Index

YEAR XVII, NUMBER 49, December 2011

European Transport \ Trasporti Europei n. 49 (2011): 3-23

Estimation of origin-destination matrix from traffic counts: the state of the art

Sharminda Bera [1], K. V. Krishna Rao [2*]

[1] *Research Scholar, Civil Engineering Department, IIT Bombay, India 400076*
[2] *Prof., Civil Engineering Department, IIT Bombay, India 400076*

Abstract

The estimation of up-to-date origin-destination matrix (ODM) from an obsolete trip data, using current available information is essential in transportation planning, traffic management and operations. Researchers from last 2 decades have explored various methods of estimating ODM using traffic count data. There are two categories of ODM; static and dynamic ODM. This paper presents studies on both the issues of static and dynamic ODM estimation, the reliability measures of the estimated matrix and also the issue of determining the set of traffic link count stations required to acquire maximum information to estimate a reliable matrix.

Keywords: Origin-Destination matrix; Static origin-destination matrix; Dynamic origin-destination matrix; Traffic count stations.

1. Introduction

In developing countries, changes in the land-use and economic state of affairs require momentous transportation planning. One of the most crucial requirements for the transportation planning is on arriving at the traffic pattern between various zones through Origin-Destination matrix (ODM) estimation. Traditional methods of estimating ODM are through large scale sampled surveys like home interview survey, roadside interview and license plate method conducted once in every 1-2 decades. But in situations of financial constraints these surveys become impossible to conduct. And by the time the survey data are collected and processed, the O-D data obtained become obsolete. In consequence, from 1970s many models have been proposed and widely applied for updating/estimating an old/sampled matrix using the current data of traffic counts collected on a set of links. The accuracy of these estimated matrix depends on

* Corresponding author: K. V. Krishna Rao (kvkrao@civil.iitb.ac.in)

the estimation model used, the input data errors, and on the set of links with collected traffic counts.

From the past studies, the ODM estimation models can be categorized as static and dynamic based on its application. In static methods the traffic flows are considered as time-independent and an average O-D demand is determined for long-time transportation planning and design purpose. Whereas from last two decades different dynamic approaches are proposed which are meant for short-term strategies like route guidance, traffic control on freeways, intersections etc. The assignment matrix which provides an approximate trip proportions based on the route choice behavior of the trip makers is the complicated part in the estimation problem. Another aspect, on which the reliability of the estimated ODM largely depends, is the optimum traffic counting locations. The traffic counts collected should provide as much traffic information as possible saving subsequent manpower requirement in data collection. Various rules are proposed in literature to select the traffic counting location points. This paper gives a vast study on the ODM estimation and the various related issues. It gives ideas about the various methodologies developed till date, the optimization algorithms used to solve the problem, convergence problems of those algorithms and most vital issue the reliability or the accuracy of the estimated ODM.

Though there are some softwares which can estimate ODM. But it is always wise to have coded algorithms which can be flexible with respect to the type of data available according to the study area. Still research is going on to estimate a reliable ODM efficiently. A comprehensive state-of-the-art review has been conveyed through this paper. Also the studies based on the optimum number of traffic counting locations required and their influences on estimated ODM are covered. This can help to acquire a good knowledge regarding the various developed models and also gives a direction for future research in this area.

The paper is organized as follows. Section 2 discusses about the static ODM estimation, its problem formulation and reviews the various methodologies developed till date. Discussions on dynamic and time-dependent ODM estimation are presented in Section 3. It encloses the models developed from past two decades. Section 4 gives a brief description of how the reliability of the estimated ODM is measured. The studies based on determining the traffic counting location points and its optimum number are presented in Section 5. Lastly Section 6 gives the final conclusions and the directions for future research.

2. Static ODM estimation from traffic counts

A static ODM estimation problem does not consider the time-dependent traffic flows and is assumed to represent a steady-state situation over a time period. The average traffic counts are collected for a longer duration to determine the average O-D trips. Let a transportation network is defined by W O-D pairs and a set of L links with $A \subseteq L$ as the subset of links with the observed traffic counts. Considering following notations:

t_w is the number of trips of O-D pair $w, w \in W$

p_w^a is the proportion of trips O-D pair $w \in W$ traversing link $a \in A$

v_a is the expected link flow for the link $a \in A$

The link flows, the trips between O-D pairs and the proportional matrix is related by the formulation given by eqn. (1).

$$\sum_{w\in W} p_w^a t_w = v_a \qquad a\in A \tag{1}$$

This estimation problem is underspecified as A is less than W ODs and there is no unique solution. So, additional information (a prior or a target or a sampled matrix) is needed to determine a unique trip matrix.

2.1 Travel Demand Model Based Methods

Initially the researchers tried to relate the trip matrix as a function of models (like the gravity models) with related parameters. Some of the researchers like Robillard (1975), Hogberg (1976) used Gravity (GR) model based approaches and some (Tamin and Willumsen, 1989; Tamin et al, 2003) used Gravity-Opportunity (GO) based models for estimating ODM. These techniques require zonal data for calibrating the parameters of the demand models. The main drawback of the gravity model is that it cannot handle with accuracy external-external trips (refer Willumsen, 1981). The fundamental model for estimating the ODM based on traffic counts in this approach is given by combining eqn. (1) for trip purpose m with the standard gravity model,

$$v_a^m = \sum_m \sum_i \sum_j .O_i^m D_j^m .A_i^m .B_j^m .f_{ij}^m .p_{ij}^{am} \tag{2}$$

where v_a^m gives the flows in particular link a for trip purpose m, O_i^m and D_j^m are the trips produced from zone i and attracted by zone j respectively for trip purpose m, f_{ij}^m here is the deterrence function, A_i^m and B_j^m are the balancing factors and p_{ij}^{am} is the proportion of trips traveling from zone i to zone j using link a for trip purpose m. Tamin and Willumsen (1989), for the calibration of unknown parameters, considered three estimation methods viz. non-linear-least-squares (NLLS), weighted-non-linear-least-squares (WNLLS) and maximum likelihood (ML). The GO models found to consume more time than GR model and does not guarantee the reliability of the estimated matrix.

2.2 Information Minimization (IM) and Entropy Maximization (EM) Approach

EM and IM techniques are used as model building tools in transportation, urban and regional planning context, after Wilson (1970) introduced the concept of Entropy in modeling. The entropy-maximizing procedure analyzes the available information to obtain a unique probability distribution. The number of micro-states $W\{t_w\}$ giving rise to meso-state t_w is given by,

$$W\{t_w\} = \frac{TN!}{\Pi_w t_w!} \tag{3}$$

where *TN* is the total number of trips and t_w is the trips of O-D pair $w \in W$. With the information contained in the observed flows and with other available information, EM is used by Willumsen (1978) and IM approach by Van Zuylen (1978) (refer Van Zuylen and Willumsen, 1980) to estimate trip matrix. Van Zuylen and Willumsen (1980) have shown that both the methods are same except the unit of observation. Van Zuylen's model uses a counted vehicle and Willumsen's a trip. Both the models results in multi-proportional problem. But the convergence of the algorithm has not been proved. Van Zuylen and Branston (1982) further extended the study of Van Zuylen and Willumsen (1980) considering the inconsistency in traffic counts for the case when there is more than one count available on some or all links of the network. Unfortunately, IM and EM based models have the disadvantage of not considering the uncertainties in traffic counts and prior matrix which can be erroneous and can influence the output.

2.3 Statistical Approaches

Several models have been presented in order to estimate or to update ODM from traffic counts for the networks without congestion and with congestion effects via parametric estimation techniques like; Maximum Likelihood (ML), Generalized Least Squares (GLS) and Bayesian Inference (BI).

The method of ***ML*** is one of the oldest and most important in estimation theory. The ML estimates are the set of parameters which will generate the observed sample most often. In ODM estimation maximum likelihood approach maximizes the likelihood of observing the target ODM and the observed traffic counts conditional on the true trip matrix. The data consist of $\hat{\mathbf{V}} = [\hat{v}_1, \hat{v}_2 .... \hat{v}_a]'$ representing a vector of a set of observed traffic counts, and a set of sampled O-D flows $\mathbf{N} = [n_1, n_2, ...n_w]'$. The likelihood of observing two sets of statistically independent data is expressed as

$$L(\mathbf{N}, \hat{\mathbf{V}} \mid \mathbf{T}) = L(\mathbf{N} \mid \mathbf{T}).L(\hat{\mathbf{V}} \mid \mathbf{T}) \tag{4}$$

The sampled O-D flows may be assumed to follow Multivariate normal distribution (MVN) or Poisson probability distribution. This is dependent on small sampling fractions *α*. Consider observed link counts on a set of *A* links and vector $\mathbf{T} = [t_1, t_2, ...t_w]'$ representing the populated trip matrix with elements t_w.

For the logarithm of the probability $L(\mathbf{N}/\mathbf{T})$ we have:

$$\ln L(\mathbf{N}/\mathbf{T}) = -\frac{1}{2}(\mathbf{N} - \alpha\mathbf{T})' \mathbf{Z}^{-1} (\mathbf{N} - \alpha\mathbf{T}) + const. \qquad \text{(MVN)} \tag{5}$$

$$\ln L(\mathbf{N}/\mathbf{T}) = \sum_{w \in W} (-\alpha t_w + n_w \ln(\alpha t_w)) + const. \qquad \text{(Poisson)} \tag{6}$$

where $\mathbf{Z}$ is the covariance matrix of $\mathbf{N}$. If the observed traffic counts are also assumed to be generated either by a MVN distribution or Poisson probability distribution, then the similar expression for the probability $L(\hat{\mathbf{V}}/\mathbf{V}(\mathbf{T}))$ is obtained:

$$\ln L(\hat{\mathbf{V}} / \mathbf{V}(\mathbf{T})) = -\frac{1}{2}(\hat{\mathbf{V}} - \mathbf{V}(\mathbf{T}))'\mathbf{W}^{-1}(\hat{\mathbf{V}} - \mathbf{V}(\mathbf{T})) + const. \qquad \text{(MVN)} \qquad (7)$$

$$\ln L(\hat{\mathbf{V}} / \mathbf{V}(\mathbf{T})) = \sum_{a \in \hat{A}} (\hat{v}_a \ln(V_a(\mathbf{T})) - V_a(\mathbf{T})) + const. \qquad \text{(Poisson)} \qquad (8)$$

where $V_a(\mathbf{T})$ denotes the flow volume on link $a \in A$, resulting from an assignment of **T** and **W** as the variance-covariance matrix. Hence ML function is the log-likelihood of the sum of eqns. (5) or (6) with (7) or (8).

Among a number of branches of regression analysis, the method of ***GLS*** estimation based on the well-known Gauss-Markov theory plays an essential role in many theoretical and practical aspects of statistical inference based model. The advantage of this approach is that no distributional assumptions are made on the sets of data and it allows the combination of the survey data relating directly to O-D movements with traffic count data, while considering the relative accuracy of these data (Bell, 1991a). Let vector $\hat{\mathbf{T}}$ denote the survey estimate of **T** obtained from the grossing up the sample counts, independently of the sampling technique used. In GLS the following stochastic system of the equations in **T** is considered:

$$\hat{\mathbf{T}} = \mathbf{T} + \eta \qquad (9)$$

$$\hat{\mathbf{V}} = \mathbf{V}(\mathbf{T}) + \varepsilon \quad (10)$$

where η is the sampling error with a variance-covariance matrix **Z**, and ε is the traffic count error with dispersion matrix **W**. Thus the GLS estimator T_{GLS} of **T** is obtained by solving:

$$T_{GLS} = \arg \min_{T \in S} (\hat{\mathbf{T}} - \mathbf{T})'\mathbf{Z}^{-1}(\hat{\mathbf{T}} - \mathbf{T}) + (\hat{\mathbf{V}} - \mathbf{V}(\mathbf{T}))'\mathbf{W}^{-1}(\hat{\mathbf{V}} - \mathbf{V}(\mathbf{T})) \qquad (11)$$

S is the feasible set of **T**. There the dispersion matrix **Z** depends on the sampling estimator adopted.

BI method has also been applied in various transportation planning problems where prior beliefs are combined with the observations to produce the posterior beliefs. For further application and details of BI, refer Dey and Fricker (1994) paper where BI has been used for updating trip generation data. The BI approach considers the target ODM as a prior probability function Pr(**T**) of the estimated ODM **T**. If the observed traffic counts are considered as another source of information with a probability $L(\hat{\mathbf{V}} / \mathbf{T})$, then Bayes theorem provides a method for combining the two sources of information. The posterior probability $f(\mathbf{T} / \hat{\mathbf{V}})$ of observing **T** conditional on the observed traffic counts is obtained as:

$$f(\mathbf{T} / \hat{\mathbf{V}}) \approx L(\hat{\mathbf{V}} / \mathbf{T}).\Pr(\mathbf{T}) \qquad (12)$$

2.3.1 Models without Considering Congestion Effects

The ODM estimation methods developed for networks with no congestion effects basically assume the route choice proportions and are independently determined outside the estimation process. For such networks Bayesians inference based approach has been first introduced by Maher (1983) for the ODM estimation. GLS estimator based approach has been studied by Cascetta (1984), Bell (1991a) etc. Bell (1991a) solved the GLS problem subject to inequality constraints and presented a simple algorithm but its application on real network has not been found in literature. Bierlaire and Toint (1995) proposed an ODM estimation method, called the Matrix Estimation Using Structure Explicitly (MEUSE), considering the information obtained from the parking surveys. Maximum likelihood based model is studied by Spiess (1987), Cascetta and Nguyen (1988), Hazelton (2000) etc. Two classical inference approaches; the ML and the GLS methods are derived and contrasted to the Bayesian method by Cascetta and Nguyen (1988). In all these studies, the link choice proportions used are estimated from the traffic assignment (TA) model and are assumed to be constant which may not estimate a dependable matrix. In consequence, Lo et al (1996) incorporated the randomness of the link choice proportions and discussed both Maximum likelihood and Bayesian approach for the estimation of the ODM by testing with a small network. Liu and Fricker (1996) introduced a stochastic logit model for calculating driver's route choice behaviour but it has certain drawback like all the link counts are considered to be known (further refer Yang et al, 2001). Lo et al (1999) extended the approach of Lo et al (1996) and developed a coordinate descent method using the partial linearization algorithm (PLA) for obtaining the optimum estimates and solved the new approach for large networks. Hazelton (2000) tested the performance of both multivariate normal (MVN) likelihood approximation and GLS techniques and found that the MVN method performed better. Hazelton (2001) studied the fundamental theoretical aspects of the ODM problem based on BI, defining the estimation, prediction and reconstruction problems as; a 'reconstruction' problem estimates the actual ODM occurring during the observation period, an expected number of O-D trips is obtained in 'estimation' problem and future O-D trips are obtained in 'prediction' problem. It has been shown that the estimation and reconstruction problems are different. There are some more studies (Hazelton, 2003; Van Aerde et al, 2003; Li, 2005, etc.) based on statistical approaches.

2.3.2 Models Considering Congestion Effects

Some authors included congestion effects in the estimation problem in which the dependence of the link costs, path choices and assignment fractions on link flows is considered. Equilibrium assignment approaches are particularly adopted for such cases. Nguyen (1977) first introduced the equilibrium based approach to estimate ODM through a mathematical programme (refer Leblanc and Farhangian, 1982). Also Yang et al (1994), Cascetta and Posterino (2001) and Yang et al (2001) solved the trip matrix estimation problem including congestion effects by considering different TA models. Cascetta and Posterino (2001) considered SUE assignment as a fixed-point problem whereas Yang et al (2001) considered the same problem of Liu and Fricker (1996) (developed without considering congestion effects) and proposed a non-linear optimization model (considering a weighted least square estimate) for the simultaneous estimation of the ODM and travel-cost coefficient based on the logit-based SUE model. For solving this non-convex optimization problem a successive quadratic programming (SQP) method (which is a descent-feasible direction algorithm solving KKT solution) has been used. Further, Lo and Chan (2003) with SUE principle (multinomial logit model) estimated both the disperion parameter θ in multinomial logit model and the trip matrix simultaneously using Quasi-Newton method. The objective function considered is a likelihood function. Most of the above developed models are found to be tested using small networks. And the models do not assure their applicability for large real size congested networks.

Combined Distribution and Assignment (CDA) Based Problem

CDA is a network equilibrium based approach combining distribution and assignment problem. The model by Erlander et al (1979) for the simultaneous prediction of the flows and the demands is modified and used for matrix estimation by some researchers. Fisk and Boyce (1983) introduced the link count data in calibrating the dispersion parameter in Erlander et al (1979) model. The observed link flows are assumed to be available for all the network links. Fisk (1989) examined that the problem of Fisk and Boyce (1983) is simplest to solve and recommended to use for the problems where link cost functions are separable. Kawakami and Hirobata (1992) extended the Fisk's model by including the mode choice behaviour and proposed an optimization problem in the form of a combined distribution, modal split, and TA method. An entropy constraint condition with respect to the traffic mode choice behaviour has been considered. The objective function is an entropy function for the traffic distribution. The constraints include a Beckmann-type user equilibrium for all traffic modes and the entropy constraint with respect to the traffic modes. An iterative convergent algorithm has been proposed and is applied for a road network in Nagoya (Japan) considering two categories of automobiles; large-sized trucks and buses, and small trucks and cars. The network consists of 16 zones, 154 nodes and 240 links. But while applying it to the road network, the estimated values found to be deviated from the observed value, might be because of generalized cost function.

Bi-level Programming Approach

Bi-level programming approach has been used for the problem of ODM estimation in case of congested network. In this approach the upper-level problem is the trip matrix estimation problem and the lower-level problem represents a network equilibrium assignment problem. Genetic Algorithm (GA, a probabilistic global searching method) can also be seen introduced to solve the bi-level programming models.

Spiess in 1990 formulated a convex minimization, gradient based model (method of steepest descent) which can be applicable for large real networks. This ODM adjustment problem, allows adjustments between the traffic flows from the assignment algorithm and counts. It has been implemented using EMME/2 transportation planning software. The convex minimization problem is the distance between the observed and the assigned volumes shown below,

$$\text{Min } F(T) = \sum_{a \in A} (v_a - \hat{v}_a)^2 \tag{13}$$

subject to

$$\mathbf{V} = \mathbf{P}(\mathbf{T}) \tag{14}$$

where v_a and $\hat{v}_a$ are the estimated and the counted flows for the link a ($a \in A$) respectively. For the convexity of the problem a set of non-decreasing link cost functions on all links of the network are assumed to be obtained from equilibrium assignment externally. The approach estimates an approximate gradient of the objective function with respect to the O-D demands. The Spiess model can estimate non-zero values for O-D pairs with zero trips considered initially, which can be a matter of concern for the planners who wants to preserve the structure of the target ODM. Thus, Doblas and Benitez (2005) further extended the above Spiess (1990) study for cases when one need to preserve the structure of the target ODM. An efficient approach has been proposed to update trip matrix for large study areas with minimum stored information. A new formulation in addition to the above formulation (eqns. 13 and 14) has been proposed which incorporates constraints based on trip generation, trip attraction, total trips in the network and trips between the O-D pairs with their upper and the lower bounds decided according to the planner. The problem has been transformed to an augmented lagrangian function and optimized using Frank-Wolfe algorithm. The approach has been tested on a real large size network.

Yang et al (1992) through a heuristic technique solved the integrated problem of GLS technique/entropy maximization with equilibrium TA model in the form of a convex bi-level optimization problem. In this study, the GLS estimate needs a matrix inversion of the size equal to the number of O-D pairs. Such an inversion requirement can have computational problem for large networks. Yang (1995) extended the bi-level programming problem by including link flow interaction and developed a model with heuristic algorithms. Kim et al (2001) discussed the problem of the bi-level models developed by Yang et al (1992). The author discussed the problem of dependency on the target ODM of the bi-level model and proposed an alternative model using GA. The upper level is solved by a combined solution method with GA for ODM estimation. The mathematical proof of the solution being optimal has not been discussed. To circumvent the difficulties with the non-differentiabilities of upper level problem Codina et al (2006) presented two alternative algorithms for solving bi-level problems and tested on small networks. First is a hybrid scheme proximal point-steepest descent method (for upper level sub-problem) and considering elastic demand TA problem and second a simple bi-level program considering fixed point mapping using linear TA problem. An iterative column generation algorithm which converges into a local minimum under the concept of continuity in path cost function is formulated by Garcia and Verastegui (2008) and applied on small networks. Recently, Lundgren and Peterson (2008) developed a heuristic bi-level problem solving it by a descent algorithm and demonstrated the algorithm using a large size network for the city of Stockholm, with 964 links and 1642 O-D pairs.

Path Flow Estimation (PFE) based Models

Recently, models based on path flow estimator which determines ODM according to the solutions of path flows have been adopted. It is a single level mathematical program in which the interdependency between O-D trip table and route choice proportion (congestion effect) is taken into account. The core component of PFE is a logit based path choice model, which interacts with link cost functions to produce a stochastic user equilibrium traffic pattern. Sherali et al (1994) proposed a linear path flow model employing user equilibrium based solution for reproducing the observed link flows (known for all links). The procedure utilizes shortest path network flow programming sub-problem and a column generation technique is applied to generate the paths out of alternate paths that will determine the optimal solution to the linear programming model. To avoid the path enumeration required in the model proposed by Sherali et al (1994), Nie and Lee (2002) solved the linear programming model considering an exogenous K-shortest-path for determining the equilibrium path flow pattern. Nie et al (2005) further extended the decoupled path flow estimator by Nie and Lee (2002) considering the generalized least squares framework in aspect of the limitations of the linear programming structure. Sherali et al (2003) enhanced the linear programming model of Sherali et al (1994) for situations where only a partial set of link volumes are available. This introduces nonlinear cost function because of the dependence of the link travel cost on link volumes and a fixed point solution is tried. Further tests using larger and real-size networks are required with these PFE based models for better assessment and efficiency checking of these models.

2.4 Multi-Vehicle ODM Estimation

In spite of single-vehicle information (aggregated information) some authors included additional information obtained from the survey of the individual vehicles types. The multiclass O-D estimation leads to eliminate the internal inconsistency of traffic flows among different vehicle classes. Baek et al (2004) used multiple-vehicle information for ODM estimation from traffic counts. The multi-vehicle ODM estimation method is given as:

$$\text{Min } F(t_w) = \frac{1}{2}\sum_{c}\sum_{a\in A}(v_a^c - \hat{v}_a^c)^2 + \gamma\frac{1}{2}\sum_{c}\sum_{w\in W}(t_w^c - \hat{t}_w^c)^2 \tag{15}$$

subject to

$$\mathbf{T} \geq 0 \tag{16a}$$
$$\mathbf{V} = \mathbf{P}(\mathbf{T}) \tag{16b}$$

where $\mathbf{T}$ is the ODM with elements t_w^c as the trips for the vehicle type c between O-D pairs w, $\hat{t}_w^c$ is the target O-D trips of vehicle type c between O-D pairs w, $\mathbf{V}$ as the vector with elements v_a^c as traffic volume of the vehicle type c on the link a, $\hat{v}_a^c$ is the observed volumes of vehicle type c on the link a, γ is the parameter reflecting the reliability of the target ODM and $\mathbf{P(T)}$ is the multi-vehicle traffic assignment map. For solving the non-convex problem, genetic algorithm has been used. The algorithm has been demonstrated only using a small network of 16 links with 8 O-D pairs and 9 nodes. It required more computing time due to involvement of genetic algorithm. An entropy maximization based ODM estimation of all vehicle classes with multiclass traffic counts observed directly from the network has been proposed by Wong et al (2005). A set of multiplication factors for adjusting the demand matrix is used which may change because of changes in land use pattern or network configuration. The estimated matrix largely depends on the reliability of the input data.

2.5 Fuzzy Based Approach

Many non-linear real-life problems in the field of transportation planning and traffic control have been solved by fuzzy logic systems. Its application can be found in ODM estimation problem too. Fuzzy logic is used to model situations in which user making decisions are so complex that it is very hard to develop a mathematical model. Xu and Chan (1993a, b) estimated ODM with fuzzy weights (refer Teodorovic, 1999). Reddy and Chakroborty (1998) proposed a bi-level optimization approach on a multipath, fuzzy inference based flow dependent assignment algorithm for generating the route choice proportions which along with the observed link flows are then used in the ODM estimation. The proposed technique has been observed to give good O-D estimates irrespective of the flow pattern. Still not much study has been carried with fuzzy logic applications. The real-size network application of fuzzy logic systems should be further studied to ensure its efficiency.

2.6 Neural Network Based Approach

Considering the highly dynamic, large scale, complex and uncertain nature of many transportation systems, neural networks are recently considered as an efficient tool in solving numerous transportation problems. Gong (1998) developed the Hopefield Neural Network (HNN) model to estimate the urban ODM from link volumes, to promote the solving speed and the precision. Though there are quite a few studies carried out by researchers using Artificial Neural Networks (ANNs), but their application still need to be studied for real networks.

3. Dynamic/Time-Dependent ODM Estimation from Traffic Counts

Dynamic/time-dependent ODM estimation is crucial estimation problem for online/offline applications such as route guidance, dynamic traffic assignment (DTA) and freeway corridor control. Also used in various microscopic simulation based studies. Dynamic/time-dependent ODM estimation got much attention due to the development of Intelligent Transportation Systems (ITS). For dynamic (online) ODM estimation traffic counts are observed for short time say 15 min interval. Time-dependent (offline) ODM estimation considers time-series of traffic counts. Mostly the developed algorithms for both dynamic and time-dependent ODM estimation are applied for "closed" networks like intersections/interchanges and small freeways. The dynamic formulation of the matrix estimation process can be expressed as

$$f_{sh} = \sum_{k} \sum_{w} p_{sh}^{wk} t_{wk} \tag{17}$$

where f_{sh} is the flow crossing sensor s in time interval h, t_{wk} is the flow between O-D pair w that departed its origin during time interval k and p_{sh}^{wk} is an assignment parameter reflecting the proportion of the demand t_{wk} crossing sensor s in time interval h. The difference of dynamic formulation from static estimations lies in parameters k and h. And very few of them are developed considering large size network application. Classical techniques of dealing these systems are State-Space Modeling and Kalman filter which are discussed below.

To develop a **State-Space model**, a state is defined first. Once a state is defined, transition and measurement equations are specified. In dynamic systems, transition equations describe the evaluation of the state over time. Measurement equations on the other hand relate the unknown state to their observed indicators.

Measurement Equation: $y_h = C_h x_h + e_h$ (18)

Transition Equation: $x_{h+1} = D_h x_h + z_h$ (19)

Where x_h is the vector that represents the latent "true state" of the system during interval h, y_h is a vector of observations made in interval h, C_h and D_h are known matrices and, e_h and z_h are vectors of random errors.

Kalman filtering is an optimal state estimation process applied to a dynamic system that gives a linear unbiased and minimum error variance recursive algorithm. ODM estimation problem is formulated with a linear mapping from the state variable to the measurement variable. The formulation is stochastic which allows for measurement and state errors, and the initial state could be described stochastically. The transition equation for Kalman Filter is of the form

$$x_h = \phi_{h-1} x_{h-1} + \Gamma_{h-1} u_{h-1} + \Lambda_{h-1} q_{h-1} \tag{20}$$

where x_h is the state vector to be determined at the time h, ϕ_h is the matrix of autoregressive coefficients, u_h is a deterministic input, Γ_h the control gain, Λ_h the disturbance gain, and q_h is the disturbance input. The measurement equation is given by

$$x_h = H_{h-1} x_{h-1} + n_h \tag{21}$$

where n_h is the error (noise) term with zero mean and positive definite covariance matrix and the matrix H_h relates the state to the measurement of x_h. The filter computes the estimate of the state while minimizing the spread of the estimate error probability distribution function.

3.1 Dynamic or On-line ODM Estimation

The dynamic ODM estimation also stated as real-time ODM estimation, characterizes the time-variant trips between each origin and destinations. It is an essential input for DTA models and used for on-line identification of travel pattern for traffic controls systems. The dynamic models can be characterized as non-DTA based models and DTA based models. The non-DTA based models like Kalman filter based models and Parameter optimization based models are basically applied for small networks like intersections, freeways etc where entry and exit flow information are available.

A state-space model with unknown O-D flows as the state vector has been first introduced by Okutani in 1987 (refer Kachroo et al, 1997). Ashok and Ben-Akiva (1993) (refer Sherali and Park, 2001) proposed a Kalman filtering approach to dynamically update an ODM. The O-D flow deviations from the prior estimates based on historical data are considered (for capturing the structural information) as the state-vectors in order to overcome inadequacy of autoregressive specification for O-D flows in Okutani's approach. Kachroo et al (1997) studied the applicability of Kalman filtering approaches for network ODM estimation from link traffic counts to explore the characteristics of the error terms in the underlying dynamic process of the O-D departures. The inconsistencies between the observed O-D flow patterns and Kalman Filter modeling assumptions is analyzed. It has been concluded that the noise is not a white Gaussian sequence. Ashok and Ben-Akiva (2000) further extended Ashok and Ben-Akiva's (1993) approach and presented a new formulation based on deviations of departure rates from origin and destination shares over time instead of destination flows.

Cremer and Keller (1987), Nihan and Davis (1987), Bell (1991b), Wu and Chang (1996), determined split parameters (averaged values) for input-output network relationships that is applicable for traffic flows at intersections or small freeway segments. Sherali et al (1997) developed a constrained optimization algorithm but with high computational cost. Li and Moor (1999) also proposed a recursive-based algorithm. The above approaches need all entry and exit information which is somewhat unrealistic. For the situations with incomplete traffic counts at some entrances and exits, Li and Moor (2002) formulated an optimization problem with linear equality constraints and non-negative inequality constraints. Van der Zijpp and Lindveld (2001) formulated a dynamic user optimal departure time and route choice (DUO-D&R) assignment problem which is used to estimate dynamic ODM with preferred departure times. There are some more studies (Van der Zijpp, 1997; Suzuki et al, 2000 etc.) on dynamic ODM estimation for intersections and freeways.

Due to disadvantages of Kalman filtering formulation (needs sufficient data and intensive matrix operations) Wu (1997) developed a real-time ODM updating algorithm based on a balancing method called multiplicative algebraic reconstruction technique (MART of Lamond and Stewart, 1981) considering entropy-maximization model. MART has been revised (RMART) by incorporating a normalization scheme, without giving any theoretical explanation of doing so. A diagonal searching technique is considered for improving the convergence speed. A numerical test has been carried out for checking the efficiency of the RMART with artificially generated database from some computer simulated problems. Zhou et al (2003) included the historical static information and ITS real-time link-level information to determine the dynamic O-D demand. The variation in day-to-day demand is studied by using multiday traffic counts. Nie and Zhang (2008) gives a brief review on dynamic ODM estimation algorithms and formulated a variational inequality problem determining the dynamic traffic assignment endogenously considering the user response to traffic congestion through a dispersion parameter θ. The problem finds out path flows denoted here as $\mathbf{f}$ which satisfy the conditions given in eqn. (22) which are transformed into a variational inequality formulation.

$$\begin{cases} f_w^{jh}(c_w^{jh} - \theta d_w^{jh}) = 0 \\ c_w^{jh} \geq \theta d_w^{jh}, f_w^{jh} \geq 0 \end{cases} \tag{22}$$

Here d_w^{jh} denotes the path derivatives between O-D pairs w for assignment interval h and c_w^{jh}, the path travel time of path j during assignment interval h. A solution for the variational inequality formulation has been proposed using a space-time expanded network (STEN, refer the journal paper for further details) to generate paths. A column generation algorithm has been proposed to solve the dynamic matrix estimation problem iterating the two sub-problems; generating paths to construct a restricted VI problem and to find an optimum solution of the restricted problem. The results depend on the initial path flows and the convergence issues still need to be studied.

3.2 Time-Dependent or Off-Line ODM Estimation

This estimation is carried out off-line, given a time-series of link counts, travel times and prior O-D information. Chang and Tao (1999) proposed a model integrating the link constraints and intersection turning flows from available DTA model, for determining the time-varying O-D trips for intersections applying Kalman filtering approach. A parametric optimization approach for off-line processing purpose is developed by Sherali and Park (2001). The proposed model seeks path flows that compromise between the least cost O-D paths and those that provide a match for the observed link flows. But with the increase in O-D paths, this model seems to be difficult to solve. Ashok and Ben-Akiva (2002) introduced the stochasticity of the assignment matrix in estimating the time-dependent O-D flows from link volumes. A GLS based solution has been studied for minimizing the error criteria for each interval and is evaluated for a case study. Tsekeris and Stathopoulos (2003) coupled multi-proportional procedure (MPP) of Murchland (1977) and multiplicative algebraic reconstruction technique (MART) of Gordon et al (1970) respectively, with a quasi-DTA model and estimated dynamic trip departure rates and ODM over a series of successive time interval. Combining the algorithms a doubly iterative matrix adjustment procedure (DIMAP) has been proposed to obtain a consistency between the trip departure rates from each origin zone and the observed link flows. The simulation based quasi-dynamic model used is based on the instantaneous link travel cost definition. The effect of congestion is incorporated and the performance of the algorithms is studied using a real network. Estimation is carried out separately using simulated link flows and real traffic flows. The DIMAP has been compared with MART and RMART (Wu, 1997). While considering simulated link flows the DIMAP algorithm found to perform better than the case using real traffic flows. Recently, BI based parsimonious parameterized model for estimating time varying ODM with traffic counts (collected on daily basis) has been recommended by Hazelton (2008).

4. The Measure of reliability of the estimated ODM

Statistical measures

The outcome of a situation which is difficult to predict is generally measured through some statistical measures. Likewise in ODM estimation problem some statistical measures are used by the authors to verify the performance of their proposed algorithms. The statistical measures only can measure the closeness of the estimated values (trips and link flows) and their true values, if known. Following are the statistical measures mostly adopted:

Relative error (RE) %:

$$\text{RE } (\%) = \sqrt{\frac{1}{2}\sum_{w\in W}\left(\frac{t_w^* - t_w}{t_w}\right)^2} \times 100\% \tag{23}$$

Total Demand Deviation (TDD) %:

$$\text{TDD } (\%) = \frac{\left|\sum_{w\in W} t_w^* - \sum_{w\in W} t_w\right|}{\sum_{w\in W} t_w} \times 100\% \tag{24}$$

Mean absolute error (MAE) %:

$$\text{MAE } (\%) = \frac{\sum_{w\in W}\left|t_w^* - t_w\right|}{N} \times 100\% \tag{25}$$

Root Mean Square Error (RMSE) %:

$$\text{RMSE } (\%) = \frac{\sqrt{\frac{1}{N}\sum_{w\in W}\left(t_w^* - t_w\right)^2}}{\frac{1}{N}\sum_{w\in W} t_w} \times 100\% \tag{26}$$

where t_w^* is the true ODM and N is the number of O-D pairs. The TDD gives the quality of the estimated ODM. The RMSE percent error quantifies the total percentage error of the estimate. The mean percent error indicates the existence of consistent under-or-over-prediction in the estimate. Smaller values of these measures will indicate the high quality of the estimated ODM. But in situations when the true values are not known these statistical measures cannot be used.

Maximum Possible Relative Error (MPRE)

Yang et al (1991) through a simple quadratic programming problem introduced a concept of maximum possible relative error (MPRE) which represents the maximum possible relative deviation of the estimated ODM from the true one and can be used only when the route choice proportions are correctly specified and the traffic counts are error free. The reliability of the estimated ODM from traffic counts is measured as,

$$\mathrm{Re}(t) = \frac{1}{1+E}, \quad E \geq 0 \tag{27}$$

where E is the measure of the error (average relative deviation) in the estimated ODM depending upon the relative deviation of the estimated O-D flows from the true ones for the O-D pairs.

Travel Demand Scale (TDS)

Based on statistical analysis the quality measure of both static and dynamic ODM models is proposed by Bierlaire (2002) by means of TDS which is independent of the estimation method and a priori matrix (say obtained from a previous study), but depends upon the network topology and route choice assumptions. The Travel Demand Scale is computed as,

$$TDS = \varphi_{\max} - \varphi_{\min} \tag{28}$$

where

$$\varphi_{\min} = \min_{t} \mathbf{T}^{'}\mathbf{e} \tag{29}$$

and

$$\varphi_{\max} = \max_{t} \mathbf{T}^{'}\mathbf{e} \tag{30}$$

subject to

$$\mathbf{PT} = \hat{v}_a \tag{31a}$$

$$\mathbf{T} \geq 0 \tag{31b}$$

where $\varphi_{\min}$ and $\varphi_{\max}$ are minimum and maximum total level, $\mathbf{e}$ is the vector only of ones and $\mathbf{P}$ is the vector notation of the assignment matrix corresponding to the links where flow observations are available. The TDA value (for values refer the journal paper) helps to optimize the resources allocated during the surveys by identifying the nature of the additional information required. It finds out the unbounded O-D pairs (O-D pairs not captured by link flow data) so that surveys can be conducted to increase the quality of the corresponding entries in the a priori ODM. Thus it helps to assess the level of investment necessary to collect data and build the a priori matrix. It is recommended to use in addition to the statistical measures.

