A FIXED-POINT ROUTE CHOICE MODEL FOR ROUTE CORRELATION

Louis de Grange (corresponding author)
Industrial Engineering Department
Diego Portales University
Phone: (56-2) 676 2400, Postal Address: Ejercito 441, Santiago, Chile
e-mail: **louis.degrange@udp.cl**

Sebastián Raveau
Department of Transport Engineering and Logistics
Pontificia Universidad Católica de Chile
Phone: (56-2) 354 4270, Postal Address: Vicuña Mackenna 4860, Santiago, Chile
e-mail: **sraveau@puc.cl**

Felipe González
Industrial Engineering Department
Diego Portales University
Phone: (56-2) 676 8119, Postal Address: Ejercito 441, Santiago, Chile,
e-mail: **felipe.gonzalezr@udp.cl**

ABSTRACT

In this paper we present a stochastic route choice model for transit networks, which explicitly addresses route correlation due to overlapping in the alternatives. The model is based on a multi-objective mathematical programming problem, which optimality conditions generate an extension to the Multinomial Logit models. The model proposed considers a fixed point problem for treating correlation between routes, which can be solved iteratively. We estimated the new model on the Santiago (Chile) Metro network and compared the results with other route choice models that can be found in literature. The new model has better explanatory and predictive power, correctly capturing the correlation factor. Our methodology can be extended to private transport networks.

Keywords: route choice, fixed point, route overlapping, route correlation.

1. INTRODUCTION

Route assignment and route choice models for transport networks, widely used in urban planning, can be divided into three major groups: deterministic equilibrium models, stochastic assignment models and dynamic assignment models.

The deterministic equilibrium group usually state optimality conditions such as minimizing transport costs or satisfying Wardrop's first principle of traffic equilibrium (Wardrop, 1952). These models assume that travelers have perfect information and seek to unilaterally minimize their travel costs (Beckmann *et al*., 1956; Dijkstra, 1959; De Cea and Fernández, 1993; De Grange and Muñoz, 2009). Typically, a mathematical programming model is formulated and solved by an iterative algorithm. If applied with care and understanding, a deterministic user-equilibrium model provides a simple but effective method of traffic assignment (Dafermos and Sparrow, 1969; Florian, 1976; Florian, 1977; Dafermos, 1980; Patriksson, 1998; Florian and Hearn, 2001; Boyce *et al*., 2005).

Stochastic or probabilistic assignment models differ from deterministic formulations in that they incorporate uncertainty, randomness and/or heterogeneity of travelers and alternative routes. Reviews of this class of models are found in Daganzo and Sheffi (1977), Hazelton (1998), Ramming (2001), and Prashker and Bekhor (2004). They are an extension of the deterministic equilibrium group.

Dynamic assignment models have little relation to the subject of the present article and will not be described here. Extensive references and developments in dynamic transportation network modeling, analysis, and computational methods are given in Ran and Boyce (1996), with further reviews in Boyce *et al*. (2001) and Szeto and Lo (2006).

The present study formulates a new route choice model for public transport networks that features significant innovations on existing models. The main enhancement is the proposed model's ability to simultaneously and explicitly integrate the traveler's lack of information (randomness or uncertainty) and the correlation between route alternatives (overlapping). The formulation is entropy-based with linear and quadratic constraints, and is extensible to networks with congestion (e.g. private transport).

In the remainder of this paper, Section 2 contains a brief review of the literature that provides context for understanding the proposed model; Section 3 sets out an analytic derivation of the new formulation; Section 4 applies the model to a medium-sized network (the Santiago Metro), comparing the results with existing models and specifying a version for private networks with congestion; and finally, Section 5 summarizes the results and presents the main conclusions.

2. LITERATURE REVIEW

The formulation of the route choice or assignment stage in transportation modeling has long followed a paradigm in which users minimize their generalized trip cost on the assumption of perfect knowledge of the transport network. Under this approach, travelers are considered to be homogeneous and each one is fully informed of the cost of each arc on the network at any level of flow (Wardrop, 1952; Beckmann *et al.*, 1956). Clearly, these assumptions are both rather strong even for a small network and the results obtained are often less than satisfactory. Yet thanks to their simplicity and availability many transport planners continue to apply such models, especially with large networks.

Since the perfect information assumption is not totally correct, there is clearly a need for models that represent users who have incomplete or imperfect information on the transport system as regards existing routes and their levels of congestion. Various route choice models in the specialized literature are based on system attributes perceived by travelers and their socioeconomic and demographic characteristics (Dial, 1971; Daganzo and Sheffi, 1977; Bell, 1995; Ramming, 2001; Prashker and Bekhor, 2004). In these models, users behave in accordance with costs as they perceive them. The socioeconomic and demographic variable data are usually obtained through user surveys or from network data records and are easily justified as integral parts of individuals' rational decision-making processes. However, because the modeler lacks information on these processes, the choices as modeled necessarily embody a degree of variability. To explain this variability it would thus be of great interest to incorporate additional information.

As regards knowledge of routes, it is widely accepted that individual users do not know (or do not perceive) all the route possibilities between a given origin-destination pair (Ben-Akiva *et al*., 1984; Cascetta *et al*., 1996; Cascetta *et al*., 1998; Cascetta *et al*., 2002; Ben-Akiva and Bierlaire, 2003). This is a reality that should be incorporated into route choice models. Cascetta *et al*. (2002) propose a logit-type model of route perception and choice similar to those based on random utility theory. Ben-Akiva *et al*. (1984) models interurban route choice as a two-stage process in which a set of route choices are defined in the first stage and the choice is made in the second.

Correlation between alternative routes can to a certain extent be indirectly addressed, according to Bovy and Hoogendoorn-Lanser (2005), by hierarchical nested logit and multi-nested GEV models in cases where it arises from overlapping segments and/or nodes. The way in which individuals construct their set of route alternatives and the implications of route similarity for user behavior are analyzed by Prato and Bekhor (2007) and Bekhor *et al*. (2006). Bliemer and Bovy (2008) study the interdependence of these two aspects.

