Dynamic Demand Estimation And Prediction For Traffic Urban Networks

Stefano Carrese [1], Ernesto Cipriani [1], Marialisa Nigro [1,*]

[1] *Department of Engineering, "Roma Tre" University, Rome, Italy*

* corresponding author: marialisa.nigro@uniroma3.it

Abstract

Availability of accurate trip demand estimates plays a key role for both long term and short term traffic planning. Additionally, on line applications of intelligent management strategies require reliable forecasts of traffic demand so that users' response to varying flow conditions, both observed and communicated, in different locations and in different time intervals can be properly taken into account and anticipated. To guarantee consistency with respect to temporal and spatial dimensions, traffic demand estimates and predictions must reflect both time variability and network patterns. This calls for a problem known in literature as Dynamic Demand Estimation problem (DDEP); it can be formulated as an off-line problem for medium to long term planning and design, or as an on-line problem for real time ATMS/ATIS applications. In both cases, traffic dynamics are dealt with at network level.

Nowadays, demand information derives from advanced traffic surveillance systems that provide updated measurements of several heterogeneous traffic data, both in fixed locations and over specific corridors or paths. Such recent technology developments suggest new promising and challenging chances, not fully addressed yet. The present paper studies some of these opportunities within the DDEP, specifically: data heterogeneity, referring to different nature of data collected on wide spatial coverage that are integrated within the estimation framework, and on line detection of non-recurrent conditions, taken into account adopting a sequential approach for short-term predictions.

Keywords

Traffic modeling, origin-destination (o-d) estimation/prediction, Floating Car Data (FCD), Local Ensemble Transformed Kalman Filter (LETKF).

1. Introduction

Nowadays modern cities are facing social and economic changes that pose several mobility concerns related to limited supply of transportation systems. Such issues cannot be tackled only with planning strategies based on capacity increasing options but have to be integrated, if not replaced, by solutions deriving from the adoption of transport management approaches. In both cases, accurate estimates on travel demand patterns represent the key information for feeding the simulation process required in the evaluation phase.

In transport engineering, estimation of travel demand based on direct observations of the network (e.g. from traffic counts and other traffic state measurements) is a classical and widely adopted procedure, both in long-term (off-line) planning and in real-time (on-line) Advanced Traffic Management and Information Systems applications.

The off-line Dynamic Demand Estimation problem (DDEP) is usually approached as a bi-level problem. The upper-level problem consists of the adjustment of a starting demand (obtained through a combination of surveys and mathematical models) using traffic measurements, which are in turn linked to the dynamic demand. This link is generated from the dynamic traffic network assignment problem at the lower level, solved by using a Dynamic Traffic Assignment (DTA) model. It results in a highly undetermined, non-linear, non-convex problem, which was object of a relevant effort in the last years (Antoniou et al, 2016, Cantelmo et al., 2014).

A relevant number of contributions exists exploring the influence on DDEP of all the available sources of information about traffic performances during the DDEP. In fact, the basic information utilized since first works, i.e. traffic counts, is not effective in discerning between congested and uncongested traffic state of a link, because of non-monotone flow-density relationship; thus, adopting count data alone in the DDEP can potentially over fit the demand to counts at the expense of traffic dynamics. Moreover, current technologies can provide a great amount of traffic data: pavement-embedded sensors, roadside radars and cameras provide measures of flows and speeds at nodes and along links; Advanced Vehicle Identification (AVI), ground-based radio navigation, cellular geo-location and GPS provide a new kind of information about travel times and route choices that integrate usual information on traffic flows at count sections.

Floating Car Data (FCD) are one of the possible new data sources. They derived from probe vehicles equipped with on board unit and able to track the position of the same vehicles along the network. One of the main features of FCD, with respect to other data derived by different technologies, is the possibility to obtain information on Origin–Destination (OD) zones of individual trips as well as on their route choices and route travel times. This is fundamental for the calibration of traffic models and even more for the estimation of OD travel demand.

Such quite recent technology suggests new promising and challenging chances for the DDEP, not fully addressed yet, that may improve the reliability of the estimated demand.

The dynamic OD trips resulting of the off-line DDEP are usually the starting point for the on-line demand estimation and prediction. In on-line applications, the dynamic OD estimation process is required to recursively provide fast estimates for recent time slices together with predictions for future time slices. The problem starts with the off-line adjusted dynamic OD matrices; then, such matrices are updated in a sequential manner in order to take into account the real-time variability of traffic conditions. The sequential update consists in fixing the demand components departing in all intervals prior to the one being considered and it is basically adopted if fast run times are required.

Several papers about on-line demand estimation focuses on the adopted solution approach. The Kalman filter (KF) algorithm (Kalman, 1960) has been widely proposed: this algorithm is mainly based on the solution of a least-square cost function in an incremental way, allowing for the update of the OD flows each time additional traffic data are available. However, the application of the KF to the on-line estimation and prediction seems to have different drawbacks: its inability to handle a large number of variables, since it requires intensive linear algebra computations, as well as the linear hypothesis behind the map connecting OD flows and traffic measurements, that cannot even more be accepted if other traffic measurements, highlighting non linearity, are included in the estimation framework.

The present contribution analyses both the off-line and the on-line DDEP in the context of recent technology developments and the related new promising and challenging chances provided by the possibility to collect several heterogeneous traffic data. Firstly, Floating Car Data (FCD) are exploited for off-line demand estimation. Information on user's dynamic route choice behaviour and route travel times have been gathered, coupled with fixed location observations, and included in the DDEP. At the same time, a recent extension of Kalman filter theory, called Local Ensemble Transformed Kalman Filter

(LETKF, Hunt et al 2007), has been experimented for on line prediction, with the objective of reducing the dimension of the problem and its computational effort, as well as to avoid any linearization of the dependency between OD flows and traffic measurements.

The remainder of this paper is structured as follows. Section 2 presents a literature review on the different sources of information usually adopted in the DDEP, deriving their limits or strengths; then an overview of on-line approaches is reported with the aim of underlining the weaknesses due to the linearity assumption, especially if new data sources are adopted, and the needs to overcome the linearity itself. Section 3 deals with the off-line DDEP and the potentialities linked with the adoption of FCD, presenting an application on a real case of study (the district of Eur, Rome,Italy). Section 4 proposes a new method for the on-line DDEP and shows first results obtained. Finally, Section 5 provides concluding discussions and statements.

2. Literature review

2.1 Sources of information for the DDEP

Traffic counts are the most common data adopted in the DDEP. However, they are not effective in discerning between congested and uncongested traffic state of a link, because of non-monotone flow-density relationship, thus generating a potential over fit of the demand to counts at the expense of traffic dynamics. In this context, appropriate selection of traffic measurements for the estimation becomes a key point of investigation in the DDEP literature.
Moreover, current technologies can provide a great amount of traffic data: pavement-embedded sensors, roadside radars and cameras provide measures of flows and speeds at nodes and along links; Advanced Vehicle Identification (AVI), ground-based radio navigation, cellular geo-location and GPS provide a new kind of information about travel times and route choices that integrate usual information on traffic flows at count sections. Examples of applications of different kind of traffic measurements in the DDEP can be found in Balakrishna et al., 2007, Ashok and Ben-Akiva, 2000, Tavana, 2001.

Cipriani et al. (2013) showed a preliminary analysis on the contribution provided by link flows, link speeds and path travel times: results demonstrated the importance of type, quality and quantity of the information, as well as how the best improvements on demand are usually obtained when a sample of path travel time measurements is considered.

Several contributions deal with the adoption of AVI data (Dixon and Rilett, 2002, Antoniou et al. 2006, Zhou and Mahmassani, 2007), bluetooth (BT) and mobile phone data (Sohn and Kim, 2008, Barceló et al., 2013).

The adoption of route choices data available by probe vehicles (Floating Car Data, FCD) is growing interest due to their potential added value in the DDEP as well as in explaining the real choice mechanism of road users: FCD have been used by Ásmundsdóttir (2008), Ásmundsdóttir et al (2010) and Zhao et al (2010) both to derive a-priori matrices and to correct them in the dynamic case. Specifically, in the first two contributes, FCD are used to derive a-priori matrices and to analyse route choices and trip length distributions to be adopted during the dynamic demand estimation by an entropy maximization approach. In the latter, static OD demand has been generated based on Remote Traffic Microwave Sensors data (RTMS), then the time-varying OD demands have been generated using the resulting time-varying splitting rates by a combined application of FCD and RTMS data.

