capturing the structural relationships present in their historical estimates.

More recently, a similar technique is used by Prakash et al. [26] and Prakash et al. [27] to enhance traditional calibration approaches of GLS and EKF (and create their counterparts PC–GLS and PC–EKF) improving their application scalability many folds. Prakash et al. [27] demonstrated the application of PC–EKF for calibrating a DTA model of the expressway network of Singapore using realistic traffic data.

Osorio [25] follows a different approach and proposes a metamodel simulation–based optimization approach that also seems to outperform SPSA and offer a scalable solution, applicable to large–scale networks.

III. METHODOLOGY

To propose a DTA model calibration approach, we formulate the calibration problem in an optimization framework, similar to [12]. The generic optimization problem formulation of DTA calibration is thus given as:

$$\underset{\beta, x}{\text{Minimize}} z(y, y', x, x^p, \beta, \beta^p) \quad (1)$$

with

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

where:

- y, y' = Observed and simulated traffic measurements
- x, x^p = OD demand vector (current and prior)
- β, β^p = DTA model parameters (current and prior)
- z = Goodness of fit function
- G = Road network

Within this research, we try to solve the traditional DODE problem by keeping the model parameters β constant and calibrating the OD demand x based on observed traffic counts y . Hence, mathematically our optimization problem can be represented as:

$$\underset{x}{\text{Minimize}} z = \sum_{h=1}^n [z_1(y_h, y'_h) + z_2(x_h^p, x_h)] \quad (2)$$

subject to:

$$y' = f(x_1, \dots, x_h; \beta; G) \quad (3)$$

$f(\cdot)$ represents a DTA model mapping OD flows to traffic data, x_h^p is the prior OD matrix for interval h and x'_h the corresponding estimated OD. It defines the difference between assignment and non–assignment based formulations. The use of assignment matrix A allows mapping to be done analytical, with $f(\cdot)$ as $y = Ax$. However, it also limits the use of any other traffic measurements than counts (although methods have been developed to extend this approach to few more types of measurements, e.g. [2]). Non–assignment–based approaches, on the other hand, provide the flexibility of using any other types of traffic measurements, like travel times, queues or even trajectory data, at the cost of additional runs of a DTA model.

A. SPSA

SPSA comes from the family of stochastic approximation (SA) algorithms. These algorithms are traditionally used for large non–linear problems with expensive objective function

evaluations and noisy measurements. A standard form of SA is given as:

$$x_{k+1} = x_k - a_k g'_k(x_k) \quad (4)$$

where x_{k+1} represents the estimated decision variable through the evaluated gradient g'_k at the k^{th} iteration by perturbing x_k at iteration k . a_k is a predefined positive scalar, specifying the minimization step size. SA algorithms find an optimum estimate x^* starting from an initial estimate x^0 iteratively tracing the path between them. In SPSA the variables are perturbed simultaneously, requiring only two objective function evaluations per iteration for evaluating the gradient (possibly averaged over multiple runs to account for stochasticity), regardless of the number of estimation variables:

$$g' = \frac{f(x_k + c_k \Delta_k) - f(x_k - c_k \Delta_k)}{2c_k} [\Delta_1 \ \Delta_2 \ \dots \ \Delta_h]^T \quad (5)$$

where h is the number of OD pairs and Δ is a h -dimensional vector generated randomly from a ± 1 Bernoulli distribution having zero mean. c_k is the perturbation coefficient predefined along a_k , both known as gain sequence.

We defined the SPSA algorithm along the modifications stated in the literature for solving DODE problems. Gain sequence parameters are defined based on the guidance given in Spall [29]. Gain sequence is based on cautiously defined parameters of c, γ, a, A and α . With k as the iteration number, these parameters define the pattern of reductions in c_k and a_k over incrementing iterations, as:

$$c_k = c/k^\gamma \quad a_k = a/(k + A)^\alpha$$

subject to:

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

Further modifications that are included in our implementation of SPSA are: the use of gradient replications for lesser biased gradient, non–negative and scaled OD flows perturbation/minimization through relative segmented perturbation/minimization based on their magnitudes as proposed by Balakrishna [10]. Equation 6 shows the expression used for evaluating relative segmented perturbation c_k and minimization a_k :

$$a_k = \frac{a_k \times i \times n}{\mu} \quad c_k = \frac{c_k \times i \times n}{\mu} \quad (6)$$

where μ is the mean of OD matrix (decision variable). The highest OD matrix magnitude is divided into segments with n as the segment size. The gain sequences c_k and a_k are scaled for each OD pair using i , which is the segment number in which this OD pair magnitude lies.

B. PC–SPSA

We propose PC–SPSA as an enhanced version of SPSA, improving its application scalability for calibrating larger and

more complex DTA models. The application of principal component analysis (PCA) is done on historical estimates to extract the structural relationships between OD pairs as PC-directions 'V', as proposed by Djukic et al. [21]. These directions are used to reduce the size of the OD estimation vector to a few uncorrelated PC-scores. These PC-scores are later calibrated using a modified SPSA algorithm.

1) *Principal Component Analysis*: We apply the methodology given in Prakash et al. [26] to reduce the dimensions of our OD vector. Singular value decomposition is applied on a data matrix X to evaluate its principal components (PCs). This data matrix X is the combination of historical OD matrices with $n_k \times n_x$ dimensions: n_k is the number of data points and n_x represents the dimension of an OD matrix in the form of a vector for each corresponding data point. Singular value decomposition results in two unitary matrices U , V and a diagonal matrix Σ :

$$X = U\Sigma V^T \quad (7)$$

Columns of the unitary matrix V present orthogonal PCs, with their corresponding PC-scores stored in the diagonal matrix Σ . Due to a property of PCA, the amount of variance captured in the PCs reduces in a descending order and thus the first few PCs can capture most of the variance present in the data matrix. Hence, V (also called PC-directions) may be reduced to \hat{V} , where only the first n_d elements are retained:

$$\hat{V} = [v_1 \ v_2 \ v_3 \ \dots \ v_{n_d}] \quad (8)$$

The reduction in the dimension of the matrix \hat{V} is based on the ratio of the cumulative summation over the summation of its corresponding PC scores vector. The new matrix \hat{V} is then used to reduce our starting OD matrix into PC scores z of dimension $n_d \times 1$, as:

$$z = \hat{V}^T x \quad (9)$$

Furthermore, the OD vector can be approximated as:

$$x \approx \hat{V} z \quad (10)$$

The data matrix used in evaluating the PC-directions V should contain previous estimates with similar characteristics temporally. This implies that different sets should be constructed e.g. for different times of day (e.g. morning and evening peak periods), as well different days of the week (e.g. work days vs. holidays). Among others, Antoniou [1, Section 5.1] and Balakrishna [10, Section 3.3] provide detailed insights for techniques to develop these historical estimates, stated as "warm-up" procedures.

