Towards a generic benchmarking platform for OD flows estimation/updating algorithms: design, demonstration and validation

Constantinos Antoniou [1], Jaume Barceló [2], Martijn Breen [3], Manuel Bullejos [2], Jordi Casas [3], Ernesto Cipriani [4], Biagio Ciuffo [5], Tamara Djukic [6], Serge Hoogendoorn [6], Vittorio Marzano [7,*], Lídia Montero [2], Marialisa Nigro [4], Josep Perarnau [3], Vincenzo Punzo [7], Tomer Toledo [8], Hans van Lint [6]

[1] *National Technical University of Athens, Greece*
[2] *Technical University of Catalonia (UPC), Barcelona, Spain*
[3] *Transport simulation systems, Spain*
[4] *"Roma Tre" University, Italy*
[5] *Institute for Energy and Transport of the European Commission, Joint Research Centre, Ispra (Varese), Italy*
[6] *Delft University of Technology, the Netherlands*
[7] *University of Napoli "Federico II", Italy*
[8] *Technion – Israel Institute of Technology, Israel*

* corresponding author: Dipartimento di Ingegneria Civile Edile e Ambientale, Via Claudio 21 – 80125 Napoli (Italy) – vmarzano@unina.it – Tel: +390817683935

Abstract

Estimation/updating of origin-destination (OD) flows and other traffic state parameters is a classical, widely adopted procedure in transport engineering, both in off-line and in on-line contexts. Notwithstanding numerous approaches proposed in the literature, there is still room for considerable improvements, also leveraging the unprecedented opportunity offered by information and communication technologies and big data. A key issue relates to the unobservability of OD flows in real networks – except from closed highway systems – thus leading to inherent difficulties in measuring performance of OD flows estimation/updating methods and algorithms. Starting from these premises, the paper proposes a common evaluation and benchmarking framework, providing a synthetic test bed, which enables implementation and comparison of OD estimation/updating algorithms and methodologies under "standardized" conditions. The framework, implemented in a platform available to interested parties upon request, has been flexibly designed and allows comparing a variety of approaches under various settings and conditions. Specifically, the structure and the key features of the framework are presented, along with a detailed experimental design for the application of different dynamic OD flow estimation algorithms. By way of example, applications to both off-line/planning and on-line algorithms are presented, together with a demonstration of the extensibility of the presented framework to accommodate additional data sources.

Keywords

Traffic modelling, origin-destination (OD) estimation/updating, benchmarking platform.

1. Background and motivation

Traffic congestion has been plaguing urban and interurban transportation systems everywhere for

1

decades. It has been widely accepted that it is not practical or feasible to attempt to counter congestion through the construction of additional capacity. Instead, smarter approaches, including traffic management and traffic control, are being followed: traffic management is a mature field with many applications and an important role in modern cities and motorway systems worldwide. One of the key inputs required for the evaluation of traffic management and policy measures are dynamic Origin-Destination (OD) demand matrices[1], providing an aggregate and concise way to describe the overall demand patterns, and a practical way to load them onto a simulation model, used for the evaluation and assessment of the considered strategies. Dynamic OD matrices are also essential inputs to a wide range of traffic models, including dynamic microscopic and mesoscopic traffic simulators (Barceló, 2010), that have experienced a significant maturation, becoming suitable tools for almost any type of traffic analysis applications, and are nowadays extensively adopted both for research and practice purposes. However, there are several challenges associated with the use of OD matrices.

Normally, a prior OD flow estimate is obtained through a combination of surveys and mathematical models: see Cascetta (2009), Ortuzar and Willumsen (2011), Simonelli et al. (2012) for thorough reviews. Unfortunately, the resulting estimate is often affected by substantial errors, mainly related to the inherent complexity of the behavioural phenomena underlying demand patterns (e.g. departure time and destination choices). These errors may be mitigated by updating prior OD flow estimates, using observed traffic counts and other traffic flow measurements: that is, updated (posterior) OD flows are found using a properly specified objective function, such that they are able to fit satisfactorily available traffic flow measurements. Estimation/updating of OD flows based on traffic counts is a classical and widely adopted procedure in transport engineering, both in off-line (e.g. for medium- to long-term planning and design) and in on-line (e.g. for designing descriptive and/or prescriptive information to be disseminated in real time through Advanced Traffic Management and Information Systems ATIS/ATMS) applications. From the literature, summarized in the next section, four key difficulties in the evaluation and cross-comparison of various dynamic OD matrix estimation/updating methods can be identified (Djukic, 2014).

The first and main difficulty, as pointed out amongst others by Yang et al. (1991), Marzano et al. (2009) and Cascetta et al. (2013), is that OD flows are normally unobservable and thus it is very difficult to detect the quality of OD estimation/updating procedures based on real datasets. In this respect, a main factor is represented by the high indeterminateness of the dynamic OD estimation/updating based on traffic flow measurements, since the number of unknowns (i.e. OD flows) is much larger than the number of equations (traffic counts and possibly other network measurements). In fact, matching closely the observed traffic measurements used for OD estimation/updating is not in general a necessary indication of a good quality of OD estimation/updating. As a result, a major contradiction is that, quite frequently, sophisticated traffic models are fed with very rough and low-quality information on the time variability of traffic patterns. This leads to situations wherein it is hard for the analyst to identify whether flaws in the intended model are due to modelling mistakes, an improperly calibrated model or an unsuitable specification of the time-varying demand.

Secondly, there is no undisputed performance indicator for estimated/updated dynamic OD matrices. Instead, many candidate statistical metrics can be used to evaluate the quality of estimated OD matrices with respect to available ground-truth OD matrices (Marzano et al., 2009). The consequence of this wide range of statistical metrics is that researchers and practitioners may use different metrics according to their needs, rather than as objective assessment criteria. Few studies have focused on the evaluation of uncertainty and quality of the estimated dynamic OD matrices in absence of ground-truth OD matrices: for instance, Yang et al. (1991) introduced the maximum possible relative error (MPRE), whilst Bierlaire (2002) has proposed the total demand scale (TDS), a measure dependent

---

[1] In the following, we will refer to them interchangeably as either OD matrices or OD flows.

2

neither on the prior OD matrix nor on the traffic data observations, which evaluates the level of arbitrariness introduced by the prior OD matrix in the estimation process. However, such performance indicators should be used in addition to, not instead of, standard statistical metrics.

The third difficulty is that the assessment of OD estimation methods is difficult to generalize, as there are many aspects in which these methods differ. In fact, various reasons, such as familiarity with alternative OD estimation methods and selection of suboptimal parameter values, could lead to a not completely fair and informative comparison between methods.

A fourth challenge associated with the comparison of dynamic OD matrix estimation methods is the lack of consistence in the experimental design and presented results. Each researcher/developer can test their algorithms and methods under different assumptions, with different networks and traffic conditions, using different traffic data and performance indicators. Thus, even if a proposed approach is contrasted with alternative approaches, the comparison might not be completely fair and informative due to various reasons, such as familiarity with the alternatives and selection of suboptimal parameter values. As a result, one of the biggest obstacles in assessing OD estimation/updating algorithms is the lack of consistency in the presented results.

In order to solve the aforementioned challenges, a considerable opportunity is offered by an approach based on synthetic data, often used to perform "controlled" experiments in order to test systems, models or algorithms. In particular, they are well suited to verify inverse problems where, typically, the inputs to the forward problem are not measurable. In case of traffic simulation, for instance, usual questions are: "given some observed traffic, find the values of model parameters that allow reproducing that traffic" (i.e. parameter calibration) or "find the input demand that has produced that traffic" (i.e. OD demand estimation). In such examples, synthetic values for parameters or input demand are first generated, and the model itself is applied to produce synthetic traffic measurements (forward problem). Such synthetic measurements are used as input to the inverse problem, whose aim is to rediscover the values of parameters or demand that generated the synthetic measurements. Since neither model parameters nor OD demand are observable, a test with synthetic data is the only feasible way to assess the performance of the algorithms applied to solve the inverse problem, or to verify whether the problem itself or its setting are well-posed, at least in a controlled experiment.

Drawing upon such premises, this paper describes the development of a common evaluation and benchmarking framework to test methodologies and algorithms for dynamic OD flows estimation/updating. Specifically, the main objective is to provide a testing framework in which a number of algorithms can be implemented and tested under equal and "standardized" conditions. Design and implementation of the benchmarking platform – available to interested parties upon request – are very flexible and allow comparing considerably different approaches in a variety of settings and conditions.

The remainder of this paper is structured as follows. Section 2 presents a general literature review on dynamic OD flow estimation/updating. Section 3 describes design and implementation of the benchmarking framework. Section 4 outlines the experimental design process for the conducted case studies, describing the algorithmic approaches considered for testing of the framework. Section 5 provides details on the setup of the platform implementing the benchmarking framework, in order to show its capabilities and characteristics, whilst Section 6 presents some selected results from the application of the platform. Finally, Section 7 provides concluding discussions and statements.

2. Literature review on relevant topics

In accordance with Cascetta et al. (2013), estimation/updating of OD flows has been studied extensively first in static networks, following four main approaches: minimum information/maximum entropy (Van Zuylen and Willumsen, 1980); maximum likelihood (Bell, 1983; Cascetta and Nguyen,

1988), Generalized Least Squares (Cascetta, 1984), Bayesian (Maher, 1983). Drawing upon these broad general approaches, a number of generalizations have been proposed: examples include incorporating the treatment of congested networks through bi-level optimization (Florian and Chen, 1995; Yang, 1995; Cascetta and Postorino, 2001), taking into account the stochastic nature of traffic counts (Lo et al., 1996; Vardi, 1996), estimating simultaneously OD flows and route choice model parameters (Lo and Chan, 2003), or dealing with the availability of traffic counts on multiple days (Hazelton, 2003).

Dynamic OD flow estimation/updating started with the seminal contributions by Okutani and Stephanedes (1984) and Cremer and Keller (1987), which were translated into a practical generic optimization problem by Cascetta et al. (1993). Since then, dynamic OD flows estimation/updating has received a lot of attention in the literature, with continuous improvement and extension, resulting in a variety of approaches: a review of early contributions in this field is reported by Balakrishna et al. (2005), whilst Toledo et al. (2015) surveys more recent research attempts. For the purposes of the paper, it is worth mentioning some key issues.

Non-linearity. Research efforts have emphasized the importance of accurate modelling of the dependence of link-flow proportions on OD flows in congested networks, and of the consistent non-linear relationship between traffic measurements and OD flows, tackled with different solution frameworks. Chang and Wu (1994) were the first to attempt to formulate a non-linear relationship between OD distributions and link traffic counts under congested conditions on highway networks, using Extended Kalman filter (EKF) estimators, subsequently extended by Chang and Tao (1996) to more general networks. The influence of non-linearity of the link-route proportion matrix on OD estimation has been studied further by several authors (e.g. Flötteröd and Bierlaire, 2009; Frederix et al., 2011). In Frederix et al. (2011) the method of marginal computation is used to efficiently calculate the sensitivity of link flows with respect to OD flows, not captured by the linear approximation of the relationship between OD flows and link traffic counts. In this respect, Frederix et al. (2010) have indicated that applying a dynamic network loading (DNL) model that represents queuing properly is only a necessary condition (i.e. not sufficient) to guarantee accurate OD flow estimation[2]. Normally, in terms of formulation, the dynamic OD flows estimation problem for congested networks has typically been specified either as a bi-level optimization problem (Tavana, 2001; Van der Zijpp and Lindveld, 2001; Zhou et al., 2003; Lindveld et al., 2003; Zhou & Mahmassani, 2006) or bypassing the complex relationship between the OD flows and traffic flow measurements by treating the assignment model as a black box (Balakrishna et al., 2007; Cipriani et al., 2011) or through proper linearization (Lundgren and Peterson, 2008; Nie et al., 2006; Flötteröd and Bierlaire, 2009). Obviously, the choice of approach implies also a consistent choice of the solution algorithm. In this respect, many more recent works (Balakrishna, 2006; Balakrishna and Koutsopoulos, 2008; Cipriani et al., 2011; Frederix et al., 2011) deal with the Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm (Spall, 1992) and its recent variants/improvements (Lu et al. 2015, Antoniou et al., 2015, Cantelmo et al., 2014b, Tympakianaki et al., 2015). SPSA efficiently approximates the gradient with only two successive evaluations of the objective function, independently of the number of parameters, thus with computational time savings for large-scale problems over traditional gradient methods (e.g. finite-differences stochastic approximation, FDSA). However, SPSA is essentially a hill-climbing method, herewith at best finding a local optimum; it only uses the aggregated error between model output and observed measurements, leading to loss of the great amount of information (e.g. location and time

---

[2] Frederix et al. (2010) note that convergence is not guaranteed when prior OD flows are not close to real OD flows. Reliability of prior OD flows has been also the starting point for further research as reported in Cantelmo et al. (2014a) and Cantelmo et al. (2015), who proposed a procedure to specify proper prior OD flows, focusing on specific OD pairs in different steps or by exploiting information on aggregated demand data such as generation data by zones.

interval of different measurements); in addition, noise from uncorrelated measurements is introduced, which grows rapidly with the size of the network and the number of intervals[3]. Other optimization techniques have been widely investigated in dynamic OD estimation, for instance Balakrishna (2006) dealt with the possible application of response surfaces techniques.