Link travel speeds and link flows derived from FCD have been adopted by Yamamoto et al (2009) and Cao et al (2013) in the dynamic OD matrices estimation solved with a bi-level generalized least squares (GLS) method.

In Morikawa and Miwa (2006) the driver's dynamic route choice behaviour using FCD have been analysed with the aim of providing some fundamental knowledge on modelling dynamic behaviour for a better representation in dynamic assignment model.

The adoption of several traffic measurements inside the DDEP is strictly correlated with the solution method of the problem. In fact, the most common approach, where only traffic counts are usually available, consists in explicitly linking OD flows to traffic counts through the assignment matrix (Frederix et al. 2011a, Toledo and Kolechkina, 2012) which is a linear approximation of the complex relationships between the two set of variables. This linear approximation clearly affects the results as demonstrated by Flötteröd and Bierlaire, 2009 and Frederix et al. (2011a). If other measurements (such as speeds, densities, travel times etc.) come inside the DDEP, approximations similar to the assignment matrix cannot exceedingly be justified, thus assignment-free approaches become the only available option (Cremer and Keller, 1984).

2.2 Background on on-line estimation and prediction

On-line estimation was firstly proposed by Okutani and Stephanades (1984). Subsequently, it has been generalized by Ashok and Ben-Akiva (1993) and Ashok (1996), which recognized the importance of structural information of OD flows inside the process, thus modelling it in terms of deviation of the demand from a previous estimate (historical or off-line estimate). The formulation results in a non-linear "state-space" model composed by the following sets of equations:

1. Transition equations, which follow the evolution of the state variables (OD flows) over time;
2. Measurements equations, mapping the state variables into the traffic measurements;
3. Analysis equations, correcting the estimate of the state variables (derived by the transition equations) with the results of the measurements equations and a Kalman gain.

The non-linearity of the formulation is in the map between OD flows and traffic measurements, required both in the measurements and in the analysis equations.

The Kalman filter (KF) algorithm has been widely proposed as solution approach for the on-line estimation (Ashok and Ben-Akiva, 2000, Chang and Wu, 1994, Van der Zijpp and Hammerslag, 1994). This algorithm is mainly based on the solution of a least-square cost function in an incremental way, allowing for the update of OD flows each time additional traffic data are available. Ashok and Ben-Akiva, 1993, modelled the within-day evolution of deviation of OD flows from historical estimates using a KF based on an autoregressive process. Zhou and Mahmassani in 2007 assumed a polynomial approximation for the deviation as an alternative to the autoregressive process.

However, the application of the KF to the on-line estimation and prediction seems to have different drawbacks. Firstly, the linear hypothesis behind the map connecting OD flows and traffic measurements using the assignment matrix; then, the need to assume the assignment matrix as fixed in the measurement equations in order to simplify the problem. Finally, the inability of KF to handle a large number of variables, since it requires intensive linear algebra computations (Bierlaire and Crittin, 2004).

About the first drawback, since the problem is highly non-linear, several extensions have been proposed in literature. Antoniou et al. (2007) tested the Extended Kalman filter (EKF) and the Unscented Kalman filter (UKF), two specific versions of KF for dealing with non-linear systems: despite the benefits obtained, they resulted to be computationally time demanding even on a small freeway network. This is mainly due to intensive elaborations both request at each time step: for computing first order derivatives of the non-linear function connecting OD flows to traffic measurements in EKF applications, for working with a set of carefully chosen sample points ("sigma points"), representing the mean and covariance of the distribution of the state variables in UKF ones. Thus, the Limiting EKF has been proposed (LimEKF), where the first derivatives to be computed at each time step for the EKF are substituted by a fixed gain matrix, usually computed off-line. With the LimEKF computational times strongly decreases; however, the way this fixed gain matrix is computed could generate excessive approximations and loss of information.

Computational issues for on-line estimation in large road networks were addressed also by Bierlaire and Crittin (2004), which focused on algorithmic enhancements: namely, they proposed the LSQR algorithm by Paige and Saunders (1982), as an alternative to KF to exploit the sparsity of the problem (not all OD flows use all sensors on the network at every departure time interval).

Recently, Marzano et al. 2015 extended a quasi-dynamic (QD) OD flows estimator, based on the assumption of constant OD shares across a reference period, whilst total flows leaving each origin varying for each sub-period within the reference period, to the on-line context. They adopt an EKF as solution approach, dealing with assignment matrices for the computation of first derivatives in the measurements equations.

Other recent research lines for on-line dynamic OD flows estimation and prediction focus on the adoption of several and different set of measurements, not only link data, but also point to point data thus exploiting information by new ICT technologies (Antoniou et al., 2004, Zhang et al 2011, Barceló et al, 2013). As reported before, in such a case any extension to include other traffic measurements with respect to the commonly adopted traffic counts collides even more with possible linearization of the problem.

3. Off-line Dynamic Demand Estimation

This section analyzes the possibility of exploiting new sources of data as FCD for the off-line DDEP. Broadly speaking, if spatial and temporal OD distribution from FCD can be representative of the whole population, they can be incorporated in the different stages of the dynamic demand estimation procedure to restrict or bind the solution space of the problem. Moreover if route choices and route travel times from FCD correctly describe behaviors of road users and the traffic state of the network, it is necessary that they also come into play in the DDEP acting as a control element for the simulation phase.

Specifically in this contribute both the information of route choices and OD travel times from FCD have been considered inside the DDEP. OD travel times can be adopted in addition to measurements on links (f.e. counts and speeds), nodes (f.e. turning movements) and starting information on demand (seed matrix, generations and attractions, distribution share etc.). Already several contributions in literature started to investigate their added value (Cipriani et al, 2013). However, although they represent an indirect measure of user's route choices and they can be easily obtained with the recent developments of ubiquitous monitoring systems, also for this measure it could happen that several configurations of dynamic demand matrices generate the same result in terms of average OD travel times (Cipriani et al, 2015). Route choice probabilities are a much more powerful information respect to OD travel times, since they are uniquely defined. Thus, fixed the route choice model and parameters, they can be inserted in the objective function of the DDEP as a measure of the distance between simulated route choices and monitored route choices. In fact, if there is a difference between monitored route choices (from FCD) and simulated route choices (results of the assignment phase and of the adopted route choice model), the DDEP will try to modify the travel demand until this difference will be minimized. Of course, if the adopted route choice model is far from the real behaviour of road users, the estimated demand will be only a distortion of the real demand, as demonstrated in (Cipriani et al, 2015).

Next subsections show the methodology adopted and the results obtained in an application related to a real road network (the Eur district, Rome, Italy).

3.1 Methodology

Considering different types of traffic measurements and by adopting a simultaneous approach the DDEP can be formulated as:

$$\left(\mathbf{d}_1^*,\ldots,\mathbf{d}_n^*\right) = \arg\min[z_1(\mathbf{l}_1,\ldots,\mathbf{l}_{n'},\hat{\mathbf{l}}_1,\ldots,\hat{\mathbf{l}}_{n'}) + z_2(\mathbf{n}_1,\ldots,\mathbf{n}_{n'},\hat{\mathbf{n}}_1,\ldots,\hat{\mathbf{n}}_{n'}) + z_3(\mathbf{x}_1,\ldots,\mathbf{x}_{n'},\hat{\mathbf{x}}_1,\ldots,\hat{\mathbf{x}}_{n'}) + \\ + z_4(\mathbf{r}_1,\ldots,\mathbf{r}_{n'},\hat{\mathbf{r}}_1,\ldots,\hat{\mathbf{r}}_{n'})] \quad (1)$$

Where

- $\mathbf{l}/\hat{\mathbf{l}}$ are respectively the simulated values and the corresponding measurements on the links;
- $\mathbf{n}/\hat{\mathbf{n}}$ are respectively the simulated values and the corresponding measurements on the nodes;
- $\mathbf{x}/\hat{\mathbf{x}}$ are respectively the estimated value and a-priori information on the dynamic demand (seed matrix);
- $\mathbf{r}/\hat{\mathbf{r}}$ are respectively the simulated values the and corresponding measurements on routes;
- $\mathbf{d}_n^*$ the estimated demand matrix for time interval n;
- z:{z_1,z_2,z_3,z_4} is the estimator represented by the deviations between the simulated/estimated and the corresponding measured/a-priori values.

