Transportation Research Record 1607 (1997)

Dynamic OD-Matrix Estimation from Traffic Counts and Automated Vehicle Identification Data

Nanne van der Zijpp
Transportation Engineering Section, Dept. of Civil Engineering
Delft University of Technology[1]
P.O.Box 5048
2600 GA Delft
The Netherlands

Email zijpp@ct.tudelft.nl

Abstract- The problem of estimating time-varying origin-destination (OD) matrices from time-series of traffic counts is extended allow for the usage of partial observations of vehicle trajectories. These can be obtained from automated vehicle identification (AVI) techniques such as automated license plate recognition, but may also originate from floating car data. The problem definition that is central to the paper allows for the use of data from induction loops and AVI equipment at arbitrary (but fixed) locations in the network, and allows for the presence of errors in all traffic counts and less then perfect recognition rates at the AVI-stations. Analysis of this problem leads to an expression for the mutual dependencies between link volume observations and AVI data, and the formulation of an estimation problem with inequality constraints. A number of traditional estimation procedures such as discounted constrained least squares (DCLS) and the Kalman filter are described, and a new procedure referred to as Bayesian updating is proposed. The advantage of this new procedure is that it deals with the inequality constraints in an appropriate statistical manner. Experiments with a large number of synthetic datasets show in all cases a reduction of the error of estimation due to usage of trajectory counts, and compared to the traditional DCLS and Kalman filtering methods, a superior performance of the Bayesian updating procedure.

Keywords- Dynamic OD-Matrix, AVI, License Plate Surveys, Bayesian Inference

1. A part of the research upon which the paper is based was done at the Centre for Transport Studies, University College London, under an EC grant in the HCM programme, contract ERBCHGCT930412

Dynamic OD-Matrix Estimation from Traffic Counts and Automated Vehicle Identification Data

Nanne van der Zijpp
Transportation Engineering Section, Dept. of Civil Engineering
Delft University of Technology

1. Introduction

This paper proposes a method to track time-varying traffic patterns from a combination of link volume counts and trajectory observations. Such a method may be applied if accurate information about Origin-Destination (OD) travel demand is needed on-line, for example as a part of a short term traffic forecasting system on motorways.

Whereas link volume counts are collected routinely by road authorities, observations of trajectories or parts of trajectories of individual vehicles are much harder to obtain in a cost effective manner. This is likely to change in the near future due to the availability of Automated Vehicle Identification (AVI) techniques. AVI may be based on various mechanisms, for example the usage of a transponder on the vehicle, the usage of automated license plate readers, or vehicles transmitting their trajectories.

In this paper license plate survey data is used as an example of trajectory information. Although observing vehicle trajectories comes close to directly observing OD-tables, a well known problem with license plate surveys is the presence of recording errors which, if no correction is applied, may lead to an overestimation of the number of trips and an underestimation of the average trip length. It is hypothesised that data from which the total number of trips can be deduced, e.g. link volume data, can be effectively used to complement license plate data.

The methods that are discussed in this paper depart from the assumption that with every entry of a transport network, a set of probabilities is associated with which the reachable destinations are selected by motorists. These probabilities are referred to as *split probabilities*.

Until now these methods have only been applied to the problem of estimating OD-matrices from time-series of traffic counts, see e.g. *Cremer and Keller* (*1981,1987*), *Nihan and Davis* (*1987, 1989*), *Cascetta et. al.*, (*1993*), and *Van Der Zijpp and Hamerslag* (*1994*). Like in above mentioned work, the considerations in this paper are limited to networks in which the route choice issue can be ignored. It should be noted though, that this is not a fundamental limitation to this type of methods.

The paper is structured as follows: Section 2 contains the problem definition and notation. Section 3 analyses the problem of estimating OD-matrices from time series of traffic counts. In this section a measurement equation is derived and a result concerning the statistical properties of the measurement error is presented. Section 4 presents a new method for the estimation of parameters in a system with linear observations and in the presence of inequality constraints, and also describes a number of traditional estimation procedures such as the discounted constrained least squares (DCLS) method and the Kalman filter. Section 5 contains a short overview of OD-estimation from license plate surveys and describes the approach taken in this paper towards the use of license plate surveys. In section 6 it is shown that the main part of the analysis presented in section 3 also applies to OD-estimation from combined traffic counts - trajectory counts. Section 7 contains evaluation results based on simulated data.

2. Problem definition

A transport network is considered in which one route exists per connected Entry-Exit (EE) pair. A typical example of such a network is a motorway corridor, see figure 1. It is assumed that time series of observations of all entering volumes and a subset of the internal link volumes are available. These data may contain observation errors. In addition to this, the presence of license plate recording stations is assumed. These stations produce lists of vehicle registrations. It is assumed that only a fraction of the vehicles are correctly identified. This fraction is referred to as the recognition rate and is assumed to be known. Typical values of this rate are in the range 50-90%. There are two remarks that need to be made at this point:

- In the present paper the recognition rate is presented as a property of only the license plate reader, implying that up- and downstream recognition rates are independent random variables. In reality the recognition rate is almost certain a combined property of a vehicle, license plate reader and other circumstances (such as weather conditions). This leads to correlation between up- and downstream recognition rates. An appropriate way to model this, would be to specify the recognition rate as a product of vehicle specific, location specific and other factors. Research into this direction will be reported on in a future paper.
- The recognition rate may not be exactly known. The derivations given in this paper depart from the assumption that the recognition rate, i.e. the *probability* of recognizing a vehicle, is exactly known. However, analytically it is feasible to make similar derivations based on the assumption that instead, the mean and the variance of the recognition rate is known.

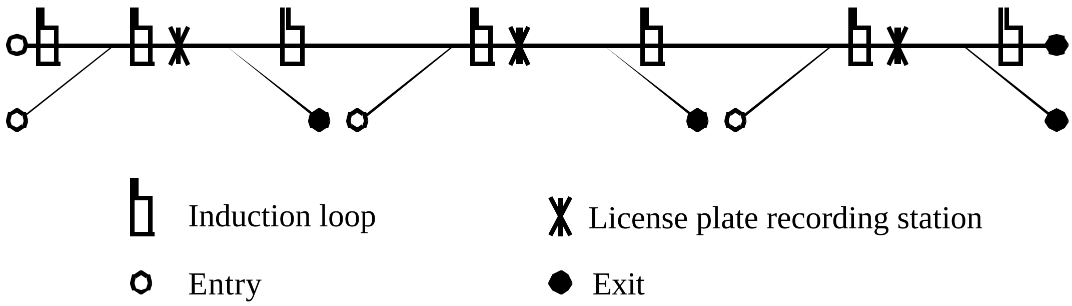

Figure 1: Sample network. Traffic data are collected with induction loops and license plate readers.

