Cost-effectiveness of cordon studies for trip matrix estimation from traffic counts

Antonio J. Torres Martínez

Abstract: Cordon studies can substantially improve estimation of origin-destination (O/D) trip matrices from link traffic counts. This paper proposes a method for integrating cordon survey results in O/D matrix estimation. This method is applied to 5 inter-urban networks in the Valencia Region, focusing on the specific formulation of the O/D matrix estimation technique, the assignment model to be implemented, the selection of traffic count locations in the cordoned area and, especially, the selection of cordon survey stations and data collection. Cost-effectiveness of cordon survey selection is addressed. The results show that cost-effectiveness ratios tend to decrease with the number of surveyed cordons in the 5 networks analyzed. This is due to non-linear reduction of origin-destination estimation errors and economies of scale on conducting cordon surveys. The results of this study can be useful to assess decision-making when conducting cordon surveys in interurban networks. Once such decisions have been taken, the method presented in this paper can be applied to the whole O/D matrix estimation procedure.

CE Database subject headings: Transportation models, Transportation networks, Traffic surveys, Traffic analysis, Data collection, Cost analysis

Introduction

The estimation of the origin-destination (O/D) matrix is essential for a wide range of transportation planning activities. Conventional survey techniques for estimating O/D matrices are often costly and time consuming, due to the large volumes of data to be collected and analyzed. O/D matrix estimation from traffic counts is comparatively faster and more economical. These advantages explain the considerable development of mathematical approaches to estimate O/D matrices from traffic counts since the late 70s.

Approaches to O/D matrix estimation from traffic counts can be classified, based on algorithmic typology, to: (a) traffic modelling based techniques; (b) statistical inference approaches; (c) gradient based solution techniques; and (d) genetic algorithm based methods. The first methods to be implemented came from traffic modelling and assumed proportional assignment, which means independence between the link traffic volumes and the proportions of the O/D trips using these links. Van Zuylen and Willumsen (1980) proposed two important models of this type: the maximum entropy model and the information minimizing model. The statistical inference approaches include generalized least squares (GLS) methods developed by Cascetta and Nguyen (1988), and Bell (1991). Both methods assumed proportional assignment. The results of the maximum entropy model are approximated by the GLS approach when the link flows in the network are known to a high degree of accuracy (Bell 1984), which may be the case for many inter-urban networks. The GLS approach developed by Bell (1991) allows the combination of survey and traffic count data, while allowing for the relative accuracy of the two data sources. Equilibrium assignment takes into account link capacity constraints and cost-flow relationships to model congestion effects. Both approaches (a) and (b) have been extended by integrating equilibrium assignment of the O/D matrix. Fisk (1988, 1989) extended the maximum entropy model to the congested case. Yang et al. (1992, 1994) showed how the GLS methods can be integrated with an equilibrium traffic assignment in congested networks in the form of a convex bi-level optimization problem. The O/D matrix estimation problem appears on the upper level and the user-equilibrium assignment problem on the lower level. This model was extended (Yang, 1995) to accommodate the case of flow interactions in determining link delay functions. Lo et al. (1996) proposed a unified statistical approach to O/D matrix estimation, treating the link choice proportions as random variables and allowing the combination of traffic count data with information related to O/D survey results and link choice proportions. In a follow-up paper, Lo et al. (1999) proposed a decomposition algorithm to efficiently solve this problem, compared to conventional numerical algorithms such as the quasi-Newton algorithm.

Gradient based solution techniques (c), which may be included in the statistical inference models, have been proposed to solve the bi-level O/D matrix estimation problem (Spiess 1990; Drissi-Kaitouni and Lundgren 1992). An advantage of this approach is its computational tractability.

Genetic algorithms (d) have recently been applied to bi-level problems with multiple vehicle O/D matrix estimation (Baek et al. 2004). The results perform well when compared with other algorithms, with respect to estimation errors.

Economic and operational advantages of O/D matrix estimation from traffic counts take place in a context of greatly increased mobility in many countries, which leads to new and frequent congestion problems. Therefore, extended application of these methods is much needed. Efficient O/D matrix estimation is the previous step to obtaining realistic traffic demand forecasts and, thus, to assessing the road network alternatives to solve or palliate such congestion problems. In order to reinforce this application, they should not be excessively complex, and should add application simplicity and comprehensibility to economy.

Therefore, it is crucial in road transport planning to have a sufficient number of error-free traffic counts, as well as other reliable information about trip distribution in the study area. Interview surveys fall within this information. They are limited due to cost and time study conditions. But O/D surveys can be conducted in cordon studies within the assumed cost and time limitations. It is evident that direct O/D trip information obtained from sampling in cordon studies can improve the O/D matrix estimation results. Inbound and outbound vehicular flows by O/D cordon stations can be estimated from real data. It is often argued that data collected by direct measurements/interviews or surveys is time consuming and costly, also involving data editing and analysis costs. Such a statement must be carefully considered case by case. For example, there are great differences between home-based interviews and cordon studies in preparation, organization, sample sizes, number and location of interviews, etc., and, therefore, estimation errors, time consumption and economic costs. Probably these economic reasons explain why integration of information obtained from cordon or screenline surveys has received little attention in the journals specialized in traffic and transport planning. A study by Landau et al. (1982) provided the solution procedure for estimating cross-cordon O/D flows from cordon studies. However, like other studies related to O/D matrix estimation from cordon surveys (Echenique and Williams 1982; Echenique 1977), the solution procedure was not integrated in a method of O/D matrix estimation from traffic counts, and the cost-effectiveness of cordon studies was not dealt with.

The main questions to be answered are how much cordon studies cost and, accordingly, how much they improve the O/D matrix estimation from traffic counts. These questions can be dealt with in terms of cost-effectiveness. The relative cost of conducting cordon studies can be compared to the relative effects of such studies, hence, to the error reduction provided by them. This is the approach of this article. Firstly, a four-step procedure is proposed. Then, essential aspects such as the formulation of the O/D matrix estimation method, the assignment model, the selection of cordon survey stations and the traffic count locations are discussed. Finally, the solution procedure is applied to 5 inter-urban networks in the Valencia Region (Spain). Cost-effectiveness ratios are obtained and discussed. Results can be useful to assess decision making for conducting cordon studies in other inter-urban networks with different interview costs. Once such decision has been made, the method proposed in this article can be applied to the whole procedure to be followed.

Solution procedure

Consider a study area divided into a number of trip generator and attractor zones *N*, represented by centroids. The road network is coded into *M* nodes and *L* links connecting pairs of nodes. A cordon line surrounds a cordoned zone inside the study area (shown in Fig. 1), which includes *Z* internal centroids. Along its boundary, there are *c* stations where vehicles entering and leaving the cordoned area are counted. In an inter-urban traffic study, the *Z* internal zones represent towns or parts of cities or metropolitan areas, and the *(N - Z)* external zones are the accesses to the study area and the towns external to the cordoned area. In an urban traffic study, the *Z* internal zones may represent not only urban zones, but also car parks, markets, educational centers, etc. In this case, the external zones are the roads entering the urban study area and the urban zones located outside the cordoned area.

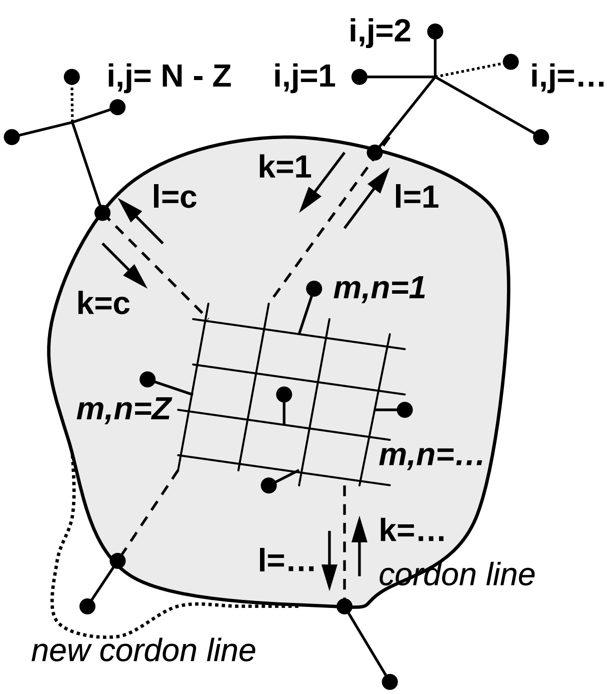

FIGURE 1. Cordon line and study area.

Let T_{kl} be the number of trips from station *k* to station *l (k, l = 1, 2, …, c).* The index "*0*" will signify an inbound trip ending inside the cordon area, or an outbound trip starting inside the cordon area. The number of inbound trips entering the cordon area during a specified period of time at station *k* is

$$O_k = \sum_l T_{kl} \qquad (k = 1, 2, \ldots, c) \tag{1}$$

Accordingly, the number of outbound trips leaving the cordon area during the same period of time at station l is

$$D_l = \sum_k T_{kl} \qquad (l = 1, 2, \ldots, c) \tag{2}$$

A random sample is selected from the inbound and outbound vehicles to allow estimation of O/D flows at each station. For the vehicles entering and leaving the cordon area, the exit and the entry stations are determined. Additional information can be obtained from the drivers of the selected vehicles, such as vehicle occupancy, frequency and purpose of travel, etc. Automatic vehicle identification can be employed along the cordon line, but only cordon surveys can provide information about the number of trips ending/starting in the specific internal zones of the cordoned area.

