A FRAMEWORK FOR THE BENCHMARKING OF OD ESTIMATION AND PREDICTION ALGORITHMS

Constantinos Antoniou*
National Technical University of Athens, Greece (e-mail: antoniou@central.ntua.gr)

Biagio Ciuffo
Institute for Energy and Transport of the European Commission, Joint Research Centre, Ispra (Varese), Italy (email: biagio.ciuffo@ jrc.ec.europa.eu)

Lídia Montero
Technical University of Catalonia (UPC), Barcelona (email: lidia.montero@upc.edu)

Jordi Casas
TSS (email: casas@aimsun.com)

Jaume Barcelò
Technical University of Catalonia (UPC), Barcelona (email: jaume.barcelo@upc.edu)

Ernesto Cipriani
"Roma Tre" University (email: ernesto.cipriani@uniroma3.it)

Tamara Djukic
Delft University of Technology (email: t.djukic@tudelft.nl)

Vittorio Marzano
"University of Napoli Federico II" (email: vmarzano@unina.it)

Marialisa Nigro
"Roma Tre" University (email: marialisa.nigro@uniroma3.it)

Manuel Bullejos
Technical University of Catalonia (UPC), Barcelona (email: manuel.bullejos@upc.edu)

Josep Perarnau
TSS (email: josep@aimsun.com)

Martijn Breen
TSS (martijn.breen@aimsun.com)

Vincenzo Punzo
Institute for Energy and Transport of the European Commission, Joint Research Centre, Ispra (Varese), Italy (email: vincenzo.punzo@ jrc.ec.europa.eu) and with the "University of Napoli Federico II" (email: vinpunzo@unina.it)

Tomer Toledo
Technion – Israel Institute of Technology (email: toledo@technion.ac.il)

*Corresponding author

Word Count: 6500 words + 0 tables + 4 figures = 7500

Submitted on August 1st, 2013
Submitted for presentation in the 93rd Annual Meeting of the Transportation Research Board.

A FRAMEWORK FOR THE BENCHMARKING OF OD ESTIMATION AND PREDICTION ALGORITHMS

ABSTRACT

In this research we describe the development of a common evaluation and benchmarking platform that has been developed within the framework of the European Union COST Action MULTITUDE. The main goal of this platform is to provide a testbed in which a number of algorithms can be implemented and tested under the same conditions. The objective is not to conclude that one approach is "best", but to provide a support comparison in a variety of settings and conditions in order to help determine the particular situations and conditions under which one approach might behave more favorably than another.

The design of the platform is presented, along with a detailed experimental design for the application of different OD estimation algorithms. The considered algorithms are then presented, along with a demonstration of the extensibility of the presented framework to accommodate additional data sources. 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 KFX2) are presented.

Keywords: traffic modeling, Origin-Destination (OD) estimation

INTRODUCTION

Dynamic traffic models, such as microscopic and mesoscopic traffic simulators [1] are traditionally used in the design and evaluation of Advanced Traffic Management and Information Systems (ATMS/ATIS). They also support real-time traffic management decisions. These models have experienced a significant maturation becoming suitable tools for almost any type of traffic analysis applications. Time-dependent Origin-Destination (OD) matrices are essential inputs to these models, both for research and practice purposes. However, a major contradiction is that quite frequently these sophisticated models have as main input a very rough and low quality information on the time variability of traffic patterns as described by the OD matrices. This results in situations in which it is hard for the analyst to identify whether flaws in the intended model are due to modeling mistakes, an improperly calibrated model or an unsuitable specification of the time varying demand. This state has fostered the interest in the estimation of time-dependent Origin-Destination (OD) matrices in the last decades. One of the biggest obstacles in assessing OD estimation and prediction algorithms is the lack of consistency in the presented results. Each researcher or developer tests their algorithms and approaches under different assumptions, with different networks and traffic conditions, using different data and goodness of fit measures. Even when a proposed approach is compared with alternatives approaches, it can be expected that due to various reason, such as familiarity with the alternatives and selection of suboptimal parameter values, the comparison might not be completely fair and informative.

In this research we describe the development of a common evaluation and benchmarking platform that has been developed within the framework of the European Union COST Action MULTITUDE. The main goal of this platform is to provide a testbed in which a number of algorithms can be implemented and tested under the same conditions. The objective is not to conclude that one approach is “best”, but to provide a support comparison in a variety of settings and conditions in order to help determine the particular situations and conditions under which one approach might behave more favorably than another.

The remainder of this paper is structure as follows. A literature review of the problems of OD estimation and prediction is presented next, followed by a description of the platform design and implementation. The experimental design process for the conducted experiments is outlined next. The following section describes the algorithmic approaches that have been considered for implementation and testing, while the next section presents some selected results from the application of the framework. The paper concludes with a discussion and concluding statements.

BACKGROUND - LITERATURE REVIEW

Estimation/updating of OD flows based on traffic counts is a classic 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 ATIS/ATMS) contexts. Normally, a prior OD estimate is obtained through a combination of surveys and mathematical models: see [2-4] for thorough reviews. Unfortunately, the resulting estimate is often affected by substantial errors, mainly related to the inherent complexity of the behavioural phenomena underlying the demand patterns (e.g. departure time and destination choices). These errors may be mitigated by updating the prior estimate of OD flows using observed traffic counts and possibly other network-based measurements. The updated posterior OD flows are found using a properly specified statistical estimator such that they are able to fit satisfactorily the available traffic measurements.

The estimation/updating of OD flows has been studied extensively in the context of static systems. Four main approaches have been proposed:

1. Minimum information/maximum entropy [5].
2. Maximum Likelihood [6,7].
3. Generalized Least Squares [8].
4. Bayesian approaches [9].

