Transportation (2014) 41:351–375
DOI 10.1007/s11116-013-9489-y

Expanding the applicability of random regret minimization for route choice analysis

Carlo Giacomo Prato

Published online: 6 July 2013

Abstract The discrete choice paradigm of random regret minimization (RRM) has been recently proposed in several choice contexts. In the route choice context, the paradigm has been used to model the choice among three routes and to formulate regret-based stochastic user equilibrium. However, in the same context the RRM literature has not confronted three major challenges: (i) accounting for similarities across alternative routes, (ii) analyzing choice set composition effects on choice probabilities, and (iii) comparing RRM-based models with advanced RUM-based models. This paper looks into RRM-based route choice models from these three perspectives by (i) proposing utility-based and regret-based correction terms to account for similarities across alternatives, (ii) analyzing the variation of choice set probabilities with the choice set composition, and (iii) comparing RRM-based route choice models with C-Logit, Path Size Logit and Paired Combinatorial Logit. The results illustrate the definition of the correction terms within the regret function, the effect of the choice set specificity of RRM-based route choice models, and the positive performance of these models when compared to advanced RUM-based models.

Keywords Random regret minimization · Route choice modeling · Route similarity · Correction factor · Path size · Paired Combinatorial Logit

Introduction

The discrete choice paradigm of random utility maximization (RUM) is the most popular and most used approach to route choice modeling over the last decades. Although fuzzy logic, artificial neural networks and cognitive psychology have been proposed as alternative approaches, researchers generally model route choice behavior while considering

C. G. Prato (✉)
Department of Transport, Technical University of Denmark, Bygningstorvet 116 B,
2800 Kgs. Lyngby, Denmark
e-mail: cgp@transport.dtu.dk

travelers as utility maximizers selecting the most preferred route for moving from their origin to their destination (see, e.g., Prashker and Bekhor 2004; Prato 2009).

Recently, the discrete choice paradigm of random regret minimization (RRM) has been proposed in several choice contexts such as travel mode, parking location, shopping destination, leisure destination, departure time, vehicle type, and road pricing policies (see, e.g., Chorus 2010, 2012a). Behavioral foundations and mathematical properties of RRM-based models have been recently discussed, and theoretical and empirical comparisons with RUM-based models have been presented (see, e.g., Chorus 2012a, b).

An RRM-based model seems suitable to represent travelers' route choice behavior, because of the plausible assumption that travelers are regret minimizers selecting the route that makes them less regretful of not having chosen an alternative route. In the route choice context, an RRM-based model has been used in order to investigate the choice among three routes (Chorus 2012a) and to formulate an RRM-based stochastic user equilibrium (Bekhor et al. 2012). However, the literature does not present a discussion about RRM-based models tackling major challenges in route choice modeling: (i) accounting for similarities across alternative routes, (ii) analyzing the effect of choice set composition on choice probabilities, and (iii) comparing RRM-based models to advanced RUM-based models.

Accounting for similarities across alternative routes has not been considered in the RRM literature so far, as an RRM-based model was estimated on stated preference data regarding choices among three distinct routes (Chorus 2012a) and the stochastic user equilibrium formulation was proposed without taking into account route similarity although acknowledging research needs in this direction (Bekhor et al. 2012). The effect of choice set composition has not been examined in the RRM literature thus far, as the RRM-based model has not been estimated on route choice revealed preference data that pose the issue of choice set generation and hence choice set composition (see, e.g., Bovy 2009; Prato 2009). The comparison between RRM-based models and RUM-based models has been mainly limited to their multinomial logit (MNL) formulations (see, e.g., Chorus 2010, 2012b) with the exception of a mixed logit formulation (Chorus and de Jong 2011), rather than extended to the comparison with more advanced models and in particular the Paired Combinatorial Logit (PCL) model that expresses the choice of an alternative conditional on the choice of a pair of alternatives (Chu 1989; Koppelman and Wen 2000).

The current study investigates RRM-based models in the route choice context from three perspectives corresponding to the three aforementioned challenges.

With respect to the first perspective, the current study proposes correction factors to the RRM-MNL model that account for similarities across alternatives analogously to commonality factors and path sizes that correct the RUM-MNL model (see, e.g., Cascetta et al. 1996; Ben-Akiva and Bierlaire 1999; Bovy et al. 2008; Frejinger et al. 2009). In fact, although not suffering from the independence from irrelevant alternatives (IIA) property as the RUM-MNL model (Chorus 2012b), the RRM-MNL model bears the same inability to account for correlation across alternatives. Accordingly, three alternative approaches to the integration of correction factors in the RRM-MNL model are proposed: (i) adding utility-based corrections to the regret function in order to correct for the similarity of a route with respect to all other alternative routes within the choice set; (ii) adding a regret-based term to the regret function in order to express the comparison of degrees of similarity of a route with each alternative route; (iii) adding a regret-based term to the regret function in order to adjust for the correlation of a route with respect to each alternative route. An experimental analysis is performed on the basis of two simple networks, namely the "overlapping network" and the "switching route network" (see Prashker and Bekhor 2004). The variation in the choice probability of routes as a function of their overlapping with other routes is examined for the proposed RRM-based models.

With respect to the second perspective, the current study investigates the effect of the choice set composition on RRM-based route choice models. Intuitively, the choice set composition affects RRM-based models because of their choice set specificity related to the inclusion of alternatives that adds terms to the regret function. An experimental analysis is proposed on the basis of a grid network (see Bliemer and Bovy 2008) and the variation in the choice probability of routes is examined as a function of the composition of choice sets used for model estimation.

With respect to the third perspective, the comparison of RRM-based and RUM-based models is investigated with the estimation of route choice models from revealed preference data collected among commuters in an urban area of the north of Italy. In particular, the experimental analysis allows comparing not only the RRM-MNL and the RUM-MNL models, but also the proposed RRM-based models with existing RUM-based models such as RUM-MNL modifications and PCL.

The remainder of the paper is structured accordingly to the three perspectives of the investigation. "Correcting for similarity across alternatives" section proposes solutions to account for similarities across alternatives in RRM-based route choice models and examines the variation of the probabilities of choosing routes as a function of the degree of similarity. "Investigating the choice set composition effect" section illustrates an experimental analysis of the effects of choice set composition on route choice probabilities. "Comparing RUM-based and RRM-based route choice models" section compares the estimation of RRM-based and RUM-based route choice models in a revealed preference case-study. "Summary and conclusions" section draws the most significant conclusions from the investigation and suggests further research directions.

Correcting for similarity across alternatives

The literature in RUM-based route choice models presents several corrections to the route utility function that account for similarities across alternatives while maintaining the simple MNL model formulation (see, e.g., Prashker and Bekhor 2004; Prato 2009). The current study proposes conceptually similar corrections for RRM-based route choice models.

RRM-based route choice models

When considering RRM-based models in the route choice context, the regret of route i is written as (see, e.g., Chorus 2010, 2012b):

$$R_i = \sum_{j \neq i} \sum_m \ln\left(1 + \exp\left(\beta_m \left(x_{jm} - x_{im}\right)\right)\right) + \varepsilon_i \tag{1}$$

where R_i is the random regret associated with route i, x_{jm} and x_{im} are the values of attribute m for alternatives j and i, β_m are parameters to be estimated for the attributes m, and ε_i is the error term representing the unobserved regret. Assuming that the negative of the error term ε_i is Gumbel distributed, the probability P_i of selecting route i within the choice set C is formulated as (see, e.g., Chorus 2010, 2012b):

$$P_i = \frac{\exp\left(-\sum_{j\neq i}\sum_m \ln\left(1+\exp\left(\beta_m\left(x_{jm}-x_{im}\right)\right)\right)\right)}{\sum_{j\in C}\exp\left(-\sum_{k\neq j}\sum_m \ln\left(1+\exp\left(\beta_m\left(x_{km}-x_{jm}\right)\right)\right)\right)} \tag{2}$$

With respect to the RUM-MNL model, the RRM-MNL model does not suffer from the IIA property because the attributes of all alternatives in the choice set enter the regret function of any alternative (Chorus 2012b). However, in the route choice context the RRM-MNL model suffers from the same inability to account for similarities across alternatives, as the regret function does not consider whether the difference in the attributes between two alternatives derives from overlapping or non-overlapping links. The current study proposes three approaches to correct for similarities across alternatives with RRM-based models.

RRM-based model with utility-based corrections

Given the probability formulation of the RRM-MNL model, a first solution to the problem of accounting for similarities across alternatives consists in adding utility-based correction measures to the argument of the exponential function in the probability formulation.

Accordingly, a hybrid RUM-RRM function combines the regret for selecting route i and the disutility for the degree of similarity or independence of route i. Hybrid functions have been discussed from a behavioral perspective (Chorus 2012b), and their application appears reasonable also in the route choice context. The probability of choosing route i is formulated as follows:

$$P_i = \frac{\exp\left(-\sum_{j\neq i}\sum_m \ln\left(1+\exp\left(\beta_m\left(x_{jm}-x_{im}\right)\right)\right)+\beta_{corr}corr_i\right)}{\sum_{j\in C}\exp\left(-\sum_{k\neq j}\sum_m \ln\left(1+\exp\left(\beta_m\left(x_{km}-x_{jm}\right)\right)\right)+\beta_{corr}corr_j\right)} \tag{3}$$

where $corr_i$ is a utility-based correction term for route i and β_{corr} is a parameter to be estimated. The correction term may be expressed by (i) commonality factors, which decrease the utility of a route because of its degree of similarity with the alternative routes, or (ii) path size measures, which indicate the fraction of a route that constitutes a "full" alternative.