The dependence between simulated traffic information in (1) and the estimated demand is obtained directly by simulation performing a dynamic traffic assignment (DTA), usually based on a dynamic user equilibrium approach (DTA-DUE); thus the DDEP results in a bi-level optimization problem where the upper level searches for dynamic demand matrices that best represent observed traffic measurements on the network and the lower level searches for the solution of the DTA-DUE problem. As reported in Antoniou et al (2016), sequential calls and integrated input/output exchange between the supply/demand interaction (DTA-DUE) and the solution algorithm of the DDEP is required to solve (1).

In addition, several constraints can be added to (1), both to avoid infeasible solutions and to bind the solution space: for example non negativity constraints or constraints on aggregated demand values (total generated trips, demand distribution shares).

3.2. Adoption of OD travel times and route choice probabilities from FCD in the DDEP

Synthetic experiments have been conducted on a specific district of the city of Rome, the Eur disctrict (54 traffic zones, 400 regular nodes, 812 road links), considering a "real" demand of 45,000 veh/h equally distributed in the whole planning horizon (one hour divided into four time slices of 15 minutes each); this "real" demand derives from a static demand corrected in order to best reproduce the known traffic volumes at the boundary links of the Eur network and uniformly split into four time slices.

Then, a starting demand of 38,000 veh/h (seed OD matrix) has been generated as initial point of the DDEP adopting a 20% of reduction of the "real" demand and perturbing each OD with an error uniformly distributed in the range {-10%;+10%}. Simulated measures for this network are:

- link flows available on 32 count sections as average value for the planning horizon, as mimicking the real traffic counts available by the Rome Mobility Agency in that network;
- 12 monitored ODs (corresponding to the ODs with the highest number of FCD signals, as resulted from the data of the Octo Telematics fleet spanning the area), for which OD travel times and route choice probabilities are collected each 15 minutes departure time interval.

All the measures have been derived by a Dynamic User Equilibrium (DUE) assignment of the "real" demand, performed by DYNASMART (DYnamic Network Assignment Simulation Model for Advanced Road Telematics, Mahmassani et al 2004). The adoption of synthetic data has been done in order to be able to evaluate the results also in terms of estimated demand, not only in terms of traffic measurements reproduction. However, it is important to underline as the settings of the experiments have been treated to best represent the network and the data available.

Accordingly, five sets of traffic data have been adopted in (1) as reported in Table 1. Then, the DDEP has been solved for each set by adopting a simulation approach where the solution algorithm is a modification of the gradient-based path search optimization method (SPSA) coupled with Asymmetric Design (AD) for gradient computation and Polynomial Interpolation (PI) of the objective function for the linear optimization (SPSA AD-PI, Cipriani et al 2011).

Table 1 Traffic data adopted in the off-line DDEP for the different experiments

	Data and prior information			
Set	*Seed OD matrix*	*Link flows*	*OD travel times*	*Route choice probabilities*
Set 1	+	+		
Set 2	+		+	
Set 3	+			+
Set 4	+	+	+	
Set 5	+	+	+	+

3.2.1 Numerical results

At the end of the experiments, the reduction of the objective function (OF) of the DDEP has been equal to 51% adopting link flows (set 1), 70% adopting the average OD travel times (set 2), 31% with the route choice probabilities (set 3), 60% mixing link flows and OD travel times (set 4) and 57% with all the measurements together. These sharp reductions reflect the ability of the procedure to better reproduce the reference measurements (Table 2), in particular the average OD travel times (Figure 1); where the OD travel times are not inside the OF, the reduction is lower (set 3).

Regarding the estimated demand, all the experiments show a proximity to the total real demand value (Table 3).

Table 2 Error on measures reproduction

		Average error [%]				
	Starting point	*Set 1*	*Set 2*	*Set 3*	*Set 4*	*Set 5*
Link flows	8	3	4	7	4	3
Average OD travel times	33	7	5	11	7	7
Route choice probabilities	40	32	27	27	28	28

Table 3 Error on total demand reproduction

	Seed OD matrix	Set 1	Set 2	Set 3	Set 4	Set 5
Total demand [veh/h]	37,874	44,067	43,720	42,905	43,718	43,529
Error respect to the real demand [%]	15	1	2	4	2	2

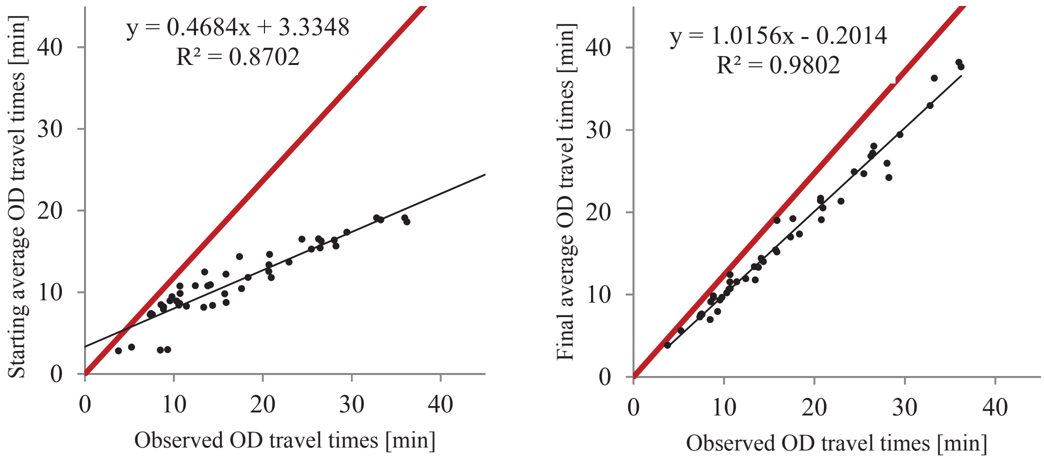

Fig. 1. Reproduction of average OD travel times at the end of the DDEP (Set 2)

The estimated demand has been analyzed also in terms of its spatial and temporal correspondence with the real demand; specifically, the Euclidean distance between each estimated OD i in each time interval t (OD_i^t) and the "real" OD for the same time interval ($\widehat{OD}_i^t$) has been computed ("Spatial and temporal error" in (2) and Table 4), as well as the Euclidean distance between the sum of each estimated OD for all time intervals (total OD value in the whole planning horizon) with the corresponding real value ("Spatial error" in (3) and Table 4). The starting point means the same statistics considering for the computation the seed and the real matrix.

$$\text{Spatial and temporal error} = \sum_t \sum_i \sqrt{\left(OD_i^t - \widehat{OD}_i^t\right)^2} \quad (2)$$

$$\text{Spatial error} = \sum_i \sqrt{\sum_t OD_i^t - \sum_t \widehat{OD}_i^t}^{\,2} \quad (3)$$

The spatial error of the estimated demand respect to the real one has been reduced by the DDEP approximately the same amount with the different measurements/Set. The higher reduction derives from the combination of all the available measurements (Set 5). Instead, the spatial and temporal error due to the distribution of the demand in the different time slices takes on different reductions depending on the adopted measurements/Set: if OD travel times are adopted together with link flows, it is possible to reach a reduction of 26% of the error on spatial and temporal reproduction. Moreover adding also route choice probabilities, the reduction reaches the 39%. Route choice probabilities alone seems to be not suitable in this case, since the reduction is equal to only 6% (Table 4).

However, this result has to be compared with the level of demand covered by the different type of measurements (Table 5).

Table 4 Error on spatial and temporal demand reproduction

	Euclidean distance					
	Starting point	*Set 1*	*Set 2*	*Set 3*	*Set 4*	*Set 5*
Spatial error	806	462	486	500	448	415
Reduction [%]	-	-43	-40	-38	-44	-48
Spatial and temporal error	414	338	350	387	306	251
Reduction [%]	-	-18	-15	-6	-26	-39

The traffic count sections collecting link flows intercept the 67% of the demand, where each vehicle is on average intercepted by two traffic count sections during its trip. Average OD travel times and route choice probabilities are instead related to only 12 ODs covering the 10% of the demand.

If the spatial and temporal investigation on estimated demand is conducted analysing separately the ODs covered by the different monitoring systems respect to the not covered ODs, results are evident (Table 5): route choice probabilities are able to accurately reproduce the monitored ODs. The improvement in the estimated demand is equal to 27% (reduction of the Euclidean distance), comparable with the achievement adopting OD travel times (28%) and higher than using link flows (19%). Of course, this improvement is strictly related to the monitored ODs, since on the not monitored ODs it decreases to 3%. Differently, link flows and OD travel times are able to work better on not monitored ODs since they provide a "global" indication of traffic conditions on the whole network, respect to the "local" indication of route choice probabilities.

