Fuzzy modelling of sensor data for the estimation of an origin-destination matrix

Krystyna Biletska[124] Sophie Midenet[4] Marie-Hélène Masson[23] Thierry Denœux[12]

1. Université de Technologie de Compiègne
2. UMR CNRS 6599 Heudiasyc,
BP 20529, 60205 Compiègne cedex, France
3. Université de Picardie Jules Verne, IUT de l'Oise
4. 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
Email: krystyna.biletska@inrets.fr (corresponding author)

Abstract— *This paper examines a short-time estimation problem of an origin-destination (OD) matrix, where each element is a volume of vehicle flow between one of the OD pair of zones of a signalised junction. The estimation is based on the use of traffic measurements provided by video sensors and on the knowledge of the traffic lights. This data is subject to redundancy, imprecision and uncertainty. The main purpose of this paper is to obtain the best estimates of the OD matrix by modelling the data imperfection, using a two-step method. First, relationships between the observed data are built in real-time using High-Level Petri Nets. Due to the imperfection of data the system obtained is underdetermined and inconsistent. Second, the fuzzy sets theory is used to model this imperfection and to overcome the inconsistency of the system.*

Keywords— Fuzzy least squares, Fuzzy linear programming, Fuzzy modelling, Origin-destination matrix

1 Introduction

Knowledge about origins and destinations (OD) of vehicles in a road network is important in most transport systems. In the case of a junction, its mathematical representation is OD matrix $\mathbf{B}$, each element b_{ij} of which is a proportion of the flow of vehicles that come from entrance (origin) i and go to exit (destination) j. Such a proportion is called the *OD flow rate*. Since the OD matrix changes in time following the changes in traffic demand, the estimation period of the OD matrix has to be as short as possible. In particular, the period should be equal to a traffic light cycle (duration of green-amber-red sequence) when we deal with a signalised junction. At INRETS/GRETIA we use such a short-time estimation of the OD matrix as part of a diagnostic system for signalised junctions [1]. This system compares the impacts of different traffic control strategies expressed in terms of CO_2 and pollutant emissions.

The OD matrix is generally deduced from vehicle counts made on each entrance and each exit 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. The estimation can be obtained from a conservation law of vehicles which is a set of relationships between exit and entrance flow counts. In general, when loops are installed on every entrance and exit of the junction, the estimation problem is underdetermined. Thus a solution is not unique and additional information such as a *prior* OD matrix is used to choose the OD matrix which corresponds best to the actual matrix. Some of the existing methods are based on the information minimisation principle [2], on maximisation of likelihood [3] or on Bayesian inference [4]. Other methods propose a recursive estimation of the OD matrix [5, 6]. Furthermore, the estimation problem can be represented as a constraint optimisation problem [7, 8]. Methods which use traffic lights are described in [9, 10].

This paper considers the problem of reconstituting the origins and destinations of vehicle flows crossing a signalised junction, at each traffic light cycle. OD flow volumes are estimated using traffic lights and traffic measurements from video cameras installed at the junction. 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 data is subject to redundancy, uncertainty and inaccuracy. Such an imperfection is linked to the reliability of video sensors, measurement conditions, traffic characteristics and drivers' behaviour.

In order to take into account all available information it is necessary to consider the nature and possible interdependence of the data. None of the cited methods takes into account the lake of imprecision of vehicle counts, the possible physical complexity of the junction and the traffic lights at the same time. Moreover they cannot be applied to the problem because the period of estimation is quite short.

The main purpose of this paper is to obtain the best estimates of the OD matrix by modelling the data imperfection, using a two-stage method. First, a conservation law of vehicles, which is represented by an inconsistent and underdetermined system of equations, is built by High-Level Petri Nets at each traffic light cycle. Second, the fuzzy sets theory is used to overcome the inconsistency of the system and to model the data imperfection. Three different approaches have been analysed to solve this system of equations: ordinary least squares, fuzzy least squares and fuzzy linear programming. A numerical study of the proposed methods has been done using the data collected in a real experimental junction fitted out with video cameras and a traffic light controller.

The rest of the paper is organised as follows. Section 2

ISBN: 978-989-95079-6-8

presents the problem and describes the real data. Sections 3 and 4 introduce the fundamental principles of High-Level Petri Nets and show how they are applied to model the OD flows through the experimental site. Section 5 proposes three methods for OD matrix estimation and the paper concludes by proposing further lines of research.