In the former case, the commonality factors may be expressed as (Cascetta et al. 1996; Cascetta 2001):

$$corr_i = CF_i = \ln\sum_{j\in C}\left(\frac{L_{ij}}{\sqrt{L_i L_j}}\right) \tag{4}$$

$$corr_i = CF_i = \ln\left[1+\sum_{\substack{j\in C\\ j\neq i}}\left(\frac{L_{ij}}{\sqrt{L_i L_j}}\right)\left(\frac{L_i-L_{ij}}{L_j-L_{ij}}\right)\right] \tag{5}$$

where L_i is the length of route i, L_j is the length of alternative route j within the choice set C, and L_{ij} is the overlapping length between routes i and j. It should be noted that the two formulations represent different concepts of similarity: the original formulation (Eq. 4) depends exclusively on the overlapping length between routes, while the modified formulation (Eq. 5) depends also on the costs of the disjoint links and suggests that the ratio between the commonality factors of two routes should increase when the overlapping between the two routes increases. Both formulations are always positive and have a null lower bound, as the arguments of the logarithms are equal to one for independent routes.

In the latter case, the path size measures may be expressed as either the original formulation (Eq. 6) derived for an application of discrete choice theory for aggregate alternatives (Ben-Akiva and Bierlaire 1999) or a modified formulation (Eq. 7) derived from an approximation of GEV models (Bovy et al. 2008):

$$corr_i = \ln(PS_i) = \ln\left(\sum_{a\in\Gamma_i}\frac{L_a}{L_i}\frac{1}{\sum_{j\in C}\delta_{aj}}\right) \tag{6}$$

$$corr_i = PSC_i = -\sum_{a\in\Gamma_i}\left(\frac{L_a}{L_i}\ln\sum_{j\in C}\delta_{aj}\right) \tag{7}$$

where PS_i is the path size of route i, PSC_i is the path size correction of route i, L_i is the length of route i, L_a is the length of link a, Γ_i is the set of links belonging to route i, δ_{aj} is the link-route incidence dummy that is equal to one if route j within the choice set C uses links a and zero otherwise. Both formulations have an upper bound for completely independent routes, but the PSC_i does not have a lower bound because there is no upper bound on the number of paths sharing a link (Bovy et al. 2008).

Both commonality factors imply that the estimated parameters β_{corr} have expected negative signs to suggest that the utility is reduced when route i overlaps with the alternative routes in the choice set. Both path size measures imply that the estimated parameters β_{corr} have expected positive signs to indicate that the utility is reduced when the route i is not an independent alternative. It should be noted that the role of commonality factors and path size measures in RRM-based models is analogous to their role in RUM-based models, and hence the expected sign of the parameters is the same.

PS-RRM

Given that the regret function accounts for pairwise comparisons of alternatives at the attribute level, a second solution to the problem of considering similarities across alternatives consists in adding to the regret function a term expressing the pairwise comparison of the degree of independence of alternatives.

Accordingly, the path size (or the path size correction) of route i is compared to the one of each alternative route j in the choice set C within the regret function. The probability of choosing route i is formulated in the PS-RRM model (where either PS_i or PSC_i may be inserted):

$$P_i = \frac{\exp\left(-\sum_{j\neq i}\sum_m \ln\left(1+\exp\left(\beta_m\left(x_{jm}-x_{im}\right)\right)\right) - \sum_{j\neq i}\ln\left(1+\exp\left(\beta_{corr}\left(PS_j-PS_i\right)\right)\right)\right)}{\sum_{j\in C}\exp\left(-\sum_{k\neq j}\sum_m \ln\left(1+\exp\left(\beta_m\left(x_{km}-x_{jm}\right)\right)\right) - \sum_{k\neq j}\ln\left(1+\exp\left(\beta_{corr}\left(PS_k-PS_j\right)\right)\right)\right)} \tag{8}$$

For an interpretation of the sign of the parameter β_{corr}, it should be noted that the regret function enters with a negative sign in the probability formulation (see, e.g., Chorus 2012a, b). Intuitively, the expected sign of the parameter β_{corr} is positive to indicate that the regret for the considered route diminishes when this route is more independent than the alternative ones, and increases when this route is less distinct than the alternative ones. The difference between path size terms is in fact negative in the former case, and positive in the latter.

C-RRM

Given that the regret function encompasses pairwise comparisons of alternatives at the attribute level, a third solution to the problem of considering similarities across alternatives consists in adding a term expressing the pairwise correlation between alternatives.

Generally, the random error variance of routes is related to some size measure of the route (see, e.g., Daganzo and Sheffi 1977):

$$\sigma(\varepsilon_i) = k \cdot L_i \tag{9}$$

where $\sigma(\varepsilon_i)$ is the variance of route i and k is a proportionality factor with the length L_i. The covariance $\sigma(\varepsilon_{i,},\varepsilon_j)$ between routes i and j within the choice set C is related to their physical overlap and, assuming the aforementioned proportionality relationship, it is related to the common route part L_{ij}:

$$\sigma(\varepsilon_i, \varepsilon_j) = k \cdot L_{ij} \tag{10}$$

The correlation $\rho(\varepsilon_{i,},\varepsilon_j)$ between the two routes is calculated as the correlation between two random variates:

$$\rho(\varepsilon_i, \varepsilon_j) = \frac{\sigma(\varepsilon_i, \varepsilon_j)}{\sqrt{\sigma(\varepsilon_i)}\sqrt{\sigma(\varepsilon_j)}} = \frac{L_{ij}}{\sqrt{L_i L_j}} \tag{11}$$

The correlation may also be weighted according to the ratios between the disjoint and the overlapping part of the two routes, in order to indicate the higher weight for the route with the smaller disjoint part:

$$\rho(\varepsilon_i, \varepsilon_j) = \frac{\sigma(\varepsilon_i, \varepsilon_j)}{\sqrt{\sigma(\varepsilon_i)}\sqrt{\sigma(\varepsilon_j)}} \frac{\left(\frac{L_i - L_{ij}}{L_{ij}}\right)}{\left(\frac{L_j - L_{ij}}{L_{ij}}\right)} = \frac{L_{ij}}{\sqrt{L_i L_j}} \left(\frac{L_i - L_{ij}}{L_j - L_{ij}}\right) \tag{12}$$

Given the definition of correlation for route i with respect to each alternative j, the correlation terms are inserted within the regret function to correct for similarities across alternatives. The probability of choosing route i is formulated in the C-RRM model:

$$P_i = \frac{\exp\left(-\sum_{j \neq i}\sum_m \ln\left(1 + \exp\left(\beta_m\left(x_{jm} - x_{im}\right)\right)\right) - \sum_{j \neq i} \ln\left(1 + \exp\left(\beta_{corr}\rho\left(\varepsilon_i, \varepsilon_j\right)\right)\right)\right)}{\sum_{j \in C} \exp\left(-\sum_{k \neq j}\sum_m \ln\left(1 + \exp\left(\beta_m\left(x_{km} - x_{jm}\right)\right)\right) - \sum_{k \neq j} \ln\left(1 + \exp\left(\beta_{corr}\rho\left(\varepsilon_j, \varepsilon_k\right)\right)\right)\right)} \tag{13}$$

where the correlation terms may be expressed by both formulations (11) and (12). Intuitively, the expected sign of the parameter β_{corr} is positive to indicate that the regret for the considered route diminishes when the correlation is high. This proposition appears logical when considering the regret for choosing route i over route j: in the case that L_{ij} is null, the regret will be proportional to the difference between L_i and L_j; in the case that L_{ij} increases, the regret will be inferior because not imputable to the common part.

Experimental analysis

Two simple networks are selected from the literature to illustrate the performances of the proposed models in terms of variation of route choice probabilities as a function of the degree of similarity. Figure 1 illustrates the two simple networks, namely the "overlapping

