ROUTE CHOICE MODELLING FOR FREIGHT TRANSPORT AT NATIONAL LEVEL

Francesco Russo, Antonino Vitetta, Agata Quattrone
Mediterranea University of Reggio Calabria
Dept. of Computer Science, Mathematics, Electronics and Transportation

1. INTRODUCTION

Freight transport plays a fundamental role in the economy of every country. The performance of the European transport sector has been in line with the expanding economy. From 1995 to 2004 total European freight transport in the 25 Member States examined grew from 2,967,000 to 3,804,000 million tkm (+28 %). Considering only inland transport, it appears that the considerable growth was almost entirely realised by road transport (tab. 1).

<table>
<thead>
<tr>
<th>Year</th>
<th>Road</th>
<th>Rail</th>
<th>Inland waterways</th>
<th>Pipelines</th>
<th>Sea</th>
<th>Air</th>
<th>Total (1,000 million tkm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995</td>
<td>1,248</td>
<td>358</td>
<td>120</td>
<td>105</td>
<td>1,133</td>
<td>1.8</td>
<td>2,967</td>
</tr>
<tr>
<td>1996</td>
<td>1,268</td>
<td>360</td>
<td>116</td>
<td>111</td>
<td>1,140</td>
<td>1.9</td>
<td>2,997</td>
</tr>
<tr>
<td>1997</td>
<td>1,317</td>
<td>380</td>
<td>124</td>
<td>110</td>
<td>1,193</td>
<td>1.9</td>
<td>3,125</td>
</tr>
<tr>
<td>1998</td>
<td>1,386</td>
<td>370</td>
<td>127</td>
<td>117</td>
<td>1,220</td>
<td>2.0</td>
<td>3,222</td>
</tr>
<tr>
<td>1999</td>
<td>1,444</td>
<td>358</td>
<td>127</td>
<td>117</td>
<td>1,270</td>
<td>2.0</td>
<td>3,317</td>
</tr>
<tr>
<td>2000</td>
<td>1,491</td>
<td>374</td>
<td>132</td>
<td>119</td>
<td>1,345</td>
<td>2.1</td>
<td>3,463</td>
</tr>
<tr>
<td>2001</td>
<td>1,521</td>
<td>359</td>
<td>130</td>
<td>124</td>
<td>1,388</td>
<td>2.2</td>
<td>3,524</td>
</tr>
<tr>
<td>2002</td>
<td>1,563</td>
<td>358</td>
<td>129</td>
<td>121</td>
<td>1,404</td>
<td>2.1</td>
<td>3,577</td>
</tr>
<tr>
<td>2003</td>
<td>1,575</td>
<td>364</td>
<td>120</td>
<td>123</td>
<td>1,435</td>
<td>2.3</td>
<td>3,619</td>
</tr>
<tr>
<td>2004</td>
<td>1,684</td>
<td>379</td>
<td>130</td>
<td>124</td>
<td>1,484</td>
<td>2.5</td>
<td>3,804</td>
</tr>
</tbody>
</table>

1995 - 2004 per year: +35%  +6%  +9%  +18%  +31%  +39%  +28%
2003 - 2004 per year: +6.9%  +4.3%  +8.7%  +0.8%  +3.4%  +8.7%  +5.1%


In particular, there were approximately 4,000,000 commercial vehicles used for road freight transport in Italy registered in 2004. Therefore the study of the behaviour of truck-drivers’ path choice is topical.

In the literature, few behavioural path choice models have been specified, calibrated and validated with real experimentation for road freight transport. The problem of path choice has been treated in two phases:

- generation of the choice set, that is the possible alternatives (Ben Akiva et al., 1984; Antonisse et al., 1985; Morikawa, 1996; Russo and Vitetta, 1995);
- path choice among the alternatives belonging to the choice set (Ben Akiva et al., 1984; Antonisse et al., 1985; Cascetta et al., 1992; Cascetta et al. 1996; Ben-Akiva and Bierlaire, 1999; Russo and Vitetta, 2003).

Regarding the generation of the choice set the extended literature review is reported in section 2.2, while the path choice model is surveyed in section 2.3. Moreover, it is worth stressing that there is no general model completely specified and calibrated for road freight transport nationwide in Italy and that
existing analyses consider only travel time, as a determining factor in path choice utility, neglecting other factors that have a strong influence on user behaviour and hence are fundamental for the study of the problem. This work therefore proposes the specification, calibration and validation of a path choice behavioural model at national scale, on the basis of observation, carried out through a survey, of the choice behaviour of a sample of truck-drivers who used the Italian road network. The main aim is to characterize and analyze the multiple factors that influence truck-driver behaviour.

