Asymmetric Multiclass Traffic Assignment: a coherent formulation

by Ph.L. Toint[1] and L. Wynter[2]

Report 95/13 October 25, 1995

[1] Department of Mathematics, Facultés Universitaires ND de la Paix
61, rue de Bruxelles, B-5000 Namur, Belgium, EU
Email : pht@math.fundp.ac.be
Current reports available by anonymous ftp from the directory
"pub/reports" on thales.math.fundp.ac.be (internet 138.48.4.14)

[2] Department of Transport Economy and Sociology, INRETS,
2, avenue du Général Malleret-Joinville, F-94114 Arceuil, France, EU
Email: laura@inrets.fr

Keywords: Asymmetric traffic assignment, multi-user class model, traffic equilibrium, multiclass cost functions.

Asymmetric Multiclass Traffic Assignment: A coherent formulation

Philippe Toint
Transportation Research Group,
FUNDP, Namur, Belgium

Laura Wynter
DEST, INRETS
Arceuil, France

October 25, 1995

Abstract

Behavioural and mathematical problems in the current formulation of the asymmetric multiclass traffic assignment model are brought to light. A new formulation is presented that resolves these difficulties. The formulation is quantified, and the result is a general class of coherent, multiclass cost functions. A numerical calibration exercise for a car and truck class model concludes the paper.

1 Introduction

The asymmetric multiclass static traffic assignment problem has witnessed a growing amount of research attention recently as the interest in advanced congestion management programs increases. Modeling reserved lanes, road pricing schemes, and electronic driver information systems all require a segmentation of traffic demand. Still, relatively few usable multiclass assignment models have been put into practice. In this paper, we will demonstrate that behavioural and mathematical incoherence in the current formulation of the problem precludes estimation of realistic cost functions based on travel times By defining a set of constraints between the different user classes, we are able to present a general class of coherent, multiclass cost functions. This methodology is then put to practice in the definition of coherent multiple value-of-time class and car-and-truck class models, and in the calibration of travel time functions for an interurban car-and-truck-class model.

The paper is organized as follows. Section 2 presents the problem and defines some notation. The incoherence of the classical formulation in terms of behaviour is exposed in Section 3, while the mathematical problems are covered in Section 4. Section 5 contains the suggested reformulation and analyzes existence and uniqueness of the solution. Two cases studies are presented in Section 6. Finally, some conclusions and perspectives are outlined in Section 7.

2 Background and notation

The classical static traffic equilibrium has proven to be a useful framework for analyzing transportation systems (see, for instance, Sheffi [25]). In this model, a single class of users is assigned routes in a network in order to fulfill a given travel demand, at the same time minimizing the cost associated with the route. This is the well-known Wardropian equilibrium [29]. However, the fact that there is only one class of users clearly introduces some limitations, for instance if one wishes to model urban networks, multi-modal networks or multiple vehicle types. These limitations can be overcome by introducing the multiclass problem: the demand is now partitioned into several classes and the cost of a route for a given class then depends on the traffic of all users' classes along the route. We are concerned, in this paper, with the behavioural and mathematical coherence of this type of model.

At this point, and in order to make the notion of trip cost more precise, it is important to make the distinction between traffic flow and density, the former being the number of vehicles passing a given point per time period, and the latter describing the number of vehicles present in a given space. As we are concerned with static, or single-time-period, traffic assignment models, we will refer herein to traffic density, rather than flow. This eliminates the need to deal with multivalued functions, as the true flow vs. travel time relation is bivalued. The density vs. travel time relation is, however, univalued, and monotonic in one variable.

We will, in this paper, consider a network with N arcs and K user classes, the cost of using arc i for a traveller of class k being defined as a given function of the travel time on this arc for this class, that is

$$c_{ik} = c_{ik}(t_{ik}(x))$$

where $x = (x_{i1}, x_{i2}, \ldots, x_{iK})$ is the vector of (nonnegative) densities on arc i for each of the K classes, and where $t_{ik} = t_{ik}(x)$ is the travel time of arc i for class k. Note that the cost functions c_{ik} may also depend on other class- or arc-specific quantities, such as toll price for the arc or value of time for the class, but that the dependence on x is uniquely via the travel time functions $t_{ik}(\cdot)$. The $c_{ik}(\cdot)$ are assumed to be continuous and increasing with respect to t_{ik} and the $t_{ik}(\cdot)$ themselves are assumed to be continuous and increasing with respect to all components of x. The multiclass model is referred to as an asymmetric traffic assignment when the influences of the different classes are not reciprocal, that is when

$$\frac{\partial c_{ik}}{\partial x_{jk}} \neq \frac{\partial c_{jk}}{\partial x_{ik}} \quad (i \neq j).$$

Behaviourally, this means that adding one more truck to an arc does not have the same effect on a car that adding another car has on a truck. In terms of priority junctions, asymmetry implies that an additional vehicle on a four-lane road does not have the same effect on the cost on a single-lane road merging into it as an additional vehicle on the small road has on the larger one. As this is often the case, the symmetry of the problem is usually lost. This more general problem is thus referred to as the *asymmetric traffic assignment problem.*

The Wardrop user equilibrium condition, as extended by Dafermos [8, 9] and Potts and Oliver

[24] to handle the multiclass traffic assignment problem is restated by Fernandez and Friesz [12] (see also Van Vliet, et al. [28] and Braess and Koch [3]) as follows :

Definition 2.1 (User Equilibrium Condition) *The routes between any origin-destination pair can be grouped into two subsets for each class, one for paths with non-zero flow and equal average (unit) travel cost, and the other for paths without flow and travel costs greater than or equal to the cost on the utilized paths.*[1]

In effect, this condition is the original Wardrop first principle applied within each class, where the term class is used in the general sense of the word. (For example, on a two-way road model, a class may be a direction of flow on the road.) We will show in this paper that such a definition alone is insufficient for ensuring a coherent multiclass traffic assignment.

In the single-class model, Definition 2.1 ensures that, at equilibrium, all travel costs are coherent since they are either at the minimum or their corresponding flows are null. However, in the multiclass assignment, this is not the case. At equilibrium, the present definition requires only that the travel costs *within* each class are equilibrated or flows are null. No constraints between the travel cost functions across classes are included in the formulation.

We begin by presenting several examples which illustrate the inconsistencies that result from solving a multiclass assignment based solely on the current problem formulation.

3 Behavioural incoherence

3.1 Example 1: Multiclass congestion

Consider two traffic classes sharing an arc but having different physical (or behavioural) characteristics. We can model both classes jointly through a multiclass assignment. In this case, we define the following generalized BPR-type travel time functions [4] for two classes[2]:

$$t_{i1}(x_i) = 3\left[1 + 0.15\left(\frac{x_{i1} + 2x_{i2}}{500}\right)^4\right] \tag{1}$$

$$t_{i2}(x_i) = 4\left[1 + 0.25\left(\frac{x_{i1} + 2x_{i2}}{500}\right)^4\right], \tag{2}$$

where t_{ik} is the travel time and x_{ik} the density of class k on arc i; the denominator of the fraction raised to the fourth power (500) represents the density in passenger car equivalent units (PCU) at which point the system performance degrades sharply (c.f. practical capacity).

[1] Potts and Oliver [24] write: For every commodity (class) k and an equilibrium flow (density) f, there exits an ordering $1, 2, \ldots p, p+1, \ldots \ell$ of the paths from O_k to D_k such that $c_{1k}(f) = c_{2k}(f) = \ldots = c_{pk}(f) \leq c_{p+1,k}(f) \ldots \leq c_{\ell k}(f)$, where the flows (densities) of commodity (class) k on path i, denoted f_{ik}, are such that $f_{ik} > 0$ for $i = 1, \ldots p$, and $f_{jk} = 0$ for $j = p+1, \ldots \ell$.

[2] The reader should note that, while the original BPR function [4] was calibrated to represent the uncongested portion of the flow vs. travel time relation, it is often applied to the congested portion as well. However, the BPR function is single-valued and increasing for positive flows, and as such more accurately represents the density vs. travel time relation when considering both congested and uncongested traffic.

