A real case application for traffic demand estimation using multi-class vehicle traffic counts

Armando Cartenì

Assistant Professor in Transportation Engineering

Department of Civil, Construction and Environmental Engineering

University of Naples Federico II

Via Claudio 21 - 80125 Naples – Italy

Abstract

All the elements of a transportation system model generally suffer from some approximation. Normally the traffic demand estimation is considered the most crucial and problematic element to be simulated, and vehicle traffic counts are generally used to "update" it so that the whole model system is able to reproduce the observed road vehicle flows. Updating demand estimation using traffic counts has received considerable attention in recent years. In this paper a methodology for updating the demand vector using traffic counts on congested networks was applied to the real case of the mobility demand estimation of the Naples city (Italy). The results of the estimations show the good capacity of the proposed methodology to reproduce the vehicle traffic counts observed. Furthermore, through the estimated traffic demand was possible to analysed the structural characteristics of the traffic demand within Naples city.

Keyword: transportation planning; origin-destination demand updating; bi-level programming problem; transportation simulation model.

1 Introduction

A transportation system can be defined as the combination of elements and their interactions, which produce the demand for travel within a given area and the supply of transportation services to satisfy this demand (Cascetta [7]).

The interactions among the various elements of a transportation system can be simulated with mathematical models. Supply models simulate the performances of the transportation services available in the different zones and demand models simulate the aspects of travel demand as a function of the activity system and of the supply performances (demand can sometimes be estimated directly through surveys). The interaction between demand and supply can be simulated through assignment models which allow estimation of link flows.

All the elements of a transport model system (zoning, graph, cost functions, demand, assignment model, etc.) are generally affected by some approximations. Normally the demand vector is considered the most crucial and problematic element to be simulated and, therefore, traffic counts are generally used in order to correct it so that the whole model system is able to reproduce the observed link flows.

Updating the origin-destination (OD) demand vector using traffic counts has received considerable attention in recent years both from the theoretical and empirical point of view. This can be easily explained by the cost and complexity of sampling surveys (disaggregate model estimation), as well as the lack of precision related to both direct and model estimators of OD flows. On the other hand, user flows on some network links (traffic counts) are cheap and easily obtainable, often automatically.

The focus of OD demand updating is on estimating and predicting aggregate values of the OD vector, i.e. traffic flows, rather than individual OD flows. It is expected that a vector capable of reproducing some of such aggregates with sufficient precision will give better predictions also after network changes.

Various methods have been developed in the literature to estimate the demand vector from traffic counts. For the OD Demand Vectors using Traffic Counts (henceforth known as UDVTC) problem on uncongested networks, many methods have been proposed in the literature. Some estimators are based on the maximum entropy principle (Van Zuylen and Willumsen [22], Willumsen [23]). Other methods are derived from classical statistical approaches such as the Generalized Least Squares (GLS) estimator (Cascetta [6], Bell [1]), or Bayesian inference (Maher [17]). Finally, a multiobjective programming formulation was

proposed by Brenninger-Göthe et al. [3], following a GLS approach, while Cascetta et al. ([9]) proposed an estimator in which equality constraints between observed and assigned link flows are explicitly considered. Other approaches and specifications can be found in the literature (McNeil [18], Nguyen [20], Ben-Akiva et al. [2], Spiess, [21]).

A comprehensive review of the various methods for uncongested networks is proposed by Cascetta and Nguyen [8]. Di Gangi [13] proposes a numerical comparison between the performance of the most widely used estimators.

With respect to the congested network approach, in the literature the problem of circular dependence between OD demand vector updating and traffic flow assignment has been widely studied. Early simple models were proposed by Nguyen [19] and Fisk and Boyce [16]. Recently, other authors formulated and solved the UDVTC problem for congested networks as a bi-level mathematical programming problem for Deterministic User Equilibrium (DUE) assignment (see for example Chen and Florian [11]; Fisk [15]; Yang [21]). Relatively few have dealt with the UDVTC problem by considering congested networks with Stochastic User Equilibrium (SUE) assignment models (see for example: Bell et al. [4]; Cascetta and Postorino [10]). In particular, in Cascetta and Postorino [10] an MSA algorithm (with some modifications) applied to a test network has been proposed.

