Available online at www.sciencedirect.com

SciVerse ScienceDirect

Transportation Research Procedia 00 (2016) 000–000

Transportation Research Procedia

www.elsevier.com/locate/procedia

International Conference on Intelligent Transport Systems in Theory and Practice, mobil.TUM 2017, 4-5 July 2017, Munich, Germany

Marginal effects evaluation with respect to changes in OD demand for dynamic OD demand estimation

Tamara Djukic [1,*], Martijn Breen [1], David Masip [1], Josep Perarnau [1], Joseph Budin [b], Jordi Casas [1]

[a]*TSS-Transport Simulation Systems, Ronda Universistat 22B, Barcelona 08007, Spain*
[b]*École Polytechnique, Département de Mathématiques Appliquées, Route de Saclay, Palaiseau 91128, France*

Abstract

Transport authorities and practitioners have long been concerned about the unavailability of reliable dynamic OD demand estimates which limits the potential for dynamic traffic assignment (DTA) deployments to analyze and alleviate traffic congestion as part of the Intelligent Transportation Systems (ITS). Modelling of non-linear relationship between traffic observations and OD flows, and its dependency on variations in OD flows has been identified by many researchers as a key challenge in the estimation and prediction of a high-quality OD matrices. Finding derivatives of link-flow proportions and traffic observation with respect to demand flows can be cumbersome task, often judged not feasible in terms of computation time.

In this paper we choose a very different method for exploring the relationship between OD flows and traffic observations, and impact of demand flow changes on traffic conditions in the network. We perform sensitivity analysis based on Random Balance Design FAST (RBD-FAST) technique proposed by Tarantola et al. (2006) to identify the OD pairs whose demand variation has a significant impact on the traffic observations without explicitly relying on the assignment matrix. Moreover, we show that on an urban, large scale network with 3249 OD pairs, RBD-FAST technique identifies the most dominant OD pairs in the range of 2% of all OD pairs whose variation has a strong impact on traffic observations. Marginal effects of changes in these OD pairs with regard to traffic observations have been compared to values from traffic assignment revealing the large inconsistency. Further, we propose a heuristic method to solve the high-dimensionality of nonlinear OD estimation problem by computing the derivatives only for the most significant OD pairs with respect to traffic observations that allows the modeler to control the trade off between simplicity of the model and the level of realism.

Keywords: Dynamic OD demand estimation; sensitivity analysis; non-linear OD demand problem, traffic simulation; computational costs reduction

* Corresponding author. Tel.: +34-933-171-693 ; fax: +34-933-171-716.
E-mail address: tamara.djukic@aimsun.com

1. Introduction

Transport authorities and practitioners have long been concerned about the unavailability of reliable dynamic OD demand estimates which limits the potential for dynamic traffic assignment (DTA) deployments to analyze and alleviate traffic congestion as part of the Intelligent Transportation Systems (ITS). In congested networks, changes in the demand affect travel times. In turn, travel times affect the route choices and travel times from trip origin nodes to traffic observation detectors that determine the assignment fractions or the relationship between OD flows and traffic observations. Modelling of non-linear relationship between traffic observations and OD flows, and its dependency on variations in OD flows has been identified by many researchers as a key challenge in the estimation and prediction of a high-quality OD matrices. For example, the dependence of link-flow proportions on the demand flows in assignment matrix should be explicitly included in the DTA process. Finding derivatives of link-flow proportions and traffic observation with respect to demand flows can be cumbersome task, often judged not feasible in terms of computation time.

From a modeling point of view, the most distinguishing difference between the OD demand estimation approaches, is how the relationship between state variables (e.g., OD flows, OD proportions) and any available traffic data (e.g., link traffic counts, speeds) is defined, calculated and re-calculated through out the estimation process. An accurate description of this relationship leads to an accurate description of traffic state reality in the network, but to more complexity as well. In the past decades, a rich body of literature, stressed the need of relaxing the fixed relationship assumptions in mapping demand flows to traffic observations through estimation process when the congestion occurs in the network. Researchers have been devoted to development of the methods to capture the impact of demand variation on traffic observations, that can be categorized in analytical derivation, simulation-, numerical- and heuristic-based approximation methods.

Typically, to calculate the weights between OD flows and link traffic counts (usually measured by loop detectors in the form of sensor counts), dynamic assignment matrices are commonly used. Theoretically, these assignment matrices can be analytically derived using network topology, path choice set, current route choice model and equilibrium travel times Ashok and Ben-Akiva (2002). However, it is recognised that the complexity of the problem at hand can quickly lead to intractable situations Ashok and Ben-Akiva (2002). Further sophisticated analytical derivations are required to capture the relationship between parameters with less direct impact and non-linear relationship.

The simulation-based approximation of the relationship between demand flows and traffic observations uses traffic simulator to uncover this relationship without the direct derivation of the assignment matrix. The most studied assignment matrix-free method is the Simultaneous Perturbation Stochastic Approximation (SPSA) method (Balakrishna (2006), Cipriani et al. (2011), Cipriani et al. (2013), Cantelmo et al. (2014), Antoniou et al. (2015)) which allows one to approximate a descent gradient direction with significantly lower computational resources than through the explicit calculation. Antoniou et al. (2015) proposed the Weighted-SPSA algorithm to overcome the deteriorated performance of the gradient calculated by SPSA algorithm. Although, the main advantage of such method is that the complex relationship between demand flows and traffic observations, not only traffic counts, is estimated by simulation model, the high number of simulation runs for large-scale networks where the DTA is computationally intensive still remains to be resolved. For example, due to stochasticity of the simulation model, for each perturbation of the SPSA or W-SPSA where gradient needs to be determined, the DTA has to be replicated R times, leading to $2R$ runs Antoniou et al. (2015).

