

On the role of bridges as anchor points in route choice modeling

Hamzeh Alizadeh¹  · Bilal Farooq¹ · Catherine Morency¹ · Nicolas Saunier¹

Published online: 2 February 2017
© Springer Science+Business Media New York 2017

Abstract This work builds upon the thought that individuals allocate higher levels of importance to some particular features of the route, so called anchor points. Previous route choice models have either ignored the effects of anchor points (route-based models), or have given an exclusive attention to their effects and ignored the behavioral accuracy and practicality of these models (anchor-based models). In this work we argue that the consideration of both route-level attributes and anchor points would enhance the behavioral aspect of route choice models as well as their estimation and prediction abilities. Global Positioning System traces have been used to investigate the effect of bridges as anchor points for trips between Montreal and its Northern suburb, Laval. A classic Nested Logit and a nested Logit Kernel model have been estimated, in which interdependencies among routes crossing the same bridge are captured through the nested structure and the adopted factor analytic approach, respectively. A Metropolis–Hastings path-sampling algorithm is applied, for the first time, on a large road network with more than 40,000 nodes and 19,000 links to provide the consideration choice set. Estimates are then compared to three alternate models, representing route-based and anchor-based formulations; namely Path-Size Logit, Extended Path-Size Logit, and Independent Availability Logit models. Empirical results showed that the proposed nested structures with MH sampling provide better estimates and also perform better in the validation step with respect to comparative models. Findings

✉ Hamzeh Alizadeh
hamzeh.alizadeh-aliabadi@polymtl.ca

Bilal Farooq
bilal.farooq@polymtl.ca

Catherine Morency
cmorency@polymtl.ca

Nicolas Saunier
nicolas.saunier@polymtl.ca

¹ Department of Civil, Geological and Mining Engineering, Polytechnique Montréal, C.P. 6079, succ. Centre-Ville, Montreal, QC H3C 3A7, Canada

underscore the importance of considering anchor points in conjunction with route level attributes in route choice decisions.

Keywords Route choice · Bridge choice · Anchor points · Discrete choice models · Nested Logit · GPS · Metropolis–Hastings

Introduction

Route choice modeling is probably one of the most complex and challenging problems in traffic assignment. It investigates the process of route selection by an individual, making a trip between predefined origin and destination (OD) pairs. The heterogeneity in travelers' behavioral characteristics, in conjunction with the complex effect of route attributes, further increases the inherent complexity of route choice modeling. Although several approaches have been proposed to tackle this problem, one of the remaining challenges in route choice modeling is the consistency of the modeling approach with the underlying behavioral process of drivers' decision making.

In general, most of the proposed route choice models focus on route related attributes of choice alternatives. This implies that in these “*route-based*” formulations, the route is perceived as an entity, and only attributes concerning the whole trajectory are used to characterize each choice alternative. From a behavioral perspective, this formulation suggests that the consideration set is formed based on route-level characteristics of trajectories and the final choice is made by selecting an entire route out of a considered set of alternatives. C-Logit (Cascetta et al. 1996) and Path-Size Logit (PSL) (Ben-Akiva and Bierlaire 1999) are among the most widely used route-based models. In these models, the similarity issue between alternatives has been addressed by adding a correction term to the deterministic part of the utility function, which alters the utility of paths¹ based on their similarities. However, the applied correction factors in these models account only for similarities between the considered set of paths. To overcome this limitation, Frejinger et al. (2009) have proposed the Extended-Path-Size Logit (EPSL) model, which accounts for correlations between sampled and non-sampled alternatives.

An alternate approach offers a “*link-based*” formulation to deal with the correlation issue among alternatives. This approach, firstly proposed by Vovsha and Bekhor (1998) and later applied by Lai and Bierlaire (2015), adopts a cross-nested logit structure in which each link of the network constitutes a nest and each route belongs to several nests. Another application of the sequential link choice method has been applied by Fosgerau et al. (2013) in the context of a recursive logit model, which can be consistently estimated without the need of path sampling. The link-based formulation suggests that drivers have a link-by-link perception of the network and their choices are based on link-level attributes. Since a real network consists of a large number of links, the estimation of these models could be very computationally expensive.

A third approach is the “*anchor-based*” formulation which gives an exclusive importance to the effect of some prominent features of the road network, so-called landmarks or anchor points. This approach basically argues that anchor points are decisive points based on which drivers choose their paths. It has been pointed out by several researchers that

¹ The terms route and path are used interchangeably in this article.

individual's perception of the road network follows a hierarchical representation (Hirtle and Jonides 1985; Holding 1994), and anchor points play a crucial role in the behavioral nature of route choice decisions (Manley et al. 2015b; Habib et al. 2013; Kazagli and Bierlaire 2015). The space hierarchy and the role of anchor points have also been found to be significant in similar decision making contexts such as location choice modeling (Elgar et al. 2009, 2015).

The importance of anchor points is emphasized in riverside cities. In these cities, the two sides, separated by the river, are usually connected through several bridges and tunnels. These infrastructures are usually prone to become traffic bottlenecks and to face recurrent traffic congestion (Woo et al. 2015; Habib et al. 2013; Sun et al. 2014). The high travel time variability of these small segments of the route, which is usually due to the large fluctuation of travel demand, turns them into influential points in the process of route selection. Accordingly, the choice of anchor points has a major effect on route selection, and routes crossing same anchor points share unobserved components such as safety, scenery, driving comfort, etc., emerging from the similarities of the road network and geographical characteristics.

In this work we argue that the three abovementioned formulations, in isolation, are not behaviorally accurate enough to represent the underlying nature of route choice behavior; rather a hybrid approach is needed. Route-based formulations suggest that route selection is mostly based on route-level attributes (Manley et al. 2015a), and ignore the influence of anchor points. Link-based formulations neglect the higher importance of anchor points by allocating the same level of importance to every link. Moreover, the anchor-based model proposed by Habib et al. (2013) incorporates a full probabilistic choice set generation, which is behaviorally inaccurate, and theoretically impractical and unmanageable in large route choice datasets and real world networks (Prato 2009).

We propose a generic “*anchor-based nested*” structure to promote the behavioral aspect of route choice models by incorporating the effect of route level attributes as well as anchor points on drivers' choices. First, we adopt a classic Nested Logit (NL) structure within a discrete choice framework, in which upper nests correspond to anchor points and lower nests include route alternatives. Second, a nested Logit Kernel (LK) model is estimated to capture the reciprocal effect of route level attributes and anchor points on route selection. In the former model, the nested structure captures the shared unobserved components of the utility function among routes crossing the same bridge, while in the latter, the adopted factor analytic approach accounts for the interdependencies and latent similarities.

Similarly to anchor-based models, these approaches allocate a distinctive importance to the selection of anchor points as crucial segments of the route. Moreover, they can handle very large datasets and real world networks; considering multiple route alternatives within each bridge is easily manageable; and route-level attributes are also considered to be decisive and influential in the final route selection.

To explore the performance of the proposed formulations, GPS traces of taxi trips between the islands of Montreal and Laval have been used. The unique aspect of these trips is that drivers have to choose among a maximum of nine bridges separating the two regions. The access to these bridges face recurrent congestion, and despite their small share in the whole route, they have a significant impact on the total travel time and hence on drivers' route choice decisions. In our application, bridges are considered as anchor points for trips between Montreal and Laval, and their effects on route choice decisions have been evaluated in conjunction with route level attributes. A very large real-world road network, with more than 40,000 nodes and 19,000 links, is used for choice set generation as well as

model estimation. Estimates are then compared to three alternate models, representing route-based and anchor-based formulations; namely PSL, EPSL, and Independent Availability Logit (IAL) models.

Taxi drivers are considered to have a more precise knowledge of the road network and its traffic conditions, due to their higher driving experience. In order to capture the effect of anchor-points, we have focused on the behavior of taxi drivers as well-informed individuals who are more familiar with travel time variations and congestion periods over bridges connecting Montreal to Laval.

This work contributes to the existing state-of-the-art through the following aspects: (1) the presented formulations capture the effect of anchor points in conjunction with route level attributes in route choice modeling, (2) they improve the behavioral aspect of anchor-based formulations by capturing shared unobserved components among route alternatives crossing the same anchor point, and (3) the MH algorithm has been employed, for the first time, on a large real world route network to generate alternative choice sets.

This paper is organized as follows. First we review earlier approaches to route choice modeling and in that context further clarify the contributions of this study. The case study and data are presented next. We then discuss in detail the proposed econometric formulations and the estimated comparative models, as well as their respective utility function specifications and choice set generation algorithms. We then discuss the results, validation process, and comparison between models. In the end, we highlight the most significant findings of this study, underscore its limitations, and suggest further research directions.

State-of-the-art

There is a large body of literature in microeconomics, behavioral science, psychology, and behavioral geography that focuses on improving the understanding of the underlying process of decision making. Accordingly, several modeling frameworks have been proposed to simulate drivers' route choice behavior. Prospect theory (Kahneman and Tversky 1979; Gao et al. 2010) and cumulative prospect theory (Tversky and Kahneman 1992; Xu et al. 2011; Connors and Sumalee 2009) have been applied by researchers to take into account the limited rationality of drivers in making decisions, by incorporating psychological and behavioral aspects. In some other studies, the uncertainty and imprecision of drivers in making route choice decisions have been taken into account using Fuzzy Logic (Henn 2003; Murat and Uludag 2008; Quattrone and Vitetta 2011; Luisa De Maio and Vitetta 2015). Artificial neural networks have also been used to take into account the non-linearity of the decision making process by imitating the human conscious structure (Dougherty 1995; Kim et al. 2005). Recently, a Random Regret Minimization (RRM) approach has been adopted by Prato (2014) in route choice modeling context. This approach argues that choice makers tend to choose the alternative which minimizes the regret of not having chosen other alternatives.