Starting from the results of Cascetta and Postorino [10], in this paper a bi-level algorithm for updating the demand vector using traffic counts on congested networks is applied to a real case. We estimated the OD demand vectors for the city of Naples (Campania, Italy) for different days (average weekday and weekend) and time slices (morning/afternoon/evening - peak hour/off-peak hour).

This application was made possible thanks to the ATENA (Ambiente Traffico Telematica Napoli) Transport Plan and especially thanks to the activities that the Department of Transportation Engineering of Naples University carried out as consultants: Initiative B (Telematic System for Traffic Management and Control) - Action 2 (Development of a Telematic System for Traffic Management and Control in Naples, Italy).

The paper is divided into three sections; in the first the methodology is presented; in the second application to a real case is described, while the third reports the conclusions.

2 Methodology

The UDVTC problem can be expressed in a general form as:

$$d^* = \underset{x \geq 0}{argmin}\left[z_1\left(x, \hat{d}\right) + z_2\left(v(x), \hat{f}\right)\right] \qquad (1)$$

The two functions $z_1(.)$ and $z_2(.)$ can be considered respectively as distance measures of the unknown demand vector x from the *a priori* estimate $\hat{d}$, and of the flows obtained by assigning x to the network, $v(x)$, from the traffic counts $\hat{f}$. In general, the functional form of the two terms $z_1(.)$ and $z_2(.)$ depends on the type of information available (experimental or non-experimental) and on the probability laws associated with such information.

The most widely used estimator is the Generalized Least Squares (*GLS*). The *GLS* estimator provides the estimate of an unknown vector d^{GLS} starting from a system of linear stochastic equations:

$$d^{GLS} = \underset{x \in S}{arg\,min}\left[(\hat{d} - x)^T Z^{-1}(\hat{d} - x) + (\hat{f} - \hat{M}x)^T W^{-1}(\hat{f} - \hat{M}x)\right] \qquad (2)$$

where:

− Z represents the variance-covariance matrix of errors between *a priori* demand and the unknown one;

− W represents the variance-covariance matrix of errors between traffic counts and assigned flows;

− $\hat{M}$ is the assignment matrix.

Under the hypothesis of congested networks (link cost functions of link flows), in equation (2) a problem of circular dependence arises, since it is possible to estimate link flows $v(d^{GLS})$, and hence link costs $c(v(d^{GLS}))$, by assigning the demand d^{GLS} solution to problem (2), which in turn is estimated from link flows and costs. The estimation problem can be formalized as a fixed-point problem. The assignment matrix $\hat{M}$ can be expressed as a function of demand flows. In fact, if we combine the relationship connecting the assignment matrix to link costs, $\hat{M} = \hat{M}(c)$, with the cost functions $c = c(f)$, and introduce the relationship between link and demand flows through the assignment model, $f = v(d)$, we get: $\hat{M} = \hat{M}(c(v(d)))$. Let $d = \delta(\hat{M})$ be the solution of the estimation problem (2) for a given assignment matrix $\hat{M}$, the UDVTC

4

problem can be expressed through a fixed-point model combining the two functions $d = \delta(\hat{M})$ and $\hat{M} = \hat{M}(c(v(d)))$:

$$d^{GLS} = \arg\min_{x \in S} \left[(\hat{d} - x)^T Z^{-1} (\hat{d} - x) + (\hat{f} - \hat{M}(c(v(d^{GLS})))x)^T W^{-1} (\hat{f} - \hat{M}(c(v(d^{GLS})))x) \right]$$

As stated above, in this paper, the bi-level algorithm proposed by Cascetta and Postorino [10] is applied to a real scale application. It is based on the iterative assignment of the OD vector and the solution of the optimisation as described in the following:

Step 0. Initialisation:

$k = 0$

initial demand vector d^0; explicit evaluation of the assignment matrix M^0 (for $d = d^0$) by using the SUE assignment model with $c^0 = c(v(d^0))$

Step 1. Update: $k = k+1$

Step 2. Solve the Problem:

$$d^k = \arg\min_{x \in S} \left[(\hat{d} - x^k)^T Z^{-1} (\hat{d} - x^k) + (\hat{f} - M^{k-1} x^k)^T W^{-1} (\hat{f} - M^{k-1} x^k) \right]$$

Step 3. Solve the SUE assignment problem for demand vector d^k, compute c^k and evaluate M^k:

$$f^k = f(d^k); \quad c^k = c(v(d^k)); \quad M^k = M(c^k)$$

Step 4. Termination test on the maximum percentage difference between elements d_i of vector d

$$\max \frac{\left| d_i^k - d_i^{k-1} \right|}{d_i^k} < \varepsilon$$

If the test is satisfied the algorithm ends, otherwise return to *Step 1*.

The OD demand vector is obtained by solving the problem at Step 2 with a fixed assignment matrix resulting from iteration $k-1$. This OD demand vector is in turn used to solve the SUE problem, and to compute the new costs and new assignment matrix.

3 Application case study

In this section UDVTC application to a real case is described. The study area is the city of Naples in southern Italy. Naples has about one million inhabitants and a very high population density.

In figure 1 the application architecture is reported. The first activity was the traffic survey from which we were able to define the traffic count sections and estimate traffic flows. Then through an OD demand model the initial demand vectors were estimated. The traffic counts and OD vectors were the input for the OD updating process as well as a SUE assignment model. The output of this procedure consisted of assignment link flows (obtained through the updated OD vector). These flows were compared with a hold-out sample consisting of some traffic sections (traffic flows) not used for OD updating.
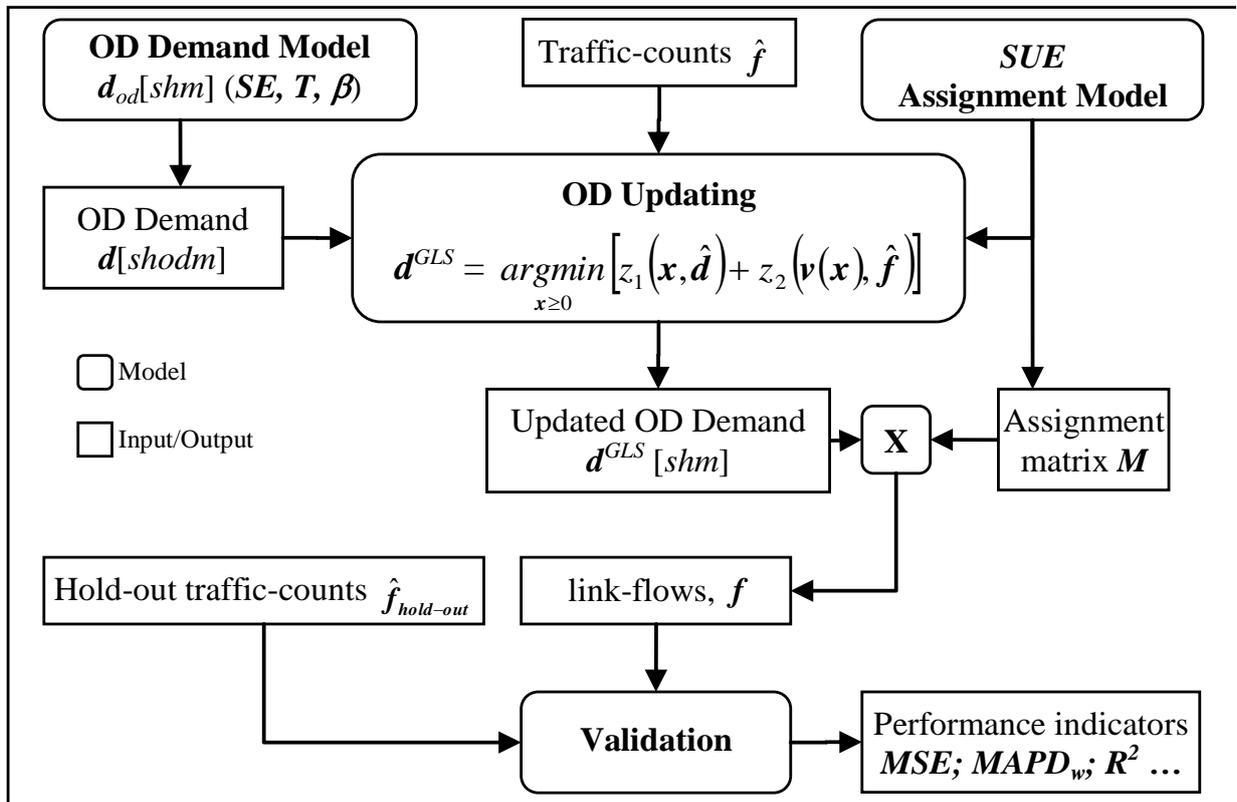