A number of generalizations and extensions of the OD updating problem have been proposed in the literature. Examples include incorporating the treatment of congested network through bi-level optimization [10-12], taking into account the stochastic nature of traffic counts [13,14], estimating simultaneously OD flows and the route choice model parameters [15], or dealing with the availability of traffic counts on multiple days [16].

The first extension to the within-day dynamic framework was provided by Cascetta et al. [17] through the proposition of two estimators: simultaneous and sequential. A recent quasi-dynamic approach showing promising results was proposed by Cascetta et al. [18]. Day-to-day dynamics also received some attention, with the proposition of methods to capture the process of traffic evolution over days, e.g. [19].

In online applications, the dynamic OD estimation process is required to recursively provide fast estimates for recent time slices together with predictions for future time slices. Online estimation was first proposed by Okutani and Stephanedes [20] and subsequently generalized by Ashok and Ben-Akiva [21, 22] and Ashok [23] that acknowledged the importance of structural information in OD flows. He modelled the within-day evolution of deviation of OD flows from historical estimates using a Kalman filter based on an autoregressive process. Zhou and Mahmassani [19] assumed a polynomial approximation for the structural deviation of the demand from the historical estimate as an alternative to the autoregressive process. Computational issues in online within-day OD estimation in large networks were addressed by Bierlaire and Crittin [24]. Notably, the Kalman filter can be used also for off-line applications, as proposed by Gelb [25] and Balakrishna et al. [26].

A number of variations on the dynamic estimation framework were proposed in the literature. For instance, Cremer and Keller [27] and Ashok and Ben-Akiva [28] introduced randomness in the dynamic assignment matrix. In order to overcome the difficulty to obtain prior knowledge of the dynamic assignment map, Cremer and Keller [29] proposed an OD estimation approach that does not use assignment matrix information. Other research direction related to the OD estimation problem have also been proposed. Liu and Fricker [30] dealt with joint estimation of demand and supply parameters. More recently, Antoniou et al. [31] and Cipriani et al. [32] developed efficient algorithms for this problem. Djukic et al. [33] explored methods to reduce the high dimensionality of OD estimation problem using principal component analysis (PCA). Several other authors studied the use of an expanded set of measurements [34-36].

PLATFORM DESIGN AND IMPLEMENTATION

A key feature for the OD estimation algorithms benchmarking exercise was the definition of a common framework to ensure equal testing conditions for various proposed methods that would support fair comparison and an understanding of their relative merits. The main elements of the common framework are:

- Traffic simulator: The mesoscopic version of the Aimsun simulation model [37] 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 [38] was used to code all algorithms. This approach allowed writing 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 work load 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 1 presents a flowchart that shows the main elements of this platform. 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.

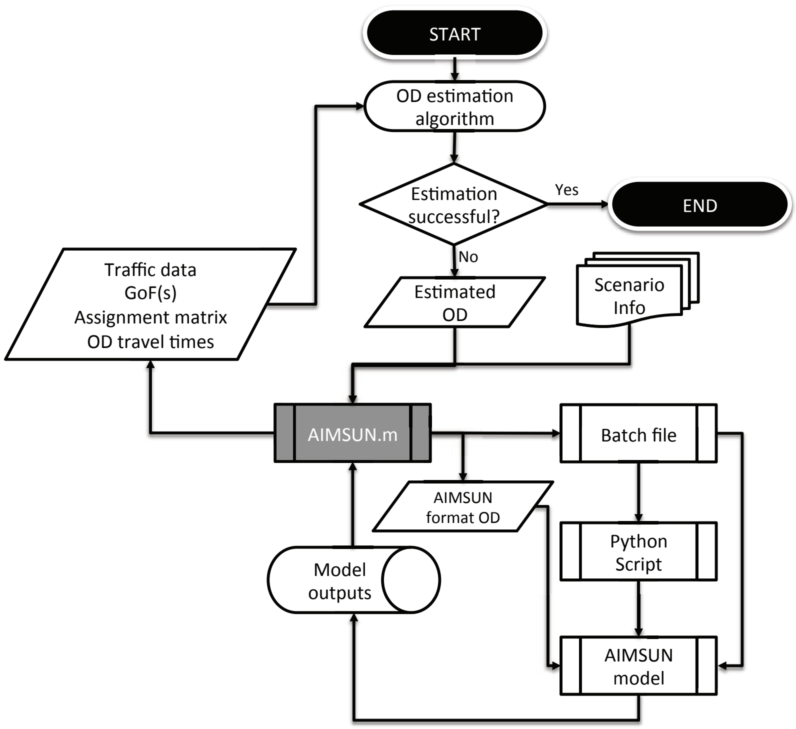


FIGURE 1. Flowchart with the main elements of the AIMSUN.m MATLAB function

EXPERIMENTAL DESIGN

Main dimensions of the experimental design

Three networks have been developed for the benchmarking effort:

1. A small network, used primarily for debugging and verification purposes
2. A medium-size network from Vittoria, Spain. This network includes 57 centroids and 2800 intersections. The road length is 600km. Traffic data is available from 389 detectors.
3. A larger network from Barcelona, Spain. This network includes 130 centroids, 1570 nodes and 2800 links.

