Route choice set generation using variational autoencoders

Rui Yao, Shlomo Bekhor

Department of Civil and Environmental Engineering,

Technion – Israel Institute of Technology, Haifa 32000, Israel

Email: andyyao@campus.technion.ac.il,

sbekhor@technion.ac.il

Abstract

Choice set generation is a challenging task, since the consideration set is generally unknown to the modelers, and the full choice set cannot be enumerated in real size networks. The proposed variational autoencoder approach (VAE) is motivated by the idea that the chosen alternatives must belong to the consideration set. The VAE approach explicitly considers maximizing the likelihood of including the chosen alternatives in the choice set, and infers the underlying generation process.

The VAE approach for route choice set generation is exemplified using a real dataset. VAE-CNL model has the best performance in terms of goodness-of-fit and prediction performance, compared to models estimated with conventionally generated choice sets.

1. Introduction

Choice set generation is particularly challenging in the context of route choice modeling, as many possible alternatives exist in the network, and enumerating all feasible alternatives is impractical. In addition, it is not possible to observe the true choice set considered by an individual before making a choice (Ben-Akiva and Boccara, 1995).

Several choice set generation models were developed in the literature, most of them deterministic, e.g., random walk, (Frejinger et al., 2009), link penalty, (De La Barra et al., 1993), which assumes individuals do consider the full choice set when making choices (Guevara and Ben-Akiva, 2013a, b; Lai and Bierlaire, 2015).

Probabilistic choice set formation approach considers the cases that, individuals may consider, for any behavioral reason, a choice set that is smaller than the full choice set, and the considered choice set is probabilistic (e.g., two-stage approach: Manski, 1977; single-stage approach: Cascetta and Papola, 2001).

The motivation of the proposed choice set generation method is that the chosen alternatives must belong to the true consideration set, otherwise they would not have been chosen. This assumption may seem rather obvious but is typically overlooked in many conventional methods.

Deep generative models (e.g. variational autoencoder, VAE, Kingma and Welling, 2014), which infer the underlying sample generation process, produce high quality samples in other domains but have not been adapted for choice set generation. We propose to adapt the VAE method as a novel choice set generation method for forming probabilistic choice sets.

2. Methodology

We believe the true consideration set is generally unknown to the modeler. Given an alternative, we assume it has some degree of fuzziness on whether it belongs to the true consideration set (Cascetta and Papola, 2001). The proposed VAE model can be interpreted as some type of "binary model" that determines whether to include the alternative into the consideration set, based on the alternative attributes.

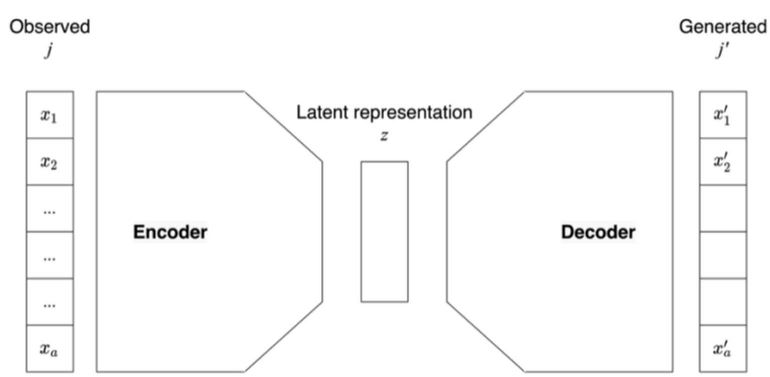

Figure 1 Illustration of an autoencoder

As shown in Figure 1, the autoencoder is composed of two parts: encoder and decoder. The encoder Φ is an inference model that maps the observed chosen alternative j to a more compact lower-dimension latent representation, z. The decoder Θ is a generative model that produces new samples, given the latent representations z. The purpose of an autoencoder model is to find the encoder and decoder such that the observation can be reproduced:

$$\Theta(\Phi(\mathrm{j})) \approx \mathrm{j} \tag{1}$$

By combining with variational Bayesian methods, the variational autoencoder model seeks to find two probability distributions: $p_\phi(z|j)$ for the inference process Φ, and $q_\theta(j|z)$ for the generative process Θ, and such that the likelihood of including the chosen alternatives in the consideration set $q_\theta(j)$, is maximized:

$$\ln q_\theta(j) = \log \int \frac{p(z|j)}{p(z|j)} p(z) q_\theta(j|z) dz = \ln \mathbb{E}_{p(z|j)} \left[\frac{p(z) q_\theta(j|z)}{p_\phi(z|j)} \right] \tag{2}$$

and this can be estimated using Monte Carlo simulation, the corresponding estimator can be derived using Jensen's Inequality as follows (Burda et al., 2015):

$$\ln q_\theta(j) = \log \mathbb{E}_{p_\phi(z|j)} \left[\frac{1}{S} \sum_{i=1}^{S} \frac{p(z) q_\theta(j|z)}{p_\phi(z|j)} \right] \geq \mathbb{E}_{p_\phi(z|j)} \left[\ln \frac{1}{S} \sum_{i=1}^{S} \frac{p(z) q_\theta(j|z)}{p_\phi(z|j)} \right] = \mathcal{L} \tag{3}$$

Where $\mathcal{L}$ is the lower bound of $\ln q_\theta(j)$. Then, by maximizing the lower bound $\mathcal{L}$, we are expecting to maximize $\ln q_\theta(j)$ as well (given the bound is tight enough). Note that, the S defines the number of random draws in the Monte Carlo, and $\mathcal{L}$ approaches $\ln q_\theta(j)$ as S goes to infinite (Burda et al., 2015).

Then, given a trained VAE model, generating an alternative means:

1. Draw z at random from the prior distribution $p(z)$
2. Draw new alternative j from the decoder Θ, $q_\theta(j|z)$, given z from step 1

3. Case study

The proposed VAE method for route choice modeling is applied as follows: we first process the raw route choice data (section 3.1), followed by the VAE choice set generation (section 3.2), and finally perform route choice model estimation and prediction (section 3.3).

3.1.Data processing

The dataset for the case study is based on the Tel Aviv household travel survey data (Nahmias-Biran et al., 2018) and map matched GPS trajectories (Yao and Bekhor, 2020). After cleaning and filtering the GPS data, 5,002 car trips are map matched. Main statistics of selected attributes are summarized in Table 1.

Table 1 Main statistics of selected route characteristic attributes

Attributes	**Mean**	**Std.**
Route length detour (ratio to shortest path length)	1.11	0.21
Route time detour (ratio to fastest path time)	1.08	0.17
Route city node percentage (ratio of num of intersections in the city center to all route intersections)	0.1	0.19
Route highway/expressway percentage (of total distance)	0.71	0.29
Route average route cost (per km)	1.66	0.26

Furthermore, the 5,002 observations are randomly split into a training subset of size 4,000 and another test subset with the remaining 1,002 observations. In order to have a fair comparison between different models, the same training subset and test subset will be used in different models.

3.2.VAE choice set generation

In this paper, the probability distributions $p_\phi(z|j)$ and $q_\theta(j|z)$ that model the underlying alternative generation process, are parameterized using neural networks (Kingma and Welling 2014; Burda et al. 2015).

The route attributes (Table 1) are in general non-negative, and the probability distribution $q_\theta(j|z)$ is often assumed to be log-normal, Gamma or truncated Normal in the literature (Nielsen and Frederiksen, 2006; Prato, 2009). Similarly, we assume $q_\theta(j|z)$ follows truncated Normal distribution as follows:

$$q_\theta(j|z) = \text{TruncatedNormal}(\mu_\theta(z), \sigma^2 \boldsymbol{I}, \boldsymbol{0}, \infty) \tag{4}$$