Let t_{kl} be the random sample of vehicles obtained at station *k* from those entering the cordon area that report leaving the area via station *1*; and t'_{kl} the corresponding random sample of vehicles obtained at station *1* from those leaving the cordon area. Obtaining t_{kl} and t'_{kl} will be discussed later in the real application to 5 inter-urban networks in the Valencia Region.

Having established the notation, the procedure adopted is described as follows:

<u>Step I</u>

The first step is to find estimates of T_{kl} *(k, l = 1, 2, …, c)* by using the cordon-count data O_k and D_l; and the sample information t_{kl} and t'_{kl}. These estimates must maximize the probability of observing the values of O_k, D_l, t_{kl} and t'_{kl}.

Let us suppose that the probability of observing a random sample is approximately multinomial, and assume that the sampling is with replacement. This assumption seems reasonable as long as the sample is a small fraction of the daily traffic. The probability of observing all random samples obtained at each of the *c* survey stations for both inbound and outbound traffic can be expressed as follows (Landau et al. 1982):

$$Max\prod_{k=1}^{c}\left\{\left[\frac{\left(\sum_{l}t_{kl}\right)!}{\prod_{l}t_{kl}!}\right]\prod_{l}\left(\frac{T_{kl}}{O_k}\right)^{t_{kl}}\right\}\prod_{l=1}^{c}\left\{\left[\frac{\left(\sum_{k}t'_{kl}\right)!}{\prod_{k}t'_{kl}!}\right]\prod_{k}\left(\frac{T_{kl}}{D_l}\right)^{t'_{kl}}\right\} \tag{3}$$

subject to the traffic count constraints (1) and (2).

The solution of this program $\{T^*_{kl}, T^*_{k0}, T^*_{0l}\}$ is obtained, using the method of Lagrange multipliers, as:

$$T^*_{kl} = (t_{kl} + t'_{kl})/(\alpha_k + \beta_l) \qquad T^*_{k0} = t_{k0}/\alpha_k \qquad T^*_{0l} = t'_{0l}/\beta_l \tag{4}$$

where the unknown Lagrange multipliers α_k and β_l are determined by a simple iterative algorithm that uses Eq. (1) and (2). This algorithm is described in the above cited paper by Landau et al. (1982). Automatic vehicle identification along the cordon line is sufficient to obtain an accurate estimate of T_{kl}, but cannot provide the information required in the following Steps II and III.

Step II

Now the initial problem has been solved, it is simple to obtain estimates of the flows T_{ijkl} from zone *i* to zone *j* external to the cordon, passing through stations *k* and *1 (i, j = 1, 2, …., N-Z)*, if the drivers sampled from the flow provide information not only about stations of entry or exit, but also about trip origin and destination.

Generalizing the previous notation for the random samples, let t_{ijkl} be the number of vehicles obtained at station *k* from those entering the cordon area that report zone *i* as origin and zone *j* as destination, and leave the cordon area by station *1*, and t'_{ijkl} be the corresponding number of vehicles obtained at station *l* from those leaving the cordon area. T^*_{ijkl} will be the best estimate of the flows from zone *i* to zone *j* via cordon stations *k* and *1*. This estimate is obtained by means of a proportional flow distribution:

$$T^*_{ijkl} = \left[(t_{ijkl} + t'_{ijkl})/(\sum_{i,j} t_{ijkl} + \sum_{i,j} t'_{ijkl})\right] T^*_{kl} \tag{5}$$

and the estimated flow between external zones *i* and *j* is:

$$T^*_{ij} = \sum_{k,l} T^*_{ijkl} \tag{6}$$

Step III

In this step, the focus is on finding estimates of the inbound trips ending in the cordon area zone *n* passing through station *k* Q^*_{ikn}, and the outbound trips starting in the internal zone *m* passing through station *1* R^*_{mlj}. The estimates can be indirectly obtained from the estimated inbound trips ending inside the cordon area T^*_{k0} and the estimated outbound trips starting inside this area T^*_{0l} as follows:

$$Q^*_{ikn} = T^*_{ik0}\, t_{ikn} / t_{ik0} \qquad (i = 1, 2, ..., .N\text{-}Z; \quad k = 1,2,\ldots c; \quad n = 1, 2, \ldots Z) \tag{7a}$$

$$R^*_{mlj} = T^*_{0lj}\, t'_{mlj} / t'_{0lj} \qquad (j = 1, 2, ..., .N\text{-}Z; \quad l = 1,2,\ldots c; \quad m = 1, 2, \ldots Z) \tag{7b}$$

where:

$$T^*_{ik0} = T^*_{k0}\, t_{ik0} / t_{k0} \qquad T^*_{0lj} = T^*_{0l}\, t'_{0lj} / t'_{0l} \tag{8}$$

and

$$t_{k0} = \sum_i t_{ik0} \qquad t'_{0l} = \sum_j t'_{0lj} \tag{9}$$

t_{ik0} is the random sample of inbound trips from origin *i* obtained at station *k* ending inside the cordon area, and t'_{0lj} is the random sample of outbound trips starting inside the cordon area that report leaving the area via station *l* to destination *j*. Similarly, t_{ikn} is the random sample of inbound trips from origin *i* obtained at station *k* ending in zone *n*, and t'_{mlj} is the random sample of outbound trips starting in zone *m* and leaving the cordon area via station *l* to destination *j*. The Eqs. (5) and (7a, 7b) take into account that trips from origin *i* or to destination *j* could enter or leave the cordon area through different cordon stations.

It has been assumed in Steps II and III that O/D samples are available at all cordon stations in both inbound and outbound directions. This is not always the case, and cost-effectiveness of reducing sample number must be assessed. When only inbound or outbound samples are missing the solution procedure described above remains unchanged. Otherwise, simplification for inter-urban traffic models is usually allowed at this stage:

- Analysis usually focuses on average annual daily traffic. It is often a justifiable assumption that trip origins and destinations are similar and, therefore, T^*_{ij} is symmetric.

- Sometimes this symmetry can be extended to T^*_{ijkl}.

- Depending on the network configuration, direct knowledge of traffic patterns can provide reasonable evidence to fix some elements of T^*_{ijkl} equal to zero. This is the case, for example, when a least cost path exists between zones i and j, fully external to the cordon area. This simplification is especially useful when there is a gap in the information at cordon stations *k* and *l*. Therefore, it can be assumed in Step II that $T^*_{ijkl} = 0$ for some O/D pairs, and also sometimes $T^*_{ij} = 0$, without significant errors. Underspecification of the problem can, thus, be reduced.

Application of the 3 aforementioned simplifications could lead to estimate (6), (7a) and (7b) directly. If this is not the case, and there is still an information gap, there are two possible strategies to follow:

- Apply the solution procedure only to cordon stations for which O/D sample information is available and, then, distribute the unallocated flow assuming equally likely outcomes.

- Modify the cordon line so that external zones without cordon station O/D samples are included in the new cordon area (see Fig. 1 “new cordon line”), and consequently become internal.

It is assumed that vehicles are counted at each cordon survey station.

Step IV

Internal flows of the cordon area still have to be estimated. The set of traffic counts conducted inside the cordon area will be used to do this. Let U_{mn} be the number of trips between internal zones *m* and *n*. Taking into account that vehicles entering and leaving the cordon area are also counted at internal stations, the fundamental equation to estimate O/D matrices from traffic counts may be written as:

$$V^s = \sum_{m,n} p^s_{mn} U_{mn} + \sum_{k,l} \overline{p}^s_{kl} T^*_{kl} + \sum_{k,n} \overline{p}^s_{kn} Q^*_{kn} + \sum_{m,l} \overline{p}^s_{ml} R^*_{ml} \tag{10}$$

(k, l = 1, …, c; m, n = 1, …, Z; s = 1,…, S)

where:

$$Q^*_{kn} = \sum_i Q^*_{ikn} \qquad\qquad R^*_{ml} = \sum_j R^*_{mlj}$$

V^s is the volume of traffic flow using link *s*, p^s_{mn} is the ratio of total trips from origin *m* to destination *n* using link *s*, and $\{\overline{p}^s_{kl}, \overline{p}^s_{kn}, \overline{p}^s_{ml}\}$ are the ratios of total trips between the corresponding zones, crossing the cordon line and using link *s*. *S* is the total number of traffic counts in the cordon area.

To determine the traffic proportions $\left\{p_{mn}^{s}, \overline{p}_{kl}^{s}, \overline{p}_{kn}^{s}, \overline{p}_{ml}^{s}\right\}$ depends on the type of traffic assignment. In the proportional assignment independence between traffic volumes and traffic proportions is assumed. The “all-or-nothing” assignment method and most stochastic methods are part of this group. Wherever congestion effects are important, capacity constraints and cost-flow relationships must be taken into account. Equilibrium assignment is, in this case, the best approach. The value of traffic proportions depends on the flow levels at all links, and cannot be determined independently of the trip matrix estimation process.