Observability. Drawing upon Yang et al. (1998) and Marzano et al. (2009), an important issue in dynamic OD flows estimation/updating refers to the high indeterminateness of the problem, since -as mentioned in Section 1- the number of unknowns (i.e. OD flows) is much larger than the number of equations (traffic flow measurements). In this respect, some authors tackled this issue with various approaches, for instance Flötteröd and Bierlaire (2009) and Djukic et al. (2012) explored methods to reduce the high dimensionality of OD estimation problem using principal component analysis (PCA). Cascetta et al. (2013) proposed recently a quasi-dynamic approach, providing a more effective unknowns/equations balance, shown to outperform traditional dynamic OD flow estimators. Cipriani et al. (2014) presented an application of a quasi-dynamic traffic assignment model that approximates the dynamic traffic model by steady-state intervals and applies approximate performance functions to reduce the computation burden of the estimation. Marzano et al. (2015) extend the quasi-dynamic framework to online applications, drawing upon Ashok (1996).

Selection of measurements. Another key issue is the selection of traffic flow measurements as inputs for dynamic OD flow estimation: indeed, in addition to link counts, also speeds, densities/occupancies and/or travel times are extensively applied (e.g. Balakrishna et al., 2007; Ashok and Ben-Akiva, 2000; Tavana, 2001). Cipriani et al. (2013) presented a preliminary analysis on the impact of different kinds of information to the estimation of dynamic OD flows, underlining the importance of using a sample of path travel time measurements, together with measures of speeds and flows on link sections. Obviously, new technologies for probe vehicle re-identification and tracking (e.g. automated vehicle identification, AVI, and automated vehicle location, AVL, systems) represent another data source of growing importance. Dixon and Rilett (2002) estimated OD flows by using OD split proportions provided by AVI data collection points. Antoniou et al. (2006) estimated dynamic OD flows by incorporating direct measurements of OD flows from AVI systems or probe vehicles. Since the direct observation of OD flows requires the location of AVI stations close to origin and destination points, Antoniou et al. (2006) relax this assumption to formulate indirect observation equation by extracting sub-paths from several connected AVI detector locations. Further, Sohn and Kim (2008) used cell-based trajectories as probe phones to estimate the average path choice proportions. Barceló et al. (2010) proposed an efficient on-line approach based on linear Kalman filtering for linear structures, subsequently extended to general networks (Barceló et al., 2013), coping with congestion effects by the use of travel times provided by new ICT technologies. Integration of route flow data into the dynamic OD demand estimation problem was tackled by researchers in terms of either turning fractions (e.g. Van Der Zijpp and Romph, 1997; Mishalani et al., 2002) and/or route flows (e.g. Sun and Feng, 2011; Asakura et al., 2000). In general, a key problem in using AVI data for OD flow estimation is the knowledge of the sampling rate, i.e. either market penetration rates or identification rates. Estimation of sampling rates, however, is a difficult problem in its own right, as these rates are essentially time-dependent and location-dependent random variables. Moreover, the inclusion of sample rates in the OD demand estimation problem could dramatically increase the number of unknowns and thus impact on the reliability of OD flow estimates. To circumvent primary difficulties associated with estimating sample rates, Zhou and Mahmassani (2007) developed an OD demand estimation framework using partial AVI data.

Online formulations. A specific research task in dynamic OD estimation research refers to online

---

[3] Recent extensions, such as W-SPSA (Lu et al., 2015; Antoniou et al., 2015; Cantelmo et al., 2014b), and c-SPSA (Tympakianaki et al., 2015) aim to overcome this particular limitation.

applications, wherein OD flow estimates for recent time slices together with predictions for future time slices should be provided, recursively and promptly. Online estimation was first studied by Okutani and Stephanedes (1984) and subsequently generalized by Cremer and Keller (1987), Nihan and Davis (1987 and 1989), Ashok and Ben-Akiva (1993 and 2000) and Ashok (1996), who acknowledged the importance of structural information in OD flows. Computational issues in online within-day OD estimation in large networks represent also a main issue, firstly addressed by Bierlaire and Crittin (2004) and more recently by Antoniou et al. (2009).

Importantly, notwithstanding significant methodological contributions to the state-of-the-art, most of the aforementioned works do not provide decisive experimental evidence of their advantages over existing methods. As mentioned in the introduction, this is primarily due to the unobservability of OD flows –except in closed-form highway networks– and thus it is very difficult to detect the quality of OD estimation/updating procedures based on real datasets. In fact, matching closely the observed measurements used for OD estimation/updating is not in general indication of a good quality of OD estimation/updating itself.

Therefore, the main motivation of this paper is to provide a common evaluation and benchmarking framework of methodologies and algorithms for OD flows estimation/updating, under equal and standardized conditions, based on a proper methodology pivoted on synthetic test beds, built upon real networks. For a description of synthetic test bed approaches in case of traffic simulation model parameters calibration or OD flows estimation, see respectively Ciuffo and Punzo (2014) and Marzano et al. (2009). Other applications of such approach to the field of traffic simulation can be found, amongst others, in Zhou and Mahmassani (2006), Ciuffo et al. (2008), Ossen and Hoogendoorn (2008), Vaze et al. (2009), Gao et al. (2010), Punzo et al. (2012) and Zheng and Van Zuylen (2013).

3. Benchmarking platform: conceptual framework and methodological approach

This section introduces the main proposition of the paper, that is a common evaluation and benchmarking framework for OD flows estimation/updating methodologies and algorithms, whose main objective is to provide a test bed in which a number of algorithms can be implemented and tested under equal and standardized conditions. Specifically, a key requirement for evaluating and comparing various OD estimation/updating algorithms is to test their performances under a range of different conditions and scenarios (e.g. network sizes, demand and thus congestion levels, surveillance system characteristics, including the number and location of sensors and the type of detected information, as well as its quality, and so on) and to ensure that these conditions are consistent across experiments. The proposed design of the benchmarking framework – encompassing three main components, conceptually illustrated in Figure 1 – is generic and allows comparing considerably different approaches in a variety of settings and conditions, and is also consistent with diverse implementation platforms.

The first step deals with the setup of the "true" case study, consistent with the forward problem in sensitivity analysis (recalled in Section 1): basically, a true test bed is defined in terms of OD flows, network characteristics, and any other relevant parameters related to demand and supply. Consistently, a true supply/demand interaction procedure leads to calculation of true link flows and traffic flow characteristics. In this respect, a wide range of possible traffic flow characteristics – in turn resulting into input measurements for the OD estimation/updating algorithm – can be calculated. By way of example, the practical case study that will be presented in Sections 4 to 6 is also based on using point-to-point travel time measurements and data collected using Bluetooth antennas. Therefore, this approach allows calculating all concerned variables and enables a variety of experiments and applications.
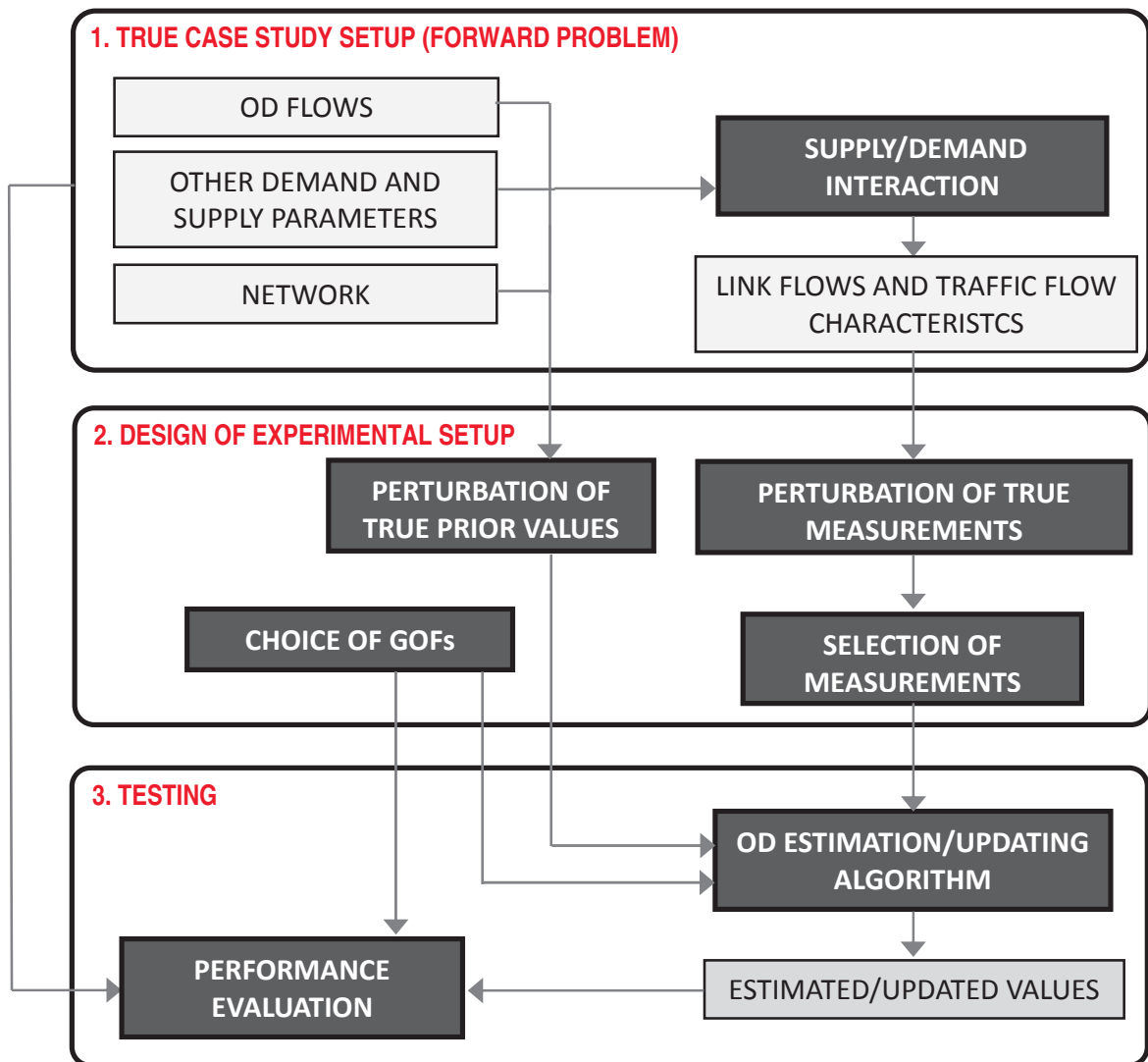
6

Figure 1 – Illustrative flowchart of the proposed methodological benchmarking framework

The second step deals with the design of the experimental setup, which consists mainly of *ad hoc* perturbations of possibly all true experimental values, i.e. OD flows, network characteristics, other demand/supply data and traffic flow measurements. This allows primarily mimicking real situations with uncertainties and/or biases in prior available estimates. In addition, a wide range of assumptions on selection of measurements can be adopted as well, so as to mimic the real and usual condition of collecting traffic flow measurements only in a few locations in the network[4]. Interestingly, both the initial true values in step one and the magnitude of the perturbations in step two can be combined, so as to cover diverse possible real conditions and test a heterogeneous and comprehensive range of situations. The experimental setup includes also the selection of appropriate indicators and goodness-of-fit (GOF) measures, allowing elucidating diverse aspects of the model and providing a more complete assessment of its performance. Indeed, OD flow estimation/updating algorithms are usually

---

[4] For details on the network sensors location problem – including the issues of detector positioning, qualitative appreciation of added detectors, and for quantifying the amount of information that a detector adds – the reader could refer to Yang and Zhou (1998), Simonelli et al. (2012) and Viti et al. (2014), whose algorithms and methodologies can be applied in order to select measurement locations for the benchmarking framework.

formulated as an optimization problem, whose objective function exhibits a functional form expressed via a GOF (or a combination of GOFs), as will be clarified in Section 4.

The third step focuses on testing OD estimation/updating algorithms under analysis. Specifically, tested methods/algorithms are fed with perturbed setup values, calculated in the second step of the procedure, and the resulting estimated/updated values are then compared with the true underlying values defined in the first step of the procedure, based on the selected set of GOFs.

Importantly, executing the OD estimation/updating algorithm implies running in turn the supply/demand interaction, in order to ensure mutual consistency between OD flows and traffic flow measurements during and at the end of the estimation/updating procedure. In other words, a dynamic communication between the OD demand estimation methods and the traffic simulator is necessary in order to execute a traffic simulation run within the OD demand estimation algorithm, with a degree of complexity and with specific implementation issues usually dependent upon the specific tested algorithm and the underlying hypotheses (e.g. congested network, type of dynamic traffic assignment, analytic solution vs. simulation, and so on). In addition, a common situation is that an OD estimation/updating algorithm runs on a different platform/tool with respect to the supply/demand interaction, thus leading to additional implementation issues, as it will be shown in the practical case study described in Section 5.

