Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Multisensor data fusion for OD matrix estimation

Krystyna Biletska*†§, Marie-Hélène Masson†‡, Sophie Midenet§ and Thierry Den¿ ux*†

* Université de Technologie de Compiègne
† UMR CNRS 6599 Heudiasyc, BP 20529, 60205 Compiègne cedex, France
‡ Université de Picardie Jules Verne, IUT de l Oise
§Université Paris-Est, Institut National de Recherche sur les Transports et leur Sécurité (INRETS), Laboratoire GRETIA
2 rue de la Butte Verte, 93166 Noisy le Grand cedex, France, E-mail: krystyna.biletska@gmail.com (corresponding author)

Abstract **Knowledge of trafŒc demand at a junction is crucial for most transport systems. Generally, it is represented by an origin-destination (OD) matrix, where each element is a volume of vehicle œow between one of the OD pair of zones of a junction. This paper introduces a new method for a short-time estimation of OD matrices at a signalised junction with a complex structure and Œtted out with video cameras. The estimation is made by taking into account the trafŒc lights and by fusion of different multisensor trafŒc data, which are subject to imprecision and uncertainty. The method proceeds in two steps. First, vehicle conservation law, expressed in terms of undertermined system of equations, is built dynamically for a short-time period. Second, four approaches, that overcome the system indetermination and model the data imperfection, are proposed to provide the best and unique estimates of the OD matrix. An experimental study has been made with data collected at the real signalised junction.**

Index Terms **Fuzzy sets, fuzzy modelling, fuzzy linear programming, origin-destination matrix estimation.**

I. INTRODUCTION

Information on trafŒc demand in a road network can be represented using the origin-destination (OD) matrix $\mathbf{B}$. For a junction, each element b_{ij} of such a matrix is a proportion of the œow of vehicles that come from entrance (origin) i and go to exit (destination) j. The matrix elements are also called the *OD œow rates*.

OD matrix estimated for a junction is a fundamental component for most of transport systems. Knowledge about trafŒc demand at a junction is generally used for trafŒc planning purposes, such as building of new links, rearrangement of turning lanes, change of the capacity of income/outcome links and of the speed limits, construction of roundabouts. Another important application of the junction OD matrix is to improve estimation of trafŒc demand at a whole road network. Thus, it is used in such trafŒc models as the model of vehicle s route choices designed for solving the trafŒc assignment problem.

Whenever a congestion occurs because of accident, meteorological disturbance or road maintenance work, it is important to assume fast and efŒcient measures to eliminate it. Therefore, information on destinations of travellers at a junction helps to develop and improve the trafŒc control strategies. The OD matrix is also employed for traveller information systems and validation of different trafŒc models.

Because of temporal variability of the trafŒc demand, most of these systems require real-time knowledge about OD matrix. Therefore, the estimation period of the OD œow rates has to be as short as possible, like a trafŒc light cycle (duration of green-amber-red sequence) at a signalised junction. Such a short-time estimated OD matrix is a key element of the diagnostic system for signalised junctions, developed at INRETS/GRETIA [1]. This system compares the impacts of different trafŒc control strategies formulated, for instance, in terms of CO_2 and pollutant emissions.

Junction OD matrix is deduced from vehicle counts made on each income and outcome link of the junction during a given time interval. These counts are usually provided by magnetic loops embedded in the road surfaces and sensitive to metallic masses. Since the loops are installed quite far from the entrances and exits of a junction, the OD matrix estimation period should be large enough in order to take into account the vehicle travel time between an OD pair of sensors and thus to respect the vehicle conservation law. In the case of a signalised junction, the travel time across the junction and the input and output œow volumes during a trafŒc cycle remain unknown. Unlike the electromagnetic loops, video sensors are able to provide the vehicle counts at any point of the junction, in particular, at the beginning of each entrance and each exit of the junction, which make possible to estimate an OD matrix at a short period of time.

The previous research works on OD matrix estimation at a junction generally yield estimates of the OD œow rates at a large-time period. Some of these methods are based on the information minimisation principle [2], on maximisation of likelihood [3] or on Bayesian inference [4]. Other methods present the estimation problem as a constraint optimisation problem [5], [6], [7]. Only a few methods can provide a short-time estimated OD œow rates, such as a recursive estimation method [8], [9] and methods based on the use of trafŒc lights [10], [11], [12]. However, they can only be applied for simple two-road-crossing junctions and, respectively, for junctions with two-phase trafŒc control strategy. Moreover, none of the cited methods considers the redundancy, uncertainty and imprecision of the data.

