LIMITATIONS OF RECURSIVE LOGIT FOR INVERSE REINFORCEMENT LEARNING OF BICYCLE ROUTE CHOICE BEHAVIOR IN AMSTERDAM

Thomas Koch
Stochastics Research Group
Centrum Wiskunde en Informatica, Amsterdam, The Netherlands 1098 XG
Email: koch@cwi.nl
ORCID: 0000-0003-2858-7010

Dr. Elenna Dugundji
Department of Mathematics, Faculty of Sciences
Vrije Universiteit Amsterdam, Amsterdam, The Netherlands, 1081 HV
Centrum Wiskunde en Informatica, Amsterdam, The Netherlands, 1098 XG
Email: e.r.dugundji@cwi.nl
ORCID: 0000-0001-9277-9485

Word Count: 6686 words + 3 table(s) × 250 = 7436 words

Submission Date: July 31, 2020

ABSTRACT

Used for route choice modelling by the transportation research community, recursive logit is a form of inverse reinforcement learning. By solving a large-scale system of linear equations recursive logit allows estimation of an optimal (negative) reward function in a computationally efficient way that performs for large networks and a large number of observations. In this paper we review examples of recursive logit and inverse reinforcement learning models applied to real world GPS travel trajectories and explore some of the challenges in modeling bicycle route choice in the city of Amsterdam using recursive logit as compared to a simple baseline multinomial logit model with environmental variables.

We discuss conceptual, computational, numerical and statistical issues that we encountered and conclude with recommendation for further research.

Keywords: recursive logit, Markov decision process, inverse reinforcement learning, GPS trajectory, bicycle route choice

INTRODUCTION

Bicycling in Amsterdam is serious business: one third of the daily movements in Amsterdam by residents and visitors is done by bike and almost half of the commute trips between work and home is done on a bike. Furthermore bicycling has seen a steady growth, not just in Amsterdam but in all (major) cities in the Netherlands. This leads to increasing congestion issues for bicyclists especially at intersections with traffic lights. To address this, policy changes are required such as new infrastructure, changes in traffic signal policies, etc. To create effective policies, policy makers need tools to gain insights in the behaviour of bicyclists, to model when and where they will bicycle.

In our previous research in Koch et al. (1) found that bicyclists in Amsterdam rarely take the shortest path and make more detours than cars. It is thus less straightforward to predict routes taken by the average bicyclist route and consequently estimate the number of bicyclists on a given street. However with the advent of new data collections techniques such as GPS on smartphones, more and more data is available on the revealed preference of bicyclists, allowing researchers to develop new ways to model influences on bicycle route choice. In this study we will to attempt to quantify the influence of environmental and spatial planning features on route choice decisions of everyday bicyclists in Amsterdam by estimating a model that can predict where and how these cyclists traverse in and around Amsterdam.

We begin this paper with a background reviewing literature relevant to bicycle route choice, by discrete choice modeling with generation of alternative routes. Next we review methods to model route choice without generating choice sets: recursive logit and inverse reinforcement learning. The case study describes the data set of observed bicycle routes. We also defined a number of environmental variables that potentially could have an effect on bicycle route choice. First we estimate a baseline multinomial logit model with the environmental variables, based on the observed routes compared with up to 16 generated route alternatives. Subsequently we describe our efforts to estimate a recursive logit model with the observed bicycle routes in our data, environmental explanatory variables, and the Amsterdam street network – without generating route choice alternatives. We discuss conceptual, computational, numerical and statistical issues that we encountered. We conclude the paper with a reflection on our experience of modeling route choice of bicyclist in Amsterdam and recommendations for further research.

BACKGROUND

Discrete choice modeling of travel routes

Since the 1970's discrete choice modeling has been a leading method to understand choice behaviour of individuals in a wide range fields such as marketing, economics and transportation. Described by McFadden et al. (2) in 1973, discrete choice modeling has subsequently been extended over the decades in order to overcome specific limitations such as overlapping alternatives and correlations over time.

The study of the specific field of route choice, is more complicated than a choice between easily enumerable distinct alternatives, as route choice is typically a sequence of choices at each intersection, each transit stop, each mode, etc. This leads to very large choice set that is theoretically infinite due to loops. Often there can also be a large overlap between different route alternatives leading to difficulties for choice modeling. We will highlight two commonly used approaches:

An approach established in the 1990s to model route choice using a collection of observed paths and for each observations a set of generated paths by a route choice generator. This approach has been used to estimate models such as multi-nominal logit (MNL) and mixed logit. This approach comes with limitations: as discussed in Koch et al. (1), these route choice generators do not necessarily create realistic routes; and Frejinger et al. (3) argues that parameter estimates can vary significantly depending on the bias of the route choice generator. To address the issues with the overlap between difference alternative paths and the resulting correlations, multiple extensions have been proposed to attempt to avoid erroneous path probabilities and substitution patterns. The most two popular are path size logit Ben-Akiva and Bierlaire (4) and C-Logit Cascetta et al. (5), which decrease the utility of overlapping paths proportional to the overlap with other paths included in the choice set.

A second approach is to achieve a consistent choice set by sampling as proposed by Frejinger et al. (3). This approach attempts to set up a sampling protocol in order to obtain unbiased parameter estimates from the route choice sets to neutralize the bias introduced by the route choice generator.

Bicycle route choice

In 2010, Menghini et al. (*6*) published a seminal route choice model for bicyclists estimated from a large sample of GPS observations in a revealed preference study with 2435 persons logging 73,493 trips in Zurich, Switzerland. Using this data they estimated a multinominal logit model and used breadth-first search link elimination (BFS-LE) to generate choice alternatives for each observed trip. They included six different variables in the choice model: length of the route, average absolute gradient change, maximum gradient change, percentage of marked bicycle paths along the route, number of traffic lights and the path size measure. Accounting for the similarity between alternatives with the path size vector, their model showed that the elasticity with respect to trip length was nearly four times larger than that with respect to the percentage of bicycle paths along the route. The only other explanatory variable that had an impact albeit small, was the product of length and the maximum gradient along the route.

In 2017, Ton et al. (*7*) reported on a route choice model for bicyclists using the same dataset as in the study in this paper. Ton et al. consider the construction of choice sets via an empirical approach, using only the observed trips in the data set to compose the choice alternatives. On basis of their specific focus (inner city travel in Amsterdam) Ton et al. selected 6 variables: distance, percentage of separate cycle paths, number of intersections, rain, sunset and sunrise times and trip purpose. Their findings suggest that bicyclists in Amsterdam are insensitive to dedicated cycle paths, attributed to an inner city characterized by a dense road network where cycling is the most prominent mode of transport. Additionally they found that cyclists in Amsterdam were found to minimize travel distance and the number of intersections per kilometer. Furthermore they found that for early morning trips there was a stronger impact of distance on route choice than outside these hours. In a subsequent paper Ton et al. (*8*) looked at a data-driven path identification approach, combining all unique routes observed for one origin-destination pair into a choice set and comparing this approach with two commonly used choice set generation methods (breadth-first search on link elimination and labelling).