In this paper the path choice model is simulated in two phases:
1. choice set generation, that is the possible alternatives;
2. path choice among the alternatives belonging to the choice set.

The choice set generation is realized with multi-criteria path generation, that defines for each \((o,d)\) pair a choice set of several paths, each of which is generated optimizing a covered function between generated and chosen paths associated to a certain criterion (e.g. minimum travel time, maximum motorway use, minimum travel cost etc.).

Path choice is simulated with a Logit and C-Logit model. The C-Logit overcomes the main shortcoming of Multinomial Logit, i.e. unrealistic choice probabilities for paths sharing a number of links, while keeping a closed analytical structure allowing calibration on disaggregate data and efficient path flow computations when paths are explicitly or implicitly enumerated. In the DC Logit version (Russo and Vitetta, 2003) it is possible to use the model within an implicit assignment procedure.

Path choice models were specified and calibrated on a truck-drivers’ road-side survey. In all, 280 interviews were conducted for the purpose of path generation and path choice modelling. The chosen path was indicated in the questionnaire through origin, destination and intermediate nodes. Computations of path generation and level-of-service attributes were carried out using the Italian national road network which consists of all the motorways and the main national roads.

For validation, formal and informal tests were carried out on parameters and statistics to assess the model’s goodness of fit. In particular, it is verified that the model is able to reproduce user choices, all the calibrated parameters are significant and correct in sign and the model hypotheses are acceptable.

The paper is structured as follows: in section 2 the literature on path choice models is surveyed: we specify the factors that affect path choice behaviour and the two main phases of generating the choice set and choosing an alternative. Section 3 treats the main features of the proposed model in terms of generation and choice; section 4 gives the experimental results, obtained by applying the proposed model to a real road transport network at national scale; in section 5 some conclusive considerations are reported.

2. PATH CHOICE MODELS

In the specification of a transport system model, assignment models assume a very important role. They concur to simulate the interaction between supply and demand transport, to calculate user flows and the performance for every element of the supply system. The flows are the result of the choice behaviours of those who use the transport system, and in particular they
depend on path choice from the zone of origin to destination. The path choice model gives the probability of every path from those perceived as admissible. Given the trip characteristics of origin, destination, trip purpose, time slice and trip mode, the problem of path choice, for a user travelling between origin-destination pair \((o,d)\), consists in characterizing "the best" path with various criteria. The best path minimizes trip disutility, which existing assignment models in literature measure through single attributes, like travel time or distance, or through formulas of generalized travel cost. The problem of path choice is very complex due to the large number of existing alternatives between each \((o,d)\) pair, also on networks of modest dimensions, and their overlapping. These difficulties are treated in various studies (Ben Akiva et al., 1984; Russo and Vitetta, 1995; Cascetta et al., 1996) in which complete model specification of path choice is articulated in two phases:

1. choice set generation, that is the possible alternatives;
2. path choice among the alternatives belonging to the choice set.

In section 2.1 the factors that affect path choice behaviour are reported; the subsequent sections treat the path choice models found in the literature distinguishing the phase of generation of the available alternatives (section 2.2) from the phase of path choice among the alternatives characterized (section 2.3).

2.1 Factors affecting path choice behaviour

The main factors that influence user behaviour in path choice can be summarized in three categories:

1) user knowledge of the network and the available paths
   - users on long-distance trips tend to restrict their route to motorways as they have insufficient acquaintance of the local roads (Tagliacozzo et al., 1973; Russo and Vitetta, 1995);
   - users are not always allowed to estimate simple path characteristics, like time or distance, and therefore it does not happen as expected that they satisfy their aim of choosing the fastest or shortest path (Burrel, 1968; Dial, 1971, Russo and Vitetta, 1995);
   - users follow a "hierarchical" path that is developed on local roads close to the origin zone, continue on higher level roads for most of the trip and then revert to local roads (Cascetta et al., 1992).

2) decisional process
   - users choose their path before beginning the trip (pre-trip choice);
   - the decisions on the path to follow are taken during the trip and are independent of those made previously (en-route choice);
   - the decisional process is intermediate to the two previous cases.