The probability of choosing route p to travel between O-D pair w $\left(P_w^p\right)$ can be estimated by discrete choice models in which the cost of a given route as perceived by a user (C_w^p) is assumed to be given by

$$C_w^p = c_w^p + \varepsilon_w^p \tag{1}$$

Cost is thus perceived as the sum of a deterministic component (c_w^p) and a random error component (ε_w^p). The latter can be interpreted in various ways, one of which is that the error reflects the user's inaccurate perception of route cost due, among other things, to the scarcity of information. In this case, the first term (c_w^p) represents the real mean cost of the route. The expression for the choice probability (P_w^p) is a function of the assumed distribution of the error terms and whether or not they are independent. If we assume the errors are i.i.d. Gumbel-distributed with a scale parameter $\theta > 0$, we have a Logit choice model in which P_w^p is given by

$$P_w^p = \frac{\exp\left(-\theta c_w^p\right)}{\sum_{r \in p^w} \exp\left(-\theta c_w^r\right)} \tag{2}$$

where p^w is the set of routes uniting the O-D pair w.

An equivalent optimization problem that generates the stochastic assignment model (2) is

$$\begin{aligned} &\min_{\{h_w^p\}} Z = \sum_w \sum_{p \in p^w} c_w^p h_w^p + \frac{1}{\theta} \sum_w \sum_{p \in p^w} h_w^p \left(\ln h_w^p - 1\right) \\ &s.t. \\ &\sum_{p \in p^w} h_w^p = T_w \qquad\qquad \forall w \end{aligned} \tag{3}$$

where T_w is the total number of trips (exogenous demand) between O-D pair w and h_w^p is the flow along route p $\left(P_w^p = h_w^p / T_w\right)$.

In problem (3) the set of routes p^w between each pair w must be predefined and maintained invariant during the assignment process. The optimality conditions lead to the assignment criterion given by (2) in which users divide up among the alternative routes according to a logit model.

The principal limitation of this model is that it does not incorporate a structure capturing correlation between routes. Possible direct extensions are either extremely simplistic or difficult to apply correctly to real-world scale networks (for example, using a hierarchical logit model) given that they cannot properly capture the various types of correlation between routes that share arcs with one other in different ways. In urban networks, the routes linking a given O-D pair will typically have many overlaps due to common arcs, with the result that the independent error assumption implicit in logit-type models such as (2) is unrealistic.

In Cascetta *et al*. (1996) and Cascetta *et al*. (2002) the authors propose a joint implicit availability/perception (IAP) and route choice (C-logit) model which also explicitly addresses the correlation issue (i.e., the lack of route independence due to common arcs) and is analytically tractable even for large-scale networks.

The basic idea behind the model is to handle route interdependence via a cost attribute called the "similarity factor" that is added to the cost of the route in a conventional logit model, instead of dealing with it in terms of error non-independence as does the probit model. The probability of choosing route p is then given by

$$P_w^p = \frac{\exp\left(-\theta c_w^p + CF_w^p\right)}{\sum_{r \in p^w} \exp\left(-\theta c_w^r + CF_w^r\right)} \tag{4}$$

where CF_w^p is the "similarity factor" for route *p* joining pair *w* and is constructed as follows:

$$CF_w^p = \beta \cdot \sum_{a \in p} \left(\frac{l_a}{L_p} \cdot \sum_{r \in p^w} \delta_{ar} \right) \tag{5}$$

where β is a parameter to be calibrated and must be negative, l_a is the length of arc *a*, L_p is the length of route *p*, and δ_{ar} is equal to 1 if arc *a* belongs to some route *r* joining *w* but 0 otherwise. Other specifications for CF_w^p may be found in Prato (2009).

Ben-Akiva and Ramming (1998) develop a model denoted path-size logit (PSL) that also aims to correct for routes which have overlaps and are therefore correlated. It attempts to incorporate behavioral theory in Cascetta's C-Logit model. In this case P_w^p is

$$P_w^p = \frac{\exp\left(-\theta c_w^p + \beta \cdot \ln PS_w^p\right)}{\sum_{r \in p^w} \exp\left(-\theta c_w^r + \beta \cdot \ln PS_w^r\right)} \tag{6}$$

where PS_w^p is the correction for route size.

The correction principle applied in this model is as follows. A route with no arcs overlapping another route needs no correction and is therefore assigned a size of 1. At the other extreme, if there are *J* duplicate routes (i.e., total overlap), each one has a size of 1/*J*. Finally, the length of routes with partial overlap is based on the sizes of the arcs, which are appropriately weighted on some criterion such as the arc's contribution to the total length of the route. Thus, PS_w^p can take the following form:

$$PS_w^p = \sum_{a \in p} \left(\frac{l_a}{L_p} \cdot \frac{1}{\sum_{r \in p^w} \delta_{ar}} \right) \tag{7}$$

where the variables are the same as those employed in (5) to define the similarity factor. Further specifications for PS_w^p may be found in Bovy *et al.* (2008).

Finally, Frejinger and Bierlaire (2007) model the correlation structure using a subnetwork, defined as a simplification of the road network that contains only roads that are easily identifiable and behaviorally relevant.

In the following section we develop a new route choice model that simultaneously incorporates (i) users with imperfect knowledge of the network, (ii) correlation between alternatives (in this case, routes), and (iii) the effect of demand or flow levels on network costs (congestion). As noted earlier, the proposed formulation is entropy-based with quadratic constraints. A variant of the model is presented in Section 4.4 to satisfy (iii).

3. ROUTE CHOICE MODEL WITH CORRELATED ROUTES

3.1 Mathematical formulation of model

The proposed stochastic equilibrium assignment model with capture of route correlation is based on the following multi-objective optimization problem:

$$\begin{aligned}
&\min \quad F_1 = \sum_{w} \sum_{p \in p^w} c_w^p h_w^p \\
&\min \quad F_2 = \sum_{w} \sum_{p \in p^w} h_w^p \left(\ln h_w^p - 1 \right) \\
&\min \quad F_3 = \sum_{w} \sum_{p \in p^w} \sum_{\substack{q \in p^w \\ q \neq p}} \left[\eta_w^{pq} \left(h_w^p - t_w \right) \left(h_w^q - t_w \right) \right] \\
&s.t.: \sum_{p \in p^w} h_w^p = T_w, \quad \forall w \qquad \qquad (\gamma_w)
\end{aligned} \tag{8}$$

Objective F_1 relates to the total system cost. Objective F_2 attempts to maximize entropy in order to determine the most likely routes; in probabilistic terms, it finds the most feasible route combinations for travelers in equilibrium. Combining F_1 and F_2 with the flow conservation constraints gives the stochastic assignment model expressed in (2).