5. Traffic Counting Location

For the ODM estimation, traffic counting or sampling survey data collection are carried out where a road or rail route crosses a cordon line and screen lines. The accuracy of the ODM estimated increases with the number of traffic counting stations adopted. But due to resource limitations, it may not be possible. Again, the traffic count at each location has different degree of influence to the ODM estimation. Hence it is necessary to determine the optimum number of counting stations and their locations on the network to intercept maximum O-D pairs. Yang and Zhou (1998) introduced four basic rules of locating traffic counting points based on the maximal possible relative error (MPRE) concept proposed by Yang et al (1991). Based on the *O-D covering rule* stated by Yang et al (1991), following are the rules proposed by Yang and Zhou (1998);
(1) *O-D covering rule*: At least one traffic counting point on the network must be located for observing trips between any O-D pair.
(2) *Maximal flow fraction rule*: The traffic counting points on a network must be located at the links between a particular O-D pair such that flow fraction ϕ_{aw} is as large as possible.

$$\phi_{aw} = \frac{p_{aw} t_w}{v_a} \tag{32}$$

where v_a is the link flow, $a \in A$ between the O-D pair $w \in W$ and p_{aw} is the proportion of trips used by link $a \in A$ for each O-D pair w.
(3) *Maximal flow-intercepting rule*: From the links to be observed, the chosen links should intercept as many flows as possible.
(4) *Link independence rule*: The traffic counts on all chosen links should be linearly independent.

For determining traffic count locations, the O-D covering rule and the link independence rule are treated as constraints and maximal flow fraction rule and maximal flow-intercepting rule are incorporated in objective function. Yim and Lam (1998) presented rules of maximum net O-D captured and maximum total O-D captured, and formulated in a linear programming model to determine the locations for ODM estimation. Bianco et al (2001) through proposed heuristic method solved the sensor location problem by proposing a two-stage procedure; determining the minimum number and location of counting points with known turning probabilities (assumed) and estimating the ODM with the resulting traffic flows. Yang et al (2003) studied the scheduling installation of traffic counting stations for long duration planning purpose and a Genetic Algorithm based sequential greedy algorithm has been proposed. Kim et al (2003) formulated two models, link-based and road-based model, to determine the location of the counting points on the link which minimizes the total cost. Three solution algorithm: Greedy Adding (GA) algorithm, Greedy Adding and Substituting

(GAS) algorithm and Branch and Bound (BB) algorithm are proposed and tested for a simple artificial network. Ehlert et al (2006) extended the problem of Chung (2001) (refer Ehlert et al, 2006) to optimize the additional counting locations assuming that some detectors are already installed increasing the O-D coverage. A software tool is developed based on mixed integer problem (MIP). With the budget restrictions for practical problems it is stated that the OD covering rule cannot always be satisfied i.e. the ODM estimation error measure MPRE cannot be applied. Gan et al (2005) studied both the traffic counting location and the error estimation measure in ODM estimation problem taking into consideration the route choice assumptions made in the TA models and the levels of traffic congestion on the networks. It has been noted that both MPRE and TDS measures are closely related to traffic counting locations. When the O-D covering rule is not satisfied by the link counting locations both the measures become infinite. Yang et al (2006) formulated an integer linear programming (ILP) problem using shortest path-based column generation procedure and branch-and-bound technique in determining screen line based traffic counting locations.

6. Conclusions

The basic goal of this paper is to explore the studies on one of the most promising topics for research which is the estimation of ODM using traffic counts on a set of links. The review shows the intricacy of the ODM estimation problem using traffic counts, the reason being the under-specification of the trip matrix estimation problem with less link count information than the number of unknowns. Till date both static and dynamic ODM estimation problems have been investigated by many researchers and models have been developed with different problem formulations, using different route choice decisions process and various solution algorithms.

The statistical approaches (ML, GLS and BI) have been mostly adopted by the researchers to solve the static problem considering congestion and without considering congestion effects. Both bi-level programming approach and simultaneous (single-level) optimization approaches have been studied in literature. Also Path flow estimation based algorithms are proposed assuming that all link costs are available, which may not be available in practical situations. Very few authors used fuzzy logic and neural network based approaches and their applicability need to be analyzed further. The review shows that most of the algorithms developed for static ODM have its own advantages and disadvantages and are implemented on small networks. However the most important consideration required is the applicability of the algorithms for real world networks which are large in size and highly congested. It is surprising to see a few realistic approaches in the literatures focused on large size network applications. Thus the developed algorithms need to be checked regarding their practical applicability for large size real networks.

The static ODM determined for long-time transportation planning and design purpose is easier to estimate as compared to dynamic ODM used for traffic management and operations for large networks because availability of real-time traffic information for all the O-D pairs required for dynamic trip matrix estimation is not possible. Compared to static ODM estimation, dynamic estimation based studies are few and mostly for intersections, freeways and small networks; as it is convenient to study the dynamic state of these networks. Some authors tried to extend the study for large networks. But for practical applications dynamic ODM still is not much in use except for performing DTA on small scale networks.

To identify the reliability of the estimated ODM the statistical measures and the MPRE needs real trip values which are not always available in practical cases. The TDS though measures quality for both static and dynamic matrices but it does not serve alone the purpose of measuring the reliability. Some authors gave emphasis on finding the optimum number and location of the traffic counting points. Quite some rules have been proposed in the literature which can help in obtaining the optimum traffic counting locations and in receiving more information of travel pattern between O-D pairs.

As indicated earlier more studies and checking (mainly regarding computational difficulties) of the developed algorithms still need to be carried out especially for the case of its application for planning and designing purpose done for large size networks.

Acknowledgement

We wish to thank the referee for his useful comments.

References

Ashok, K. and Ben-Akiva, M. E. (1993) "Dynamic origin-destination matrix estimation and prediction for real-time traffic management systems", In: Daganzo, C. F. (Ed.), *Transportation and Traffic Theory*, Elsevier, Amsterdam.

Ashok, K. and Ben-Akiva, M. E. (2000) "Alternative approaches for real-time estimation and prediction of time-dependent origin-destination flows", *Transportation Science* 34: 21-36.

Ashok, K. and Ben-Akiva, M. E. (2002) "Estimation and prediction of time-dependent origin-destination flows with a stochastic mapping to path flows and link flows", *Transportation Science* 36: 184-198.

Baek, S., Kim, H. and Lim, Y. (2004) "Multiple-vehicle origin-destination matrix estimation from traffic counts using genetic algorithm", *Journal of Transportation Engineering*, ASCE, May/June: 339-347.

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

Bell, M. (1991b) "The real time estimation of origin-destination flows in the presence of platoon dispersion", *Transportation Research* 25B: 115-125.

Bianco, L., Confessore, G. and Reverberi, P. (2001) "A network based model for traffic sensor location with implications on o/d matrix estimates", *Transportation Science* 35(1) : 49-60.

Bierlaire, M. and Toint, Ph. L. (1995) "Meuse: an origin-destination matrix estimator that exploits structure", *Transportation Research*, *Part B: Methodological* 29: 47-60.

Bierlaire, M. (2002) "The total demand scale: a new measure of quality for static and dynamic origin-destination trip tables", *Transportation Research*, *Part B: Methodological* 36: 837-850.

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

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

Cascetta, E. and Postorino, M. N. (2001) "Fixed point approaches to the estimation of O/D matrices using traffic counts on congested networks", *Transportation Science* 35: 134-147.

Chang, G. and Tao, X. (1999) "An integrated model for estimating time varying network origin-destination distributions", *Transportation Research, Part A: Policy and Practice* 33: 381-399.

Chung, I. H. (2001) An optimum sampling framework for estimating trip matrices from day-to-day traffic counts, *Ph.D. Thesis*, University of Leeds.

Codina, E., Garcia, R. and Marin, A. (2006) "New algorithm alternatives for the O-D matrix adjustment problem on traffic networks", *European Journal of Operational Research* 175: 1484-1500.

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

Dey, S. S. and Fricker, J. D. (1994) "Bayesian updating of trip generation data: combining national trip generation rates with local data", *Transportation* 21(4): 393-403.

Doblas, J. and Benitez, F. G. (2005) "An approach to estimating and updating origin-destination matrices based upon traffic counts preserving the prior structure of a survey matrix", *Transportation Research*, *Part B: Methodological* 39: 565-591.

Ehlert, M., Bell, M. G. H. and Grosso, S. (2006) "The optimization of traffic count locations in road networks", *Transportation Research*, *Part B: Methodological* 40: 460-479.

Erlander, S., Nguyen S. and Stewart N. (1979) "On the calibration of the combined distribution/assignment model", *Transportation Research*, *Part B: Methodological* 13: 259-267.

Fisk, C. S. and Boyce, D. E. (1983) "A note on trip matrix estimation from link traffic count data", *Transportation Research*, *Part B: Methodological* 17: 245-250.

Fisk, C. S. (1989) "Trip matrix estimation from link traffic counts: the congested network case", *Transportation Research*, *Part B: Methodological* 23: 331-356.

Gan, L., Yang, H. and Wong, S. C. (2005) "Traffic counting location and error bound in origin-destination matrix estimation problems", *Journal of Transportation Engineering*, ASCE/July: 524-534.

Garcia, R. and Verastegui, D. (2008) "A column generation algorithm for the estimation of origin-destination matrices in congested traffic networks", *European Journal of Operational Research* 184: 860-878.

Gong, Z. (1998) "Estimating the urban o-d matrix: a neural network approach," *European Journal of Operational Research* 106: 108-115.

Gordon, R., Bender, R. and Hermann, G. T. (1970) "Algebraic Reconstruction Techniques (ART) for three-dimensional electron microscopy and X-ray photography", *Journal of Theoretical Biology*, Vol. 29: 471-481.

Hazelton, M. L. (2000) "Estimation of origin–destination matrices from link flows on uncongested networks", *Transportation Research*, *Part B: Methodological* 34: 549-566.

Hazelton, M. L. (2001) "Inference for origin-destination matrices: estimation, prediction and reconstruction", *Transportation Research*, *Part B: Methodological* 35: 667-676.

Hazelton, M. L. (2003) "Some comments on origin-destination matrix estimation", *Transportation Research*, *Part A: Policy and Practice* 37: 811-822.

Hazelton, M. L. (2008) "Statistical inference for time varying origin-destination matrices", *Transportation Research*, *Part B: Methodological* 42: 542-552.

Hogberg, P. (1976) "Estimation of parameters in models for traffic prediction: a non-linear regression approach", *Transportation Research*, Vol. 10: 263-265.

Kachroo, P., Ozbay, K. and Narayanan, A. (1997) "Investigating the use of Kalman filtering for dynamic origin-destination trip table estimation", Southeastcon'97, 'Engineering NewCentury', Proceedings, *IEEE*: 138-142.

Kawakami, S. and Hirobata, Y. (1992) "Estimation of origin-destination matrices from link traffic counts considering the interaction of the traffic modes", *Regional Science Association International*, Vol. 71, No. 2: 139-151.

Kim, H., Beak, S. and Lim, Y. (2001) "Origin-destination matrices estimated with a genetic algorithm from link traffic counts", *Transportation Research Record* 1771, Transportation Research Board, Washington, D.C.: 156-163.

Kim, H. J., Chung, H. I. and Chung, S. Y. (2003) "Selection of the optimal traffic counting locations for estimating origin-destination trip matrix", *Journal of the Eastern Asia Society for Transportation Studies*, Vol. 5: 1353- 1365.

Lamond, B. and Stewart, N. F. (1981) "Bregman's balancing method", *Transportation Research*, *Part B: Methodological* 15: 239-248.

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

Li, B. and Moor, B. D. (2002) "Dynamic identification of origin-destination matrices in the presence of incomplete observations", *Transportation Research*, *Part B: Methodological* 36: 37-57.

Li, B. (2005) "Bayesian inference for O-D matrices of transport networks using the EM algorithm", *Technometrics* 47(4): 399-408.

Liu, S. and Fricker, J. D. (1996) "Estimation of a trip table and the Θ parameter in a stochastic network", *Transportation Research*, *Part A: Policy and Practice* 30: 287-305.

Lo, H., Zhang, N. and Lam, W. (1996) "Estimation of an origin–destination matrix with random link choice proportions: a statistical approach", *Transportation Research*, *Part B: Methodological* 30: 309-324.

Lo, H. P., Zhang, N. and Lam, W. H. K. (1999) "Decomposition algorithm for statistical estimation of OD matrix with random link choice proportions from traffic counts", *Transportation Research*, *Part B: Methodological* 33: 369-385.

Lo, H. P. and Chan, C. P. (2003) "Simultaneous estimation of an origin-destination matrix and link choice proportions using traffic counts", *Transportation Research*, *Part A: Policy and Practice* 37: 771-788.

Lundgren, J. T. and Peterson, A. (2008) "A heuristic for the bilevel origin-destination matrix estimation problem", *Transportation Research*, *Part B: Methodological* 42: 339-354.

Maher, M. J. (1983) "Inferences on trip matrices from observations on link volumes: a bayesian statistical approach", *Transportation Research*, *Part B: Methodological* 17: 435-447.

Murchland, J. D. (1977) *The Multi-Proportional Problem*, Research Note JDM-263, University College London, London.

Nguyen, S. (1977) "Estimating an OD matrix from network data: a network equilibrium approach", *Publication No. 87*, Center de recherché sur les transports, Universite de Montreal, Montreal, Quebec H3C 3J7.

Nie, Y. and Lee, D. (2002) "Uncoupled method for equilibrium-based linear path flow estimator for origin-destination trip matrices", *Transportation Research Record* 1783, Transportation Research Board, Washington, D.C.: 72-79.

Nie, Y., Zhang, H. M. and Recker, W. W. (2005) "Inferring origin-destination trip matrices with a decoupled GLS path flow estimator", *Transportation Research*, *Part B: Methodological* 39: 497-518.

Nie, Y. and Zhang, H. M. (2008) "A variational inequality formulation for inferring dynamic origin-destination travel demands", *Transportation Research*, *Part B: Methodological* 42: 635-662.

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

Okutani, I., (1987) *The Kalman filtering approach in some transportation and traffic problems*, International symposium on transportation and traffic theory, N. H. Gartner. and N. H. M. Wilson (eds), Elveiser Science Publishing Company Inc.: 397-416.

Reddy, K. H. and Chakroborty, P. (1998) "A fuzzy inference based assignment algorithm to estimate O-D matrix from link volume counts", *Comput., Environ. and Urban Systems*, Vol. 22, No. 5: 409-423.

Robillard, P. (1975) "Estimating the O-D matrix from observed link volumes", *Transportation Research*, Vol. 9: 123-128.

Sherali, H. D., Sivanandan, R. and Hobeika, A. G. (1994) "A linear programming approach for synthesizing origin-destination trip tables from link traffic volumes", *Transportation Research*, *Part B: Methodological* 28: 213-233.

Sherali, H. D., Arora, N. and Hobeika, A. G. (1997) "Parameter optimization methods for estimating dynamic origin-destination trip-tables", *Transportation Research*, *Part B: Methodological* 31: 141-157.

Sherali, H. D. and Park, T. (2001) "Estimation of dynamic origin-destination trip tables for a general network", *Transportation Research*, *Part B: Methodological* 35: 217-235.

Sherali, H. D., Narayanan, A. and Sivanandan, R. (2003) "Estimation of origin-destination trip-tables based on a partial set of traffic link volumes", *Transportation Research*, *Part B: Methodological* 37: 815-836.

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

Spiess, H. (1990) "A gradient approach for the O-D matrix adjustment problem", EMME/2 Support Center, CH-2558 Aegerten, Switzerland, http://www.spiess.ch/emme2/demadj/demadj.html.

Suzuki, H., Nakatsuji, T., Tanaboriboon, Y. and Takahashi, K. (2000) "Dynamic estimation of origin-destination travel time and flow on a long freeway corridor", *Transportation Research Record* 1739, Transportation Research Board, Washington, D.C.: 67-75.

Tamin, O. Z. and Willumsen, L. G. (1989) "Transport demand model estimation from traffic counts", *Transportation* 16: 3-26.

Tamin, O. Z., Hidayat, H. and Indriastuti, A. K. (2003) "The development of maximum-entropy (ME) and bayesian-inference (BI) estimation methods for calibrating transport demand models based on link volume information", *Proceedings of the Eastern Asia Society for Transportation Studies*, Vol. 4: 630-647.

Teodorovic, D. (1999) "Fuzzy logic systems for transportation engineering: the state of the art", *Transportation Research*, *Part A: Policy and Practice* 33: 337-364.

Tsekeris, T. and Stathopoulos, A. (2003) "Real-time dynamic origin-destination matrix adjustment with simuation and actual link flows in urban networks", *Transportation Research Record* 1857, Transportation Research Board, Washington, D.C.: 117-127.

Van Aerde, M., Rakha, H. and Paramahamsam, H. (2003) "Estimation of origin destination matrices", *Transportation Research Record* 1831, Transportation Research Board, Washington, D.C.: 122-130.

Van der Zijpp, N. J. (1997) "Dynamic origin-destination matrix estimation from traffic counts and automated vehicle information data", *Transportation Research Record* 1607, Transportation Research Board, Washington, D.C.: 87-94.

Van der Zijpp, N. J. and Lindveld, C. D. R. (2001) "Estimation of origin-destination demand for dynamic assignment with simultaneous route and departure time choice", *Transportation Research Record* 1771, Transportation Research Board, Washington, D.C.: 75-82.

Van Zuylen, J. H. (1978) "The information minimizing method: validity and applicability to transport planning", In: *New Developments in Modelling Travel Demand and Urban Systems* (edited by G. R. M. Jansen et al).

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

Van Zuylen, J. H. and Branston, D. M. (1982) "Consistent link flow estimation from counts", *Transportation Research*, *Part B: Methodological* 16: 473-476.

Willumsen, L. G. (1978) "Estimation of O-D matrix from traffic counts: a review", *Working Paper* 99, Institute for Transport Studies, University of Leeds.

Willumsen, L. G. (1981) "Simplified transport models based on traffic counts", *Transportation* 10: 257-278.

Wilson, A. G. (1970) *Entropy in Urban and Regional Modeling*, Methuen, Inc., New York.

Wong, S. C., Tong, C. O., Wong, K. I., Lam, W. H. K., Lo, H. K., Yang, H. and Lo, H. P. (2005) "Estimation of multiclass origin-destination matrices from traffic counts", *Journal of Urban Planning and Development*, ASCE, March: 19-29.

Wu, J. and Chang, G. (1996) "Estimation of time-varying origin-destination distributions with dynamic screenline flow", *Transportation Research*, *Part B: Methodological* 30: 277-290.

Wu, J. (1997) "A real-time origin-destination matrix updating algorithm for on-line applications", *Transportation Research*, *Part B: Methodological* 31: 381-396.

Xu, W. and Chan, Y. (1993a) "Estimating an origin-destination matrix with fuzzy weights. Part 1: Methodology", *Transportation Planning and Technology* 17: 127-144.

Xu, W. and Chan, Y. (1993b) "Estimating an origin-destination matrix with fuzzy weights. Part 2: Case Studies", *Transportation Planning and Technology* 17: 145-164.

Yang, H., Iida Y. and Sasaki T. (1991) "An analysis of the reliability of an origin-destination trip matrix estimated from traffic counts", *Transportation Research*, *Part B: Methodological* 25: 351-363.

Yang, H., Sasaki, T., Iida, Y. and Asakura, Y. (1992) "Estimation of origin-destination matrices from link traffic counts on congested networks", *Transportation Research*, *Part B: Methodological* 26: 417-433.

Yang, H., Sasaki, T. and Iida, Y. (1994) "The Equilibrium-based origin-destination matrix estimation problem", *Transportation Research*, *Part B: Methodological* 28: 23-33.

Yang, H. (1995) "Heuristic algorithms for the bi-level origin-destination matrix estimation problem", *Transportation Research*, *Part B: Methodological* 29: 231-242.

Yang, H. and Zhou, J. (1998) "Optimal traffic counting locations for origin-destination matrix estimation", *Transportation Research*, *Part B: Methodological* 32: 109-126.

Yang, H., Meng, Q. and 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: 107-123.

Yang, C., Chootinan, P. and Chen, A. (2003) "Traffic counting location planning using genetic algorithm", *Journal of the Eastern Asia Society for Transportation Studies*, Vol.5: 898-913.

Yang, H., Yang, C. and Gan, L. (2006) "Models and algorithms for the screen line-based traffic-counting location problems", *Computers & Operations Research* 33: 836-858.

Yim, P. K. N. and Lam, W. H. K. (1998) "Evaluation of count location selection methods for estimation of O-D matrices", *Journal of Transportation Engineering*, ASCE, July/August: 376-383.

Zhou, X., Qin, X. and Mahamassani, H. S. (2003) "Dynamic origin-destination demand estimation with multiday link traffic counts for planning applications", *Transportation Research Record* 1831, Transportation Research Board, Washington, D.C.: 30-38.

European Transport \ Trasporti Europei n. 49 (2011): 24-37

Quantifying the performance improvement potential of Foliated Transportation Networks

Joakim Kalantari [1*]

[1] *Division of Logistics and Transportation, Chalmers University of Technology, Göteborg, Sweden*

Abstract

The purpose of this paper is to quantify the performance improvement potential of foliated transportation networks (FTN) compared to a traditional direct shipment network (DS) with respect to key performance indicators (KPI). The network models are evaluated through discrete event simulation. Input data and parameters have been drawn from a case company. The concept of FTN is shown to reduce the negative environmental impact of a transportation network through its implementation. This is the first study that quantifies the potential of FTN. Furthermore, the study is based on empirical data.

Keywords: Mixed model transportation; Discrete event simulation; Direct shipment; Hub and spoke; Transportation network performance evaluation.

1. Introduction

In the context of inter-city general cargo freight transportation networks, there are two predominant network structures, namely, direct shipment (DS) and hub-and-spoke (HS) networks (Woxenius, 2007). A DS network is best utilized when the number of nodes in the network is limited, the demand for transportation in every connection is sufficiently high and the primary optimization parameter is time and transport work. On the other hand, a HS setup is preferred in a network with a vast number of nodes, where aggregation of demand is necessary to attain adequate flows and the primary parameter of optimization is resource utilization and coverage (Crainic, 2003).

Aside from the impact of empty haulage on the resource utilization of a transport network (McKinnon and Ge, 2006), the match between the distribution of transport need and network structure and the discrete property of freight transportation also significantly impact the performance of a transportation system (Lumsden et al., 1999). These technical inefficiencies in the road bound general freight transportation systems have significant negative external effects which make efficient use of the physical

* Corresponding author: Joakim Kalantari (joakim.kalantari@chalmers.se)

resources of interest for society as a whole (European Commission, 2003). The negative impacts referred to include pollution (e.g., NO_X, HC, CO_2, particles etc.), noise pollution, congestion and traffic hazards.

A conceptual model that attempts to address these concerns is Foliated Transportation Networks (FTN). FTN is a hybrid model that aims to improve the physical performance of a transportation system by combining the two predominant network structures, i.e., direct shipment (DS) and hub and spoke (HS) (Bjeljac and Lakobrija, 2004, Persson and Lumsden 2006). It is stipulated that by foliating the two structures (i.e., DS and HS), and by dynamically planning, controlling and optimizing the distribution of goods and resources between the two sub structures, strengths of the individual setup will be amplified at the same time as their weaknesses diminish, resulting in better system performance than any single one on its own (Kalantari and Sternberg, 2009; Persson and Lumsden, 2006; Persson and Waidringer, 2006). However, to date, no study has actually quantified this hypothesized potential.

The purpose of this paper is to quantify the performance improvement potential of foliated transportation networks (FTN) compared to a traditional direct shipment network (DS) with respect to selected key performance indicators (KPI) that are identified to express the physical performance of a transportation network. The KPI include fill rate, number of trucks, transport work and traffic work. This research takes the perspective of the transportation service provider and is based on empirical data from the domestic Swedish general cargo freight transportation service provider.

2. Transportation networks

Transportation networks are commonly described in terms of nodes and links (Lumsden, 2006). The transportation network designs can be divided into two principle categories: direct shipment or hub and spoke (Crainic, 2002; Lumsden, 2006). In reality, one only seldom finds any pure systems (Crainic, 2002) and, also in theory, variations of the same two themes exist (Lumsden, 2006; Woxenius, 2007).

2.1 Direct Shipment vs. Hub and Spoke

In a direct shipment network all nodes are interconnected with direct relations. Direct relation means that the transport is dedicated (Crainic, 2003; Lumsden, 2006). The only nodes involved in the coordination of the transport and the consolidation of goods are the origin and destination nodes. Goods are not consolidated along the way and the transportation is independent of other O/D pairs (Woxenius, 2007). The DS network is best utilized when the number of nodes in the network is limited, the demand for transportation in every connection is sufficient and the primary optimization parameter is time and flexibility.

A DS setup by default affords the shortest time in transit since the goods always travel directly, taking the shortest way and without any additional stops, consolidation operations or handling. A DS setup requires a greater number of resources, e.g., trucks in the system, leads to a lower transportation frequency and its performance is dependent on sufficiently large volumes, i.e., the demand in each relation must match the capacity reasonably well in order to achieve acceptable levels of the resource utilization rate.

Conversely, in a HS setup, all the nodes are only interconnected with a hub resulting in an additional coordination and consolidation/deconsolidation step (Crainic, 2002) which means that deliveries seldom run the shortest way. A hub network creates a larger spatial coverage and high transport frequency for the network since the volume that flows between the O/D pairs does not need to be very large to be included (Woxenius, 2007). Other advantages of the HS setup are a high resource utilization rate, both regarding time and capacity, a lower number of resources in the network and more leveled flows (Lumsden et al., 1999). A HS setup is preferred in a network with a vast number of nodes, where aggregation of demand is necessary to attain adequate flow and the primary parameter of optimization is resource utilization and coverage (Bryan and O'Kelly, 1999).

2.2 Foliated Transportation Networks

Bjeljac and Lakobrija (2004) describe FTN as a DS network where only “full” units are sent directly and all units that are not full are consolidated in the HS layer of the system (Figure 1). In this setup, even though they regard the network as consistent of two different layers in its management, the physical network is the same for both layers.

Lumsden et al. (1999) describe a similar approach where the basic structure is a HS setup that can be shortcut with occasional direct relations when possible, i.e., when there is enough demand to fill a whole unit in any O/D pair relation.

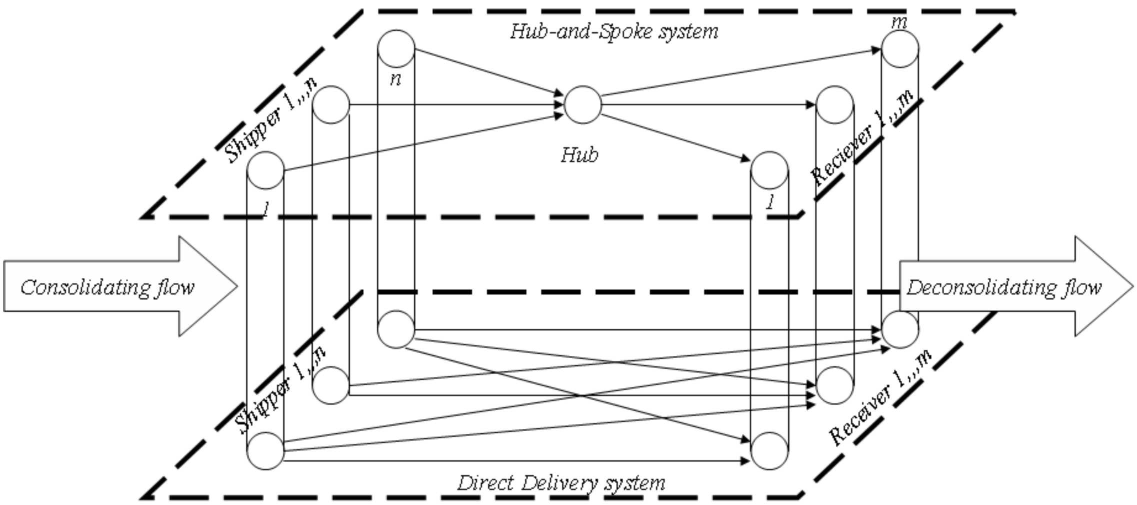


Figure 1: Foliated Transportation Network model.
Source: Persson and Lumsden, 2006.

In the Bjeljac and Lakobrija (2004) model the hub volumes are identified after the departure of all DS trucks, i.e., the decision to send a sub set of the total volume in a O/D pair to the hub is made after it is operationally apparent that the remaining volume is not sufficiently large to fill a whole unit. In contrast, Persson and Lumsden (2006) suggest a setup where the hub volumes are identified in advance and are sent first in order to improve the system’s performance from the perspective of total time in transit. The argument is that because the required transit time through a HS network is inherently longer compared to a DS setup, it would, on a system level, be beneficial to afford that portion of the goods the longest time window by shipping it first. This approach would, however, require an unprecedented level of accuracy regarding operations planning and control when it comes to identifying the hub volumes in advance (Kalantari and Sternberg, 2009).

3. Methodology

This study makes use of empirical data collected from a representative company to serve as input and configuration parameters in an experimental discrete event simulation model. The Swedish general cargo freight market is almost entirely dominated by two major companies, each covering the country with their own network of terminals. This offers support for the argument to use empirical data from one company in a general model, i.e., without actually modeling the specific system of that company. In this study, one of these companies has been chosen based on availability.

The models represent an ideal typical direct shipment network and a foliated transportation network (FTN) in order to experiment and compare the performance of the different network setups. FTN, as described by Bjeljac and Lakobrija (2004), will be modeled and compared with DS in order to quantify its potential. A model of FTN as described by Persson and Lumsden (2006) will also be used in order to investigate the sensitivity of the identified potential with regard to planning and control precision.

The choice of simulation was an obvious one based on the fact that the systems in question are complex and dynamic where experimentation in a real system would be too expensive at the same time that no satisfactory analytical solution could be found (Law and Kelton, 1991). The concept of FTN has not left its conceptual cradle, which means that an implementation of the system cannot yet be found to study. Simulation allows for testing new procedures, policies or methods (Banks *et al.* 2001). In this study a discrete event simulation model is used.

This study relies on both quantitative as well as qualitative empirical data. The qualitative data has been collected by the corresponding author through semi-structured interviews and observations. The data has mainly been used for purposes of system configuration, guidance in the collection of quantitative data, input data analysis and validation of the models. A number of informants have been interviewed, ranging from senior managers, middle managers, internal consultants and senior sales representatives to operative personnel as well as ERP-system designers and maintenance staff.

The quantitative data has been provided by the company on request. In order to improve the reliability of the study, different data sets with different resolutions have been obtained from different sources within the company. The primary sources of data used in the models have been provided by a data management specialist from a special unit in the company. The data set has contained the daily volumes (payload expressed in kg) in each relation of the network. A sample of eight consecutive weeks (40 working days) has been drawn from the provided data. The data has been verified through a comparison with aggregated (i.e., total volume in each relation aggregated for the entire period) data provided by senior management for the same period of time, and historic data provided by middle management at one of the terminals. The data has also been verified through cross check with qualitative data collected in interviews and during on site observation.

The chosen period of time is considered representative and the total flow exhibits signs of high stability. This is in line with the other sources of empirical evidence i.e. interview materials and historical quantitative data aside from the sample. The empirical data in the sample has been used for deciding the physical network setup, i.e., the number and position of terminals, and to create theoretical distributions for the daily volume of goods in each of the relations of the network. The performance of the two different network models under study has been measured through the set of KPI presented in Table 1.

Table 1: Key Performance Indikators (KPI).

KPI	*Description*
Number of trucks	The minimum number of trucks required to fulfill the transportation need of each cycle.
Average system fill rate	The ratio of the total amount of goods and the total loading capacity (C) of the fleet of trucks deployed.
Traffic work	Truck capacity multiplied by the distance traveled.
Transport work	Amount of goods shipped multiplied by the distance traveled.

3.1 Delimitations

The focus of this study is on the long-haul transport portion of the domestic Swedish general cargo freight transportation operation of the reference company. The system limit for this study is drawn at the departing gate of the origin terminal and the receiving gate of the destination terminal. System boundaries are drawn around the overnight guaranteed domestic network of the company. Aspects regarding order reception, goods collection and distribution to end consignee have been excluded. An introduction of the mentioned additional aspects would also be at the cost of the generality of the models because the specific company's business model, product range and customer profiles would play a major role on the output of the system.

4. The simulation study

The empirical data collected from/provided by the reference company has been used not to simulate the operations of the specific firm in and of itself but rather to provide representative input parameters for the comparison of two general models, i.e., a direct shipment network and a FTN. Also a third model was developed for purposes of sensitivity analysis. The simulation study follows the steps proposed by Law (2001a) and the model is created in Simul8[1] simulation software. The input data analysis, model

[1] Simul8 version 15.0 © copyright 1993-2008 Simul8 corpration. http://www.simul8.com Free student version.

descriptions and validation aspects of the simulation study will be discussed in detail below.

4.1 Input data analysis

The sample data (8 working weeks of detailed data) was checked for extreme values, autocorrelation, variance and other anomalies that would be in contrast to the other sources of data, i.e., the aggregated data for the same period, historic or qualitative data (Leemis, 2004). None of the tests revealed any reason to doubt the correctness of the data provided. Using the criteria of overnight service guarantee as defined by Bjeljac and Lakobrija (2004), 15 terminals were selected from the network, yielding an O/D matrix of 210 relations.

A randomly selected sample of 10% of the relations were selected and tested for fit for theoretical distributions. These tests for goodness of fit were performed using a statistical software called Stat::Fit[2]® (Law, 2001b). In the same fashion the theoretical distribution that best fit the size of the single shipments was estimated. The tests provided no reasons for rejecting any of the selected distributions.

4.2 Model description

Three different models were developed: a direct shipment model, a FTN model as described by Bjelajc and Lakobrija (2004) and a FTN model as described by Persson and Lumsden (2006). The first two models are used to quantify the potential impact of the FTN on the KPI as compared to DS and the third model is used to quantify the impact of the planning and control error on the identified potential of FTN. This error is the result of attempting to identify the hub volumes first as Persson and Lumsden (2006) prescribe as opposed to the setup used by Bjeljac and Lakrobrija (2004) where the hub volumes are identified and handled last. Prognosis error refers to the variation in fill rate due to the discrepancies between forecasted need for capacity and the actual outcome on a truck-by-truck basis (Kalantari and Sternberg, 2009). The prognosis, i.e., the identification of the hub volumes, created based on placed orders and not projected demand.

In all three cases the network structure, i.e., number of terminals and the distance in between them, has been mirrored from sampled data without revision. Goods in all setups have been modeled as weight measures with direction, i.e., an origin and a final destination. The weight dimension has been used to model the truck capacities. However, the empirical data is a composite measure of weight and volume expressed only in kg, i.e., payload (see 3). This has allowed for the assumption that truck capacity can/should reasonably be utilized up to 100%. This notion is supported by qualitative data and respondent validation.

In all setups, goods are served by the FIFO principle, thus there is no sophisticated selection criteria that would make optimization in grouping different consignments to a particular loading unit possible. This is in contrast to a real world system where some selection, however unsystematic or arbitrary, is bound to be feasible.

---

[2] Stat::Fit® version 2 © copyright 1995-2001 Geex mountain software corp. http://www.geerms.com Free student version.

4.3 Direct shipment model specifics

In the DS setup, each run starts with the random generation of total volume per relation and day based on the theoretical distributions extracted from the empirical data. The volume in each relation is subsequently divided in individual consignments according to the distribution of the consignment size.

Trucks are then loaded with consignments according to the FIFO principle and are shipped to their destinations. During the run, total number of trucks, average fill rate of the trucks, average system fill rate, total volume of goods, total transport work and total traffic work are collected.

4.4 FTN Comparison Model Specifics

Bjeljac and Lakobrija's (2004) FTN has been modeled for the main experiment. In the FTN models, there is one terminal designated as the hub. For the decision of which terminal will act as hub, a gravity model (Lumsden, 2006) is used. The process of random generation of goods volumes in every direction and its division to individual consignment is identical in all three setups. The difference between the FTN and DS setup appears first when the goods are to be loaded on trucks and shipped. In every relation before a truck is to be loaded, a test is performed to see whether the remaining amount of goods is more or less than what would fit on one truck. If the amount is more than the capacity of one truck, the goods are loaded and shipped directly, identical to the DS setup.

However, when the remaining amount of goods in one relation, after shipping the amount referred to above, is less than one truck's capacity, the goods are loaded onto trucks that have the hub terminal as the intermediate destination. Also, all the goods destined for the hub terminal as their final destination are loaded on the same trucks without discrimination or separation. The "hub-trucks" depart then after all DS trucks have left their terminals of origin.

The contents of the hub-trucks are upon arrival to the hub terminal unloaded, sorted and then reloaded onto trucks and sent to their final destinations. The goods that have the hub terminal as their final destination are not further handled in the model. The trucks arriving at the hub terminal are also utilized for shipping goods from the hub terminal to the final destinations.

4.5 FTN Model created for sensitivity analysis

In the model that, among others, Persson and Lumsden (2006) present, the hub volumes are to be identified and shipped first, before the DS volumes depart. This sequence is in reverse of the one defined in the model outlined by Bjeljac and Lakobrija (2004) where the DS volumes are shipped first and the hub volumes are identified by default. This reversal created at least two new sources of uncertainty that need handling.