In prior research with the same specific sub-selection of the data, we found in Koch et al. (*1*) that bicyclists in Amsterdam often deviate from the shortest path, more than car drivers, indicating that there are different and possibly also more factors that have an effect on the routes bicyclists in Amsterdam take. In Koch et al. (*1*) we focused on the concept of route complexity: counting the number of locations where people deviate from the shortest path, in the interest of improving route choice generation techniques and potentially get more insight into the motivations for the route choice for bicyclists. In this study we explore other effects on route choice using different methodologies, without looking at route complexity or where people deviate from the shortest path. In future research we intend to combine both streams of work.

Zimmermann et al. (9) showed that is possible to estimate bicycle route choice without the restrictiveness of pre-generated route choice sets and model route choice as a sequence of choices via recursive logit. For comparison with the recursive logit model, we will estimate a simple baseline multinomial logit model using a synthetically generated choice set. The synthetic approach allows us to generate additional plausible route alternatives outside the set of observed routes. This means that we can include all observations even between origin destinations pairs with only a single observations, unlike the study by Ton et al. (*7*) that is limited to trips traversing the inner city of Amsterdam, due to a insufficient density of trips in the suburbs for this empirical approach to work there.

METHODS

In this section we will review two variants of performing choice modeling without choice sets: an analytical solutions via recursive logit and a computational approximation via inverse reinforcement learning.

Dynamic discrete choice modeling of travel link sequences

An alternative approach uses link-based Markov decision process to model route choice as a series of sequential decisions. First proposed by Fosgerau et al. (10) it uses a linear system of equations to efficiently compute choice probabilities by using a solver to solve Bellman equations.

An incidence matrix is established that defines the exponential utility to perform action a from state k:

$$M_{ka} = \left\{ \begin{matrix} \delta\left(a|k\right) e^{\frac{1}{\mu}v(a|k)}, & a \in A\left(k\right) \\ 0 & \text{otherwise.} \end{matrix} \right\} \tag{1}$$

The size of the incidence matrix is given by $|\widetilde{A}|$ describing the number of states A and the number of dummy links d representing termination states of destination. As the dummy links d have no successors, the row $k = d$ will be zero. Secondly Fosgerau et al. (10) define a vector z of size $|\widetilde{A}x1|$ vector where $z_k = e^{\frac{1}{\mu}V(K)}$ and a vector b of size $|\widetilde{A}x1|$ where $b_k = 0, k \neq d$ and $b_d = 1$. Now given the identity matrix I, Fosgerau et al. (10) write the linear equation:

$$z = Mz + b \iff (I - M)z = b \tag{2}$$

This system has a solution if $I - M$ is invertible, which might not be the case. As Fosgerau et al. (10) note this is highly dependent on the balance between the number of paths that connect the nodes in the network and the size of instantaneous utilities $\frac{1}{\mu}v(a|k)$. They note that this issue is particularly important to consider when estimating a model, as depending on the value of β, $I - M$ can be ill-conditioned or even singular. Fosgerau et al. (10) note that this limits the possible values of parameters, as when equation 2 does not yield a valid solution for at least one observation, the log likelihood function is not defined. They suggest to deal with this issue by starting at a feasible point (meaning a large enough magnitude in the parameters) and then being conservative in the initial step size of the line search algorithm at the price of an increased number of iterations.

Mai et al. (11) proposed a nested recursive logit that relaxes the independence from irrelevant alternatives property of the logit model by allowing scale parameters to be link specific. Zimmermann et al. (9) subsequently look at bicycle route choice problem in the city of Eugene, Oregon. By using 648 observations of bike trips collected from 103 users. They test a long list of 14 potential parameters: length; link constant to penalize paths with many constants; length interacted separately with upslope, medium traffic, heavy traffic, regional multi-use path, bicycle boulevard, bike lane; bridge; bridge interacted with bike facilities; no turn; no turn interacted with crossroad; left turn interacted with crossroad separately for medium traffic and for heavy traffic.

In Mai et al. (12) an improvement is proposed to Fosgerau et al. (10) by reducing the numbers of linear systems that need to be solved. By adding all observed destinations in vector b of size $|\widetilde{A}x|D||$ it becomes possible to solve the problem one iteration instead of solving the system for each destination separately, allowing for 30 times performance gain in their example. They use this performance gain to propose a mixed recursive logit, which allows for random taste variation by adding a random value to the utility function and running the model n draws each iteration to allow for a random variation. They perform a case study in two cities. First a car route choice model in the Swedish city of Borlänge, with 466 destinations, 1832 observations and a bicycle route choice model in Eugene, Oregon with 286 destinations with a unknown number of observations.

In de Freitas et al. (13), recursive logit is used to model inter-modal travel based on a static network that describes various connections in Zurich, Switzerland. The street network consists of 30,372 links and 13,828 nodes and the transit network consists of 10,298 transit links and 1585 nodes.

Inverse reinforcement learning on real world travel trajectories

From the field of computer science but similar to Recursive logit, inverse reinforcement learning (IRL) aims to find reward function parameters θ by observing the behaviour of each agent in a Markov decision process (MDP) with a finite set S of N states. The reward function $R(\zeta)$ for trajectory $\zeta = \{s, a\}$, performing action a at state s with f_s the feature vector of state s, is given by:

$$R(\zeta) = \theta^T \mathbf{f}_\zeta = \sum_{s \in \zeta} \theta^T \mathbf{f}_s \tag{3}$$

In the computer science literature there are several studies performing IRL on real world problems. Ziebart et al. (14) introduced Maximum entropy inverse reinforcement learning in 2008 based on the principle of maximum entropy by Jaynes (15) that the probability of a trajectory ζ with higher reward is exponentially higher than that of a smaller reward: $P(\zeta) \propto e^{R(\zeta)}$. In order to learn from observed behaviour, the maximum entropy IRL algorithm maximizes the likelihood of the observed trajectories under the maximum entropy distribution T

$$\theta^* = \underset{\theta}{\operatorname{argmax}} \sum_{\zeta} \log P(\widetilde{\zeta}|\theta, T) \tag{4}$$

The maximum entropy distribution T is derived using

$$P(\zeta|\theta) = \frac{1}{Z(\theta)} e^{\sum s_j \in \zeta_i^{\theta \top \mathrm{f}_{s_j}}} \tag{5}$$