This paper focuses on the short-time estimation of OD matrices at a signalised junction with a complex structure and Œtted out with video cameras. In particular, we are interested in the use of trafŒc lights and imprecise and uncertain trafŒc measurements issued from video cameras. These measurements, provided every second, are the vehicle counts made on each entrance and exit of the junction and the number of vehicles stopped at each inner section of the junction. The main idea of the proposed method consists in constructing at each trafŒc light cycle a set of (fuzzy) relationships between the sequences of exit and entrance œow counts. A model of the vehicle œows crossing the junction is built via High-Level Petri Nets (HLPN) [13]. This model indicates every second a set of vehicle œows which could be present at the junction, the beginning and the end of these œows. It makes possible to draw dynamically a conservation law of vehicles, represented by an underdetermined system of equations whose unknowns are the OD œow rates.

Four different approaches, that overcome the inconsistency of the system and model data imperfections, are proposed to estimate the OD œow rates: ordinary least squares and linear programming (LP) methods based on the use of historical knowledge about the OD œow rates, LP and fuzzy LP methods based on a modelling of the imprecision of the data using, respectively, intervals or fuzzy numbers. An experimental study of the proposed methods has been made with the real data collected in an experimental site equipped with video cameras and trafŒc light controller. These tests demonstrated that the presented dynamic method yields accurate results.

The rest of the paper is organised as follows. Section 2 describes the real data. Section 3 brieœy introduces the model of OD œows crossing the junction and states the problem. Section 4 and 5 proposes four methods for OD matrix estimation and shows the results. The last section concludes the paper and proposes the further lines of research.

II. DATA DESCRIPTION

The experimental site is an isolated signalised junction of two double-lane roads situated in the south suburb of Paris (cf Fig. 1). The main road B-D which connects the suburbs to Paris has a high trafŒc volume, whereas the road A-C has a lower trafŒc volume. TrafŒc lights control four incoming links and four inner zones. Note that right-turning vehicles use special lanes and are not taken into account in this study. Only eight OD œows are statistically signiŒcant: $(AC, AD), (BD, BB), (CB), (DB, DC, DD)$.

Eight video cameras are installed at the junction in order to capture all the entrance and exit links and the inner zones. The location, height and angle of each camera depend on the geometry of the junction and are chosen to favour the measurement of space trafŒc parameters such as queue length on incoming links. The camera views are analysed in real time using image processing techniques developed at INRETS [14]. They provide several measurements every second:

- $\mathcal{X}_i(\tau)$ vehicle counts measured at the end of an entrance i at second τ $(i = 1, \ldots, n)$,
- $\mathcal{Y}_j(\tau)$ number of vehicles that have passed through the beginning of the exit j at second τ $(j = 1, \ldots, m)$,
- $\mathcal{Z}_k(\tau)$ number of stopped vehicles at inner zone k at second τ $(k = 1, \ldots, p)$,

where n, m and p are the numbers of entrances, exits and inner zones respectively. Here, a *trafŒc light cycle* is a period of time between two sequential onsets of the red light on the entrance Bf.

The temporal series can be represented graphically. Figure 2 presents an example of measurements of the vehicle œow arrived from the entrance Af and crossing some successive zones toward exits Cf and Df. Here vertical lines (solid and stippled) indicate the onsets of the red and green light respectively. The grey squares point out the vehicles counts at the corresponding entrance and exits. The white squares represent the number of vehicles stopped at inner zones. It can be seen that the volumes of incoming and exit œows are equal, that is to say $4/5$ of vehicles arrived from the entrance Af have leaved the junction by exit Cf and $1/5$ by exit Df. Actually, one vehicle from OD œow ADf has stopped twice in the inner zones 1f and 2f because of the red light in zone 2f. The Figure 2 demonstrates that the measurements are captured in inner zones with a delay with regard to the actual trafŒc situation.

It is worth noting that generally the vehicle counts are imprecise and erroneous so that the vehicle conservation law is not respected. The vehicles counts are more imprecise than the measure of occupation rates of inner zones by stopped vehicles.

The low quality of the measurement of trafŒc parameters is due to many phenomena. TrafŒc conditions (peak or off-peak periods) are the reason for many trafŒc count errors. If the trafŒc œow is heavy, the gaps between vehicles are small and it is difŒcult to distinguish these gaps on the video images. Thus the number of vehicles measured is lower than the actual number.