If travel-times cannot be neglected, a first approximation is to let the partitioning of the time-axis depend on the location, resulting in *moving time coordinates*, see *Van Der Zijpp* (*1996*). Applying moving time coordinates requires no fundamental modification to the methods presented in this paper. Further refinements are obtained by taking travel-time *dispersion* into account, see e.g. *Bell* (*1991*b) and *Chang and Wu* (*1994*), but such approaches require the estimation of extra parameters and hence do not lead to a reduced error of estimation by default. In the present paper, travel-time dispersion is not explicitly modelled but is accounted for by the introduction of *assignment errors*, i.e. vehicles being counted in two different periods.

A notation is used in which most quantities are expressed as vectors:

m, n, l, h	Number of entries, number of exits, number of link volume observations (excluding entry volumes), and number of license plate readers
i, j, k	Indices corresponding to entries, exits, and link volume observations
$\tilde{q}(t)$	Entry flow vector (length m). This vector represents the true entry volumes in period t, t=1,2,…
$q(t)$	Vector of entry volume observations in period t
$f_{ij}(t)$	Number of trips for EE-pair i-j with departure period t, i=1,…m, j=1,…n.
$\tilde{y}(t)$	Idealized link flow vector (length l). Element k of this vector contains the number of trips with departure period t that traverse location k.
$y(t)$	Vector of link volume observations in period t
t	Path-link observation incidence map. τ_{ijk}=1 if route i-j uses link k and zero otherwise, k=1,2,…l
$e_{ab}(t)$	Trajectory count. Number of vehicles observed during period t at location a and (not necessarily in period t) at location b, but not at any site upstream of a or downstream of b. The indexing of the license plate readers is chosen in such a way that $a<b$ implies that a is not reachable from b.
α_k	recognition rate at site k.
k	Path-license plate reader incidence map. κ_{ija}=1 if route i-j traverses license plate reader a and zero otherwise, a=1,2,…h

The errors in the entry volume and link volume observations are represented by $r(t)$ and $s(t)$ respectively. These errors are assumed to be zero mean random variables of which the second moments are given. The observation mechanism is described by the following equations:

$$q(t) = \tilde{q}(t)+r(t)$$
$$\mathrm{E}[r(t)] = \mathbf{0}$$
$$\mathrm{E}[r(t)\, r(p)'] = \Phi\delta_{tp} \tag{1}$$

and:

$$y(t) = \tilde{y}(t)+s(t)$$
$$\mathrm{E}[s(t)] = \mathbf{0}$$
$$\mathrm{E}[s(t)\, \mathrm{s}(p)'] = \Theta\delta_{tp} \tag{2}$$

where Φ and Θ are matrices of appropriate size.

The objective is to estimate the EE-flows $f_{ij}(t)$ on the basis of the observations that are available up to and including period t.

3. Dynamic OD-estimation from time series of traffic counts

Analysis

The problem of determining the OD-flows from the observations described in the previous section is an *under-specified* problem, i.e. many OD matrices exist that would reproduce the observations. In literature various mechanisms are described that could be used to select a 'best' matrix, e.g. maximizing entropy or calibrating a model of travel demand. In this paper the assumption of slowly varying split probabilities is used. According to this idea, with every entry a set of probabilities is associated with which destinations are selected. These probabilities are represented by $b_{ij}(t)$, i.e.: $b_{ij}(t)$ is the probability that a vehicle that enters at entry i in period t will leave the network at exit j. The split probabilities may vary slowly over time, satisfying the following random walk model:

$$b(t+1)=b(t)+w(t) \tag{3}$$

where $b(t)$ represents the *vector* of split probabilities. Within each period the split probabilities are assumed to remain constant. The vector $w(t)$ represents a zero mean random variable of which the covariance matrix is denoted by S_t, i.e.:

$$\begin{gathered} \mathrm{E}[w(t)]=\mathbf{0} \\ \mathrm{E}[w(t)\ \mathrm{w}(p)'] = S_t\delta_{tp} \end{gathered} \tag{4}$$

The split probabilities that correspond to one origin should be non-negative, not larger then unity, and add up to unity. This is expressed in the following equations:

$$\mathbf{0}\le b(t)\le\mathbf{1} \tag{5}$$

$$\mathrm{p}'b(t)=\mathbf{1} \tag{6}$$

where p is a matrix of which the nonzero elements are defined by:

$$\mathrm{p}_{\varphi(i,j),i} = 1 \qquad i=1,\ldots m,\ \ j=1,\ldots n \tag{7}$$

In this equation $\varphi(i,j)$ represents the order of the split probabilities $b_{ij}(t)$ in the vector $b(t)$, i.e. $b_{\varphi(i,j)}(t)=b_{ij}(t)$. An example of a mapping that can be used for this purpose is:

$$\varphi(i,j)=(n\text{-}1)i+j \tag{8}$$

This mapping corresponds to $b(t) = \mathrm{vec}(\ [b_{ij}(t)]\ )$, i.e. the elements of the matrix $[b_{ij}(t)]$ are rearranged column by column in a vector. The assumption of slowly varying split probabilities imposes a minimum of restrictions on the OD-flows but at the same time constitutes a sufficient basis for the statistical estimation of OD-flows from time series of traffic counts. Applications of the assumption go back to *Cremer and Keller (1981).* The assumption of piecewise-constant split probabilities implies that the distribution of the OD-flows with departure period t is given by the following multinomial distribution, see e.g. *Nihan and Davis (1989)*:

$$\mathrm{P}[f(t)|\tilde{q}(t),b(t)]=\prod_{i=1}^{m}\tilde{q}_i(t)!\prod_{j=1}^{n}\frac{b_{ij}(t)^{f_{ij}(t)}}{f_{ij}(t)!} \tag{9}$$

In this equation $f(t)$ represents the vector of EE-flows with departure period t. The elements of the idealized link flow vector are the sum of all flows with departure period t that pass a given location. This is expressed in with:

$$\tilde{y}(t) = U' f(t) \tag{10}$$

where U is a matrix of which the elements are defined by:

$$U_{\varphi(i,j),k} = \tau_{ijk}$$
$$i=1,\dots m,\ \ j=1,\dots n,\ \ k=1,\dots l$$

From equations (9) and (10) no tractable probability distribution of the idealized link volume vector follows. Instead the statistical properties of this vector are described in terms of their first and second moments. Combining equations (9) and (10) with the assumptions about observation errors given in equations (1) and (2) can be shown to result in:

$$\mathrm{E}[y(t)|q(t)] = H'(t)b(t) \tag{11}$$

where $H(t)$ is a matrix of which the non-zero elements are defined by:

$$H_{\varphi(i,j),k}(t) = \tau_{ijk}\, q_i(t)$$
$$i=1,\dots m,\ \ j=1,\dots n,\ \ k=1,\dots l \tag{12}$$