Table 5 Error on reproduction of monitored and not monitored ODs

	Starting point	**Set 1**	**Set 2**	**Set 3**	**Set 4**	**Set 5**
Intercepted demand [%]		67	10	10	67	67
Euclidean distance (monitored ODs)	*Set 1, 4 and 5: 374*	304	-	-	281	223
	Set 2 and 3: 167	-	120	121	-	-
Reduction [%]	-	-19	-28	-27	-25	-40
Euclidean distance (not monitored ODs)	*Set 1, 4 and 5: 177*	148	-	-	120	115
	Set 2 and 3: 378	-	329	368	-	-
Reduction [%]	-	-16	-13	-3	-32	-35

Nonetheless, it can be stated that route choice probabilities are full of information respect to classical link flows, against a large difference in terms of coverage, thus making this potential data from FCD particularly suitable for the DDEP.

If OD travel times are added to link flows, improvements on monitored ODs is comparable with the adoption of single measurements at network level (compare set 4 with set 2 and 3); but, if both the information at network level are combined in the OF (set 5), the improvement on monitored ODs reaches the 40%.

Finally, the possibility to combine different information, network data with link data as well as "global" indicators with "local" indicators, remains fundamental to be able to reproduce also not monitored ODs (set 4 and set 5).

4. On-line Dynamic Demand Estimation and Prediction

Given the potentialities associated with the new source of data, as seen in the previous section adopting data from FCD in the off-line case, here a recent extension of KF, the Local Ensemble Transformed Kalman Filter (LETKF, Hunt et al 2007) has been proposed for the on-line estimation and prediction of dynamic OD flows. LETKF has the following main strengths: 1) it starts with an ensemble of the state variables (OD flows). Then, it moves from the space of the state variables to a lower dimension space (space of the ensemble), thus reducing the dimension of the problem and its computational effort; 2) it avoids any linearization of the dependency between OD flows and traffic measurements, which is not considered in an analytic formula, but implicitly captured by traffic simulation; 3) it permits to break down the problem into sub-problems to be solved in parallel fashion.

Next subsections show the mathematical principles of LETKF, providing remarks for its implementation in the case of on-line travel demand estimation and prediction. Then, a simple example to validate the reliability of the LETKF compared with another well-known non-linear KF approach is proposed; finally, numerical tests carried out by applying LETKF on a simple, but not trivial road network, investigate the behavior of the method for several conditions of the input parameters.

4.1 Methodology

LETKF has been proposed in meteorological sciences by Hunt et al 2007 as a "refinement" of previous approaches as the Local Ensemble Kalman Filter (LEKF) and the Ensemble Transform Kalman Filter (ETKF). The aim of LETKF is to deal with non-linear problems, large-scale models and data sets.

As in the basic Ensemble Kalman Filter (EnKF), LETKF deals with non-linear systems approximating the state estimate and its uncertainty at a given time interval by an ensemble of system states. Thus, the ensemble is used to parametrize the distribution of the state variables.

EnKF starts with an ensemble $\mathbf{E}^a$ of the state variables $\mathbf{x}^a$ in a specific time interval *n-1*:

$$\mathbf{E}^a = \{\mathbf{x}^{a(i)} : i = 1 \ldots . k\} \text{ with } k = \text{number of elements in the ensemble} \quad (4)$$

Then, each element of $\mathbf{E}^a$ is propagated from a time interval to the next adopting transition equations, thus generating the "background" ensemble $\mathbf{E}^b$:

$$\mathbf{E}^b = \left\{\mathbf{x}^{b(i)} = \mathbf{M}_{n-1,n}(\mathbf{x}^{a(i)})\right\} \quad (5)$$

with $\mathbf{M}_{n-1,n}$= propagation map between time intervals *n-1* and *n*

When the estimation step is performed through the analysis equations, the method determines which linear combination of the ensemble members generates the best estimate of the current state, given the current observations $\mathbf{y}^0$.

LETKF outperforms EnKF, since 1) it permits to minimize the Kalman filter cost function in a space whose dimension is the number of elements of the ensemble (it is "transformed"); 2) it provides a framework for data assimilation that permits a system-dependent localization strategy, breaking down the problem into sub-problems to be solved in parallel fashion (it is "local").

About the first strength of LETKF, Hunt et al 2007 demonstrated that solving the minimization of the Kalman filter cost function J during the analysis step for non-linear system in the space of the state variables (6) is equivalent to solve the minimization of (7) in the space of the ensemble.

$$min\, J(\mathbf{x}) = min\{(\mathbf{x} - \bar{\mathbf{x}}^b)^T (\mathbf{P}^b)^{-1} (\mathbf{x} - \bar{\mathbf{x}}^b) \quad + [\mathbf{y}^0 - H(\mathbf{x})]^T \mathbf{R}^{-1} [\mathbf{y}^0 - H(\mathbf{x})] \ \} \tag{6}$$
$$min\, \tilde{J}(\mathbf{w}) = min\{(k-1)\mathbf{w}^T\mathbf{w} + [\mathbf{y}^0 - H(\bar{\mathbf{x}}^b + \mathbf{X}^b\mathbf{w})]^T \mathbf{R}^{-1} [\mathbf{y}^0 - H(\bar{\mathbf{x}}^b + \mathbf{X}^b\mathbf{w})] \ \} \tag{7}$$

where: $\bar{\mathbf{x}}^b$=mean of the background ensemble; $\mathbf{P}^b$=covariance of the background ensemble; $\mathbf{y}^0$= observed data; H = non-linear model correlating state variables to data; $\mathbf{R}$ = covariance of the data; $\mathbf{w}$ = vector relating to the transformation of $\mathbf{x}$ in the subspace $\tilde{S}$ (space with a dimension equal to the number of elements of the ensemble); $\mathbf{X}^b$= deviation of each member of the background ensemble with respect to the mean of the background ensemble.

Solving (7) instead of (6) is mainly possible for the property of $\mathbf{P}^b$ (namely, $(\mathbf{P}^b)^{-1}$is well-defined on $\tilde{S}$). In order to perform the analysis on $\tilde{S}$ a change of coordinate system is required adopting the eigenvectors of $\mathbf{P}^b$ as basis. From the practical point of view, the information on how to perform this change of coordinate is reported in the next subsection. Instead, mathematical details of the property of the model can be found in Hunt et al 2007.

About the second strength of LETKF, it permits a "local" implementation, exploiting the concept of spatial localization. Spatial localization means that the modelled system has a "correlation distance" and the analysis should ignore ensemble correlations over larger distance. As reported in Hunt et al 2007, localization can be achieved considering only specific observations for each area or considering a distance-dependent function that decays to zero beyond a specific distance. This function could multiply $\mathbf{P}^b$, so that the data do not affect the state variables beyond such distance. In dynamic travel demand estimation, the spatial localization can be seen as the "OD coverage" concept: the location of traffic sensors, as well as the adoption of specific measurements, are directly correlated to the OD flows to be corrected. Working "locally" is convenient for computational purposes, but also to reduce the uncertainty and the errors in the estimation and forecast process. First attempts on this research line have been recently proposed by Etemadnia and Abdelghany (2009) and Frederix et al (2011b, 2014) for the off-line case.

In this paper, we focus on investigating the first strength of LETKF, i.e. the reliability of solving the minimization (7), while the "local" implementation is postponed to further research.

The implementation of LETKF for the on-line estimation and prediction of dynamic demand matrices is formalized in the following subsections.

4.1.1 LETKF for on-line dynamic demand estimation and prediction

The road traffic network is modelled by a directed graph G={N, A}, where N is the set of nodes and A is the set of links. The whole simulated horizon is T, divided into equal intervals h=1,……,t.