Among the proposed approaches, Random Utility Maximization (RUM) models have received considerable attention. In this approach, individuals' preferences are represented by a value called "*utility*", which captures the effect of different factors on the actual choice, and decision makers tend to maximize their perceived utilities. Since the decision maker may not have a perfect knowledge about these factors, an error component has been introduced to take into account the stochasticity and imprecision caused by uncertainty and

behavioral randomness (Ben-Akiva and Bierlaire 2003). This paper is built upon the random utility maximization framework.

In the context of route choice, the perceived utility can be attributed to factors such as travel time, distance, congestion, safety, scenery, route complexity, number of traffic signals, trip purpose, fuel consumption, toll, road type, anchor points, etc. Behavioral and mental factors such as the inertia of taking the same route, memory, spatial abilities, driving experience, and the learning process may also play a role and add to the complexity of the modeling process (Prato 2009).

It has been suggested that individuals have a hierarchical planning strategy following the hierarchical representation of space and its connectivity (Manley et al. 2015b; Wiener and Mallot 2003), which yields an anchor-based navigation in which individuals orient themselves based on distinguished features of the route (Foo et al. 2005). Anchor points are defined as being important focal points and cognitively salient cues with prominent features, with applications in cognitive tasks, comprising way-finding, distance assessment, and direction estimation. Major route infrastructures, such as bridges, highways, eminent road interchanges, intersections and roundabouts, can be considered as anchor points in route choice modeling.

Several studies have argued that anchor points influence route choice decisions (Lynch 1960; Prato et al. 2012; Kaplan and Prato 2012; Golledge et al. 1985; Couclelis et al. 1987; Habib et al. 2013; Prato and Bekhor 2007). Manley et al. (2015a) studied minicab drivers in London, using GIS and statistical analysis, and concluded that their route choice behavior is poorly described by shortest-path algorithms and is improved when the role of anchor points is considered. To emphasize the significant role of anchor points, they have also proposed a conceptual subjective anchor-based route choice modeling schema. Similarly, Kazagli and Bierlaire (2015) argue that drivers describe their routes using a short sequence of Mental Representation Items (MRIs) such as anchor points or pieces of infrastructures instead of using a link-sequence representation. Moreover, a recent study by Manley et al. (2015b) confirms that the mental representation of the spatial hierarchy influences route choices. They propose a coarse to granular hierarchical representation of space, represented from top to bottom by *Regions*, *Nodes*, and *Roads*, where *Regions* represent clusters of nodes sharing a common characteristic; *Nodes* represent certain road junctions, landmarks, and anchor points; and *Roads* form the basis of the hierarchy, defining the route between consecutive *Nodes*. In this hierarchical schema, route choice is made through the selection of a sequence of regions, nodes across subsequent regions, and eventually, roads between successive nodes.

Despite the undeniable importance of anchor points on drivers' route choice decisions, relatively little attention has been given to *anchor-based* route choice models. An interesting approach has been investigated by Habib et al. (2013) in which the authors applied an Independent Availability Logit (IAL) model, originally proposed by Swait and Ben-Akiva (1987), in a route choice context to emphasize the role of bridge choice in route choice decisions; the case study was the Greater Montreal Area. The IAL model follows the probabilistic two-stage choice model proposed by Manski (1977) in which the selection probability of an alternative depends on the selection probability of all subsets of the universal choice set containing that particular alternative. The IAL model jointly estimates the final choice and choice sets among all the possible combinations.

The study by Habib et al. (2013) was based on data collected from the OD survey of Montreal, in which the authors had only access to declared chosen bridges. They have considered a shortest path algorithm, based on segments' speed limits, to generate one path per bridge for each OD pair, comprising the choice set of route alternatives. A noticeable

limitation, pointed out by Prato (2009), of the fully probabilistic choice set generation approach adopted by the IAL approach, is its immense calculation burden when the size of the choice set increases. Considering m as the total number of possible alternatives, the total number of possible non-empty choice sets ($2^m - 1$) increases exponentially with m , which makes it impractical to apply this model in route choice problems, where the considered choice set is large. In Habib et al. (2013), the inability of the IAL model in handling large choice sets might be an additional reason to their data limitation issue for considering only one alternative per bridge. In behavioral terms, the consideration of a single shortest path per bridge scales down the route choice problem to a bridge choice problem. Moreover, it is behaviorally unrealistic to assume that decision makers would consider every possible subset of the consideration choice set, before making a choice. Also, it is worth mentioning that the problem of shared segments between alternatives is not addressed in this recent application of the IAL model.

Context and dataset

Nowadays, the prevalent use of GPS technology provides researchers with an abundance of high-resolution geospatial data, which allows obtaining continuous and detailed (link-by-link) information on drivers' travel paths, accompanied by possible additional information such as travel direction and speed. A relatively new source of GPS data is recorded by taxi companies around the world, mainly for operational purposes.

This study is based on GPS traces of taxi drivers, collected by a taxi company, in the context of the metropolitan region of Greater Montreal, depicted in Fig. 1a. The data was collected by a taxi company that constitutes around 25% of the Montreal Island taxi fleet, and its operation is restricted to trips starting or ending in the central part of the island. Data has been stored in a PostgreSQL database, and the PostGIS spatial extension has been added to support geographical datatypes and queries. A direction-based nearest link point-to-curve map matching algorithm has been adopted to associate each GPS record to the road network. In point-to-curve map matching algorithms, every GPS point is matched

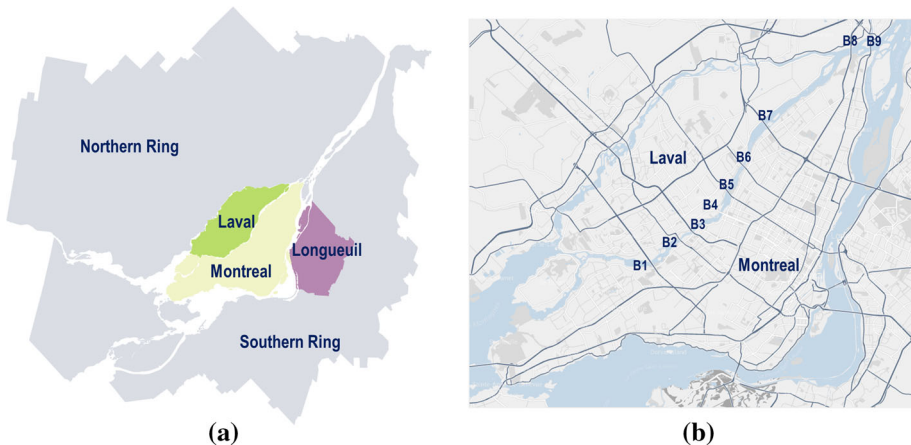


Fig. 1 Context of the studied region. **a** The metropolitan region of Greater Montreal; **b** locations of bridges connecting Montreal to Laval

onto the closest link in the network. The major shortcoming of these algorithms is that they do not produce reliable results in high road density networks and specially at intersections due to directional problems (White et al. 2000; Zhou and Golledge 2006). To overcome this issue, we associated each GPS record to its nearest link on the network with respect to its azimuth, so that it ensures that GPS points are not incorrectly matched to closer links with incorrect directions. A distance-based shortest-path algorithm has then been applied between consecutive GPS records to deduce the entire path for each trip.

Montreal is an island city, separated from its suburbs by two rivers. This means that drivers entering or exiting Montreal need to cross one of the sixteen bridges connecting Montreal to its suburbs. These bridges face recurrent congestion and act as bottlenecks. For trips heading to or exiting Montreal, bridge choice can have a significant importance on route choice decisions (Habib et al. 2013). Considering bridges as anchor points, the geographical context of Montreal allows us to study the effect of anchor points in conjunction with route level attributes in route choice decisions.

In this study we focus on trips taking place between the Islands of Montreal and Laval, the largest suburb of Montreal located on the north of the city. These two islands are directly connected through seven bridges; B1–B7. Since bridges B8 and B9 might also provide convenient alternatives for trips between the East of Montreal and Laval, they have been included in this study. Figure 1b depicts the location of these nine bridges and Table 1 summarizes some of the pertaining properties of these bridges.

Montreal and Laval cover a total surface of 632.3 km² containing a population of roughly 2.3 million inhabitants (Communauté métropolitaine de Montréal 2012). Their road networks comprise more than 40,000 nodes and 19,000 links. The network data has been extracted from OpenStreetMap project in the format of geographical layers (shapefiles).

GPS records for the month of October 2014 have been extracted for this study. The dataset includes two tables: (a) *GPS* table containing information for every recorded point such as ID, time, geographical position, speed, and etc.; (b) the *Events* table, including information regarding the state of the taxi such as when the passenger has boarded and got off the taxi, etc. These two tables are related through a unique trip identifier (Ride ID). Table 2 shows the schema of the dataset.

Table 1 Description of bridges connecting Montreal to Laval

ID	Name	No. of lanes	Length* (km)	Road type	Speed limit	Toll
B1	Pont Louis-Bisson	4	0.535	Highway (13)	100	No
B2	Pont Lachapelle	3	0.264	Arterial	50	No
B3	Pont Mederic-Martin	4	0.361	Highway (15)	100	No
B4	Pont Viau	2	0.658	Arterial	50	No
B5	Pont Papineau-Leblanc	3	0.425	Highway (19)	100	No
B6	Pont Pie-IX	3	0.658	Arterial	80	No
B7	Pont Olivier-Charbonneau	3	1.200	Highway (25)	100	Yes
B8	Pont Charles-De-Gaulle	3	1.450	Highway (40)	100	No
B9	Pont Le Gardeur	2	0.500	Arterial	100	No

* Bridge lengths have been measured in QGIS software

Table 2 Internal structure of the database

GPS table	Events table
ID	Ride ID
Time	Events
Object ID	Flag
Longitude	Accept
Latitude	Reject
Status	Cancel
Speed	Meter
Direction	ON/OFF
Employer ID	Payment
Ride ID	–

Our dataset consists of 4409 GPS records comprising 543 journeys with an average length of around 11 km and a standard deviation of 9 km. The average recorded travel time is around 15 min with a standard deviation of 12 min. The dataset comprises weekdays, as well as weekends and trips made in peak hours as well as off-peak trips. Eighty percent of the dataset (434 records) was randomly selected for calibration purposes, while the remainder 20% (109 records) was used for result validation.