where, the mean $\mu_\theta(z)$ of the truncated normal distribution is the decoder neural network, and θ are the weights of the neural network, variance σ is a fixed hyperparameter, and $q_\theta(j|z)$ is bound from below at 0. And the posterior distributions $p_\phi(z|j)$ is assumed to follow a parametric normal distribution:

$$p_\phi(z|j) = N\left(\mu_\phi(j), \text{diag}\ \sigma_\phi(j)\right) \tag{5}$$

where, the mean $\mu_\phi(j)$ and variance $\sigma_\phi(j)$ of the normal distribution are the encoder neural networks, and ϕ are the weights of the neural network. Based on this assumption, the latent variable z can be interpreted as cluster scores, capturing the membership of the alternative belonging to which cluster (nest); or as latent factors for the random constraints (Swait and Ben-

Akiva, 1987; Ben-Akiva and Boccara, 1995). For simplicity, we assume that the prior distribution follows a standard normal distribution:

$$p(z) = N(\mathbf{0}, \boldsymbol{I}) \quad (6)$$

Note that, the proposed VAE method and the objectives are general to handle different distributions, we assume these distributions (e.g., $q_\theta(j|z)$ as truncated normal distribution) according to our dataset. There is no restriction to the assumption of distributions, apart from being continuous. The proposed neural network VAE model is shown in Figure 2. Herein, $j = \{x_1, \dots, x_a\}$ are the characteristics attributes of an alternative j, $z = \{z_1, \dots, z_b\}$ are the attributes of the latent variable.

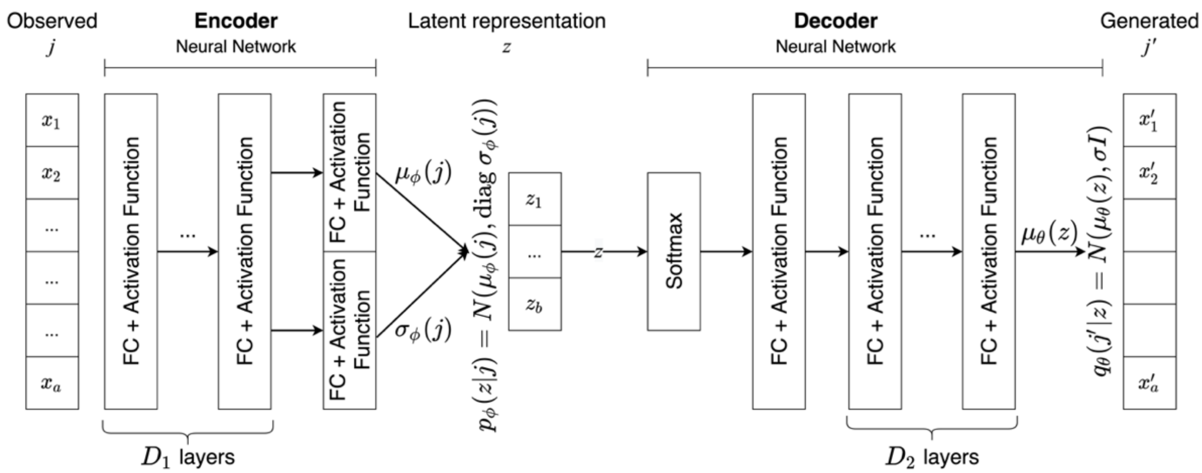

Figure 2 VAE neural network for route choice modeling

In the training phase of this neural network, the alternative attributes j undergo transformation in the encoder neural network first, which estimates the mean $\mu_\phi(j)$ and variance $\sigma_\phi(j)$ of the posterior distribution $p_\phi(z|j)$. Next, the latent attributes z are sampled from this posterior distribution and passed to the decoder neural network.

On receiving the latent attributes, the decoder network first normalizes these attributes using the Softmax (MNL) function, and then use these attributes to estimate the mean $\mu_\theta(z)$ of $q_\theta(j|z)$. Finally, a new sample is draw from $q_\theta(j|z)$. The VAE model parameters are estimated by maximizing the log-likelihood function in equation (3).

