Operational route choice methodologies for practical applications

Evanthia Kazagli [*]   Michel Bierlaire [*]
Matthieu de Lapparent [*]

May 26, 2017

Report TRANSP-OR 170526
Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne
transp-or.epfl.ch

[*]École Polytechnique Fédérale de Lausanne (EPFL), School of Architecture, Civil and Environmental Engineering (ENAC), Transport and Mobility Laboratory, Switzerland, {evanthia.kazagli,michel.bierlaire,matthieu.delapparent}@epfl.ch

1

Abstract

Understanding the route choice of individuals is a key question in travel behavior research and fundamental to transportation demand management. Modeling route choice involves several challenges, pertaining to both data availability and accuracy, and model structure and computational efficiency. The main challenge related to the route choice problem, namely the generation of the choice set, has been circumvented by Fosgerau et al. (2013) through a Markovian link choice problem that is equivalent to a path-based logit model. It is called the recursive logit (RL) model and it is directly applicable for analysis at a disaggregate level. Recent methodologies are designed to simplify the structure of the model and reduce the data requirements. Kazagli et al. (2016) propose a representation of routes that is based on aggregate elements, inspired by the environmental images of the physical world formed by the individuals (Lynch, 1960), instead of links. It is called the mental representation items (MRI) model and is useful for analysis at an aggregate level. From an application point of view, both disaggregate and aggregate route choice indicators are useful. In this paper, we describe how the RL and the MRI models can be used in practice to derive link and route flows, that are the most relevant indicators for route choice applications. Our analysis identifies the advantages and the limitations of each model and allows to draw insights into the use of a specific model, depending on the needs of the application and the data availability. Finally, we investigate the performance of the two models on real data and discuss specific features of the MRI model. The results demonstrate the validity of the MRI model for the purposes of an aggregate analysis.

1   Introduction

Route choice is one of the key travel demand problems. Given an origin (o), a destination (d) and a transportation mode, the objective is to develop a model that describes how travelers select their itineraries between o and d. Route choice models are used to predict the distribution of travelers on the network, allowing for the identification of congestion and supporting a variety of transportation analyses for planning and management, as well as real-time operations, such as route guidance. In this paper we focus on car trips, although the methodology can be extended to other models of transportation.

Classical route choice models assume a mathematical representation of the network $G = (A, \mathcal{V})$, defined in terms of nodes (vertices) $\mathcal{V}$ and links (oriented arcs) A, and a number of link additive attributes associated with each link. Following this network representation, a route is defined as a sequence of links connecting the o and the d of the trip. It is commonly referred to as a *path*.

The need to go beyond the initial shortest and fastest path models (Dijkstra, 1959) has triggered a great deal of research. Route choice analysis is commonly performed within the framework of discrete choice modeling. The primary element for the development of an operational random utility model consists in the definition of the choice set $\mathcal{C}_n$ of each individual n in the data sample. Identifying the path alternatives that form the $\mathcal{C}_n$ of a route choice model in a real network is a difficult task. Different heuristics have been proposed to generate subsets of paths, commonly referred to as consideration sets (e.g. Ben-Akiva et al., 1984; Azevedo et al., 1993; de la Barra et al., 1993; van der Zijpp and Fiorenzo Catalano, 2005; Friedrich et al., 2001; Hoogendoorn-Lanser, 2005; Prato and Bekhor, 2006), to which the observed path is subsequently added (Bekhor et al., 2006). The values of the parameter estimates are dependent on the definition of the choice sets (Frejinger et al., 2009) and may vary significantly for different instances of sampled sets. Assuming all feasible paths in the choice set is hence important to avoid biases in the model (Frejinger et al., 2009). Sampling approaches considering all feasible paths between each *od* pair and correction of paths utilities for sampling are proposed for this reason (Frejinger et al., 2009; Flötteröd and Bierlaire, 2013; Lai and Bierlaire, 2015).

Triggered by previous works developed in the context of traffic assignment (e.g. Akamatsu, 1996; Baillon and Cominetti, 2008), Fosgerau et al. (2013) present the recursive logit (RL) model, where the path choice problem is formulated as a sequential link choice problem in a dynamic framework. The proposed technique avoids the full enumeration of paths and does not require sampling of alternatives. It is further extended by Mai et al. (2015), Mai (2016) and Mai et al. (2016) to account for correlation of alternatives by means of a nested, cross-nested and mixed logit formulation, respectively.

Investigating in a different direction, Kazagli et al. (2016) propose an approach

that simplifies the structure of the model. The authors develop an aggregate route choice model that describes the strategic decisions of people, represented by the *mental representations* of their itineraries, instead of the physical ones, represented by *links* or *paths*. The concept of the *Mental Representation Item* (MRI) is introduced as a modeling element and the methodology for the definition and operationalization of route choice models based on MRIs is presented. The advantage of this approach is that it provides a behaviorally realistic representation of the routes with a lower structural model complexity, as the use of the MRI as a modeling element allows to obviate the need for sampling of alternatives for the generation of the choice set.

The second challenge pertaining to route choice analysis concerns the structural correlation among the path alternatives that overlap. Tackling the correlation of alternatives in the stochastic part of the utility function entails higher model complexity and estimation times. For route choice analysis in particular, the estimation of the nested or mixed logit models is cumbersome. Vovsha and Bekhor (1998) are the first to use a cross-nested logit (CNL) model, also known as the *link-nested model*, in the context of route choice analysis. According to their formulation, each link of the network corresponds to a nest ℓ and each alternative p, i.e. each path, is allocated to nests according to $\alpha_{p\ell}$. The $\alpha_{p\ell}$ parameters are fixed and assumed to be equal to $\frac{\text{length}_{\text{link}}}{\text{length}_{\text{path}}}$. This nesting structure entails a very big number of nests and renders the estimation of the model a challenging task. Nest specific coefficients cannot be estimated.

Lai and Bierlaire (2015) tackle sampling of alternatives within a Multivariate Extreme Value (MEV) model. They present a cross-nested logit model and adopt the Metropolis-Hastings algorithm proposed by Flötteröd and Bierlaire (2013) with an expansion factor inspired by Guevara and Ben-Akiva (2013) allowing to avoid the enumeration of paths. The estimation of this model in large networks remains a challenging task. Recently, Mai et al. (2017) have presented a dynamic programming approach allowing to estimate large network-based MEV models in reasonable computational times.

Finally, a major challenge related to route choice modeling with revealed preference concerns the high requirements in data. The primary data elements for the specification and estimation of a route choice model are (i) a network model, (ii) information about its attributes at the link level, and (iii) route choice observations. Network models are often incomplete though, and detailed information about their attributes, e.g. link travel times, is missing. Obtaining the resolution of information required for disaggregate route choice analyses, that treat the route alternatives as link-by-link sequences, is a challenging task.

Disaggregate route choice analysis[1] entailing higher data availability and resolution –and more generally greater model complexity– may not always be necessary.

---

[1] We refer the reader to Section 2 for a detailed definition of disaggregate and aggregate route choice.

From an application point of view in particular, both disaggregate and aggregate indicators are of interest. Macroscopic traffic assignment and route guidance are typical examples of route choice applications that do not require detailed data. Aggregate approaches, which are not hinged upon detailed data and allow for a lower model complexity, are useful in such cases.

The route choice behavior of individuals manifests itself as flow in the transportation network. Therefore, the link and route flows are the most important *route choice indicators*. In this paper, we focus on these two indicators and consider their representation both at the disaggregate and the aggregate level. For a given demand and network configuration, the only ingredient for the computation of the link and route flows are the *choice probabilities*. These are proportionate to the flows and can be obtained from the route choice model on hand. In what follows, we identify the most appropriate disaggregate and aggregate models to derive these indicators.

Based on the literature review, tractability is achieved at the disaggregate level by the RL model (Fosgerau et al., 2013), and at the aggregate level by the MRI model (Kazagli et al., 2016). They are selected in this paper as the representative ones of each level of aggregation. The recursive logit (RL) model is used as a benchmark representing the disaggregate models, given its advantage over the classical route choice models, i.e. (i) it can be consistently and efficiently estimated on the universal set of paths without sampling of alternatives and (ii) it is equivalent to a logit model. The objective of the MRI approach is to give to modelers the flexibility to decide between modeling on the disaggregate or the aggregate level of analysis, depending on the needs of the application and the data availability.

This paper investigates these tractable models and evaluates how they perform in deriving disaggregate and aggregate route choice indicators. The objectives of the paper are: (i) the identification of the advantages and disadvantages of the two available methods in the literature, namely the *link/ path* and the MRI approach, and (ii) the elaboration on specific features of the MRI model. Both models can be extended to account for correlation among the alternatives. This has been done for the RL model (Mai et al., 2015; Mai, 2016 and Mai et al., 2016). We describe how this generalization applies to the MRI model.

Section 2 recapitulates the RL model and outlines the MRI modeling approach and its extension, that is the definition of a graph of MRI elements. Section 3 delineates the derivation of the indicators using the *link-/ path-* and the MRI-based choice models. Section 4 presents (i) an evaluation process for the performance of the models at the aggregate level of analysis and (ii) the incorporation of the correlation of alternatives in their specification. Section 5 applies the models and methods presented in the previous Sections and presents estimation and prediction results using real data. The last section summarizes the findings of the present study and identifies directions for further research.

2 Disaggregate and aggregate route choice models

We start by presenting the RL and the MRI models. We proceed with the definition MRI graph that serves as the basis for the specification and application of the MRI-based models in the remainder of the paper.

2.1 The recursive logit model

The recursive logit model decomposes the route choice problem into a sequential link choice problem that is equivalent to a logit model. The model has the Markov property and is solved in a dynamic discrete choice framework, given a directed connected graph $G = (A, \mathcal{V})$, where A denotes the set of links and $\mathcal{V}$ the set of nodes. At each state (link) k the individual n chooses the next state (next link) a that maximizes the sum of the instantaneous utility $u_n(a \mid k)$ and the expected downstream utility $V^{d'}(a)$ to the destination[2] d', where

$$u_n(a \mid k) = v_n(a \mid k) + \mu \varepsilon_n(a), \tag{1}$$

and ε_n is independently and identically distributed (i.i.d.) extreme value type I with zero mean. μ is the scale parameter of the model. $v_n(a \mid k) = v(x_{n,a \mid k}; \beta)$, where $x_{n,a \mid k}$ is the vector of attributes associated with the link pair (k, a) and β the vector of parameters to be estimated. $v(d' \mid k) = 0$, $\forall k$ that are connected to d'. The next link choice probability is given by the logit model

$$P_n^{d'}(a \mid k) = \frac{e^{\frac{1}{\mu}(v_n(a \mid k) + V^{d'}(a))}}{\sum_{a' \in A(k)} e^{\frac{1}{\mu}(v_n(a' \mid k) + V^{d'}(a'))}}, \tag{2}$$

where $A(k)$ is the set of outgoing links from link k. $V^{d'}(a)$ are value functions computed using the Bellman equation (Bellman, 1957)

$$V^{d'}(k) = \mu \ln \sum_{a \in A} \delta(a \mid k) e^{\frac{1}{\mu}(v_n(a \mid k) + V^{d'}(a))} \forall k \in A. \tag{3}$$

As $A(d') = \varnothing$, the value function at the destination is 0. The link choice probabilities are organized in $P^{d'}$ destination specific matrices. The model is estimated under the assumption of the universal set of paths $\mathcal{U}$. A path p is defined as a sequence of links and its probability is given by the product of the link probabilities in the sequence:

$$P_n(p \mid \mathcal{U}) = \prod_{k=o}^{d} P_n^{d'}(a \mid k). \tag{4}$$

---

[2] d' refers to the destination dummy link, that is an absorbing state with no successors.

2.2 The mental representation items model

The definition of a MRI model is based on the identification and definition of the MRI elements for the case study of interest. They are the building blocks of the alternatives, in the same way that the links are for the paths. A route is defined as a sequence of MRIs that connects the origin to the destination of the trip. The choice set C_n of an individual n consists of MRI sequences.

A MRI is associated with the strategic decisions of travelers. It captures perceived items associated with a route. It may correspond to an area (e.g. the center of a city), a road segment (e.g. a major arterial) or any other prominent element in the transportation network that can be traversed, such as a bridge or a cordon. The definition of the items is context and application dependent (see Section 5 for specific examples). Each MRI is characterized by (Fig. 1):

1. a name,

2. a description,

3. a geographical span, and

4. a list of representative points.

