IMPROVING THE RELIABILITY OF DEMAND ESTIMATION USING TRAFFIC COUNTS BY INCLUDING INFORMATION ON LINK FLOW OBSERVABILITY

F. VITI[a], G. CANTELMO [a], F. CORMAN[b] and M. RINALDI[c]

[a]*Research Unit of Engineering Science, Faculty of Science, Technology and Communication, University of Luxembourg, Luxembourg, Email: francesco.viti@uni.lu*

[b]*Transport Engineering and Logistics Section, Maritime and Transport Technology Department, Delft University of Technology, Delft, the Netherlands*

[c]*Centre for Industrial Management / Traffic & Infrastructure, KU Leuven, Heverlee, Belgium*

ABSTRACT

Demand estimation problems are traditionally solved by an optimisation problem, where some distance between measured and simulated flows is minimised. To limit the effects of solution under-determinedness in such problems, a-priori assumptions on the OD matrix structure are adopted.

In this paper we study the impact of sensor locations on the quality of OD estimations. We show that linear correlations between link flow data may negatively affect the estimation reliability. By contrast, efficient sensor location models, able to identify sensor positions that collect linearly independent link flows allow improving the reliability of the estimation process, especially when no good prior information is available on the OD matrix structure.

In addition, we show that by selecting sensor locations according to partial observability metrics introduced by the authors in recent research, the contribution of the additional information in the objective function also improves with respect to alternative metrics introduced in the literature.

Keywords: OD Estimation, Partial Observability, Seed Matrix, Link Flow Correlations

1. INTRODUCTION

The importance of traffic models is widely acknowledged. They form a central component in a variety of traveller information systems and traffic management systems, and they are indispensable to estimate and predict traffic flows at (congested) networks. Such models require several inputs, e.g. road capacities, route choice proportions, etc. A critical component among these is the demand, or origin-destination (OD) matrix, which captures the spatial (and temporal) distribution of traffic demand. Errors in this matrix lead to errors in the representation of congestion and wrong predictions of the future evolution of the demand. For this reason, estimating OD-flows correctly is of paramount importance.

Traditional solution approaches aim at finding the most likely OD flows that reproduce the observed traffic condition which are observed on a subset of the links in the network. The most common information are the traffic counts, measured through counting stations (e.g. loop detectors). These OD flows are found by solving an optimisation problem, in which some distance between measured and simulated flows is minimised. This work deals mainly within this framework. Alternative approaches, which rely on information taken from other data sources (license plate recognition cameras, tolled gantries, floating car data, etc.), and offering other type of information (speeds, densities, route flows, travel times, vehicle trajectories, etc.) are out of our scope.

The non-unique mapping between OD flows and link flows causes typical solution under-determinedness, i.e., often, an infinite number of solutions may reproduce equally well the observed flows, even when sensors are installed on all links in a network (Cascetta, 2009). To partly limit the effects of solution under-determinedness, a-priori assumptions on the matrix structure or on the OD flows are adopted. For instance, the deviation between estimated OD flows and an available seed matrix can be derived by an outdated solution, or a travel survey, or by using any opportune initial starting point for exploring the solution space. This often is referred to as matrix correction or adjustment procedure, and its solution will directly depend on the given a-priori (or seed) matrix (Marzano et al. 2009).

Different importance can be given to prior information and observations. By contrast, data taken from the same sensor type is usually not differentiated, i.e. any observed link is assumed to share equal impact. This results in larger observed flows providing a larger impact on the search directions (i.e. the marginal change in the OD flows that causes the largest decrease in the goal function) when deviations between assigned and observed flows are taken in terms of absolute distances, while the opposite holds when relative errors are used. Hence, the efficiency of the solution will depend on the choice of the objective function. If weights are given, these are normally chosen pragmatically based on experience or on practical rules (e.g. using road hierarchy, or assumed importance in a network), and not on their relative contribution to the overall information in the network. In addition, in such formulations, it is common practice to consider observed link flows from each sensor as (linearly) independent, i.e. the component of the goal function related to the error between observed and estimated link flows is a simple (squared) sum of each residual (Cascetta, 2009). This is clearly not true in most of the cases, due to the topological and connectivity relations.

Driven by the above motivations, in this paper we study the impact of:

1) the size of the solution space of the objective function in OD estimation problems as function of the available information on the network, and its relation to OD estimation reliability[1];
2) the different information that each link count may contain[2], and what is the overall information on the network based on the number and position of the sensors, and

[1] The definition of reliability in this paper follows the one introduced in Yang et al. (1991).

[2] A formal definition of information in this paper is given in Viti et al. (2014) and will be repeated in Section 3.

3) the impact of link flow correlations in the optimisation problem, and the benefit of selecting sensors characterised by small correlations and being able to uniquely determine indirectly observed flows[3].

To analyse the above features, we explore the potentials of strategic sensor placement. Recent advancements in network topology analysis, in the context of flow estimation problems, have revealed the importance of exploiting the concept of flow observability. This relates to how much information on unmeasured flows can be determined by using information on the measured ones, and, in turn, which is the most efficient subset of network links to monitor in order to infer flow information to the rest of the network (see, e.g., Viti et al. 2008; Castillo, 2008; Hu et al., 2009; Gentili and Mirchandani, 2012).

In accordance with definitions introduced in Viti et al. (2014), in observability problems we talk about *observed* flows when these are directly measured by a sensor, while we make a distinction between *observable* (or *indirectly observed*) and *non-observable* or *partially observable* flows if they can be fully expressed by a linear combination of observed flows or not, respectively.

In this paper we show how both directly and indirectly observed flows can be used to systematically reduce the demand estimation error in a network. Information stemming from flow observability approaches can in fact be translated into how much each sensor's information contributes to the estimation quality, and this contribution is original (i.e. linearly independent) if compared to any other sensor available in the network. We also show that we can improve the estimation process by extending the objective function to also include unmonitored, but indirectly observed flows.

The contributions of this paper are therefore threefold:

1) We show that, while OD estimation quality depends on the number of unknown variables and the number of links counted, its reliability strongly depends on the positioning of the sensors. In this respect, this analysis extends early findings by Lam and Lo (1990) and Yang et al. (1991). In particular, we show that the reliability of OD flows can be measured by the partial observability metric introduced recently in Viti et al. (2014), given the same amount of sensors on the network;
2) Using the partial observability metric and the greedy algorithm introduced in Viti et al. (2014), we can identify partial and full flow observability solutions that contain a large set of observable unmonitored links, which therefore can be used as additional terms into the objective function of the demand estimation method;
3) We test the proposed methodology on a toy network as proof of concept, and on a mid-sized network using synthetic data to prove the increased accuracy of the OD estimation results when compared to a traditional OD estimation procedure based on traditional Generalised Least Squares estimation (Cascetta et al., 1984). Other estimators have been tested (e.g. Simultaneous Perturbation Stochastic Approximation, Finite Difference Stochastic Approximation) yet leading to analogous results.

This paper is structured as follows. Section 2 introduces the demand estimation problem and its inverse, the network sensor location problem, and gives a short review of the relevant literature. For a more extensive review of OD estimation problems based on traffic counts one can refer to e.g. Abrahamsson (1998) or Cascetta et al. (2013), while for sensor location problems we refer to Gentili and Mirchandani (2012) or Viti et al. (2012; 2014). Section 3 presents the strategy for choosing sensor location, based on the Null-Space (NSP) metric and the greedy algorithm developed in Viti et al. (2014), and casts this methodology into the demand estimation problem. In particular, we present a

[3] A formal definition of 'indirectly observed flows is given later in this paper and follows the one introduced by Gentili and Mirchandani (2012).

modified formulation for the objective function, which includes the deviation between estimated and observable variables, and show the influence of network topology on the solution space shape. Section 4 shows how the new methodology contributes to improving the OD estimation performances and the accuracy of the solution found with well-established optimisation techniques. Section 5 presents an analysis of the results on a test case based on the well-known Sioux Falls network. Finally, Section 6 provides conclusions, recommendations and future research directions.

2. LITERATURE REVIEW

We divide the literature review in works discussing OD estimation from traffic counts, and on sensor location problems, i.e. on which variables should be measured in order to describe all network flows (full observability) or only a subset of them (partial observability).

2.1 OD Estimation Based on Traffic Counts

Origin-destination matrix estimation problems use given information on traffic and/or travel data to estimate what is the most likely distribution of OD flows in a traffic network. These problems are usually seen as a reverse problem with respect to the traffic assignment problem (Cascetta, 2009), since the estimated OD flows are chosen such that, when assigned to the network, they simulate traffic flows that are as close as possible to the observed ones. Thus, a key issue in the estimation of the demand matrix from traffic counts is the identification of the origin-destination pairs whose trips use a particular link in which traffic is monitored.