When the training phase is finished, we can generate new alternatives by drawing z from the prior distribution $p(z)$, and pass it to the decoder network. In this paper, the choice set size is assumed to have 20 alternatives (as suggested in Bekhor et al., 2006). This means, for each observation, the process of sampling z and passing it the decoder network is conducted 20 times. The alternative attributes generated using the procedure described above are shown in Figure 3.

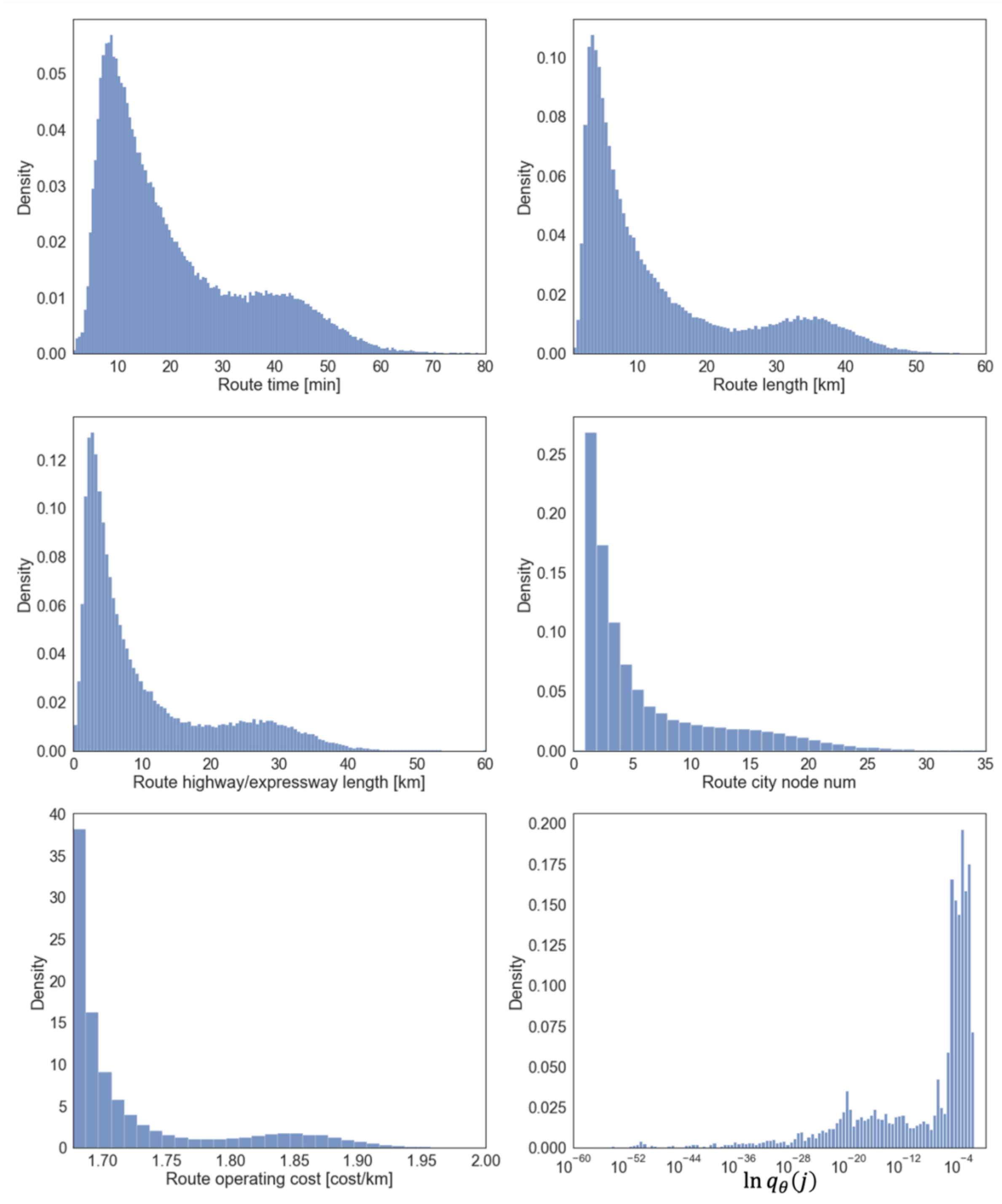

Figure 3 VAE generated alternative attributes for the case study

3.3.Route choice model estimation and prediction

In this subsection, we present estimation results of the route choice models using the training subset, and prediction performance of the estimated model using the testing subset. Given the generated choice set D_n, we specify the systematic part of the utility function as follows:

$$V_{jn} = \sum_{a} \beta_a \cdot x_{a,jn} \tag{7}$$

where, β_a is the generic parameter for the explanatory variable a, and $x_{a,jn}$ is the attribute value for alternative j of individual n. Note that, the explanatory variables are listed in Table 1, and are converted from normalized attributes to absolute attributes.

Two types of route choice models with VAE choice set generation are estimated: 1) MNL; 2) CNL. In addition, 10 replications of the generation–estimation procedure are performed for these models. The estimation results are summarized in Table 2, in which the empirical mean parameter estimates $\hat{\beta}$, mean standard deviations $\hat{\sigma}$, standard deviations across replications σ_{rep}, and their t-test values against zero (in brackets) are presented.

Table 2 Parameter estimates of route choice models with VAE choice set generation

		MNL	CNL
Route time	$\hat{\beta}$	-0.0936	-0.0665
	$\hat{\sigma}$	0.0083	0.0058
	t-test(0)	(-11.29)	(-11.52)
	σ_{rep}	0.0058	0.0039