The first two components capture the conceptual aspects of the items, while the last two are necessary for the model to be operational, by enabling the association of the MRIs with objective representations, such as a map or a network model (Kazagli et al., 2016).
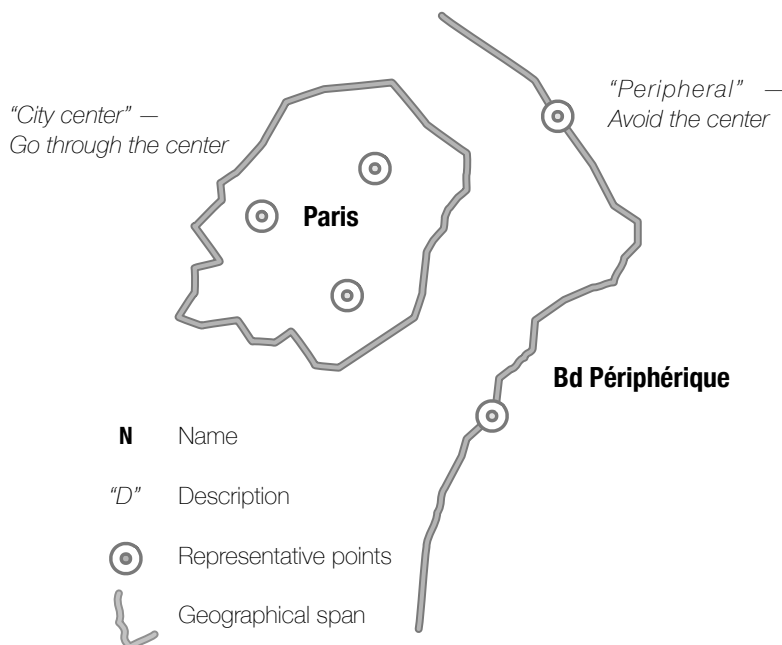

Figure 1: Example of two MRIs and their components.

The role of the geographical span is to relate the observations to MRI elements. Each observation is associated with a MRI if it traverses its geographical span. The geographical span can be defined as (i) a list of links −if a detailed network model is available, or (ii) a geomarked element on a map, otherwise. The representative points signify characteristic locations, such as important junctions or landmarks for each MRI. As described in Kazagli et al. (2016), they are useful for the specification of the utility functions of the MRI model, by associating each MRI alternative with a representative path, e.g. the shortest or fastest path, connecting, the o, the MRIs in the sequence and the d of the trip. This specification assumes that a detailed network model is available. A discussion about the operationalization of the model in lack of a detailed network can be found in Kazagli et al. (2016).

As for any random utility model, the utility of a MRI alternative r is specified on the basis of its attributes x_{rn}. If available, the characteristics z_n of the individual can be considered in the specification. Subsequently, the probability $P(r|\mathcal{C}_n, x_{rn}, z_n)$ of MRI sequence r can be estimated from the data −considering any choice model− using maximum likelihood estimation. The development of a measurement equation is necessary if the available data is not reported in terms of MRI sequences. Let y be an observation, e.g. a trajectory reconstructed by GPS traces. Then

$$\sum_{r \in \mathcal{C}_n} P(y \mid r) \cdot P(r \mid \mathcal{C}_n, x_{rn}, z_n), \tag{5}$$

where $P(y|r)$ is the measurement equation that specifies the probability to observe y given that the MRI sequence r is actually chosen by n.

2.2.1 The MRI graph

In Kazagli et al. (2016), the alternatives of the MRI model involve one MRI in the sequence. The choice set consists of four labeled alternatives, common for all travelers. This paper extends the framework by defining a network of MRI elements that, same as for any route choice model specified on a detailed graph $G = (A, \mathcal{V})$, serves as the basis for (i) the generation of the MRI sequences comprising the choice set of the aggregate model and (ii) the application of the model.

The MRI network is a directed connected graph. It is defined in terms of aggregate nodes m, that represent the MRIs and the origin and destination zones of the case study, and virtual links ℓ, that represent the possible transfers among the nodes m. A link in the MRI graph does not correspond to a single piece of infrastructure, as in the path-based approach defined on G. This is why we refer to it as a virtual link. Its main function in the model concerns the generation of the possible sequences of MRIs that define the choice set, by determining the connectivity of the graph. For the specification of the model though, e.g for the identification of the attributes and for the structure of the correlation among the alternatives, and unlike the path-based approach, the most important elements are

the nodes of the graph that represent MRIs. We further discuss the implications of these features of the MRI model in Sections 4 and 5.

We denote the MRI graph as $G^{\mathcal{M}} = (\mathcal{L}, \mathcal{M})$, where $\mathcal{L}$ is the set of links and $\mathcal{M}$ is the set of nodes. Its definition observes the following steps:

1. Identify and define the MRIs and the origin O and destination D zones for the case study of interest.

2. For each MRI element add a node in the MRI graph.

3. For each O and D zone add a node in the MRI graph.

   The number of nodes of the MRI graph is the sum of the number of MRIs and zone centroids.

4. For each pair of nodes in the MRI graph create a link if the transfer from one node to another is possible.

We use capital letters to differentiate between the aggregate OD pair on the MRI graph (zone level) and the od pair on the detailed network model (node level).

3 Deriving indicators

The focus of this Section is on the application of route choice models to derive useful indicators. Both disaggregate and aggregate indicators are considered. We assume that a RL and/ or a MRI model are available, and present a set of tools aiming at bringing together and analyzing the results from the various levels of aggregation.

The derivation of disaggregate indicators from a disaggregate model, and of aggregate indicators from an aggregate model, is straightforward. On the other hand, the derivation of indicators from a model that does not correspond to the intended level of analysis is not. Applications of the aggregate MRI model to calculate disaggregate indicators (top-down approach) are described in Kazagli et al. (2016). We summarize them here and, in order to complete the toolbox, we present methods to assimilate the output of the disaggregate models to calculate aggregate indicators (bottom-up approach). Although the RL model is link-based, we also present the derivation for a path-based formulation.

We begin by delineating the disaggregate and aggregate route choice indicators of interest. We then describe their derivation. Table 1 summarizes some important notations for the remainder of this Section.

3.1 Disaggregate and aggregate route choice indicators

The flow is the most important *route choice indicator*, as its distribution reflects the route choice behavior of travelers. At the disaggregate level, the *links* and the

Table 1: Notations

notation	element | quantity
p	disaggregate route representation, i.e. path in G
r	aggregate route representation, i.e. MRI sequence in $G^{\mathcal{M}}$
k, a	disaggregate element, i.e. link of G
m	aggregate element, i.e. MRI node of $G^{\mathcal{M}}$

paths, defined on a detailed network model represented by G, are the reference elements for which the flows shall computed. At the aggregate level, it may not always be relevant to talk about link flows. It rather makes sense to refer to element flows, where elements fall under the MRI definition. The scale of the application determines the relevant reference elements for which flows shall be computed. For instance, route flows may be distributed among highways between two cities or, among combinations of major arterials between two areas in one city. For high level of aggregation in the analysis, the route flows may actually degenerate to element flows, i.e. flows across single elements. As an example, an aggregate analysis of the route flows between the cities of Geneva and Lausanne, in Switzerland, may only consider the split of flow between (i) the highway and (ii) the lake road. In this case, route flows are self-same to single element flows.

Table 2 presents the relevant indicators at the disaggregate and aggregate level. For a given origin-destination demand G and network configuration, the only ingredient for the computation of the element and route flows are the *choice probabilities*. These are proportionate to the flows and the outcome of the route choice model on hand.

3.2 Deriving disaggregate indicators

3.2.1 Application of the RL model

The application of the RL model allows to derive disaggregate link flows in a straightforward way. The output of the model consists of destination specific link transition probability matrices: $P^{d'}$ (see Section 2.1). The expected link flow $F(a)$ on link a is the summation of the flow originating at link a and the expected incoming flow. It is given by

$$F(a) = G(a) + \sum_{k \in A} P^{d'}(a \mid k) \cdot F^{d'}(k). \tag{6}$$

where $G(a)$ is the demand originating at link a and ending at destination d'.

Link flows can be computed *for a specific destination* and multiple origins using the link transition probabilities by solving the following system of linear equations (Fosgerau et al., 2013)

$$(I - P^T)F = G, \tag{7}$$

Table 2: Route choice indicators

	route flows		link flows	
disaggregate level	$F(p) = G^{od} \cdot P(p)$		$F(a) = \sum_p \delta_{ap} \cdot P(a \mid p) \cdot F(p)$	
aggregate level	$F(r) = G^{OD} \cdot P(r)$		$F(m) = \sum_r \delta_{mr} \cdot P(m \mid r) \cdot F(r)$	

$\delta = \{0, 1\}$ element-route incidence

where I is the $A \times A$ identity matrix, P is the d' specific $A \times A$ link transition probability matrix, F is a $A \times 1$ vector representing the expected link flow given by (6) and G is a $A \times 1$ vector representing the demand $G(a)$. Link flows for multiple destinations can be computed by summation over the elemental link flows for specific destinations. The route flows $F(p)$ for a specific *od* pair are the product of the demand G^{od} and the route probabilities

$$F(p) = G^{od} \cdot P(p \mid \mathcal{U}), \tag{8}$$

where $P(p \mid \mathcal{U})$ is given by (4).

3.2.2 Application of the MRI model

Given the choice probabilities $P(r \mid \mathcal{C}_n)$ of the MRI model, the probability that traveler n traverses link a can be computed using the following equation (Kazagli et al., 2016)

$$\mathrm{Prob}(a) = \sum_{r \in \mathcal{C}_n} P(a \mid r) \cdot P(r \mid \mathcal{C}_n), \tag{9}$$

where $P(a \mid r)$ is the probability of using link a given MRI sequence r. An example of model specification for $P(a \mid r)$ is proposed in Kazagli et al. (2016). It relies on a path-based representation, given an *od* pair, and is specified as

$$P(a \mid r) = \sum_p \delta_{ap} \cdot P(p \mid r), \tag{10}$$

where δ_{ap} is the zero/ one link-path incidence and $P(p \mid r)$ is the probability of traveling along path p given MRI sequence r. The specification of $P(p \mid r)$ incorporates two factors: (i) a consistency score of each path p for every MRI sequence r and (ii) a cost-dependency term favoring faster paths. As the number of paths with nonzero probability given a MRI r is too high to be enumerated, this specification requires sampling. The Metropolis-Hastings algorithm proposed by Flötteröd and Bierlaire (2013) is used for this purpose. We refer the reader to Kazagli et al. (2016) for the technical details regarding the specification of the assignment model.

The same specification of $P(p \mid r)$ applies to the computation of the disaggregate route choice probabilities:

$$\mathrm{Prob}(p) = \sum_{r \in \mathcal{C}_n} P(p \mid r) \cdot P(r \mid \mathcal{C}_n). \tag{11}$$

3.3 Deriving aggregate indicators

3.3.1 Application of the MRI model

The application of the MRI model allows to directly compute aggregate flows. The aggregate route flow is given by

$$F(r) = G^{OD} \cdot P(r \,|\, \mathcal{C}_n), \tag{12}$$

and the aggregate element flow by:

$$F(m) = \sum_r P(m \,|\, r) \cdot F(r), \tag{13}$$

where $P(m \,|\, r)$ is a zero/ one MRI element-sequence incidence.

3.3.2 Application of the RL model

Consider the probability $P(r|\mathcal{C}_n)$ of the MRI sequence r as obtained from the MRI model. We are interested in reconstructing this probability using the output of a disaggregate model and use it to compute aggregate flows. This can be formulated as

$$\mathrm{Prob}(r) = \sum_p P(r \,|\, p) \cdot P(p \,|\, \mathcal{U}), \tag{14}$$

where $P(r \,|\, p)$ is the probability of going through r given a path p, and $P(p \,|\, \mathcal{U})$ is the probability of p given by (4). The sum span over all paths for a given *od* pair. In the case when the MRIs are associated with a well identified geographical span, it is convenient to specify $P(r \,|\, p)$ as a deterministic zero/ one mapping of p to r: $P(r|p) = 1$, if p traverses the geographical span of the MRIs in the sequence r, and zero otherwise.