Now, let us examine what occurs at congestion. Consider a traffic density of

$$x_1 = 1000 \qquad x_2 = 250$$

on a particular arc. Note that this level represents three times the maximum (practical) density of the link. Then, the ratio of the travel time of class 1 to that of class 2 at congestion is the following:

$$\frac{t_{i1}(x_i)}{t_{i2}(x_i)} = \frac{3\left[1 + 0.15\left(\frac{x_{i1}+2x_{i2}}{500}\right)^4\right]}{4\left[1 + 0.25\left(\frac{x_{i1}+2x_{i2}}{500}\right)^4\right]} = 0.46 \tag{3}$$

This means that, even in high congestion, class 1 can traverse the arc in 0.46 the time needed by class 2 to traverse the same arc during the same time period! This is clearly a non-physical result. In fact, in this example, $t_1/t_2 \to 0.45$ as $x \to \infty$. Clearly we would not expect this to be so.[3]

Now, let us consider the following quadratic travel time functions taken from [9].

$$\begin{aligned} t_{i1}(x_i) &= 2x_{i1}^2 + x_{i1}x_{i2} + 1500x_{i1} \\ t_{i2}(x_i) &= 4x_{i2}^2 + x_{i1}x_{i2} + 1700x_{i2} \end{aligned} \tag{4}$$

where, as before, t_{ik} is the travel time and x_{ik} the density of class k on arc i. We note that $t_{ij}(0) = 0$ in the above four travel time functions. We may suppose that $t_{ij} = t_{ij} - t_{ij}^0$ where t^0 represents a free flow travel time. Furthermore, as link capacity is not mentioned, we may assume that the units are already appropriately scaled.

Suppose that we have, as before, traffic densities on arc 1 given by

$$x_{i1} = 1000 \qquad x_{i2} = 250.$$

Then, we obtain the following travel time ratio for the two classes:

$$\frac{t_{i1}(x_i)}{t_{i2}(x_i)} = \frac{2x_{i1}^2 + x_{i1}x_{i2} + 1500x_{i1}}{4x_{i2}^2 + x_{i1}x_{i2} + 1700x_{i2}} = 2.59. \tag{5}$$

So that the travel time of class 1 on arc 1 is 2.59 times that needed by class 2 to traverse the same arc during the same time period, even in the presence of heavy congestion. For traffic composed mainly of class 1 vehicules, the results are yet more incoherent. Let

$$x_{i1} = 1200 \qquad x_{i2} = 50.$$

Then,

$$\frac{t_{i1}(x_i)}{t_{i2}(x_i)} = \frac{2x_{i1}^2 + x_{i1}x_{i2} + 1500x_{i1}}{4x_{i2}^2 + x_{i1}x_{i2} + 1700x_{i2}} = 21.55 \tag{6}$$

[3] Defining the generalized BPR-type function for car and truck classes as $t_k(x_i) = \alpha_{ik}\left[1 + \beta_{ik}\left(\frac{x_{i1}+\eta x_{i2}}{C_i}\right)^{\gamma_{ik}}\right]$, so that the exponent can vary across the two classes, we observe that, as exponents γ_{ik} differ across each class, the incoherence demonstrated above will increase. Indeed, letting Equation (1) remain unchanged and changing only the exponent of Equation (2) be equal to 5 (rather than 4), the ratio of the travel times of class 1 to class 2 is given by $t_1/t_2 = 0.16$. Thus, the travel time for class 2 at congestion is 6.26 times that of class 1! Increasing the difference between the exponents for class 1 and 2 will increase further this unphysical result.

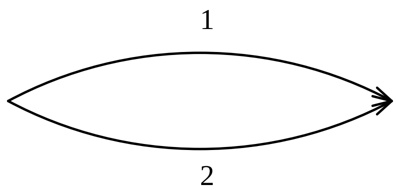

Figure 1: Two arc network

so that class 1 takes 22.55 times as long to traverse arc 1 in heavy congestion as class 2. In fact, in the limit, as x_{i1} increases and x_{i2} decreases, the time taken by one vehicule of class 1 on arc 1 increases polynomially with class 1, whereas the time taken by a vehicule of class 2 on the same arc approaches zero!

Clearly, these examples illustrate a non-physical result of the current formulation of the multiclass assignment problem. They occur because very important constraints between the classes were never included in the problem formulation. Indeed, the concept of *multiclass congestion* has not been defined. To do so, we should impose that, as the limit is approached (in terms of multiclass density), all of the classes on the arc tend to travel at the same speed. Technically, this means that the *ratios of the travel time functions of each class approach 1.*

3.2 Example 2: Inconsistent class definitions

Consider now a further example. Using the network of Figure 1, we take the following two linear travel time functions from [20] :

$$t_{i1}(x_i) = 5x_{i1} + 2x_{i2} + 500 \tag{7}$$

$$t_{i2}(x_i) = x_{i1} + 4x_{i2} + 200. \tag{8}$$

We consider the following traffic demands for each class: $d_1 = 50$ and $d_2 = 300$. Solving for user equilibrium according to Definition 2.1, we obtain the following solution:

$$x^* = \begin{pmatrix} 25 & 25 \\ 150 & 150 \end{pmatrix} \tag{9}$$

At equilibrium, the corresponding travel time ratio for the two classes are

$$t_{i1}(x_i) = 1.12\, t_{i2}(x_i) \tag{10}$$

for both arcs, $i = 1, 2$. Thus, class 1 is the slower class, both on arcs 1 and 2. This behaviour may be the desired result of modeling class 1 as heavy vehicles or as car drivers with a lower value of time. Note that x^* is only one of three equilibrium solutions for this network; the other two equilibrium solutions, however, do not include both classes on both arcs.

Now, let us change the composition of the traffic in terms of the two classes. For example, letting travel demands for each class be given by $d_1 = 50$ and $d_2 = 500$, we obtain the following

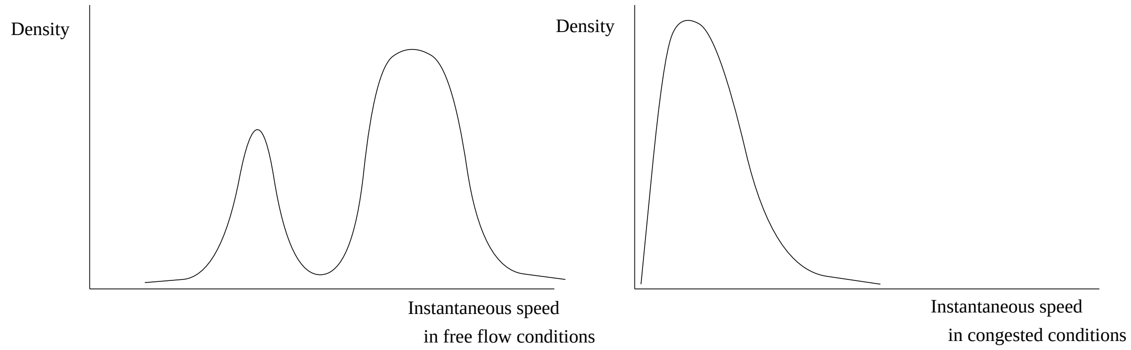

Figure 2: Typical speed profiles at free flow and at capacity

equilibrium solution, again solving for the equilibrium solution with both classes on both arcs:

$$x^* = \begin{pmatrix} 25 & 25 \\ 250 & 250 \end{pmatrix} \tag{11}$$

At this equilibrium point, the corresponding travel time ratio for the two classes is

$$t_{i1}(x_i) = 0.92\, t_{i2}(x_i) \tag{12}$$

for $i = 1, 2$. So, in this case, by simply modifying the composition of the travel demand, we have reversed the roles the classes. Whereas class 1 represented the slower class in the first case, class 1 is now the faster of the two classes, on both arcs.

The quadratic travel time functions of Section 3.1 give the same incoherent results. Indeed, let $d_1 = 50$ and $d_2 = 500$. Then, $t_{i1} = 0.07\, t_{i2}$. That is, class 2 is *much slower* than class 1. But changing the composition of the traffic, so that $d_1 = 500$ and $d_2 = 50$, we obtain, $t_{i1} = 9.9\, t_{i2}$. That is, class 2 now travels almost 10 times faster than class 1 on the same arc. Clearly, the definition of the classes is incoherent if such results can occur.

