ICT BASED ESTIMATION OF TIME-DEPENDENT ORIGIN-DESTINATION MATRICES

Authors:

J.Barceló, L.Montero, L. Marqués, and C. Carmona

Department of Statistics and Operations Research and CENIT
(Center for Innovation in Transport)
Technical University of Catalonia
Campus Nord, Building K2M, Office 306
Jordi Girona 29
08034 Barcelona
Spain

(jaume.barcelo, lidia.montero, laura.marques, carlos.carmona)@upc.edu
Tel: +34 93 405 4659

Paper submitted for presentation and publication to
90th Transportation Research Board 2011 Annual Meeting
Washington, D.C.
July 2010

WORDS: 7293

ABSTRACT

Time-Dependent Origin-Destination (OD) matrices are a key input to Dynamic Traffic Models, microscopic and mesoscopic traffic simulators are relevant examples of such models, traditionally used to assist in the design and evaluation of Traffic Management and Information Systems (ATMS/ATIS). Dynamic traffic models are also starting to be used to support real-time traffic management decisions. The typical approaches to the time-dependent OD estimation have been based either on ad hoc heuristics using mathematical programming approaches, or on Kalman-Filtering. The advent of the new Information and Communication Technologies (ICT), as for example Automatic Vehicle Location, License Plate Recognition, detection of mobile devices, Vehicle to Infrastructure (V2I) and so on, makes available new types of traffic data of higher quality and accuracy allowing for new modeling hypothesis leading to more computationally efficient algorithms. This paper extends the previous research on Kalman Filtering approaches for Freeway OD estimation using these data, to more complex topologies of urban networks were alternative path choices between origins and destinations are available. Ad hoc procedures based on Kalman Filtering have been designed and implemented successfully and the numerical results of the computational experiments are presented and discussed.

Keywords: Time-Dependent Origin Destination Matrices, Estimation, Prediction, Kalman Filter, ICT, ATIS, ATMS

INTRODUCTION

The relevance of the estimation of Time-Dependent OD matrices

Time-Dependent Origin-Destination (OD) matrices are a key input to Dynamic Traffic Models, microscopic and mesoscopic traffic simulators are relevant examples of such models. Florian et al. (1), propose a conceptual scheme on the logics of most those dynamic models based on Dynamic Traffic Assignment (DTA) or Dynamic User Equilibrium (DUE) approaches.The logic diagram highlights that all these computational schemes assume that the main input is a Time-Dependent Origin-Destination matrix, modeling the time variability of traffic demand, whose consequences on traffic behavior will be captured in the model by the time dependent path flow rates and the flow dynamics emulated by the network loading process. The main outputs will be the time dependent path flows, travel times and queue and congestion dynamics. These outputs will support the evaluation and the impact analysis of management policies. A common drawback to all these models is that if the key Time-Dependent OD input inappropriately reproduces the time variability of the demand then, independently of the quality of the modeling approaches, the outputs will not be as good as expected. Therefore a question of crucial importance, both for researchers and practitioners, is how to produce acceptably good estimates of the, so far unobservable, Time-Dependent OD matrices.

This problem has usually been addressed resorting to the formulation of the problem in terms of mathematical programming approaches, especially those based on a bilevel optimization model, which upper level minimizes an objective function measuring the quality of the estimate, while at the lower level link flow estimates are the output of either, a static user equilibrium assignment, (2), (3), (4) or a heuristic based on traffic simulation (5), (6). The objective function is usually defined in terms of a distance between observed and estimated link flow counts on a subset of links in the network and, in some cases, a complementary term measuring the distance between an a priory OD matrix and the adjusted OD matrix.

Kalman Filter: a modeling approach that captures time dependencies and inherent randomness of traffic phenomena

Other researchers have tried to capture the time dynamics of the traffic system formulating the problem in terms of Kalman Filtering approaches, (7). Kalman Filter can be considered as a State Space Model Approach, to estimate the dynamics of a system whose state at each instant k in time is defined by the values of a set of unobserved state variables, represented by a vector $x(k)\in\Re^p$ (where p is the number of state variables). The system state transitions evolve in time governed by the stochastic linear difference equation:

$$x(k) = \Phi x(k-1) + w(k) \qquad (1)$$

where Φ is the transition matrix and w(k) represents the process noise, assumed to be white, Gaussian, with zero mean and covariance matrix Q. The system is observed at time k with measurements $y(k)\in\Re^q$ (where q is the number of observations) related to the state by the linear measurement equation:

$$y(k) = Ax(k) + v(k) \qquad (2)$$

with a measurements noise v(k) also assumed to be white, Gaussian, with zero mean and covariance matrix R. Process and measurement noises are assumed to be independent with covariance matrices Q and R which may change at every step. The discrete Kalman Filter cycles recursively between a time update, which projects the current state and covariance estimates ahead in time, from time step k-1 to time step k, to provide an a priori estimate:

$$\begin{aligned} \hat{x}^-(k) &= \Phi\hat{x}(k-1) \\ P_k^- &= \Phi P_{k-1}\Phi^T + Q \end{aligned} \qquad (3)$$

and a measurement update, that adjusts the projected estimate by the available measurements at that time. The measurement update starts by computing the Kalman gain G_k, and then generates an a posteriori estimate by incorporating the measurements y(k) at that time step and calculating the a posteriori error covariance estimate:

$$\begin{aligned} G_k &= P_k^- A^T\left(AP_k^- A^T + R\right)^{-1} \\ \hat{x}(k) &= \hat{x}^-(k) + G_k\left[y(k) - A\hat{x}^-(k)\right] \\ P_k &= \left(I - G_k A\right)P_k^- \end{aligned} \qquad (4)$$

Recursive Kalman Filter approaches to estimate time-dependent OD based on traffic counts were proposed by Nihan and Davis (8) for the case of intersections, using the OD proportions between an entry and all possible destination ramps as state variables, and the exit flows at off ramps for each time interval as observation variables. The relationship between the state variables and the observations includes a linear transformation, where the numbers of departures from entries during time interval k are explicitly considered. Sensors are assumed in all origins and destinations and provide time-varying traffic counts. OD travel times are considered negligible and non-negativity constraints on the proportions and row sums equal to 1 are imposed.

Similar models were proposed by other researchers to estimate time-dependent OD in freeway corridors, networks in which path choices are irrelevant. Van Der Zijpp and Hamerslag (9) proposed a space-state model assuming for each OD pair a fixed and non negligible OD travel time distribution, and time-varying OD proportions (between one entry and all possible destination ramps) as state variables, The main section flow counts for each time interval are the observation variables, no exit ramp counts are available and the relationship between the state variables and the observations includes a linear transformation that explicitly accounts for the number of departures from each entry during time interval k, and a constant indicator matrix detailing OD pairs intercepted by each section detector. Suggestions for dealing with structural constraints on state variables were proposed.

Chang and Wu (10) proposed a space-state model considering for each OD pair a non fixed OD travel time estimated from time-varying traffic measures and including implicitly traffic flow models in the state variables which are are time-varying OD proportions and fractions of OD trips that arrive at each off-ramp m interval after their entrance at interval k. The observation variables are main section and off-ramp counts for each interval and the relationship between the state variables and the observations is complex and nonlinear. An Extended Kalman-filter approach is proposed. Work et al (11) propose the use of an Ensemble Kalman Filtering approach as a data assimilation algorithm for a new highway velocity model proposal based on traffic data from GPS enabled mobile devices.

