MOST LIKELY TRAFFIC EQUILIBRIUM ROUTE FLOWS — ANALYSIS AND COMPUTATION

T. Larsson [1]
J. T. Lundgren [1]
M. Patriksson [2]
C. Rydergren [1]

Abstract

When analyzing equilibrium traffic flows it is usually the link flows and link travel demands that are of interest, but in some certain cases analyses require the knowledge of route flows. It is well known that the route flows are non-unique in the static and deterministic cases of traffic equilibrium. Furthermore, different assignment methods can generate different route flow output. We show how this non-uniqueness can affect the results in applications such as in the O–D estimation/adjustment problem, in the construction of induced O–D matrices, exhaust fume emission analyses and in link toll usage analyses. We state a model for finding, uniquely, the most likely route flows given the equilibrium link flows, and propose a solution algorithm for the problem based on partial dualization. We present computational results for the proposed algorithm and results from an application to exhaust fume emissions.

Keywords: Traffic equilibrium, route flows, entropy maximization.

1 Introduction

Equilibrium traffic assignment based on the Wardrop conditions is frequently used in urban traffic planning to predict and analyze traffic flows. The equilibrium conditions are based on the assumption of rational route choice behaviour by the individual user and define a situation where the routes used by the travellers between each origin and destination are the shortest.

There are many assignment methods available for finding the unique link flow pattern that satisfies the user equilibrium condition. Standard solution methods for the traffic assignment problem often operate in the link flow space. One class of such methods is based on the method of Frank and Wolfe (see, e.g., Sheffi, 1985, Chapter 5, and Patriksson, 1994, Section 4.2, for overviews). Other methods, more recently applied to the assignment model, operate in the route flow space and therefore provide a route flow solution consistent with the equilibrium link flow solution (e.g., Larsson and Patriksson, 1992).

As is well known, the route flows are usually not uniquely determined in the deterministic assignment model; trips assigned to routes between origins and destinations can almost always be swapped to other routes such that the link flows are unchanged. The route flow solution obtained depend therefore on the solution method applied, and for a given solution method also on the initiation of the solution procedure, parameter choices in the method, etcetera. Nevertheless, a route flow solution is needed as input to methods used in several applications.

2 Illustrative examples and applications

In this section we present a comparison of the route flow output from two assignment algorithms and the result of the usage of these route flows in four potential applications. The assignment

[1]Department of Mathematics, Linköping University, SE-581 83 Linköping, Sweden
[2]Department of Mathematics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden

procedures used to compute equilibrium route flows are: (1) the Frank–Wolfe (FW) method, supplied with a procedure to generate an equilibrium route flow and (2) the route flow based disaggregate simplicial decomposition (DSD) method of Larsson and Patriksson (1992). In all the examples, the equilibrium flows have been determined to a (very) high accuracy with both methods.

2.1 Illustrative examples

We here illustrate the differences in the route flow output between the FW and the DSD algorithms. We first study in particular the topology of the two sets of route flow solutions in O–D pair $(14, 10)$ in the Sioux Falls network (LeBlanc et al., 1975). Both the FW method and the DSD method generate two routes in this O–D pair. The travel demand from 14 to 10 is 2100 vehicles (measured in vehicles per day). The route flows in the FW solution are 837 for the route 14–11–10 and 1263 for the route 14–15–10 (see Figure 1). The route flows in the DSD solution are 368 and 1732, respectively.

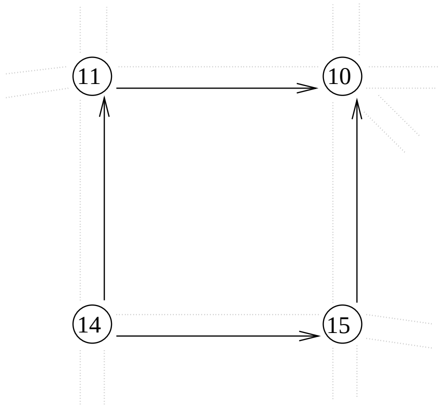

Figure 1: Sioux Falls O–D pair (14, 10).

We next give aggregated measures of the difference between the route flow solutions for applications of the FW and DSD methods to the network of Winnipeg (Florian and Nguyen, 1976). We first restrict the analysis to the 65 largest origin–destination (O–D) pairs (which provide 5% of the total travel demand) where the route flow solution from the DSD method includes five or more routes in each pair. Comparing this route flow solution to the one from the FW method for these O–D pairs we have that, on average, 53% of the O–D demand use identical routes. If we calculate the same measure for the 270 largest O–D pairs (which provide 21% of the total travel demand) which have two or more equilibrium routes in each pair, then the corresponding number is 74%. If we instead analyze the O–D disaggregated link flow solution for the seven links with the largest flow and with 258 O–D pairs or more present on each of the links we have that 92% of the flow originate from identical O–D pairs. This high percentage can be explained by the fact that most of the O–D pairs present on these links only have one route between the origin and destination in both of the route flow solutions.

2.2 Applications

We give four examples of potential applications where a route flow solution is used as input and present results based on the route flow output from the DSD and the FW method. In the examples the test network of Sioux Falls is used.

Origin–destination matrix estimation The estimation or adjustment of O–D trip matrices is a well known problem in traffic planning. The estimation is often based on traffic count observations (e.g., Cascetta and Nguyen, 1988, and Fisk, 1988). Assume that we wish to estimate an O–D matrix consisting of N elements (O–D pairs) where we have observed traffic counts $\bar{v}_i$ for M links. The fundamental equation in a standard O–D matrix estimation problem can be expressed as

$$\sum_{j=1}^{N} p_{ij} t_j = \bar{v}_i, \quad i = 1, ..., M,$$

where p_{ij} is the proportion of trips in O–D pair j using link i and t_j is the number of trips in O–D pair j which is to be determined. In the case of a congested network the coefficients p_{ij} are functions of t_j and can only be determined in an iterative procedure. Usually, an equilibrium assignment for a given t_j is made alternatively with an adjustment of the O–D matrix based on the coefficients p_{ij}. In the procedure the coefficients p_{ij} are extracted from a route flow solution (or an O–D disaggregated link flow solution) and the result is therefore dependent of the route flow solution obtained.

Construction of an induced origin–destination matrix Most urban areas can be divided into a number of subareas, each with different characteristics, for example, a heavily congested core area and its surrounding arterials. To analyze each of the subareas separately we need consistent induced O–D matrices for each of the subareas. A transfer decomposition approach for calculating an induced O–D matrix is presented in Hearn (1984) and Barton et al. (1989). The induced O–D matrix obtained by this approach is non-unique. Another approach (described in Drive project v1054), which uses route flow information, to obtain O–D matrices for each of the subnetworks is the following two step procedure: (1) perform an equilibrium assignment in the complete network using an algorithm that provides route flow information, and (2) use the route flow information to calculate the induced O–D matrices. Also by this approach the induced O–D matrices are non-unique and are dependent of the route flow solution used.

We give an example using the Sioux Falls network where we are interested in a consistent O–D matrix for the sub-network of nodes 14, 15, 23 and 22. We introduce the new nodes A, B, C, D, E and F to represent the area outside the defined sub-area (see Figures 2 and 3). Two route flow solutions are studied, the route flow output from the FW method and the DSD method. The number of trips from the node A to the nodes $\{14, 15, 23, 22, \text{B}, \text{C}, \text{D}, \text{E}, \text{F}\}$ are

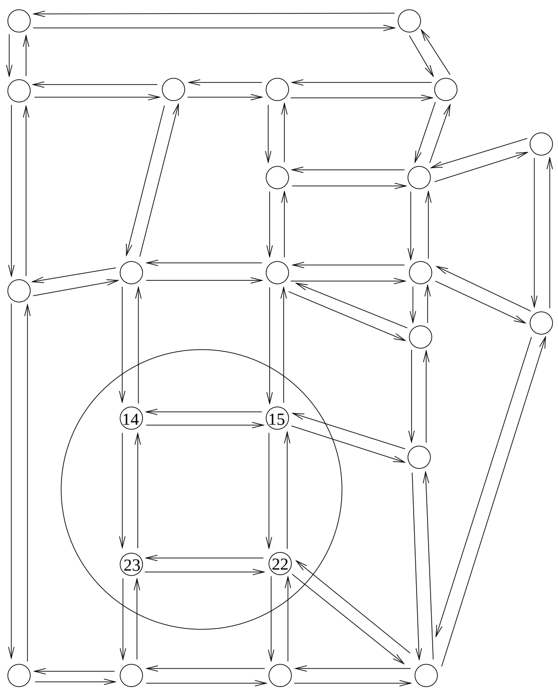

Figure 2: The Sioux Falls network.