Figure 1: The application methodology

3.1 Traffic vehicle survey

Sound planning of the traffic counts survey appears to be critical for this kind of application. The traffic count sections were identified through a screen-line method.

The survey was conducted by Elasis sc.p.a.. In all, 131 traffic count sections were identified and monitored (see figure 2) for two average days: weekday (working day) and weekend (holiday). The reference period measured was 06:00-21:00 for the average weekday and 09:00-24:00 for the average weekend. The flows measured were divided by vehicle type: car, motorcycle, light goods vehicles and heavy goods vehicles.

For each traffic count section it was possible to estimate the peak and off-peak hours for morning, afternoon and evening periods.

For data analysis, the study area was divided into two basins (see figure 2): city centre and an outer band (west basin $\cup$ north basin $\cup$ east basin).

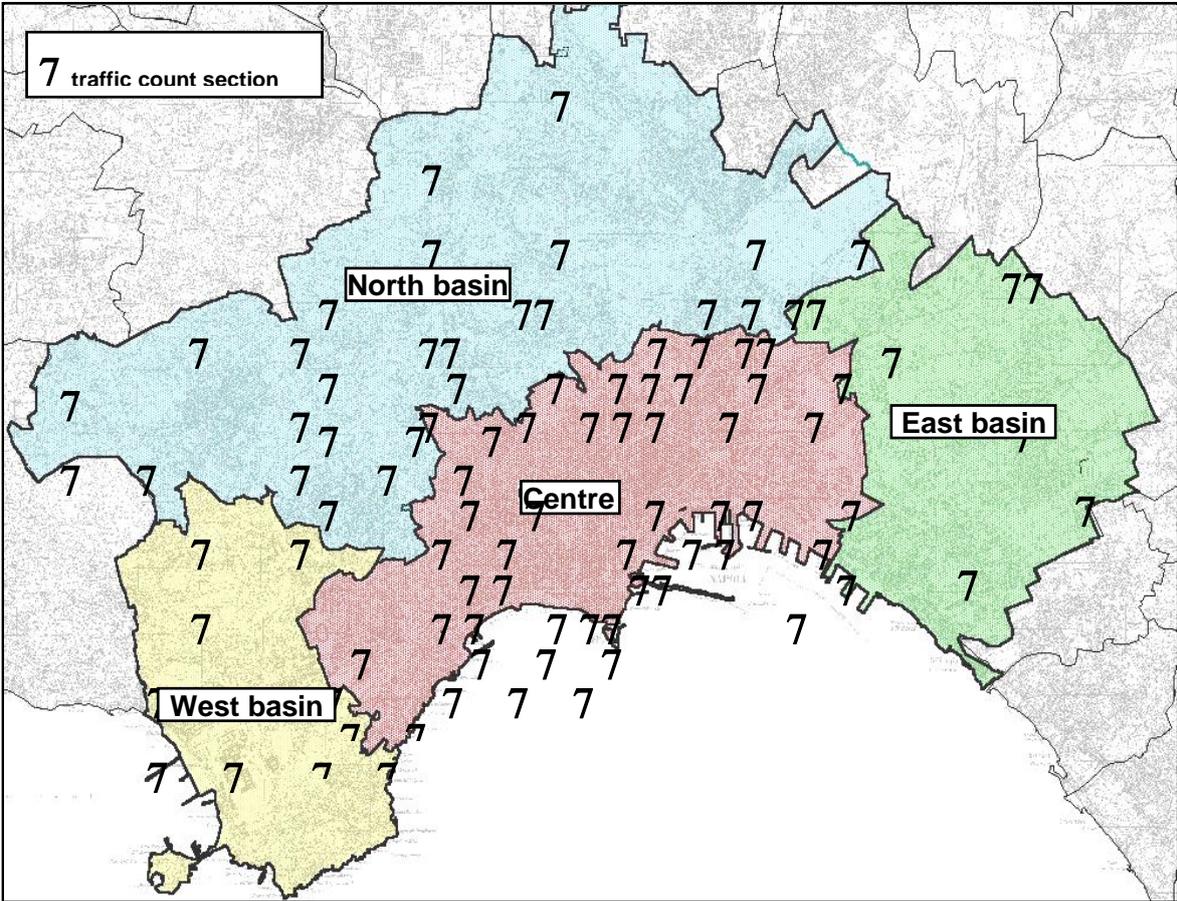**Figure 2: The study area of Naples city (Italy)**