This gives rise to specification of the following *measurement equation*:

$$y(t)=H'(t)b(t)+v(t) \tag{13}$$

This equation can be used when estimating the split probabilities $b(t)$. The vector $v(t)$ represents the zero mean measurement error that accounts for mis-specification of the matrix $H(t)$, i.e. observation errors contained in $q(t)$, random effects due to the uncoordinated choices of motorist, and observation errors, i.e. the difference between the idealized link volume vector $\tilde{y}(t)$ and the observation $y(t)$. The fact that the problem definition poses no restrictions on the locations on which traffic is counted and a number of other factors imply that elements of the vector $v(t)$ are dependent. The covariance matrix of $v(t)$ is denoted with R_t, i.e.:

$$\mathrm{E}[v(t)]=\mathbf{0}$$
$$\mathrm{E}[v(t)\ \mathrm{v}(p)'] = R_t \tag{14}$$

Statistical analysis of the random variable $v(t) \equiv y(t)-H'(t)b(t)$, as presented in *Van Der Zijpp* (*1996*) has revealed that:

$$R_t=\mathrm{E}[\ v(t)\ v(p)'\ |b(t),q(t)] =(\ U'\mathrm{cov}[f(t),f(t)|b(t),q(t)]U + \Theta\ )\delta_{tp},$$
$$\mathrm{cov}[f(t),f(t)|b(t),q(t)]=\boldsymbol{Q_t}(\boldsymbol{B_t}-\boldsymbol{B_t}\mathrm{pp}'\boldsymbol{B_t}')+\boldsymbol{B_t}\mathrm{p}\Phi\mathrm{p}'\boldsymbol{B_t} \tag{15}$$

where $\boldsymbol{Q_t}$=diag(p$q(t)$), $\boldsymbol{B_t}$=diag($b(t)$) and diag(.) is the operator that takes a vector as its input and outputs a diagonal matrix with the elements of this vector on its diagonal. Evaluation of (15) requires knowledge of $b(t)$ which is generally not available in a procedure aimed at estimating $b(t)$. Instead the expression can be used as a way to approximate the covariance matrix of $v(t)$ that applies in absence of knowledge of $b(t)$, replacing $b(t)$ with its most up to date estimate. Such a matrix is hence referred to as a *point-estimate based approximation*.

As an alternative to usage of an approximation based on (15), tests are also performed with the following simplified approximation for R_t (see section 7):

$$R_t=\mathrm{diag}(\bar{y}) \tag{16}$$

where $\bar{y}$ is the average of the observations $y(t)$. Usage of this matrix is equivalent to assuming that the elements of the observation error vector are mutually independent. Estimators based on (15) and (16) will be compared in section 7.

4. Estimation procedure

Various approaches are possible to estimate the split probabilities $b(t)$ in the system described by equations (3), (4), (5), (6), (13) and (14). In this section we describe methods based on discounted constrained least squares, Bayesian updating and the Kalman filter.

<u>Discounted constrained least squares (DCLS)</u>

The DCLS estimate for the split probabilities (*Nihan and Davis*, *1987* and *Cremer and Keller*, *1987*) is given by:

$$\bar{b}(t) \equiv \underset{\substack{b \\ \mathbf{0} \leq b(t) \leq \mathbf{1} \quad p' b(t) = \mathbf{1}}}{\operatorname{argmin}} J(b, t) \tag{17}$$

where:

$$J(b, t) = \sum_{k=1}^{t} \lambda^{t-k} \| y(k) - H'(k) b \|^2 \tag{18}$$

The parameter λ, $0 \leq \lambda \leq 1$, determines the weight that is put on older observations. The advantage of the DCLS method is its ease of implementation. A disadvantage is its weak theoretical basis: it is not possible to exploit any of the details specified in the previous section, such as equations (1) or (2), nor can knowledge of the joint probability distribution of the EE-flows given by equation (9) be used.

<u>Bayesian approach</u>

The *subjective probability distribution*, given by $p[b(t)|y(1), y(2) \ldots y(t)]$, is central to the Bayesian approach. This distribution encapsulates all knowledge about $b(t)$ available at time t. Point estimates for $b(t)$ may be derived by maximizing this expression, resulting in a *Maximum APosteriori* (MAP) estimate, or better, by taking the expectation, resulting in the *subjective expectation*. Given the subjective probability distribution, the subjective expectation can be shown to be a minimum variance unbiased estimate.

Information contained in a new observation may be incorporated in the subjective probability distribution using *Bayes'* rule:

$$p[b(t)|y(1) \ldots y(t)] = \frac{p[y(t)|b(t), y(1) \ldots y(t-1)] \cdot p[b(t)|y(1) \ldots y(t-1)]}{p[y(t)|y(1) \ldots y(t-1)]} \tag{19}$$

The denominator of expression (19) is referred to as the normalisation constant as it is invariant for $b(t)$. The numerator consists of the product of the Likelihood function (left) and the prior distribution (right). It is well known that if both the prior distribution and the likelihood function are Multivariate Normal (MVN), the posterior distribution is also MVN. Moreover if $v(t)$ satisfies (14) and is MVN, the expectation and covariance matrix characterising the posterior distribution are given by:

$$\begin{aligned}
\bar{b}(t)^{+} &= \bar{b}(t)^{-} + K_t[y(t) - H(t)' \bar{b}(t)^{-}] \\
K_t &= \Sigma_b(t)^{-} H(t)[H(t)' \Sigma_b(t)^{-} H(t) + R_t]^{-1} \\
\Sigma_b(t)^{+} &= \Sigma_b(t)^{-} - \Sigma_b(t)^{-} H(t)[H(t)' \Sigma_b(t)^{-} H(t) + R_t]^{-1} H(t)' \Sigma_b(t)^{-}
\end{aligned} \tag{20}$$

where $\overline{b}(t)^{-}$ and $\Sigma_b(t)^{-}$ are the *apriori* mean and covariance, and $\overline{b}(t)^{+}$ and $\Sigma_b(t)^{+}$ are the *aposteriori* mean and covariance. This measurement update, supplemented with time extrapolation equations is known as the Kalman filter (*Kalman*, *1960*).

However, the MVN distribution is not a very suitable way to represent the knowledge about split probabilities, as it does not reflect the inequality constraints (5). Also, usage of (20) could easily lead to negative estimates. This problem was already recognised in *Nihan and Davis* (*1989*), where several heuristic algorithms were proposed to constrain the estimates to the feasible region. A similar observation was made in the context of static OD-estimation in *Bell* (*1991*a). A remedy for this problem presented in *Van Der Zijpp and Hamerslag* (*1994*) is the usage of a *Truncated* Multivariate Normal (TMVN) distribution.