3.3 Interpreting Examples 1 and 2

In Example 1, the incoherent result occurred at congestion. We would expect that, in the physical limit of the arc the speeds of all classes approach a single speed. This is a fact well known by traffic engineers. Consider car and truck classes on a link. At free flow conditions, the speed profile on the link is "bimodal" since trucks travel at a lower speed than cars. As number of vehicles is increased towards congestion, the speed profile becomes unimodal as it shifts towards the left. (See Figure 2). To model this very intuitive behavioural phenomenon, it is necessary to include constraints linking the travel time functions of all classes sharing an arc at the point of congestion. That is, it is necessary to define *multiclass congestion.*

In Example 2, the inconsistency concerns the class definitions. When we change the travel demand, in this case of class 2, the roles of the two classes are interchanged. The equilibrium

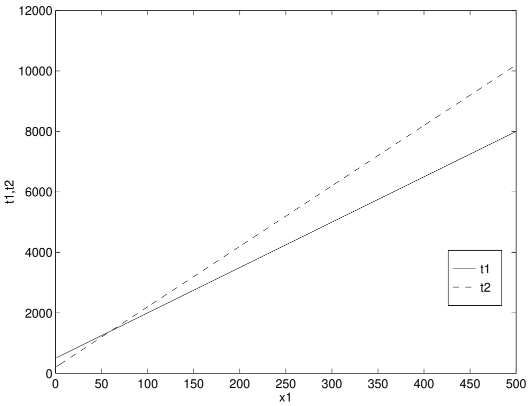

Figure 3: Travel time function that intersects before capacity (class ratio $x_2/x_1 = 5$)

solution is reversed when class 2 travel demand is increased to $d_2 = 500$, in that class 2 (which was the faster class) is now slower than class 1. It is easy to see that, at $d_2 = 400$, the travel times of the two classes on each arc are identical. However, this point ($d_1 = 50$, $d_2 = 400$) has no particular physical meaning. Whether we consider that the different classes represent different vehicle types (cars and trucks) or vehicles of the same type with different values of time, this cost function behaviour is clearly unrealistic. Indeed, in Example 2, the inconsistency arises in the equilibrium solution because the cost functions intersect at an arbitrary point (see Figure 3). This result is quite intuitive as well. Daganzo [10] noted that "one would expect c_i^1 [the travel time of a car on arc i] to be smaller than c_i^2 [the travel time of a truck on the same arc] for any combination of the x_i^ms [the class flows]." In short, class definitions should be constant for all traffic levels.

Examples 1 and 2 show clearly, on the one hand, that the existing formulation of the multiclass traffic assignment problem is incomplete. We will propose a new formulation of the problem in the following section. The examples show furthermore that all applications of the asymmetric traffic assignment problem are not identical.

Example 1 concerns all multiclass models where the vehicles included in the multivariate cost functions *share the same space during the same time period.* Indeed, we may identify certain asymmetric traffic assignment models for which this is not the case. Multiclass assignments of cars and bicycle traffic are one such example; in this case, heavy congestion resulting from a high density of cars on the link would not presuppose that congestion exists for the bicycles, if they use special lanes. Another such example is the modeling of two-way streets and some junction interactions. Again, congestion on a link in one direction, while it may affect the speed of the vehicles in the other direction, does not imply that both classes must be experiencing congestion. The interactions between the classes, in these examples, are not binding.

Example 2 provides us with constraints that are more difficult to interpret. Indeed, for models of multiple-vehicle-type or multiple-driver-class assignments, we would expect that the constraint should be valid: if one class has a faster travel time at traffic density level x, then so it should at density level $x + \Delta x$ on that link. Thus, a faster, more efficient, or better informed class remains so for all traffic densities before the limit. Any other circumstance would call into question the meaning of the different user class definitions.

However, an asymmetric traffic assignment problem designed to treat two-way streets and junction interactions in urban networks need not always be subject to such a constraint. Consider a low capacity, but fast road intersecting a traditional multi-lane road. At low volumes, the former would permit faster travel times. At high volumes, however, the performance of the fast, low capacity road would degrade before that of the traditional road, in which case the road which was faster at low volumes is slower at high volumes. Clearly, then, our second example requires a clarification as did the first: for Example 2 to make sense, the interactions being modeled must exist *on the same space during the same time period.* In short, this means that the interactions concern vehicles on a single arc.

Thus, we see that representing junction interactions through an asymmetric assignment, while similar to modeling multiple modes, classes, and vehicle types, is not identical. The constraints that must be taken into account in the problem definition and cost function formulation are not the same in the former as in the latter models. Indeed, we have seen that there does not exist one single asymmetric traffic assignment problem, but several.

3.4 Empirical Evidence

We have shown through examples the incoherent results that occur in a number of multiclass traffic assignments, as a result of defining the multiclass cost functions within each class only. It was illustrated schematically why joint constraints, linking the cost functional behaviour of all of the classes, must be satisfied.

In this section, we make use of existing traffic data from the French highway network to confirm that which we have demonstrated through examples on test networks. The data is obtained through SIREDO[4] machines, implanted by the National Road Directorate's technical service, SETRA, and their regional subdivisions, the Centres d'Études Techniques de l'Équipement (CETE). The SIREDO machines operate in conjunction with the magnetic loop detectors already embedded in most of the French national road network. These machines permit the storing of individual vehicle information, notably the length, instantaneous speed, and the time of presence (on the

[4] Système Informatisé de Receuil de Données

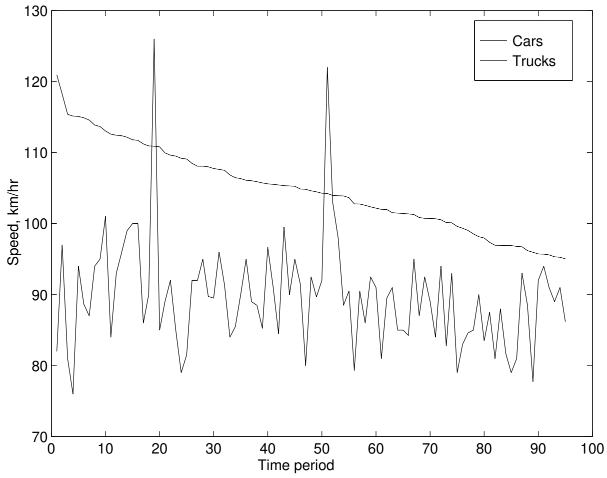

Figure 4: Car and truck speeds in freeflow conditions-Friday late morning

loop) of the vehicle. Data of this type was obtained for several of the highways in the Provence region, with the help of the CETE Mediterranée, CRICR Traffic Center in Marseille, and the Autoroutes du Sud, and for the A6 highway south of Paris through a SIREDO machine located at INRETS.

The Figures 4 and 5 correspond to data from the A6 highway south of Paris, which had the most significant congestion levels of all of the data sets. All data sets were collected during weekdays under normal traffic conditions (free from accidents or other incidents, and particularly averse weather conditions).

Each data point corresponds to a within-class (car or truck) speed averaged over 30 second intervals All averages include three single-direction lanes of the highway. The classes were identified by the length of the passing vehicle; greater than seven meters signified a truck, and less than seven meters a car. While large buses would also be classified as trucks according to this method, the incidence of mis-identification should be insignificant, as the number of buses on the highway during the week is quite low. Vehicles of less than one and a half meters or with speeds of more than 175 kilometers per hour were taken as misread data, and thus eliminated from the data set. The remaining data points give an indication of the joint speed patterns in multiple vehicle traffic. The reader may note that the car speeds are given by a smooth curve while the truck speed curve is quite rough. This is due to the large number of cars in each time period and the relatively small number of trucks. As class speeds are averaged over all vehicle speeds in that class, it is expected that a class having a large number of observations per period would give a smooth curve.

In Figure 4, we can see that, with the exception of two data points in the sample, the speed (and thus the travel time) of the slower, truck class is bounded by that of the faster, car class.