The problem of estimating U_{mn} can be expressed in the following general form:

$$\min\left\{\gamma_1 F_1(U_{mn}, U_{mn}^0) + \gamma_2 F_2(V^S, V_0^S)\right\} \tag{11}$$

subject to Eq. (10), where U_{mn}^0 is a target O/D matrix, which is typically assumed to come from an old O/D matrix or a sample survey, and V_0^S is the observed set of traffic counts, which may be regarded as an observation of the real flows V^S to be estimated. F_1 y F_2 are functions measuring distance between U_{mn}^0 vs. U_{mn}, and V_0^S vs. V^S respectively. γ_1 and γ_2 are the relative weights to be calibrated depending on the reliability of U_{mn}^0 and V_0^S.

Having established the solution procedure, a decision must be made for practical applications concerning:

- the specific formulation (11) of the O/D matrix estimation method to be chosen,
- the assignment model to be implemented in the procedure, in coherence with the O/D matrix estimation method,
- the selection of traffic count locations in the cordon area,
- the selection of cordon survey stations and data collection.

The next section is devoted to applying the solution procedure to 5 inter-urban networks. The aforementioned questions will be tackled with special focus on the cordon survey selection and its cost-effectiveness.

Application to 5 inter-urban networks in the Valencia Region

Cordon surveys and traffic counts were conducted in 2004-2006, and traffic studies have been carried out in 2005-2006, for the following inter-urban networks in the Valencia Region (Spain):

1°) Alicante Centro

2°) Valencia CV 50

3°) Castellón CV 10

4°) Area Metropolitana de Valencia AMV-V 30

5°) La Safor

Former automatic vehicle identification studies have been used in the AMV - V30 and Valencia CV 50 networks to obtain prior matrices (1998). Outstation cameras identified vehicles using automatic number-plate recognition at several cordon stations. Traffic counts in the period 1994-2005 have been utilized in order to obtain estimation of older O/D matrices and, therefore, justified medium-term traffic forecasts (2010-2012).

Selection of the O/D matrix estimation model

A traffic modelling based approach is selected for the formulation (11): the maximum entropy model (Van Zuylen and Willumsen 1980). This model is extended to include congestion effects and, therefore, equilibrium assignment. (see *Implementation of the assignment model* below). It can be integrated within the four-step methodology defined in the preceding chapter. The link flows are known at a relatively high number of traffic counting stations in the 5 inter-urban networks, and the rules for optimal location of traffic counting stations have been respected (see *Selection of traffic counting locations* below). Therefore, the entropy approach approximates the GLS approach, which also enables survey and traffic count data to be combined in an integrated way, but estimates in practice fitted values which frequently contravene non-negativity or other constraints (Bell, 1991). The algorithm used to solve the entropy model is efficient enough in computational time for inter-urban networks. Moreover, it is conceptually simple compared to other algorithms, which are more complicated and difficult to implement with real traffic engineering data. This simplicity is essential in a context where many data need to be corrected. Also, links, nodes, centroids and initial assumptions are usually modified numerous times, as well as travel costs, equivalent value of time for toll systems, etc. Complicated methods are much more time-consuming and expensive in human resources under real normal O/D study conditions, without any guarantee of giving better results.

Finally, our experience at the Dept. of Transportation Engineering gained through numerous traffic studies in the Valencia Region has taught us that when the decision-maker understands the algorithm used to solve the model, this affords a big advantage.

This model can be obtained from Eq. (10) and (11) by making $F_2 = 0$ and following an entropy maximizing approach:

$$\max\left\{-\sum_{m,n}\left(U_{mn}\ln U_{mn} - U_{mn}\right)\right\} \tag{12}$$

subject to

$$V_0^{\,s} - \sum_{m,n} p^s_{mn}U_{mn} - \sum_{k,l}\overline{p}^s_{kl}T*_{kl} - \sum_{k,n}\overline{p}^s_{kn}Q*_{kn} - \sum_{m,l}\overline{p}^s_{ml}R*_{ml} = 0 \qquad \text{and} \qquad U_{mn} \geq 0 \tag{13}$$

This formulation implies that $\{T*_{kl}, Q*_{kn}, R*_{ml}\}$ are known and, thus, can be fixed along the O/D matrix estimation procedure. This assumption is acceptable as long as sampling errors are sufficiently small (see "*Selection of cordon survey stations and data collection*" below).

The solution to the program defined by Eq. (12) and (13) is obtained, using Lagrangian methods, such as

$$U_{mn} = \prod_{S} X_S^{p^s_{mn}} \tag{14}$$

where $X_S = e^{-\lambda_S}$ and λ_S is the Lagrangian multiplier associated with the count on link *s*.

This problem can be solved by an algorithm described by Van Zuylen and Willumsen (1980). In this algorithm, it is easy to fix O/D elements $U_{mn} = 0$ if there is sufficient evidence for this assumption. These elements should be annulled in the prior O/D matrix before the first iteration of the algorithm. Prior O/D matrices for the 5 networks have been obtained from a trip distribution model adjusted in 2001.

The solution (14) reproduces the observed flows. Consistency in estimation results has been studied by Van Zuylen and Willumsen (1980), Willumsen (1981), Bell (1983), Van Zuylen and Branston (1982) and, lately, by Jörnsten and Wallace (1993). As average annual daily traffic (AADT) has been utilized in the 5 inter-urban studies, inconsistency in traffic counts arise because of counting errors and AADT calculation procedure. AADT has been obtained with different degrees of reliability (primary, secondary and coverage stations), from the 3 public administrations responsible for the different networks. As Jörnsten and Wallace (1993) conclude, consistency in estimation results should not be expected given the way traffic counts are made. Their approach uses the counts directly by trying to find an O/D matrix that fits the data as well as possible. Following the same approach here, the Eq. (13), which compares observed and modeled flows, has been applied as convergence criteria. These criteria must be verified for each link where a count is available, for adequate levels of relative and absolute errors:

$$\varepsilon_R \geq \left|V^S - V_0^S\right| / V_0^S \qquad \text{whether} \qquad V_0^S \geq LIM \tag{15}$$

$$\varepsilon_A \geq \left|V^S - V_0^S\right| \qquad \text{whether} \qquad V_0^S \leq LIM \tag{16}$$

where LIM is the AADT threshold for deciding whether to apply the relative or the absolute convergence criteria. Usually, both criteria requirements are strictly applied (equal sign) for a small percent of traffic counts.

Implementation of the assignment model

Equilibrium assignment is selected to take congestion effects into account, which are relatively important, especially in the AMV - V 30 network. The assignment and O/D estimation models are interdependent. An iterative program is implemented to incorporate both models. Firstly, link costs are treated as stochastic variables, and estimated consistently with traffic counts. The proportions p_{mn}^s are calculated by a Monte Carlo algorithm; where each simulation is computed with the Dijkstra's shortest path algorithm (Dijkstra, 1959).

The first O/D matrix U_{mn} is estimated applying the maximum entropy model, with link counts V_0^S. U_{mn} is assigned to the network and the cost – flow relationships are verified. The Highway Capacity Manual (2000) cost – flow relationships have been implemented in the model.

The link costs are consequently modified and the new proportions $\left\{p_{mn}^s\right\}^1$ are calculated. The second O/D matrix $\left\{U_{mn}\right\}^1$ is then estimated with the same link counts. This procedure is repeated until the difference between two iterations $\left\{p_{mn}^s\right\}^I$ and $\left\{p_{mn}^s\right\}^{I+1}$ is sufficiently small. The matrix $\left\{U_{mn}\right\}^{I+1}$ is the proposed solution to the equilibrium assignment problem.

Selection of traffic counting locations

The traffic counting location is critical to the accuracy of the estimated O/D matrix. The research of Yang and Zhou (1998) derived four rules to locate traffic counting stations: the O/D covering rule, the maximal flow proportion rule, the maximal flow-intercepting rule and the link independence rule. Recently, Gan et al. (2005) investigated traffic counting location, the O/D estimation method, the assignment model and the error bound in an integrated way. In the context of this application, traffic counts are collected at all the cordon stations and at a certain number of stations in the cordoned area. Traffic counts are available at a relatively high number of traffic counting stations in the 5 inter-urban networks considered (see Table 1). The O/D covering rule, the maximal flow proportion rule and the maximal flow-intercepting rule have been successfully verified for the 5 networks. Concerning the link independence rule, as many traffic counts are available, consistency is not perfect and some links with collected traffic flows are dependent. The convergence criteria (15) and (16) have been checked to ensure a level of traffic count consistency (see Table 1 for the error limits that have been applied).

Selection of cordon survey stations and data collection.

Within the established theoretical framework, several practical questions about the cordon surveys are yet to be answered. The number and location of cordon surveys has to be decided, depending on cost/effectiveness ratios, within the integral procedure of O/D matrix estimation. Conducting many cordon surveys could be too expensive, compared to the relative improvement of the O/D matrix estimates. This is especially the case when inbound and outbound traffic flows are small compared to interior traffic. Survey costs and the quality of the results are also affected by sample sizes. Finally, traffic variability must be taken into account.