Objective F_3, the novel element in the proposed model, explicitly incorporates correlation of flows between different routes, whether or not they join the same O-D pair w. The objective is constructed as a weighted sum of the divergences of the flows on the individual defined routes joining w from the average flow on those routes, where $t_w = \frac{T_w}{N_w}$ is the average flow, N_w is the number of defined routes (i.e. the cardinality of p^w) and T_w is the (fixed) total number of trips between w. Parameters η_w^{pq} are exogenous and determine the degree of correlation (0 to 1) between routes p and q of w. The values of the parameters can be defined in a number of ways (see Cascetta *et al.*, 1996; Yai *et al.*, 1997; Ramming, 2001).

As with F_2 , objective F_3 is an information criterion (Golan, 2002). If, for example, the flows on all traveled routes were uniform (that is, if $h_w^p = t_w, \forall p$), the value of F_3 would be 0 and thus contain no information. F_2 also takes its lowest possible value if $h_w^p = t_w, \forall p$.

A substitute optimization problem (Marler and Arora, 2004; De Cea *et al.*, 2008) for (8) that generates the proposed stochastic equilibrium model is the following:

$$\begin{aligned}
&\min \quad F = \sum_{w} \sum_{p \in p^w} c_w^p h_w^p + \frac{1}{\theta} \sum_{w} \sum_{p \in p^w} h_w^p \left(\ln h_w^p - 1 \right) + \frac{1}{\rho} \sum_{w} \sum_{p \in p^w} \sum_{\substack{q \in p^w \\ q \neq p}} \left[\eta_w^{pq} \left(h_w^p - t_w \right) \left(h_w^q - t_w \right) \right] \\
&s.t.: \sum_{p \in p^w} h_w^p = T_w, \quad \forall w \qquad \qquad (\gamma_w)
\end{aligned} \tag{9}$$

The terms $1/\theta$ and $1/\rho$ are the respective relative weights of the two information criteria F_2 and F_3 with respect to the reference objective F_1. θ and ρ are parameters to be estimated.

The first-order conditions for (9) are

$$c_w^p + \frac{1}{\theta} \ln h_w^p + \frac{1}{\rho} \sum_{\substack{q \in p^w \\ q \neq p}} \left[\eta_w^{pq} \left(h_w^q - t_w \right) \right] + \gamma_w = 0 \tag{10}$$

where c_w^p could, for a public transport application, be replaced by $L_w^p = \sum_k \beta_k X_{w,k}^p$.

From (10) we get

$$h_w^p = \exp\left(-\theta c_w^p - \frac{\theta}{\rho} \sum_{\substack{q \in p^w \\ q \neq p}} \left[\eta_w^{pq} \left(h_w^q - t_w \right) \right] - \theta \gamma_w \right) \tag{11}$$

$$\sum_{p \in p^w} h_w^p = T_w = \exp\left(-\theta \gamma_w\right) \cdot \sum_{p \in p^w} \exp\left(-\theta c_w^p - \frac{\theta}{\rho} \sum_{\substack{q \in p^w \\ q \neq p}} \left[\eta_w^{pq} \left(h_w^q - t_w \right) \right] \right) \tag{12}$$

Dividing (11) and (12) we have

$$h_p^w = T_w \frac{\exp\left(-\theta c_w^p - \frac{\theta}{\rho} \sum_{\substack{q \in p^w \\ q \neq p}} \left[\eta_w^{pq} \left(h_w^q - t_w \right) \right] \right)}{\sum_{r \in p^w} \exp\left(-\theta c_w^r - \frac{\theta}{\rho} \sum_{\substack{q \in p^w \\ q \neq p}} \left[\eta_w^{rq} \left(h_w^q - t_w \right) \right] \right)} \tag{13}$$

Since η_w^{pq} is an exogenous model parameter (defined by the modeler) and t_w is assumed to be constant (for calibration and modeling purposes), the term $\frac{\theta}{\rho} \sum_{\substack{q \in p^w \\ q \neq p}} \left[\eta_w^{pq} t_w \right]$ is also constant. Letting $\frac{\theta}{\rho} \sum_{\substack{q \in p^w \\ q \neq p}} \left[\eta_w^{pq} t_w \right] = \alpha_w^q$, an intercept or modal constant, (13) can then be rewritten as

$$h_w^p = T_w \frac{\exp\left(\alpha_w^p - \theta c_w^p - \frac{\theta}{\rho} \sum_{\substack{q \in p^w \\ q \neq p}} \left[\eta_w^{pq} h_w^q \right] \right)}{\sum_{r \in p^w} \exp\left(\alpha_w^r - \theta c_w^r - \frac{\theta}{\rho} \sum_{\substack{q \in p^w \\ q \neq p}} \left[\eta_w^{rq} h_w^q \right] \right)} \tag{14}$$

This non-linear expression is similar in structure to the model specified by Ben-Akiva and Ramming (1998), given here as (6). The main difference is that the right-hand side of (14) includes the endogenous variable h_w^q and is therefore a fixed-point function whereas in (6), the right-hand side contains only the model's exogenous variables.

A substitute can be specified for F_1 that can model service levels in public transport networks. An example of such a replacement is $\tilde{F}_1 = \sum_w \sum_{p \in p^w} L_w^p h_w^p$ where $L_w^p = \sum_k \beta_k X_{w,k}^p$ is a generalized cost given by the weighted sum of attributes or explanatory variables $X_{w,k}^p$ that could represent, for example, the trip time, wait time, cost, etc. of route p between pair w.