2) *Application of PCA on SPSA*: To incorporate PCA in SPSA, we propose a modified version of SPSA. In PC-SPSA, we calibrate PC-scores instead of actual parameters along the modifications required, due to the difference in characteristics between OD flows and their evaluated PC-scores. To transform SPSA into PC-SPSA we propose the following modifications:

- Addition of a data matrix consisting of previous estimates to evaluate PC-directions V through Equation 7.

- Transformation of OD flow vectors to PC-scores for the steps of perturbation and minimization using Equation 9.
- Instead of having a relative segmented change, perturbation and minimization are applied as a percentage increase or decrease. Gain sequences c_k and a_k are also redefined, as they specify a percentage change in PC scores z . The expressions for perturbation and minimization thus become:

$$\text{Perturbation:} \quad z^\pm = z_k \pm z_k \times c_k \Delta \quad (11)$$

$$\text{Minimization:} \quad z_{k+1} = z_k - z_k \times a_k g' \quad (12)$$

where z_{k+1} are the new estimated PC-scores at k th iteration by perturbing and evaluating z_k .

- Conversion of PC scores to OD matrices after perturbation and minimization using Equation 10, to evaluate the objective function through a DTA model simulation.

Modifications for the steps of perturbation and minimization are proposed because of the variations present in the magnitudes of the PC-scores. Each PC-score represents the amount of variance captured by its corresponding principal component, hence their magnitudes can vary significantly. In case of calibrating OD flow vectors directly, the variation is far less significant and may be addressed considering the relative segmented change using Equation 6; however, for PC-scores the variation can be much larger and requires further modifications.

Figure 1 describes the main steps of PC-SPSA. The basic structure of this algorithm is similar to SPSA, with the additional steps of evaluating PC-directions (using equation 7) in the beginning and steps to transform the OD flow vectors and PC-scores into each other (using Equations 9 and 10), before perturbation, minimization and objective function evaluations.

C. Objective function

PC-SPSA is defined as a non-assignment matrix-based approach, requiring a DTA model simulation to map OD flows into traffic measurements. Figure 2 describes the basic structure behind the objective function evaluation we used in both SPSA and PC-SPSA. It requires an OD matrix as a basic input to simulate and outputs different traffic measurements like counts and travel times. These simulated outputs are then compared with the observed traffic measurements using a goodness of fit (Gof) measure, hence resulting in a Gof value as the output of our objective function.

DODE is a non-linear and underdetermined problem and its simulated outputs cannot be directly related back to OD flows (since OD flows are in general still not observable). Hence, we use the root mean square normalized (RMSN) error as our goodness of fit (Gof) criteria for the measurable network performance measures, as shown in Equation 13 [as used e.g. in 3, 9]:

$$RMSN = \frac{\sqrt{n \sum_{i=1}^n (\hat{y}_i - y_i)^2}}{\sum_{i=1}^n y_i} \quad (13)$$

where y are the observed traffic flows and \hat{y} are their simulated counterparts, and n represents the total number of values.

The use of RMSN as Gof provides the practical advantage of having a single output from our objective function evaluation.