2 The problem

The experimental site is an isolated signalised junction of two double-lane roads situated in the south suburb of Paris (Fig. 1). The main road B-D which connects the suburbs to Paris has a high traffic volume, whereas the road A-C has lower traffic volume. Traffic 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 flows are statistically significant: $AC, AD, BD, BB, CB, DB, DC, DD$.

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 traffic parameters such as queue length on incoming links. The camera views are analysed in real time using image processing techniques developed at INRETS [11]. 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 *traffic light cycle* is a period of time between two sequential onsets of the red light on the main road B-D .

Many phenomena influence the quality measurement of traffic parameters. Traffic conditions (peak or off-peak periods) are the reason for many traffic count errors. If the traffic flow is heavy, the gaps between vehicles are small and it is difficult 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 measurement 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.

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*

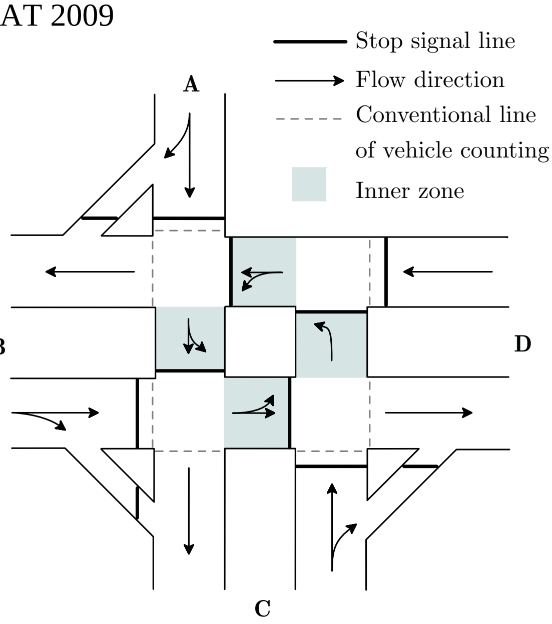

Figure 1: The experimental junction

b_{ij} 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} ($\forall i \in [\![1, n]\!], \forall j \in [\![1, m]\!]$) 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 when the vehicle flow $x_i(c)$ leaves the junction or stops at inner zones.

This paper proposes a method to obtain the values of $y_j(c)$ and $z_{kij}(c)$ from $\mathcal{Y}_j(\tau)$ and $\mathcal{Z}_k(\tau)$ respectively and then to estimate the OD flow rates. First, a tool for vehicle flow segmentation is built via two High-Level Petri Nets (HLPN). The first net indicates the set of vehicle flows which may be present in each zone at any given second. It makes possible to associate these flows to the measurements taken in the corresponding zones and provides the beginnings of the flows. The second net provides the ends of the flows. For a given cycle we thus know the duration of the presence of the flows in each zone

ISBN: 978-989-95079-6-8

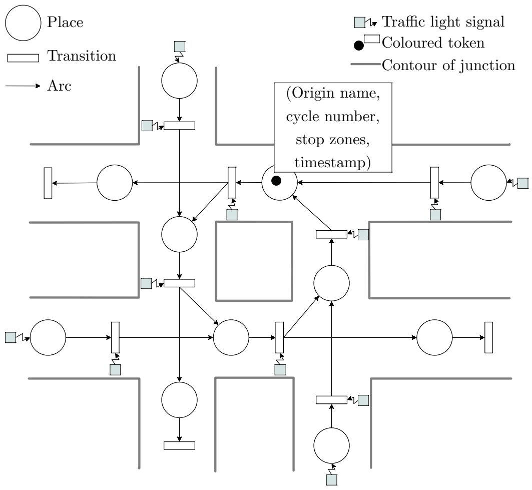

Figure 2: High Level Petri Net 1

and can collect the corresponding measurements. A consistent underdetermined system of equations, whose unknowns are the OD flow rates, can thus be built dynamically and solved at each traffic light cycle.