For one, the identification of the hub volumes is done on the basis of a prognosis. The impact of the prognosis error is not known and quantified. In order to investigate the sensitivity of the results of the potential of FTN to these sources of uncertainty, a third model was created in accordance with the model description presented by Persson and Lumsden (2006). Secondly, the load composition on a specific truck, i.e., the order of loading, the mixture of different individual consignment and the physical properties of those consignments create an uncertainty regarding the operational outcome of the utilization rate of the physical capacity. This is true both for the hub volumes as well as, and primarily regarding, the DS volumes. This discrepancy is labeled and modeled as *loading error* in all three simulation models.

In this setup, the hub volumes are identified through calculation on the placed orders, i.e., the generated total volumes in each relation. The hub volumes are picked out based on the FIFO principle, and are consolidated and loaded on trucks and shipped to the hub terminal. The goods are unloaded, sorted and reloaded for further transport to their final destinations. Finally, the remaining volumes are processed and shipped in a fashion identical to the DS model description above.

4.6 Validation

Banks (1998) outlines eight different validation strategies for a simulation model. Some of the strategies are not applicable because they require testing the results in comparison to a historic outcome or a specific existing physical system. The models in this study are to be considered generally and are not a representation of a specific system, and two of the model setups simulate systems that are not yet extant. This limits the validation strategies that are applicable to face validation, sensitivity analysis and validation of conceptual model assumptions.

Face validation is implicitly performed throughout the modeling and execution of the simulation in iterative fashion and is largely due to the different role, expertise and level of involvement in the different aspects of the study on the part of the two authors. In addition, transportation and transportation research professionals have been consulted in the validation process. These experts are primarily not affiliated with the company that has provided the empirical data, because of the fact that the models are meant to be general and not a representation of a specific real world system.

The most extensive validation effort has been invested in the sensitivity analysis presented below (section 5). The results of the sensitivity analysis offer details about the sensitivity of the results to different major factors and conditions. The sensitivity of the results regarding the prognosis error was examined. Based on some counterintuitive results of the original runs, the impact of the fill rate of the last trucks was also explored.

Furthermore, the system input, conceptual model and run-model behavior has been compared to and confirmed by the broad range of qualitative data. The extensive input data analysis also provides support for the validity if the models (Sargent, 2004). The combined effort offers confirmation for the validity of the models. Any validation effort becomes more complicated when it regards system setups or policies that are yet not in effect. The implication of this condition will be discussed in detail in the sensitivity analysis part of this paper.

5. Experiment design

The main experiment designed for determining the potential of FTN via comparison with a DS setup is performed through a paired t-test of the results of 20 runs of the two models i.e. FTN according to Bjeljac and Lakobrija (2004) and a corresponding DS model. 20 replicates were deemed more than sufficient through iteration and with regard to the confidence interval of the results sought. It is likely that a smaller number of runs would also be enough. However, due to the fact that each set of replicates did not take more than one minute, no attempts were made to reduce the number of runs to the bare minimum. The aim of the experiment was to identify differences between the results of the two compared systems at the level of $p<0.01$ and to quantify those differences at the confidence interval of 99%.

5.1 Factorial Design for Sensitivity Analysis

The impact of the factors system setup, the number of terminals, variation of demand, loading error and truck capacity were examined in a two-level five-factor complete factorial design. The five factors were selected based on qualitative reasoning and previous results presented in the literature.

The system setup is introduced as a factor in order to test the validity of the results of the original experiment and to exclude the possible dominating impact of other factors on the results. The number of terminals was selected because the size of a network fit for an implementation of FTN has been discussed in previous research though without any definitive results (Hakimian and Zandi, 2009). The input data analysis revealed stable levels of demand. This has also been corroborated by quantitative historical as well as qualitative data. By including this as a factor, the importance of stable demand for obtaining the results from the original experiment could be considered. This is important in order to explore the validity of the outcome with regard to the stable demand that might be company specific. Because of the simplifying assumption of uniform fleet the truck capacity was deemed to warrant its inclusion in the factorial design (Kalantari and Lumsden, 2007). The factors and their levels are found in Table 2.

Table 2: Factorial design factors.

Factor	*High (1)*	*Low (-1)*
System setup (A)	FTN	DS
No. Of terminals (B)	15	8
Demand (C)	Empirical standard deviation	Empirical standard deviation +50%
Loading error (D)	U[0.95 , 1]	U[0.85 , 1]
Truck capacity (E)	40 ton	25 ton

The low levels for factors B and C are somewhat arbitrary. The values have been decided based on a notion of what is reasonable for the purpose of this experiment. The levels of the other factors all have been decided based on empirical inputs or existing theory. Factor D has been described with a uniform distribution, which Banks et al. (2001) names "the distribution of maximum ignorance". No quantitative data has been able to be obtained in order to model the distribution. The selected distribution is not unreasonable even though no specific quantitative justification can be found in its support. Uniform distribution is recommended in cases such as these (Law and Kelton, 1991).

The result variables for this experiment are the ratio of total volume and the number of trucks (R1), the average fill rate (R2) and the ratio of total transport work and the total traffic work (R3). The reason for including two ratios is the fact that the number of terminals (factor B) affects the total amount of goods, the number of trucks and total distance traveled, rendering the two levels of that factor incomparable. By introducing a ratio, the relative indirect impact on the number of trucks and the distance traveled could possibly be revealed.

5.2 Last Truck Fill Rate

The FTN setups modeled here are based on the existing cited literature. In these no specific level of fill rate is determined for the decision to send the last truck through the different network layers, i.e., DS or HS. Based on this, any truck that is not 100% filled has been sent through the HS layer, i.e., via the hub terminal. Reviewing the results of the original experiment, this reveals a conceptual gap in the model. In reality, one would probably prefer to send a, e.g., 90% full truck directly instead of shipping it via the hub. The unutilized capacity may not warrant the additional time, distance and terminal handling operations that a hub detour would entail.

Based on this, the sensitivity of the results to the impact of different levels' fill rate would mean a truck's rerouting through the hub terminal has been investigated.

5.3 Prognosis Error Impact

In the model description of FTN presented by Persson and Lumsden (2006), it is required on the basis of prognosis to identify and ship the hub volumes before the direct volumes depart from their origin terminals. This approach inserts an uncertainty in the setup, the effects of which are difficult to foresee. Therefore, this aspect has been included in the sensitivity analysis of the results.

The outcome of the FTN model according to Persson and Lumsden (2006) has been compared to the DS model outcome with an incremental fix prognosis error. The error has been increased by increments of 1% from -4% to +15%. With this approach the impact of just the size of the prognosis error on the results has been clarified. It is, however, unrealistic to assume that a fixed systematic error would be sustained without correction. Therefore, the same sets of tests have been run with a randomly distributed error with an incrementally increasing standard deviation. A triangle distribution has been used for this end where the extremes have been increased in 20 steps to go from trig [0.9525, 0.95, 0.935] to trig [1, 0.95, 0.65]. The result of this analysis is meant to shine a light on the robustness of the results of the FTN with regard to the size and spread of the prognosis error.

6. Results

Throughout the presentation of the results, if nothing else is explicitly expressed, the number of replicates is 20. The results referred to are true at a significance level of $p<0.01$ and are quantified at the confidence interval of 99%. The main experiment shows that in the FTN setup the system level average fill rate of the trucks increased by 14.5% (±0.2), the minimum number of trucks required was reduced by 10.5% (±0.4), the total transport work increased by 15.2% (±0.5) and the total distance traveled was not affected compared to the DS setup.

Table 3: Results of the experiment.

	FTN potential (DS=100%)	*FTN potential (DS>75%)*	*Delta*
Average fill rate	+14.5% ± 0.2	+13.65% ± 0.23	-0.75% ± 0.11
Number of trucks	-10.5% ± 0.4	-9.60% ± 0.41	+0.95% ± 0.08
Transport work	+15.2% ± 0.5	+12.57% ± 0.45	-2.29% ± 0.08
Traffic work	No significant difference	-13.62% ± 0.44	-14.17% ± 0.16

However, when the fill rate of the trucks to be allocated to the DS layer was reduced to more than 75% instead of 100% the results were affected. The improvement potential regarding the number of trucks required and the average fill rate of trucks was marginally diminished at the same time as the total distance traveled was drastically reduced and total transport work was also marginally reduced (Table 3).

Table 4: Results of factorial design experiment.

Factor	*Volume/no. trucks (R1)*	*Average truck fill rate (R2)*	*Transport work/traffic work (R3)*
A	0,103911	0,072164	0,074589
D	0,047858	0,045098	0,045364
AB	0,026314	0,022663	0,020262
± K	0,003436	0,000706	0,000702

In the factorial design experiment factors A, D and AB have the highest impact on all three result variables consecutively. Factor A, i.e., the systems setup has by far the highest impact in all three cases. The system setup has twice the impact of factor D and almost five times the impact of factor AB (Table 4).

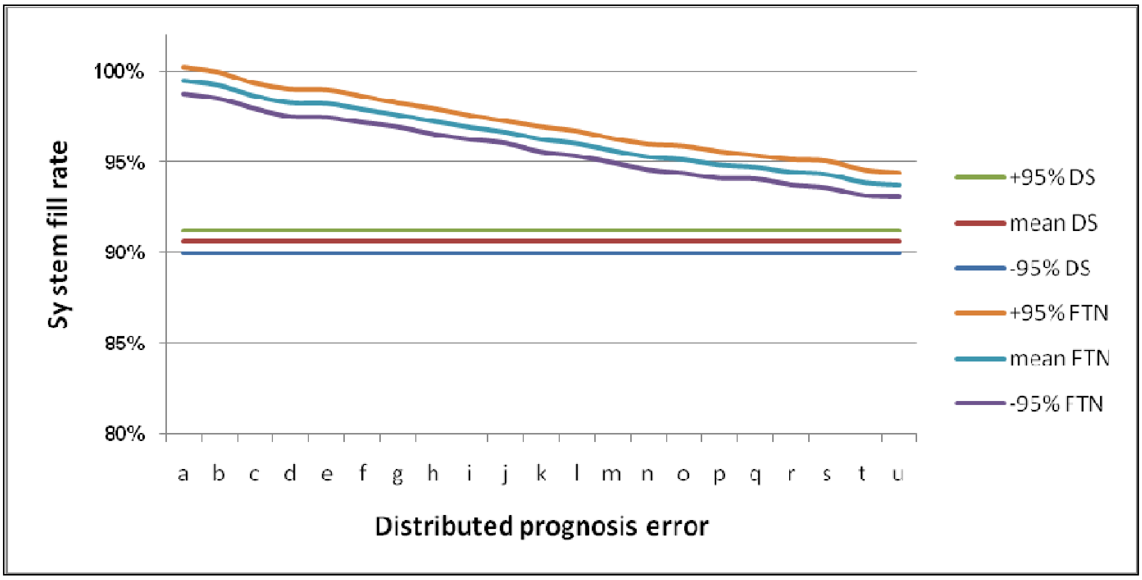


Figure 2: Impact of the distributed prognosis error on the average system fill rate.

The FTN system is shown to be fairly robust regarding the effects of the prognosis error where the error needs to reach unrealistic levels of size and spread before the FTN performance is lowered to the same level as the DS setup (Figure 2). This is true for all KPI except transport work where the FTN performance, in compliance with theory, is actually worse than DS. This robustness is identified even though the relation between prognosis precision and system performance is near linear. The robustness can be explained by the mere size of the improvement potential of FTN. This observation is in line with the main result of the factorial design analysis.

7. Conclusion

It is concluded that the performance improvement potential of FTN compared with DS with regard to the above identified KPI is relatively large. The sensitivity analysis confirms that the observed results are the impact of the introduction of the FTN setup and that this setup is fairly robust with regard to prognosis error. However, because of the relative strength of the impact of factor D, i.e., the loading error, it can be concluded that any setup is sensitive to how well truck capacity can be utilized operationally.

It is further concluded that the governing rules of FTN need to be developed in finer detail as the sensitivity analysis showed that simple manipulation with cut off values results in significant changes in the end results.

8. Discussion and future research

In this section the results are to be put in context as well as for general implications of them to be drawn out. It must be observed that there always exists a trade-off between fill rates and service level. In these experiments, this trade-off has not been addressed by means of ensuring 100% service level i.e. every transport order fulfilled over night. It is likely, and indeed the qualitative data from the reference company concurs, that in reality the transport provider would opt to reduce its service level in order to increase that utilization rates. The approach cannot be said to have biased the results, because the same conditions were set for all three of the models. However, this must still be taken into consideration because for systems were this tradeoff has yielded a better system performance it can be anticipated the introduction of FTN would increase the level of service whilst perhaps the impact would be less pronounced regarding the physical KPI.

Furthermore, cited previous studies as well as the study at hand explain the performance improvement potential of FTN by the load consolidation of the last trucks the cutoff fill rate cannot be achieved in every relation. This indicates that FTN would probably be best suited for systems where the volume of goods handled is neither very low or very high as compared to the size of the load carriers. For systems where the volumes are lower than at least one load carrier in every direction a pure HS system would be more appropriate. Conversely, as the size of the flows start to require a sufficiently large number of load carriers, the impact of consolidation aspect of FTN on the system level would start to dwindle. The upper boundary of this range for the

"Goldilocks system" that would benefit most from FTN is not treated in this study and should get some attention in the future.

On the same note, it was demonstrated that a simple adjustment to the static cut-off value for the fill rate of the last truck, yielded considerable positive dividends for the end results. This indicates that development and implementation of dynamic governing rules such as cut-off values for the last truck or other adapted network optimizing tools would presumably generate an even greater potential for FTN. In the opinion of the authors, this is the premier track for future research.

An incidental observation during the performing of the simulations were that roughly a quarter of the fleet was send to the hub, and subsequently redistributed in the network during the same cycle, in the FTN setups. The implications of this, if any, on the fleet management aspects of the system should generate future research attention.

Finally the reduction of the traffic work coupled with the reduction in the number of trucks in the system should lead to decreased negative environmental impact on the system level. This reduction is supposedly not linear, because of the discrete nature of a transportation system. The ramifications of FTN for a transportation systems environmental footprint ought to be studied in greater detail in future work.

Acknowledgements

The authors would like to extend their gratitude to VINNOVA (Swedish Governmental Agency for Innovation Systems) for the funding of this project. Thanks are due to Schenker DB and Volvo Logistics for co-funding this project. A great many thanks to Professor Kent Lumsden for his insightful input throughout this study.

References

Banks, J. (1998) "Principles of Simulation", In: Banks, J. (Ed.) *Handbook of Simulation - Principles, Methodology, Advances, Applications and Practice*, John Wiley & Sons, Inc., New York, NY.

Banks, J., Carson, J. S. I., Nelson, B. L. and Nicol, D. M. (2001) *Discrete-Event System Simulation*, Upper Saddle River, NJ, Prentice Hall, Upper Saddle River, NJ.

Bjeljac, R. and Lakobrija, D. (2004) Foliated Hub Systems, the Necessary Requirements for Implementation at Schenker?, *M. Sc. thesis*, Chalmers University of Technology, Gothenburg.

Bryan, D. and O'Kelly, M. (1999) "Hub-and-Spoke Networks in Air Transportation: An Analytical Review", *Journal of Regional Science* 39: 275-295.

Crainic, T. (2003) "Long-haul Freight Transportation", In: Hall, R. W. (Ed.) *Handbook on Transportation Science*, Kluwer Academic Publishers, New York.

Crainic, T. G. (2002) "A Survey of Optimisation Models for Long-Haul Freight Transportation", In: Hall, R. W. (Ed.) *Handbook of Transportation Science*, Kluwer.

European Commission, D. E. a. T. (2003) Energy and Transport: main figures, *Energy and Transport Outlook to 2030 EU-15*.

Hakimian, H. and Zandi, M. S. (2009) Foliated Transportation Systems, *M. Sc. thesis*, University Collage of Borås, Borås.

Kalantari, J. and Lumsden, K. R. (2007) "Stepwise Replacement of a Direct Shipment Network with a Hub and Spoke System", *The 19th Annual Nofoma Conference*, (Halldórsson, Á. and Stefansson, G. (eds.)), Reykjavik, Iceland.

Kalantari, J. and Sternberg, H. (2009) "Research outlook on a mixed model transportation network", *European Transport* 14: 61-76.

Law, A. (2001a) "How to Build Valid and Credible Simulation Models", *Winter Simulation Conference*, (Peters, B. A., Smith, J. S., Medeiros, D. J. and Rohrer, M. W. (eds.)), Arlington, Va.

Law, A. and Kelton, D. (1991) *Simulation Modeling & Analysis*, Singapore, McGraw-Hill, Singapore.

Law, A. M. (2001b) "How the Expertfit Distribution -Fitting Software Can Make Your Simulation Models More Valid", *Winter Simulation Conference*, (Peters, B. A., Smith, J. S., Medeiros, D. J. and Rohrer, M. W. (eds.)), Arlington, VA.

Leemis, L. M. (2004) "Building Credible Input models", *Winter Simulation Conference*, (Ingalls, R. G., Rossetti, M. D., Smith, J. S. and Peters, B. A. (eds.)), Washington, DC.

Lumsden, K. (2006) *Fundamentals of Logistics*, Lund, SWE, Studentlitteratur, Lund, SWE.

Lumsden, K., Dallari, F. and Ruggeri, R. (1999) "Improving the Efficiency of the Hub and Spoke System for the SKF European Distribution Network", *International Journal of Physical Distribution & Logistics Management* 29: 50-64.

McKinnon, A. C. and Ge, Y. (2006) "The potential for reducing empty running by trucks: a retrospective analysis", *International Journal of Physical Distribution & Logistics* 36: 391-410.

Persson, P.-O. and Lumsden , K. R. (2006) Foliated Transportation Networks, *Logistics Research Network*, Newcastle, United Kingdom.

Persson, P.-O. and Waidringer, J. (2006) "Foliated Transportation Networks", In: Jahre, M. (Ed.) *18th Annual NOFOMA Conference*, Oslo, Norway.

Sargent, R. (2004) "Validation and Varification of Simulation Models", *Winter Simulation Conference*, (Ingalls, R. G., Rossetti, M. D., Smith, J. S. and Peters, B. A. (eds.)), Washington, D.C.

Woxenius, J. (2007) "A Generic Framework for Transportation Network Design: applications and treatment in intermodal freight transport literature", *Transport Reviews* 27: 733-749.

Defining level of service criteria of urban streets in Indian context

Prasanta Kumar Bhuyan [1], K. V. Krishna Rao [2*]

[1] *Former Research scholar, Department of Civil Engineering, Indian Institute of Technology Bombay, Powai, Mumbai - 400076, Maharashtra, India*
[2] *Professor, Civil Engineering Department, IIT Bombay, India 400076*

Abstract

Speed ranges of Level of Service (LOS) categories of urban streets are not well defined for highly heterogeneous traffic flow condition on urban streets in Indian context. In this respect, a study was carried out in the city of Mumbai, India and the result was tested on two major corridors in Kolkata City. Average travel speed on street segments is used as the measure of effectiveness, which in this case has been derived from second by second speed data collected using Global Positioning System (GPS) receiver fitted on mobile vehicles. Hierarchical Agglomerative Clustering (HAC) is applied on average travel speeds to define the speed ranges of urban street and LOS categories. Applying this methodology it is found that urban street speed-ranges of LOS categories valid in Indian context are different from that values specified in HCM (2000). The application of this procedure is that in a simple manner with the application of GPS it can be applied in the evaluation of level of service of urban streets in different environment.

Keywords: Level of service; Urban streets; Hierarchical agglomerative clustering; GPS; Heterogeneous traffic flow.

---

1. Introduction

The *Highway Capacity Manual* (HCM 2000) designates six levels of service for each type of facility, from A to F, with LOS "A" representing the best operating conditions and LOS "F" the worst. It uses distinct values as boundaries for the various levels of service, each of which represents a range of operating conditions. The classification of urban streets into number of street classes and speeds into different levels of service categories is well defined in HCM 2000 are well applicable in homogenous traffic flow condition. Hence, in this study it has been attempted to classify urban streets and speeds into number of categories that is applicable in the prevailing context of heterogeneous

* Corresponding author: K. V. Krishna Rao (kvkrao@civil.iitb.ac.in)

traffic flow. Traffic travels at reduced speed in heterogeneous traffic flow condition in India; whereas speed ranges defined in HCM 2000 are higher since it was developed in homogeneous traffic flow condition. Hence the speed ranges for urban streets in Indian context need to be defined. However if this methodology is applied in developed countries, the result may indicate LOS boundaries similar to the HCM 2000 values.

This study emphases on the easy and accurate way of speed data collection by the use of GPS as well as its applicability in defining speed ranges of LOS categories applicable in the present context. Though speed ranges are not well defined for LOS categories, it has been suggested by Indian Road Congress (1990) that on urban roads, the levels of service are strongly affected by factors like the heterogeneity of traffic, speed regulations, frequency of intersections, presence of bus stops, on-street parking, roadside commercial activities, and pedestrian volumes etc. Limited studies have carries out for heterogeneous traffic condition in India. Marwah and Singh (2000) have classified level of service into four groups (LOS I-IV). In another study for heterogeneous traffic condition on urban roads Maitra et al. (1999) redefined the LOS boundaries by quantifying congestion as measure of effectiveness.

Kikuchi and Chakroborty (2006) have reviewed the definitions of LOS categories that have been followed traditionally. The authors have examined the uncertainty associated with the measuring and mapping of existing six LOS categories. And six frameworks were formulated to address the uncertainty lies within six levels of service categories. Zolnik and Cromley (2006) have developed a poissioned- multilevel bicycle level of service methodology using the bicycle-motor vehicle collision frequency and severity in the GIS environment. This new methodology complements bicycle level of service methodologies on mental stressors by incorporating the characteristics of cyclists involved in bicycle - motor vehicle collisions as well as the physical stressors where bicycle-motor vehicle collisions occurred to assess bicycle, level of service for regional road network. Romana and Perez (2006) have used a threshold speed to assess LOS. The definition of threshold speed used "the minimum speed users consider acceptable in traveling on a uniform road section under heavy flows and platooning traffic". The method used the same Measure of Effectiveness (MOEs) proposed by HCM, one reflecting percent time spent following (PTSF) and the second reflecting speed. However it has been suggested a threshold speed would be used to decide which MOE governs LOS in each period: if average travel speed (ATS) is higher than the threshold value, only PTSF would be examined, implying user consider the speed reasonable. If ATS is below the threshold speed, platooning would be behind speed in importance in the view of drivers.

Jensen and Trafitec (2007) have developed pedestrian and bicyclist satisfaction models using cumulative logit regression of ratings and variables. The model includes variables, which relate significantly to satisfaction ratings. Motorized traffic volume and speed, urban land uses, rural landscapes, type and width of pedestrians and bicyclist facilities, number and width of drive lanes, volume of pedestrians, bicyclists and parked cars, and also presence of median, tress and bus stops significantly influence the level of satisfaction. Models return percentage splits of six levels of satisfactions. These splits are then transferred into a level of service. Muraleetharan and Hagiwara (2007) have developed a methodology for estimating the overall LOS of pedestrian walkways and cross walks based on the total utility value, which came from a stated preference survey. Each sidewalk and crosswalk link was assigned with an overall LOS according to its operational and geometrical characteristics. The model result indicates that pedestrians choose route not only for distance, but also for the overall LOS of sidewalks and crosswalks.

Xu et al. (1999) found that neural network-based macro taxi model gives much more accurate information on taxi services than the simultaneous equations model for Hong Kong urban areas. Madanat et al. (1994) have applied ordered probit model to find threshold values for each LOS categories using user perceptions. Cheol and Stephen (2002) used *K*-means, Fuzzy and SOM clustering in a real-time signalized intersection surveillance system for the determination of LOS categories. Lingra (1995) compared grouping of traffic pattern using the Hierarchical Agglomerative Clustering and the Kohonen Neural Network methods in classifying traffic patterns. In this study it has been observed the advantage of hierarchical grouping on a small subset of typical traffic patterns to determine the appropriate number of groups. Also, Lingra (2001) applied Hierarchical Agglomerative Clustering technique and an evolutionary Genetic Algorithms approach for the classification of highway sections. It has been pointed out that hierarchical approach tends to move farther away from the optimal solution for smaller number of groups.

Generally, average travel speed on street segments is used as the measure of effectiveness for defining speed ranges of LOS categories. The use of moving observer method is the most commonly used technique for the collection of travel speed data. However, limitation to the use of moving observer method is that accuracy in data collection varies from technician to technician. The use of GPS in this study helps in gathering large amount of speed data collected second wise which is better than traditional techniques. Accurately collected speed data were easily managed in the GIS environment. The collected data are used to find the free-flow speed ranges of urban street classes and speed ranges of levels of service categories as these are not well defined for heterogeneous traffic flow condition in Indian context.

While carrying out a thorough study in this context, this paper presents a methodology for defining speed ranges of level of service categories of urban street classes. Urban streets are classified into four classes using the free flow speeds data averaged over each street segments comes under the study area. Also, the travel speeds averaged on each street segment were classified into six levels of service categories. Consequently this study demonstrates the potential use of average travel speed as a measure of effectiveness for assessing the level of service of urban streets. The average travel speed on segments, is derived from second by second speed data collected using Global Positioning System (GPS) receiver fitted on mobile vehicles traveled on major corridors in the city of Mumbai, India. . However, in order to justify the applicability of this level of service criteria, similar set of data were collected for Kolkata City. Speed ranges of level of service categories developed for Mumbai city are also well applicable to Kolkata City. With the capability to obtain average travel speed by direct field measurement method using GPS fitted with mobile vehicle system, the possibility of obtaining new LOS criteria that can be used for urban street analysis emerges. The methodology for applying hierarchical agglomerative clustering and validation parameter ‘Silhouette’ has been demonstrated by finding cohesiveness that lies among speed data points while fixing speed ranges for LOS categories. Five major road corridors comprising of 100 street segments in Mumbai city, operating under mixed traffic are taken for this study. The following section provides brief review of the Hierarchical Agglomerative Clustering technique and its validation measure. The remaining sections describe the application of hierarchical agglomerative clustering methodology, study corridors and methods of data collection, analysis of the collected data, results obtained, summary and conclusions.

2. Hierarchical Agglomerative Clustering (HAC) Methodology

Cluster analysis, also called data segmentation, has a variety of goals. All relate to grouping or segmenting a collection of objects (also called observations, individuals, cases, or data rows) into subsets or "clusters", such that those within each cluster are more closely related to one another than objects assigned to different clusters. Central to all of the goals of cluster analysis is the notion of degree of similarity (or dissimilarity) between the individual objects being clustered. Hierarchical agglomerative cluster analysis is a statistical method for finding relatively homogeneous clusters of cases based on measured characteristics. It starts with each case in a separate cluster and then combines the clusters sequentially, reducing the number of clusters at each step until only one cluster is left. The following three steps were followed to perform Hierarchical Agglomerative Clustering (HAC) on free flow speed data using the statistics toolbox in MATLAB to classify urban streets into four classes and on average travel speeds into six levels of service categories.

2.1 Step 1: Finding the Similarity between Objects

The similarity between objects (speeds) is calculated by the use of distance function. For a data set made up of m objects, there are $m \times (m-1)/2$ pairs in the data set. For example, consider a sample data set, X, made up of six objects {say average free flow speed (ffs) values in kmph on six street segments}. The data set can be defined as a matrix:

X = [ffs_1; ffs_2; ffs_3; ffs_4; ffs_5; ffs_6] = [85.00; 92.56; 72.85; 50.66; 39.89; 38.58].

While applying HAC, the distance function calculates the distance between ffs_1 and ffs_2, ffs_1 and ffs_3, and so on until the distances between all the pairs have been calculated. The distance function returns the information such that each element contains the distance between pair of objects (ffs) can be represented by a distance vector say Y.

Y = [$ffs_{1\text{-}2}$ $ffs_{1\text{-}3}$ $ffs_{1\text{-}4}$ $ffs_{1\text{-}5}$ $ffs_{1\text{-}6}$ $ffs_{2\text{-}3}$ $ffs_{2\text{-}4}$ $ffs_{2\text{-}5}$ $ffs_{2\text{-}6}$ $ffs_{3\text{-}4}$ $ffs_{3\text{-}5}$ $ffs_{3\text{-}6}$ $ffs_{4\text{-}5}$ $ffs_{4\text{-}6}$ $ffs_{5\text{-}6}$] =
[7.56 12.15 34.34 45.11 46.42 19.71 41.90 52.67 53.98 22.19 32.96 34.27 10.77 12.08 1.31].

The distance vector can also be reformatted into a matrix to make it easier to see the relationship between the distance information generated by the distance function and the objects in the original data set. In this study, the distance function is applied two times to compute the dissimilarity between every pair of speed data. Firstly, on the free-flow speed data averaged over each street segments and secondly on the congested speed data. There are several distance functions like City block or Manhattan, Minkowski, Cosine, Correlation, Hamming, Jaccard and Euclidean to calculate the dissimilarity matrix. In this study "Euclidean distance" was used as the measuring distance.

2.2 Step 2: Defining the Links between Objects

Once the proximity between objects in the data set has been computed, it can be determined how objects in the data set should be grouped into clusters, using the linkage function. The linkage function takes the distance information generated by the distance function and links pairs of objects that are close together into binary clusters. The linkage function then links these newly formed clusters to each other and to other objects to create bigger clusters until all the objects in the original data set are linked together in a hierarchical tree. For example, given the distance vector *Y* generated by distance function from the sample data set of ffs, the linkage function generates a hierarchical cluster tree, returning the linkage information in a matrix, *Z.*

Z =

5	6	1.31
1	2	7.56
4	7	10.77
3	8	12.15
9	10	22.19

In this output, each row identifies a link between objects or clusters. The first two columns identify the objects that have been linked. The third column contains the distance between these objects. For the sample data set of ffs*,* the linkage function begins by grouping objects 5 and 6, which have the closest proximity (distance value = 1.31). The linkage function continues by grouping objects 1 and 2, which have a distance value of 7.56.

The third row indicates that the linkage function grouped objects 4 and 7. If the original sample data set contained only six objects, which is the object 7? Object 7 is the newly formed binary cluster created by the grouping of objects 5 and 6. When the linkage function groups two objects into a new cluster, it must assign the cluster a unique index value, starting with the value *m*+1, where *m* is the number of objects in the original data set. (Values 1 through *m* are already used by the original data set.) Similarly, object 8 is the cluster formed by grouping objects 1 and 2. Linkage uses distances to determine the order in which it clusters objects. The distance vector *Y* contains the distances between the original objects 1 through 6. But linkage must also be able to determine distances involving clusters that it creates, such as objects 7, 8, 9 and 10. By default, linkage uses a method known as single linkage. However, there are a number of different methods available. The linkage function grouped object 8, the newly formed cluster made up of objects 1 and 2, with object 3 from the original data set. Similarly, the linkage function grouped object 9, the newly formed cluster made up of objects 4 and 7, with object 10, the newly formed cluster made up of objects 3 and 8.

Valid cluster links objects in the cluster tree that has strong correlation with the distances between objects in the distance vector. And, the cophenet function was used to compare these two sets of values and compute their correlation, returned a value called the cophenetic correlation coefficient. The closer the value of the cophenetic correlation coefficient is to 1, the better the clustering solution.

The cophenetic correlation between Z and Y is defined as

$$C = \frac{\sum_{i<j}(Y_{ij} - y)(Z_{ij} - z)}{\sqrt{\sum_{i<j}(Y_{ij} - y)^2 \sum_{i<j}(Z_{ij} - z)^2}} \tag{1}$$

where:
Y_{ij} is the distance between objects i and j in Y,
Z_{ij} is the distance between objects i & j in Z,
y and z are the average of Y and Z respectively.

The cophenetic correlation coefficient is highest for 'average' linkage type using the average travel speed data hence used in this study.

2.3 Step 3: Creating Clusters

The hierarchical, binary cluster tree created by the linkage function is most easily understood when viewed graphically. The Statistics Toolbox function dendrogram plots the tree using the example data set is shown in the Figure 1. In the figure, the numbers along the horizontal axis represent the indices of the objects (ffs_1; ffs_2 etc.) in the original data set. The links between objects are represented as upside-down U-shaped lines. The height of the U indicates the distance between the objects. For example, the link representing the cluster containing objects 5 and 6 has a height of 1.31. The link representing the cluster that groups object 3 together with objects 1 and 2 (which are already clustered as object 8) has a height of 12.15. In this step, the hierarchical tree was cut at the desired point to form the required number of clusters using the cluster function. For example, in Figure 1 when the dendogram is cut at a height of 15, two clusters will be formed. The cluster 1 comprises of 3 objects (ffs_1; ffs_2; ffs_3) and cluster 2 comprises of 3 objects (ffs_4; ffs_5; ffs_6). In this study, the cluster function was applied five times; once for the free-flow speed data and four times on the speed data of urban street classes.

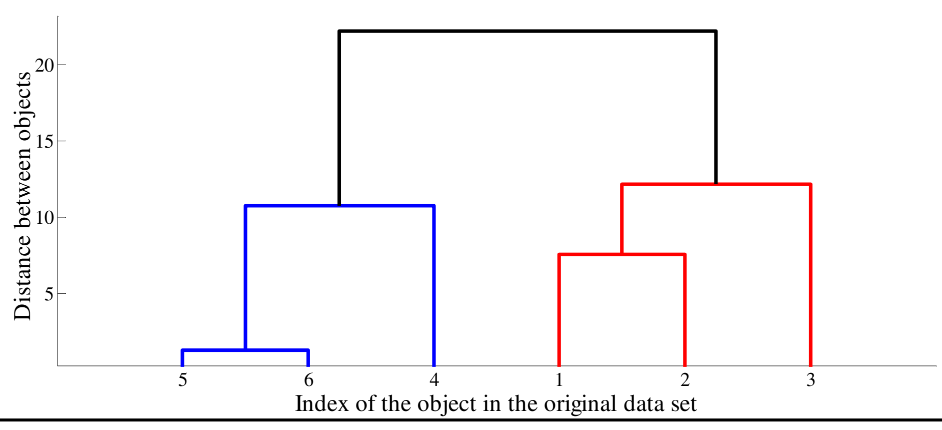


Figure1: Typical Dendogram using sample free flow speeds data.

3. Study Corridors and Data Collection

3.1 Study Corridors

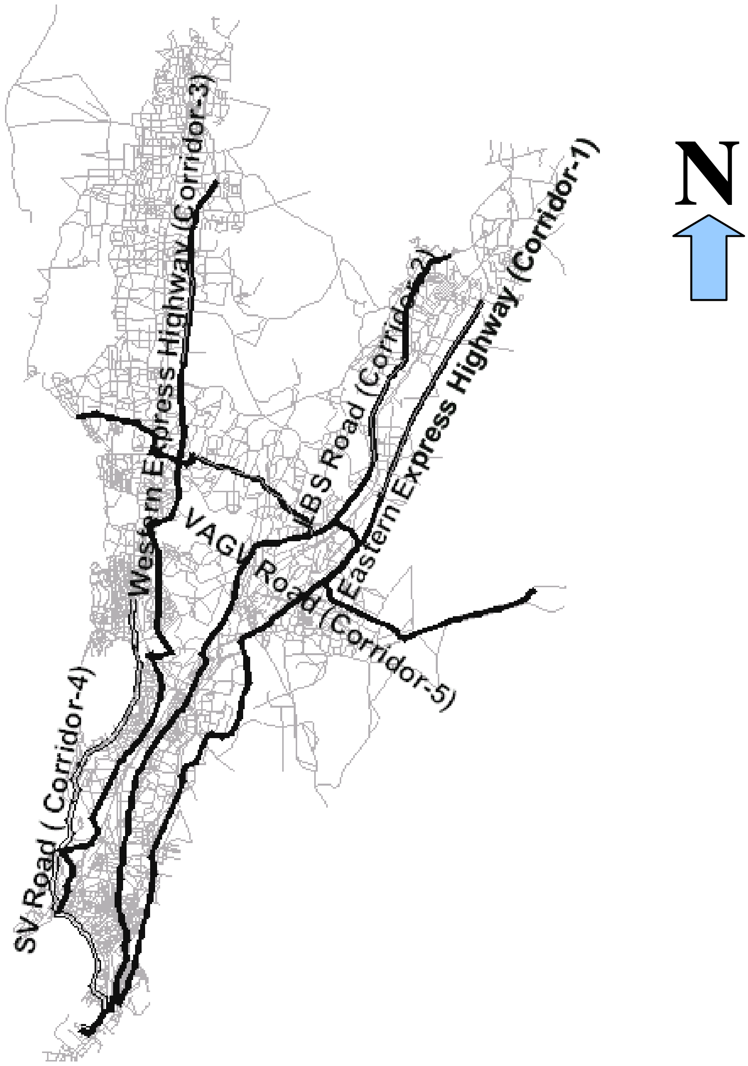


Figure 2: Map showing selected corridors of Greater Mumbai.