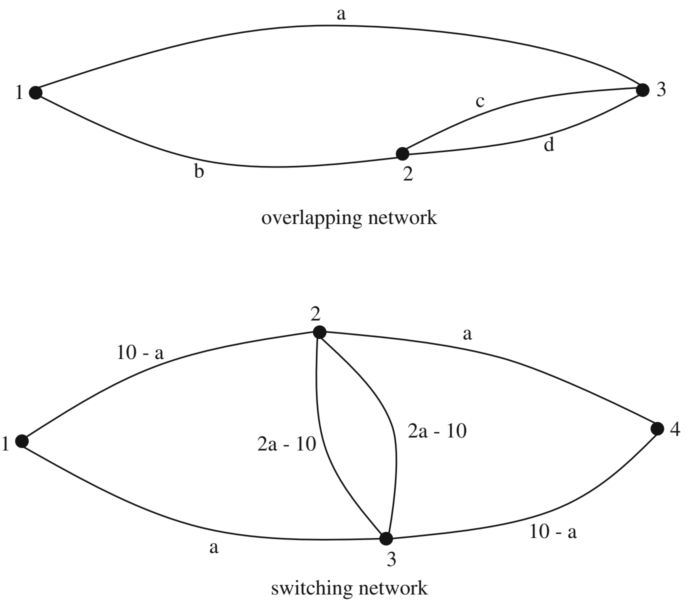


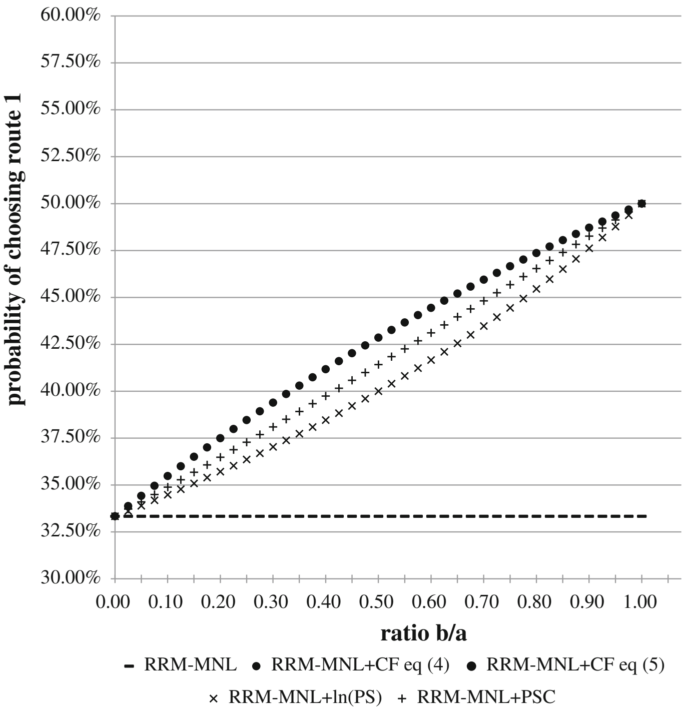
Fig. 1 Experimental networks

network" (see Daganzo and Sheffi 1977; Prashker and Bekhor 2004) and the "switching route network" (see Prashker and Bekhor 2004).

The "overlapping network" allows identifying three alternative routes, with one being independent and the remaining two being overlapping because of the common link *b*. It should be noted that the three routes share the same cost of 10 units regardless of the cost of link *b*. The "switching route network" allows identifying four alternative routes, with three routes (1-2-4, 1-3-4, 1-2-3-4) sharing the same cost of 10 units regardless of the cost *a*, while the fourth route 1-3-2-4 has cost ($4a - 10$) that is equal to the other routes when $a = 5$ and becomes increasingly less attractive when *a* increases. It should be noted that the cost *a* varies between 5 and 10 in order not to have negative link costs.

Probabilities for choosing the alternative routes are calculated for the proposed RRM-based route choice models. Generalized route cost is the only attribute *m* for the calculation of the regret of the alternative routes, and the cost parameter β_m is assumed to be equal to -1. The parameters β_{corr} are assumed according to the expected signs previously discussed: for utility-based corrections, β_{corr} is assumed to be equal to -1 when $corr_i$ represents a commonality factor and $+1$ when $corr_i$ represents a path size; for regret-based corrections, β_{corr} is assumed to be equal to $+1$ for both PS-RRM and C-RRM models.

Figure 2 illustrates the probability of selecting route 1 (i.e., link *a*) as a function of the ratio between the common link *b* and the route cost. In the two extreme cases, this probability should be equal to 33.3 % when the three alternative routes are completely independent and to 50.0 % when the remaining two routes are completely overlapping. It should be noted that, even though the RRM-MNL model does not exhibit the IIA property, the regret function accounts for the (null) difference between route costs and ignores the cost of link *b*, and hence always leads to equal choice probabilities for the three alternative routes.



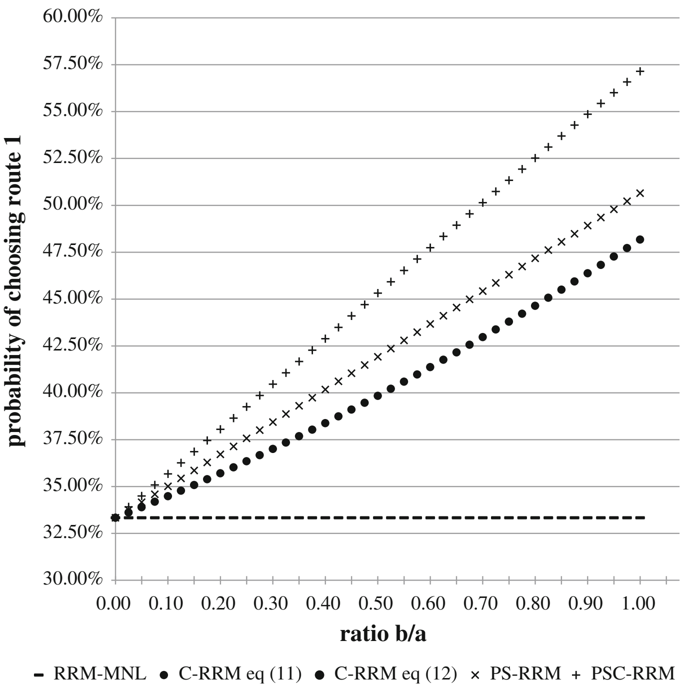


Fig. 2 Similarity effect in the "overlapping network"

Similarly to RUM-MNL modifications (see Prashker and Bekhor 2004), the utility-based corrections to the RRM-MNL model allow retrieving observing the expected probabilities at the two extreme cases. The curves are either convex or concave in the middle section, with the definition of similarity playing a role since convex curves result from commonality factors and concave curves result from path sizes. The regret-based corrections do not allow observing the expected probabilities at the two extreme cases. PSC-RRM highly overestimates and PS-RRM slightly overestimates the probability of choosing route 1 when the two alternative routes are equivalent. Attempts to mitigate this effect by reducing the parameter β_{corr} have been performed without obtaining better results. C-RRM models exhibit the same curve regardless of the formulation considered for the pairwise correlation between alternatives, slightly underestimate the probability in the extreme case of two independent routes, and show a concave shape which is different from the convex one when commonality factors are added as utility-based terms. Overall, utility-based corrections appear more successful in obtaining the expected probabilities at the extreme cases in this simple network, but C-RRM formulations appear promising and outperform more complex model approaches such as GEV models (see Prashker and Bekhor 2004).

Figure 3 shows the probability of selecting route 3 (i.e., 1-2-3-4) as a function of cost a. In the two extreme cases, this probability should be equal to 25 % when the four routes are completely independent and to 33.3 % when three routes are independent and the fourth is highly inconvenient. It should be noted that the RRM-MNL model manages to reproduce the expected probabilities at the extreme cases.

With the exception of the modified commonality factor that accounts for the disjoint part of the route, the utility-based corrections do not allow retrieving the expected probabilities at the two extreme cases and largely overestimate the probability of route 3. Similar results and hence similar problems are presented by Prashker and Bekhor (2004) for RUM-MNL modifications. Regret-based corrections illustrate similar behavior, with PS-RRM and PSC-RRM largely overestimating the probability of route 3 at the extreme case. C-RRM shows correct values for the weighted formulation of the correlation between alternative routes (Eq. 12).

The purpose of the analysis with the two simple networks presented above is to illustrate the importance of the specification of similarity measures in the RRM-based route choice models. Similarly to RUM-MNL modifications and GEV models (see Prashker and Bekhor 2004), models perform positively with one network and unsatisfactorily with the other. From this experimental analysis, it appears that the most promising correction is the regret-based correction accounting for the correlation between alternative routes weighted according to the disjoint parts of the routes.

Investigating the choice set composition effect

The literature in RUM-based route choice models has shown growing interest in comprehending the role of choice set composition on model estimates and traffic flow predictions (e.g., Bliemer et al. 2007; Prato and Bekhor 2007; Bekhor et al. 2008; Bliemer and Bovy 2008). The same question applies to RRM-based route choice models, considering that the calculation of regret functions and route choice probabilities requires the definition of a choice set. The current study presents an experimental analysis to help evaluating the effects of choice set composition on the proposed RRM-based models.