We propose two ways to derive $\mathrm{Prob}(r)$, given (i) the link- (2) and (ii) the path-based (4) output of the RL model.

From link to MRI *sequence probabilities* The link transition probability matrices obtained from the RL model can be exploited to directly compute MRI sequence probabilities. The method to derive the MRI sequence probabilities from the destination specific link transition probability matrices builds upon the routine for the computation of link flows given by (7) and does not require sampling of paths.

The toy network presented in Figure 2 is used to illustrate the derivation of the MRI sequence probabilities. The derivation is done under the assumption of a demand of *one* vehicle at the origin link, $G(o) = 1$, and *zero* at any other link of the network, $G(k) = 0$ for all $k \neq o$. Under this scenario, the flow $F(d)$ at the destination link of this observation is proportionate to a probability. The numbers above the edges denote link transition probabilities given by the RL model. The

numbers below the edges denote link flows. Dashed arrows signify links that are connected to the destination.
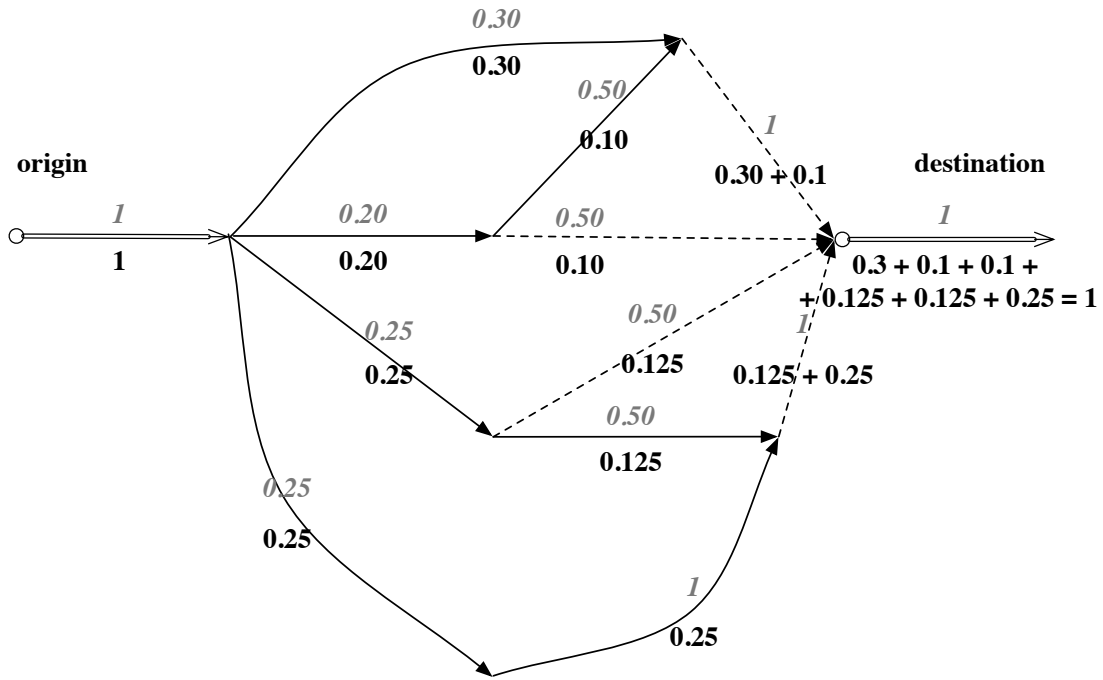
(a) The unit flow reaches d.
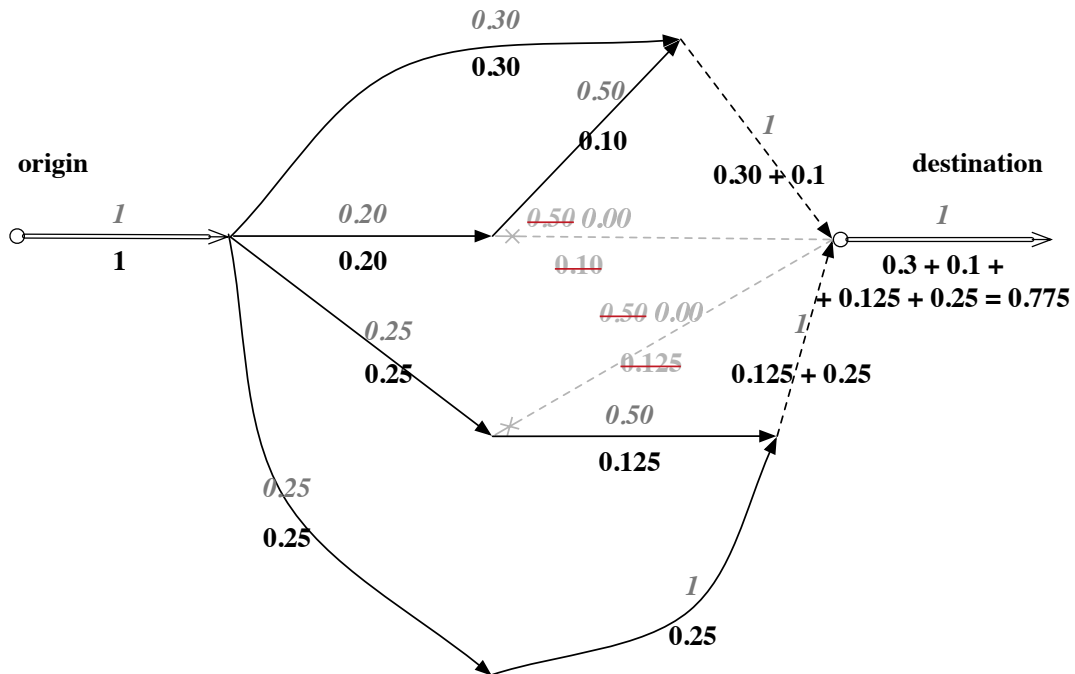
(b) Setting $P(a \mid k) = 0$, $\forall a \in \{\text{city center}\}$ results in $F(d) < 1$.

Figure 2: Illustrative example.

Assume now that the toy network is demarcated in two MRIs, namely the *city center* and the *avoid* the city center. Each link is assigned to the geographical span of one of the two MRIs. We are interested in the probability of each MRI to be traversed. In this example, not passing through the *city center* entails using links belonging to the *avoid* alternative[3].

The probabilities of the two MRIs can be computed as follows:

1. Set the transition probabilities of the links belonging to the geographical span of the *city center* to 0:

   $$P(a \mid k) = 0, \ \forall a \in \{\text{city center}\}.$$

2. Calculate $F(d)$ using (7):

   It corresponds to $F^{\text{avoid}}(d)$.

3. Set $\text{Prob}(\text{avoid}) = F^{\text{avoid}}(d)$.

4. Compute $\text{Prob}(\text{city center})$:

   $$\text{Prob}(\text{city center}) = 1 - \text{Prob}(\text{avoid}).$$

This process does not affect the link transition probabilities that have been previously estimated from the model. It only results in a loss of flow entering the *city center* area (Fig. 2b). This flow does not reach the destination and therefore $F(d)$ is lower than 1. The flow that reaches d, corresponds to the probability of *avoid* the city center. In this example $\text{Prob}(\text{avoid}) = 0.775$. Then, the probability of the *city center* is simply $\text{Prob}(\text{city center}) = 1 - 0.775 = 0.225$. This process is readily extendable to more MRIs and it is context dependent. A simple instance of it is summarized in Algorithm 1. The geographical spans of the MRIs should be disjoint for this method to be operational.

The proposed method gives direct access the $\text{Prob}(r)$. They can be used for the computation of aggregate route and element flows for a given demand between each *OD* pair.

Algorithm 1 Derive the probability of each $r \in \mathcal{C}_n$, for individual n with destination d, given the universal set of paths $\mathcal{U}$ and the RL P^d

---

1: Identify the MRIs to be handled in succession: handle $= \{h_1, h_2, ..., h_H\}$.
2: **for** $\forall h \in \{\text{handle}\}$ **do:**
3:     Set $P(a \mid k) = 0, \ \forall a \in \{h\}^*$
4:     Calculate $F^h(d)$ using (7)
5:     Set $\text{Prob}(h) = 1 - F^h(d)$

---

*h is used to denote the MRI and {h} to denote the geographical span of h.

---

[3]From an operational perspective, in this specific example, the links constituting the span of *avoid* must be used to reach the *city center*, or any additional MRI that may be added to the configuration.

From path to MRI *sequence probabilities*   The method that we propose in this paragraph applies to any path-based model. It relies on the fact that several paths in the road network may be associated with the same MRI sequence[4]. The path choice probabilities are aggregated to MRI sequence probabilities using (14). The $P(r \mid p) = 0$ or 1 mapping of paths to MRI sequences is assumed to hold.

For a given *od*, the number of MRI sequences in the MRI choice set $\mathcal{C}_n$ can be enumerated, while the number of paths in $\mathcal{U}$ cannot. Hence, sampling of paths is required to calculate the sum in (14). The procedure observes the following steps:

1. For each observation, a large number of paths is sampled.

   - Simulation of a random walk on the network is used to draw from the RL link transition probabilities a large number of paths for each *od* pair (Section 5).

2. The probability $P(p \mid \mathcal{U})$ of each sampled path p is computed and corrected for sampling.

   - $P(p \mid \mathcal{U})$ are computed using (4).
   - The expansion factor w_p^{Lai} proposed by Lai and Bierlaire (2015) is used to correct for sampling (see Section 5): $w_p^{Lai} = \frac{k_p}{k_s} \frac{b(s)}{b(p)}$, where $k(p)$ is the number of times a path p is sampled and $b(p)$ is the sampling probability. s denotes the most sampled path, and each p is deterministically assigned to a MRI alternative.

3. The $P(r \mid p) = 0$ or 1 incidences are computed.

4. The $Prob(r)$ are computed using (14).

Remarks   The flow-at-the-destination approach is significantly faster than the sampling approach. The latter requires sampling of a large number of paths and further processing to assign them to MRI sequences. Both tasks are computationally expensive. For the former, matrix multiplications suffice. Both methods are applied to the RL model in Section 5. We show that, for a large number of sampled paths, the two methods provide equivalent results and we establish the former for the remaining steps of the analysis.

4   Additional considerations

This Section identifies issues towards the consolidation of the aggregate route choice approach and proposes ways to address them.

---

[4]An equivalent statement would be that the choice of an MRI sequence may be implemented using various paths in the road network.

4.1 Evaluation at the aggregate level

The aim of the process presented in this paragraph is to evaluate the adequacy of the MRI model that is intended for aggregate analyses. Along these lines, we are interested in whether or not the proposed aggregate model gives a better result than the one of the disaggregate model that is subjected to aggregation. The objective is to justify that, for an aggregate analysis, an aggregate model should be preferred against a disaggregate one.

An important performance indicator for any choice model is the probability of the chosen alternative. It may reveal limitations in the specification of the model and is used to compute the likelihood of the sample in order to evaluate the model fit. Note that both models that we investigate in this paper are estimated using maximum likelihood estimation. In Section 5, we discuss the issue of small probabilities and test the performance of the models at the aggregate level of analysis. We perform the following steps:

1. The output of the RL model is aggregated to compute the probabilities $\mathrm{Prob}(r)$. They are used to calculate an aggregate log likelihood value for the RL model. We denote this likelihood as $\mathcal{LL}^{\mathrm{RL}}$.

   The aggregate route choice probabilities $\mathrm{Prob}(r)$ are computed using the flow-at-the-destination approach (see Section 3.3.2).

2. The log likelihood $\mathcal{LL}^{\mathrm{MRI}}$ of the MRI model is directly calculated from the choice probabilities $P(r \,|\, \mathcal{C}_n)$.

3. $\mathcal{LL}^{\mathrm{RL}}$ and $\mathcal{LL}^{\mathrm{MRI}}$ are compared and the effect of small probabilities on the log likelihood values is identified and analyzed.

4. A cross-validation approach is elaborated on the basis of the $\mathcal{LL}^{\mathrm{RL}}$ and the $\mathcal{LL}^{\mathrm{MRI}}$.