OD estimation from traffic counts can be formulated as a constrained optimisation problem, where the error between observed and simulated flows is minimised, subject to equilibrium of traffic flows, which in its general form is the solution of a traffic assignment problem. A conventional OD estimation formulation (1) can be expressed as follows:

Objective Function: $$\mathbf{x}^* = \underset{\mathbf{x}}{\operatorname{argmin}}[z_1(\mathbf{d},\mathbf{x}) \cdot w_1 + z_2(\mathbf{f^o},\mathbf{f^s}) \cdot w_2] \tag{1a}$$

Subject to: $$\mathbf{f^s} = \mathbf{A}(\mathbf{c}) \cdot \mathbf{x} = \Delta \cdot \mathbf{B}(\mathbf{c}) \cdot \mathbf{x} \tag{1b}$$

Where:

- $\mathbf{d}, \mathbf{x}$ are respectively the a-priori (seed) and estimated demand matrices;
- $\mathbf{f^o}, \mathbf{f^s}$ are respectively the observed and simulated flows vectors;
- $\mathbf{A}(\mathbf{c})$ is the assignment matrix;
- Δ is the link-routes incidence matrix;
- $\mathbf{B}(\mathbf{c})$ is the routes-OD incidence matrix;
- z_1, z_2 are the estimators measuring the deviation between previous/estimated and observed/measured data ;
- w_1, w_2 are the weights for the error between previous/estimated and observed/measured data.

The main difficulty of estimating OD flows from traffic counts using eqn (1) is the under-specification of the problem, i.e. the existence of multiple solutions, which map a set of link counts to OD flows. To reduce the negative effects of the solution under-determinedness, a good deal of research proposed ways of reducing the solution space of eqn (1a) either by selecting number and location of sensors in an opportune way, as it will be discussed more in detail in the next sub-section, or by specifying the objective functions z_1, z_2 and the weights w_1, w_2opportunely. Taking additionally into account a term specifically tailored to demand, in the form of OD flows derived from an a-priori matrix, also leads to a reduction in the solution space size and under-determidness of the problem. Finally, the use of equation (1b) allows further reduction of the solution space, as solution is then constrained to the assigned OD flows.

Robillard (1975) proposed to constrain the solution of (1a) through a proportional assignment in (1b). In proportional assignment, the total assigned flow on a link equals the sum of the assigned flows obtained when the estimation method is applied to each OD pair separately. Proportional assignment was adopted in Maximum Entropy/Minimum Information approaches (Van Zuylen and Willumsen, 1980; Bell, 1983), anhd in statistical inference approaches such as Bayesian updating (Maher, 1983) and Maximum Likelihood techniques (Spiess, 1987). Cascetta (1984) proposed to use the Generalized Least Squared (GLS) estimation method to combine target trip matrix, model prediction and traffic counts within a single framework. GLS estimation was also extended by Bell (1991) and Bielaire and Toint (1994) to include non-negativity and flow conservation constraints, and to incorporate congestion effects.

In case of congested networks, drivers choose which routes to take in proportion to the current (and past) information of route travel times. If this effect is explicitly considered, the OD matrix estimation problem becomes more complicated as the assignment matrix $\mathbf{A}(\mathbf{c})$ becomes flow-dependent: since path costs are not constant anymore, the route choices will depend on the current route costs, which in turn depend on the OD matrix. Thus, the relationship between route flows and the OD matrix can only be implicitly defined. A few equilibrium assignment-based OD estimation approaches in literature have followed a bi-level structure to combine trip distributions with route choice distributions (Nguyen, 1977). LeBlanc and Fahrangian (1982) proposed to formulate the upper level as minimization of the Euclidean distance between simulated link flows, obtained by assigning the solution matrix, and the target matrix (a-priori OD flows), while Fisk (1988) extended the maximum entropy approach to include user optimal assignment. OD estimation formulated as bi-level problems have the inherent advantage of considering both target OD matrices (in the upper level) and the traffic counts (in the lower level), and to deal with the effects of congestion.

In OD estimation, solution accuracy relies on a good estimation of the gradient of the upper level objective function, which needs to be approximated, given the mutual interaction between upper level and lower level. In gradient-based solution techniques, the target OD matrix can be taken as an initial solution to the OD matrix estimation problem. The target OD matrix is therefore “adjusted” to reproduce the traffic counts by iteratively calculating directions based on the gradient of the objective function. The link volumes are implicit functions of OD flows and obtained through the assignment procedure, transforming the problem into a one-level problem. In the past, different solution methods using a gradient-based approach were proposed, both in the static (e.g. Yang et al., 1992) and in the dynamic case (e.g. Frederix et al., 2011).

The complexity of the estimation problem (1) has therefore warranted different simplifications, from the linearization of the objective function to the approximation of the response function represented by the equilibrium assignment. These simplifications affect the quality of the estimates, together with the choice of the a-priori information and of the location of traffic sensors.

Marzano et al. (2009) showed that, because of the structure of the OD estimation problem, the available solution procedures are generally unable to provide an effective correction of the OD matrix. This is because the number of (stochastic) equations, determined by the number of independent observed link flows is usually far lower than the number of unknowns (OD flows). In a more recent study it was found that the level of under-determinedness can be broadly assessed by simply measuring the ratio between the number of unknown variables (origin-destination pairs, times the number of time slices in the case of dynamic OD estimation problems) and the number of link counts (Cascetta et al, 2013). The study found that a minimum number of sensors should be allocated, which should be near the number of unknown OD flow variables. The study however did not explore the impact that different sensor locations can have on this relation, which is partly the aim of this paper.

2.2 Network Sensor Location Problems using Link Flow Information

Earlier studies on the reliability of OD matrix estimation highlighted the importance that sensor locations have on defining and possibly limiting an estimation error (Lam and Lo, 1990; Yang et al., 1991). In this stream of the literature, different metrics were introduced to quantify the error bound, given a specified set of sensors. Yang et al. (1991) proposed the Maximum Possible Relative Error

(MPRE), which represents the largest possible error that can be made when estimating OD flows, given number and position of traffic counts (assumed error free). Gan et al. (2005) extended the MPRE concept to account for the expectation value of the error, the Expected Relative Error (ERE), which was calculated using Monte-Carlo simulations. Conceptually different metrics were proposed by Chen et al. (2005), who applied the Total Demand Scale (TDS) metric of Bierlaire (2002), which is calculated as the difference between the maximum and minimum possible total OD demand estimates in a polyhedron constrained by traffic measurements. More recently, Zhou and List (2010) and Simonelli et al. (2012) introduced measures of OD matrix estimation variability, which are found by using the trace of the covariance matrix of the posterior OD estimate. These metrics were then integrated in the optimisation of sensor locations.

The problem of finding the optimal set of locations where to observe traffic flows is known as the Network Sensor Location Problem (NSLP). NSLP problems can be distinguished between those focusing on the algebraic and topologic properties of the network structure and connections (*observability problems*), and those relating observed traffic states (usually, flows) with the ones derived using estimation techniques (*flow-estimation problems*).

2.2.1 Observability Problems

Sensor location problems based on flow observability exploit the relations between flows, represented by the assignment matrix $\mathbf{A}(\mathbf{c})$. As explained above, this matrix can relate link, route and OD flows through conservation of vehicle equations:

$$\sum_{r\in\mathcal{R}} \delta_{lr} h_r = v_l \qquad \forall l \in \mathcal{L} \tag{2a}$$

$$\sum_{r\in\mathcal{R}} \rho_{wr} h_r = x_w \qquad \forall w \in \mathcal{W} \tag{2b}$$

with

- x_{w} the OD flow for an origin-destination pair w in the set of OD pairs $\mathcal{W}$;
- h_r the route flow on a route r belonging to the routes set $\mathcal{R}$;
- v_l the link flow on link l belonging to the link set $\mathcal{L}$.

The matrices $\boldsymbol{\delta}$ and $\boldsymbol{\rho}$, consist of elements δ_{lr} and ρ_{wr}, having values 0 or 1, depending on whether route r contains or not link l, and whether route r connects OD-pair w, respectively.