Cordon surveys were conducted in August 2004 – March 2006. Previously, permission was requested and planned activities communicated to the 3 public administrations responsible for the different networks (local, regional, national). The highway concessionary companies and gas stations where the surveys were to be conducted were also informed. Two types of costs can be distinguished, therefore, in the surveys: management and operation. Survey preparation time was similar to the time needed to complete them. Management and survey preparation and organization costs involve: requests made to the different administrations, interviews, data collection, accreditations, travel and identification of optimal places for the surveys, as well as computerization of data. Direct survey costs are interview costs and interviewers' travel costs. Random sampling without replacement has been selected. The reason for this choice was that there is an element that stops the traffic in all the survey stations. This element allows the different types of cordon surveys to be classified:

- Urban arterials: vehicles stopping at a traffic-light used as a control point.
- Gas stations: vehicles stopping to refuel

- Freeway tolls: vehicles stopping to take ticket or to pay the toll.

These elements have guaranteed random sampling. The surveyed variables have been, in addition to trip origin and destination, exit or entry of cordon station, travel purpose and frequency, vehicle type and car occupancy.

The distribution of the population sampled in a cordon station is multinomial. A formula to estimate sample size for multinomial populations is derived in Thompson (1987), requiring no prior knowledge of the proportion of each category (T_{ijkl} / O_k and T_{ijkl} / D_l). This formula is useful in this approach, because no efficient prior estimates of T_{ijkl} are available. The size of the multinomial population (the traffic volume counted at the cordon station) is large enough in all the cordons considered (>2500 vehicles/day). Therefore, in this case, the finite population correction factor can be ignored and normal approximation can be applied. Thompson (1987) showed that the sample size depends only on the simultaneous level of significance for all the categories and the width of the confidence intervals selected for the observed proportions of T_{ijkl} / O_k and T_{ijkl} / D_l. Selecting a level of significance equal to 90 % and a width of the confidence intervals equal to 0.10, sample size is 403, independent of the number of categories. As shown later in Table 2 sample sizes for the cordon stations are bigger; only in 1 cordon out of 16 the local restrictions prevented the interviewers from obtaining the 403 interviews (inbound + outbound). The strategy was to determine first the number of interviewers to reach 403 interviews/inbound - outbound station (O/D matrix is approximately symmetric for 4 networks out of 5). In cordon stations with high traffic volumes (> 25000 vehicles/day) this number was increased, where possible, to reach a relative sample size of, at least, 7 %. Once the interviewer was selected, the instruction given was to make as many interviews as possible. The highest sample sizes were obtained at freeway tolls, followed by urban arterials and, finally, gas stations. Thus, the level of significance and the width of the confidence intervals are smaller at gas-station surveys; in which doubts also arise regarding the randomness of the sample. On the other hand, cordon surveys have been conducted in summer and autumn to take into account the seasonal character of traffic along the Mediterranean coast. This article shows the results of the autumn surveys, closer to the average annual traffic values. Concerning weekly variability, Tuesdays, Wednesdays and Thursdays were considered as average working days. For that reason all the surveys were limited to these days. Finally, as far as daily variability, surveys were conducted from 8:00 to 20:00 to grasp possible variations. More than 29000 valid interviews were conducted, of which 22200 correspond to the 5 networks presented in this article.

Table 1 shows the main network characteristics, the aggregated data collection variables and the main parameters for method application, such as convergence limits and total number of prior unknown O/D matrix elements. Results for the cordon studies are summarized in Table 2. Absolute and relative sample sizes are shown for inbound and outbound traffic for each network and cordon number.

TABLE 1. Case studies and traffic characteristics

	Road Network Case Study				
	Alicante Central	Valencia CV50	Valencia V30	Castellón CV10	La Safor
Links (*L*)	134	201	320	128	185
Nodes (*M*)	118	163	222	104	131
Centroids (*N*)	42	55	49	36	39
Internal centroids (*Z*)	34	46	42	30	32
Centroid connectors	53	74	66	50	55
Total traffic counts ($2\,c + S$)	59	59	77	57	68
Cordon traffic counts ($2\,c$)	16	18	14	12	14
Cordon surveys conducted	8	8	10	6	4
Average absolute sample size	518	770	575	629	592
Average relative sample size	4,1 %	5,7 %	3,7 %	5,2 %	13,8%
Inbound+outbound trips/day	205556	233472	351388	160700	113655
O/D matrix elements equal to zero	313	430	555	88	89
Unknown O/D matrix elements	1409	2540	1797	1172	1393
Relative error	3%	3%	3%	4%	3,5%
Absolute error	300	300	300	400	400
AADT threshold (*LIM*)	7000	7000	7000	8000	6000

TABLE 2. Cordon surveys

		Road Network Case Study				
		Alicante Central	Valencia CV50	Valencia V30	Castellón CV10	La Safor
Absolute sample sizes (inbound traffic) $\sum_l t_{kl}$	*k=1*	1565	1412	293	1147	455
	k=2	163	293	1412	363	425
	k=3	109	1530	426	242	
	k=4	304	235	235		
	k=5			581		
Absolute sample sizes (outbound traffic) $\sum_k t'_{kl}$	*l=1*	1227	1147	292	1412	797
	l=2	164	292	1147	363	689
	l=3	307	896	428	242	
	l =4	306	356	356		
	l =5			580		
Relative sample sizes (inbound traffic) $\sum_l t_{kl} / O_k$	*k=1*	8.10%	8.80%	2.00%	7.60%	10,26%
	k=2	1.20%	2.00%	9.30%	3.50%	10,49%
	k=3	1.20%	11.20%	1.30%	2.30%	
	k=4	3.50%	2.40%	2.40%		
	k=5			11.10%		
Relative sample sizes (outbound traffic) $\sum_k t'_{kl} / D_l$	*l=1*	6.30%	7.10%	2.00%	9.30%	17,81%
	l=2	1.20%	2.00%	7.60%	3.50%	16,62%
	l=3	3.30%	6.50%	1.30%	2.10%	
	l =4	3.50%	3.70%	3.70%		
	l =5			11.20%		

Deviation measurements of O/D matrix estimates

In order to ease the notation let θ_w be the global O/D matrix including all pairs $w \in W$; and W the set of all O/D pairs considered in Steps I to IV. The true O/D matrix θ_w is unknown and, therefore, the error of the estimation $\theta*_w$ cannot be obtained in real problems. However, effectiveness of cordon surveys may be assessed comparing the estimates $\theta*_w$ with the best O/D matrix $\overline{\theta_w}$, which is obtained with the maximum number of cordon surveys. Various deviation measurements may be used for this purpose (Gan et al., 2005). In order to get the maximum information, three measurements will be computed:

- Total demand deviation: $$TD(\%) = 100\left|\frac{\sum_{w\in W}\theta *_w - \sum_{w\in W}\overline{\theta_w}}{\sum_{w\in W}\theta *_w}\right| \quad (17)$$

- Weighted relative error: $$WR(\%) = 100\sqrt{\sum_{w\in W}\left(\frac{\theta *_w - \overline{\theta_w}}{\theta *_w}\right)^2 \frac{\theta *_w}{\sum_{w'\in W}\theta *_{w'}}} \quad (18)$$

- Root mean square error: $$RM(\%) = 100\frac{\sqrt{\frac{1}{n}\sum_{w\in W}\left(\theta *_w - \overline{\theta_w}\right)^2}}{\frac{1}{n}\sum_{w\in W}\theta *_w} \quad (19)$$

These error measurements refer only to the global O/D matrix. They will also be computed for the cordoned area, excluding the estimated flow between external zones $\{TDC(\%), WRC(\%), RMC(\%)\}$, and for the internal trips of the cordoned area $\{TDI(\%), WRI(\%), RMI(\%)\}$

Relative error measurements are excluded to avoid excessive importance of small values of $\theta *_w$ in the deviation. In order to compute these three deviation measurements, the estimates $\theta *_w$ must be obtained step by step, from the "0" situation without cordon studies to the final situation when all cordon studies have been taken into account and, therefore, the best O/D matrix $\overline{\theta_w}$ has been estimated. For each phase of this procedure, the 4-step method described above should be applied. The computational results for all networks are summarized in Table 3. From the results obtained the following remarks can be made.