$\{4112, 740, 527, 2914, 0, 459, 0, 0, 1024\}$ and $\{4592, 767, 586, 2353, 0, 420, 0, 0, 1058\}$ for the DSD and FW methods, respectively. The number of O–D trips from node 14 to the nodes $\{\text{A}, \text{B}, \text{C}\}$ are $\{3974, 2226, 2900\}$ and $\{4582, 1618, 2900\}$ for the DSD and FW methods, respectively. (The

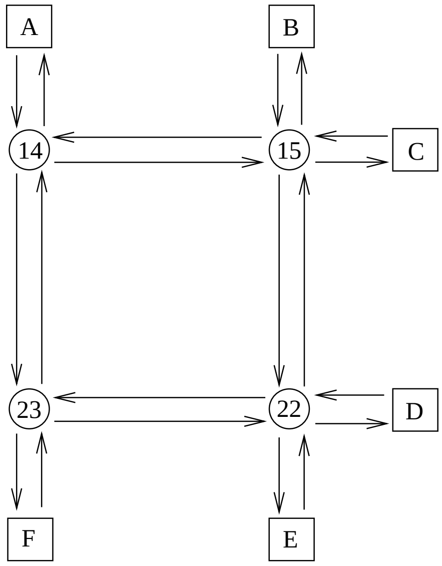

Figure 3: Sub-area in the Sioux Falls network.

number of trips to nodes 15, 23, 22 are unchanged from the original O–D matrix and the number of trips to the other nodes are zero.) Clearly, the two route flow solutions induce significantly different O–D matrices.

Exhaust fume emission analysis The application that gave the first inspiration for this work was the application of exhaust fume emission analysis and especially the analysis of exhaust fume emission from vehicles starting with cold engines. Experiments have shown that cold engines in combination with a cold catalytic converter emit large amounts of, for example, nitrogen oxides (NO_x). When analyzing the spatial distribution of emissions from cold starts the emission distribution is given as a function of the travelled distance, for which an equilibrium route flow information is needed (e.g., Sérié and Joumard, 1997 and Edwards, 1998). An example of a simplified exhaust fume emission analysis is given using the network in Figure 4. Assume that a group of 10 vehicles travel from node A to C and another group of 10 vehicles travels from node B to node C. Assume further that the equilibrium link flows are 10 on every link. One set of route flows consistent with the equilibrium link flows is where group one uses the route that includes the upper link between B and C and group two uses only the lower link. Another set of consistent route flows is where group one uses only the lower link between B and C and group two use the upper link. Let us assume that each link has the length of 200 meters and that the effects of the cold starts result in an emission only during the first 200 meters travelled. We observe that the first set of route flows results in an emission corresponding to 10 cold starts on the lower link between B and C and for the second set of route flows the lower link between node B and C is free from emission originating from cold starts. The conclusion

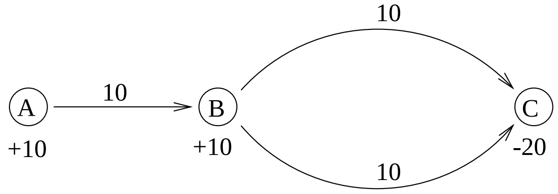

Figure 4: Example network for emission analysis.

is that the analysis made in this network is very sensitive to the set of route flows used.

Link toll usage analysis The old remedy for traffic congestion was to expand the traffic network. Experiments have shown that this approach is undermined by the level of suppressed travel demand as new roads may result in even more traffic. Congestion pricing and especially link tolls is a viable alternative for improving the efficiency of the road system. Other motives for congestion pricing are environmental protection and/or infrastructure financing (see The Economist, December 6, 1997 for a discussion on these subjects). When implementing a congestion pricing scheme where link tolls are introduced, there may be several reasons for analyzing which traveller groups are affected by the link tolls. One example is in the case of reimbursement of link toll charges to the travellers (see e.g., Small, 1992). To target these reimbursements an analysis of the origin of the travellers using the tolled links is made (Rossi et al., 1989). In such analyses O–D disaggregated link flows (or route flows) are needed as input.

We give an illustrative example where we analyze the origin of the travellers on a specific link in the network of Sioux Falls. Let us assume that we would like to determine the origin of the travellers using the link $(14, 15)$, marked in Figure 5. The equilibrium link flow on $(14, 15)$ is 9035 vehicles. The FW and the DSD methods are used to obtain two route flow solutions. According to both route flow solutions the vehicles that use the link are O–D pairs with origin nodes in the set $\{11, 12, 13, 14, 23, 24\}$. In the FW route flow solution the flows from these origins on link $(14, 15)$ are $\{1043, 356, 95, 6266, 1156, 119\}$, respectively and in the DSD route flow solution the corresponding flows are $\{1356, 344, 59, 6643, 536, 96\}$.

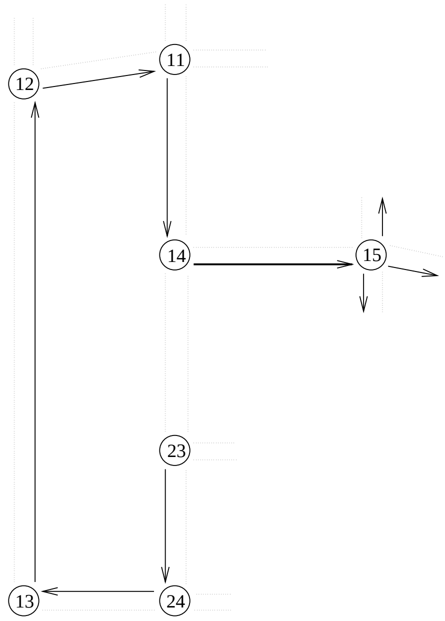

Figure 5: Link example network.

In this section we have given four examples of applications where a more detailed description of the equilibrium flow than the link flow solution is needed as input. Two of the applications, the O–D matrix estimation and the link toll usage analysis, need to have the link flow disaggregated by commodity, that is, O–D link flows. The other two applications, the construction of an induced O–D matrix and the exhaust fume emission analyses need an even more disaggregated flow description, that is, a route flow solution. For the construction of an induced O–D matrix an O–D link flow solution gives us unique row and column sums of the O–D matrix, but a route flow solution is necessary to determine (uniquely) the actual O–D matrix. In the exhaust fume emission analysis we need the route flow information to calculate the travelled distance. However, if the exhaust fume emissions are given as a function of the time travelled, it is sufficient to have an O–D link flow solution.

These applications, together with the numerical examples, illustrate how the non-uniqueness of the route flows can affect the results in each of the applications and illustrate the need for a procedure that provides a specified route flow solution (or in some cases a specified O–D link flow solution) that is reasonable according to some specified criterion.

In this paper we state an entropy problem (Rossi et al., 1989), which as its optimal solution gives a unique set of reasonable route flows (see Section 3). For this problem, we propose an exact solution method based on a partial dualization approach.

In Section 3 necessary notation for the analysis of equilibrium flows is introduced, and previous work on finding a reasonable and uniquely defined route flow solution (the most likely route flow solution) is reviewed. We establish that the solution to the model stated for finding the most likely route flows is a limiting case of the stochastic user equilibrium route flows when the travellers' cost perception dispersion tends to zero. The relations between the solution to the problem of finding the most likely O–D link flow according to Akamatsu (1997) and the solution to the most likely route flow problem are discussed. We also give a procedure for finding the most likely route flow solution from the most likely O–D link flow solution. In Section 4, we present a solution procedure for the problem of finding the most likely route flows. In Section 5 promising experimental results for medium scale networks are reported, and, in Section 6, results from an application to exhaust fume emissions is presented.

3 Most likely equilibrium flows

3.1 Preliminaries

Consider a network describing an urban region with sets of nodes, $\mathcal{N}$, and directed links, $\mathcal{A}$. Between every O–D pair of nodes, $(p,q) \in \mathcal{C}$, there is a fixed trip demand, d_{pq}. Each link, $a \in \mathcal{A}$, is assumed to be associated with a travel cost (or time) which depend on the total number of travellers, f_a, on link a, defined by a positive, continuous, and strictly increasing function, c_a. The link travel costs are assumed to be additive, that is, each route cost is assumed to be the sum of the travel cost for the links defining the route. The solution to the assignment model defines a situation where each traveller minimizes his/her own travel cost, and results in the Wardrop equilibrium where the travel cost on the routes used are equal and equal to or less than those which would be experienced by a single vehicle on any unused route.