The majority of the literature on observability problems focuses on finding efficient full observability solutions, i.e. find the minimum sets of observed link (or route, OD) flows in order to make the system of equations describing all traffic states uniquely determined, i.e. such that all unobserved flows become observable. Through opportune matrix operations, it is possible to identify a subset of (link, route, OD) flows that is able to describe linearly all other flows (Castillo et al., 2008; Castillo et al., 2009). Alternative approaches can provide efficient solutions without a priori knowledge of the routes by using information at nodes (Ng, 2012; He, 2013), or by using the concept of holes and non-planar networks (Castillo et al., 2014).

One of the main advantages of observability solutions is that the resulting observed variables are linearly independent. This is an agreeable property in the context of OD estimation, since the link flow terms in eqn (1a) would be linearly independent. Thus, it is possible to have a lower degree of under-determinedness compared to a solution with the same number of sensors, but where some link flow is linearly dependent on other observed flows.

Recently, the concept of partial observability was introduced to identify the subset of non-measured flows that are on the other hand indirectly observed. Gentili and Mirchandani (2012) provided an intuitive definition of partial observability problems as to “find the (minimum) subset of state variables such that the system is partially observable at level h”. In this definition, the number of indirectly observed variables defines the level h. An analogous definition was proposed by Castillo et al. (2012), where, those h links that, at least, should be observable are specified a priori. This means that the problem is reformulated as to "find the minimum number of observed links such that a pre-selected set of h link flows becomes observable". This means that any solution of this problem would be at least of level h according to Gentili and Mirchandani (2012). Another intuitive definition is to define solution of partial observability, solving the problem of "determining the optimal locations for given number of sensors" (see e.g. Hu et al., 2009; Ng, 2012; and He, 2013). Weights or priorities must in this case be pre-specified to indicate the different importance that some links can have with respect to the others (e.g., motorways that are likely to carry larger flows).

All of the above definitions, and related efficient methods for minimising the number of observed variables that guarantees a partial observability level h, do not consider any information that solutions may still provide on the non-fully observable variables. Viti et al. (2014) introduced a new metric that considers the ensemble of information from both observable and non-observable link flows. This is done by exploring the Null Space of the link-route matrix, which describes the degrees of freedom resulting from solution under-determinedness. This metric is employed in a greedy algorithm to identify an efficient sequence of link flows to measure in order to maximise the information on the network. This algorithm will be adopted in this paper and it will therefore be summarised in Section 3.

2.2.2 Flow-Estimation Problems

Observability problems consider networks purely from a topological perspective, i.e. they require only information on network connectivity and (eventually) an explicit enumeration of routes. Flow-estimation problems use instead extra information or assumption on the flows, i.e. they assume the availability of a prior OD Matrix (e.g., Yang and Zhou, 1998), or prior knowledge and reliability of the route flows (e.g., Wang et al., 2012), or known turning proportions at nodes (e.g. Bianco et al., 2001), or information on link usage proportions (e.g. Gan et al., 2005), or on shortest routes (Yang et al., 2006).

In the absence of prior information, flow-estimation problems use principles analogous to observability problems as they also exploit the relations (2) to limit the search space for an efficient solution. To do so, a number of intuitive rules are introduced in relation to the type of estimator used in (1).Yang and Zhou (1998) exploited the MPRE earlier introduced in Yang et al. (1991) and starting from its definition they proposed four rules, namely:

1) *OD-coverage*, i.e. there must be at least one sensor installed on any of the routes connecting each OD pair,
2) *Maximum flow fraction*, i.e. the portion of OD flow measured by a sensor w.r.t. all other OD pairs measured by the same sensor is maximized,
3) *Maximum flow intercepting*, i.e. sensors should be positioned to maximise the total flow captured, and
4) *Link independence*, i.e. locations should be chosen such that information extracted is linearly independent.

Analogous rules were introduced by Larsson and Peterson (2010), Cipriani et al. (2006) and Yang et al. (2006) to account for route flow information independence and maximum total flow captured.

In the attempt to capture some of the above rules, different solution algorithms have been proposed. Yang and Zhou (1998) proposed an efficient greedy algorithm seeking to minimize the MPRE metric. Both OD coverage rule and maximum flow fraction rules are considered in this procedure to guarantee that the solution space is compact and the MPRE metric takes only finite values. Yim and Lam (1998) proposed to specify a weight to the link-OD proportions using the results of a traffic assignment model and a prior OD matrix, and formulated a linear programming model to maximize both net and total captured flows. Bianco et al. (2001) developed a two-stage procedure that aims at determining sensor locations that best estimate turning fractions. The objective is to maximize the OD coverage and reduce the estimation error on the turning flows. A similar approach was proposed by Elhert et al. (2006) for finding a second-best solution, given a set of existing sensors preinstalled on the network.

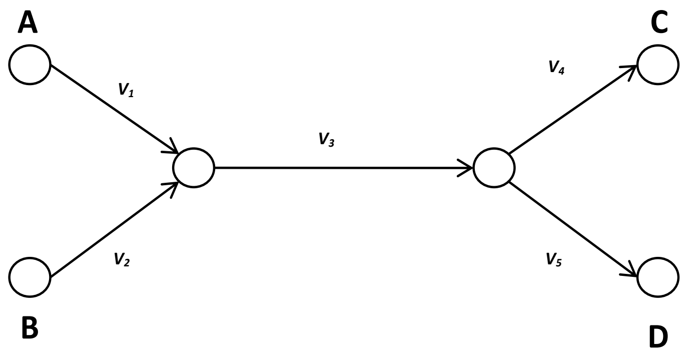
3. METHODOLOGY

3.1 The Impact of Link Flow Correlations

Equation (1) assumes that a certain number of sensors are installed in the road network; these sensors are responsible for the quality and reliability of the second term of eqn. (1a). The first term, $z_1(\mathbf{d}, \mathbf{x}) \cdot w_1$, assumes that an a-priori, or seed, matrix is available. If such a matrix is not available, or any (whatsoever) information on the OD flows is used just as a starting point for exploring the solution space, it might be more relevant to drop the first term $z_1(\mathbf{d}, \mathbf{x}) \cdot w_1$ in (1a), e.g. by assuming a weight $w_1$value of zero. Since this paper focuses on the contributions provided by the link flows in the overall OD optimisation problem, this first term will indeed be omitted.

In equation (1a) the value of the OD flow matrix $\mathbf{x}$ is responsible of the simulated flow values $\mathbf{f^s}$, on all links, calculated using equation (1b). However, only those flows corresponding to measured links appear in the objective function (1a). For example, z_2 can be specified in terms of deviation between observed (measured) and the corresponding simulated link flows (e.g. using some kind of metric such as Euclidean distances, or squared distances). This means that eqn (1a) consists of additive terms, counting as many as the number of sensors. Hence, the objective function as expressed in (1a) does not exploit any information on all other link flows that are not explicitly measured.

To quantify the impact of sensor positioning in OD estimation, we propose an illustrative example hereafter based on classical toy network shown in Figure 1. The network has 5 links, 4 OD pairs {(A, C), (A, D), (B, C), (B, D)} and 4 routes, which are specified in the attached table. Due to the structure of the network, each OD pair corresponds to a single route, and vice versa; estimating route flows (or fractions) is equivalent to estimating OD flows. In this example, even if all links were equipped with flow sensors, route flows would not be uniquely determined, because the rank of the link-route matrix $\boldsymbol{\delta}$ is 3 rather than 4 (see Viti et al., 2014 for more details), yielding an under-determined problem. This means that three links at best contain linearly independent flow information and therefore make the other two unobserved flows fully observable (e.g. link flows v_1, v_2 and $v_4$uniquely determine the flows in v_3 and v_5). It is therefore clear that if the three link flows to be measured would be selected arbitrarily, there would be a chance that information from one is linearly dependent from the others. Whilst flows v_1, v_2 and v_4 determine all other flows in the network, choosing a different subset of three links, such as v_1, v_2 and v_3 would result in an information from the traffic counts which is not of the highest quality, since v_3 can be expressed as the simple sum of v_1 and v_2 and therefore does not contribute to increase the overall information about the OD flows, neither of the other link flows.



Routes / OD	links
h_1	$\{v_1, v_3, v_4\}$
h_2	$\{v_2, v_3, v_4\}$
h_3	$\{v_1, v_3, v_5\}$
h_4	$\{v_2, v_3, v_5\}$

Figure 1: Simple network example

Looking for optimal sensor positions so that measured link flows are linearly independent wrt all other measured flows is one of the main motivations that have sparked research in the area of network sensor location problems, as already discussed in Section 2.2. By contrast, in the following of this paper we also show that discarding completely the potential of linearly dependent information, and in particular neglecting the contribution of indirectly observed flows, may also be detrimental in procedures such as OD estimation. The reason is that indirectly observed flows, i.e. a linear combination of available link flows may still contain original information about the OD flows. This extra information may help at having a better knowledge of the sensitivity of link flows to changes in the OD flows. We show this by means of a numerical example in the following sub-section.