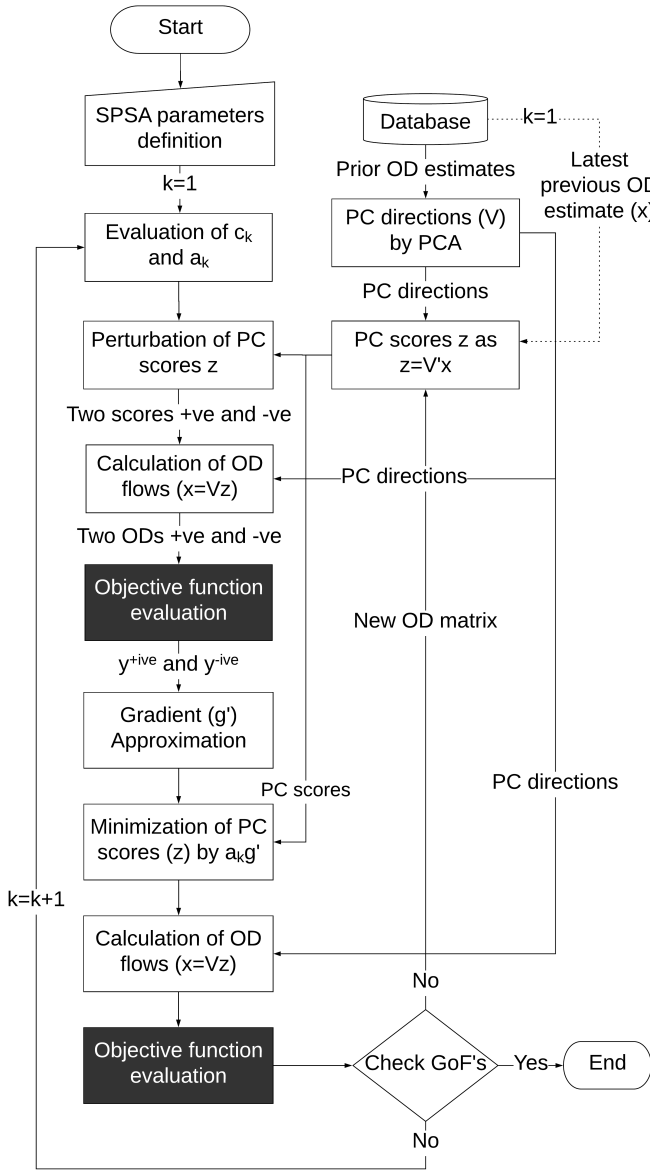


Fig. 1: Flow chart of PC-SPSA

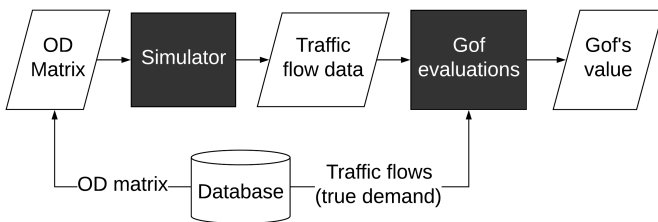


Fig. 2: Structure within an objective function evaluation

As the inputs and outputs of our objective function are in general neither equal nor relatable, a single RMSN value output helps in formulating a generic SPSA formulation. RMSN has been chosen previously in many similar researches including [9], [27].

IV. CASE STUDY ON VITORIA NETWORK

We use a simulator based case study to validate our approach. A network representing the city of Vitoria, Spain, is used (Figure 3). Vitoria represents a medium-sized urban city having a dense city center with radial highway network and a motorway bypass. The network provides a reasonable level of congestion and route choice, and consists of 5799 links (about 600km in length) with 2884 nodes. There are 395 detectors installed (with their location shown in (Figure 3)). For demand modeling, the network is divided into 57 zones, leading to an OD matrix with the dimensions of 3249 OD pairs.



Fig. 3: Network of Vitoria-Spain, with detector locations

A. Experimental Setup

We adopt a commercial software Aimsun [16] as our simulator. Aimsun provides the ability to simulate in multiple resolutions, including a mesoscopic resolution providing the desired numerical stability against smaller variations in simulation parameters and detailed driving behavior patterns with faster simulations than microscopic simulation [8]. The simulations are set up with stochastic route choice based scenarios using 10 replications to generate statistically significant results. Each replication provides varying stochasticity based on its starting random seed and every 10 generated random seeds are normally distributed in Aimsun.

Route choice parameters are set constant with fixed traffic assignment generated through dynamic user equilibrium (DUE) using the method of successive averages. Fixed traffic assignment helps to remove the stochasticity of route choice and DUE ensures better congestion mitigation over constantly changing demand during the iterative process. Appropriate traffic path assignment is crucial for the experimental setup, as perturbing the OD flows can create congested situations on the network, causing the simulation outputs to become less reliable for our objective function evaluations.

We set up a true network state considering its demand as true demand and its simulation outputs as observed traffic measurements. An evening peak hour demand is used with four 15-min intervals and a 30 min warm-up period. To solve the DODE problem, we calibrate each of the four demand intervals separately using traffic counts as the measurement to formulate our objective function described in Section III-C. In this specific case study, we do not consider the second term of the objective function, i.e. the distance from the starting OD. Due to its generic objective function formulation, this experimental setup could later be easily extended with travel times or other trajectory data.

B. Historical Demand

PC-SPSA requires a historical demand data matrix to evaluate PC-directions 'V' for the reduction of the dimension of the OD flow vector. We formulate this demand using the expression given in Equation 14. Multiple data points are generated through perturbing the (calibrated) true demand. r and q are the coefficients of reduction and randomization, while δ is a vector of Gaussian random numbers with zero mean and standard deviation equal to $\sigma = \frac{1}{3}$ resulting to 99.7% of the values between -1 and 1 .

$$x_{hist} = (r + q\delta) \times x_{true} \quad (14)$$

Randomness added by q and δ ensures the creation of different historical demand patterns from the true demand to replicate actual conditions, where actual demand patterns are unknown.

C. Demand scenarios

We use two different demand scenarios to evaluate the performance of SPSA and PC-SPSA. To create a demand scenario we follow the guidelines of Antoniou et al. [8]. The true demand is perturbed using the expression given in Equation 14 with two different r and q coefficients for the creation of historical estimates. These demand scenarios are:

Scenario 1: $x_{hist} = (0.75 + 0.15\delta) \times x_{true}$

Scenario 2: $x_{hist} = (0.65 + 0.20\delta) \times x_{true}$

Both scenarios have the same target demand with different reduced historical demand data matrices. Reduced demand scenarios are created to avoid artificial over-saturation of the DTA model and create an increased target demand estimation problem. The first scenario represents a mild demand change, while the other scenario includes a more severe change in reduction and randomization. The latest previous historical estimate from the historical demand data matrix is taken as the seed OD matrix for calibration.

D. Estimation of PC-directions

To evaluate the PC-directions for this experimental setup, 25 historical estimates are generated: each of one hour demand, having four OD vector corresponding to its four 15 minutes intervals. The historical data matrix X contains 25×4 i.e. 100 OD vectors, each of 3249 OD pairs (total dimensions 100×3249). The application of single value decomposition on

X results in PC directions matrix V and a diagonal matrix Σ containing its corresponding PC scores. It is clarified that the selection of 25 instances (per OD interval) is indeed arbitrary, and future research is needed to determine guidelines for choosing the appropriate number for each application. Figure

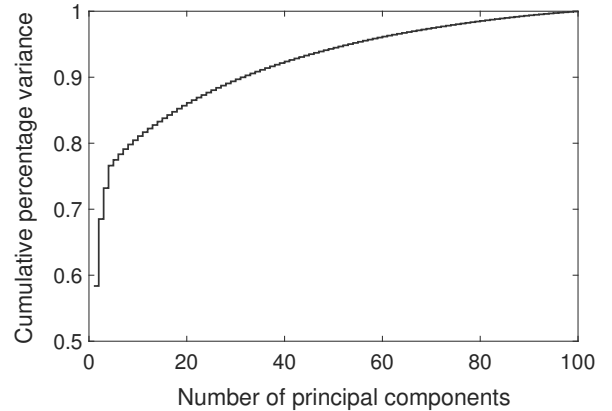


Fig. 4: Cumulative percentage variance captured by PCs

4 shows the graph of the cumulative percentage variance captured over increasing number of principal components, computed using the Σ 's diagonal values. We use this cumulative percentage variance to reduce the number of PCs till the retained PCs explain at least 95% of the variance present in the historical data matrix X . From Figure 4 we perceive that more than 95% of the variance is captured by the first 55 PCs; hence a reduction factor of $3249/55$ i.e. 59 is achieved with the application of PCA.