3 Fundamentals of High-Level Petri Nets

A High-Level Petri Net (HLPN) is a graphical and mathematical tool used to model discrete-event systems [12, 13]. In this paper, the HLPN is a classical Petri Net [14] complemented by the notions of time and colour. In graphic representation, an HLPN is a directed graph composed of two kinds of nodes: *places*, drawn by circles, and *transitions* represented as bars. The arcs connect either a place to a transition or, inversely, a transition to a place. A place may be *marked* by coloured *tokens* endowed with timestamps and characterises the system state. A transition represent an event that can change the state of the system modelled. The set of input and output places of a transition is also interpreted as a set of pre- and post-conditions of an event. When the event occurs the transition is *enabled* and the marking of input and output places changes. Thereby, the dynamic behaviour of the system is expressed by means of time-varying marking.

A 7-tuple $N = (\Sigma, \mathcal{P}, \mathcal{T}, \mathcal{A}, \mathcal{C}, \mathcal{W}, \mathcal{F})$ is a formal representation of HLPN, where:

- Σ is a finite non-empty colour set,
- $\mathcal{P}$ and $\mathcal{T}$ are the finite sets of places and transitions respectively, such that $\mathcal{P} \cap \mathcal{T} = \emptyset$ and $\mathcal{P} \cup \mathcal{T} \neq \emptyset$,
- $\mathcal{A} \subset (\mathcal{P} \times \mathcal{T}) \cup (\mathcal{T} \times \mathcal{P})$ is the finite set of arcs,
- $\mathcal{C} : \mathcal{P} \rightarrow 2^{\Sigma}$ is a *colour function* which associates a set of colours to a place,
- $\mathcal{W} : \mathcal{A} \rightarrow 2^{\Sigma}$ is an *arc function*.
- $\mathcal{F}$ is a delay associated to an arc.

A token is represented by a triple (p, s, θ), where $p \in \mathcal{P}$ is the place marked by the token, $s \subseteq \mathcal{C}(p)$ is a colour set which describes the system state in p, called *colour*, and θ is the time at which the token becomes available and the system state changes, called *timestamp*. Marking $\mathcal{M}(p, \tau)$ of place $p \in \mathcal{P}$ at time τ is a multiset represented by the pair $(S, \mathbf{n})$ where S is the set of colours of tokens arrived in p at $\tau = \theta$ and $\mathbf{n}$ is the vector every element of which is the number of occurrences of the token's colour s in set S.

Function $\mathcal{W}(p, t)$, corresponding to the arc from output place p to t, is a multiset of colours that are elements of the set $\mathcal{C}(p)$. The function $\mathcal{W}(t, p)$ is defined similarly.

Transition $t \in \mathcal{T}$ is enabled if $\mathcal{M}(p, \tau) \supseteq \mathcal{W}(p, t)$. It removes $\mathcal{W}(p, t)$ tokens from each of its input places p of t and adds $\mathcal{W}(t, p)$ tokens to its output places p with delay $\mathcal{F}(t, p)$ associated to an arc from t to p.

4 Modelling the traffic flows crossing a signalised junction

In this paper two HLPNs are used to model and segment the OD flows of vehicles crossing the junction in order to built the vehicle conservation law in a dynamic way. In the first net, a place stands for a zone of the experimental junction (Fig. 2). A token in a place is represented by a colour set $s = (s_1, s_2, s_3)$, where $s1 \in \{A, B, C, D\}$ is the name of flow origin, $s2 \in \mathbb{N}$ is a cycle number and $s3$ is a set of stop zones. In addition, it has a timestamp θ. A token in place p means the possible presence of a flow of origin $s1$ in the zone corresponding to p. Some transitions are related to the change of traffic lights in entrances and inner zones, other transitions represent the flow departures. An arc function $\mathcal{W}(t, p)$ is defined over a set of colours $s1$ representing the origins of the flows that pass through this arc. Delay $\mathcal{F}(t, p)$ is an average vehicle travel time between two successive zones of the junction corresponding to the input and output places of transition t.