Interestingly, the proposed benchmarking framework is able to accommodate a wide range of applications, so as to check the performance of the tested dynamic OD flows estimation methods with respect to theoretical, practical or computational challenges known in the literature (see Section 2 also for a recap of key issues), for instance:

- *Impacts of biased prior OD flows.* Several prior OD matrices can be tested with different degree of errors and biases, mimicking the common approach of deriving prior dynamic OD flow estimates by splitting a static prior OD matrix using an exogenously superimposed time profile. Thus, biases can occur, first of all, as a consequence of errors in estimation of the total number of trips and spatial distributions resulting from static demand modelling. Moreover, prior static OD matrices are also adjusted with traffic counts, possibly resulting in underestimation of demand when in congested conditions, because of the misinterpretation of low flow values due to traffic congestion (i.e. erroneously considered as resulting from low demand values). Also the dynamic splitting of the static prior OD matrix is source of errors, because of difficulties in observing reliable dynamic OD time profiles;

- *Influence of other demand parameters.* Diverse route choice options can be implemented, thus testing the effects of possible errors in route choice modelling, because route choice models are normally calibrated based on either aggregated data (e.g. link counts) and/or disaggregate path observations provided by monitored vehicles, leading in both cases to possible estimation errors;

- *Effects of combined types of measurements.* A variety of traffic data types can be adopted as inputs for OD flows estimation, and the proposed framework allows to analyze the combined effects of their fusion, in terms of measurement distortions/errors, collinearity, and so on;

- *Inherent properties of the OD flows estimation problem.* As recalled in Section 2, dynamic OD flow estimation is a highly non-linear, non-convex problem. The shape of its objective function clearly depends on diverse factors, e.g. the type of estimator adopted, the boundary conditions, and so on. In this respect, the benchmarking framework allows investigating how the structure of the objective function – and therefore the performance of the solution algorithm – changes with respect to these input factors, with important consequences in assessing the practical tractability of OD flows estimation.

- *Impacts of traffic simulation/assignment.* Performance of OD estimation methods depends

8

highly upon the structure and the properties of the supply/demand interaction – that is of the assignment map, defined either explicitly in analytical method or implicitly through simulation – and also upon the congestion level of the network (Frederix et al, 2013; Flötteröd and Bierlaire 2009). In this respect, the proposed benchmarking framework enables testing, on one side, a given OD estimation method with respect to different congestion levels through a proper definition of the true case study setup and of the experimental setup, and, on the other side, to compare different OD estimation methods under the same congestion level. A specific instance of this feature is proposed in Section 5.2.1.

In general, the conceptual benchmarking framework presented in this Section should be operationalized in a benchmarking platform, capable to perform comparisons between different OD estimation/updating methods in a remarkably wide range of case studies and experimental setups. Therefore, the paper presents a practical instance of implementation of a benchmarking platform based on the proposed framework, discussing the contrasted OD estimation/updating methods (Section 4), the implementation of the platform (Section 5), and the main testing results (Section 6).

4. Illustration of tested OD flow estimation/updating methods

This section reports on the OD estimation/updating algorithms tested using the benchmarking framework described in Section 3 and practically implemented in the platform illustrated in Section 5. Specifically, Section 4.1 presents a general formulation of the tested (dynamic) OD estimation/updating methods, whilst Section 4.2 provides details on the specifications and on the solution algorithms.

4.1 General formulation of the tested dynamic OD flow estimation methods

This subsection presents a fairly general formulation of the dynamic OD flow estimation problem. First, definitions that will be used further in formulation of selected dynamic OD demand methods are introduced. The modelling horizon has T time steps, and is divided in time intervals t, with $t = 1, 2, ..., T$. The traffic demand between origins and destinations is stored in the OD vector $\boldsymbol{x} = \{x_i | i \in I\}$ wherein I is the set of all OD pairs; the corresponding link flows are stored in the vector $\boldsymbol{y} = \{y_l | l \in L\}$, being L the set of links in the network. The historical or prior or seed OD vector $\tilde{\boldsymbol{x}} = \{x_i | i \in I\}$ is a vector that needs to be updated. Normally, a wide range of traffic flow measurements are available, ranging from traditional link flow measurements (e.g., speed, density, occupancy) to innovative data collected by a variety of sensors (e.g. travel times collected from Bluetooth sensors, partial OD flows and travel times from GPS traces, point-to-point various traffic flow measurements between pairs of coupled sensors). Thus, let $\hat{L} \subseteq L$ the set of monitored links, i.e. with available link traffic counts $\hat{\boldsymbol{y}} = \{\hat{y}_l | l \in \hat{L}\}$ and link-based traffic flow measurements denoted as $\hat{\boldsymbol{c}} = \{\hat{c}_l | l \in \hat{L}\}$, and let $\hat{P} \subseteq P$ be the subset of pairs of nodes/points in the network with available additional traffic flow measurements[5] denoted as $\hat{\boldsymbol{z}} = \{\hat{z} | p \in \hat{P}\}$.

From a fairly general standpoint, the dynamic OD estimation problem, combining observed link flow data and traffic flow measurements, can be expressed as:

---

[5] Notice that **z** includes possibly node based, point-to-point based and (partial) OD based measurements. Notably, when travel time measurement are included in the estimation method, Barceló et al. (2012) show that a detection layout based on node covering instead of the usual link covering, provides richer and more sensitive data, as this layout captures higher flows and allow to exploit the identification of partial paths between detectors, as later in this paper will be shown in the case of Vitoria`s network. In-depth discussions/analyses on path flow estimators have been proposed by various authors, and a basic reference could be found in Bell (1997).

$$x^* = \text{argmin}_{x \in S_x} \; \varphi(x, \widetilde{x}, \widehat{y}, \widehat{c}, \widehat{z}) \qquad \text{or} \qquad x^* = \text{argmax}_{x \in S_x} \; \psi(x, \widetilde{x}, \widehat{y}, \widehat{c}, \widehat{z}) \qquad (1)$$

wherein $x^* = [x_1^*, ..., x_T^*]$ is the unknown OD demand vector[6] for time intervals $t \in 1, 2, ..., T$, $\varphi(\cdot)$ and $\psi(\cdot)$ general functional forms and S_x the feasibility set of the problem, embedding non-negativity constraints for OD flows and also other possible constraints (e.g. total production/attraction, geographical correlations, trip length distributions, and so on).

The dynamic OD estimation methods considered in this benchmark study is a particular, yet very common, instance of the generic formulation (1), given by:

$$x^* = \text{argmin}_{x \geq 0}[f_1(x, \widetilde{x}) + f_2(y(x), \widehat{y}) + f_3(c(x), \widehat{c}) + f_4(z(x), \widehat{z})] \qquad (2)$$

wherein the four functions f_1, f_2, f_3 and f_4 express the distance between measured and modelled values as a function of different goodness-of-fit measures. An intuitive interpretation of the problem given in equation is that it searches the vector x^* that is closest to the a priori estimate $\widetilde{x}$, and, once assigned to the network, produces the traffic data, $y(x)$, $c(x)$ and $z(x)$ closest to their observed values. At each iteration step or time interval t, $y(x)$, $c(x)$ and $z(x)$ could be extracted from inputs of the traffic simulator and calculated based on running supply/demand interactions: this aspect will be described in detail in Section 5.3, and further details can be found in Yang and Bell (1998), Tavana and Mahmassani (2001), Flötteröd and Bierlaire (2009) and Frederix et al. (2011). The set of constraints depends on application of the problem as well as the desired level of accuracy, and it can include non-negativity constraints, initial condition constraints, lower and upper bound constraints to avoid infeasible solutions and restrict search space, and so on. Also, traveller's route choice or traffic assignment rules are often obtained by optimizing an objective function, which can be explicitly included in the set of constraints. This formulation results in a bi-level optimization problem, where the upper level is a non-linear non-differentiable optimization problem, whose objective function is usually defined in terms of GOFs, and the lower level is an equilibrium assignment problem. In some of the approaches tested in this paper (see Section 4), the optimization problem has been solved with an *ad hoc* version of optimization-simulation approach based on SPSA, which also evaluates the function by a mesoscopic simulation based Dynamic User Equilibrium (DUE).

A key aspect for the practical implementation of the generic formulation (2) refers to the choice of the functional form of the component functions f_1 … f_4. Drawing upon Cascetta (1984), a natural and straightforward choice is to assume a least square approach. However, since different traffic measurements contain very diverse values in terms of order of magnitude, a normalized mean error (NME) specification can be applied as well. Specifically, the choice of a particular NME performance measure comes from the analysis of the trend of each goodness-of-fit measure between the considered prior OD matrix and the true OD matrix, obtaining that the NME best represents the descendent direction through the desired solution. This effect can be explained by the fact that the NME is an algebraic rather than an absolute error measure, capable to balance possible instabilities of link quantities (pertaining to the assignment phase), and to capture the improvement that can be detected at network level, when the estimated matrix moves along the descent direction.

---

[6] Importantly, working with OD flows is not a strict requirement. For instance, there might be situations wherein it is more effective to separate generation profiles and distribution shares – such as in the quasi-dynamic approach by Cascetta et al. (2013) or the method proposed by Cipriani et al. (2010) (see section 4.2.2) – or to work directly with OD path flows, such as in Barceló et al. (2013) (see section 4.2.5).

4.2 Description of the selected dynamic OD estimation methods for the benchmark case study

Consistent with the general formulation reported in Section 4.1, Table 1 summarizes the main properties of the five selected dynamic OD estimation methods for the benchmarking case study presented in the paper, showing their diversity of input data, objective function and solution algorithm. A detailed description for each of them is reported in the following subsections.

Table 1 - Properties of selected dynamic OD demand estimation methods

Method	Input data				Objective function **	Solution algorithm
	prior OD	*link flow*	*link density*	*travel times*		
Method 1	+	+			LS	LSQR
Method 2	+ *	+	+		LS + NME	SPSA AD-PI
Method 3	+	+	+	+	NME	enhanced SPSA AD-PI
Method 4		+	+	+	NME	SPSA CG-TR
Method 5	+	+		+	LS (state - space)	Kalman filter

* two different versions of Method 2 are implemented: the former, labelled Method 2-TD, encompasses trip distribution shares as prior OD information; the latter, labelled Method 2-OD, encompasses OD flows as prior information,
** LS=least square, NME=normalized mean error.

4.2.1 Method 1: the LSQR method

The first tested method is the least square formulation of the dynamic OD demand estimation model (2) used by Bierlaire and Crittin (2004), given by considering only the two functions f_1 and f_2. Furthermore, exploiting Ashok and Ben-Akiva (1993), deviation of OD flows from their historical estimates are assumed as state variables and deviations of link flows from historical counts as traffic measurements. The least square problem is solved through the LSQR algorithm, exhibiting computational efficiency for very large networks, analytically equivalent to a conjugate gradient method, based on bi-diagonalization procedures (Paige and Saunders, 1982). A key property of the LSQR algorithm is that the assignment matrix – very sparse in large scale networks, such as in the practical case study implemented in Section 5 – does not need to be explicitly constructed or stored. For a more detailed explanation of this algorithm the reader may refer to Bierlaire and Crittin (2004). Summarizing, Method 1 assumes the following:
- Input data: prior OD demand and link flow data;
- Solution algorithm: LSQR algorithm.

4.2.2 Method 2: the SPSA AD-PI method

The second tested method is formulated by adding traffic flow densities as measurements in the problem (2), with the aim of providing additional information on the traffic regime. Drawing upon the several variants of the SPSA algorithm proposed by Cipriani et al. (2010, 2011) –where the off-line dynamic OD demand estimation problem is formulated as a bi-level nonlinear optimization program without introducing assignment matrices– the proposed solution approach is a modification of the gradient-based path-search optimization method (SPSA) coupled with Asymmetric Design (AD) for gradient computation and Polynomial Interpolation (PI) of the objective function for the linear optimization. SPSA AD-PI permits to reduce the computational efforts with respect to the usual gradient-based methods: for a more detailed explanation of this algorithm the reader may refer to

11

Cipriani et al. (2011).

Preliminary results using SPSA AD-PI show an interesting trade-off in the choice of the type of prior OD information in the problem (2). Indeed, on the one hand, including the distance from the prior OD matrix values in the function f_1 of Equation (2) may bind too strongly the search of the solution within a neighbourhood of the prior OD matrix itself. On the other hand, removing prior demand (i.e. neglecting f_1) would give importance only to the remaining terms of the problem (2), i.e. leading to satisfactory results only with respect to matching both link measurements and possibly to the total amount of generated trips by each origin, whilst OD flow estimates are usually poor. Thus, it is worth exploring inclusion of "structural properties" of the OD matrix as prior estimates, e.g. production/attraction by zones, geographical correlations, temporal smoothness, trip length distributions, and so on. Therefore, Method 2 will be specifically applied in the following by assuming trip distribution as prior OD values, thus labelled as "Method 2-TD". In addition, in order to provide a direct comparison with the subsequent Method 3 (Section 4.2.3), Method 2 has been applied also considering prior OD demand instead of prior OD trip distributions, thus labelled as "Method 2-OD".

Summarizing, Methods 2 assumes:
- Input data: prior OD trip distributions ("Method 2-TD") or prior OD flows ("Method 2-OD"), link flow data and density data;
- Solution algorithm: SPSA AD-PI algorithm.

4.2.3 Method 3: the enhanced SPSA AD-PI method

The previous Method 2 (Section 4.2.2) can be enriched assuming the availability of point-to-point travel times detected using Bluetooth sensors along main paths/corridors in the network. In particular, Barceló et al. (2013) has proved that a suitable Bluetooth sensor layout allows identification of the paths between sensors and therefore measurement of the associated travel times. Thus, the corresponding dynamic OD estimation problem is defined by normalized mean error functions f_1, f_2, f_3 and f_4 in equation (2) and is solved by the modified SPSA AD-PI approach used in Method 2 and already described in Section 4.2.2.