In Figure 5, taken from the same station on the A6 highway a Friday (early) morning, shows now the two class speed patterns during heavy traffic congestion. There is a significant amount of noise, due to the stop-and-go nature of traffic during rush hour, as well as to the fact that the data

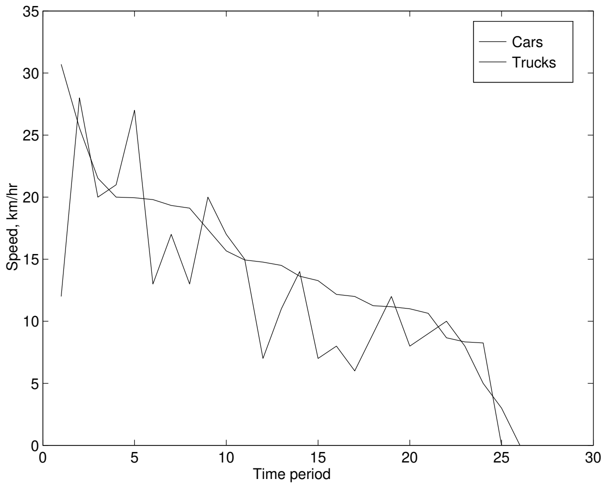

Figure 5: Car and truck speeds in congestion-Friday early morning

are averaged over three lanes, some of which move while others are at a standstill. Nevertheless, it is easy to observe that a smoothed curve for each class would confirm the existence of a bound on the speeds (and thus travel times) from the higher-speed (or inversely, lower travel-time) classes. Furthermore, in heavy congestion, the difference between the speeds of the two classes approaches zero, as expected.

4 Mathematical Incoherence

We have indicated that the commonly accepted multiclass formulation leads not only to behavioural inconsistency but that there is mathematical incoherence in the model results as well.

It is often the case in travel time functions that the multiclass dependencies are reduced to a single-class dependency, where the new variable is defined in terms of all of the classes. Passenger car-equivalent units, or PCU, are one such way of reducing the complexity of a multiclass assignment. In this case, the new variable x_i is defined as $x_i = x_{i1} + \sum_{j=2}^{K} E_{ij} x_{ij}$, where E_{ij} is the equivalence weight, for arc i, of the jth class (in terms of the first class), and K the total number of classes. Then, the capacity or maximum arc density, C_i, can be defined in terms of this single variable x_i as well. We will show that it is in most cases impossible to obtain meaningful results from a multiclass assignment in which this simplification is used.

Dropping the arc subscript i, let us define a single variable for two classes, $x = x(x_1, x_2)$, where $x(\cdot, \cdot)$ is an arbitrarily defined (not necessarily linear) functions of two variables, and two bivariate cost functions, $c_1(x)$ and $c_2(x)$, one for each class. Then, we examine the partial derivative, or Jacobian, matrix of $F = (c_1 \; c_2)$, $\nabla F(x)$ for one arc, and obtain that

$$\nabla F(x) = \begin{pmatrix} \frac{\partial c_1}{\partial x_1}(x) & \frac{\partial c_1}{\partial x_2}(x) \\ \frac{\partial c_2}{\partial x_1}(x) & \frac{\partial c_2}{\partial x_2}(x) \end{pmatrix}. \tag{13}$$

Then, we have that

$$\det[\nabla F(x)] = \left(\frac{\partial c_1}{\partial x}\frac{\partial x}{\partial x_1}\frac{\partial c_2}{\partial x}\frac{\partial x}{\partial x_2} - \frac{\partial c_2}{\partial x}\frac{\partial x}{\partial x_1}\frac{\partial c_1}{\partial x}\frac{\partial x}{\partial x_2}\right)(x) = 0, \tag{14}$$

and thus the Jacobian matrix $\nabla F(x)$ is singular for each x. For a general arc in the case of K classes, we obtain :

$$\det[\nabla F(x)] = \begin{vmatrix} \frac{\partial c_1}{\partial x} & \left(\frac{\partial x}{\partial x_1} \quad \frac{\partial x}{\partial x_2} \quad \cdots \quad \frac{\partial x}{\partial x_K}\right) \\ \frac{\partial c_2}{\partial x} & \left(\frac{\partial x}{\partial x_1} \quad \frac{\partial x}{\partial x_2} \quad \cdots \quad \frac{\partial x}{\partial x_K}\right) \\ \vdots & \\ \frac{\partial c_K}{\partial x} & \left(\frac{\partial x}{\partial x_1} \quad \frac{\partial x}{\partial x_2} \quad \cdots \quad \frac{\partial x}{\partial x_K}\right) \end{vmatrix} \tag{15}$$

whose rows are clearly linearly dependent.

We now verify that, in most cases, this causes an undesirable indeterminacy of the results in terms of repartition across classes. Let $\{x^*_{i1}\}$ and $\{x^*_{i2}\}$ be the equilibrium point for each arc i of the netweok of Figure 1. Suppose that the single variable, x is given by $x = \alpha_1 x_1 + \alpha_2 x_2$. (We are assured of the existence of at least one such equilibrium by the continuity of $c_1(x)$ and $c_2(x)$.) Assume furthermore that both arcs are used by both classes. Then, any set satisfying

$$\alpha_1 x_{i1} + \alpha_2 x_{i2} = \alpha_1 x^*_{i1} + \alpha_2 x^*_{i2}, \tag{16}$$

for each arc $i = 1, 2$, and

$$\sum_{i=1}^{N} x_{i1} = d_1 \quad \text{and} \quad \sum_{i=1}^{N} x_{i2} = d_2 \tag{17}$$

is also a solution. It is clear that the rank of this system of equations is less than the number of variables, as both the rows on the left-hand side and on the right-hand side of Equation (16), for all the links, will be equal to $\alpha_1 d_1 + \alpha_2 d_2$. Thus, there are an infinite number of $\{x_{i1}\}$ and $\{x_{i2}\}$ that satisfy the system. (See also Netter, [21]). This result for two arcs may of course be extended to more complex route choices in larger networks. Thus, if each class' cost function depends only on a single variable regrouping the class variables, where the weights between the class variables are pre-defined, then the multiclass problem is trivially reduced to a sum of single-class problems.

Clearly, this calls into question a number of commonly accepted multiclass model formulations. Imagine, for example, a model of multiple driver classes, where each class is distinguished by a value of time (VOT) unique to that class. One would suppose, intuitively, that the value-of-time classes all have the same travel time function, say $t(x_i)$, where $x_i = x_{i1} + x_{i2} + \cdots + x_{iK}$ for all K value of time classes. Then, the cost functions of each class differ linearly by their values of time. A typical example of this formulation is $c_{ik} = p_{ik} + V_k t(x_i)$, where c_{ik} is a generalized cost experienced by class k on arc i, p_{ik} is a fixed (possibly class-specific) toll price, and V_k is the value of time of class k. The result is then, as we have shown, an infinite number of possible class splits on each arc. That is, the results of our assignment will be entirely inconclusive. This is an undesirable situation if we hope to make policy recommendations (optimal toll prices, or the effects of road pricing, reserved lanes, etc.) on the basis of the model findings.

Finally, note that not all equilibrium solutions satisfy the implicit assumption made above that the classes share at least some of the routes. Indeed, it may formally happen that disjoint, class-specific routes are available for each origin-destination pair, in which case the relative share of each class for each route is of course well specified. This situation is however highly unlikely in realistic networks, and singularity of the Jacobian is thus better avoided if one wishes to use multiclass techniques in practical applications.

5 Reformulating the multiclass model and the associated cost functions

We have seen that the asymmetric traffic assignment problem is more complex and diverse than a straightforward extension of the single-class assignment, and that the current formulation of the problem is not always sufficient for guaranteeing a coherent solution. Let us restate the following assumption:

Assumption 5.1 *The travel time functions on any given arc depend only on the vehicle classes on that arc, where all of the vehicle classes in the function occupy the same road space during the same time period.*

Assumption 5.1 is satisfied for models of different vehicle types and multiple driver classes, as well as some models of detailed urban networks that consider delays due to different classes on the same link (e.g. the effect of turning traffic on through traffic). Our new problem formulation is then given below.