A truncated probability distribution is derived from its original by redistributing the probability mass that was originally assigned to points outside the truncation interval proportionally over the points inside this interval. The use of a truncated distribution is indicated if one beliefs that a distribution provides a reasonable model for a phenomenon inside the truncation interval while at the same time one knows that the phenomenon can never take values outside this interval. A truncated distribution is characterized by a subset of the parameters that define the original. For example, the truncated MVN distribution is characterized by a vector and a matrix. Unlike the non-truncated MVN distribution these parameters do not correspond directly to the mean and variance of the truncated distribution.

If in (19) the prior distribution is TMVN and the likelihood function is MVN, it can be shown that the posterior distribution remains in the class of TMVN distributions. Moreover, equation (20) still defines the parameters that characterise the posterior distribution (*Van Der Zijpp and Hamerslag*, *1994*).

This is illustrated graphically in figure 2. The posterior distribution that is obtained by multiplying a normally distributed prior distribution and likelihood function and normalizing the result is a normal distribution, see top graph. If the prior distribution is replaced with a TMVN distribution, see bottom graph, the shape of the resulting posterior distribution remains unchanged inside the truncation interval, and hence is characterised by parameters identical to those obtained when using an MVN distributed prior.

A practical difficulty is the computation of the mean associated with a TMVN distribution. No analytical solution exists for the multidimensional integral that needs to be solved, nor is numerical integration an option due to CPU time constraints. However the mean may be approximated by taking the average value of a large number of random numbers drawn from a TMVN distribution. TMVN random numbers may be generated simply by generating MVN random numbers and rejecting all outcomes that do not satisfy (5). Details of this technique are described in *Van Der Zijpp* (*1996*)

In the above, only the measurement update has been discussed. In addition to this a time extrapolation is needed to account for changes of the split probabilities over time such as described by (3). An approximation based on the Kalman time extrapolation equations is used:

$$\begin{gathered} \mathrm{p}[b(t+1)|Y(t)] \approx \mathrm{TMVN}[\,\overline{b}(t+1)^{-}, \Sigma_b(t+1)^{-}\,]\Big|_{b(t+1)} \\ \text{with:} \\ \overline{b}(t+1)^{-}=\overline{b}(t)^{+} \\ \Sigma_b(t+1)^{-}=\Sigma_b(t)^{+}+S_t \end{gathered} \tag{21}$$

The error that is contained in this approximation vanishes if $w(t)$ is zero. A discussion on the sizes of this error can be found in *Van Der Zijpp (1996)*.

A final remark concerns the usage of the equality constraints (6). These constraints are dealt with by considering them as perfect measurements, see *Anderson and Moore* (*1979*) and *Van Der Zijpp and Hamerslag* (*1994*).

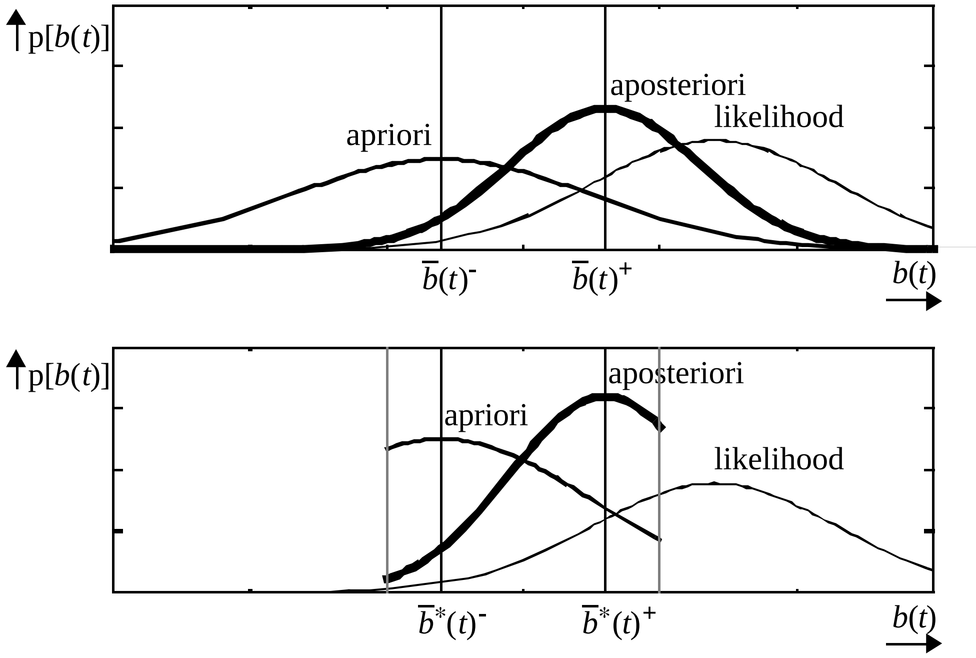

Figure 2: Bayesian update: aposteriori=likelihood.apriori/normalising constant. Top graph: prior distribution is MVN. Bottom graph: prior distribution is TMVN.

<u>*Kalman filter*</u>

The Bayesian updating approach described above results in a recipe that closely resembles a statistical approach known as the Kalman filter. The Kalman filter consists of a recursion for $\bar{b}(t)$ identical to (20) and (21), and was applied by *Cremer and Keller (1987)* and *Nihan and Davis (1987)*. The differences between the Bayesian and the Kalman methods stem from the interpretation of $\bar{b}(t)$. In the Kalman approach $\bar{b}(t)$ represents the minimum variance unbiased estimate for $b(t)$ while in the Bayesian approach $\bar{b}(t)$ is interpreted as a parameter characterising a TMVN distribution. This leads to two practical differences:

1. The Bayesian approach admits solutions for $\bar{b}(t)$ violating (5), while the Kalman filter does not. The following statement is therefore included after the measurement update in the Kalman filter used in the experiments (see section 7):

$$\bar{b}(t)=\max[\min[\bar{b}(t),\mathbf{1}],\mathbf{0}] \tag{22}$$

2. In the Bayesian approach a special postprocessing algorithm is needed to derive a point estimate from $\bar{b}(t)$, while in the Kalman filter $\bar{b}(t)$ itself is the point estimate.