The level of demand is a key element affecting the performance of OD estimations. It is well known that the problem becomes harder under congestion. The experimental design considers three different demand levels. These levels are grounded in the base demand level *D* for each of the networks. In constructing the three "true" demand levels, each entry in the demand matrix was perturbed randomly according to the following patterns, which capture various conditions around the base demand *D* for each network (i.e. both higher and lower):

1. Low demand (denoted as D7 in the presented results in section VI): D*[0.7+0.3*rand( )]. The mean value of this demand is 85% of the base demand, with a range of +/- 15%.
2. Medium demand (D8): D*[0.8+0.3*rand( )]: The mean value of this demand is 95% of the base demand, with a range of +/- 15%.
3. High demand (D9): D*[0.9+0.3*rand( )]: The mean value of this demand is 105% of base demand, with a range of +/- 15%.

A similar approach was followed in developing scenarios regarding the characteristics of the surveillance system. For each of the demand levels described above, using the "true" demand within the simulation model, the "true" measurements Y were calculated. These measurements need to be corrupted with noise to mimic measurement errors in the real world. In all cases, the duration of the time intervals were set to 15 minutes, for both the OD matrices and the surveillance measurements.

The scenarios are also defined in terms of average or realization conditions. A realization is based on the results from a single replication of the simulation. Average conditions are based on calculating the average values from a set of ten replications

Measures of goodness-of-fit and measures of performance

A number of goodness-of-fit measures can be used to evaluate the overall performance of OD estimation and prediction algorithms. For a thorough review, see [55, 56]. 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.

CONSIDERED ALGORITHMS: SELECTION, IMPLEMENTATION AND TESTING

The considered algorithms include (i) Kalman filter variants (in which case the problem is formulated as a state-space model), such as the Extended Kalman Filter (EKF), the Limiting EKF and quasi-dynamic Kalman Filter and special linear versions of Kalman Filter and linear state-space formulations, as well as (ii) direct optimization algorithms (in which case the problem is formulated as a standard optimization problem), such as SPSA, GLS, and LSQR.

This section provides an overview of the algorithms that have been considered in this process and implemented in the developed platform. Different types of data are considered for different algorithms. Besides the conventional loop detectors, counts from Bluetooth (BT) detectors and travel time information between detectors (e.g. Bluetooth sensors) are also considered. Due to the special requirements of some algorithms utilizing this type of detectors, a modification of the developed framework is required. This modification is described in this section (and applied in the case studies, along-side the unmodified framework), thus demonstrating the flexibility and extensibility of the approach and implementation.

In Barceló et al. [39], a linear state-space formulation has been adopted for real-time estimation and short-term prediction of OD trip matrices. The approach exploits the travel times and counts collected, respectively, by tracking Information and Communication Technologies (ICT) equipped vehicles and conventional detection technologies. Time-varying dependencies between measurements (sensor counts) and state variables (deviations of equipped OD path flows) are modeled by updating discrete approximations to travel time distributions that exploit the travel ICT time measurements from equipped vehicles. Keeping a linear relationship between state-variables and measurements is computationally advantageous and reduces the number of state variables in KF formulation. State variables are defined as deviations of OD path flows in a subset defined as the most likely OD path flows identified from a DUE assignment (based on the Historic time-sliced OD matrix).

The approach fits the needs of real-time applications but it has proved to be very sensitive to the quality of the initialization in terms of the Historic OD matrix [40]. To feed the KF short-term prediction module an off-line approach to the estimation of Historic time-sliced OD matrices according to day-to-day variability has been developed in UPC. The proposal is a gradient approximation approach for adjusting time-sliced origin–destination matrices based on a bilevel formulation where the lower level is a DUE problem and the upper level aims at minimizing the "distance" between actual and estimated observations (counts, travel times, speeds) and the “distance” between estimated and a-priori Historic O-D matrix. The method is an adaptation and extension of [32]. The extension adds a new term to the objective function, and therefore to the computation of the gradient, to account for the available travel times between Bluetooth sensors along the main paths connecting them in the network, defined by a suitable layout [41] that allow their identification.

The Kalman filter is the optimal minimum mean square error (MMSE) estimator for linear state-space models [42]. However, the OD estimation is usually nonlinear (due to the indirect measurement equation). The most straightforward extension is the Extended Kalman Filter (EKF), in which optimal quantities are approximated via first- order Taylor series expansion (linearization) of the appropriate equations [25, 42]. The EKF has found several applications in the field of ITS, including, e.g., on-line calibration of traffic dynamics models, short-term travel time prediction, and modeling of car-following driver behavior. A special case of the EKF with very favorable computational properties is the limiting EKF (LimEKF) [43, 44]. Another variance of the EKF that has been implemented is the SP-EKF, which uses the SPSA [45, 46] algorithm for the linearization step [47].

Different variants of the SPSA algorithm have been proposed in [32, 48], where the offline dynamic OD demand estimation problem is formulated as a bi-level nonlinear optimization program and solved with an assignment-matrix-free method dealing with the Asymmetric Design (AD) for gradient computation and the Polynomial Interpolation (PI) of the objective function for the linear optimization. In a recent work [49], a second order SPSA AD-PI has been investigated: the proposed “Adaptive SPSA” (ASP) method derives by an analogue of the Newton-Raphson one [50, 51]; as the latter, it allows to overcome the difficulty in optimizing variable components that present substantial magnitude differences. This scaling property is obtained computing the inverse of the estimation of the Hessian matrix of the objective function. In applications where the assignment matrix is available, the second order approach has inspired some developments [49] that consist in using information deriving from OD path proportions on each sensor, in place of the Hessian matrix, to weigh the approximated gradient, so speeding the convergence of solution procedure.

