Australasian Transport Research Forum 2017 Proceedings
27 – 29 November 2017, Auckland, New Zealand
Publication website: http://www.atrf.info

The effect of sample size and priors on the estimation of path choice parameters: A Bayesian perspective

Mohadeseh Rahbar*, Mark Hickman, Mahmoud Mesbah, Ahmad Tavassoli
School of Civil Engineering, the University of Queensland, Australia
*Email for correspondence: m.rahbar@uq.edu.au

Abstract

This paper examines the effect of sample size on the accuracy of a path choice model in a Bayesian framework. In this study, we examine parameters associated with travel time. Furthermore, different priors (uninformative, informative, and overly informative) are chosen to achieve a sequence of posteriors at each given sample size. For comparison purposes, the root-mean-square errors (RMSE) between each posterior and actual observation and between each posterior and General Transit Feed Specification data are computed. The posterior minimizing the %RMSE defines the effective sample size (ESS). The suggested approach is applied to automatic fare collection data from South East Queensland, Australia for one day, and quantifies the effects of sample size and prior information level. The results show that for an uninformative case detecting the ESS is not possible. For the model with an informative prior, the ESS is 50% of the sample (12,450 OD pairs and 39,270 journeys) and, for the model with an overly informative prior, the ESS is 10% of the sample (2,490 OD pairs and 7,270 journeys).

1. Introduction

Public transport planners use transit assignment models to predict passenger loads and levels of service in order to evaluate existing and future scenarios. Transit assignment is the method of assigning a given set of passenger flows for an origin-destination (OD) pair to the current transit network based on particular travellers' path choice principles. Recently, some studies using Bayes' theory have been done for detecting used paths and traffic and transit assignment (Hazelton, 2008, Hazelton, 2010, Wei and Asakura, 2013, Sun et al., 2015, Rahbar et al., 2017). Modifying one's initial probability statements about the parameters (known as "priors") using the data at hand is the learning process involved in Bayesian inference (Congdon, 2001). From Bayes' theorem (Equation 1), prior knowledge is shown by the density $p(\Theta)$ of the parameters Θ, the likelihood of observations y given the parameters Θ is $p(y|\Theta)$, and the posterior density of the parameters Θ is $p(\Theta|y)$, which contains the updated knowledge from the observations y.

$$p(\Theta|y) \propto p(y|\Theta)\, p(\Theta) \tag{1}$$

The amount of data y and the distribution of the prior(s) Θ are two fundamental questions in any Bayesian analysis (Morita et al., 2008). For many parametric Bayesian models, the answers do not seem straightforward. For example, with a small to moderate sample size of y, the prior's role would be substantial. If the prior is a technically acceptable choice, then detecting the effective sample size (ESS) would become more important (Berger et al., 1994).

On the other hand, by computing ESS, the use of an overly informative prior may be avoided in a sense that inference is mostly affected by the prior rather than the data. The ESS is the minimum sample size needed to estimate a process parameter, precisely. In fact, the sample size is an important feature of any study in which the objective is to make inferences for a population from a sample.

In this paper, we conduct an approach for computing the ESS for the data *y* by adjusting a variety of settings. First, we select the transit assignment model suggested by Rahbar et al. (2017). Rahbar et al, taking advantage of high-quality travel data provided by smart cards, presented a transit assignment framework using a Bayesian inference approach. Second, we select different sample sizes from one weekday data of the South East Queensland (SEQ) bus, train, and ferry modes. At each sample size level, a number of iterations are generated to overcome the bias of a randomly selected sample size. Third, priors are generated using three definitions: 1) uninformative, 2) informative, and 3) overly informative. The combination of sample size levels and priors gives a particular scenario. The results of different scenarios are compared with the actual data and with the General Transit Feed Specification (GTFS) data (Google Inc., 2013). Since GTFS feeds let public transit agencies publish their transit data, it is a good basis for comparison in this study. Finally, the value of sample size minimizing the %RMSE is the ESS.

The remaining sections of this paper are organized as follows. Section 2 presents the application of Bayesian inference in detecting used paths and passenger flow assignment. The methodology including data description, sample size selection, and definition of priors is presented in Section 3. Applications are described in Section 4, including discussions of comparisons between the results of the model and actual data, as well as GTFS. We close with a brief discussion in Section 5.