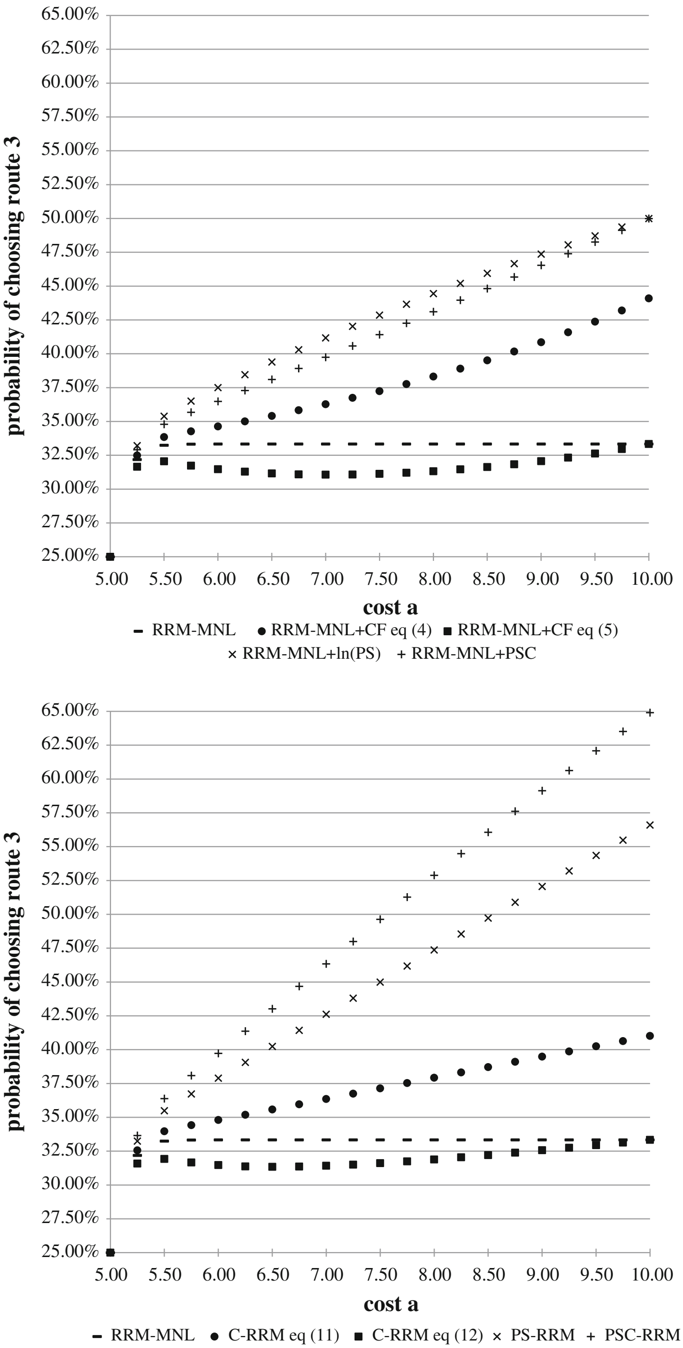


Fig. 3 Similarity effect in the "switching route network"

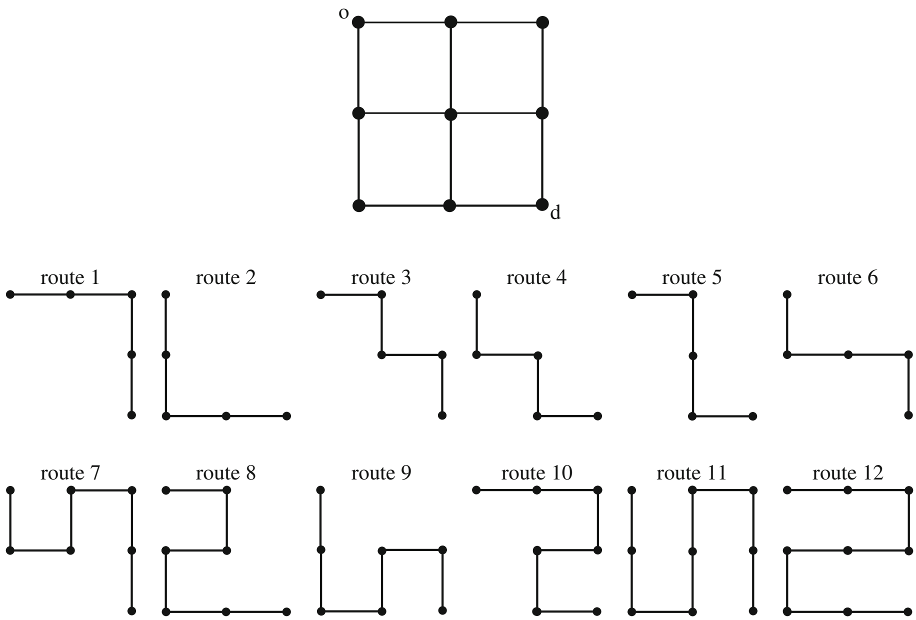


Fig. 4 Experimental grid network

Experimental design

A simple network is selected from the literature to illustrate the performance of the proposed models in terms of variation of route choice probabilities as a function of choice set composition. The network selection is motivated by the desire to maximize the control of the size and composition of the route choice set through the complete knowledge about relevant routes, irrelevant routes and choice set properties (Bliemer and Bovy 2008). Figure 4 illustrates the grid network with bidirectional links of unit cost (see Bliemer and Bovy 2008). The network contains 12 possible alternatives for the considered o–d pair: routes 1-6 have minimum cost equal to 4, routes 7-10 have cost equal to 6, and routes 11-12 have maximum cost equal to 8.

A probit model is assumed to represent the true route choice behavior. Links have a deterministic component c_a equal to their cost and a stochastic component ε_a that is normally distributed with null mean and variance σ_a^2. For comparability reasons, the variances at the route level of the probit and the RRM-based models are constructed equal. Accordingly, the variance $var(\varepsilon_i)$ of route i is equal to $\pi^2/6\mu^2$ and σ_a^2 is equal to 1/4 of this value while assuming cost equal to 4 (for details, see Bliemer and Bovy 2008). The true probit route choice probabilities for the twelve alternative routes are computed by 10,000 Monte Carlo simulations, and are presented in Table 1. Interestingly, the six shortest routes are not equally preferred because of different degrees of similarity. Specifically, the more the routes overlap, the lower their probability is in favor of independent alternatives. Routes 1-2 have a higher probability of being selected because of the lower overlap with other alternatives than routes 3-4. The longer routes are far less relevant, and they have only a limited probability of being selected.

Table 1 Route choice probability as a function of choice set composition

Choice set	Model	β_m	β_{corr}	RMSE	1	2	3	4	5	6	7	8	9	10	11	12
	MNP				0.1955	0.1955	0.1451	0.1451	0.1534	0.1534	0.0030	0.0030	0.0030	0.0030	0.0000	0.0000
Route probability with estimates from 6 alternative routes																
12 alt.	RRM-MNL			0.0156	0.1646	0.1646	0.1646	0.1646	0.1646	0.1646	0.0031	0.0031	0.0031	0.0031	0.0000	0.0000
	RRM-MNL + CF eq(4)			0.0268	0.1368	0.1368	0.1621	0.1621	0.1621	0.1621	0.0184	0.0184	0.0184	0.0184	0.0022	0.0022
	RRM-MNL + CF eq(5)			0.0070	0.1863	0.1863	0.1393	0.1393	0.1667	0.1667	0.0038	0.0038	0.0038	0.0038	0.0001	0.0001
	RRM-MNL + ln(PS)			0.0284	0.1335	0.1335	0.1596	0.1596	0.1596	0.1596	0.0221	0.0221	0.0221	0.0221	0.0030	0.0030
	RRM-MNL + PSC			0.0291	0.1317	0.1317	0.1607	0.1607	0.1607	0.1607	0.0220	0.0220	0.0220	0.0220	0.0030	0.0030
	C-RRM eq(11)			0.0331	0.1302	0.1302	0.1853	0.1853	0.1798	0.1798	0.0024	0.0024	0.0024	0.0024	0.0000	0.0000
	C-RRM eq(12)			0.0074	0.1861	0.1861	0.1450	0.1450	0.1684	0.1684	0.0002	0.0002	0.0002	0.0002	0.0000	0.0000
	PS-RRM			0.0341	0.1271	0.1271	0.1829	0.1829	0.1829	0.1829	0.0036	0.0036	0.0036	0.0036	0.0000	0.0000
	PSC-RRM			0.0361	0.1232	0.1232	0.1849	0.1849	0.1849	0.1849	0.0035	0.0035	0.0035	0.0035	0.0000	0.0000
10 alt.	RRM-MNL			0.0156	0.1637	0.1637	0.1637	0.1637	0.1637	0.1637	0.0044	0.0044	0.0044	0.0044	–	–
	RRM-MNL + CF eq(4)			0.0192	0.1560	0.1560	0.1514	0.1514	0.1514	0.1514	0.0206	0.0206	0.0206	0.0206	–	–
	RRM-MNL + CF eq(5)			0.0070	0.1863	0.1863	0.1394	0.1394	0.1667	0.1667	0.0038	0.0038	0.0038	0.0038	–	–
	RRM-MNL + ln(PS)			0.0229	0.1496	0.1496	0.1496	0.1496	0.1496	0.1496	0.0256	0.0256	0.0256	0.0256	–	–
	RRM-MNL + PSC			0.0227	0.1500	0.1500	0.1500	0.1500	0.1500	0.1500	0.0250	0.0250	0.0250	0.0250	–	–
	C-RRM eq(11)			0.0131	0.1693	0.1693	0.1628	0.1628	0.1591	0.1591	0.0044	0.0044	0.0044	0.0044	–	–
	C-RRM eq(12)			0.0053	0.1906	0.1906	0.1441	0.1441	0.1647	0.1647	0.0003	0.0003	0.0003	0.0003	–	–
	PS-RRM			0.0158	0.1624	0.1624	0.1624	0.1624	0.1624	0.1624	0.0064	0.0064	0.0064	0.0064	–	–
	PSC-RRM			0.0158	0.1625	0.1625	0.1625	0.1625	0.1625	0.1625	0.0062	0.0062	0.0062	0.0062	–	–
6 alt.	RRM-MNL	−1.000		0.0158	0.1667	0.1667	0.1667	0.1667	0.1667	0.1667	–	–	–	–	–	–
	RRM-MNL + CF eq(4)	−1.000	−1.210	0.0033	0.1979	0.1979	0.1511	0.1511	0.1511	0.1511	–	–	–	–	–	–
	RRM-MNL + CF eq(5)	−1.000	−2.035	0.0073	0.1892	0.1892	0.1415	0.1415	0.1693	0.1693	–	–	–	–	–	–
	RRM-MNL + ln(PS)	−1.000	0.575	0.0033	0.1979	0.1979	0.1510	0.1510	0.1510	0.1510	–	–	–	–	–	–
	RRM-MNL + PSC	−1.000	0.779	0.0033	0.1979	0.1979	0.1510	0.1510	0.1510	0.1510	–	–	–	–	–	–
	C-RRM eq(11)	−1.000	0.951	0.0029	0.1982	0.1982	0.1499	0.1499	0.1520	0.1520	–	–	–	–	–	–