For parameters θ the partition function $Z(\theta)$ will always converge for the problem with finite horizons and infinite horizon problems with discounted reward weights. Since function 4 is convex for a deterministic MDP, gradients for optimizers can be obtained by taking the difference between the observed feature counts and the expected feature counts based on a given set of parameters θ, that can be formulated as the expected state visitation frequencies D_{s_i}. To compute the gradients Ziebart et al. (14) uses:

$$L(\theta) = \widetilde{\mathrm{f}} - \sum_{\zeta} P(\zeta|\theta,T) f_{\zeta} = \widetilde{\mathrm{f}} - \sum_{s_i} D_{s_i} \mathrm{f}_{s_i} \tag{6}$$

To efficiently compute the expected state frequencies for parameters θ, Ziebart et al. (14) has proposed an algorithm that approximates the state frequencies by recursively backing up from each possible terminal state, computing each probability mass of each branch along the way, computing partition function Z at each action and state. The branching values give the local action probability that can be used to compute state frequencies and summed up for total frequency counts. Ziebart et al. (14) apply the maximum entropy IRL model to learn the reward function of taxi drivers on the road network of Pittsburgh, Pennsylvania. To do so, GPS logging of approximately 7403 trajectories are used to determine the cost of different road type, speed, number of lanes and turn costs. The MDP modeled from the road-network of Pittsburg is assumed to be deterministic with over 300,000 states (street segments) and 900,000 actions (transitions at intersections).

In Nguyen et al. (16) a generalization of the IRL problem is proposed that allows multiple locally consistent reward functions to generate the trajectories. By representing the IRL problem with a probabilistic graph model, an expectation-maximization (EM) algorithm can be devised to iteratively learn different reward functions and the stochastic transitions between them, in order to improve the likelihood of the observed trajectories. As a result, the EM algorithm can be used to derive locally consistent reward functions. Nguyen et al. (16) empirically evaluated their algorithm with a small real world network and GPS data of 59 taxis in Singapore. In this evaluation the road network is modelled as a simplified grid world with 193 states.

Mai et al. (17) proposes a generalized version of the causal entropy maximization problem, allowing the possibility to generate a class of maximum entropy IRL models. Their proposed generalized model has the advantage of being able to recover an expert function that would (partially) capture the impact of the connecting structure of the states on experts' decision. Their empirical evaluation on a real-world dataset and a grid-world dataset shows that their generalized model outperforms classical approaches in terms of recovering reward functions and demonstrated trajectories.

Mai et al. (18) proposes a tractable approach to compute directly a log-likelihood of observed trajectories with incomplete/missing data. By performing the training by solving a sequence of linear equations that does not depend on the number of missing segments it is efficient at handling a large number of missing segments. Their empirical evaluation showed that their approach outperforms other approaches.

Mo (19) looks at bicycle route choice applying the maximum entropy IRL approach. To achieve multi-reward functions an extension is used known as Behaviour Clustering IRL (BCIRL). He performs multiple experiments to investigate the applicability of these methods in the context of bicycle route choice. In this study it was found that a low number of demonstrated trajectories, short trajectory lengths, large number of Markov decision processes to be solved, and class imbalance were problematic issues for the methods. An application was performed on a dataset of GPS trajectories in Amsterdam, but no factors other than distance were found to be relevant.

CASE STUDY

Collecting data on bicycle movements

For this study we used the 2016 FietsTelweek ("Bicycle Counting Week") data set (Bikeprint (*20*)) that is available at their website. During the week of the 19th of September 2016 approximately 29,600 bicyclists volunteered to track their bicycle movements using a smartphone app. For this case study we limited the study to bicycle trips to and/or from the city of Amsterdam, Diemen, Amstelveen and Ouder-Amstel, leaving around 29,684 trips. In figure 1 we visualized all observed trajectories.

This app ran in the background collecting all movements by the bicyclists using the phone's GPS and acceleration sensors. The cyclists used their bike in a way as often seen in the Netherlands, using their bike as transportation from and to work, supermarket, school, etc. For privacy reasons the resulting data was anonymized by the data provider before making it publicly available (i) by the removal of user information to make it impossible to trace multiple trips to a single person and (ii) by rounding of the trip departure time into one-hour bins to the nearest hour and (iii) removal of the random number between 0 and 400 meters from the start and the end of the trip to obfuscate the true origin and destination of each trip.

In prior research based on this data we found in Koch et al. (1) that bicyclists in Amsterdam often deviate from the shortest path, more than car drivers, indicating that there are different factors at play in the route choice of bicyclists in Amsterdam. In Koch et al. (1) we focused on the concept of route complexity: counting the number of locations where people deviate from the shortest path, in the interest of improving route choice generation techniques and potentially get more insight into the motivations for the route choice for bicyclists.

Generation of alternatives

To find out what kind of alternatives exist for each observed path we applied synthetic route choice generation using the Double Stochastic Generation Function (DSGF) method described by Nielsen (*21*). The DSGF approach produces heterogeneous routes because both the cost and parameters used in the cost function for the links are drawn from a probability function. This way it can generate random paths, just by calculating the shortest path since the cost of each route is based on random factors. Halldórsdóttir et al. (*22*) showed that DSGF has a high coverage level of replicating routes taken by bicyclists and that it performs well up to 10 kilometer. Furthermore Bovy and Fiorenzo-Catalano (*23*) state that the method guarantees, with high probability, that attractive routes are included in the choice set, while unattractive routes are left out.

We used an existing implementation of DSGF, specifically POSDAP by ETH-Zurich (*24*) working on a street network provided by the data collection team of the Fietstelweek, that they imported from OpenStreetMap. We slightly modified POSDAP to execute at most a given number of $M = 128$ iterations (instead of running for a given duration) so that it behaves identically on different machines. For some origin destination pairs POSDAP was not able to find as many as N_0 routes in M iterations, in which case we will use all found routes. The choice sets are written to CSV files for further processing.

Additionally we also run an implementation of Breadth First Search Link Elimination (BFS-LE) by Rieser-Schüssler et al. (*25*), but we opted to leave out these alternatives since there was not much variance in the route choice set. In (*1*) we published on the coverage and consistency of both synthetic choice sets.