Summarizing, the main properties of Method 3 are given as follows:
- Input data: prior OD demand, link flow data, density data and point-to-point (i.e. between Bluetooth sensors) travel time data;
- Solution approach: SPSA AD-PI algorithm.

4.2.4 Method 4: the SPSA CG-TR method

As mentioned in Section 4.2.2, the computational experience showed that the presence of prior OD flow information in the general problem (2) has a twofold negative influence, because it prevents exploring OD solutions significantly different from the (biased) prior matrix and at the same time it implies high computational burden. Based on these tenets, Bullejos et al. (2014) proposed a specification of the problem (2) excluding prior information on OD flow data. Consistently, the dynamic OD estimation problem is defined by normalized mean error performance functions f_2, f_3 and f_4 in expression (2).

Furthermore, this case study focuses on exploring solution approaches capable to reduce the computational burden. For this aim, Bullejos et al. (2014) proposed a modified SPSA approach with a twofold enhancement: i.) to include a conjugate gradient strategy in the computation of the approximated average gradient, as suggested in Cantelmo et al. (2014b), because it is known that conjugate directions permit to reach the solution faster than using the basic gradient method; ii.) to use

12

of a trust region scheme as in Osorio and Bierlaire (2010), whose main idea is to set implicitly at each iteration a neighbourhood around the current solution: indeed, avoiding matrices outside the trust region is essential to reduce the computational burden. For more detail explanation of this algorithm the reader may refer to Bullejos et al. (2014). Summarizing, Method 4 is based on the following:

- Input data: link flow data, occupancy data and travel time data;
- Solution approach: SPSA CG-TR algorithm.

4.2.5 Method 5: Kalman filter method

Barceló et al. (2013) adopted a linear state-space formulation for real-time estimation and short-term prediction of OD matrices, assuming the availability of travel times between Bluetooth sensors along the main paths connecting a subset of pairs of Bluetooth sensors in the network. Thus, Method 5 exploits the travel times and counts collected, respectively, by tracking Bluetooth equipped vehicles and conventional loop detectors. Consistently, the dynamic OD estimation problem is defined as a state-space model solved by a linear Kalman filter algorithm with no assignment matrices, as proposed by Ashok (1996) and Ashok and Ben-Akiva (2000). Keeping a linear relationship between state-variables and measurements is computationally advantageous and reduces the number of state variables in the Kalman filter formulation. Further details on the mathematical formulation and on the corresponding solution algorithm are reported in Barceló et al. (2013).

The main properties of Method 5 are:

- Input data: prior OD flows, link flow data from loop detectors, and traffic counts at Bluetooth sensors and travel times data between entry Bluetooth sensors and inner Bluetooth sensors;
- Solution approach: a linear Kalman filter, whose state variables are defined as OD path flows in the set of the most likely used OD paths[7].

5. Case study: implementation of the benchmark platform and experimental setup

The practical implementation of a benchmarking platform, capable to contrast the performances of the OD estimation/updating algorithms described in Section 4, should follow the methodological tasks identified in the general framework illustrated in Section 3. This section follows exactly the three main steps depicted in Figure 1, that is the true case study setup (Section 5.1), the design of the experimental setup (Section 5.2) and illustrates how the interplay between the supply/demand interaction and the OD estimation/updating algorithm has been developed for testing purposes (Section 5.3).

This setup is simply for exposition of the platform and in no way an integral (or "hard-coded") part of it. Instead, the user applying the platform for a specific application would be free to use any type of traffic simulator, network and traffic measurement that is suitable for that application. Furthermore, the demand structure adopted by the user could be perturbed in a different way, to e.g. further obscure the structure of the initial demand, or perhaps make the demand structure perturbations dependent on the traffic state.

5.1 True case study setup

First, a specific supply/demand interaction approach should be defined in terms both of theoretical

---

[7] Usual approaches require a total path enumeration resulting impractical for large network, but a detailed analysis of equilibrium solutions shows that only a reduced number of paths are significantly used and that most of them show a stability across time intervals in a DUE, therefore the reformulation of the problem in terms of these paths, what we call the most likely used paths, provides a formulation of reasonable size even for large networks.

assumptions (e.g. macro/meso/micro, equilibrium/network loading, and so on) and practical implementation (e.g. commercial software vs. in-house built functions). In this respect, without loss of generality, the mesoscopic Aimsun traffic simulator (TSS, 2013) has been chosen in this study. The choice of simulator is not an inherent part of the presented platform, and indeed only a factor in the case study. Any type of traffic simulators can be used, including open-source traffic simulators, whose open-source nature might actually simplify the process of developing the interfaces to the platform. In this respect, an important feature that any simulator has to exhibit for inclusion in the benchmarking platform is the numerical stability of its results, i.e. small OD flow perturbations should lead to small consistent variations in simulation outputs: this is indeed a necessary condition to ensure fair comparison amongst results using different methods/algorithms.

In terms of test beds, three different networks have been modelled:

- A small network used primarily for debugging and verification purposes;
- A medium-size network from Eixample, the central business district of Barcelona, consisting of 120 origin centroids, 1227 nodes (grouped in intersections), 130 destination centroids, 2800 links and 1469 OD pairs;
- A larger network from the city of Vitoria (Spain), depicted in Figure 2, including 57 centroids, 3249 OD pairs, 2884 nodes and a modelled network of about 600 km, corresponding to 5799 links. This network resembles a reasonable sized real-life network, with representative congestion levels and route choice dimension, as found in many large urban areas.

For the sake of brevity, the paper reports results related only to the Vitoria network; however, other case studies using this benchmarking platform are reported in Barceló et al. (2013), Bullejos et al. (2014), and Cantelmo et al. (2014b). Consistent with the true case study setup described in Section 3, a true model has been assumed as ground-truth together with a true OD matrix. The corresponding true link flows and all other concerned traffic flow characteristics have been derived from assignment of true OD matrix using the aforementioned mesoscopic Aimsun traffic simulator. Specifically, a Dynamic User Equilibrium (DUE) has been applied, based on a mesoscopic simulation for the network loading phase and on a Method of Successive Averages (MSA) for OD paths and OD path flows computation; convergence is defined upon a gap function, and results rely on a proper calibration of the model. Additionally, applying a DUE instead to a DTA has the advantage of smoothing out the inherent stochasticity of route choice through the iterative process.

The simulation period refers to the afternoon hourly peak period, divided into 15-minutes time intervals plus another additional warm-up time interval, leading to 5 time intervals overall. The trips for vehicles that departed at a given time interval between some of the OD pairs are not necessarily completed within the same time interval, due to congestion on network or to the distance between OD pairs, thus resulting in 4 lagged time intervals and very sparse assignment matrices. All these decisions are user-defined input parameters to the platform, so they would be determined by the user specifically for each application. Similarly, the complication of stochasticity (e.g. due to the stochastic route choice models in a specific traffic simulation algorithm) should be treated by the specific calibration algorithms (which are exogenous to the platform) that are being considered and is not a factor in the operation of the platform itself. Anyway, if a DUE is used instead of a DTA, these effects would also be smoothed out, as noticed above.

5.2 Design of experimental setup

The second main implementation task is to provide a comprehensive design of the experimental setup in terms of perturbation of true OD flows (Section 5.3.1), choice of traffic flow measurements and

locations with corresponding initial perturbations (Section 5.3.2), and choice of GOFs (Section 5.3.3).

5.2.1 Perturbation of true OD flows

In general, and consistent with the general formulation (2) of the tested OD estimation/updating algorithms, a prior OD flow estimate should be provided as input. Generally, the quality of such prior estimate, in terms of both demand level and patterns, is a key element affecting the performance of the estimation/updating algorithms, even more relevant for congested networks. Therefore, selecting proper perturbations of the true OD flows in order to derive the historical OD flow estimates to be provided as inputs for the OD estimation/updating algorithm is a crucial task. In this respect, three different prior OD flow scenarios have been chosen for the case study presented in this research, mimicking various sources of errors, specifically:

- **Low-demand scenario (D7):** this scenario addresses situations wherein, for instance, prior OD flow estimates result from out-of-date surveys. In detail, the low-demand scenario is defined as the 85% of the true OD demand level (Section 5.1), plus uniformly distributed random fluctuations over each OD pair and departure time interval in the range of $\pm$ 15% of the true value, that is:

$$\tilde{x}_{ij,t}^{LD} = x_{ij,t} \cdot \left[0.7 + 0.3\alpha_{ij,t}\right] \qquad \alpha_{ij,t} \sim U(0,1)$$

wherein the generic OD pair is indexed by *i,j* and *t* is the time interval;

- **Medium-demand scenario (D8):** in this scenario, the prior OD demand pattern corresponds for each OD pair to the 95% of the true OD demand level, plus an uniform random component in the range of $\pm$ 15%:

$$\tilde{x}_{ij,t}^{MD} = x_{ij,t} \cdot \left[0.8 + 0.3\alpha_{ij,t}\right] \qquad \alpha_{ij,t} \sim U(0,1)$$

- **High-demand scenario (D9):** this scenario addresses for instance situations where the prior OD demand reflects travel demand in peak-hours, when congestion occurs on network. The prior OD demand pattern is thus pivoted on 105% of the true OD demand level plus an uniform random component in the range of $\pm$ 15%:

$$\tilde{x}_{ij,t}^{HD} = x_{ij,t} \cdot \left[0.9 + 0.3\alpha_{ij,t}\right] \qquad \alpha_{ij,t} \sim U(0,1)$$

5.2.2 Selection of measurement locations and perturbation of true traffic flow measurements

In order to identify realistic scenarios in terms of measurement types and locations, two different sets of sensors have been identified in the Vitoria network:

- 389 standard loop detectors, located as depicted in the left-hand side of Figure 2, providing flows, speeds, occupancies, and so on, related to all vehicles at the detected loop;
- 50 Information and Communication Technologies (ICT) sensors, located as shown in the right-hand side of Figure 2, providing automatic signature identification for a subset of the vehicles, for instance through suitable antennas to catch Bluetooth devices in discovery mode. Notably, the Bluetooth sensors are deployed following a layout strategy to optimize the capture of equipped vehicles and provide effective travel time measurements between Bluetooth sensors.

The simulation of the true scenario leads to almost 90% of the trips collected twice at least in the

15

afternoon-peak demand scenarios, which account for 95% of the number of OD pairs and 86% of the most likely used paths identified in a DUE assignment with the true prior OD matrix. The procedure proposed by Barceló et al. (2012) returns the location for Bluetooth sensors that provide simulated travel times on these predefined and stored routes satisfying the former properties.

In terms of perturbation of the true link flows and other traffic flow measurement, the same approach adopted for OD flow perturbation (Section 5.2.1) can be followed in developing scenarios regarding the characteristics of the surveillance system, so as to mimic measurement errors in the real world. Without loss of generality, for the sake of simplicity, error-free measurements have been hypothesized in this paper.

Figure 2 – The Vitoria network, Basque Country, Spain: loop detector sensors layout (left) and ICT sensors layout (right)

5.2.3 Choice of GOFs

The final component of the benchmarking framework processes the outcomes of the OD demand estimations and calculates a series of performance indicators. The choice of performance indicators plays an important role in benchmark analysis of OD demand estimation methods under uncertain inputs. In this case study, we distinguish two levels of performance indicators:

- *Statistical performance indicators*: a number of statistical measures can be used to evaluate the overall performance of OD estimation methods, measuring the divergence between estimated outputs (e.g. OD flows, traffic counts, travel times) and their ground-truth values. In the context of this platform, the following goodness-of-fit measures have been implemented: RMSE, RMSNE, NRMSE, GEH1, MAE, MANE, NMAE, SE, U, ME, MNE, and NME. The reader may refer for details to Hollander and Liu (2008), Ciuffo and Punzo (2010), Antoniou et al. (2014) and Buisson et al. (2014);
- *Computational efficiency*: one of the research objectives is improvement in the computational efficiency of the dynamic OD flow estimation/updating problem for real time applications. To quantify the computational efficiency, the implemented platform calculates also the CPU computational time.

5.3 Practical implementation of the OD estimation/updating algorithms in the benchmarking platform

For the purposes of the benchmarking platform, it is crucial to design how the supply/demand interaction and the OD estimation/updating algorithm interact with each other in running the experiment (i.e. step 3 of the general framework in Figure 1). Indeed, the practical implementation of

the general benchmarking framework illustrated in Section 3 should take into account that the optimization of the problem (2) requires sequential calls and integrated input/output exchange between the supply/demand interaction and the OD estimation/updating algorithm[8]. For instance, calculation of the descent direction at a given iteration i of the OD estimation/updating algorithm requires normally several evaluations of the objective function, which in turn implies running the supply/demand interaction in order to calculate model-based link flows and traffic flow characteristics to compare with observed measurements in the terms f_2 to f_4 of expression (2). The practical declination of this interplay requires also defining the specific characteristics of the supply/demand interaction: as mentioned at the beginning of Section 5, for the purposes of the paper, and without loss of generality, a mesoscopic traffic simulation model has been chosen.