An extension and adaptation of the basic version proposed by Cipriani et al. [32] has also been studied by the UPC team assuming the availability of travel times between Bluetooth sensors along the main paths connecting them in the network. The previous research reported in Barceló et al. [41] had proved that a suitable Bluetooth sensor layout allows the identification of the paths between sensors and therefore the measurement of the associated travel times. Consequently, to implement the proposed approach, we needed that the lower level DUE conducted with Aimsun Meso generates not only the simulated flows and speeds at traditional detection stations, as in [32], but also the simulated travel time estimates from Bluetooth antennas along the corresponding paths.

The problem is then formulated as:

$$\left(d_1^* \ldots d_{n_h}^*\right) = arg \quad min_{(x_1 \ldots x_{n_h}) \geq 0} [f_1\left(x_1 \ldots x_{n_h}, \hat{d}_1 \ldots \hat{d}_{n_h}\right) + f_2\left(v_1 \ldots v_{n_h}, \hat{v}_1 \ldots \hat{v}_{n_h}\right) + f_3\left(s_1 \ldots s_{n_h}, \hat{s}_1 \ldots \hat{s}_{n_h}\right) + f_4\left(t_1 \ldots t_{n_h}, \hat{t}_1 \ldots \hat{t}_{n_h}\right)]$$

where

x_i estimated matrix for departing time interval i, i = 1… n_h

v_i simulated volumes on links ∈ S for departing time interval i, i = 1… n_h

s_i simulated speeds on links ∈ S for departing time interval i, i = 1… n_h

$\hat{d}_i$ seed matrix for departing time interval i, i = 1… n_h

$\hat{v}_i$ traffic volumes on links ∈ S for departing time interval i, i = 1… n_h

$\hat{s}_i$ measured speeds on links ∈ S for departing time interval i, i = 1… n_h

The three first terms in the objective function, as in [32] represents the “distance” between observed and simulated flows and speeds and the “distance” between the seed matrix and the resulting estimated demand. Additionally a new term has been incorporated in the objective function. This term refers to the travel times between pair of Bluetooth sensors in predefined paths. Therefore, the objective function now has four different terms and the SPSA gradient calculations are modified accordingly. The sensor layout and the most likely used paths between them have been defined on basis to the procedures already used in Barcelo et al. [41]. 50 Bluetooth sensors have been additionally located in Vitoria’s network and the Aimsun microscopic model has been modified accordingly to include the sensors, emulate the Bluetooth detection and retrieve the partial paths travel times as depicted in Figure 2 (a).

From the Aimsun model one can extract the complete underlying graph of the urban network including all turnings and their associated penalties. Running the Aimsun model for an estimated OD matrix $\hat{\mathbf{g}}(\mathbf{t_i})$ for a given time interval $\mathbf{t_i}$ one can generate a Data Base with the estimated link travel times for that time interval. The link travel times, the graph of the urban network, the detection layout and the defined paths between pairs of Bluetooth antennas are the data to calculate the measured path travel times **tt** which will be one of the input data sets to the Aimsun.m function. The logical diagram of the modified testing process is depicted in Figure 2(b).

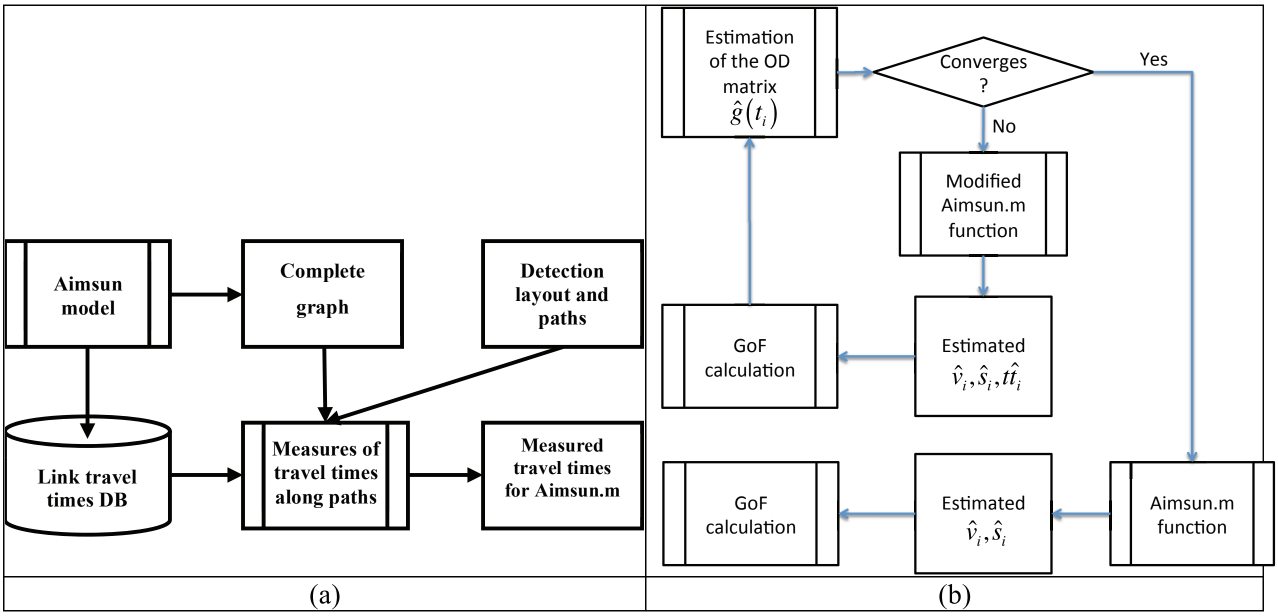


FIGURE 2: Framework extensions needed for ICT sensor data: (a) Additional intermediate step in the bi-level procedure to estimate path travel times between pairs of Bluetooth antennas, (b) Modified testing process.