By the route choice probability function in (14), the marginal utility of the attribute $X_{w,k}^p$ can be written according to (15), where V_w^p is the utility of route p joining pair w.

$$\frac{\partial V_w^p}{\partial X_{w,k}^p} = \beta_k + \frac{\partial V_w^p}{\partial X_{w,k}^p} \cdot \frac{1}{T_w} \cdot \frac{\theta}{\rho} \sum_{\substack{q \in p^w \\ q \neq p}} \eta_w^{pq} \cdot h_w^p \cdot h_w^q \tag{15}$$

Clearing, we get

$$\frac{\partial V_w^p}{\partial X_{w,k}^p} = \frac{\beta_k}{1 - \frac{1}{T_w} \cdot \frac{\theta}{\rho} \sum_{\substack{q \in p^w \\ q \neq p}} \eta_w^{pq} \cdot h_w^p \cdot h_w^q} \tag{16}$$

Since the denominator of (16) is independent of attribute k, the marginal rates of substitution are generic and have the same functional form as traditional discrete choice models.

$$\frac{\partial V_r \big/ \partial X_{w,k_1}^p}{\partial V_r \big/ \partial X_{w,k_2}^p} = \frac{\beta_{k_1}}{\beta_{k_2}} \tag{17}$$

Thus, in the proposed fixed-point model with spatial correlation, the marginal rates of substitution between attributes can be obtained without any additional complexity.

To implement model (14), the parameters $\left(\alpha_w^p, \theta, \rho^*\right)$, where $\rho^* = \frac{\theta}{\rho}$, must first be estimated and the fixed-point equation (14) then solved given that the variable h_w^p appears on both sides of the equation.

However, since the endogenous variable (h_w^p) appears on the right-hand side of the model, the parameters $\left(\alpha_w^p, \theta, \rho^*\right)$ cannot be estimated using maximum likelihood because the presence of the endogenous variable violates the assumption of independent marginal probability functions necessary for defining the likelihood function.

To get around this problem, we resort to the use of an instrumental variable to replace h_w^p as explained below.

3.2 Estimation of model parameters

An instrument or instrumental variable, (Greene, 2008) is an exogenous variable that is highly correlated with an explanatory variable exhibiting endogeneity, and can therefore be used as a replacement for the latter without loss of asymptotic properties in the estimated parameters. In our case, a suitable instrument h_w^{p0} to stand in for h_w^p is

$$h_w^{p0} = T_w \frac{\exp\left(-\lambda c_w^p\right)}{\sum_{r \in p^w} \exp\left(-\lambda c_w^r\right)} \tag{18}$$

This formula is a classic multinomial logit model and can be estimated easily. Once this is done, the value of h_w^{p0} is substituted into the right-hand side of (14):

$$h_w^p = T_w \frac{\exp\left(\alpha_w^p - \theta c_w^p - \rho^* \sum_{\substack{q \in p^w \\ q \neq p}} \left[\eta_w^{pq} h_w^{q0}\right]\right)}{\sum_{r \in p^w} \exp\left(\alpha_w^r - \theta c_w^r - \rho^* \sum_{\substack{q \in p^w \\ q \neq p}} \left[\eta_w^{rq} h_w^{q0}\right]\right)} \tag{19}$$

Unlike (14), the parameters $\left(\alpha_w^p, \theta, \rho^*\right)$ in (19) can be estimated directly by maximum likelihood (Raveau *et al*., 2011) since h_w^{q0} does not exhibit endogeneity. Using these estimates, which we now denote $\left(\alpha_w^{p0}, \theta^0, \rho^{0*}\right)$, we can estimate h_w^{p1} by

$$h_w^{p1} = T_w \frac{\exp\left(\alpha_w^{p0} - \theta^0 c_w^p - \rho^{0*} \sum_{\substack{q \in p^w \\ q \neq p}} \left[\eta_w^{pq} h_w^{q0}\right]\right)}{\sum_{r \in p^w} \exp\left(\alpha_w^{r0} - \theta^0 c_w^r - \rho^{0*} \sum_{\substack{q \in p^w \\ q \neq p}} \left[\eta_w^{rq} h_w^{q0}\right]\right)} \tag{20}$$

The original model (14) is thus estimated iteratively from the following recursive relation (which represents the equilibrium of the fixed-point function):

$$h_w^{p(n+1)} = T_w \frac{\exp\left(\alpha_w^{p(n)} - \theta^{(n)} c_w^p - \rho^{(n)*} \sum_{\substack{q \in p^w \\ q \neq p}} \left[\eta_w^{pq} h_w^{q(n)} \right] \right)}{\sum_{r \in p^w} \exp\left(\alpha_w^{r(n)} - \theta^{(n)} c_w^r - \rho^{(n)*} \sum_{\substack{q \in p^w \\ q \neq p}} \left[\eta_w^{rq} h_w^{q(n)} \right] \right)} \tag{21}$$

If the estimator of parameter $\rho^{(n)*}$ in (21) is statistically different from 0, the null hypothesis of no correlation between route *p* and any of the other routes joining O-D pair *w* is rejected.

In analogous fashion to the derivation of (15), upon iteratively solving model (14) the marginal utility of attribute $X_{w,k}^{p}$ is then

$$\frac{\partial V_w^{p(n+1)}}{\partial X_{w,k}^{p}} = \beta_k^{(n)} + \frac{\partial V_w^{p(n)}}{\partial X_{w,k}^{p}} \cdot \frac{1}{T_w} \cdot \rho^{(n)*} \sum_{\substack{q \in p^w \\ q \neq p}} \eta_w^{pq} \cdot h_w^{p(n)} \cdot h_w^{q(n)} \tag{22}$$

This expression generates the marginal utilities recursively in successive iterations. The model converges to the equilibrium of the fixed-point function as follows:

$$\frac{\partial V_w^{p(n+1)}}{\partial X_{w,k}^{p}} \approx \frac{\partial V_w^{p(n)}}{\partial X_{w,k}^{p}} \qquad \rightarrow \qquad \frac{\partial V_r \big/ \partial X_{w,k_1}^{p(n+1)}}{\partial V_r \big/ \partial X_{w,k_2}^{p(n+1)}} = \frac{\beta_{k_1}^{(n)}}{\beta_{k_2}^{(n)}} \tag{23}$$

The marginal rates of substitution thus continue to be generic and equal in their functional form to those of the traditional discrete choice models.