In addition, it should be considered that normally the OD estimation/updating algorithm and the traffic simulation model run on different environments, e.g. an in-house built programming code for the former and a commercial software package for the latter. Therefore, illustratively, from a fairly general standpoint, the conceptual diagram depicting the interplay between OD estimation/updating algorithm and traffic simulator is reported in Figure 3.
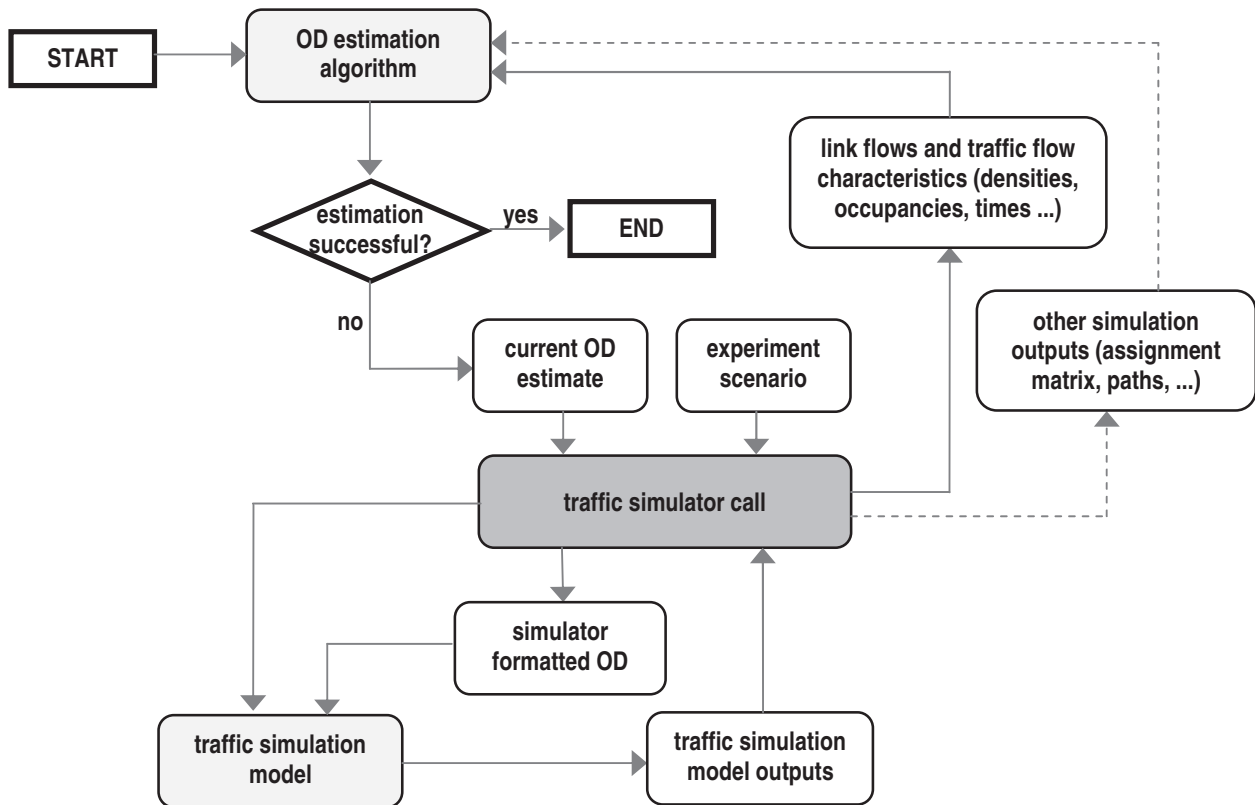

Figure 3 – Conceptual diagram of interplay between OD estimation/updating algorithm and traffic simulator within the benchmarking platform

Notably, within the OD estimation main function, whenever a simulation run is needed the Traffic Simulator Call Function is initiated. This function converts the demand to be simulated to the Simulator

---

[8] Rigorously, this does not apply to Method 5 described in Section 4, which evaluates the GoF at the end of the simulation period and does not requires a DUE simulation call each iteration, just counts and travel times between Bluetooth sensors.

17

Format, generates all the inputs necessary to properly run the traffic simulation models and finally calls and executes the Traffic Simulator with these inputs. From this logical diagram the main requirements of a Traffic Simulator to be used in this process become evident. The exchange of information between the proposed OD estimation procedure and the Simulator requires that it can be called in run time from the estimation algorithm as a subprogram. It also requires that the simulation output provides the inputs required by the OD estimation process and the GOFs calculations (e.g. traffic data, link flows, speeds, occupancies, OD travel times or travel times along the paths between OD pairs of ICT sensors), depending on the requirements of the specific method to test. In addition, as represented by the dashed lines in Figure 3, specific methods[9] to be tested could also require further additional traffic modelling data, e.g. estimates of the assignment matrices and/or set of chosen/feasible paths for each OD pair.

From a practical standpoint, the benchmarking platform adopted in this paper is based on a combination of Aimsun as traffic simulator and Matlab (Matlab, 2013) as programming language for running the OD estimation/updating algorithms, with the calls between the two defined through Python: details of the implementation are given in the Appendix to the paper.

6. Results and discussion

This section provides some[10] illustrative results obtained from the use of the benchmark platform described in the previous sections, with the main goal of demonstrating its potential application to various algorithms and using different traffic flow measurements. Specifically, the following applications of the methods illustrated in Section 4 are presented:

- Method 1 applied to demand scenarios D7 and D9 (Section 5.1), showing the effect on the estimated demand matrices and predicted link flows;
- Method 2-TD (i.e. the method illustrated in Section 4.2.2 with OD trip distributions as prior information) applied to the demand scenario D7;
- Method 2-OD (i.e. the method illustrated in Section 4.2.2 with OD flows as prior information) compared with Method 3 using the demand scenario D7;
- Methods 3 and 4 applied to the demand scenarios D7 and D9, showing the progress of the algorithms along the iterations when point-to-point travel times are considered in both methods;
- Method 5, demonstrating the potential effect of using new sensing capabilities, as travel times between the entry points to the system and the ICT internal sensors, beyond point counters for an on-line application, applied to the scenario D9.

Figure 4 presents the performance of Method 1 using prior OD demand scenarios D7 and D9 and error-free link flow data. As discussed above, both demand scenarios are biased: on average, the D7 seed matrix is biased downwards by 15%, and the D9 seed matrix is biased upwards by 5%. The figure shows the fit of the estimated OD flows, both to the true OD demand and to the link flows: as expected, a slightly better fit is obtained when the average bias in the seed OD matrix compared to the true demand is smaller.

Scenario D7 was also used with Method 2-TD, which includes several parameters (step size, number of replications and use of information from previous iterations in the gradient calculations), set based on the results of a sensitivity analysis conducted by Cantelmo et al. (2014b). Figure 5a presents

---

[9] This is for instance the case of Methods 3 and 4, described respectively in Sections 4.2.3 and 4.2.4, for which the traffic simulator used must also be able of performing a DUE assignment and of exporting some internal information of the network structure and additional link data.

[10] As mentioned in Section 5, the presented platform is being used by several of the participants in active research, that has already led to the following publications, where the interested reader can find case studies and research aspects not covered here for the sake of brevity: Barceló et al. (2013), Bullejos et al. (2014), Cantelmo et al. (2014b), Barceló et al. (2014).

results in terms of estimated OD flows for Method 2-TD, that are comparable with those reported for Method 1 at convergence; instead Figure 5b indicates a worse fit for link flows with Method 2.
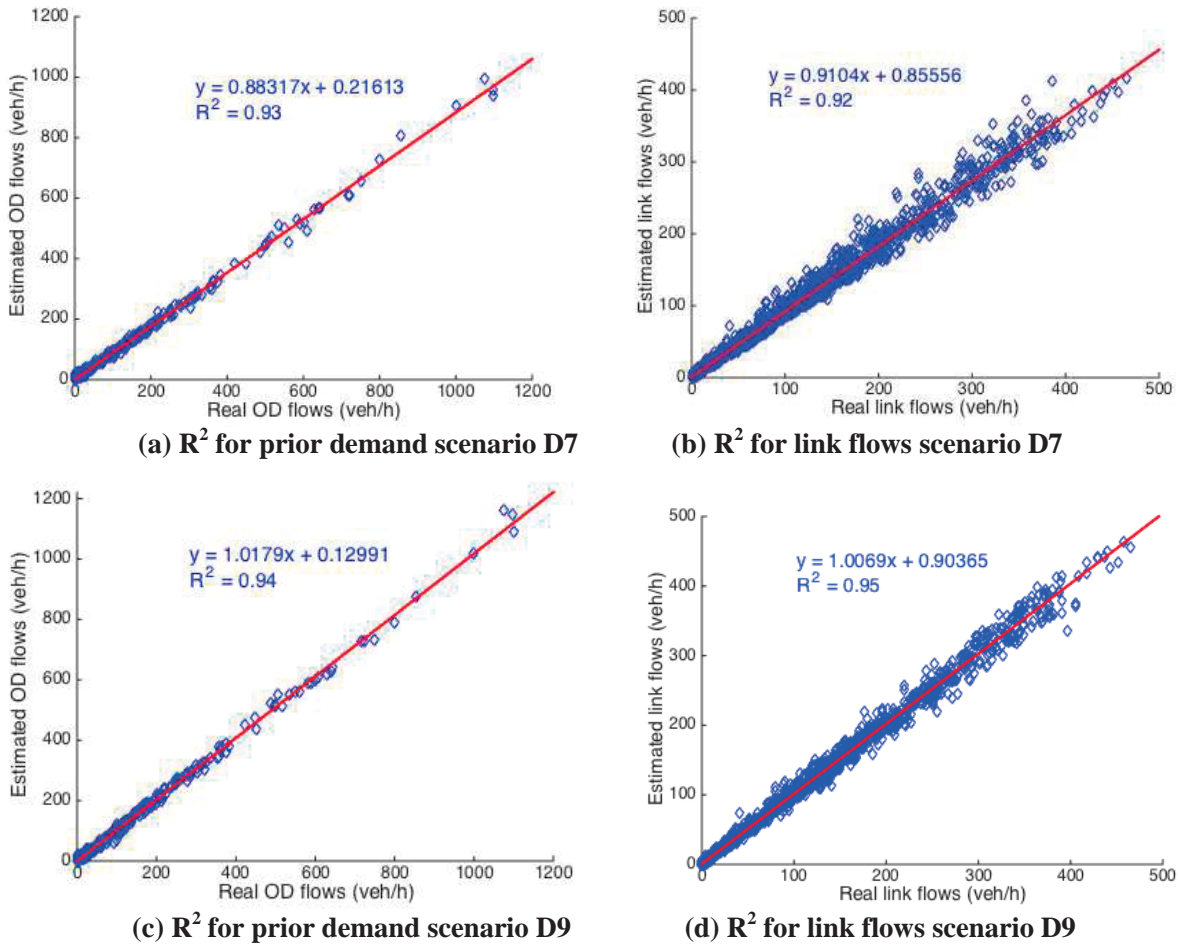

(a) R^2 for prior demand scenario D7

(b) R^2 for link flows scenario D7

(c) R^2 for prior demand scenario D9

(d) R^2 for link flows scenario D9

Figure 4 - Method 1 results: R^2 for prior demand and for link flows - scenarios D7 (top) and D9 (bottom)

(a) R^2 for prior demand scenario D7

(b) R^2 for link flows scenario D7
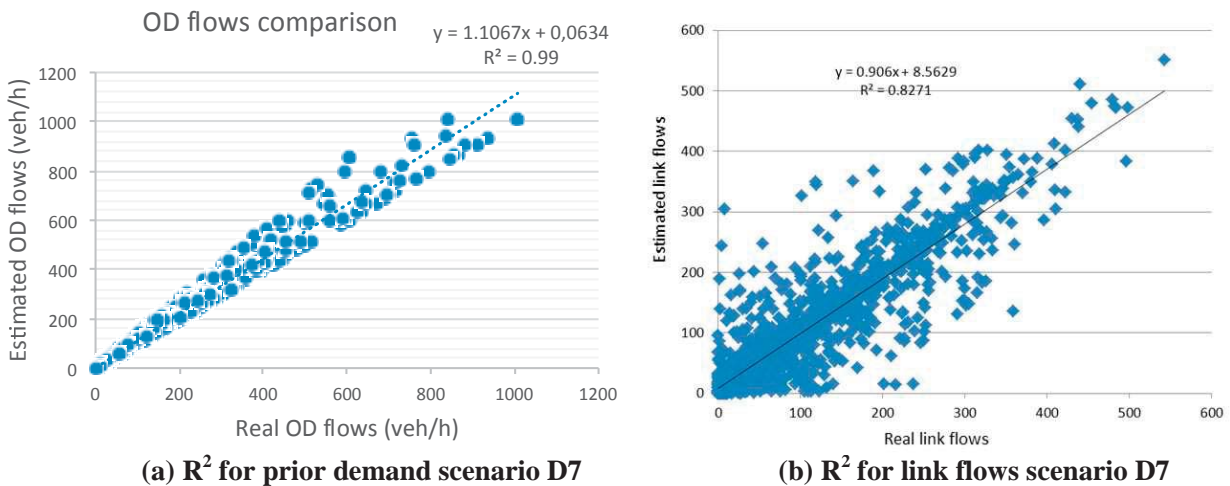
Figure 5 - Method 2-TD results: R^2 for prior demand and for link flows - scenario D7