Methodology

The principal aim of this research is to provide a behavioral framework, which explicitly takes into account the effect of anchor points as well as route-level attributes in route selection. To highlight the importance of anchor points in conjunction with route-level attributes, we adopt the following nested structures. First, we adopt a classic Nested Logit model in which the effect of anchor points is addressed in upper nests while route level decisions are represented in lower nests. Then, we adopt a nested Logit Kernel model, which accounts for the interdependencies of route alternatives crossing the same anchor point through the specification of its error structure.

In behavioral terms, our application of these nested structures suggests that individual taxi drivers, travelling between Montreal and Laval, consider bridges as crucial elements, along with other route level attributes, which affect their route choices. This section delineates the formulations of the abovementioned econometric models, presents the comparative route-based and anchor-based models, and describes their respective utility functions and choice set generation algorithms.

Econometric model formulation

Nested Logit

The Nested Logit (NL) formulation was proposed by Ben-Akiva (1973) and proved to be consistent with the stochastic utility maximization theory by McFadden (1978). It is an extension of the Multinomial Logit (MNL) model, and captures some of the unobserved similarities among alternatives by dividing the choice set into several nests. These nests are considered to be collectively exhaustive and mutually exclusive in covering the considered

alternatives. Every nest contains a subset of alternatives sharing a particular characteristic, independent from other subsets of alternatives in other nests. In other words, the probability of choosing an alternative from a nest is considered to be independent from alternatives in other nests, which is known as the Independence of Irrelevant Alternatives (IIA) property. Therefore, the probability of choosing an alternative can be expressed as the product of the conditional probability of choosing that alternative given a particular nest and the choice probability of that respective nest (Guevara and Ben-Akiva 2013; Ben-Akiva 1973). It is worth mentioning that in our case, the nine considered bridges between Montreal and Laval are located relatively far apart from each other, and it is not far-fetched to assume that the IIA property holds within each nest of the model.

NL is a member of the Generalized Extreme Value (GEV) models (McFadden 1978), also known as Multivariate Extreme Value (MEV) models (Guevara and Ben-Akiva 2013). Within this framework, the probability of choosing alternative i by individual n within the true choice set C_n is given by

$$P(i|C_n) = \frac{e^{V_{in}} G_i(e^{V_{1n}}, \dots, e^{V_{J_n n}})}{G(e^{V_{1n}}, \dots, e^{V_{J_n n}})} \tag{1}$$

$$G_i(e^{V_{1n}}, \dots, e^{V_{J_n n}}) = \frac{\partial G}{\partial e^{V_{in}}} (e^{V_{1n}}, \dots, e^{V_{J_n n}}) \tag{2}$$

where G is a non-negative differentiable MEV generating function, G_i is its partial derivative with respect to $e^{V_{in}}$, V_{in} specifies the systematic part of the utility function, and J_n is the number of alternatives in C_n . The probability of choosing alternative i from the true choice set can be written as:

$$P(i|C_n) = \frac{e^{V_{in} + \ln G_i(e^{V_{1n}}, \dots, e^{V_{J_n n}})}}{\sum_{j=1}^{J_n} e^{V_{jn} + \ln G_j(e^{V_{1n}}, \dots, e^{V_{J_n n}})}} \tag{3}$$

where the partial derivative of the MEV generating function for Nested Logit G_i , for the true choice set C , is:

$$G_i(C) = G_i(e^{V_1}, \dots, e^{V_j}) = \mu e^{V_i(\mu_m - 1)} \left(\sum_{i=1}^{J_m} e^{\mu_m V_i} \right)^{\frac{\mu}{\mu_m} - 1} \tag{4}$$

in which μ and μ_m are scale parameters for the model and its nests, respectively, where $\mu/\mu_m \leq 1$, and m is the nest including alternative i . Since it is not feasible to enumerate the true choice set, a subset D has to be sampled, which must include the chosen alternative i . To consistently estimate this model on a subset of alternatives, the correction approach proposed by McFadden (1978) can be adopted, in which an alternative specific correction term is added to the utility function.

$$P(i|D) = \frac{e^{V_{in} + \ln G_i(C) + \ln \pi(D|i)}}{\sum_{j=1}^J e^{V_{jn} + \ln G_j(C) + \ln \pi(D|j)}} \tag{5}$$

$\ln \pi(D|i)$ is the sampling correction factor, and $\pi(D|i)$ is the conditional probability of choosing subset D given the alternative i has been chosen. The approach developed by McFadden (1978) has been adopted by Bierlaire et al. (2008) to demonstrate that the maximization of the quasi-log-likelihood function of Eq. (5) yields consistent parameter estimates. It is worth mentioning that in our case, since a finite set of anchor points (bridges

B1–B9, connecting the two regions) is considered for the upper nest level, the application of this correction factor is found to be superfluous and unneeded.

Although this function leads to the conditional probability of choosing alternative i given the subset D , its application is not valid with sampling of alternatives because $\ln G_j(C)$ is still dependent on the true choice set (Guevara and Ben-Akiva 2013). In order to compensate for the loss of information due to sampling in each nest, an expansion factor w should be considered to approximate the generating function based on the considered sample D' . Moreover, to take into account the physical overlap between routes crossing the same anchor point, the Extended Path-Size factor (see Eq. (17)) has been added to the deterministic part of the utility function, so that:

$$P(i|D, D', w) = \frac{e^{V_i + \ln EPS_i + \ln \hat{G}_i(D', w) + \ln \pi(D|i)}}{\sum_{j=1}^J e^{V_j + \ln EPS_j + \ln \hat{G}_j(D', w) + \ln \pi(D|j)}} \tag{6}$$

The closed form quasi-log-likelihood function has the following structure:

$$QL_{MEV, D, D', w} = \sum_{n=1}^N \ln P(i|D, D', w) = \sum_{n=1}^N \ln \frac{e^{V_i + \ln EPS_i + \ln \hat{G}_i(D', w) + \ln \pi(D|i)}}{\sum_{j=1}^J e^{V_j + \ln EPS_j + \ln \hat{G}_j(D', w) + \ln \pi(D|j)}} \tag{7}$$

Guevara and Ben-Akiva (2013) demonstrated that in order to achieve unbiasedness and consistency, the expansion factor should have the following structure:

$$w_j = \frac{k_j}{E[k_j]} \tag{8}$$

in which k_i is the number of times alternative i has been sampled, and $E[k_j]$ denotes its expected value or its sampling probability. In this study, we adopt the formulation proposed by Lai and Bierlaire (2015) to approximate Eq. (5) over the choice set D' :

$$w_j = \frac{k_j b_{(s)}}{k_s b_{(j)}} \tag{9}$$

where s denotes the path that has been sampled the most, k_s is the number of times alternative s has been sampled ($k_s \geq k_i \forall i \in D'$), and $b_{(s)}$ and $b_{(j)}$ are the theoretical frequencies of the most sampled path and path j , respectively.

Logit kernel (LK)

Logit Kernel, which is a combination of Probit and Logit models, was first proposed by Bolduc and Ben-Akiva (1991). The random component of its utility function is composed of a Probit-like term, which captures the interdependencies among alternatives, and an i.i.d. Gumbel distributed random component. The interdependencies between alternatives can be explicitly specified using a factor analytic approach, proposed by McFadden (1984). This approach accommodates different error structures and reduces the estimation complexity of the model (Bekhor et al. 2002; Bierlaire and Frejinger 2005). The utility function for individual n is defined as below:

$$U_n = X_n \beta + F_n T \zeta_n + v_n \tag{10}$$

where $U_n - (J_n \times 1)$ is the utility vector, and J_n is the number of alternative in the choice set C_n ; $X_n - (J_n \times K)$ is the matrix of explanatory variables; $\beta - (K \times 1)$ is the vector of

unknown parameters; $F_n - (J_n \times M)$ is the factor loading matrix; $T - (M \times M)$ is a diagonal matrix of the standard deviation of each factor; $\zeta_n - (M \times 1)$ is the vector of i.i.d. random variables with zero mean and unit variance; and $v_n - (J_n \times 1)$ is the vector of i.i.d. Gumbel distributed random term with zero location, a scale equal to μ , and a variance equal to $(\pi^2/6\mu^2)$.

The LK model can replicate any error structure and approximate any random utility model (Walker et al. 2004; McFadden and Train 2000; Ben-Akiva et al. 2001). In a Nested Logit analog of the LK model, also known as the nested LK model, F_n is defined to be the alternative-nest incident matrix and is obtained by defining a dummy variable for each nest that equals 1 if an alternative belongs to that particular nest, and 0 otherwise. Moreover, ζ_n is usually assumed to be normally distributed $N(0, 1)$, and T captures the amount of correlation between alternatives belonging to the same nest (Train 2009; Walker et al. 2004). In this study, the correlation related to the physical overlap between routes crossing the same bridge is expressed through the Extended Path-Size factor (see Eq. (17)), which is added to the utility function (ln *EPS*). If the factors ζ_n are known, the probability of choice i given ζ_n is estimated using the MNL formulation:

$$\Lambda(i|\zeta_n) = \frac{e^{\mu(X_{in}\beta + \ln EPS_{in} + F_{in}T\zeta_n)}}{\sum_{j=1}^{J_n} e^{\mu(X_{jn}\beta + \ln EPS_{jn} + F_{jn}T\zeta_n)}} \tag{11}$$

Since ζ_n is unknown, the unconditional probability takes the following form:

$$P(i) = \int_{\zeta} \Lambda(i|\zeta_n) \prod_{m=1}^M \phi(\zeta_m) d\zeta \tag{12}$$

where $\phi(\zeta_m)$ is the standard univariate normal density function, and $\prod_{m=1}^M \phi(\zeta_m)$ represents the joint density function of ζ . Since the probability function does not have a closed form, it is approximated through simulation:

$$\hat{P}(i) = \frac{1}{D} \sum_{d=1}^D \Lambda(i|\zeta_n^d) \tag{13}$$

where D is the number of simulation draws and ζ^d denotes draw d from the distribution of ζ . In this study, the factor analytic specification takes into account the effect of anchor points on route choice decisions and corresponds to bridges connecting Montreal to Laval. These factors capture the unobserved similarities among routes crossing the same anchor points. Accordingly, the F_n matrix is defined to be the route-bridge incident matrix with a dummy variable for each bridge, equal to 1 if a route crosses that particular bridge, and 0 otherwise.