3.3 Existence and uniqueness of the fixed-point problem solution

To demonstrate the existence of a solution to the fixed-point problem expressed by (21), we apply the Brouwer fixed-point theorem. This theorem may be stated as follows:

Let H be a non-empty compact convex set of a finite-dimensional Euclidean space k and $f : H \rightarrow H$ be a continuous function. Then f has a fixed point, i.e., $\exists \boldsymbol{h} \in H : \boldsymbol{h} = f(\boldsymbol{h})$.

Below it is shown that the system of equations describing the equilibrium in model (21) satisfies the theorem's existence and completeness hypotheses.

Let

$$H_w := \left\{ \boldsymbol{h} \in \mathbb{R}^k : \sum_p h_w^p = T_w, \ h_w^p \geq 0 \right\} \tag{24}$$

where in this case the dimension is given by $k = \left| p^{w} \right| < \infty$. If we define the affine hyperplane $A_w = \left\{ \boldsymbol{h} \in \mathbb{R}^n : \sum_p h_w^p = T_w \right\}$, then $H_w = \left[0, T_w\right]^k \cap A_w$ and is a non-empty compact convex set given that it is the intersection of $\left[0, T_w\right]^k$, which is compact convex, and A_w, which is closed convex. A point belonging to H_w can therefore be easily found.

The function $f : H_w \to H_w$ is defined as

$$f\left(\boldsymbol{h}_w\right) = \left(T_w \frac{\exp\left(\alpha_w^p - \theta c_w^p - \rho^* \sum_{q \in p^w} \left[\eta_w^{pq} h_w^q \right] \right)}{\sum_{r \in p^w} \exp\left(\alpha_w^r - \theta c_w^r - \rho^* \sum_{q \in p^w} \left[\eta_w^{rq} h_w^q \right] \right)} \right)_p \tag{25}$$

This expression is a construction of two continuous functions and is therefore itself continuous. The parameters $\left(\alpha_w^p, \theta, \rho^*\right)$ are estimated using maximum likelihood given an $\boldsymbol{h}_w$ and thus are continuous functions with respect to the latter, thereby satisfying the Brouwer theorem hypotheses and proving the existence of at least one fixed point. In formal terms,

$$\exists \boldsymbol{h}^* \in H : f\left(\boldsymbol{h}^*\right) = \left(T_w \frac{\exp\left(\alpha_w^p - \theta c_w^p - \rho^* \sum_{q \in p^w} \left[\eta_w^{pq} h_w^{q*} \right] \right)}{\sum_{r \in p^w} \exp\left(\alpha_w^r - \theta c_w^r - \rho^* \sum_{q \in p^w} \left[\eta_w^{rq} h_w^{q*} \right] \right)} \right) \tag{26}$$

We now prove uniqueness for the case of $\rho^* \geq 0$. Assume that there are 2 different equilibrium points $h^*, g^* \in H$. It is therefore the case that

$$\left(\boldsymbol{h}^* \neq \boldsymbol{g}^* \quad \wedge \quad \sum_p h_w^{p*} = \sum_p g_w^{p*} = T_w \right) \tag{27}$$

$$\Rightarrow \exists p^+, p^- \in \mathbb{N} : \left(p^+ \neq p^-; 1 \leq p^+, p^- \leq n; h_w^{p*-} < g_w^{p*-}; h_w^{p*+} > g_w^{p*+} \right) \tag{28}$$

$$\frac{h_w^{p*}}{g_w^{p*}} = \frac{\sum_{p \in p^w} \exp\left(\alpha_w^p - \theta c_w^p - \rho^* \sum_{q \in p^w} \left[\eta_w^{pq} g_w^{q*} \right] \right)}{\sum_{p \in p^w} \exp\left(\alpha_w^p - \theta c_w^p - \rho^* \sum_{q \in p^w} \left[\eta_w^{pq} h_w^{p*} \right] \right)} \times \exp\left\{ \rho^* \left(\sum_{q \in p^w} \left[\eta_w^{pq} h_w^{p*} \right] - \sum_{q \in p^w} \left[\eta_w^{pq} g_w^{q*} \right] \right) \right\} \tag{29}$$

Without loss of generality we can assume that

$$\sum_{p\in p^w} \exp\left(\alpha_w^p - \theta c_w^p - \rho^* \sum_{q\in p^w}\left[\eta_w^{pq} h_w^{p^*}\right]\right) \geq \sum_{p\in p^w} \exp\left(\alpha_w^p - \theta c_w^p - \rho^* \sum_{q\in p^w}\left[\eta_w^{pq} g_w^{q^*}\right]\right) \tag{30}$$

For the case of p^+ we have

$$h_w^{p^{*+}} > g_w^{p^{*+}} \Rightarrow \frac{h_w^{p^{*+}}}{g_w^{p^{*+}}} > 1 \tag{31}$$

$$\Rightarrow \underbrace{\frac{\sum_{p\in p^w} \exp\left(\alpha_w^p - \theta c_w^p - \rho^* \sum_{q\in p^w}\left[\eta_w^{pq} g_w^{p^{*+}}\right]\right)}{\sum_{p\in p^w} \exp\left(\alpha_w^p - \theta c_w^p - \rho^* \sum_{q\in p^w}\left[\eta_w^{pq} h_w^{p^{*+}}\right]\right)}}_{\leq 1} \times \underbrace{e^{-\rho^*\left(\sum_{q\in p^w}\left[\eta_w^{pq} h_w^{p^{*+}}\right] - \sum_{q\in p^w}\left[\eta_w^{pq} g_w^{p^{*+}}\right]\right)}}_{<1} < 1 \tag{32}$$

This is true given assumption (30) and provided that $\rho^* > 0$. From this we deduce that $-\rho^*\left(h_w^{p^*} - g_w^{p^*}\right) < 0 \Rightarrow e^{-\rho^*\left(h_w^{p^*} - g_w^{p^*}\right)} < 1$. There is thus a contraction and we conclude that a fixed point not only exists but is unique.

4. NUMERICAL RESULTS AND EXTENSIONS

In 4.1 we present a route choice case study comparing the results of a real application of the proposed model to those generated by a classic multinomial model, the C-logit model in (4) above and the path-size logit model in (6) above. Then, in 4.2, we develop an extension of the proposed formulation with route correlation to traffic assignment on congested networks (private transport).