3.1.1 Illustrative Example

We use again the network of Figure 1 to illustrate the impact of information correlation in the sensors data. As shown in Section 3.1, this network is a typical case where a unique solution for the OD estimation problem does not exist, even when measuring all link flows.

Figure 2 provides the input information used for this illustrative example.

A-priori Matrix $\begin{pmatrix} AC \\ AD \\ BC \\ BD \end{pmatrix} = \begin{pmatrix} 0 \\ 75 \\ 50 \\ 50 \end{pmatrix}$ *Simulated Link Flows* $\begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \end{pmatrix} = \begin{pmatrix} 75 \\ 100 \\ 175 \\ 50 \\ 125 \end{pmatrix}$

Real Matrix $\begin{pmatrix} AC \\ AD \\ BC \\ BD \end{pmatrix} = \begin{pmatrix} 50 \\ 50 \\ 50 \\ 50 \end{pmatrix}$ *Real Link Flows* $\begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \end{pmatrix} = \begin{pmatrix} 100 \\ 100 \\ 200 \\ 100 \\ 100 \end{pmatrix}$

Figure 2: Input data used for generating the illustrative example

Matrix 1

$$\begin{pmatrix} AC \\ AD \\ BC \\ BD \end{pmatrix} = \begin{pmatrix} 25 \\ 75 \\ 50 \\ 50 \end{pmatrix}$$

Simulated Flows

$$\begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \end{pmatrix} = \begin{pmatrix} 100 \\ 100 \\ 200 \\ 75 \\ 125 \end{pmatrix}$$

Error in Flows

$$\begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -25 \\ 25 \end{pmatrix}$$

Matrix 2

$$\begin{pmatrix} AC \\ AD \\ BC \\ BD \end{pmatrix} = \begin{pmatrix} 0 \\ 100 \\ 50 \\ 50 \end{pmatrix}$$

Simulated Flows

$$\begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \end{pmatrix} = \begin{pmatrix} 100 \\ 100 \\ 200 \\ 50 \\ 150 \end{pmatrix}$$

Error in Flows

$$\begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -50 \\ 50 \end{pmatrix}$$

Matrix 3

$$\begin{pmatrix} AC \\ AD \\ BC \\ BD \end{pmatrix} = \begin{pmatrix} 0 \\ 75 \\ 75 \\ 50 \end{pmatrix}$$

Simulated Flows

$$\begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \end{pmatrix} = \begin{pmatrix} 75 \\ 125 \\ 200 \\ 75 \\ 125 \end{pmatrix}$$

Error in Flows

$$\begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \end{pmatrix} = \begin{pmatrix} -25 \\ 25 \\ 0 \\ -25 \\ 25 \end{pmatrix}$$

Matrix 4

$$\begin{pmatrix} AC \\ AD \\ BC \\ BD \end{pmatrix} = \begin{pmatrix} 0 \\ 75 \\ 50 \\ 75 \end{pmatrix}$$

Simulated Flows

$$\begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \end{pmatrix} = \begin{pmatrix} 75 \\ 125 \\ 200 \\ 50 \\ 150 \end{pmatrix}$$

Error in Flows

$$\begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \end{pmatrix} = \begin{pmatrix} -25 \\ 25 \\ 0 \\ -50 \\ 50 \end{pmatrix}$$

Figure 3: different starting matrices, respective simulated link flows, and corresponding error in the simulated flows wrt real ones.

Let us assume an a-priori matrix, reported as the vector of the 4 OD pairs considered, as in Figure 2. The simulated links flows, which the a-priori matrix generates, are also displayed. In the same figure we also report the real OD matrix, and the real link flows, i.e. our target values. Note that for simplicity's sake the error in the seed matrix is, in this example, only related to the origin zone A, namely OD flows AC is 0 instead of 50, and AD is 75 instead of 50. The prior estimates of the OD flows from origin B are assumed to be correct. The prior OD flows are hence characterized by an error in the OD that will result in (50, 25, 0, 0) for each OD flow, respectively.

We then calculate the link flow changes when the prior OD flows are changed of 25 units. These are presented as four possible OD solutions in Figure 3 (top matrices), which are simply built by adding 25 units respectively on each component of the a-priori matrix. This step is intended to illustrate a search direction starting from the a-priori matrix with an arbitrary step size. Simulated link flows are then calculated accordingly (Figure 3, center row) for the four cases.

We now study all possible subsets of measured links. Table 1 reports for each subset of sensors (first column): the vector Γ in the table, containing only {0, 1} values, indicates which link flows are observed for the calculation of the objective function. The second column reports the *h* value, i.e. how many non-observed variables can be determined completely by looking at the observed ones, i.e. they are indirectly observed. The other 4 columns report the error (in term of Absolute Error) between the link flows corresponding to each of the 4 solution matrices reported in Figure 3 (in the same order). The error is calculated in terms of Absolute Error:

$$z_2(\mathbf{f^o}, \mathbf{f^s}(\mathbf{x})) = \sum_{l \in \tilde{L}} |\mathrm{f}_l{}^{\mathrm{s}}(\mathrm{x}) - \mathrm{f}_l{}^{\mathrm{o}}| \qquad (3)$$

As each of the 4 matrices is a finite variation wrt the seed matrix, the values reported in Table 1 represent the sensitivity (in terms of link flows AE) of the objective function to a finite movement in the solution space. This is intended to simulate one iteration of a gradient-based approach.

Starting from the last row, for the solution with full observability, i.e. the full set of 5 links, the variation in the observed flows corresponds to full extent of the deviation between the simulated and the real link flows, i.e. the estimation of the descent direction is correct, since the lowest error is observed for the *Matrix 1*, which means that the gradient information is correctly influencing the computed matrix towards the real one. When using a lower number of sensors, the estimation of the direction may significantly deviate from the correct one. Looking at the table, in some conditions two detectors are enough to determine the correct gradient, in other cases using three detectors out of five there is still error in estimating the variation of the objective function. More specifically, when only the detector v_3 is used (third row), it is impossible to determine a descent direction, since the AE for each of the 4 matrices is zero. This corresponds to the impossibility for any method to determine a descent direction for the objective function by imposing a finite variation to the seed matrix. The optimization problem then returns the a-priori matrix as final solution. The same applies if the initial matrix chosen is either 1 or 2, and the triple (v_1, v_2, v_3) is used as set of observed variables: since no error is observed on the direction of OD pairs AC and AD, in the next iteration the OD flow on both will remain the same, so the estimation will not get closer to the correct values and on the contrary will tend to move the correct ones (BC and BD).

This may suggest that if detectors with strongly aggregated measurements (i.e. v_3) are used, the sensitivity of objective function may be incorrectly estimated. Looking in fact at the case with 4 detectors, if vector v_3 is used, the objective function underestimates the total deviation for matrices 3 and 4 wrt the case where v_3 is not used. This is also in line with expectations since v_3 was shown in Section 2 to be formulated as linearly dependent on v_1 and v_2, for example. This is one possible relation to identify link flow v_3, which would then be defined as a dependent variable with respect to observed data v_1 and v_2, which are, in these settings, independent variables.

Exploiting this concept, it is possible to define efficient and not-efficient detector samples. For instance the solution Γ ={1, 1, 0, 0 0} is efficient for a 2-detector case, since we can easily derive the flow on v_3 as $v_1 + v_2$, and thus obtain the same results as for {1, 1, 1, 0, 0}.