5. OD-estimation from license plate surveys

<u>*Historical overview*</u>

It has long been recognized that license plates can be a valuable source of (static) OD-information (*Kryger and Ottesen, 1956, Brenner et al., 1957*). Until recently, the recording of license plate registrations had to be done manually. To economize on manpower as well as to minimize errors of recording it has become usual to only record a part of the license plate, for example a combination of the last three digits or letters. Such a survey is known as a *partial license plate survey.*

A problem that occurs if only a part of the license plate is recorded is that of the *spurious matches* which occur if two different vehicles have identical partial registrations. The extent to which spurious matches influence the OD-estimates was studied by *Makowski and Sinha (1976)* and *Hauer (1979),* who also proposed approximate statistical procedures to estimate OD-matrices from partial license plate surveys in the presence of spurious matches. These procedures were later improved by *Maher (1985).*

Another area of work concerned the *elimination* of spurious matches. These procedures were usually heuristic and require not only the partial registration, but also other data, such as time stamp, vehicle type etc. Spurious matches are then eliminated by imposing extra conditions to a 'match', for example by requiring corresponding vehicle types, arrival within a certain time window, etc. Methods proposed by *Shewey (1983)* and *Evans et al. (1993)* can be classified in this category. Also most commercially available license plate matching programs are of this category, e.g. *Buchanan and Partners (1986)*, *Lucas (1986)*, and *MVA Systematica (1987)*.

Statistical procedures based on partial registration data *and* time stamp information were proposed in *Watling and Maher (1988, 1992) and Watling (1990, 1994).*

Some effort has also been spent on the recovery from observation errors. Some types of errors are frequently made by human observers, for example mixing up the characters '3' and '8' or transposing characters, e.g. writing down 'xx-34-np' instead of 'xx-43-np'. Heuristic procedures to compensate or correct for this type of recording errors are common practice, see e.g. the NOPCOP (*Lucas, 1986*) and MicroMatch computer programs (*Buchanan and Partners, 1986*), and proposed improvements by *Evans et al. (1993).*

Reviewing the literature on this subject, it catches the eye that there is an emphasis on the statistical treatment of the phenomenon of spurious matches, while very little attention is paid to the statistical treatment of recording errors (a noted exception is *Geva et al., 1982*).

<u>Analysis</u>

The problem that is considered in the present paper deviates from the traditional one in a number of ways:

- The license plate surveys will be used in the context of *dynamic* OD-estimation rather than static OD-estimation;
- The possibility of spurious matches may safely be ignored. As *complete* rather than partial registrations will be read, the probability that a misread registration is matched with another registration is negligible;
- The issue of *recording errors* can not be ignored as reading full registrations with automated equipment implies that a significant fraction of the registrations is misread (researchers report recognition rates during field tests in the range 65%-90% (*Kanayama et al., 1991*));
- Recording stations need not correspond with network entries or exits;
- Any number of recording stations may be present on an entry-exit path.

Figure 1 shows a simple network configuration that expresses all these properties. In the analysis it is assumed that each image processor produces a list of, partly erroneous, registrations. Only the *number* of matches is used in the estimation process. These numbers are summarized in a vector $e(t)$ consisting of *trajectory counts* $e_{ab}(t)$, $a<b$, which denote the number of vehicles that were recognized at site a during period t and at site b, but not at any site upstream of a or downstream of b. An extra category $e_{00}(t)$ accounts for those vehicles that are detected correctly once or not at all.

Whether or not a vehicle travelling on EE-pair i-j contributes to $e_{ab}(t)$ depends on a number of nested selections. In order to contribute, it is required that no recognition(NR) occurs upstream of a or downstream of b, and recognition(R) occurs at both a and b, so the probability that a trip in flow $f_{ij}(t)$ contributes to $e_{ab}(t)$ equals:

$$P[(\text{NR upstream of } a) \wedge (\text{R at } a) \wedge (\text{R at } b) \wedge (\text{NR downstream of } b)] \tag{23}$$

This probability is given by:

$$p_{ab}^{ij} = \left(\prod_{c=1}^{a-1} (1-\alpha_c)^{\kappa_{ijc}}\right) \alpha_a^{\kappa_{ijr}} \alpha_b^{\kappa_{ijs}} \left(\prod_{d=b+1}^{h} (1-\alpha_d)^{\kappa_{ijd}}\right) \tag{24}$$

for $1 \le a < b \le h$, and complemented with p_{00}^{ij}. As an aid in deriving the joint probability distribution of $\{e_{ab}(t), 1 \le r < s \le h\}$, *trajectory count contributions*, denoted by $g_{ab}^{ij}(t)$, are introduced which denote the contribution of flow $f_{ij}(t)$ to trajectory count $e_{ab}(t)$. The probability that an arbitrary vehicle entering at i will contribute to trajectory count $g_{ab}^{ij}(t)$ is equal to $b_{ij}(t).p_{ab}^{ij}$. Hence, the joint distribution of the trajectory count contributions is given by the following multinomial distribution:

$$P[g(t)|\tilde{q}(t), b(t)] = \prod_{i=1}^{m} \tilde{q}_i! \prod_{j=1}^{n} \prod_{r<s \vee ab=0} \frac{(p_{ab}^{ij} b_{ij}(t))^{g_{ab}^{ij}(t)}}{g_{ab}^{ij}(t)!} \tag{25}$$

Now let $\theta(a,b)$ denote the position of the element $e_{ab}(t)$ in the vector of trajectory counts $e(t)$, then the expectation of $e(t)$ is given by:

$$E[e(t)|q(t)] = G'(t)b(t) \tag{26}$$

where $G(t)$ is defined by:

$$G_{\varphi(i,j),\theta(a,b)}(t) = q_i(t).p_{ab}{}^{ij}$$
$$i=1,2\ldots m,\ \ i=1,2\ldots n,\ \ a=1,2\ldots h,\ \ b=a+1,r+2\ldots h \tag{27}$$

Analogous to (13), equation (26) gives rise to the specification of a measurement equation:

$$e(t) = G'(t)b(t) + z(t) \tag{28}$$

where $z(t)$ is a zero mean random variable that accounts for random effects and mis-specification of $G(t)$ due to observation errors contained in $q(t)$. Parallel to (13-16), one can derive the covariance matrix of $z(t)$ from the multinomial distribution of the trajectory count contributions, or one can choose a more simple approach, replacing this matrix with a diagonal matrix derived from the average observed values.

Instead of analysing this issue to the last detail at this stage we will first consider the more general case of estimating split probabilities from combined link-volume counts - trajectory counts.

6. OD-estimation from multiple data sources

Analysis

The combined usage of volume counts and trajectory information is a yet unexplored possibility to collect dynamic OD-information. It is expected that the two sources of data effectively complement each other, provided that an appropriate statistical estimation procedure is used.

Equations (13) and (28) can be combined in the following measurement equation:

$$\begin{bmatrix} y(t) \\ e(t) \end{bmatrix} = \begin{bmatrix} H'(t) \\ G'(t) \end{bmatrix} b(t) + \begin{bmatrix} v(t) \\ z(t) \end{bmatrix} \tag{29}$$

Dependencies between elements of the measurement error vector $[v'(t)\, z'(t)]'$ follow from the fact that both $y(t)$ and $g(t)$ can be written as linear combinations of trajectory count contributions which are multinomially distributed according to (25). The estimates of the split probabilities may benefit from knowledge of these dependencies, summarized in a covariance matrix $R^*(t)$.