Hu et al (12), (13) propose an Extended Kalman Filtering algorithm for the on-line estimation dynamic OD matrices incorporating time-varying model parameters provided by simulation, or included as state variables in the model formulation. The approach takes into account temporal issues of traffic dispersion. Lin and Chang (14) proposed an extension of Chang and Wu (10) to deal with traffic dynamics assuming travel time information is available.

Other researchers address the problem of the estimation of time-dependent OD matrices for urban networks, a more complex problem given the existence of alternative paths between each OD pair making that route choice becomes relevant. In a seminal paper Ashok and Ben-Akiva (15) propose a Kalman Filter formulation in which the state variables are the deviations of the OD flows with respect to a priori historical OD matrices, the Kalman Filter is modeled as an autoregressive process that models the temporal relationships among deviations in OD flows. The measurements are the link flow counts in a subset of links in the network where detectors are located. An additional input is the assignment matrix which describes the mapping between OD flows and link flows. This assignment matrix is provided off-line by the DynaMIT supply simulator (16). The autoregressive process is characterized by a set of coefficients that capture the effect of the deviations during one time interval, on the deviations during another subsequent time interval. These coefficients are estimated off-line using a linear regression model for each time interval. The matrix form of the transition equation in terms of the autoregressive process is:

$$g_{k+1} - \hat{g}_{k+1} = \sum_{j=k-r}^{k} f_k^j \left(g_j - \hat{g}_j\right) + w_k \qquad (5)$$

g_k is the vector of flows departing each OD pair at time interval k, $\hat{g}_k$ is the corresponding historical estimate f_k^j is the matrix of effects of $\left(g_j - \hat{g}_j\right)$ on $\left(g_k - \hat{g}_k\right)$ and r is the order of the autoregressive process, that is the number of past intervals influencing the current one, w_k is the usual Gaussian white noise as in equation (1). In a similar way the measurements equation is stated as follows:

$$\hat{v}_k - v_k = \sum_{j=k-p'}^{k} a_k^j \left(g_j - \hat{g}_j\right) + \xi_k \qquad (6)$$

Where v_k is the vector of the link flows measured at time interval k, $\hat{v}_k$ are the link flows obtained from the assignment of the historical OD onto the network, p' is the number of time intervals corresponding to the longest trip, a_k^j is the assignment matrix of contributions of g_j to v_k, and ξ_k is, as in equation (2) a Gaussian white noise. Revised versions of this model can be found in (17) and (18) in the context of calibration of dynamic traffic models. Similar approaches can also be found in (19) and (20).

Drawbacks of current approaches and how ICT traffic data provide the ground for more efficient formulations

All these models for the time-dependent OD estimation share in common various modeling hypothesis:

- Only link flow counts are measured
- Travel times between origins and destinations and between origins and detector locations are assumed to be constant in the simpler cases, or are estimated on basis to certain models, either based on stochastic assumptions about flow propagation as in Chang and Wu (10) and Lin and Chang (14), or on explicit traffic flow models. These approaches imply that speeds are implicit state variables estimated by additional models with the corresponding increase in complexity and computational burden. Estimations add an additional error factor that is not always captured by the filter.

- In the case of networks, where alternative paths are available, the effects of the influence depend on an off line correlation analysis based on historical data, and predetermined assignment matrices that are independent of traffic congestion. This could also add a distortion effect depending on the type and quality of the assignment (i.e. DTA versus DUE)

Taking into account the limitations and lack of quality of measurements provided by the traditional technologies such as loop detectors, we decided to investigate what could be achieved if measurements provided by the new ICT technologies were also available. Thinking of those ICT technologies that are currently available, with different degrees of penetration, or that will become available in short according to most of the technological forecasts (as for example Automatic Vehicle Location (AVL), License Plate Recognition (LPR), detection of Bluetooth mobile devices onboard vehicles, Electronic Toll Collection (ETC also known as TAG systems) or the forthcoming Vehicle to Infrastructure (V2I) systems). These technologies may provide two classes of data: primary data, which can be considered almost a standard, and complementary data still object of controversy. Among the complementary data candidates, depending on the technology, could be the speed, origin and destination, route, and so on. The primary data on which we can relay are: the identity of the device (not necessarily the vehicle in all cases), the position at which the device is detected and the detection time. The basic principles on how these technologies operate are depicted in Figure 1:

A mobile device is captured by one of the Road Side Units (RSU) or sensors implementing the corresponding technology, i.e. Bluetooth of Wi-Fi detection, at a given position, RSU-5 in Figure 1 at time t_1 and later on it is captured again downstream by other RSU, RSU-6 and RSU-10 in the example, at times t_2 and t_3 respectively, allowing to estimate directly the travel times between RSU locations. In the example the travel time between RU-5 and RSU6, $\tau_{5,6} = t_2 - t_1$.