2. The Bayesian Approach

Recently, some studies using Bayes' theory have been done for detecting used paths and traffic flow assignment. One of the first Bayesian applications in a transport context belongs to Maher (Maher, 1983) in estimating an origin-destination (OD) matrix. Eventually, the estimation of an OD matrix using the Bayesian approach has been more widely accepted in some other studies (Li, 2005, Hazelton, 2008, Li, 2009, Yamamoto et al., 2009, Hazelton, 2010, Perrakis et al., 2012, Perrakis et al., 2015). In general, the overall uncertainty of the estimates by delivering posterior distributions for the parameters, as well as of the predictive distributions for future OD flows, was reduced using Bayesian methodology.

Given used routes for each OD pair, Fu et al. (2014) presented a Gaussian mixture model to find the route shares and the mean and variance of travel times for each route. Later, an approach including the number of used routes as an unknown parameter into a Bayesian framework based on a reversible-jump MCMC algorithm was proposed by Lee and Sohn (2015). They compared the performance of this new approach with the existing method, which depends on the Bayesian information criterion (BIC). The new method showed flexibility in recognizing route-use patterns through the marginal posterior distribution of other unknown parameters. A Bayesian statistical inference framework to model passenger flow assignment in a metro network was proposed by Sun et al. (2015). By observing passenger travel times provided by smart card data and prior knowledge about the Singapore Metro network, they built the posterior density of path choice parameters. The results showed that the disutility of transfer time is about twice of that of in-vehicle travel time in the Singapore Metro system. Rahbar et al. (2017) developed a Bayesian statistical framework to model passengers' path choice behaviour and to estimate travel time attributes of the network. The travel time on each link, transfer time, the impact of in-vehicle time, and the impact of transfer time on passenger path choice behaviour were the important variables and parameters of their model. That model, with respect to its successful application on the Brisbane, Australia transit network, is chosen for analysis in this paper. The model is written as follows:

$$\pi(c,\Theta|T) \propto \prod_{od\in OD}\prod_{t\in T_{od}}\left(\sum_{r\in PT_{od}} \frac{\exp\left(\frac{-(t-\sum_{i\in r} c_i)^2}{2(\sum_{i\in r} {c_i}^2+\sigma^2)}\right)}{\sqrt{2\pi(\sum_{i\in r} {c_i}^2+\sigma^2)}} \times \frac{\exp(\Theta_1\sum_{a\in r} c_a+\Theta_2\sum_{f\in r} c_f)}{\sum_{\acute{r}\in PT_{od}}\exp(\Theta_1\sum_{a\in\acute{r}} c_a+\Theta_2\sum_{f\in\acute{r}} c_f)} \times \pi(c)\times\pi(\Theta)\right) \tag{2}$$

where c_a, c_f, Θ_1, and Θ_2 show travel time on link *a*, transfer time on link *f*, the impact of in-vehicle time, and the impact of transfer time on passenger path choice behaviour, respectively. Based on this model, the probability of parameters given the observed travel times $\pi(c,\Theta|T)$ is equal to the probability of observing travel time *t* on path *r* from path set PT_{od}, multiplied by the probability of selecting path *r*, mulitiplied by the prior probability of parameters $\pi(c)\times\pi(\Theta)$. As can be seen in Equation 2, the likelihood function contains two parts, a Multinomial Logit Model (Equation 3) which is a function of route attribute (link travel time) and the probability of observing *t* given path *r* and other parameters (Equation 4).

$$p_r = \frac{\exp(\Theta_1\sum_{a\in r} c_a+\Theta_2\sum_{f\in r} c_f)}{\sum_{\acute{r}\in PT_{od}}\exp(\Theta_1\sum_{a\in\acute{r}} c_a+\Theta_2\sum_{f\in\acute{r}} c_f)} \tag{3}$$