V. RESULTS

Based on the generation of historical estimates, two demand scenarios are developed. We demonstrate the comparison of the performance obtained by both SPSA and PC-SPSA in calibrating each scenario. First, we compare the rate of convergence using RMSN values. Then, we use so-called 45-degree plots to depict the comparison of the initial and target OD estimates with the calibrated OD estimate, and also similarly the comparison between observed and calibrated traffic counts, from both SPSA and PC-SPSA.

A. Convergence

Figure 5 illustrates the RMSN error convergence, while calibrating scenario 1 and 2, for each of the four time intervals. Both algorithms have ran to a maximum number of 80 iterations. It is clear that PC-SPSA not only converges to much better solutions, but it achieves the majority of its gains in a very small number of iterations (less than ten).

Scenario 1 replicates a mild change of target demand with a base decrease of 25% and randomized perturbation of $\pm 15\%$. Results from the first scenario show that PC-SPSA has a far better rate of convergence than SPSA, reducing the overall RMSN error values to almost 5% within the first 10 iterations.

Scenario 2 replicates a more severe change in target demand with a base decrease of 35% and a randomized increase/decrease of $\pm 20\%$. Results from this scenario show

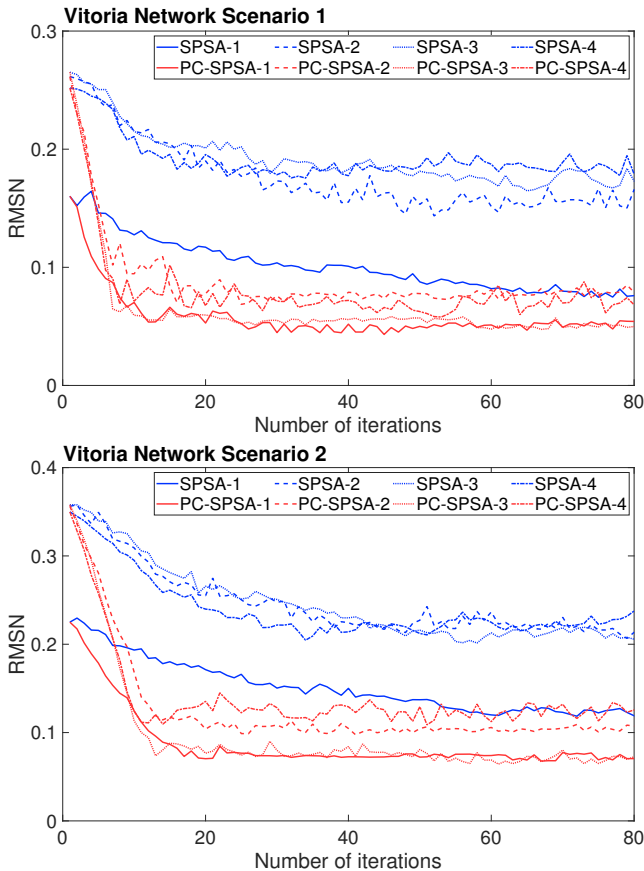


Fig. 5: RMSN comparison for demand scenario 1 (top) and 2 (bottom)

similar patterns as in scenario 1 as PC-SPSA converges to an RMSN error less than 12% for all intervals within the first 13 iterations. Since the overall base increase in target demand is more in this scenario than the previous, the number of iterations required for overall error reduction has also increased. SPSA also shows the similar rate of convergence as in the previous scenario, always significantly outperformed by PC-SPSA.

B. Traffic counts

Figure 6 illustrates 45-degree plots to compare the observed and calibrated traffic counts of all four demand intervals from both scenarios. The left subfigures depict the calibrated counts from SPSA, which are more scattered in comparison to the calibrated counts from PC-SPSA, shown in the right subfigures. The scatteredness in all the figures replicates and validates the RMSN value achieved by both algorithms during calibration.

C. Validation using OD travel times

Figure 7 illustrates 45-degree plots to compare the observed and calibrated OD travel times of all four demand intervals from both scenarios, to further validate the calibration results. As travel times were not used for the calibration, their parallel improvement provides additional evidence that the superior PC-SPSA performance is not related with issues, such as

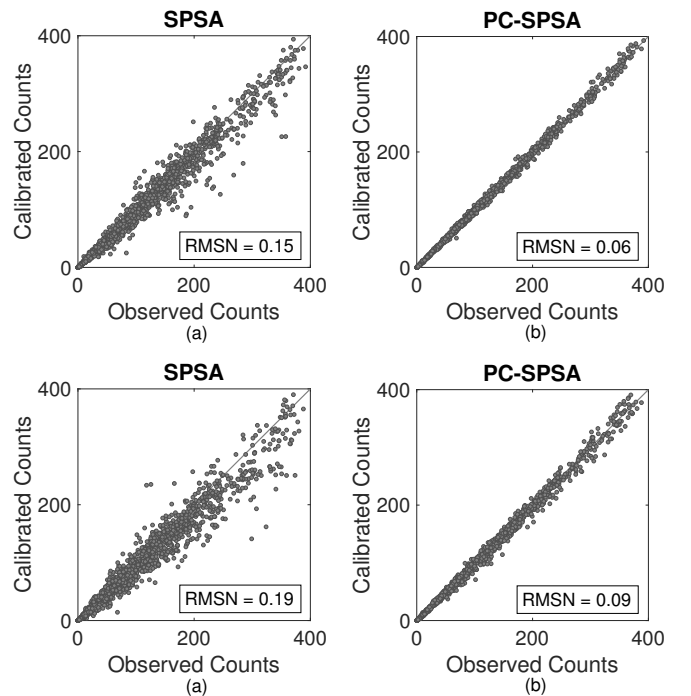


Fig. 6: Comparison of observed and calibrated traffic counts (top: scenario 1, bottom: scenario 2)

overfitting. The left subfigures again depict the calibrated OD travel times from SPSA, which are more scattered in comparison to the calibrated OD travel times from PC-SPSA, shown in the right subfigures. The RMSN between the observed and calibrated OD travel times validate the calibration results shown in section V-B.