Recent studies by Toledo and Kolechkina (2013), Frederix et al. (2013) and Shafiei et al. (2017) rely on linear approximation of the assignment matrix with non-separable response in every iteration, which relaxes the assumption of constant assignment proportions and explicitly accounts the congestion effects. This definition requires the computation of the marginal effects of demand flow change on the assignment proportions at the current solution of each iteration. It is possible to use the finite deferences approach to numerically approximate the Jacobian matrix by using traffic simulator, but it would require in every iteration of the gradient solution to perturb each element in the OD demand vector, one at a time, leading to $2RDK$ runs, where D the number of OD pairs in the network and K the number of time intervals for the simulation period. To overcome computational overhead, authors proposed heuristic-based approaches. Toledo and Kolechkina (2013) neglected the effect of changes in one OD pair over the other OD pairs in the assignment matrix, Frederix et al. (2013) implemented space decomposition of the network in the congested and non-congested sub-networks, where derivatives were computed only for congested area. Shafiei et al. (2017) reduces computation costs through iterations progress and computing derivatives on OD pairs whose flows have higher tendency to vary during dynamic OD demand estimation process. However, all these approaches rely on strong heuristic assumptions such as ignoring the effect of OD demand changes outside of congested area or have been tested on relatively medium sized networks. Further research is necessary to develop solution approach for nonlinear OD estimation problem that will guaranty its reliability and applicability in large scale networks.

In this paper we choose a very different method for exploring the relationship between OD flows and traffic observations, and impact of demand flow changes on traffic conditions in the network. In other fields where exact analytical expression between input and output data is almost impossible to derive, sensitivity analysis (SA) is a commonly used method to reveal the impact of the input parameters on the model predictions. The obvious choice for a SA approach is to use some sort of affordable model-free computational method to estimate the first-order sensitivity Sobol indices, which does not scale with input dimensionality. In this paper we perform sensitivity analysis based on Random Balance Design FAST (RBD-FAST) technique proposed by Tarantola et al. (2006) to identify the OD pairs whose demand variation has a significant impact on the traffic observations without explicitly relying on the assignment matrix. RBD-FAST technique belongs to group of frequency-based SA methods, which is computationally cheaper than Monte Carlo-based techniques. In the sensitivity analysis, mesoscopic traffic simulation model in Aimsun is employed as a black-box which realistically captures the congestion phenomena that is more adequate in developing dynamic OD estimation algorithms. We show that on a urban, large scale network with 3249 OD pairs, RBD-FAST technique identifies the most dominant OD pairs in the range of 2% of all OD pairs whose variation has a strong impact on traffic observations. Marginal effects of changes in these OD pairs with regard to traffic observations have been compared to values from traffic assignment revealing the large inconsistency. Further, we propose a heuristic method to solve the high-dimensionality of nonlinear OD estimation problem by computing the derivatives only for the most significant OD pairs with respect to traffic observations that allows the modeler to control the trade off between simplicity of the model and the level of realism.

The paper is organized as follows. In the first part of the paper we will outline the motivation for the sensitivity analysis and theoretical background of RBD-FAST technique and explain its main properties. In the second part of the paper we will first define the structure of the experimental data set and the way in which we derive set of OD matrices per departure time interval for sensitivity analysis on a real urban network. Next we demonstrate how RBD-FAST technique can be used to reveal the most dominant OD pairs and consequences of assumption on fixed relationship between OD flows and traffic observations. In the third part of paper we propose the application of RBD-FAST technique as a pre-processing tool in estimation of dynamic OD demand matrices to reduce the computational costs dramatically, without significant loss of accuracy. The paper closes with a discussion on further application perspectives of RBD-FAST technique in estimation and prediction of OD demand and further research.

2. The problem formulation

This section describes the most critical issue in OD matrix estimation, whether static or dynamic, the relationship (mapping) of the observed traffic condition data with unobserved OD flows. From a modeling point of view, the most distinguishing difference between the OD demand estimation approaches presented in the literature, is how the relationship between OD flows and any available traffic data (e.g. link traffic counts, speeds, densities, etc.) is defined, calculated and re-calculated through out the estimation process. This relationship is often accomplished by means of an assignment matrix. In the dynamic problem, the assignment matrix depends on link and path travel times and traveler route choice fractions - all of which are time-varying, and the result of dynamic network loading models and route choice models. These dynamics are reflected in travel times between each origin and destination trips on a network, influenced by traffic link flow. While a vast body of literature has been developed in this area over the past two decades, this section focuses on some of the efforts that highlight the basic problem dimensions.

The general OD estimation problem is to find an estimate of OD demand matrix by effectively utilizing traffic and demand observations. Let $\Omega \subseteq U \times U$ be set of all d OD pairs in the network, and $\hat{L} \subseteq L$ be the set of r links where traffic data observations are available. The time horizon under consideration is discretized into K time intervals of equal duration, indexed by $k = 1, 2, ..., K$. If $\mathbf{x} \in \mathbb{R}^n$ represents the OD demand for each OD pair in Ω, the $\mathbf{x}_k$ represents the OD demand at departure time interval k_i, $i = 1, ..., K$. In this chapter the dynamic OD demand is represented by a vector, rather than a matrix. It is also important to define κ, the maximum number of time intervals needed to travel between any OD pair in the network. For instance, in dynamic context, depending on the size of the network and its complexity (travel times and distance from the origin o to the destination d), some vehicles could need more than one time interval to reach their destination d or pass traffic sensor at link l. The vector $\mathbf{y}_{k,\hat{L}} = A(\mathbf{x}_h) \in \mathbb{R}^r$, for time interval $h = k, k-1, ..., k-\kappa$, represents the observed link traffic data at time interval k (e.g. link traffic counts) for each link in $\hat{L}$.