Table 1. Absolute Error for all possible subsets of observed flows

	h	*Matrix 1*	*Matrix 2*	*Matrix 3*	*Matrix 4*
k=1 Detector					
Γ ={1, 0, 0, 0, 0}	0	0	0	25	25
Γ ={0, 1, 0, 0, 0}	0	0	0	25	25
Γ ={0, 0, 1, 0, 0}	0	0	0	0	0
Γ ={0, 0, 0, 1, 0}	0	25	50	25	50
Γ ={0, 0, 0, 0, 1}	0	25	50	25	50
k=2 Detectors					
Γ ={1, 1, 0, 0, 0}	1	0	0	50	50
Γ ={1, 0, 1, 0, 0}	1	0	0	25	25
Γ ={1, 0, 0, 1, 0}	0	25	50	50	75
Γ ={1, 0, 0, 0, 1}	0	25	50	50	75
Γ ={0, 1, 1, 0, 0}	1	0	0	25	25
Γ ={0, 1, 0, 1, 0}	0	25	50	50	75
Γ ={0, 1, 0, 0, 1}	0	25	50	50	75
Γ ={0, 0, 1, 1, 0}	1	25	50	25	50
Γ ={0, 0, 1, 0, 1}	1	25	50	25	50
Γ ={0, 0, 0, 1, 1}	1	50	100	50	100
k=3 Detectors					
Γ ={1, 1, 1, 0, 0}	1	0	0	50	50
Γ ={1, 1, 0, 1, 0}	**2**	**25**	**50**	**75**	**100**
Γ ={1, 0, 1, 1, 0}	**2**	**25**	**50**	**50**	**75**
Γ ={0, 1, 1, 1, 0}	**2**	**25**	**50**	**50**	**75**
Γ ={1, 1, 0, 0, 1}	**2**	**25**	**50**	**75**	**100**
Γ ={1, 0, 1, 0, 1}	**2**	**25**	**50**	**50**	**75**
Γ ={0, 1, 1, 0, 1}	**2**	**25**	**50**	**50**	**75**
Γ ={1, 0, 0, 1, 1}	**2**	**50**	**100**	**75**	**125**
Γ ={0, 1, 0, 1, 1}	**2**	**50**	**100**	**75**	**125**
Γ ={0, 0, 1, 1, 1}	1	50	100	50	100
k=4 Detectors					
Γ ={0, 1, 1, 1, 1}	**1**	**50**	**100**	**75**	**125**
Γ ={1, 0, 1, 1, 1}	**1**	**50**	**100**	**75**	**125**
Γ ={1, 1, 0, 1, 1}	**1**	**50**	**100**	**100**	**150**
Γ ={1, 1, 1, 0, 1}	**1**	**25**	**50**	**75**	**100**
Γ ={1, 1, 1, 1, 0}	**1**	**25**	**50**	**75**	**100**
k=5 Detectors					
Γ ={1, 1, 1, 1, 1}	**0**	**50**	**100**	**100**	**150**

It is finally interesting to observe that only the case with Γ ={1, 1, 0, 1, 1}, i.e. the flow v_3 is not observed, the gradient calculation returns exactly the same results as with all 5 links observed. This means that for the calculation of the gradient, and only in this case, there would be no gain in adding information on the indirectly observed flow v_3 while in all other cases it would allow calculating the correct total extent of the AE, i.e. the one with all 5 link flows measured.

Hence, we modify equation (1a) to include extra terms related to the indirectly observed link flows. These will then appear as additional terms. By also omitting the distance from the a-priori matrix as mentioned above, our OD estimation formulation (4) reads as follows:

Objective Function: $$\mathbf{x}^* = \underset{\mathbf{x}}{\operatorname{argmin}}\left[z_2(\mathbf{f^o}, \mathbf{f^s}(\mathbf{x}))w_2 + z_3\left(\mathbf{f^d}(\mathbf{f^o}), \mathbf{f^s}(\mathbf{x})\right)w_3\right] \tag{4a}$$

Subject to: $$\mathbf{f^s} = \mathbf{A}(\mathbf{c})\mathbf{x} = \Delta\mathbf{B}(\mathbf{c})\mathbf{x} \tag{4b}$$

where we denote with $\mathbf{f^d}$ the set of indirectly observable flows. Note that we only consider those links whose flow can be uniquely determined by algebraic relations of other link flows, i.e. derived from equations (2); link flows that have still some limited degree of freedom are not included. Adding this last possibility may bring additional issues, as information on those links is not complete and therefore is less reliable. This problem will be tackled in future research.

In the following we make use of the insight acquired with the illustrative example to study how to identify efficient sensor locations in terms of estimation of its gradient.

3.2 Exploiting the Concept of Partial Observability in the OD Estimation Process

The observability problem described in Section 2.2.1 can be exploited to identify linear relations between measured link flows and non-measured ones. We here adopt the approach presented by Castillo et al. (2008) to determine all those relations. We will refer to this approach as the pivoting procedure. The procedure determines all the relations between a subset of (observed) sensors and all other variables; in other terms, it determines a subset of sensor locations which allows to fully observe the network as all other link flows become indirectly observed.

Figure 4 provides an example of a pivoting procedure calculated again using the example in Figure 1.

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{pmatrix}$$

(a) Initial link-route relations: $\boldsymbol{v} = \boldsymbol{\delta h}$

$$\begin{pmatrix} h_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{v_1} \\ h_2 \\ h_3 \\ h_4 \end{pmatrix}$$

(b) First iteration: swapping v_1 with h_1

$$\begin{pmatrix} h_1 \\ h_2 \\ \boldsymbol{v_3} \\ v_4 \\ v_5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ \mathbf{1} & \mathbf{1} & 0 & 0 \\ 1 & 1 & -1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{v_1} \\ \boldsymbol{v_2} \\ h_3 \\ h_4 \end{pmatrix}$$

(c) Second iteration: swapping v_2 with h_2

$$\begin{pmatrix} h_1 \\ h_2 \\ \boldsymbol{v_3} \\ h_3 \\ \boldsymbol{v_5} \end{pmatrix} = \begin{pmatrix} 0 & -1 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ \mathbf{1} & \mathbf{1} & \mathbf{0} & 0 \\ 1 & 1 & -1 & -1 \\ \mathbf{1} & \mathbf{1} & \mathbf{-1} & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{v_1} \\ \boldsymbol{v_2} \\ \boldsymbol{v_4} \\ h_4 \end{pmatrix}$$

(d) Last iteration: swapping v_4 with h_3

Figure 4: Initial matrix and three iterations of the pivoting procedure on the network in Fig. 1 (illustration reprinted from Viti et al., 2014)

The pivoting procedure simply operates through performing opportune matrix manipulations, i.e. by swapping or sorting rows or columns, while carefully maintaining the original relations between the variables as in the original set of relations (2). One of the main advantages of the pivoting procedure is that in any pivoting solution, the observed variables are by construction linearly independent, which is the first property we aim for in our methodology. In Figure 4(a) an initial specification of eqn (2a) is provided, together with three iterations of the pivoting process as described by Castillo et al. (2008). Figure 4(b) shows, for instance, that variables h_1 and v_1 can be swapped, which means that also the other link flow variables on the left side can be related directly to v_1. We mark v_1 in bold in this case to indicate that it is selected as an observed variable (i.e. we decide to install a sensor on link 1).

In the process of finding full observability solutions, some variables may become indirectly observed before completing the pivoting process. In fact, in the second iteration shown in Figure 4(b), we swap v_2 with h_2. By doing so, v_3 can be described as linear combination of the observed variables v_1 and v_2. Hence, link 3 becomes already indirectly observed by simply observing links 1 and 2; we underline this by denoting v_3 in red bold. Therefore this partial observability solution is of level h=1 using the definition of Gentili and Mirchandani (2012). In the estimation process, we can add h indirectly observable variables in equation (3a) in a partial observability solution of level h. Full observability is finally reached by also observing v_4, as v_5 becomes then indirectly observed as well (hence h=2), also shown in red bold in Figure 4(d). Thus the full observability solution is reached in this case (as was also indicated in Table 1 in bold): v_3 and v_5 are fully described by the other three links. The triple of links in the right-hand side of the Figure 4(d) (i.e. v_1, v_2, v_4) are the observed variables of the pivoting solution, while v_3 and v_5 are the observable flows. For further details on the pivoting procedure, one can refer to e.g. Castillo (2008). It is easy to observe that observability solutions are not unique, as well as there are many combinations leading to a partial observability level h, as one can see looking back at Table 1.

Therefore, our interest is in finding a metric able to differentiate the different partial and full observability solutions, such as the estimation of the gradient in equation (4) gets as close as possible to the one obtained by measuring all flows in the network. We study in the following section the opportunities offered by the recent Null-Space (NSP) metric introduced in Viti et al. (2014).

3.3 The NSP Metric and a Greedy Heuristic for Selecting Sensor Locations

Multiple pivoting can be performed given an initial expression (2a), depending on the order in which link variables are swapped with route variables; this implies that different partial as well as full observability solutions can be found, depending on the sequence in which rows and columns are swapped. In addition, different solutions can be reached by specifying a different route set (this is analysed more extensively and demonstrated in Rinaldi et al, 2015). To make a distinction between these pivoting solutions, in Viti et al. (2014) a metric was proposed that enables one to rank full observability solutions based on a limited number of "families". These families depend on the total amount of information contained by the different full observability solutions, in terms of quantity and quality of independence-dependence relationships. The term "information" is defined in Viti et al. (2014), and is adopted in this paper as a measure related to the extent of the solution space.