Comparative models specification

The two presented anchor-based nested formulations are compared with three other models representing route-based and anchor-based formulations, namely the PSL, EPSL and IAL models. A concise introduction to these models follows.

Path-size Logit

First, a PSL model is estimated as an instance of route-based models (Prato 2009; Ben-Akiva and Bierlaire 1999). This model uses a correction factor in the deterministic part of the utility function to account for the correlation among sampled paths; however, it ignores the correlation with non-sampled paths:

$$P_{PSL}(i|C_n) = \frac{e^{\mu(V_{in} + \ln PS_{in})}}{\sum_{j \in C_n} e^{\mu(V_{jn} + \ln PS_{jn})}} \tag{14}$$

where $P_{PSL}(i|C_n)$ is the conditional probability of user n choosing alternative i from the universal choice set C_n , μ is a scale factor, and V is the deterministic part of the utility function. The Path-Size factor $\ln PS$ is added in a logarithmic scale to the deterministic part and is calculated as below:

$$PS_{in} = \sum_{a \in \Gamma_i} \frac{L_a}{L_i} \frac{1}{\sum_{j \in \varphi_n} \delta_{aj}} \tag{15}$$

where L_a and L_i represent the length of link a and path i , Γ_i is the set of road segments in path i , φ_n denotes the considered choice set, and δ_{aj} is the link-path incident binary variable which is 1 if link a is on path i , and 0 otherwise. In other words, $\sum_{j \in \varphi_n} \delta_{aj}$ indicates the total number of alternatives in the choice set sharing link a , for observation in φ_n .

Extended-path-size Logit

Second, an EPSL model (Frejinger et al. 2009) has been estimated which is also an instance of route-based models, in which the PS factor has been extended to take into account the correlation of each alternative with all the possible paths in the true choice set. However, the structure of the conditional probability stays the same:

$$P_{EPSL}(i|C_n) = \frac{e^{\mu(V_{in} + \ln EPS_{in}) + \ln \left(\frac{k_{in}}{q(i)} \right)}}{\sum_{j \in C_n} e^{\mu(V_{jn} + \ln EPS_{jn}) + \ln \left(\frac{k_{jn}}{q(j)} \right)}} \tag{16}$$

and the EPS factor is defined by

$$EPS_{in} = \sum_{a \in \Gamma_i} \frac{L_a}{L_i} \frac{1}{\sum_{j \in \varphi_n} \delta_{aj} \omega_{jn}} \tag{17}$$

where ω_{jn} is an extension factor with a value equal to 1 if $\delta_{aj} = 1$ or $q(j) R_n \geq 1$, and $1/(q(j)R_n)$ otherwise; where R_n denotes the total number of paths drawn with replacement from the universal choice set, $q(j)$ is the sampling probability of path j , and k_{jn} is the empirical frequency or the actual number of times path j is drawn.

Independent availability Logit

Third, an IAL model (Swait and Ben-Akiva 1987) has been estimated to illustrate the performance of anchor-based models. In this formulation the choice set is latent and the

probability of considering any combinations of alternative as the final choice set is calculated. The conditional probability of alternative i being chosen is calculated as

$$P_i^{IAL} = \sum_{D \subseteq \Gamma} P_D P_{i|D} = \sum_{D \subseteq \Gamma} \frac{\prod_{i \in D} A_i \prod_{i \in C-D} (1 - A_i)}{1 - \prod_{i \in C} (1 - A_i)} \left(\frac{\exp(\mu V_i)}{\sum_{j \in D} \exp(\mu V_j)} \right) \tag{18}$$

where P_D is the probability of drawing the choice set D from a set of all possible non-empty choice sets Γ of the universal choice set C ; $P_{i|D}$ denotes the probability of choosing alternative i from the choice set D ; and $A_i = (1 + \exp(-\alpha x))^{-1}$, where x denotes attributes and α refers to parameters to be estimated. In order to achieve the proposed formulation for P_D , it is assumed that the IIA property holds for alternatives in the considered choice set.

Utility function specification

Four attributes are used to specify the systematic part of the utility function:

- Mtl_Len specifies the portion of trip length made on the island of Montreal,
- Lvl_Len denotes the portion of trip length made on the island of Laval,
- Hgw_Len stands for the portion of trip length made on highways, and
- Seg_Len indicates the average length of road segments.²

The minimum, maximum, average and median values of these attributes over the whole dataset are reported in Table 3.

In estimating NL, LK, PSL, and EPSL models, the utility function for observation i is defined to be:

$$V_i = \beta_{Mtl_Len} \times Mtl_Len_i + \beta_{Lvl_Len} \times Lvl_Len_i + \beta_{Hgw_Len} \times Hgw_Len_i + \beta_{Seg_Len} \times Seg_Len_i \tag{19}$$

For the IAL model, the length of the trip made on the island of Montreal, which practically specifies the distance from the origin to the bridge, is used in the first part of the model to determine the selection probability of each choice sets:

$$A_i = (1 + \exp(-\beta_{Mtl_Len} \times Mtl_Len_i))^{-1} \tag{20}$$

The other three variables are used to define the systematic part of the utility function:

$$V_i^{IAL} = \beta_{Lvl_Len} \times Lvl_Len_i + \beta_{Hgw_Len} \times Hgw_Len_i + \beta_{Seg_Len} \times Seg_Len_i \tag{21}$$

Choice set generation

The consideration set should include attractive alternatives. Since random sampling of alternatives in large universal choice sets is not efficient in terms of providing information, an *importance sampling* method would be more convenient and favorable (Hess and Daly 2010). Several deterministic and probabilistic path generation methods which are mostly based on repeated shortest path algorithms have been proposed in the literature to form the consideration set. Among them are link labelling (Ben-Akiva et al. 1984), link elimination

² Road segments are defined to be the portion of a road between two consecutive junctions.

Table 3 Statistics on attributes included in the systematic part of the utility function

Attributes	Min.	Max.	Avg.	Med.
Mtl_Len (m)	266.6	22,846.3	7459.9	6347.6
Lvl_Len (m)	117.5	23,474.8	3729.0	2414.6
Hgw_Len (m)	0.0	33,694.9	7381.6	3956.1
Seg_Len (m)	89.0	621.3	231.2	195.5

(Azevedo et al. 1993), and link penalty (de la Barra et al. 1993) methods. A thorough review of these methods is presented in (Prato 2009) and (Frejinger and Bierlaire 2010). The major downside of using these methods is that they do not provide researchers with sampling probabilities of the generated alternatives. Model estimates based on these path generation methods are biased, unless the sampling probability of every alternative in the universal set is equal, which is not the case in route choice modeling.

Several alternative approaches have been proposed. Cascetta et al. (2002) adopted a two stage process, where in the first step, a complete set of alternatives is generated for all the observations by maximizing a coverage factor between the generated set and the set of routes perceived as available. Then, in the second step, a binomial Logit model is adopted to estimate the probability of including a given route in the users' consideration set. Frejinger et al. (2009) applied a biased random walk to sample a subset of paths and derived a sampling correction to obtain unbiased parameter estimates. More recently, Flötteröd and Bierlaire (2013) used a Metropolis–Hastings (MH) algorithm to generate sample sets based on an arbitrary distribution providing the sampling probability of each alternative. This algorithm requires a road network and a definition of path weight as an input. It uses an underlying Markov Chain process to sample alternatives and calculates its sampling probability without the need of normalizing it over the full choice set.

For the NL model estimated in this study, the MH algorithm has been adopted to generate nine alternatives per nest. In order to apply this algorithm on the large road network of Montreal and Laval, 100 separate input files have been prepared to provide the possibility of parallel calculation. Files have been imported into a cluster of 26 computers (2 processors Intel(R) Xeon(R) X5675 @ 3.07 GHz) which took about 104 h (4 days and 8 h) to generate the output files. For the LK, PSL, and EPSL models, the MH algorithm has been adopted to draw 19 choice alternatives from the universal choice set. Similarly to the NL model, 100 separate input files have been prepared to provide the possibility of parallel calculation, which resulted in a calculation time of 46 h (1 day and 22 h).

For the IAL model, a shortest path algorithm, using segments' speed limits as travel cost, has been adopted. The same algorithm was used in the original application by Habib et al. (2013) and provides the possibility of comparison between the outputs. In the IAL formulation, the feasible choice set built in the choice generation step is considered to be the equivalent of the universal choice set C . Since the algorithm calculates the choice probability of every non-empty subset of the universal choice, the number of considered alternatives should be restricted for computational purposes. The chosen alternative and eight shortest-distance paths, crossing the eight alternative bridges, comprise the nine feasible alternatives for each observation. Although Habib et al. (2013) showed that the IAL performs well in anchor choice prediction, it is very computationally expensive for applications involving large choice sets, such as route choice modeling applications.

In this study, different sizes of choice sets have been generated for practicality reasons. Considering larger choice sets would have increased the computational time dramatically;

however, similarly to Elgar et al. (2009) the estimation gain would have probably been minor. In all the above mentioned applications, the chosen alternative has been added to the choice set, where it was not generated by the adopted choice set generation algorithm (McFadden 1978; Elgar et al. 2009; Arifin 2012; Prato et al. 2012; Frejinger et al. 2009; Habib et al. 2013; Dhakar and Srinivasan 2014; Hess et al. 2015).