Route length	$\hat{\beta}$	-0.7235	-0.5009
	$\hat{\sigma}$	0.0187	0.0130
	t-test(0)	(-38.73)	(-38.39)
	σ_{rep}	0.0145	0.0103
Route highway/expressway length	$\hat{\beta}$	0.9755	0.6811
	$\hat{\sigma}$	0.0154	0.0108
	t-test(0)	(63.29)	(63.01)
	σ_{rep}	0.0107	0.0075
Route count of city nodes	$\hat{\beta}$	0.0605	0.0406

	$\hat{\sigma}$	0.0035	0.0025
	t-test(0)	(17.09)	(16.22)
	σ_{rep}	0.0036	0.0026
Route average cost	$\hat{\beta}$	-0.1166	-0.0805
	$\hat{\sigma}$	0.0132	0.0092
	t-test(0)	(-8.82)	(-8.76)
	σ_{rep}	0.0144	0.0098

Results indicate that all parameter estimates have significant explanatory power, in terms of the t-test values (with critical value of 1.96). Moreover, the VAE model produces consistent coefficient estimates across different replication runs, as the standard deviations across replications σ_{rep} are relatively small for all the parameter estimates.

We further compare the estimation and prediction performances of our proposed VAE models, against two conventional models with choice sets constructed using link penalty method. The model prediction performances, in terms of log-likelihood, false positive rate (FPR) and F1-score, are evaluated using the testing subset with the estimated models. The FPR provides probability of falsely rejecting the hypothesis for a particular test, and it is calculated as follows:

$$FPR = \frac{False\ Positive}{False\ Positive + True\ Negative} \tag{8}$$

we report the average FPR across different alternatives in Table 3, while the F1-score is calculated using:

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \tag{9}$$

where, $\text{precision} = \frac{True\ Positive}{True\ postive + False\ Postive}$ and $\text{recall} = \frac{True\ Positive}{True\ Postive + False\ Negative}$. We report the goodness-of-fit and prediction performance results in Table 3. Remind that, the training subset has 4,000 observations, and the testing subset has 1,002 observations.