4.2 Addressing the correlation of alternatives

The choice models presented in Fosgerau et al. (2013) and Kazagli et al. (2016) are logit. Both the RL and the MRI models can be extended to account for the correlation among the alternatives. We refer the reader to Mai et al. (2015), Mai (2016) and Mai et al. (2016) for the extensions of the RL model accounting for correlation using nested, cross-nested and mixed logit formulations, respectively. In the following paragraphs, we describe how the MRI model addresses the correlation of alternatives.

4.2.1 The nested and cross-nested logit with MRIs

Consider a MRI graph as defined in Section 2.2.1. The choice set consists in MRI sequences. The correlation can be captured at the MRI level by a nested logit
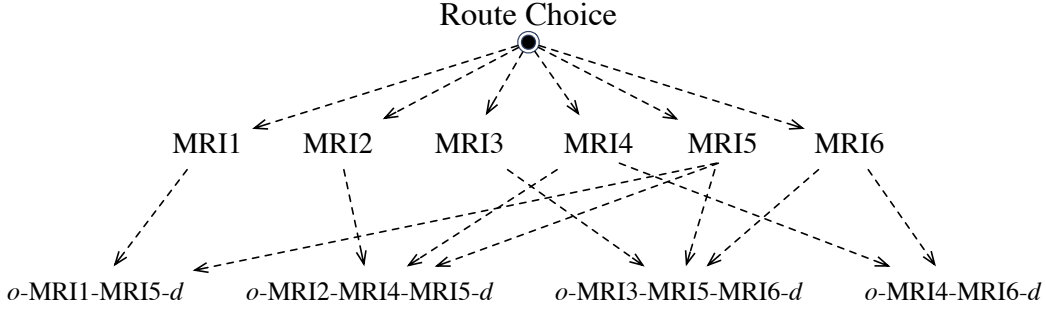
Figure 3: The underlying MRI nesting structure.

(NL) or a cross-nested logit (CNL) after noting that each MRI may correspond to a nest. An alternative r belongs to a nest h if the MRI h appears in the sequence r (Fig. 3). This specification is similar to the link nested model proposed by Vovsha and Bekhor (1998). The difference is that the nests correspond to MRIs instead of links, allowing for a much lower level of model complexity. The estimation of nest specific scales is not precluded by the MRI nesting structure, as the number of nests is low and under the control of the modeler. In addition, it is possible to estimate the inclusion coefficients α_{rh} using, for instance, a parametrization as the one presented below:

$$\alpha_{rh} = \frac{\delta_{rh} \cdot e^{w_h}}{\sum_m \delta_{jh} \cdot e^{w_h}}, \tag{15}$$

where δ_{rh} is 1 if alternative r uses MRI h, and zero otherwise, and w_h is a nest-specific parameter to be estimated. As only $h - 1$ nest-specific parameters can be estimated, one has to be normalized to zero.

4.2.2 Error component model with MRIs

The error component (EC) model is a mixture of logit models incorporating elements that cause correlations among the utilities of the alternatives. Frejinger and Bierlaire (2007) introduced the concept of subnetworks within a factor analytic specification of an error component model. The subnetwork components capture the perceptual correlation of alternatives passing through the same part of the network.

This framework applies particularly well to the MRI model, where the correlation can be captured on the basis of aggregate elements. Each MRI is associated with an error component, and an alternative r is correlated with alternative j if they have at least one MRI in common. This is modeled by introducing MRI specific constants with which the random coefficients ζ_h are associated. The underlying assumption is that alternatives using the same MRI m share the random terms ζ_h. The random terms are assumed to be normally distributed $\sim \mathcal{N}(0, \sigma_h^2)$ and the σ_h of each component is estimated from the data (see Section 5.5).

16

5 Case study

In this case study, we discuss specific features of, and compare, the RL and MRI models on the basis of the indicators presented in Section 3. We evaluate the performance of the models at the aggregate level and test for correlation of the MRI alternatives, following the process described in Section 4. We conclude with some remarks regarding critical aspects of the MRI approach. The MRI-based models presented in this Section are estimated using `Biogeme` (Bierlaire, 2003).

5.1 Dataset and MRI graph

This case study exploits a GPS dataset from the city of Borlänge in Sweden. It is used as a proof of concept in Kazagli et al. (2016)[5]. It is extended by including two additional MRIs in the choice context and defining the MRI graph. The detailed network model consists of 3077 nodes and 7459 unidirectional links (Fig. A1, Appendix A). A sample of 239 observations, with a minimum length of 2 km, is used for the estimation and validation of the model.

The MRI graph depicted in Figure 4 is defined as follows:

1. The city is divided into seven *OD* zones.

   The natural barriers of the city are used to divide the zones. These are (i) the river, which splits the city in two parts and entails a bridge choice, and (ii) the main national roads (trunk highways R70, R50) passing through the city and granting its connection with the rest of the country (see Fig. A2, of Appendix A).

2. Six MRIs are identified. These are:

   1: the *city center* (CC), **2**: the *clockwise movement around the center* (CL), **3**: the *counter-clockwise movement around the center* (CO), **4**: the possibility to *avoid the center* (AV), **5**: the *secondary bridge* (B1), **6**: the *main bridge* (B2). The CC serves as an *OD* zone (zone 7), and a MRI.

3. For each of the six MRIs a node is added in the graph.

4. For each of the five aggregate *O* and *D* zones a node is added in the graph.

5. Links are defined according to the possible transfers between the MRIs themselves and the MRIs and the zones.

   (a) A map of the city associating the *OD* zones and the MRIs is provided in Fig. A3, of Appendix A.

---

[5]The same dataset is used for the application of the RL model and its extension in Fosgerau et al. (2013), Mai et al. (2015) and Mai (2016).
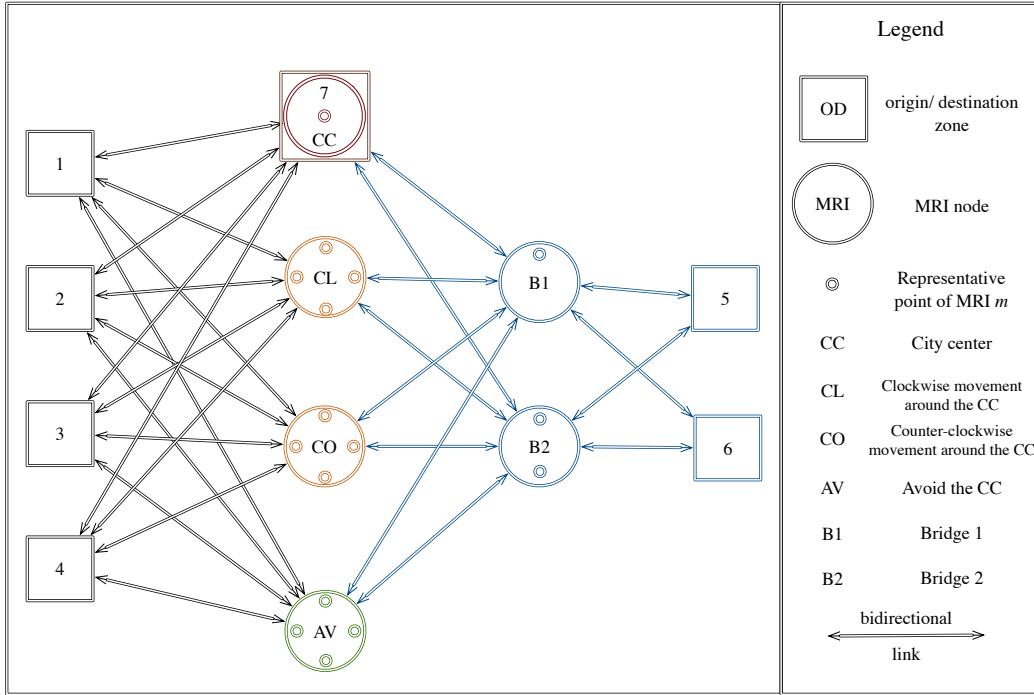
Figure 4: The MRI graph.

(b) MRIs 1 to 4 that are associated with the city center are mutually exclusive, i.e. there are no links among them. The same counts for the two bridges.

(c) A decision among the four MRIs associated with the city center occurs for all trips originating or ending in zones 1 to 4.

(d) A bridge choice occurs for all trips originating or ending in zones 5 and 6.

In summary, the MRI graph consists of 12 nodes and 56 links, allowing for the enumeration of all the routes on the graph for each OD pair. An indicative example of a MRI sequence, describing the route of a trip originating at zone 5 and ending at zone 3, is the *main bridge* (B2) followed by the *counter-clockwise movement around the center* (CO). The detailed list of alternatives is presented in Table B1, of Appendix B.

5.1.1 Operationalization of the MRI graph

The geographical span of the MRIs in Borlänge is depicted in Figure A3, of Appendix A. Given that in this case study a detailed network model is available, the span is defined, for each MRI, as a list of links of the detailed network model. Each observation y is deterministically associated with a MRI alternative r: $P(y|r) = 1$, if y traverses the geographical span of r, and 0 otherwise.

18

The representative points of the MRIs are used to connect each pair of MRI nodes through the shortest path on the detailed graph. This approach adapts the methodology proposed in Kazagli et al. (2016). An example of representative paths associated with MRI sequences is presented in Fig. A4, of Appendix A.

5.2 Model specifications and estimation results

We first present the specifications and the estimation results for the two models. We further elaborate on them in the remainder of the Section.

5.2.1 RL model

The specification of the RL model is similar to the one proposed by Fosgerau et al. (2013), including link travel times $TT(a)$, a constant $LC(a)$ which is equal to 1 for all links in order to penalize paths with many crossings, a left turn dummy $LT(a|k)$ that is equal to one if the turning angle from link k to link a is larger than 40 degrees and less than 177 degrees, and zero otherwise, and a u-turn dummy $UT(a \mid k)$ if the turning angle is larger than 177 degrees. We let the parameter associated with the u-turn to be estimated, instead of fixing it to -20 (assumption in Fosgerau et al. (2013) to strongly penalize u-turns).

The instantaneous utility $v(a|k)$ of a link pair is given by the following expression

$$
\begin{aligned}
v(a \mid k) = {} & \beta_{\text{TravelTime}} TT(a) + \beta_{\text{Penalty}} LC(a) \\
& + \beta_{\text{LeftTurn}} LT(a \mid k) + \beta_{\text{UTurn}} UT(a \mid k).
\end{aligned}
\tag{16}
$$

The estimation results are presented in Table 3. The estimated parameters are significant, the signs are as expected and magnitudes are reasonable.

5.2.2 MRI-based logit model

The utility functions of the MRI model are specified as follows:

$$
V_{rn} = \beta_{\text{TravelTime}} TT_r + \text{dummy}_{\text{AV}} \delta_{rAV} + \text{dummy}_{\text{Around}} \delta_{r\text{Around}},
\tag{17}
$$

$$
\begin{aligned}
V_{rn}^{\text{cross}} = {} & \beta_{\text{TravelTime}} TT_r + \text{dummy}_{\text{AV}} \delta_{iAV} + \text{dummy}_{\text{Around}_{\text{CL}}} \delta_{r\text{Around}_{\text{CL}}} + \\
& + \text{dummy}_{\text{Around}_{\text{CO}}} \delta_{r\text{Around}_{\text{CO}}} + \text{dummy}_{\text{B2}} \delta_{rB2},
\end{aligned}
\tag{18}
$$

where δ_{rn} is 1 if alternative i includes MRI m, and 0 otherwise. V_{rn}^{cross} is specified for alternatives entailing a bridge choice. Apart from the necessary normalizations, dummy variables are associated with each MRI. dummy_{CC} and dummy_{B1} are normalized to zero. Parameters related to the number of left turns and number of intersections were not found to be significant. The estimation results are presented in Table 4 and further discussed in Section 5.5.2.

Table 3: Estimation results for the RL model

Parameter	value (rob. t-test)
$\beta_{\text{TravelTime}}$	-2.91 (-13.15)
β_{Penalty}	-0.32 (-12.60)
β_{LeftTurn}	-1.00 (-16.10)
β_{UTurn}	-7.46 (-7.48)
Number of observations	239
Number of parameters	4
$\mathcal{LL}(\hat{\beta})$	-1255.538

Table 4: Estimation results of the MRI logit model

Parameter	value (rob. t-test)
$\beta_{\text{TravelTime}}$	-0.991 (-7.47)
dummy_{AV}	3.93 (5.81)
dummy_{Around}	1.80 (3.37)
$\text{dummy}_{Around_{CL}}$	2.50 (3.16)
$\text{dummy}_{Around_{CO}}$	3.27 (4.48)
dummy_{B2}	-1.43 (-1.94)
Number of observations	239
Number of parameters	6
$\mathcal{LL}^{\text{MRI}}(\hat{\beta})$	-112.440

5.2.3 Main remarks

Before proceeding with the detailed analysis, we highlight some important features of the two models, regarding how they behave when deriving indicators of their intended level of aggregation. These are related to the nature and the specification of the models. It is useful to keep them in mind for the remainder of the analysis.