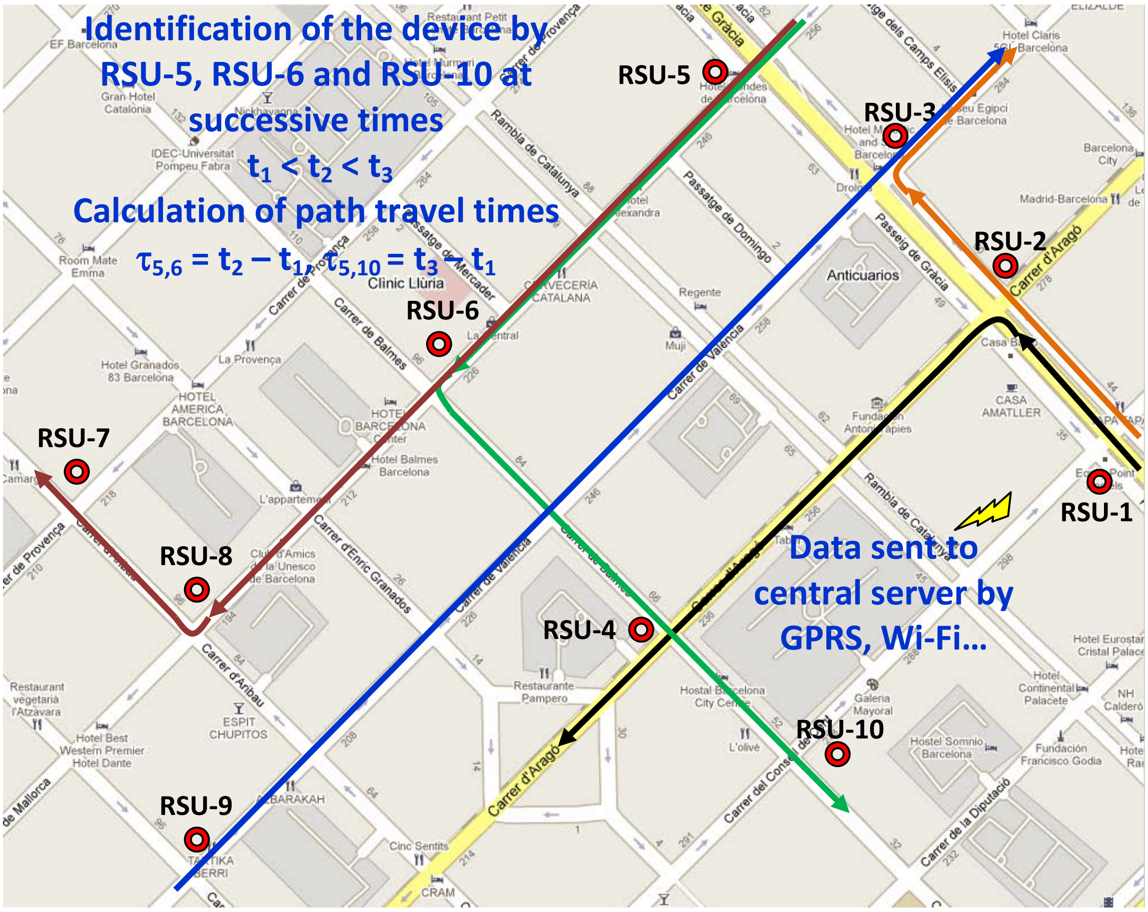

Figure 1: Vehicle monitoring with ICT based sensors

In (21) we explored the possibility of exploiting this new data for the time-dependent OD estimate in freeway corridors. We proposed a space-state formulation for dynamic OD matrix estimation in corridors, considering congestion, combining elements of the Chang and Wu (10), Hu et al. (13) and Van Der Zijpp and Hamerslag (9) proposals. A recursive linear Kalman-Filter approach for state variable estimation was implemented. Tracking of the vehicles is undertaken by processing Bluetoooth and WiFi signals whose sensors are located as described above. Traffic counts for every sensor and OD travel time from each entry ramp to the other sensors (main section and ramps) are available for any selected time interval length higher than 1 second. Travel time delays between OD pairs or between each entry and sensor locations are directly provided by the detection layout and are no longer state variables but measurements, which simplify the approach and make it more reliable. A basic hypothesis, that requires a statistical contrast for test site applications, is that equipped and non-equipped vehicles are assumed to follow common OD patterns; we assume that this hypothesis holds true in the computational experiments.

Although the results obtained improved those of the previous references, in uncongested as well as in congested conditions, we found some drawbacks with the initialization and the updating of the covariance matrices as a consequence of using proportions as state variables. In (22) the model was reformulated with an ad hoc version of time dependencies derived from Lin and Chang (14) and two new formulations using OD flows and OD deviates as state variables as in (17) and (18). The computations experience showed substantial improvements in the results, robustness with respect to initializations and no drawbacks wit covariance matrices.

This paper explores the reformulation of the Kalman Filter approaches in (21) and (22) extending it to the case of urban networks where alternative paths are available and route choice is relevant.

FORMULATION OF THE EXTENDED MODEL FOR NETWORKS

As a consequence of the experience gained in (21) and (22) we propose a formulation of the Kalman Filtering approach that uses deviations of OD flows as state variables as in (17) and (18), model the time-varying dependencies between measurements and state variables adapting the Lin and Chang approach (14) but replacing estimates by sampling experiments that use the tracking of vehicles made available by the ICT technologies (In particular detection of Bluetooth mobile devices on board vehicles, the technology available in our test). According to Ben Akiva *et al* (17) and Antoniou *et al* (18) formulations where state variables are defined as deviates of OD flows with respect to best historical values present several benefits with respect to the use OD flows as state variables:

- OD flows have skewed distributions, but Kalman Filter theory is developed for normal variables and thus symmetric distributions. Deviates from historical values would have more symmetric distributions and thus fit better approximations to normality.
- OD flow deviates from the best historical values allow incorporating more historical data in the model formulation *as a priori* structural information.
- According to our experience in (21) and (22) for the dynamic OD flow estimation on corridors, KF iterations at the filtering stage, require an optimization step to satisfy the non negativity constraints that must be imposed to OD flow proportions , in which the step size α becomes 0 very often to prevent creating unfeasibility on state variables; non-negativity constraints on state variables become critical in the evolution of the KF estimates as far as they not always fit the normality of state variables requested by KF hypothesis. Reformulation in which the state variables are deviates from historical values fits better a scheme closer to normality.
- And last, but not least, a higher performance in the prediction process seems a good value for Advanced Traffic Management Systems (ATMS). Historical data from the previous type of day will be easily available.

The approach assumes a time horizon split in M+1 time intervals of equal length Δt, with M the maximum number of time intervals required by vehicles to traverse the entire network considering a high congestion scenario, and that sensor data are available for equipped vehicles at all time intervals 1 to k for any time interval within the time horizon. The solution provides estimations of the OD matrices for each time interval up to the k-th interval. If historical OD matrices are not available then the formulation reduces to the case when the state variables are the OD flows instead of their deviations. Figure 2 depicts the assumed sensoring layout in which flow counting detectors (i.e. loop detectors) configure a cordon around the network, measuring entry flows from origins for all origins, and a cordon plus some interior sensors (as Sensor q in the Figure) for ICT sensors.

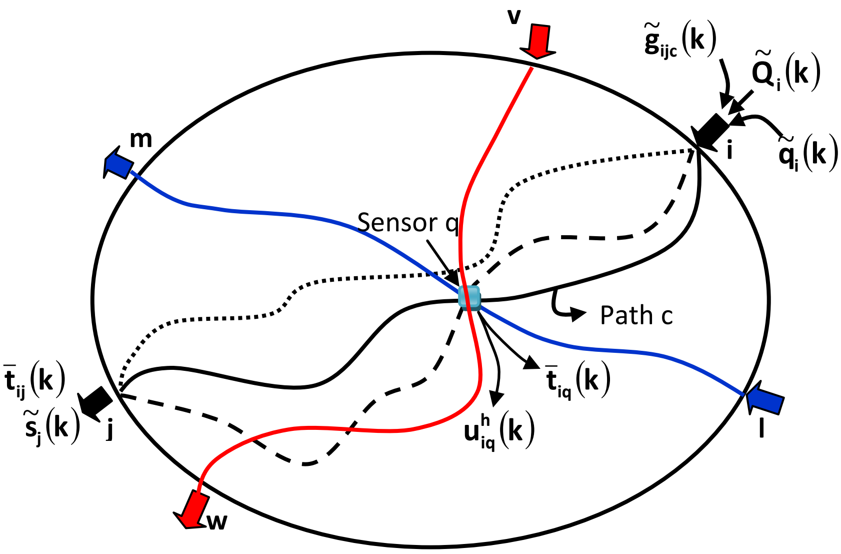

Figure 2: Assumed sensor layout

The total number of flow origins is I, origin centroids are identified by index *i*, *i* = 1,..,I; the total number of destination centroids J, destinations are identified by index *j*, *j* = 1,…,J; the total number of ICT sensors is Q, ICT sensors are identified by index *q*, *q* = 1,…Q, where Q = I+J+P, I, ICT sensors located at origins, J, ICT sensors at destinations and P, ICT sensors located in the inner network; and the total number of likely used paths between origins and destinations is K. The notation used in this paper is the following:

$\tilde{Q}_i(k)=\tilde{Q}_i$: Historic number of vehicles entering the network from centroid *i* at time interval k

$\tilde{q}_i(k)=\tilde{q}_i$: Historic number of equipped vehicles entering the network from centroid *i* at time interval k

$Q_i(k)$: Number of vehicles entering the network from centroid *i* at time interval k

$q_i(k)$: Number of equipped vehicles entering the network from centroid *i* at time interval k

$\tilde{s}_j(k)=\tilde{s}_j$: Historic number of equipped vehicles leaving the network at centroid *j* at time interval k

$s_j(k)=s_j$: Number of equipped vehicles leaving the network at centroid *j* at time interval k

$\tilde{y}_q(k)=\tilde{y}_q$: Historic number of equipped vehicles crossing main section sensor *q* and at time interval k

$y_q(k)=y_q$: Number of equipped vehicles crossing main section sensor *q* and at time interval k

$G_{ijc}(k)$: Number of vehicles entering the network at centroid *i* during interval *k* with destination to centroid *j* using path *c*.