Results and discussion

This section presents, compares and discusses the estimation and prediction abilities of the aforementioned models. The BIOGEME software package (Bierlaire and Feterarison 2009; Bierlaire 2003) has been used for all model estimations.

Estimation

Eighty percent of the observations, that is 434 trips, were randomly selected for estimation purposes. A heat-map of origin and destination points, presented in Fig. 2, illustrates higher density regions, located mostly around metro stations, airport, downtown Montreal, and commercial centers, which have an expectedly higher taxi demand. A heat-map of chosen routes between OD pairs is illustrated in Fig. 3. Around 40% of the whole network size, considered for choice set generation, has been covered by drivers' route choices.

A detailed description of models' estimates is provided in Table 4. Presented models are fully identified according to the smallest singular value approach implemented in BIOGEME (Bierlaire 2015). The scale parameter in PSL, EPS and LK models were estimated while they have been fixed to 1 in IAL and NL models for identification purposes. Scale factors for nests in the NL model have been estimated and $\mu \leq \mu_m$ holds for every nest. Since the composition of the choice set differs from a model to another, not much can be inferred from the comparison of their scale factors. However, an out-of-sample validation method has been used to properly compare the models' performances, which will be presented and discussed in the subsequent section.

Intuitively, taxi drivers are apt to minimize their travel distance by choosing a shorter route, and their travel time by riding on segments with higher speed limits. This behavior is confirmed by the obtained results from all the estimated models. Coefficients β_{Mtl_Len} and β_{Lvl_Len} are negative while β_{Hgw_Len} has a positive sign, meaning that taxi drivers are more willing to take shorter alternatives and are more inclined to ride on highways. These findings are in agreement with results reported by Duan and Wei (2014) who claimed that most taxi drivers tend to minimize their travel time. The effect of the average length of the segment is expectedly positive for PSL, EPSL, LK, and NL models, implying that taxi drivers tend to avoid intersections and prefer to take routes with a longer average segment length. However, this estimate has a negative sign for the IAL model, which might be attributed to the fact that alternatives are assumed to be completely independent from each other and their correlations have been neglected. The positive signs of β_{PS} and β_{EPS} are a negative correction of the utility for overlapping routes, giving a higher chance to less similar alternatives to be chosen. Similar findings are reported by (Dhakar and Srinivasan 2014; Prato and Bekhor 2006, 2007; Bierlaire and Frejinger 2008).

Note that the anchor-based nested models, namely LK and NL models, result in significantly higher Rho-square values compared to the route-based and anchor-based models. This emphasizes the importance of bringing the concept of anchor points in route choice

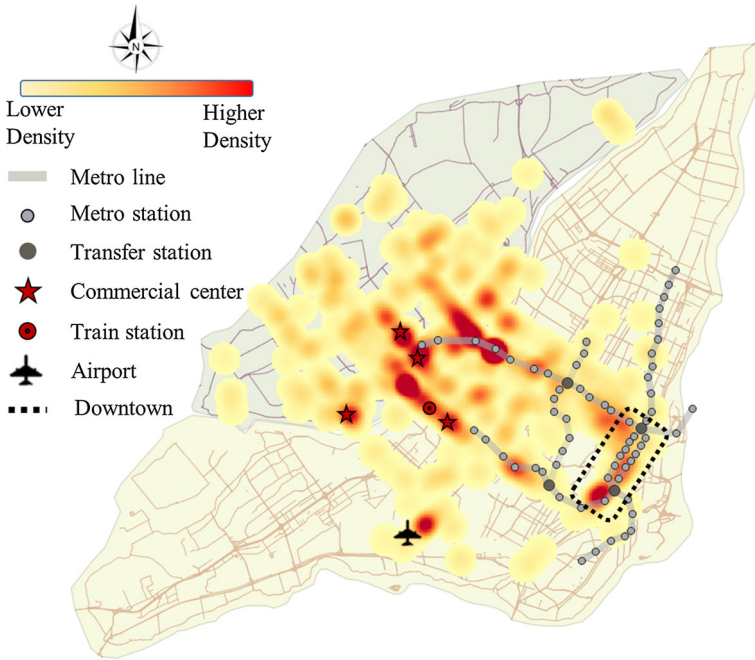


Fig. 2 Heat-map of origin and destination points for trips between Montreal and Laval

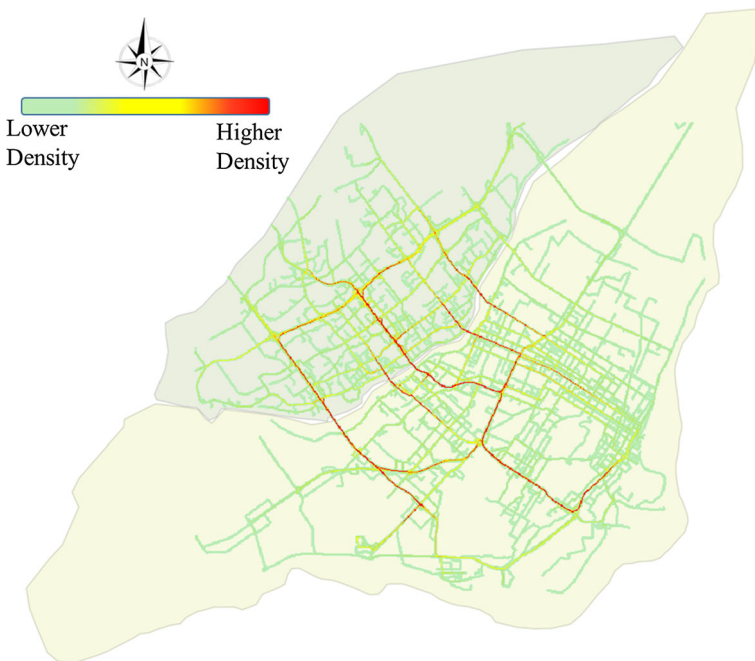


Fig. 3 Heat-map of chosen routes between OD pairs for trips between Montreal and Laval

Table 4 Model estimates for taxi trips between Montreal and Laval

Parameters	PSL		EPSL		IAL		LK		NL	
	Est. ^a	tt. ^b	Est.	tt.	Est.	tt.	Est.	tt.	Est.	tt.
$\beta_{Mut_Len}^c$	-0.0009	-10.9	-0.0048	-4.5	-0.0177	-8.4	-0.0012	-5.7	-0.0016	-9.2
$\beta_{Lst_Len}^c$	-0.0002	-3.1	-0.0037	-5.1	-0.0006	-8.8	-0.0003	-2.2	-0.0015	-7.0
$\beta_{Hgw_Len}^c$	0.0003	2.4	0.0028	4.8	0.0078	9.5	0.0005	2.7	0.0008	4.3
$\beta_{Srg_Len}^c$	0.1070	6.2	0.1910	6.8	-0.1970	-8.3	0.1020	11.6	0.0455	6.2
β_{PS}	4.1	6.2								
β_{EPS}			3.57	4.5			2.76	11.5	0.887	3.4
ASC_Brg2					33.6	11.2				
ASC_Brg3					4	-				
ASC_Brg4					18.3	13.2				
ASC_Brg5					-18.8	-4.5				
ASC_Brg6					25	-				
Mu	0.66	7.1	0.38	6.9	1	-	0.65	7.4	1	-
σ_{Brg1}							0.48	2.7		
σ_{Brg2}							-1.12	-2.8		
σ_{Brg3}							-1.81	-4.8		
σ_{Brg4}							0.96	2.9		
σ_{Brg5}							4.35	2.3		
σ_{Brg6}							1.76	1.5		
Mu_Brg2									3.78	17.1
Mu_Brg3									1.08	4.9
Mu_Brg4									3.73	10.6
Mu_Brg5									1.52	6.1
Mu_Brg6									1.42	5.1
Initial LL	-2079.8		-2180.6		-1151.2		-5201.0		-2404.7	

Table 4 continued

Parameters	PSL		EPSL		IAL		LK		NL	
	Est. ^a	tt. ^b	Est.	tt.	Est.	tt.	Est.	tt.	Est.	tt.
Final LL	-623.9		-606.7		-341.4		-747.7		-336.8	
Rho-square	0.697		0.719		0.697		0.854		0.855	
Est. time (s)	<1		<1		514		2417		123	

^a Estimated value

^b Robust *t* test

^c Length has been measured in meters

modeling, so that it becomes more consistent with the actual behavior of drivers. The σ estimates are highly significant (except for bridge 6) for the LK model, implying that the factor analytic structure captures a significant correlation structure between routes crossing the same bridges. It is also noted that this effect is statistically different from the effect captured by the EPS factor. This is consistent with findings in Bekhor et al. (2002) and Bierlaire and Frejinger (2005), where the authors presented a LK route choice model considering subpath components. The better fit of the LK model with an EPS attribute over the PSL and EPSL models is in line with findings reported by Ramming (2001) and Bierlaire and Frejinger (2005). It is worth mentioning that the travel time-based shortest path algorithm, used in the choice set generation step of the IAL model, does not provide the researcher with the sampling probability of paths in order to correct the sampling effect. The better fit of all other models can be partially attributed to the application of MH algorithm, which provides the possibility of considering the sampling correction factor.

The estimation time has also been reported in Table 4. All model estimations were conducted on a machine with a core i7-4720HQ CPU running at 2.6 GHz and a Random Access Memory (RAM) of 16.0 GB. Concerning PSL and EPSL models, the computational time was found to be less than 1 s which is probably due to their simple multinomial logit structure. The small number of considered alternatives per observation may be a further reason for the simplicity of their calculation. However, there is a large difference in computational costs between these models and the IAL model. Although the number of feasible alternatives for the IAL model is limited to 9, the high computational cost might be related to the fact that every possible non-empty subset ($2^9 - 1 = 511$ subsets) must be considered for every observation. Implementing a NL structure reduces the calculation time substantially, compared to the IAL model, by providing a more realistic structure to represent drivers' route choice behavior. It is behaviorally not realistic and computationally not feasible to assume that drivers consider 511 non-empty subsets in order to make a choice from a set of 9 nine alternatives. However, the increase in computational time compared to the PSL and EPSL models might be explained by the more complex structure of NL compared to MNL, and the greater number of alternatives considered for the NL model (82 alternatives per observation compared to 20 alternatives for PSL and EPSL models). Expectedly, the largest estimation time is recorded for the LK model. The normally distributed portion of the disturbance requiring large number of draws in model estimation, leads to a computationally demanding model (Ben-Akiva et al. 2001; Walker et al. 2004).