4.1 Application to a medium-sized network: the Santiago Metro

This case study applies the various models to route choice on the Metro rail transport system in the Chilean city of Santiago. Successive transfer points are chosen between origin and destination pairs that in many instances are joined by more than one feasible route (see Figure 1). The analysis focuses on the morning and evening peak periods (7 am to 9 am and 6 pm to 8 pm) when approximately 790,000 trips are taken across the system, 44% of which include transfers.

Trip data was obtained through an O-D survey conducted on the Metro in which 92,800 system users, or about 12% of the total, participated. Since only those trips for which there existed more than one alternative route were retained in the data set, the number of individuals or observations finally employed was 16,029, or about 40% of users who transferred.

When the survey was taken in October 2008, the Santiago Metro consisted of 5 lines and 85 stations, 7 of which were transfer points. Of the 7,140 O-D pairs on the network, 4,985 (70%) of them required making a transfer. The reasonability criterion applied in including a route as a possible alternative for a given pair was that at least one surveyed traveler was observed to have used it. Data on the alternative routes for O-D pairs across the system that had more than one route on this criterion are given in Table 1. Although in the majority of cases there were only two alternatives, some pairs had as many as four. Denser networks than this one would no doubt have a greater proportion of pairs with alternative routes.

Figure 1.- Santiago Metro in 2008.

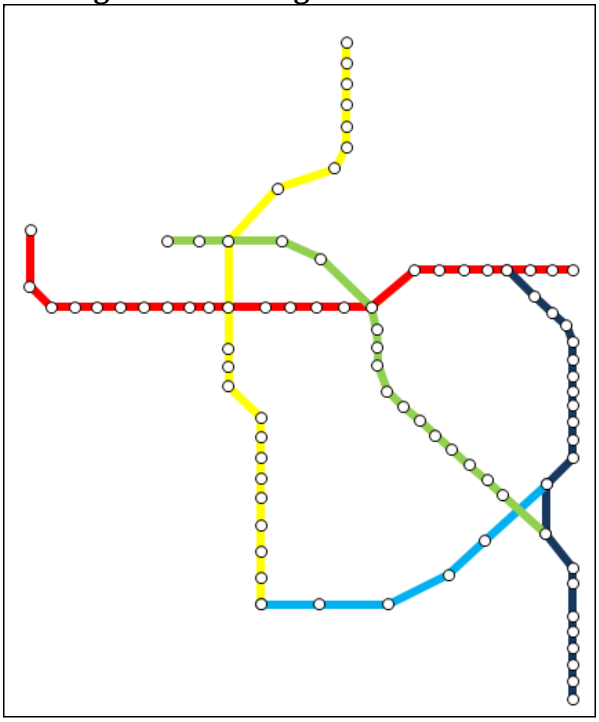

Table 1.- Origin-destination pairs and alternative routes, 2008.

No. of observed alternative routes	% of all O-D pairs	% of all trips
2	97 %	93 %
3	3 %	7 %
4	< 1 %	< 1 %

For the proposed model (14), the generalized cost of traveling between pair *w* on route *p* is defined as $L_w^p = \beta_{time} X_{time,w}^p + \beta_{trans} X_{trans,w}^p$, where $X_{time,w}^p$ is total trip time and $X_{trans,w}^p$ is the number of transfers along the route. Total trip time is the sum of in-vehicle time, wait time and transfer time (walking between platforms). The number of transfers on a given route alternative is considered as an indicator of the disutility of transferring. Since the Metro uses a flat fare system, the fare is the same regardless of route and can therefore be omitted from the cost function.

The η_w^{pq} term, the spatial correlation factor due to route overlap between routes *p* and *q* joining pair *w*, is defined as suggested by (Yai *et al.*, 1997):

$$\eta_w^{pq} = \frac{D_w^{pq}}{\sqrt{D_w^p \cdot D_w^q}} \qquad (33)$$

where D_w^{pq} is the length of overlap between routes p and q, and D_w^p and D_w^q are the respective lengths of routes p and q.

This expression stems from the traditional definition of the similarity factor in the C-logit model. As with the parameter β in that model (see (5) above), the parameter $-\rho^*$ must be negative, ensuring that ρ^* is positive.

The proposed model was compared with a conventional multinomial logit that included only network service level variables (trip time and number of transfers). In other words, ρ^* was set to 0. The comparison was performed using statistical tests.

Thus, the four route choice models estimated for the Santiago Metro were:

a) Multinomial logit (MNL). This is the base model since it does not account for correlation and is also used to construct the route choice proxy variables.
b) C-logit (Cascetta *et al*., 1996).
c) Path size (Ben-Akiva and Ramming, 1998).
d) Fixed point with spatial correlation (FPM), the proposed model.

In all four models, the explanatory variables were trip time and number of transfers. Given its high collinearity with the number of transfers, the frequency of service (the inverse of wait time) was excluded along with the flat fare variable.

The estimation results are set out in Table 2. The proposed FPM converged in only 4 iterations at the 0.01% tolerance level. As can be seen, the trip time and number of transfers parameters had the correct sign and were statistically significant. In both the C-logic and the path size models the correlation parameter was weakly significant, implying that the two formulations were little different statistically from the MNL base model with no correlation factor. On the other hand, in the proposed FPM model the parameter not only had the expected sign but was highly statistically significant, indicating that it was the only one which captured the correlation between routes.

Table 2.- Route choice model estimation results.

Variable	**MNL**	**C-logit**	**Path size**	**FPM**
Trip time	- 0.141 (- 47.20)	- 0.142 (- 44.68)	- 0.141 (- 43.96)	- 0.117 (- 37.15)
No. of transfers	- 1.140 (- 48.44)	- 1.144 (- 48.43)	- 1.144 (- 48.42)	- 0.924 (- 35.36)
Spatial correlation	n. a.	- 0.081 (- 0.57)	- 0.101 (- 0.68)	0.062 (14.98)
Log-likelihood	- 7.381	- 7.380	- 7.380	- 7.235
Adjusted rho squared	0.362	0.362	0.362	0.374

The proposed model also showed significant improvements in goodness of fit (log likelihood). For comparative purposes the FPM can therefore be contrasted exclusively with the MNL model, the C-logit and path size formulations being in effect statistically equivalent to the latter.