If the symbol $o(t)$ is used to denote the combined observation vector $[y'(t)\ e'(t)]'$, then parallel to (2) and (10) it holds that:

$$o(t)=U^{*\prime}\, g(t)+\begin{bmatrix} s\,(t) \\ \mathbf{0} \end{bmatrix} \tag{30}$$

for some matrix U^* of which each element is either one or zero.

The traditional problem of estimating EE-matrices from link volumes is characterized by the dependencies $b \rightarrow f \rightarrow y$, where the split probabilities b need to be estimated, the conditional distribution of the EE-flows f given b is multinomial, and the vector of traffic counts y is a linear combination of f.

The problem of estimating EE-matrices from combined data can be characterized in a similar way. In fact for each network with combined observations one can define an imaginary network with link volume observations only, in which the observations behave equivalently. For this purpose divide each flow $f_{ij}(t)$ into subflows $g^{ij}_{ab}(t)$ that each travel over their own imaginary path with a probability $b_{ij}(t)\,p^{ij}_{ab}$ and suppose that the observations $o(t)$ satisfy (30).

Now $o(t)$ has the statistical properties that belong to a combined observation but at the same time is a linear combination of (notional) subflows. The statistical properties can hence be captured in equation (15).

Using this recipe results in the following point-estimate based approximation for R_t^*:

$$\begin{gathered} R_t^* =(U^{*\prime}\mathrm{cov}[g(t),g(t)|b(t),q(t)]U^* + \Theta^*)\delta_{tp} \\ \mathrm{cov}[g(t),g(t)|b(t),q(t)]=\boldsymbol{Q_t}^*(\boldsymbol{B_t}^*\text{-}\boldsymbol{B_t}^*\mathrm{p}^*\mathrm{p}^{*\prime}\boldsymbol{B_t}^{*\prime})+\boldsymbol{B_t}^*\mathrm{p}^*\Phi\mathrm{p}^{*\prime}\boldsymbol{B_t}^* \end{gathered} \tag{31}$$

where $\boldsymbol{Q_t}^*$, $\boldsymbol{B_t}^*$, p^* and Θ^* are defined analogous to $\boldsymbol{Q_t}$, $\boldsymbol{B_t}$, p and Θ in equation (15).

Again, as an alternative to the above method, a simplified approach will also be tested (see section 7):

$$R_t^*=\mathrm{diag}(\bar{o}) \tag{32}$$

where $\bar{o}$ represents the average observed value of $o(t)$.

Estimation procedure

The estimation of the split probabilities from combined data poses no new problems. The DCLS, Kalman and Bayesian updating methods that were proposed in section 4 to estimate split probabilities from link volume observations can be used unaltered when equations (13), (15) and (16) are replaced with (29), (31) and (32) respectively.

7. Experimental results

The objective of the experiments is to evaluate the effect of the input data and estimation method on the error of estimation. In particular the usage of link volume data is compared with the usage of combined data, and the traditional estimation methods such as the DCLS method and the Kalman filter are compared with the Bayesian updating method described in section 4. To versions of the latter method are tested: the first version uses a covariance matrix derived from the average observations while the second version uses the point-estimate based matrix described in equations (15) and (31).

Test-data

Test-data are generated according to a number of specifications that can be given in advance. The parameters that define these specifications are given in table 1A, and their meaning is explained in the text below. Travel-time dispersion has not explicitly been taken into account when generating the test-data Instead an error term was added to all traffic counts, allowing for assignment errors (see section 2). The test-data comprise EE-flows as well as synthesized traffic counts and trajectory counts. The topology of the network and the locations of induction loops and license plate readers are shown in figure 1. Possibly, more efficient configurations involving the same number of license plate readers exist for this network. The primary objective has however been to illustrate end test the theory presented earlier.

The generation of the test-data involves the following steps:

1. *Generation of the split probabilities*. The split probabilities are initialized with a randomly generated vector $b(0)$. Subsequently the vectors $b(0)\ldots b(T)$ are generated using the random walk model (3) while 'mirroring' the elements of b that violated the inequality constraints (5). The following specification for S_t is used:

$$S_t = \mathrm{b}I \tag{33}$$

2. *Generation of the entry flow volumes*. The entry volumes are sampled from a normal distribution of which mean and variance equal a parameter $\bar{\boldsymbol{q}}$.
3. *Generation and assignment of the EE-flows*. The EE-flows are sampled from the multinomial distribution (9), using the probabilities generated under step -1-. The idealized link volumes follow from applying (10).
4. *Generation of the entry- and link volume observation errors*. These errors are generated in accordance with equations (1) and (2), using the following specifications for Φ and Θ:

$$\Phi = \mathrm{s}_{\boldsymbol{q}}^{2} I \tag{34}$$

$$\Theta = \mathrm{s}_{y}^{2} I \tag{35}$$

5. *Generation of the trajectory counts*. For each individual vehicle its contribution to the vector of trajectory counts is determined by simulating its detection status at the

license plate readers on its path. The recognition rates that are used in this process are given by:

$$\alpha_r = a \qquad r = 1,2 \ldots h \tag{36}$$

In order to reduce random effects, for every set of parameters described in table 1A, ten independent datasets are generated. Each estimation method will be applied to all of these sets, after which the errors of estimation will be averaged. Each set consist of 48 periods and was generated bearing in mind a period length of ten minutes. Network 1 represents the default configuration. Each of the specifications 2-6 differs only in one parameter from this configuration.

Estimation methods

The methods that are compared are divided in the following categories (see also table 1B):

1. The DCLS method, see equations (17) and (18). After some experimenting, the discounting parameter λ was set to 1- b.
2. The Bayesian Updating method (BU), see equations (20) and (21). In accordance with the way the test-data are generated the matrix S_t is set to bI. The observation error covariance matrix R_t is either derived from the average observed value (methods 3 and 4) or from theoretical considerations (methods 5 and 6).
3. The Kalman filter method (methods 7 and 8). This method is identical to the Bayesian updating method, except for remarks -1- and -2- at the end of section 4.

Each method is first applied to a dataset consisting of entry volumes and link volumes only, and subsequently to a dataset consisting of entry volumes, link volumes and trajectory counts.

Evaluation criterion

The following measure is used as an evaluation criterion:

$$\text{RMSE}(t) = \sqrt{\frac{1}{N}\sum_{i,j}\left(q_i(t)\,\bar{b}_{ij}(t) - f_{ij}(t)\right)^2} \tag{37}$$

where N represent the number of connected EE-pairs. For every period and network specification in table 1A the measure is averaged over the number of datasets that are generated. The averages over the last 40 periods of the measure are summarized in table 1C.