The RL model is specified and estimated on the detailed network model. Its specification allows to capture the influence of local features, such as left turns, and minor differences among the alternatives. In general, the RL is characterized by very high probabilities for specific itineraries and links, resulting in low probabilities for the remaining alternatives.

The specification of the MRI model is a product of aggregation. It cannot capture minor differences among alternatives. Therefore, in contrast with the RL, the MRI model tends to distribute the probabilities among the aggregate alternatives.

5.3 Route choice indicators

In the following paragraphs, we elaborate on the results of the models with respect to the derivation of the route choice indicators of interest. We particularly focus on the bottom-up approaches, proposed in Section 3 to derive aggregate indicators. In lack of detailed demand information, we either omit it or make the assumption of a unit demand.

5.3.1 Disaggregate link flows

The RL model is directly applicable to obtain link flows (Section 3.2.1), while the application of the MRI model is subjected to the model presented in Section 3.2.2. Indicative links flows, assuming a demand of one vehicle for a given *od* pair, are presented in Fig. 5 for (a) the RL and (b) the MRI model. The example concerns a trip from the south to the north of the city. Both models indicate that the bulk part of the choice probability (dark shade) is concentrated in the west perimeter of the center. It corresponds to the *clockwise movement around the center* MRI.

Although the two models agree on which part of the network is the most probable to be traversed, the RL model appears to be more conservative on the distribution of the link flows in comparison with the MRI model. The high probabilities of the RL model restrict the flows on specific links, while the aggregate MRI probabilities, that are translated to disaggregate probabilities based on the proposed mapping, allow for the flow to distribute among the possible itineraries.

More extensive analysis, which is out of the scope of this paper, incorporating a detailed demand input –i.e. an origin-destination matrix for the city– is necessary to draw a clearer picture of the disaggregate flows.

(a) Application of the RL model using (6) - (7).   (b) Application of the MRI model using (9).

Figure 5: Link flows of a single traveler.

5.3.2   Aggregate route flows

We now focus on the probabilities of the MRI sequences that are proportionate to aggregate route flows. Their computation is straightforward for the MRI model. For the RL model, the two methods presented in Section 3.3.2 are employed to compute $\text{Prob}(r)$. These are (i) the flow-at-the-destination and (ii) the sampling of paths approaches[6].

Figure 6 illustrates the probabilities $\text{Prob}(r)$ of the *chosen* MRI alternatives for each observation in our data, as obtained from the two methods that handle the RL output. A deviation in the result as derived from the two methods is noted for three observations. Regressing the probabilities of the first approach against the ones of the second, and including an intercept, gives a coefficient estimate of 1.004 and an intercept ~ 0. That is, the two methods give an equivalent result.

The sampling approach provides an approximation of the result. The bigger the sample size is, the better this approximation is; the $\sum_r \text{Prob}(r)$ over all MRI sequences should then be close to 1 and $\text{Prob}(r) \sim P(r \mid \mathcal{C}_n)$. A remark with respect to this, concerns a few observations for which the $\sum_r \text{Prob}(r)$, before correcting for sampling, is in the order of *0.5-0.7*, indicating that more paths must have been sampled. Finally, not all MRI alternatives are represented by sampled paths for every observation. That is, there are paths with very low probability, close to zero, that equivalently correspond to MRI alternatives with very low probability. Even with a huge sample it is not possible to sample these instances[7].
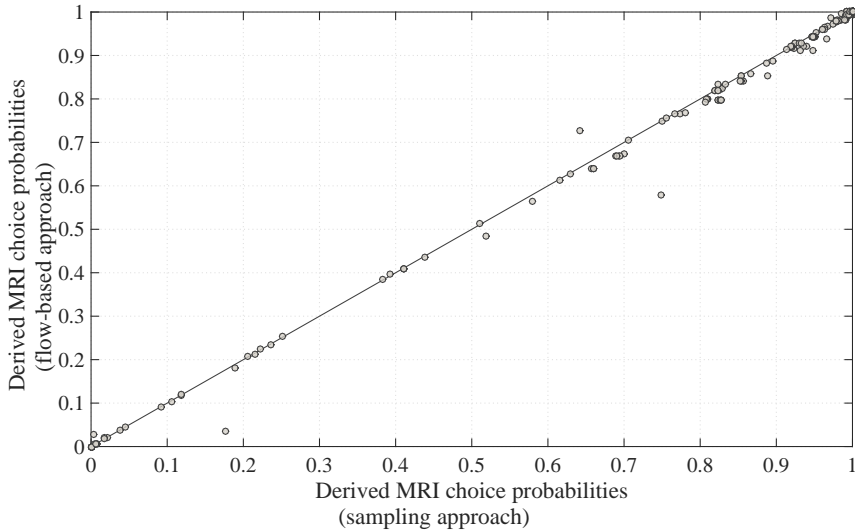

Figure 6: MRI choice probabilities as derived from the two methods.

---

[6]500000 paths are sampled for each *od* pair, by simulating a random walk on the network to draw from the RL link transition probabilities.

[7]This would justify the use of importance sampling. Nevertheless, this problem does not preclude our analysis, hence we only keep this as a remark.
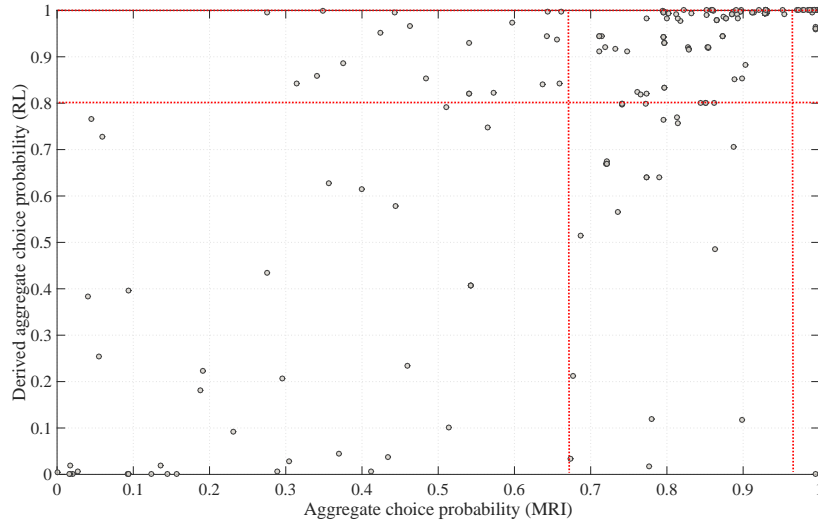
Figure 7: MRI choice probabilities.

To conclude, the flow-at-the-destination approach is not subject to the limitations of the sampling approach. It is used for the remainder of the analysis as, in contrast with the sampling approach, it is very fast to compute and it provides an exact result.

The probability of the chosen alternative is used as a reference to investigate the behavior of the two models in details. Figure 7 depicts the probabilities of the chosen MRI alternatives $P(r|\mathcal{C}_n)$ against $\text{Prob}(r)$, as derived from the RL. The highlighted lines represent the 25% and the 75% percentiles. The RL model forecasts, on average, marginally higher probability for the chosen alternatives. Both models are faced with low probabilities for the chosen alternative (bottom-left of the chart). Examining the instances for which low probabilities occur reveals that, in comparison with the MRI model, the RL model is faced with much lower probabilities, that for some observations are ~ 0. In summary, the RL model is characterized by more extreme instances –concerning either high or low probabilities– while the MRI model appears to be generally moderate. This is in accordance with the expectations based on the features of the models discussed in Section 5.2.3.

For a few instances the RL model is dominated by the MRI model (bottom-right of the chart). Figure 8 shows an indicative example. The dashed line denotes the observed path. The pale-color is an indicator of a low probability. The highlighted, dark line denotes the most probable path. Evidently, the RL model assigns a low probability to the chosen alternative at the disaggregate level.

Let's now examine what happens at the aggregate level. The observed path corresponds to choosing the AV MRI alternative. Its probability according to the RL model is 0.117, against 0.454 for the CC, 0.429 for the CO, and ~ 0 for the CL. This is consistent with the output at the disaggregate level. The MRI model assigns a probability of 0.899 to the chosen alternative. The reason why the RL model is

Figure 8: Indicative example of an observation for which the RL is faced with low probability for the chosen alternative.

faced with low probabilities for the chosen alternative, both at a disaggregate and an aggregate level, is the low transition probability to the link (encircled on Fig. 8) that grants access to (i) the observed path and (ii) to paths realizing the chosen MRI alternative, respectively for the disaggregate and aggregate levels. The low transition probability of specific links explains all of the instances for which the RL model is faced with low choice probabilities, even at an aggregate level when it is expected to perform better. The reason why these low link transition probabilities occur is not clear. One explanation is the small sample of data that is used for the estimation of the models. It is possible that the RL model needs more observations to ensure coverage of the physical network and avoid

Let's now focus on the instances for which the MRI model is dominated by the RL, or faced with low probability for the chosen alternative. The source of the problem in this case is the limited specification of the model. Travel time is the only attribute included in the utility function. Consequently, the dummy variables associated with the MRIs play an important role in the specification (see Table 4). If the travel times of two or more alternatives do not differ considerably, the dummy variables take over their choice probabilities. One such instance is depicted in Figure 9. The chosen MRI alternative is the CC (dashed line). Its travel time does not differ considerably from the ones of the *clockwise* and CO. As $\text{dummy}_{\text{Around}} = 1.80$, CL and CO are favored in expense of the CC alternative, for which the dummy variable is normalized to zero. The MRI model assigns a probability of 0.045 to CC, against 0.421, 0.533 and 0.001, accordingly for the CL, the CO and the AV alternatives. The RL model that is able to capture minor differences dominates the MRI model in this case. It assigns 0.766, 0.133, 0.101, ~ 0, for the CC, the CL, the CO and the AV alternatives.
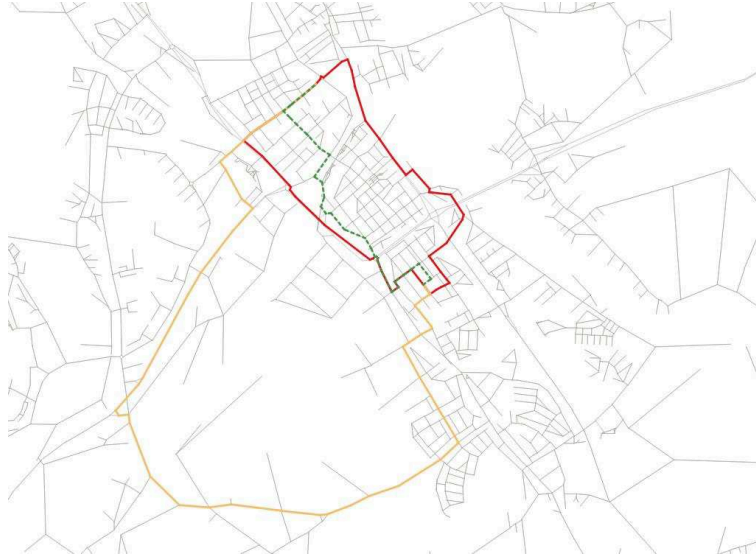
Figure 9: Indicative example of an observation for which the MRI model is faced with low probability for the chosen alternative.

We further elaborate on the effect of the small probabilities on the performance of the models in Section 5.4.

5.3.3 Aggregate element flows

Table 5 reports aggregate probabilities for the MRI elements. They are obtained from the two models by aggregating the probabilities of the alternatives that traverse the corresponding MRI element. These are relevant for the derivation of aggregate element flows. The first column of the table reports the number of times each MRI is actually observed in the data sample.

Obviously, the MRI model reproduces the MRI probabilities as its result is the output of maximum likelihood on the basis of MRI alternatives. The result of the RL model, which is estimated of the universal set of paths, is a product of aggregation and gives an approximation of the shares.