The characteristics of vehicles are also a source of mea-

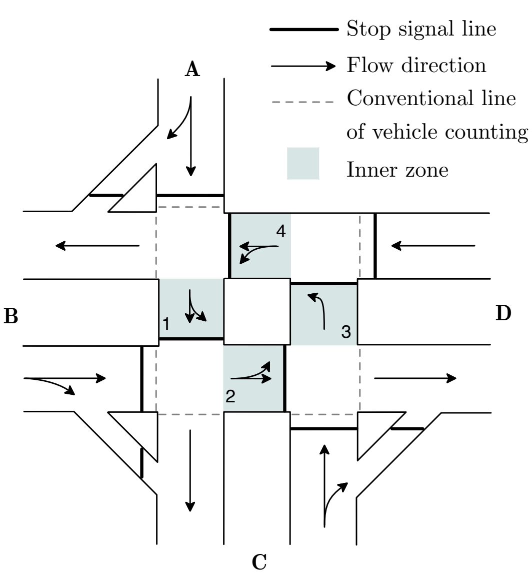

Fig. 1. The experimental junction

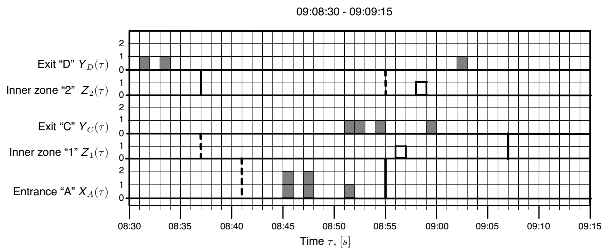

Fig. 2. Graphical representation of OD vehicle flows ACf and ADf.

surement errors. High vehicles passing in front of camera will hide the smaller vehicles or the whole camera field, i.e. they will produce a *masking effect*. Two-wheeled vehicles are only seldom counted because they are small. The heterogeneous colours of vehicle roofs also add to the problem of detection.

Meteorological conditions inevitably have an influence on all types of traffic measurements and the video are blurred: the wind shakes the posts the cameras are fixed to, the sun s rays cause the reflections on the vehicle surfaces and camera lenses, rain, snow and fog obscure a camera field. Changes in brightness caused by the position of the sun, clouds and headlights at night also determine the reliability of the measurements.

III. Problem statement

Let $x_i(c)$ be the flow volume at entrance i during a traffic light cycle c and $y_j(c)$ be the flow volume which entered the junction during cycle c and leaves it by exit j. *OD flow rate* $b_{ij}(c)$ is the proportion of the flow of vehicles that come from entrance i and go to exit j. The problem is to estimate OD flow rates $b_{ij}(c)$ $(\forall i \in [\![1,n]\!], \forall j \in [\![1,m]\!])$ at the end of each traffic light cycle c, such that

$$y_j(c) = \sum_{i=1}^{n} b_{ij}(c) x_i(c), \tag{1a}$$

$$x_i(c) b_{ij}(c) \geq z_{kij}(c) \quad \forall k \text{ s.t. } \delta_{kij} = 1, \tag{1b}$$

$$\sum_{j=1}^{m} b_{ij}(c) = 1, \tag{1c}$$

$$b_{ij}(c) \geq 0, \tag{1d}$$

where $z_{kij}(c)$ is the number of vehicles which cross the junction from i to j and stop at inner zone k during cycle c, $\delta_{kij} = 1$ if OD flow from i to j can pass through inner zone k and is 0 otherwise. For a given cycle c the value of variable $x_i(c)$ can be obtained from instantaneous vehicle counts

$$x_i(c) = \sum_{\tau=1}^{\mathcal{G}_i(c)} \mathcal{X}_i(\tau),$$

where $\mathcal{G}_i(c)$ is a duration of the green light of cycle c in entrance i. The values of $y_j(c)$ and $z_{kij}(c)$ cannot be obtained directly from traffic measurements, because it is impossible to know the period of time during which the vehicle flow $x_i(c)$ leaves the junction or stops at inner zones.

In order to obtain the values of $y_j(c)$ and $z_{kij}(c)$ from $\mathcal{Y}_j(\tau)$ and $\mathcal{Z}_k(\tau)$ respectively, we have proposed a dynamical model of vehicle flows crossing the junction using High-Level Petri Nets [13]. This model provides the onsets of the beginnings and the ends of the flows. For a given cycle c we thus know the possible duration of the presence of the flows in each zone and can collect the corresponding measurements during the same period. Therefore it is possible to put into a one-to-one correspondence the flows crossing the junction zones and the measurements taken at this zones. Consequently, at each cycle c, we can build dynamically a vehicle conservation law, represented by the following system of equations whose unknowns are the OD flow rates:

$$\mathbf{y} = \mathbf{X_1} \mathbf{b}, \tag{2}$$

such that

$$\mathbf{X_2} \mathbf{b} \geq \mathbf{z}, \tag{3a}$$

$$\mathbf{I} \mathbf{b} = \mathbf{1}, \tag{3b}$$

$$\mathbf{b} \geq 0, \tag{3c}$$

where $\mathbf{y}$ and $\mathbf{z}$ are vectors composed of J and K elements respectively, $\mathbf{b}$ is a vector containing the I unknowns b_{ij}, $\mathbf{X_1}$ is an $J \times I$ matrix rearranged so that (2) is equal to (1a), $\mathbf{X_2}$ is an $K \times I$ matrix built so that (3a) is equivalent to (1b), $\mathbf{I}$ is an indicator $n \times I$ matrix organised in such a way that (3b) is equal to (1c) and $\mathbf{1}$ is an identity n-vector. Note that for the clarity of notation the cycle c has been omitted.