The equilibrium conditions are formulated using route flows, h_{pqr}, that is, the flow between origin p and destination q on route $r \in \mathcal{R}_{pq}$, where $\mathcal{R}_{pq}$ denotes the set of routes in this specific O–D pair. A set of non-negative route flows satisfying the O–D demand can be described by the following constraints which define the route flow polytope, $\mathcal{H}^r$:

$$\sum_{r \in \mathcal{R}_{pq}} h_{pqr} = d_{pq}, \qquad \forall (p,q) \in \mathcal{C}, \tag{1a}$$

$$h_{pqr} \geq 0, \qquad \forall r \in \mathcal{R}_{pq}, \quad \forall (p,q) \in \mathcal{C}. \tag{1b}$$

To relate routes and links we define $\delta_{pqra} = 1$ if route r between p and q includes link a, and 0 otherwise. We can now express the relation between link and route flows as

$$f_a(h) = \sum_{(p,q) \in \mathcal{C}} \sum_{r \in \mathcal{R}_{pq}} \delta_{pqra} h_{pqr}, \quad \forall a \in \mathcal{A}. \tag{2}$$

The cost of using route r between p and q can, by the additivity assumption and (2), be expressed as

$$c_{pqr}(h) = \sum_{a \in \mathcal{A}} \delta_{pqra} c_a(f_a).$$

The Wardrop equilibrium condition can now be expressed as

$$\begin{aligned} h_{pqr} > 0 &\implies c_{pqr} = \pi_{pq}, \quad \forall r \in \mathcal{R}_{pq}, \\ h_{pqr} = 0 &\implies c_{pqr} \geq \pi_{pq}, \quad \forall r \in \mathcal{R}_{pq}, \end{aligned}$$

where π_{pq} denotes the minimum route cost in O–D pair (p,q). Denote the equilibrium routes in O–D pair (p,q) by

$$\mathcal{R}^*_{pq} = \{r \in \mathcal{R}_{pq} | c_{pqr}(h) = \pi_{pq}\}. \tag{3}$$

The equilibrium link flow solution, f^*, can be shown to be found as the optimal solution to the problem (e.g., Patriksson, 1994, Theorem 2.1)

$$\min_{h \in \mathcal{H}^r} \sum_{a \in \mathcal{A}} \int_0^{f_a(h)} c_a(s) ds. \tag{4}$$

The relation between equilibrium link and route flows is

$$\sum_{(p,q)\in\mathcal{C}} \sum_{r\in\mathcal{R}^*_{pq}} \delta_{pqra} h_{pqr} = f^*_a, \quad \forall a \in \mathcal{A},$$

where the set of equilibrium route flow solutions is defined by the polytope, $(\mathcal{H}^r)^*$,

$$\sum_{r\in\mathcal{R}^*_{pq}} h_{pqr} = d_{pq}, \qquad \forall (p,q) \in \mathcal{C}, \tag{5a}$$

$$\sum_{(p,q)\in\mathcal{C}} \sum_{r\in\mathcal{R}^*_{pq}} \delta_{pqra} h_{pqr} = f^*_a, \qquad \forall a \in \mathcal{A}, \tag{5b}$$

$$h_{pqr} \geq 0, \qquad \forall r \in \mathcal{R}_{pq}. \quad \forall (p,q) \in \mathcal{C}. \tag{5c}$$

This set is usually not a singleton, as was demonstrated in Section 2. We consider the following model for finding a unique set of "reasonable" equilibrium route flows. The problem is defined using route flow variables and the optimal solution defines the set of *most likely route flows*, denoted h^*.

$$\max -\sum_{(p,q)\in\mathcal{C}} \sum_{r\in\mathcal{R}^*_{pq}} h_{pqr} \ln h_{pqr} \tag{6a}$$

subject to

$$\sum_{r\in\mathcal{R}^*_{pq}} h_{pqr} = d_{pq}, \qquad \forall (p,q) \in \mathcal{C}, \tag{6b}$$

$$\sum_{(p,q)\in\mathcal{C}} \sum_{r\in\mathcal{R}^*_{pq}} \delta_{pqra} h_{pqr} = f^*_a, \qquad \forall a \in \mathcal{A}, \tag{6c}$$

$$h_{pqr} \geq 0, \qquad \forall r \in \mathcal{R}^*_{pq}, \quad \forall (p,q) \in \mathcal{C}. \tag{6d}$$

In the problem (6) the route flow entropy is maximized under the restriction that the route flows are consistent with the unique equilibrium link flows.

A derivation of the problem (6) was presented in Rossi et al. (1989). The derivation of the model is based on the minimum information principle in Snickars and Weibull (1977). The equilibrium link flow solution is defined as a macro state which arises as a result of the route choices of individual travellers. The route choices between minimum, and equal, cost routes of a set for individual travellers define a set of micro states. All micro states consistent with the macro state are argued to be equally probable, since the travellers are indifferent to which shortest route they use. Using well known principles from information theory they argue that the most likely route flow solution is the one which corresponds to the largest number of micro states among those which are consistent with the observed macro state. After an approximation using the Stirling formula the problem (6) is obtained for determining this (unique) micro state.

The indifference of the travellers and the assumption that all micro states are equally probable can be shown to also follow from the *efficiency principle* on which an alternative derivation based on behavioural assumptions can be made (Smith, 1978 and Smith, 1983).

In Rossi et al. (1989) a procedure to obtain the most likely route flows described as a two-step procedure is presented. First, the construction of the entropy program requires that the link flow solution is calculated and all equilibrium routes are enumerated. Second, the entropy program is solved. Rossi et al. present a numerical result for a small-scale problem where the general-purpose nonlinear programming code MINOS was used. The solution to problem (6) is analyzed for a small scale network in Bell and Iida (1997, Section 5.5).

3.2 An alternative derivation

The definition of a stochastic user equilibrium extends the Wardrop conditions (the deterministic user equilibrium). To reflect a variation in the user perception of cost a random component, here assumed to be Gumbel distributed, is added to the route cost. The stochastic user equilibrium can be defined as the optimal solution to the following optimization model (Fisk, 1980),

$$\min \sum_{a \in \mathcal{A}} \int_0^{f_a(h)} t_a(s)ds + \frac{1}{\theta} \sum_{(p,q)\in\mathcal{C}} \sum_{r \in \mathcal{R}_{pq}} h_{pqr} \ln h_{pqr}, \tag{7a}$$

subject to

$$\text{(1a), (1b) and (2),} \tag{7b}$$

where $\theta > 0$ is the dispersion parameter. Denote the optimal route flow solution to problem (7) for a fixed θ by $h^*(\theta)$. The following theorem tells us that the optimal solution to the stochastic user equilibrium problem when $\theta \to \infty$ coincides with the most likely route flow solution, h^*.

Theorem 1. $\lim_{\theta\to\infty} h^*(\theta) = h^*$.

Proof. See Appendix A. □

Akamatzu et al. have in a set of papers developed a *stochastic user equilibrium* model stated in O–D disaggregated link flows using an "entropy decomposition" technique (Akamatsu et al., 1990, Akamatsu, 1996, Akamatsu, 1997). The "entropy decomposition" technique is applicable also to the problem of finding the most likely O–D link flows; such a model is given in Akamatsu (1997). The model is stated in O–D link flow variables, f_{pqa}, that is, the O–D flow in pair $(p, q) \in \mathcal{C}$ on link $a = (i, j) \in \mathcal{A}$:

$$\min \sum_{(p,q)\in\mathcal{C}} \left(\sum_{a\in\mathcal{A}} f_{pqa} \ln f_{pqa} - \sum_{j\in\mathcal{N}} \left(\sum_{i\in\mathcal{T}(j)} f_{pq(i,j)} \right) \ln \left(\sum_{i\in\mathcal{T}(j)} f_{pq(i,j)} \right) \right) \tag{8a}$$

subject to

$$\sum_{(p,q)\in\mathcal{C}} f_{pqa} = f_a^*, \qquad \forall a \in \mathcal{A}, \tag{8b}$$

$$\sum_{i\in\mathcal{T}(j)} f_{pq(i,j)} - \sum_{i\in\mathcal{S}(j)} f_{pq(j,i)} = \begin{cases} d_{pq}, & \text{if } j = q, \\ -d_{pq}, & \text{if } j = p, \\ 0, & \text{otherwise,} \end{cases} \qquad j \in \mathcal{N}, \quad (p,q) \in \mathcal{C}, \tag{8c}$$

$$f_{pqa} \geq 0, \qquad \forall a \in \mathcal{A}, \quad \forall (p,q) \in \mathcal{C}. \tag{8d}$$

where $\mathcal{T}(j)$ denotes the set of nodes i such that $(i, j) \in \mathcal{A}$ and $\mathcal{S}(j)$ denotes the set of nodes i such that $(j, i) \in \mathcal{A}$. The problem (8) is convex and has a strictly convex objective function (Akamatsu, 1997), and the optimal solution to the problem defines uniquely the most likely O–D link flow solution. The relation between the most likely O–D link flow solution and the most likely route flow solution is made clear by the following theorem.

Theorem 2. *The most likely O–D link flows defined by the optimal solution to problem (8), are equal to the O–D link flows defined by* $\sum_{r\in\mathcal{R}_{pq}^*} \delta_{pqra} h_{pqr}^*$ *where* h_{pqr}^* *is the optimal solution to problem (6).*

Proof. See Appendix B. □

The model (8) can in certain situations be more favourable than (6). Such situations may arise when an equilibrium solution disaggregated by origin and destination (or just by origin or by destination) is sufficient. This is the case for the O–D matrix estimation and link toll usage analysis applications (cf. Section 2.2). The efficiency of the solution method for (8) sketched in Akamatsu (1997) is an open question.

Given the most likely O–D link flow solution from (8) it is possible, with a simple procedure, to obtain the most likely route flow solution. Such a procedure includes the enumeration of a restricted set of routes in the network and is discussed at the end of Appendix B.