The rules for changing colours of tokens, enabling transitions and marking places are defined as follows. During the green light at entrance i of the junction, the corresponding place is marked by one token with colour $s = (s1, s2, s3)$, $s_3 = \emptyset$, every second. Transition t associated to entrance i becomes enabled and the token is transmitted to output places of t. The transition associated with an inner traffic light is enabled if its input place contains at least one token and the corresponding traffic light is green. The enabled transition transmits the token downstream according to $\mathcal{W}(t, p)$ and $\mathcal{F}(t, p)$. If the light in inner zone k is red, tokens are stacked in the corresponding place. Their colours are changed by adding the name of inner zone k to the colour set s_3. At the beginning of the green light these tokens are shifted to the exit place p with predefined delay $\mathcal{F}(t, p)$. The transition related to the exit of junction is enabled when each of its input places contains at least one token.

Marking $\mathcal{M}(p, \tau)$ of place p characterizes a set of entrance flows that can be present at time τ in a junction zone corresponding to p. When the marking of p does not change during certain period of time $\Delta\tau$, all measurements taken at the corresponding zone during $\Delta\tau$ are associated to the set of flows. Thus, for a given cycle c, each exit flow $y_j(c)$ can be seg-

ISBN: 978-989-95079-6-8

mented into L_j vehicle platoons, such that

$$y_j^l(c) = \sum_{\tau \in \Delta\tau_l(c)} \mathcal{Y}_j(\tau) \qquad \forall l \in [\![1, L_j(c)]\!],$$

$$z_{kij}^l(c) = \max_{\tau \in \mathcal{R}_k(c)} \mathcal{Z}_k(\tau)\delta_{kij} \qquad \begin{array}{l} \forall l \in [\![1, L_j(c)]\!], \\ \forall i \in \mathcal{M}(p, \Delta\tau_l), \\ \forall k \text{ s.t. } \delta_{kij} = 1, \end{array}$$

where $\Delta\tau_l(c)$ is a period of time during which platoon l has left the junction, $z_{kij}^l(c)$ is a maximum number of vehicles that belong to platoon l and have stopped at inner zone k, $\mathcal{R}_k(c)$ is a duration of red light in inner zone k, δ_{kij} indicates if zone k is situated between OD pair of zones (i, j), $\mathcal{M}(p, \Delta\tau_l)$ is a marking of place $p \in \mathcal{P}$ corresponding to exit j. The number of platoons $L_j(c)$ depends on the number of entrances of the junction, and on traffic volume and traffic light command.

Note that the first occurrence of token $(p, \{s1, s2, s3\}, \theta)$ denotes that the beginning of the flow with origin $s1$ crosses the zone corresponding to place p. The ends of the flows are provided by the second HLPN, which has the identical topology and the same representation of places and transitions as the first HLPN. However, the meaning of the tokens is different. A token stands for the end of a flow and is represented by a set of two colours without a timestamp: the name of flow origin $s1 \in \{A, B, C, D\}$ and a cycle number $s2 \in \mathbb{N}$.

The onsets of the beginning and the end of a flow in place p allow us to determine the duration of the flow presence in zone related to p. Thus we can collect the measurements made in this zone and, at each cycle c, can built the following set of equations

$$y_j^l = \sum_{i \in \mathcal{M}(p_j, \Delta\tau_l)} b_{ij}^l x_i \qquad \forall j \in [\![1, m]\!], \forall l \in [\![1, L_j]\!], \tag{2a}$$

$$x_i b_{ij}^l \geq z_{kij}^l \qquad \begin{array}{l} \forall j \in [\![1, m]\!], \forall l \in [\![1, L_j]\!], \\ \forall i \in \mathcal{M}(p_j, \Delta\tau_l), \\ \forall k \text{ s.t. } \delta_{kij} = 1, \end{array}, \tag{2b}$$

$$\sum_{j,l} b_{ij}^l = 1 \qquad \forall i \in [\![1, n]\!], \tag{2c}$$

$$b_{ij}^l \geq 0 \qquad \begin{array}{l} \forall j \in [\![1, m]\!], \forall l \in [\![1, L_j]\!], \\ i \in \mathcal{M}(p_j, \Delta\tau_l), \end{array} \tag{2d}$$

where the cycle c has been omitted to simplify the notations.

Let J and K be the numbers of constraints (2a) and (2b) respectively, and I be the number of unknowns b_{ij}^l, such that

$$J = \sum_{j=1}^{m} j L_j,$$

$$K = \sum_{j=1}^{m} \sum_{l=1}^{L_j} \sum_{i \in \mathcal{M}(p_j, \tau_l)} \sum_{k=1}^{p} \delta_{kij},$$

$$I = \sum_{j=1}^{m} \sum_{l=1}^{L_j} \sum_{i \in \mathcal{M}(p_j, \tau_l)} i.$$

Since in general $I \geq J$ (in this paper $I \approx 10$), the obtained system of equations (2a) is underdetermined.