For the average weekday, it emerges that the morning peak hours ($\subset$ [06:30-09:00]) for the outer band are earlier than the central ones ($\subset$ [07:30-10:00]). This is explained by the fact that trips towards the city centre need more time to reach their destination and so have to anticipate their departure time. Besides, the flows towards the

centre have their peak in the early hours of the morning; while, the flows leaving the city centre are, on average, less variable (constant).

With respect to afternoon peak hours in the north-west basin, maximum flows occur between 16:30 and 20:00, during which there are presumably "work to home" trips. By contrast, in the north-east basin of the city, the afternoon peak hours are strongly anticipated and occur between 13:00 and 16:30.

As for the weekend, the most significant traffic flows occur after 12:00. In the afternoon, traffic flows are about 50% of those on weekdays. In the evening the peak hour occurs at about 20:00.

From the data monitored we were able to estimate the network peak (off-peak) hours (see table 2).

3.2   Traffic demand updating and results analysis

As stated above, after estimation of the traffic count flows (called updating sample), the initial demand vectors were estimated. The study area was divided into 167 zones. The topological supply model consists of a graph with 3,070 nodes and 5,748 links; the generalized transport cost c associated to each link was estimated through the relation: $c = b_1 t^r(f) + b_2 t^w(f)$; where: $t^r(f)$ is the running time, a function of the link flows vector f (estimated using the function proposed by Carten ìand Punzo [5]); $t^w(f)$ is the waiting time at the intersection, a function of the link flows vector f (estimated using the function proposed by Doherty [14]);

The partial share Multinomial Logit model calibrated by de Luca and Papola [12] was used for initial OD vector estimation. For link flow vector estimation a Stochastic User Equilibrium (SUE) assignment model was used.

The traffic counts and the initial OD vectors were the input of the bi-level *GLS* models described in section 2 used for the OD updating process. The output of this procedure was the updated (corrected) assignment link flows.