Table 1 continued

Choice set	Model MNP	β_m	β_{corr}	RMSE	1 0.1955	2 0.1955	3 0.1451	4 0.1451	5 0.1534	6 0.1534	7 0.0030	8 0.0030	9 0.0030	10 0.0030	11 0.0000	12 0.0000
	C-RRM eq(12)	−1.000	0.951	0.0029	0.1982	0.1982	0.1499	0.1499	0.1520	0.1520	–	–	–	–	–	–
	PS-RRM	−1.000	0.193	0.0033	0.1979	0.1979	0.1511	0.1511	0.1511	0.1511	–	–	–	–	–	–
	PSC-RRM	−1.000	0.262	0.0033	0.1979	0.1979	0.1511	0.1511	0.1511	0.1511	–	–	–	–	–	–
Route probability with estimates from 12 alternative routes																
12 alt.	RRM-MNL	−0.348		0.0156	0.1646	0.1646	0.1646	0.1646	0.1646	0.1646	0.0031	0.0031	0.0031	0.0031	0.0000	0.0000
	RRM-MNL + CF eq(4)	−0.364	1.924	0.0024	0.1955	0.1955	0.1492	0.1492	0.1492	0.1492	0.0030	0.0030	0.0030	0.0030	0.0001	0.0001
	RRM-MNL + CF eq(5)	−0.295	−2.038	0.0070	0.1869	0.1869	0.1397	0.1397	0.1672	0.1672	0.0030	0.0030	0.0030	0.0030	0.0001	0.0001
	RRM-MNL + ln(PS)	−0.337	−0.872	0.0024	0.1955	0.1955	0.1492	0.1492	0.1492	0.1492	0.0030	0.0030	0.0030	0.0030	0.0001	0.0001
	RRM-MNL + PSC	−0.339	−1.059	0.0024	0.1955	0.1955	0.1492	0.1492	0.1492	0.1492	0.0030	0.0030	0.0030	0.0030	0.0001	0.0001
	C-RRM eq(11)	−0.364	−0.982	0.0039	0.1935	0.1935	0.1525	0.1525	0.1477	0.1477	0.0031	0.0031	0.0031	0.0031	0.0000	0.0000
	C-RRM eq(12)	−0.263	1.084	0.0066	0.1873	0.1873	0.1400	0.1400	0.1665	0.1665	0.0031	0.0031	0.0031	0.0031	0.0000	0.0000
	PS-RRM	−0.336	−0.147	0.0024	0.1955	0.1955	0.1491	0.1491	0.1491	0.1491	0.0031	0.0031	0.0031	0.0031	0.0000	0.0000
	PSC-RRM	−0.338	−0.178	0.0024	0.1955	0.1955	0.1491	0.1491	0.1491	0.1491	0.0031	0.0031	0.0031	0.0031	0.0000	0.0000
10 alt.	RRM-MNL			0.0156	0.1637	0.1637	0.1637	0.1637	0.1637	0.1637	0.0045	0.0045	0.0045	0.0045	–	–
	RRM-MNL + CF eq(4)			0.0182	0.1597	0.1597	0.1676	0.1676	0.1676	0.1676	0.0026	0.0026	0.0026	0.0026	–	–
	RRM-MNL + CF eq(5)			0.0070	0.1869	0.1869	0.1398	0.1398	0.1673	0.1673	0.0030	0.0030	0.0030	0.0030	–	–
	RRM-MNL + ln(PS)			0.0156	0.1650	0.1650	0.1650	0.1650	0.1650	0.1650	0.0025	0.0025	0.0025	0.0025	–	–
	RRM-MNL + PSC			0.0156	0.1650	0.1650	0.1650	0.1650	0.1650	0.1650	0.0026	0.0026	0.0026	0.0026	–	–
	C-RRM eq(11)			0.0178	0.1600	0.1600	0.1682	0.1682	0.1641	0.1641	0.0038	0.0038	0.0038	0.0038	–	–
	C-RRM eq(12)			0.0047	0.1917	0.1917	0.1385	0.1385	0.1617	0.1617	0.0041	0.0041	0.0041	0.0041	–	–
	PS-RRM			0.0156	0.1641	0.1641	0.1641	0.1641	0.1641	0.1641	0.0038	0.0038	0.0038	0.0038	–	–
	PSC-RRM			0.0156	0.1641	0.1641	0.1641	0.1641	0.1641	0.1641	0.0039	0.0039	0.0039	0.0039	–	–
6 alt.	RRM-MNL			0.0158	0.1667	0.1667	0.1667	0.1667	0.1667	0.1667	–	–	–	–	–	–
	RRM-MNL + CF eq(4)			0.0375	0.1228	0.1228	0.1886	0.1886	0.1886	0.1886	–	–	–	–	–	–
	RRM-MNL + CF eq(5)			0.0073	0.1892	0.1892	0.1415	0.1415	0.1693	0.1693	–	–	–	–	–	–

Table 1 continued

Choice set	Model	β_m	β_{corr}	RMSE	1	2	3	4	5	6	7	8	9	10	11	12
	MNP				0.1955	0.1955	0.1451	0.1451	0.1534	0.1534	0.0030	0.0030	0.0030	0.0030	0.0000	0.0000
	RRM-MNL + ln(PS)			0.0366	0.1246	0.1246	0.1877	0.1877	0.1877	0.1877	–	–	–	–	–	–
	RRM-MNL + PSC			0.0346	0.1286	0.1286	0.1857	0.1857	0.1857	0.1857	–	–	–	–	–	–
	C-RRM eq(11)			0.0267	0.1444	0.1444	0.1765	0.1765	0.1791	0.1791	–	–	–	–	–	–
	C-RRM eq(12)			0.0041	0.2034	0.2034	0.1469	0.1469	0.1496	0.1496	–	–	–	–	–	–
	PS-RRM			0.0268	0.1444	0.1444	0.1778	0.1778	0.1778	0.1778	–	–	–	–	–	–
	PSC-RRM			0.0256	0.1467	0.1467	0.1767	0.1767	0.1767	0.1767	–	–	–	–	–	–

The RRM-based route choice models are estimated with a regret function containing only the cost attribute m for two choice set compositions consisting of (i) only the relevant routes and (ii) all the relevant and the irrelevant routes. Observed route choices correspond to the minimum cost routes for each of the 10,000 observations from the Monte Carlo simulation of the postulated model. After estimates for β_m and β_{corr} are obtained from one choice set, choice probabilities for the alternative routes are calculated by applying the estimated parameters to different choice sets, and RMSE are computed with respect to the probit probabilities in order to evaluate the ability to reproduce the true route choice probabilities (see, e.g., Nerella and Bhat 2004; Bliemer and Bovy 2008). It should be noted that the model is unidentified when only the relevant routes are considered since they share the same minimum cost, and hence the parameter β_m is fixed equal to -1.

Experimental results

Table 1 presents the route choice probabilities as a function of choice set composition. The table is divided in two parts referring to the route choice probabilities calculated after the application of the model estimated with respectively, 6 and 12 alternative routes. Each of the two parts is further divided into three sections referring to the route choice probabilities calculated after the application of the estimated model to the choice sets with respectively 6, 10 and 12 alternative routes. Estimates of the RRM-based route choice models are reported for each of the two parts (i.e., for 6 and 12 alternative routes), and the RMSE are presented alongside the choice probabilities of the 6, 10 or 12 routes for each section of the two parts.

Regardless of the choice set composition used for model estimation, the sign of parameter β_m is negative to correctly imply that minimizing regret corresponds to minimizing travel time. When considering the choice set with only the 6 relevant routes, the signs of the parameters β_{corr} correspond to the expected signs to indicate that the similarity is captured as expected. When considering the choice set with all the 12 alternative routes, the signs of the parameters β_{corr} are often the opposite of the expected ones. Only the modified commonality factor (Eq. 5) and the weighted correlation (Eq. 12) allow obtaining the expected signs, suggesting that considering also the portion of disjoint routes is beneficial in the correction of the regret function. It should be noted that counterintuitive results, which indicate the route overlap being weighted positively so that more utility is added to a route with a higher overlap, are also observed for RUM-MNL modifications (Bliemer and Bovy 2008). The composition of the choice set causes these results, as for example routes 1-2 overlap significantly with irrelevant routes and hence estimated parameters correct for their utility in the direction opposite to the one expected. Clearly, the choice set composition affects model estimates and the insertion of irrelevant route changes the purpose of utility-based and regret-based corrections.