Validation

In order to further compare these models, their ability to predict should also be evaluated in the final stage. An out of sample validation has been performed to evaluate the prediction capability of the estimated models. Twenty percent of the observations (109 trips), which have not been used for model estimation, have been randomly sampled for this purpose. The validation has been performed based on models' abilities to correctly predict:

1. The chosen bridge,
2. The chosen route, and
3. The total overlapping percentage with chosen alternatives (coverage rate).

The aforementioned three indicators have been calculated and results are illustrated in Fig. 4. Part A of Fig. 4 compares models' performances in terms of correctly predicting the taken route. It clearly shows that LK and NL perform better than the three other models

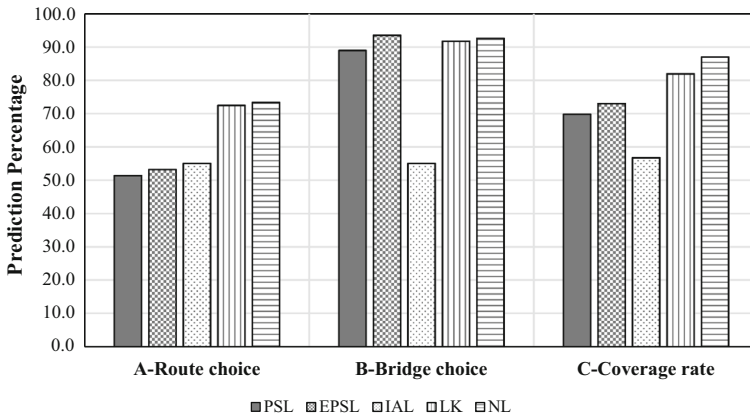


Fig. 4 Comparison of models' prediction abilities

with a prediction rate of 72.5 and 73.4% compared to 51.4, 53.2 and 55% for the PSL, EPSL and IAL models, respectively. The ability to correctly predict the chosen bridge is compared in part B of Fig. 4. Similarly, LK and NL outperform PSL and IAL with prediction rates of 91.8 and 92.7% compared to 89 and 55%, respectively. The performance of EPSL is roughly similar to the anchor-based nested formulations with a small difference of around 1.0%, which may be attributed to simulation errors. In order to understand how closely each model has predicted the chosen route, we have calculated the length percentage that has been correctly predicted by each model. This coverage rate is compared in part C of Fig. 4, and clearly demonstrates the superiority of anchor-based nested formulations.

Since IAL model is an anchor-based model, it was expected to perform better in bridge choice prediction with respect to studied route-based models. The poor performance of this model, in this study, in terms of both estimation and prediction might be attributed to the fact that a fixed utility function has been used to compare the four models. Using more explanatory variables describing bridge characteristics in Eq. (18) might have improved both the model's fit over data and its prediction abilities.

Based on the abovementioned results, we conclude that the proposed anchor-based nested approaches outperform route-based as well as anchor based models. This conclusion is in line with findings reported by (Manley et al. 2015a; Kazagli and Bierlaire 2015; Habib et al. 2013; Manley et al. 2015b; Prato and Bekhor 2007) who claimed that anchor points have an important effect on individuals' route selection behavior. However, the important aspect of this study, which is the estimation of the comprehensive effect of both route-level attributes and anchor points, clearly demonstrates that considering the role of bridges as anchor points in conjunction with route-level attributes, for trips between Montreal and Laval, enhances both the estimation and prediction abilities of the model.

Conclusions

In this paper we have explored the application of a nested structure to improve the behavioral aspect of route choice modeling by incorporating the effect of space hierarchy in drivers' decision making process. We have argued that current approaches have either

neglected the effect of anchor points by considering only route level attributes (route-based formulations), or have adopted a full probabilistic approach for choice set generation, which is behaviorally inaccurate, and theoretically impractical to implement in real world networks (anchor-based formulations).

The anchor-based nested approaches, proposed in this paper, attempt to improve the behavioral aspect of route choice modeling by incorporating the effects of anchor points and route level attributes at the same time. As previous studies have shown (Habib et al. 2013; Sun et al. 2014; Woo et al. 2015), in riverside cities such as Montreal, route choice decisions are highly influenced by their respective bridge choices. Moreover, routes crossing a same bridge share unobserved components such as safety, scenery, driving comfort, etc., which is mainly because they share the same network and geographical characteristics. In this study, we capture this reciprocal effect through a nested structure.

GPS data provided by a taxi company in Montreal has been used to study trips made between the island cities of Montreal and Laval. For these trips, drivers have to choose between a maximum of nine bridges, providing plausible route alternatives between these cities. These bridges face recurrent congestion and play an important role in drivers' route decisions due to their high travel time variability. To address the role of bridges as anchor points, a discrete choice utility maximization framework has been adopted. First, a NL formulation has been proposed, in which upper nests represent bridges and lower nests consist of route alternatives crossing respective bridges. Second, a nested LK with a factor analytic structure is specified. The EPS factor has been added to the deterministic part of the utility functions of these models to account for physical overlap among routes crossing the same bridge. The unobserved similarities among these routes are captured through the nested structure and the factor analytic structure in NL and LK models, respectively.

To evaluate the performance of the proposed anchor-nested formulations, they are compared to the recent route- and anchor-based models, namely PSL, EPSL and IAL models. For the sake of simplicity in estimation and comparison, four most important variables have been selected to define a common utility function between models. Findings revealed that the nested structures provided better model fits and underscored the importance of considering the comprehensive effect of anchor points and route level attributes in route choice decisions. Results have expectedly illustrated that taxi drivers are more likely to drive on highways and tend to decrease their travelled distance. It has also been found that they prefer to avoid intersections and tend to drive on routes with a higher average segment length. The predictive ability of these models has also been compared by an out-of-sample validation approach. Three indicators have been used for this purpose, namely the number of correctly predicted routes, the number of correctly predicted bridges, and the overlap percentage between the predicted and chosen alternatives. The overall results suggest that LK and NL outperform the other three models in the validation step.

In short, incorporating the effects of anchor points in a nested structure has the following advantages over the previously studied anchor based model: (1) the nested structure improves the behavioral aspect of decision making process. In the conventional anchor based model (IAL), a probability is assigned to every subset of the universal choice set, and the conditional probability of an alternative being chosen implies that the decision maker has considered every possible combinations of alternatives as his final choice set, which is behaviorally unrealistic, (2) LK and NL models are easily manageable and practical, even by considering a large number of alternatives (McFadden 1978). IAL incorporates a full probabilistic choice set generation approach which is inapplicable in route choice modeling (Prato 2009). For instance, considering a small dataset of 10 alternatives, a selection probability has to be calculated for every 1023 non-empty subsets of alternatives which is

very time-consuming and impractical, and (3) the nested structure allows the consideration of multiple anchor points and their effects on route choice decisions. It is worth mentioning that the inclusion of multiple landmarks and anchor-points, and the consideration of several forms of heterogeneity, such as decision makers' taste variations, are much more manageable and can be accommodated more easily in LK than in the NL mode. This is due to the flexible structure of the error term, which can approximate almost any desirable error structure (Walker et al. 2004).

This study contributes to the existing literature in two ways. First, it improves the behavioral, theoretical, and practical aspects of anchor-based route choice models by capturing the effects of both anchor points and route level attributes within a nested choice model framework, and clearly underscores the importance of considering the effects of anchor points in conjunction with route-level attributes. Second, a large real-world road network, consisting of over 40,000 nodes and 19,000 links, has been studied and a MH algorithm has been adopted to generate a set of considered alternatives. To the best of the authors' knowledge, the largest network previously tested on this algorithm was composed of about 8000 nodes and 17,000 links (Flötteröd and Bierlaire 2013). The major advantage of MH sampling algorithm over conventional methods (e.g. link labelling, link elimination, etc.) is that it provides researchers with path sampling probabilities, so that model estimates based on these sets are not biased. It is noted that in route-based models and most of the link-based formulations, the consideration set is commonly generated using shortest-path algorithms with some pre-defined impedance function, which do not provide path sampling probabilities, and hence do not account for the correlation between sampled and non-sampled paths, resulting in biased estimates.

It is worth mentioning that since this paper is based on a dataset covering a fraction of taxi fleets operating in Montreal, results may not be directly transferable to other car drivers in Montreal or any other similar regions in the world. Additional datasets from different contexts and population segments are needed to provide more insights on this subject. However, it reveals the undeniable effect of anchor points on the decision of the whole path, and provides valuable insights regarding drivers' route selection behavior.

As the core of traffic assignment methods, a more realistic route choice model improves the travel demand assessment on the road network. An application instance of the proposed models would be their utilities in predicting drivers' behavior under hypothetical situations, and the way drivers react to different policies and changes. For example, the effect of a temporary lane closure on one of the bridges can be assessed on other bridges and route segments.

For future works, it would be interesting to investigate the spatial and temporal transferability of the proposed structure for different datasets on similar case studies. Also, more interesting structures such as multilevel nested models can be estimated to explore the effects of multiple anchor points on route choice decisions. In this work, we have neglected the effect of shared segments between routes crossing different bridges (alternatives in different nests), to accommodate the IIA property of the NL model. The inclusion of a correction factor accounting for this shared similarity might improve estimation results.