Another least-square modeling approach for solving the OD estimation and prediction problem proposed in [24] is also considered. The linear state-space formulation where the state variables are the deviations between historical and actual OD flows is solved by efficient LSQR algorithm [52] for large-scale real time applications.

In Djukic et al. [33], a linear state-space formulation has been developed for real-time estimation and prediction of high dimensional OD demand matrices. The approach exploits the idea of dimensionality reduction and approximation of OD demand based on principal component analysis (PCA) [53] to linearly transform the high dimensional OD matrices into the lower dimensional space, where a new transformed set of variables represents the OD demand. These new variables are used as state variables in a novel reduced state space model formulation that are then updated on-line from traffic counts for real time estimation of OD demand. The state space model is solved recursively using the so-called colored noise Kalman filter [54].

EXAMPLE RESULTS

This section provides some indicative results obtained from the use of this framework, aimed at demonstrating the feasibility and flexibility of this approach. In particular, besides an application of the unmodified framework (outlined in Figure 1), a case study that incorporates Bluetooth data is also presented. As this requires custom functionality, an extension and modification of the developed framework is applied (outlined in the previous section and shown in Figure 2). Both case studies are implemented on the (medium-size) Vitoria road network.

A case using the unmodified framework

Case study results

The SPSA AD PI algorithm has been applied considering a seed matrix lower than the real one, with a total number of about 43,000 trips. Excluding the first time slice, which is used as warm-up and therefore kept fixed, the total amount of trips for the remaining intervals is equal to about 32.600 trips. Due to a large number of OD pairs with very low flow, only OD pairs with more than 5 trips have been considered as variable for the adjustment procedure, thus lowering the number of variable components from 12.996 to 918; while these represent only 7% of the total number of OD pairs, they account for nearly the entire amount of starting demand (32.500 trips, 99% of total demand).

The parameters adopted for the algorithm are: 1) a value of the step to compute the approximated gradient equal to 0.23; 2) a number of gradient replications, to compute the average gradient, equal to the 5% of the variable components; 3) only the current gradient is considered, i.e. no information of the past iterations is taken into account. All these choices derive from the sensitivity analysis conducted by Cipriani *et al.* [49] on the parameters of the first order SPSA AD PI.

Preliminary results demonstrate that the algorithm works mainly on the small ODs (lower than 20 veh/15 min): specifically, at the end of the optimization the OD variation is greater than 20% for pairs with values up to 10 veh/15 min (reaching nonetheless OD variation of also 32%), greater than 15% for pairs with values up to 12 veh/15 min and greater than 10% for pairs with values up to 20 veh/15 min. Moreover, among these ODs that vary, 55% of them are moving towards the real matrix (along the “right direction”), with a maximum of 60% for low ODs (with values up to 10 veh/15 min): this is a promising result considering that the average approximated gradient is computed with only the 7% of the number of variables.

A case using an extended framework

The on-line and short-term prediction approach developed at UPC is based on a Kalman filter. State variables are OD flows in the set of most likely used paths (MLU paths) according to a dynamic user equilibrium (DUE) using the available historic OD matrix. Observations of equipped vehicles according to the detection layout provide counts and travel times between BT antennas. The approach has been tested successfully by simulation in corridors and medium-size networks [39,40].

The Bilevel-DUE off-line approach is designed to feed the on-line/short-term forecasting (KFX2) with a reliable seed matrix provides promising results. In this section, we present results from both models, starting from the same historical OD flows in both cases. After analyzing the behavior of different goodness-of-fit measures described above, NME (Normalized Mean Error) seems to be the most suitable to be used in the objective function definition. The objective function is composed of three terms: flows, densities and travel times. NME in each component (between simulated and observed data) returns a very similar value, so the weight of each part is considered equal to one.

As detailed in Cipriani et al. [32] the choice of the gain sequences (a_k and c_k) is critical to the performance of SPSA. Before running the experiments, a scan between the seed and real matrices is realized. The goal of this scan is to understand the evolution (or sensitivity) of the objective function’s value as the input ranges from the “seed” matrix to the “true” matrix. After studying the results and the descent direction, a suitable value was chosen for the gain sequence c_k, which is used for the gradient update. The value of the gain sequence a_k , which is used for the solution update, was chosen depending on the average of the calculated gradient approximations. The dimension of the gradient is related to the statistical goodness of fit measure used in the objective function (in this case NME).

Figure 3(a) depicts the results of the first twenty iterations of the Bilevel-DUE algorithm for demand level D7. In these first twenty iterations, the objective function (as a whole, but also each individual component) shows approximately a 25% reduction. The number of trips in the estimated matrix is very close to the real case (estimated trips: 47677, actual trips: 48642). Figure 3(b) shows a “45-degree” plot, indicating that the estimated OD flows are very close to the true OD flows. Ideally, all points would fall in the solid line. In this figure, the points are very close to the line and well divided around it, indicating a lack of bias. The fit is also quantified by the R^2 of 94%.