D. Quality of calibrated OD matrix

Figure 8 illustrates 45-degree plots to compare the initial and target OD estimates with the calibrated OD estimate from only the first demand interval of both scenarios. The left subfigures show the comparison of the SPSA-calibrated OD matrix with the target OD matrix. Both plots show scattered data points proposing that the patterns of calibrated OD matrix are very different from the true OD matrix. SPSA being a random search algorithm changes the variance patterns of calibrated OD matrices drastically from their initial values, because it randomly perturbs its values directly, without any search directions. Even with a scaled perturbation, it can really change the patterns from their initial values over a larger number of iterations.

The right subfigures show the comparison of PC-SPSA's calibrated OD matrix with the target OD matrix. Both plots show minimal scatteredness in comparison to other SPSA plots with all the data points closer to the 45-degree plot line. This illustrates that PC-SPSA is able to find better and more accurate solutions, closer to the true OD estimates. Due to the application of PCA, SPSA's search area is restricted only within the variance captured over the historical estimates.

Of course, as Djukic [20] underlines, this is only one way of controlling the quality of the estimated OD matrices. The

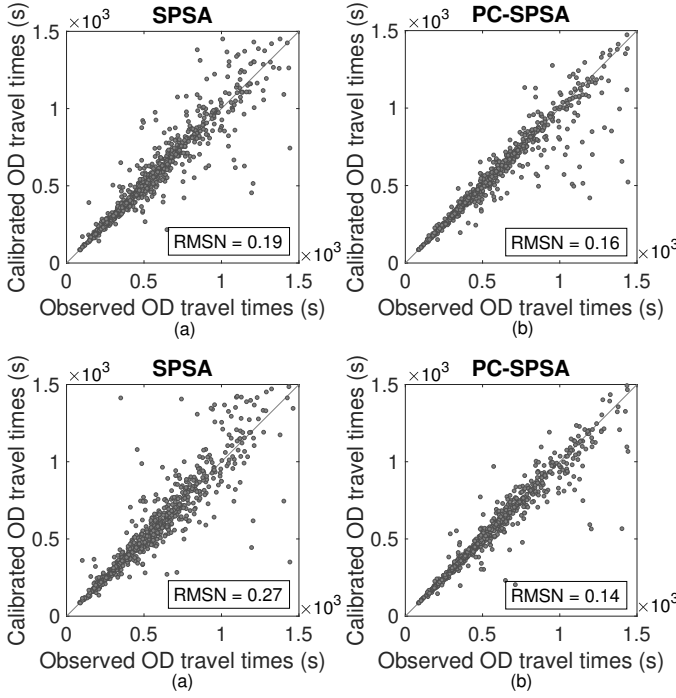


Fig. 7: Comparison of observed and calibrated OD travel times (top: scenario 1, bottom: scenario 2)

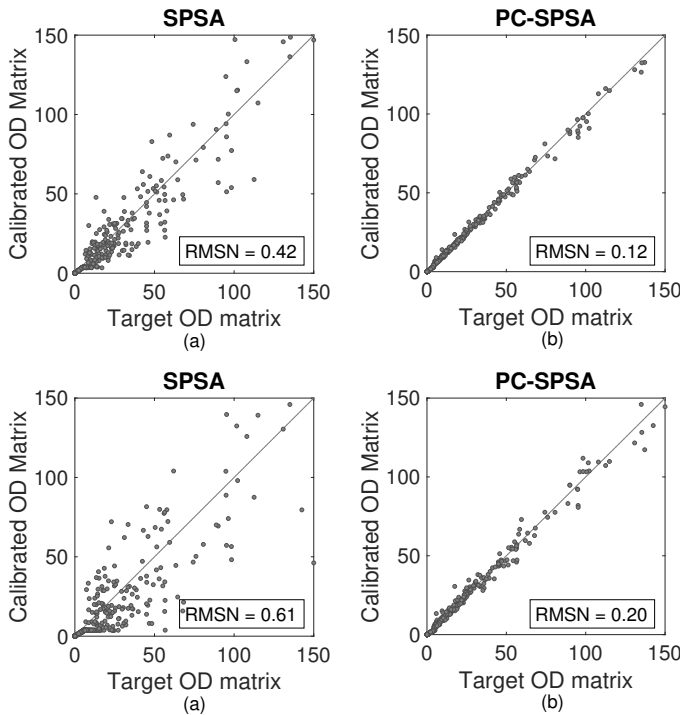


Fig. 8: Comparison of target and calibrated OD matrices (top: scenario 1, bottom: scenario 2)

author also suggested the use of a similarity metric, SSIM, to check the structural similarity of the estimated OD matrices with their targets.

The results from this case study confirm the enhancements in PC-SPSA due to the application of PCA. The improvements in the new algorithm are evident as firstly we perturb 54 times lesser set of variables from 3249 to a mere 60. Secondly, it reduced the noisy effect of SPSA random search. Being a random search algorithm it finds the solution by perturbing the parameters directly and continuously changes its solution patterns without any search directions. But by reducing the model parameters using PCA, SPSA's search gets the required directions to move faster closer towards the true solution.

VI. DISCUSSION

Based on the simulator case study, PC-SPSA proves to be an enhanced version of SPSA having faster convergence and more appropriate solutions. In this section, we use a synthetic experimental design by replacing the simulator with a non-linear synthetic function to further discuss and explore different aspects regarding PC-SPSA's performance. In particular, we perform a robustness analysis upon different dimensions, discussing its major strengths and weaknesses, its comparison with previously developed approaches and how it can be proposed as an online calibration algorithm for DTA model calibration.