$\tilde{g}_{ijc}(k)=\tilde{g}_{ijc}$: Historic flow of equipped vehicles entering the network at centroid *i* at time interval k that are headed towards destination *j* using path *c*. *Paths are assumed to be the output of a DUE defined on paths, and the number of used paths for each OD pair (i,j)is limited to a maximum of* $K_{ij}^{\max}$. *Proportions assigned to active paths constitute an output of the linear KF proposal.*

$\Delta g_{ijc}(k)$	:	Deviate of equipped vehicles entering the network at centroid *i* during interval *k* that are headed towards centroid *j* (no intrazonal trips are considered) using path *c* with respect to historic flow $\Delta g_{ijc}(k) = g_{ijc}(k) - \tilde{g}_{ijc}(k)$.
$\tilde{z}(k)$	:	The *historic observation variables* during interval k, a column vector of dimension J+P+I, whose structure is $\tilde{z}(k)^T = (\tilde{s}(k) \quad \tilde{y}(k) \quad \tilde{q}(k))^T$
$z(k)$	:	The current *observation variables* during interval k, a column vector of dimension J+P+I, whose structure is $z(k)^T = (s(k) \quad y(k) \quad q(k))^T$.
IJ	:	Number of feasible OD pairs depending on the zoning system defined in the network. This is the maximum number of IxJ
IJK	:	Number of considered OD path flows on the zoning system defined in the network and the level of congestion. This is the maximum number of IxJx K_{max}
$\bar{t}_{ij}(k)$	:	Average measured travel time for equipped vehicles entering at centroid *i* and leaving at centroid *j* during interval *k*
$\bar{t}_{iq}(k)$	:	Average measured travel time for equipped vehicles entering at centroid *i* and crossing sensor *q* during interval *k*
$u_{iq}^{h}(k)$	:	Fraction of vehicles that require *h* time intervals to reach sensor *q* at time interval k that entered the system at centroid i (during time interval $[(k-h-1)\Delta t \quad , \quad (k-h)\Delta t])$.
$u_{ijcq}^{h}(k)$	:	Fraction of equipped vehicles, according time-varying model parameters (updated from measured travel times of equipped vehicles), that during interval k are detected since their trip from centroid *i* to sensor *q* using path c takes *h* time intervals, where $i = 1,\ldots,I$, $j = 1,\ldots,J$, $h = 1\ldots M$, $q = 1\ldots Q$

The state variables are OD path flow deviates from historic OD path flows for each time interval k of length *Δt*; and the number of equipped vehicles entering the network at centroid *i* during interval *k* with destination centroid *j* using path *c*. The model also assumes that

- For equipped vehicles, the entry point to the network and the time entering the network are known and thus $q_i(k)$ is also known. Internal trips are assumed to enter at the centroid representing the zone where it has been detected for the first time. Interzonal trips are considered as measurement noise.
- Conservation equations from entry points are explicitly considered.
- When the total number of vehicles entering the study area is available for each time interval k, $Q_i(k)$ then $Q_i(k)/q_i(k)$ can be considered an estimate of the expansion factor and $G_{ijc}(k)$ can be directly estimated from (non-deviated) state variables $g_{ijc}(k)$, assuming that the expansion factor from equipped to population is common for all OD path flows sharing the same entry point. Without $Q_i(k)$, a generic expansion factor has to be applied to $g_{ijc}(k)$ to get $G_{ijc}(k)$. For realistic applications $Q_i(k)$ are available at gates, not for internal zones.

Let $\Delta g(k)$ be a column vector of dimension IJK containing the state variables $\Delta g_{ijc}(k)$ for each time interval k for all feasible OD pair paths (i,j,c) ordered by OD pair. The state variables $\Delta g_{ijc}(k)$ are assumed to be stochastic in nature and OD path flows deviates at current time k are related to the OD path flow deviates of previous time intervals by an autoregressive model of order r <<M:

$$\Delta g(k+1) = \sum_{w=1}^{r} \mathbf{D}(w) \Delta g(k-w+1) + \mathbf{W}(k) \qquad (7)$$

Where the $w_{ijc}(k)$'s are assumed to be independent Gaussian white noise sequence with zero mean and covariance matrix $\mathbf{W_k}$, and $\mathbf{D}(w)$ are IJKxIJK transition matrices which describe the effects of previous OD path flow deviates $\Delta g_{ijc}(k-w+1)$ on current flows $\Delta g_{ijc}(k+1)$ for w = 1,…,r and all feasible OD path flows (i,j,c). In the preliminary results in this paper we assume r=1. The structural constraints that should be satisfied by the state variables are:

$$\Delta g_{ijc}(k) \geq -\tilde{g}_{ijc}(k) \qquad i=1\ldots I, \quad j=1\ldots J \; c=1\ldots K_{ij}^{\mathbf{max}}$$
$$q_i(k)=\sum_{j=1}^{J}\sum_{c=1}^{K_{ij}^{\mathbf{max}}} g_{ijc}(k) \rightarrow \sum_{j=1}^{J}\sum_{c=1}^{K_{ij}^{\mathbf{max}}} \Delta g_{ijc}(k) = q_i(k) - \tilde{q}_i(k) \qquad i=1\ldots I \qquad (8)$$

Equality constraints are explicitly considered in the observation equations through the definition of dummy sums of centroid entries where measurement errors are allowed. The observation equations are counts of vehicles entering the network, leaving the network and detected in sensors (located in access lanes to intersections) for each interval k. The relationship between the state variables and the observations involves a time-varying linear transformation that considers:

- The number of equipped vehicles entering from each entry centroid during time intervals k, k-1, k-M, $q_i(k)$.
- $H<M$ time-varying model parameters in form of *fraction matrices*, $\left[U_{ijcq}^{h}(k)\right]$.

Structural constraints should also be satisfied for the time-varying model parameters $u_{ijcq}^{h}(k)$ reflecting temporal traffic dispersion, they account for congestion,

$$u_{ijcq}^{h}(k) \geq 0 \quad i=1\ldots I, \quad j=1\ldots J, c=1\ldots K_{ij}^{\mathbf{max}}, q=1\ldots Q, h=1\ldots H$$
$$\sum_{h=1}^{H} u_{ijcq}^{h}(k)=1 \qquad i=1\ldots I, \quad j=1\ldots J, c=1\ldots K_{ij}^{\mathbf{max}}, q=1\ldots Q, \qquad (9)$$

Structural constraints for $u_{ijcq}^{h}(k)$ have not been explicitly considered in the observation equations since can be guaranteed by the updating process of time-varying model parameters from travel-time data on equipped vehicles at current time k. Since the time varying travel times have to be taken into account to be able to model congestion, then time varying delays from entries to sensor positions are considered and thus entry volumes per centroid for M+1 intervals $k, k-1, \ldots, k-M$. State variables for intervals $k, k-1, \ldots, k-M$ are required to model interactions between time-varying OD patterns on paths, counts on sensors and distribution of travel time delays (traffic dispersion) from entry centroids to sensor positions. Measures provided by ICT sensors are direct samples of travel times that allow the updating of discrete approximations of travel time distributions, making unnecessary to incorporate models for traffic dynamics. This model simplification due to the availability of the new ICT is one of the major novelties in the proposed reformulation of Kalman Filter. However, we must be aware that the final destination j is unknown.