Mean Square Error (*MSE*), the weighted Mean Absolute Percent Deviation (*MAPD$_w$*) and R^2 are the performance indicators used to assess estimation quality:

$$MSE = \frac{\sum_{i=1}^{nc}\left(f_i^a - f_i^c\right)^2}{n_c - 1}$$

$$MAPD_w = 100 \cdot \frac{\sum\limits_{i=1}^{nc} f_i^c \cdot \dfrac{\left| f_i^a - f_i^c \right|}{f_i^c}}{\sum\limits_{i=1}^{n} f_i^c} \; ;$$

$$R^2 = 1 - \frac{\sum\limits_{i=1}^{nc} \left(f_i^a - f_i^c \right)^2}{\sum\limits_{i=1}^{nc} \left(f_i^a - \bar{f}^c \right)^2}$$

where:

- f_i^c is the generic traffic count (flow) on link i;

- f_i^a is the generic assigned traffic flow on link i;

- n_c is the number of traffic count sections;

- $\bar{f}^c = \Sigma_i f_i^c / n_c$ is the average traffic count.

The quality of the estimation was validated through a *hold-out sample*, consisting of 30 traffic sections (flows) not used for the OD updating procedure.

As shown by the results (Table 1), the performance indicator values were statistically significant both for the updating sample and for the hold-out. In figure 3 traffic counts are measured against assigned flows.

The updated OD vectors are studied from both the temporal and spatial point of view. From a temporal point of view, different results were obtained for different periods of time (see table 2). For the average weekday (business day), the peak hour is 7:30-8:00 with more than 127,000 vehicles/hour within Naples (home to work trips). For the rest of the day the demand level is quite constant with about 85,000 vehicles/hour.

With respect to the average weekend (holiday) the demand level increases during the first hours of the day reaching its peak between 12:00 and 13:00 with about 71,500 vehicles/hour. After 13:00 the demand level decreases till 20:00 when the evening peak hour occurs.

Table 1: Results in terms of performance indicator values

| OD			Calibration sample			Hold-out sample		
			$MAPD_w$	R^2	MSE	$MAPD_w$	R^2	MSE
weekday	morning	peak hour	11.41%	0.91	83,453	15.69%	0.88	102,499
		off-peak hour	15.52%	0.90	111,061	17.07%	0.89	107,845
	afternoon	peak hour	9.17%	0.96	63,106	10.47%	0.94	65,093
		off-peak hour	12.37%	0.94	85,056	14.07%	0.93	82,101
	evening	peak hour	17.59%	0.89	128,220	16.15%	0.91	101,744
		off-peak hour	11.02%	0.93	80,082	13.17%	0.91	79,420
weekend	morning	peak hour	16.35%	0.89	113,208	18.64%	0.89	116,377
		off-peak hour	13.92%	0.93	99,403	15.49%	0.92	96,666
	afternoon	peak hour	8.42%	0.97	60,344	12.79%	0.93	81,408
		off-peak hour	12.59%	0.93	86,616	14.55%	0.91	87,710
	evening	peak hour	14.46%	0.92	98,437	12.64%	0.93	76,349
		off-peak hour	12.28%	0.91	86,971	15.14%	0.89	88,359
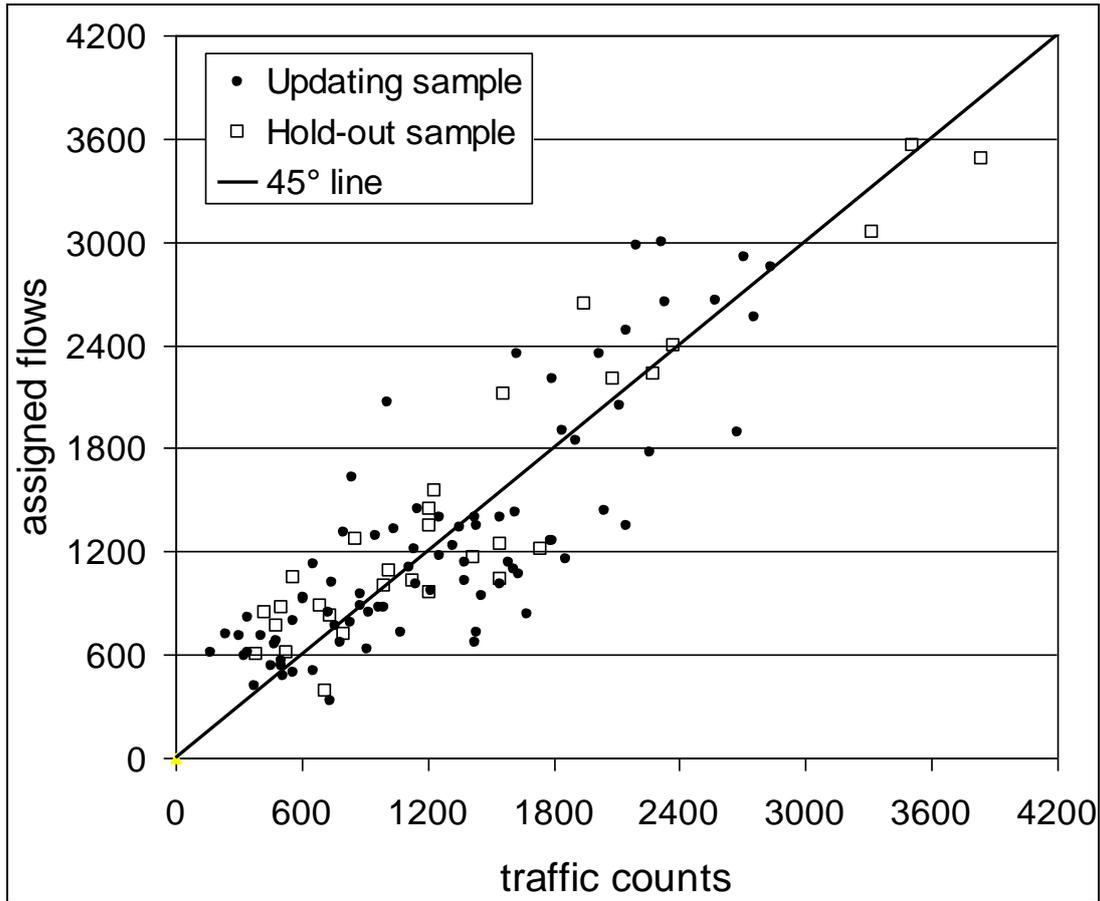
Figure 3: Traffic counts measured against assigned flows (updating and hold-out sample) – average weekday, morning peak hour