When considering choice probabilities calculated by using the estimates from the choice set containing only the 6 relevant routes, models exhibit a relatively low RMSE when replicating the route choice probabilities of the probit model. Among utility-based corrections, the difference in the route choice probabilities of routes 1-2, 5-6 and 3-4 is captured only by the modified commonality factor (Eq. 5). None of the other corrections distinguishes between routes 5-6 and routes 3-4 in terms of choice probabilities. Among regret-based corrections, the same behavior is observed as C-RRM models perform better than the PS-RRM and the PSC-RRM models, and the weighted correlation (Eq. 12) produces the best results. Adding irrelevant alternative routes for calculating the probabilities with the estimates obtained from the choice set with only relevant alternatives is not beneficial. Adding routes 7-10 leads to obtaining very similar results conceptually, although the probabilities of these same routes are highly overestimated. Further adding routes 11-12 drives not to attain the same results, as routes 1-2 become significantly less preferable because of their high degree of similarity with the maximum cost routes. This same effect has been observed for RUM-MNL modifications (Bliemer and Bovy 2008).

When considering choice probabilities computed by using the estimates from the choice set including all the 12 alternative routes, models exhibit a comparably low RMSE when replicating the route choice probabilities of the probit model. Among utility-based corrections, the difference in the route choice probabilities of routes 1-2, 5-6 and 3-4 is captured only by the modified commonality factor (Eq. 5). However, this same commonality factor produces the highest RMSE and hence the worst reproduction of the probit probabilities. Among regret-based corrections, the same behavior is observed as C-RRM models reproduce the correct differentiation, but the PS-RRM and the PSC-RRM models better reproduce the probit probabilities overall. The inability to replicate the postulated behavior increases when removing irrelevant alternative routes. Removing routes 11-12 causes routes 1-2 not being preferable with respect to the other relevant routes, while removing routes 7-10 further increases this tendency. The C-RRM performs significantly better than the other models, both from the raw probability perspective and the total error perspective.

Similarly to RUM-MNL modifications and GEV models (Bliemer and Bovy 2008), RRM-based route choice models can quite accurately replicate route choice probabilities only for the choice set on which they are estimated, but probabilities are quite different when smaller or larger choice sets are considered. These empirical results are hardly surprising when considering that RRM-based models are choice set specific. The additive regret function implies obtaining smaller (larger) parameter estimates for RRM-based models estimated on larger (smaller) choice sets, and consequently predictions made on the basis of an estimated RRM-based model should refer to a choice set of the same size as the one used for model estimation (Chorus 2012a). Being choice set specific is an advantage of RRM-based models, which are able to capture choice set composition effects and hence cannot be transferred seamlessly to different choice sets. Also, being choice set specific reminds route choice modelers that the definition of choice sets for individual travelers plays a major role for RRM-based models as for RUM-based route choice models.

Comparing RUM-based and RRM-based route choice models

The literature in RRM has shown comparisons between RRM-MNL and RUM-MNL in several choice contexts. As RRM-based models have been proposed in the route choice context, the current study answers the question of their comparison with advanced RUM-based models (i.e., C-Logit, PS-Logit, PSC-Logit, PCL) through the estimation for revealed preference data collected in an urban network.

Estimation data and methods

A web-based survey allowed to observe route choice behavior of car drivers commuting in the morning in the metropolitan area of Torino (Italy). The considered network consists of 23 districts, 92 zones, 417 nodes, and 1427 links, and covers an area containing roughly 900,000 inhabitants within the city limits. The network includes major arterials crossing the city from north to south and from east to west, main roads that connect different districts of the city, minor roads that link points within the same district, and some local roads.

The sample for model estimation consists of 575 routes from the 236 survey participants, as some recorded more than one chosen route. Routes are in average 4.8 km and 15.4 min long, with standard deviations of 2.0 km and 5.9 min. Alternative routes are generated with a modified branch and bound algorithm (Prato and Bekhor 2006) that produces 2–19 alternatives per observation with a median value of 11 alternatives. The coverage of the observed routes (see Ramming 2002) is equal to 85.4 % with a 100 % overlap threshold and 91.3 % with an 80 % overlap threshold, thus showing high realism of the implemented path generation technique with respect to the observed behavior. It should be noted that all the observations were covered at least at the 65.4 % level, and the observed routes not reproduced at the 80 % overlap threshold were added to the generated choice set. Further details on the dataset and the choice set generation are presented by Prato et al. (2012).

Estimated models are the proposed RRM-based models and their RUM-based counterparts. For RUM-MNL modifications, the probability of choosing route i is formulated as follows:

$$P_i = \frac{\exp\left(\sum_m \beta_m x_{im} + \beta_{corr} corr_i\right)}{\sum_{j \in C} \exp\left(\sum_m \beta_m x_{jm} + \beta_{corr} corr_j\right)} \tag{14}$$

where x_{im} is the value of attribute m for route i, and the sum over the attributes of the product between parameters β_m and values x_{im} represents the deterministic part of the utility function. The correction terms $corr_i$ may correspond to the commonality factors in Eqs. (4) and (5) to have the C-Logit model formulation (Cascetta et al. 1996; Cascetta 2001), or to the path size measures in Eqs. (6) and (7) to have the PS-Logit (Ben-Akiva and Bierlaire 1999) and the PSC-Logit (Bovy et al. 2008). In addition to RUM-MNL modifications, the PCL model is estimated in order to compare the RRM-MNL model with the RUM-based model considering pairwise comparison of alternatives. For the PCL model, the probability of choosing route i is formulated as follows (see Prashker and Bekhor 1998; Koppelman and Wen 2000):

$$P_i = \sum_{i \neq j} P(ij) P(i|ij) \tag{15}$$

where $P(ij)$ is the marginal probability of choosing pair (i,j) among the $n(n-1)/2$ possible pairs of alternative routes, and $P(i|ij)$ is the conditional probability of choosing route i given the chosen pair (i,j). Conditional and marginal probabilities depend on the degree of similarity between routes within the pair:

$$P(i|ij) = \frac{\exp\left(\frac{\sum_m \beta_m x_{im}}{1-\sigma_{ij}}\right)}{\exp\left(\frac{\sum_m \beta_m x_{im}}{1-\sigma_{ij}}\right) + \exp\left(\frac{\sum_m \beta_m x_{jm}}{1-\sigma_{ij}}\right)} \tag{16}$$

$$P(ij) = \frac{\left(1-\sigma_{ij}\right)\left(\exp\left(\frac{\sum_m \beta_m x_{im}}{1-\sigma_{ij}}\right) + \exp\left(\frac{\sum_m \beta_m x_{jm}}{1-\sigma_{ij}}\right)\right)^{1-\sigma_{ij}}}{\sum_{k=1}^{n-1} \sum_{l=k+1}^{n} \left(1-\sigma_{kl}\right)\left(\exp\left(\frac{\sum_m \beta_m x_{km}}{1-\sigma_{kl}}\right) + \exp\left(\frac{\sum_m \beta_m x_{lm}}{1-\sigma_{kl}}\right)\right)^{1-\sigma_{kl}}} \tag{17}$$

The similarity coefficient σ_{ij} between routes i and j is defined with a parameterization similar to the commonality factor presented in Eq. 4 (Prashker and Bekhor 1998):

$$\sigma_{ij} = \left(\frac{L_{ij}}{\sqrt{L_i L_j}}\right)^{\gamma_\sigma} \tag{18}$$

where γ_σ is a parameter to be estimated.

All route choice models are estimated with GAUSS matrix programming language and are compared in terms of model performances and value-of-time.

Goodness-of-fit of a model M is measured as McFadden's corrected likelihood ratio index ρ_M^2:

$$\rho_M^2 = 1 - \frac{LL_\beta - k_{M\partial}}{LL_0} \tag{19}$$

where LL_β is the log-likelihood at the estimates, LL_0 is the log-likelihood at zero, and k_M is the number of parameters estimated. The comparison of model performances requires to consider whether models are nested in a statistical sense, namely whether a model may be considered a special case of an alternative model by assigning appropriate values to some subset of the alternative model's parameters (see, e.g., Horowitz 1983). Likelihood ratio (LR) tests are used to compare nested models, and the likelihood ratio index (LRI) test is applied to compare non-nested models by evaluating the probability of the difference between corrected likelihood ratio indices being significant (Horowitz 1983).

Value-of-time (VOT) is measured for RUM-based and RRM-based models according to the formulations (see, e.g., Chorus 2012a, b):

$$VOT_{RUM} = \frac{(\partial U_i/\partial time_i)}{(\partial U_i/\partial \mathrm{cost}_i)} \tag{20}$$

$$VOT_{RRM} = \frac{(\partial R_i/\partial time_i)}{(\partial R_i/\partial \mathrm{cost}_i)} \tag{21}$$

where U_i is the utility of route i, R_i is the regret of route i, and $time_i$ and $cost_i$ are respectively, time and cost of route i. Although there is a difference between RUM and RRM formulation of VOT since the latter is alternative and choice set specific (Chorus 2012a, b), in the current study the calculation of the VOT according to the presented formulations is performed for the same choice set and hence provides additional elements to the comparison between RUM-based and RRM-based models.

Estimation results

The estimated models consider the following attributes m in the utility and the regret functions: (i) cost of travel calculated as proportional to the distance covered; (ii) travel time; (iii) number of right turns; (iv) number of left turns; (v) number of traffic lights crossed. The parameters β_m of these variables are expected of negative sign to indicate that utility is maximized or regret is minimized with cheaper, faster and straighter routes.