A. Robustness

Although PC-SPSA provides promising results in calibrating a moderately sized urban network, we explore the robustness of PC-SPSA and compare it with SPSA in calibrating multiple problems with different dimensions, based on a non-linear synthetic function. This function is used to replace the simulator for mapping the OD flows into counts, so that we can perform a sensitivity analysis on both algorithms with problems of varying dimensions. The non-linear synthetic function is formulated as:

$$y = Wx + W_s x^2 \quad (15)$$

where y represents the traffic counts and x the OD flow vector. W and W_s are two randomly generated weight matrices of dimensions $[x \times y]$ acting as the correlation matrices to map the OD flows x into counts y . A secondary weight matrix W_s is used to add non-linearity into the function. We create eight sets of problems with eight different dimensions ' d ', indicated by the number of zones, ranging from 20 to 90 zones. The size of OD flows vector x will be d^2 and the number of counts y is set to one-fifth of OD flows x . We randomly generate x using a uniform distribution and the correlation matrices W and W_s using Bernoulli distributions having 80% outputs as zero and 20% as one. The target demand is created with a base decrease of 30% and perturbation of $\pm 15\%$.

Figure 9 illustrates the RMSN values for calibrating the problems based on the non-linear synthetic function using both algorithms. Each problem, indicated by its dimensions ' d ', is shown with a unique line pattern. It is evident that SPSA's performance deteriorates significantly as the problem

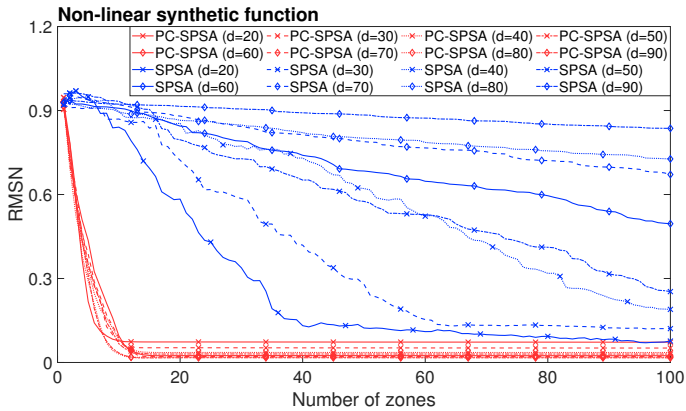


Fig. 9: Comparison of RMSN values for SPSA and PC-SPSA over multiple dimensions

dimensions increase, while the performance of PC-SPSA remains similar regardless of the increase in problem dimensions. It is noted that the performance for all PC-SPSA instances outperforms the performance of even the best SPSA instance.

B. Characteristics of PC-SPSA

Through the exploration of PC-SPSA performance on a real-life network and a multidimensional robustness analysis, we come up with some basic characteristics regarding PC-SPSA and its application feasibility:

- PC-SPSA iteratively searches for different patterns within the variance of historical estimates. We perturb the magnitude of its PC-scores, each of which corresponds to an orthogonal PC, capturing a variance orthogonally to all other PCs. Hence, we perturb and minimize the contributions of each orthogonal variance towards the overall solution instead of capturing the behavior of each individual OD pair towards the true solution (as in SPSA).
- PC-SPSA's performance is not directly affected by the increase in problem size; instead, it is affected by the increase in its complexity or non-linearity. The reason is that we reduce the PC direction matrix V based on the cumulative variance captured over principal components (PCs) and with the increase in problem complexity the variance will get distributed over a higher number of PCs, due to richer structural relationships between model parameters, hence reducing the scale of dimension reduction and increasing the number of iterations required for calibration.
- The performance of PC-SPSA depends to a large degree on the quality of the historical estimates as, if the patterns of the target demand are not similar to its historical estimates, the overall reduction achieved in RMSN error will be lower. However, its rate of convergence should not be affected, as it must depend on the scale of dimension reduction. Good historical estimates should ideally include most of the variation possible in the target estimate.

- The aspect of having good quality historical estimates is also affected by the scale of non-linearity present in the problem, as highly non-linear problems will include richer structural relationships, which increase the uncertainty of having similar patterns between historical and target estimate, affecting the lowest RMSN achieved.

C. Comparison with previously developed approaches

As discussed in the literature review section, previous extensions to the SPSA framework tried to address the issues of stochasticity and non-linearity in DTA model calibration. W-SPSA [7] proposed the use of a correlated weight matrix to perturb and minimize the variables. This weight matrix captures the correlations between model parameters, reducing the noise of uncorrelated parameters. C-SPSA [35], on the other hand, tries to reduce the effect of non-linearity by creating clusters, in which the least correlated parameters are combined. PC-SPSA using PCA produces similar effects more sophisticatedly by assessing the structural correlation statistically between model parameters in the form of PC directions V and later using them to evaluate PC-scores, which are orthogonal to each other, almost removing the effect of non-linearity. In addition, the number of variables to be estimated are far less in PC-SPSA than in its predecessors.

D. Online calibration

Online calibration approaches tend to estimate the model parameters for real-time estimation and prediction in DTA systems [3]. For these approaches, computational time ends up being the major constraint. Hence, SPSA is not intended to be an online calibration approach, due to its lengthy computation time (large number of iterations). However, PC-SPSA could conceivably be used as an online calibration algorithm, as it obtains the majority of the loss function reduction after only a handful of iterations. Furthermore, multiple simulations (e.g. to alleviate the effects of stochasticity) can be run in parallel within each iteration. Since an online calibration approach tends to fine tune the offline calibrated parameters, PC-SPSA can be feasible due to smaller errors required to minimize.