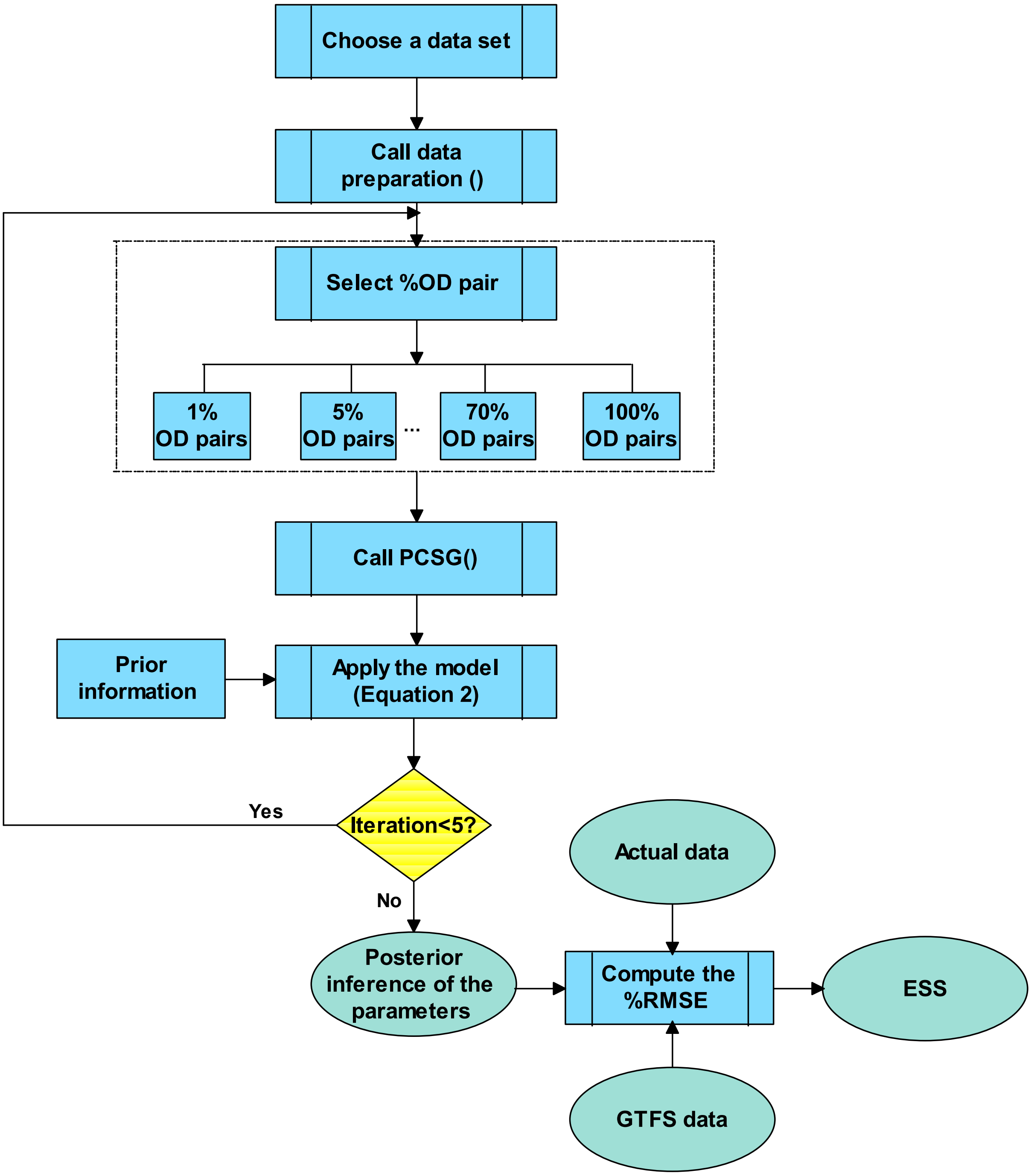
$$p(t|r,c,\Theta) = \frac{\exp\left(\frac{-(t-\sum_{i\in r} c_i)^2}{2(\sum_{i\in r} {c_i}^2+\sigma^2)}\right)}{\sqrt{2\pi(\sum_{i\in r} {c_i}^2+\sigma^2)}} \tag{4}$$

Also, several studies have been conducted to improve the transferability of mode choice models and travel demand models using the Bayesian updating method (Talvitie and Kirshner, 1978, Galbraith and Hensher, 1982, Santoso and Tsunokawa, 2005, Santoso and Tsunokawa, 2010, Rashidi et al., 2013, Karasmaa, 2007, Mei et al., 2005, Molla, 2017). One primary reason for this is that the transferability of models indicates that a model developed in one location at one point in time can be applied to another location at another point in time, thereby considerably reducing data collection and model estimation requirements in the application context (Karasmaa and Pursula, 1997). The ultimate objective of this paper is also enhancing the transferability of the model (Equation 2).