The number of unknowns I depends on the junction structure and on the number of platoons that compose an OD flow. Depending on the traffic control strategy and the flow volume, an OD flow can leave the junction as a single platoon or as several platoons. The number of equations J depends on the number of exits and, in the same manner, on the control strategy. Since $I \geq J$ (in this paper $I \approx 10$), the system of equations (2) is underdetermined.

In order to overcome the problem of the system indetermination and to guarantee the uniqueness of the solution it is necessary to use some additional information about the values of the OD flow rates. In this paper, for choosing a solution among a solution set, we propose to use the mean value of the OD flow rates estimated during a time period preceding the estimation, called *sliding mean value*. Moreover, to model the uncertainty and imprecision of the real data, we suggest to introduce slack variables and to represent the values of $\mathbf{b}$, $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{y}$ and $\mathbf{z}$, first, by intervals, and, second, by triangular fuzzy numbers.

In fine, four methods are proposed in this paper. To estimate the crisp unknowns $\mathbf{b}$ when the elements of model (2) and of constraint (3c) are crisp, we apply the ordinary least squares (*LS*) and the linear programming (*CLP*) methods using the sliding mean value of $\mathbf{b}$. Linear programming (*ILP*) and fuzzy linear programming (*FLP*) approaches are formulated for the cases where these elements are represented by intervals and fuzzy numbers respectively. We introduce in the sequel the four methods and show their results on our application.

IV. OD FLOW RATES ESTIMATION

A. Least squares method with crisp variables and coefficients (CLS)

The estimation problem of OD flow rates $\mathbf{b}$ can be considered as an estimation of coefficients of the regression model equivalent to the relationship (2). Let the input and output variables $\mathbf{X_1}, \mathbf{X_2}, \mathbf{y}, \mathbf{z}$ and the coefficients $\mathbf{b}$ be non-negative crisp numbers. A first approach to estimate $\mathbf{b}$ is to use a least squares method. A sliding mean values $\overline{\mathbf{b}} = \{\overline{b}_l\}_{l=1,\ldots,8}$, available for eight OD pairs which are described in Section II, are used to provide a unique solution. To insure the existence of feasible solutions, we introduce slack variables, $e_{1,j}$ ($\forall j = 1, \ldots, J$) and $e_{1,k}$ ($\forall k = 1, \ldots, K$), and we propose to solve the following problem:

$$\min_{\mathbf{b}^*, \mathbf{b}, \xi, \mu, \eta} \left(\| \mathbf{b}^* - \overline{\mathbf{b}} \|^2 + \gamma_1 \sum_{j=1}^{J} e_{1,j} + \right.$$
$$\left. + \gamma_2 \sum_{j=1}^{J} e_{2,j} + \gamma_3 \sum_{k=1}^{K} e_{3,k} \right) \tag{4}$$

s.t.

$$\begin{cases} \mathbf{A}\mathbf{b} = \mathbf{b}^*, \\ \mathbf{I}\mathbf{b}^* = \mathbf{1}, \\ \mathbf{X_1}\mathbf{b} = \mathbf{y} + \mathbf{e_1} - \mathbf{e_2}, \\ \mathbf{X_2}\mathbf{b} + \mathbf{e_3} \geq \mathbf{z}, \\ \mathbf{b} \geq 0, \\ \mathbf{e_1} \geq 0, \mathbf{e_2} \geq 0, \mathbf{e_3} \geq 0, \end{cases} \tag{5}$$

where γ_1, γ_2 and γ_3 are given weight coefficients, $\mathbf{b}^*$ is the vector of the eight OD flow rates, considered in Section II, and $\mathbf{A}$ is a matrix rearranged in such a way that $\mathbf{A}\mathbf{b} = \mathbf{b}^*$.

B. Linear programming problem with crisp variables and coefficients (CLP)

A linear programming approach can be applied to estimate the crisp elements of the regression model (2). Using the same notations, we can build the following minimisation problem:

$$\min_{\mathbf{b}^*, \mathbf{b}, \mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}, \mathbf{e_4}, \mathbf{e_5}} \left(\gamma_1 \sum_{j=1}^{J} e_{1,j} + \gamma_2 \sum_{j=1}^{J} e_{2,j} + \right.$$
$$\left. + \gamma_3 \sum_{k=1}^{K} e_{3,k} + \gamma_4 \sum_{i=1}^{I} e_{4,i} + \gamma_5 \sum_{i=1}^{I} e_{5,i} \right) \tag{6}$$

s.t.