Definition 5.1 (Multiclass travel time functions) *The travel time for each class on a given arc is bounded by those of the slower classes for all density levels, and, the ratios of the travel times of all classes approach 1 as congestion is approached.*

Example 1 showed that, in the limit (in terms of the maximum density on the link), the travel time ratios should equal one. If congestion is defined in such a way that it may be exceeded (for the purpose of avoiding explicit constraints on the links), then beyond this point, all vehicles on an arc, regardless of their class, should travel at the same speed (which approaches zero). Example 2 showed the effect of arbitrary intersections of the travel time functions of the different classes before capacity. In order to avoid such non-physical behaviour, as was observed in Example 2, the travel time functions of each class must be bounded from above by the functions of all other classes with longer travel times, for all density levels up to the point of congestion. Thus we obtain the following constraints on the multiclass cost functions:

$$t_{ik}(0) \quad = \quad t^0_{ik}, \qquad \forall i, k \tag{18}$$

$$x_i \to C_i \quad \Rightarrow \quad \frac{t_{ik}(x_i)}{t_{i\hat{k}}(x_i)} \to 1 \qquad \forall k, \hat{k} \in K \tag{19}$$

$$t_{i1}(x_i) \;\leq\; t_{i2}(x_i) \;\leq \cdots \leq\; t_{i,k-1}(x_i) \;\leq\; t_{ik}(x_i), \qquad \forall i,\, \forall x \tag{20}$$

$$\det[\nabla t(x)] \quad \neq \quad 0, \tag{21}$$

where t^0_{ik} is the freeflow travel time for class k on arc i, C_i is the multiclass density at congestion on arc i, and $\nabla t(x)$ is, as before, the Jacobian matrix of the travel time functions $t(x)$. We will examine the implications of this formulation in the following section.

5.1 Defining multiclass cost functions

We now proceed to put together the different elements for the purpose of proposing coherent multiclass cost functions. We propose to consider a cost structure composed of three terms. The first term ensures that Equation (18) holds, so that, at freeflow conditions, the travel time for each class equals a class-specific constant. The second term in the cost function must ensure that Equation (19) holds at congestion, C, defined as a unique point in terms of all of the classes. This second term should ensure that the ratios of the travel times of all classes, at this point, approaches 1. That is, the second term must *dominate* the third term (the first is a constant) at congestion. In addition, this term should be written in terms of a single variable so that the definition of congestion is constant across all of the classes. The third and final term of the cost function would then represent the individual reaction of each class, where these reactions are not written only in terms of a single variable, to avoid the Jacobian matrix singularity that we saw in the preceding section. Furthermore, this third term will not dominate at or before congestion. The cost function for each class would thus take on the following form, where we have supressed the arc indices for the sake of clarity:

$$t_k(x_1, x_2, \ldots, x_K) = t^0_k + J(x) + S_k(x_1, x_2, \ldots, x_K). \tag{22}$$

We will now look more closely at some possible functional forms for the second and third terms, $J(x)$ and $S_k(x_1, x_2, \ldots, x_K)$.

5.1.1 Multiclass congestion

The second term serves to describe the effect of multiclass congestion and ensures that, at the congestion, the ratios of the travel times of the different classes sharing the link approach 1. Thus, this term should blow up at that point, or at least increase much faster than $S_k(x)$ near this point. If congestion can be defined in terms of a single variable $x = x_{i1} + \sum_{k=2}^{K} E_{ik} x_{ik}$ (for a particular arc i), then $J(x)$ can be defined in a manner that contributes very little to the total cost for low densities, but increasingly so as congestion is approached. Again, dropping arc subscripts, one such example is

$$J(x) = \frac{\gamma}{C - x}, \tag{23}$$

where γ and C are arc-specific constants.

If congestion cannot be identified as a single point, then a simplification must be made; this is often done through the use of high-order polynomials. In this case, however, the condition (19) is satisfied only in the limit as x tends to infinity, and as such is only an approximation (within a few percent) to the behavior that we wish to model. As such, if $J(x)$ is defined, for example, by a fourth-order polynomial (as is the single-class BPR function, in many applications). Another possible form for $J(x)$ when congestion cannot be identified as a single point was proposed by Leurent [14] following Spiess [26] and is based on the two-dimensional sections of an obtuse three-dimensional cone:

$$J(x) = 1 + \eta + \sqrt{\gamma^2(1-x)^2 + \beta^2} - \gamma(1-x) - \beta, \tag{24}$$

where $\beta = \frac{\eta(2\gamma-\eta)}{2(\gamma-\eta)}$, and $\gamma > \eta$.

5.1.2 Class-specific reactions

The third term ensures that class definitions are constant over all traffic levels. In order to prevent the phenomena of Example 1 (Section 3.1) from occuring, and thus satisfy Definition 5.1 (and, equivalently, Equation (19)), this term must not dominate before or at congestion, and must therefore increase less rapidly than the second term. Furthermore, the particular form of this term, and the constants, must be chosen carefully to ensure that the travel time functions for each class are bounded by those of the slower classes, as stated in Equation (20), as well as to prevent singularity, thus satisfying Equation (21). If, for example, $J(x)$ is given by a quartic, we may define $S_k(x)$ as a linear, quadratic, or cubic term of the $\{x_k\}$.

Let us renumber the classes so that the fastest class is class 1 and the slowest is class K. Then, we want to ensure that $t_1(x_1,\ldots,x_K) \leq t_2(x_1,\ldots,x_K) \leq \cdots \leq t_K(x_1,\ldots,x_K)$, for all feasible values of x_i, $i = 1,\ldots K$. Since $J(x)$ is constant across all classes, this reduces to defining

$$t_1^0 + S_1(x_1,\ldots,x_K) \leq t_2^0 + S_2(x_1,\ldots,x_K) \leq \cdots \leq t_K^0 + S_K(x_1,\ldots,x_K). \tag{25}$$

Clearly, the freeflow travel times of each class k should be in accordance with this constraint, $t_1^0 \leq t_2^0 \leq \cdots \leq t_K^0$. So, the only difficulty is to define class-specific cost function terms $S_k(x_1,\ldots,x_K)$ that do not diverge at congestion, that *do* satisfy inequality (25) for appropriate freeflow travel time constants, and that do not lead to Jacobian matrix singularity.

Let $S_1(x_1,\ldots,x_K)$, for the fastest class, be defined as a function of all K classes:

$$S_1(x_1,\ldots,x_K) = S_1(x). \tag{26}$$

So that $S_1 \leq S_2$, $\forall(x_1,\ldots,x_K) \in X^f$, we add an additional positive function $\tilde{S}_2(x_1,\ldots,x_K)$ of some or all of the classes. Then,

$$S_2(x) = S_1(x) + \tilde{S}_2(x_1,\ldots,x_K). \tag{27}$$

As such, we define the second term of the travel time function (22) by the following set of equations:

$$\begin{aligned} S_1(x) &= S_1(x), && (28)\\ S_k(x) &= S_{k-1} + \tilde{S}_k(x_1,\ldots,x_K), \qquad \forall\, k = 2,\ldots K. && (29) \end{aligned}$$

We are interested in conditions on the functions $\tilde{S}_k(x)$ that guarantee that Jacobian matrix singularity (Constraint (21)) is avoided. Defining $\tilde{S}_1(x) = J(x) + S_1(x)$, we see from (29) that

$$J(x) + S_k(x) = \sum_{\ell=1}^{k} \tilde{S}_\ell(x). \tag{30}$$

Recalling now that

$$\nabla t(x) = \left(\begin{array}{ccc} \frac{\partial J}{\partial x_1}(x) + \frac{\partial S_1}{\partial x_1}(x) & \cdots & \frac{\partial J}{\partial x_K}(x) + \frac{\partial S_1}{\partial x_K}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial J}{\partial x_1}(x) + \frac{\partial S_K}{\partial x_1}(x) & \cdots & \frac{\partial J}{\partial x_K}(x) + \frac{\partial S_K}{\partial x_K}(x) \end{array} \right), \tag{31}$$

we obtain from (30) that

$$\det[\nabla t(x)] = \left| \begin{array}{ccc} \frac{\partial \tilde{S}_1}{\partial x_1}(x) & \cdots & \frac{\partial \tilde{S}_1}{\partial x_K}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1} \sum_{\ell=1}^{K} \tilde{S}_\ell(x) & \cdots & \frac{\partial}{\partial x_K} \sum_{\ell=1}^{K} \tilde{S}_\ell(x) \end{array} \right| = \left| \begin{array}{ccc} \frac{\partial \tilde{S}_1}{\partial x_1}(x) & \cdots & \frac{\partial \tilde{S}_1}{\partial x_K}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial \tilde{S}_K}{\partial x_1}(x) & \cdots & \frac{\partial \tilde{S}_K}{\partial x_K}(x) \end{array} \right|. \tag{32}$$