The base year road network of the study area is prepared in GIS environment. Additional attributes on the road network were added to the base map by conducting a road inventory survey. Five important urban road corridors of the city of Mumbai of Maharastra State, India are selected for the present study. Grater Mumbai is an Island city with a linear pattern of transport network having predominant North-South commuter movements. Passengers move towards south for work trip in the morning hours and return back towards north in the evening hours. Hence four north-south corridors and one east-west corridor comprising of 100 street segments have been chosen for the present study. Major roads like Eastern express highway extending up to south (Corridor-1), LBS Road extending up to south via Ambedkar road (Corridor-2), Western express highway extending up to marine drive (Corridor-3), SV road extending up to south via Veer Savarkar road (Corridor-4) and Versova- Andheri- Ghatkopar-

Vashi (VAGV) (Corridor-5) are included. These are shown on the GIS base map in Figure 2. In order to show the applicability of this study in other cities of India a similar survey was carried out in Kolkatta City. Two corridors having varying geometric and surrounding environmental characteristics were taken into considerations i.e. one corridor was Airport to Joka and the other corridor was Airport to Ulberia. These two corridors are approximately 80 kilometer length; comprised of 50 street segments and the data were collected during August, 2009. The interesting fact on selecting these two cities for this study is that traffic composition and road geometric characteristics along with functionality brings the true variation that was required for this purpose.

3.2 Data Collection

The probe vehicle used in this study was mid-sized car. This vehicle was fitted with Trimble Geo-XT GPS receiver, capable of logging speed data continuously at time intervals of one second. The survey was conducted during March and April of the year 2005. GPS provides both spatial and time/distance based data from which various traffic parameters were derived, including travel time and travel speeds. In order to get unbiased data sets three mid-sized cars were used and help of three drivers on different days of the survey work was taken. Basically three types of data sets were collected. The first type is roadway inventory details, for which a data dictionary was prepared using Pathfinder office 3.0. During the collection of inventory details proper segmentation technique was applied, which is just after signalized intersection to just after next signalized intersection. Details on segments like segment number, number of lanes, median types, pedestrian activity, road side development, access density, construction activity, speed limit, separate right turn lane, number of flyovers, date and day of data collection and segment length were collected. From the inventory data, it was found that most of urban street segments are four lane divided type. On street parking of vehicles, pedestrian activities coupled with commercial activity of vendors are affecting the smooth movement of the vehicles considerably. Inadequate road infrastructure with lack of enforcement in traffic rule and regulations are bringing traffic into a chaotic condition during peak hours on these corridors.

The second type of survey conducted was to find the free flow speed. Before going for the free flow speed data collection, we should know when the traffic volume is less than or equal to 200 vehicles per lane per hour. A detailed 24 hour traffic volume count survey was conducted by this group for Western Freeway Sea- Link (WFSL) project in the month of April 2005.The traffic volume data were collected on 45 stations on seven screen lines. From this survey data, traffic volume per lane per hour was calculated for roads comings under this study area. It was found that free flow traffic condition (less than 200veh/ln/hr) is approaching at 12 mid-night and all road sections are having free flow traffic conditions from 1 AM to 5 AM. Hence free flow speed for all these corridors were collected during these hours. The third type of data colleted was congested travel speed. Congested travel speed survey was conducted during both peak and off-peak hours on both directions of all corridors. Number of trips covered for each direction of travel and for the study hours (peak, off-peak and free-flow) is at least 3 and sometimes it is up to six trips. After data has been collected in the field, it has been transferred back to the office computer by using Pathfinder office version 3.00. Accuracy of field data were significantly improved through a process called *differential correction.* This requires a set of base files those are collected at a known location at the same time that field data (rover) files were collected. For this study, we collected the base files from remote sensing division, department of civil engineering, IIT Bombay. Data were visually checked before exported to a GIS or spatial database. This was to confirm that all the expected data were there, and was looked for any unwanted positions. Further, GPS data were incorporated into a GIS based transportation software called TransCAD. Similar method was applied to the data collected from Kolkata city also.

4. Validation measure

Cluster validity refers to the problem whether a given partition fits to the data at all. The clustering algorithm always tries to find the best for a fixed number of clusters and the parameterized cluster shapes. However this does not mean that even the best fit is meaningful at all. Either the number of clusters might be wrong or the cluster shapes might not correspond to the groups in the data, if the data can be grouped in a meaningful way at all (Bensaid et al. 1996; Bezdek and Pal 1998). Silhouette is used in testing the cohesiveness lies among speed data points those falls under same category.

4.1 Silhouette

The graphical representation of each clustering is provided displaying the silhouettes introduced by Rousseeuw (1987). The entire clustering is displayed by plotting all silhouettes into a single diagram, allowing us to compare the quality of the clusters. Therefore, the silhouette shows which objects lie well within their cluster and which ones are merely somewhere in between clusters. A wide silhouette indicates large silhouette values and hence a pronounced cluster. The other dimension of a silhouette is its height, which simply equals the number of objects in a group. In order to obtain an overview, the silhouettes of the different clusters are printed below each other. In this way the entire clustering can be displayed by means of a single plot, which enables us to distinguish clear-cut clusters from weak ones. The average of the silhouettes for all objects in a cluster is called the *average silhouette width* of that cluster. The average of the silhouettes for the entire data set is called the *average silhouette width for the entire data set*. The choice of optimal number of clusters is one of the most difficult problems of cluster analysis, for which no unique solution exists. This *average silhouette width for the entire data set* should be as high as possible, is used for the selection of optimal number of clusters. For application this maximum value of average *silhouette width for the entire data set* is called the *silhouette coefficient.* The silhouette coefficient is a dimensionless quantity which is at most equal to 1.

5. Results and Analysis

Average travel speed was calculated direction-wise on each segment. Hierarchical Agglomerative Cluster analysis was applied in two stages. First on average free flow speed on all segments and corresponding range of free-flow speed were found for each urban street class. Second on congested average travel speed on each segment for both peak and off-peak hours and corresponding speed ranges were found for LOS categories. While applying Hierarchical Agglomerative Clustering method on speed data "Euclidian distance" was used as the measuring distance on speed values. Various linkage types were tested through the statistical measuring parameter Cophenetic Coefficient (CC) to find out the best linkage type whose value approaches to one. Cophnetic Coefficient values for various linkage types are calculated and "Average linkage" type was selected because of maximum CC value among all. Hierarchical tree of binary clusters was divided into larger clusters using the cluster function. The hierarchical cluster tree (dendogram) formed out of free flow speed data was cut off at a level where it formed five clusters.

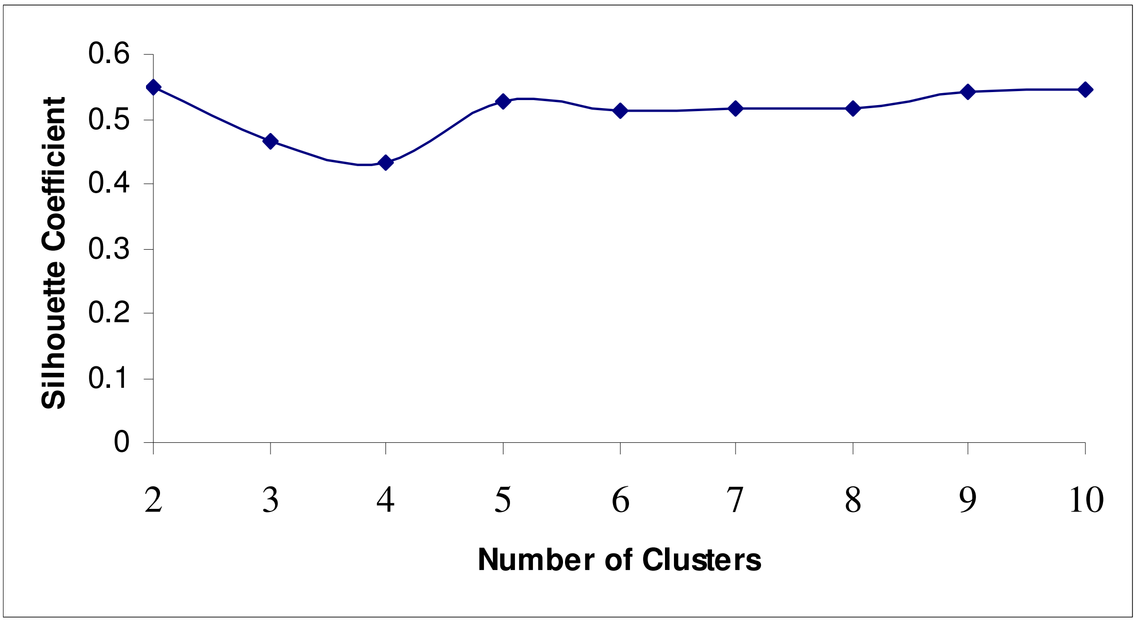
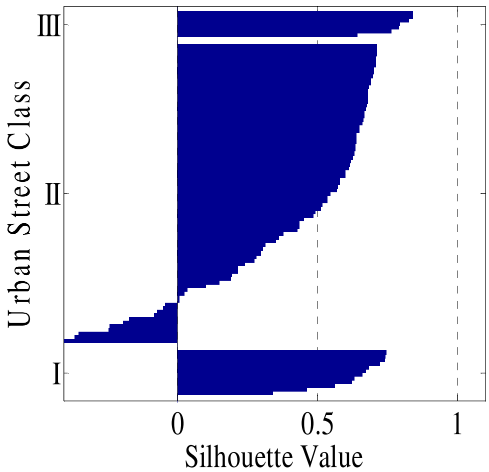
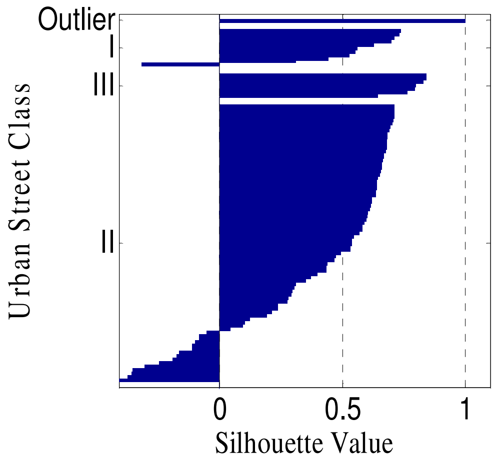


Figure 3: Silhouette Coefficient for finding Optimal Clusters used in Urban Street Classification.

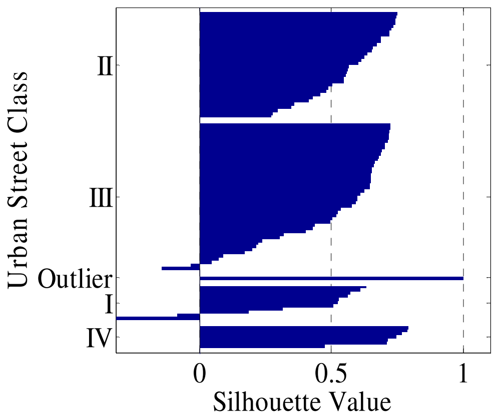
The heuristic reasoning is that the silhouettes should look best for a "natural" number of clusters. Hence, silhouette is used as the validation measure in finding optimal number of clusters and degree of cohesiveness that exist in inter-cluster and intra-cluster data points. One way to choose appropriate number of clusters is to select that value of cluster for which silhouette coefficient is as large as possible. At the same time it is considered that partitions with lesser clusters are better, when the differences between the values of validation measures are minor as number of cluster increases. While finding optimal number of clusters using silhouette coefficients as shown in Figure 3, it turns out that the best choice is when number of clusters is 5. The silhouette coefficient equals 0.527 at this cluster number five and according to previous discussion, it is interpreted that a reasonable good structure has been exist while clustering the free-flow speed data points to find urban street classes. One should never merely accept a high overall average silhouette width as its face value, but also look at the graphical output itself to find out what caused the large value. In order to look into degree of cohesiveness that lies within urban street classes, silhouettes were plotted for each free-flow speed. In Figure 4 (A) and 4 (B) for three and four urban street classes it is found that urban street class II is not well classified in both cases because they contains large number of negative silhouette values. Comparing Figure 4 (C) and 4 (D) it is interpreted that best cluster partition lies when urban street is classified into five classes which was further justified by Figure 3. In conclusion, the silhouette plots tell us that a partition into five clusters is probably most natural. But out of five classes one is considered as an outlier because in this group only a single free flow speed data point lies. Depending on the subject matter and task at hand, we wanted to put the outlier aside for further investigation and classified urban street into four classes.



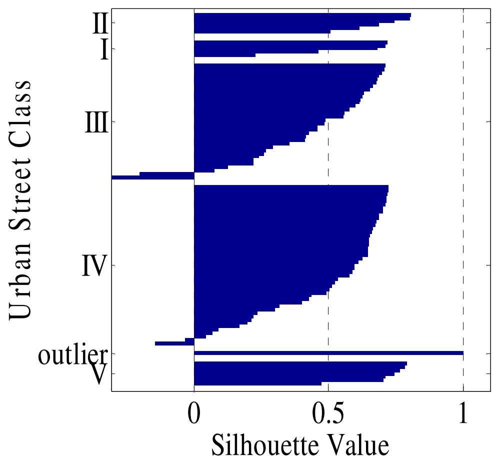
A: Silhouette Values for Three Urban Street Classes



B: Silhouette Values for Four Urban Street Classes



C: Silhouette Values for Five Urban Street Classes



D: Silhouette Values for Six Urban Street Classes

Figure 4: Silhouette Values used in Urban Street Classification.

As in HCM (2000), further justified by optimal clustering criteria it has been attempted to categorize the urban streets into four classes based on free flow speed, geometric and surrounding environmental characteristics in the present context. Free flow speed ranges for four urban street classes are obtained in the present context. Speed rages for level of service categories of urban street classes (I-IV) obtained in Indian context are shown in Figure 5. These results are also shown in Table 1. Silhouette Values for Urban street classes are plotted and it was found that average Silhouette width lie between 0.35 and 0.75 for all street classes. This indicates that reasonable classifications have been presented.

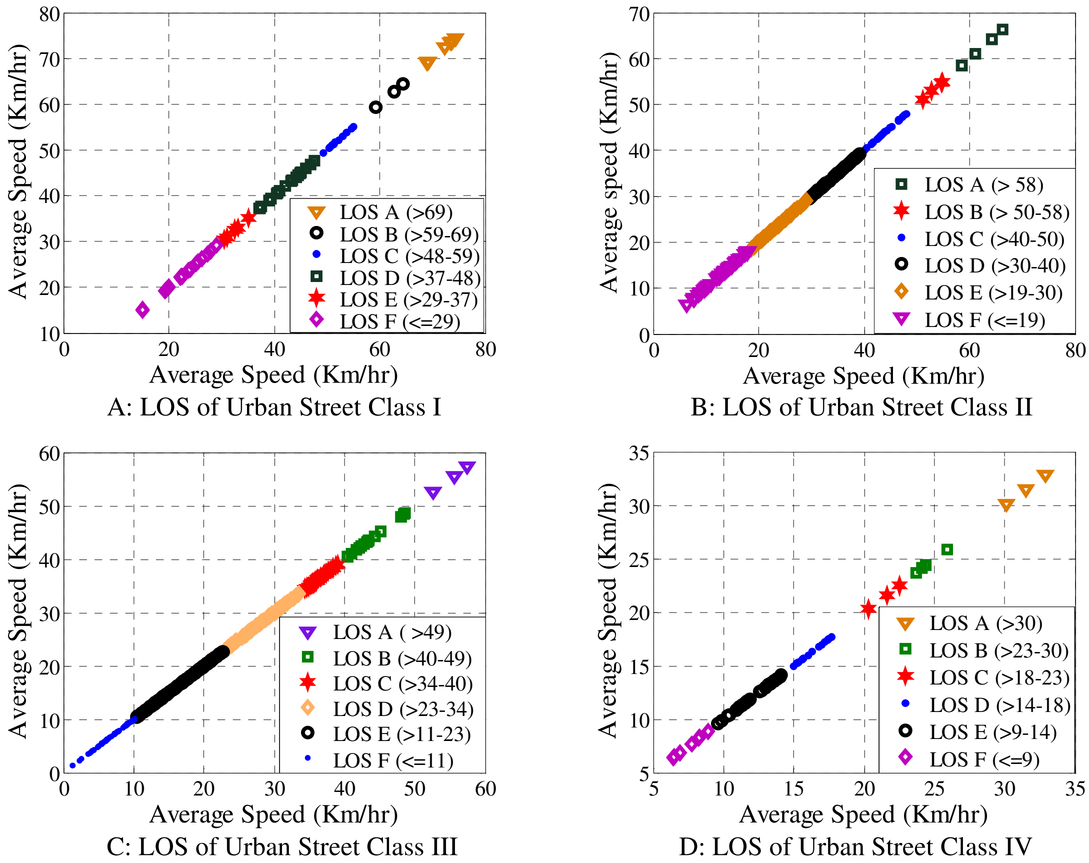
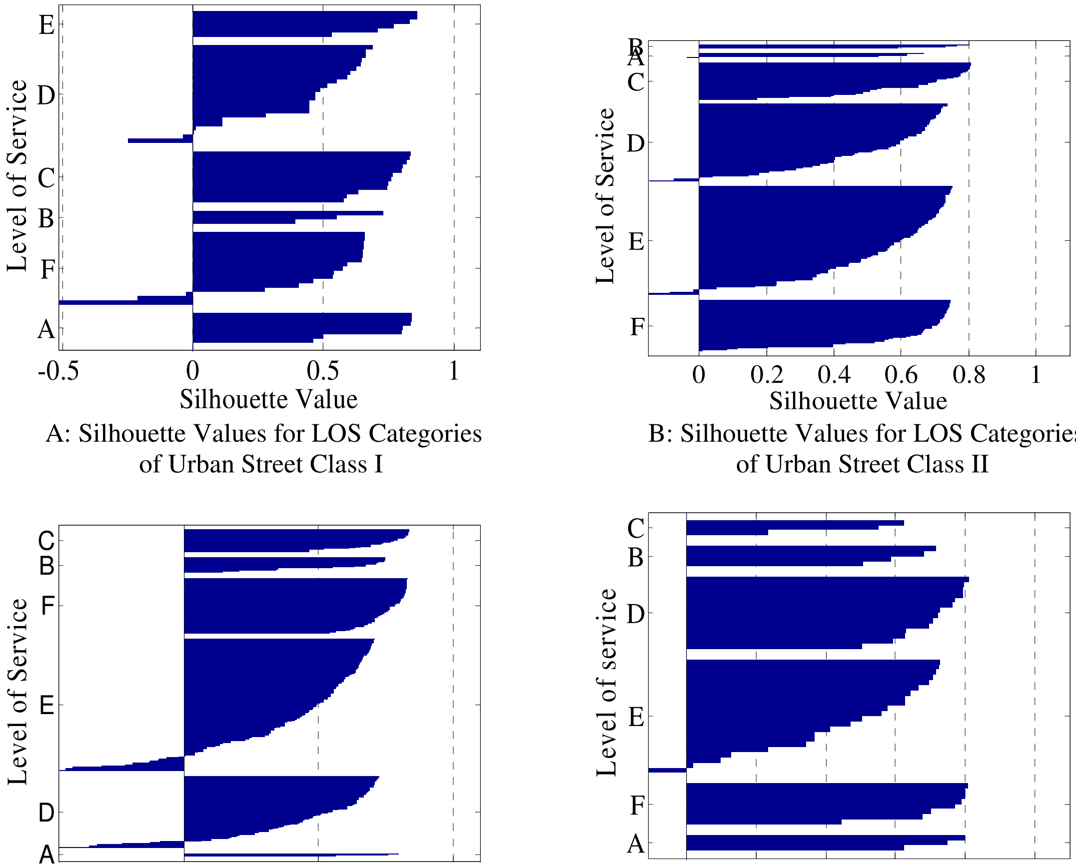


Figure 5: HAC of Average Travel Speeds for Level of Service of Urban Street Classes (I-IV).

Table 1: Urban Street Speed Ranges for different LOS Proposed for Indian Conditions using Hierarchical Agglomerative Clustering.

Urban Street Class	*I*	*II*	*III*	*IV*
Range of Free Flow Speed (FFS)	68 to 85 km/h	54 to 68 km/h	38 to 54 km/h	25 to 38 km/h
Typical FFS	75km/h	60km/h	50km/h	35 km/h
LOS	Average Travel Speed (Km/h)			
A	>69	>58	>49	>30
B	>59-69	>50-58	>40-49	>23-33
C	>48-59	>40-50	>34-40	>18-23
D	>37-48	>30-40	>23-34	>14-18
E	>29-37	>19-30	>11-23	>9-14
F	≤29	≤19	≤11	≤9

It can be inferred from Figure 6 that reasonable good classification has been found for urban street classes and level of service categories. Looking at this figure it is interpreted that vehicles move at congested level of service of “D”, “E” and “F” on all these street classes more frequently because silhouette heights are more.



A: Silhouette Values for LOS Categories of Urban Street Class I

B: Silhouette Values for LOS Categories of Urban Street Class II

C: Silhouette Values for LOS Categories of Urban Street Class III

D: Silhouette Values for LOS Categories of Urban Street Class IV

Figure 6: Silhouette Values for LOS Categories of Urban Street Classes (I-IV).

In order to check the application of this level of service criteria; data collected in Kolkata city were tested. Free flow speed and average travel speed during both peak and off-peak hours on each of segments on both corridors were calculated. The street segments were classified into four classes based on free-flow speed, geometric and surrounding environmental characteristics. Also, levels of service provided by the street segments during peak and off peak hours were estimated using Table 1 shown above. The percentage of travel runs under different levels of service categories found for urban street classes in Kolkata city during the survey period are shown in Table 2. From this table it has been observed that the observed vehicle traveled at better quality of service under urban street class I, whereas under other urban street classes the observed vehicle traveled at medium quality of service during the observed period. Average travel speeds for level of service categories expressed in percentage of free- flow speeds are calculated and the same are shown in Table 3. Values found using HAC method for the present context is compared with those values mentioned in HCM (2000) and IRC (1990). From this table it is found that there lies reasonable difference in the percentage values valid for Indian condition and those values mentioned in HCM (2000) and IRC (1990).

Table 2: Percentage of travel runs under different levels of service categories for urban street classes in Kolkata city.

Level of Service	*Urban Street Class*			
	I	*II*	*III*	*IV*
A	52.78	0.00	3.92	15.56
B	30.56	22.22	11.76	24.44
C	8.33	22.22	21.57	22.22
D	2.78	22.22	31.37	17.78
E	2.78	27.78	19.61	6.67
F	2.78	5.56	11.76	13.33

Table 3: Comparison of Percent FFS Values for each LOS Categories as obtained using HAC Method.

Level of Service	*% FFS (HCM, 2000)*	*% FFS (IRC, 1990)*	*% FFS (HAC Method)*	*Typical % FFS (HAC Method)*
A	90	90	85-95	90
B	70	70	75-90	80
C	50	50	60-75	70
D	40	40	45-60	55
E	33	33	35-45	40
F	25-33	25-33	20-35	20-35

6. Conclusion

Similar to the classifications adopted in HCM (2000) and satisfying to local conditions, the urban streets in the present context are classified into four classes and free-flow speed ranges were fixed for each urban street class. Secondly, cluster analysis is applied on both peak and off-peak hour traffic speed data and speed ranges for level of service categories were found out. It has been found from this study that urban street speed-ranges valid in Indian context are different from the corresponding values mentioned in HCM (2000). These differences are due to the heterogeneous nature of traffic flow along with varying geometric characteristics of road sections. From this result it has been found that typical average travel speed for LOS categories expressed in terms of the percentage of free flow speed for level of services “A”, “B” and “C” are comparatively higher to that values mentioned in HCM (2000). But these good qualities of service happen for lesser period of time on these study corridors. The probe vehicle travelled at level of service of “D”, “E” and “F” for more frequently. The limitation of this study is that we need large number of speed data points in order to get better result in classifications.

References

Bensaid, A.M., Hall, L.O., Bezdek, J.C., Clarke, L.P., Silbiger, M.L., Arrington, J.A. and Murtagh, R.F. (1996) "Validity - guided (re) clustering with applications to image segmentation", *IEEE Transactions on Fuzzy Systems* 4(2): 112-123.

Bezdek, J.C. and Pal, N.R. (1998) "Some new indexes of cluster validity", *IEEE Transactions on System, Man, and Cybernetics-Part B: Cybernetics* 28(3): 301-315.

Cheol, O. and Stephen, G. R. (2002) "Real-time inductive-signature-based level of service for signalized intersections", *Transportation Research Record 1802*, TRB, National Research Council, Washington, D.C.: 97-104.

"Highway Capacity Manual" (2000) *Special Rep. 209*, Transportation Research Board, Washington, D.C.

Indian Road Congress (1990) *Guidelines for capacity of urban roads in plain areas*, IRC: 106, New Delhi.

Jensen, S.U. and Trafitec, A. (2007) "Pedestrian and bi-cycle level of service on roadway segments", *86th annual meeting of the Transportation Research Board*, National Research Council, Washington, D.C.: 1-18.

Kikuchi, S. and Chakroborty, P. (2006) "Framework to represent uncertainty when level of service is determined", *Transportation Research Record 1968*, TRB, National Research Council, Washington, D.C.: 53-62.

Lingras, P. (1995) "Classifying highways: Hierarchical grouping versus kohonen neural networks", *Journal of Transportation Engineering*, ASCE, 121(4): 364-368.

Lingras, P. (2001) "Statistical and genetic algorithms classification of highways", *Journal of Transportation Engineering*, ASCE, 127(4: 237-243.

Madanat, S.M., Cassidy, M.J. and Ibrahim, W.H.W. (1994) "Methodology for determining level of service categories using attitudinal data", *Transportation Research Record 1457*, TRB, National Research Council, Washington, D.C.: 59-62.

Maitra, B., Sikdar, P.K. and Dhingra, S.L. (1999) "Modeling congestion on urban roads and assessing level of service", *Journal of Transportation Engineering*, ASCE, 125(4): 508-514.

Marwah, B.R. and Singh, B. (2000) "Level of service classification for urban heterogeneous traffic: A case study of Kanpur Metropolis", Paper presented at the fourth international symposium on Highway Capacity, Hawaii, June-July 2000.

Muraleetharan, T. and Hagiwara, T. (2007) "Overall level of service of the urban working environment and its influence on pedestrian route choice behavior: analysis of pedestrian travel in Sapporo, Japan", *86th annual meeting of the Transportation Research Board*, National Research Council, Washington, D.C.: 1-16.

Romana, M.G. and Perez, I. (2006) "Measures of effectiveness for level-of-service assessment of two-lane roads: An alternative proposal using a threshold speed", *Transportation Research Record 1988*, TRB, National Research Council, Washington, D.C.: 56-62.

Xu, J., Wong, S.C., Yang, H. and Tong, C.O. (1999) "Modeling level of urban taxi services using neural network", *Journal of Transportation Engineering*, ASCE, 125(3): 216-223.

Zolnik, E.J. and Cromley, E.K (2007) "A Poission-multilevel, bicycle level-of-service methodology for rural road networks", *86th annual meeting of the Transportation Research Board*, National Research Council, Washington, D.C: 1-13.

European Transport \ Trasporti Europei n. 49 (2011): 53-74

Does the transport industry gain from manufacturing internationalization? An empirical investigation on the Italian regions

Stefano Elia [1], Elena Maggi [2], Ilaria Mariotti [3*]

[1] DIG-Politecnico di Milano, P.zza L. da Vinci, 32 – 20133 Milan (Italy)
ph. +39 02 2399 2756, fax +39 02 2399 2710; stefano.elia@polimi.it
[2] Università del Molise, Faculty of Economics, SEGeS, via De Sanctis, Campobasso (Italy)
ph. +39 0874 404992; elena.maggi@unimol.it
[3]DiAP-Politecnico di Milano, P.zza L. da Vinci, 32 – 20133 Milan (Italy)
ph. +39 02 2399 3928, fax +39 02 2399 4105; ilaria.mariotti@polimi.it

Abstract

The present paper deals with the impact of manufacturing internationalisation, in the forms of international trade, cooperation agreements – measured by inward and outward processing trade (IPT and OPT, respectively) - and FDI, on the transport industry employment. Descriptive statistics and econometric analysis are carried out at the "regional-industry" level (20 NUTS2 regions and 8 transport sub-industries) with reference to Italy in the period 1996-2001. Results show that export, FDI and the components of IPT (temporary import and re-export) positively affect the transport employment variation in 1996-2001, while import and the components of OPT (temporary export and re-import) display a negative impact.

Keywords: Transport industry; Employment; Regional-industry; Trade; FDI; IPT; OPT.

1. Introduction

The globalisation of the economy, which has been fostered by the trade barriers reduction and the falling transport, communication and co-ordination costs (Krugman *et al.*, 1995; Glaeser and Kohlhase, 2004), has changed the structure of the production processes from being concentrated in one plant to being fragmented in different plants and in different countries. This has fostered a tight increase of exchange flows, which do not only include final goods, but also intermediate and unfinished goods, being transferred from one country to another in order to be processed either by an affiliate or

* Corresponding author: Ilaria Mariotti (ilaria.mariotti@polimi.it)

an external firm. The impact of internationalisation of manufacturing activities on the home country is highly debated in literature, both on the internationalised firm and on the national and – in a few cases - regional or local economic systems.

The present paper belongs to this literature focusing on a specific industry (the transport one) at the regional economic unit. Specifically, it aims at adding to this literature in three ways. Firstly, it focuses on the effects of manufacturing internationalisation on the transport industry, by looking at changes in the labour demand in Italy in the period 1996-2001. The main hypothesis to be tested is that manufacturing internationalisation, inducing a high increase of goods' flows to be moved, implies a reorganization of the supply chain, leading the manufacturing firm to outsource transport and other logistics services to specialised operators.

Secondly, since different forms of internationalisation may give birth to different effects on the labour demand in transport industry, the present paper investigates simultaneously the impact of manufacturing activities in terms of trade, cooperation agreements and FDI. International trade (i), which consists of import and export, is the most common form of internationalisation and the first entry mode adopted by a firm facing the global scenario, because it implies low involvement and risk degree for the internationalised firm. It consists of the exchanges of final and intermediate goods and services through the national borders. Cooperation agreement (ii), a more advanced and risky strategy than international trade, is mainly adopted by small and medium size enterprises (SME) because it does not require capital investment and is of short – medium term. The cooperation agreement is entered into by a firm and a foreign partner operating backward, forward or in the same stage of a value chain (Ietto-Gillies, 2005). This agreement refers to the development, distribution, and/or manufacture of goods to be sold in the foreign market. It is a non-equity strategy because it is developed through agreements (licensing, franchising, alliances, subcontracting) between a firm and one (or more) of its suppliers or distributors in order to supply, manufacture, or distribute goods and/or services without equity sharing. Finally, FDI (iii) represents the most articulated and binding mode to enter the foreign markets, because it requires a significant capital investment through greenfield or mergers and acquisitions and imply a medium-long term obligation. FDI is the main tool adopted by medium and large sized firms, which aim to share the capital of a foreign firm, eventually with one or more partners. FDI gives birth not only to intra-firm trade, but also to further exchanges between the host and the home country since the affiliates of the parent company establish economic relationships with home and host countries' suppliers and distributors.

Finally, a third contribution arises from the geographic unit of our analysis. Several authors in the literature (among the others, Cusmano *et al.*, 2009; Illeris, 2005) have stressed that outsourcing has a clear and predominant regional dimension. Italian SME tend, indeed, to organise the supply chain on specific local systems – named industrial districts – allowing the exploitation of agglomerative advantages and the capture of the efficiency of proximity between suppliers and users (Boix and Galletto, 2009; Mariotti *et al.*, 2008). Moreover, the geographical proximity has a strong influence on the selection of the transport providers also for the not-district manufacturing firms (Isfort, 2003; Razzaque and Sheng, 1998; Peters *et al.*, 1998), even because it helps in cutting the logistics costs. Besides, the cooperation between the shipper (manufacturing firm) and the external company (transport provider) is considered as a strategic issue (European Commission, 2000), and this is easier when the two firms are closely located.

Therefore, the analysis of the present paper is run at "regional-industry" level, that is 20 NUTS2 regions (from now on "regions") and 8 transport sub-industries. Regional dimension can be considered, indeed, the most appropriate geographical unit taking into account such strategic role of geographical proximity between the manufacturing and the transport firms. The region is large enough to capture the district dimension, given that industrial districts are frequently located within regions.

The paper is structured into six sections. The introduction is followed by the literature review on the impact of internationalisation on transport employment. Section three focuses on data and methodology. Descriptive statistics and econometric analysis are presented and discussed in sections four and five, respectively. Conclusions and policy recommendations follow.

2. The impact of internationalisation on the home country employment

The issue of the impact of manufacturing internationalisation on the home country employment is highly debated and analysed by several theoretical and empirical studies (among the others, Dunning and Lundan, 2008; Molnar *et al.*, 2007). In contrast with the general public view, the main findings of the works focusing on the OECD countries show that the impact of internationalisation on aggregate labour market is small and the domestic job losses are slight, although particular skill and occupational groups (especially low skill level) have been affected more strongly (Crinò, 2009; Barry and Walsh, 2008).

However, to the best of our knowledge, lack of attention has been devoted to the regional impact as well as to the simultaneous impact of the three above described internationalisation forms. Besides, there is no evidence at regional level on the internationalisation effects on the employment of specific industries supplying services to manufacturers, such as the transport industry.

Within this context, it is possible to identify three different strands of literature analysing separately the impact of trade, cooperation agreements and FDI on labour demand. As concerns the trade literature, many studies focus on the employment and wage impact on a national scale, generally finding changes in the labour composition (high skilled and low skilled workers) and, in some circumstances, a decline in the relative demand for low skilled employees, especially in the industries facing import competition (Krugman *et al.*, 1995, Addison *et al.*, 2000). Nevertheless, recent studies stress that while the effect of increased imports on jobs is generally negative and the impact of increased export is positive, the overall effects of increased trade are positive (Kletzer, 2002). In particular, trade has had no significant effect on the overall unemployment rate of the OECD countries (Hill *et al.*, 2008). Conversely, the few studies on the regional scale highlight that the magnitude of trade's impact on the labour markets remains widely contrasting (Richardson, 1995; Kapstein, 2000). Indeed, the effects of export growth on the employment tend to be mixed, either positive or negative, depending on different regional characteristics and dynamics, such as the regional size, the industrial structure and the trade patterns (Baldwin and Brown, 2004; Leichenko and Silva, 2004; Markusen *et al.*, 1991).

The literature on cooperation agreements stresses the impact on the national employment, mainly focusing on its composition and wage level. For example, a study on two European Countries (Italy and Germany) (Helg and Tajoli, 2005) estimates that, during the 1990s, this strategy has increased the high skilled-to-low skilled labour ratio in Italy, while it has not affected the German demand for high skilled labour. Egger and Egger (2001), by focusing on cooperation agreements between manufacturing industries of EU and non EU-countries during the period 1995–1997, find that they have reduced the skill-to-low-skill ratio in EU exporting industries, while have had more ambiguous effects in import-competing industries.

The literature on FDI, instead, has mainly investigated the effects on the employment at firm level (among the others, Castellani *et al.*, 2008 on the Italian case; Head and Ries, 2002 on the Japanese MNE) or at domestic industry level (among the others, Slaughter, 2000 on the USA; Falzoni and Grasseni, 2003 on Italy). Conversely, few studies have focused on the effects on both the internationalised firm and its economic environment, adopting the NUTS2 region or the NUTS3 province as scale of analysis (Mariotti *et al.*, 2003; Elia *et al.*, 2009; Federico and Minerva, 2008). Specifically, Mariotti *et al.* (2003) find that in the period 1985-1995 the Italian FDI have significantly affected the labour intensity of the domestic production at a "regional-industry" scale, defined as the ensemble of firms operating in the same industrial macro-industry – composed of interdependent sectors belonging to the same filière – and localised in the same geographical region. In particular, the impact is negative for investments undertaken in less developed countries and positive for market-seeking investments in advanced countries. Elia *et al.* (2009), by adopting the same unit of analysis (regional-industry), investigate the impact of outward FDI upon the demand for high and low skilled workers in Italy throughout the period 1996-2002. It results that all outward FDI - regardless of the country of destination - have significant negative effects on the demand for low skilled workers, while outward FDI towards OECD countries negatively affect the demand for high skilled workers. Federico and Minerva (2008), who assess the impact of Italy's outward FDI on local employment growth in 1996-2001 for 12 manufacturing industries and 103 administrative provinces, find that net effect of FDI on the employment of the whole local area is positive. Employment growth in local areas investing more abroad appears to be stronger than the industry average growth, especially in some sectors.

At least to our knowledge, only Mariotti and Piscitello (2007) devote attention to the effects on the tertiary sector, by investigating the impact of the manufacturing FDI on the labour demand for services between 1996 and 2003 in the industrial districts of Veneto region in the North-East of Italy. The finding is that the more internationally involved an industrial district, the higher the employment growth in the service sector.