Table 5: MRI element shares (independent of OD).

MRI	Observed	RL model	MRI model
CC	4	22.8	4.0
CL	54	47.2	53.3
CO	60	57.6	60.7
AV	63	53.4	63.0
B1	26	21.7	26.0
B2	74	78.2	74.0

5.4 Performance at the aggregate level

The evaluation process outlined in Section 4.1 is conducted here. We start by commenting on the final aggregate log likelihood values of the two models, and identifying the impact of small probabilities on them. We complete the analysis with a cross-validation. The objective is to show that for analysis at the strategic level, an aggregate model should be preferred.

5.4.1 Final aggregate log likelihood

The final log likelihood of the MRI model is $\mathcal{LL}^{\text{MRI}} = -112.440$ (Table 4). The corresponding final log likelihood given by the RL model, computed based on Prob(r), is $\mathcal{LL}^{\text{RL}} = -271.118$. We elaborate on this result and explain why one cannot claim that the MRI model has a better fit or predictive performance based on it.

Following up on the analysis of Section 5.3.3 and the discussions about Figure 7, the instances for which low probabilities occur in the case of the RL model are, in comparison with the MRI, more extreme, that is ~ 0 for some observations. This, results in a significant decrease of the log likelihood. Several instances of the RL model are characterized by decreases of more than 10, with the worst one being ~ 26 degrees. On the other hand, the worst instance of the MRI model is characterized by a decrease of ~ 7.5. The extreme instances concern 9 observations. Summing the logarithms of the probabilities of these 9 observations gives a downgrade of ~ 165 in the $\mathcal{LL}^{\text{RL}}$.

Excluding these 9 observations from the computation of the log likelihood values results in $\mathcal{LL}^{\text{MRI}} = -91.222$ and $\mathcal{LL}^{\text{RL}} = -94.446$. This is an important result as it demonstrates two things: (i) the two models give compatible results and (ii) the efficacy of the MRI model to produce meaningful results intended for an aggregate analysis.

5.4.2 Forecasted and predicted aggregate likelihood

A cross-validation approach is used to evaluate the performance of the MRI and RL models with respect to the predicted MRI choice probabilities. 80% of the observations are randomly drawn and used to estimate the models, while the remaining 20% is used to apply them. 100 training and test samples are generated for this purpose. The same samples are used across all models. The resulting log likelihood values are conditional on the samples. The average of the log likelihood values over the samples is computed in order to obtain unconditional values as

$$\overline{\mathcal{LL}}_{\text{t}} = \frac{1}{\text{t}} \sum_{s=1}^{\text{t}} \mathcal{LL}_s, \quad \forall \text{t} = 1, \cdots, 100 \tag{19}$$

where s is a sample. The values of $\overline{\mathcal{LL}}_t$ are plotted in Figures 10-13[8]. Figures 10 and 11 depict the final and the predicted log likelihood values of each replication for the training and the test samples, respectively. The MRI LL0 line shows the log likelihood of the null model corresponding to the MRI choice set[9] and serves as a reference for the performance of the models. The box plots in Figures 12 and 13 summarize the output of the 100 samples for each model.
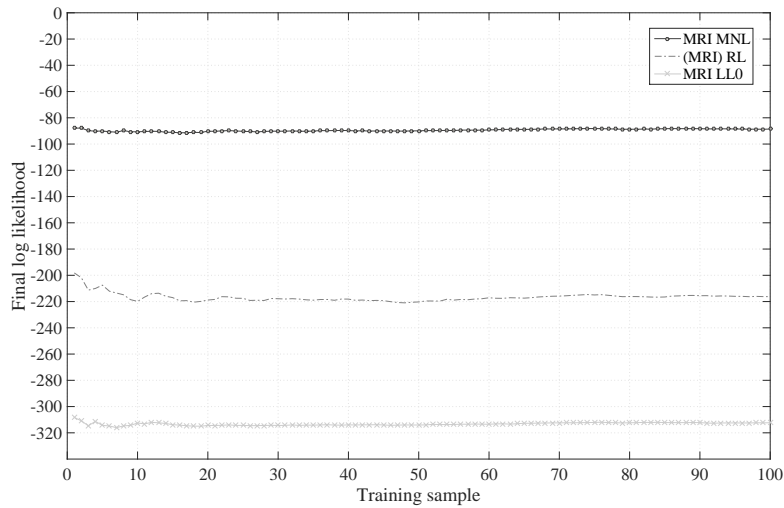

Figure 10: Log likelihood of the training samples.
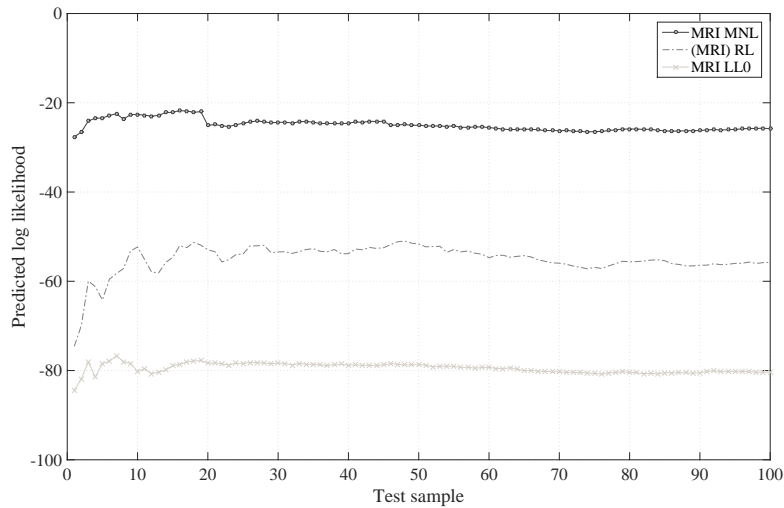
Figure 11: Predicted log likelihood of the test samples.

---

[8]The log likelihood converges to the unconditional value as long as t → ∞. 100 samples are used so that the analysis can be conducted in a reasonable computational time. The corresponding figures with the conditional values are presented in Figures A5-A8 of Appendix A.

[9]The null model predicts equal probabilities for all alternatives, as all its parameters are fixed to zero.
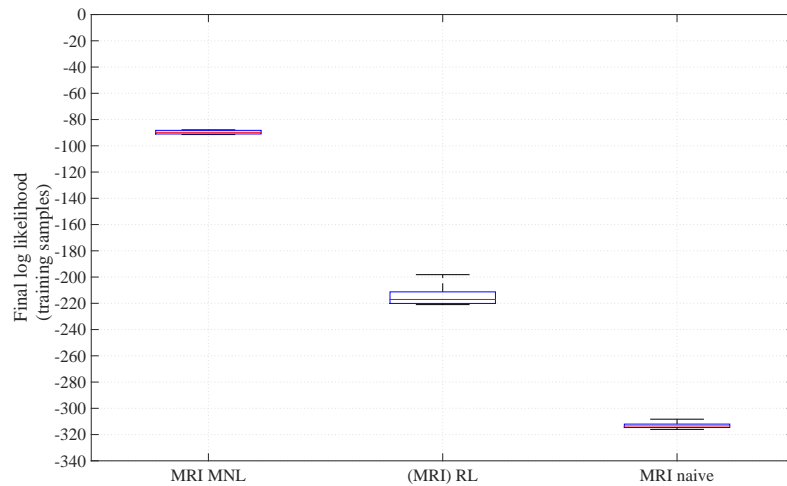
Figure 12: Box plot of the log likelihood for the training samples.
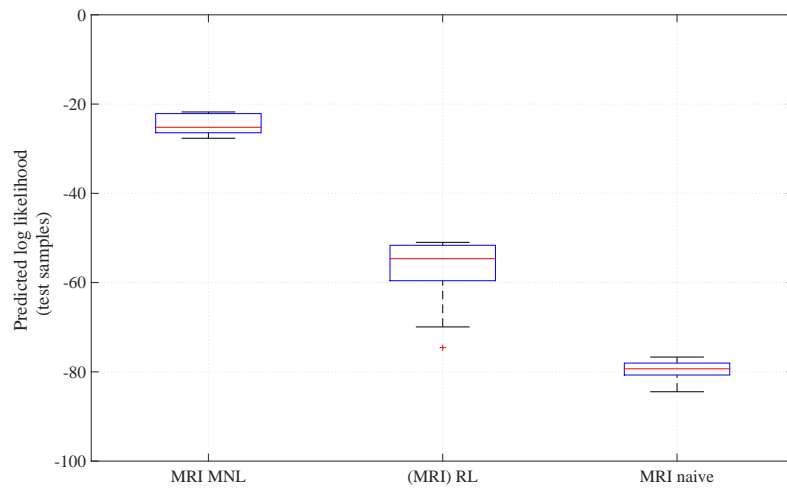

Figure 13: Box plot of the predicted log likelihood for the test samples.

Due to the small sample size, it is not certain that the estimated coefficients have converged to their true values, hence the log likelihood value. The main remark with respect to Figures 10 and 11 is that the log likelihood values of the RL model appear to be worse than the ones of the MRI model. This is attributed to the issues discussed in the previous paragraph –note that inside the validation framework we do not exclude the extreme instances. The output of the RL model is also characterised by more variance. This can be explained by the fact that the RL model is estimated on the detailed network, and due to the small data sample, it is more sensitive to network attributes than the MRI model.

5.5 Capturing correlation with the MRI model

5.5.1 Model specifications

The logit model presented in Section 5.2.2 serves as a benchmark for testing for correlation within the MRI framework. We make the reasonable assumption that MRI sequences using the same MRI are correlated. It is tested using a NL and an EC specification. One nest, or component, is assumed accordingly for the NL and the EC models. It corresponds to the option of *going around the city center*, either using the *clockwise* or the *counter-clockwise* movement. Regarding the EC model, the underlying assumption is that alternatives using the same MRI m share the random terms ζ_h. ζ_h are normally distributed $\sim \mathcal{N}(0, \sigma_h^2)$ and the σ_m of each component is estimated from the data. The utility functions are specified as follows

$$
\begin{aligned}
V_{rn} = \beta_{\text{TravelTime}} TT_r + dummy_{AV} \delta_{rAV} + \\
+ dummy_{Around} \delta_{rAround} + \zeta_{Around} \delta_{rAround},
\end{aligned}
\tag{20}
$$

$$
\begin{aligned}
V_{rn}^{\text{cross}} = \beta_{\text{TravelTime}} TT_r + dummy_{AV} \delta_{rAV} + \\
+ dummy_{Around_{CL}} \delta_{rAround_{CL}} + dummy_{Around_{CO}} \delta_{rAround_{CO}} + \\
+ dummy_{B2} \delta_{rB2} + \zeta_{Around} \delta_{rAround}.
\end{aligned}
\tag{21}
$$

This specification is the same as the one of the logit model (see Section 5.2.2) with the inclusion of the error components. Various CNL specifications were estimated as well. It was not possible to find a specification for which none of the parameters corresponding to the scales μ_h of the nests was not hitting the bounds. This can be related to the small sample that is used for the estimation.

5.5.2 Estimation results

Table 6 presents the parameter estimates for the three models. The parameters for travel time are significant and have the expected signs.

The parameters μ_{Around} and σ_{Around} are significant, indicating that the alternatives using the MRIs of *clockwise movement around the center* and the *counter-clockwise movement around the center* are correlated. Consistent with this result, there is an improvement in the final log likelihood of the NL and EC models, as compared to the logit model.

5.5.3 Forecasting results and validation

The cross-validation approach (see Section 5.4.2) is also performed for the NL model. Figures 14 and 15 demonstrate the improvement in the final and the predicted log likelihoods due to the NL specification, as opposed to the logit model.