Route choice is also influenced by travel time and congestion. A shortcoming of this study is that travel time related attributes have been neglected in this study due to data availability issues. In order to improve models' estimation and prediction abilities, it is recommended to consider travel time related attributes on bridges and route segments. Furthermore, including physical characteristics of bridges can also be interesting and can provide useful insights on their effects on the attractiveness of an alternative. It is expected that the inclusion of these factors will enhance models' estimation and prediction abilities.

Another appealing aspect to investigate would be the incorporation of some socio-demographic and behavioral factors which were not available in our dataset. Also, incorporating the role of dynamic information, knowledge, and level of experience would add an interesting aspect to this modeling process. A recent study by Vitetta (2016) explores a new realm of route choice models called Quantum Utility Model (QUM), which captures the effect of intermediate (during the trip) decisions, where decision makers are uncertain about their final choices. A comparison study between the proposed approach in this study and the study by Vitetta (2016) might also be interesting as a future expansion of this work. Another interesting area to explore would be the extension of the Recursive Logit (RL) model proposed by Fosgerau et al. (2013) by incorporating the effects of anchor points.

Acknowledgement The authors are grateful to Gunnar Flötteröd for his invaluable help with the implementation of MH algorithm. We also acknowledge collaborators from Taxi Diamond who provided access to data for research purpose, as well as the anonymous reviewers for their helpful comments and suggestions. We would also thank the Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT) for providing the possibility of using their cluster computer for our research funded by the Natural Sciences and Engineering Council of Canada (NSERC) Research Tools and Instruments grant.

References

- Arifin, Z.N.: Route choice modeling based on GPS tracking data. Diss., Eidgenössische Technische Hochschule ETH Zürich, Nr. 20400, (2012)
- Azevedo, J., Costa, M.E.O.S., Madeira, J.J.E.S., Martins, E.Q.V.: An algorithm for the ranking of shortest paths. *Eur. J. Oper. Res.* **69**(1), 97–106 (1993)
- Bekhor, S., Ben-Akiva, M., Scott Ramming, M.: Adaptation of logit kernel to route choice situation. *Transp. Res. Rec. J. Transp. Res. Board* **1805**, 78–85 (2002)
- Ben-Akiva, M.: Structure of Passenger Travel Demand Models. Massachusetts Institute of Technology, Cambridge (1973)
- Ben-Akiva, M., Bergman, M., Daly, A.J., Ramaswamy, R.: Modeling inter-urban route choice behaviour. In: Volmuller, J., Hamerslag, R. (eds.) *Proceedings of the 9th International Symposium on Transportation and Traffic Theory*, pp. 299–330. VNU Press, Utrecht (1984)
- Ben-Akiva, M., Bierlaire, M.: Discrete choice methods and their applications to short term travel decisions. In: Hall, R.W. (ed.) *Handbook of Transportation Science*, pp. 5–33. Springer (1999)
- Ben-Akiva, M., Bierlaire, M.: Discrete choice models with applications to departure time and route choice. In: Hall, R.W. (ed.) *Handbook of transportation science*, p 32 (2003)
- Ben-Akiva, M., Bolduc, D., Walker, J.: Specification, Identification and Estimation of the Logit Kernel (or Continuous Mixed Logit) Model. Department of Civil Engineering Manuscript, MIT, Cambridge (2001)
- Bierlaire, M.: BIOGEME: a free package for the estimation of discrete choice models. In: *Swiss Transport Research Conference*, vol. TRANSP-OR-CONF-2006-048 (2003)
- Bierlaire, M.: BisonBiogeme 2.4: estimating a first model. *Series on Biogeme TRANSP-OR* **150720** (2015)
- Bierlaire, M., Bolduc, D., McFadden, D.: The estimation of generalized extreme value models from choice-based samples. *Transp. Res. Part B Methodol.* **42**(4), 381–394 (2008). doi:[10.1016/j.trb.2007.09.003](https://doi.org/10.1016/j.trb.2007.09.003)
- Bierlaire, M., Fetiariison, M.: Estimation of discrete choice models: extending BIOGEME. In: *Swiss Transport Research Conference (STRC)* (2009)
- Bierlaire, M., Frejinger, E.: Route choice models with subpath components. In: *Swiss Transportation Research Conference*, vol. TRANSP-OR-CONF-2006-032 (2005)
- Bierlaire, M., Frejinger, E.: Route choice modeling with network-free data. *Transp. Res. Part C Emerg. Technol.* **16**(2), 187–198 (2008)
- Bolduc, D., Ben-Akiva, M.: A multinomial probit formulation for large choice sets. In: *Proceedings of the 6th International Conference on Travel Behaviour*, pp. 243–258 (1991)
- Cascetta, E., Nuzzolo, A., Russo, F., Vitetta, A.: A modified logit route choice model overcoming path overlapping problems: specification and some calibration results for interurban networks. In:

- Proceedings of the 13th International Symposium on Transportation and Traffic Theory, pp. 697–711. Pergamon Oxford, NY, USA (1996)
- Cascetta, E., Russo, F., Viola, F.A., Vitetta, A.: A model of route perception in urban road networks. *Transp. Res. Part B Methodol.* **36**(7), 577–592 (2002)
- Communauté métropolitaine de Montréal, C. (2012) An attractive, competitive and sustainable greater Montréal. In: Library and Archives Canada Montréal, Québec
- Connors, R.D., Sumalee, A.: A network equilibrium model with travellers' perception of stochastic travel times. *Transp. Res. Part B Methodol.* **43**(6), 614–624 (2009). doi:[10.1016/j.trb.2008.12.002](https://doi.org/10.1016/j.trb.2008.12.002)
- Couclelis, H., Golledge, R.G., Gale, N., Tobler, W.: Exploring the anchor-point hypothesis of spatial cognition. *J. Environ. Psychol.* **7**(2), 99–122 (1987)
- de la Barra, T., Perez, B., Anez, J.: Multidimensional path search and assignment. In: PTRC Summer Annual Meeting, 21st, 1993, University of Manchester, United Kingdom (1993)
- Dhakar, N., Srinivasan, S.: Route choice modeling using GPS-based travel surveys. *Transp. Res. Rec. J. Transp. Res. Board* **2413**, 65–73 (2014)
- Dougherty, M.: A review of neural networks applied to transport. *Transp. Res. Part C Emerg. Technol.* **3**(4), 247–260 (1995). doi:[10.1016/0968-090X\(95\)00009-8](https://doi.org/10.1016/0968-090X(95)00009-8)
- Duan, Z., Wei, Y.: Revealing taxi driver route choice characteristics based on GPS data. In: CICTP 2014@ sSafe, Smart, and Sustainable Multimodal Transportation Systems, pp. 565–573 (2014)
- Elgar, I., Farooq, B., Miller, E.: Modeling location decisions of office firms: introducing anchor points and constructing choice sets in the model system. *Transp. Res. Rec. J. Transp. Res. Board* **2133**, 56–63 (2009)
- Elgar, I., Farooq, B., Miller, E.J.: Simulations of firm location decisions: replicating office location choices in the Greater Toronto Area. *J. Choice Modell.* **17**, 39–51 (2015). doi:[10.1016/j.jocm.2015.12.003](https://doi.org/10.1016/j.jocm.2015.12.003)
- Flötteröd, G., Bierlaire, M.: Metropolis-Hastings sampling of paths. *Transp. Res. Part B Methodol.* **48**, 53–66 (2013)
- Foo, P., Warren, W.H., Duchon, A., Tarr, M.J.: Do humans integrate routes into a cognitive map? Map-versus landmark-based navigation of novel shortcuts. *J. Exp. Psychol. Learn. Mem. Cogn.* **31**(2), 195 (2005)
- Fosgerau, M., Frejinger, E., Karlstrom, A.: A link based network route choice model with unrestricted choice set. *Transp. Res. Part B Methodol.* **56**, 70–80 (2013). doi:[10.1016/j.trb.2013.07.012](https://doi.org/10.1016/j.trb.2013.07.012)
- Frejinger, E., Bierlaire, M.: On path generation algorithms for route choice models. In: Hess, S., Daly, A. (eds.) *Choice Modelling: The State-of-the-Art and the State-of-Practice*, pp. 307–315 (2010)
- Frejinger, E., Bierlaire, M., Ben-Akiva, M.: Sampling of alternatives for route choice modeling. *Transp. Res. Part B Methodol.* **43**(10), 984–994 (2009)
- Gao, S., Frejinger, E., Ben-Akiva, M.: Adaptive route choices in risky traffic networks: a prospect theory approach. *Transp. Res. Part C Emerg. Technol.* **18**(5), 727–740 (2010). doi:[10.1016/j.trc.2009.08.001](https://doi.org/10.1016/j.trc.2009.08.001)
- Golledge, R.G., Smith, T.R., Pellegrino, J.W., Doherty, S., Marshall, S.P.: A conceptual model and empirical analysis of children's acquisition of spatial knowledge. *J. Environ. Psychol.* **5**(2), 125–152 (1985)
- Guevara, C.A., Ben-Akiva, M.E.: Sampling of alternatives in multivariate extreme value (MEV) models. *Transp. Res. Part B Methodol.* **48**, 31–52 (2013)
- Habib, K.N., Morency, C., Trépanier, M., Salem, S.: Application of an independent availability logit model (IAL) for route choice modelling: considering bridge choice as a key determinant of selected routes for commuting in Montreal. *J. Choice Modell.* **9**, 14–26 (2013)
- Henn, V.: Route choice making under uncertainty: a fuzzy logic based approach. In: Verdegay, J.-L. (ed.) *Fuzzy Sets Based Heuristics for Optimization*, vol. 126. *Studies in Fuzziness and Soft Computing*, pp. 277–292. Springer, Berlin Heidelberg (2003)
- Hess, S., Daly, A.J.: *Choice Modelling: The State-of-the-Art and the State-of-Practice*. In: *Proceedings from the Inaugural International Choice Modelling Conference: The State-of-the-art and the State-of-practice: Proceedings from the Inaugural International Choice Modelling Conference*. Emerald Group Publishing (2010)
- Hess, S., Quddus, M., Rieser-Schüssler, N., Daly, A.: Developing advanced route choice models for heavy goods vehicles using GPS data. *Transp. Res. Part E Logist. Transp. Rev.* **77**, 29–44 (2015)
- Hirtle, S.C., Jonides, J.: Evidence of hierarchies in cognitive maps. *Mem. Cognit.* **13**(3), 208–217 (1985)
- Holding, C.S.: Further evidence for the hierarchical representation of spatial information. *J. Environ. Psychol.* **14**(2), 137–147 (1994). doi:[10.1016/S0272-4944\(05\)80167-7](https://doi.org/10.1016/S0272-4944(05)80167-7)
- Kahneman, D., Tversky, A. (1979) Prospect theory: An analysis of decision under risk. *Econom. J. Econom. Soc.* 263–291