TABLE 3. O/D matrix estimation computational results and deviation measures

Road Network Case Study	Variables	Number of cordon studies utilized in the O/D matrix estimation process					
		X=0	*X=1*	*X= 2*	*X=3*	*X=4*	*X=5*
Alicante Central	T(X)	526.4	528.6	520.8	518.4	517.1	
	TE(X) TI(X)	39.1 369.6	28.9 361.9	22.7 347.9	22.2 344.7	21.2 342.3	
	N(X)	0	38.5	65.3	82.8	101.4	
	NM(X)(%)	0.0	18.7	31.8	40.3	49.3	
	TD(%)	1.76	2.17	0.70	0.24	0.00	
	RM(%)	107.7	71.9	52.2	25.7	0.0	
	WR(%)	59.8	42.8	30.9	12.9	0.0	
	TDC TDI (%)	1.77 7.41	0.74 5.4	0.42 1.61	0.04 0.71	0.0 0.0	
	RMC RMI (%)	94.7 65.0	65.9 44.5	53.6 38.5	25.6 12.8	0.0 0.0	
	WRC WRI (%)	59.1 35.8	42.4 27.1	31.1 23.6	12.5 6.7	0.0 0.0	
Valencia CV 50	T(X)	514.2	511.7	513.2	509.4	513.4	
	TE(X) TI(X)	4.2 291.9	11.8 297.1	16.7 301.3	22.6 303.6	25.4 310.8	
	N(X)	0	32.1	61.7	89.1	113.3	
	NM(X)(%)	0.0	13.8	26.4	38.2	48.5	
	TD(%)	0.15	0.34	0.04	0.79	0.00	
	RM(%)	222.8	170.5	96.9	43.1	0.0	
	WR(%)	128.1	45.4	35.9	19.0	0.0	
	TDC TDI (%)	4.31 6.50	2.36 4.63	1.70 3.17	0.26 2.40	0.0 0.0	
	RMC RMI (%)	227.9 163.3	179.4 131.2	101.9 80.1	45.2 51.5	0.0 0.0	
	WRC WRI (%)	106.3 62.2	45.9 44.1	36.4 53.2	19.4 45.9	0.0 0.0	
Valencia V 30	T(X)	843.0	858.2	854.4	854.9	851.3	850.9
	TE(X) TI(X)	24.8 517.6	29.2 527.0	40.8 522.5	41.9 523.0	44.5 530.0	44.9 530.5
	N(X)	0	79.2	145.4	211.6	231.0	248.2
	NM(X)(%)	0.0	22.5	41.4	60.2	65.7	70.7
	TD(%)	0.94	0.84	0.41	0.46	0.04	0.00
	RM(%)	184.6	159.7	120.3	36.9	12.3	0.0
	WR(%)	94.6	77.5	37.7	16.7	6.3	0.0
	TDC TDI (%)	1.49 2.48	2.77 0.66	0.94 1.52	0.87 1.43	0.10 0.09	0.0 0.0
	RMC RMI (%)	184.2 73.8	161.9 60.7	126.5 56.4	37.7 22.9	12.9 13.2	0.0 0.0
	WRC WRI (%)	56.7 37.4	48.0 33.9	36.9 30.7	17.0 12.7	6.2 5.1	0.0 0.0
Castellón CV 10	T(X)	246.0	245.7	253.2	253.1		
	TE(X) TI(X)	22.5 136.5	23.2 138.0	23.0 142.7	23.3 143.1		
	N(X)	0	66.2	86.9	108.8		
	NM(X)(%)	0.0	41.2	54.1	67.7		
	TD(%)	2.90	3.00	0.04	0.00		
	RM(%)	146.7	91.1	59.7	0.0		
	WR(%)	55.7	36.1	25.8	0.0		
	TDC TDI (%)	2.8 4.87	3.28 3.69	0.16 0.32	0.0 0.0		
	RMC RMI (%)	152.9 106.1	100.3 112.8	65.7 72.3	0.0 0.0		
	WRC WRI (%)	57.8 38.8	37.9 37.3	26.8 29.5	0.0 0.0		
La Safor	T(X)	182.2	186.1	186.0			
	TE(X) TI(X)	25.9 94.4	24.4 96.9	25.3 97.7			
	N(X)	0	27.9	52.6			
	NM(X)(%)	0.0	24.5	46.3			
	TD(%)	2.10	0.06	0.00			
	RM(%)	86.6	25.2	0.0			
	WR(%)	46.7	23.9	0.0			
	TDC TDI (%)	2.8 3.5	0.64 0.84	0.0 0.0			
	RMC RMI (%)	97.6 104.6	25.0 15.0	0.0 0.0			
	WRC WRI (%)	48.3 39.0	21.7 9.9	0.0 0.0			

Note:T(X)=total number of trips(thousands)/day; TE(X)= total number of trips (thousands)/day between external zones; TI(X)= total number of trips (thousands)/day between internal zones; N(X)=surveyed inbound & outbound trips (thousands/day); NM(X)=surveyed/total inbound& outbound trips; TD=total demand deviation; RM=root mean square error; WR=weighted relative error; TDC, RMC, WRC: the former 3 error measures calculated for the cordoned area; TDC, RMC, WRC: the former 3 error measures calculated for the cordoned area; TDI, RMI, WRI = Id. for the cordoned area internal trips.

Firstly, the initial O/D matrix is estimated exclusively with traffic counts. Then, the cordon is selected where there is maximum inbound + outbound traffic; 1 row and 1 column of the O/D matrix are directly estimated from these 2 surveys by factoring to the corresponding 2 traffic counts. Row and column elements are fixed and the second O/D matrix is estimated with traffic counts. In the third phase, the 2° cordon with maximum traffic is added to the first one, and O/D matrix elements (2 rows and 2 columns) must be estimated by applying the 4-step methodology described in this article. In the last phase of this procedure the cordon with minimum traffic is added to the rest, and the best possible O/D matrix is estimated using the same method. Each phase is characterized by the number of cordon surveys utilized and the total number of inbound and outbound trips which can be fixed in Steps I to III of the methodology before the complete O/D matrix is estimated with traffic counts (Step IV).

Let us suppose that *X* cordon surveys (inbound + outbound) have been employed. The total number of fixed inbound and outbound trips is *N(X)*:

$$N(X) = \sum_{k=1}^{X}\sum_{l=1}^{c} T*_{kl} + \sum_{k=X+1}^{c}\sum_{l=1}^{X} T*_{kl} + \sum_{k=1}^{X} T*_{k0} + \sum_{l=1}^{X} T*_{0l} \tag{20}$$

If *N(X)* is divided by the total number of inbound and outbound trips estimated in the phase number "*X*" of the method, we have

$$NM(X) = N(X) / \sum_{k=0}^{c}\sum_{l=0}^{c} T*_{kl}(X) \tag{21}$$

NM(X) measures the % of determination of the O/D matrix provided by the cordon surveys. In order to assess the effectiveness of the cordon surveys, it is useful to know the correlation between NM(X) and the deviation measurements. Numerical results of WR, RM, WRC, RMC, WRI and RMI depending on NM(X) are shown for all networks in Fig. 2. The estimated flow between external zones, compared to the total traffic, ranges from 4,1% (Alicante Central) to 13,6% (La Safor). As this percentage is relatively small, the differences between {WR, RM} and {WRC, RMC} are also small. The estimated internal traffic of the cordoned area compared to the total traffic ranges from 52,5% (La Safor) to 66,2% (Alicante Central). The differences between {WR, RM} and {WRI, RMI} are significant.