In addition to the log-likelihood function, the following indicators were used to compare the MNL and fixed-point models on goodness of fit (Copas, 1989; Wasserman, 2006):

i. Percent of Correct Predictions (*PCP*):

ii. Residual Sum of Squares (*RSS*): $S' = \sum_i (Y_i - P_i)^2$, where Y_i is 1 if alternative i is chosen and 0 otherwise, and P_i is the probability predicted by the model of choosing alternative i.

iii. Weighted Residual Sum of Squares (*WRSS*): $S' = \frac{\sum_i (Y_i - P_i)^2}{P_i(1-P_i)}$.

The results of the three indicators for the proposed FPM and the MNL model are given in Table 3, showing clearly that the fixed-point formulation performed better.

Table 3.- Goodness-of-fit indicators for MNL vs. FPM

Indicator	MNL	FPM
PCP	80.60	80.72
RSS	2,328.6	2,292.5
WRSS	30,976.2	19,803.1

As regards the valuation of the attributes by system users, the marginal rate of substitution between number of transfers and trip time for the FPM is 7.9 minutes. In other words, users are willing to travel up to 7.9 minutes more to avoid transferring. For the MNL model, the marginal rate of substitution is 8.1 minutes. The difference between the two models is thus very small, a sign that the additional explanatory capacity of the proposed model afforded by the capture of route correlation cannot be attained through traditional time and transfer variables.

4.2 Extension of proposed fixed-point model to traffic assignment with route correlation

The proposed fixed-point model can be extended to the traffic assignment problem in transport networks with congestion. The multi-objective optimization problem would then take the following form:

$$
\begin{aligned}
&\min \quad F_1 = \sum_a \int_0^{f_a} c_a(x)\,dx \\
&\min \quad F_2 = \sum_w \sum_{p \in p^w} h_w^p \left(\ln h_w^p - 1\right) \\
&\min \quad F_3 = \sum_w \sum_{p \in p^w} \sum_{\substack{q \in p^w \\ q \neq p}} \left[\eta_w^{pq} \left(h_w^p - t_w\right)\left(h_w^q - t_w\right) \right] \\
&s.t.: \sum_{p \in p^w} h_w^p = T_w, \quad \forall w \qquad\qquad (\gamma_w) \\
&\qquad \sum_{\substack{p \in p^w \\ a \in p}} h_w^p = f_a, \quad \forall a
\end{aligned}
\tag{34}
$$

Objective F_1 is the classic Beckmann transformation of the traffic assignment problem (Beckmann *et al*., 1956). Objectives F_2 and F_3 were described above in 3.1 as information criteria. Combining F_2 and F_3 together with flow conservation constraints gives the Fisk stochastic assignment model (Fisk, 1980).

A substitute optimization problem for (8) that defines a new stochastic traffic assignment model with route correlation is the following:

$$
\begin{aligned}
&\min \quad Z = \sum_a \int_0^{f_a} c_a(x)\,dx + \frac{1}{\theta} \sum_{p \in p^w} \sum_w h_w^p \left(\ln h_w^p - 1\right) + \frac{1}{\rho} \sum_w \sum_{p \in p^w} \sum_{\substack{q \in p^w \\ q \neq p}} \left[\eta_w^{pq} \left(h_w^p - t_w\right)\left(h_w^q - t_w\right) \right] \\
&s.t.: \sum_{p \in p^w} h_w^p = T_w, \quad \forall w \qquad\qquad (\gamma_w) \\
&\qquad \sum_{\substack{p \in p^w \\ a \in p}} h_w^p = f_a, \quad \forall a
\end{aligned}
\tag{35}
$$

The first-order conditions for (35) are

$$
h_w^p = T_w \frac{\exp\left(-\theta C_w^p - \dfrac{\theta}{\rho} \sum_{q \in P_w} \left[\eta_w^{pq} \left(h_w^q - t_w\right) \right] \right)}{\sum_{p \in P_w} \exp\left(-\theta C_w^p - \dfrac{\theta}{\rho} \sum_{q \in P_w} \left[\eta_w^{pq} \left(h_w^q - t_w\right) \right] \right)}
\tag{36}
$$

where $C_w^p = \sum_{\substack{p \in p^w \\ a \in p}} c_a\left(f_a^*\right)$.

5. CONCLUSIONS

A new transport network route choice model was developed that explicitly incorporates the phenomenon of correlation between routes. The model is applicable to public transport networks and extensible to the traffic assignment problem. The presence of correlation between route alternatives was contrasted empirically by means of classical econometric techniques.

A multi-objective problem was stated and a substitute problem then formulated whose optimality conditions yielded a logit specification with an endogenous variable constituting a fixed-point model that is estimated and solved iteratively. The estimation was performed by maximum likelihood and the use of instrumental variables due to the presence of endogeneity in the model's explanatory variables. The functional form of the proposed fixed-point model combined with the utilization of instrumental variables guaranteed both the existence and uniqueness of the solution.

The proposed model was compared with other route choice models reported in the literature in a case study of the Santiago Metro. The results obtained were both satisfactory and superior to the existing formulations. Unlike the latter, the proposed fixed-point model was able to capture the correlation between routes and provided better goodness of fit.

Although the proposed formulation is more complex to estimate due to the iterative process that must be used to obtain the parameters, in practice the iterations converge rapidly. Furthermore, the additional estimation complexity does not complicate the derivation of project evaluation indicators such as marginal rates of substitution, which retain the simple functional form of MNL models.

Finally, future research should address the apparent advantages of the proposed model on larger networks and traffic assignment problems.

REFERENCES

Beckmann, M.J., McGuire, C.B. and Winsten, C.B. (1956). *Studies in the Economics of Transportation.* Yale University Press, New Haven, Connecticut.

Bell, M. G. H. (1995) Alternatives to dial's logit assignment algorithm, Transportation Research, 29B, 287–295.

Bekhor, S., M. E. Ben-Akiva and Ramming, M. S. (2006) Evaluation of choice set generation algorithms for route choice models, Annals of Operations Research, 144 (1) 235–247.