Given a vector of observed traffic data at time interval k, $\mathbf{y}_k \in \mathbb{R}^r$, the dynamic OD estimation problem consists of finding an OD demand for departure time k, $\mathbf{x}_k$, such that $\hat{y}_{k,\hat{L}}(\mathbf{x}_h)$ is as close as possible to observed values $\mathbf{y}_k$. Therefore, the dynamic OD estimation problem is formulated as:

$$\hat{\mathbf{x}}_k = \underset{x \geq 0}{\operatorname{argmin}}[\alpha f(\mathbf{x}_k, \tilde{\mathbf{x}}_k) + (1-\alpha) f(\hat{\mathbf{y}}_k, \mathbf{y}_k)] \tag{1}$$

Regardless of the function f used, the purpose is to obtain an OD matrix that yields OD flows and traffic data as closely as possible to their observed values. When solving the OD problem in Eq.(1), the relationship between traffic observations and OD demand has to be defined, implicitly or explicitly. Most dynamic OD demand estimation methods, define this relationship implicitly by the assignment matrix that can be expressed as:

$$\hat{\mathbf{y}}_k = \sum_{h=k-\kappa}^{k} \mathbf{A}_k^h \mathbf{x}_h \tag{2}$$

There are two main drawbacks of relationship defined in Eg.(2):

- Separability of traffic count observations: it assumes that the traffic flow observed at the link l during time interval k can always and only be changed by changing one of the OD flows that passes link l in time interval h when $\mathbf{x_h}$ is assigned. This assumption of separability is incompatible with some typical phenomena in congested networks, such as congestion spillback between links and time lags due to the delay during congestion. In these cases it is very likely that increasing an OD flow will cause delays to other flows that do not pass that time-space interval, hereby altering the amount of flow passing the link in the considered time interval. This issue has been addressed in past studies (Yang and Zhou (1998), Tavana and Mahmassani (2001), Lundgren and Peterson (2008)). Frederix et al. (2013) suggested using the Taylor approximation to specify the linear approximation of Eg.(2) using non-separable response function, given by

$$\hat{\mathbf{y}}_k = \sum_{h=k-\kappa}^{k} \mathbf{A}_k^h(\mathbf{x}_0)\mathbf{x}_h + \sum_{h=k-\kappa}^{k} (\mathbf{x}_h - \mathbf{x}_0) \left[\sum_{h'=k-\kappa'}^{k} \frac{d(\mathbf{A}_k^{h'}(\mathbf{x}_{h'}))}{d\mathbf{x}_{h'}} \mathbf{x}_{h'} \right] \tag{3}$$

- Limited only to one data source: formulation of relationship by assignment matrix in Eg.(2) and Eg.(3) restricts dynamic OD demand estimation problem to use of traffic count data only, which can potentially over-fit to counts at the expense of traffic dynamics. Relationship between traffic condition data, such as speeds and densities, and OD flows are expected to be non-linear and approximations similar to the assignment matrix can not be justify (Balakrishna and Koutsopoulos (2008)). This issue has been addressed in the past studies (Balakrishna and Koutsopoulos (2008), Cipriani et al. (2013), Cantelmo et al. (2014), Antoniou et al. (2015)) who proposed use of traffic simulation models to capture the nonlinear relationship between OD flows and traffic observations instead of assignment matrix.

Although presented solutions significantly contributed to quality improvement of dynamic OD demand estimates, they still share a common challenge to overcome high computational costs. A complicating factor in utilizing these methods for estimation or prediction purposes, is that OD matrices are very large data structures, that grows rapidly in large networks. Even in case such high-dimensional OD flows can be reduced (see e.g. Djukic et al. (2012)), and this is not entirely unlikely, there are serious methodological difficulties in finding optimal solutions (e.g. getting stuck in local minima, slow convergence, high number of simulation runs, etc.), aside from the computational and memory requirements for such a procedure on the basis of thousands (to millions) of traffic observations. For example, computing the exact Jacobian vector in the second term of Eg.(3) with respect to changes in OD flows for each OD pair remains intractable even when efficient, well calibrated, DTA model is used. In the next section, we propose a different method for exploring the relationship between OD flows and traffic observations by applying sensitivity analysis. We relax assignment matrix dependence of Eg.(3) to evaluate marginal effect of demand flow changes on traffic conditions in the network.

3. The concept of sensitivity analysis based on RBC

One way to explore the marginal impact of the changes in the OD demand on traffic observations, is to identify the most sensitive input parameters, i.e., OD pairs, through sensitivity analysis. Sensitivity analysis (SA) of model output investigates how the predictions of a model are related to its input parameters. In doing so, the traffic simulation software is considered a black box in function form $Y = f(X)$, providing a certain outcome Y (traffic observations) given certain input (OD demand), X. The obvious choice for a SA approach is to use some sort of affordable model-free computational method to estimate the first-order sensitivity Sobol indices, which does not scale with input dimensionality. There are various solutions to limit the large number of samples needed without fundamentally changing the overall idea. In this sensitivity analysis, the solution strategy is to use so-called RBD-FAST (Random balance design) technique from frequency-based SA methods, which is computationally cheaper than Monte Carlo-based.