Given the a-priori OD flows $\mathbf{x}_{n-1}^a$ containing all trips departing during time interval *n-1*, the a-priori ensemble is generated:

$\mathbf{E}_{n-1}{}^{a} = \{\mathbf{x}_{n-1}^{a(i)}: i = 1 \ldots. k\}$ (8)

Transition Equations

Transition equations are computed to capture the temporal dynamic of the system from *n-1* to *n*. For each element *i* of $\mathbf{E}_{n-1}{}^{a}$, a forecast of the travel demand (background state estimate) for the following time interval *n* is obtained as:

$$\mathbf{x}_{n}^{b(i)} = \mathbf{M}_{n-1,n}(\mathbf{x}_{n-1}^{a(i)}) \quad (9)$$

obtaining a background ensemble $\mathbf{E}_{n}{}^{b}$ with average value $\overline{\boldsymbol{x}}_{n}^{b}$, deviation $\mathbf{X}_{n}^{b} = \mathbf{x}_{n}^{b(i)} - \overline{\boldsymbol{x}}_{n}^{b}$ and covariance matrix $\mathbf{P}_{n}^{b} = (k-1)^{-1}\mathbf{X}_{n}^{b}(\mathbf{X}_{n}^{b})^{T}$.

Measurements equations

Measurements equations map the state variables until time interval *n* onto the simulated traffic measurements $\mathbf{y}$ at time *n*:

$$\mathbf{y}_{n}^{b(i)} = H(\mathbf{x}_{1}^{a(i)}, \ldots \ldots \mathbf{x}_{n-1}^{a(i)}, \mathbf{x}_{n}^{b(i)}) \quad (10)$$

Then, the average value of measurements $\overline{\boldsymbol{y}}_{n}^{b}$ and the deviation with respect to the observed data at time *n* ($\mathbf{y}_{n}^{o}$), are computed as: $\mathbf{Y}_{n}^{b} = \mathbf{y}_{n}^{o} - \overline{\boldsymbol{y}}_{n}^{b}$. Covariance matrix of measurements is called $\mathbf{R}$ and it is usually assumed constant across time intervals. *H* is the non-linear model, i.e. the dynamic traffic assignment (DTA) model, connecting OD flows to traffic measurements.

Analysis equations

Finally, the analysis equations correct the background state estimate $\overline{\boldsymbol{x}}_{n}^{b}$, generating the a-priori ensemble $\mathbf{E}_{n}{}^{a}$ and its mean $\overline{\mathbf{x}}_{n}^{a}$ for time interval *n*.

The first step is to perform a coordinate change from the travel demand space to the ensemble space, i.e. from a n_{OD}-dimension space to a *k*-dimension space. With this coordinate change, the covariance matrix for the analysis state in the *k*-dimension space is:

$$\widetilde{\mathbf{P}}_{n}^{a} = [(k-1)\mathbf{I} + (\mathbf{Y}_{n}^{b})^{T}\mathbf{R}^{-1}\mathbf{Y}_{n}^{b}]^{-1} \quad (11)$$

with $\mathbf{I}$ identity matrix of dimension (k × k) and the average of the analysis state at time interval *n* in the *k*-dimension space:

$$\overline{\boldsymbol{w}}_{n}^{a} = \widetilde{\mathbf{P}}_{n}^{a}(\mathbf{Y}_{n}^{b})^{T}\mathbf{R}^{-1}\mathbf{Y}_{n}^{b} \quad (12)$$

where $\widetilde{\mathbf{P}}_{n}^{a}(\mathbf{Y}_{n}^{b})^{T}\mathbf{R}^{-1}$ is the Kalman gain. It has to be noticed that this transformation allows dealing with a Kalman gain that is not function of the non-linear operator *H*, neither of its Jacobian, as usually required in other linear and non-linear KF approaches.

Now it is possible to go back to the travel demand space (n_{OD}-dimension), computing:

1. the average of the analysis state at time interval *n* in the n_{OD} -dimension space:

$$\overline{\mathbf{x}}_{n}^{a} = \overline{\mathbf{x}}_{n}^{b} + \mathbf{X}_{n}^{b}\overline{\boldsymbol{w}}_{n}^{a} \quad (13)$$

2. the ensemble of the analysis state at time interval *n* in the n_{OD} -dimension space:

$$\mathbf{x}_{n}^{a(i)} = \overline{\mathbf{x}}_{n}^{b} + \mathbf{X}_{n}^{b}[\overline{\boldsymbol{w}}_{n}^{a} + \left[(k-1)\widetilde{\mathbf{P}}_{n}^{a}\right]^{\frac{1}{2}}] \quad (14)$$

3. the covariance matrix of the analysis state at time interval n in the n_{OD} -dimension space:

$$\mathbf{P}_n^a = \mathbf{X}_n^b \widetilde{\mathbf{P}}_n^a (\mathbf{X}_n^b)^T \quad (15)$$

4.1.2 Implementing LETKF

In order to apply the LETKF approach for the on-line demand estimation and prediction in a within day dynamic framework, it is necessary to:

1. Appropriately define the a-priori OD matrix $\mathbf{x}_{n-1}{}^a$ as the starting point of the procedure;
2. Appropriately select the elements of the ensemble (8);
3. Define the propagation map $\mathbf{M}_{n-1,n}$ (9);
4. Define the non-linear operator H (10).

For the a-priori OD matrix, a previous information on travel demand for each time interval is supposed to be available (historical or off-line estimate).

About the selection of the elements of the ensemble, the aim is to have possible matrices for each time interval as a sample of the travel demand generating average traffic conditions on the network. We may select the last matrices estimated during iterations of an off-line adjustment process or several matrices obtained from off-line adjustment conducted for several days. In such a way, an "off-line" ensemble can be created, representing the starting point of LETKF for any time interval in which the process has to be started.

About the definition of the propagation map, different formulations can be found in literature. Which one is the best is outside the scope of this paper and only for operative purposes, a polynomial approximation has been here adopted: it interpolates each OD flow derived by the average of the off-line ensemble from one time step to the next.

Finally H, that it is commonly replaced in KF approaches by the assignment matrices, is here considered as a "black box" calling for a dynamic traffic assignment (DTA) able to link the demand to all type of traffic measurements, not only traffic counts, without introducing any linear approximation. Thus, the DTA is called in order to estimate simulated measurements on the network that is the output of the measurements equations (10). For each time update, LETKF requires a number of DTAs equal to the number of matrices included in the ensemble. Assuming to halve the computational time of the DTA, with the same number of matrices in the ensemble, we could halve the time of prediction. This goal could be reached working on: 1) the acceleration of the algorithms adopted by the DTA; 2) the type of DTA adopted during the on-line procedure: i.e. dynamic network loading or user equilibrium approach; 3) the reduction of the dimension of the problem ("local" implementation exploiting the concept of spatial localization).

In the next section, several computational experiments have been carried out. Firstly, a comparison of LETKF with a standard non-linear KF is presented for a simple closed network in stationary conditions. Then, the sensitivity of LETKF to varying input parameters has been investigated considering a not trivial road network and dynamic conditions.

4.2 Computational experience

4.2.1. Comparison of LETKF with EnKF

The proposed LETKF has been applied in the case of a mono directional corridor (Fig. 1) in order to correct and predict the travel demand for three time slices, given the historical demand and traffic counts collected in real-time on links {7,8} and {9,10}.

Travel demand is assigned on the corridor considering stationary conditions for each time slice and results obtained with LETKF have been compared with results derived by the application of a standard non-linear EnKF.

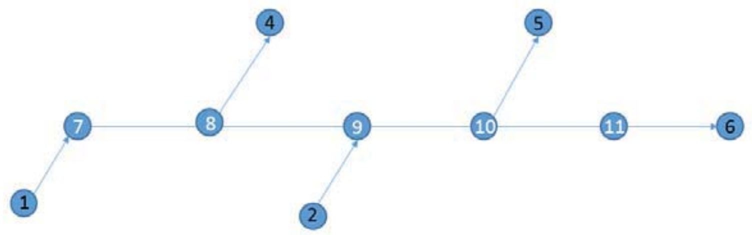

Fig. 2. Mono directional corridor

In order to evaluate the results of each procedure, the starting error on reproducing traffic counts given the historical demand has been computed adopting the Mean Absolute Error (MAE) statistic. Then, at the end of the on-line estimation, this reference value (Ref., Table 10) has been compared with the final MAE, representing the error on reproducing traffic counts given the estimated matrices; i.e., a negative number suggests a reduction of the MAE, thus an improvement on traffic counts reproduction due to the estimation and prediction process.

Since the on-line estimation starts to provide a correction of the OD flows from the second time interval (T=2, Table 10), the ability in reproducing traffic counts is evaluated starting from this interval until the end (T=3, Table 10); moreover, an overall evaluation is conducted on the whole planning horizon (Table 10, Tot).