- Kaplan, S., Prato, C.G.: Closing the gap between behavior and models in route choice: the role of spatiotemporal constraints and latent traits in choice set formation. *Transp. Res. Part F Traffic Psychol. Behav.* **15**(1), 9–24 (2012). doi:[10.1016/j.trf.2011.11.001](https://doi.org/10.1016/j.trf.2011.11.001)
- Kazagli, E., Bierlaire, M.: A route choice model based on mental representations. In: *Proceedings of the 15th Swiss Transport Research Conference*, vol. EPFL-CONF-208979 (2015)
- Kim, J., Sung, S., Namung, M., Jang, Y.: Development of dynamic route choice behavioral applied intelligent system theory. In: *Proceedings of the Eastern Asia Society for Transportation Studies*, pp. 1615–1630 (2005)
- Lai, X., Bierlaire, M.: Specification of the cross-nested logit model with sampling of alternatives for route choice models. *Transp. Res. Part B Methodol.* **80**, 220–234 (2015). doi:[10.1016/j.trb.2015.07.005](https://doi.org/10.1016/j.trb.2015.07.005)
- Luisa De Maio, M., Vitetta, A.: Route choice on road transport system: a fuzzy approach. *J. Intell. Fuzzy Syst.* **28**(5), 2015–2027 (2015)
- Lynch, K.: *The Image of the City*, vol. 11. MIT press, Cambridge (1960)
- Manley, E., Addison, J., Cheng, T.: Shortest path or anchor-based route choice: a large-scale empirical analysis of minicab routing in London. *J. Transp. Geogr.* **43**, 123–139 (2015a)
- Manley, E., Orr, S., Cheng, T.: A heuristic model of bounded route choice in urban areas. *Transp. Res. Part C Emerg. Technol.* **56**, 195–209 (2015b)
- Manski, C.F.: The structure of random utility models. *Theor. Decis.* **8**(3), 229–254 (1977)
- McFadden, D.: *Modelling the Choice of Residential Location*. Institute of Transportation Studies, University of California, Berkeley (1978)
- McFadden, D., Train, K.: Mixed MNL models for discrete response. *J. Appl. Econom.* **15**(5), 447–470 (2000)
- McFadden, D.L.: Econometric analysis of qualitative response models. *Handb. Econom.* **2**, 1395–1457 (1984)
- Murat, Y.S., Uludag, N.: Route choice modelling in urban transportation networks using fuzzy logic and logistic regression methods. *J. Sci. Ind. Res.* **67**(1), 19 (2008)
- Prato, C., Bekhor, S.: Applying branch-and-bound technique to route choice set generation. *Transp. Res. Rec. J. Transp. Res. Board* **1985**, 19–28 (2006)
- Prato, C., Bekhor, S.: Modeling route choice behavior: how relevant is the composition of choice set? *Transp. Res. Rec. J. Transp. Res. Board* **2003**, 64–73 (2007)
- Prato, C.G.: Route choice modeling: past, present and future research directions. *J. Choice Modell.* **2**(1), 65–100 (2009). doi:[10.1016/S1755-5345\(13\)70005-8](https://doi.org/10.1016/S1755-5345(13)70005-8)
- Prato, C.G.: Expanding the applicability of random regret minimization for route choice analysis. *Transportation* **41**(2), 351–375 (2014)
- Prato, C.G., Bekhor, S., Pronello, C.: Latent variables and route choice behavior. *Transportation* **39**(2), 299–319 (2012)
- Quattrone, A., Vitetta, A.: Random and fuzzy utility models for road route choice. *Transp. Res. Part E Logist. Transp. Rev.* **47**(6), 1126–1139 (2011). doi:[10.1016/j.tre.2011.04.007](https://doi.org/10.1016/j.tre.2011.04.007)
- Ramming, M.S.: *Network Knowledge and Route Choice*. Massachusetts Institute of Technology, Cambridge (2001)
- Sun, H., Wu, J., Ma, D., Long, J.: Spatial distribution complexities of traffic congestion and bottlenecks in different network topologies. *Appl. Math. Model.* **38**(2), 496–505 (2014). doi:[10.1016/j.apm.2013.06.027](https://doi.org/10.1016/j.apm.2013.06.027)
- Swait, J., Ben-Akiva, M.: Incorporating random constraints in discrete models of choice set generation. *Transp. Res. Part B Methodol.* **21**(2), 91–102 (1987)
- Train, K.E.: *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge (2009)
- Tversky, A., Kahneman, D.: Advances in prospect theory: cumulative representation of uncertainty. *J. Risk Uncertain.* **5**(4), 297–323 (1992)
- Vitetta, A.: A quantum utility model for route choice in transport systems. *Travel Behav. Soc.* **3**, 29–37 (2016). doi:[10.1016/j.tbs.2015.07.003](https://doi.org/10.1016/j.tbs.2015.07.003)
- Vovsha, P., Bekhor, S.: Link-nested logit model of route choice: overcoming route overlapping problem. *Transp. Res. Rec. J. Transp. Res. Board* **1645**, 133–142 (1998)
- Walker, J., Ben-Akiva, M., Bolduc, D. (2004) Identification of the logit kernel (or mixed logit) model. In: *10th International Conference on Travel Behavior Research*, Lucerne, Switzerland
- White, C.E., Bernstein, D., Kornhauser, A.L.: Some map matching algorithms for personal navigation assistants. *Transp. Res. Part C Emerg. Technol.* **8**(1), 91–108 (2000)
- Wiener, J.M., Mallot, H.A.: ‘Fine-to-coarse’ route planning and navigation in regionalized environments. *Spat. Cogn. Comput.* **3**(4), 331–358 (2003)

- Woo, C.K., Cheng, Y.S., Li, R., Shiu, A., Ho, S.T., Horowitz, I.: Can Hong Kong price-manage its cross-harbor-tunnel congestion? *Transp. Res. Part A Pol. Pract.* **82**, 94–109 (2015). doi:[10.1016/j.tra.2015.09.002](https://doi.org/10.1016/j.tra.2015.09.002)
- Xu, H., Zhou, J., Xu, W.: A decision-making rule for modeling travelers' route choice behavior based on cumulative prospect theory. *Transp. Res. Part C Emerg. Technol.* **19**(2), 218–228 (2011). doi:[10.1016/j.trc.2010.05.009](https://doi.org/10.1016/j.trc.2010.05.009)
- Zhou, J., Golledge, R.: A three-step general map matching method in the GIS environment: travel/transportation study perspective. *Int. J. Geogr. Inf. Syst.* **8**(3), 243–260 (2006)

Hamzeh Alizadeh received his Bachelor degree in Civil Engineering, from Imam Khomeini International University, Ghazvin, Iran (2008). He pursued higher education in Traffic and Pavement Engineering at Sharif University of Technology, Tehran, Iran (2011), and ranked second among all the graduate students admitted to that program. Since 2012, he is a PhD candidate in Transportation Engineering at Polytechnique, Montréal. His current research focuses on improving the understanding of drivers' route choice behaviours and their decision making process using advanced econometrical models and machine learning methods.

Bilal Farooq received his Bachelor of Engineering degree from University of Engineering and Technology, Pakistan, Master in Computer Science degree from Lahore University of Management Sciences, Pakistan, and Ph.D. degree from University of Toronto, Canada in 2001, 2004, and 2011, respectively. From 2003 to 2005, he worked at Bentley Systems Inc. as a Software Developer and Team Lead and from 2005 to 2006 he worked as a Chief Architect at xCube Innovations, a start-up he cofounded. In 2011 to 2013 he worked at École Polytechnique Fédérale de Lausanne, Switzerland as a Research Fellow. Since 2013 he is working as an Assistant Professor and Founding Director of Laboratory of Innovations in Transportation at École Polytechnique Montréal, Canada. His current research goal is to develop multidisciplinary and highly intelligent solutions for sustainable planning, design, and operations of urban infrastructure systems of tomorrow.

Catherine Morency Catherine Morency is a civil engineer, graduated from Laval University (Quebec). She has a master and doctorate degree in transportation engineering from Polytechnique Montreal. She is full professor at the department of Civil, Geological and Mining Engineering at Polytechnique Montreal. Since 2010, she is Head of the Mobility Research Chair on the implementation of sustainability in transportation and recently was awarded a Canada Research Chair on Personal Mobility. She conducts research on data collection and processing methods and travel behaviour analysis including active and alternative modes of transportation such as carsharing, bikesharing, taxi, ridesharing and multimodal trips.

Nicolas Saunier Nicolas Saunier has an engineering degree and a Doctorate (Ph.D.) in computer science from Telecom ParisTech (2005). He worked for four years at the French institute of science and technology for transport (IFSTTAR) for his doctorate, then for four years as postdoctoral fellow and research associate at the Department of Civil Engineering of the University of British Columbia (UBC) with Prof. Tarek Sayed. He has spent a four-months sabbatical in 2015 in the Road and Transport Department of Lund University. Since 2009, he is working in the Civil, Geological and Mining Engineering Department of Polytechnique Montréal and is currently an associate professor in Transportation Engineering. He is a member of the Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT) and co-heads its ITS lab. His interests include intelligent transportation systems (ITS), road safety, and data analysis for transportation (data collection, storage, processing, machine learning, and visualization).