The original formulation of FAST (Fourier amplitude sensitivity test) technique is introduced by Cukier et al. (1978). Sample points of input parameter space are chosen such that the indices can be interpreted as amplitudes obtained by Fourier analysis of the function. Further extension of this method provided by Tarantola et al. (2006) to avoid the problem of interferences, is RBD-FAST, where the RBD stands for random balanced design. RBD-FAST is a group of modifications of FAST technique, which use random permutations of design points to avoid interferences.The original idea by Tarantola et al. (2006) is to assign the same frequency to all input variables, but to randomize their values independently before evaluating f. The first-order Sobol index of X_j, for $j = 1, ..., N$ can then be estimated by reordering the evaluations in the way X_j was permuted, so that the amplitude at the frequency returns the sensitivity of X_j only.

3.1. Theoretical background of RBD

The RBD-FAST approach adopted in this paper to compute the entire first-order sensitivity indices can be summarized in the following points:

1. Generate a (d, N) permutation matrix P of random numbers where d represents number of OD pairs and N number of experiments for sensitivity analysis. Each row of the permutation matrix P represents the sequence of quasi-random numbers between 0 and 1 by step $1/N$. Randomly permute the rows of this matrix to generate a set of scrambled points that cover the factor space.
2. Generate a (d, N) frequency matrix U of N points over $(-\pi, \pi)$ using the parametric equation 4:
$$u_{d,n} = 1 + \frac{1}{\pi}(arcsin(sin(2\pi p_{d,n}))) \tag{4}$$
The elements of frequency matrix will have values in range between 0.5 and 1.5
3. Define a (d, N) demand matrix X of perturbed OD demand vectors, by multiplying OD demand vector by frequency matrix.
$$X^{'} = \mathbf{x} \times U \tag{5}$$
4. The traffic simulation is run and evaluated for all the $[Nx1]$ combinations of input demand variables as given by matrix X to produce the $(Nx1)$ vector of outputs $y_{\mathbf{x}'} = f(\mathbf{x}^{'})$. This vector is sufficient for the evaluation of all the first-order sensitivity indices. The total cost of this approach is kept down to N simulation runs, instead of $d \times N$ like in FAST as it is independent of the input parameter size, i.e. the number of OD pairs. Since N is usually not lower then 1000, the number of runs required by this efficient approach is not, in any case, negligible, particularly for complex and expensive traffic simulation models and high-dimensional demand vectors. For this reason, in the common practice, the approach presented here can be considered relevant.

The values of model output Y_{x_j}, $j = 1, ..., N$, are then re-ordered such that the corresponding values of $X_1(x_{1j})$ are ranked in increasing order. By doing this, the harmonic content of X_1 propagates through f to $Y^R(x_j)$. The sensitivity of Y to X_1 is determined by the harmonic content of Y^R, which is quantified by normalized Fourier coefficients. For a weight ω, the Fourier coefficients for each OD pair can then be numerically estimated by

$$A_\omega = \frac{1}{n}\sum_{j=1}^{n} f(x(s_j)cos(\omega s_j) \qquad \text{and} \qquad B_\omega = \frac{1}{n}\sum_{j=1}^{n} f(x(s_j)sin(\omega s_j) \tag{6}$$

and the first-order sensitivity of the i^{th} OD pair, D_i, can be estimated by the sum of the Fourier coefficients up to the order M, where in practice $M = 4$ or $M = 6$ are used

$$D_i = 2\sum_{m=1}^{M}(A^2_{m\omega_i} + B^2_{m\omega_i}). \tag{7}$$

Note that this method is very fast to compute sensitivity when there are a lot of variables (compared to other SA methods), but it is limited to find the correlation of input and output variables. However, we use this method only to identify the OD pairs whose variance has the strongest impact on the traffic observations that will be used further to compute partial derivatives.

There are no universal recipes for the choice of N. It can vary from few hundreds to several thousands. In order to assess if the sensitivity indices calculated for a given N are sufficiently stable, one can perform calculation of their confidence intervals or repeating the SA by increasing sample sizes in order of $N = 500$, $N = 1.000$ and $N = 2.000$, etc. For this paper, the results of the indices calculations and analysis will be presented in the graphical form in the following section.

4. Numerical experiment design

This section describes how the variations of demand in OD matrices per departure time were set up for the experiments and how these data were organized for sensitivity analysis with mesoscopic simulation model in Aimsun.

4.1. Sensitivity analysis with mesoscopic simulation model

In this case study, an advanced traffic simulation model Aimsun, TSS-Transport Simulation Systems (2013), is used as a black-box to perform the proposed sensitivity analysis. Aimsun is the leading traffic modelling software environment, that stands out for the exceptionally high speed of its simulations and for integrating travel demand modelling, hybrid microscopic-mesoscopic traffic simulation and dynamic traffic assignment all within a single software application. In our case study, we use the mesoscopic event-based demand and supply models in Aimsun, each synthesising microscopic and macroscopic modelling concepts. They couple the detailed behaviour of individual drivers' route choice behaviours with more macroscopic models of traffic dynamics. The travel demand in Aimsun is represented by dynamic OD demand matrices. Vehicle generation is done for each OD pair separately with arrival times that follow an exponential distribution. The iterative interaction between demand and supply models allows the system to update the set of routes and the travel times after each iteration leading to robust estimation and prediction of traffic conditions in the network.