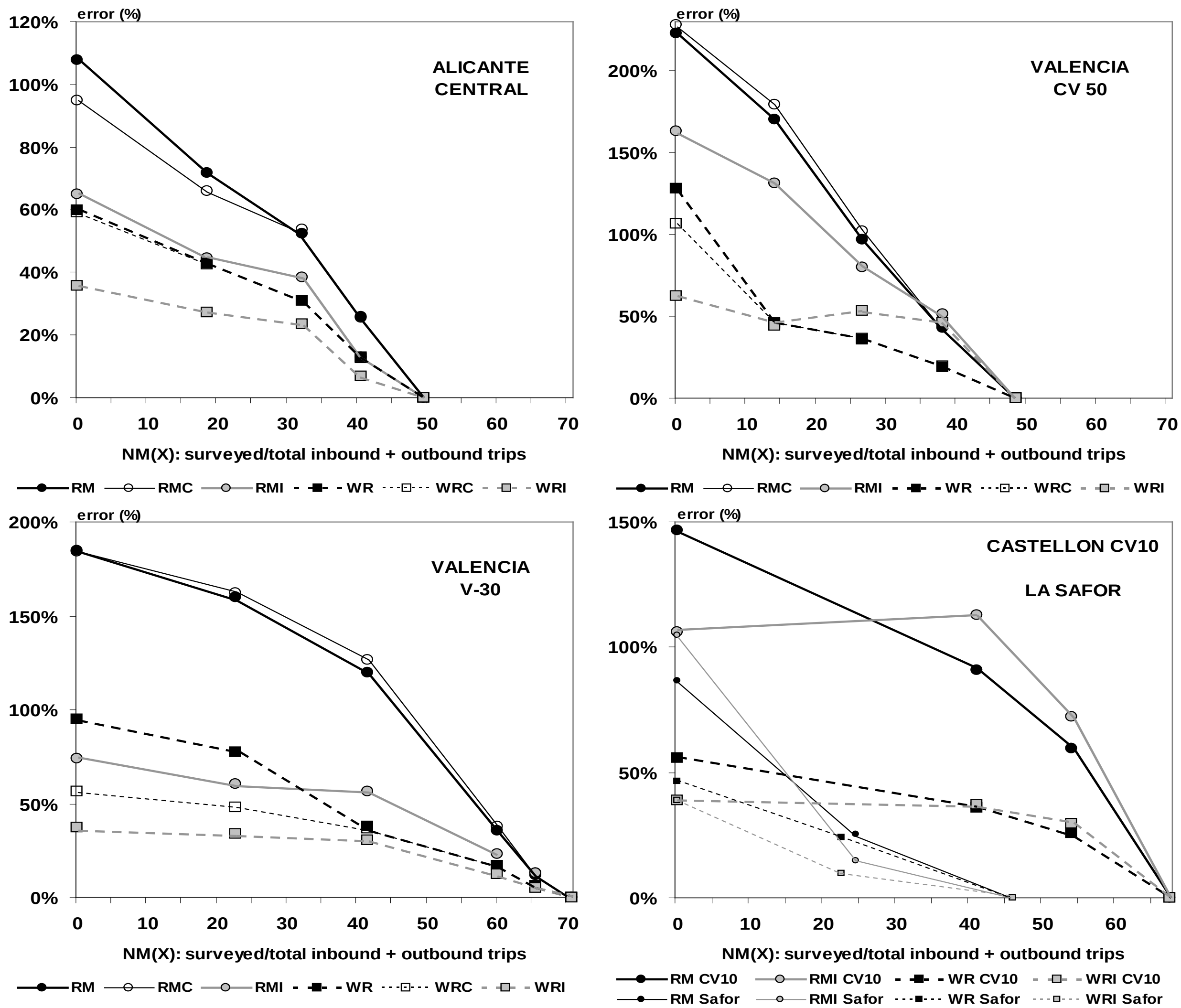

FIGURE 2. Deviation measures RM and WR with variable ratio of surveyed / inbound and outbound trips NM(X)

Percentages of reduction in root mean square and weighted relative errors are significant when *NM(X)* increases, being greater for RM and for the total O/D matrix. These percentages depend on particular network characteristics and traffic patterns; however, Fig. 2 shows behavior is similar for all networks, except in the CV 50 case. There are larger error reductions for the CV 50 due to the greater O/D matrix indetermination degree for this network. Percentages of reduction in RM and WR are not linear with *NM(X).* Consequently, marginal utility of conducting more cordon surveys tends to increase with *NM(X)*. Therefore, information provided by cordon surveying is of great value in reducing O/D estimation errors.

Variation in traffic generation and attraction per centroid when *NM(X)* increases is an interesting model outcome, because traffic forecasts are based on future generation and attraction estimates via distribution models. Fig. 3 compares, for each network and per centroid, trip generation and attraction between O/D estimation without cordon surveys and O/D estimation with the maximum number of cordon surveys conducted. Estimation errors are relatively important for a significant number of centroids when O/D is estimated exclusively from traffic counts.

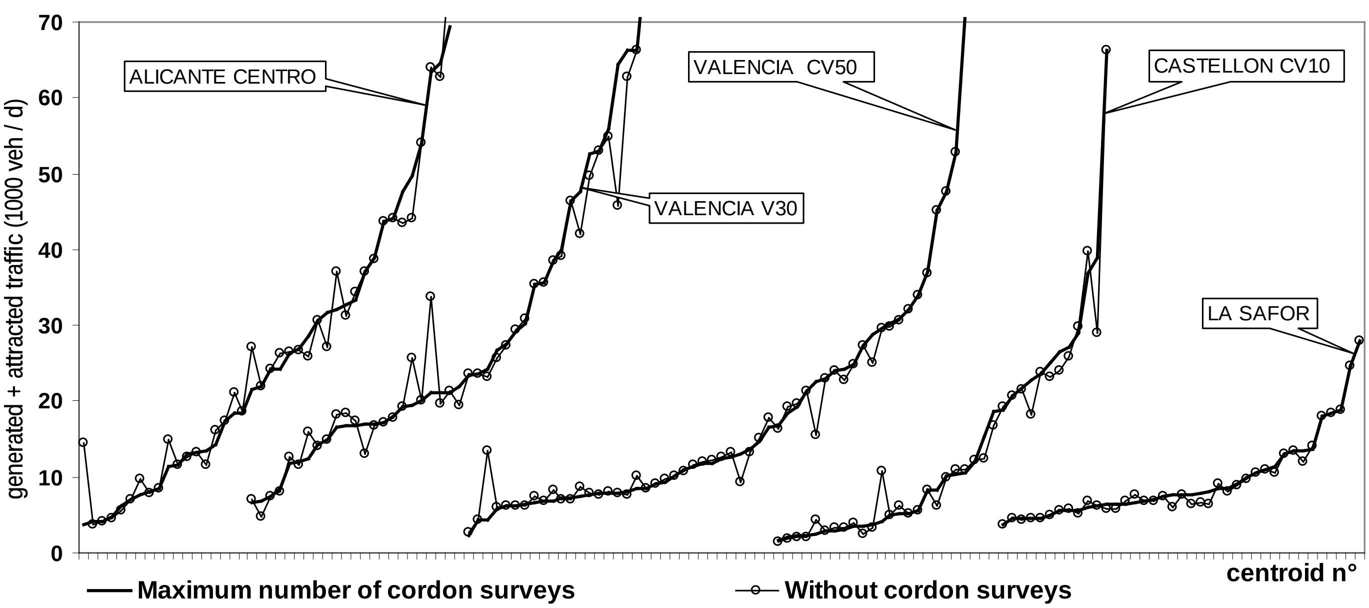

FIGURE 3. Trip generation and attraction estimated for the global O/D matrix: a) without cordon surveys and b)using the maximum number of cordon surveys

Cost-effectiveness. Discussion of numerical results

Cost effectiveness is a technique for comparing relative costs and effects associated with two or more courses of action. Cost-effectiveness is typically expressed as an incremental cost-effectiveness ratio (*CER*): change in costs divided by change in effects. Let $Cost(\theta*_w)$ be the total economic cost of conducting cordon surveys in order to estimate $\theta*_w$. Comparing the costs of obtaining the estimates $\theta*_w$ and $\overline{\theta_w}$ the result is:

$$CER = \frac{Cost(\bar{\theta}_w) - Cost(\theta*_w)}{ERM(\bar{\theta}_w, \theta*_w)} \tag{23}$$

where $ERM(\bar{\theta}_w, \theta*_w)$ is a measure of deviation between $\theta*_w$ and $\overline{\theta_w}$. In the 5 real cases studied measures RM and WR are utilized. TD is discarded because it is much less pertinent as an indicator than WR or RM. The result may be considered as “the price” of obtaining the best estimate $\overline{\theta_w}$ from $\theta*_w$ by eliminating deviation between them. Conducting cordon surveys is cost-effective if this price is low enough. Therefore, a value judgment about this ratio is required. These CER measures refer only to the global O/D matrix. In order to obtain information about the cost performance of the estimation of the cordoned area O/D matrix, the corresponding CER measures for the cordoned area (CERC) and for the internal trips (CERI) have been computed. The calculations of the cost-effectiveness ratios are summarized in Table 4. The following observations are made from the results.

TABLE 4. Cost – effectiveness ratios

Road Network Case Study	Variables	Cordon studies utilized in the O/D matrix estimation process				
		1	*1,2*	*1,2,3*	*1,2,3,4*	*1,2,3,4 5*
Alicante Central	SMC (€)	1389	1531	1682	1853	
	SMI (€)	395	705	1015	1325	
	TSC (€)	1784	2236	2697	3178	
	CER RM (€ / %)	49.9	40.3	32.9	29.5	
	CERI RMI (€ / %)	87.0	84.4	51.7	48.9	
	CER WR (€ / %)	104.5	77.3	57.4	53.1	
	CERI WRI (€ / %)	205.0	183.3	92.7	88.8	
Valencia CV 50	SMC (€)	1015	1183	1538	1707	
	SMI (€)	308	530	838	1060	
	TSC (€)	1323	1713	2375	2767	
	CER RM (€ / %)	25.3	13.6	13.2	12.4	
	CERI RMI (€ / %)	41.3	20.6	21.2	16.9	
	CER WR (€ / %)	16.0	18.6	21.8	21.6	
	CERI WRI (€ / %)	-	191.7	146.3	44.5	
Valencia V 30	SMC (€)	774	1141	1337	1506	1732
	SMI (€)	212	509	721	933	1145
	TSC (€)	986	1650	2058	2439	2877
	CER RM (€ / %)	39.5	25.6	13.9	14.1	15.6
	CERI RMI (€ / %)	-	94.9	40.4	40.2	39.0
	CER WR (€ / %)	59.8	51.9	35.2	29.0	29.4
	CERI WRI (€ / %)	281.7	246.3	83.3	75.5	76.9
Castellón CV 10	SMC (€)	1085	1268	1426		
	SMI (€)	325	565	805		
	TSC (€)	1410	1833	2231		
	CER RM (€ / %)	25.3	21.1	15.2		
	CERI RMI (€ / %)	-	54.2	21.0		
	CER WR (€ / %)	71.9	61.3	40.1		
	CERI WRI (€ / %)	-	199.2	57.5		
La Safor	SMC (€)	953	1175			
	SMI (€)	224	443			
	TSC (€)	1177	1618			
	CER RM (€ / %)	19.2	18.7			
	CERI RMI (€ / %)	13.1	15.5			
	CER WR (€ / %)	51.6	34.7			
	CERI WRI (€ / %)	40.5	41.5			

Note: SMC=Survey management cost; SMI=Survey interviews cost; TSC=Total survey cost; CER RM=Cost-effectiveness ratio - root mean square error for the global O/D matrix; CERI RMI= Id. for the cordoned area internal trips; CER WR=Cost-effectiveness ratio - weighted relative error for the global O/D matrix; CERI WRI= Id. for the cordoned area internal trips

Cost-effectiveness ratios decrease, in general, with the number of surveyed cordons and the total traffic surveyed. Cordon stations with relatively low traffic intensities (8000 veh/day) show cost-effectiveness ratios lower than those of cordon stations with high traffic intensities (> 30000 veh/day). For the global O/D matrix, cost of RM 1 % reduction varies between 12 and 50 €, while cost of WR 1 % reduction varies between 16 and 105 €. Cordon studies are less cost-effective for estimating internal trips, especially for the WRI measures, because cost of RMI 1 % reduction varies between 13 and 95 €, while cost of WRI 1 % reduction varies between 40 and 282 €.