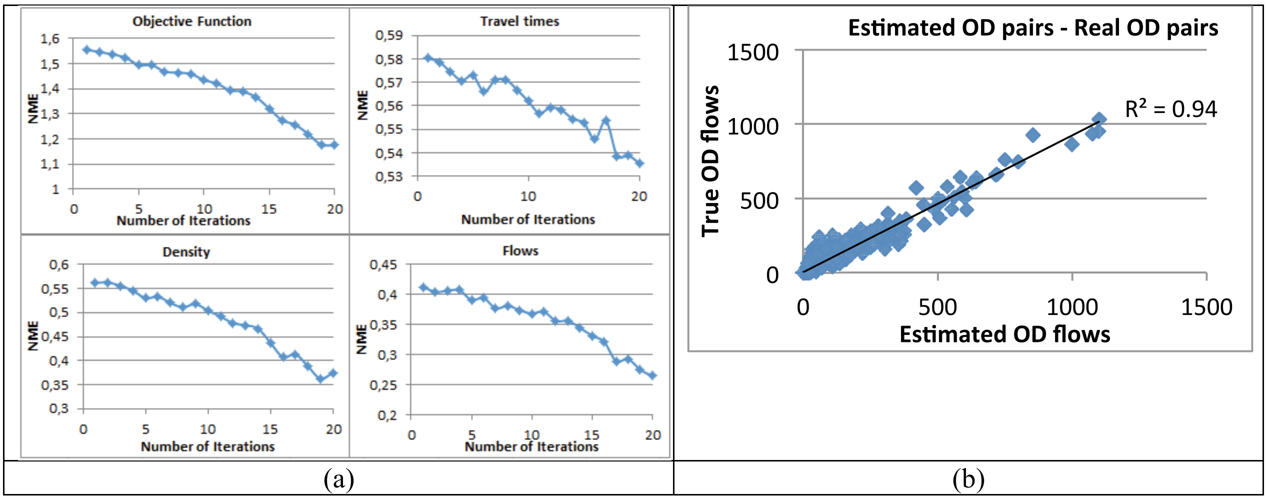


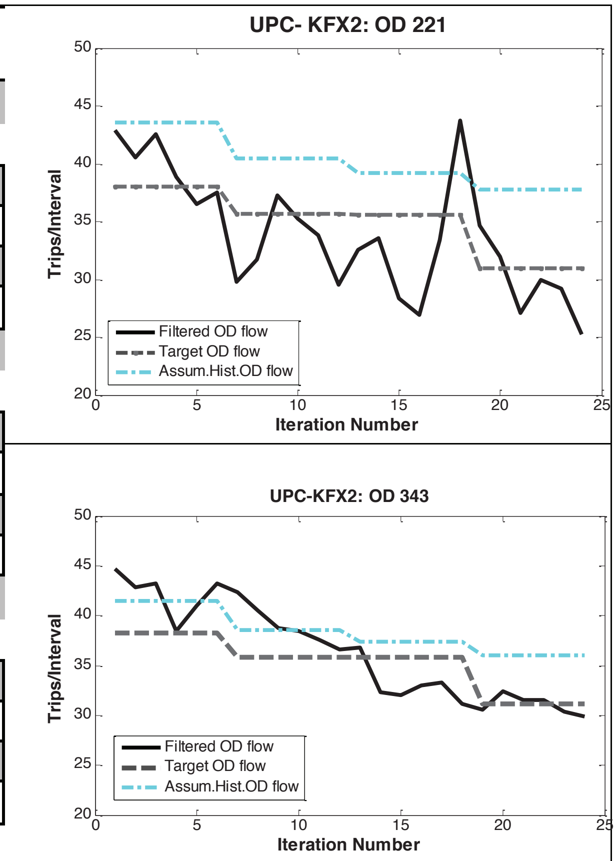
FIGURE 3. Bilevel-DUE results: (a) Evolution of the objective function´s value and its components: flows, densities and travel times, (b) Global OD Fit – Demand level D7.

The remainder of this section describes the application of the on-line KFX2 for the Vitoria network assuming 100% equipped vehicles. Figure 4(a) presents results for several statistics (Normalized Root Mean Squared Error, Theil's U coefficient and R^2) when assuming as an initial point the three different demand levels (D7 through D9) and the Original Demand. An excellent R^2 fit (above 85%) is obtained for the overall OD pairs and demand levels, but mostly for the most important OD flows (i.e. those in the 4^{th} and 3^{rd} quantiles). Figure 4(b) presents the evolution of the OD flow throughout subintervals for two of the most important OD flows when initialized with demand level D9. This figure demonstrates the capability for recovering from an initial point showing greater flow (than the target flow).

The fit of "true" versus estimated OD flows for all considered OD pairs (for the aggregated 1 hour period) for a scenario initialized with demand D7 and 100% BT equipped vehicles is also considered. The coefficient of determination of the simple regression line is almost 90%, i.e. lower than the value achieved by the Bilevel-DUE (off-line proposal presented before, which led to an R2 of 95%), but the computational burden decreases from hours (in Bilevel-DUE) to minutes in KFX2.

Initial Matrix	UPC - KFX2 - 100% BT Equipped Vehicles –Average Theil's Coefficient, RMSEN and R^2 for quartile groups of OD pairs according to Target OD flows				
	Theil's Coefficient				
	Q1	Q2	Q3	Q4	all
Target	0.14	0.13	0.12	0.15	0.13
D7 (100%)	0.16	0.15	0.14	0.16	0.15
D8 (100%)	0.15	0.13	0.12	0.15	0.14
D9 (100%)	0.15	0.13	0.13	0.15	0.14
	RMSEN				
	Q1	Q2	Q3	Q4	all
Target	37.9%	34.8%	23.4%	25.7%	30.5%
D7 (100%)	45.1%	39.6%	27.4%	28.3%	35.1%
D8 (100%)	40.8%	36.3%	24.3%	26.5%	31.9%
D9 (100%)	39.1%	36.0%	24.4%	26.5%	31.5%
	R^2				
	Q1	Q2	Q3	Q4	all
Target	26.3%	17.8%	75.5%	74.8%	87.2%
D7 (100%)	23.4%	15.9%	68.1%	77.4%	88.5%
D8 (100%)	25.3%	17.0%	74.5%	76.5%	88.0%
D9 (100%)	26.8%	18.0%	75.6%	72.8%	86.3%

(a)



(b)

FIGURE 4. KFX2 results: (a) Some Performance Indicators, (b) Filtered values throughout 1h. Assumed Initial Matrix D9

CONCLUSION AND DISCUSSION

In this paper, we motivate and present a flexible platform for the implementation, validation and comparison of different OD estimation and prediction algorithms. The structure of the platform is presented, along with an experimental design that has been developed within the European COST Action MULTITUDE. A long list of algorithms that are currently implemented and being applied using this platform are presented and preliminary results of three algorithms are presented. 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 KFX2) 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).