Table 6: Estimation results

Parameter	Logit value (rob. t-test)*	NL value (rob. t-test)*	EC value (rob. t-test)*
dummy_{AV}	3.93 (5.81)	3.77 (6.45)	6.70 (3.10)
dummy_{Around}	1.80 (3.37)	2.24 (4.08)	3.99 (2.26)
$\text{dummy}_{Around_{CL}}$	2.50 (3.16)	2.72 (4.26)	4.77 (2.40)
$\text{dummy}_{Around_{CO}}$	3.27 (4.48)	3.09 (4.97)	5.40 (2.77)
dummy_{B2}	-1.43 (-1.94)	-0.824 (-1.33)	-1.16 (-1.21)
$\beta_{TravelTime}$	-0.991 (-7.47)	-0.790 (-7.52)	-1.43 (-4.88)
μ_{Around}	$\times$	2.36 (2.75)	$\times$
σ_{Around}	$\times$	$\times$	3.13 (2.71)
# draws	$\times$	$\times$	1000
# observations	239	239	239
# parameters	6	7	9
$\mathcal{LL}(0)$	-392.742	-392.742	-392.742
$\mathcal{LL}(\hat{\beta})$	-112.440	-105.053	-106.867

*t-test against 0 for all the parameters except for μ_m, for which the reported t-test is against 1.
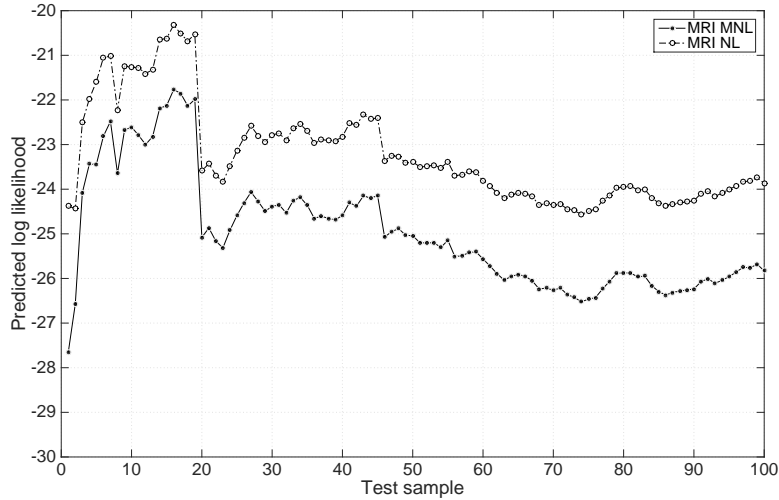
Figure 15: Predicted log likelihood of the test samples.

5.6 Discussion

We conclude this Section with a discussion about specific features of the MRI model. Despite its limited specification, the results indicate that it can adequately capture the high level route choice behavior of individuals, and be used for a meaningful aggregate analysis. In comparison with the RL model, its performance at the aggregate level is found to be at the same level of sufficiency, and slightly better. In that sense, the MRI approach may prove to be a useful tool for the design and assessment of congestion charging schemes. In addition, it is especially relevant
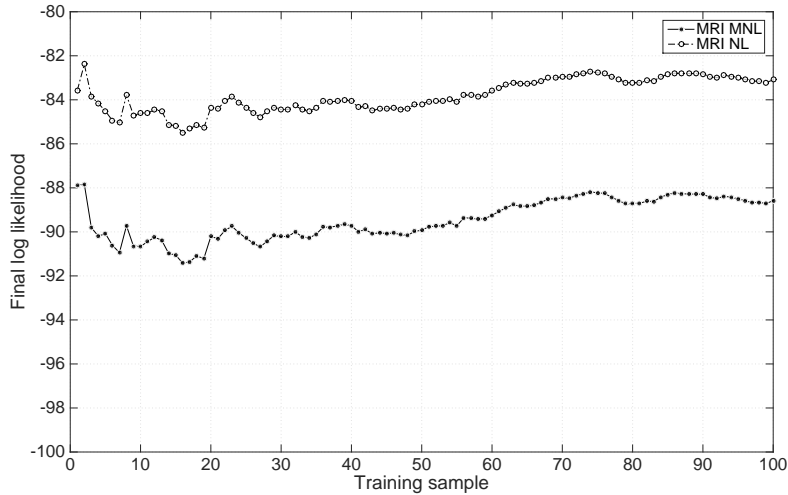
Figure 14: Log likelihood of the training samples.

with the needs of traffic assignment at a macroscopic level and the development of route guidance systems. These are typical applications that do not require a detailed model.

Although we refer to the potential of the MRI model under limited data availability and towards a network-free approach, in this paper we make use of the detailed network model to make the choice model operational. This is the main drawback of the current specification with respect to complexity and sensitivity. That is, it relies on the definition of representative paths on the detailed graph, for the computation of the attributes of the MRI alternatives. The identification of the representative points that are necessary for the generation of these paths is straightforward for the small case study presented in this paper. For a bigger and more complex network though, this feature brings back the complexity of the path-based approach. The sensitivity of the model to the selection of these points needs to be tested in that case, e.g. through Monte Carlo simulations to sample various instances of representative points and re-estimation of the model parameters.

This limitation can be circumvented by eliminating the role of the representative paths in the operationalisation of the model. Along these lines, the representative points can be removed from the definition of the MRIs and the geographical span shall be the only important component for the specification of the model. The definition of the geographical span is not hinged upon a detailed network model and allows to incorporate aggregate attributes that are consistent with the level of aggregation of the model. Aggregate attributes may be (i) congestion indices for specific areas or segments –represented in the model by the MRIs and their span, (ii) the distances of the first and last MRIs in the sequence accordingly from the origin and destination of the trip, or (iii) measures of entropy of a route as an indicator of complexity of the navigation task (see e.g. Gallotti et al., 2016).

31

6 Conclusion

The paper presents a practical scheme for the application and assessment of tractable route choice models. It builds upon the current state-of-the-art and identifies a modeling paradigm, where different levels of aggregation, in the representation of the routes, are targeted for different purposes of analysis. The elaboration on the results elicited from this paradigm is intended to give insights into the use of specific models, depending on the needs of the application and the data availability.

The RL and the MRI models are investigated. We demonstrate the derivation of relevant disaggregate and aggregate indicators for the application of these two models. The derivation of indicators from a model corresponding to a different level of aggregation than the one required by the application is not straightforward. We fill this gap by presenting methods to derive aggregate indicators from a disaggregate model and vice versa. We further investigate and compare the performance of the models on real data. According to the foundings of the case study, an aggregate model should be preferred, versus a disaggregate one, at the strategic level of analysis. The validity of the MRI approach, as demonstrated by the results, is an important step towards the simplification of route choice models intended for an aggregate analysis, and can facilitate various applications in practice.

Various extensions of the RL have been proposed (Mai et al., 2015; Mai, 2016; Mai et al., 2016). Currently, we are working on extending and generalizing the MRI model. The objective is to apply it in big and complex networks, in order to investigate its advantage in dealing with the extra complexity. Furthermore, as network models are often incomplete, we are working on a definition that is less dependent on detailed network data for the operationalization of the model.

Acknowledgements

References

Akamatsu, T. (1996). Cyclic flows, markov process and stochastic traffic assignment, *Transportation Research Part B: Methodological* **30**(5): 369–386.

Azevedo, J., Santos Costa, M., Silvestre Madeira, J. and Vieira Martins, E. (1993). An algorithm for the ranking of shortest paths, *European Journal of Operational Research* **69**(1): 97–106.

Baillon, J.-B. and Cominetti, R. (2008). Markovian traffic equilibrium, *Mathematical Programming* **111**(1–2): 33–56.

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

Bellman, R. (1957). *Dynamic Programming*, 1 edn, Princeton University Press, Princeton, NJ, USA.

Ben-Akiva, M., Bergman, M. J., Daly, A. and Ramaswamy, V. (1984). Modeling interurban route choice behavior, *Proceedings of the 9th International Symposium on Transportation and Traffic Theory*, Utrecht, The Netherlands, pp. 299–330.

Bierlaire, M. (2003). BIOGEME: A free package for the estimation of discrete choice models, *3rd Swiss Transportation Research Conference*, Ascona, Switzerland.

de la Barra, T., Perez, B. and Anez, J. (1993). Multidimensional path search and assignment., *Proceedings of the 21st PTRC Summer Annual Meeting*, Manchester, England.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs, *Numer. Math.* **1**(1): 269–271.
URL: *http://dx.doi.org/10.1007/BF01386390*

Flötteröd, G. and Bierlaire, M. (2013). Metropolis-Hastings sampling of paths, *Transportation Research Part B: Methodological* **48**: 53–66.

Fosgerau, M., Frejinger, E. and Karlstrom, A. (2013). A link based network route choice model with unrestricted choice set, *Transportation Research Part B: Methodological* **56**(0): 70 − 80.

Frejinger, E. and Bierlaire, M. (2007). Capturing correlation with subnetworks in route choice models, *Transportation Research Part B: Methodological* **41**(3): 363–378.

Frejinger, E., Bierlaire, M. and Ben-Akiva, M. (2009). Sampling of alternatives for route choice modeling, *Transportation Research Part B: Methodological* **43**(10): 984–994.

Friedrich, M., Hofsaess, I. and Wekeck, S. (2001). Timetable-based transit assignment using branch and bound techniques, *Transportation Research Record: Journal of the Transportation Research Board* pp. 100–107.

Gallotti, R., Porter, M. A. and Barthelemy, M. (2016). Lost in transportation: Information measures and cognitive limits in multilayer navigation, *Science Advances* **2**(2).

Guevara, C. A. and Ben-Akiva, M. E. (2013). Sampling of alternatives in multivariate extreme value models, *Transportation Research Part B: Methodological* **48**(0): 31 − 52.

Hoogendoorn-Lanser, S. (2005). *Modelling Travel Behavior in Multi-modal Networks*, PhD thesis, Technology University of Delft.

Kazagli, E., Bierlaire, M. and Flötteröd, G. (2016). Revisiting the route choice problem: A modeling framework based on mental representations, *Journal of Choice Modelling* **19**: 1–23.

Lai, X. and Bierlaire, M. (2015). Specification of the cross-nested logit model with sampling of alternatives for route choice models, *Transportation Research Part B: Methodological* **80**: 220 − 234.

Lynch, K. (1960). *The Image of the City*, M.A.: MIT Press, Cambridge.

Mai, T. (2016). A method of integrating correlation structures for a generalized recursive route choice model, *Transportation Research Part B: Methodological* **93, Part A**: 146 − 161.

Mai, T., Bastin, F. and Frejinger, E. (2016). A decomposition method for estimating complex recursive logit based route choice models, *EURO Journal on Transportation and Logistics* pp. 1 − 23.

Mai, T., Fosgerau, M. and Frejinger, E. (2015). A nested recursive logit model for route choice analysis, *Transportation Research Part B: Methodological* **75**: 100 − 112.

Mai, T., Frejinger, E., Fosgerau, M. and Bastin, F. (2017). A dynamic programming approach for quickly estimating large network-based {MEV} models, *Transportation Research Part B: Methodological* **98**: 179 − 197.

Prato, C. G. and Bekhor, S. (2006). Applying branch and bound technique to route choice set generation, *Transportation Research Record* pp. 19–28.

van der Zijpp, N. J. and Fiorenzo Catalano, S. (2005). Path enumeration by finding the constrained K-shortest paths, *Transportation Research Part B: Methodological* **39**(6): 545–563.

Vovsha, P. and Bekhor, S. (1998). Link-Nested Logit Model of Route Choice: Overcoming Route Overlapping Problem, *Transportation Research Record: Journal of the Transportation Research Board* **1645**: 133–142.