For this study, a route choice set is pre-computed in Aimsun and used as fixed for all the simulation runs in sensitivity analysis. In this way, dependence of rerouteing effects on the changes in the OD demand is ignored. Here we focus to investigate effects of travel time variation and congestion spill-back on traffic observations in the network.

Figure 1 presents a sensitivity analyses flowchart with the main elements implemented in Aimsun. For each OD demand vector generated by RBD-FAST technique, the Aimsun call function (AIMSUN.m) is initiated. This function converts the demand to be simulated to the Aimsun format, creates the batch file to execute the requested simulations, generates the Python file with the Aimsun run flags and finally calls and executes mesoscopic traffic simulation in Aimsun with these inputs and fixed paths. After the simulation runs have been completed, it imports the observed traffic data and the simulation outputs and calculates the GoF measures that were defined within the algorithm and

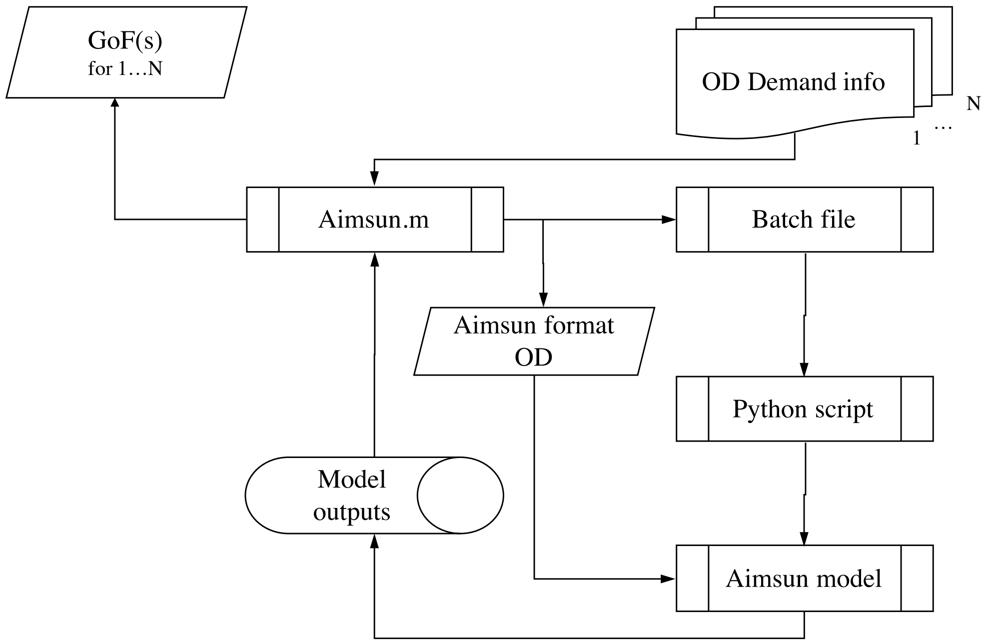


Fig. 1: Sensitivity analysis flowchart with the main elements.

assignment matrices. Since the possible measures of performance are all the time series of counts at the existing detectors (see Figure 2), a strategy to aggregate them in a single measure needs to be put in place. In order to assess the dependence from the GoF measure selected, we used three of them, namely, the root-mean-square error, the mean absolute error and GEH statistics.

4.2. Network and traffic data

Prior to demand variation modelling, a Vitoria network has been selected, consisting of 57 centroids, 3249 OD pairs with a 600km road network, 2800 intersections and 389 detectors, presented as black dots in Figure 2. This network is available in the mesoscopic version of the Aimsun traffic simulation model for the reproduction of traffic propagation over the network. The true OD demand is available for this network, which allows analysts to assess the performance of the proposed method. The true assignment matrix and traffic counts on detectors are derived from the assignment of true OD matrix in Aimsun for the evening period from 19:00 to 20:00 reflecting a congested state of the network. The simulation period is divided into 15 minute time intervals with an additional warm-up time interval, K = 5. The link flows resulting from the assignment of the true OD demand are used to obtain the traffic count data per observation time interval.

The trips between some of the OD pairs are not completed within one time interval due to congestion on the network or the distance between OD pairs. In this way, a vehicle entering the network during a particular departure time interval might need more than one time interval to reach a traffic detector, where the departure time interval and detection time are different. In the chosen study network, the maximum travel time between OD pairs observed on the network takes three time intervals, which leads to very sparse assignment matrices, and the number of lagged time intervals k = 3.

5. Results of the sensitivity analysis

Steps from 1 to 3 of the RBD-FAST technique presented in Section 3 have been implemented on the ground truth OD demand vector to obtain the set of OD demand vectors for sensitivity analysis. Equation (4) produces frequency depended data such that each OD pair distribution will have zigzag line shape with height between 0.5 and 1.5. In

Fig. 2: The Vitoria network, Basque Country, Spain.

this way, the mean value of demand per OD pair is 100% of the nominal demand value, with a variance of +/- 50%. In addition heuristic assumption on OD pairs selection is employed to reduce the computation time for the sensitivity analysis. In large-scale dynamic OD matrices, many OD pairs have small demand values. The impact of these individual OD pairs are insignificant on the network and hence, the sensitivity analysis for each of them is unnecessary. Therefore, a set of candidate OD pairs whose demand is higher then 30 trips per time interval is defined on which the sensitivity analysis is applied. This assumption led to selection of 125 out of 3249 OD pairs for the analysis. A total of $N = 2500$ simulations were carried to compute the sensitivity indices by RBD-FAST technique.