Another approach is taken in Janson (1993) where a procedure for finding an approximate most likely O–D link flow solution is presented. The approach is based on the fact that different O–D link flow solutions are obtained when varying the initialization of the Frank–Wolfe method. A fixed number of user equilibria are obtained from different initiations in order to obtain a set of O–D link flow solutions. Each of the O–D link flow solutions is incorporated as a column in a restricted entropy problem. The solution to the restricted entropy problem is a set of O–D link flows which are more likely than any solution found by solving only one equilibrium problem. Computational results for a small scale network are presented. Note, however, that the most likely O–D link flows defined by the optimal solution to the model stated in Janson (1993) are *not* consistent with the most likely route flows obtained as the optimal solution to problem (6). This is pointed out in Akamatsu (1997).

4 Solution procedure for the entropy program

In this section we present a solution procedure for the problem of finding the most likely route flows, problem (6). We assume that the assignment problem (4) is solved and that the equilibrium link flow f^* is known. The procedure for finding the most likely route flows is a two phase procedure: (1) generate the equilibrium routes, $\mathcal{R}^*_{pq}$, from which an instance of the entropy model is constructed; and (2) solve the entropy program instance.

To generate the equilibrium routes, that is, to determine the sets $\mathcal{R}^*_{pq}$, we use the link flow solution obtained from the assignment problem and the equilibrium link costs $c_a(f^*_a)$. Construct sub-networks $\mathcal{G}^*_p = (\mathcal{N}, \mathcal{A}^*_p)$ of equilibrium links, one for each origin p, in the following way. Identify the set of equilibrium node potentials y^*_i, $i \in \mathcal{N}$ by solving a shortest route problem from source node p where the link costs are defined as $c_a(f^*_a)$. The set of equilibrium links $\mathcal{A}^*_p$ then consists of links a such that $y^*_{t(a)} - y^*_{h(a)} = c_a(f^*_a)$, where $t(a)$ and $h(a)$ denote the tail and head of arc a, respectively. The set of equilibrium routes $\mathcal{R}^*_{pq}$, for all O–D pairs $(p,q) \in \mathcal{C}$ is obtained by an enumeration of the routes in the sub-network $\mathcal{G}^*_p$. The route enumeration is made using a depth-first search procedure with root node p. From the sets of equilibrium routes and the set of equilibrium link flows we construct an instance of the entropy program.

In phase 2, we solve the entropy program. An overview of solution algorithms for entropy programs can be found in, for example, Lamond and Stewart (1981) and a method that work on the dual problem directly is found in, for example, Fang and Tsao (1994).

We present a solution approach for the entropy program (6) that works on the dual side of the problem. The approach is based on the dualization of the constraints (6c) in problem (6). Let $\beta \in \mathbb{R}^{|\mathcal{A}|}$ be the vector of dual variables for the link flow equality constraints (6c). After transforming problem (6) into a minimization problem we construct the dual objective function

$$\varphi(\beta) = \min \; L(h,\beta) \tag{9a}$$
$$\text{subject to (6b) and (6d),} \tag{9b}$$

with

$$L(h,\beta) = \sum_{(p,q)\in\mathcal{C}} \sum_{r\in\mathcal{R}^*_{pq}} h_{pqr} \ln h_{pqr} + \sum_{a\in\mathcal{A}} \beta_a \left(\sum_{(p,q)\in\mathcal{C}} \sum_{r\in\mathcal{R}^*_{pq}} h_{pqr}\delta_{pqra} - f^*_a \right).$$

The problem (9) separates in one problem for each pair $(p,q) \in \mathcal{C}$. We define

$$\varphi_{pq}(\beta) = \min \sum_{r\in\mathcal{R}^*_{pq}} \left[h_{pqr} \ln h_{pqr} + h_{pqr} \sum_{a\in\mathcal{A}} \delta_{pqra}\beta_a \right] \tag{10a}$$

subject to

$$\sum_{r\in\mathcal{R}^*_{pq}} h_{pqr} = d_{pq}, \tag{10b}$$

$$h_{pqr} \geq 0, \qquad \forall r \in \mathcal{R}^*_{pq}. \tag{10c}$$

This problem is explicitly solved by applying its Karush–Kuhn–Tucker conditions, which gives the solution

$$h_{pqr}(\beta) = d_{pq} \frac{e^{-\sum_{a \in \mathcal{A}} \beta_a \delta_{pqra}}}{\sum\limits_{r \in \mathcal{R}^*_{pq}} e^{-\sum_{a \in \mathcal{A}} \beta_a \delta_{pqra}}}.$$

This expression is recognized as a logit model for route choice (restricted to equilibrium links) where the link costs are β_a. The dual problem then is

$$\sup \varphi(\beta) = -\sum_{a \in \mathcal{A}} \beta_a f_a^* + \sum_{(p,q) \in \mathcal{C}} \varphi_{pq}(\beta). \tag{11}$$

The dual objective function φ is continuously differentiable and concave and the dual problem, thus, is an unconstrained differentiable, concave maximization problem. Under the assumption that problem (6) has a positive feasible solution, the supremum in problem (11) will be attained for some vector $\beta = \beta^*$ of Lagrange multipliers (see, e.g., Bazaraa et al., 1993, Theorem 5.3.1). Further, strong duality holds, and the optimal solution to problem (6) is

$$h^*_{pqr} = d_{pq} \frac{e^{-\sum_{a \in \mathcal{A}} \beta^*_a \delta_{pqra}}}{\sum\limits_{r \in \mathcal{R}^*_{pq}} e^{-\sum_{a \in \mathcal{A}} \beta^*_a \delta_{pqra}}}. \tag{12}$$

If a route flow is enforced to be zero in an optimal solution, this route can (in principle) be excluded from the set $\mathcal{R}^*_{pq}$ (see, e.g., Erlander and Stewart, 1990). Otherwise, the supremum in problem (11) will not be reached and some β^*_a will attain 'infinite values'.

The dual problem is solved using a conjugate gradient method. The gradient of φ at β is

$$\nabla \varphi(\beta) = \left[\sum_{(p,q) \in \mathcal{C}} \sum_{r \in \mathcal{R}^*_{pq}} \delta_{pqra} h_{pqr}(\beta) - f^*_a \right]_{a \in \mathcal{A}}$$

that is, the i:th component of $\nabla \varphi(\beta)$ is the difference between the link flow solving (10), calculated as a sum of the route flows, and the equilibrium link flow for link a. The norm of the gradient is used in the termination criterion for the algorithm. The route flows are, for numerical reasons, calculated as

$$h_{pqr}(\beta) = d_{pq} \frac{e^{-\sum_{a \in \mathcal{A}} \beta_a \delta_{pqra} - \gamma_{pq}}}{\sum\limits_{r \in \mathcal{R}^*_{pq}} e^{-\sum_{a \in \mathcal{A}} \beta_a \delta_{pqra} - \gamma_{pq}}},$$

where

$$\gamma_{pq} = \min_{r \in \mathcal{R}^*_{pq}} \sum_{a \in \mathcal{A}} \beta_a \delta_{pqra}.$$

Now, we introduce an iteration counter k. The search direction is computed in iteration $k = 0$ as $d^k = \nabla \varphi(\beta^k)$ and for $k > 0$ as

$$d^k = \nabla \varphi(\beta^k) + \frac{\nabla \varphi(\beta^k)^T \nabla \varphi(\beta^k)}{\nabla \varphi(\beta^{k-1})^T \nabla \varphi(\beta^{k-1})} d^{k-1}.$$

An approximate step length is computed based on a one dimensional Newton-step as

$$t^k = -\frac{\nabla \varphi(\beta^k)^T d}{{d^k}^T \nabla^2 \varphi(\beta^k) d^k},$$

where the Hessian is (see, e.g., Bertsekas, 1995, Chapter 6, Equation 1.11)

$$\begin{aligned} \nabla^2 \varphi(\beta) &= -\nabla g(h(\beta))^T \nabla^2_{hh} L(h, \beta)^{-1} \nabla g(h(\beta)) \\ &= -\nabla g(h(\beta))^T \operatorname{diag}[h_{pqr}(\beta^k)] \nabla g(h(\beta^k)) = -\Delta \operatorname{diag}[h_{pqr}(\beta^k)] \Delta^T, \end{aligned}$$

where g is the vector function of the left-and-sides in (6c) and $\Delta^T = (\delta_{pqra})$. A sufficient ascent is ensured by the use of an Armijo backtracking line search (with an acceptance parameter 0.2). The dual feasible solution $\beta = 0$ is used as a starting solution. The solution method is terminated when $||\nabla\varphi(\beta)|| \leq \varepsilon$ for some value of $\varepsilon > 0$.

5 Experimental results

Three test networks have been used for the evaluation of the solution method described in Section 4 and for the analysis of the properties of the most likely route flow solution. The networks studied are the networks of Sioux Falls (SD, USA), Winnipeg (Canada) and Linköping (Sweden), see Table 1 for their sizes. In the computational experiments, the route flow solutions obtained from the FW and the DSD methods are compared to the most likely route flow solution, the number of routes in each of the route flow solutions are analysed, and the topological differences between the route flow solutions are studied. The test results are presented separately for each of the three test networks.

Table 1: Traffic equilibrium networks.