$$\begin{cases} \mathbf{X_1}\mathbf{b} = \mathbf{y} + \mathbf{e_1} - \mathbf{e_2}, \\ \mathbf{X_2}\mathbf{b} + \mathbf{e_3} \geq \mathbf{z}, \\ \mathbf{A}\mathbf{b} = \mathbf{b}^*, \\ \mathbf{b}^* = \overline{\mathbf{b}} + \mathbf{e_4} - \mathbf{e_5}, \\ \mathbf{I}\mathbf{b} = \mathbf{1}, \\ \mathbf{b} \geq 0, \\ \mathbf{e_1} \geq 0, \mathbf{e_2} \geq 0, \mathbf{e_3} \geq 0, \mathbf{e_4} \geq 0, \mathbf{e_5} \geq 0, \end{cases} \tag{7}$$

where γ_1, γ_2, γ_3, γ_4 and γ_5 are given weight coefficients and $e_{4,i}$, $e_{5,i}$ ($\forall i = 1, \ldots, I$) are slack variables.

C. Linear programming problem with variables and coefficients represented as intervals (ILP)

In order to take into account the inherent imprecision of the data, we choose to represent the elements of the model (2) and constraint (3a) by intervals, denoted as $a = [a^-, a^+]$. The values of $\mathbf{X_1^-}$ and $\mathbf{X_1^+}$ ($\mathbf{X_2^-}$ and $\mathbf{X_2^+}$) was inferred from empirical distributions of the error counts (cf Fig. 3). In most cases, counts of vehicles at the entrances are smaller than the true value of $\mathbf{X}_1$ ($\mathbf{X}_2$). Here, the equation (2) is interpreted as $\mathbf{X_1}\mathbf{b} \subseteq \mathbf{y}$ and inequality constraint (3a) as the least conservative inequality $\mathbf{X_2^+}\mathbf{b}^+ \geq \mathbf{z}^-$. Since there is no available histograms of error counts for $\mathbf{y}$ and $\mathbf{z}$, intervals $[\mathbf{y}^-, \mathbf{y}^+]$ and $[\mathbf{z}^-, \mathbf{z}^+]$ was constructed by adding to the known values of $\mathbf{y}$ and $\mathbf{z}$ the symmetrical spreads chosen experimentally. With respect to constraints (2)-(3c), we propose to estimate $\mathbf{b}$ such that $\mathbf{b}^- \leq \mathbf{b} \leq \mathbf{b}^+$ and the width of intervals $[\mathbf{b}^-, \mathbf{b}^+]$ tend toward zero. Considering that all variables and coefficients of the regression model are non-negative, the minimisation problem is written as follows:

$$\min_{\mathbf{b}, \mathbf{b}^+, \mathbf{b}^-, \mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}} \left(\sum_{i=1}^{I} \left(\mathbf{b_i^+} - \mathbf{b_i^-} \right) + \gamma_1 \sum_{j=1}^{J} e_{1,j} + \right.$$
$$\left. + \gamma_2 \sum_{j=1}^{J} e_{2,j} + \gamma_3 \sum_{k=1}^{K} e_{3,k} \right) \tag{8}$$

s.t.

$$\begin{cases} \mathbf{X_1^-}\mathbf{b}^- \leq \mathbf{y}^- + \mathbf{e_1}, \\ \mathbf{X_1^+}\mathbf{b}^+ + \mathbf{e_2} \geq \mathbf{y}^+, \\ \mathbf{X_2^+}\mathbf{b}^+ + \mathbf{e_3} \geq \mathbf{z}^-, \\ \mathbf{I}\mathbf{b} = \mathbf{1}, \\ 0 \leq \mathbf{b}^- \leq \mathbf{b} \leq \mathbf{b}^+ \leq 1, \\ \mathbf{e_1} \geq 0, \mathbf{e_2} \geq 0, \mathbf{e_3} \geq 0, \end{cases} \tag{9}$$

where γ_1, γ_2 and γ_3 are given weight coefficients.

D. Linear programming problem with fuzyy variables and coefficients (FLP)

Let $\tilde{\mathbf{X}}_\mathbf{1}, \tilde{\mathbf{X}}_\mathbf{2}, \tilde{\mathbf{y}}, \tilde{\mathbf{z}}$ and $\tilde{\mathbf{b}}$ be triangular fuzzy numbers represented as triple $\tilde{\mathbf{a}} = (\mathbf{a}^-, \mathbf{a^m}, \mathbf{a}^+)$. The form of fuzzy numbers $\tilde{\mathbf{X}}_1$ was derived from empirical distributions of the error counts as shown in Fig. 3. The fuzzy number $\tilde{\mathbf{X}}_2$ is determined similarly to $\tilde{\mathbf{X}}_1$. Because of the lack of histograms of error counts for $\mathbf{y}$ and $\mathbf{z}$, fuzzy numbers $\tilde{\mathbf{y}}$ and $\tilde{\mathbf{z}}$ are supposed to be symmetrical with the spreads chosen experimentally.