Table 3 Goodness-of-fit and prediction performance results

		Link penalty		**VAE approach***	
		MNL	CNL	MNL	CNL
Training subset	$LL(\beta = 0)$	-10010.79	-9947.46	-11982.93	-11997.04
	$LL(\hat{\beta})$	-6585.70	-6565.65	-5294.72	-5283.47

	$\bar{\rho}^2$	0.3416	0.3390	0.5577	0.5592
Testing subset	$LL(\beta = 0)$	-2504.43	-2483.43	-3001.72	-3007.87
	$LL(\hat{\beta})$	-1589.01	-1586.06	-1301.34	-1299.13
	$\bar{\rho}^2$	0.3635	0.3593	0.5648	0.5664
	FPR	2.76%	2.72%	2.19%	2.20%
	F1-score	43.81%	45.27%	58.26%	58.07%

**: VAE goodness-of-fit and prediction performance results are averaged over 10 replications*

The VAE models outperform models with conventional link penalty, in terms of final model log-likelihoods, False-Positive-Rate (FPR) and F1-score. This indicates that the VAE models could generate diverse alternatives that improve the estimated model generalizability.

In terms of prediction performances, the VAE-MNL and VAE-CNL models have better FPR and F1-score compared to Link penalty-MNL and Link penalty-CNL models. Moreover, the VAE-CNL model outperforms all other models in terms of model estimation and prediction, because of its capability of capturing the similarities between alternatives.

Summary and Conclusions

In this paper, a novel approach adapting variational autoencoders for probabilistic choice set generation is proposed. The VAE approach combines variational Bayesian methods with autoencoder to approximate the probability distribution for the underlying choice set generation process, by maximizing the likelihood of perceiving the chosen alternatives in the choice set.

The VAE first infers the alternatives to a more compact lower-dimension latent representation using the encoder model, then generates new alternatives with the decoder model given the latent representations. The choice set is implicitly generated by sampling latent representations from the latent space and passing them to the decoder model. The proposed VAE choice set generation model is parameterized using neural networks.

The capability of the VAE approach is illustrated with real data, using MNL and CNL models. Estimation results show that the models with choice set generated using the VAE approach outperform models with conventional choice sets, both in terms of model goodness of fit and prediction performance. In particular, the VAE-CNL model that captures the correlation between alternatives, has the best performance.

References

Ben-Akiva, M., & Boccara, B. (1995). Discrete choice models with latent choice sets. International journal of Research in Marketing, 12(1), 9-24.

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

Burda, Y., Grosse, R., & Salakhutdinov, R. (2015). Importance weighted autoencoders. arXiv preprint arXiv:1509.00519.

Cascetta, E., & Papola, A. (2001). Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand. Transportation Research Part C: Emerging Technologies, 9(4), 249-263.

De La Barra, T., Perez, B., & Anez, J. (1993). Multidimensional path search and assignment. In PTRC Summer Annual Meeting, 21st, 1993, University of Manchester, United Kingdom.

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

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

Guevara, C. A., & Ben-Akiva, M. E. (2013b). Sampling of alternatives in logit mixture models. Transportation Research Part B: Methodological, 58, 185-198.

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

Manski, C. F. (1977). The structure of random utility models. Theory and decision, 8(3), 229.

Nahmias-Biran, B. H., Han, Y., Bekhor, S., Zhao, F., Zegras, C., & Ben-Akiva, M. (2018). Enriching activity-based models using smartphone-based travel surveys. Transportation Research Record, 2672(42), 280-291.

Nielsen, O. A., & Frederiksen, R. D. (2006). Optimisation of timetable-based, stochastic transit assignment models based on MSA. Annals of Operations Research, 144(1), 263-285.

Prato, C. G. (2009). Route choice modeling: past, present and future research directions. Journal of choice modelling, 2(1), 65-100.

Swait, J., & Ben-Akiva, M. (1987). Incorporating random constraints in discrete models of choice set generation. Transportation Research Part B: Methodological, 21(2), 91-102.

Kingma, D. P., & Welling, M. (2014). Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114.

Yao, R., & Bekhor, S. (2020). Experiments on route choice set generation using a large GPS trajectory set. arXiv preprint arXiv:2006.045