Network	# Nodes	# Links	# O–D pairs
Sioux Falls	24	76	528
Winnipeg	1052	2836	4344
Linköping	335	882	12372

The network equilibrium problems are solved to a relative error between lower and upper bounds of the optimal objective value of 10^{-6} for the DSD algorithm and to $5 \cdot 10^{-5}$, $3 \cdot 10^{-5}$ and $7 \cdot 10^{-6}$ for the Sioux Falls, Winnipeg and Linköping networks, respectively for the FW method. The lower required accuracy for the FW algorithm is chosen since the method has a slower convergence rate. The number of routes generated by the DSD method is highly dependent of how accurate each of the master problems (see, Larsson and Patriksson, 1992) are solved. In our experiments, we have solved each of the master problems to a relative error between lower and upper bounds of the optimal objective value of 10^{-6}.

In order to implement the two-phase procedure, we introduce two tolerances: (1) a link is included in the subnetworks $\mathcal{A}_p^*$ of equilibrium links if the reduced cost for a link divided by the link cost is smaller than 0.01. This tolerance is crucial for the equilibrium route enumeration, if the tolerance is zero, only the shortest route generated in the last iteration of the equilibrium assignment method will be an equilibrium route. If however the tolerance is large, routes considered as equilibrium route may include non-equilibrium links, and this may lead to cycles in the sub-networks, thus making the enumeration impossible. (2) In the enumeration procedure, an enumerated route is considered to be an equilibrium route if the sum of reduced costs divided by the sum of link travel costs is smaller than 0.1.

Precautions have been made to ensure that problem (6) have a feasible solution; with the tolerances used, a small number of routes supplied by the DSD method are excluded from the resulting equilibrium sub-network, $\mathcal{A}_p^*$. These routes are however added to the set of equilibrium routes during the route enumeration procedure. The number of added routes are 0, 80, 93 for the evaluation networks of Sioux Falls, Winnipeg and Linköping, respectively. (This number is highly dependent of the solution accuracy required in the assignment solution.)

In the tables below a route is considered significant in the equilibrium solution if

$$h_{pqr} > \frac{d_{pq}}{100|\mathcal{R}_{pq}^*|},$$

that is, if the route flow is larger than one percent of the flow obtained if the demand were uniformly spread over all equilibrium routes in the O–D pair (p, q).

All problems are solved on a SUN Ultrasparc 170E with 192 MB memory.

5.1 The Sioux Falls network

For the Sioux Falls network the termination criterion for the entropy problem is set to $\varepsilon = 10^{-6}$. The computational time for the route enumeration and the solution of the entropy problem is 0.3 and 20 CPU seconds, respectively. In Table 2 the number of equilibrium routes, the number of equilibrium routes per O–D pair, and the entropy value for the route flow solution are presented for the DSD, FW, and the most likely route flow solution (the number in the parenthesis is the number of enumerated equilibrium routes).

Table 2: Solution data for the Sioux Falls network.

	Entropy value	# Routes		# Routes/O–D pair
DSD	42.17	718		1.37
FW	54.75	785		1.49
EP	54.80	762	(770)	1.45

The entropy objective function in problem (6) gives a measure of probability for a given route flow solution. The entropy values for the three route flow solutions indicate that the route flow solution generated by the FW method is close to the most likely route flow relative to the solution to the DSD method. Note however, that in the route flow solution from the FW method a number of non-equilibrium routes are included, and the entropy value may therefore be deceptive.

The entropy value obtained as the objective function value in problem (6) is an aggregated measure. In an individual O–D pair the most likely route flow need not have a higher entropy value than a route flow solution from another method. This is the case in O–D pair $(16, 8)$ were the route flow solution from the FW method have two routes with significant flow and the entropy value for this pair is -1.19 while the most likely route flow solution also has two routes with significant flow and the entropy value is -1.24. In the example in Section 2.2 two route flows for O–D pair $(14, 10)$ for the DSD and the FW method are presented. The entropy in this O–D pair is -0.146 (-0.583) for the route flow solution from the FW (DSD) method. The most likely route flow solution has 838 and 1262 vehicles for the routes 14–11–10 and 14–15–10, respectively. This route flow set has an entropy value of -0.145 (calculated in thousands of vehicles).

The maximal number of significant routes in any O–D pair is 2 and 8 in the DSD and most likely route flow solution, respectively. The largest difference between the number of significant O–D link flows on one link in the DSD and the most likely route flow solution is 7 and the number of O–D pairs with a significant O–D link flow on a link is on average 28.72 and 29.53 in DSD and the most likely route flow solutions, respectively. The small difference in the averaged numbers can be explained by the fact that a large number of the O–D pairs only use one route between the origin and the destination.

We also compute an aggregated measure of the difference between the route flow solutions. The analysis is restricted to the 72 O-D pairs where the DSD method has generated two or more routes in each pair (these pairs provide 21% of the total travel demand). When comparing the most likely route flow solution with the route flow solution from the DSD method for these O–D pairs, on average, 82% of the O–D demand use identical routes. The corresponding number for the most likely route flow and the solution to the FW method is 97%.

5.2 The Winnipeg network

For the Winnipeg network the termination criterion for the entropy program is set to $\varepsilon = 1.0$. This termination criterion implies that the left-hand-side in (6c) have a relative difference less than 0.2% from the equilibrium link flows when considering equilibrium link flows with a flow larger then 10 units (less than 10% of the links have a flow smaller than 10 units). The computational time for the route enumeration and the solution of the entropy problem is 125 and 638 CPU seconds, respectively. The entropy values and route data are given in Table 3 (the number in parenthesis is the number of enumerated equilibrium routes).

Table 3: Solution data for the Winnipeg network.

	Entropy value	# Routes		# Routes/O–D pair
DSD	−178355	8957		2.06
FW	−168212	16494		3.80
EP	−166648	18111	(22338)	4.17

Next, we compare the topological differences between the most likely route flow and the route flow solution from the DSD method for one O–D pair with a relatively large travel demand. The set of equilibrium links in O–D pair number $(1, 15)$ in the Winnipeg network is given in Figure 6 (the links in the network are aggregates of sequences of links in the original network). The route enumeration procedure gives that 110 routes in this network are equilibrium routes. The DSD method gives a route flow solution in which 6 equilibrium routes have a significant flow. These equilibrium routes can be found in the network in Figure 7 and in Table 4 we list the routes and the route flows from the DSD method and the corresponding route flows in the most likely route flow solution. Comparing the two route flow solutions we observe that there is a large difference in the route flows for the two solutions in this O–D pair. We note that in the most likely route flow solution, only about 60% of the O–D demand use the routes that carry all the flow in the DSD route flow solution. The routes in the most likely route flow solution

Table 4: Two set of equilibrium routes in the order they are generated in the DSD scheme.

Equilibrium DSD routes	Flow, DSD	Flow, EP
1 – 2 – 3 – 6 – 12 – 13 – 14 – 15	33.52	5.52
1 – 2 – 4 – 5 – 7 – 9 – 11 – 10 – 14 – 15	0.98	0.03
1 – 3 – 6 – 9 – 12 – 13 – 14 – 15	0.67	1.95
1 – 3 – 6 – 12 – 13 – 14 – 15	1.68	13.77
1 – 2 – 3 – 4 – 5 – 8 – 10 – 14 – 15	2.86	0.19
1 – 3 – 4 – 5 – 7 – 9 – 11 – 13 – 14 – 15	0.29	1.93
Sum	40.00	23.40

are depicted in Figure 8. The most likely route flow solution uses 51 significant routes.

An aggregated measure of the difference between the route flow solutions is given where we restrict the analysis to the 270 largest O–D pairs (these pairs provide 21% of the total travel demand) where the DSD method have generated two or more routes in each pair. Comparing the most likely route flow solution to the route flow solution from the DSD method, we have that, on average, 75% of the O–D demand in these O–D pairs use identical routes. The corresponding number for the most likely route flow compared with the solution computed with the FW method is 94%.

The largest number of routes used in one O–D pair is 9 and 396 in the DSD and most likely route flow solutions, respectively. The largest difference between the number of significant O–D link flows on one link in the DSD and the most likely route flow solution is 113. The number of O–D pairs with a significant O–D link flow on a link is on average 47.61 and 47.78 in the DSD and the most likely route flow solutions, respectively. Also for the Winnipeg network we conclude that the small difference in the averaged numbers can be explained by the fact that a large number of the O–D pairs have only one route between the origin and the destination.