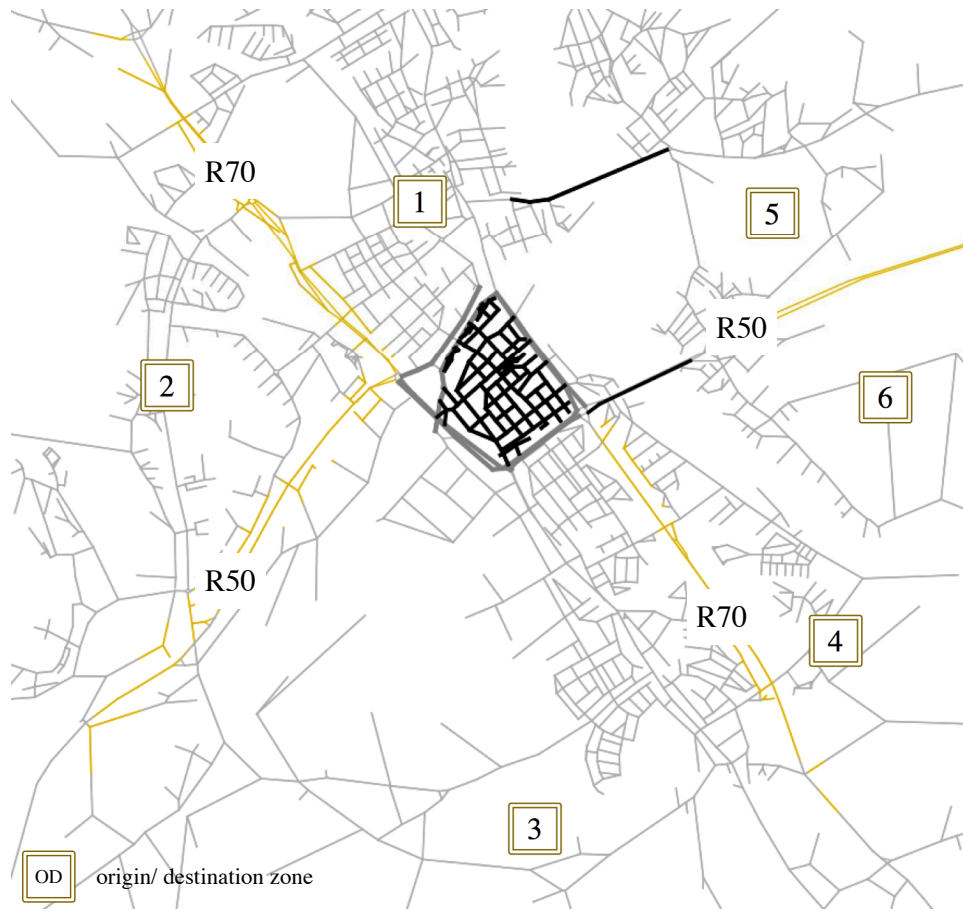# A Additional figures

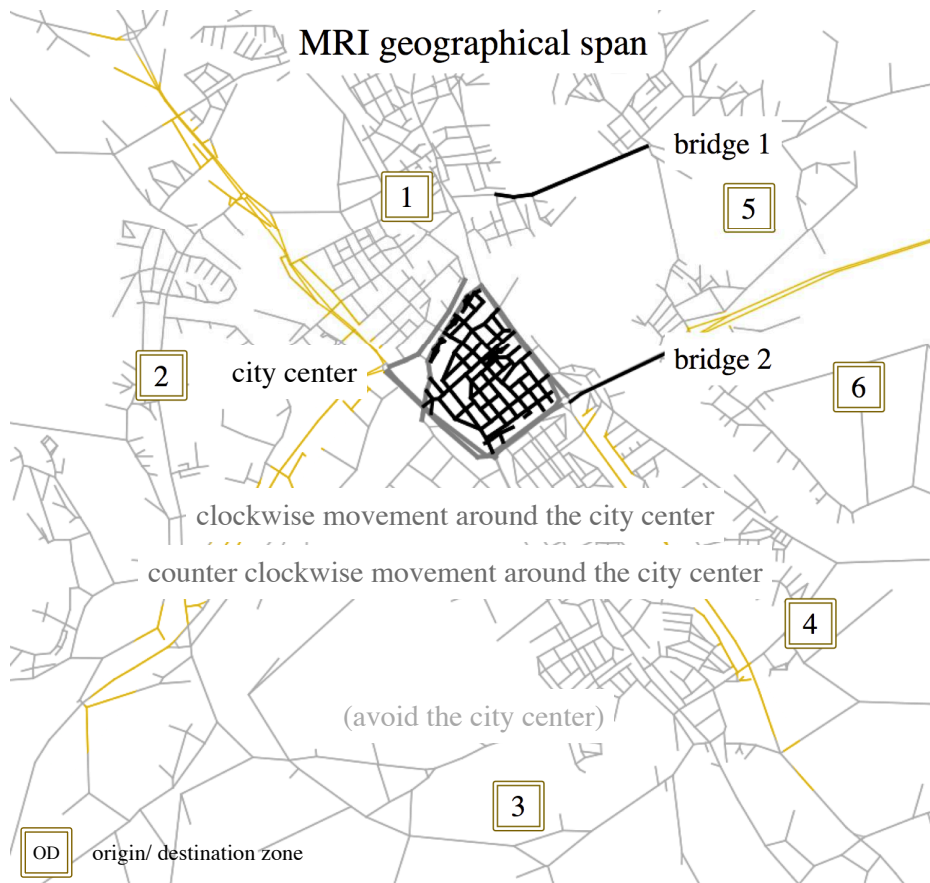Figure A1: The network of Borlänge.

Figure A2: The *OD* zones in Borlänge.

Figure A3: The geographical span of the MRIs in Borlänge.
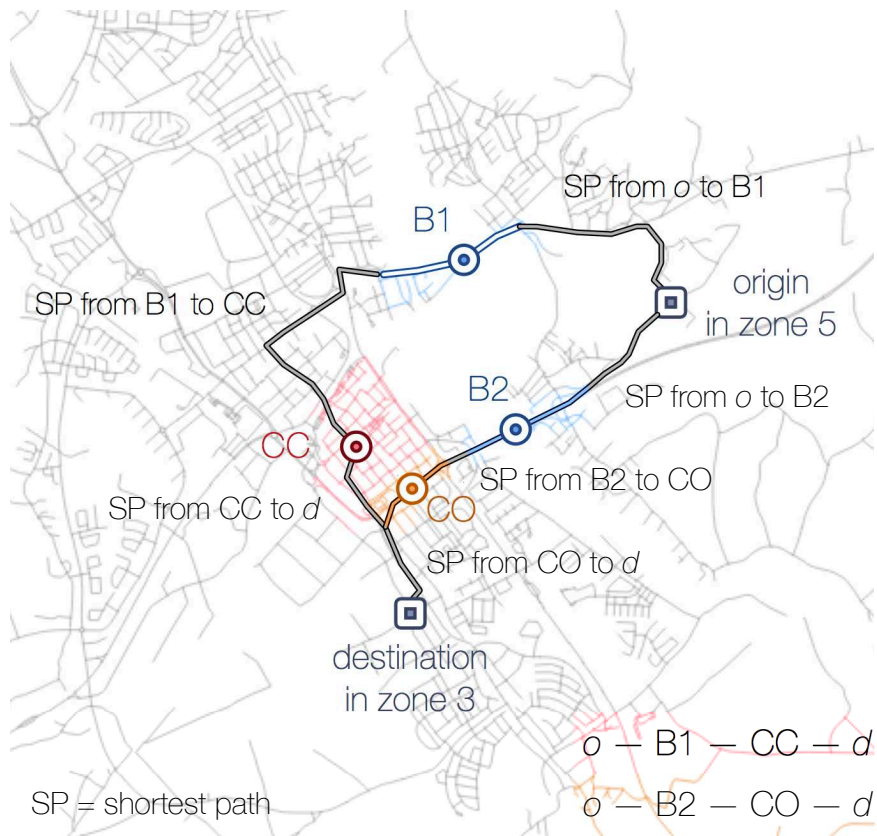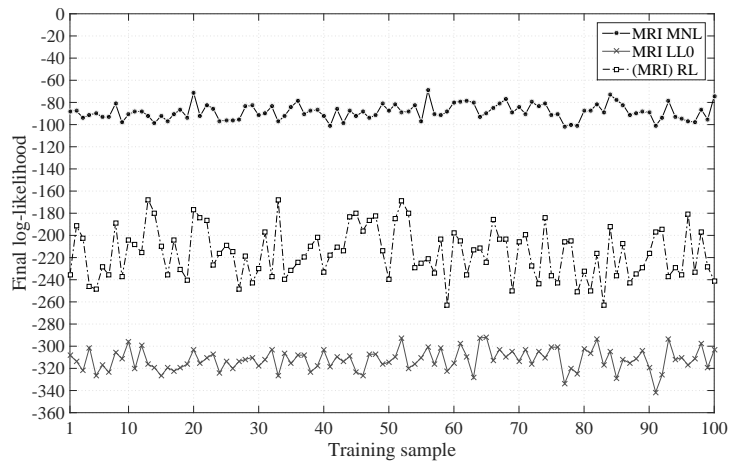
Figure A4: Example of MRI sequences in Borlänge.

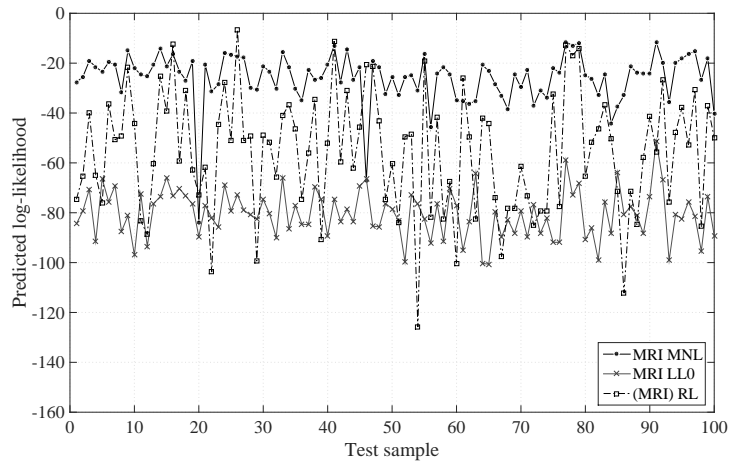Figure A5: Log likelihood of the training samples.

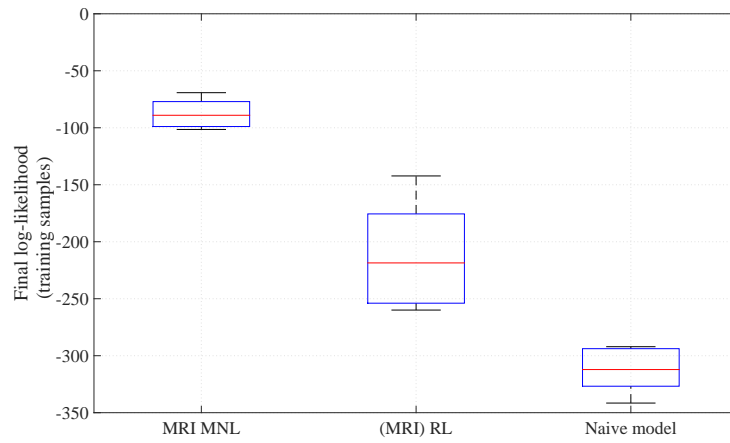Figure A6: Predicted log likelihood of the test samples.

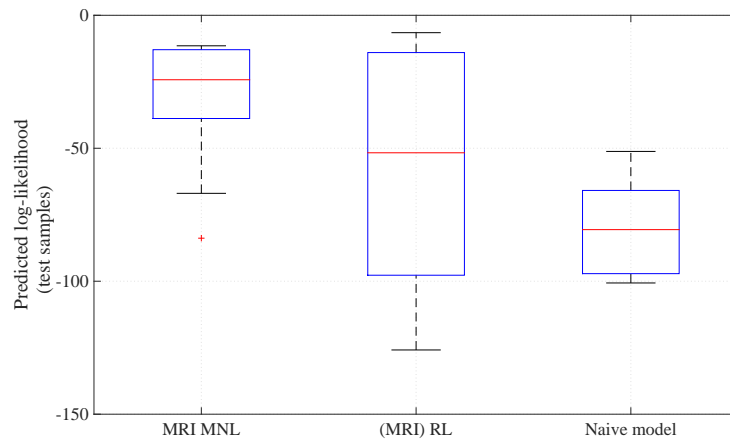Figure A7: Boxplot of the log likelihood for the training samples.

Figure A8: Boxplot of the predicted log likelihood for the test samples.

B   Additional tables

Table B1: List of alternatives of the MRI model

a/a	MRI choice set	OD pairs
1	{CC, CL, CO, AV}	$[12, 13, 14, 21, 23, 24, 31, 32, 34, 41, 42, 43]$
2	{CC, CL, CO, AV}[1] $\rightarrow$ {B1, B2}[2]	$[15, 25, 35, 45, 16, 26, 36, 46]$
3	{B1, B2}[1] $\rightarrow$ {CC, CL, CO, AV}[2]	$[51, 52, 53, 54, 61, 62, 63, 64]$
4	{B1, B2}	$[57, 67, 75, 76]$

{MRI1, MRI2}[1]: first step decision, {MRI3, MRI4}[2]: second step decision