Model parameters to account for temporal traffic dispersion are again fractions u_{iq}^{h}, but in fact they also account for variable traffic conditions and therefore the time-varying model parameters $u_{iq}^{h}(k)$ have to satisfy structural constraints, where $H<M$:

$$u_{iq}^{h}(k) \geq 0 \quad i=1\ldots I, \quad q=1\ldots Q, h=1\ldots H$$
$$\sum_{h=1}^{H} u_{iq}^{h}(k)=1 \qquad i=1\ldots I, \quad ,q=1\ldots Q, \qquad (10)$$

As in Lin and Chang (14) we assume that the travel time T_{iq} from origin i to sensor q is a random variable that depends on the time evolution of traffic conditions, with a non stationary probability density function $t_{iq}(t)$. This probability distribution can be approximated by $T_{iq}(k)$, the discrete travel time distribution for vehicles reaching sensor *q* at time interval *k* that entered the network from centroid *i*, its density function $t_{iq}^{(k)}(t)$ can be approximated in terms of $u_{iq}^{h}(k)$ whose updated from the (assumed random) sample of on-line travel time data of equipped vehicles, as shown in Figure 3. This is again one of the modeling assumptions in which hypothesis on the dynamics of traffic flows is replaced by measures of travel times provided by ICT sensors. A discretization in $H<M$ time intervals

will be initially assumed, but it is still an open question if a (i,q) (centroid, sensor) dependent horizon is more suitable.

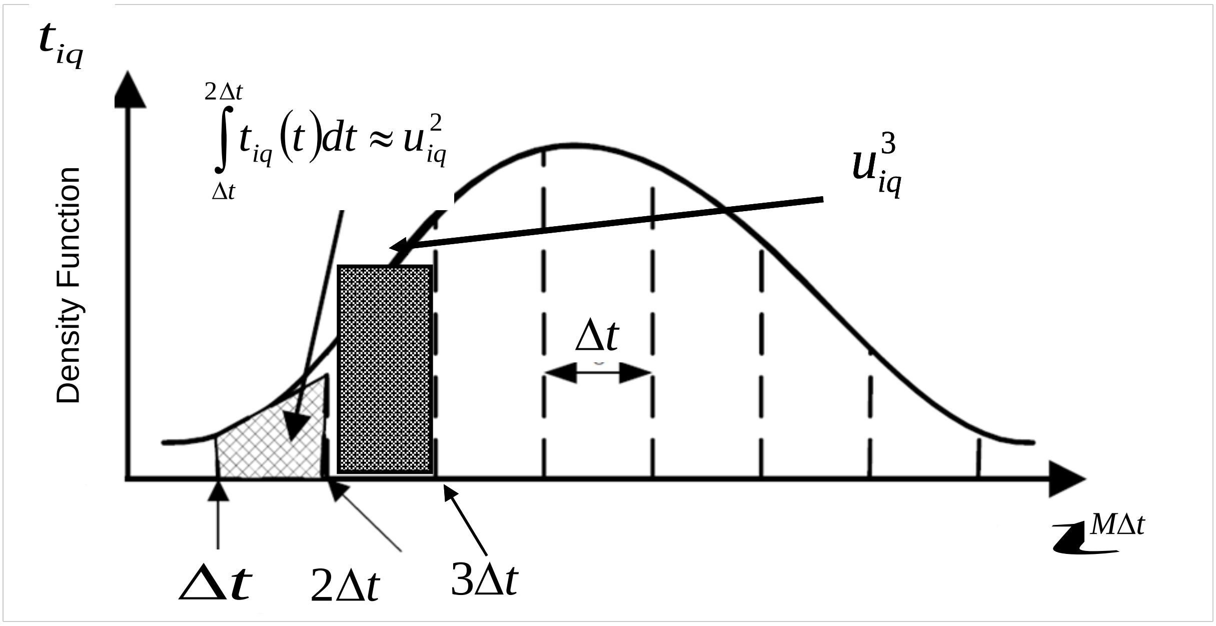

Figure 3: Approximation to the discrete travel time density function t_{iq} in terms of the sampled u_{iq}^{h}

At each time interval k, the average travel time $\bar{t}_{iq}(k)$ experienced by the equipped vehicles is available and the discrete travel time distributions can be be updated.

To complete the model formulation we define the auxiliary matrices:

E	:	Matrix of row dimension I containing 0 for columns related to state variables in time intervals $k-1,\ldots,k-M$ and B for time interval k .
B	:	Matrix of dimension I x IJK defining equality constraints (sum to 1 in OD path proportions for each entry) for state variable in time interval *k*.
U(*k*)	:	Matrix of dimensions (1+M)IJK x (1+M)Q consisting on diagonal matrices $U(k),\ldots,U(k-M)$ containing $u_{ijcq}^{h}(k)$. For $U(k-h)$ is a matrix of dimensions IJKxQ containing the estimated proportion of equipped vehicles whose travel-time from entry *i* to a given sensor *q* takes *h* intervals for vehicles captured by the *q* sensor at time interval *k*. Average measured time – varying travel times are critical and the clue for taking into account congestion effects.
A	:	Matrix of dimensions Q x (1+M)Q that adds up for a given sensor *q* (regular section or exit connector) traffic flows from any feasible entry centroid included in the most likely used paths (according to historic DTA or DUE) arriving to sensor at interval *k* assuming their travel times are $t_{iq}(k)$

The measurement equations are defined in terms of the vector z(k) of state variable defined formerly. At time interval k, their values are determined by those of the state variables at time intervals k, k-1, …k-M, where M is the maximum number of intervals necessary to cross the network.

$$\Delta z(k) = \begin{pmatrix} s(k)-\tilde{s}(k) \\ y(k)-\tilde{y}(k) \\ q(k)-\tilde{q}(k) \end{pmatrix} = \begin{pmatrix} \mathbf{AU(k)}^{\mathbf{T}} \\ \mathbf{E} \end{pmatrix} \Delta\mathbf{g(k)} + \begin{pmatrix} v_1(k) \\ v_2(k) \\ v_3(k) \end{pmatrix} = \begin{pmatrix} \mathbf{F_1\,(k)} \\ \mathbf{F_2\,(k)} \\ \mathbf{F_3\,(k)} \end{pmatrix} \Delta\mathbf{g(k)} + \begin{pmatrix} v_1(k) \\ v_2(k) \\ v_3(k) \end{pmatrix} = \mathbf{F(k)}\Delta\mathbf{g(k)} + v(k) \quad (11)$$

Where v_i, i = 1,2,3 are respectively white Gaussian noises with covariance matrices R_i. **F(k)** maps the state vector $\Delta\mathbf{g(k)}$ onto the current blocks of measurements at time interval k: counts of equipped vehicles on exit points, link sensors and entry centroids, accounting for time lags and congestion effects. Tilde counts at time interval k mean the observed counts when the historical demand $\tilde{g}_{ijc}(k)=\tilde{g}_{ijc}$ (if available) is assigned to the network given the current traffic conditions (included in the block diagonal matrix $\mathbf{U(k)}$ of dimensions QxM*IJK).