5.3 The Linköping network

For the Linköping network the termination criterion for the entropy program is set to $\varepsilon = 2.0$. The computational time for the route enumeration and the solution of the entropy problem is 64 and 970 CPU seconds, respectively. The number of equilibrium routes with significant route flow, the number of routes per O–D pair and the entropy value are given in Table 5. Although

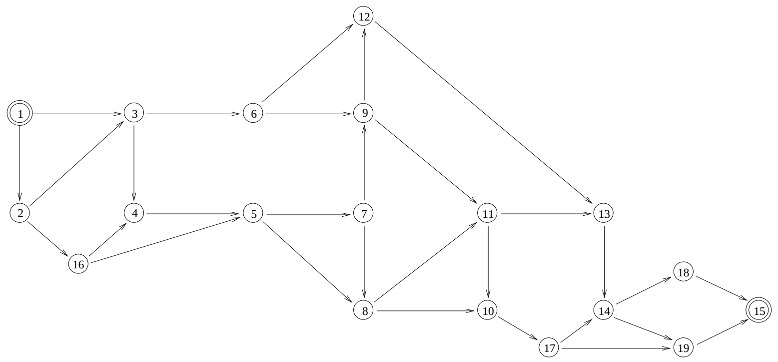

Figure 6: Network in which equilibrium routes are enumerated

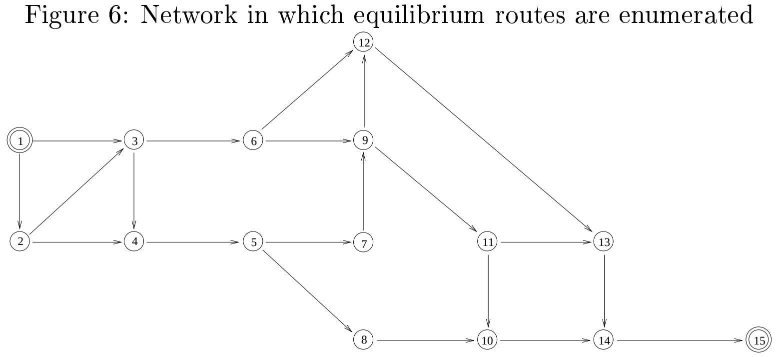

Figure 7: Network of generated routes in DSD solution.

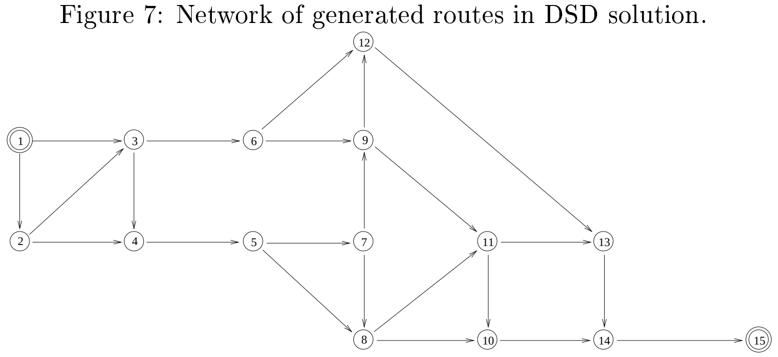

Figure 8: Network of significant equilibrium routes in max entropy solution.

the network is, in some areas, highly congested the number of routes per O–D pair indicates that the number of possible route alternatives in the network is low.

We restrict the analysis of the difference of the route flow solutions to the 382 largest O-D pairs where the DSD method have generated two or more routes in each pair (these pairs provide 13% of the total travel demand). When comparing the most likely route flow solution with the route flow solution from the DSD method for these O–D pairs we have that, on average, 81% of the O–D demand use identical routes. For the same O–D pairs and the route flow solution from the FW method and the most likely route flow solution we found that 97% of the demand travel on identical routes.

The maximal number of used routes in any O–D pair is 5 and 19, respectively, in the DSD and the most likely route flow solutions. The largest difference in the number of significant O–D link flows on one link is 165 between the DSD and the most likely route flow solutions. The number of O–D pairs with a significant O–D link flow on a link is 238.6 and 241.2 in the DSD and the max entropy solutions, respectively, on average.

From the numerical experiments we conclude that the solution procedure can be used to obtain the most likely route flow solution in reasonable computing time for medium sized networks. The difference between the route flow solution from the DSD method and the most

Table 5: Solution data for the Linköping network.

	Entropy value	# Routes		# Routes/O–D pair
DSD	-757620	18672		1.51
FW	-743075	21193		1.71
EP	-743069	21447	(21894)	1.73

likely route flow solution is significant in certain O–D pairs of our test networks. The FW method produces, for our test networks, a route flow solution very similar to the most likely route flows in most O–D pairs.

With respect to this conclusion, we must however also remark that in our experiments, the instance of the DSD algorithm used—given by (among others) certain parameter choices governing the termination of each restricted master problem—was one which minimizes the total CPU time; from the experiments conducted in Larsson and Patriksson (1992) with the method, we then concluded that a minimal CPU time most often means that each restricted master problem should be solved to a relatively high accuracy in every iteration, thus, in effect, minimizing the total number of iterations. The effect of such a choice is of course that of a small number of shortest route calculations. Its consequences here (as observed) are that the number of equilibrium routes generated are among the smallest possible, and that the entropy of the route flow solution obtained is rather small. It thus seems that there is a conflict between the goal of high efficiency in computing an equilibrium route flow solution and the generation of a probable equilibrium route flow.

For large networks, especially those with a grid structure, there is a possibility that the number of routes found in the path enumeration procedure becomes very large. To avoid this and still use the approach presented the networks need to be aggregated. The establishment of a proper principle for aggregating the network is a subject for future studies.

6 An application: Exhaust fume emission analysis

In this section an exhaust fume emission analysis is made for the Linköping network. In the experiments made we analyze the spatial distribution on the links of cold start dependent excess from vehicles starting with cold engines and catalytic converters. The data for vehicle emissions, the travel demand and the link delay functions used in the analysis was provided by the Swedish National Road and Transport Research Institute (VTI). Figure 9 shows the central part of Linköping, which is the set of links studied in detail. The average exhaust fume

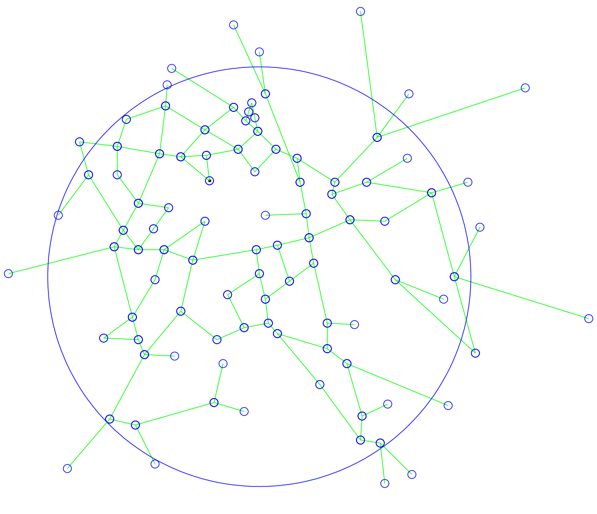

Figure 9: The central part of Linköping.

emission of nitrogen oxides (NO_x) arising from a cold engine (in contrast to that from a warm engine) is assumed to be described by an exponential function of the travelled distance only, see Figure 10.

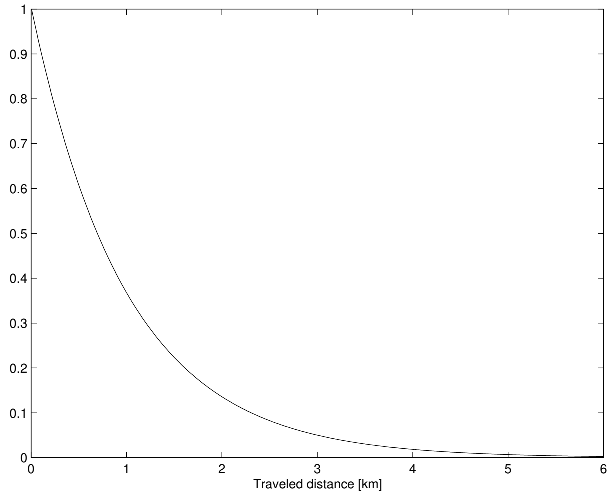

Figure 10: The NO_x emission distribution.

In Figure 11 the emission distribution of the links in the central part of Linköping is shown. The emission distribution is computed based on the most likely route flow solution. From this figure we may observe that some links have a very high level of NO_x compared to other links. Experiments where the NO_x emission calculation is based on the route flow output from

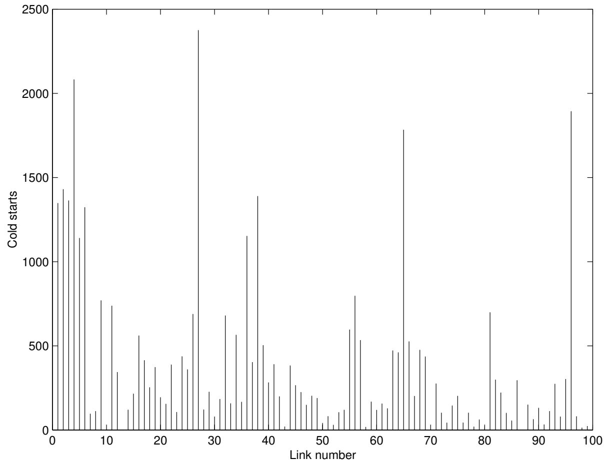

Figure 11: NO_x emission on the links in the central part of Linköping according to the most likely route flow solution.

the DSD method have shown that the difference to the emission obtained by the most likely route flow is very small. (Significant differences are found only on links where the emission are relatively small compared to other links.) Due to the structure of the network the results do not differ very much. However, in the analysis made, there are differences in the resulting link emission and these differences are likely to be larger in a network where the traveller has a larger number of alternative routes to choose from, although lack of sufficient data prevent us from making analyses of other city centres.