3) path attributes and choice preferences
   - travel time;
   - distance;
   - number of stops and traffic-light regulated intersections;
   - scenic path attraction;
   - safety;
   - commercial development along the path;
   - congestion;
2.2 Choice set generation

There are different approaches to path choice set generation in the literature: (figure 1):

- **exhaustive approach**: all the analytical paths loop less on the network are available and belonging to the choice set that is only one for all users;
- **selective approach**: only some available paths represent attractive choice alternatives.

In the second case, the generation of available paths can be obtained following three different approaches: with the monocriterion approach, multi-criteria approach and a probabilistic model of belonging to the choice set.

Regarding the monocriterion approach the available paths are obtained by the satisfaction of a single criterion. In order to generate the minimum paths for the criterion it is necessary to minimize trip disutility measured with a single attribute (time, distance, number of traffic-light regulated intersections, etc.) or construct a covered function of the criterion that is a weighted sum of the travel time and the descriptive attribute, in order that the generated paths do not differ considerably from the minimum paths. The parameters of the covered function must be calibrated, maximizing, for example, the overlapping factor between the generated and the chosen paths, such that the choice set constructed actually includes the paths chosen by users.

Alternatively, it is possible to use the multi-criteria approach according to which the available paths are obtained by satisfying some criteria calibrated by maximizing the overlapping factor between the generated and the chosen paths, similar to the monocriterion approach.

![Fig. 1 – Approaches for choice set generation](image)

The main objective of the multicriteria approach is to define the choice set, for a specific origin-destination pair \((o,d)\), replacing the large number of physically available paths with few paths, called “label paths” (Ben Akiva et al., 1984), each of which represents the optimum with respect to a certain criterion. The
criteria used in literature (Ben Akiva et al., 1984; Antonisse et al., 1985; Russo and Vitetta, 1995) for choice set generation are:

- to minimize travel time;
- to minimize distance;
- to maximize the "view" along the path;
- to minimize the number of traffic-light regulated intersections;
- to maximize motorway use;
- to minimize the use of congested paths;
- to maximize the presence of commercial areas along the path;
- to travel on high-quality roads, in terms of traffic sign and paving quality, visibility and a small number of tunnels, bridges and viaducts along the path and low level of winding;
- to travel on high-capacity roads;
- to follow a "hierarchical" path that is developed on local roads close to the origin zone, continue on higher level roads for most of the trip and then revert to local roads for the remainder of the journey;
- to travel on safe roads.

The possible alternatives can also be obtained with a more rigorous approach from the behavioural point of view, which demands the specification and calibration of a probabilistic model of belonging to the choice set. As regards generation with probabilistic model of belonging to the choice set, two different approaches can be followed:

- implicit approach: the perception/availability of each alternative is simulated with a model that includes in the function of systematic utility the model used for path choice. It is assumed that each alternative can have various degrees of perception/availability in the interval [0, 1], such that the choice set is represented by a set of continuous variables that supply the degree of inclusion of each alternative in the choice set. This approach is used within the DC Logit model that allows implicit calculation of choice probabilities using a specific algorithm (Russo and Vitetta, 2003).
- explicit approach: choice set generation is achieved with an adjunctive model. This approach is not very widely used, since the models that derive from it are not so efficient in computational terms: the size of the choice set grows exponentially with the increase in possible alternatives (Antonisse et al., 1985; Morikawa, 1996), making it difficult to calculate the choice probabilities and calibrate the parameters.

2.3 Choice model

As regards path choice, most of the models proposed in the literature belong to the family of random utility models (RUMs). Based on assumptions on the random residual of the utility perceived by users, the models can have different specifications, among which those most commonly used for path choice are Multinomial Logit and Probit (Ben Akiva and Lerman, 1985; Sheffi, 1985; Ortuzar and Williansen, 1990; Russo and Vitetta, 1995; Cascetta, 2001). In the field of non-RUMs we may recall the Fuzzy model (Cantarella and Fedele, 2003). In the Logit model the random residuals related to the alternatives are assumed to be independent and identical Gumble variates of zero mean and
This model has the benefit of a closed analytical structure, allowing efficient calibration on disaggregate data. Usually the Logit is used in order to simulate path choice starting from a choice set generated with the selective approach and implicit (Dial, 1971) or explicit enumeration of paths (Ben Akiva et al., 1984; Antonisse et al., 1985; Cascetta et al., 1992).