The literature concerning the specific impact of internationalisation of manufacturing activities on the transport industry is even more scanty. The internationalisation of production and the growing global trade rate highly affect the logistics activities, especially transport (Maggi *et al.*, 2008). Transport plays a key role in connecting the different import and export markets and the vertically disaggregated components of production system, which are widespread in the world (Yieming *et al.*, 2002). Besides, the increasing large share of goods flows, fostered by internationalisation, must be managed by transport functions at the level of both the manufacturing firm and the transport providers. In the first case, when transport is carried out by the manufacturing firm (insourcing), there is a rise of the labour demand for workers specialised in planning, managing and controlling the transport activities (Blomström *et al.*, 1997). In the second case, that is when transport is outsourced, the transport operators restructure themselves becoming larger and changing into the so-called Third-Party Logistics (3PLs) or Fourth Party Logistics (4PLs) Providers, by modifying their supply from single based services into an increasing number of integrated high value-added services (for example, transport, storage, inventory management, tracking and tracing, packaging, labelling and secondary assembly of products) (Brewer *et al.*, 2001).

Only in the second case it is possible to observe an increase of labour demand in the transport industry, which is the object of analysis of the present paper. In fact, while the rate of outsourcing of integrated logistics as a whole is very low within the Italian boundaries (13% in 1997 and 16% in 2004 - Commission Européenne, 2001 on data AT Kearney-ELA), the transportation is outsourced by the majority of the Italian manufacturing firms that become international through one of the three forms identified above, i.e. trade, cooperation agreements and FDI (Confetra, 2002).

In literature it is possible - at least to our knowledge - to identify only two papers providing some evidences on the impact of internationalisation on transport services. The first one (Savona and Schiattarella, 2004) investigates the impact of cooperation agreements, measured by a specific index based on trade data, on different services' labour demand within the Italian NUTS3 provinces over the period 1991-1996. The authors conclude that internationalisation towards low wage countries shows a positive significant impact on the more traditional services' employment, such as transport.

The second (Maggi *et al.*, 2008) provides evidence on the relationship between FDI, undertaken by the industrial district firms located in Veneto region, and the employment change in the transport and other logistics industries, occurred in the same industrial districts between 1996 and 2003. Results show that, although all the internationalised industrial districts exhibit an increase in the logistics labour demand, only in a few of them an internationalisation degree above the average is positively correlated to a logistics' employment growth.

3. Data and methodology

3.1 Data

The impact of internationalisation on the labour demand in transport industry has been investigated in the present paper by using the employment variation in the 8 transport sub-industries of each Italian region as dependent variable (Table 1). Data about the employees are provided by the Manufacturing and Services Activities' (MSA) Census of the Italian Statistical Institute (ISTAT) at six digit level and refer to the category "I – Transport, warehousing, communications"[1]. The sub-industries considered in the

[1] Data on the MSA national Census for the years 1996 and 2001 are classified following the NACE REV 1. The NACE six digit structure allows to distinguish the share of employees working in the freight transport activities from the share involved in passengers transport, with the exception of rail, maritime and air transport modes. For these three sub-industries, the rate of freight employees has been estimated by using the Amadeus database. The authors thank Aleid Brouwer of the University of Groningen for the cooperation in collecting the data.

analysis concern the different transport modes (rail, road, sea and air[2]) and their supplying and supporting transport activities (i.e. goods loading and unloading and stevedoring) (Table 1). The employment variation is computed between the years 2001 and 1996, which correspond to the last two MSA Census provided by ISTAT.

Table 1: Italian NUTS2 regions and the NACE REV. 1 transport sub-industries.

Regions	*Transport sub-industries*	*NACE Codes*
Abruzzo	**Land transport; transport via pipelines**	**60**
Basilicata	Transport via railway	60.10.1
Calabria	Auxiliary activities to transport via railway	60.10.2
Campania	Transport by road	60.25.0
Emilia Romagna	**Water transport**	**61**
Friuli Venezia Giulia	Sea transport	61.11.0
Lazio	**Air Transport**	**62**
Liguria	Non-scheduled air transport	62.20.0
Lombardy	**Supporting and auxiliary transport activities; activities of travel agencies**	**63**
Marche	Cargo handling – Air transport	63.11.1
Molise	Cargo handling – Water Transport	63.11.2
Piedmont	Cargo handling – Land Transport	63.11.3
Puglia		
Sardinia		
Sicily		
Tuscany		
Trentino Alto Adige		
Umbria		
Valle d'Aosta		
Veneto		

As explained in the previous sections, the three forms of internationalisation taken into account are trade, cooperation agreements and FDI. Data on trade come from the dataset on the Italian Trade (Coeweb) provided by ISTAT and are expressed in terms of kilograms by transportation modes. Data refer to the years 1996 and 2001 and are expressed in terms of difference (Δ_{96}^{01}) (further details are provided in Appendix 1).

Cooperation agreements are difficult to measure: several papers use, as proxy, imports of intermediate inputs, estimated by combining Input–Output tables and final import data (among the others, Feenstra and Hanson, 1996; Minondo and Rubert, 2006); few papers use the data on processing trade (PT) (Helg and Tajoli, 2005; Egger and Egger,

[2] The scheduled air transport industry (62100) has been excluded from the analysis because it mainly concerns passenger air transport.

2001). Because of data availability, in the present paper cooperation agreements are measured in terms of PT, which is trade in goods being exported (or imported) for reason of processing abroad and subsequently re-imported (or re-exported) with favorable tariff treatment. PT is composed by four elements: (a) temporary exports of goods exported by a EU country to be processed in a non-EU member and (b) re-imports by the EU of the processed goods, on the one hand; (c) temporary imports of goods to be processed in the EU and (d) re-exports of those goods to the country of origin outside the EU, on the other hand. The first two flows measure the so-called Outward Processing Trade (OPT); the last two measure Inward Processing Trade (IPT) (Baldone *et al.*, 2006). Data about the four PT components are provided by ISTAT as total amount of kg but not by transport mode. Therefore, they have been distributed among the 8 transport sub-industries by employing the ratios of trade for each transport mode. Data refer to the years 1996 and 2001, and are expressed in terms of differences (Δ_{96}^{01})[3]. Therefore, the final dataset appears as a cross section where each observation refers to a specific combination of a NUTS 2 regions and a transport sub-industry, for a total of 160 observations (20 regions*8 transport sub-industries) (further details are provided in Appendix 1).

Outward FDI undertaken by the Italian manufacturing MNE are measured as the cumulated sum of the employees in their foreign affiliates from 1994 to 2000. The lag between FDI and employment stands on the hypothesis that foreign affiliates need one-two years time to fully work and this delay is necessary to the transport suppliers to reorganise their activities in order to satisfy the new customers' demand. Since the goods related to FDI are transported by the different transport modes, FDI data have been distributed across the 8 transport sub-industries by using the same percentages of export and import. The data source for the Italian outward FDI is Reprint dataset, which is developed by the Department of Management, Economics and Industrial Engineering of the Politecnico di Milano and is sponsored by the Italian Institute for Foreign Trade – ICE (further details are provided in Appendix 1).

Finally, the production of manufacturing firms has been considered as explicative variable that may also affect the transport industries' employment. Indeed, a higher production is likely to imply a higher demand for transport. The production of manufacturing firms is measured in terms of value added (Berman *et al.*, 1994), which is expressed at current prices for the years 1996 and 2001 and comes from ISTAT. Besides, data about value added have been distributed among the eight transportation sub-industries according to the same percentages calculated for the components of IPT and OPT. Finally, value added has been expressed in terms of Balassa index for each pair of industry/region in order to account for the relative amount of manufacturing

[3] The components of IPT and OPT have been analysed separately because each component may involve different transport operators located either in the domestic or in the foreign country, with different effects on labour demand in the Italian transport industry. Furthermore, both IPT and OPT typically require between six months and more than one year to be completed, since the processed goods need first to be temporary imported (or temporary exported), then to be transformed and finally re-exported (or re-imported). The data of the present paper refer to two single years (1996 and 2001) and are distributed among different months. As a consequence, most of data concerning re-import and re-export of 1996 and 2001 are likely to refer to processes that have started in 1995 and 2000, respectively, while data on temporary export and temporary import are likely to refer to processes that have started in 1996 and 2001 and have finished in 1997 and 2001, respectively.

goods transported by each transportation sub-industry in each region with respect to the other transport sub-industries and to the other regions (further details are provided in Appendix 1).

3.2 Methodology

The methodology employed to test the impact of internationalisation on the transport employment consists of an econometric analysis, which aims at estimating the following equation:

$$\Delta_{96}^{01} Emp_{r,s} = \alpha_{s,r} + \beta_1 \Delta_{96}^{01} Y_{r,s}^m + \beta_2 \sum_{t=94}^{00} FDI_{r,s}^m + \beta_3 \Delta_{96}^{01} I_{r,s}^m + \beta_4 \Delta_{96}^{01} X_{r,s}^m + \beta_5 \Delta_{96}^{01} T\_I_{r,s}^m + \beta_6 \Delta_{96}^{01} R\_X_{r,s}^m + \beta_7 \Delta_{96}^{01} T\_X_{r,s}^m + \beta_8 \Delta_{96}^{01} R\_I_{r,s}^m + \varepsilon_{r,s}$$

where s are the 8 transport sub-industries, r the 20 NUTS2 regions, m the whole manufacturing industry[4].

$\Delta_{96}^{01} Emp_{s,r}$: is the variation of the number of employees between 2001 and 1996, by each sub-industry s and region r;

$\Delta_{96}^{01} Y_{r,s}^m$: is the difference between 2001 and 1996 of the value added of manufacturing production (m) transported by each sub-industry s of each region r (computed as Balassa Index);

$\sum_{t=94}^{00} FDI_{r,s}^m$: is the sum of FDIs cumulated between 1994 and 2000 by manufacturing industries (m), distributed among the sub-industries s of each region r;

$\Delta_{96}^{01} I_{r,s}^m$: is the difference between 2001 and 1996 of the manufacturing goods' import flows (m) transported by each sub-industry s of each region r ;

$\Delta_{96}^{01} X_{r,s}^m$: is the difference between 2001 and 1996 of the manufacturing goods' export flows (m) transported by each sub-industry s of each region r;

$\Delta_{96}^{01} T\_I_{r,s}^m$: is the difference between 2001 and 1996 of the manufacturing goods' temporary import flows (m) transported by each sub-industry s of each region r;

$\Delta_{96}^{01} R\_X_{r,s}^m$: is the difference between 2001 and 1996 of the manufacturing goods' re-export flows (m) transported by each sub-industry s of each region r;

$\Delta_{96}^{01} T\_X_{r,s}^m$: is the difference between 2001 and 1996 of manufacturing goods' temporary export flows (m) transported by each sub-industry s of each region r;

$\Delta_{96}^{01} R\_I_{r,s}^m$: is the difference between 2001 and 1996 of the manufacturing goods' re-import flows (m) transported by each sub-industry s of each region r.

The equation suggests that the variation of the demand for transport workers between 2001 and 1996 within a region, is related to: (i) the change of the manufacturing industries' value added, (ii) the cumulated sum of the manufacturing FDI undertaken in 1994-2000, (iii) import and export change, and (iv) change in the components of IPT

[4]The manufacturing industry is identified by the NACE codes from 15 to 45.

and OPT. Finally, region and industry dummies have been included in order to control for fixed effects.

4. Descriptive statistics on the transport industry employment growth

Between 1996 and 2001, the employees in the transport sub-industries grew of 22.32%, which corresponds to an absolute increase of 78,785 units (see Table 2). All the transport sub-industries increased in terms of employment, with the exception of rail and air transport supporting activities, which are the less used because of the low speed and reliability of the Italian railway transport network, and because of the high cost of air with respect to the other transport modes. Furthermore, the non-scheduled air transport also presents a growth (2%), even if smaller than the other sectors.

The highest increase has occurred in the cargo handling for land transport sub-industry (63113), which displays both the highest percentage (+109.97%) and absolute employment variation (+47,885). Given that the railway sectors show negative values (-44.10% and -8,768 in 60101; -56.91%; and -707 in 60102), the employees' increase in the cargo handling for land transport (63113), which is vertically integrated with both railway and road transport, must be linked to the positive performance of road transport (60250), which displays an increment of +39,240 employees, corresponding to +14.58%.

Table 2: Employment change in terms of percentages and absolute values in the transport sub-industries (1996-2001).

Transport sub-industries	*NACE codes*	*Growth rate 96-01 (%)*	*Absolute variation*
Transport via railway	60.10.1	-44.10	-8,768
Auxiliary activities to transport via railway	60.10.2	-56.91	-707
Transport by road	60.25.0	14.58	39,240
Sea transport	61.11.0	37.53	2,091
Non-scheduled air transport	62.20.0	2.08	5
Cargo handling– Air transport	63.11.1	-62.41	-2,240
Cargo handling– Water Transport	63.11.2	13.14	1,279
Cargo handling– Land Transport	63.11.3	109.97	47,885
Total	(All)	22.32	78,785

The predominance of road transport is due to the fact that more than 50% of the manufacturing activities' internationalisation (with the exception of re-export) occurs towards Europe (Table 6 in the Appendix 2). As a consequence, since the Italian railway system is less efficient for goods transport, most of the land freight flows are moved by road. According to Confetra (2004), in 2002 about the 70% of export towards Europe has been managed by road transport. On the other hand, when internationalisation is addressed towards non-EU destinations, the most preferred mode becomes the maritime transport, which involves the sea transport (61110) and cargo handling for water transport (63112). These sub-industries exhibit a positive performance in terms of employment growth and absolute variation (+37.53% and +2,091 employees in 61110; +13.14% and +1,279 employees in 63112). Conversely, air transport is less used to support internationalisation: non-scheduled air transport (62200) shows a very low relative and absolute variations (2.08% and +5, respectively), and cargo handling for air transport (63111) even shows a strong employment reduction (-62.41%, which corresponds to -2,240 workers).

Summarizing, descriptive statistics on employment change in transport sub-industries provides a contrasting picture. While water and road transport have strongly increased their employees, the opposite is true for air and railway transport. This picture does not help to formulate any insights on the relationship between internationalisation of manufacturing activities and the employment in transport industry. Therefore, some further descriptive statistics, taking into account the geographical dimension of transport employment, follow.

Table 3: Employment change in transport sub-industries in terms of percentage and absolute values in the Italian regions (1996 and 2001).

Regions	*Employment Growth Rate*	*Employment Absolute Variation*
Centre	**24.91**	**18,210**
ABRUZZO	37.61	2,347
LAZIO	29.75	7,813
MARCHE	16.16	1,742
TUSCANY	25.01	5,845
UMBRIA	7.19	463
North-East	**20.80**	**20,227**
EMILIA ROMAGNA	15.90	6,966
FRIULI VENEZIA GIULIA	9.81	917
TRENTINO ALTO ADIGE	20.85	1,519
VENETO	29.44	10,825
North-West	**28.27**	**29,996**
LIGURIA	6.02	892
LOMBARDY	37.54	22,790
PIEDMONT	20.86	6,253
VALLE D'AOSTA	6.27	32
South and islands	**13.54**	**10,382**
BASILICATA	28.36	564
CALABRIA	20.47	1,327
CAMPANIA	24.65	5,829
MOLISE	25.45	316
PUGLIA	9.85	1,663
SARDINIA	9.65	887
SICILY	-1.18	-204
Total	**22.32**	**78,785**

Table 3 shows the employment change in terms of percentage and absolute variation across the 20 different Italian regions, while Table 6 in the Appendix 2 shows the distribution of each internationalisation form among the 20 Italian regions. In both tables, the Italian regions have been grouped into four macro-areas (North-West, North-East, Centre, and South and Islands), according to their economic characteristics. Figure 1 in the Appendix provides a visualization of the total employment change of the transport sub-industries, which has occurred between 1996 and 2001 in each Italian region.

Table 3 shows that the north-western regions report the best performance in terms of employment growth: Lombardy (the Italian leading economic region) displays the highest absolute variation (+22,790) and the second highest percentage variation (+37.54%), while Piedmont presents a noticeable increase, especially in terms of percentages (+20.86). The significant transport employment increase in the North-West, and specifically in Lombardy, is strictly related to the regional dimension of outsourcing and to the significant internationalisation degree of the area. A recent firm level study on Lombardy (Cusmano *et al.*, 2009) has, indeed, shown that, on average, more than 40% of the firms refer to a regional supplier for some of the functions they have decided to contract out. Besides, the 53.46% of outward FDI originates in the North-West, with 34.71% in Lombardy, and 17.94% in Piedmont (Table 6). The North-West is also the most active as regards export and import, even if the gap with respect to the other three areas is low. Besides, most of the internationalisation originating in the North-West is undertaken in Europe (63.45% of FDI, 78.61% of export and 73.25% of import), and this may explain the high increase of transport by road displayed by Table 2.

As concerns the north-eastern regions a high employment increase occurs in Veneto, Emilia Romagna and Trentino Alto Adige (+29.44%, +15.90% and +20.85%, respectively), with Veneto exhibiting the second highest absolute variation (+10,825) (Table 3). Indeed, the North-East is highly internationalised, being placed in the second position as regards FDI, export, import and temporary export and in the first position as concerns re-import (Table 6). The significant internationalisation degree and the high employment increase in this area can be explained by the presence of about one third of the Italian industrial districts, which are specialised in the made in Italy sectors (fashion, home furniture and furnishing, food and traditional mechanical engineering products), and which have started a strong internationalisation process since the middle of the ‘80’s. Furthermore, most of these flows occurs towards (or from) Europe, especially as regard re-import (84.99 %) (Table 6), and this trend may contribute to explain the land transport employment increase (Table 2).

The Centre displays high employment growth in terms of percentages (Abruzzo +37.61%, Lazio +29.75%, Tuscany +25.01% and Marche 16.16%) (Table 3). However, the absolute variations are, on average, lower than those of the north-western and north-eastern regions. Specifically, Lazio exhibits a high performance in absolute value (+7,813) probably because it hosts the Italian capital city Rome, where several transport firms choose to locate their headquarters. Firms from the Centre, which are mainly SME, show low levels of internationalisation, since they are always placed in the third or fourth positions in all the internationalisation modes (Table 6). Therefore, in this case low levels of internationalisation are associated to middle-low employment increase.

Finally, the regions from the South and Islands display high transport employment growths in terms of percentage, but low increases in terms of the absolute variations, with the exception of Campania (+5,829) (Table 3). Southern regions have a limited manufacturing economic activity with respect to the rest of the country and are specialized in primary industries. The main internationalisation forms are export (especially of primary goods), whose percentage (25.49%) is similar to that of north-eastern regions (27.78%), and IPT, whose components display the highest percentages in the country (36.45% for temporary import and 69.41% for re-export) (Table 6).

Summarizing, the descriptive statistics underlines that, on average, the most internationalised geographical areas (i.e. North-West and North-East) and regions (i.e. Lombardy and Veneto), report the highest employment growth, especially as regards transport by road because the majority of the flows are addressed (or come from) Europe (Tables 2, 3 and 6). However, also geographic areas (e.g. Centre) and regions (e.g. Campania) that display a lower level of internationalization exhibit a middle increase of employees. At the same time, it is possible to observe some regions with middle levels of some internationalization forms (e.g. export of Sicily and Sardinia) that show low employment increase or even a decrease.

However, given that employment growth may be (is) the result of the variation of all internationalization forms, an econometric analysis is required in order to understand more punctually the relationship between each form of internationalisation and the variation of employment in transport industry.

5. Econometric findings

This section aims at verifying whether there is a positive relationship between the transport employment and the internationalisation process, in all its forms, undertaken by the Italian regions. In order to reach this goal, the equation presented in section 4 has been estimated through econometric regressions. Table 5 of the Appendix 2 provides the descriptive statistics and correlations of the explanatory variables that have been employed in the econometric analysis[5].

Results are reported in Table 4, which displays four columns corresponding to four different models: (i) without dummies, (ii) with regional dummies, (iii) with industrial dummies, (iv) with both regional and industrial dummies. Given that the Breusch-Pagan test displays that OLS regressions, which are not reported in this paper, is affected by heteroskedasticity, Weighted Least Square regressions have been used in order to assign a lower weight to those observations with high variance.

WLS regressions show that manufacturing production is not significant for the transport industries' labour demand. This might be due to the fact that most of Italian firms are SME, which mainly produce for local market. As a consequence, they are more likely to manage transportation inside (insourcing), since they have to transfer small-medium volumes of freight and have to cover short-medium distances (Confetra, 2002).

Conversely, firms that internationalise are typically larger than domestic uni-national firms (Barba Navaretti and Venables, 2004) and need to cover longer distances and to deal with customs and duties issues. Therefore, as the empirical literature stresses,

[5] In the econometric analysis all the variables have been standardized because they are expressed in different units of measure.

internationalised firms are more likely to outsource transport (Confetra, 2002). This is the reason why FDI and export turn out to be positively and strongly correlated with the transport industries' employment in all the four models. Transport is needed both to reach the final market, especially in case of export of final goods, and to reach the other affiliates or the suppliers or distributors, especially in case of FDI.

Besides, a negative relationship between import and employment in transport industry occurs (Table 4); this may be related to the fact that the imports of goods towards Italy are mainly managed by foreign suppliers (Confetra, 2002), thus generating a substitution effect. In other words, the increase of import may be partially related to a substitution of national goods with foreign goods. Considering that national goods are mainly transported by national firms while foreign goods are transported by foreign firms, the substitution of national goods with foreign goods is likely to generate a decrease of labour demand in national transport industry and an increase of labour demand in foreign transport industries.

As regards the IPT and OPT components, the econometric analysis finds that temporary import and re-export show a positive impact on the labour demand, while temporary export and re-import exhibit a negative coefficient, even if very weak and often not significant. Given that the former are the components of IPT and the latter of OPT, it seems to emerge a complementary impact of this internationalisation forms on the labour demand. Indeed, as concerns the IPT, the positive sign reveals that transportation of intermediate and unfinished goods is mainly undertaken by Italian firms, which temporary import raw material and/or semi-components to be processed in Italy and later re-export the processed goods abroad. The opposite is true when the Italian firms send the intermediate goods abroad to be processed, i.e. in case of OPT. Indeed, the negative signs of temporary export and re-import suggest that transportation is still managed by the firms of the country where the intermediate goods are processed. The final impact of OPT is, however, not very significant probably because there is not a substitution effect. Indeed, when firms do not internationalise through OPT, goods are entirely produced by the Italian firms and, hence, transport of intermediate goods abroad are not required. Conversely, when the OPT process starts, the production is fragmented generating a new flow of goods (managed by foreign firms) that did not exist before. Therefore, the final impact on the national transport industry employment is not (or not very much) significant. The same consideration holds for the strong positive impact of IPT: after starting IPT, a new flow of goods that did not exist before takes place by generating a labour demand in national transport industry.

Besides, the analysis has been run at regional-industry level because, as stated in the introduction, the assumption is that the logistics activities in Italy are mainly carried out at the regional scale. As shown in Table 4, the results are confirmed both without and with industrial and regional dummies, meaning that the relationship between internationalization and employment growth produces effects both at national and regional/industry levels. However, some differences arise in terms of intensity and significance of coefficients when considering data without and with dummies, meaning that the relationship between each form of internationalization and the labour demand in transport industry may loose or gain intensity when regional and industrial dimensions are taken into account.

Table 4: Results of the WLS estimations.

	$\Delta_{96}^{01} Emp_{s,r}$	*p-value*	$\Delta_{96}^{01} Emp_{s,r}$	*p-value*	$\Delta_{96}^{01} Emp_{s,r}$	*p-value*	$\Delta_{96}^{01} Emp_{s,r}$	*p-value*
	(1)		**(2)**		**(3)**		**(4)**	
$\Delta_{96}^{01} Y_{r,s}^{m}$	0.010	0.529	0.012	0.738	-0.030	0.207	-0.027	0.439
	(0.63)		(0.33)		(-1.26)		(-0.77)	
$\sum_{t=94}^{00} FDI_{r,s}^{m}$	1.273	0.000	1.400	0.000	1.018	0.000	1.177	0.000
	(17.23)		(8.51)		(12.33)		(7.16)	
$\Delta_{96}^{01} I_{r,s}^{m}$	-0.252	0.000	-0.561	0.000	-0.211	0.001	-0.269	0.005
	(-4.73)		(-5.97)		(-3.35)		(-2.78)	
$\Delta_{96}^{01} X_{r,s}^{m}$	0.993	0.000	0.911	0.000	0.688	0.000	0.780	0.000
	(18.43)		(7.86)		(9.90)		(6.99)	
$\Delta_{96}^{01} T\_I_{r,s}^{m}$	0.667	0.000	0.583	0.000	0.413	0.000	0.424	0.001
	(8.41)		(4.35)		(4.90)		(3.31)	
$\Delta_{96}^{01} R\_X_{r,s}^{m}$	-0.030	0.634	0.410	0.000	0.186	0.005	0.243	0.005
	(-0.48)		(4.59)		(2.81)		(2.79)	
$\Delta_{96}^{01} T\_X_{r,s}^{m}$	-0.168	0.000	-0.120	0.241	-0.239	0.000	-0.306	0.003
	(-3.50)		(-1.17)		(-4.08)		(-3.02)	
$\Delta_{96}^{01} R\_I_{r,s}^{m}$	-0.225	0.002	-0.005	0.966	-0.101	0.159	-0.088	0.452
	(-3.06)		(-0.04)		(-1.41)		(-0.75)	
_cons	0.561	0.000	0.488	0.000	-0.057	0.314	-0.001	0.994
	(19.02)		(3.66)		(-1.01)		(-0.01)	
DummyIndustry	No		No		Yes		Yes	
DummyRegion	No		Yes		No		Yes	
n. obs.	160		160		160		160	
Chi2	5183.62	0.000	638.29	0.000	1117.03	0.000	478.52	0.000

Note: Z-statistics between brackets.

6. Conclusions

In an era where internationalisation has become a strategic challenge that firms have to face in order to remain competitive, the analysis of the effects of this phenomenon has assumed a key role. It is widely acknowledged that the impact of internationalisation highly differs, according to the industry and the geographical context, which is taken into consideration. The present paper enriches the literature by focusing on the impact of manufacturing internationalisation on the transport industry labour demand, an industry which manages the growing flows at international scale. This is the first empirical work, at least to our knowledge, considering all the three internationalisation strategies (trade, cooperation agreements, in the form of IPT and

OPT, and FDI), undertaken by the manufacturing industry, and testing their specific impact on transport employment both through a descriptive statistics and an econometric analysis at the regional scale.

Our analysis shows that FDI and export are the main responsible for the increase of labour demand in transport industries, while IPT components have a weaker positive impact. Conversely, import and OPT have a negative (but weaker in terms of coefficients) effect. These results are very interesting also in terms of policy implications: internationalisation, at least as concerns export and FDI, fosters the growth of the transport services' demand and, therefore, an increase of the logistics providers' potential market. This means that when transport becomes more complex and strategic, as it happens within the internationalisation process, manufacturing firms tend to outsource the freight movements to specialised suppliers which can reach economies of scale and scope. Therefore, the industry gains in terms of employment and competitiveness. As a consequence, in order to satisfy the growing transport demand, there is a strong need to draw policies able to remove the existing obstacles that limit the development of the transport industry in Italy. We refer, in particular, to the inefficient and insufficient transport networks and the lack of culture about logistics within the Italian SME.

Investments in transport labour training, particularly at the scholar level, may be useful in order to improve the transport firms' competitiveness. Furthermore, policies aiming at increasing the rate of transport outsourcing, like the recent edition of the "Piano Nazionale della Logistica" (National Logistics Plan by the Italian Minister of Infrastructure and Transport - *Ministero delle Infrastrutture e dei Trasporti*, 2011) suggests, may be able to better exploit the opportunities offered by the manufacturing internationalisation process to the transport labour market. Transport outsourcing might be encouraged by policies aiming at reducing the transport providers' labour cost by fiscal or social security relief, and aiming at stimulating the development of the "network contracts" (*contratto di rete*), in order to aggregate the demand and supply, therefore overcoming the transport providers' small size and more easily gaining economies of scale and scope.

However, since the growth of the road transport produces significant negative externalities affecting the environment, specific policies aiming, from one side, at reducing those externalities and on the other side, at promoting environmentally friendly transport modes and logistics organisation, are advocated. For example, in order to gain competitiveness, the reliability of rail and road-rail combined transport should be improved by eliminating the bottle-necks and investing in infrastructure (railways and terminal); also the maritime transport should be promoted, by improving the efficiency of ports, reducing the length of the duty controls and the level of bureaucracy.

Further research is needed to investigate these issues, providing an answer to the following questions. Firstly, in countries where the foreign logistics MNE own a high share of the logistics market, such as Italy, does the employment growth concern both the domestic and foreign firms or mainly the last ones, which are larger and more competitive and offer integrated and high value added services?[6] Secondly, if a part of the internationalised manufacturing firms makes transport or other logistics activities inside, an investigation on the impact of internationalisation even on the manufacturing

[6] See, among the others, Brouwer and Mariotti (2009).

employment is advocated. Finally, given that industries are sometimes vertically integrated and display different characteristics, it would be interesting to distinguish the impact of internationalisation of different groups of manufacturing sectors on the transport industries' labour demand.

Acknowledgements

We thank the referees for the helpful comments and suggestions on drafts of this paper.

References

Addison, J.T., Douglas, A.F. and Ruhm, C.J. (2000) "Technology, Trade Sensitivity, and Labor Displacement", *Southern Economic Journal* 66(3): 682-699.

Baldone, S., Sdogati, F. and Tajoli, L. (2006) "On some effects of international fragmentation of production on comparative advantages, trade flows, and the income of countries", *CESPRI working paper*, 187, Milan.

Baldwin, J.R. and Brown, W.M. (2004) "Regional manufacturing employment volatility in Canada: The effects of specialisation and trade", *Regional Scienc* 83: 519-541.

Barba Navaretti, G. and Venables, A.J. (2004) *Multinational Firms in the world economy*, Princeton University Press, Princeton.

Barry, F. and Walsh, F. (2008), "Gains and losses from sectoral relocation: A review of theory and empirics", *Structural Change and Economic Dynamics* 19: 4-16.

Berman, E., Bound, J. and Griliches, Z. (1994) "Changes in the demand for skilled labor within U.S. manufacturing: evidence from the annual survey of manufacturers", *Quarterly Journal of Economics* 109(2): 367-397.

Blomström, M., Fors, G. and Lipsey, R.E. (1997) "Foreign direct investment and employment: home country experience in the United States and Sweden", *The Economic Journal* 445(107): 1787-1797.

Boix, R. and Galletto, V. (2009) "Innovation and Industrial Districts: A First Approach to the Measurement and Determinants of the I-District Effect", *Regional Studies* 43(9): 1117-1133.

Brewer, A.M., Button, K.J. and Hensher, D.A. (2001) (Eds) *Handbook of logistics and supply-chain management*, Pergamon, Amsterdam.

Brouwer, A., and Mariotti, I. (2009) "Evaluating firms' ownership differences using propensity score estimation and discrete choice modelling. The case of the logistics sector in Italy", *49th Congress of the European Regional Science Association (ERSA)*, 25-29 August, Łódź, Poland.

Castellani, D., Mariotti, I. and Piscitello, L. (2008) "The impact of outward investments on parent company's employment and skill composition. Evidence from the Italian case", *Structural Change and Economic Dynamics* 19(1): 81-94.

Commission Européenne (2001) *Impact des NTIC sur la logistique des enterprises commerciales*, Rapport final, DG Entreprise, Bruxelles.

Confetra (2002) *La fattura Italia dei servizi logistici*, Quaderno N° 98/2, Centro Studi Confetra, Milan.

Confetra (2004) *Commercio estero e trasporti*, Quaderno N° 126, Centro Studi Confetra, Milan.

Crinò, R. (2009) "Offshoring, multinationals and labour market: a review of the empirical literature", *Journal of Economic Surveys* 23(2): 197–249.

Cusmano, L., Mancusi, M. and Morrison, A. (2009) "Globalization of Production and Innovation: How Outsourcing is Reshaping an Advanced Manufacturing Area", *Regional Studies,* First published on 12 January 2009 (iFirst): 1-17.

Dunning, J. H. and Lundan, S. (2008) *Multinational enterprises and the global economy*, 2nd ed. Edward Elgar, Cheltenham.

Egger, H. and Egger, P. (2001) "Cross-border sourcing and outward processing in EU manufacturing", *The North American Journal of Economics and Finance* 12(3): 243-256.

Elia, S., Mariotti, I. and Piscitello, L. (2009) "The impact of outward FDI on the home country's labour demand and skill composition", *International Business Review* 18: 357-372.

European Commission (2000) *Protrans: analysis of third-party logistics market*, Deliverable No. 1, October, Competitive and Sustainable Growth Programme of the 5th Framework programme, Bruxelles.

Falzoni, A.M. and Grasseni, M. (2003) "Outward FDI and skill upgrading in the Italian labor market", Paper presented at AIEL, *23rd Convegno Nazionale di Economia del Lavoro*, Messina, 25-26 September.

Federico, S. and Minerva, G.A. (2008) "Outward FDI and Local Employment Growth in Italy", *Review of World Economics* 144(2): 295-324.

Feenstra, R.C. and Hanson, G.H. (1996) "Globalization, outsourcing, and wage inequality", *American Economic Review* 86(2): 240–245.

Glaeser, E.L. and Kohlhase, J.E. (2004) "Cities, regions and the decline of transport costs", *Papers in Regional Science* 83: 197-228.

Head, K. and Ries, J. (2002) "Offshore production and skill upgrading by Japanese manufacturing firms", *Journal of International Economics* 1: 81-105.

Helg, R. and Tajoli, L. (2005) "Patterns of international fragmentation of production and the relative demand for labor", *North American Journal of Economics and Finance* 16: 233-254.

Hill, S., Lesher, M. and Nordås, H.K. (2008) "Trade and Labour Market Adjustments", *OECD Trade Policy Working Papers*, No. 64, OECD Publishing.

Ietto-Gillies, G. (2005) *Transnational Corporations and International Production, Concepts, Theories and Effects*. Edward Elgar, Cheltenham.

Illeris, S. (2005) "The role of services in regional and urban development: a reappraisal of our understanding", *Services Industries Journal* 25: 447-460.

Isfort (2003) *Osservatorio Nazionale sul Trasporto e la Logistica, Le dinamiche della domanda e dell'offerta*, Rapporti Periodici, n. 1, Rome.

Kapstein, E.B. (2000) "Winners and Losers in the Global Economy", *International Organization* 54(2): 359-384.

Kletzer, L.G. (2002) *Imports, Exports, and Jobs. What Does Trade Mean for Employment and Job Loss?*, W.E. Upjohn Institute for Employment Research, Kalamazoo, Michigan.

Krugman, P., Cooper, R.N. and Srinivasan, T.N. (1995) "Growing World Trade: Causes and Consequences", *Brookings Papers on Economic Activity* 1995 (1), 25th Anniversary Issue: 327-377.

Leichenko, R.M. and Silva, J. (2004) "International Trade, Employment and Earnings: Evidence from US Rural Counties", *Regional Studies* 38(4): 355-374.

Maggi, E., Mariotti, I. and Boscacci, F. (2008) "The indirect effects of manufacturing internationalisation on logistics. A focus on the industrial districts of Veneto", *International Journal of Logistics Economics and Globalisation* 1(2): 205-224.

Mariotti, I. and Piscitello, L. (2007) "The impact of outward FDI on local employment. Evidence from the Italian case", In: Arauzo, M., Liviano, D. and Martín, M. (eds) *Entrepreneurship, Economic Growth and Industrial Location*, Edward Elgar, London: 299-320.

Mariotti, S., Mutinelli, M. and Piscitello, L. (2003) "Home country employment and foreign direct investment: evidence from the Italian case", *Cambridge Journal of Economics* 27: 419-431.

Mariotti, S., Mutinelli, M. and Piscitello, L. (2008) "The Internationalization of Production by Italian Industrial Districts' Firms: Structural and Behavioral Determinants", *Regional Studies* 42(5): 719-735.

Markusen, A.R., Noponen, H. and Driessen, K. (1991) "International trade, productivity and US regional job growth: a shift-share interpretation", *International Regional Science Review* 14: 15-39.

Ministero delle Infrastrutture e dei Trasporti (2011), *Piano nazionale della logistica*, linee guida, Rome.

Minondo, A. and Rubert, G. (2006) "The effect of outsourcing on the demand for skills in the Spanish manufacturing industry", *Applied Economics Letters* 13(9): 599-604.

Molnar, M., Pain, N. and Taglioni, D. (2007) "The internationalisation of production, international outsourcing and employment in the OECD", *OECD Economics Department Working Papers*, No. 561, OECD Publishing.

Peters, M.J., Lieb, R.C. and Randall, H.L. (1998) "The use of third-party logistics services by European industry", *Transport Logistics* 1(3): 167-179.

Razzaque, M.A. and Sheng, C.C. (1998) "Outsourcing of logistics functions: a literature survey", *International Journal of Physical Distribution & Logistics Management* 28(2): 89-107.

Richardson, J.D. (1995) "Income inequality and trade: how to think, what to conclude", *Journal of Economics Perspectives* 9: 33-55.

Savona, M. and Schiattarella, R. (2004) "International relocation of production and the growth of services: the case of the Made in Italy industries", *Transnational Corporations* 13(2): 57-76.