The ad hoc version of the iterative Kalman Filter algorithm described in equation (3) and (4) becomes in this case:

KF Algorithm	:	Let K be the total number of time intervals for estimation purposes and M maximum number of time intervals for larger trip
Initialization	:	k=0; Build constant matrices and vectors: A, B, C, D, E. Inicialize as an identity matrix by an scalar parameter properly tuned $W_{\mathbf{k}}$. Inicialize $R_{\mathbf{k}}$ as a diagonal matrix proportional to the variance of historic measument variables (z(k)). $\Delta\mathbf{g}_{\mathbf{k}}^{\mathbf{k}}=\mathbf{0}\quad \mathbf{P}_{\mathbf{k}}^{\mathbf{k}}=\mathbf{V}[\mathbf{W}(\mathbf{0})]$
Prediction Step	:	$\Delta\mathbf{g}_{\mathbf{k+1}}^{\mathbf{k}}=\mathbf{D}\Delta\,\mathrm{g}_{\mathbf{k}}^{\mathbf{k}}$ and $\mathbf{P}_{\mathbf{k+1}}^{\mathbf{k}}=\mathbf{D}\mathbf{P}_{\mathbf{k}}^{\mathbf{k}}\mathbf{D}^{\mathbf{T}}+\mathbf{W}_{\mathbf{k}}$
Kalman gain computation	:	Get observations of counts and update fractions in travel times bins: $q(k+1),\quad s(k+1),\quad y(k+1),\quad u_{ij}^{h}(k+1)$. Build $\boldsymbol{\Delta}\,\mathrm{z}(\mathrm{k}+1)$ and $\mathbf{U}(\mathbf{k}+\mathbf{1})$. Build $\mathbf{F}_{\mathbf{k+1}}=\mathbf{F}(\mathbf{k}+\mathbf{1})$. Compute $\mathbf{G}_{\mathbf{k+1}}=\mathbf{P}_{\mathbf{k+1}}^{\mathbf{k}}\mathbf{F}_{\mathbf{k+1}}^{\mathbf{T}}\left(\mathbf{F}_{\mathbf{k+1}}\mathbf{P}_{\mathbf{k+1}}^{\mathbf{k}}\mathbf{F}_{\mathbf{k+1}}^{\mathbf{T}}+\mathbf{R}_{\mathbf{k}}\right)^{-}$
Filtering	:	Compute $\mathbf{d}_{\mathbf{k+1}}=\mathbf{G}_{\mathbf{k+1}}\left(\boldsymbol{\Delta}\,\mathrm{z}(\mathrm{k}+1)-\mathbf{F}_{\mathbf{k+1}}\Delta\mathbf{g}_{\mathbf{k+1}}^{\mathbf{k}}\right)$ filter for state variables and errors $\boldsymbol{\varepsilon}_{\mathbf{k+1}}=\left(\boldsymbol{\Delta}\mathrm{z}(k+1)-\mathbf{F}_{\mathbf{k+1}}\Delta\mathbf{g}_{\mathbf{k+1}}^{\mathbf{k}}\right)$. Search maximum step length α such that $\Delta g_{\mathbf{k+1}}^{\mathbf{k+1}}=\Delta g_{\mathbf{k+1}}^{\mathbf{k}}+\alpha\,\mathbf{d}_{\mathbf{k+1}}\geq-\tilde{\mathbf{g}}(\mathbf{k})$ and $\mathbf{P}_{\mathbf{k+1}}^{\mathbf{k+1}}=\left(\mathbf{I}-\mathbf{G}_{\mathbf{k+1}}\mathbf{F}_{\mathbf{k+1}}\right)\mathbf{P}_{\mathbf{k+1}}^{\mathbf{k}}$
Iteration	:	*k*=*k*+1 if *k*=K EXIT otherwise GOTO Prediction Step
Exit	:	*Output results*

The symbol $()^{-}$ refers to to computation of the pseudoinverse of a square matrix.

COMPUTATIONAL EXPERIMENTS

In the preliminary computational experiments conducted to test the KF algorithms we have focused our attention in the quality of the results and not in the computational efficiency. A prototype has been implemented in MATLAB and tested computationally. From a practitioner's point of view, assuming that the MATLAB code is available, the input data to apply the method to a network of interest is the following:

- Network topology: nodes and directed links. The graph of the road network can be obtained exporting the network model, if available, from most of the professional software for Geographic Information Systems or for Transport Planning analysis
- Indicator for nodes that are considered centroids, entry and/or exit points of the network. This is information directly available when the network model comes from professional Transport Planning software.

- List of OD pairs. For each origin centroid a pointer to the vector containing the list of destination centroids. A data structure that can be easily generated when the previous information is available.
- For each OD pair, the considered paths as a list of pointers defining the sequence of links that define each path. The natural order of state variables is defined for (origin centroid, destination centroid, OD path id). This is the procedural step that requires the resort to specific traffic assignment software. In our case we have performed a DUE assignment with Dynameq (23) to generate the set of most likely used paths from which we have generated the corresponding data structures.
- For each sensor (access/leaving centroid or link sensor) the list of origin centroids and the list of OD paths affected by the measure should be easily available. This means, in other words, matching the OD-paths structures to the detection layout.

The detection layout raises some methodological concerns. As we have specified it consists of two components: the cordon component encircling the network with sensoring at input-output gates (as for instance is currently available is most of the urban pricing systems), and the detection layout at the interior of the encircled area, as graphically illustrated in figure 2. From the point of view of the estimation of time-dependent OD matrices the optimal sensor location is a problem strongly related to the observability of the network, namely when a space state approach is used to solve the problem, as it is the case of the Kalman Filter approach. The observability of a system from a state space representation means that for any possible sequence of state and control vectors the current system state can be determined using only the outputs. In other words, that the behavior of the system can be determined from the system outputs. Gentilli and Mirchandani (24) propose generic formulations of the problem in terms of various types of traffic detectors, solutions for the case when the detectors are located at links and measure traffic volumes can be found in (25) and (26). We have slightly modified the algorithm for the set covering formulation of the problem in (26) to find an approximate solution that captures the 100% of the traffic demand, ensuring in this way the complete observability of the system.

After the pilot project in Freeways with Bluetooth sensors reported in (21) and (22) a pilot project has been proposed in a urban network of limited size in the Business Central District of the city of Barcelona, as part of a Research & Development project funded by the Spanish Ministry of Research and Innovation. A simulation experiment has been conducted prior to deploying the technology for the pilot project. The selected site has been the network of the Amara District in the city of San Sebastian in Spain. The network has 232 links, 142 nodes and 85 OD pairs and a rich structure of alternative paths between OD totalizing 358 paths according to the DUE with Dynameq. Figure 4(a) displays a snapshot of the microsimulation model used to emulate the RSU and the Centroids. Figure 4(b) displays, highlighted as red dots, the detection layout of the 48^{th} RSU according with the proposed specifications. Once a vehicle is generated in the simulation model, it is randomly identified as an equipped vehicle depending on the proportion of penetration of the technology, 30% in our case, according to the available information on the penetration of the technology in the Metropolitan Region of Barcelona. The simulation emulates the logging and time stamping of this random sample of equipped vehicles.