However the use of a Logit model can lead to non-realistic choice probabilities for paths sharing a number of links (Florian and Fox, 1976; Daganzo and Sheffi, 1982). This stems from the property of independence of the irrelevant alternatives that underlies Logit models and that derives from the assumptions made on random residuals. The Probit model (Sheffi, 1985), that derives from the assumption that the random utility residuals of the alternatives are distributed as a Multivariate Normal of zero, takes into account the similarity among paths which have links in common by introducing a random residual covariance proportional to the cost attributes of shared links (Daganzo and Sheffi, 1982). Although this helps to overcome the shortcomings of the Logit model, it increases the complexity of the problem in analytical terms. The Probit model can indeed entail greater difficulties than the Logit in the explicit calculation of choice probabilities.

To calculate path choice a modified specification of the MNL, the C-Logit model, which overcomes the main shortcoming of the Logit, has been proposed (Cascetta et al., 1996; Bhat, 2002). The C-Logit model takes into account similarity among the alternatives by considering an additional cost attribute, called “commonality factor”, that reduces the systematic utility of a generic path according to its degree of similarity (or overlapping) with the alternative paths. The C-Logit overcomes the main shortcoming of MNL, i.e. unrealistic choice probabilities for paths sharing a number of links, while keeping a closed analytical structure, thereby allowing calibration on disaggregate data and efficient path flow computations when paths are explicitly enumerated.

The Path Size model (Ben-Akiva and Bierlaire, 1999; Frejinger and Bierlaire, 2006) is similar to C-logit in that a correction term, the Path Size (PS) attribute, is added to the deterministic part of the utility. However, PSL has a different theoretical basis. The notion of size comes from the theory of aggregate alternatives, which was first employed for destination and residence choice (Ben Akiva and Lerman, 1985). However, unlike destination choice, in which zones may have a size representing thousands of elemental destinations (e.g., jobs), the largest size a path may have is one. Such a path shares no links with other paths and may be called a distinct or disjoint path. The path-size term may be calculated on the basis of the length of links in a path and the relative lengths of the paths that share a link.

The C-Logit and the Path Size models entail explicit calculation of the choice probabilities. The DC Logit model (Russo and Vitetta, 2003) overcomes this limit since it allows implicit calculation of choice probabilities using a specific algorithm.

3. THE PROPOSED MODEL

The proposed model is used to analyze the behaviour of a sample of truck-drivers in terms of path choice on a road network at national level. The path choice model is simulated in two phases:
1. choice set generation, that is the possible alternatives;
2. path choice among alternatives included in the choice set.

Concerning the generation of the choice set, the followed selective approach with multicriteria, described in section 2.1, is used to define for every \((o,d)\) pair a choice set that consists of several paths, each of which is generated by optimizing a covered function associated to a certain criterion (i.e. minimum travel time, maximum motorway use, minimum travel cost etc.). The criteria used depend on the factors that affect the truck-drivers’ behaviour. The covered function is calibrated by maximizing the degree of overlapping of the chosen paths by the sample with the generated paths.

Regarding the path choice among the alternatives belonging to the generated set, the model used is the C-Logit, whose calibration is carried out with the Maximum Likelihood Method that gives the values of the parameters on which the utility of every alternative depends, maximizing the probability of observing the choices made by the user sample.

The proposed specification can be used with a DC Logit approach inside an assignment procedure.

In section 3.1 the factors that are assumed to affect the user sample in the specific case are described; in section 3.2 the generation process of the set of the possible alternatives is defined; in section 3.3 the specification and calibration of the path choice model proposed are treated.

### 3.1 Factors affecting path choice behaviour

In the phase of choice set generation and hence of the available alternatives, reference was made to previous findings reported in the literature with respect to the behaviour assumed by users in path choice and the factors that determine them. We took into account the fact that travel time and distance are not the only determining factors in the choice process: the user is often induced to choose a path for other reasons, such as trip cost, the level of road safety (i.e. users can choose a certain path as it has low road accident levels), the scenic attraction of the path or the commercial development along it (i.e users may prefer roads along which there are service areas, bars and restaurants, etc.), road quality (a path can be attractive due to low winding levels etc.).

Moreover, users on long distances are inclined to prefer the use of motorways, on which they are likely to have more information from personal experience and from any advanced information systems (e.g. VMS – Variable Message Signal), compared with the local road network of which knowledge may be scant.

Finally, meteorological conditions can affect path choice. At certain times of the year, users may prefer coastal and level roads, rather than winding roads that cross mountain zones affected by adverse meteorological conditions.

The main factors taken into account for choice set generation are:

- travel time;
- trip cost;
- total length of path and length on motorway;
- scenic attraction of path and commercial development along it;
- road quality;
- road safety;
− meteorological conditions.