Slaughter, M.J. (2000) "Production transfer within multinational enterprises and American wages", *Journal of International Economics* 2: 449-472.
Yieming, Z., Han Swee, L. and Seah Kiat, K. (2002) "The Third-Party Logistics Services and Globalization of Manufacturing", *International Planning Studies* 7(1): 89-104.

Appendix 1: Data and variables

Export and import:

Coeweb database provides data on import and export (in kg and euro) by transport modes (rail, road, sea and air) for both 1996 and 2001. We decided to employ kg rather than Euros for the following reasons. First, a higher amount of kilograms implies a higher amount of freight to be transferred, while a higher amount of Euros may simply imply that the freight is more valuable or that it is strongly demanded by the market. A second reason is that the value in Euros may be subject to fluctuations of the exchange rate when the freight is imported (or re-imported or temporary imported) by countries that do not adopt the Euro currency. Finally, considering that exchanges may occur intra-firm, values in Euros may be subject to transfer pricing and, hence, they would not reflect the real amount of freight transferred by transportation firms.

As regards the distribution of the kg of export and import of each transport mode (rail, road, sea and air) among the 8 transport sub-industries, the following strategy has been employed. The kg of export and import rail has been adopted for both railway transport (60101) and its auxiliary activities (60102), given that the two sub-industries are vertically integrated. The sum of kg of rail and road has been attributed to land transport cargo handling (63113), because of the vertical integration between this sub-industry and both railway transport (60101) and road transport (60250). For the same reason, the kg of sea have been attributed to both water transport (61110) and its cargo handling (63112); similarly, the kg of air have been attributed to both air transport (62200) and its cargo handling (63113).

IPT and OPT:

Data on IPT and OPT are collected since 1988 at the EU Member country level. The reason for data collection about this special type of trade is that goods re-imported and re-exported after processing abroad are subject to customs treatment particularly advantageous to final imports and exports (Egger and Egger, 2001). Actually, the data on these flows are underestimated mainly because firms are not obliged to declare the processing trade.

Also data concerning IPT and OPT components have been expressed in kg. In order to distribute the kg of each IPT and OPT component among the different transport modes, the rates of export and import of each transport mode (rail, road, sea and air) over total trade have been employed. Specifically, the kg of export and import have been summed both in 1996 and in 2001. Then, the amount of export + import transported by each transport mode, with respect to the total amount of export + import transported by all transport modes, has been computed for each year. The resulting percentages have been employed to distribute the kg of the four components of IPT and OPT among the four transport modes (i.e. railway, road, water and air). The allocation of the four resulting values across the eight transport sub-industries has been implemented according to the guidelines described above for export and import.

FDI:

Outward FDI, undertaken by the Italian manufacturing MNE, are measured as the cumulated sum of the employees in their foreign affiliates from 1994 to 2000. FDI have been included as cumulated sum instead of flows because their effect on the employment is continuous along the time, given that investments occurred in the past need transport services also at the present time. The same does not hold for export, import and IPT and OPT components, which manifest and exhaust their effect in the moment they occur.

The workers of the foreign affiliates of each MNE have been distributed among the NUTS2 regions where one or more MNE's plants are located. The share of foreign affiliates' workers has been attributed to the NUTS2 regions where the MNE's plants are located proportionally to the plants' size in terms of number of employees. This method makes it possible to look at the impact of outward FDI not only on the employment of the region the MNE's headquarter belongs to, but also on the employment of all the regions where the firm is present with its plants (Elia *et al.*, 2009).

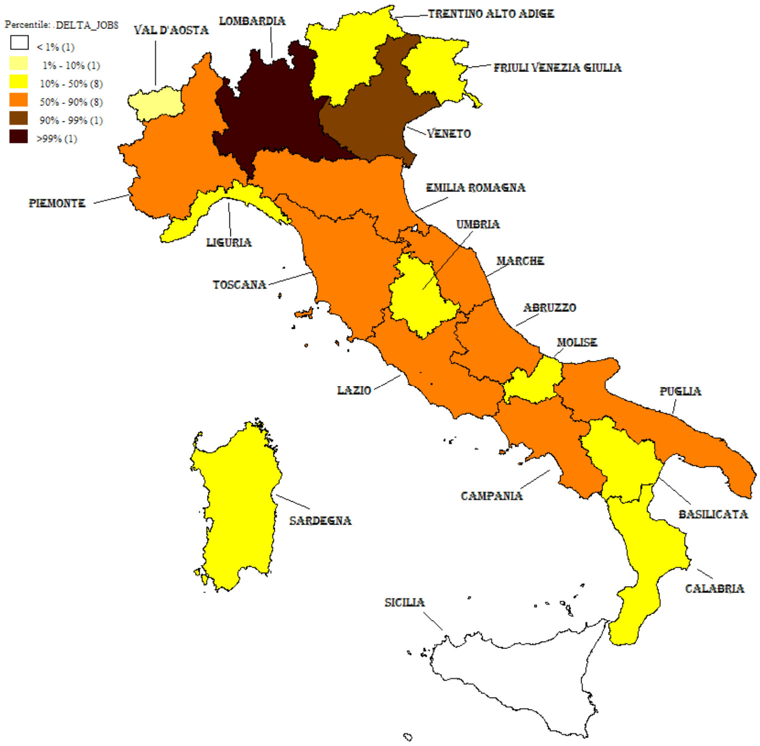
In order to distribute FDI among the four different transport modes (rail, road, sea and air), we employed the rates computed for IPT and OPT (see above). However, given that data about FDI are expressed as cumulated sum, i.e. they are expressed as a single value rather than as a difference between 2001 and 1996, we employed, for each transport mode, the mean between the rate of 1996 and the rate of 2001. The values of each transport modes have been then distributed across the eight transport sub-industries according to the guidelines discussed for export and import.

Production of manufacturing firms

The Balassa index of the value added ($y^B_{r,s}$) has been computed as the ratio between the share of value added transported by the transport sub-industry s in region r ($y_{r,s}$) over the value added transported by sub-industry s in all regions ($\sum_r y_{r,s}$) and the share of value added transported by all transport s in region r ($\sum_s y_{r,s}$) over the total value added transported by all transport s of all regions r ($\sum_{r,s} y_{r,s}$), according to the following formula:

$$y^B_{r,s} = \frac{y_{r,s}}{\sum_r y_{r,s}} \Bigg/ \frac{\sum_s y_{r,s}}{\sum_{r,s} y_{r,s}}$$

Appendix 2



Percentile	***Observations***	***Value of Delta Jobs***
<1%	1	N<0
1% - 10%	1	0<=N<200
10% - 50%	8	200<=N<1,600
50 % - 90 %	8	1,600<=N<8,000
90 % - 99 %	1	8,000<=N<15,000
> 99%	1	N>15,000

Figure 1: Distribution of total employment variation of the 8 transport sub-industries occurred between 1996 and 2001 in Italian regions.

Table 5: Correlation matrix and descriptive statistics of the explicative variables.

	$\Delta_{96}^{01} Y_{r,s}^{m}$	$\sum_{t=94}^{00} FDI_{r,s}^{m}$	$\Delta_{96}^{01} I_{r,s}^{m}$	$\Delta_{96}^{01} X_{r,s}^{m}$	$\Delta_{96}^{01} T\_I_{r,s}^{m}$	$\Delta_{96}^{01} R\_X_{r,s}^{m}$	$\Delta_{96}^{01} T\_X_{r,s}^{m}$	$\Delta_{96}^{01} R\_I_{r,s}^{m}$
Mean	0.0121	4396	419000000	189000000	-6838989	-10400000	1009981	6468925
Std. Dev.	0.6334	10559	1470000000	734000000	124000000	91800000	4399682	27900000
Min	-1.1177	0	-4610000000	-740000000	-1040000000	-764000000	-12300000	-3858230
Max	5.0890	66584	7200000000	5450000000	320000000	69700000	22500000	191000000
Obs	160	160	160	160	160	160	160	160
$\Delta_{96}^{01} Y_{r,s}^{m}$	1							
$\sum_{t=94}^{00} FDI_{r,s}^{m}$	-0.0298	1						
$\Delta_{96}^{01} I_{r,s}^{m}$	0.0059	0.4688	1					
$\Delta_{96}^{01} X_{r,s}^{m}$	0.0047	0.3512	0.4584	1				
$\Delta_{96}^{01} T\_I_{r,s}^{m}$	0.0038	0.0519	-0.3643	-0.6957	1			
$\Delta_{96}^{01} R\_X_{r,s}^{m}$	0.0214	0.1193	0.2379	-0.0755	-0.2012	1		
$\Delta_{96}^{01} T\_X_{r,s}^{m}$	0.0087	-0.0425	0.2374	0.5195	-0.5189	0.1521	1	
$\Delta_{96}^{01} R\_I_{r,s}^{m}$	-0.0160	0.3811	0.2274	0.1816	0.0036	0.0614	0.0522	1

Table 6: Contribution of the Italian regions to the internationalization forms(*) and percentage of flows undertaken towards EU.

	FDI		***Export***		***Import***		***IPT***				***OPT***			
							Temp. Import		**Re-export**		**Temp. Export**		**Re-import**	
Regions	**% of total**	**% to EU**	**% of total**	**% to EU**	**% of total**	**% from EU**	**% of total**	**% from EU**	**% of total**	**% to EU**	**% of total**	**% to EU**	**% of total**	**% from EU**
Centre	**12.17**	**66.52**	**11.47**	**60.99**	**16.64**	**58.63**	**12.32**	**79.04**	**7.96**	**11.30**	**14.79**	**46.47**	**6.60**	**58.57**
ABRUZZO	1.09	37.52	1.16	75.64	0.91	75.19	0.28	76.81	0.43	2.74	1.93	62.09	1.14	58.86
LAZIO	3.66	52.40	2.03	70.03	7.09	48.05	2.43	48.96	0.74	40.75	1.39	54.30	1.39	71.29
MARCHE	2.67	71.46	1.81	57.24	1.19	64.01	0.42	73.89	0.20	43.73	3.29	50.33	1.91	56.63
TUSCANY	4.37	82.16	5.49	55.29	6.56	66.03	7.80	87.39	6.34	4.46	7.03	40.81	1.50	56.19
UMBRIA	0.38	71.10	0.98	63.72	0.88	64.21	1.39	86.80	0.25	84.98	1.15	34.28	0.66	42.27
North-East	**24.45**	**49.72**	**27.78**	**67.06**	**28.45**	**65.73**	**18.33**	**58.34**	**10.98**	**24.39**	**28.25**	**44.46**	**57.75**	**84.99**
EMILIA R.	13.59	39.12	12.16	64.20	10.90	63.60	3.81	69.95	1.88	44.45	8.49	49.19	46.05	95.42
FRIULI V. G.	1.80	76.37	3.63	66.45	4.56	61.94	2.83	62.91	2.53	15.82	2.15	82.61	1.07	80.30
TRENTINO A.A.	1.09	64.98	1.41	85.64	1.82	83.23	0.93	96.94	0.15	80.57	1.83	88.29	0.94	86.15
VENETO	7.97	59.69	10.58	68.07	11.16	66.50	10.75	49.68	6.42	20.57	15.78	31.64	9.69	35.81
North-West	**53.46**	**63.45**	**35.27**	**78.61**	**36.53**	**73.25**	**32.90**	**61.04**	**11.64**	**49.86**	**44.40**	**48.76**	**30.96**	**60.58**
LIGURIA	0.80	47.41	7.70	90.05	3.29	52.03	2.76	36.67	0.43	9.75	1.48	76.43	0.27	52.97
LOMBARDY	34.71	65.32	18.97	73.40	26.22	74.09	17.53	67.88	5.24	45.04	27.53	45.57	20.91	58.98
PIEDMONT	17.94	60.56	8.49	79.77	6.94	80.03	12.61	56.83	5.97	56.99	13.84	46.41	9.02	61.28
VALLE AOSTA	0.01	23.02	0.11	86.41	0.08	80.35	0.01	100.00	0.01	100.00	1.55	100.00	0.75	99.34
South and Islands	**9.92**	**46.00**	**25.49**	**51.07**	**18.38**	**55.89**	**36.45**	**28.16**	**69.41**	**12.08**	**12.56**	**88.01**	**4.69**	**68.97**
BASILICATA	1.13	25.61	0.20	71.88	0.15	86.24	1.70	99.76	0.15	90.17	0.76	98.19	0.09	20.33
CALABRIA	0.43	24.91	0.09	65.48	0.21	87.73	0.02	72.54	0.01	6.91	0.03	75.26	0.02	96.05
CAMPANIA	4.06	49.02	2.77	59.76	3.49	59.01	9.47	14.37	3.09	8.38	1.86	87.30	0.64	77.45
MOLISE	0.43	44.19	0.17	78.89	0.14	74.55	0.16	76.55	0.10	21.25	0.22	89.93	0.16	88.31
PUGLIA	2.50	55.41	6.14	59.98	3.73	47.94	7.96	19.17	1.44	3.62	4.38	77.65	2.91	80.27
SARDINIA	0.23	68.36	11.05	41.65	6.96	59.84	10.82	50.94	16.88	0.27	5.15	95.23	0.26	43.55
SICILY	1.14	39.14	5.06	54.05	3.70	49.75	6.32	0.59	47.75	16.50	0.16	98.65	0.61	18.23
Total	**100.00**	**58.74**	**100.00**	**66.36**	**100.00**	**65.48**	**100.00**	**50.78**	**100.00**	**17.77**	**100.00**	**52.13**	**100.00**	**74.94**

Note: (+) Average between the years 1996 and 2001 for all internationalization modes with the exception of FDI, which has been computed by considering the cumulated sum between 1996 and 2001.

Factors facilitating intermodal transport of perishable goods - transport purchasers viewpoint

Thor-Erik Sandberg Hanssen [1*], Terje Andreas Mathisen [1]

[1] *Bodø Graduate School of Business, University of Nordland, NO-8049, Norway*

Abstract

The aim of this article is to identify factors that facilitate increased use of intermodal transport for perishable products based on a survey of firms exporting fresh fish from Norway to Continental Europe. The experiences in the Norwegian aquaculture industry indicate that intermodal transport solutions must be expanded and that the long haul by rail must run all the way to a central hub in Europe. This can only be achieved with a balanced flow of goods and if processors coordinate transport to deliver sufficiently large volumes to fill trains at an acceptable frequency.

Keywords: Intermodal transport; Rail; Fresh fish; Perishable goods; Export; European regions.

1. Introduction

The promotion and development of intermodal transport solutions, defined as the movement of goods in a single loading unit or vehicle that successively uses two or more modes of transport without handling the goods themselves in changing modes (UN/ECE, 2001, p. 17), is considered by the European Commission (2009) to be an important contribution to achieving a sustainable European transport sector.

The main advantage of intermodal transport solutions is their comparatively low external costs. It has been estimated that the total external cost of an intermodal train per tonne-km, including the cost of accidents, air pollution, greenhouse gases and noise, is only 28% of the external cost of a general freight truck (Forkenbrock, 2001). If focusing purely on greenhouse gases it is estimated, using transport between Basel and Rotterdam as an example, that CO_2 emissions from transport by waterways are four times higher and by lorry eight times higher compared to that of rail (UIC, 2008). Naturally, the results of any comparative study of the external costs of transport solutions will be affected by the types of external effects taken into consideration

* Corresponding author: Thor-Erik Sandberg Hanssen (thh@uin.no)

(Quinet, 2004). One obvious omission made by Forkenbrock (2001) is the cost of congestion. Given that the external cost of congestion is higher for road-only transport than for intermodal transport, the actual difference between the external costs for intermodal and road-only solutions may be even greater. Moreover, it has been argued that intermodal transport reduces transport costs because the most suitable transport mode is used on each part of a trip, thereby increasing national competitiveness through increased economic productivity and efficiency (OECD, 2001).

In Europe, intermodal freight transport can be divided into two groups: ocean liner trade and inland trade. Whereas intermodal ocean liner trade in Europe has been deemed a success, this has not been the case for inland intermodal transport (Stone, 2008). Between 1996 and 2006 the total freight volume (measured in tonne-kilometres) transported on the roads of the European Union (EU) grew by 45%, whereas transport by rail and by sea increased by 11% and 33%, respectively (European Environment Agency, 2009). This increase rate indicates that the policies implemented to promote intermodal transport have had limited success (Janic, 2007; Konings et al., 2008) and that the potential for transport by rail and water has barely been exploited (van Reeven, 2005). The unsatisfactory development of intermodal rail transport has been linked to substandard infrastructure, lack of interoperability, fragmented operational control, a lack of connection between operational control and responsibility, and unclear and rapidly changing institutional arrangements (Tsamboulas, 2008), which have generated concerns about the reliability, capacity, speed and flexibility of the rail network (European Commission, 2007).

Distance and the type of goods to be transported are the most important determinants of the transport mode chosen (Blauwens et al., 2006). The attractiveness of an intermodal road-rail transport solution depends on the extent to which the relatively low cost of rail transport offsets the extra cost of pre- and post-haulage and necessary transhipments (Bärthel and Woxenius, 2004). A decade ago, intermodal transport solutions were attractive for distances in excess of 500 km (van Klink and van den Berg, 1998). However, during the last decade Tsamboulas (2008) suggests that the minimum distance has fallen to 400 km. The break-even distance will vary both with the properties of the consignment and the transport services (Janic, 2007).

Specific challenges for implementation of intermodal transport between Norway and Europe are the heavy, and steadily increasing, unbalanced flow of goods (Hovi et al., 2008) and the high number of small loading units (Ludvigsen and Klæboe, 2010). The unbalanced flow, following the high imports relative to exports, would provide low average capacity utilization in intermodal transport solutions and make it difficult to achieve profitability. The problems related to small quantities could, according to Ludvigsen and Klæboe (2010), be approached by designating a number of terminals in Europe as “rail ports” where goods to and from Norway are handled to generate larger volumes and fewer rail stops. Also Wichser et al. (2007) emphasize the benefits for intermodal transport by reducing the number of participants in the supply-chain when studying east-west corridors in Europe. Sufficiently high service frequency is important and presupposes substantial and regular demand, see e.g. Janic (2007). This can be more easily achieved if players are consolidated.

One category of goods that requires special attention during transport is “perishable” or “deteriorating” goods such as fresh fish, flowers and fruits, which are particularly time-sensitive and fragile. Because intermodal transport solutions rely on time-consuming transhipment, they are at a disadvantage compared with all-road transport

(Sommar and Woxenius, 2007). In addition, transhipment requires movements that may damage fragile goods. Currently, the intermodal transport of perishable goods is very rare (Bontekoning and Priemus, 2004); namely, there is a substantial growth potential if the obstacles caused by sensitivity to time and fragility can be handled.

The aim of the article is to study the experience of the Norwegian aquaculture to advice policymakers regarding measures that would increase the use of intermodal transport solutions for fresh fish specifically and perishable goods in general.

The article is organized as follows. Section 2 briefly presents the literature on intermodal transport and identifies a number of factors particularly relevant for fresh fish. Then, in Section 3, the empirical data from the Norwegian aquaculture industry are presented. The data are analysed in Section 4. Finally, concluding remarks including notes on strategies for achieving increased use of intermodal transport of perishable goods are provided in Section 5.

2. Intermodal transport of perishable goods

The literature defines intermodal transport as the movement of cargo/products using more than one mode of transport (Bontekoning et al., 2004). Hence, the transfer of goods between transport modes is required, and the transfer must be conducted efficiently at the terminals to make intermodal transport solutions competitive (Woxenius and Barthel, 2008). Improved cooperation and interaction between the organizations that operate and use the terminals can help operators to plan terminal operations better and thereby run them more efficiently (Stokland et al., 2010).

According to Stone (2008), improved interconnectivity and interoperability are critical success factors for intermodal growth. However, interconnectivity and interoperability are difficult to achieve because the current European transport system is the result of 150 years of development and is affected by perceived needs, private and public initiatives, wars, funding, and national policies of support and intervention. This long trajectory has resulted in 37 different combinations of rail gauge, tunnel clearance and power systems across Europe (Tsamboulas, 2008).

Dedicated information on the specific challenges associated with the intermodal transport of perishable goods is hard to find, as are recommendations regarding how such transport can be facilitated. Irrespective of the length of the journey, perishable goods are typically transported by road (Blauwens et al., 2006). The predominance of this method is in part due to the flexibility that it offers regarding “door-to-door” deliveries (Blauwens et al., 2006) and the time factor involved due to the possibility of deterioration of the goods in question. The importance of time when transporting perishable goods is visible in rough assessments of price reductions of between 20 and 25% for fresh fish delayed in transit by 48 hours (Lervåg et al., 2001).

Although improved refrigeration technology has made the time factor somewhat less critical (Nordtvedt, 2009), perishable goods continue to be particularly sensitive with regard to on-time reliability (Patterson et al., 2008). Temperature largely determines the rate of microbial activity in fish (Giannakourou et al., 2005) and thereby the quality and safety of fish products. Therefore, cold chain management is important to keep fish products safe and to maintain their economic value (Tingman et al., 2010). However, it has been recognised that transfer points are weak links in chilled perishable food management (Nychas et al., 2008). This means, ceteris paribus, disadvantages for intermodal transport as compared to unimodal transport.

To realize real growth in the perishable goods market, which, according to Bontekoning and Priemus (2004), is among those with the highest growth potential, shorter transport time will be necessary on the rail haul segment of the transport chain, as will shorter processing time at the terminals, delay control, more and better transport relationships and increased service frequency (Bontekoning and Priemus, 2004).

Table 1 presents the main factors affecting the successful use of intermodal transport to be studied further in the empirical context of exported Norwegian aquaculture products. The list is based on factors identified in the EU project PROMOTIC (2000) and challenges that are particularly relevant for the export of fresh fish from Norway based on market knowledge.

Table 1: Facilitators for intermodal success with transport of fresh fish.

Factor (abbreviated title)	***Comments and explanations***
Properties related to interconnectivity	
Terminal proximity	Terminal must be located close to the farming plant
Terminal capacity	Terminal must have sufficient capacity
Terminal access	Terminal must be easily accessible from the main road network
Properties of the transport solutions offered by the transport/export companies	
Include rail	Rail must be included in the transport solutions offered by the export companies
Rail to Europe	The long haul trip by rail must go all the way to a central hub in Europe
Continuous cooling	A continuous cooling chain is required to maintain the freshness of the goods
Properties of close substitutes	
Road price	Increased price on close substitutes will make rail more attractive
Sea price	
Quality aspects of the rail haul	
Rail speed	Increased speed, frequency, punctuality will make rail-based transport solutions more attractive
Rail frequency	
Rail punctuality	
Rail price	Price must be competitive compared to other transport solutions
Rail capacity	Capacity on rail must be increased to make rail a better alternative
Rail customer service	Suppliers of rail-based transport solutions must improve their customer support
Rail environmental friendly	Rail provides more environmental friendly transport compared to pure road transport solution

The factors are categorized into four groups focusing on the properties of (1) the terminals where the cargo is loaded, (2) the transport solutions offered, (3) close substitutes and (4) quality aspects of the rail haul. The 15 factors identified in Table 1 form the basis for the subsequent analysis in Section 4 of means to increase the use of intermodal transport for the export of farmed fresh salmon from Norway to Continental Europe.

3. Empirical Data

According to the Norwegian National Rail Administration (2008), the railway network had in 2007 a track length of 4114 km, of which only 227 km (5%) is double track. Over the next decade, there are plans for investments in more double tracks and improved traffic control systems to ensure improved efficiency (Norwegian Ministry of Transport and Communications, 2009). The railway network illustrated in Figure 1 shows several connections from selected terminals in the coastal areas in the Western and Northern parts of Norway, where the aquaculture industry operates, to the national terminal in Oslo. The main provider of transport capacity at these terminals is CargoNet AS with more than 99% of the domestic "tonne-kilometers" in 2007. Transport of iron ore by Malmtrafikk AS from Northern Sweden to Narvik for further shipment by ship makes up the major portion of cross-border traffic.

Figure 1: The Norwegian railway network with selected important terminals for domestic transport of aquaculture products.
Source: Mathisen et al. (2009).

3.1 The Norwegian Salmon Export Industry

Norway is one of the largest exporters of fish commodities (FAO, 2009), second only to China, and Continental Europe represents one of the main markets. In 2008, the export value of Norwegian salmon exceeded 18 billion NOK, and almost all of that salmon was produced by the aquaculture industry and sold fresh (Norwegian Seafood Export Council, 2009). Transport patterns for exported fish can partly be found in statistics gathered by the Norwegian Customs Department and prepared for analysis by Statistics Norway. The data includes county of origin, border crossing and expected destination. With respect to the border crossing, details are given regarding name of customs station, transport mode, type of goods, weight and value. However, these data are based on planned transport route prior to departure of the shipment. Hence, redirections of transports on-route are not captured in this data set. Mathisen et al. (2009) corrected for this weakness by interviewing fish farmers about their transport patterns and concludes that salmon exported in 2007 mainly crossed the border on articulated vehicles (77%), with smaller quantities on boats/ferries (15%) and airplanes (8%). In contrast, no fish crossed the Norwegian border by rail.

Even though most transport of fresh fish currently are carried out by lorries or articulated vehicles, a considerable proportion of domestic transport, especially from the Northern part of Norway, are transported by rail. Mathisen et al. (2009) estimate that about 70000 tonnes of the total exported volume of 520000 tonnes were transported domestically by rail in 2007. The majority originate from the Northern part of the country and end in the Southeastern part of Norway, where the fresh fish is reloaded to trucks for further transport to the markets in Continental Europe. All fresh fish exported from Norway to Continental Europe is transported longer than the 400 km considered to be required for intermodal transport solutions to be competitive. Thus, it could be expected that at least some of the exported fresh fish might be eligible for efficient intermodal transport during the international portion of the trip.

3.2 Data Set

Information about the use of rail-based intermodal transport was gathered from managers at Norwegian farming plants and exporters of aquaculture products using a web-based questionnaire. The questionnaire was designed to reveal the managers' experiences with rail-based intermodal transport and which factors they consider most important to further increasing their use of such transport solutions.

The email addresses used to distribute the questionnaire to the respondents were collected from two sources. The Norwegian Seafood Federation, an interest organization for the aquaculture industry, provided a list of email addresses for all member farms. The email addresses for the exporters, on the other hand, come from the Norwegian Seafood Export Council, which maintains a publicly available register of all companies allowed to export sea food, including their email addresses. Because all exporters of seafood must by law be enrolled in this register, it represents the total population, and there is in practice no bias involved because all firms have email addresses. As presented in Table 2, the questionnaire was sent to 665 firms, including both farmers and exporters of seafood.

Table 2: Details about response rate and products.

	Selection (no. of firms)	*Responses*	*Response rate*	*Products [a]*			
				Salmon	*Cod*	*Trout*	*Halibut*
Farmers	250	43	17,2 %	58%	33%	26%	12%
Exporters	415	70	16,8 %	23%	24%	11%	16%
Total	665	113	17,0 %	36%	27%	17%	14%

Note: [a] indicates the fish products handled by the respondents. Each respondent could have more than one product.

Web-based surveys have long been recognized as taking less time and costing less money than more conventional survey methods (e.g. Schmidt, 1997), but they typically have low response rates (Cook et al., 2000). As shown in Table 2, a total of 113 respondents answered the questionnaire, yielding a response rate of 17%. However, because the respondents were not obliged to answer all of the questions, there are some missing values for some of the questions. A large portion of both the firms selected and the responders were exporting firms.

These firms often handle more than one product. Fresh salmon, provided by 36%, is most frequently indicated as one of the respondents' main products. Also, other fresh aquaculture products such as cod, trout and halibut are important products for both farmers and exporters. A mixed group of different frozen fish, shellfish and other white fish also makes up a sizeable portion of the total product portfolio, particularly for exporters. With respect to geographical distribution, the sample includes respondents located in all counties along the coast from the Southwest of Norway to the Northeastern border with Russia. The data set captures the variation in firm size found in the industry, with annual volumes ranging from 6 tonnes for small seasonal operators to 170000 tonnes for multinational companies listed on the stock exchange (with an average of about 13000 tonnes).

4. Analysis

4.1 Deciding on Transport Mode and Transport Route

Transport by road from Norway to Europe has increased with rising exports and imports (Hovi et al., 2008). In the case of aquaculture products, several actors are involved in choosing transport modes and transport routes from processing plants in Norway to customers in Continental Europe. Hence, an important question to address when seeking to identify policy measures that can contribute to the increased use of intermodal transport is who decides on the transport mode and route.

The value chain for farmed salmon starts with the production of fish feed and continues with the cultivation and breeding of the fish at the hatchery before they are farmed and processed. Then the fish is sold by exporting companies and transported (primarily by road) to customers abroad. It is not unusual for large companies in this industry to vertically integrate parts of the logistical chain, from hatcheries to exporters, with further vertical coordination with supermarkets (e.g. Kvaløy and Tveterås, 2008).

The respondents were asked to indicate to what degree six different actors in the logistical chain influence the chosen transport mode and route. Because they are not directly related to the export of fresh fish, the actors involved in the early fodder and hatchery stages were not considered. Still, one option was to indicate whether other participants in the logistics chain influence the choice of route and transport mode. The actors are as follows (sorted according to the transport chain for aquaculture products):

- Farmers – The companies producing the fish (breeder)
- Processor – The companies processing the fish (slaughterhouses)
- Exporters – Sales companies
- Transporters – The companies offering transport solutions to the exporters
- Railway companies – The suppliers of rail services (firms that own and run the trains)
- Customers – The retailers located in Continental Europe (fish buyers)

Responses were provided on a scale from 1 (no influence) to 5 (a decisive influence). The respondents' impression of the influence each actor has on mode and route choice is presented using mean values in Table 3 accompanied by the number of responses and standard deviations. In the survey, the rating for each factor was accompanied by an explanation stating the meaning of each number, with the value 3 representing neither a low nor a high degree of influence. This ordinal scale does have its limitations with respect to econometric analysis in that it produces non-metric data (e.g. Hair et al., 1998). It is, however, clear that a rating of 3 is better than a rating of 2. In the following analyses, it is assumed that the respondents perceive the differences between the grades as equal so that average values can be calculated. The varying number of responses for the role played by each actor, ranging from 68 (other) to 98 (exporter), occurred because the respondents were not forced to answer all of the questions. Standard deviations are used to indicate the variation in the answers.

Table 3: Respondents view of the influence of actors on choice of transport mode and route.

Actor	*Transport mode*			*Transport route*		
	Responses	*Mean* [a]	*Std. dev.*	*Responses*	*Mean* [a]	*Std. dev.*
Exporter	*98*	*4.6*	*0.8*	*91*	*4.1*	*1.2*
Customer	95	3.5	1.4	91	3.5	1.5
Transporter	90	3.4	1.4	92	4.1	1.3
Processor	96	2.5	1.5	95	2.2	1.4
Farmer	85	2.4	1.4	84	2.0	1.2
Railway firm	84	2.2	1.3	83	2.2	1.3
Other	68	1.5	0.9	68	1.4	0.9

Note: [a] 1= no influence, 2= low influence, 3= neither low nor high influence, 4= high influence, 5= decisive influence.

It is evident from Table 3 that exporters have the most influence on choice of transport mode (score 4.6). This is reasonable because exporters often choose from among transport solutions offered by several transporters. The relatively low standard deviation for the exporters indicates that there is a high degree of agreement among the respondents that this actor decides on the transport mode. Also, customers (score 3.5) and transporters (score 3.4) greatly influence the choice of transport mode. Even though the same three actors have the main influence on the chosen transport route, the values presented indicate that the influence on choice of transport route is more evenly distributed than that on choice of transport mode. Processors, farmers, and railway firms have little influence regarding the choice of mode and route. The fact that the "other" category is given close to no influence indicates that the six other actors wield virtually all of the influence.

Hence, if policymakers want to alter the chosen mode of transport for aquaculture products, their efforts should be primarily aimed at exporters. If the aim is to influence route choice decisions, however, the transport companies are equally important to address.

4.2 Means to Increase the Use of Rail in Transport of Fresh Fish

The respondents were asked to state to what degree the 15 factors identified in Table 1 influenced their choosing rail as the long-haul transport mode. The scale ranged from 1 (no importance) to 5 (high importance) with a score of 3 representing "some influence". The statements are ranked in Table 4 according to mean value, with the number of respondents varying between 34 and 38 because the respondents were not forced to answer all questions. The respondents were about 80% farmers and 20% exporters.

Table 4 indicate that almost all factors are considered to have a substantial influence on the use of rail-based intermodal transport solutions. The exception is the price of close substitutes such as sea and road transport, which is assessed to be of little importance. Such a conclusion with respect to pricing of other transport modes is supported by Ivaldis' (2007) theoretical discussion of the effect on rail transport by introducing a road toll. The low importance, however, could be due to tactical answers by respondents hoping to avoid increased prices on current transport solutions.

As argued in Section 2, the need for an unbroken cooling chain (continuous cooling) is a decisive factor in making rail-based transport a viable alternative for the transport of fresh fish and other perishable goods. Equally important is the fact that too few rail-based intermodal transport solutions are offered and that none of them provide a long-haul service to Continental Europe by rail. Also, certain aspects of the terminals where the fish is loaded on the rail are considered important. The respondents claim that the use of rail services will increase for this type of goods if the terminals are located closer to the farmers and if access from the main roads improves. With respect to the terminals themselves, it should be noted that farmers and processors of aquaculture products are often located on islands or in other rural areas along the coast, whereas the railway infrastructure is located much further inland. This explains why improved terminal proximity and accessibility will be necessary to make rail transport a genuine alternative. However, it is also worth noting that intermodal terminals need a critical catchment area for efficient operations (Bergqvist et al., 2010). Improved proximity might reduce average catchment area for the terminals and might thus also reduce their efficiency. Aspects of rail-based service such as speed and frequency are also considered to have a considerable degree of influence.

Table 4: Respondents assessment of statements regarding increased use of rail-based intermodal transport of fresh fish from Norway to central Europe.

Statement	*Responses*	*Mean value*	*Std. dev.*
Continuous cooling	36	4.2	1.2
Include rail	38	4.2	1.0
Terminal proximity	38	3.9	1.2
Rail to Europe	36	3.9	1.0
Rail speed	37	3.8	1.1
Terminal access	36	3.7	1.1
Rail frequency	37	3.6	1.2
Rail punctuality	36	3.4	1.2
Rail environmental friendly	35	3.4	1.1
Terminal capacity	35	3.4	1.3
Rail customer service	36	3.3	1.1
Rail price	36	3.1	1.3
Rail capacity	34	3.0	1.2
Road price	36	2.5	1.1
Sea price	35	2.3	1.1

Note: [a] 1= no influence, 2= low influence, 3= some influence, 4= considerable influence, 5= high influence.

Based on the responses given by farmers and exporters, it becomes evident that farmers tend to consider each factor as having a greater influence on the use of rail-based intermodal transport solutions than do exporters. The average value that farmers give to each factor is 3.5, whereas that figure is 3.1 for exporters. The two only factors that exporters consider more important than farmers do are the need for lower prices for rail transport and for road transport to be more expensive. However, the differences between the mean values for the two groups are not statistically significant. Hence, it cannot be argued that other factors are more important for exporters than they are for farmers.

The respondents were then asked to prioritize three of the 15 factors that, from their point of view, are most crucial to increasing the use of rail-based intermodal transport. A total of 130 priorities were indicated by the respondents. In contrast to the ranking presented in Table 4, in which all factors could be assessed as important, the prioritization illustrated in Figure 2 forced the respondents to state which factors they considered most vital. The factors in Figure 2 are ranked according to the total number of priorities.

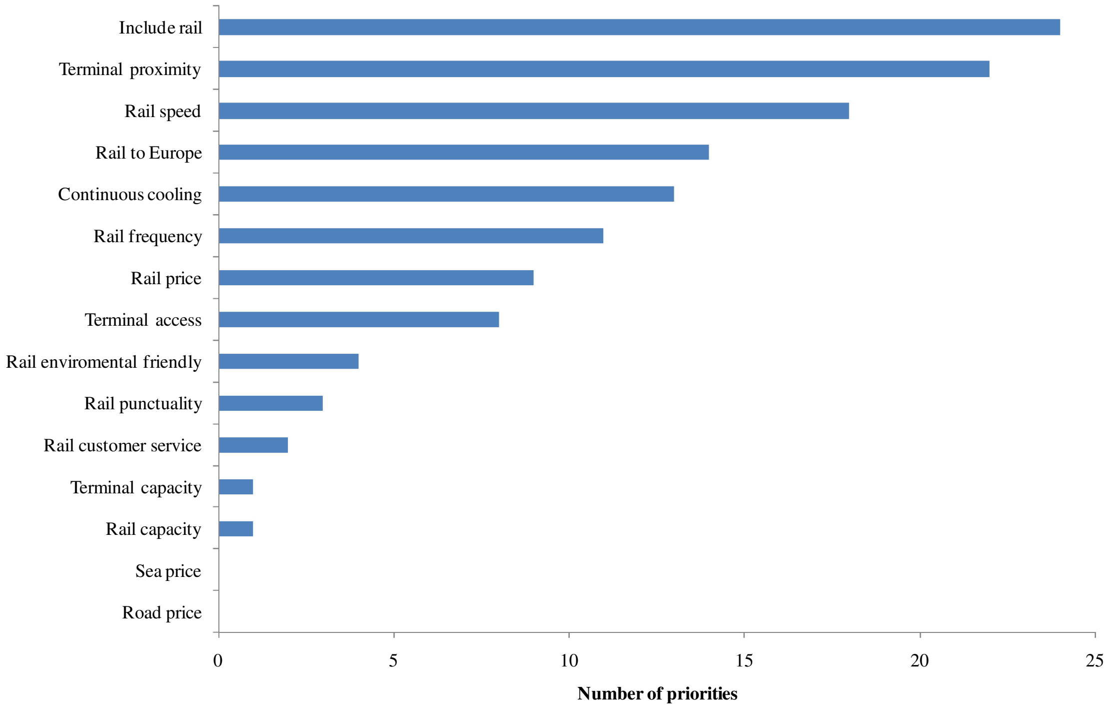


Figure 2: Prioritization of the most important factors to promote the use of rail-based intermodal transport of fresh fish from Norway to Continental Europe.