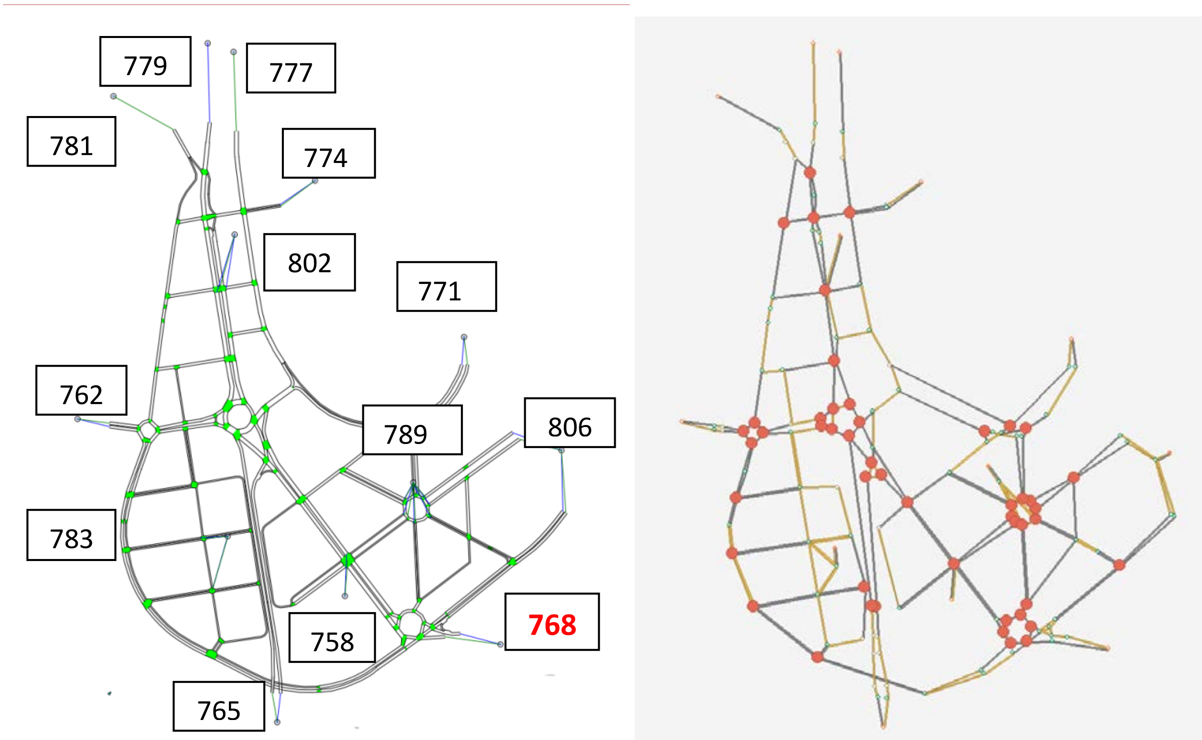

Figure 4: 4(a) Microsimulation Model 4(b) Detection layout

COMPUTATIONAL RESULTS AND PRELIMINARY CONCLUSIONS

The length of the time interval has been set to 90 seconds and the simulation period to 1 hour, that is 40 intervals. The extended state vector is set to M=10 according to the maximum time used to complete an OD trip. The dimension of the extended state vector is 358x(10+1).

In the preparation of the input special attention has been paid to the specification of the data structures defining the lists of emerging links from each node, OD pairs sharing the same origin centroid, OD path identifiers for each OD pair (which will depend on the DUE results), and the lists of OD paths and OD pairs intercepted by each sensor q detecting vehicles arriving from origin centroid i. As mentioned earlier, from a practical point of view this an important cumbersome task, to make it useful to practitioners it would be necessary to develop an application automatically building these data structures form the network topology, the detection layout and the path structure from a DUE. A task that we think is affordable since this information can be automatically extracted from most of the available software, i.e. Dynameq.

A design factor in computational experiments is the number H of bins used to update the discrete approximations of travel time distributions from entry *i* to sensor *q*. This affects the data structures and the updating algorithm. A discretization in H=3 bins appears to be satisfactory, according to our previous experience, for most of the tests.

The parameters of the KF approach, as constants affecting variance-covariance matrices of sensor measures and state variables have been tuned. The behavior consistency according to the previous experience has been checked. Conservation of entry flow from origin centroids has been explicitly considered as a measure equation with near zero variance leading to a pseudo inverse computation in Kalman gain step. A diagonal variance-covariance matrix has been initially tested, but a multinomial variance-covariance performs better.

The preliminary tests have been conducted implementing the filter code in Matlab, but the computing performance shows that, for a networks of the size of Amara, it is almost at the limits of what Matlab can support and, therefore, new versions implemented with more appropriate programming tools will be developed in the future to efficiently deal with larger networks. Computational tests have been conducted with different initializations for state variables and historic data:

- An initial OD pattern stable for the simulation horizon (OD proportions from origin to destination centroids of OD pairs) as the true OD pattern (used in the simulation), with input flows from origin centroids for each interval increased/decreased in percentages for an hour 15, 25, 30 and 30% and distribution of OD path flows according to:

    1. Proportional distribution of OD trips to OD available paths.
    2. Allocation of OD trips to only 1 of the available OD paths for each OD pair.

- An initial OD pattern stable for the simulation horizon (OD proportions from origin to destination centroids of OD pairs) as a non informative OD pattern (not used in the simulation) with input flows from origin centroids for each interval increased/decreased in percentages for an hour 15, 25, 30 and 30% and distribution of OD path flows according to:
    1. Proportional distribution of OD trips in OD available paths.
    2. Allocation of OD trips to only 1 of the available OD paths for each OD pair.

Figure 5 summarizes the values of the RMSE for OD flows centroid 768 to all at the end of the process for the computational experiment with **v**ariable OD flows in 4 time slices. Empirical discrete travel time distributions approximated with H=3 are considered. In order to check the robustness of the new approach we have assumed the worst case initialisation with equiprobabilities in the OD pattern from every entry centroid to all destinations and in the OD path selection for all OD pairs .

Els 2 gràfics per 4 slices…..,

		OD pairs from origin centroid 768 to all – Fix OD Pattern – 4 Sliced OD flows											
		758	**762**	**765**	**768**	**771**	**774**	**777**	**781**	**783**	**789**	**802**	**806**
Interval length 90 sec	**Target OD Pattern (%)**	7.0%	1.3 %	19.7 %	0.0%	2.9%	0.5%	2.3%	0.0%	41.6 %	21.6 %	1.9%	1.1%
	RMSE ODflows												
OD flows: Target veh per slice	**Slice 1: 15 min**	10	2	29	0	4	1	3	0	61	32	3	2
	Slice 2: 15 min	17	3	48	0	7	1	6	0	102	53	5	3
	Slice 3: 15 min	21	4	58	0	8	2	7	0	122	64	6	3
	Slice 4: 15 min	21	4	58	0	8	2	7	0	122	64	6	3

Figure 5: OD Pairs from centroid 768 to all (in table). In graphics, filtered OD flows (left) and (right) Target OD flows (continuous line) and historic OD flows (discontinuous line) for OD pairs 768-765 (blue), 768-783 (green) and 768-789 (red). Y Scale are OD flows per interval (90seg). X axis is the iteration number.

Computational experiments show that tuning of KF parameters is directly related to the historic OD flows considered: each initialization requires a specific tuning of the parameters but this not always guarantees the convergence. This is related to the structure of the variance-covariance matrices capturing the relations amongst OD path flows. To overcome these convergence problems more elaborated relationships have to be developed.

A complementary set of experiments accounts for the effect of historic flows in the convergence to the true values of OD flows in similar conditions. Convergence to target OD flows (in black) is improved depending the considered historic entry flows and tuning of the parameters has to be considered for practical purposes (Figure 6).

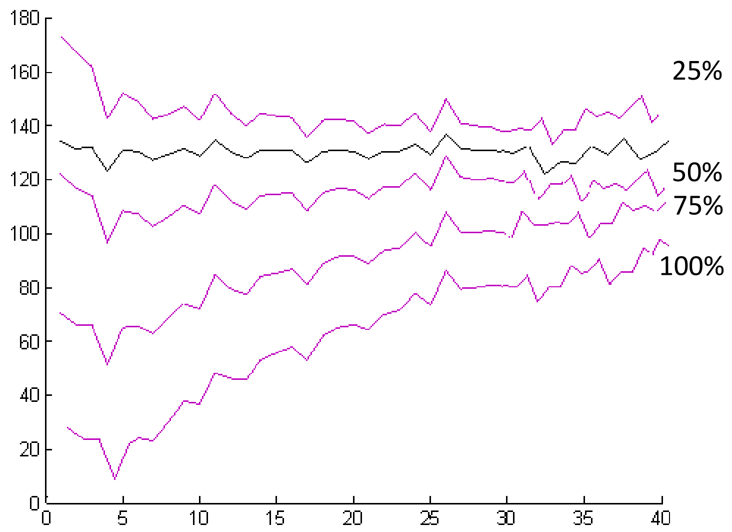

Figure 6: Effect of historic data in the convergence for 1 Slice entry flows and constant OD pattern without congestion (time horizon 1h). Discrete travel time distributions in H=3 bins.