Table 10 Numerical results (mono directional corridor, stationary conditions)

Reduction of MAE [%] with respect to Ref.			
	T=2 [Ref. 100 veh/h]	*T=3 [Ref. 215 veh/h]*	*Tot [Ref. 158 veh/h]*
EnKF	-98	-10	-38
LETKF	-98	-36	-55

Results (Table 10) highlight that LETKF is able to outperform EnKF. Moreover, the following considerations have to be reported:

1. The EnKF is a non-linear approach as the LETKF, but in the former, the computation of the Kalman gain requires the explicit knowledge of the non-linear map between OD flows and traffic measurements; in the latter, this map is no longer required. It means the need for EnKF of the assignment matrix in the stationary condition, while the calculation, storage and inversion of large, augmented assignment matrices moving to dynamic conditions;
2. Finally, even if the assignment matrices were somehow tractable, in the dynamic case we would miss the possibility to analytically explicit the map between the state variables and other

traffic measurements other than traffic counts.

The LETKF overcomes the previous weaknesses, being a non-linear approach that does not require assignment matrices for the computation of the Kalman gain. Moreover, joining the LETKF with a dynamic traffic simulation, other traffic measurements respect to traffic counts can be adopted, making the LETKF suitable for dynamic applications.

4.2.2. Numerical Experiments in a Dynamic Context

Now, LETKF has been applied for the on-line estimation and prediction of dynamic demand matrices in the road network reported in Fig.3. It consists of 22 nodes (14 are signalized intersections), 68 links, 6 traffic zones, a whole planning horizon of 35 minutes discretized into 5 minutes intervals, 12 monitored directional links providing traffic counts for each time interval.

The objective is to perform a first assessment of the LETKF applied to this specific problem (on-line estimation and prediction). Thus, laboratory experiments have been conducted considering (Tab.11, Fig.4):

1. Different starting matrices (seed matrices) with different degree of reliability with respect to the "real on-line demand" (i.e. the demand which generates on-line traffic measurements); the starting matrices have been obtained perturbing the "real on-line demand" (15,104 vehicles);
2. An increasing number *k* of elements in the ensemble (according to the number of variables);
3. Different levels of error ε between the elements in the ensemble (ensemble matrices) and the starting matrices (seed matrices). Specifically the error for generating each element has been derived from a uniform distribution between [$-\varepsilon$, $+\varepsilon$].

LETKF has been implemented in Matlab (The Mathworks, Inc.2013) and it calls for DYNASMART simulator, here adopted as a black box in order to obtain simulated measurements to be compared with real traffic measurements in the measurements equations.

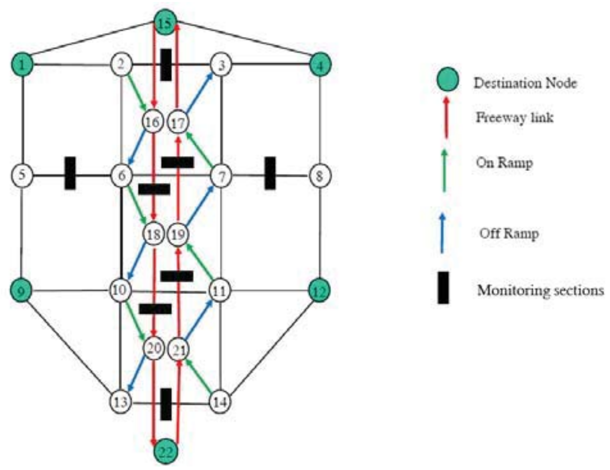

Fig. 3. Test network

Table 11 Experiments layout

Seed matrices [total demand value]	*k*	*ε* [%]
Seed_1 [14,668 veh]	{3, 6, 9, 12, 24, 36}	{10; 30; 50}
Seed_2 [15,145 veh]		
Seed_3 [14,015 veh]		

As in the stationary application, the starting error on traffic counts given by the different seed matrices has been computed in terms of Mean Absolute Error (MAE). Results report the total reduction (on all time slices) of MAE with respect to its reference starting value (Table 12).

Starting from Seed_1 (Table 12), it is possible to notice that LETKF can obtain satisfactory results, if the ensemble is maintained in a neighborhood of the starting matrix (ε=10%), regardless of the number of elements in the ensemble. This is related to the reliability of Seed_1, since it is a good representation of the real demand both in terms of total trips and OD flows distribution in space and time (Fig.4). If the ensemble is generated in a wider space (ε=30% or ε=50%), the only chance to get a satisfactory result, i.e. at least comparable with results obtained adopting ε=10%, is to increase the number of elements in the ensemble. In fact, it can be noticed a stabilization of the MAE reduction with increasing *k*.

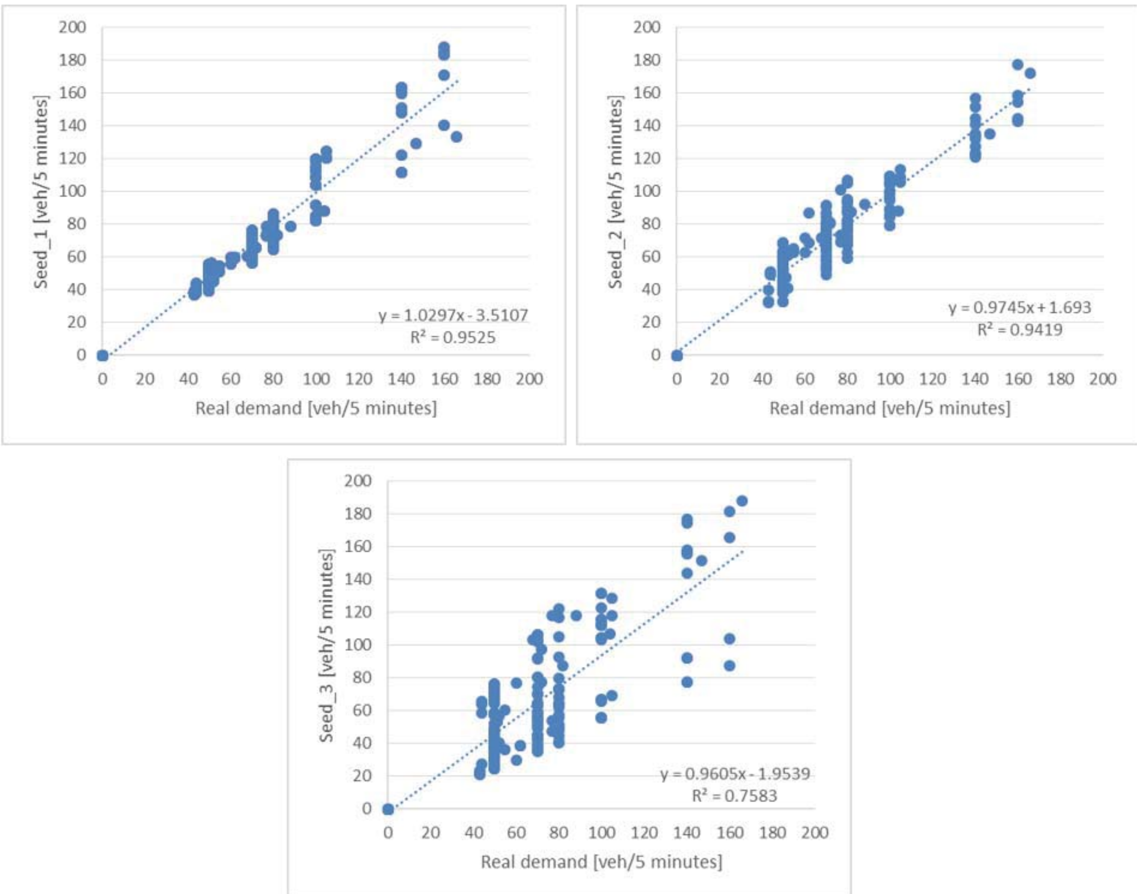

Fig. 4. Scatter plots Real OD flows VS Seed OD flows (for the different seed matrices adopted)

Considering Seed_2 (Table 12), we are facing with a starting demand that reports approximately the same amount of total trips as the real one, but with respect to the previous analyzed case (Seed_1) some problems can be detected in its distribution, especially for low-medium OD flow values (Fig.4). LETKF shows a general difficulty in achieving substantial improvements. This is true especially if a low error of the matrices in the ensemble with respect to the seed matrix is considered (ε=10%). If this error is high (ε=50%), the variability of the results for the different *k* compromises their validity. On the other hand, it should be noted as in such a starting condition, finding a correct "space" dimension for the elements in the ensemble becomes crucial to gain as much information as possible during the process. This is what happens considering ε=30%, where the improvements are significant especially with the increase of *k*.