5.1. Detection of the most dominant OD pairs

A number of approaches are available to determine which outputs are mainly driven by the input parameters. In general, the choice of the OD pairs that have the highest impact on traffic observations is made by the visual examination of a number of different criteria. The simplest criterion relates to plotting of the sensitivity indices sorted by size, that is, by preparation of a so-called scree plot and evaluation of the plot for an elbow. Figure 3 presents the scree plot of sensitivities obtained from the GEH values as GoF measure calculated on traffic counts data from all the loop detectors for visual examination of the sensitivity indices magnitudes.

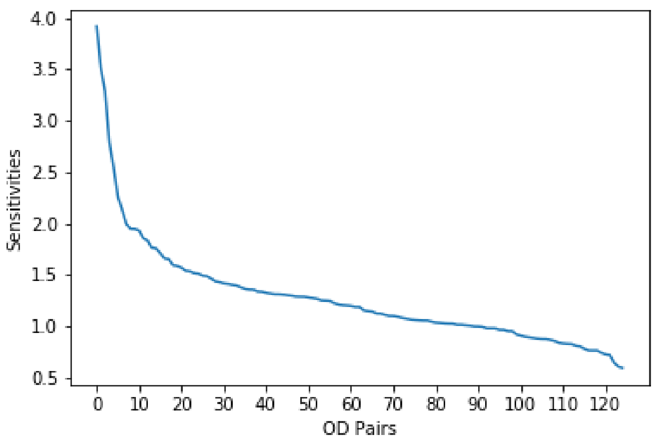
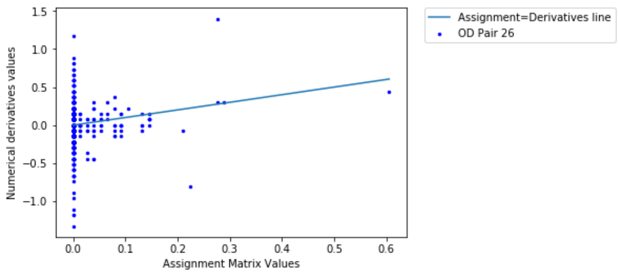


Fig. 3: Scree-plot of first-order indices on the GEH values using traffic counts from all the detectors.

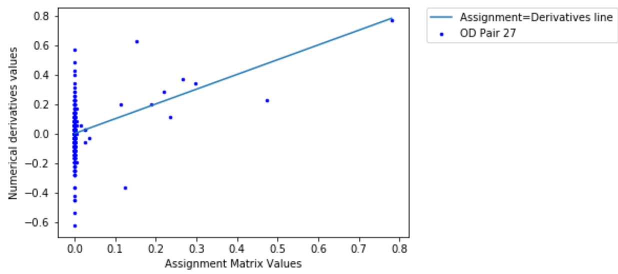
The figure here presented is based on GEH measure since it provides the most consistent and stable results. A sharp elbow appears at about the seventh OD pair, which shows that the effective variance in the time series of OD demand data is far smaller than the apparent dimensionality of the total number of OD pairs in the network.

5.2. Evaluation of marginal effects with respect to changes in OD flows

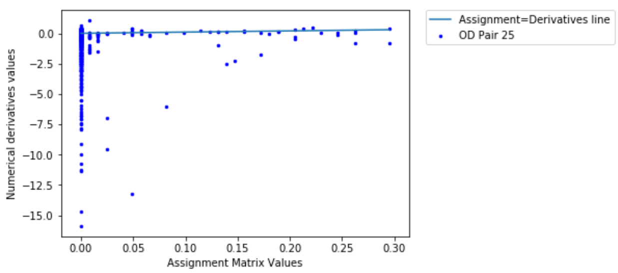
To investigate consequence of ignoring the marginal effect with respect to changes in OD demand, i.e. by keeping the assignment matrix fixed through dynamic OD demand estimation, effect of the first 7 OD pairs has been selected for analysis. In this analysis, we have computed the numerical derivatives using finite differences method of the traffic counts with respect to changes in these 7 OD pairs. Numerical derivatives are computed by increasing and decreasing the demand for each OD pair by 5% over 10 replications, leading to 140 simulation runs. Figure 4 shows the comparison of the assignment matrix and derivatives values for the most dominant OD pairs. It is clear from the figures 4c, 4a, 4b, and 4d that the values in the assignment matrix are not always a good approximation of the derivative of the traffic counts with respect to changes in OD demand.



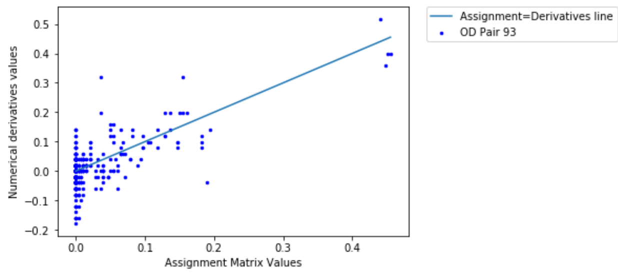
(a) OD pair: 26



(b) OD pair: 27



(c) OD pair: 25



(d) OD pair: 93

Fig. 4: Comparison of the values of the assignment matrix and the values of the derivatives

Another common feature that dominates in the Figure(s) 4c, 4a, 4b, and 4d is that there are many loop detectors with assignment value equal to 0 but non-zero derivative. These deviation points reflect loop detectors that are not located on the routes of investigated OD pair but still affected when the demand on the OD pair changes, mainly due to congestion spillback between links, time lags due to delay during congestion, and interdependencies between crossing (or opposing) flows encountered at intersections. For this reason the derivatives have a bias towards negative values along the line with assignment values equal to zero. This was expected since it is very likely that increasing an OD flow will cause delays to other flows that do not pass time-space interval, hereby altering the amount of flow passing the detector in the considered time interval and detecting many less vehicles. Figure 5 depicts this phenomena for the same OD pair with departure time at 19:00 and 19:30. These results are consistent with earlier observations in the literature (e.g.Tavana and Mahmassani (2001), Lindveld et al. (2003), Lundgren and Peterson (2008), Frederix et al. (2013)), where assumption on link flows separability in congested networks has been criticised.