3.2 Choice set generation

3.2.1 Specification

In order to define the path choice set the selective multicriteria approach was followed. Starting from the factors that were assumed to most influence user behaviour in the dimension of path choice, the criteria were defined and optimized to generate suitable paths for the choice set. These criteria \( h \) are:

1. minimum path travel time;
2. minimum path length;
3. minimum winding;
4. minimum motorway path with bridges and viaducts;
5. maximum motorway path with service areas;
6. maximum motorway path with parking areas;
7. maximum motorway path with bars and restaurants;
8. minimum path monetary cost;
9. maximum motorway path;
10. minimum path with high levels of road accidents;
11. minimum path with adverse meteorological conditions.

3.2.2 Calibration

For all the criteria, relative to every link \( i \) of the network, a covered function equal to the weighted sum of travel time and an attribute that characterizes the same criterion are defined. The function is dependent on unknown parameters that must be calibrated in order to maximize the degree of cover of the paths chosen with those generated (label paths).

The general structure characterized for the covered functions for every criterion \( h \) is:

\[
\begin{align*}
t_{ih} &= \beta_{Ah} \cdot A_i + \alpha_{Ah} \cdot Z_{ih} & \text{if the link is motorway} \\
t_{ih} &= \beta_{Si} \cdot A_i + \alpha_{Sh} \cdot Z_{ih} & \text{otherwise}
\end{align*}
\]

where

\[
\begin{align*}
\beta_{Ah}, \beta_{Si}, \alpha_{Ah}, \alpha_{Sh} &= \text{parameters calibrated for every criterion } h \\
Z_{ih} &= \text{attribute of every criterion } h
\end{align*}
\]

For some criteria (3–7) the value of this attribute is estimated with the following relation:

\[
Z_{ih} = \frac{V_{ih}}{L} \cdot L_i
\]

where:

\[
\begin{align*}
V_{ih} &= \text{number of tunnels } (N_{Gall}), \text{ bridges and viaducts } (N_{PV}), \text{ service areas } (N_{AS}), \text{ parking areas } (N_{Parch}) \text{ or bars and restaurants } (N_{Rist}), \text{ existing on the motorway to which link } i \text{ belongs, with respect to the type of specific criterion;} \\
L &= \text{total length of motorway to which link } i \text{ belongs;} \\
L_i &= \text{length of link } i.
\end{align*}
\]
Table 2 describes the parameters relative to the covered functions which have the above general structure.

<table>
<thead>
<tr>
<th>Criterion (h)</th>
<th>$\beta_A$</th>
<th>$\beta_S$</th>
<th>$A_i$</th>
<th>$\alpha_{Ah}$</th>
<th>$\alpha_{Sh}$</th>
<th>$Z_{ih}$</th>
<th>$V_{ih}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum path travel time</td>
<td>1</td>
<td>1</td>
<td>$t_{0i}$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Minimum path length</td>
<td>1</td>
<td>1</td>
<td>$L_i$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Minimum path winding</td>
<td>1</td>
<td>10</td>
<td>$t_{0i}$</td>
<td>to calibrate</td>
<td>$V_{ih}$</td>
<td>$L_i/L$</td>
<td>$N_{Gal}$</td>
</tr>
<tr>
<td>Minimum motorway path with bridges and viaducts</td>
<td>1</td>
<td>10</td>
<td>$t_{0i}$</td>
<td>to calibrate</td>
<td>$V_{ih}$</td>
<td>$L_i/L$</td>
<td>$N_{PV}$</td>
</tr>
<tr>
<td>Maximum motorway path with service areas</td>
<td>1</td>
<td>10</td>
<td>$t_{0i}$</td>
<td>to calibrate</td>
<td>$V_{ih}$</td>
<td>$L_i/L$</td>
<td>$N_{AS}$</td>
</tr>
<tr>
<td>Maximum motorway path with parking areas</td>
<td>1</td>
<td>10</td>
<td>$t_{0i}$</td>
<td>to calibrate</td>
<td>$V_{ih}$</td>
<td>$L_i/L$</td>
<td>$N_{Parch}$</td>
</tr>
<tr>
<td>Maximum motorway path with bars and restaurants</td>
<td>1</td>
<td>10</td>
<td>$t_{0i}$</td>
<td>to calibrate</td>
<td>$V_{ih}$</td>
<td>$L_i/L$</td>
<td>$N_{Rist}$</td>
</tr>
<tr>
<td>Minimum path monetary cost</td>
<td>1</td>
<td>10</td>
<td>$t_{0i}$</td>
<td>to calibrate</td>
<td>$C_{i,Carb}+Ped_i$</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Maximum motorway path</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>to calibrate</td>
<td>$t_{0i}$</td>
<td>-</td>
</tr>
<tr>
<td>Minimum path with high levels of road accidents</td>
<td>1</td>
<td>1</td>
<td>$t_{0i}$</td>
<td>to calibrate</td>
<td>$t_{0i}$</td>
<td>$l_i$</td>
<td>-</td>
</tr>
</tbody>
</table>