It is evident from Figure 2 that the most important factor in facilitating the rail-based intermodal transport of aquaculture products from processor to consumer is that service offerings including rail must be expanded. At the time the survey was conducted, there were virtually no cross-border transport solutions for fresh products including rail transport, and none offered long-haul rail transport all the way to a central hub in Europe, which is ranked as the fourth most prioritized characteristic. The second most important factor is that terminals must be located in close proximity to fish processors. The third most important factor is rail speed. Other characteristics of the rail and terminal services are assessed as less important.

The prioritization of factors in Figure 2 largely corresponds with the ranking of the mean values in Table 4. That is, the amount of intermodal transport solutions including rail must be increased, and they must run all the way to a hub in Continental Europe in a continuous (unbroken) thermo chain. Moreover, the aquaculture industry states that capacity on the rails and in the terminals are currently not a problem and that increasing the price of other transport modes is not a desirable way to make rail-based transport solutions more attractive.

5. Conclusions and Policy Implications

The substantial external costs associated with freight transport on the road are both due to it being a major source of pollution and an important contributor to congestion on the road network. As a result, it has become European policy on both the regional and the national level to shift freight from the road to other transport modes like rail and boat.

The most critical factor in achieving sustainable transport is the successful promotion of intermodal transport. Given the limited success of previously implemented policies related to perishable goods, the need for additional knowledge regarding how to increase the use of intermodal transport becomes obvious. This article identifies the factors that must be addressed to make the intermodal transport of perishable products from Northern to Southern European countries more attractive. The export of fresh fish from Norway to Continental Europe is used as an example, but the findings are also relevant to intermodal transport for other categories of perishable goods.

The analysis is based on a survey of 113 managers in the Norwegian aquaculture industry. The industry states that if policymakers want to alter the mode of transport chosen for aquaculture products, they should primarily aim their attention at exporters. If the aim is to influence route choice decisions, however, the preferences of transporters are equally important to address.

It is most critical that intermodal transport solutions including rail service be increased and that these routes run all the way to a hub in Continental Europe in a continuous (unbroken) thermo chain. The challenge of creating an unbroken thermo chain across countries could be approached via equipment standardization. Other obstacles to developing such intermodal services include creating a better balanced flow of goods from north to south and sufficiently large volumes.

The problem of unbalanced flow has to do with the need for the refrigerated containers used to transport fish southbound to Continental Europe to be filled with other commodities that require cooling when they return northbound to Norway. This will increase overall capacity utilization so that the rail-based transport solution becomes competitive with respect to price and simultaneously remains profitable for operators. Currently, imports to Norway that require cooling, such as fruit and vegetables, are transported primarily by road, though they are also transported by sea.

The scattered structure of the Norwegian aquaculture industry poses a challenge to the establishment of sufficiently high volumes. There are many relatively small companies that do not individually produce sufficiently large quantities to make rail-based transport solutions practical. Today, these companies procure transport services separately for their own production. However, the total export of aquaculture products is sufficiently large to fill trains so that they can depart at a satisfactory frequency. Hence, if the industry can cooperate in procuring transport services, they will be able to achieve intermodal transport solutions involving long-haul transport to Continental Europe by rail. This requires a commitment among purchasers of transport solutions to providing predetermined volumes of commodities or establishing of dedicated “rail ports” for freight of goods to and from Norway. However, the structural changes that are currently taking place within the industry, encouraging the creation of fewer and larger companies, is paving the way for such agreements because fewer companies need to coordinate.

The respondents from the aquaculture industry also point out the need for investments in the intermodal terminals where their products are loaded onto the trains. Although current capacity is considered acceptable, there is a need to reduce the transport barriers separating farmers from the terminals. This can be achieved by establishing terminals closer to processors and by improving the road network so that the terminals will become more easily accessible. Such investments require public subsidies, possibly in cooperation with companies within the logistics chain for aquaculture products if they find it sufficiently attractive. In addition to the positive effects of a general upgrade in infrastructure via such investments, the main argument for public financing is that it reduces the external cost of transport (e.g., it is more environmentally friendly, causes fewer accidents and decreases congestion on the roads).

Most of the characteristics of long-haul transport by rail that are singled out as important relate to the time sensitive nature of deliveries of perishable goods. Higher frequency will reduce waiting time and, as previously mentioned, can be improved if processors commit to delivering sufficiently large volumes at certain deadlines. Moreover, the punctuality of trains, which has been a problem on the Norwegian railway network, although most prominent for passenger transport, must be at an acceptable level. The expected transport time will be further reduced by investments in centralized traffic control (CTC), double tracks or side crossings and increased prioritization of freight trains. Further research could reveal how the policy actions to make intermodal transport more attractive relate to supply chains for different types of products.

Facilitating intermodal transport by implementing policies based on the factors identified in this study will require political will and improved competitiveness for intermodal transport of perishable goods will not be achieved unless policymakers invest in and prioritize rail transport. Moreover, by addressing these factors, European politicians will reduce the external cost of transport, improve European integration by reducing export barriers and support the rural regions of Europe that often rely on the production and sale of perishable agricultural products.

Acknowledgements

The article is based on a research project financed by the SMARTRANS program of The Research Council of Norway with financial support from the Norwegian Public Roads Administration and the Norwegian Seafood Federation. Moreover, the authors wish to thank the anonymous referees of this journal for their valuable comments.

References

Bergqvist, R., Falkemark, G. and Woxenius, J. (2010) "Establishing intermodal terminals", *World Review of Intermodal Transportation Research*, 3, 3: 285-302.

Blauwens, G., De Baere, P. and Van de Voorde, E. (2006) *Transport economics*, De Boeck, Antwerpen.

Bontekoning, Y. M., Macharis, C. and Trip, J. J. (2004) "Is a new applied transportation research field emerging? -A review of intermodal rail-truck freight transport literature", *Transportation Research Part A: Policy and Practice*, 38, 1: 1-34.

Bontekoning, Y. M. and Priemus, H. (2004) "Breakthrough innovations in intermodal freight transport", *Transportation Planning and Technology*, 27, 5: 335 - 345.

Bärthel, F. and Woxenius, J. (2004) "Developing intermodal transport for small flows over short distances", *Transportation Planning and Technology*, 27, 5: 403-424.

Cook, C., Heath, F. and Thompson, R. L. (2000) "A meta-analysis of response rates in Web- or internet-based surveys", *Educational and Psychological Measurement*, 60, 6: 821-836.

European Commission (2007) "Towards a rail network giving priority to freight", *EU COM (2007) 608 final*, Brussels.

European Commission (2009) "A sustainable future for transport: Towards an integrated, technology-led and user friendly system", *EU COM (2009) 279 final*, Luxembourg.

European Environment Agency (2009) "Transport at a crossroads TERM 2008: indicators tracking transport and environment in the European Union", *EEA Report No 3/2009*, Copenhagen.

FAO (2009) *Fishery and Aquaculture Statistic*s, Food and Agriculture Organization of the United Nations, Rome.

Forkenbrock, D. J. (2001) "Comparison of external costs of rail and truck freight transportation", *Transportation Research Part A: Policy and Practice*, 35, 4: 321-337.

Giannakourou, M. C., Koutsoumanis, K., Nychas, G. J. E. and Taoukis, P. S. (2005) "Field evaluation of the application of time temperature integrators for monitoring fish quality in the chill chain", *International Journal of Food Microbiology*, 102, 3: 323-336.

Hair, J. F., Anderson, R. E., Tatham, R. L. and Black, W. C. (1998) *Multivariate data analysis* (5 ed.). Prentice Hall, Upper Saddle River, NJ.

Hovi, I. B., Madslien, A., Askildsen, T. C., Andersen, J. and Jean-Hansen, V. (2008) "Globaliseringens effekt på transportmidel- og korridorvalg til og fra Norge" - "Effect of globalization on mode of transportation and choice of transport corridor to and from Norway", *TØI report 970/2008*, The Institute of Transport Economics, Oslo, (in Norwegian).

Ivaldi, M. (2007) "The European Market for Freight Services: Towards A Simulation Model of Competition", *IDEI Report*, No. 12.

Janic, M. (2007) "Modelling the full costs of an intermodal and road freight transport network", *Transportation Research Part D: Transport and Environment*, 12, 1: 33-44.

Konings, R., Priemus, H. and Nijkamp, P. (2008) "The future of intermodal freight transport: an overview", In: R. Konings, H. Priemus & P. Nijkamp (Eds.), *The Future of Intermodal Freight Transport: Operations, Design and Policy*: 1-10, Edward Elgar Publishing Limited, Cheltenham.

Kvaløy, O. and Tveterås, R. (2008) "Cost Structure and Vertical Integration between Farming and Processing", *Journal of Agricultural Economics*, 59, 2: 296-311.

Lervåg, L.-E., Meland, S. and Wahl, R. (2001) "Utvikling av NEMO/REGO - Parameterverdier" - "The development of NEMO/REGO - Parameter values", *SINTEF Bygg og miljø*, Trondheim, Norway, (in Norwegian).

Ludvigsen, J. and Klæboe, R. (2010) "Green Handshake Sustainable rail freight connections between Norway and Europe", *TØI report 1118/2010*, The Institute of Transport Economics, Oslo.

Mathisen, T., Nerdal, S., Solvoll, G., Jørgensen, F. and Hanssen, T.-E. S. (2009) "Ferskfisktransporter fra Norge til kontinentet - Dagens transportstrømmer og utfordringer ved bruk av intermodale transportopplegg" - "Transport of fresh fish from Norway to Continental Europe - Current transport patterns and challenges for intermodal transport", *SIB-report 2/2009*, Bodø, Norway, (in Norwegian).

Nordtvedt, T. S. (2009) "Superkjøling av fisk: en litteraturstudie og prosjektoversikt" - "Superchilling of fish: a litterature review and project overview", TR A6797, *SINTEF energiforskning AS*, Trondheim, Norway, (in Norwegian).

Norwegian Ministry of Transport and Communications (2009) *Nasjonal transportplan 2010-2019*, (in Norwegian).

Norwegian National Rail Administration (2008) "Jernbanestatistikk 2007", *Norwegian railway statistics*, (in Norwegian).

Norwegian Seafood Export Council (2009) "Eksportstatistikk 2008", *Export statistics 2008*, (in Norwegian).

Nychas, G. J. E., Skandamis, P. N., Tassou, C. C. and Koutsoumanis, K. P. (2008) "Meat spoilage during distribution", *Meat Science*, 78, 1-2: 77-89.

OECD (2001) *Intermodal freight transport: institutional aspects*, Organisation for Economic Co-operation and Development (OECD), Paris.

Patterson, Z., Ewing, G. O. and Haider, M. (2008) "The potential for premium-intermodal services to reduce freight CO_2 emissions in the Quebec City-Windsor Corridor", *Transportation Research Part D: Transport and Environment*, 13, 1: 1-9.

PROMOTIQ (2000) "Conditions for the promotion of a new generation of intermodal transport services and operators", *Contract No: IN-98-RS.5054*, European Commission.

Quinet, E. (2004) "A meta-analysis of Western European external costs estimates", *Transportation Research Part D: Transport and Environment*, 9, 6: 465-476.

Schmidt, W. C. (1997) "World-Wide Web survey research: Benefits, potential problems, and solutions", *Behavior Research Methods, Instruments, and Computers*, 29, 2: 274-279.

Sommar, R. and Woxenius, J. (2007) "Time perspectives on intermodal transport of consolidated cargo", *European Journal of Transport and Infrastructure*, 7, 2: 163-182.

Stokland, Ø., Sund, A. S. and Netland, T. (2010). Challenges in intermodal logistics networks and terminals - a Norwegian viewpoint. Paper presented at the 12th World Conference on Transport Research Society, Lisbon, Portugal.

Stone, B. (2008) "Critical success factors: interconnectivity and interoperability". In: R. Konings, H. Priemus & P. Nijkamp (Eds.), The Future of Intermodal Freight Transport - Operations, Design and Policy. Edward Elgar Publishing, Cheltenham.

Tingman, W., Jian, Z. and Xiaoshuan, Z. (2010) "Fish product quality evaluation based on temperature monitoring in cold chain." African Journal of Biotechnology, 9, 37, pp. 6146-6151.

Tsamboulas, D. (2008) "Development strategies for intermodal transport in Europe". In: R. Konings, H. Priemus & P. Nijkamp (Eds.), The Future of Intermodal Freight Transport - Operations, Design and Policy. Edward Elgar Publishing, Cheltenham.

UIC (2008) "Rail Transport and Environment - Facts and Figures". International Union of Railways.

UN/ECE (2001) "Terminology on Combined Transport". United Nations (UN) & Economic Commission for Europe (ECE), New York and Geneva.

van Klink, H. A. and van den Berg, G. C. (1998) "Gateways and intermodalism." Journal of Transport Geography, 6, 1, pp. 1-9.

van Reeven, P. (2005) "Transport Policy in The European Union". In: K. J. Button & D. A. Hensher (Eds.), Handbook of Transport Strategy, Policy and Institutions. Elsevier, Amsterdam.

Wichser, J., Weidmann, U., Fries, N. and Nash, A. (2007). Strategies for increasing intermodal freight transport between Eastern and Western Europe. Paper presented at European Transport Conference 2007,

Woxenius, J. and Barthel, F. (2008) "Intermodal road-rail transport in the European Union". In: R. Konings, H. Priemus & P. Nijkamp (Eds.), The Future of Intermodal Freight Transport - Operations, Design and Policy. Edward Elgar Publishing.

European Transport \ Trasporti Europei n. 49 (2011): 90-111

Public tendering of ferry services in Europe

Alfred J. Baird [1*], Gordon Wilmsmeier [1]**

[1] *Transport Research Institute (TRI), Edinburgh Napier University*

Abstract

This paper outlines the EU regulatory framework and analyses current practice in four European countries in respect of public procurement and tendering of ferry services. Tender management for major ferry services resides with national government agencies, while tenders for smaller volume regional and/or inter-isles services are generally managed by local authorities and/or regional transport authorities. Operator selection criteria increasingly emphasises service quality aspects, and environmental impacts, as well as price (i.e. amount of subsidy required). There is a continued trend towards privately-owned operators providing and investing in essential ferry services, with an increasing role played by larger international integrated transport organisations. For transport authorities, ferry service procurement involves a continuous evolutionary process of specifying, offering, selecting, monitoring and reviewing services supplied. Based on the information collected and analysed, the authors have developed a 'tender route map' which explains the different stages and key issues concerning public procurement of ferry services.

Keywords: Ferry services; Public tendering; Procurement; Europe.

1. Introduction

This paper outlines the regulatory framework concerning the tendering of ferry services by public transport authorities in the European Union (EU) and thereafter analyses the tendering regime and tender processes employed in four European countries - Denmark, Greece, Sweden and Norway - based on a case study approach. The research formed part of a project undertaken by the authors on behalf of the Scottish Government during its Scottish Ferries Review (Transport Research Institute, 2010). The aim of the work was to review the delivery and operation of subsidised ferry services.

From a public sector perspective, the delivery of essential ferry services poses a dilemma. The public sector may seek to guarantee a minimum level of service to a

* Corresponding author: Alfred J. Baird (a.baird@napier.ac.uk)
** Corresponding author: Gordon Wilmsmeier (g.wilmsmeier @napier.ac.uk)

particular community even though such a service might not be commercially viable, and thus requiring of public subsidy. At the same time the public sector needs to secure an operator that provides a continued service at an efficient cost level, but with the lowest subsidy rate possible, whilst also ensuring that important aspects of service quality are maintained or improved.

Based on the results from the case study analyses the authors develop a tendering route map which is considered to offer useful guidance to public authorities tendering ferry services in future.

2. EU Regulatory Framework

2.1. Regulatory overview

Any solution for state subsidised delivery of ferry services in the EU must be compatible with national and EU laws. Of particular relevance to the design and implementation of any solution are the following:

- Council Regulation (EEC) No 3577/92, also known as the 'Cabotage Regulation' (European Commission, 2006)
- Public Service Obligations (PSO)
- Public Service Contracts (PSC)
- EU and national procurement law; and
- limitations on the provision of state aid.

These legal and regulatory requirements must be taken into account by transport authorities when developing procurement strategies and finalising decisions on how ferry services should be procured.

2.2. Cabotage Regulation

The 'Cabotage Regulation' (European Commission, 2006) regulates the transportation of passengers and goods by sea between two points within Member States of the European Union ("Member States"). The 'Cabotage Regulation' applies the principle of free movement of services to maritime transport and obliges Member States to allow community shipowners to operate freely in the European market.

The 'Cabotage Regulation' recognises that maritime transport of passengers and goods is vital for the inhabitants of island and peninsular communities. As such, exceptions to the principle of free movement of services are allowed where, owing to special circumstances, market forces would not provide a satisfactory level of service. Recognising the need to protect such ferry services to islands and remote peninsular communities, the 'Cabotage Regulation' allows Member States to intervene in particular markets by imposing PSOs and/or providing funding to undertakings who take on public service obligations through entering into PSCs.

2.3. Public Service Obligations

Member States may impose public service obligations (PSOs) in order to ensure provision of an adequate regular ferry service to and from a given location where community shipowners, in considering their own commercial interests, would not provide an adequate level of service or under the same conditions. It is for individual Member States to determine on which routes the market would otherwise fail to provide adequate ferry service and, thus, which require PSOs to ensure that the required ferry services are provided. The nature of PSOs that a Member State may impose is limited to the following aspects:

- requirements concerning ports to be served;
- requirements in relation to regularity, continuity, frequency and capacity to provide the service in question;
- rates to be charged for the services; and
- manning of vessels.

PSOs can be imposed by Member States in a variety of ways, although mainly this will involve one of the following approaches being adopted for the route in question:

- entering into PSCs with individual operators for a given route; or
- applying PSOs to all operators for a given route by way of a declaration regime, a licensing system or authorisation system.

PSOs may be implicit within PSCs, by virtue of the specified service obligations included in the contracts (European Commission, 2009). It has been argued that the PSO is a separate matter from PSCs, and that a PSO must first be established for each route which is considered to require intervention in some form or other in order to meet the needs of the community concerned (Kay, 2008). However, the Commission has concluded in line with the Altmark criteria that PSOs will be clearly defined within PSCs (European Commission, 2009). A PSC can therefore be used to impose public service obligations, with the obligations set out in the contact.

In a European Court Judgement it was stated that the combined provisions of Article 1 and Article 4 of Regulation No 3577/92 applying the principle of freedom to provide services to maritime transport within Member States (maritime cabotage) permit the provision of regular maritime cabotage services to, from and between islands to be made subject to prior administrative authorisation only if (European Court, 2001):

- a real public service need arising from the inadequacy of the regular transport services under conditions of free competition can be demonstrated;
- it is also demonstrated that a prior administrative authorisation scheme is necessary and proportionate to the aim pursued;
- such a scheme is based on objective, non-discriminatory criteria which are known in advance to the undertakings concerned.

2.4. Public Service Contracts

Public service contracts (PSCs) are the instrument generally used to impose PSOs where the imposition of a PSO on all ship owners would not support an area’s transport needs. A PSC will be used where compensation is payable for providing PSOs. A PSC can cover broader issues than those covered by PSOs and may include a requirement to satisfy fixed standards of continuity, regularity, capacity and quality and require services to be provided at specified rates.

PSCs can be used to ensure that ferry operators provide year-round services, where the market would not otherwise provide such services. In limited circumstances, exclusivity to provide a service may be appropriate. The PSC:

- is shipowner specific;
- is concluded between the State/Region and a specific operator on a given route or routes.

A Member State may impose PSOs that affect some shipping companies operating on the route(s) concerned and, at the same time, conclude PSCs with others for the same route(s) to ensure appropriate service levels are in place for the carriage of traffic to, from or between islands. This is possible provided:

- a real public service need can be demonstrated; and
- the application of the two methods concurrently is not discriminatory and is justified in relation to the public interest objective being pursued.

The ‘Cabotage Regulation’ requires that, for both imposing PSOs and concluding PSCs, the Member State shall do so on a non-discriminatory basis in respect of all community ship owners. Any subsidy for ferry services must be available to all community ship owners. All community ship owners are entitled to apply for compensation in exchange for accepting PSOs and (given that subsidies are paid via a PSC) all community ship owners should be entitled to tender for the provision of services for which subsidies will be paid.

2.5. State Aid

Where any support granted falls within the following conditions, it is deemed to constitute state aid and will need to be assessed to establish whether it is compatible with EU rules (European Commission, 2004):

- the measure is granted by the state or through state resources;
- the measure confers an advantage to an “undertaking” – “undertaking” covers any entity engaged in economic activity, and the legal status of the undertaking is not important;
- the measure is liable to distort competition by favouring certain undertakings – the competitive position of the undertaking being selectively targeted is improved by the giving of the aid in relation to other competitors; and
- the measure must have the potential for affecting trade between Member States.

Where a member state intends to give lawful state aid, it should notify the Commission and seek authorisation for the aid. Otherwise, it may be treated as unlawful unless it is deemed to be *de minimis* or otherwise exempt. *De minimis* relates to the assessment of state aid and not the ‘Cabotage Regulation’s. Issues raised by each of the ‘Cabotage Regulations’ procurement regulations and state aid rules need to be considered.

The “Altmark” decision confirmed that a State measure will not be treated as state aid where it compensates an undertaking for discharging PSOs in a way that the undertaking does not enjoy a real financial advantage in exchange for discharging PSOs

(European Court, 2003). To fall outwith the definition of state aid, the following four conditions must be satisfied:

- the receiving undertaking must actually have PSOs to discharge and these must be clearly defined;
- the basis of compensation must be calculated in an objective and transparent manner;
- compensation cannot exceed what is necessary to cover the costs in discharging the PSO, taking into account relevant receipts and a reasonable profit;
- if the undertaking concerned is not chosen under a public procurement procedure, then the level of compensation needed must be determined on the basis of an analysis of costs which an efficient undertaking would have incurred.

If a subsidy does not comply with one or more of the four Altmark criteria it must be regarded as state aid. In such cases, the measure will have to be notified to and approved by the Commission before it can be implemented. In any case involving public service compensation where the measure is within Article 87(1) of the Treaty, it may nevertheless be exempt under Article 86(2) of the Treaty where the conditions of that exemption are satisfied.

3. Public tendering of ferry services

3.1. *Methodology*

This section of the paper investigates, based on a case study approach, the practical functioning and evolution of tender procedures in four European countries - Denmark, Greece, Sweden and Norway. For each country a brief introduction to the respective ferry sector and the delivery of PSO/PSC ferry services is given. This is followed by discussion and analysis of the procurement procedures, the methods/types of support given to ferry services, the public agencies involved, and the outcomes of tendering experiences as appropriate.

Much of the information presented has been gained through direct contact with the tendering authorities and ferry operators in each of the respective countries. This has been supplemented by publicly available information. The four countries were selected because they each have recent experience of ferry service tendering and procurement in compliance with the above regulations.

The aim is to identify the main strengths and weaknesses of each procurement approach, and to review the outcomes, in an effort to facilitate development of our conceptualisation of what could comprise an 'optimal' approach to tendering/procurement of subsidised ferry services within the context of EU regulations.

3.2. Tendering of ferry services in Denmark

Denmark has more than 400 islands, some of which are connected to the mainland or other islands by bridges while others are connected by ferries. There are a total of 65 domestic ferry routes in Denmark (and approximately 20 international routes).

The Public Transport Authority (Trafikstyrelsen), an agency within the Danish Ministry of Transport, is engaged in national as well as international commitments. Its remit includes the administration of public procurement of railway and ferry transport services through organising tenders for operating contracts in accordance with Government decisions and to monitor the contractors' performance (Trafikstyrelsen 2009).

Figure 1: Danish Ferry network.
Source: http://www.trafikstyrelsen.dk.

Until 2005, many ferries in Denmark were operated by the Danish state. In 2005, limited liability companies were created. The delivery of services today is summarised in Table 1, also giving an overview on the type of contract in place on different routes. Previously, most of the routes were both owned and run by local authorities. Consequently, many of the ferry companies have now been split into two entities: one company (often owned by the 'state', mostly at municipal level) which owns the ferry (or ferries) and berth facilities and another "operating" company which runs the service. However, on certain routes the operator also owns the ferries.

Subsidies paid by the Danish State in accordance with PSCs are based on "net cost contracts", where the shipping company receives the income from tickets. The tendering authority in Denmark regulates the maximum and average ticket price for all tickets per ticket group. The level of operating subsidies for three PSCs in 2009 was as follows:

- Rønne-Køge, Rønne–Ystad (ferry services Bornholm) – about €18.4 million p/a.
- Samsø–Kalundborg - About €1.2 million p/a.
- Bøjden-Fynshav - About €1 million p/a.

Table 1: Ferry service delivery in Denmark.

Route/area	*PSO/PSC*	*Duration*	*Comments*
2 routes (Rønne-Ystad and Rønne-Køge)	PSO/PSC	5 years	PSC operated by Bornholmstrafikken A/S
18 routes	PSC/PSO	On average 5 years	PSC not exclusive, with the exception of route Spodsbjerg-Tårs.
16 routes	PSC	-	Operated by local authorities; 16 tender procedures cancelled due to lack of interest from operators.
11 routes	PSC	-	Operated by local authorities; never submitted to tender procedures.

Source: Derived from Trafikstyrelsen.

Table 2 illustrates route characteristics for the Bornholm ferry services, including key service requirements, capacity and other relevant information.

Subsidies in Denmark come out of the state budget and are split in terms of running costs and investment grants (for investments in port facilities, new ferries etc.). In the case of local government's investments in ships and port facilities, the state, the regional government and the local government each contribute 1/3rd of the annual repayments.

There is no protection for the PSO operator as such, but the local authorities can reduce fares on, for example, cross-fjord ferry lines, in order to attract more passengers and to change transport patterns. Any operator is free to establish a ferry line on a commercial basis as long as it is not in a 'monopolised area'. Monopolised areas are concession areas where operators have the exclusive rights to operate. A few ferry lines (e.g. public service lines to the islands Ærø and Anholt) and the line Sælvig (on the island Samsø)-Hou (in Jutland) are monopolised. Tendering is based on small bundles of routes (as in Bornholm's two routes tendered together) or for single route operations.

The first round of Danish tenders run according to EU regulations was realised in 1997. Tendering in the Danish context aims at benefitting users and communities and at decreasing subsidy levels. Public tendering is used as the means to ensure cost efficient transport services for citizens in all sectors of public transport in Denmark where today buses, trains, planes and ferries are all tendered. By 2009 the 3rd round of tenders was under way with the 4th round being prepared for. Specific tender contract requirements are shown in Figure 2.

Table 2: Route & service characteristics for Bornholm ferry services.

	Ronne Koge	*Ronne -Ystad*
Distance	91.6 nm	38.8 nm
Type of traffic	Primarily freight + passenger	Mainly passengers with cars + freight
Time	Night sailing (1130 – 0600) in both directions 1 weekly trip for dangerous cargo	High speed traffic, 2-5 sailings a day all year round Additional round trip with Ro-Pax per week in low season Extra sailings on key holidays with conventional car ferry
Passenger	Max. ca. 200, rarely > 100 in low season, rarely > 50 with berths	Up to 7000 per day and direction Minimum ca. 3-400 per day in winter
Freight	Up to 75 trailers/night in each direction	Up to 20 lorries a day/each direction
Effects		Introduction of high speed craft has almost doubled passenger numbers in summer

Source: Trafikstyrelsen.

Service requirements differ by contract/route and an evolution of criteria can be observed when comparing the 2nd and 3rd round of tenders. The set requirements refer to the specific characteristics of each service. One change from the 2nd to the 3rd round is the contract length from 5 years (2003/2004) to 5 years plus a one year extension option.

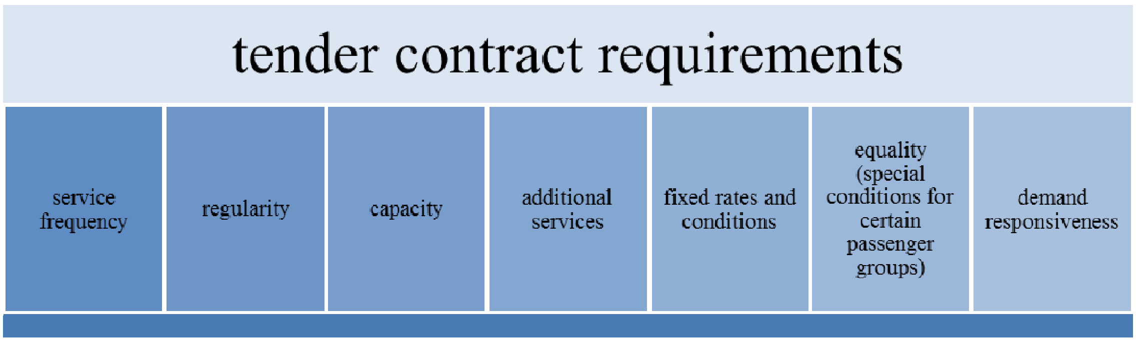


Figure 2: Tender contract requirements in Denmark.
Source: Authors.

The selection criteria have evolved from "lowest price" to a selection of the "economically most advantageous application" in the 3rd round. The criteria as used in the latter are:

- price, quality (70%),
- flexible frequency and security of supply (20%) and
- ferry quality (10%).

Additionally, the technical and financial conditions of the bidding company are taken into account. This includes turnover in the shipping business for the last 3 years and a reference list regarding the most important shipping business service contracts for the last 3 years as well as a declaration regarding debt.

Operators are required to 'bring their own ships' and, in the case of the new Bornholm contract a new catamaran currently under construction in Australia is entirely funded by the company, financed by Nordea Bank. Bornholmstrafikken is however 50% owned by the state and 50% by the private sector, so at least some of the commercial risk still rests with the state.

Procurement procedures differ by route, depending on the political focus and if the tender asks for deployment of new ferries. The current set up requires other bidders to bid against a government controlled operator (Bornholmstraffiken) that is also backed in the Nordic Ferry Services (NFS) joint venture by the 2nd largest Danish shipping company (Clipper A/S).

Table 3: Bornholm ferry tender process, service requirements and selection criteria.

	Tender process	
	Current contract	*Upcoming contract*
Contract signed	06/04	02/09
Traffic start	05/05	09/11
4 qualified operators: Scandlines Sydfynske A/S, Bornholmstrafikken/Nordic Ferry Services A/S, Mols-Linien A/S, Rederi Gotland AB		
1 application – Bornholmstrafikken (annual turnover ca. 70.6 mill €)		
Contract price/year:	ca. 17.9 mill. €, 2004 prices	ca. 27.0 mill. €, 2009 prices
	Service requirements	
	Current contract	*Upcoming contract*
Ronne – Ystad route		
Capacity (max)	32 d/a with full capacity – 8 275 pax/d and 1 575 cars/d in each direction	At least 32-38 d/a with min. 10 000 pax/d and 2 200 cars/d in each direction
Ship requirements	84 days with 1 high speed and 1 conventional car ferry	84 days with two high speed ferries
Frequency and crossing time	Min. 3 trips/d – 2 @ 75 min. and 1 @ 150 min.	Min. 3 trips/d @ 80 min.
Ronne – Koge route		
Capacity (max)	Min. 1235 lm, 400 pax/d per direction	Min 1500 lm, 400 pax/d per direction
	Selection criteria	
Criteria	Lowest price	Economically most advantageous application
Negotiation	No negotiation	Negotiation
Capacity	Number of trips/ship, ship capacity, crossing times etc.	Capacity/day and crossing time
Risk	Operator takes all economical risks, including consequences of rising oil prices	Possibility of adding fuel-surcharges as a consequence of rising oil prices
Duration	5 years	5 years with option for 1 year

Source: Trafikstyrelsen.

The specific service requirements for the 2nd and 3rd tender round of the Bornholm service are shown in Table 3. Besides minimum daily capacity, seasonal adjustment of crossing time, frequency and freight capacity, the requirement for an increase in capacity for the upcoming contract are included.

Figure 3 provides an example of the timescale involved in the procurement procedure for the major ferry services to the island of Bornholm. The whole process took about 1.5 years from time of publication of contract notices to contract signing. In the case of Bornholm the contract was signed in 2009, an order for an additional vessel was signed by the operator also in 2009, and this ship was due for delivery in 2011 in time for the new contracted service to commence.

It is important to reflect that there was then an additional 2.5 years until the contract started – so a total of 4 years lead in time. This effectively means that 1 year into a 5 year contract the tendering authority would be starting the tendering of the next contract, and that 2.5 years into a 5 year contract the existing operator might find he will no longer be the ferry operator. This creates a certain planning security for the winning operator and also provides time to effectively source tonnage for the service, if required. If the incumbent operator loses his contract to a competitor he then also has the possibility to prepare for a smooth exit and to look for alternative business.

In the latest tender round, shipping companies are to a certain extent protected against the increasing price of oil as the operator is allowed to partly raise the price of the tickets in accordance with rising costs of oil. That is, the service users pay any oil surcharge, not the state.

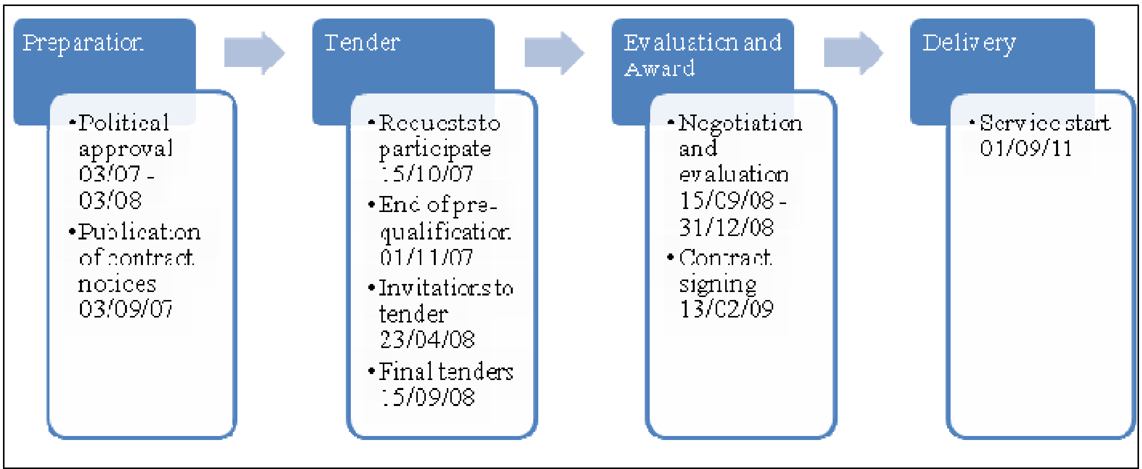


Figure 3: Procurement procedure/timescale for Bornholm ferry services.
Source: Authors.

3.3. Tendering of ferry services in Greece

Domestic ferries link the economies of 114 inhabited islands in Greece to the mainland. There are approximately 150 ports at which a total of about a thousand ferry connections are made each day. The Greek ferry market is composed of 7 regions and 58 ferry operators operating 365 itineraries. These can be split into:

- 84 routes subsidised by the Greek State;
- 30 short haul routes covering distances of up to 15 nautical miles; and
- 242 free itineraries where competitive conditions prevail (i.e. no subsidy is provided).

About 70% of the coastal passenger ships operating in Greece are deployed in the Aegean Sea area. The Aegean network has a mono-hub structure based upon the Port of Piraeus. Additionally, 70% of the national passenger transport demand is served through the Aegean network.