One of the way for interpreting the equality of two fuzzy numbers $\tilde{a}$ and $\tilde{b}$ (cf [15]) is to consider ${}^{\alpha}A \subseteq {}^{\alpha}B$ ($\forall \alpha \in [0,1]$), where the notation ${}^{\alpha}A = \{x | \mu_A(x) \geq \alpha\}$ stands for the α-cut of fuzzy number $\tilde{\mathbf{a}}$. In the same manner we define $\tilde{a} \geq \tilde{b}$ as ${}^{\alpha}A \geq {}^{\alpha}B$, $\forall \alpha \in [0,1]$.

According to (2)-(3c) and supposing the unknown OD flow rates $\tilde{\mathbf{b}}$ to be triangular fuzzy numbers of the form $(\mathbf{b}^-, \mathbf{b}^m, \mathbf{b}^+)$, we propose to solve the previous problem (*ILP*) for each α-cut. For this reason a new constraint is introduced:

$$[{}^{\alpha_i}\mathbf{b}^-, {}^{\alpha_i}\mathbf{b}^+] \subseteq [{}^{\alpha_j}\mathbf{b}^-, {}^{\alpha_j}\mathbf{b}^+],$$

where $i < j$, $\forall i, j \in [\![1, N]\!]$, and N is the finite number of α-cuts. The following fuzzy linear programming is written:

$$\min_{\mathbf{b}^m, \mathbf{b}^+, \mathbf{b}^-, \mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}} \left(\sum_{\alpha=1}^{N} \sum_{i=1}^{I} \left({}^{\alpha}\mathbf{b}_i^+ - {}^{\alpha}\mathbf{b}_i^- \right) + \gamma_1 \sum_j e_{1,j} + \right.$$

$$\left. + \gamma_2 \sum_j e_{2,j} + \gamma_3 \sum_k e_{3,k} \right) \quad (10)$$

s.t.

$$\begin{cases} {}^{\alpha}\mathbf{X}_1^- {}^{\alpha}\mathbf{b}^- \leq {}^{\alpha}\mathbf{y}^- + \mathbf{e_1}, \\ {}^{\alpha}\mathbf{X}_1^+ {}^{\alpha}\mathbf{b}^+ + \mathbf{e_2} \geq {}^{\alpha}\mathbf{y}^+, \\ {}^{\alpha}\mathbf{X}_2^+ {}^{\alpha}\mathbf{b}^+ + \mathbf{e_3} \geq {}^{\alpha}\mathbf{z}^-, \\ 0 \leq {}^{\alpha_i}\mathbf{b}^- \leq \mathbf{b}^m \leq {}^{\alpha_i}\mathbf{b}^+ \leq 1, \quad \forall i \in [\![1, N]\!] \\ {}^{\alpha_i}\mathbf{b}^- \leq {}^{\alpha_j}\mathbf{b}^- \leq {}^{\alpha_j}\mathbf{b}^+ \leq {}^{\alpha_i}\mathbf{b}^+, \ i < j, \forall i, j \in [\![1, N]\!], \\ \mathbf{I}\mathbf{b}^m = \mathbf{1}, \\ \mathbf{e_1} \geq 0, \mathbf{e_2} \geq 0, \mathbf{e_3} \geq 0, \end{cases} \quad (11)$$

where γ_1, γ_2 and γ_3 are given weight coefficients and the α-cuts of the fuzzy numbers $\tilde{\mathbf{X}}_1, \tilde{\mathbf{X}}_2, \tilde{\mathbf{y}}$ and $\tilde{\mathbf{z}}$ are calculated as follows:

$${}^{\alpha}\mathbf{X}_1^- = \alpha(\mathbf{X}_1^m - \mathbf{X}_1^-) + \mathbf{X}_1^-, \quad {}^{\alpha}\mathbf{X}_1^+ = \alpha(\mathbf{X}_1^m - \mathbf{X}_1^+) + \mathbf{X}_1^+,$$
$${}^{\alpha}\mathbf{X}_2^+ = \alpha(\mathbf{X}_2^m - \mathbf{X}_2^+) + \mathbf{X}_2^+, \quad {}^{\alpha}\mathbf{y}^- = \alpha(\mathbf{y}^m - \mathbf{y}^-) + \mathbf{y}^-,$$
$${}^{\alpha}\mathbf{y}^+ = \alpha(\mathbf{y}^m - \mathbf{y}^+) + \mathbf{y}^+, \quad {}^{\alpha}\mathbf{z}^- = \alpha(\mathbf{z}^m - \mathbf{z}^-) + \mathbf{z}^-.$$

V. EXPERIMENTAL RESULTS

These methods have been tested using real data collected at the experimental junction during 30-minutes period with peak traffic conditions. The estimation is made on 25 consecutive traffic light cycles. The actual values of OD flow rates β_i ($\forall i = 1, \ldots, 8$), calculated manually from video images, are available for all cycles.