Consequently, we conclude that the desired nonsingularity of the Jacobian is guaranteed whenever

$$\left| \begin{array}{ccc} \frac{\partial \tilde{S}_1}{\partial x_1}(x) & \cdots & \frac{\partial \tilde{S}_1}{\partial x_K}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial \tilde{S}_K}{\partial x_1}(x) & \cdots & \frac{\partial \tilde{S}_K}{\partial x_K}(x) \end{array} \right| \neq 0. \tag{33}$$

This condition is easily verifiable in the simple case where there are two classes ($K = 2$). In this case, (33) simply states that $\tilde{S}_1(x)$ and $\tilde{S}_2(x)$ are not identical up to a constant. If they were, then one would obtain that $J(x) + S_1(x)$ is the same as $J(x) + \alpha S_1(x) + \beta$, for some constants α and β. This would mean that we could again describe the cost functions associated to the two classes using a single variable x, inducing the singularity discussed above. In the case of more than two classes, we must impose that the $\tilde{S}_k(x)$ are linearly independent of $\tilde{S}_j(x)$ for all $j, k \in K, j \neq k$

5.2 Existence and Uniqueness Implications

What are the implications of imposing equations (18) through (21) on the cost functions in the asymmetric traffic assignment problem? To respond to this question, we must briefly review the existence and uniqueness results of the asymmetric problem as it has existed to date.

We remind the reader of the existence and uniqueness results for the separable traffic assignment model. It is both necessary and sufficient, for a unique equilibrium to exist, that the travel cost functions $t_i(x_i)$ are strictly monotonically increasing in the feasible domain, X^f. The extension to a multiclass model then required a multivariate version of this condition, so as to permit defining a similar constraint. This amounts to requiring that the total cost operator $t_{ik}(x)$ (where x is now a vector) is strictly monotone for all classes, K. This condition is equivalent to requiring that the Jacobian matrix of the cost operator $\nabla t(x)$ is positive definite. It was demonstrated that this condition ensures the uniqueness of the solution (see Stampacchia, [27]).

The existence of a solution was proven under more lenient requirements (continuity of $t(x)$) for the traffic assignment problem through use of a fixed-point theorem).

The uniqueness condition will be satisfied if the Jacobian matrix $\nabla t(x)$ is positive definite. (The condition for existence being satisfied automatically, for continuous travel time functions.) Since Assumption 5.1 is satisfied, and thus the interactions between vehicle classes occur on the same link, this Jacobian matrix will be block diagonal. Then, we require each block to be positive definite. This reduces to ensuring positive definiteness for each set of arc cost functions.

For $\nabla t(x)$ to be positive definite, it is sufficient that the terms on the diagonal are much greater than the off-diagonal terms. From (32), this condition reduces to

$$\left|\frac{\partial \tilde{S}_i}{\partial x_i}(x)\right| > \max\left[\sum_{i\neq j}\left|\frac{\partial \tilde{S}_j}{\partial x_i}(x)\right|, \sum_{i\neq j}\left|\frac{\partial \tilde{S}_i}{\partial x_j}(x)\right|\right]$$

for all $i = 1, \ldots, K$. This means that the class-specific reactions would have to be stronger with respect to that particular class than with respect to the other classes. (As expected, this condition seldom holds in reality and positive definiteness as well as solution uniqueness cannot be guaranteed in practice.)

The condition for existence, which relies only on the continuity of the operator, has not been altered by our formulation. Furthermore, the cost function formulation of Equation (22) does not significantly change the requirements for solution uniqueness.

In the deterministic multiclass model, uniqueness of the class densities is ensured for strictly monotone cost operators, or equivalently for positive definite cost function Jacobians. This property has not been altered by our formulation. We have shown that it is often the case, in articles concerning multiclass models, that the cost function Jacobian is singular. Clearly, then, the total cost operator is not strictly monotone, and the result is an infinity of possible solutions in terms of the classes. This situation can be avoided through use of the cost function formulation presented in this paper.

Our formulation does not, however, ensure that the cost function Jacobian is positive definite, but only that it is non-singular. For non-monotone cost operators, the asymmetric assignment solution will not be unique. However, so long as singularity of the Jacobian matrix is avoided, the number of solutions to the problem will be *finite*[5]. This will hold in the deterministic, arc-based, models. This will hold, as well, for the paths in the path-based assignment methods (stochastic assignments, equilibration algorithms).

As cost operator monotonicity cannot generally be satisfied in practice, we are still often faced with non-uniqueness of the equilibrium solutions. This issue raises further questions that will be addressed elsewhere. (See for example [30].)

We conclude this section by noting that, when the maximum density C_i is approached on a particular arc i, our theory requires that the class-independent functions $J(x)$ dominate all

[5] The number of solutions is finite when the Jacobian is non-singular as long as the feasible region is compact. This is the case in the traffic assignment problem because the densities may all be bounded above by the total demand per unit length.

other components of the travel time. Theoretically, the Jacobian will increase with increasing density, and will thus not approach zero. On the other hand, if the difference between the travel time functions of each class differs by less than machine precision, then numerical difficulties will result. Fortunately, this should not be the case for reasonable values of $J(x)$ and $S_k(x)$.

6 Case Studies

We conclude with case studies for two common applications of the problem: first, a model of multiple driver classes. Then, we present a model of car and truck classes on an interurban network. In this example, we calibrate car- and truck-class cost functions from real network data, obtained from the SIREDO system on the French highways. We again drop the arc index throughout this section.

6.1 A model of multiple driver classes

Models of multiple driver classes, while they have existed (notably in France) for decades, have recently received newfound attention. (See, for example, references in Mahmassani and Peeta [17] or Peeta and Mahamassani [23] within the context of dynamic assignment models, Labbé et al. [13], Arnott, et al. [1] and Chen and Bernstein [5] within the context of congestion pricing, and Bifulco [2], Leurent [15], Marcotte and Zhu [18] and Dial [11], who discuss potential models and algorithms.) With the advent of innovative traffic control strategies, distinguishing between different classes of users often becomes a prerequisite for a useful network equilibrium model.

The most frequently cited driver class distinction is in terms of values of time (VOT). The distinction among drivers according to class-specific values of time permits the testing of road pricing scenarios. Without such a distinction between drivers, all travelers (in the model) would react in an identical manner to toll, or other, price increases (fuel prices, taxes, etc.). Through explicit treatment of different driver classes, each valuing the ratio of money to travel time to a different extent, it is possible to analyze to what degree the price increases change the current equilibrium. From here on, we drop the arc indices for the sake of clarity.

Ideally, we would like to define a single variable x, $x = x_1 + x_2 + \cdots + x_K$. Then, the multiclass travel cost to class k could be expressed as

$$c_k(x) = p_k + V_k t(x), \tag{34}$$

where V_k is the value of time of class k, p_k is the toll for class k and $t(x)$ is a travel time in terms of the total density, x. This formulation was used by, among others, Van Vliet [28] and Bifulco [2]. However, the reader should now be able to observe the difficulty that arises from such a formulation. Indeed, the Jacobian matrix of the cost functions is singular, as justified above. We have, once again, an infinite number of possible solutions in terms of the classes. This result is expected since the travel time dependence was defined only in terms of a single variable; we saw earlier that such a definition leads to an infinity of possible splits across the classes.

We now show that it is nonetheless possible to develop a multiple driver class model able to handle different perceptions of the cost-to-time ratio across the population.

Let us recall the constraints we have imposed on the multiclass cost functions, Equations (18) through (21). Since the multiple driver classes *share the same space during the same time period*, thus satisfying Assumption 5.1, these constraints apply. We can thus reinterpret the multiple value of time model within the context of these cost function restrictions needed to insure discrete, coherent solutions.

To do so requires making some assumptions about the travel times of each class and, in particular, about the class-specific interaction terms. Returning to the cost function formulation of Equation (22), we see that it will be necessary to define differently $S_k(x)$ for each class; that is, the class-specific part of the cost function must vary across the classes in order for the cost function Jacobian to be non-singular.

Let us define class 1 as those drivers with the highest value of time and class K as those having the lowest value of time. Then, we will assume that class 1 drivers prefer to arrive at their destination faster than drivers of classes 2 through K. Similarly, class 2 drivers prefer to arrive at their destination before those of class 3 through class K, and so on.