Table 2 presents the estimates for the RUM-based route choice models. The PCL model outperforms the RUM-MNL and its modifications ($LR_{PCL-RUM\text{-}MNL} = 11.25$, $p = 0.0008$; $LRI = \rho^2_{PCL} - \rho^2_{C\text{-}Logit\ \ eq(4)} = 0.0024$, $p < 0.0001$; $LRI = \rho^2_{PCL} - \rho^2_{PS\text{-}Logit} = 0.0032$, $p <$ 0.0001; $LRI = \rho^2_{PCL} - \rho^2_{PSC\text{-}Logit} = 0.0010$, $p = 0.0019$). Within RUM-MNL modifications, the PSC-Logit outperforms the remaining models using alternative expressions of the

Table 2 Estimates of RUM-based route choice models

Model	RUM-MNL		C-Logit eq(4)		C-Logit eq(5)		PS-Logit		PSC-Logit		PCL	
Cost of travel	−1.4293	−3.36	−1.4990	−3.54	−1.3665	−3.11	−1.6932	−3.97	−1.6766	−3.93	−1.3349	−3.44
Travel time	−0.3591	−7.09	−0.3764	−7.29	−0.3631	−7.20	−0.3777	−7.41	−0.3804	−7.42	−0.3377	−7.22
Right turns	−0.0912	−1.85	−0.0965	−1.97	−0.0975	−1.98	−0.0984	−1.98	−0.0992	−2.01	−0.0799	−1.75
Left turns	−0.6874	−8.77	−0.6090	−7.83	−0.6223	−7.96	−0.4221	−5.40	−0.5348	−6.87	−0.6156	−8.31
Traffic lights	−0.5500	−3.45	−0.4880	−3.08	−0.4810	−3.03	−0.2940	−1.84	−0.3650	−2.30	−0.5160	−3.66
Correction factor			−0.6168	−2.27	−0.4638	−1.78	0.6170	3.63	0.6183	3.29		
Gamma											0.9939	4.99
Null log-likelihood	−1298.38		−1298.38		−1298.38		−1298.38		−1298.38		−1298.38	
Final log-likelihood	−1050.29		−1047.81		−1049.31		−1048.78		−1046.03		−1044.67	
ρ^2	0.1911		0.1930		0.1918		0.1922		0.1944		0.1954	
Adjusted ρ^2	0.1872		0.1884		0.1872		0.1876		0.1897		0.1908	
Value-of-time (€/h)	15.07		15.07		15.94		13.38		13.61		15.18	

correction factor ($LRI = \rho^2_{PSC\text{-}Logit} - \rho^2_{C\text{-}Logit\ \ eq(4)} = 0.0014$, $p < 0.0001$; $LRI = \rho^2_{PSC\text{-}Logit} - \rho^2_{PS\text{-}Logit} = 0.0014$, $p < 0.0001$). All parameters β_m are negative as expected to represent the utility maximization behavior of morning commuters in the urban network of Torino (Italy). All parameters β_{corr} exhibit the expected signs, with estimates for commonality factors being negative and for path size measures being positive to reduce the utility on the basis of the degree of similarity of the routes in the choice set. When looking at the VOT, the PCL model shows the lowest value at 14.19 €/h, while the RUM-MNL modifications vary from 15.07 €/h of the C-Logit with original commonality factor to 15.94 €/h of the C-Logit with modified commonality factor. These values appear slightly high when considering the commuting purpose (see Shires and De Jong 2009), but over 50 % of the sample are faculty members with higher income and potentially higher value-of-time.

Table 3 shows the estimates for the RRM-based route choice models with utility-based correction terms. As observed in other choice contexts (see, e.g., Chorus 2012a, b), the RRM-MNL model has goodness-of-fit similar to the RUM-MNL model ($LRI = \rho^2_{RRM\text{-}MNL} - \rho_{RUM\text{-}MNL}{}^2 = 0.0002$, $p = 0.3036$). Adding a utility-based correction term leads to some models with performances inferior to the corresponding RUM-MNL modifications ($LRI = \rho^2_{C\text{-}Logit\ \ eq(4)} - \rho^2_{RRM\text{-}MNL+CF\ \ eq(4)} = 0.0015$, $p < 0.0001$; $LRI = \rho^2_{PSC\text{-}Logit} - \rho_{RRM\text{-}MNL+PSC} = 0.0021$, $p < 0.0001$) and others with performances superior to the corresponding RUM-MNL modifications ($LRI = \rho^2_{RRM\text{-}MNL+ln(PS)} - \rho^2_{PS\text{-}Logit} = 0.0021$, $p < 0.0001$). Interestingly, the RRM-MNL model is slightly outperformed by the PCL model ($LRI = \rho^2_{PCL} - \rho_{RRM\text{-}MNL}{}^2 = 0.0034$, $p < 0.0001$). All parameters β_m are negative as anticipated to characterize the regret minimization behavior of morning commuters in the studied urban network. All parameters β_{corr} exhibit the expected signs as their RUM-based counterparts. When looking at the VOT, the values are generally lower than the RUM-based models with values ranging between 11.76 €/h for the RRM-MNL + PSC model and 13.25 €/h of the RRM-MNL + C-Logit with original commonality factor. These values appear more in line with commuters value-of-time in Europe (see Shires and De Jong 2009), and it should be noted that they depend on the relative performance of the chosen routes with respect to the alternative routes in the choice set (Chorus 2012b).

Table 4 illustrates the estimates for the RRM-based route choice models with regret-based correction terms. Interestingly, adding a regret-based correction term appears more beneficial than summing a utility-based correction term from the perspective of the goodness-of-fit of the route choice models ($LRI = \rho^2_{C\text{-}RRM\ \ eq(11)} - \rho^2_{RRM\text{-}MNL+CF\ \ eq(4)} = 0.0037$, $p < 0.0001$; $LRI = \rho^2_{C\text{-}RRM\ \ eq(12)} - \rho^2_{RRM\text{-}MNL+CF\ \ eq(5)} = 0.0039$, $p < 0.0001$; $LRI = \rho^2_{PSC\text{-}RRM} - \rho^2_{RRM\text{-}MNL+PSC} = 0.0015$, $p < 0.0001$), with the exception of the original path size measure ($LRI = \rho^2_{RRM\text{-}MNL+ln(PS)} - \rho^2_{PS\text{-}RRM} = 0.0007$, $p < 0.0269$). When comparing the best performing RUM-based and RRM-based models, the C-RRM model performs similarly to the PCL model ($LRI = \rho^2_{C\text{-}RRM\ \ eq(12)} - \rho^2_{PCL} = 0.0002$, $p < 0.2700$), which is remarkable when considering that the latter takes 7 times more to converge. All parameters β_m are negative as expected, and all parameters β_{corr} are positive as expected to indicate that the correction terms capture similarity as hypothesized in "Correcting for similarity across alternatives" section. When looking at the VOT, the values are similar to the other RRM-based models with variation from 10.53 €/h for the PSC-RRM model to 12.37 €/h for the C-RRM with weighted correlation.

Overall, the proposed RRM-based models are comparable to the RUM-based models, and the C-RRM model is comparable with the more sophisticated, but far more demanding computationally, PCL model.

Table 3 Estimates of RRM-based route choice models with utility-based correction factors

Model	RRM-MNL		RRM-MNL + CF eq(4)		RRM-MNL + CF eq(5)		RRM-MNL + ln(PS)		RRM-MNL + PSC	
Cost of travel	−0.2656	−3.69	−0.2706	−3.79	−0.2540	−3.43	−0.2855	−4.00	−0.2922	−4.07
Travel time	−0.0584	−6.73	−0.0604	−6.93	−0.0587	−6.85	−0.0587	−6.82	−0.0599	−6.91
Right turns	−0.0222	−2.25	−0.0221	−2.26	−0.0213	−2.17	−0.0226	−2.32	−0.0225	−2.30
Left turns	−0.1434	−9.14	−0.1464	−9.33	−0.1506	−9.60	−0.1392	−8.95	−0.1434	−9.14
Traffic lights	−0.1110	−3.73	−0.1030	−3.49	−0.0914	−3.10	−0.0614	−2.05	−0.0831	−2.79
Correction factor			−0.5739	−2.17	−0.5684	−2.30	0.5172	3.19	0.5217	2.89
Null log-likelihood	−1298.38		−1298.38		−1298.38		−1298.38		−1298.38	
Final log-likelihood	−1050.05		−1049.76		−1049.48		−1047.02		−1048.76	
ρ^2	0.1913		0.1915		0.1917		0.1936		0.1923	
Adjusted ρ^2	0.1874		0.1869		0.1871		0.1890		0.1876	
Value-of-time (€/h)	12.62		12.79		13.25		11.80		11.76	

Table 4 Estimates of RRM-based route choice models with regret-based correction factors

Model	RRM-MNL		C-RRM eq(11)		C-RRM eq(12)		PS-RRM		PSC-RRM	
Cost of travel	−0.2656	−3.69	−0.2888	−4.04	−0.2680	−3.59	−0.2905	−4.08	−0.3254	−4.53
Travel time	−0.0584	−6.73	−0.0602	−6.98	−0.0577	−6.70	−0.0587	−6.83	−0.0596	−6.94
Right turns	−0.0222	−2.25	−0.0212	−2.17	−0.0217	−2.21	−0.0227	−2.33	−0.0219	−2.26
Left turns	−0.1434	−9.14	−0.1494	−9.52	−0.1500	−9.56	−0.1404	−9.02	−0.1440	−9.25
Traffic lights	−0.1110	−3.73	−0.0906	−3.10	−0.0986	−3.31	−0.0577	−1.92	−0.0645	−2.17
Correction factor			0.4048	3.46	0.4142	3.60	0.2561	3.82	0.1282	4.57
Null log-likelihood	−1298.38		−1298.38		−1298.38		−1298.38		−1298.38	
Final log-likelihood\	−1050.05		−1044.96		−1044.38		−1047.93		−1046.85	
ρ^2	0.1913		0.1952		0.1956		0.1929		0.1937	
Adjusted ρ^2	0.1874		0.1906		0.1910		0.1883		0.1891	
Value-of-time (€/h)	12.62		11.97		12.37		11.60		10.53	

Summary and conclusions

As the discrete choice paradigm of RRM has been recently introduced in several choice contexts (see, e.g., Chorus 2012a, b), and the suitability to represent travelers' route choice behavior appears logical, this paper looks into the RRM paradigm in the route choice context from three perspectives: (i) similarity across alternative routes; (ii) choice set composition effects on route choice probabilities; (iii) comparison of RUM-based and RRM-based models.