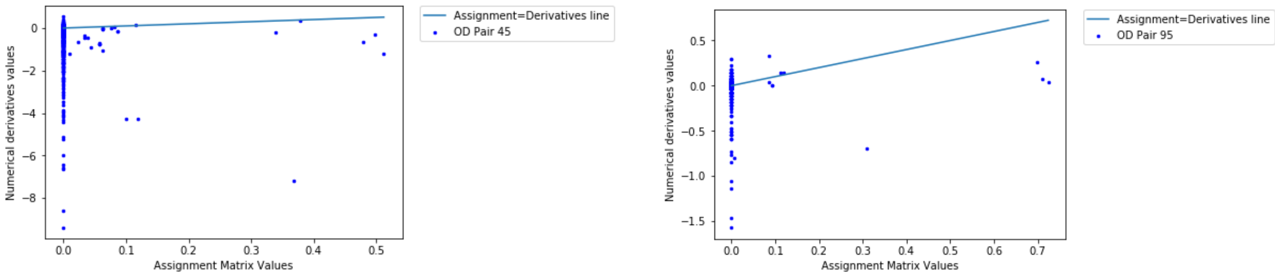


Fig. 5: Comparison of the values of the assignment matrix and the values of the derivatives for OD pair 45. (a) OD pair:45 with departure time at 19:00; (b) OD pair:45 with departure time at 19:30.

5.3. Application example: a heuristic approach to estimate and predict OD matrices

Dynamic OD demand estimation and prediction methods consist in predicting unknown dynamic OD matrices based on past known high dimensional historical OD matrices data and traffic observations. Solution algorithms proposed in literature to solve the problem given in Eq. (1), incorporate computing the marginal effects of demand changes on traffic observations that lead to high computational costs for medium- or large-scale networks. In this situation, dimensionality reduction of simulation runs required to capture these marginal effects is necessary leading to improve prediction computational performance.

In order to overcome problem related to dimensionality of OD demand we propose the application of sensitivity analysis based on RBD-FAST technique as a pre-processing tool with historical OD demand data before estimation and prediction process. As we shown in previous sections, RBD-FAST technique provides tools for identification and selection of the OD pairs whose changes significantly dominate changes in traffic conditions and observations in the network. Here we propose a heuristic approach that uses RBD-FAST technique in the following steps:

1. **Identification step**: first in this step we run RBD-FAST technique to explore impact of the variability patterns in OD demand matrices on traffic observations using traffic simulation model as a black-box, as we described in previous sections. In this manner, we can identify the OD pairs that significantly dominate changes in traffic conditions in the network.
2. **Reduction step**: in this step we propose the selection criteria of OD pairs based on sensitivity indices magnitudes. For example, only the OD pairs having a sensitivity indices above the specified threshold value will be selected.
3. **Estimation step**: finally, OD demand estimation and prediction methods are applied such that numerical derivatives of changes in OD demand on traffic observations are computed only for the chosen subset of OD pairs to estimate OD demand.

Proposed heuristic solution approach can be employed to reduce the computation time for the dynamic OD demand estimation and prediction. There are several reasons why the variations of many OD pairs may not have a significant effect on the dynamic OD demand estimation process:

- In large-scale dynamic OD demand matrices, many OD pairs have small values. The impact of these individual OD pairs are insignificant on the network and hence, the sensitivity analysis for each of them is unnecessary.
- Flows associated with many OD pair do not cross the congested area. As a result, the new demand can be accurately estimated using the traditional assignment matrix-free method.
- Within a few iterations a number of OD demand flows converge and would not change considerably in the subsequent iterations. Thus, these OD pairs are excluded from the sensitivity analysis.

6. Conclusions

The common approach usually adopted in dynamic OD demand estimation and prediction consists in solving an optimization problem in which the distance between observed and simulated traffic conditions is minimized by assuming the relationship between OD flows and traffic observations independent of traffic conditions in the network. This approach has a sever shortcoming as it does not take into account the impact of demand flows variation on traffic observations in congested networks. Modelling of non-linear relationship between traffic observations and OD flows, and its dependency on variations in OD flows has been identified by many researchers as a key challenge in the estimation and prediction of a high-quality OD matrices.

In this paper, the entire non-linear relationship has been formulated and evaluated under the light of sensitivity analysis. The SA is crucial to individuate the most important input, OD flows, to predict traffic observations. The Random Balance Design FAST (RBD-FAST) technique for the SA is presented to identify the OD pairs whose demand variation has a significant impact on the traffic observations without explicitly relying on the assignment matrix. The SA has been performed on an urban, large scale network with 3249 OD pairs, where using RBD-FAST technique we identified the most dominant OD pairs in the range of 2% of all OD pairs whose variation has a strong impact on traffic observations. Marginal effects of changes in these OD pairs with regard to traffic observations have been compared to values from traffic assignment revealing the large inconsistency. These deviation reveals that traffic observations on loop detectors that are not located on the routes of investigated OD pair are still affected when the demand on the OD pair changes, mainly due to congestion spillback between links, time lags due to delay during congestion, and interdependencies between crossing (or opposing) flows encountered at intersections. These findings may help further improve development of heuristic approaches for dynamic OD demand estimation or used to evaluate assumptions in these models.