Moreover, full observability solutions are rather impractical in real-sized networks. Several studies report that about 60 to 80% of the links should be equipped with sensors in order to obtain full observability (Castillo, 2012; Ng, 2012; Viti and Corman, 2012; He, 2014). For this reason Viti et al. (2014) introduced a new metric, which allows one to assess the quality of a partial observability solution (in relation to a full observability solution) if only a subset of the links of a full observability solution is taken. It does so by quantifying the maximum information loss on the unobserved link flows in relation to a full observability solution containing the same subset of flows. This error thus ranges between 0 (full observability) and 1 (no sensor in the network) and is related to the Null space of the pivoted matrix, i.e. one obtained by using the pivoting process by Castillo et al. (2008).

In more technical terms, the Null space describes the degrees of freedom resulting from under-determinedness of the subspace of $\mathbb{R}^l$of solutions that are possible given the pivoted matrix, i.e., in case of full observability, there are no degrees of freedom left to the link flows. For all other cases, the extent of the solution space is clearly an uncertainty measure for the choice of not measuring one or more independent variables. The goal of the metric is to assign a single scalar value to vector space defined by any possible subset of the measured variables in the Null space. For a detailed formulation, description and graphical interpretation of the NSP metric we refer to Viti et al. (2014).

In the same paper, greedy algorithms were proposed to determine partial observability solutions (i.e. some set of link flows to be measured). The *Add* heuristic starts from an empty set of observed variables, and adds, as variable (link) to be observed, (the) one that results in the largest increase of information. Iteratively, the chosen variable is then added in the set of observed variables. Alternatively, a *Remove* heuristic removes variables from the full observability solution in order of least information contents.

In this paper we therefore try to tackle our problem (4) by exploiting the partial observability information, as obtained by applying the aforementioned heuristic algorithms on the different case study networks.

This methodology was tested on different toy networks and a real-sized network; we compare the quality of OD estimation solutions when dealing with either randomly chosen sensor locations or those yielded by our partial observability metric, and show how even few links might contain a large deal of information, and therefore reduce estimation uncertainty.

It seems natural to think that the properties found by the introduced metric, and the heuristic algorithms, are rather desirable in light of the extended OD estimation formulation (4). Employing a locating strategy that tends to reduce the solution space, while, at the same time, providing linearly independent information contributing to find *h* observable other flows is expected to improve the quality and the reliability of the OD estimates.

4. CASE STUDIES

4.1 The Sioux Falls Network

We test our methodology on the Sioux Falls network, shown in Figure 5. This classical toy network consists of 76 links and 24 nodes. We assumed 30 OD pairs to have non-zero flows and for each OD pair we used a k-shortest path (k=5) algorithm to generate the path flows and construct the relations (2). The number of routes was sufficient to obtain a rather homogeneous distribution of the OD flows in the network. Link cost functions have been chosen using the classical BPR function (BPR, 1964) with α=0.5 and β=4. Stochastic route choice behaviour has been modeled following a Multinomial Logit model (Ben Akiva and Lerman, 1985).

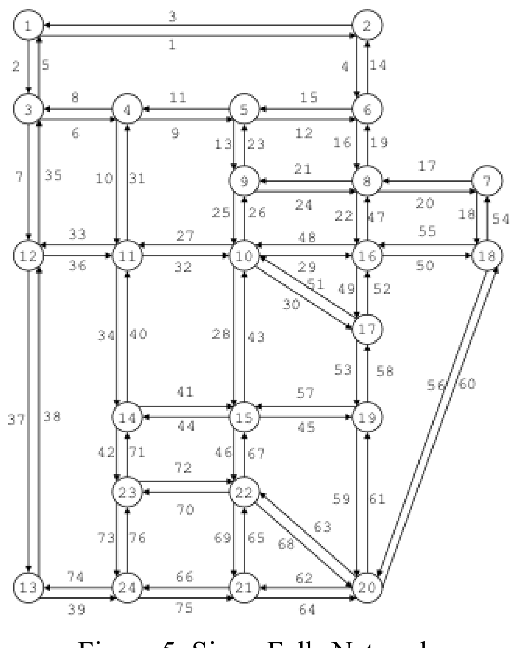


Figure 5: Sioux Falls Network

Five seed matrices are considered in this study. Each matrix is derived by imposing a deterministic or a random perturbation on a reference matrix, with which we have computed the observed link flows. Each chosen seed matrix presents a different error:

- Matrix 1: Deterministic Perturbation
- Matrix 2-3: Random perturbation, OD flows=integer number (i.e. 90);
- Matrix 4-5: Random perturbation, OD flows= not integer number (i.e. 68.89);

The OD estimation problem compares the solution based on sensors obtained through solving the Sensor Location Problem with a set of randomly chosen sensor locations.

In total 75 sets of sensors locations and 15 sets of sensor locations generated from the greedy heuristic procedure with an incremental amount of sensors are considered in this work. This leads to a total of 90 different sensor locations with an amount of sensors ranging from 0 to full observability, which are analysed in this case study, for their impact to OD estimation. The values reported in the remainder of the paper (under the label "random sets") refer to the average over the random sets. For the sake of clarity, we will refer to the random generated sets as "random" solutions, while the one derived from the pivoting as "efficient" solution.

The OD estimation problem is then solved using the following traditional approach:

1) The Generalised Least Squares (GLS) method was used as estimator;
2) The Objective Function (OF) used to obtain these results was the Mean Absolute Error (MAE). To focus on the contribution of our methodology we used only the link flows terms and omitted the term related to the deviation between an a-priori matrix and the estimated one.
3) The Assignment Method was a stochastic Method of Successive Averages (MSA).

One can refer to Cascetta (2009) for the details of these three specifications.

4.1.1 Analysis of Results

We analyse the results of the different OD estimations if only the set of sensors is used to calculate the deviations, i.e. equation (1), and we compare it with the same calculation made including the term concerning the indirectly observed flows, i.e. equation (4).

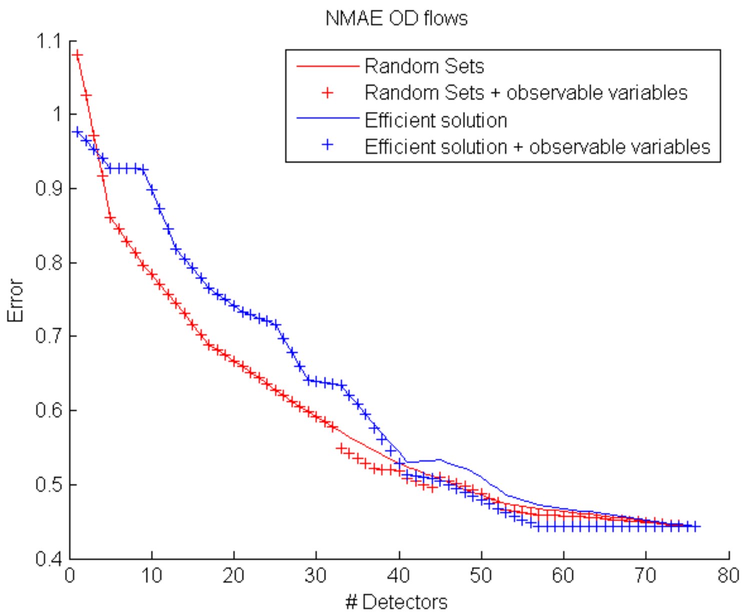


Figure 6: NMAE average values for a growing number of observed flows and between estimated and true OD matrices using eqns (4)

Figure 6 shows the Normalised Mean Absolute Error (NMAE) with and without considering the extra term related to indirectly observed flows. The lines represent the solutions of problem (1), while the crosses indicate the solutions corresponding to problem (4). The effect of including indirectly observed variables becomes visible for more than 30 randomly chosen sensors (red curves) and, respectively, for more than 40 sensors selected following the greedy heuristic (blue curves). The gain to be found in terms of mean error reduction is not too evident. Only a small reduction of the estimation error is achieved, and this appears to be more prominent for the efficient solution, although for a higher number of sensors. Indeed, the efficient solution is not capable of performing better than the randomly chosen cases before reaching 40 sensors.