Note:
$t_{0i}$ = travel time calculated by using functions reported in the literature. The functions proposed by the Italian National Research Council (1983) were used for motorways, while TRRL functions (1980) were used for extra-urban roads. $P_{Ped_i}$ = monetary cost of tolls estimated as a product of the link's length times the specific unit cost per kilometre assumed equal to 0.05 €/km. $C_{i,Carb}$ = cost of fuel consumption (Russo, 2005). $l_i$ = number of incidents taking place in 2001 (the number of incidents per kilometre divided by the length in km of the road to which link $i$ belongs).

### 3.3 Choice model

#### 3.3.1 Specification

Path choice is simulated with a C-Logit model (Cascetta et al., 1996). A generic user $n$, travelling between an origin-destination pair $(o, d)$, associates to each path $k$ belonging to the set $I_{od}$ of available paths connecting that $(o, d)$ pair a perceived utility $U_k^n$ which may be expressed as:

$$U_k^n = V_k^n + \varepsilon_k^n \quad \forall k \in I_{od}$$

where

- $V_k^n$ is the systematic utility of path $k$
- $\varepsilon_k^n$ is the random residual usually assumed to include perception errors of the decision maker as well as modelling approximation of the analyst.

The C-Logit model, with respect to the Multinomial Logit model, introduces a modified systematic disutility as follows:

$$\tilde{V}_k^n = V_k^n - CF_k \quad \forall k \in I_{od}$$

Hence the path choice probabilities $p_{od}(k/n)$ can be expressed as:
The “Commonality Factor” \( CF_k \) of path \( k \), is directly proportional to the degree of similarity (or overlapping) of path \( k \) with other paths belonging to \( L_{od} \). Heavily overlapping paths have larger commonality factors and thus a smaller systematic utility (larger generalised cost) with respect to similar, but independent, paths. The commonality factor can be specified in different ways, giving rise to different path choice probabilities and, ultimately, different C-Logit model specifications.

The general structure of the commonality factor is:

\[
CF_k = \beta_0 \cdot \phi(L_{hk}, L_h, L_k)
\]

where

- \( \beta_0 \) is a calibration parameter
- \( L_{hk} \) is the “length” (generalized cost) of links common to paths \( h \) and \( k \)
- \( L_h \) and \( L_k \) are the overall “lengths” (sum of link lengths) of paths \( h \) and \( k \) respectively
- \( \phi \) is a continuous and monotone function that increases with \( L_{hk} \) and decreases with \( L_h \) and \( L_k \)

One possible way to specify the commonality factor is as follows:

\[
CF_k = \beta_0 \cdot \ln \sum_{i \in k} w_{ik} N_i
\]

where

- \( w_{ik} \) is the proportional weight of link \( i \) for path \( k \)
- \( N_i \) is the number of paths, connecting the same \((o,d)\) pair, which share the link \( i \).

Coefficients \( w_{ik} \) can be specified in different ways, expressing different hypotheses on the perceived relevance of an individual link in a path. One possibility is to assume \( w_{ik} \) as the fraction of total path “travel cost” which can be attributed to each link \( i \).

Terms \( N_i \) can be seen as the summation of the 0/1 elements of the link-path incidence matrix relative to link \( i \) and to all paths connecting the \((o,d)\) pair:

\[
N_i = \sum_{h \in L_{od}} a_{ih}
\]

### 3.3.2 Calibration

Calibration of the model consists in obtaining the estimation of the parameters on which the same model depends starting from the choices made by the user sample. The C-Logit model can be calibrated with the classic method of ML - Maximum Likelihood, (Ben Akiva and Lerman, 1985). This method supplies the values of the unknown parameters that maximize the probability of observing the choices made by the users.

If link “lengths” or weights used in \( CF_k \) specifications are expressed through link generalised costs including unknown coefficients, the utility function is
non-linear with respect to coefficients, and specialised ML models and algorithms have to be used.