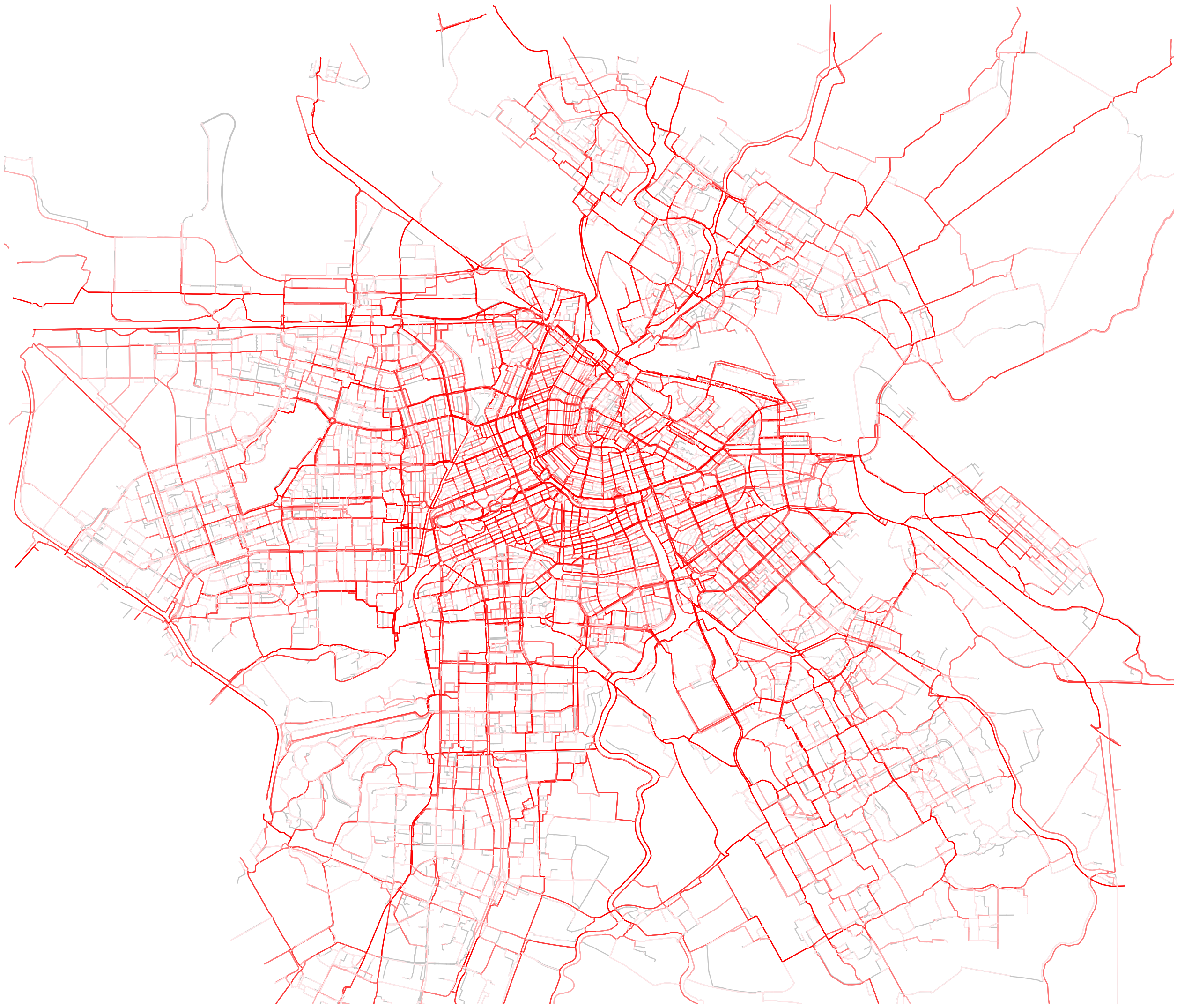


FIGURE 1: Map visualizing the trajectories observed in Amsterdam in the case study

Environmental variables

To collect a set of variables that would reasonably impact route choice of bicyclists we collected and processed open data sources to compute various explanatory variables describing each route. The procedure for the generation of the variables is described below. Descriptive statistics of the variables in the Amsterdam region and in the city center are given in Table 1.

First of all for each link in the network we include the length of that link as ***distance*** and if that link is a dedicated cycle-way, we include the length as ***oncycleway***. Additionally we have a variable ***traveltime*** based on the length and an estimated speed based on the GPS observations.

To include data about the environment of each link we extracted information of data made openly available by the city of Amsterdam. Firstly we pulled potentially relevant variables from a geographical data-set with land-use zones. To combine the street-network with other relevant geographical data-sets, we cut each street link into small segments of 5 meters and determined the distance of that segments to a geographical feature in the land use data-set. The variable ***nearwater*** measures the distance of street situated close to water bodies such as the canals of Amsterdam,

TABLE 1: Descriptive statistics for variables in the region of Amsterdam and in the smaller selection in the city center of Amsterdam

Amsterdam region	min	max	median	mean	std-dev	kurtosis
travel time	0.001	54.038	0.555	0.978	1.52	159.464
length	0	3.257	0.034	0.059	0.086	145.122
tree covered length	0	0.505	0	0.008	0.021	52.74
cycleway length	0	2.96	0	0.024	0.066	182.591
55 db exposure length	0	1.889	0	0.014	0.041	182.787
70 db exposure length	0	1.025	0	0.006	0.026	177.314
Near water length	0	2.96	0	0.009	0.046	718.247
Near green length	0	2.95	0	0.009	0.050	531.531
Amsterdam city center	**min**	**max**	**median**	**mean**	**std-dev**	**kurtosis**
travel time	0.010	21.361	0.503	0.861	0.993	25.905
length	0.001	0.517	0.028	0.050	0.055	5.707
tree covered length	0	0.154	0	0.006	0.016	15.043
cycleway length	0	0.42	0	0.011	0.032	28.507
55 db exposure length	0	0.467	0	0.014	0.033	22.767
70 db exposure length	0	0.346	0	0.008	0.023	34.888
Near water length	0	0.517	0	0.012	0.035	30.569
Near green length	0	0.258	0	0.002	0.012	129.199

(small) lakes, rivers and other water bodies wider than 6 meters. To determine a preference for routes through parks and forests we did the same thing with the variable ***neargreen***, measuring distance of street situated within a 25 meter radius of 'green' land used for parks, forests and meadows.

For a more fine-grained indication of the level of green and trees along a route, we used a data-set of the location of each individual tree in Amsterdam to determine what portion of each street segment is covered by trees. Our reasoning is that the number of trees has an influence on route choice as they can provide shade on hot days and function as a cover against the wind in storm conditions. To determine the variable ***neartree*** we measured the distance of street within 30 meters left or right from one or more tree(s). This way a street along a row of trees would have the full distance. We determined the distance of 30 meters between road and tree based on various situations where a rows of trees are situated along bicycle roads in Amsterdam.

To measure the effect of residential areas the variable ***nearresidential*** measures the distance of streets in residential areas. . The variable ***nearretail*** describes distance within areas purposed for *'Shops, malls and hotels-restaurants-pubs'*, *'Public offices and services'* and *'Cultural, social, medical, educational'*.

To see if the vicinity of busy roads, a major source of noise and pollution, has any impact on route choice we used a data set with the noise contours map of road traffic in Amsterdam as shown in Figure 2. This data-set is produced by a model that estimates the level of exposure to traffic noise in this map there are four noise levels with respectively at least 55, 60, 65 or 70 decibels of noise. The variables ***nearXdb*** represent the distance of the street passing through these exposure zones.