The width of intervals $\mathbf{y}$ and $\mathbf{z}$ of the *ILP* method and the support width of fuzzy numbers $\tilde{\mathbf{y}}$ and $\tilde{\mathbf{z}}$ of the *FLP* method were experimentally fixed to 2. The weight coefficients in all models has been fixed as follows: $\gamma_1 = 0.1$, $\gamma_2 = 0.1$, $\gamma_3 = 0.2$, $\gamma_4 = \gamma_5 = 1$.

The estimation error has been calculated for the OD flow rates (Fig. 4): $E = \hat{\mathbf{b}}\mathbf{X}^* - \beta\mathbf{X}^*$, where $\mathbf{X}^*$ is a vector of actual vehicle counts at the entrances of the experimental junction, $\hat{\mathbf{b}}$ is a vector of estimated OD flow rates. For *FLP* method we consider $\hat{\mathbf{b}} = \mathbf{b}^m$. The best results are obtained with *CLP* and *FLP* methods for which the median error is about zero for all OD flows. The prediction of the exit flows with *ILP* method is less accurate than those of the *FLP* method (cf Fig. 5). Note that the estimation error E of all methods is higher, if the flow volume is lower, like for the OD flows ADf and BBf.

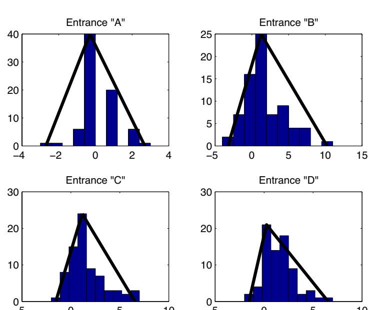

Fig. 3. Empirical distributions of error counts for the entrances of the junction

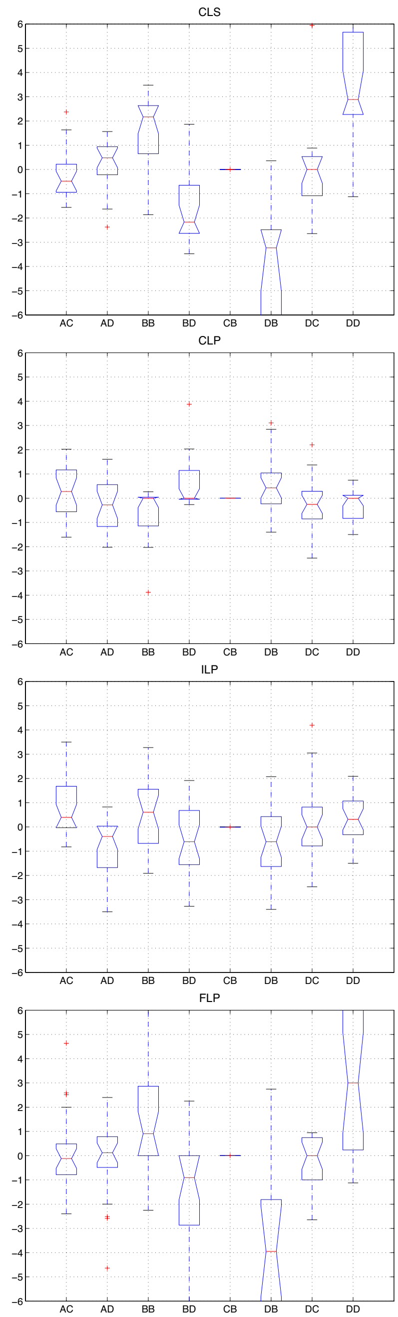

Fig. 4. Estimation errors (in number of vehicles) for 8 OD flow rates for the 25 cycles

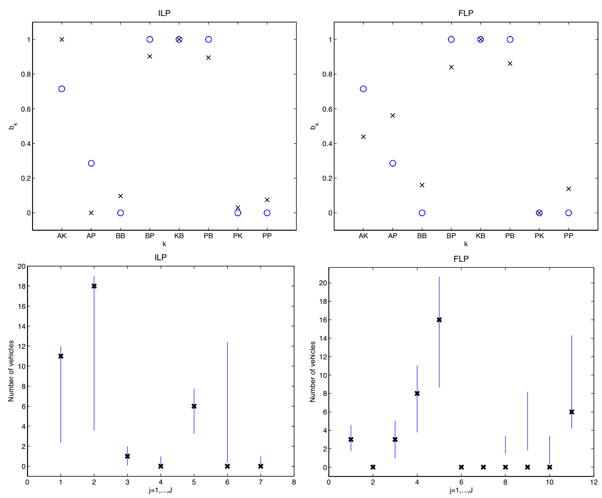

Fig. 5. Estimations $\hat{\mathbf{b}}$ ($\times$) and true values β (circles) for 8 OD flow rates and model prediction $[\mathbf{X}_1^-\hat{\mathbf{b}}, \mathbf{X}_1^+\hat{\mathbf{b}}]$ of exit flow $\mathbf{y}$ for one particular cycle

VI. Conclusion

A new short-time estimation method of the OD matrix for a signalised junction have been proposed in this paper. The method is founded on the construction of a conservation law of vehicles at each traffic light cycle, represented by the underdetermined system of equations. This system is obtained in a dynamical way from the model of traffic flows built using the High-Level Petri Nets (HLPN) [13]. Real data collected at the experimental signalised junction fitted out with video cameras are used to estimate the OD flow rates.