Particular attention is given to the issue of selecting the appropriate objective function to determine the fit of the algorithms. It is well-known that different measures of performance may give different support to specific conclusions and as such, both individual measures are considered (e.g. flows, densities, speeds) and also compound measure that are incorporate all of these. Furthermore, the issue of selection of the OD pairs on which to focus the attention of the algorithms is discussed. Essentially, very small OD flows are susceptible to high volatility and as such two approaches are considered: first, very small OD flows are held constant to their original values and, second, the estimation results are analyzed in subgroups (quantiles) of the OD pairs.

From an algorithmic point of view, the Linear Kalman Approach called KFX2 in this paper in which the non-linearity to model flow dynamics and estimate travel time between detectors has been replaced by travel time measurement provided by ICT sensors has proven to be computationally efficient both in terms of the quality of results and the computational effort required to achieve the desired convergence, paving the path to real time applications. However, the computational experiments showed that this convergence could strongly depend on the quality of the target matrix, therefore our research attention was driven to find sound initial estimates for each time slice. Taking into account the superior performance of bilevel procedures with respect to other mathematical programming approaches to adjust OD matrices from measurements of traffic variables it was quite natural to investigate how this could be implemented if the usual assignment problem of the lower level was replaced by a dynamic user equilibrium assignment to account for the demand variability. The extended bilevel approach based on DUE combined with he SPSA to solve the upper level optimization has also proved to achieve good results when in the extension the available travel time between detectors is included. The next step will be combining both procedures in such a way that the bilevel approach provides a sound, efficient initialization to the KFX2 procedure.

ACKNOWLEDGEMENTS

Research contained within this paper benefited from participation in EU COST Action TU0903 MULTITUDE – Methods and tools for supporting the Use caLibration and validaTIon of Traffic simUlation moDEls (www.multitude-project.eu).

REFERENCES

(1) Barceló, J. (Ed.). (2010). *Fundamentals of traffic simulation* (Vol. 145). Springer.
(2) Cascetta, E. (2009). Transportation systems analysis: models and application, 2nd Edition, Springer.
(3) Ortuzar, J. And L. Willumsen (2011). Modelling transport. 4th edition. Wiley ed.
(4) Simonelli F., V. Marzano, A. Papola, I. Vitiello (2012). A network sensor location procedure accounting for o-d matrix estimate variability. Transportation Research part B, vol. 46, p. 1624-1638.
(5) Van Zuylen, H.J. and L. G. Willumsen, "The most likely trip matrix estimated from traffic counts," Transportation Research Part B: Methodological, vol. 14, pp. 281-293, 1980.
(6) Bell, M.G.H. (1983). The estimation of origin-destination matrix from traffic counts. Transportation Science 10, 198-217.
(7) Cascetta, E. and S. Nguyen (1988). A unified framework for estimating or updating origin/destination matrices from traffic counts. Transportation Research 22B, 437–455.
(8) Cascetta, E. (1984). Estimation of trip matrices from traffic counts and survey data: a generalized least squares estimator. Transportation Research 18B, 289–299.
(9) Maher, M. (1983). Inferences on trip matrices from observations on link volumes: a Bayesian statistical approach. Transportation Research 20B, 435–447.
(10) 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.
(11) Yang, H. (1995). Heuristic algorithms for the bi-level origin-destination matrix estimation problem. Transportation Research 29B, 231-242.
(12) Cascetta E. and Postorino M.N. (2001). Fixed point models for the estimation of O-D matrices using traffic counts on congested networks. Transportation Science 35(3).
(13) 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.
(14) Vardi, Y. (1996). Network tomography: estimating source–destination traffic intensities from link data. Journal of the American Statistical Association 91, 365–377.
(15) 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.
(16) Hazelton, M. (2003). Some comments on origin–destination matrix estimation. Transportation Research 37A, 811–822.
(17) Cascetta E., Inaudi D. and G. Marquis (1993). Dynamic Estimators of Origin-Destination Matrices using Traffic Counts. Transportation Science 27, 363-373.