Ending with Seed_3 (Table 12), this starting demand represents a bad approximation of the real demand, not just in terms of total demand value, but most in terms of its spatial and temporal distribution (Fig. 4). In such a case, good results have been also detected (higher than 20%). However, the variability for a fixed ε can be quite high. Thus, assessing the capability of LETKF in the case of Seed_3 is not possible and it confirms the need of starting the on-line process with a reliable off-line demand estimate.

Table 12 Numerical results and computational times

Reduction of MAE [%] with respect to Ref.							
Seed_1	*k3*	*k6*	*k9*	*k12*	*k24*	*k36*	**MAE Ref. [veh/5 minutes]**
ε 10%	-13	-15	-12	-10	-18	-17	23
ε 30%	-8	-4	-8	-7	-6	-11	
ε 50%	+3	-6	-9	-13	-16	-15	
Seed_2	*k3*	*k6*	*k9*	*k12*	*k24*	*k36*	**MAE Ref. [veh/5 minutes]**
ε 10%	-6	-7	-9	-11	-8	-7	21
ε 30%	-1	-2	-5	-25	-21	-19	
ε 50%	+4	-3	-10	-13	-3	-6	
Seed_3	*k3*	*k6*	*k9*	*k12*	*k24*	*k36*	**MAE Ref. [veh/5 minutes]**
ε 10%	-18	-20	-13	-7	-4	-18	29
ε 30%	-15	-20	-21	-7	-10	-16	
ε 50%	-13	-18	-7	-7	-10	-7	
Total Runtime	*k3*	*k6*	*k9*	*k12*	*k24*	*k36*	
[min:sec]	00:58	01:56	02:54	03:45	7:40	11:00	

Figure 5 shows some examples, for the three seed matrices adopted and different values of ε, of the trend of the LETKF behavior with the advance of time intervals. The MAE between on-line traffic counts and simulated flows, derived both by the assignment of the starting matrix and by the assignment of the resulting matrix of LETKF, has been reported. It could be noticed that in conditions where LETKF is generally able to work well, it can start to improve the correspondence with traffic conditions since the first time intervals. Instead, if starting conditions are not suitable, LETKF hardly follows what is happening on the road network and it seems to require a starting training phase, before being able to make substantial improvements.

The runtime of the experiments (Table 12, on an Intel Core i7 920, 2.67GHz, 12GB RAM) refers to the entire simulation period (35 minutes), with the estimation and prediction done every 5 minutes. These times are suitable for on-line applications. However, they can increase with the dimension of the network, thus they can be accepted only if the "local" approach, for which the LETKF is designed, can be effectively exploited.

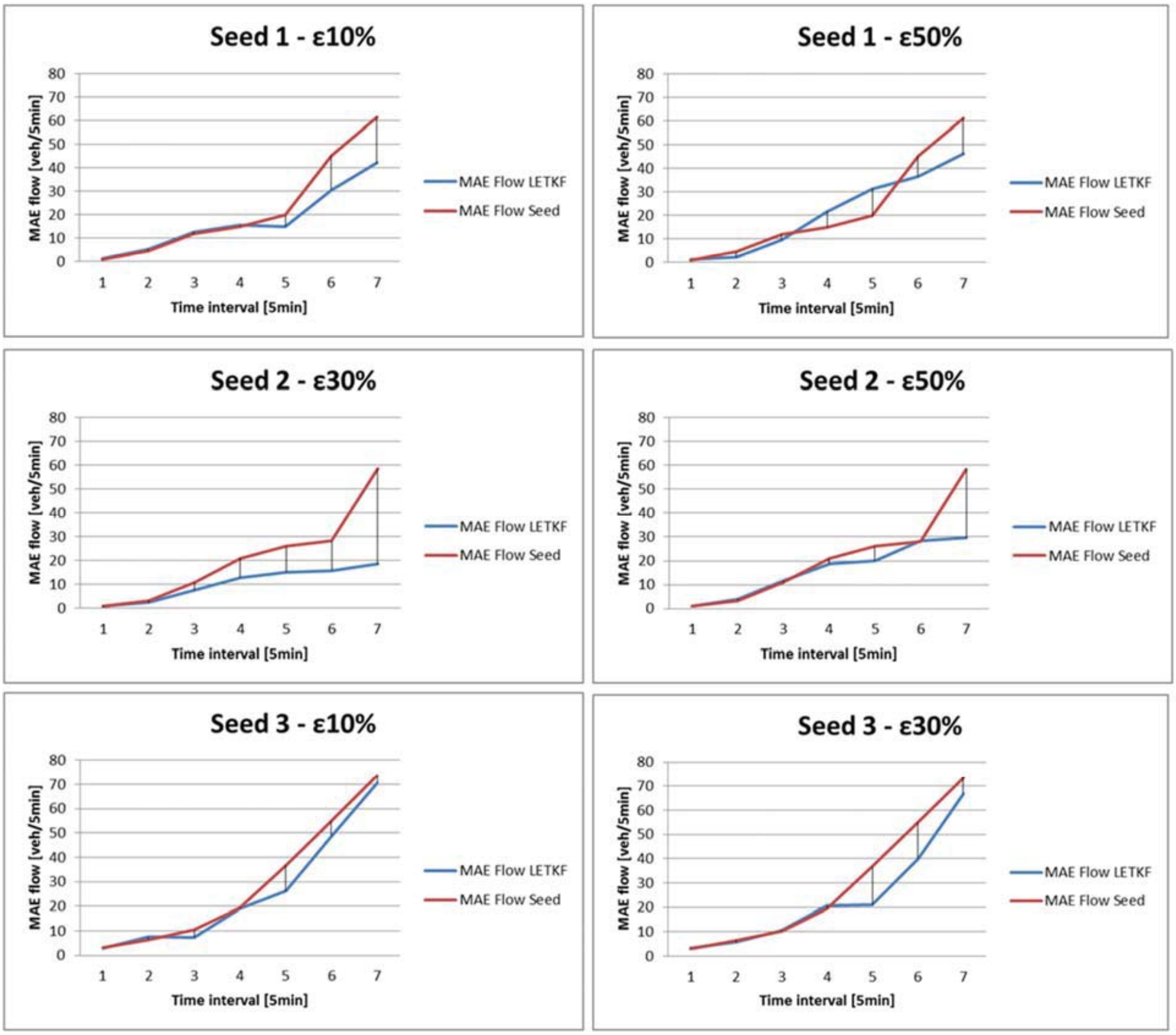

Fig. 5. Examples of MAE trend during time intervals between on-line traffic counts and simulated flows (from seed matrices and results of LETKF).

6. Conclusion and discussion

This paper faces with both the off-line and the on-line DDEP in the context of recent technology developments and the related new promising and challenging chances provided by the possibility to collect several heterogeneous traffic data.

Firstly, Floating Car Data (FCD) are exploited for off-line demand estimation. Information on user's dynamic route choice behavior and route travel times have been gathered, coupled with fixed location observations, and included in the DDEP.

Synthetic experiments have been conducted on the real network of Eur. These experiments analyzed the difference of adopting standard information as link flows during the DDEP respect to the new sources of data derived by FCD, considering the different sets of data both alone as well as their combined effects. Results demonstrated the strength and robustness associated to network based data as OD travel times and path choices. Link measurements are not able alone to define the real traffic pattern and this is reflected in solutions that do not represent the real users' behaviors. If information on route choices on the road network are available (and it could be possible with the large development of advanced monitoring systems), these information could assure a very refined solution. In fact, adopting both the information of OD travel times and route choice probabilities together during the estimation, the spatial and temporal reliability of the estimated demand consistently increases: this is true on the intercepted and the not intercepted ODs, despite a low value of the FCD penetration rate.

At the same time, a recent extension of the Kalman filter theory has been experimented for on line demand estimation and prediction, the Local Ensemble Transformed Kalman Filter (LETKF). The objective is to propose a method able to overpass any linearization of the dependency between OD flows and traffic measurements, thus able to englobe new sources of data as the FCD itself.

LETKF is easy to implement and computationally efficient. It permits to minimize the Kalman filter cost function in the ensemble space, thus reducing the dimension of the problem and its computational times. Operatively, this transformation permits also to deal with a Kalman gain in the analysis step that is not directly function of the non-linear DTA model. Thus, DTA is only required for the measurements equations.