4 EXPERIMENTATION

The proposed model, described in the previous section, is specified, calibrated and validated for a road truck users’ segment, studied with regard to route choice on a road transport network at national scale. The data base is described in section 4.1, the choice set generation method is presented in section 4.2 and the path choice model specification, calibration and validation results are reported in section 4.3.

4.1. The data base

Path choice models were specified and calibrated on a truck-drivers’ road-side survey. In all, 280 interviews were held for modelling path generation and path choice. The surveys, carried out by the Transport System Analysis Laboratory (LAST), were conducted partly in Sicily and partly in Calabria, both regions in southern Italy, in the period between September 1998 and December 1999. The chosen path was indicated in the questionnaire by means of a map. Computations of path generation and level-of-service attributes were carried out using the national road network, which consists of all the motorways and the main roads (the network consists of 4480 nodes and 16029 links). The data² for each network link concern the physical and functional characteristics (i.e. length, width and number of lanes, capacity, type of link, slope etc.). Once the number of service areas, galleries, bridges and viaducts, parking areas, bars and restaurants on the motorway network are known, the attributes $Z_{ik}$, required to specify the covered functions, were calculated as described in section 3.2.

4.2. Choice set generation

The set of possible alternatives was defined on the basis of the satisfaction of some criteria, among these defined in the previous section, calibrated by maximizing the overlapping factor between the generated and the chosen paths. The latter were obtained with software, developed in the LAST of the Mediterranea University of Reggio Calabria, which generates, by optimizing one determined covered function, a set number of paths for each $(o,d)$ pair and calculates the percentage $\lambda$ of similarity among the paths stated by the users in the survey and those generated, with the relation:

$$\lambda = \frac{2L_c}{L_d + L_g}$$

where

$L_c$ = length of link shared by both paths (stated and generated) relative to the same $(o,d)$ pair
$L_d$ = total length of the stated path
$L_g$ = total length of the generated path
The overlapping of a stated path was assumed when its similarity reached $\lambda = 90\%$ with the generated path. Considering five of the criteria specified in section 3.2 (tab. 3) we obtained, through the first and second path generated, a overlapping for 216 of the 280 stated paths.

<table>
<thead>
<tr>
<th>Criterion (h)</th>
<th>Covered Paths</th>
<th>$\alpha_{sh}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Minimum path travel time</td>
<td>141</td>
<td>-</td>
</tr>
<tr>
<td>4. Minimum motorway path with bridges and viaducts</td>
<td>50</td>
<td>12 min</td>
</tr>
<tr>
<td>8. Minimum monetary cost</td>
<td>2</td>
<td>10 min/€</td>
</tr>
<tr>
<td>9. Maximum motorway path</td>
<td>6</td>
<td>1.5</td>
</tr>
<tr>
<td>10. Minimum path with high levels of road accidents</td>
<td>17</td>
<td>$\alpha = 1$ on motorways $\alpha = 2$ otherwise</td>
</tr>
</tbody>
</table>

Total covered paths | 216 |
Total stated paths | 280 |
Total non-covered paths | 64 |

### 4.3 Choice model

Having determined the choice set for each (o,d) pair, we defined the model of path choice among the possible alternatives. The model used is the Logit and C-Logit described in section 3.3 in whose functional form the systematic utility $V^*_k$ associated to the generic alternative $k$ is a linear combination of the attributes $X_i$:

$$V^*_k = \sum_{i} \beta_i X_{ik}$$

The attributes considered for calibration are:
- path travel time ($TT$);
- monetary cost ($MC$), equal to the sum of the fuel and toll cost;
- the length on motorway ($ML$);
- a vector of specific label variables ($X_i$).

The overall set of label variables used is:
- minimum travel time (MinT);
- minimum motorway path with bridges and viaducts (MinB&V);
- minimum monetary cost (MinC);
- maximum motorway path length (MaxM);
- minimum path with high levels of road accidents (MinA)

Preliminary calibrations were carried out using the maximum Likelihood method.

Table 4 reports the calibrated parameters and the relative statistical indicators, considering for the five criteria the variables of specific labels. Calibrated models include models using only level-of-service attributes and models with
the addition of label variables. Two different models were proposed: a model without labels and one with four labels. In all cases the models were specified with and without the $CF_k$ attribute. Specification of the $CF_k$ attribute was adopted as the similarity of generalised costs of enumerated paths, as described in section 3.3.1.