(18) Cascetta, E., Papola, A., Marzano, V., Simonelli, F. and I. Vitiello (2013). Quasi-dynamic estimation of o-d flows from traffic counts: formulation, statistical validation and performance analysis on real data. Transportation Research Part B, doi 10.1016/j.trb.2013.06.007.
(19) 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
(20) Okutani I. and Y. Stephanedes (1984). Dynamic Prediction of Traffic Volume through Kalman Filtering Theory. Transportation Research B, 18(2).
(21) 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.
(22) 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.
(23) Ashok, K. (1996). Estimation and prediction of time-dependent origin-destination flows. PhD Thesis, Massachusetts Institute of Technology, Cambridge, MA
(24) Bierlaire, M., and Crittin, F. (2004). An efficient algorithm for real-time estimation and prediction of dynamic OD tables. *Operations Research*, *52*(1), 116-127.
(25) Gelb A. (1974). Applied Optimal Estimation. M.I.T. Press.
(26) Balakrishna, R., Koutsopoulos, H. N. and M. Ben-Akiva (2005). Calibration and validation of dynamic traffic assignment systems. Proceedings of 16th ISTTT, 407-426.
(27) Cremer, M. and H. Keller (1981). Dynamic identification of o-d flows from traffic counts at complex intersections. Proceedings of 8th ISTTT.
(28) Ashok K. and Ben-Akiva M. (2002). Estimation and Prediction of Time-Dependent Origin-destination Flows with a Stochastic Mapping to Path Flows and Link Flows. Transportation Science 36, 184-198.
(29) Cremer, M. and H. Keller (1984). A systems dynamics approach to the estimation of entry and exit o-d flows. Proceedings of 9th ISTTT.
(30) Liu S.S., and J.D. Fricker. (1996). Estimation of a trip table and the theta parameter in a stochastic network. Transportation Research 30 A, 287-305.
(31) 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.
(32) Cipriani E., Florian M., Mahut M. And M. Nigro (2011). A gradient approximation approach for adjusting temporal origin–destination matrices. Transportation Research C, 19(3), 270-282
(33) Djukic, T., G. Flotterod, H. van Lint, and S. Hoogendoorn, Efficient real time OD matrix estimation based on Principal Component Analysis. Proceedings of the IEEE Intelligent Transportation Systems Conference, 2012, Vol. 2, pp. 115-121.
(34) 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.
(35) Barcelò J., Montero L., Marqués L. and C. Carmona (2011). ICT based estimation of time dependent o-d matrices. Proceedings of the 90th TRB, Washington D.C.
(36) Zhang H.M., Nie Y. and Z. Qian (2011). Estimating Time-Dependent Freeway Origin–Destination Demands with Different Data Coverage. Proceedings of the 90th TRB, Washington D.C.
(37) TSS-Transport Simulation Systems (2013). Aimsun 7 Dynamic Simulator User's Manual, TSS, Barcelona, Spain.
(38) Mathworks Inc. 2010. www.mathworks.com
(39) Barcelò J., Montero L., Bullejos M., Serch O. and C. Carmona (2013). A Kalman Filter Approach for Exploiting Bluetooth Traffic Data When Estimating Time-Dependent OD Matrices. JITS Journal of Intelligent Transport Systems 17(2): 1-19 .
(40) Barceló, J., L.Montero, M.Bullejos, M.P. Linares, O. Serch (2013), Robustness and Computational Efficiency of a Kalman Filter Estimator of Time-Dependent OD Matrices Exploiting ICT Traffic Measurements. Paper 13-3919, accepted for publication in TRR Transportation Research Records: Journal of the Transportation Research Board.

(41) Barceló, J., F. Gilliéron, M.P. Linares, O. Serch, L.Montero (2012), Exploring Link Covering and Node Covering Formulations of Detection Layout Problem. Transportation Research Records: Journal of the Transportation Research Board, No. 2308, pp.17-26

(42) Kalman, R.E. “A new approach to linear filtering and prediction problems,” J. Basic Eng. (ASME), vol. 82D, pp. 35–45, 1960.

(43) Antoniou, C., M. Ben-Akiva, and H. N. Koutsopoulos (2007). Non–linear Kalman Filtering Algorithms for On–line Calibration of Dynamic Traffic Assignment Models. IEEE Transactions on Intelligent Transportation Systems, Volume 8, Issue 4, Page(s): 661 - 670, Dec. 2007.

(44) Antoniou, C., A. Kondyli, G.M. Lykogianni and E. Gyftodimos (2013). Exploratory assessment of the Limiting Extended Kalman filter properties. Transport and Telecommunication, 14(1), pp. 1-12.

(45) Spall, J.C. “Multivariate Stochastic Approximation Using a Simultaneous Perturbation Gradient Approximation”. *IEEE Transactions on Automatic Control*, Vol. 37, pp. 332-341. 1992

(46) Spall, J.C. “Implementation of the Simultaneous Perturbation Algorithm for Stochastic Optimization”. *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 34, pp. 817-823. 1998

(47) Antoniou, C., H. N. Koutsopoulos, and G. Yannis (2007). An efficient non-linear Kalman filtering algorithm using simultaneous perturbation and applications in traffic estimation and prediction. Proceedings of the IEEE Intelligent Transportation Systems Conference, Seattle, USA.

(48) Cipriani E., Florian M., Mahut M., Nigro M. (2010). Investigating the Efficiency of a Gradient Approximation Approach for Solution of Dynamic Demand Estimation Problem. In: Chris M.J. Tampere;Francesco Viti;Lambertus H. (Ben) Immers. New Developments in Transport Planning: Advances in Dynamic Transport Assignment (Transport Economics, Management and Policy Series). Edward Elgar Publishing, ISBN: 1848449631

(49) Cipriani E., Gemma A., Nigro M. (2013). A bi-level gradient approximation method for dynamic traffic demand estimation: sensitivity analysis and adaptive approach. Proceedings of the IEEE Conference on Intelligent Transportation Systems, 16th IEEE ITSC, 2013

(50) Spall, J.C. “Adaptive Stochastic Approximation by the Simultaneous Perturbation Method”. *IEEE Transactions on Automatic Control*, Vol. 45, pp. 1839-1853. 2000

(51) Spall, J.C. “Introduction to Stochastic Search and Optimization. Estimation, Simulation and Control”. Wiley-Interscience. Hoboken, New Jersey. 2003

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

(53) Jolliffe, I. T. (2002). Principal Component Analysis, Springer.

(54) Bryson A.E. Jr, & Henrikson L.J. (1968) Estimation using sampled data containing sequentially correlated noise, Journal of Spacecraft and Rockets, 5(6), 662-665.

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

(56) 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.