FIGURE 2: Noise contour map of Amsterdam, used for the variable that indicates the distance of a trajectory along roads with noisy traffic

Based on the idea that tramlines in Amsterdam form a radial artery towards the heart of the city, we construct the variable ***neartram*** indicating the portion of the route that is situated 100 meter from tram rails either to the left or right of the path, measured using segments of 10 meters.

Finally we wanted to see if the number and frequency of traffic signals has a measurable effect on route choice. We included this in two ways: first the exact number of traffic signals with ***ntrafficsignal*** and secondly the frequency of traffic signals ***trafficsignalfreq*** where the number of signals is divided by the length of the route.

Since Amsterdam has no elevation changes beyond the occasional bridge, we did not include any elevation changes as a variable.

BASELINE MULTINOMIAL LOGIT MODEL

Based on MNL models with a single variable per model, we removed some variables with high correlation between them. Firstly we decided to include only the highest and lowest level of traffic noise exposure. While the four variables were significant on their own, there was not much difference for the estimated coefficient values between 60 and 70 decibels. Secondly we only included the absolute number of traffic signals as the frequency of traffic signals per kilometer had a lower t-test score. The choice model was estimated both using PandasBiogeme(*26*) the estimation results are reported in Table 2.

TABLE 2: Estimated parameters of multinomial logit model

Utility parameter	Estimated value	t-Test
Length in meters	0.00186	62.1
Number of traffic signals	-0.0496	-9.84
Fraction of distance with exposure to traffic noise = 55db	-0.324	-3.32
Fraction of distance with exposure to traffic noise = 70db	-2.09	-26.3
Fraction of distance along green landuse	2.47	28.9
Fraction of distance along residential landuse	-0.372	-5.51
Fraction of distance along retail landuse	0.639	6.9
Fraction of distance along water	2.11	36
Fraction of distance along trees	1.46	11.5
Fraction of distance along tram lines	1.3	19.2
Fraction of distance on dedicated cycleways	4.3	78
Log likelihood	-79817.19	

The results of the choice model in Table 2 are what we expected to see and saw before in other studies. Bicyclists are prepared to travel longer to travel over nicer and safer routes. In the model we see a very significant effects of dedicated cycle-way infrastructure and the number of parks, meadows and forests along the route. The effect of just trees along the route is less significant yet still positive. We also see the result we expected for noise exposure: a significant avoidance of very heavy traffic noise exposure of 70 decibels or higher and smaller yet still significant effect from noise exposure to lower noise exposure of 55 decibels or higher. The effect we expected of attraction for routes along tram lines possibly being a landmark for navigation could be true based on the positive beta. For the effects of land-use along the route we see a negative value along routes with more residential land-use, likely because it is easier to navigate around residential neighbourhoods than go through them. Bicyclists also seem to prefer routes along areas with retail land use.

RECURSIVE LOGIT MODEL EXPERIMENTS

In this section we describe a series of experiments in modeling bicycle route choice in the city of Amsterdam using recursive logit model with the observed bicycle routes in our data, environmental explanatory variables, and the Amsterdam street network – without generating route choice alternatives.

Recursive logit with environmental variables

Our initial attempt was to model the Amsterdam network with each intersection as a node and the streets as actions, following example in Zimmermann and Frejinger (27). This resulted in a network with approximately 46,000 links and 30,000 observations, which we carefully controlled for full connectivity and no isolated graphs. Our motivation to model intersections as states instead of links as states was driven to lower the number of total states to be modeled, under the assumption that turn angles might have a low influence on bicycle route choice in Amsterdam. We tested the recursive logit model with the five variables *length*, *oncycleway*, *nearwater*, *neargreen* and *near55db*. However we were unable to get the solver to give plausible results for equation 1 as the solver would return incorrect results.

Python re-implementation of Recursive logit
For the purpose of a better understanding of the algorithm we also implemented our own version of the original recursive logit model in Fosgerau et al. (10) and the significantly faster decomposition recursive logit model Mai et al. (12) in Python with SciPy and NumPy. We were indeed able to successfully replicate models with data available online using our Python re-implementation. However also our Python re-implementation was not effective in giving plausible estimation results for the Amsterdam bicycle case.

Reduction of study area to Amsterdam city center only
Subsequently we simplified the study area to just the Amsterdam city center area containing only about 4500 links, excluding the entire municipality and surrounding suburbs. Again we carefully controlled for full connectivity and no isolated graphs. This too did not lead to plausible estimation results. We visualized this specific selection in Figure 3.

Reduction of street network complexity
Based on the remark by Fosgerau et al. (10) on dense networks and the number of alternative paths, our next action was to simplify the street network in the Amsterdam city center and remove all footpaths to reduce the complexity of the network. Again we carefully controlled for full connectivity and no isolated graphs. We accordingly also removed all observations of GPS trajectories cycling over footpaths. This too did not lead to plausible estimation results.

Edge-based network versus intersection-based network
Finally to transform our model to a model more similar to the studies in the literature, we instead created an edge-based network, instead of the intersection-based network. In the adapted implementation, each state is a street-segment and each action is a move to another street segment. This link-link approach allows the possibility to create new features with a boolean to indicate turns, left turns and u-turns, similar to the Borlänge model in Fosgerau et al. (10) and Mai et al. (12). For the entire city of Amsterdam this model contains 40063 links as states with 137724 transitions between states; for the city center area only it consists of 4204 links as states with 15234 transitions.

Edge-based network on entire city of Amsterdam: turn variables
With this network we were now able to solve the linear system to obtain a solution of z without (invalid) negative values for the entire city of Amsterdam. However even when setting the maximum number of links per observation at 30 links, we are still unable to calculate a log likelihood due to values of $z_{origin} == 0$ for one or more of the observations.