VII. CONCLUSION

This paper proposes a modified SPSA algorithm referred to as PC-SPSA. The motivation behind this research is to enhance the application scalability of SPSA on DTA model calibration by reducing its search noise. As a random search algorithm, SPSA searches the solution by perturbing the variables randomly. Its performance is sensitive against a set of predefined gain sequence parameters and problem characteristics and deteriorates significantly with the increase in problem dimensions and non-linearity. PC-SPSA uses a dimension reduction technique (PCA) to evaluate structural relationships between the historical estimates as PC-directions and later use these PC-directions to reduce the set of estimation variables into PC-scores. These PC-scores are then calibrated using SPSA. PCA restricts the search area of SPSA within the variance captured in its historical estimates.

Modifications in SPSA are developed to incorporate PCA and the required methodology behind PC-SPSA is presented. Due to the greater variations present between PC-scores, scaling in the steps of perturbation and minimization is done in a modified manner along the usage of PC-scores instead of the actual model parameters.

Performance evaluations are done for PC-SPSA on an urban network of Vitoria, Spain, in comparison with SPSA showing promising results. PC-SPSA outperforms SPSA in both the rate of convergence and quality of the solution. PC-SPSA is not only computationally efficient but it also finds the solution with patterns very close to the true values. Since it only requires a handful of iterations, PC-SPSA is not very sensitive against the predefined gain sequence, skipping the efforts required to fine tune these parameters as in SPSA.

Future research includes extending the scope of the calibration to demand and supply parameters, as well as validating the performance of PC-SPSA in other networks and conditions.

ACKNOWLEDGMENT

The authors would like to thank Aimsun for providing us with the calibrated Vitoria network, as well as valuable support in using the Aimsun software. This research has been partly supported by the German Research Foundation - Deutsche Forschungsgemeinschaft (DFG) [Project number 392047120, Research grants in collaboration with China] and the International Graduate School of Science and Engineering - IGSSE (MO3 Project).

REFERENCES

- [1] C. Antoniou. *On-line calibration for dynamic traffic assignment*. PhD thesis, Massachusetts Institute of Technology, 2004.
- [2] C. Antoniou, M. Ben-Akiva, and H. N. Koutsopoulos. Incorporating automated vehicle identification data into origin-destination estimation. *Transportation Research Record*, 1882(1):37–44, 2004.
- [3] C. Antoniou, M. E. Ben-Akiva, and H. N. Koutsopoulos. Nonlinear Kalman filtering algorithms for on-line calibration of dynamic traffic assignment models. *IEEE Transactions on Intelligent Transportation Systems*, 8(4): 661–670, 2007.
- [4] C. Antoniou, R. Balakrishna, H. N. Koutsopoulos, and M. Ben-Akiva. Off-line and on-line calibration of dynamic traffic assignment systems. In *IFAC Proceedings Volumes (IFAC-PapersOnline)*, number 15, pages 104–111. IFAC, 2009.
- [5] C. Antoniou, R. Balakrishna, and H. N. Koutsopoulos. A synthesis of emerging data collection technologies and their impact on traffic management applications. *European Transport Research Review*, 3(3):139–148, Nov 2011. ISSN 1866-8887. doi: 10.1007/s12544-011-0058-1. URL <https://doi.org/10.1007/s12544-011-0058-1>.
- [6] C. Antoniou, R. Balakrishna, H. N. Koutsopoulos, and M. Ben-Akiva. Calibration methods for simulation-based dynamic traffic assignment systems. *International Journal of Modelling and Simulation*, 31(3):227–233, 2011. doi: 10.2316/Journal.205.2011.3.205-5510.
- [7] C. Antoniou, C. L. Azevedo, L. Lu, F. Pereira, and M. Ben-Akiva. W-spsa in practice: Approximation of weight matrices and calibration of traffic simulation models. *Transportation Research Part C: Emerging Technologies*, 59:129–146, 2015.
- [8] C. Antoniou, J. Barceló, M. Breen, M. Bullejos, J. Casas, E. Cipriani, B. Ciuffo, T. Djukic, S. Hoogendoorn, V. Marzano, L. Montero, M. Nigro, J. Perarnau, V. Punzo, T. Toledo, and H. van Lint. Towards a generic benchmarking platform for origin-destination flows estimation/updates algorithms: Design, demonstration and validation. *Transportation Research Part C: Emerging Technologies*, 66:79–98, 2016.
- [9] K. Ashok and M. E. Ben-Akiva. Alternative Approaches for Real-Time Estimation and Prediction of Time-Dependent Origin Destination Flows. *Transportation Science*, 34(1):21–36, 2000.
- [10] R. Balakrishna. *Off-line calibration of dynamic traffic assignment models*. PhD thesis, MIT, 2006.
- [11] R. Balakrishna, M. Ben-Akiva, and H. N. Koutsopoulos. Time-Dependent Origin-Destination Estimation Without Assignment Matrices. In *Second International Symposium of Transport Simulation (ISTS06)*, 2006.
- [12] R. Balakrishna, M. Ben-Akiva, and H. Koutsopoulos. Offline calibration of dynamic traffic assignment: Simultaneous demand-and-supply estimation. *Transportation Research Record*, 2003:50–58, 2007.
- [13] M. Broadie, D. M. Cicek, and A. Zeevi. Multidimensional stochastic approximation: Adaptive algorithms and applications. *ACM Trans. Model. Comput. Simul.*, 24(1): 6:1–6:28, 2014. ISSN 1049-3301.
- [14] G. Cantelmo, E. Cipriani, A. Gemma, and M. Nigro. An adaptive bi-level gradient procedure for the estimation of dynamic traffic demand. *IEEE Transactions on Intelligent Transportation Systems*, 15(3): 1348–1361, June 2014. ISSN 1524-9050. doi: 10.1109/TITS.2014.2299734.
- [15] G. Cantelmo, F. Viti, C. M. Tampère, E. Cipriani, and M. Nigro. Two-step approach for correction of seed matrix in dynamic demand estimation. *Transportation Research Record*, 2466(1):125–133, 2014.
- [16] J. Casas, J. L. Ferrer, D. Garcia, J. Perarnau, and A. Torday. *Traffic Simulation with Aimsun*, pages 173–232. Springer New York, New York, NY, 2010.
- [17] E. Cascetta, D. Inaudi, and G. Marquis. Dynamic Estimators of Origin Destination Matrices Using Traffic Counts. *Transportation Science*, 27(4):363–373, 1993.
- [18] E. Cipriani, M. Florian, M. Mahut, and M. Nigro. A gradient approximation approach for adjusting temporal origin-destination matrices. *Transportation Research Part C: Emerging Technologies*, 2011.
- [19] E. Cipriani, A. Gemma, and M. Nigro. A bi-level gradient approximation method for dynamic traffic demand estimation: Sensitivity analysis and adaptive approach. In *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2013.