Four approaches have been proposed to estimate the OD matrix. The inherent imprecision of the data was modelled in three different ways: by introducing slack variables, by representing the data by intervals and by triangular fuzzy numbers. An additional historical knowledge about the values of the OD flow rates is used in two of the proposed approaches to guarantee the uniqueness of the solution.

The best results were obtained with two linear programming approaches, one of which is based on the use of the sliding mean value of the OD flow rates and the second is founded on fuzzy modelling of the data. The latest method is applied if the lack of precision of traffic measurements is important. However, this method does not guarantee the uniqueness of the solution.

Our future lines of research will be centred, first, on the improvement of the FLP method in order to provide the best and unique estimation of the OD flow rates and, second, on the modelling of the temporal uncertainty of the data by introducing the notion of a fuzzy token in the HLPN model. The application of our method should be also extended to a sequence of junctions.

References

[1] S. Midenet, F. Boillot, and J.-C. Pierrelée, "Signalized intersection with real-time adaptive control: on-field assessment of CO_2 and pollutant emission reduction," *Transportation Research Part D*, vol. 9, pp. 29–47, 2004.

[2] H. J. Van Zuylen, "The estimation of turning flows on a junction," *Traffic Engineering Control*, vol. 20, no. 11, pp. 539–541, 1979.

[3] E. Hauer, E. Pagitsas, and B. T. Shin, "Estimation of turning flows from automatic counts," *Transportation Research Record*, no. 795, pp. 1–7, 1981.

[4] M. J. Maher, "Inferences on trip matrices from observations on link volumes: a Bayesian statistical approach," *Transportation Research Part B*, vol. 17, no. 6, pp. 435–447, 1983.

[5] H. D. Sherali, N. Arora, and A. Hobeika, "Parameter optimization methods for estimating dynamic origin-destination trip-tables," *Transportation Research Part B*, vol. 31, no. 2, pp. 141–157, 1997.

[6] B. Li and B. De Moor, "Recursive estimation based on the equality-constrained optimization for intersection origin-destination matrices," *Transportation Research Part B*, vol. 33, no. 3, pp. 203–214, 1999.

[7] A. P. Tarko and R. Lyles, "Using flow conservation and data redundancy to estimate turning flows at intersections," *A paper submitted for presentation and subsequent publication at the 80th Annual Transportation Research Board Meeting*, January 2001.

[8] M. Cremer and H. Keller, "A new class of dynamic methods for the identification of origin-destination flows," *Transportation Research Part B*, vol. 21, no. 2, pp. 117–132, 1987.

[9] N. L. Nihan and G. A. Davis, "Recursive estimation of origin-destination matrices from input/output counts," *Transportation Research Part B*, vol. 21, no. 2, pp. 149–163, 1987.

[10] H. Keller and G. Ploss, "Real-time identification of O-D network flows from counts for urban traffic control," in *Proceedings of the 10th International Symposium on Transportation and Traffic Theory*, N.H.Gartner and N.H.M.Wilson, Eds. Elsevier Science Publishing Co., 1987, pp. 267–289.

[11] I. Matschke and B. Friedrich, "Dynamic OD estimation using additional information from traffic signal lights timing," *Proceedings of the Triennial Symposium on Transportation Analysis TRISTAN IV*, June 13-19, Sao Liguel - Azores, Portugal 2001.

[12] J. Tian, M. R. Virkler, and C. Sun, "Field testing for automated identification of turning movements at signalized intersections," *TRB, Transportation Research Record*, no. 1867, pp. 210–216, 2004.

[13] K. Biletska, S. Midenet, M. Masson, and T. Denoeux, "Fuzzy modelling of the sensor data for the estimation of an origin-destination matrix," 2009, to appear in Proceedings of IFSA-EUSFLAT 2009, 20-24 July, Lisbon, Portugal.

[14] D. Aubert, S. Bouzar, F. Lenoir, and J. M. Blosseville, "Automatic vehicle queue measurement at intersections using image-processing," *8th IEE International Conference on Road Traffic Monitoring and Control, London*, vol. 422, pp. 100–104, 1996.

[15] D. Dubois and H. Prade, Eds., *Fuzzy sets and Systems: Theory and Applications*. Academic press, Inc., 1980.