Results

The test results are summarized in table 1C. One should relate the outcomes in table 1C to an average entry volume of 100 or 200 (see table 1A). When comparing the results, the methods should be divided in two groups: those using link volume counts only (method 1, 3, 5 and 7), and those using combined data (method 2, 4, 6 and 8). In both groups it is clear that the best performance is obtained from the Bayesian updating methods (3-4-5-6) followed by the Kalman methods (7-8) and the DCLS method (1-2). The difference between the use of a point-estimate based covariance matrix (5-6) and the use of a matrix derived from the average observations (3-4) is marginal but observable. The performance of both groups of methods for network 1 can be viewed in more detail in figure 3.

The most striking result is the superior performance of the BU methods relative to the DCLS and Kalman methods. This is illustrated by the fact that the BU methods produce better

Table 1A: Specification of test-data

	Network specification					
	1	**2**	**3**	**4**	**5**	**6**
b	0.0001	*0.01*	*0*	0.0001	0.0001	0.0001
$\bar{\boldsymbol{q}}$	100	100	100	*200*	100	100
s_q^2	20	20	20	20	*100*	20
s_y^2	20	20	20	20	20	*100*
a	0.5	0.5	0.5	0.5	0.5	0.5

Table 1B: Specification of estimation methods

Method	Type	Input data	Remarks
1	DCLS	link volumes	λ=1- b
2	DCLS	combined	λ=1- b
3	BU	link volumes	R_t=diag($\bar{y}$)
4	BU	combined	R_t*=diag($\bar{o}$)
5	BU	link volumes	R_t, see (15)
6	BU	combined	R_{t*}, see (31)
7	Kalman	link volumes	R_t=diag($\bar{y}$)
8	Kalman	combined	R_t*=diag($\bar{o}$)

Table 1C: Error of estimation (RMSE veh./period)

Method	Network specification					
	1	**2**	**3**	**4**	**5**	**6**
1	18.24	22.16	17.50	34.81	18.12	19.22
2	17.02	20.73	16.40	33.31	17.00	17.70
3	16.66	16.65	16.44	32.01	17.17	17.00
4	16.09	15.74	15.68	30.95	16.22	16.33
5	16.52	16.51	16.31	31.89	16.57	16.71
6	16.07	15.71	15.67	31.05	16.02	16.21
7	17.60	18.31	21.33	33.38	17.91	18.68
8	16.70	17.15	18.42	32.44	16.62	17.24

results using link volume counts only than the DCLS and Kalman methods using both link counts and trajectory information.

Another conclusion is that the difference between BU methods that attempt to take dependencies of observations into account (method 5 and 6) barely produce better results than BU methods that implicitly assume independence (method 3 and 4).

For network specification 1, the main difference between the performance of the methods is the number of periods needed to reach a certain level of accuracy (see figure 3). In this respect the advantages of the BU methods are evident. This also makes the Bayesian updating the preferred method under less favourable circumstances, where the network may involves a larger number of EE-pairs, the measurements contain larger errors, or if the rate of change in the split probabilities is higher.

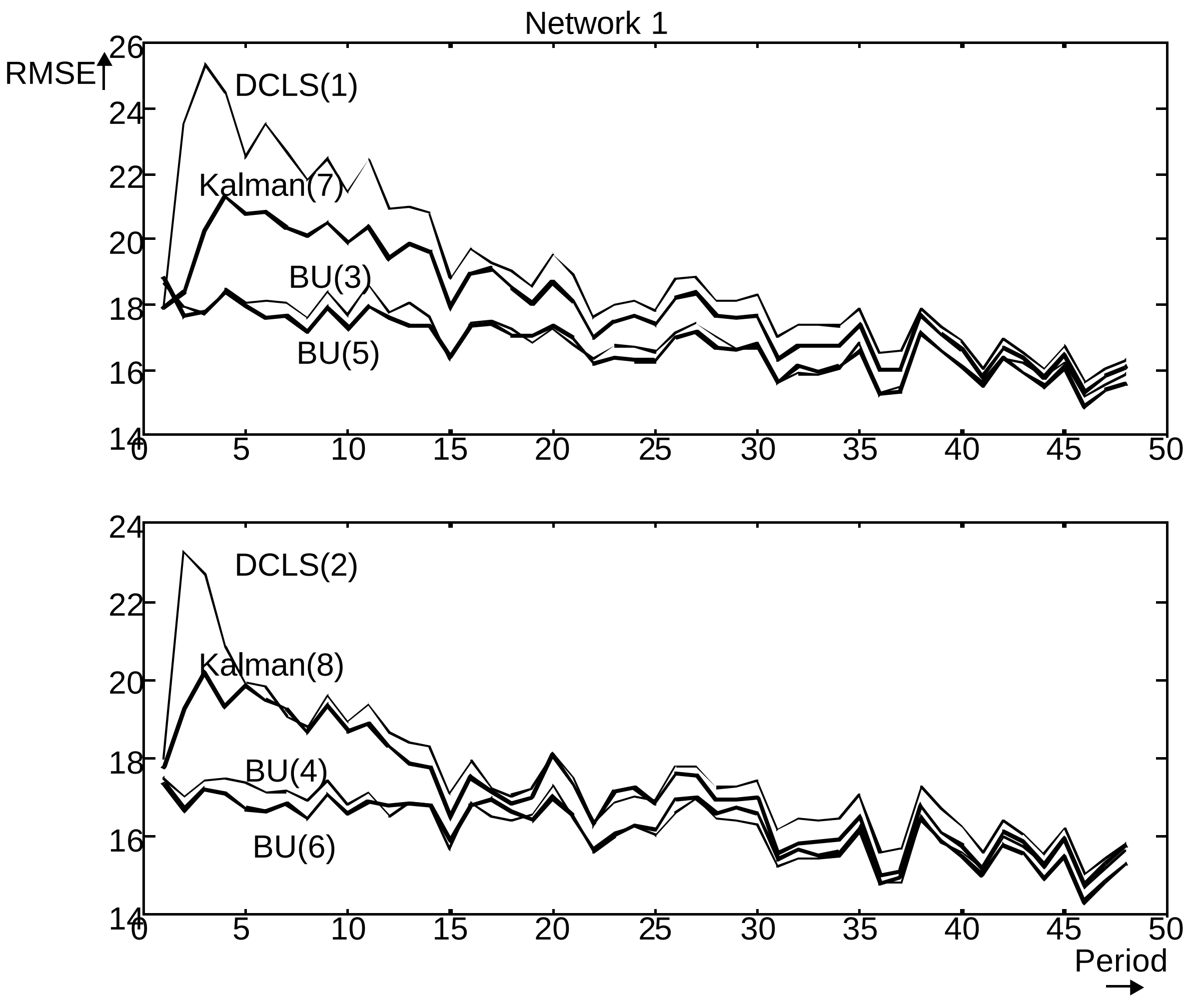

Figure 3: Average error of estimation compared for four types of methods applied to networks generated according to specification 1 (see table 1A), using link volume counts (top graph) and combined data (bottom graph).