Edge-based network on Amsterdam city center only: turn variables
In the smaller area of the city center of Amsterdam it is possible to estimate a model but also with a relatively low limit of 30 links per observation, as a higher limit would again return zero values for z_{origin}. This meant we are able to process only 987 observations and 681 destinations. We listed the results of this model in Table 3 where we would describe the betas to be plausible. An increase travel time would be a obvious cost. The negative value of an intersection, especially in the city center where almost all road/bicycle intersections are equipped with traffic lights is also expected. The positive reward for for left-turn seems as expected to avoid crossing traffic. The

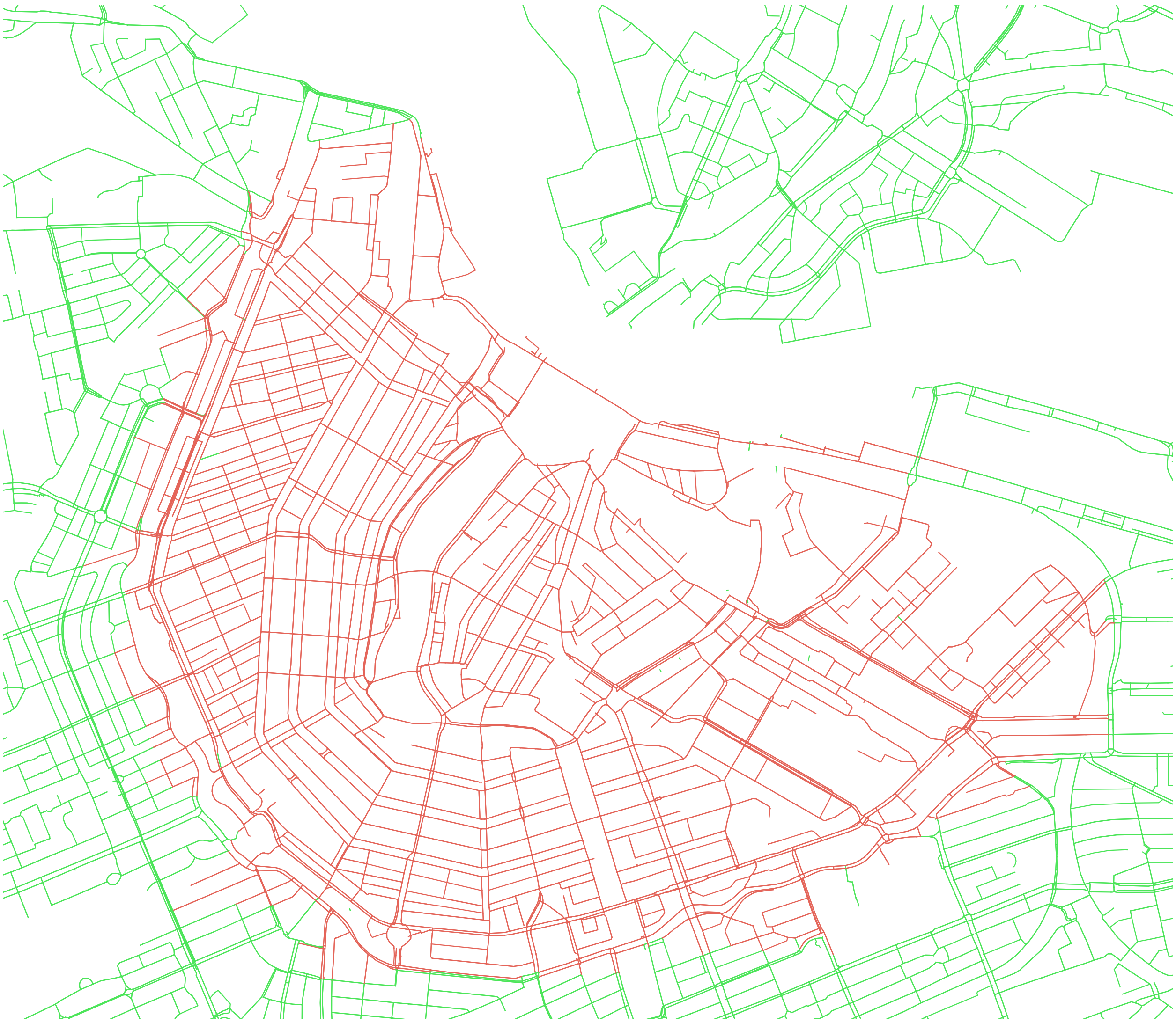


FIGURE 3: Map showing the selected area for city center of Amsterdam in red

positive reward for u-turn may seem odd, but u-turn costs on a bicycle should be less costly than in a motorized vehicle.

Edge-based network on Amsterdam city center: environmental variables

Subsequently we modelled travel-time with our five variables separately: 70db traffic noise distance, near green distance, near water distance, tree covered distance, cycle way distance. While all these models did converge, it was only able to invert the hessian to calculate a standard error for models: travel-time and length with tree cover; travel-time and length along water; travel-time and length with traffic noise. We listed these results in Table 3. The $\beta length - treecover$ for is not significant, possibly because tree cover is less of an issue in the city center which is shielded by buildings. The result for traffic noise is not significant either. The only significant effect found besides travel-time was the distance travelled near water, possibly due to cycling along the cobble-stone paved narrow canals navigating between cars, trucks and tourists being perceived as disadvantageous to persons who cycle for daily activities.

TABLE 3: Results from models that were estimated on 987 observations in the city center of Amsterdam.

Utility parameter	Estimated value	t-Test
Travel time (minutes)	-2.9119	-19.7751
Intersections	-0.6356	-13.8377
Left turn	-1.5717	-20.8542
U-turns	0.4205	7.5003
Log likelihood	-4.489078	
Travel time (minutes)	-18.0368	-2587.03
Log likelihood	-10.279184	
Travel time (minutes)	-18.03681	-2587.035
Distance exposed to traffic noise >= 55db	-1.9330	-0.1354
Log likelihood	-10.279184	
Travel time (minutes)	-18.0369	-2580.7
Distance along water (km)	-2566.333	-3.7053
Log likelihood	-10.274874	
Travel time (minutes)	-18.0368	-2587.0
Distance along trees (km)	-1.9330	-0.0098
Log likelihood	-10.279184	

DISCUSSION

Given our experience with the Amsterdam model, we highlight several challenges during the estimation of the recursive logit model. We reflect on why our initial plan for estimating the bicycle choice behavior in the entire city of Amsterdam with environmental variables was feasible with the baseline multinomial logit model, but faced numerical complications with the recursive logit model.

Negative reward formulation

In the original paper by Fosgerau et al. (10) on recursive logit it is mentioned that to formulate the path choice problem as a dynamic discrete choice model with the utility maximization problem consistent with a dynamic programming problem, the deterministic utility component is required to have negative value: $v_n(a|k) = v(x_{n,a|k;\beta}) < 0$.