From a spatial point of view, the study area is divided into four basins (see figure 2): west basin, north basin, east basin and historical city centre. Most of the daily trips occur inside the historical centre and the north basin. On an average weekday a significant number of trips are made towards the east basin (high business density), while on an average weekend a significant number of trips are made towards the city centre where many cultural and free-time activities are concentrated.

Table 2: Peak /off-peak hours and OD demand level (in vehicles/hour)

		Average weekday		*Average weekend*	
		peak hour	*off-peak hour*	*peak hour*	*off-peak hour*
morning	time interval	07:30-08:30	12:00-13:00	12:00-13:00	09:00-10:00
	OD level	*127,411*	*80,160*	*71,483*	*63,625*
afternoon	time interval	15:45-16:45	16:45-17:45	13:45-14:45	16:00-17:00
	OD level	*85,738*	*85,615*	*57,287*	*53,297*

evening	time interval	20:00-21:00	-	20:00-21:00	23:00-24:00
	OD level	*84,037*	-	*17,886*	*57,033*

4 Conclusions

In this paper the methodology proposed by Cascetta and Postorino [10] for updating the demand vector using traffic counts on congested networks was applied to the real case of the city of Naples (Italy). The results show that this method is suitable for reproducing the traffic counts measured. Furthermore, through the estimated OD vectors it was possible to make a structural analysis of transportation demand in Naples. The estimated OD demand vectors for different time slices will be useful for future transportation planning applications.

References

[1] Bell, M., *The estimation of origin-destination matrices by constrained generalized least squares*. Transp. Research vol.25B(2), pp.13-22, 1991.

[2] Ben Akiva, M.E., Macke, P.P., Hsu, P.S., *Alternative methods to estimate route-level trip tables and expand on-board surveys.* TRB1037, 1-11, 1985.