Three approaches are proposed to consider similarities across alternatives within the regret function: (i) adding utility-based corrections; (ii) adding a regret-based term that compares the degree of similarity of a route with the one of each alternative; (iii) adding a regret-based term that adjusts for the correlation of each route with any other alternative. The behavior of the proposed RRM-based route choice models when overlapping routes are considered is investigated with an experimental analysis of the "overlapping network" and the "switching route network" (see Prashker and Bekhor 2004). The analysis of the variation of choice probabilities as a function of the overlapping quantity illustrates that the first approach leading to a hybrid RUM-RRM model is preferable to represent overlap when considering the "overlapping network", but produces incorrect results when examining the "switching routes network". This behavior is analogous to the one exhibited by RUM-MNL modifications that account for the same correction terms (Prashker and Bekhor 2004). The analysis also shows that the C-RRM model introduces less error than PS-RRM and PSC-RRM, especially when the correlation is weighted according to the disjoint portions of the routes.

The behavior of the proposed RRM-based route choice models for different choice set compositions is investigated with an experimental analysis of choice probability differences in the "grid network" proposed by Bliemer and Bovy (2008). The proposed models are estimated and then choice probabilities are calculated from the application of estimated parameters to different choice sets and compared to the postulated true behavior. Firstly, model estimation while considering only relevant routes in the choice set allows obtaining parameters of the expected sign. On the contrary, model estimation while including also irrelevant routes in the choice set guides the parameters to show the opposite signs because the degree of similarity of relevant routes is very high with respect to irrelevant routes. Secondly, RRM-based route choice models are choice set specific and present the expected feature of being able to reproduce accurately true route choice probabilities only when estimation and prediction are performed for the same choice set composition. Last, calculated choice probabilities show that the C-RRM model is able to differentiate between relevant routes according to the correlation level proposed by the postulated probit model.

The comparison of the RRM-based models with the RUM-based models is investigated with model estimation from revealed preference data collected among commuters in an urban network. From the goodness-of-fit perspective, the RRM-MNL model is comparable to the RUM-MNL model as shown in the literature about contexts other than route choice (Chorus 2012a, b). In addition, models with regret-based correction terms that account for the pairwise correlation between alternatives are slightly better than models with utility-based corrections. In particular, the C-RRM model performs comparably to the PCL model. From the VOT perspective, RRM-based route choice models produce values generally lower than their RUM-based counterparts by roughly 3 €/h and approximately in line with existing knowledge about commuter trips (Shires and De Jong 2009).

In conclusion, this paper contributes to the literature in route choice by verifying that it is plausible to consider travelers as regret minimizers and proposing corrections for similarity across alternatives analogously to RUM-based models. This paper contributes also to the literature in RRM by further expanding the applicability of RRM-based models for another choice context, illustrating the similarity in the performances with respect to the RUM-based models, and showing the preference with respect to enhanced RUM-based models that are neither more parsimonious, nor easier to code, nor faster to estimate.

All three perspectives open the possibility for further research directions. From the similarity perspective, additional measures might be derived theoretically or at least hypothesized conceptually in order to correct the regret function. From the choice set composition perspective, a deeper investigation could be performed in order to understand to what extent RRM-based route choice models suffer from misspecification of choice sets. Moreover, dynamic approaches with RRM-based models instead of RUM-based models could be considered to solve the lack of robustness with respect to the choice set composition (see Fosgerau et al. 2011). From the comparative perspective, additional datasets can be used to expand the literature in RRM with respect to both model performances and value-of-time calculations, and synthetic experiments can be set up to evaluate which models better reproduce route choice behavior.

Acknowledgments The significant contribution of three anonymous reviewers, who provided knowledgeable and insightful comments that greatly contributed to the final version of the manuscript, is gratefully acknowledged. The financial support of the Danish Council for Strategic Research for the project "Analyses of activity-based travel chains and sustainable mobility" is appreciatively acknowledged.

References

Bekhor, S., Chorus, C.G., Toledo, T.: Stochastic user equilibrium formulation for random regret minimization-based route choice model. Transp. Res. Rec. **2284**, 100–108 (2012)

Bekhor, S., Toledo, T., Prashker, J.N.: Effects of choice set size and route choice models on path-based traffic assignment. Transportmetrica **4**(2), 117–133 (2008)

Ben-Akiva, M.E., Bierlaire, M.: Discrete choice methods and their applications to short term travel decisions. In: Hall, R.W. (ed.) Handbook of transportation science, pp. 5–34. Kluwer, Dordrecht (1999)

Bliemer, M.C.J., Bovy, P.H.L.: Impact of route choice set on route choice probabilities. Transp. Res. Rec. **2076**, 10–19 (2008)

Bliemer, M.C.J., Bovy, P.H.L., Li, H.: Some properties and implications of stochastically generated route choice sets. In: Proceedings of the 6th Tristan conference, Phuket (2007)

Bovy, P.H.L.: On modelling route choice sets in transportation networks: a synthesis. Transp. Rev. **29**(1), 43–68 (2009)

Bovy, P.H.L., Bekhor, S., Prato, C.G.: The factor of revised path size: an alternative derivation. Transp. Res. Rec. **2076**, 132–140 (2008)

Cascetta, E.: Transportation systems engineering: theory and methods. Kluwer Academic Publishers, Dordrecht (2001)

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: Lesort, J.B. (ed.) Proceedings of the thirteenth international symposium on transportation and traffic theory, pp. 697–711. Pergamon Press, Lyon (1996)

Chorus, C.G.: A new model of random regret minimization. Eur. J. Transp. Infrastruct. Res. **10**(2), 181–196 (2010)

Chorus, C.G.: Random regret minimization: an overview of model properties and empirical evidence. Transp. Rev. **32**(1), 75–92 (2012a)

Chorus, C.G.: Random regret-based discrete choice modeling: a tutorial. Springer, Heidelberg (2012b)

Chorus, C.G., de Jong, G.C.: Modelling experienced accessibility for utility-maximizers and regret-minimizers. J. Transp. Geogr. **19**, 1155–1162 (2011)

Chu, C.: A paired combinatorial logit model for travel demand analysis, pp. 295–309. Proceedings of the 5th world conference on transportation research, Ventura (1989)

Daganzo, C.F., Sheffi, Y.: On stochastic models of traffic assignment. Transp. Sci. **11**(3), 253–274 (1977)

Fosgerau, M., Frejinger, E., Karlström, A.: A dynamic discrete choice approach for consistent route choice model estimation. Proceedings of the Swiss transport research conference, Ascona (2011)

Frejinger, E., Bierlaire, M., Ben-Akiva, M.E.: Sampling of alternatives for route choice modeling. Transp. Res. B **43**(10), 984–994 (2009)

Horowitz, J.L.: Statistical comparison of non-nested probabilistic discrete choice models. Transp. Sci. **17**(3), 319–350 (1983)

Koppelman, F.S., Wen, C.H.: The paired combinatorial logit model: properties, estimation and application. Transp. Res. B **34**(2), 75–89 (2000)

Nerella, S., Bhat, C.R.: Numerical analysis of effect of sampling of alternatives in discrete choice models. Transp. Res. Rec. **1894**, 11–19 (2004)

Prashker, J.N., Bekhor, S.: Investigation of stochastic network loading procedures. Transp. Res. Rec. **1645**, 94–102 (1998)

Prashker, J.N., Bekhor, S.: Route choice models used in the stochastic user equilibrium problem: a review. Transp. Rev. **24**(4), 437–463 (2004)

Prato, C.G.: Route choice modeling: past, present and future research directions. J Choice Model **2**(1), 65–100 (2009)

Prato, C.G., Bekhor, S.: Applying branch & bound technique to route choice set generation. Transp. Res. Rec. **1985**, 19–28 (2006)

Prato, C.G., Bekhor, S.: Modeling route choice behavior: how relevant is the choice set composition? Transp. Res. Rec. **2003**, 64–73 (2007)

Prato, C.G., Bekhor, S., Pronello, C.: Latent variables and route choice behavior. Transportation **39**(2), 299–319 (2012)

Ramming, S.: Network knowledge and route choice. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge (2002)

Shires, J.D., de Jong, G.C.: An international meta-analysis of values of travel time savings. Eval. Program Plan. **32**(4), 315–325 (2009)

Author Biography

Carlo Giacomo Prato is Associate Professor at the Department of Transport of the Technical University of Denmark. His main research interests are transport models with particular focus on route choice behavior, travel behavior and road safety analysis.