5 OD matrix estimation

The estimation problem of OD flow rates b_{ij}^l from the relationship (2a) can be viewed as an estimation of coefficients of the regression model

$$\mathbf{y} = \mathbf{X_1 b}, \tag{3}$$

where $\mathbf{y}$ is a vector composed of J output variables y_j^l, $\mathbf{b}$ is a vector containing the I unknowns b_{ij}^l, $\mathbf{X_1}$ is an $J \times I$ matrix rearranged in such a way that (2a) corresponds to (3).

The elements of model (3) must satisfy the constraints:

$$\mathbf{X_2 b} \geq \mathbf{z}, \tag{4a}$$

$$\mathbf{I b} = \mathbb{1}, \tag{4b}$$

$$\mathbf{b} \geq 0, \tag{4c}$$

where $\mathbf{z}$ is a vector composed of K variables z_{ijk}^l, $\mathbf{X_2}$ is an $K \times I$ matrix built so that (2b) is equal to (4a), $\mathbf{I}$ is an indicator $n \times I$ matrix organised in such a manner that (4b) is equivalent to (2c) and $\mathbb{1}$ is an identity n-vector.

Considering $\mathbf{b}$ to be an m-vector of crisp elements, three situations are studied in this section of the paper:

A) values of $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{y}$ and $\mathbf{z}$ are crisp

B) values of $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{y}$ and $\mathbf{z}$ are fuzzy

C) values of $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{y}$ and $\mathbf{z}$ are crisp, relationships (3) and (4a) are fuzzy, where the strict mathematical relations are replaced by fuzzy equivalents $\lessapprox$ and $\gtrapprox$ that are read “essentially smaller or equal to” and “essentially greater than or equal to”.

In the first case, we use the least squares method to estimate the crisp regression coefficients $\mathbf{b}$. The fuzzy least squares (FLS) model based on commonly used Diamond ρ_2-metric [15] has been applied to the second situation. For the last case, we formulate a fuzzy linear programming problem (FLP) following the approach proposed by Zimmernann [16]. We present in the sequel the three methods and their results on our application.

5.1 *Least squares method*

Let the input and output variables $\mathbf{X}_1$ and $\mathbf{y}$ of the regression model (3) be non-negative crisp numbers. The unknown coefficients $\mathbf{b}$ of the model and the elements of $\mathbf{X}_2$ and $\mathbf{z}$ are also considered to be crisp. A first approach to estimate $\mathbf{b}$ is to use a least squares method. To insure the existence of feasible solutions, we introduce slack variables, e_i ($\forall i = 1, \ldots, K$), and we propose to solve the following problem:

$$\min_{\mathbf{b},\mathbf{e}} \| \mathbf{y} - \mathbf{X_1 b} \|^2 + \sum_{i=1}^{K} e_i \tag{5}$$

subject to

$$\begin{cases} \mathbf{X_2 b} + \mathbf{e} \geq \mathbf{z}, \\ \mathbf{I b} = \mathbb{1}, \\ \mathbf{b} \geq 0, \\ \mathbf{e} \geq 0. \end{cases} \tag{6}$$

5.2 *Fuzzy least squares method*

To take into account model errors and the inherent imprecision of data measurements, we choose to represent the input and output variables of the regression model (3) by triangular fuzzy numbers $\tilde{\mathbf{X}}_1 = (\underline{\mathbf{X}}_1, \mathbf{X}_1^m, \overline{\mathbf{X}}_1)$, $\tilde{\mathbf{y}} = (\underline{\mathbf{y}}, \mathbf{y}^m, \overline{\mathbf{y}})$ and