References

Akamatsu, T. Tsuchiya, Y. and Shimazaki, T., "Parallel distributed processing on neural network for some transportation equilibrium assignment problems," *In Proceedings of the 11th International Symposium on the Theory of Traffic Flow and Transportation*, Ed: Koshi, M., 369–386, Yokohama, Elsevier, NY, (1990)

Akamatsu, T., "Cyclic flows, Markov process and stochastic traffic assignment," *Transportation Research* 30, 369–386 (1996)

Akamatsu, T., "Decomposition of path choice entropy in general transport networks," *Transportation Science* 31, 349–362 (1997)

Barton, R.R. Hearn, D.W. and Lawphongpanich, S., "The equivalence of transfer and generalized Benders decomposition methods for traffic assignment," *Transportation Research* 23B, 61–73 (1989)

Bazaraa, M.S. Sherali, H.D. and Shetty, C.M., Nonlinear Programming: Theory and Algorithms John Wiley & Sons, New York, NY, 1993

Bell, M.G.H. and Iida, Y., Transportation Network Analysis. John Wiley & Sons, Chichester, England, 1997

Bertsekas, D.P., Nonlinear Programming. Athena Scientific, Belmont, MA., 1995

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

Drive project v1054, 1st. Deliverable, Report on model requirements, 1989

Erlander, S. and Stewart, N.F., The Gravity Model in Transportation Analysis. VSP BV, Utrecht, The Netherlands., 1990

Edwards, H., "Estimation of excess cold start emissions on links in traffic networks," (In Swedish) Working paper, VTI, S-581 95 Linköping, Sweden, 1998

Fang, S.C. and Tsao, H.S.J., "A quadratic convergent global algorithm for the linearly-constrained minimum cross-entropy problem," *European Journal of Operations Research* 79, 369–378 (1994)

Fisk, C.S., "Some developments in equilibrium traffic assignment," *Transportation Research* 14B, 243–255 (1980)

Fisk, C.S., "On combing maximum entropy trip matrix estimation with user optimal assignment," *Transportation Research* 22B, 69–79 (1988)

Florian, M. and Nguyen, S., "An application and validation of equilibrium trip assignment methods," *Transportation Science* 10, 374–390 (1976)

Hearn, D.W., "Practical and theoretical aspects of aggregation problems in transportation models," *In Transportation Planning Models, Proceedings of the course given at the International Center for Transportation Studies, Amalfi, Italy, October 11–16, 1982*, Ed: Florian, M., 257–287, North-Holland, Amsterdam (1984)

Janson, B.N., "Most likely origin–destination link uses from equilibrium assignment," *Transportation Research* 27B, 333–350 (1993)

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

Larsson, T. and Patriksson, M., "Simplicial decomposition with disaggregate representation for the traffic assignment problem," *Transportation Science* 26, 4–17 (1992)

LeBlanc, L.J. Morlok, E.K. and Pierskalla, W.P., "An efficient approach to solving the road network equilibrium traffic assignment problem," *Transportation Research* 9, 309–318 (1975)

Patriksson, M., The Traffic Assignment Problem—Models and Methods. VSP BV, Utrecht, The

Netherlands, 1994

Polyak, B.T., Introduction to Optimization. Optimization Software, Inc., Publications Division, N.Y., 1987

Rossi, T.F. McNeil S. and Hendrickson C., "Entropy model for consistent impact fee assessment," *Journal of Urban Planning and Development/ASCE* 115, 51–63 (1989)

Sérié, E. and Joumard, R., "Modelling of cold start emissions for road vehicles," INRETS report LEN 9731, 1997

Small, K.A., "Using revenues from congestion pricing," *Transportation* 19, 359–381 (1992)

Smith, T.E., "A cost-efficiency of spatial interaction behaviour," *Regional Science and Urban Economics* 8, 137–168 (1978)

Smith, T.E., "A cost-efficiency approach to the analysis of congested spatial-interaction behavior," *Environment and Planning* 15A, 435–464 (1983)

Snickars, F. and Weibull, J.W., "A minimum information principle: Theory and practice," *Regional Science and Urban Economics* 7, 137–168 (1977)

Sheffi, Y., Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods. Prentice–Hall, Englewood Cliffs, NJ, 1985

The Economist, December 6, 21–24 (1997)

Van Vliet, D., "Selected node-pair analysis in Dial's assignment algorithm," *Transportation Research* 15B,, 65–68 (1981)

Appendix A: Relation between the stochastic user equilibrium and the most likely route flows

Proof. (Theorem 1) The proof is stated for two arbitrary functions, f and g, and is based on the proof in Polyak (1987, Chapter 6.1, Theorem 4), modified to the case where g is strictly convex.

Assume that $f : \mathbb{R}^n \to \mathbb{R}$ is convex and $g : \mathbb{R}^n \to \mathbb{R}$ is strictly convex, and let $X_f^* = \text{argmin}_{x \in X} f(x)$, where $X \subset \mathbb{R}^n$ is a nonempty, closed and bounded set. Define the trajectory $\{x_\varepsilon | \varepsilon > 0\}$ by $x_\varepsilon = \text{argmin}_{x \in X}\{f(x) + \varepsilon g(x)\}$ and let $\{x^*\} = \text{argmin}_{x \in X_f^*} g(x)$. To prove the theorem we show that $x_\varepsilon \to x^*$ when $\varepsilon \to 0^+$.

For any fixed $\varepsilon > 0$, $f + \varepsilon g$ is strictly convex. Since X is compact and convex, x_ε exists and is unique, and the trajectory $\{x_\varepsilon | \varepsilon > 0\}$ is well-defined and bounded.

From the definition of x_ε, and since $x^* \in X$ we have that

$$f(x_\varepsilon) + \varepsilon g(x_\varepsilon) \leq f(x^*) + \varepsilon g(x^*). \tag{13}$$

Further, since $x^* \in X_f^*$ and $x_\varepsilon \in X$ we also have the relation $f(x^*) \leq f(x_\varepsilon)$. From these two inequalities, $g(x_\varepsilon) \leq g(x^*)$, follows.

Extract a subsequence $\{x_{\varepsilon_i}\}$, $\varepsilon_i \downarrow 0$, converging to a point $\tilde{x} \in X$. Since g is continuous, $\lim_{i \to \infty} g(x_{\varepsilon_i}) = g(\tilde{x})$ holds, which gives that,

$$g(\tilde{x}) \leq g(x^*). \tag{14}$$

Taking the limit corresponding to the subsequence $\{x_{\varepsilon_i}\}$ in (13) gives that $f(\tilde{x}) \leq f(x^*)$, and therefore $\tilde{x} \in X_f^*$. By relation (14), the definition of x^*, and the strict convexity of g, we then obtain that $\tilde{x} = x^*$. It follows that $x_{\varepsilon_i} \to x^*$.

Since the subsequence was arbitrary, $\{x_\varepsilon\} \to x^*$ follows.

Now, identify x with the link flow variables $f_a(h)$, f with $\sum_{a \in \mathcal{A}} \int_0^{f_a(h)} t_a(s) ds$ and g with $\sum_{(p,q) \in \mathcal{C}} \sum_{r \in \mathcal{R}_{pq}} h_{pqr} \ln h_{pqr}$ to complete the proof. □

Appendix B: Relation between the models for finding the most likely O–D link flows and the most likely route flows

Proof. (Theorem 2) The proof is made in two steps. In the first we show that problems (6) and (8) have the same feasible set with respect to O–D link flows, that is there can not exist a set of O–D link flows that is feasible in problem (8) and that has a positive O–D link flow on a non-equilibrium route. In the second step we show that the optimal solutions to the two problems (6) and (8) coincide in terms of the O–D link flows.

We show that the link flow in O–D pair (p,q) on link $a \in \mathcal{A}$ is zero whenever link a is not included in an equilibrium route from p to q. The Wardrop conditions can be stated as (e.g., Patriksson, 1992, Theorem 3.14a)

$$\bar{c}(h^*)^T(h - h^*) \geq 0, \quad \forall h \in \mathcal{H}^r,$$

where $\bar{c}$ denotes the vector of route costs. Let $\bar{h} \in \{h \in \mathcal{H}^r \,|\, \sum_{(p,q)\in\mathcal{C}} \sum_{r\in\mathcal{R}_{pq}} \delta_{pqra} h_{pqr} = f_a^*\}$. From the additivity assumption we have that

$$\bar{c}(h^*)^T(\bar{h} - h^*) = \sum_{a\in\mathcal{A}} c_a(f_a^*) \left(\sum_{(p,q)\in\mathcal{C}} \sum_{r\in\mathcal{R}_{pq}} \delta_{pqra}\bar{h}_{pqr} - f_a^* \right) = c(f^*)^T(f^* - f^*) = 0.$$

Hence,

$$\bar{c}(h^*)^T(h - \bar{h}) = \bar{c}(h^*)^T(h - h^*) - \bar{c}(h^*)^T(\bar{h} - h^*) = \bar{c}(h^*)^T(h - h^*) \geq 0, \quad \forall h \in \mathcal{H}^r.$$

This implies that $\bar{h} \in (\mathcal{H}^r)^*$, that is, if $c_{pqr} > \pi_{pq}$ then $\bar{h} = 0$. We conclude that a link not included in any equilibrium route from p to q has a zero O–D link flow in pair (p,q), and that the feasible set of problems (6) and (8) therefore coincide with respect to O–D link flows.