Despite the importance of reducing the average estimation error, the difference between the solutions is not significant. Hence, the solution using the greedy heuristic does not improve the overall estimation accuracy. The opposite does not hold either. This is to be expected since the problem we are studying represents a static assignment problem with a convex solution space. So in this case having a better estimation of the OF gradient does not effectively have an impact in the long run.

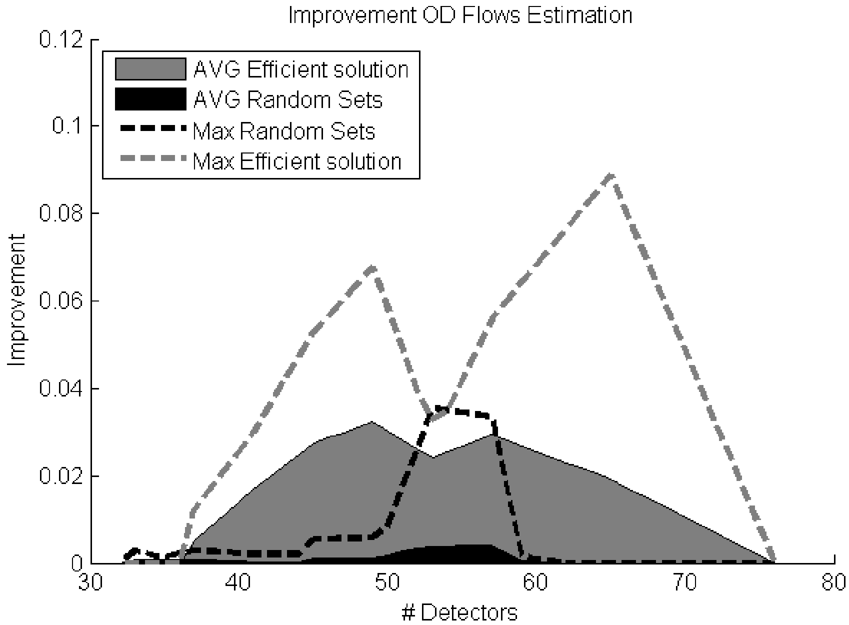


Figure 7: % improvement in the NMAE values for a growing number of observed flows and between estimated and true OD matrices using eqns (4)

Figure 7 shows how much link flow error is captured by the efficient solution wrt the random cases. The grey and black-shaded areas indicate the average gain, while the dashed lines indicate the maximum gain. The values on the y-axis mean that the sensitivity of the OF using the detector positions is significantly better estimated if a higher improvement is achieved.

As one can see, the gain becomes significant when more than 37 sensors are installed, while it reduces after 65 sensors and becomes zero when all link flows are observed. The reason for these results is more easily understandable once the partial observability value of h is shown, as in Figure 8. The greedy algorithm for selecting the efficient locations starts by looking for locations that contain the largest deal of information, i.e. it gives priority to those links that contribute the most to reduce the level of under-determinedness in problem (4). After 40 sensors, the greedy algorithm quickly looks for ways to get observable variables, until the full observability solution is reached, i.e. at 56 sensors (Figure 8). Clearly there is no gain to be found after that point, since the information offered by link sensors is at its maximum.

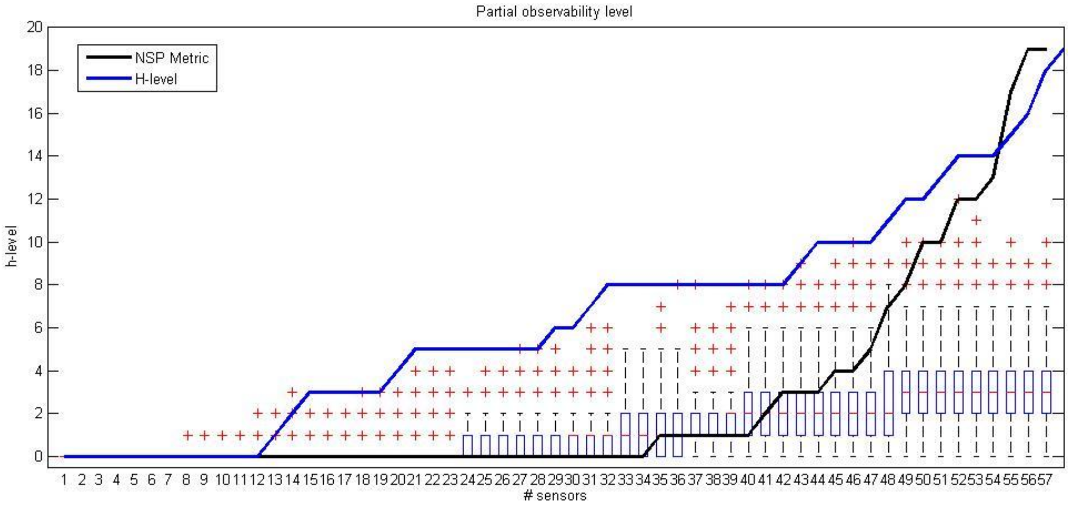


Figure 8: Partial observability level h using the definition of Gentili and Mirchandani (2012)

One could question whether directly looking at the value of h instead of using a more elaborated metric such as NSP would further improve the results. Figure 8 shows however the growing h-level using the greedy algorithm (black line) in comparison with a method that would look for sensor positions that directly maximize h (blue line). A box-plot represents, instead, the results of the randomly chosen sensor locations. Despite the gain in terms of h values is evident, as h starts being a positive number already after 12 sensors placed (i.e. a first indirectly observable variable is added), results in the OD estimation are less efficient in terms of NMAE (Figure 9), where the h-level case exhibits a higher error than the other sets of links.

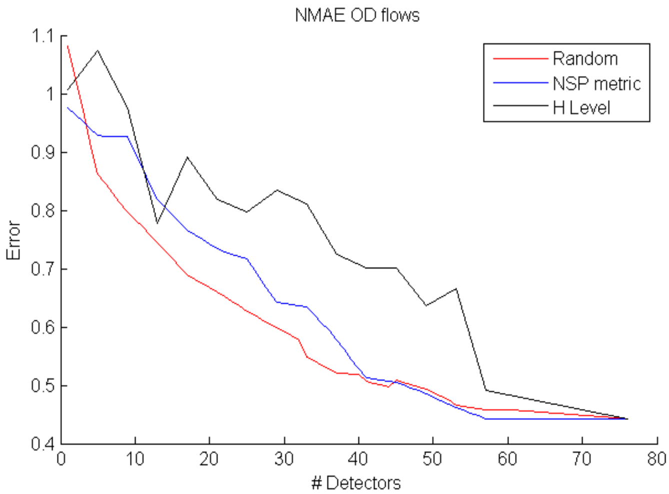


Figure 9: NMAE values for a growing number of detectors installed and between estimated and true OD matrices using eqns (4). Maximisation of h is added as comparison.

In conclusion, even if the greedy algorithm is not explicitly maximizing the value of h, it seems to provide the most effective outcomes in terms of overall error reduction, while allowing to estimate more efficiently the gradient of the estimation error associated to the OD flows.

Again, it is important to stress out that the apparent good performance of the random solutions is not surprising, since in such mid-sized networks is obvious that the first half of the detectors installed would be nearly all equally good, especially if one considers the convexity of the problem adopted. In future research we will test how this gain can become more significant with the growing size of the network.

5. CONCLUSIONS AND FUTURE RESEARCH

The positioning of sensors in a network has been widely considered a crucial problem for network estimation problems, and in particular for estimating origin-destination matrices.

This paper has provided extra insights into the impact of using linearly independent information for the link counts, and in turn the potential gains of adding linearly dependent information that observed flows can provide on the non-observed ones.

Using recent findings in Network Sensor Location problems, and in particular in the area of observability problems, we have extended the traditional objective function used in OD estimation by adding extra terms that are related to fully observable flows, i.e. those non-observed flows that however can be derived uniquely by the observed variables.

Results on a mid-sized network such as the Sioux Falls are promising, as the contribution of the extra term is systematic. Moreover, we show that locations found using the NSP metric and the greedy algorithm proposed in Viti et al. (2014) seem to outperform solutions that directly look for the minimum number of sensors that provide a certain partial observability level *h*.

In future research we will analyse whether this gain is scalable with the size of the network, its topological heterogeneity and the application in a dynamic setting. Furthermore, we will compare the adopted methodology with alternative flow-estimation solutions such as the greedy heuristic of Yang and Zhou (1998).