LETKF has been demonstrated to outperform common non-linear KF. Moreover, preliminary results about the applicability of LETKF for a not trivial road network have been presented. Satisfying reproduction of traffic data can be obtained and these are closely related to the distance between the starting matrix and the "real on-line demand", combined with the space of the ensemble both in terms of number of elements and error with respect to the starting conditions.

These results lead the way for further investigations of LETKF. Firstly, other analyses are needed to evaluate the capability of LETKF regardless of the complexity of the DTA model. Moreover, the conducted experiments have been based on the adoption of only traffic counts as traffic data, while the structure of LETKF easily permits to add other type of measurements, which can help in a better reproduction of the correct state of the network.

Finally, LETKF provides a framework for data assimilation that permits a localization strategy, i.e. if different data can be referred to specific OD flows, the problem can be broken down into sub-problems to be solved in parallel fashion ("local" approach). The "local" approach that this article has only mentioned remains one of the main and most interesting research lines in order to verify the applicability of LETKF on large-scale road networks.

Acknowledgments

This paper contains a revisited version of Nigro et al 2016, presented at the TRB Annual Meeting 2016. The authors acknowledge the Standing Committee on Transportation Demand Forecasting, which reviewed it.

References

Antoniou, C., Ben-Akiva, M., Koutsopoulos, H.N. (2004). Incorporating Automated Vehicle Identification Data into Origin–Destination Estimation. Transportation Research Record: Journal of the Transportation Research Board, No. 1882, pp. 37–44.

Antoniou, C., Ben-Akiva, M., Koutsopoulos, H.N., 2006. Dynamic traffic demand prediction using conventional and emerging data sources. IEE Proc. Intell. Transp. Syst. 153 (1), 97–104.

Antoniou, C., Ben-Akiva, M., Koutsopoulos, H.N. (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.

Antoniou C, Barceló J, Breen M, Bullejos M, Casas J, Cipriani E, Ciuffo B, Djukic T, Hoogendoorn S, Marzano V, Montero L, Nigro M., Perarnau J, Punzo V, Toledo T, van Lint H (2016). "Towards a generic benchmarking platform for OD flows estimation/updating algorithms: design, demonstration and validation". Transportation Research. Part C, Emerging Technologies, ISSN: 0968-090X, doi: doi:10.1016/j.trc.2015.08.009

Ashok, K., M. Ben-Akiva (1993). Dynamic origin-destination matrix estimation and prediction for real-time traffic management systems. C. Daganzo, ed. Transportation and Traffic Theory, Proc. 12th ISTTT. Elsevier Science Publishing Company, Inc., Berkeley, CA.

Ashok, K. (1996). Estimation and prediction of time-dependent origindestination flows. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Ashok, K., M. Ben-Akiva (2000). Alternative approaches for real-time estimation and prediction of time-dependent origin-destination flows. Transportation Sci. 34(1) 21–36.

Ásmundsdóttir R: 'Dynamic OD matrix estimation using floating car data'. Msc thesis, Delft University of Technology, 2008

Ásmundsdóttir, R., Chen, Y. and Zuylen H. J.: 'Dynamic Origin-Destination Matrix Estimation Using Probe Vehicle Data as a Priori Information'. Traffic Data Collection and Its Standardization, International Series in Operations Research & Management Science, 2010, 144, pp 89-108

Balakrishna, R., Ben-Akiva, M., Koutsopoulos, H.N., 2007. Off-line calibration of dynamic traffic assignment: simultaneous demand and supply estimation. Transport. Res. Rec. 2003, 50–58.

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

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.

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

Cao, P., Miwa, T., Yamamoto, T., et al.: 'Bilevel Generalized Least Squares Estimation of Dynamic Origin-Destination Matrix for Urban Network with Probe Vehicle Data', Transportation Research Record: Journal of the Transportation Research Board, No. 2333, Transportation Research Board of the National Academies, Washington, D.C., 2013, pp. 67–73

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

Cipriani, E., Florian, M., Mahut, M., Nigro, M. (2011): 'A gradient approximation approach for adjusting temporal origin-destination matrices', Transportation Research Part C: Emerging Technologies, 2011, 19, (2), pp. 270–282

Cipriani, E., Nigro, M., Fusco, G., Colombaroni, C., (2013). Effectiveness of link and path information on simultaneous adjustment of dynamic o–d demand matrix. Eur. Transp. Res. Rev. http://dx.doi.org/10.1007/s12544-013-0115-z, ISSN: 1867-0717.

Cipriani, E., del Giudice, A., Nigro, M., Viti, F., Cantelmo, G. (2015): 'The Impact of Route Choice Modeling on Dynamic OD Estimation'. IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, Volume 2015-October, 30 October 2015, Article number 7313334, Pages 1483-1488

Cremer, M. and H. Keller (1984). A systems dynamics approach to the estimation of entry and exit o-d flows. Proceedings of 9th ISTTT.

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

Etemadnia H. & Abdelghany K. (2009). Distributed approach for estimation of dynamic origin-destination demand. Transportation research Record vol. 2105, pp 127-134.

Flötteröd, G., Bierlaire, M., 2009. Improved estimation of travel demand from traffic counts by a new linearization of the network loading map. In: Proceedings of 2009 European Transport Conference.

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

Frederix, R., Viti, F., Tampère, C.M.J. (2011b). A hierarchical approach for dynamic origin-destination matrix estimation on large-scale congested networks. IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC

Frederix, R., Viti, F., Himpe, W.W.E., Tampère, C.M.J. (2014) Dynamic origin-destination matrix estimation on large-scale congested networks using a hierarchical decomposition scheme. Journal of Intelligent Transportation Systems: Technology, Planning, and Operations.

Hunt, B. R.,Kostelich, E. J. and Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos: a Local Ensemble Transform Kalman Filter. Physica D, 230, 112–126.

Mahmassani, H. S., Sbayti, H., and Zhou, X.: DYNASMART-P Version 1.0 User's Guide. Maryland Transportation Initiative, College Park, Maryland, 2004, 137.

Marzano, V., Papola, A., Cascetta, E., Simonelli, F. (2015). A Kalman filter for on-line quasi-dynamic o-d flows estimation/updating. Presented at 2015 Models and Technologies for Intelligent Transportation Systems (MT-ITS) 3-5. June 2015. Budapest, Hungary.

Morikawa, T., and Miwa, T.: 'Preliminary Analysis on Dynamic Route Choice Behavior: Using Probe Vehicle Data', Journal of Advanced Transportation, 2006, 40, (2), 2006, pp. 140-163

Nigro, M., Mannini, L., Cipriani, E. (2016). A local ensemble transformed Kalman filter for the on-line estimation and prediction of dynamic demand matrices. Presented at TRB 95th Annual Meeting, January 2016.

Okutani, I., Stephanedes, Y., 1984. Dynamic prediction of traffic volume through Kalman filtering theory. Transport. Res. Part B 18 (2).

Paige, C. C., M. A. Saunders (1982). LSQR: An algorithm for sparse linear equations and sparse least squares. ACM Trans. Math. Software 8, 43–71.

Sohn, K., Kim, D., 2008. Dynamic origin destination flow estimation using cellular communication system. IEEE Trans. Veh. Technol. 57 (5), 2703–2713.

Tavana, H., 2001. Internally Consistent Estimation of Dynamic Network Origin–Destination Flows from Intelligent Transportation Systems Data using Bilevel Optimization. PhD Thesis, University of Texas, Austin.

Toledo, T., Kolechkina, T. (2012). Estimation of Dynamic Origin–Destination Matrices Using Linear Assignment Matrix Approximations. IEEE Transactions On Intelligent Transportation Systems. Digital Object Identifier 10.1109/TITS.2012.2226211.

Van der Zijpp, N., R. Hammerslag (1994). Improved Kalman filtering approach for estimating origin-destination matrices for freeway corridors. Transportation Res. Record 1443 54–64.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. J. Basic Engrg., Trans. ASME, Series D 82(1) 33–45.

Yamamoto, T., Miwa, T., Takeshita, T., et al.: 'Updating Dynamic Origin-Destination Matrices Using Observed Link Travel Speed by Probe Vehicles'. Transportation and Traffic Theory 2009: Golden Jubilee, 2009, pp. 723-738.

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.

Zhao, H., Yu, L., Guo, J., et al.: 'Estimation of time-varying OD demands incorporating FCD and RTMS data', Journal of Transportation Systems Engineering and Information Technology, 2010, 10, (1), pp. 72-80

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