CER and CERI decrease with NM(X) for all the case studies, excepting the CV 50 case using WR measure. In this particular case, CER is almost stable with NM(X). The reason has been mentioned above; WR is strongly reduced by the first cordon due to its high traffic volume and the high number of unknown O/D matrix elements (2540), compared to the other networks. For all study cases, CERI strongly decreases with NM(X) for the WRI measures.

The decrease in cost-effectiveness ratios with the number of cordon surveys is due to the following reasons:

- Economies of scale on conducting surveys. Management and survey preparation and organization costs are mostly fixed costs for a given network. Conducting more cordon surveys adds interview costs but only a part of the costs related to management, survey preparation and data treatment. Therefore, to some extent, in terms of cost-effectiveness economies of scale compensate conducting more cordon surveys.
- Marginal utility of conducting more cordon surveys does not decrease (Fig. 2)

Therefore, in general, conducting more cordon surveys is cost-effective compared to restricted survey data collection. If conducting just one cordon survey is deemed cost-effective, then more cordon surveys must be carried out unless financial restrictions prevent extending data collection. Extending cordon surveys to 100 % of inbound/outbound traffic seems to be cost-effective provided that traffic volume of all cordon stations is above 8000 veh/day. If traffic volume of cordon stations is below 2500 veh/day, relative sample size of cordon surveys must be increased. CER will probably increase, because of the extra cost of surveys which could not significantly reduce RM or WR.

CERs are strongly dependent on case studies, from approximately 1 to 3 for RM and from 1 to 4 for WR for the global O/D matrix (Fig. 4). These differences are due to the distance from the network case study and the regional capital Valencia. Concerning absolute CER values, a value judgment is always required, but the costs of reducing errors do not seem extreme for the case studies analyzed (between 2300 and 3300 € to be invested for each network; the total amount being 13500 €). It has to be pointed out that these traffic studies are not financed every year. In the Valencia Region, experience has taught us that this kind of study is repeated at a minimum of every 5 years.

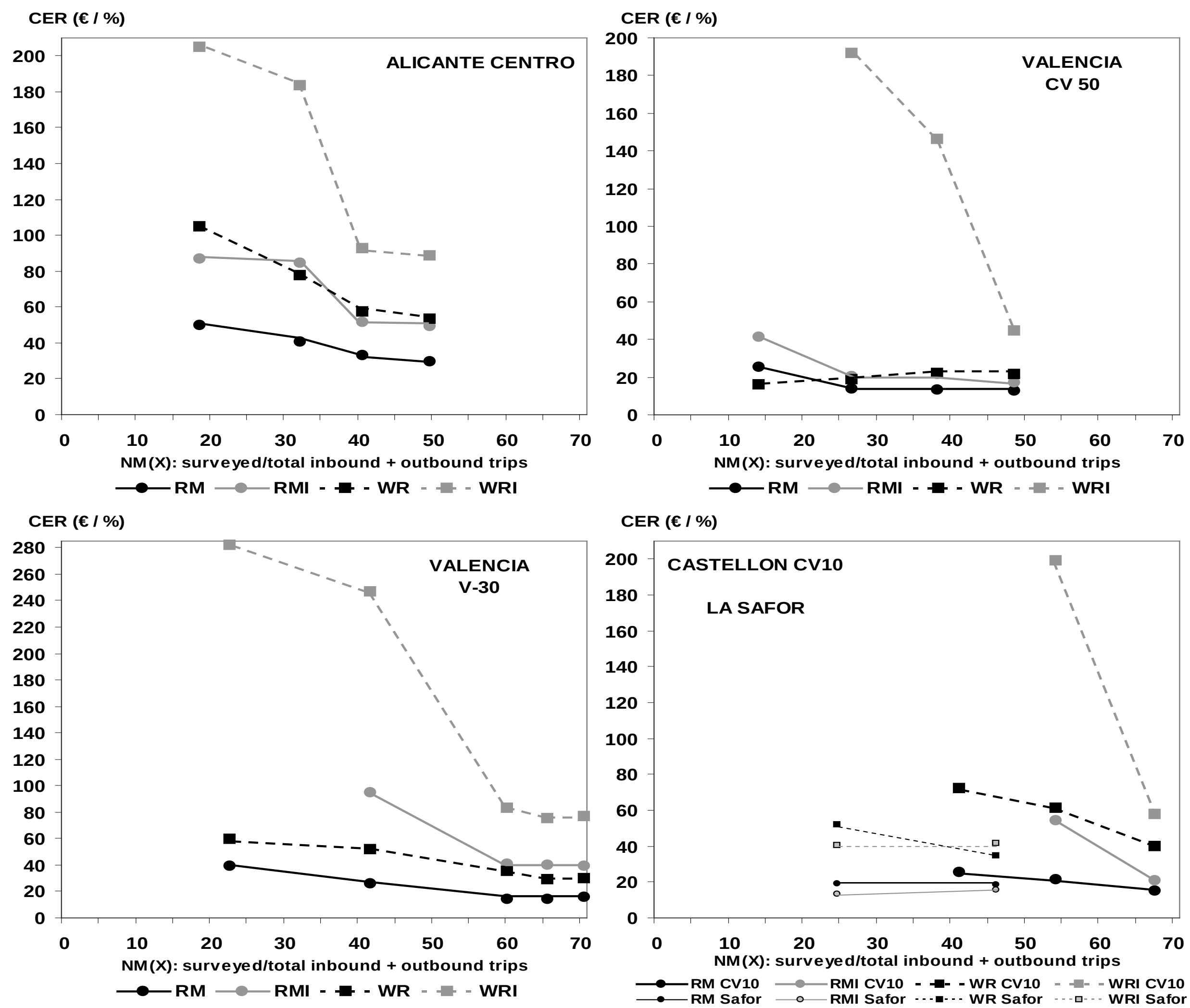

FIGURE 4. Cost-effectiveness ratios (CER) of reducing deviation measures RM and WR, with variable ratio of surveyed / inbound and outbound trips NM(X)

Remarks

Errors in data collection strongly affect outcome quality. In the 5 case studies, risks of inconsistency have been observed between surveyed trip origin and destination and model centroids. A careful centroid definition is necessary, which must be well understood by the interviewers. On the other hand, achieving random sampling is not always guaranteed. Results may also be distorted by traffic variability, which is not always easy to tackle. Errors in traffic counts have been detected. Inconsistency has been corrected iteratively, verifying source data, counting stations location, traffic coefficients, etc.

Model simplifications could affect the results. On applying this method to the 5 networks, single vehicle O/D information has been employed. However, the percentage of trucks varies between 10 % and 25 % at cordon stations. Link costs for calculating the proportions p_{mn}^{s} have been approximated by travel time for passenger cars. Tolls have been reduced to time savings by means of an average time value for all vehicles.

It is easier to apply this method to inter-urban models than to urban traffic networks. The comparative advantages could be summarized as follows.

- Degrees of determination of O/D matrices provided by cordon surveys *NM(X)* are usually greater for inter-urban studies. In a cordoned urban study area the percentage of inbound and outbound traffic over the total is assumed to be approximately 20%. By contrast, this percentage tends to be greater in a cordoned inter-urban study area. In the 5 case studies, this percentage ranges from 33.8% (Alicante Centro) to 47.5% (La Safor), with 37,7% for Valencia AMV-V30, 39.5% for CV 50 and 43.5% for Castellón-CV10.

- One cordon station can usually capture the flow of one external centroid in inter-urban networks. However, urban zones are connected to the network through street links. Therefore, more cordon stations are needed (sometimes a great many) to conduct O/D flow surveys with an origin or destination in a particular urban zone.

- Conducting traffic surveys is more difficult at urban accesses than at inter-urban stations.

- Urban studies require peak-hour analysis. This approach complicates the O/D traffic study; more data and analysis work are needed. Moreover, obtaining a sample size for a sufficient level of significance and confidence intervals may be extremely difficult in some links at peak hours. Finally, possible simplification from O/D matrix symmetry is discarded in peak-hour analysis.