Method 3 extends Method 2 by the use of additional travel time information in the estimation process. Results in Figure 6 refer to comparisons with Method 2-OD (i.e. including OD flows as prior information) in order to make it comparable with Method 3. Specifically, Figure 6a shows how the flow term in the objective function evolves through iterations for both Method 3 (considering all demand scenarios D7, D8 and D9) and Method 2-OD (considering the demand scenario D7). In addition, Figure 6b compares the solutions of Method 3 and Method 2-OD at convergence for the demand scenario D7, evidencing the benefit of including a travel time term in the objective function.
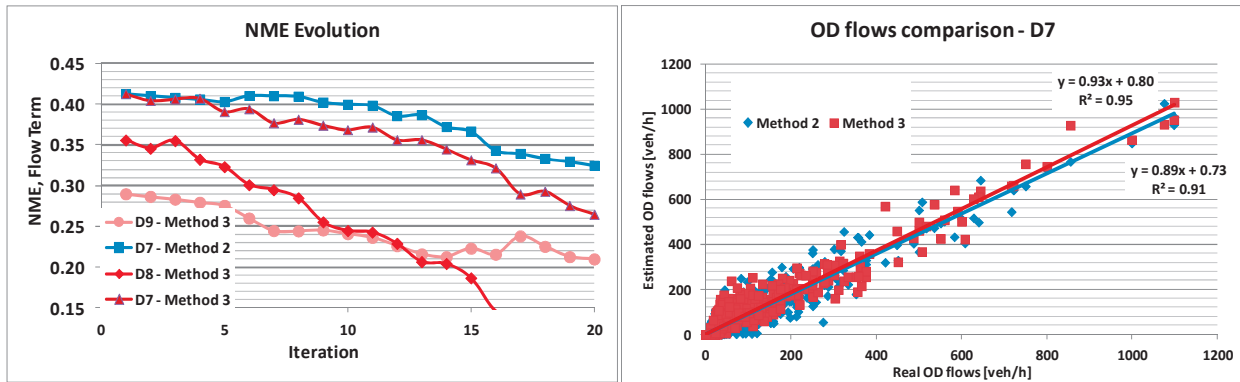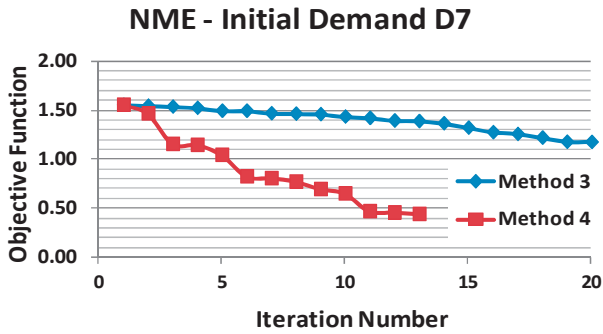

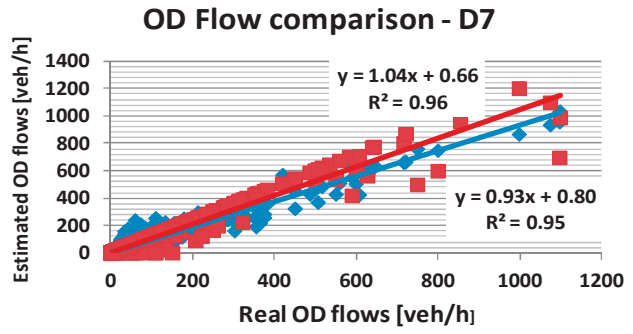**(a) evolution of function f_1 for Method 2-OD and Method 3**

(b) true vs. estimated OD flows for Method 2-OD and Method 3

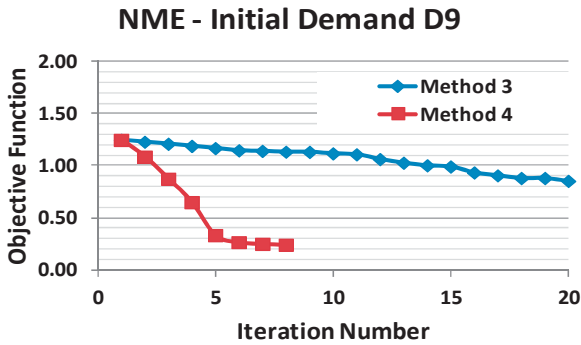**Figure 6 – Results for Method 2-OD and Method 3**

Method 4 is another variant of the basic algorithm underlying Methods 2 and 3. Its main characteristics are: i. it does not use a prior OD matrix, and ii. it exhibits a major algorithmic enhancement, consisting of approximating a stochastic conjugate gradient descent direction combined with a trust region scheme. In this respect, Figure 7 reports the results of a comparison of Method 3 and Method 4 under equal experimental conditions. Specifically, Figure 7a and 7c report the change of the objective function through iterations, respectively for demand scenarios D7 and D9. Similarly, Figure 7b and Figure 7d compare Methods 3 and 4 based on the R^2 goodness-of-fit of the OD flows solution at convergence, again respectively for demand scenarios D7 and D9. In general, Method 4 shows a faster convergence than Method 3 in terms of reduction in the value of the objective function. However, in both cases, the quality of the solution in terms of coefficient of determination is similar for both methods, as shown by comparing Figures 7b and 7d, with estimation accuracy slightly magnified when the prior demand is close to the true one. Further numerical details can be found in Bullejos et al. (2014). Consequently, as shown in Figures 7a and 7c, Method 4 significantly reduces the number of iterations.
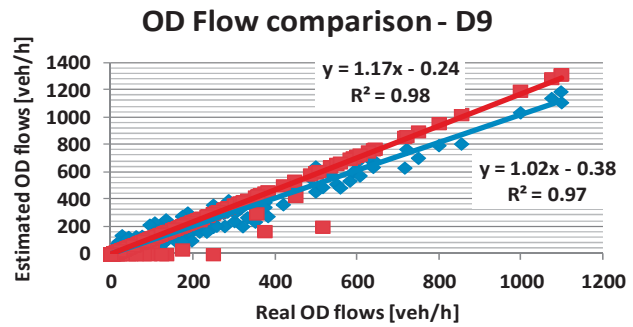
(a) OD flow evolution for Methods 3 and 4, demand scenario D7

(b) true vs. estimated OD flows for Method 2-OD and Method 3, demand scenario D7

(c) OD flow evolution for Methods 3 and 4, demand scenario D9

(d) true vs. estimated OD flows for Method 2-OD and Method 3, demand scenario D9

Figure 7 - Results for Method 2-OD and Method 3

Method 5 is a Kalman filtering method suitable for on-line applications. Results for Method 5, assuming a 100% penetration of Bluetooth equipped vehicles and prior OD scenarios D7, D8 and D9, are presented in Figure 8a for selected GOF measures: Normalized Root Mean Squared Error, Theil's U coefficient and R^2. Good performance of Method 5 is reflected in high R^2 fit (above 85%) that is obtained for the overall OD pairs and demand levels, but mostly for the most important OD flows (i.e. those in the 4th and 3rd quantiles). The fit of true versus estimated OD flows for all considered OD pairs (for the aggregated 1 hour period) for a scenario initialized with prior demand levels D7, D8 and D9 show coefficients of determination of the simple regression line of almost 90%, i.e. lower than the value achieved by Method 4, but the computational burden decreases from hours in Method 4 to minutes in Method 5. Figure 8b presents the evolution of estimated OD flows per departure time interval for the prior high OD demand scenario D9, demonstrating also how Method 5 is able to recover from an initial point characterized by overestimated prior OD flow with respect to the true OD flow.

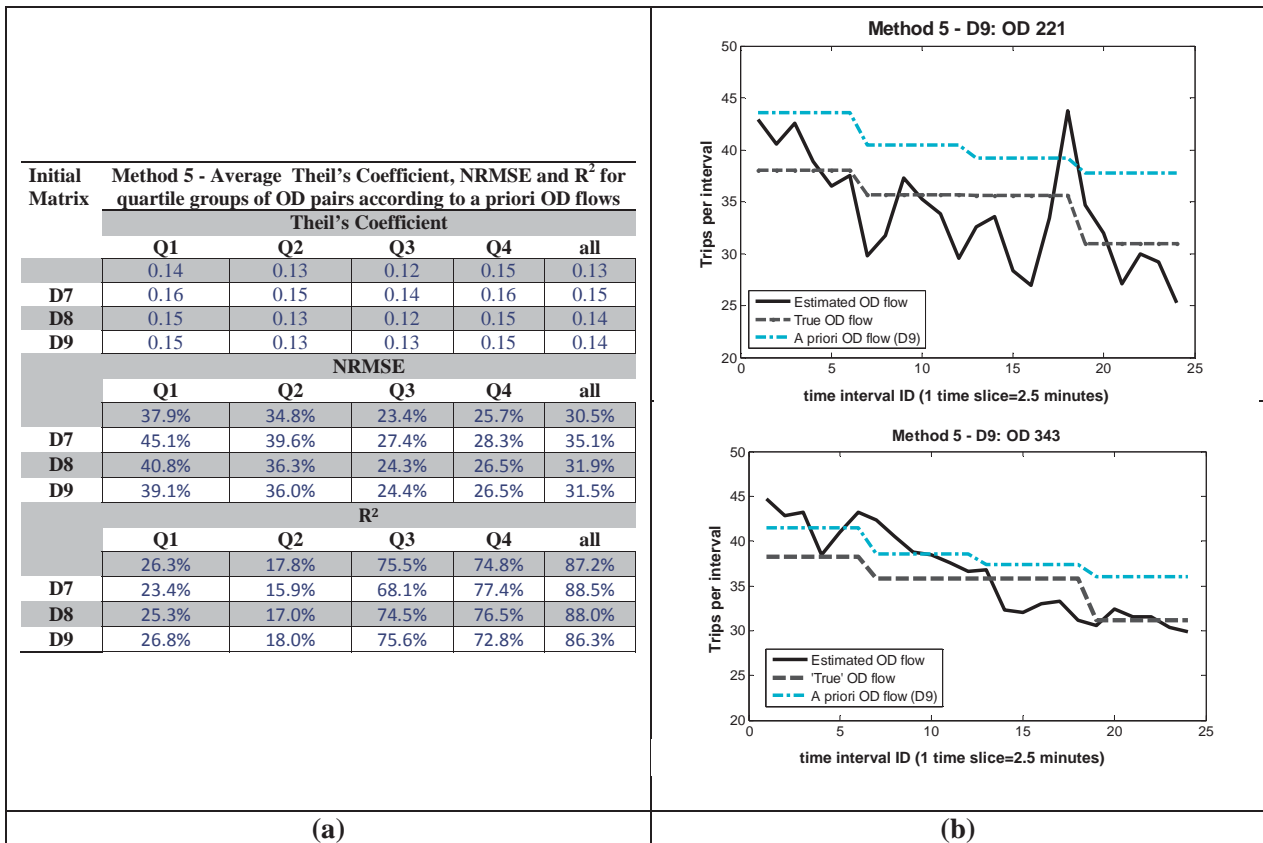| Initial Matrix	Method 5 - Average Theil's Coefficient, NRMSE and R^2 for quartile groups of OD pairs according to a priori OD flows				
	Theil's Coefficient				
	Q1	**Q2**	**Q3**	**Q4**	**all**
	0.14	0.13	0.12	0.15	0.13
D7	0.16	0.15	0.14	0.16	0.15
D8	0.15	0.13	0.12	0.15	0.14
D9	0.15	0.13	0.13	0.15	0.14
	NRMSE				
	Q1	**Q2**	**Q3**	**Q4**	**all**
	37.9%	34.8%	23.4%	25.7%	30.5%
D7	45.1%	39.6%	27.4%	28.3%	35.1%
D8	40.8%	36.3%	24.3%	26.5%	31.9%
D9	39.1%	36.0%	24.4%	26.5%	31.5%
	R^2				
	Q1	**Q2**	**Q3**	**Q4**	**all**
	26.3%	17.8%	75.5%	74.8%	87.2%
D7	23.4%	15.9%	68.1%	77.4%	88.5%
D8	25.3%	17.0%	74.5%	76.5%	88.0%
D9	26.8%	18.0%	75.6%	72.8%	86.3%

(a)   **(b)**

Figure 8 - Method 5 results: (a) performance indicators, (b) filtered values throughout 1h for prior OD scenario D9 (horizontal axis shows time intervals IDs, with a duration of 2.5 minutes each)

Overall, results presented in this section show that the potential of information from advanced traffic measurements to improve OD demand estimation can be fully exploited only using OD estimation methods capable to correct for biases in spatial and temporal OD pattern given by a prior OD matrix, especially in congested networks. In this respect, the computational experiments presented in this paper prove the robustness and quality of the OD estimates exploiting AVI measurements. The computational performance of the SPSA CG-TR proposal (Method 4) solved without prior OD information and using gradients and a trust region has been substantially increased by significantly reducing the number of function evaluations and the number of iterations, converging faster in this way to better demand estimates. This OD estimation method provides an effective tool for off-line pre-processing of prior OD data producing an initial OD matrix suitable for on-line applications of methods based on, for example, Kalman filtering (Method 5). According to the published research with Method 5 using Vitoria and Eixample networks, the performance of the method in terms of goodness of fit is highly dependent on the quality of the a priori OD matrix (Barceló et al., 2013).