Then, $J(x)$ is defined in terms of the total density $x = \sum_{k=1}^{K} x_k$, where we assume here that all K classes of vehicles are of similar type and thus weighted equally[6], and $J(x)$ is the dominating term at congestion. We may define $J(x)$ according to any of the functional forms presented in Section 5.1.

The important feature of these multiple driver class travel time functions, which permits us to ensure a finite number of coherent solutions to the multiclass problem, is that we insist that:

$$S_1(x) < S_2(x) < \cdots < S_K(x), \tag{35}$$

for $x_k \neq 0 \ \forall k$. These terms may include the effects of the higher VOT-class vehicles passing the slower vehicles of lower VOT class on the road. Thus, the class K drivers, with the lowest value of time, are the slowest in that they are passed by all $K-1$ vehicle-classes that travel faster than them on the road. The vehicles of class $K-1$ are slower than those vehicles of classes 1 through $K-2$, and so on. So, the class-specific reaction term of the travel time function is of the form:

$$\begin{aligned}
S_1(x) &= S_1(x_1, \ldots, x_K) \\
S_2(x) &= S_1(x_1, \ldots, x_K) + \tilde{S}_2(x_1) \\
S_3(x) &= S_1(x_1, \ldots, x_K) + \tilde{S}_2(x_1) + \tilde{S}_3(x_2) \\
&\vdots \\
S_K(x) &= S_1(x_1, \ldots, x_K) + \tilde{S}_2(x_1) + \cdots + \tilde{S}_K(x_{K-1}),
\end{aligned} \tag{36}$$

where, for all x,

$$\frac{\partial \tilde{S}_k}{\partial x_{k-1}}(x_{k-1}) > 0, \tag{37}$$

[6] Extending this model to multiple vehicle types is straightforward; see Section 6.2.

for $k = 2, \ldots, K$. Note that one might reasonably assume that the travel time of the fastest class is increasing with the density in each class, that is

$$\frac{\partial S_1}{\partial x_k}(x) > 0 \tag{38}$$

for $k = 1, \ldots, K$ and all x. The freeflow travel times should be defined in accordance with the premise that the higher value of time classes prefer to travel faster than those of lower value of time. Consequently, the freeflow travel times of each class should be strictly inferior to those of all classes having a lower value of time:

$$t_1^0 < t_2^0 < \cdots < t_K^0. \tag{39}$$

Consequently,

$$t_1^0 + J(x) + S_1(x) < \cdots < t_K^0 + J(x) + S_K(x). \tag{40}$$

Thus, Equations (18) through (21) are satisfied by this model.

The multiclass cost is then defined as in Equation (34), with the exception that the travel time function is unique to each class,

$$c_k(x) = p_k + V_k t_k(x). \tag{41}$$

Equations (36)–(38) and the increasing nature of $J(x)$ then imply that condition (33) holds, thus guaranteeing discrete solutions.

6.2 Multiple vehicle types on an interurban network

For modeling vehicles of different sizes, defining congestion in terms of a single point becomes an important issue. We will thus assume that, based on vehicle size, a set of equivalency weights can be obtained. For example, to convert trucks to passenger car unit (PCU) equivalents, a weight of 2 or 3 is often used, though it may vary by road type (see Cohen and Liger [7] and McShane and Roess [19]).

Cohen [6] and Cohen and Liger [7] showed furthermore that this equivalence ratio *is not a constant* but varies with the percentage of trucks in the traffic stream. (The relationships examined were total vehicular flow (and thus practical capacity) versus percentage of occupation (in terms of all vehicles), this for different percentages of trucks in the traffic stream. The results indicates clearly that an increase in the percentage of trucks in the traffic stream affects the flow on the link *differently* than an “equivalent” increase in cars. This implies, then, that a simple constant equivalence is not valid for accounting for the effect of heavy vehicles sharing the roadspace with cars. In addition, Cohen and Liger [7] demonstrated that the functional relationship that links the percentage of heavy vehicules in the traffic stream to capacity approaches a constant value (of approximately 2) as the number of vehicles present on the link aproaches the physical limit of that link. This means, then, that before congestion, a constant equivalence factor between cars and trucks is not valid; the term $S_m(x_1, \ldots, x_K)$ is thereby confirmed empirically.[7]

We thus define $J(x)$, the second term of the multiclass cost function from Equation (22), in terms of this set of equivalency weights. The methodology for the multiple vehicle type model then follows directly from Section 5.1. If congestion can be identified as a single point, then $J(x)$ can be of the form

$$J(x) = \frac{\gamma}{C - (x_c + E x_t)}$$

where C and γ are arc-specific constants, E is the (possibly arc-specific) truck-car equivalency weight, and x_c and x_t are the densities of cars and trucks, respectively. The freeflow travel times of cars and trucks t_c^0 and t_t^0, are easy to measure from traffic data.

If C cannot be identified, then, as discussed in Section 5.1, we can make use of higher order polynomials or a conic cost function as shown in Equation (24). to ensure that at congestion, the ratios of the travel times of cars and trucks approach 1.

It thus remains only to estimate and calibrate the class-specific reactions to car and truck densities. Assuming that cars are the faster class (so that $t_c^0 < t_t^0$), we must define first the car-specific travel time term $S_c(x_c, x_t)$ as a function of both car and truck densities. Then, the truck-specific travel time term is defined as a sum of the car-specific term plus an additional positive term.

The travel time function calibration made use of the data from the SIREDO machines located on the A6 highway towards Paris near INRETS. The individual vehicle information is collected through the use of two magnetic loops underneath the road surface, and the data is sent by telephone line to a computer at INRETS. The data in which we are interested must contain a sufficiently high percentage of truck traffic, as well as both the fluid and congested states. The data sets were thus collected before and at the morning peak, as congestion is heavier in the morning than the evening on this section. Furthermore, individual vehicle data is essential as no aggregated statistics (that we aware of) calculate the average truck and car speeds independently. From the individual vehicle data, such class-specific parameters can be obtained. The major difficulty encountered was the low percentage of trucks in the traffic stream during periods of significant congestion.

Class flows and average speeds can be calculated easily from the data. Then, class densities and class (unit) travel times follow directly. In order to capture the transient phenomena, we took averages over one minute intervals. Clearly, the disadvantage of short averaging intervals is the ever-present noise (notably, the stop-and-go traffic on a three-lane highway). However, cruder averaging would serve to mask the critically congested periods, as they often last no more than a few minutes at a time.

For the calibration process, we made use of the nonlinear Gauss-Newton method programmed into the SAS package of statistical software. The basic function was of the form given by Equa-

[7] However, to confirm empirically our constraint requiring the terms $S_j(x_1, \ldots, x_K) \neq S_m(x_1, \ldots, x_K)$ before congestion, it would be necessary to examine the *individual* effects of trucks on both car and truck travel times, for example. Demonstrating that these two relationships *are not identical* would thus confirm that the singularity arising from most multiclass model formulations is, not only mathematically undesirable, but behaviourally incorrect as well. The results of such an experimental study would be both interesting and useful.

tions (22) and (23), where the independent variables were car and truck densities (x_c and x_t, respectively). The class-specific terms were calibrated both as linear and as quadratic terms from a data set of 368 points. We obtained:

$$t_c = 0.046\left[1 + \frac{1}{C - x_c - 2x_t} + 0.86x_c + 1.56x_t\right] \tag{42}$$

$$t_t = 0.067\left[1 + \frac{1}{C - x_c - 2x_t} + 0.80x_c + 1.25x_t\right] \tag{43}$$

for the linear pre-congestion increase in travel time. These parameters provided the optimal sum of squares residual error for the car and truck functions while satisfying Constraints (18) through (21). The functions based on a quadratic increase of the travel time before congestion are as follows

$$t_c = 0.046\left[1 + \frac{1}{C - x_c - 2x_t} + 0.019x_c^2 + 0.13x_t^2\right] \tag{44}$$

$$t_t = 0.067\left[1 + \frac{1}{C - x_c - 2x_t} + 0.020x_c^2 + 0.11x_t^2\right] \tag{45}$$

The quadratic functions are more appealing theoretically in that they more accurately replicate the parabolic "fundamental diagram" of traffic engineering.