[3] Brenninger-Göthe, M., Jornsten, K. O., Lundgren, J. T, *Estimation of origin/destination matrices from traffic counts using multiobjective programming formulation*. Transportation Res. 23B 257–269, 1989.

[4] Bell, M.G.H., Lam, W.H.K., Iida, Y., *A time-dependent multiclass path flow estimator*. Proc. 13th Symposium on Transportation and Traffic Theory. Lyon, France, 1996.

[5] Cartenì, A. and Punzo, V., *Travel time cost function for urban roads: a case study in Italy*. WIT Transactions on the Built Environment, 96, pp. 233-243, 2007.

[6] Cascetta, E., *Estimation of trip matrices from link traffic counts and survey data: a generalized least square estimation.* Trans. Research vol. 18B, 1984.

[7] Cascetta, E., *Transportation systems engineering: theory and methods*. Kluwer Academic Publishers, 2001.

[8] Cascetta, E., Nguyen, S., *A unified framework for estimating or updating Origin-Destination matrices from traffic counts.* Transportation Research, vol. 22B, pp. 437-455, 1988.

[9] Cascetta, E., Papola, A. and Cartenì A., *Prediction reliability of the transport simulation models: a before and after study in Naples*. Proceedings of the European Transport Conference (ETC), Strasbourg, France, 2005.

[10] Cascetta, E., Postorino, M.N., *Fixed point models for the estimation of O-D matrices using traffic counts on congested networks*. Transportation Science, Vol. 35, 2001.

[11] Chen, Y., Florian, M., *O/D demand adjustment problem with congestion. Part I: model analysis and optimality conditions*. CRT Publication No. 94-56, Centre de Recherche sur les Transport, Univ. de Montreal, Canada, 1994.

[12] de Luca, S. and Papola, A., *The evaluation of Travel demand management policies in the urban area of Naples*. 8th "Urban Transport and the Environment" Proceedings, Lemnos, Greece, WIT Press, pp.185-194, 2001.

[13] Di Gangi, M., *Una valutazione delle prestazioni statistiche degli estimatori della matrice O/D che combinano i risultati di indagini e/o modelli con i conteggi di flusso di traffico*. Ricerca Operativa No. 51, pp. 23-59, 1988.

[14] Doherty, A.R., *A comprehensive junction delay formula*, LTR 1Working Paper, Department of Transport, 1977.

[15] Fisk, C. S., *On combining maximum entropy trip matrix estimation with user optimal assignment*. Transportation Res. 22B 69–73, 1988.

[16] Fisk, C. S. and Boyce, D. E., *A note on trip matrix estimation from link traffic counts data*. Transportation Res. 17B 245–250, 1983.

[17] Maher, M.J., *Inference on trip matrices from observations on link volumes. A Bayesian statistical approach*. Transp. Research vol.17B(2),435-447, 1983.

[18] McNeil, S., *Quadratic matrix entry estimation methods*. Ph.D. thesis, Dept. of Civil Engineering, Carnegie-Mellon University, Pittsburgh, PA, 1983.

[19] Nguyen, S., *Estimation on O/D matrix from network data: A network equilibrium approach*. Publication 87, Centre de Recherche sur les Transport, Université de Montreal, Canada, 1977.

[20] Nguyen, S., *Modèle de distribution spatiale tenant compte des itineraire*. INFOR21(4), 1983.

[21] Spiess, H., *A gradient approach for the O/D matrix adjustment problem*. Publication No. 693, CRT Université de Montréal, Montréal, Canada, 1990.

[22] Van Zuylen, J.H., Willumsen, L.G., *The most likely trip matrix estimated from traffic counts*. Transportation Research vol. 14B, 1980.

[23] Willumsen, L.G., *Estimating time-dependent trip matrices from traffic counts.* Proceeding of 9th International Symposium on Transportation and Traffic Theory, VNU Science Press, Utrecht, 1984.

[24] Yang, H., *Heuristic algorithms for the bi-level origin/destination matrix estimation problem.* Transportation Research 29B 2, 1995.