- [20] T. Djukic. Reliability assessment of dynamic od estimation methods based on structural similarity index. 01 2013. doi: 10.13140/RG.2.1.4174.1929.
- [21] T. Djukic, G. Flotterod, J. Lint, and S. Hoogendoorn. Efficient real time od matrix estimation based on principal component analysis. *Conference Record - IEEE Conference on Intelligent Transportation Systems*, pages 115–121, 2012.
- [22] R. Frederix, F. Viti, and C. M. Tampère. Dynamic origindestination estimation in congested networks: theoretical findings and implications in practice. *Transportmetrica A: Transport Science*, 9(6):494–513, 2013. doi: 10.1080/18128602.2011.619587.
- [23] L. Lu, Y. Xu, C. Antoniou, and M. Ben-Akiva. An enhanced SPSA algorithm for the calibration of Dynamic Traffic Assignment models. *Transportation Research Part C: Emerging Technologies*, 51:149–166, 2015.
- [24] I. Okutani and Y. J. Stephanedes. Dynamic prediction of traffic volume through Kalman filtering theory. *Transportation Research Part B*, 18(1):1–11, 1984.
- [25] C. Osorio. Dynamic origin-destination matrix calibration for large-scale network simulators. *Transportation Research Part C: Emerging Technologies*, 98:186–206, 2019.
- [26] A. Prakash, R. Seshadri, C. Antoniou, F. Pereira, and M. Ben-Akiva. Reducing the Dimension of Online Calibration in Dynamic Traffic Assignment Systems. *Transportation Research Record: Journal of the Transportation Research Board*, pages 96–107, 2017.
- [27] A. Prakash, R. Seshadri, C. Antoniou, F. Pereira, and M. Ben-Akiva. Improving scalability of generic online calibration for real-time dynamic traffic assignment systems. *Transportation Research Record Journal of the Transportation Research Board*, 2017. doi: 10.1177/0361198118791360.
- [28] J. C. Spall. An Overview of the Simultaneous Perturbation Method for Efficient Optimization. *Johns Hopkins Apl Technical Digest Volume: 19, Issue: 4*, 19(4):482–492, 1998.
- [29] J. C. Spall. Implementation of the simultaneous perturbation algorithm for stochastic optimization. *IEEE Transactions on Aerospace and Electronic Systems*, 34(3):817–823, 1998.
- [30] J. C. Spall. Adaptive stochastic approximation by the simultaneous perturbation method. *IEEE Transactions on Automatic Control*, 2000.
- [31] J. C. Spall. Simultaneous Perturbation Stochastic Optimization. *Introduction to Stochastic Search and Optimization: Estimation, Simulation and Control*, pages 176–207, 2003.
- [32] H. Spiess. A gradient approach for the O-D matrix adjustment problem. *Centre for research on transportation, University of Montreal, Canada*, 1990.
- [33] T. Toledo and T. Kolehkhina. Estimation of dynamic origindestination matrices using linear assignment matrix approximations. *Intelligent Transportation Systems, IEEE Transactions on*, 14:618–626, 06 2013. doi: 10.1109/TITS.2012.2226211.
- [34] A. Tympakianaki, H. N. Koutsopoulos, and E. Jenelius. c-SPSA: Cluster-wise simultaneous perturbation stochastic approximation algorithm and its application to dynamic origin-destination matrix estimation. *Transportation Research Part C: Emerging Technologies*, 55:231–245, 2015.
- [35] A. Tympakianaki, H. N. Koutsopoulos, and E. Jenelius. Robust SPSA algorithms for dynamic OD matrix estimation. *Procedia Computer Science*, 130:57–64, 2018.



Moeid Qurashi received the B.E. degree in civil engineering from National University of Science and Technology, Islamabad, Pakistan in 2014, and the M.S. degree in transportation systems from Technical University of Munich, Munich, Germany in 2018. He is currently working as a Ph.D. candidate for transportation engineering at the Technical University of Munich, Munich, Germany.

His research interests include transport modeling and simulation with specifics in DTA model calibration and modeling dynamic mobility services.



Tao Ma is a Postdoctoral Instructor and Project Team Leader at the Chair of Transportation Systems Engineering of the Technical University of Munich. He has received his PhD at the University of Toronto, Canada. His research interests include Intelligent transportation systems, modeling and simulation of transportation systems; Big data analytics, time series modeling, advanced forecast methods; Statistical and machine learning; heuristic optimization; Traffic operations and management, transportation planning, public transit.



Emmanouil Chaniotakis is a Research Associate at the Technical University of Munich. Until September 2016, he was a Research Associate at the Hellenic Institute of Transport of the Centre for Research and Technology Hellas. He holds a Diploma in Rural and Surveying Engineering (2011) from the Aristotle University of Thessaloniki and an MSc in Transportation Infrastructure and Logistics (2014) from the Delft University of Technology (TU Delft). His research focuses on modelling and simulation of transportation systems, demand modeling and Big

Data Analysis in Transportation.



Constantinos Antoniou is Full Professor in the Chair of Transportation Systems Engineering at the Technical University of Munich (TUM), Germany. He holds a Diploma in Civil Engineering from NTUA (1995), a MS in Transportation (1997) and a PhD in Transportation Systems (2004), both from MIT. His research focuses on data analytics, modelling and simulation of transportation systems, Intelligent Transport Systems (ITS), calibration and optimization applications, road safety and sustainable transport systems.