As an experiment we set up a network based on a simple grid layout, with 625 intersections, allowing the user to move left, right, up, down. There is one diagonal connection across from the top left corner to the bottom right corner. We included each segment between the intersections as a single unit of distance. See Figure 4 for a visualization of 10 by 10 grid. We set up 4 different variables: $\beta_{distance}$ for the unit distance, $\beta_{intersection}$ that counts each intersection passed, β_{left} that counts each move towards the left side of the grid, $\beta_{diagonal}$ that counts each diagonal move. We included two observations across the top and right of the grid and a observation across the left

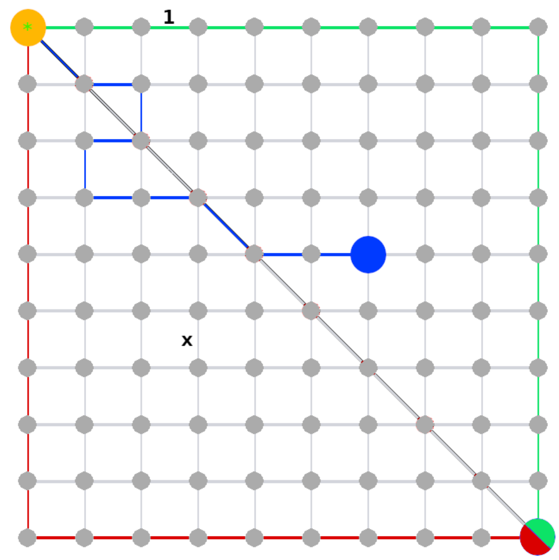


FIGURE 4: Two fixed paths to same destination along the boundaries of the graph (in red and green), plus example of one randomly generated path (in blue). All paths start at the top left corner and end respectively at the large red/green circle and the large blue circle.

and bottom of the grid and a series of 10 random observations that have a strong preference to move diagonally when possible. This model estimated with a log likelihood of -10 and $\beta_{distance} = -1.54467129, \beta_{intersection} = -2.04467129, \beta_{diagonal} = -2.09161539, \beta left = -81.34025$.

What we observed is that altering the attribute of a single link of this model to make the utility of that link positive lead to the inability of the linear solver to return a valid solution and thus not being able to find a log likelihood or estimate a model.

An implication that when using recursive logit you should aim for only including costs in your function u. In practice this might turn out tricky as cost variables may turn out to be correlated with reward variables not included in your model. For example heavy traffic near a bicycle path may seem like a cost variable at first, but as such roads are likely equipped with street lights in contrast to a path through a dark and empty park, such variable may turn out to have a negative cost.

Valid initial parameters and length of observations

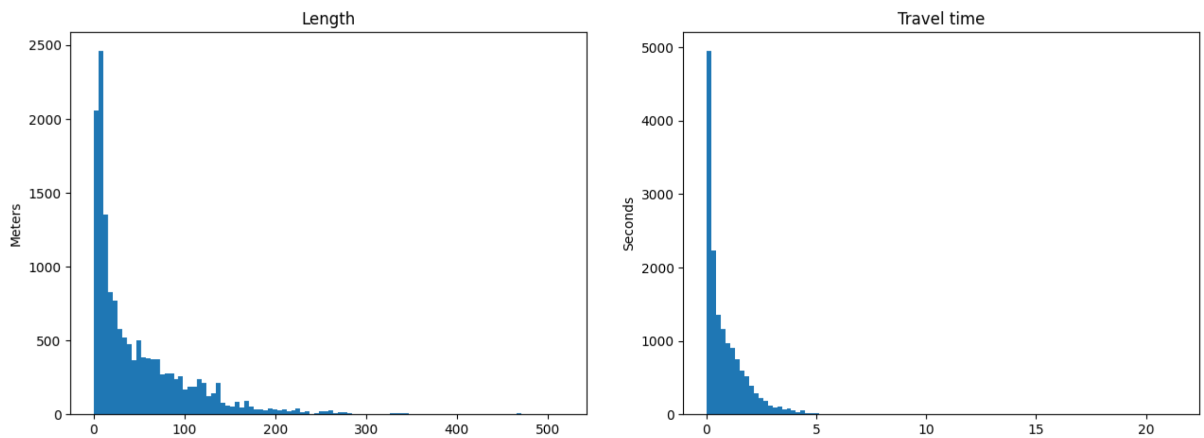
To take a closer look at how difficult it can be to determine a valid initial parameter prior to iterative solution of the system, we proceeded to look at solely at travel-time without any other features in the model. To do so, we manually computed the log-likelihood function for a range of the $\beta_{travel-time}$ parameter in the range between -1 and -25. We saw that only in a small window of $\beta_{travel-time}$ between approximately -18.02 and -21.01 a valid log likelihood function exists. For a $\beta_{travel-time} <= -18$ the equation system would return an invalid sign for the log likelihood, for $\beta_{travel-time} >= 21.01$ at least one of the observations would return a $exp(V) = 0$ for at the starting value.

This narrow range was achieved with a number of links in each observations limited at 40.

If we allowed observations with more links we were unable to find a window of initial parameters where the log likelihood function is valid at all. We see numerical issues as the root cause of this. As a long recursion will be a sum of each link utility, with high values due to the exponential, we expect these results to be caused by overflows and under flows in the solver.

The distribution of values of features and network degree centrality

Subsequently we attempted a similar experiment with the only feature in the model being β_{length}, which is correlated with $\beta_{travel-time}$. We were unable to find an exact parameter of β_{length} that is valid, but deduce it is somewhere between -413.6 and -413.7, based on where the solver returns a valid solution but $exp(V) = 0$. To look at the difference between both variables we refer to descriptive statistics presented earlier in Table 1 and histograms of the distributions plotted in Figure 5.



(a) Length **(b) Travel time**

FIGURE 5: Histogram of the variables length and travel time of the bicycle network in the city center of Amsterdam

Based on the descriptive statistics we expect the same root cause that makes it difficult to find a valid starting parameter. The lower kurtosis in the distribution of the length indicates a fatter right-tailed distribution presenting more possibility for a significant number of relatively large values to end up added together in the recursion on links. This too can lead to overflows and under flows making it difficult to find starting values betas due to numerical issues.

The number of alternative choice options

Another difference with existing studies in the literature that due to the complexity of bicycle infrastructure in Amsterdam, the number of possible options is higher than we would see in car route choice or in a city without two cycle-paths on both sides of major roads or two roads in both directions (for cyclists) along the canals.

CONCLUSION

Recursive logit is a promising solution for inverse reinforcement learning on specific route choice problems. However when designing your model and variables it is very important to keep the limitations of the linear equation system in mind. These limitations can make it impossible to estimate your model or lead to wrong estimations.

As recursive logit may fail to converge if even a single link has a (high) reward instead of cost, it is important to think through whether your variables are always costs for all links in the network. This can be hard in practice, as assumptions can be deceiving. For example you might model a bridge as a cost, as there is a small slope involved, however in reality people might prefer a route over a bridge as a form of sight seeing opportunity. Furthermore preferences can differ by person or vary over the time of day. For example a park might be a beneficial detour during the day, but during the night an empty badly lit park that feels unsafe might be worth a detour around instead.

RECOMMENDATIONS

For future study we are interested in the precise computational details that lead to the invalid estimates by the solver when faced with numerical overflow and underflow issues.

We are also looking into how well extensions of algorithms based on maximum entropy IRL of Ziebart et al. (14) will function with the Amsterdam bicycle network given the successful implementation of inverse reinforcement learning for bicycle paths in the work by Mo (19), however as in this study similar limitations were noted regarding the size of the state space.