Many aspects of this proposed approach need further research. The evaluation of different GoF measures in SA should be carried out. For example, we used a scalar GEH value in SA: the sum of all distances between traffic simulation outputs and corresponding observed traffic measurements. This value is valid in case of highly correlated input parameters, OD flows, and traffic observations. However, in a real-world traffic system, correlations between OD flows and traffic observations could be sparse which implies further need for selection of appropriate performance measures. In addition, alternative SA methods in proposed methodology should be studied, particularly when RBD-FAST design still rely in an excessive number of simulations and limits the application of proposed methodology for off-line OD demand estimation. Finally, for computational efficiency assessment, the proposed methodology should be applied to traffic simulation model without fixed paths. Exploring the impact of OD demand variation on route choice selection still remains a part of future work.

Since our final objective was to show application of RBD-FAST technique as a preprocessing tool in estimation of dynamic OD matrices, we proposed a heuristic method to solve the high-dimensionality of nonlinear OD estimation problem by computing the derivatives only for the most significant OD pairs with respect to traffic observations. We suggest further development of dynamic OD demand estimation models based on these findings.

Acknowledgements

We acknowledge the support of the SETA project funded from the European Unions Horizon 2020 research and innovation program under grant agreement No 688082.

References

Antoniou, C., Azevedo, C.L., Lu, L., Pereira, F., Ben-Akiva, M., 2015. W-spsa in practice: Approximation of weight matrices and calibration of traffic simulation models. Transportation Research Part C: Emerging Technologies 59, 129 – 146. Special Issue on International Symposium on Transportation and Traffic Theory.

Ashok, K., Ben-Akiva, M.E., 2002. Estimation and prediction of time-dependent origin-destination flows with a stochastic mapping to path flows and link flows. Transportation Science 36, 184–198.

Balakrishna, R., 2006. Off-Line Calibration of Dynamic Traffic Assignment Models. Ph.D. thesis. Department of Civil and Environmental Engineering, Massachusetts Institute of Technology.

Balakrishna, R., Koutsopoulos, H., 2008. Incorporating within-day transitions in the simultaneous off-line estimation of dynamic origin-destination flows without assignment matrices. Transportation Research Record , 31–38.

Cantelmo, G., Cipriani, E., Gemma, A., Nigro, M., 2014. An adaptive bi-level gradient procedure for the estimation of dynamic traffic demand. Intelligent Transportation Systems, IEEE Transactions on 15, 1348–1361.

Cipriani, E., Florian, M., Mahut, M., Nigro, M., 2011. A gradient approximation approach for adjusting temporal origin-destination matrices. Transportation Research Part C: Emerging Technologies 19, 270–282.

Cipriani, E., Gemma, A., Nigro, M., 2013. A bi-level gradient approximation method for dynamic traffic demand estimation: sensitivity analysis and adaptive approach. Proceedings of the IEEE conference on Inteligent TRansport Systems, 16th IEEE ITSC 1.

Cukier, R., Levine, H., Shuler, K., 1978. Nonlinear sensitivity analysis of multiparameter model systems. Journal of Computational Physics 26, 1 – 42.

Djukic, T., Flotterod, G., van Lint, H., Hoogendoorn, S., 2012. Efficient real time od matrix estimation based on principal component analysis, in: Proceedings of the IEEE Intelligent Transportation Systems Conference, IEEE. pp. 115–121.

Frederix, R., Viti, F., Tampére, C.M., 2013. Dynamic origin-destination estimation in congested networks: theoretical findings and implications in practice. Transportmetrica A: Transport Science 9, 494–513.

Lindveld, K., Netherlands Research School for Transport, I., Logistics, 2003. Dynamic O-D matrix estimation: a behavioural approach. TRAIL thesis series, Netherlands TRAIL Research School.

Lundgren, J.T., Peterson, A., 2008. A heuristic for the bilevel origin?destination matrix estimation problem. Transportation Research Part B: Methodological 42, 339 – 354.

Shafiei, S., Saberi, M., Zockaie, A., Sarvi, M., 2017. A sensitivity-based linear approximation method to estimate time-dependent origin-destination demand in congested networks. Proceedings of Transportation Research Board - 96th Annual Meeting , 1–16.

Tarantola, S., Gatelli, D., Mara, T., 2006. Random balance designs for the estimation of first order global sensitivity indices. Reliability Engineering System Safety 91, 717 – 727.

Tavana, H., Mahmassani, H., 2001. Estimation of dynamic origin-destination flows from sensor data using bi-level optimization method. Presented at the 80th annual meeting of the Transportation Reasearch Board, Washington DC, USA .

Toledo, T., Kolechkina, T., 2013. Estimation of dynamic origin-destination matrices using linear assignment matrix approximations. IEEE Transactions on Intelligent Transportation Systems 14, 618–626.

TSS-Transport Simulation Systems, T., 2013. Aimsun 7 dynamic simulator User's Guide. version 1.1 ed., Transport Simulation Systems, Barcelona, Spain.

Yang, H., Zhou, J., 1998. Optimal traffic counting locations for origin-destination matrix estimation. Transportation Research Part B: Methodological 32, 109–126.