Ben-Akiva, M., and Bierlaire, M. (2003). Discrete choice models with applications to departure time and route choice. In Hall, R. (ed) Handbook of Transportation Science, 2nd edition, 7-38. Kluwer.

Ben-Akiva, M., Bergman, M.J., Daly, A.J. and Ramaswamy, R. (1984). Modeling inter urban route choice behavior. In: Ninth International Symposium on Transportations and Traffic Theory, VNU Science Press.

Ben-Akiva, M. and Ramming, S. (1998). Lecture notes: Discrete choice models of traveler behavior in networks. Prepared for Advanced Methods for Planning and Management of Transportation Networks. Capri, Italy.

Bliemer, M. C. J. and Bovy, P. H. L. (2008) Impact of route choice set on route choice probabilities, Transportation Research Record, 2076, 10–19.

Bovy, P. and Hoogendoorn-Lanser, S. (2005). Modeling Route Choice Behavior in Multi-Modal Transport Networks. Transportation, 32, 341-368.

Bovy, P. H. L., Bekhor, S. and Prato, C. G. (2008) The factor of revisited path size: Alternative derivation, Transportation Research Record, 2076, 132–140.

Boyce, D.; Lee, D. H. and Ran, B. (2001). Analytical Models of the Dynamic Traffic Assignment Problem. Networks and Spatial Economics, 1, 377-390.

Boyce D., Mahmassani H. and Nagurney A. (2005). A Retrospective on Beckmann, McGuire and Winsten's Studies in the Economics of Transportation. Papers in Regional Science 84, 85-103.

Cascetta, E., Nuzzolo, A., Russo, F. and Vitetta, A. (1996). A modified logit route choice model overcoming path overlapping problems. Specifications and some calibrations results for interurban network. In: Proceedings of ISTTT conference, Lyon, France.

Cascetta, E., Papola, A., Russo, F. and Vitetta, A. (1998). Implicit availability/perception logit models for route choice in transportation networks. Preprint 8 WCTR, Antwerp, Belgium.

Cascetta, E., Russo, F., Viola, F.A. and Vitetta, A. (2002). A model of route perception in urban road networks. Transportation Research, 36B, 577-592.

Copas, J. B. (1989). Unweighted sum of squares test for proportions. Applied Statistics; 38,71 -80.

Dafermos, S. and Sparrow, F.T. (1969). The traffic assignment problem for a general network. *J. Res. Nat. Bur. Standard* 73B, 91-117.

Dafermos, S. (1980). Traffic equilibrium and variational inequalities. *Transportation Science* 14, 42-54.

Daganzo, C. F. and Sheffi, Y. (1977). On stochastic models of traffic assignment. Transportation Science, 11, 253-274.

De Cea, J. and Fernández, J. E. (1993) Transit assignment for congested public transport systems: an equilibrium model. Transportation Science, 27(2), 133 - 147.

De Cea, J. Fernández, J. E., and De Grange, L. (2008). Combined Models with Hierarchical Demand Choices : a Multi-Objective Entropy Optimization Approach. Transport Reviews, 28, 415-438.

De Grange, L. and Muñoz, J. C. (2009). An Equivalent Optimization Formulation for the Traffic Assignment Problem with Asymmetric Linear Costs. Transportation Planning and Technology, 32, 1-25.

Dial, R. B. (1971). A probabilistic multipath traffic assignment model which obviates path enumeration. Transportation Research, 5, 83-113.

Dijkstra, E. W. (1959), A note on two problems in connexion with graphs, in Numerische Mathematik, 1, 269–271.

Fisk, C. (1980). Some Developments in Equilibrium Traffic Assignment. Transportation Research, 14B, 243-255.

Florian, M. (1976). Traffic Equilibrium Methods, Lecture Notes in Economics and Mathematical, Systems 18, Springer-Verlag, Berlin.

Florian, M. (1977). A traffic equilibrium model of travel by car and public transit modes, Transportation Science, 11, 169-179.

Florian, M. and Hearn, D. W. (2001). Traffic assignment: Equilibrium models, Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering, 215, 305–315.

Frejinger, E., and Bierlaire, M. (2007). Capturing correlation with subnetworks in route choice models, Transportation Research, 41B,363-378.

Golan, A. (2002). Information and entropy econometrics — Editor's view. Journal of Econometrics, 107, 1-15.

Greene, W. H. (2008). Econometric Analysis, 6th Edition.

Hazelton, M.L. (1998). Some remarks on Stochastic User Equilibrium. Transportation Research, 32B, 101–108.

Marler, T. and Arora, J. (2004). Review of multi-objective optimization concepts and algorithms for engineering. Technical Report Number ODL 01-04, University of Iowa.

Patriksson, M. (1998): Nonlinear Programming and Variational Inequality Problems: A Unified Approach, vol. 23 of *Applied Optimization*, Kluwer, Dordrecht

Prashker, J. N. and Bekhor, S. (2004). Route Choice Models Used in the Stochastic User Equilibrium Problem: A Review. Transport Reviews, Vol. 24, No. 4, 437–463.

Prato, C. G. and Bekhor, S. (2007) Modeling route choice behavior: How relevant is the composition of choice set?, Transportation Research Record, 2003, 64–73.

Prato, C. G. (2009) Route choice modeling: past, present and future research directions. Journal of Choice Modeling, 2, 65–100.

Ramming, M. S. (2001) Network knowledge and route choice, unpublished dissertation, MIT, Cambridge, MA.

Ran, B., and Boyce, D. E. (1996), Modeling Dynamic Transportation Networks, Springer-Verlag, Berlin, Germany.

Raveau, S., Muñoz, J. C. and De Grange, L. (2011). A Topological Route Choice Model for Metro Transportation Research, 45A, 138 - 147.

Szeto, W.Y. and Lo, H. K. (2006). Dynamic traffic assignment properties and extensions, Transportmetrica, 2, 31-52.

Wardrop, J.G. (1952) Some theoretical aspects of road traffic research. Proceedings of the Institute of Civil Engineering, Part II 1, 325-378.

Wasserman, L. (2006). All of Nonparametric Statistics. Berlin: Springer-Verlag.

Yai, T., Iwakura, S. and Morichi S. (1997) Multinomial probit with structured covariance for route choice behavior, Transportation Research, 31B, pp. 195–207.