- More links are heavily congested in urban networks. Although the Highway Capacity Manual cost – flow relationships for urban streets and arterials could easily be implemented in the model, the equilibrium assignment problem is more difficult to solve.

Conclusions

This paper proposes a method to integrate cordon survey results in a complete procedure of O/D matrix estimation from traffic counts. Once this method has been established, the main issue to be addressed is of a practical nature, i.e. the cost-effectiveness of conducting cordon surveys. In this study the method was applied to 5 inter-urban networks in the Valencia Region (Spain). Important aspects of the application were dealt with, such as cordon survey data collection, formulation of O/D matrix estimation and assignment models and selection of traffic count locations. Results can be usefully applied to assess decision making when conducting cordon studies in other inter-urban networks with different cost conditions and traffic patterns. Cost-effectiveness ratios decrease with the number of cordon surveys. This is explained by the increasing marginal utility and the economies of scale on conducting more cordon surveys. A prior value judgment about this ratio is required, but if the first cordon survey is deemed cost-effective, then more cordon surveys must be carried out unless financial restrictions prevent extending data collection. The case studies analyzed in this article indicate that cost-effectiveness ratios decrease until 70% of inbound and outbound trips have been surveyed, provided that traffic volume at all cordon stations surveyed is above 8000 veh/day. Values of ratio {cost / reduction of O/D estimates deviation measurements} are dependent on particular networks. Comparing the cost of reducing 1% root mean square error (RM), if 50% of cordoned trips are surveyed, costs range from 12 to 30 € for the global O/D matrix, and from 15 to 49 € for the internal trips of the cordoned area. For all networks considered, with the cost conditions that prevailed, conducting cordon surveys was considered clearly cost-effective.

For future research, cost-effectiveness of surveying more than 70% of inbound and outbound traffic must be examined. Model simplification effects should be checked. Multiple vehicle O/D matrix estimation could be implemented; in this case the O/D matrix estimation and assignment models should be modified. Cost-effectiveness of cordon studies for estimating urban O/D matrices from traffic counts should be addressed. Finally, data from screenline surveys are available for the 5 analyzed networks. This method must be adapted to these data in order to improve O/D estimation and address cost-effectiveness on conducting screenline surveys compared to cordon surveys.

Acknowledgements

The author wishes to thank Professors José V. Colomer and Julia I. Real for helpful discussions on data collection management and costs. English text corrected by F.Barraclough.

Notation

The following symbols are used in this paper:

CER = cost-effectiveness ratio for the study area

CERC = cost-effectiveness ratio for the cordoned area (excluding the trips between external zones)

CERI = cost-effectiveness ratio for the cordoned area (only internal trips)

N(X) = total number of inbound and outbound trips determined by *X* cordon surveys

NM(X) = *N(X)* divided by the total number of inbound and outbound trips estimated using *X* cordon surveys

O_k, D_l = inbound trips entering the cordon area at station *k* and outbound trips leaving the cordon area at station l

p_{mn}^{s} = ratio of total trips from origin *m* to destination *n* using link *s*

$Q*_{ikn}, R*_{mlj}$ = estimates of the inbound / outbound trips ending / starting in the cordon area zone *n / m* passing through station *k / l*

T_{kl} = total trips from cordon station *k* to cordon station *l*

$T*_{ijkl}$ = estimate of the flows from zone *i* to zone *j* via cordon stations *k* and *1*.

t_{kl}, t_{kl}' = random samples of vehicles obtained at station *k / l* from those entering/leaving the cordon area

U_{mn} = total trips between internal zones *m* and *n*

V^{s} = volume of traffic flow using link *s*

θ_w = global O/D matrix including all O/D pairs considered in Steps I to IV

$\theta*_w$ = estimate of the global O/D matrix θ_w

$\overline{\theta_w}$ = estimate of the global O/D matrix θ_w obtained with the maximum number of cordon surveys

References

Abrahamsson, T. (1998). "Estimation of origin-destination matrices using traffic counts – A literature survey". *IIASA Interim Report nº IR-98-021/May*, pp. 1 – 27. Laxenburg. Austria.

Baek, S., Kim, H., Lim, Y. (2004). "Multiple vehicle O/D matrix estimation from traffic counts using genetic algorithm". *Journal of Transportation Engineering*, Vol. 130, Issue 3, pp. 339 – 347

Bell, M.G.H. (1983). "The estimation of an origin-destination matrix from traffic counts". *Transportation Science*, Vol. 17, pp. 198 – 217

Bell, M.G.H. (1984). "Log-linear models for the estimation of origin-destination matrices from traffic counts : An approximation". Proceedings of the 9th International Symposium on Transportation and Traffic Theory. The Netherlands, Delft, July, pp. 451-470

Bell, M.G.H. (1991). "The estimation of origin-destination matrices by constrained generalised least squares". *Transportation Research*, Vol. 25B, pp. 13 – 22

Cascetta, E., Nguyen, S. (1988). "A unified framework for estimating or updating origin/destination matrices from traffic counts". *Transportation Research*, Vol. 22B, pp. 437 - 455

Dijkstra, E.W. (1959). "A note on two problems in connection with graphs". *Numerische Math*., 1, pp. 269-271.

Drissi-Kaitouni, O., Lundgren, J. (1992). "Bilevel OD matrix estimation using a descent based approach". Technical report LiTH-MAT-R-92-49, Department of Mathematics, Linkoping Institute of Technology. Sweden.

Echenique, M.H. (1977). "Methodology for the construction of origin-destination matrices from a roadside interview". *SPT Working Paper, Applied Research of Cambridge Ltd*.

Echenique, M.L., Williams, I.N. (1982). "O/D matrix production from cordon survey data" *Traffic Engineering and Control*, December 1982, pp. 584 – 589

Fisk, C.S. (1988). "On combining maximum entropy trip matrix estimation with user optimal assignment". *Transportation Research*, Vol. 22B, pp. 69 – 73

Fisk, C.S. (1989). "Trip matrix estimation from link traffic counts: the congested network case". *Transportation Research*, Vol. 23B, pp. 331 – 336

Gan, L., Yang, H., Wong, S.C. (2005). "Traffic counting location and error bound in origin-destination matrix estimation problems". *Journal of Transportation Engineering*, Vol. 131, Issue 7, pp. 524 – 534

Jörnsten, K., Wallace, S. W. (1993). “Overcoming the (apparent) problem of inconsistency in origin-destination matrix estimations”. *Transportation Science*, Vol. 27, pp. 374 – 380

Landau, U., Hauer, E., Geva, I. (1982). “Estimation of cross-cordon origin-destination flows from cordon studies”. *Transportation Research Record n° 89*, pp. 5 – 10. Transportation Research Board.

Lo, H.P., Lam, W.H.K., Zhang, N. (1996). “Estimation of an origin-destination matrix with random link choice proportions from traffic counts”. *Transportation Research*, Vol. 30B, pp. 309 – 324

Lo, H.P., Zhang, N., Lam, W.H.K. (1999). “Decomposition algorithm for statistical estimation of OD matrix with random link choice proportions: a statistical approach”. *Transportation Research*, Vol. 33B, pp. 369 – 385

Spiess, H. (1990). “A descent based approach for the OD matrix adjustment problem”. Publication n° 693 – Centre de Recherche sur les Transports. Université de Montréal, Canada.

Thompson, S.K. (1987). “Sample size for estimating multinomial proportions”. *The American Statistician*. Vol. 41, No. 1, February. American Statistical Association, U.S.A. pp. 42-46.

Transportation Research Board (2000). “Highway Capacity Manual”. *National Research Council*, Washington D.C.

Van Zuylen, H.J., Branston, D. M. (1982). “Consistent link flow estimation from counts”. *Transportation Research*, Vol. 16B, pp. 473 – 476.

Van Zuylen, H.J., Willumsen, L.G. (1980). “The most likely trip matrix estimated from traffic counts”. *Transportation Research*, Vol. 14B, pp. 281 – 293.

Willumsen, L.G. (1981). “Simplified transport models based on traffic counts”. *Transportation*, Vol. 10, pp.257-278.

Yang H., Iida Y., Sasaki T. (1991). “An analysis of the reliability of an origin-destination trip matrix estimated from traffic counts”. *Transportation Research*, Vol. 25B, pp. 351 – 363.

Yang H., Sasaki T., Iida Y., Asakura Y. (1992). “Estimation of origin-destination matrices from link traffic counts on congested networks”. *Transportation Research*, Vol. 26B, pp. 1 – 18

Yang H., Iida Y., Sasaki T. (1994). “The equilibrium-based origin-destination matrix estimation problem”. *Transportation Research*, Vol. 28B, pp. 23 – 33

Yang H. (1995). “Heuristic algorithms for the bilevel origin-destination matrix estimation problem”. *Transportation Research*, Vol. 29B, pp. 231 – 242

Yang, H., Zhou, J. (1998). “Optimal traffic counting locations for origin-destination matrix estimation”. *Transportation Research*, Vol. 32B, pp. 109 – 126