Tab 4 – Calibration results for the Logit and C-Logit model

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Unit of measurement</th>
<th>Without $CF_k$</th>
<th>With $CF_k$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>TT</td>
<td>[min]</td>
<td>-0.153</td>
<td>-0.164</td>
<td>-0.039</td>
<td>-0.429</td>
<td>(-2.9)</td>
<td>(-3.1)</td>
</tr>
<tr>
<td>MC</td>
<td>[€]</td>
<td>-0.012</td>
<td>-0.015</td>
<td>-0.021</td>
<td>-0.025</td>
<td>(-9.2)</td>
<td>(-8.3)</td>
</tr>
<tr>
<td>ML</td>
<td>[km]</td>
<td>-0.025</td>
<td>-0.046</td>
<td>-0.019</td>
<td>-0.028</td>
<td>(-7.6)</td>
<td>(-7.2)</td>
</tr>
<tr>
<td>CF_k</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-1.154</td>
<td>-1.466</td>
<td>(-5.1)</td>
<td>(-4.0)</td>
</tr>
<tr>
<td>MinT</td>
<td>0/1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MinB&amp;V</td>
<td>0/1</td>
<td>-</td>
<td>2.449</td>
<td>-</td>
<td>1.717</td>
<td>(4.5)</td>
<td>(3.6)</td>
</tr>
<tr>
<td>MinC</td>
<td>0/1</td>
<td>-</td>
<td>-3.739</td>
<td>-</td>
<td>-2.639</td>
<td>(-5.8)</td>
<td>(-4.5)</td>
</tr>
<tr>
<td>MaxM</td>
<td>0/1</td>
<td>-</td>
<td>1.135</td>
<td>-</td>
<td>0.388</td>
<td>(2.7)</td>
<td>(1.1)</td>
</tr>
<tr>
<td>MinA</td>
<td>0/1</td>
<td>-</td>
<td>-2.655</td>
<td>-</td>
<td>-2.453</td>
<td>(-5.1)</td>
<td>(-5.1)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Final value of likelihood $\rho^2$</th>
<th>-225.081</th>
<th>-191.392</th>
<th>-199.535</th>
<th>-187.371</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho^2$</td>
<td>0.328</td>
<td>0.428</td>
<td>0.404</td>
<td>0.440</td>
</tr>
</tbody>
</table>

Note:
Initial value of likelihood = -334.763; Total number of observations = 208; (t) = Value of t-Student variable

The global indicators of calibration provide some optional goodness-of-fit values for the models, with parameter $\rho^2$ between 0.43 and 0.44 if there are specific variables, and between 0.33 and 0.40 if such variables are not considered. Moreover, from analysis of the data reported in the table, it emerges that all the calibrated parameters are correct in sign. The $CF_k$ coefficients are highly significant, with the expected negative sign and close to the desirable value of (-1). Furthermore, t-Student and Likelihood Ratio Tests reject the null hypothesis of zero $\beta_0$ coefficients in all cases.

It also emerges that the inclusion of the $CF_k$ variable generally increases the value of the $\beta$ coefficient relative to travel time. It is also evident from the results that inclusion of label variables increases the goodness of fit of the calibrated models without significantly modifying the values of the $\beta_0$ coefficient.
5 CONCLUSIONS

In this paper the problem of truck-drivers’ path choice at a nation-wide scale was treated. Path choice behavioural models were specified and calibrated for the Italian road transport system based on a truck-drivers’ road-side survey. These calibrations supplied some valid results and provide sound indications on the user path decisional process in the extra-urban context. Preliminary calibrations for the Italian national network also show that significant improvements can be obtained by adding the commonality factor and label variables accounting for path perception. Moreover, alternative specifications of $CF_i$ should be compared with calibration results.

In the future, the model should be further analysed through the use of a more extensive database. In order to obtain more information about user path choice behaviour, the vehicles of some truck-drivers will be equipped with on-board Intelligent Transportation Systems (ITS) for monitoring good transport. The utilization of ITS supplies new real-time data about path choice. The database obtained from the truck-drivers’ road-side survey will be updated with these new data. Finally, the proposed model and its further advances will be compared with other similar models (i.e. Path Size and Probit models for RUM, and Fuzzy model for non-RUM).

REFERENCES


© Association for European Transport and contributors 2006


NOTES
1Source: ACI – Automobile Club d’Italia
2Source: ANAS S.p.A. - Ente nazionale per le strade; AISCAT - Associazione Italiana Società Concessionarie Autostrade e Trafori.