7. Conclusion

The problem of OD flow estimation/updating is one of the main problems in modern traffic management systems and Intelligent Transportation Systems deployments and it receives a large amount of attention from researchers and practitioners alike. Demand is very volatile, in the sense that

it can be influenced by a number of factors, both intentional and unintentional, such as weather conditions, special events and public announcements. On the other hand, demand is in general unobservable, considering current technology. Therefore, it is rarely possible to know the "true" demand, and therefore it is difficult to know how well each approach performs in reality.

The nature of the OD estimation problem makes it a problem that is particularly difficult to solve. In particular, OD estimation tends to deal with a large number of unknowns (thousands or even tens or hundreds of thousands) and a limited number of observations (from a few hundreds of point measurements to several hundred or a few thousands point-to-point observations, using modern data collection techniques). These properties make it particularly difficult to benchmark and compare algorithms, as the underlying conditions are often very different.

In this paper, we motivate and present a flexible and software-agnostic platform for the implementation, validation and comparison of different OD flow estimation/updating algorithms. The structure of the platform is first presented in a general way, and then a specific implementation is described, as a blueprint for other implementations. This platform is general and can be used with any type of algorithm and traffic simulation software; furthermore, it is freely available to other researchers. A long list of algorithms are currently implemented and are being tested with this platform in a systematic and objective way, and in this paper we report on the results from some of these. We foresee (and hope) that more researchers will join the authors of this paper in contributing algorithms and results.

In particular, the presented results demonstrate that the developed framework is capable of supporting the development, application and testing of a wide range of algorithms. First, both off-line/planning level algorithms (like the Bilevel-DUE) and on-line algorithms (like that SPSA AD-PI and the LSQR) are presented. Furthermore, the modular and open design of the framework allows its extension so that it can accommodate other characteristics of the data and the algorithms considered by the researchers. In particular, an extension of the framework is presented, which allows the consideration of additional sources of data (in this case, data from Bluetooth sensors).

This research does not intend to claim or demonstrate the superiority of any algorithm or type of traffic data over the others, but was intended to show the potential of different algorithms and types of traffic data for dynamic OD estimation. The use of advanced traffic data to model dynamic OD demand is relatively new, and the literature still shows a lack of empirical experiments to validate their use for dynamic OD estimation. The benchmark study presented here indicates that advanced traffic data require more research efforts and new techniques to turn them into usable information.

Like many other aspects of research, dynamic OD estimation keeps evolving, taking advantage of occurring technological and methodological advances. In this context, it is now becoming likely that demand estimation will be at some point possibly supported by big data (e.g. from social networks or cell-phone traces). Furthermore, several existing traffic simulator models are using different demand representations, e.g. explicitly describing the mobility patterns of the travellers (Balmer et al., 2008; Flötteröd et al., 2011) or taking directly input from activity based or urban simulation models, such as the coupling presented in Nicolai et al. (2011). The presented platform is extensible and can accommodate all these developments –with some straightforward extension, which has been already demonstrated in this research– thus providing a common facility for the demonstration of the incremental benefits of moving to the new paradigms.

Acknowledgments

References

Antoniou, C., M. Ben-Akiva, and H. N. Koutsopoulos (2006). Dynamic Traffic Demand Prediction Using Conventional and Emerging Data Sources. IEE Proceedings Intelligent Transport Systems, Vol. 153, No. 1, pp. 97-104.

Antoniou C., Balakrishna R., Koutsopoulos H.N. and M. Ben-Akiva (2009). Off-Line and On-Line Calibration of Dynamic Traffic Assignment Systems. Presented at the 12th IFAC Symposium on Control in Transportation Systems.

Antoniou, C., J. Barceló, M. Brackstone, H.B. Celikoglu, B. Ciuffo, V. Punzo, P. Sykes, T. Toledo, P. Vortisch and P. Wagner (2014). Traffic Simulation: Case for Guidelines. Brackstone and Punzo (Eds.). JRC Scientific and technical Report, EUR 26534 EN. ISBN 978-92-79-35578-3. Publications Office of the European Union, Luxembourg.

Antoniou, C., C. L. Azevedo, L. Lu, F. Pereira and M. Ben-Akiva (2015). W-SPSA in practice: Approximation of weight matrices and calibration of traffic simulation models. Transportation Research Part C (accepted).

Asakura, Y., E. Hato, M. Kashiwadani (2000). Origin-destination matrices estimation model using automatic vehicle identification data and its application to the Han-Shin expressway network. Transportation, 27(4), pp. 419–438

Ashok, K. (1996). Estimation and prediction of time-dependent origin-destination flows. PhD Thesis, Massachusetts Institute of Technology, Cambridge, MA

Ashok K. and M. Ben-Akiva (1993). Dynamic origin-destination matrix estimation and prediction for real-time traffic management systems. Proceedings of the 12th ISTTT, 465-484.

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

Balakrishna, R. (2006) Off-Line Calibration of Dynamic Traffic Assignment Models. PhD thesis, Massachusetts Institute of Technology.

Balakrishna, R., Koutsopoulos, H. N. and M. Ben-Akiva (2005). Calibration and validation of dynamic traffic assignment systems. Proceedings of 16th ISTTT, 407-426.

Balakrishna, R., M. Ben-Akiva, H. N. Koutsopoulos (2007). Off-line calibration of dynamic traffic assignment: Simultaneous demand and supply estimation. Transportation Research Record, (2003), pp. 50–58.

Balakrishna, R., and H. N. Koutsopoulos (2008). Incorporating Within-Day Transitions in the Simultaneous Offline Estimation of Dynamic Origin–Destination Flows Without Assignment Matrices. Transportation Research Record: Journal of the Transportation Research Board 2085: 31–38.

Balmer, M., K. Meister, M. Rieser, K. Nagel and K.W. Axhausen (2008). Agent-based simulation of travel demand: Structure and computational performance of MATSim-T. paper presented at the 2nd TRB Conference on Innovations in Travel Modeling, Portland, June 2008

Barceló, J. (Ed.). (2010). Fundamentals of traffic simulation, vol. 145. Springer.

Barceló, J., F. Gilliron, M. Linares, O. Serch, and L. Montero (2012) Exploring Link Covering and Node Covering Formulations of Detection Layout Problem. Transportation Research Record:

Journal of the Transportation Research Board, Vol. 2308, No. 1, 2012, pp. 17–26.

Barceló, J., L. Montero, M. Bullejos, M. Linares, and O. Serch (2013). Robustness and computational efficiency of a Kalman Filter estimator of time dependent OD matrices exploiting ICT traffic measurements. Transportation Research Record: Journal of the Transportation Research Board, 2344 (4), pp. 31–39.

Barceló J., Montero L., Bullejos M., M.P. Linares (2014). A practical proposal for using origin-destination matrices in the analysis, modeling and simulation for traffic management. Proceedings of the 93rd TRB Annual Meeting

Bell, M.G.H. (1983). The estimation of origin-destination matrix from traffic counts. Transportation Science 10, 198-217.

Bell, M.G.H. (1987). Transportation Network Analysis. John Wiley Ed.

Bierlaire, M. (2002) The total demand scale: a new measure of quality for static and dynamic origin-destination trip tables, Transportation Research Part B: Methodological, 36(9), pp. 837–850.

Bierlaire, M. and F. Crittin (2004). An Efficient Algorithm for Real-Time Estimation and Prediction of Dynamic OD Tables. Operations Research, Vol. 52 (1), pp. 116–127.

Buisson C., Daamen W., Punzo V, Montanino M. (2014). Calibration and Validation Principles, in Traffic Simulation and Data editors: Daamen W. Buisson C. and Hoogendoorn S.P., pp. 89-117, CRC Press, Taylor & Francis Group.

Bullejos, M., J. Barceló, and L. Montero (2014). A DUE based bilevel optimization approach for the estimation of time sliced OD matrices. International Symposium of Transport Simulation, Ajaccio, Corsica, Procedia Social and Behavioral Sciences, Ajaccio, Corsica.

Cantelmo G, Viti F, Tampere CMJ, Cipriani E, Nigro M (2014a). Two-Step approach for correction of seed matrix in dynamic demand estimation. Transportation Research Record, vol. 2466, p. 125-133

Cantelmo G, Cipriani E, Gemma A, Nigro M (2014b). An adaptive bi-level gradient procedure for the estimation of dynamic traffic demand. IEEE Transactions On Intelligent Transportation Systems, ISSN: 1524-9050, doi: 10.1109/TITS.2014.2299734

Cantelmo, G., Viti, F., Cipriani, E., Nigro, M. (2015). Improving the reliability of a two-step dynamic demand estimation approach by sequentially adjusting generations and distributions. Transportation Research Board, 2015 Annual Meeting Compendium of Papers.

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

Cascetta, E. (2009). Transportation systems analysis: models and application, 2nd Edition, Springer.

Cascetta, E. and S. Nguyen (1988). A unified framework for estimating or updating origin/destination matrices from traffic counts. Transportation Research 22B, 437–455.

Cascetta, E., Inaudi, D., Marquis, G. (1993). Dynamic Estimators of Origin-Destination Matrices using Traffic Counts. Transportation Science 27(4), 363-373.

Cascetta, E. and Postorino M.N. (2001). Fixed point models for the estimation of OD matrices using traffic counts on congested networks. Transportation Science 35(3).

Cascetta, E., Papola A., Marzano V., Simonelli F., Vitiello I. (2013). Quasi-dynamic estimation of o-d flows from traffic counts: formulation, statistical validation and performance analysis on real data. Transportation research Part B, vol. 55, pp.171-187

Chang, G.-L., X. Tao (1996). Estimation of dynamic od distributions for urban networks, in: Transportation and traffic theory: Proceeding of the 13th international symposium on transportation and traffic theory.

Chang, G.-L., J. Wu (1994). Recursive estimation of time-varying origin-destination flows from traffic counts in freeway corridors, Transportation Research Part B: Methodological, 28(2), pp. 141–160.

Cipriani, E., M. Florian, M. Mahut, and M. Nigro (2010). Investigating the efficiency of a gradient approximation approach for the solution of dynamic demand estimation problems. In New

Developments in Transport Planning: Advances in Dynamic Transport Assignment (V. F. I. L. Tampere, C.M.J., ed.), Edward Elgar, Cheltenham UK and Northampton, MA, USA

Cipriani, E., M. Florian, M. Mahut, and M. Nigro (2011). A gradient approximation approach for adjusting temporal origin-destination matrices. Transportation Research Part C: Emerging Technologies, Vol. 19, No. 2, 2011, pp. 270–282.

Cipriani, E, Nigro M, Fusco G, Colombaroni C (2013). Effectiveness of link and path information on simultaneous adjustment of dynamic o-d demand matrix. European Transport Research Review, ISSN: 1867-0717, doi: 10.1007/s12544-013-0115-z

Cipriani, E, Gemma A, Nigro M, Colombaroni C, Fusco G (2014). Reliability of SPSA AD-PI method for dynamic demand estimation: a case of study. In: ACE 2014 Proceeding Book. Istanbul, Turkey

Ciuffo and Punzo (2010). Verification of Traffic Micro-simulation Model Calibration Procedures: Analysis of Goodness-of-Fit Measures. Proceeding of the 89th Annual Meeting of the Transportation Research Record, Washington, D.C.

Ciuffo B., Punzo V. (2014). "No free lunch" theorems applied to the calibration of traffic simulation models. IEEE Transactions on Intelligent Transportation Systems vol. 15, No. 2, p.553-562.

Ciuffo B., Punzo V., and V. Torrieri (2008). A comparison between simulation based and model-based calibrations of traffic flow micro-simulation models. Transp. Res. Rec., vol. 2088, pp. 36–44.

Cremer, M., H. Keller (1987) A new class of dynamic methods for the identification of origin-destination flows, Transportation Research Part B: Methodological, 21(2), pp. 117–132.

Dixon, M. P., L. R. Rilett (2002) Real-time od estimation using automatic vehicle identification and traffic count data, Computer-Aided Civil and Infrastructure Engineering, 17(1), pp. 7–21.

Djukic, T., G. Flotterod, H. van Lint, and S. Hoogendoorn (2012). Efficient real time OD matrix estimation based on Principal Component Analysis. Proceedings of the IEEE Intelligent Transportation Systems Conference, Vol. 2, pp. 115-121.

Djukic, T. (2014). Dynamic OD demand estimation and prediction for dynamic traffic management. PhD thesis, TU Delft, Delft University of Technology.

Florian, M. and Y. Chen. (1995). A coordinate descent method for the bi-level O–D matrix adjustment problem. International Transportation Operations Research 2, 165–179.

Flötteröd, G. and M. Bierlaire (2009). Improved Estimation of Travel Demand from Traffic Counts by a New Linearization of the Network Loading Map, Proceedings of 2009 European Transport Conference.

Flötteröd, G., M. Bierlaire and K. Nagel (2011). Bayesian demand calibration for dynamic traffic simulations. Transportation Science, 45 (4) 541-561

Frederix, R., C. M. Tampere, F. Viti, L. H. Immers (2010). The effect of dynamic network loading models on DTA-based od estimation, New developments in transport planning: advances in dynamic traffic assignment, Edward Elgar, Cheltenham UK and Northampton, MA, USA.