ISBN: 978-989-95079-6-8

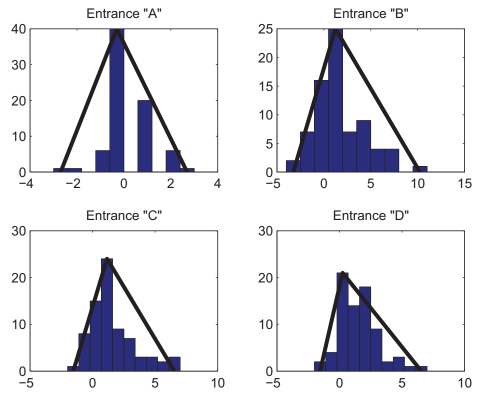

Figure 3: Empirical distributions of error counts for the entrances of the junction

$\tilde{\mathbf{z}} = (\underline{\mathbf{z}}, \mathbf{z}^m, \overline{\mathbf{z}})$, where $[\underline{a}, \overline{a}]$ is the support and a^m is the mode of fuzzy number $\tilde{a}$. The form of fuzzy numbers $\tilde{\mathbf{X}}_1$ was derived from empirical distributions of the error counts as shown in Fig. 3. In most cases, counts of vehicles at the entrances are smaller than the true value of $\mathbf{X}_1$. The fuzzy number $\tilde{\mathbf{X}}_2$ is determined similarly to $\tilde{\mathbf{X}}_1$. Since there is no available histograms of error counts for $\mathbf{y}$ and $\mathbf{z}$, the fuzzy numbers $\tilde{\mathbf{y}}$ and $\tilde{\mathbf{z}}$ are supposed to be symmetrical with the spreads chosen experimentally. Note that $\underline{\mathbf{X}}_1, \underline{\mathbf{X}}_2, \underline{\mathbf{y}}, \underline{\mathbf{z}} \in \mathbb{R}^+$ because of the nature of the data.

According to the fuzzy least squares (FLS) method proposed by Diamond [15, 17] and supposing the unknown coefficients $\mathbf{b}$ to be crisp, the following minimisation problem is written:

$$\min_{\mathbf{b}} \left(\| \underline{\mathbf{y}} - \underline{\mathbf{X}}_1 \mathbf{b} \|^2 + \| \mathbf{y}^m - \mathbf{X}_1^m \mathbf{b} \|^2 + \right. \\ \left. + \| \overline{\mathbf{y}} - \overline{\mathbf{X}}_1 \mathbf{b} \|^2 \right) \tag{7}$$

subject to

$$\begin{cases} \overline{\mathbf{X}}_2 \mathbf{b} \geq \underline{\mathbf{z}}, \\ \mathbf{I}\mathbf{b} = \mathbb{1}, \\ \mathbf{b} \geq 0. \end{cases} \tag{8}$$

Note that the first constraint in (8) is the least conservative translation of the strong relation (4a).

5.3 *Fuzzy linear programming approach*

Another approach to estimate the OD flow rates is to model the constraints (3) and (4a) by fuzzy sets. Assuming that the values of $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{y}$ and $\mathbf{z}$ are strict numbers, we consider the estimation problem of $\mathbf{b}$ that satisfies:

$$\mathbf{X}_1 \mathbf{b} \gtrapprox \mathbf{y} \quad -\mathbf{X}_1 \mathbf{b} \gtrapprox -\mathbf{y}, \tag{9a}$$
$$\mathbf{X}_2 \mathbf{b} \gtrapprox \mathbf{z}, \tag{9b}$$
$$\mathbf{I}\mathbf{b} = \mathbb{1}, \tag{9c}$$
$$\mathbf{b} \geq 0, \tag{9d}$$

where the sign $\gtrapprox$ means that we accept a small violation of the strict relation $\geq$ in a sense that is described below..

Denoting

$$\mathbf{H} = (-\mathbf{X}_1, \mathbf{X}_1, -\mathbf{X}_2)^\top,$$
$$\mathbf{h} = (-\mathbf{y}, \mathbf{y}, -\mathbf{z})^\top,$$

we write (9a)-(9b) as

$$\mathbf{H}\mathbf{b} \lessapprox \mathbf{h} \tag{10}$$

Each row i of (10) is assumed to be a fuzzy set with a membership function $\mu_i(\mathbf{b})$:

$$\mu_i(\mathbf{b}) = \begin{cases} 1 & H_i\mathbf{b} \leq h_i \\ 1 - \frac{H_i\mathbf{b} - h_i}{\xi_i} & h_i < H_i\mathbf{b} \leq h_i + \xi_i \\ 0 & H_i\mathbf{b} > h_i + \xi_i \end{cases} \tag{11}$$

where each $\xi_i > 0$ is a given constant reflecting an acceptable degree of constraint violation, $i = 1, \ldots, 2J + K$. Here $\mu_i(\mathbf{b})$ is defined as the degree to which $\mathbf{b}$ satisfies the relationship i.