Els 2 gràfics per 1 slice, no sé si podem augmenter en 1 gràfics el paper, sinó suggereixo suprimir l'actual Figure 6 i deixar la 7 (que passaria a numerar-se com a 6).

		OD pairs from origin centroid 768 to all – Fix OD Pattern – 1 Sliced constant OD flows											
		758	**762**	**765**	**768**	**771**	**774**	**777**	**781**	**783**	**789**	**802**	**806**
Interval length 90 sec	**RMSE ODflows**												
	Target OD flows per interval	2,3	0,5	6,4	0,0	1,0	0,2	0,8	0,0	13,6	7,1	0,6	0,4
	Average Travel times (min)	2.2	4.4	1.2	0.0	3.8	6.1	4.4	0.0	0.6	0.8	4.9	2.2

Figure 7: OD Pairs from centroid 768 to all. Convergence for 1 Slice entry flows and constant OD pattern without congestion (time horizon 1h). Discrete travel time distributions in H=3 bins. In graphics, filtered OD flows (left) and (right) Target OD flows (continuous line) and historic OD flows (discontinuous line) for OD pairs 768-765 (blue), 768-783 (green) and 768-789 (red). Y Scale are OD flows per interval (90seg). X axis is the iteration number.

REFERENCES

(1) Florian, M., Mahut, M. and Tremblay, N., (2001), A Hybrid Optimization-Mesoscopic Simulation Dynamic Traffic Assignment Model, Proceedings of the 2001 IEEE Intelligent Transport Systems Conference, Oakland, pp. 118-123.

(2) Spiess H., (1990), A Gradient Approach for the O/D Matrix Adjustment Problem, Publication No. 693, Centre de Recherche sur les Transports, Université de Montréal.

(3) Florian M. and Chen Y., (1995), A Coordinate Descent Method for the Bi-level O/D Matrix Adjustment Problem, International Transactions in Operations Research, Vol. 2, No. 2, pp. 165-175.

(4) Codina E., Barceló,J. (2004) Adjustment of O-D matrices from observed volumes: an algorithmic approach based on conjugate gradients, European Journal of Operations Research, Vol. 155, pp. 535-557.

(5) Choi, K., Jayakrishnan, R., Kim, H., Yang, I., Lee, J., (2009), Dynamic OD Estimation using Dynamic Traffic Simulation Model in an Urban Arterial Corridor, Paper presented at the 88th TRB Meeting, Washington D.C.

(6) Jintanakul, K., Jayakrishnan, R., Chu, L., (2009), Dynamic Demand Input Preparation for Microscopic Traffic Simulation Analysis under Planning Applications, Paper presented at the 88th TRB Meeting, Washington D.C.

(7) Kalman, R.E. (1960). A new approach to linear filtering and prediction problem. *Journal of Basic Engineering* **82** (1): 35–45.

(8) Nihan, N.L. and Davis, G.A., (1987). Recursive estimation of Origin. In: *Destination matrices from Input/Output counts. Transportation Research-B* **21B**, pp. 149–163 (2) .

(9) Van Der Zijpp, N.J. and Hamerslag, R., (1994). An improved Kalman filtering approach to estimate Origin—Destination matrices for freeway corridors. In: *Transportation Research Records 1443*, National Research Council, pp. 54–64.

(10) Chang, G.L. and Wu, J., (1994). Recursive estimation of time-varying Origin—Destination flows from traffic counts in freeway corridors. Transportation Research B Volume 28B No.2 pp. 141-160

(11) Work, D., Toossavainen, O-P. Blandin, S., Bayen, A, Iwuchukwu and Tracton, T. (2008), An Ensemble Kalman Filtering Approach to Highway Traffic Estimation Using GPS Enabled Mobile Devices, Proceedings of the 47th IEEE Conference on Decision and Control, pp 5062-5068, Cancun, Mexico, Dec. 9-11, 2008, DOI: 10.1109/CDC.2008.4739016

(12) Hu, S., Chen, C. (2004) Dynamic Estimation of Freeway Origin-Destination Demand and Travel Time Using using Extended Kalman filtering, Proceedings of the 2004 IEEE IC on Networking, Sensing and Control-Taiwan.

(13) Hu, S., Madanat S.M., Krogmeier J., Peeta S. (2001) Estimation of Dynamic Assignment Matrices and OD Demands using Adaptative Kalman filtering, Intelligent Transportation Systems Journal vol. 6, pp 281-300.

(14) Lin, P., Chang (2007), G. A generalized model and solution algorithm for estimation of the dynamic freeway origin-destination matrix. Transportation Research B 41 554-572.

(15) Ashok, K. and Ben-Akiva, M., (1993), Dynamic origin-destination matrix estimation for real-time traffic management systems. In Transportation and Traffic Theory, C.F. Daganzo (Ed.), Elsevier Science Publishers, Oxford, UK, pp. 465-484

(16) Ben-Akiva, M., Bierlaire, M., Burton, D., Kotsopoulos, H.N. Mishalani, R., (2001), Network State Estimation and Prediction for Real-Time Traffic Management, Networks and Spatial Economics, 1, pp. 293-318.

(17) Ashok, K. and Ben-Akiva, M., (2000) Alternative Approaches for Real-Time Estimation and Prediction of Time-Dependent Origin-Destination Flows, Transportation Science 34(1), 21-36.

(18) Antoniou, C., Ben-Akiva, M. and Koutsopoulos H.N., (2007), Nonlinear Kalman Filtering Algorithms for On-Line Calibration of Dynamic Traffic Assignment Models, IEEE Transactions on Intelligent Transportation Systems Vol. 8 No.4.

(19) Zhou, X., Mahmassani, H.S., (2005). Recursive approaches for online consistency checking and O–D demand updating for real-time dynamic traffic assignment operation. Journal of the Transportation Research Board 1923, 218–226.

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

(21) Barceló J., Montero L., Marqués L. and Carmona C., (2010), Travel time forecasting and dynamic OD estimation in freeways based on Bluetooth traffic monitoring, TRB Paper 10-3123 Accepted for presentation and publication to the 89th Transportation Research Board 2010 Annual Meeting Washington, D.C. To appear in TRR. Barceló J., Montero L., Marqués L. and Carmona C., (2010),

(22) Barceló J., Montero L., Marqués L. and Carmona C., (2010), A Kalman-Filter Approach For Dynamic OD Estimation In Corridors Based On Bluetooth And Wifi Data Collection, Paper number 01387, presented at the 12th World Conference on Transportation Research, Lisbon July 11-15, included in the CD-Rom.

(23) Dynameq, (2010), Version 1.4.6, INRO www.inro.ca.

(24) Gentili M., Mirchandani P. B., (2005), Location of Active Sensor on Traffic Network, Annals of Operations Research, 136, 1, 229-257.

(25) Fei, X., Eisenman, S. M. and Mahmassani, H.,(2007) Sensor Coverage and Location for Real-Time Traffic Prediction in Large-Scale Networks, Paper presented at the 86th Annual Meeting of the Transportation Research Board, Washington D.C..

(26) Gilliéron F. and Barceló J., (2008), Optimization of road traffic counts location in a network, Paper 2728, presented at the ITS 2008, European ITS Conference, Genève, Included in the Conference CD-ROM