Frederix, R., Viti F., Corthout R., and C. M. J. Tampère (2011). New Gradient Approximation Method for Dynamic Origin-Destination Matrix Estimation on Congested Networks. Transportation Research Record: Journal of the Transportation Research Board 2263, 19-25.

Frederix, R., F. Viti, and C. M.J. Tampère (2013). Dynamic Origin-destination Estimation in Congested Networks: Theoretical Findings and Implications in Practice. Transportmetrica A: Transport Science 9, no. 6 (2013): 494-513

Gao, S., Frejinger, E., Ben-Akiva, M. (2010). Adaptive route choices in risky traffic networks: A prospect theory approach. Transportation Research Part C: Emerging Technologies, 18(5), 727-740

Hazelton, M. (2003). Some comments on origin–destination matrix estimation. Transportation Research 37A, 811–822.

Hollander, Y., Liu, R. (2008). The principles of calibrating microsimulation models. Transportation, Vol. 35, pp. 347-362.

Lindveld, K., I. (2003). Dynamic OD matrix estimation: a behavioural approach, TRAIL thesis series, Netherlands TRAIL Research School.

Lo, H.P. and C.P. Chan. (2003). Simultaneous estimation of an origin-destination matrix and link choice proportions using traffic counts. Transportation Research 37A, 771–788.

Lo, H., Zhang, N. and H. Lam (1996). Estimation of an origin-destination matrix with random link choice proportions: a statistical approach. Transportation Research 30B, 309–324.

Lu, L., Y. Xu, C. Antoniou and M. Ben-Akiva (2015). An Enhanced SPSA Algorithm for the Calibration of Dynamic Traffic Assignment Models. Transportation Research Part C, 51, pp. 149-166.

Lundgren T. J., Peterson A. (2008). A heuristic for the bilevel origin–destination-matrix estimation problem, Transportation Research Part B: Methodological, 42(4), 339-354

Maher, M. (1983). Inferences on trip matrices from observations on link volumes: a Bayesian statistical approach. Transportation Research 20B, 435–447.

Marzano V., Papola A., Simonelli F. (2009). Limits and perspectives of effective o-d matrix correction using traffic counts. Transportation Research Part C, 17(2), pp. 120-132

Marzano V., Simonelli F., Papola A., Cascetta E. (2015). A Kalman filter for online quasi-dynamic o-d flows estimation/updating. Proceedings of 2015 MT-ITS conference, Budapest.

MATLAB (2013). MATLAB and Statistics Toolbox Release 2012b, The MathWorks, Inc., Natick, Massachusetts, United States.

Mishalani, R., B. Coifman, D. Gopalakrishna (2002). Evaluating Real-Time Origin-Destination Flow Estimation Using Remote Sensing-Based Surveillance Data, chap. 80, pp. 640–647.

Nicolai T.W., L. Wang, K. Nagel and P. Waddell (2011) Coupling an urban simulation model with a travel model – A first sensitivity test, paper presented at Computers in Urban Planning and Urban Management (CUPUM), Lake Louise, Canada.

Nie, Y., H. M. Zhang, W. W. Recker (2005). Inferring origin-destination trip matrices with a decoupled GLS path flow estimator, Transportation Research Part B: Methodological, 39(6), pp. 497–518.

Nihan N. L. and Davis G. A. (1987) Recursives estimation of origin-destination matrices from input/output counts. Transpn. Res., 21B, 149-163.

Nihan N. L, and Davis G. A. (1989) Application of prediction-error minimization and maximum likelihood to estimate intersection O-D matrices from traffic counts. Transpn. Sci., 23, 77-90.

Okutani I. and Y. Stephanedes (1984). Dynamic Prediction of Traffic Volume through Kalman Filtering Theory. Transportation Research B, 18(2).

Ortuzar, J. And L. Willumsen (2011). Modelling transport. 4th edition. Wiley ed.

Osorio, C. and M. Bierlaire (2010). A simulation-based optimization approach to perform urban traffic control. in Proceedings of the Triennial Symposium on Transportation Analysis (TRISTAN), Troms, Norway.

Ossen S. J. L. and S. P. Hoogendoorn (2008). Validity of trajectory-based calibration approach of car-following models in the presence of measurement errors. Transp. Res. Rec., vol. 2088, pp. 117–125

Paige, C. and M. A. Saunders (1982). LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares. ACM Trans. Math. Softw. 8(1), 43-71.

Punzo V., Ciuffo B. and Montanino M. (2012). Can results of car-following model calibration based on trajectory data be trusted? Transp. Res. Rec., vol. 2315, p. 11-24.

Simonelli F., Marzano V., Papola A., Vitiello I. (2012). A network sensor location procedure accounting for o-d matrix estimate variability. Transportation Research part B, vol. 46, pp. 1624-1638.

Sohn, K., D. Kim (2008). Dynamic origin destination flow estimation using cellular communication system, Vehicular Technology, IEEE Transactions on, 57(5), pp. 2703–2713.

Spall, J.C. (1992). Multivariate Stochastic Approximation Using a Simultaneous Perturbation Gradient

Approximation, *IEEE Transactions on Automatic Control*, vol. **37**, pp. 332-341.

Sun, J., and Y. Feng (2011). A novel od estimation method based on automatic vehicle identification data, in: Chen, R., ed., Intelligent Computing and Information Science, vol. 135 of Communications in Computer and Information Science, Springer Berlin Heidelberg, pp. 461–470.

Tavana, H. & Mahmassani, H.S. (2001). Estimation of Dynamic Origin-Destination Flows from Sensor Data using Bi-level Optimization Method. Presented at the 80th annual meeting of the Transportation Research Board, Washington DC, USA.

Tavana, H. (2001). Internally consistent estimation of dynamic network origin–destination flows from intelligent transportation systems data using bilevel optimization. PhD thesis. Austin: University of Texas.

Toledo T., Kolechkina T., Wagner P., Ciuffo B., Azevedo C., Marzano V., Flötteröd G. (2015). Network model calibration studies. In: Daamen W., Buisson C. and S. Hoogendoorn (eds). Traffic simulation and data: validation methods and applications, CRC Press, Taylor and Francis, London, pp. 141-162

TSS (2013). Aimsun 7 dynamic simulator User's Guide. Transport Simulation Systems, Barcelona, Spain, version 1.1.

Tympakianaki, A., Koutsopoulos, H. N., and Jenelius, E. (2015). 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.

Van Der Zijpp, N. J. and E. Romph (1997). A Dynamic Traffic Forecasting Application on the Amsterdam Beltway. In: Modelstudies Verkeersbeheersingsmaatregelen, Stichting PAO, P. Bovy and N. J. Van der Zijpp (editors)

Van der Zijpp, N. J. and C. Lindveld (2001). Estimation of origin-destination demand for dynamic assignment with simultaneous route and departure time choice, Transportation Research Record: Journal of the Transportation Research Board, 1771, pp. 75–82.

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

Vardi, Y. (1996). Network tomography: estimating source–destination traffic intensities from link data. Journal of the American Statistical Association 91, 365–377.

Vaze V., C. Antoniou, Y.Wen, and M. Ben Akiva (2009). Calibration of dynamic traffic assignment models with point-to-point traffic surveillance," Transp. Res. Rec., vol. 2090, pp. 1–9

Viti F., Rinaldi M., Corman F. and C.M.J. Tampere (2014). Assessing partial observability in network sensor location problems. Transportation Research Part B: Methodological, Volume 70, 65-89

Yang, Hai, and M. G. H. Bell (1998). Models and Algorithms for Road Network Design: A Review and Some New Developments. Transport Reviews 18, no. 3 (1998): 257-78.

Yang, H., Y. Iida, T. Sasaki (1991) An analysis of the reliability of an origin-destination trip matrix estimated from traffic counts, Transportation Research Part B: Methodological, 25(5), pp. 351–363.

Yang, H., Zhou, J. (1998). Optimal traffic counting locations for origin-destination matrix estimation. Transportation Research Part B 32(2), 109-126.

Yang, H. (1995). Heuristic algorithms for the bi-level origin-destination matrix estimation problem. Transportation Research 29B, 231-242.

Zheng F., Van Zuylen H. (2013).Urban link travel time estimation based on sparse probe vehicle data. Transportation Research Part C 31 (2013) 145–157

Zhou, X., X. Qin, H. S. Mahmassani (2003). Dynamic origin-destination demand estimation with multiday link traffic counts for planning applications, Transportation Research Record: Journal of the Transportation Research Board, 1831(1), pp. 30–38.

Zhou, X., H. S. Mahmassani (2006). Dynamic origin-destination demand estimation using automatic vehicle identification data, Intelligent Transportation Systems, IEEE Transactions, 7(1), 105–114

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

Appendix. Implementation of the benchmarking platform in an integrated Matlab+Aimsun environment

This appendix provides details of the benchmarking platform described in Section 5, whose two main elements are:

- *traffic simulator*: the mesoscopic version of the Aimsun simulation model (TSS, 2013) was used as the common traffic model. The mesoscopic model was used because it is substantially faster than the microscopic one. Thus, it allows for more elaborate testing and a richer experimental design. A default set of parameters was used in all cases;
- *OD estimation algorithm codes*: MATLAB was used to code all algorithms. This approach allowed the development of common functions to write inputs for the simulation, execute it and read simulation outputs, thus reducing the differences in run times that stem from the efficiency of these functions and lowering the workload for using the framework.

A dynamic communication between the MATLAB and Aimsun software was necessary in order to execute a traffic simulation run within the OD estimation algorithm. To this aim a MATLAB function was created allowing the following logical steps:

- it receives as input a traffic demand matrix generated by the estimation algorithm;
- it calls Aimsun for a new traffic simulation run with the new traffic demand and waits until the simulation ends. The actual communication of the instructions for the Aimsun call is done through Python;
- it imports the result of the simulation run as matrices in the MATLAB environment.

The MATLAB function that forms the engine for Aimsun execution and communication within the OD estimation algorithms uses the following inputs:

- the demand pattern to be simulated in the form of OD flows per time interval;
- the time series of traffic data to be compared with the outputs of the new traffic simulation. The user may choose the types of traffic data among counts, speeds, densities and occupancies at detectors, and the intervals to be considered;
- a subset of the OD pairs for which the average travel times are requested;
- the number of replications to carry out with the specific input.

The Aimsun scenario is then simulated by creating and executing a batch file, which launches the Aimsun executable and the python script with the relevant information. When the simulation is finished, the MATLAB function collects and organizes all the outputs, and produces several relevant outputs:

- an array of 12 measures of goodness-of-fit (GoF) resulting from the comparison between reference and simulated traffic data;
- a matrix with the time-dependent set of simulated traffic data;
- a matrix with the dynamic assignment matrix resulting from the simulation;
- a matrix with the average travel times between the OD pairs defined in input.

Figure 7 presents a flowchart that shows the main elements of this platform, i.e. representing a

29

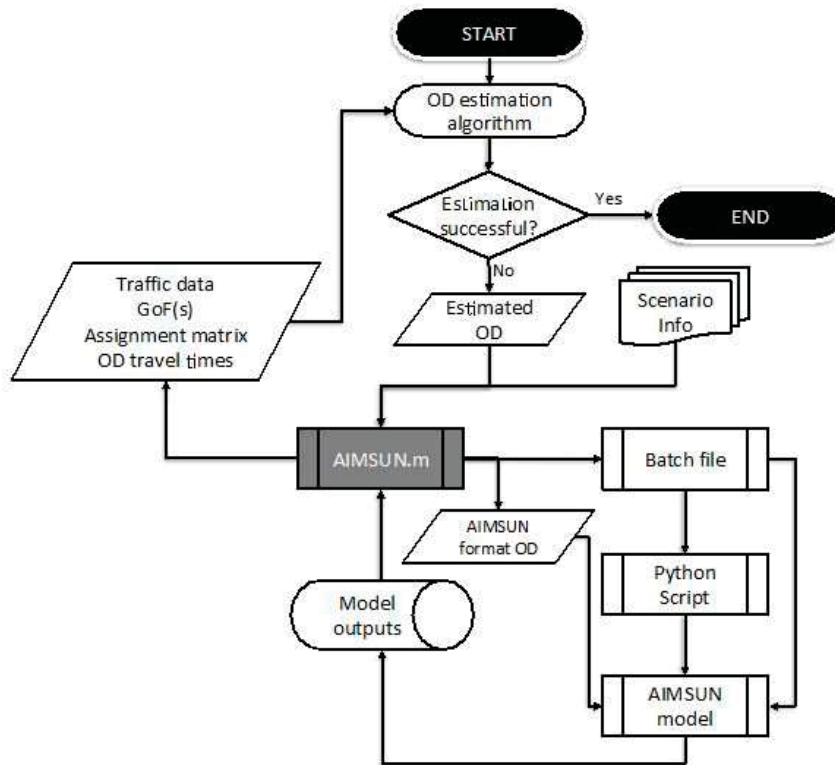practical instance of the general flowchart depicted in Figure 3.

Figure 7 - Flowchart with the main elements of the AIMSUN.m MATLAB function

Within the OD estimation main function, whenever a simulation run is needed 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 Aimsun with these inputs. 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, assignment matrices and travel times.

The AIMSUN.m function has been developed to be compatible with AIMSUN 7 and to operate under Microsoft Windows (32 and 64 bit) and Mac OSs. The function has been developed in the framework of the MULTITUDE Project (www.multitude-project.eu) and is freely available from the authors of the present paper.