With respect to Bellman-Zadeh decision-making rule for fuzzy sets [18], the membership function of the fuzzy decision set $\tilde{B}$ is defined as follows

$$\mu_{\tilde{B}}(\mathbf{b}) = \min_i (\mu_i(\mathbf{b})) = \min_i \left(1 - \frac{H_i\mathbf{b} - h_i}{\xi_i} \right), \\ \forall i = 1, \ldots, 2J + K. \tag{12}$$

According to the symmetric FLP method proposed by Zimmermann [16, 19], the crisp optimal solution can be given by

$$\max_{\mathbf{b} \geq 0, \mathbf{I}\mathbf{b} = \mathbb{1}} \mu_{\tilde{B}}(\mathbf{b}) \tag{13}$$

which can be obtained by solving the following problem of linear programming:

$$\max_{\lambda, \mathbf{b}} \lambda \tag{14}$$

subject to

$$\begin{cases} \lambda < 1 - \frac{H_i\mathbf{b} - h_i}{\xi_i} & \forall i = 1, \ldots, 2J + K, \\ \mathbf{I}\mathbf{b} = \mathbb{1}, \\ \mathbf{b} \geq 0. \end{cases} \tag{15}$$

5.4 *Results*

These methods have been tested using the *raw data* collected at the experimental junction. The estimation is made on 25 consecutive traffic light cycles, equivalent to 30 minutes. In addition, the actual values of OD flow rates b_i^* ($\forall i = 1, \ldots, 8$), derived manually from video images, are available for all cycles.

Since the measure of the number of stopped vehicles is more accurate than the measure of exit flow volume, the following parameters of the FLS method where chosen: $\mathbf{y}^m - \underline{\mathbf{y}} = \overline{\mathbf{y}} - \mathbf{y}^m = 3$, $\mathbf{z}^m - \underline{\mathbf{z}} = \overline{\mathbf{z}} - \mathbf{z}^m = 2$. The acceptable degree of constraint violation ξ_i ($\forall i = 1, \ldots, 2J + K$) in FLP method has been fixed to 5 for constraints (9a) and to 3 for constraints (9b).

The estimation error has been calculated for the OD flow rates (Fig. 4): $E = \mathbf{b}\mathbf{X}^* - \mathbf{b}^*\mathbf{X}^*$, where $\mathbf{X}^*$ is a vector of actual vehicle counts at the entrances of the experimental junction. The best results are obtained with the FLS method for which the median error is almost zero for all OD flows. The results of the least squares method are less accurate than those of the FLS method. Surprisingly, the FLP method, although it seems to be a more natural fuzzy translation of the crisp initial problem, does not provide very good results. Note that the estimation error E of all methods is higher, if the flow volume is lower, like for the OD flows “AD” and “BB”.