Both sets of functions give the travel time in minutes per kilometer. The maximum density C used in the calibration, obtained from the data, was 50 pcu per kilometer per lane, or 150 vehicules per kilometer for the three-lane section. The functions were calibrated so as to give a freeflow speed for cars of approximately 130 kilometers per hour, and approximately 90 kilometers per hour for trucks. Then, the constants in front of the linear and quadratic terms were calibrated with a Gauss-Newton nonlinear estimation procedure. Convergence of the method was achieved in all cases.

To calibrate the functions, it was necessary to know the maximum density of the road section. We calculated this value empirically by examining the density when the travel speeds were very low. It is desirable to calibrate the travel time functions this way, when the data is available, as they thereby correspond more closely with the physical characteristics of the road sections. Certainly, there will be instances in which this data cannot be gathered for the network as a whole, and as such, an identification of the maximum densities may be unfeasible. In this case, the asymptotic part of the travel time functions can be calibrated according to a high-order polynomial, as discussed earlier in the text; however, as such the multiclass conditions are satisfied only in the limit as $x \to \infty$. In practice, the few percentage points of error resulting from this approximation may be acceptable.

The reader will observe that Conditions (18) through (21) are satisfied by the calibrated functions. Furthermore, it is clear that they are not monotonic.

As far as the practical use of functions (42) to (45) is concerned, it is important for numerical reasons to prevent the load x on any arc from approaching too closely the maximum density value, C. To do so, each travel time function can simply be set to a fixed value for loads greater than or equal to the value of C, or to a linear function whose derivative matches that of the original function just before the singular point.

7 Conclusions and perspectives

We have thus demonstrated, both through examples and through empirical data taken from the French highways, that the multiclass traffic assignment problem, as it has been defined to date, is incomplete, at least in the case where the cost functions depend on travel times. Examples on test networks permitted us to show the effects of this incomplete formulation on traffic assignment results–incoherence in the results may occur in these cases. Furthermore, a common problem in many multiclass assignment model formulations was Jacobian matrix singularity. This leads to an infinite number of class solutions, and consequently to a multiclass assignment which gives no information about the dispersion of the individual classes.

We reformulated the asymmetric traffic assignment problem, and particularly the functional forms used to define travel time in the model definition, so as to avoid these behavioural inconsistensies and mathematical difficulties. Our definition was then quantified and a general class of functional forms proposed. This travel time function class was examined within the context of mathematical results for problem existence, uniqueness. It was shown that our new formulation did not drastically change these results (uniqueness is still not guaranteed in practice), but it did ensure that an infinite number of class solutions would not result.

The final part of the paper included two case studies. Coherent and mathematically rigourous models were presented for a multiple driver-class model as well as an interurban multiple-vehicle-type model. Car and truck class travel time functions were calibrated for this latter case study from real highway network data.

It is hoped that the new problem formulation will lead to a more widespread use of coherent multiclass equilibrium models in practice.

Acknowledgements: The authors wish to thank Michel Bierlaire for many useful discussions. Helpful comments from Piet Bovy, Gautier Chatelus, Fabien Leurent, Hani Mahmassani, and Marcos Papageorgiou are gratefully acknowledged.

References

[1] Arnott, R., DePalma, A. and Lindsey, R (1994) The Welfare Effects of Congestion Tolls with Heterogeneous Commuters, *J. of Transp. Econ. and Policy* **28**(2).

[2] Bifulco, C. (1993) A multimodal and multiuser assignment model, Proceedings of the 2nd EURO Working Group Meeting on Transportation, INRETS, Arcueil, France.

[3] Braess and Koch (1979) On the Existence of Equilibria in Asymmetrical Multiclass-User Transportation Networks, *Transpn. Sci.* **13** (1), 56–63.

[4] Bureau of Public Roads (1964) U.S. Highway Capacity Manuel, US Dept of Commerce, Washington D.C.

[5] Chen, O.J., and Bernstein, D. (1995) AM/PM Congestion Pricing with a Single Toll Plaza, TRB Preprint 950623, Washington DC.

[6] Cohen, S. (1983) Méthodes d'estimation de la capacité des autoroutes, IRT Research report 62, Arcueil.

[7] Cohen, S. and Liger, M. (1993) The taking into account of different vehicle categories in the operation of motorways, Proceedings of the World Conference on Transportation Research, Lyon.

[8] Dafermos, S.C. (1971) An Extended Traffic Assignment Model with Applications to Two-way Traffic, *Transpn. Sci.* **5**, 366–389.

[9] Dafermos, S.C. (1972) The Traffic Assignment Problem for Multiclass-User Transportation Networks, *Transpn. Sci.* **6**, 73–87.

[10] Daganzo, C.F. (1981) Equilibrium Model for Carpools on an Urban Network, Trans. Res. Rec. 835, 74–79.

[11] Dial, R. (1995) T2: A Bicriteria Model, paper presented at the Transportation Research Board Annual Meeting.

[12] Fernandez, J.E.L. and Friesz, T.L. (1983) Equilibrium Predictions in Transport Markets: The State of the Art. *Transpn. Res.* **17B**(2), 155–172.

[13] Labbé, M., Marcotte, M. and Savard, G. (1994) A bilevel model of taxation and its application to optimal highway pricing. Unpublished article, SMG, Université Libre de Bruxelles.

[14] Leurent, F. (1991) *Notes sur l'utilisation des modèles statiques d'affectation de trafic routier urbain*, Memoire de DEA de l'Ecole Nationale des Ponts et Chaussées et de l'Université de Paris XII, Creteil.

[15] Leurent, F. (1994) Cost versus Time Equilibrium over a Network, *Transpn. Res. Record* **1443** 84-91.

[16] Mahmassami, H.S., and Mouskos, K.C. (1988) Some Numerical Results on the Diagonalization Algorithm for Network Assignment with Asymmetric Interactions Between Cars and Trucks. *Transpn. Res.* **22B**(4), 275–290.

[17] Mahmassani, H.S., Peeta, S.(1992) Network Performance under System Optimal and User Equilibrium Dynamic Assignments: Implications for ATIS, *Transp. Res. Rec.* 1408, 83–93.

[18] Marcotte, P. and Zhu, D.(1994) An Efficient Algorithm for a Bicriterion Traffic Assignment Problem, Proceedings of the Capri Seminar.

[19] McShane, W.R. and Roess, R.P. (1990) *Traffic Engineering* Prentice Hall, Englewood Cliffs (NJ).

[20] Nagurney, A.B. (1984) Comparative Tests of Multimodal Traffic Equilibrium Methods, *Transpn. Res.* **18B**(6), 469-485.

[21] Netter, M. (1972) Affectations de Trafic et Tarification au Coût Marginal Social: Critique de Quelques Idées Admises. *Transpn. Res.* **6** 411–429.

[22] Ortuzar J.D. and Willumsen, L.G. (1990) *Modelling Transport* John Wiley & Sons, New York.

[23] Peeta, S and Mahmasani H.S. (1995) Multiple User-Classes Real-Time Traffic Assignment for On-Line ATIS/ATMS Operations: A Rolling Horizon Solution Framework, TRB Preprint 951066, Washington DC.

[24] Potts, R.B. and Oliver, R.M. (1972) *Flows in Transportation Networks*, Academic Press, New York.

[25] Sheffi, Y. (1985) Urban Transportation Networks, Prentice-Hall.

[26] Spiess, H. (1990) Conical Volume-Delay Functions, *Trans. Sci* **24**(2) 153–158.

[27] Stampacchia, G. (1969) Variational Inequalities, in Theory and Applications of Monotone Operators, Proceedings of the NATO Advanced Study Institute, 101–192.

[28] Van Vliet, D., Bergman, T., and Scheltes, W.H. (1986) Equilibrium Traffic Assignment with Multiple User Classes. *Proceedings of the PTRC Summer Annual Meeting.*

[29] Wardrop, J. (1952) Some theoretical aspects of road traffic research, Proceedings of the Institute of Civil Engineers, part II, volume 1, 325–378.

[30] Wynter, L. (1995) *Advances in the Theory and Application of the Multiclass Traffic Assignment Model*, unpublished Ph.D. thesis, Ecole Nationale des Ponts et Chaussées, Paris.