In our initial experiments with implementing maximum entropy IRL by Ziebart et al. (14) on even the simplified problem of the Amsterdam city center with observations with a length of less than 30 links, we encountered overflows when calculating the local action probabilities as the expected reward for each state grew exponentially even when applying a discount factor.

Our recommendation would consist of searching for a way to dramatically simplify the state space by find ways to abstract the decision process. One possibility could be for example applying principles of path complexity such as in Koch et al. (1).

ACKNOWLEDGEMENTS

This research has been conducted in the framework of the Impact Study North/Southline, funded in part by the Municipality of Amsterdam and the regional transportation authority of Amsterdam. We would gratefully like to acknowledge the Vrije Universiteit Amsterdam legal affairs office, IT department, and university library research data management team as well as the IT and Facilities department of CWI for support in the design and implementation of the computing infrastructure. Hereby we also thank CWI/VU guest researcher Luk Knapen for useful discussions related to the research topic and assistance with the data preparation.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: TK, ED; data collection: TK, ED; analysis and interpretation of results: TK, ED; draft manuscript preparation: TK, ED. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. Koch, T., L. Knapen, and E. Dugundji, Path complexity and bicyclist route choice set quality assessment. *Personal and Ubiquitous Computing*, 2019, pp. 1–13.
2. McFadden, D. et al., Conditional logit analysis of qualitative choice behavior, 1973.
3. Frejinger, E., M. Bierlaire, and M. Ben-Akiva, Sampling of alternatives for route choice

modeling. *Transportation Research Part B: Methodological*, Vol. 43, No. 10, 2009, pp. 984–994.

4. Ben-Akiva, M. and M. Bierlaire, Discrete choice methods and their applications to short term travel decisions. In *Handbook of transportation science*, Springer, 1999, pp. 5–33.
5. Cascetta, E., A. Nuzzolo, F. Russo, and A. Vitetta, A modified logit route choice model overcoming path overlapping problems. Specification and some calibration results for interurban networks. In *Transportation and Traffic Theory. Proceedings of The 13th International Symposium On Transportation And Traffic Theory, Lyon, France, 24-26 July 1996*, 1996.
6. Menghini, G., N. Carrasco, N. Schüssler, and K. W. Axhausen, Route choice of cyclists in Zurich. *Transportation research part A: policy and practice*, Vol. 44, No. 9, 2010, pp. 754–765.
7. Ton, D., O. Cats, D. Duives, and S. Hoogendoorn, How Do People Cycle in Amsterdam, Netherlands?: Estimating Cyclists' Route Choice Determinants with GPS Data from an Urban Area. *Transportation research record*, Vol. 2662, No. 1, 2017, pp. 75–82.
8. Ton, D., D. Duives, O. Cats, and S. Hoogendoorn, Evaluating a data-driven approach for choice set identification using GPS bicycle route choice data from Amsterdam. *Travel behaviour and society*, Vol. 13, 2018, pp. 105–117.
9. Zimmermann, M., T. Mai, and E. Frejinger, Bike route choice modeling using GPS data without choice sets of paths. *Transportation research part C: emerging technologies*, Vol. 75, 2017, pp. 183–196.
10. Fosgerau, M., E. Frejinger, and A. Karlstrom, A link based network route choice model with unrestricted choice set. *Transportation Research Part B*, Vol. 56, 2013, pp. 70–80.
11. Mai, T., M. Fosgerau, and E. Frejinger, A nested recursive logit model for route choice analysis. *Transportation Research Part B: Methodological*, Vol. 75, 2015, pp. 100–112.
12. Mai, T., F. Bastin, and E. Frejinger, A decomposition method for estimating recursive logit based route choice models. *EURO Journal on Transportation and Logistics*, Vol. 7, No. 3, 2018, pp. 253–275.
13. de Freitas, L. M., H. Becker, M. Zimmermann, and K. W. Axhausen, Modelling intermodal travel in Switzerland: A recursive logit approach. *Transportation Research Part A: Policy and Practice*, Vol. 119, 2019, pp. 200–213.
14. Ziebart, B. D., A. L. Maas, J. A. Bagnell, and A. K. Dey, Maximum entropy inverse reinforcement learning. In *Aaai*, Chicago, IL, USA, 2008, Vol. 8, pp. 1433–1438.
15. Jaynes, E. T., Information theory and statistical mechanics. *Physical review*, Vol. 106, No. 4, 1957, p. 620.
16. Nguyen, Q. P., B. K. H. Low, and P. Jaillet, Inverse reinforcement learning with locally consistent reward functions. In *Advances in neural information processing systems*, 2015, pp. 1747–1755.
17. Mai, T., K. Chan, and P. Jaillet, Generalized Maximum Causal Entropy for Inverse Reinforcement Learning. *arXiv preprint arXiv:1911.06928*, 2019.
18. Mai, T., Q. P. Nguyen, K. H. Low, and P. Jaillet, Inverse Reinforcement Learning with Missing Data. *arXiv preprint arXiv:1911.06930*, 2019.
19. Mo, M., *Cyclist's route choice prediction with inverse reinforcement learning*. Master's thesis, Universiteit van Amsterdam, the Netherlands, 2019.
20. Bikeprint, *Download bestanden Nationale Fietstelweek 2015 en 2016*, 2017.

21. Nielsen, O. A., A stochastic transit assignment model considering differences in passengers utility functions. *Transportation Research Part B: Methodological*, Vol. 34, No. 5, 2000, pp. 377–402.
22. Halldórsdóttir, K., N. Rieser-Schüssler, K. W. Axhausen, O. A. Nielsen, and C. G. Prato, Efficiency of choice set generation techniques for bicycle routes. *European journal of transport and infrastructure research*, Vol. 14, No. 4, 2014, pp. 332–348.
23. Bovy, P. H. and S. Fiorenzo-Catalano, Stochastic route choice set generation: behavioral and probabilistic foundations. *Transportmetrica*, Vol. 3, No. 3, 2007, pp. 173–189.
24. ETH-Zurich, *Position Data Processing*. `https://sourceforge.net/projects/posdap/`, 2012.
25. Rieser-Schüssler, N., M. Balmer, and K. W. Axhausen, Route choice sets for very high-resolution data. *Transportmetrica A: Transport Science*, Vol. 9, No. 9, 2013, pp. 825–845.
26. Bierlaire, M., *PandasBiogeme: a short introduction*. Technical Report TRANSP-OR 181219, Transport and Mobility Laboratory, EPFL, 2018.
27. Zimmermann, M. and E. Frejinger, A tutorial on recursive models for analyzing and predicting path choice behavior. *EURO Journal on Transportation and Logistics*, 2020, p. 100004.