REFERENCES

Ben-Akiva, M., Lerman, S.R. (1985). *Discrete Choice Analysis*, MIT Press.

Bianco L., Confessore G., Reverberi P. (2001). A Network Based Model for Traffic Sensor Location with Implications on O/D Matrix Estimates. *Transportation Science* 35(1), pp. 50–60.

Bell M.G.H.(1983). The estimation of an origin destination matrix from traffic counts. *Transportation Science* 17(2), pp. 198-217.

Bell M.G.H.(1991). The estimation of origin-destination matrices by constrained generalized least squares. *Transportation Research Part B: Methodological* 25(1), pp. 13-22.

Bierlaire M., Toint P.L. (1994), Meuse: An origin-destination matrix estimator that exploits structure. *Transportation Research Part B: Methodological* 29(1), pp. 47-60.

BPR (1964). Traffic Assignment Manual. U.S DOT.

Cascetta E. (1984). Estimation of trip matrices from traffic counts and survey data: a generalized least squares approach. *Transportation Research Part B: Methodological* 18, pp. 289-299

Cascetta E. (2009). *Transportation Systems Analysis: Models and Applications.*Springer (2nd Edition).

Cascetta E., Papola A., Marzano V., Simonelli F., Vitiello I. (2013). Quasi-dynamic estimation of o–d flows from traffic counts: Formulation, statistical validation and performance analysis on real data. *Transportation Research Part B: Methodological* 55, pp. 171-187

Castillo E., Calviño A., Lo H.K., Menendez, J.M., Grande Z. (2014). Non-planar hole-generated networks and link flow observability based on link counters. *Transportation Research Part B: Methodological* 68, pp. 239-261.

Castillo E., Conejo A., Menendez J.M., Jimenez P. (2008). The Observability Problem in Traffic Network Models. *Computer Aided Civil and Infrastructure Engineering*, 23, pp. 208–222.

Castillo E., Gallego I., Menéndez J.M., Jiménez P. (2012) Link Flow Estimation in Traffic Networks on the Basis of Link Flow Observations. *Journal of Intelligent Transportation Systems* 15(4), pp. 205-222.

Castillo E., Jimenez P., Menendez J.M., Conejo J.A. (2009). The Observability Problem in Traffic Models: Algebraic and Topological Methods. *IEEE Transactions on Intelligent Transportation Systems* 9(2), pp. 275-287.

Chen A., Chootinan P., Recker W. (2005). Examining the Quality of Synthetic O-D Trip Table Estimated by Path Flow Estimator. *Journal of Transportation Engineering*, 131, 506–513.

Cipriani E., Fusco G., Gori S., Petrelli M. (2006). Heuristic Methods for the Optimal Location of Road Traffic Monitoring Stations. Proceedings of the IEEE-ITS Conference, September 2006, Toronto, Canada.

Ehlert A., Bell M.G.H., Grosso, S. (2006). The Optimization of Traffic Count Locations in Road Networks. *Transportation Research Part B: Methodological* 40, pp. 460-479.

Fisk C.S. (1988). On combining maximum entropy trip matrix estimation with user optimal assignment. *Transportation Research Part B: Methodological* 22(1), pp. 69-73.

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

Gan L., Yang H., Wong S.C. (2005) Traffic Counting Location and Error Bound in Origin-Destination Matrix Estimation Problems. *Journal of Transportation Engineering* 131(7), pp. 524–534.

Gentili M., Mirchandani P.B. (2012). Locating Sensors on Traffic Networks: Models, Challenges and Research Opportunities. *Transportation Research Part C: Emerging Technologies*24, pp. 227-255.

He S. (2013). A Graphical Approach to Identify Sensor Locations for Link Flow Inference. *Transportation Research Part B: Methodological* 51, pp. 65-76.

Hu S., Peeta S., Chu C. (2009) Identification of Vehicle Sensor Locations for Link-Based Network Traffic Applications. *Transportation Research Part B: Methodological* 43, pp. 873-894.

Lam W.H.K., H.P. Lo (1990). Accuracy of O-D estimates from traffic counts. *Traffic Engineering and Control* 31, pp. 358-367.

Larsson T., Peterson J.T.L. (2010) Allocation of Link Flow Detectors for Origin-Destination Matrix Estimation - A Comparative Study. *Computer-Aided Civil and Infrastructure Engineering* 25, pp. 116–131

LeBlanc L.J., Fahrangian K. (1982). Selection of a trip table which reproduces observed link flows. *Transportation Research Part B: Methodological* 16, pp. 83-88.

Marzano V., Papola A., Simonelli F. (2009). Limits and perspectives of effective O–D matrix correction using traffic counts. *Transportation Research Part C: Emerging Technologies* 17(2), pp. 120-132.

Maher M.J., 1983. Inferences on trip matrices from observations on link volumes: a Bayesian statistical approach. *Transportation Research Part B: Methodological*17, No. 6, pp. 435-447.

Ng M. (2012). Synergistic Sensor Location for Link Flow Inference without Path Enumeration: A Node-based Approach. *Transportation Research Part B: Methodological*, 46(3), pp. 781-788.

Nguyen S., 1977. Estimating an OD Matrix from network data: a network equilibrium approach. Publication No. 60, Centre de Recherche sur les Transports, Universite de Montreal, Quebec.

Rinaldi M., Viti F., Corman F. (2013). A Null-space Metric for the Analysis of Partial Network observability in Sensor Location Problems. In Proceedings of the 92nd TRB Annual Meeting, January 11-15 2013, Washington DC.

Rinaldi M., Corman F., Viti F. (2015). Assessing the effect of route information on network observability applied to sensor location problems. *Transport Research Procedia* (in press).

Robillard P. (1975). Estimating the O-D Matrix from observed link volumes. *Transportation Research* 9, pp. 123-128.

Simonelli F., Marzano V., Papola A., Vitiello I. (2012). A Network Sensor Location Procedure Accounting for O-D Matrix Estimate Variability. *Transportation Research Part B: Methodological*, 46, pp. 1624-1638.

Spiess H. (1987). A maximum-likelihood model for estimating origin-destination matrices, *Transportation Research Part B: Methodological* 21, pp. 395-412.

Van Zuylen H.J., Willumsen L.G. (1980). The most likely trip matrix estimated from traffic counts. *Transportation Research Part B: Methodological* 14, pp. 281-293.

Viti F. Corman F. (2012). A Novel Approach to the Sensor Location Problem for Measuring the Observed Network Flow Variability. Proceedings of the 5th International Symposium of Traffic Network Reliability, 18-19 December, Hong Kong, China

Viti F., Rinaldi M., Corman F., Tampere C.M.J. (2014). Assessing partial observability in network sensor location problems. *Transportation Research Part B: Methodological* 70, pp. 65-89.

Wang N., Gentili M., Mirchandani P. (2012). A Model to Locate Sensors for Estimating Static OD Volumes Given Prior Flow Information. *Transportation Research Record* 2283, pp. 67-73.

Willumsen, L.G. (1978) Estimation of an O-D Matrix from Traffic Counts – A Review. Working Paper. Institute of Transport Studies, University of Leeds, Leeds, UK.

Yang H., Iida Y., Sasaki T. (1991) An Analysis of the Reliability of an Origin-destination Trip Matrix Estimated from Traffic Counts. *Transportation Research Part B: Methodological* 25, pp. 351–363.

Yang H., Meng, Q., Bell, M.G.H. (2001). Simultaneous estimation of the origin–destination matrices and travel-cost coefficient for congested networks in a stochastic user equilibrium. *Transportation Science* 35, pp. 107–123.

Yang H., Sasaki T., Iida Y., Asakura Y. (1992). Estimation of origin-destination matrices from link traffic counts on congested networks. *Transportation Research Part B: Methodological* 26, pp. 417-434.

Yang H., Yang C., Gan. L. (2006). Models and Algorithms for the Screen Line-based Traffic-Counting Location Problems. *Computers and Operations Research* 33(3), pp. 836–858.

Yang H., Zhou J. (1998). Optimal Traffic Counting Location for Origin-Destination Matrix Estimation. *Transportation Research Part B: Methodological* 32(2), pp. 109-126.

Yim K.N., Lam W.H.K. (1998) Evaluation of Count Location Selection Methods for Estimation of OD Matrices. *Journal of Transportation Engineering* 124(4), pp. 376–383.

Zhou X., List G. (2010). An Information-Theoretic Sensor Location Model for Traffic Origin-Destination Demand Estimation Applications. *Transportation Science*, 44(2), pp. 254–273.