Problem (8) can thus be equivalently stated as

$$\min \sum_{(p,q)\in\mathcal{C}} \left(\sum_{a\in\mathcal{A}} f_{pqa} \ln f_{pqa} - \sum_{j\in\mathcal{N}} \left(\sum_{i\in\mathcal{T}'(j)} f_{pq(i,j)} \right) \ln \left(\sum_{i\in\mathcal{T}'(j)} f_{pq(i,j)} \right) \right) \tag{15a}$$

subject to

$$\sum_{(p,q)\in\mathcal{C}} f_{pqa} = f_a^*, \qquad \forall a \in \mathcal{A}, \tag{15b}$$

$$\sum_{i\in\mathcal{T}'(j)} f_{pq(i,j)} - \sum_{i\in\mathcal{S}'(j)} f_{pq(j,i)} = \begin{cases} d_{pq}, & \text{if } j = q, \\ -d_{pq}, & \text{if } j = p, \\ 0, & \text{otherwise,} \end{cases} \qquad j \in \mathcal{N}, \quad (p,q) \in \mathcal{C}, \tag{15c}$$

$$f_{pqa} \geq 0, \qquad \forall a \in \mathcal{A}, \quad \forall (p,q) \in \mathcal{C}. \tag{15d}$$

where $\mathcal{T}'(j)$ denote nodes i such that $(i,j) \in \mathcal{A}$ and is included in an equilibrium route from (p,q) and $\mathcal{S}'(j)$ denote nodes i such that $(j,i) \in \mathcal{A}$ and is included in an equilibrium route from (p,q).

In step two we show that the optimal solutions to problems (6) and (15) coincide in terms of the O–D link flows. One O–D link flow solution, feasible in problem (15), can be calculated from the most likely route flows, h^*, as

$$f_{pqa}(h^*_{pqr}) = \sum_{r\in\mathcal{R}^*_{pq}} \delta_{pqra} h^*_{pqr}, \qquad \forall (p,q) \in \mathcal{C}, \quad \forall a \in \mathcal{A},$$

where

$$h^*_{pqr} = d_{pq} \frac{e^{-\sum_{a\in\mathcal{A}} \beta_a^* \delta_{pqra}}}{\sum_{r\in\mathcal{R}^*_{pq}} e^{-\sum_{a\in\mathcal{A}} \beta_a^* \delta_{pqra}}}, \qquad \forall r \in \mathcal{R}^*_{pq}, \quad \forall (p,q) \in \mathcal{C} \tag{16}$$

with β^* being the vector of Lagrange multipliers for constraints (6c). Consider an arbitrary node j, and define $\mathcal{R}^j_{pq}$ as the set of routes in $\mathcal{R}^*_{pq}$ that pass node j. Any route in $\mathcal{R}^j_{pq}$ passing node $i \in \mathcal{T}'(j)$ can be partitioned into three parts, from p to i, from i to j, and from j to q. Hence,

$$\mathcal{R}^j_{pq} = \mathcal{R}^j_{pi} \bigcup \left(\bigcup_{i \in \mathcal{T}'(j)} (i,j) \right) \bigcup \mathcal{R}^j_{jq}.$$

Define k as the denominator in expression (16) and define

$$k_j = \sum_{r \in \mathcal{R}^j_{pq}} e^{-\beta_r} = \sum_{i \in \mathcal{T}'(j)} \sum_{r_1 \in \mathcal{R}^j_{pi}} \sum_{r_2 \in \mathcal{R}^j_{jq}} e^{-\beta_{r_1}} e^{-\beta_{ij}} e^{-\beta_{r_2}} =$$
$$\left(\sum_{r \in \mathcal{R}^j_{jq}} e^{-\beta_r} \right) \sum_{i \in \mathcal{T}'(j)} \left(e^{-\beta_{ij}} \sum_{r \in \mathcal{R}^j_{pi}} e^{-\beta_r} \right),$$

where $\beta_r = \sum_{a \in r} \beta_a$. The last equality in (17) is adopted from Van Vliet (1981). Define $\mathcal{A}(j) = \{(i,j) | i \in \mathcal{T}(j)\}$ and $\mathcal{A}'(j) = \{(i,j) | i \in \mathcal{T}'(j)\}$. The total O–D flow in pair (p,q) into node j then is

$$\sum_{a \in \mathcal{A}'(j)} f_{pqa}(h^*) = \sum_{a \in \mathcal{A}'(j)} \sum_{r \in \mathcal{R}^j_{pq}} \delta_{pqra} h^*_{pqr} = \frac{k_j}{k} d_{pq}.$$

Choose an arbitrary node $\bar{i} \in \mathcal{T}'(j)$. The flow in O–D pair (p,q) on link $\bar{a} = (\bar{i}, j)$ is,

$$f_{pq\bar{a}}(h^*) = \sum_{r \in \mathcal{R}^j_{p\bar{i}} \times \{(\bar{i},j)\} \times \mathcal{R}^j_{jq}} \delta_{pqr\bar{a}} h^*_{pqr} = \frac{d_{pq}}{k} \left[\left(\sum_{r \in \mathcal{R}^j_{p\bar{i}}} e^{-\beta_r} \right) e^{-\beta_{\bar{i}j}} \left(\sum_{r \in \mathcal{R}^j_{jq}} e^{-\beta_r} \right) \right].$$

Introducing the notation $\mu_i = -\ln \left(\sum_{r \in \mathcal{R}^j_{pi}} e^{-\beta_r} \right)$, we now have that

$$\frac{f_{pq\bar{a}}(h^*)}{\sum_{a \in \mathcal{A}'(j)} f_{pqa}(h^*)} = \frac{\left(\sum_{r \in \mathcal{R}^j_{p\bar{i}}} e^{-\beta_r} \right) e^{-\beta_{\bar{i}j}}}{\sum_{i \in \mathcal{T}'(j)} \left(e^{-\beta_{ij}} \sum_{r \in \mathcal{R}^j_{pi}} e^{-\beta_r} \right)} = \frac{e^{-\mu_{\bar{i}}} e^{-\beta_{\bar{i}j}}}{\sum_{i \in \mathcal{T}'(j)} e^{-\beta_{ij}} e^{-\mu_i}}. \tag{17}$$

By observing that

$$1 = \sum_{l \in \mathcal{A}'(j)} \frac{f_{pql}}{\sum_{a \in \mathcal{A}'(j)} f_{pqa}} = \frac{1}{e^{-\mu_j}} \sum_{i \in \mathcal{T}'(j)} \frac{e^{-\beta_{ij}}}{e^{-\mu_i}},$$

and by taking the logarithm of expression (17) we have that,

$$\ln f_{pq\bar{a}} - \ln \left(\sum_{a \in \mathcal{A}'(j)} f_{pqa} \right) + \beta_{\bar{i}j} + \mu_{\bar{i}} - \mu_j = 0$$

which holds for all $j \in \mathcal{N}$ and all $\bar{i} \in \mathcal{T}'(j)$.

This expression is however exactly the Karush–Kuhn–Tucker conditions for the problem (15) with $\beta \in \mathbb{R}^{|\mathcal{A}|}$ and $\mu \in \mathbb{R}^{|\mathcal{N}|}$ interpreted as the Lagrange multipliers for the constraint sets (15b) and (15c).

Since the O–D link flows $f_{pqa}(h^*_{pqr}) = \sum_{r \in \mathcal{R}^*_{pq}} \delta_{pqra} h^*_{pqr}$, for all $(p,q) \in \mathcal{C}$, and $a \in \mathcal{A}$ is feasible in problem (15), and fulfill the Karush–Kuhn–Tucker conditions for this problem, we conclude that they are optimal. □

We conclude this appendix by describing a procedure for finding the most likely route flows given the most likely O–D link flows. The procedure is similar to a backward pass in the Dial assignment method. Given the subnetwork of equilibrium links for origin p, denoted $\mathcal{G}_p^*(\mathcal{N}, \mathcal{A}_p^*)$ (see Section 4), perform an enumeration of the routes between p and q, starting with node q. In each intermediate node calculate the route choice portions as

$$\frac{f_{pqa}^*}{\sum\limits_{a \in \mathcal{A}_p^*} f_{pqa}^*},$$

where f_{pqa}^* is the optimal solution to problem (8). The most likely route flow on a specific route is then obtained as the product of these quantities along the route.