8. Conclusions

AVI technologies which are currently being developed and implemented are likely to give rise to information that can be effectively used to supplement time series of traffic counts as a source of data for dynamic OD-estimation. An example of an AVI technology is license plate recognition based on image processing. This technology can be implemented at relatively low cost, especially if a limited number of license plate readers are placed at strategic locations.

Allowing data to be used from induction loops and license plate readers at arbitrary locations gives rise to mutual dependencies between observations. These dependencies can be described by considering both the traffic counts and the trajectory counts as sums of trajectory count contributions. A similar argument can be put forward if the usage of traffic counts is to be combined with other sources of trajectory information, even if the latter apply to only one category of vehicles (e.g. probe vehicles).

The estimation problem in which above analysis has resulted is characterized by the presence of nonnegativity constraints applying to split probabilities which are to be estimated from linear measurements. This problem can be solved with traditional estimation procedures such as DCLS or the Kalman filter, but on theoretical grounds a newly developed method referred to as the Bayesian updating method is preferred.

Experiments clearly show the benefits of this new method and also show that the error of estimation reduces if trajectory counts derived from license plate survey data are used in addition to traffic counts. The experimental results are less conclusive about the practical need for the analysis of the dependencies between observations, as the results show only a marginal improvement if these dependencies are taken into account.

9. References

Anderson, B.D.O. and Moore, J.B. (1979) *Optimal Filtering*, Prentice Hall

Bell, M.G.H. (1991a) The Estimation of Origin-Destination Matrices by Constrained Generalised Least Squares, *Transportation Research B*, Vol.25B, No.1, pp.13-22

Bell, M.G.H. (1991b) The Real Time Estimation of Origin-Destination Flows in the Presence of Platoon Dispersion, *Transportation Research-B*, Vol. 25-B, pp. 115-125

Brenner, R. , Mathewson, J.H. and Gerlough, D.L. (1957) A General Method for Estimating Through Traffic in a Road Network, *Highway Research Abstracts*, 27 (8)

Buchanan, C. and Partners (1986) Micromatch Number-plate Matching Suite, *User Guide*

Cascetta, E., Inaudi, D. and Marquis, G. (1993) Dynamic Estimators of Origin-Destination Matrices Using Traffic Counts, *Transportation Science*, Vol.27, No.4, pp.363-373

Chang, G. and Wu, J. (1994) Recursive Estimation of Time-Varying Origin-Destination Flows from Traffic Counts in Freeway Corridors, *Transportation Research-B*, Vol. 28B, pp. 141-160

Cremer, M. and Keller, H. (1981) Dynamic Identification of Flows from Traffic Counts at Complex Intersections, In: *Proc. 8th Int. Symposium on Transportation and Traffic Theory*, University of Toronto Press, Toronto Canada

Cremer, M. and Keller, H. (1987) A New Class of Dynamic Methods for the Identification of Origin-Destination Flows, *Transportation Research-B*, Vol.21B, No.2, pp.117-132

Geva, I., Hauer, E. and Landau, U. (1982) *Maximum Likelihood and Bayesian Methods for the Estimation of O-D Flows*, University of Toronto, Publication 82-11, ISBN 0-7727-7038-7

Evans, R., Martin, P.T. and Bell, M.C. (1993) A Method for Analysing Partial Registration-plate Data, *Traffic Engineering and Control*, Vol.34, pp.76-79

Hauer, E. (1979) Correction of License Plate Surveys for Spurious Matches, *Transportation Research A*, Vol.13A, pp.71-78

Kalman, R.E. (1960) A New Approach to Linear Filtering and Prediction, *Journal of Basic Engineering (ASME)*, 82D, pp.35-45

Kryger, P. and Ottesen K.A. (1956) Van the License Plate Method be Used for Traffic Studies?, *Traffic Quarterly*,

Vol.10, pp.377-386
Lucas, C.F. (1986) The NOPCOP Range of Number Plate Comparison Programs, *User Guide*, Transport & Road Research Laboratory
Maher, M.J. (1985) The Analysis of Partial Registration-plate Data, *Traffic Engineering and Control*, Vol.26, pp.495-497
Makowski, G.G. and Sinha, K.C. (1976) A Statistical Procedure to Ananlyze Partial License Plate Numbers, *Transportation Research*, Vol.10, pp.131-132
Mood, A.M., Graybill, F.A. and Boes, D.C. (1963) *Introduction to the Theory of Statistics*, 3rd edition, McGraw W-Hill International Book Company
MVA Systematica (1987) MVMACH license plate matching program, *User Guide*
Nihan, N.L. and Davis, G.A. (1987) Recursive Estimation of Origin-Destination Matrices from Input/Output Counts, *Transportation Research B*, Vol. 21B, No.2, pp. 149-163
Nihan, N.L. and Davis, G.A. (1989) Application of Prediction-Error Minimization and Maximum Likelihood to Estimate Intersection O-D Matrices from Traffic Counts, *Transportation Science*, Vol 23, No. 2
Shewey, P.J.H. (1983) An Improved Algorithm for Matching Partial Registration Numbers, *Transportation Research B*, Vol.17B, No.5, pp.391-397
Van Der Zijpp, N.J. and Hamerslag, R. (1994) An Improved Kalman Filtering Approach to Estimate Origin-Destination Matrices for Freeway Corridors, *Transportation Research Records*, No.1443, pp.54-64
Van Der Zijpp, N.J. (1996), *A Bayesian Framework for Dynamic Origin-Destination Estimation on Motorway Networks*, Ph.D. Thesis, Civil Engineering Department, Delft University of Technology
Watling, D.P. and Maher, M.H. (1988) A Graphical Procedure for Analysing Partial Registration-plate Data, *Traffic Engineering and Control*, Vol.29, pp.515-519
Watling, D.P. (1990) *The Statistical Analysis of Partial Registration Plate Data*, Ph.D. Thesis, Dep. of Probability and Statistics, University of Sheffield
Watling, D.P. and Maher, M.J. (1992) A Statistical Procedure for Estimating A Mean Origin-Destination Matrix from a Partial Registration Plate Survey, *Transportation Research B*, Vol.26B, No.3, pp.171-193
Watling, D.P. (1994) Maximum Likelihood Estimation of an Origin-Destination Matrix from a Partial Registration Plate Survey, *Transportation Research B*, Vol. 28B, No. 4, pp. 289-314