ISBN: 978-989-95079-6-8

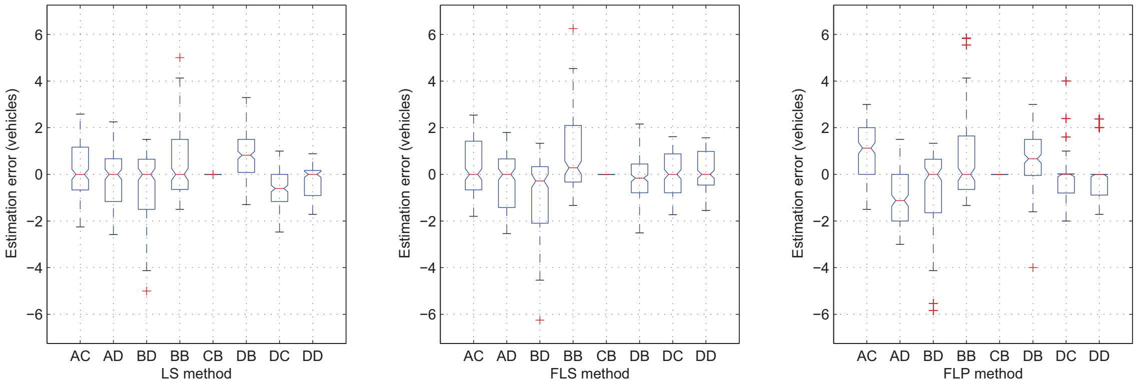

Figure 4: Estimation errors for 8 OD flow rates

6 Conclusion and future lines of research

In this paper we have proposed a new short-time estimation method of the OD matrix for a signalised junction. We have built a model of the segmentation of traffic flows crossing the junction in order to draw up a vehicle conservation law, represented by a underdetermined system of equations, for each traffic light cycle. The model, made using High-Level Petri Nets, is based on the use of traffic lights.

Real traffic measurements collected at the real experimental junction are used to estimate the OD flow rates. The experimental comparison shows that a small gain is obtained when the data imprecision and uncertainty is taken into account using a fuzzy modelling. Among the two proposed fuzzy approaches, the fuzzy least squares yields the most accurate results. In spite of the important imprecision of the real data, the first results are very encouraging to continue the research on improvement of the estimation accuracy.

The main future line of research is to fuzzify the token timestamps of the first HLPN in order to model the temporal imprecision of data measurements. The application of our method should be also extended to a sequence of junctions.

References

[1] S. Midenet *et al.* Signalized intersection with real-time adaptive control: on-field assessment of CO_2 and pollutant emission reduction. *Transportation Research Part D*, 9:29–47, 2004.

[2] H. J. Van Zuylen. The estimation of turning flows on a junction. *Traffic Engineering Control*, 20(11):539–541, 1979.

[3] E. Hauer, E. Pagitsas, and B. T. Shin. Estimation of turning flows from automatic counts. *Transportation Research Record*, (795):1–7, 1981.

[4] M. J. Maher. Inferences on trip matrices from observations on link volumes: a Bayesian statistical approach. *Transportation Research Part B*, 17(6):435–447, 1983.

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

[6] N. L. Nihan and G. A. Davis. Recursive estimation of origin-destination matrices from input/output counts. *Transportation Research Part B*, 21(2):149–163, 1987.

[7] B. Li and B. De Moor. Recursive estimation based on the equality-constrained optimization for intersection origin-destination matrices. *Transportation Research Part B*, 33(3):203–214, 1999.

[8] H. D. Sherali *et al.* Parameter optimization methods for estimating dynamic origin-destination trip-tables. *Transportation Research Part B*, 31(2):141–157, 1997.

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

[10] I. Matschke and B. Friedrich. Dynamic OD estimation using additional information from traffic signal lights timing. *TRISTAN IV Conference*, 2001.

[11] D. Aubert *et al.* Automatic vehicle queue measurement at intersections using image-processing. *8th IEE International Conference on Road Traffic Monitoring and Control, London*, 422:100–104, 1996.

[12] M. Dotoli and M. P. Fanti. An urban traffic network model via coloured timed Petri nets. *Control Engineering Practice*, 14:1213 – 1229, 2006.

[13] K. Jensen *et al.* Coloured Petri nets and CPN tools for modelling and validation of concurrent systems. In *International Journal on Software Tools for Technology Transfer*, pages 98–132. Springer Verlag, 1998.

[14] T. Murata. Petri nets: properties, analysis and applications. In *Proceedings of the IEEE*, volume 77, pages 541–580, April 1989.

[15] P. Diamond and R. Korner. Extended fuzzy linear models and least squares estimates. *Computers and Mathematics with Applications*, 33(9):15–32, May 1997.

[16] H.-J. Zimmermann. Description and optimization of fuzzy systems. *International Journal of General Systems*, 2:209–215, 1976.

[17] R. Slowinski, editor. *Fuzzy sets in decision analysis, operation research and statistics*. The handbooks of fuzzy sets. Kluwer Academic Publishers, 1998.

[18] R. Bellman and L. A. Zadeh. Decision-making in a fuzzy environment. *Management Science*, 17:141–164, 1970.

[19] H.-J. Zimmermann, editor. *Fuzzy set theory and its applications*. Kluwer Academic Publishers, second edition, 1991.

ISBN: 978-989-95079-6-8