One of the main developments in the Greek ferry sector over the past 10 years has been the restructuring and modernisation of the fleet. Relatively new ships account for about half the fleet. *On several high-volume routes fast ferries (i.e. ships with service speed > 24 knots) are now operating*. Examples of ferry renewal include:

- *in mid-1999 Strintzis (now Blue Star Ferries) introduced two domestic routes in the Cyclades and Northern Aegean served by new high-speed ships;*
- *Minoan introduced new fast ships on the Piraeus-Crete route in 2000;*
- *new high-speed ships began to operate for Hellenic in the Aegean from 2000; and,*
- *NEL introduced the first of three fast ships in 2000.*

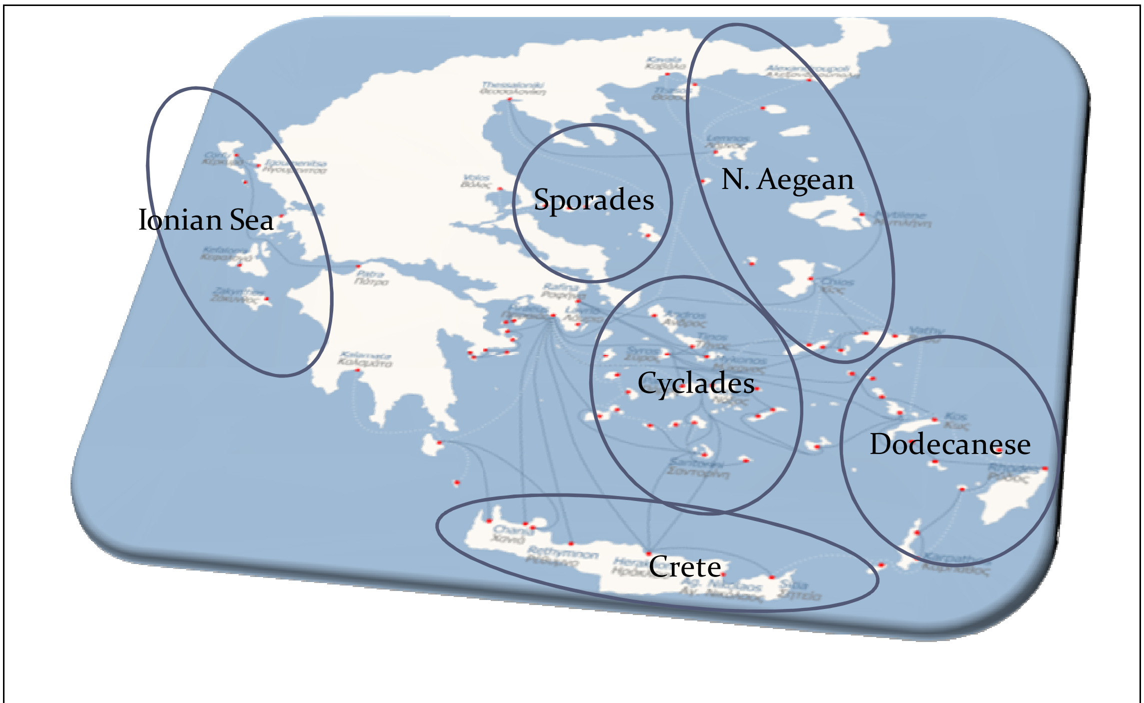


Figure 4: Greek ferry system.
Source: XTRC 2008.

Investment in new ships has also been closely linked with the development of the Athens Stock Exchange as a source of capital. The financing of Greek ferry companies has been made mainly via syndication loans (Lekakou & Vitsounis, 2008). All non-listed companies have preferred bilateral loans. The new, faster vessels have in many cases greatly reduced travel time, e.g.:

- between Piraeus to Rhodes via Kos, down to 10 hours from 18 hours;
- Piraeus to Chania, down to 5¼ hours from 10 hours.

Corresponding with a move towards faster and larger capacity ships, the number of ships employed has significantly decreased over time. However, fleet renewal has slowed over the past several years. The average age of the fleet in 2008 was 16.8 years.

In the Greek market all ferry operators are privately owned. The state does not operate any ferry services in Greece, and does not own ferries. However ports are in public ownership. Ferry companies specialise in particular routes and sea areas, often reflecting their historical roots on particular islands. For example, Minoan (now majority owned by Grimaldi, Naples) and ANEK, both based in Crete, historically held all the licenses for routes to Crete from the mainland (OECD, 2000). But the situation has since changed with new entrants such as Blue Star permitted to come into the market. This results in a challenge for authorities when deciding to what extent entry barriers should be lowered, thereby allowing for greater competition on particular routes.

Domestic routes in Greece can be categorised as follows (Lekakou & Vitsounis, 2008):

- local, profitable and competitive routes
- local, remote and with limited traffic routes
- local, of average profitability routes

Profitable and competitive routes carry significant tourist and local traffic, as well as commercial traffic. Examples include the East Cyclades complex of islands (e.g. Siros, Tinos, Mykonos). Expensive, modern, fast, high quality vessels are deployed on these routes. Access for new entrants is restricted to the minimum and competition and innovation effectively “drives” non-performing operators out of the market (Lekakou & Vitsounis, 2008).

Vessels operating in the Aegean are mainly Ro–Pax type. The market has seen a strong concentration of operators over the last several years (Lekakou & Vitsounis, 2008), albeit a fringe of smaller companies remains. Figure 5 indicates the percentage share split for passenger volumes carried by the top 5 operators; the sum of small operators is considered to account for less than 10% of the overall market.

Demand for ferry services is highly seasonal and demand among the routes is highly skewed. August accounts for some 23% of total annual passengers carried, February for only 2%. This disparity is caused not only because of differences in tourism at the island destinations but also by differences in permanent population on the islands.

The routes with low, highly seasonal traffic volumes and limited commercial viability are subsidised by the state, e.g. North East Aegean routes. Operators selected from the bidding process provide services based on a fixed income for a certain period, with set itineraries and destinations. A tendency of “lower” quality vessels on subsidised routes can be observed.

Domestic ferry services are the responsibility of the Ministry of Mercantile Marine (MMM) and the Ministry of Aegean and Island Policy (MAIP), which were merged in 2007. They are responsible for regulating and overseeing the sector, including: market entry, licensing, pricing, route scheduling, manning (hotel as well as engineering), imposition of PSOs, determination of and tendering for unprofitable routes, enforcement of licence terms, certification, control, vetting and inspection of ferries for navigational and environmental safety. The Minister issues inter alia licenses and decrees controlling prices, albeit only after a tender on a specific route that has been identified as being one that requires subsidy.

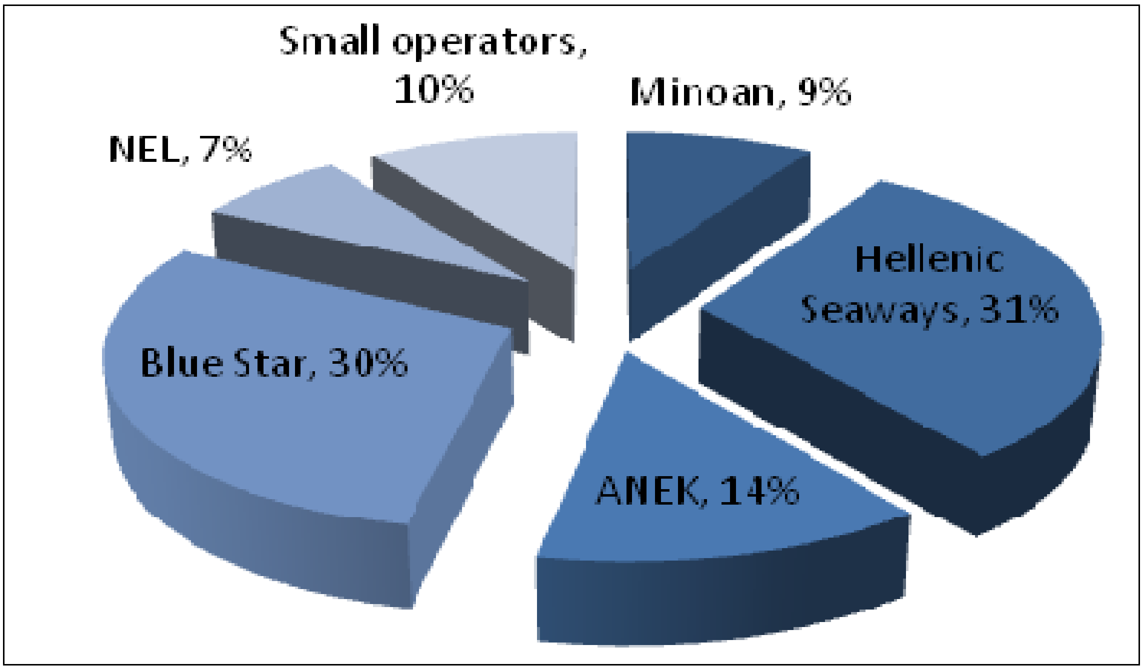


Figure 5: Market share of passengers, by ferry operator, Greece, 2005.
Source: Lekakou & Vitsounis, 2008.

Figure 6 indicates the estimated development of total aid for non-commercially viable ferry services in Greece. While the aid appears to be increasing continuously, not all this aid is paid by the State. Each ferry ticket sold in the major competitive/profitable ferry sector includes a 3% passenger surcharge which is then used (i.e. hypothecated) to help subsidise ferry services that cannot be operated in an economically viable way. This effectively means that economically viable and profitable routes help to subsidise the least profitable routes. Lines that are not subsidised understandably have a tendency to resent this 'tax'.

The overall subsidy level recorded an average annual increase of 8% between 2005-2008. The amount of public funds reached a level of over €35 million for the coverage of 84 service itineraries in 2008 (subsidised services being termed 'thin lines'). As the total aid amounted to €70 million, this means that about €35 million of aid (i.e. 50%) was raised from the competitive, profitable operators/routes. Of the 84 routes, 31 were subsidised by MMM and 53 by MAIP. These routes were served by 40 ferry operators. In comparison with other states the level of subsidy paid per service in Greece appears very low. This in part helps to explain the relatively poor quality of service provided on subsidised routes. Moreover, it is evident that the vast majority of ferry services in Greece are not subsidised at all. Figure 7 outlines the development of subsidised routes since 2002.

The Greek State Law 2932/2001 anticipated the protection of public services in order to encourage the principle of competition. It was a first attempt to harmonize the political and state framework. However, the Greek ferry industry has not yet completely conformed to EU regulation 3577/1992 (Chlomoudis et al., 2007).

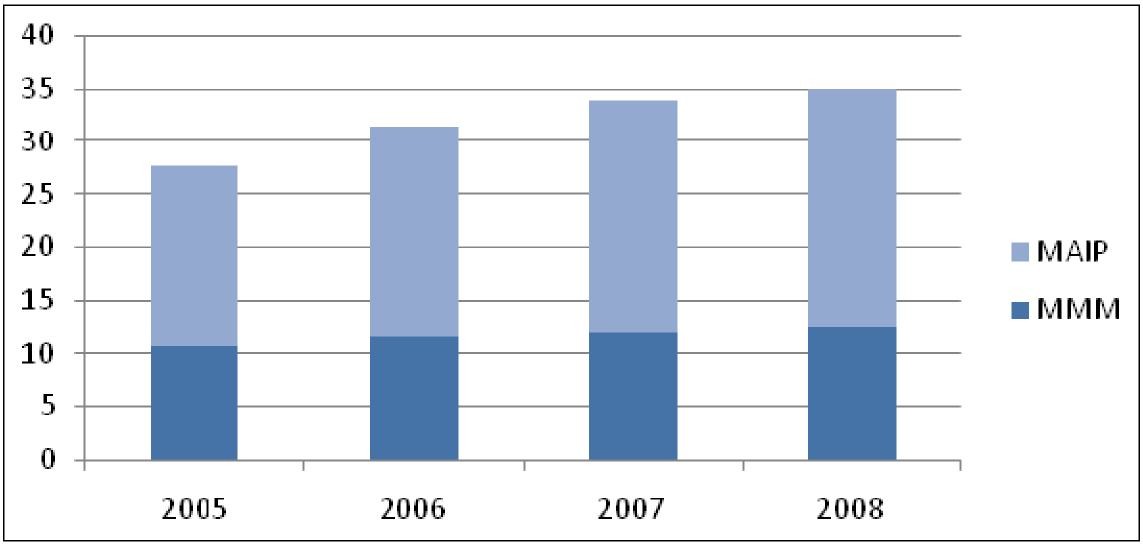


Figure 6: Greek Ferry Subsidies by number of routes (2005–2008), € Million.
Source: Annual Greek budget reports.

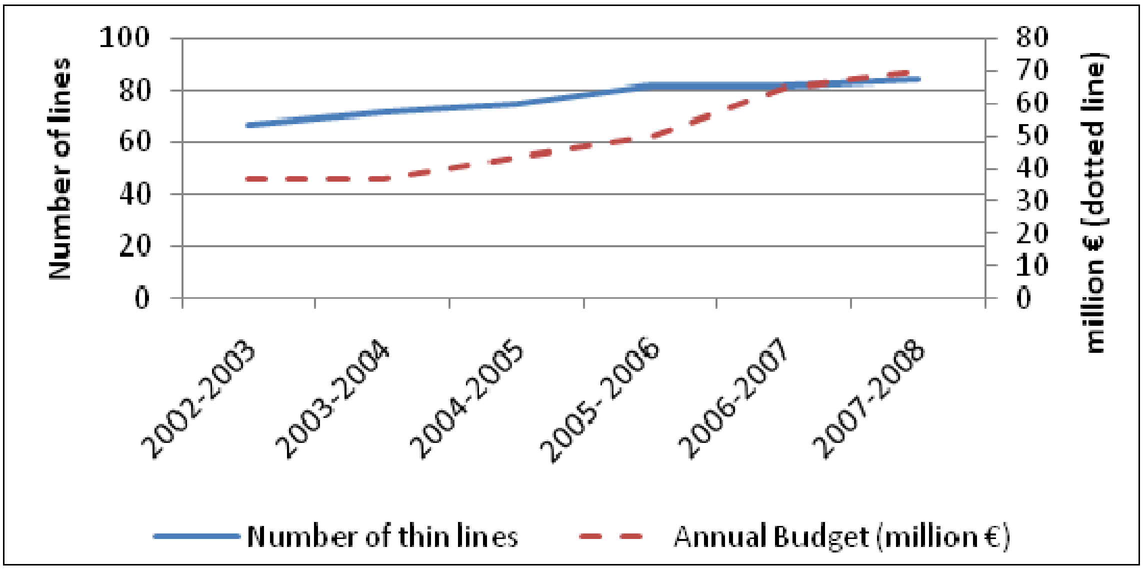


Figure 7: Aid for subsidised ferry services in Greece (2002-2008), € Million.
Source: MMM.

Procurement of services starts with analysis of whether a route requires subsidy. This is identified if no operator asks for operating permission on a specific route. If no operator can operate that route profitably, competitive tenders are run that identify the bidder requesting the minimum subsidy. Government sets the maximum subsidy level for that specific route and the minimum requirements for "essential services" (i.e. effectively the PSO) which are shown in Figure 8.

Lekakou (2007) concluded that oligopolistic market features and ineffective state policies have resulted in an unstable coastal market in Greece and this has undermined the quality of the supplied shipping services. She stresses the essential characteristics that a passenger transportation network with a strong public interest has to fulfil (i.e. system accessibility, affordability, safety/security, quality requirements, etc.) and highlights the limited extent to which these criteria have been applied as pre-conditions for developing modern and efficient Greek coastal shipping. Fulfilling these criteria would nevertheless seem to imply the need for a significant increase in public subsidies which, especially considering the current economic crisis, seems unlikely.

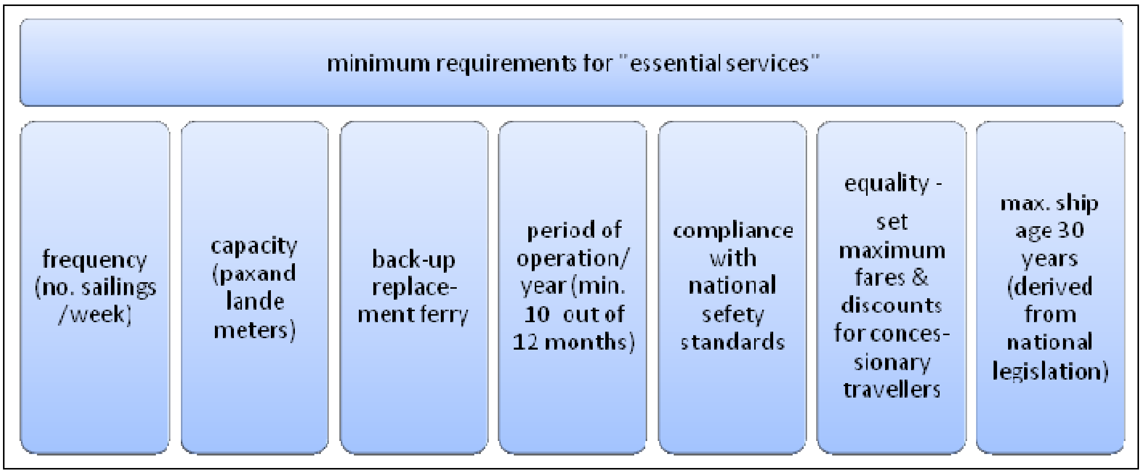


Figure 8: Minimum requirements for essential ferry services in Greece.
Source: Authors.

3.4. Tendering of ferry services in Sweden

The major subsidised ferry service connecting an island with the mainland in Sweden relates to the island of Gotland. Ferry services to the island of Gotland are tendered by Swedish National Public Transport Agency Rikstrafiken. Gotland is situated in the Baltic Sea, about 100 km from the Swedish mainland. The island has over 57,000 inhabitants, or 0.7% of the population of Sweden (Commission des Isles, 2003).

Gotland has regular ferry services on two routes guaranteed via a PSC (Figure 9): Visby–Nynäshamn and Visby–Oskarshamn, plus summer traffic to Öland, all operated by the same ferry company – Destination Gotland.

Since 1998, the private operator-owned and financed fleet serving Gotland has been completely renewed (ShipPax, 2009). Destination Gotland carried 1.5m passengers, 450,000 cars, and 700,000 lane meters of Ro-Ro cargo (i.e. over 50,000 trailers) in 2008.

In June 1998 the Swedish Parliament decided on a new transport policy, resulting in a new agency – the National Public Transport Agency, or Rikstrafiken – being established. Rikstrafiken's transport policy objective is to achieve a socio-economically efficient and long-term transport supply for the entire country (Rikstrafiken, 2009). Working together with the rail, road, sea and aviation administrations and authorities, Rikstrafiken has two main tasks:

- to procure public transport services (air, rail, sea, bus); and
- to develop and coordinate public transport throughout the country

PSCs for ferry services to and from Gotland initially came into force towards the end of the 1980s, when services were first subjected to competitive procurement procedures. The current legislation (SFS 1996:19) on limiting the right to operate ferry services to Gotland dates from 1996 and the government's decree (1997:748) that specifies the regulation came out the following year.

Figure 9: Ferry routes to Gotland.
Source: Destination Gotland.

Anyone who carries out regular sea transport between Gotland and the Swedish mainland must make a call at a mainland port at least five times a week year-round; this is a requirement of the route PSO. The purpose behind these provisions is to hinder anyone from carrying traffic only during periods when there is a great demand (i.e. summer tourists), to the detriment of the contracted services, or in other words 'cherry picking'. Requirements in the regulation of year-round traffic means that the procured ferry services are, in practice, carried out without competition from other ferry traffic (i.e. they enjoy exclusivity).

The Swedish Government has developed specific support measures for the island with the objective of ensuring an affordable price for all types of transport. In Sweden, gross contracts are the dominant contract form. This means the operator provides an agreed amount of transport services, with compensation then paid to the operator while ticket income goes to the Public transport authority.

The tendering process that preceded the current contract for the Gotland traffic was a disappointment for Rikstrafiken. In the first round, no bids were received. In the second round, which took place in early 2006, Destination Gotland, the incumbent operator, placed the only bid that was received.

Rikstrafiken has since stated its intention to alter the next bidding process. Under the new regime, there may be two separate tenders for the Gotland service – one will concern the supply of the vessels and the other will concern the operation of the ships on the route (ShipPax, 2010). Rikstrafiken also proposes shorter routes and new vessels with lower speeds and emissions.

Stakeholders (i.e. groups mainly representing islanders and island businesses) have pointed out that there are several aspects of the proposed changes that would result in deteriorations viewed from the Gotland standpoint. Some basic elements include:

- if slower crossings are introduced this may put at risk development of tourism and the travelling possibilities for firms and Gotlanders, therefore swift crossings must be considered of high value in the bidding process;
- present capacity is believed to be inadequate for freight, for the peak travelling season and to take into account any increase of cargo, cars and passengers – this is crucial for development of the Gotland economy and community; and,
- already ferry charges are considered to be too high.

State subsidies for Gotland passenger traffic have grown substantially from about €3.1 million/year in the mid-1980s to the 2009 level of approximately €46.4 million/year. Rikstrafiken's compensation for the total term of the current contract (2009-2015) is €290 million for Gotland's two routes. The proceeds from ferry services in 2005 was approximately €79.8 million, divided 53% for passenger ticket proceeds, 9% for freight traffic proceeds and 38% for Rikstrafiken's compensation.

Subsidies paid to support Gotland's two routes therefore amount to more than half the total subsidy allocated in Greece, the latter supporting some 84 routes involving a subsidy of €70 million.

3.5. Tendering of ferry services in Norway

Whilst Norway is not a member of the EU the country tends to adopt a similar approach to procurement of public services as countries within the EU.

There are estimated to be some 300 ferries operating in Norway today. Most of the ferries are owned and operated by private firms. But, like almost all other forms of public transport, many services operate at a loss and therefore require subsidy. The total annual subsidy for ferries in Norway is above €73.8 million (Minken & Killi, 2000).

Ferry links are considered to be a part of the Norwegian trunk road system. The National Public Roads Administration is responsible for supplying vehicle ferry services connecting trunk roads, and for regulating both prices and service parameters (e.g. operating hours, frequencies etc.). The Norwegian Public Roads Administration is divided into five departments each based in different parts of the country (North, South, East, West and Mid). In addition, tendering responsibilities have also been devolved to road agencies for local ferry services (mostly passenger ferries) operating in their areas (since 1994).

The traditional subsidy scheme has provided for improvements in services (e.g. frequencies, operating hours etc.), albeit at an increasing cost. Since the subsidy system was not considered sufficient in providing incentives to achieve cost efficiency, tender competitions have been introduced (Odeck & Bråthen, 2009).

It is argued that tendering has delivered substantial cost savings to the public sector in Norway (Bråthen et al., 2004). It is further suggested that where operating costs are perceived to be particularly high, or where there is a need for new vessels, tendering can prove successful as a correcting mechanism.

Ferry services are maintained by private ferry companies, each operating a monopoly franchise on a single route or across a relatively small bundle of ferry routes within a single region/area. Bundles/routes are therefore decided on a geographic area basis. A key distinction is that major trunk road ferry connections are tendered from the national level, whereas local/regional ferry services typically connecting outlying areas with the main regional towns in that area are tendered by local authorities (LAs) and/or local transport agencies.

While some single routes have been tendered (e.g. out of Tromso), several tenders have involved small bundles involving between 2-5 routes linking outlying areas and islands to the same major regional town/city at which essential services for health, recreation, education, business, and government are accessed.

Until 1990 subsidies were paid ex post on a cost-plus basis, but this was considered a weak incentive in cost efficiency terms. Since 1990, subsidies have been paid ex ante to encourage cost efficiency (Bråthen et al., 2004).

Tender documents are designed within the local area transportation departments as sufficient expertise exists from tenders which these same departments run in other areas (e.g. for buses and airlines). Evaluation is undertaken by a group from the transport department and lawyers using a pre-specified template. The administrative management of the transport department makes the final choice of operator.

In the calls for tender (first six tenders after 1994), six to nine bids were received for each route. Significant differences in the required subsidy levels were found in the tenders for each route. In five of six cases, the incumbent operators won the tender.

All tenders were evaluated on the basis of multi criteria, with cost being only one of the criteria. Non-monetary criteria included: environmental impacts, safety, quality/functionality, capacity, ferry, and option of services (Table 4). Experience in Norway indicates that costs should be valued at no more than 40% overall.

Table 4: Møre - Romsdal tender evaluation criteria.

Criterion	*Weight (%)*	*Basis*
Price	35	Annual compensation and compensation by modified route production and consideration of changed hours
Age of vessels built	35	Age of reserve vessel also attributed
Delivery-quality	30	Emergency procedures for the reserve of supply vessels (cleaning programme, maintenance of passenger amenities)

Source: Sunde et al. 2008.

In recent years, trunk route ferry tenders have resulted in investment by private operators in more than 8 new large ferries (covering 6 routes) with an estimated capex of €91 million, as well as additional quality improvements (improved frequency etc.). And on local/regional routes several operators have introduced new vessels after winning tenders (e.g. Veolia, Tide).

The duration of tendered contracts varies between 5-8 years, depending on the need for asset-specific investments. The contract duration is not seen as a huge matter for the operator: if a private company invests in a vessel with a lifetime of 20-30 years it is expected to be able to use its ferry for more than one contract period, the lower depreciation cost affording some advantage in subsequent bidding rounds. Decisions on specific ferry sizes/capacity and ferry acquisitions are left to the operating companies themselves.

Service timetabling is perceived as crucial in the tender processes, especially by ferry companies. The Norwegian Government sets quality standards. This includes retaining control of maximum prices and subsidies. A frequency premium is seen as a good incentive to overcome varying conditions (e.g. varying unit costs). Operators able to offer higher frequency are rewarded accordingly in the evaluation. Selection is based on the economically most advantageous offer, pursuant to Regulations Section Bidding. All criteria are assessed on a scale from 1 to 5, where 5 is the best score. Bidding criteria are almost evenly split between three aspects – price, vessel and quality.

Price is undoubtedly a legitimate award criterion. The vessels' age is basically a legal award criterion. It is questionable whether it is expedient to use age as the only award criterion linked to the vessels as it does not necessarily refer to a vessel's quality and suitability.

4. Discussion and conclusions

Our analysis of experiences in contracting ferry services indicates that ferry service procurement involves a continuous evolutionary process of specifying, offering, selecting, monitoring and reviewing services supplied. A number of common patterns are repeated across the range of countries under study.

Reflecting this commonality, the authors have developed on the basis of the research a 'ferry tendering route map' (Figure 10). The route map sets out what we consider to be the fundamentals of the ferry tendering process. In line with the aims of the paper, the case studies have helped facilitate development of our conceptualisation of what we consider should comprise an 'optimal' approach by transport authorities to tendering/procurement of subsidised ferry services taking into account EU regulations

The case studies suggest that Sweden, Norway and Denmark are continuously revising their procurement strategies. EU regulation is in general not seen as a barrier for development, but rather one more argument that supports the idea to secure efficient ferry services via tendering.

The procurement of ferry services is based on the same principles as procurement of transport services across other modes. Transport authorities at local, regional and national levels are managing ferry tenders, as well as managing tenders for local bus services, air services, and at the national level rail services.

In Denmark, the public procurement strategy has been successful on a number of routes, albeit less successful on some small-scale routes. The tender exercise from a public sector perspective is seen as a tool to drive operations to higher efficiency levels. However, out of 27 ferry route tender exercises undertaken in Denmark, only 2 routes actually changed operator. Nevertheless, the number of bidders – when comparing the 2nd and 3rd round – has increased, and foreign bidders are now appearing in the process.

A move towards multi-criteria evaluation in order to identify the economically most advantageous bid reflects the need to consider service quality as well as cost. A high weighting given to cost may reduce the impact of quality in the selection process. Evidence suggests that the duration of the contract (5 to 6 years) is not a barrier to the deployment of new vessels, notwithstanding the fact that tendering authorities do not necessarily require brand new tonnage at every round (e.g. used vessels may be acceptable).

The Greek approach provides for a number of interesting observations. The Greek ferry industry was able to significantly improve fleet structure, especially driven by the threat of external entrants and helped by sophisticated private finance strategies. While these improvements primarily reflect experience on higher volume profitable routes, they do not seem to have reached the more peripheral and less densely populated islands. For the latter, incentives given through the procurement process do not seem to be sufficient enough to stimulate service improvements (Lekakou & Vitsounis, 2008). Moreover, the one year duration of contracts is clearly insufficient to allow operators to make necessary investments in new or replacement vessels; this should be changed to longer periods of at least 5-6 years, as is now common practice in other states.

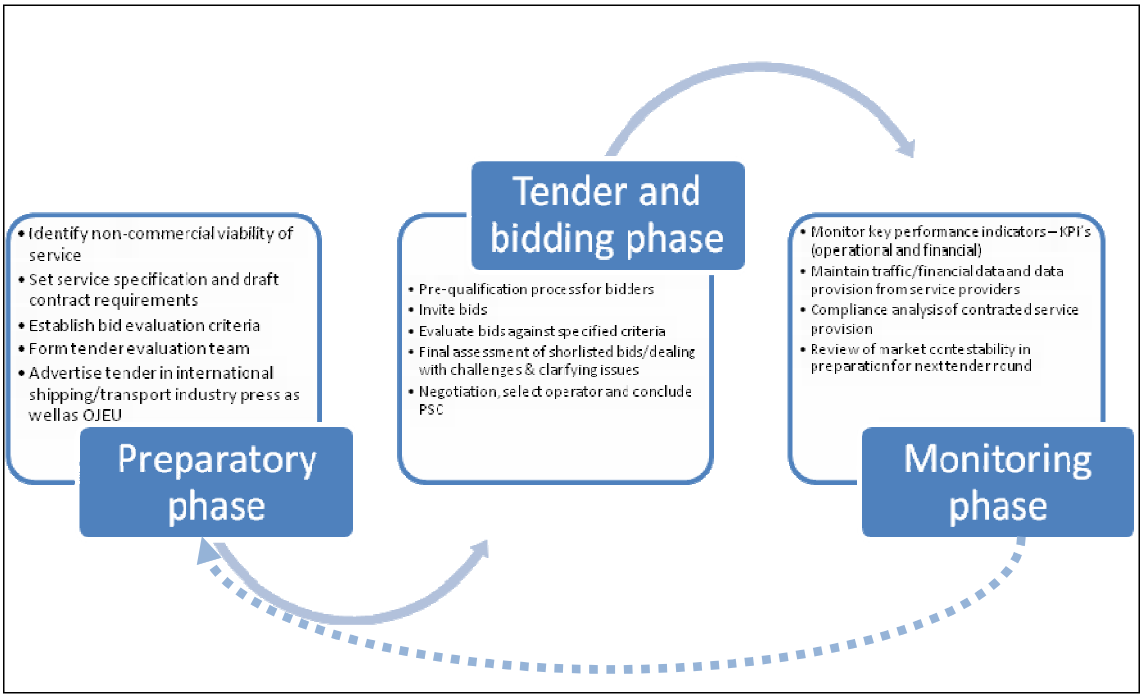


Figure 10: Ferry tendering route map.
Source: Authors.

The level of customer satisfaction in respect of subsidised routes/services in Greece is rather low, with complaints about the limited frequency of service, the poor quality of vessels and slow speed (Chlomoudis et al., 2007). The evidence further suggests that overall subsidy levels in Greece may be inadequate in terms of improving service quality. Moreover, the strategy of forcing profitable ferry operators (and their customers) to help subsidise unrelated non-profitable services also seems questionable, primarily as this effectively penalizes/taxes successful non-subsidised services.

In Sweden, the case of Gotland demonstrates that the contracted operator has delivered improvements to the service in terms of modern ships and service quality, and achieved growing passenger numbers, albeit at rather high ongoing cost to the public sector On the other hand the responsible agency, Rikstrafiken, is continuously pushing for more competition in the market, albeit the latest move to open up the market seems somewhat confused (i.e. the proposal to have separate tenders for ships and operations).

The lack of interest from other bidders (aside from the incumbent operator) suggests the possibility of barriers to entry (e.g. access to vessels).

The Norwegian experience in procurement of ferry services illustrates that: competitive tendering of individual routes has been perceived as a means to improve cost efficiency; it is a challenge to design the right measures that enhance allocative efficiency (allocative efficiency occurs when there is an optimal distribution of goods and services, taking into account consumer preferences), and; tender specification and procurement expertise needs to evolve over time and requires constant adjustment based on market development and experience. The Norwegian case gives clear evidence of improvements achieved in terms of service quality through tendering, including private investment in new ships. It also defies the myth that subsidies necessarily increase after deploying tendering procedures. We consider the Norwegian experience to represent probably the most successful and best developed approach to procurement of ferry services in Europe.

A tendency towards the further expansion of pan-European integrated transport providers into ferry markets via tenders can be observed. These entities are increasingly familiar with tender procedures, across different transport modes, throughout Europe. This is mainly because such firms are now frequently involved in tendering for public transport bus, rail, as well as ferry services, throughout Europe.

Submitting an operator to a tender compels it to reassess the value for money it is providing, and the efficiency of its services, thus leading to better services; this is especially true when there is the "threat" of effective competition. Also noticeable is the increasing role of private operators, with private firms willing to invest in new vessels, in turn limiting the need for state managed/owned operations.

Tendering is clearly a continuous process which is repeated over set periods. Arguments that the process is found to be too complex, long in duration and too expensive may be related to procurement in countries with a weak or discontinued procurement strategy. The timing of tendering procedures is decisive, since only strict timelines can provide the necessary framework for operators to prepare attractive bids.

Finally, it should be remembered that public tendering of ferry services is only necessary when the market is unable or unwilling to provide an adequate service on a given route. There are many routes throughout Europe where the market does indeed provide an adequate service, without subsidy, and which therefore negates in these circumstances any need for tenders.

Acknowledgements

An earlier version of this paper was presented at the International Association of Maritime Economists (IAME) Conference in Lisbon, July 7-9, 2010. The authors thank the reviewers for helpful suggestions on improving the paper. Any errors are the responsibility of the authors.

References

Bråthen, S., Hervik, A., Odeck, J. & Sunde, Ø. (2004) "Competitive tendering versus subsidy schemes in public transport services: comparing efficiency in the provision of ferry services", *Proceedings of the European Transport Conference,* Strasbourg, France, September.

Chlomoudis, C.I., Pallis, P.L., Papadimitriou, S. and Tzannatos, E.S., (2007) "The liberalisation of maritime transport and the island regions in EU, Evidence from Greece", *European Transport* 37: 1-15.
Commission des Isles (2003) http://www.commissiondesiles.org/pub/docs2/58_eng_-_conclusions_seminar_on_transports_western_isles.pdf.
European Commission (2004) "Commission communication C (2004) 43 - Community guidelines on State aid to maritime transport. (2004/C 13/03)", European Commission, Brussels, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2004:013:0003:0012:EN:PDF.
European Commission (2006) "Council Regulation (EEC) No 3577/92 of 7 December 1992 applying the principle of freedom to provide services to maritime transport within Member States (maritime cabotage), OJ L 364, 12.12.1992", European Commission, Brussels, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31992R3577:EN:HTML.
European Commission (2009) "Commission Decision of 28.10.2009 o the State aid No C 16/2008 (ex NN 105/2005 and NN 35/2007) implemented by the United Kingdom of Great Britain and Northern Ireland, Subsidies to CalMac and NorthLink for maritime transport services in Scotland, C(2009)8117 final", European Commission, Brussels, http://ec.europa.eu/energy/coal/state_aid/doc/decisions/2008/2008_0016_uk_c.pdf.
European Court (2001) http://eur-lex.europa.eu/smartapi/cgi/sga_doc?smartapi!celexplus!prod!CELEXnumdoc&lg=en&numdoc=699J0205.
European Court (2003) "Case C-280/00, Altmark Trans GmbH and Regierungspräsidium Magdeburg v Nahverkehrsgesellschaft Altmark GmbH", European Court, Luxembourg, 24 July 2003, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:62000J0280:EN:HTML.
Kay, N (2008) www.brocher.com/Ferries/Barrot%20answer.doc.
Lekakou, M. B. (2007) "The eternal conundrum of Greek coastal shipping", In Pallis A.A. (ed) Maritime Transport: The Greek Paradigm, *Research in Transportation Economics* 21: 257-296.
Lekakou, M. B. and Vitsounis, T. K. (2008) "Measuring market concentration in Greek coastal waters", *Proceedings of the International Association of Maritime Economists (IAME) Conference*, Dalian, China.
Minken, H. & Killi, M. (2000) "Productivity growth in Norwegian ferry links 1988-1996, and implications for regulation", Institute of Transport Economics, Oslo, http://www.thredbo.itls.usyd.edu.au/downloads/thredbo7_papers/thredbo7-workshopC-Minken-Killi.pdf.
Odeck, J. and Brathern, S. (2009) "The efficiency of Norwegian ferries in providing public transport services", *International Journal of Transport Economics* XXXVI, 1.
OECD (2000) http://www.rekabet.gov.tr/word/uluslararasi/CD/OECD2008/CountryStudies/Greece-ElectricityDomesticFerriesTrucking.pdf.
Rikstrafiken (2009) http://www.rikstrafiken.se.
ShipPax (2009) "ShipPax Finance 09", *ShipPax*, Halmstad, Sweden.
ShipPax (2010) "Back to the old days for Gotland", *Cruise & Ferry Info*, Halmstad: 24.
Trafikstyrelsen (2009) www.trafikstyrelsen.dk.
Sunde, Ø., Husdal, J., and Svein Bråthen, S. (2008) "Anbud I Fylskommunale Bartruter - Vurdering av dagens praksis og forslag til endringer", *Rapport 0810,* Møreforsking Molde AS, Molde, Norway.
Transport Research Institute (2010) "Methods of Ferry Service Delivery and Operation, Competition and Procurement and Environmental Issues", *Final Report for Scottish Government*, TRI-Edinburgh Napier University, February.