3. Methodology

In-vehicle time and transfer time are two main parameters of the path choice model (Equation 2) used by Rahbar et al. (2017). The values of these parameters (c_a and c_f) are based on the estimated values of individual link travel times (c), which cannot be observed from the data. In this paper, to study the effect of different sample sizes on the accuracy level of the estimated link travel times, different sample sizes are chosen randomly each time from the full data set. Figure 1 shows the procedure of sample size selection and ESS detection. This procedure starts with data set selection. After selecting a data set, the data preparation algorithm (Figure 2) is called. The main functions of this algorithm, explained in the next section, are OD matrix estimation, travel time calculation, and path choice set generation. Then, data sampling is done randomly, based on the OD pairs; for example, 1% of OD pairs, 5% of OD pairs, and so on. Since different OD pairs are selected for each sample size, specific path choice sets need to be generated. For this purpose, the Path Choice Set Generation (PCSG) task of the data preparation algorithm is called for each sample size. In the PCSG, the actual used paths by passengers are extracted from the dataset to generate the path choice sets. In the next stage, using the prior information on travel times and path choices, the Bayesian model (Equation 2) can be applied on the prepared data. For each sample size, a number of iterations are generated to overcome the bias of the randomly selected sample. The first output of this procedure is the posterior distributions of in-vehicle time and transfer time. The mean of the posterior distributions are compared with the actual data and GTFS data and the %RMSE between each posterior distribution and actual data (OD travel time) and between the posterior distribution and GTFS data are computed. The value of sample size minimizing the %RMSE is the ESS, as the last output of this procedure.

Figure 1: Sample size selection and estimation algorithm



3.1. Data preparation and sample size selection

In this study, the main focus is on those journeys which start within a weekday morning peak (7–9 AM) and finish within or after this period. Using the regional smart card fare payment system in SEQ, Australia, the 'Go Card' (TRANSLink, 2017), all essential data such as date, time, boarding and alighting location, route number, and direction of travel for each transaction are provided. Table 1 shows an example of four transactions of the selected data on 20 March 2013.

Table 1: Sample transactions of the selected data

Card ID	Boarding stop	Alighting stop	Boarding time	Alighting time	Route
1	Logan Rd at Swain Street	Elizabeth Street Stop 81 near George St	8:22	8:39	P189
2	RBWH station	Cultural Centre station	7:38	7:50	66
2	Cultural Centre station	UQ Lakes station	7:53	8:03	109
3	Maundrell Tce at John Goss Reserve	King George Square station	7:59	8:35	P343

Figure 2 shows the procedure of extracting required information from the smart card data. In the first step (data pre-processing), all records in which boarding or alighting stops are missing, boarding or alighting times are missing, the boarding stop is the same as the alighting stop, and/or the boarding time is later than the alighting time, are removed (Tavassoli et al., 2016). At the data pre-processing step, about 30% of data has been removed. Also, Bureau of Meteorology (2017) data is used to be sure that the selected day is a typical day in terms of the weather condition, rainfall, temperature, humidity, and wind speed and direction. In the second step, based on the boarding time and alighting time of a unique card ID on the selected day, the sequence of all trip legs are extracted. Also, the network coordinates (nodes and links) from the strategic transport model are used. In the third step, the in-vehicle time and the number of transfers for each unique card ID are calculated. For ferry and bus, the difference between the alighting time and boarding time is defined as the in-vehicle time. For train, this difference in “boarding” and “alighting” times also includes waiting time and egress time, as the smart card transactions occur on platforms, rather than on-board the vehicles. In the fourth step, the trip leg sequence of each unique card ID helps us to create the journeys. Then, using the journey’s trip legs and the number of passengers on each route, the stop-level OD matrix is estimated. Using the route number, the involved links are extracted from South East Queensland’s GTFS. GTFS includes stops, routes, and timetables across SEQ. GTFS data expresses transit schedule information in a format that is a routable spatiotemporal network graph with stops as nodes, scheduled travel between stops as edges, and estimated travel times as the cost. After detecting all the paths used by passengers, the path sets for all OD pairs are generated in the PCSG phase. Each path can include one or more than one route. Finally, the travel time on each path and the path choice sets are the outputs of this algorithm.

For the selected typical day (20 March 2013), there are 24,878 OD pairs and 80,767 journeys during the AM peak period. Table 2 shows the characteristics of the selected data set. For example, row 4 indicates that each path set has at least one path and at most nineteen paths.

Figure 2: The data preparation algorithm

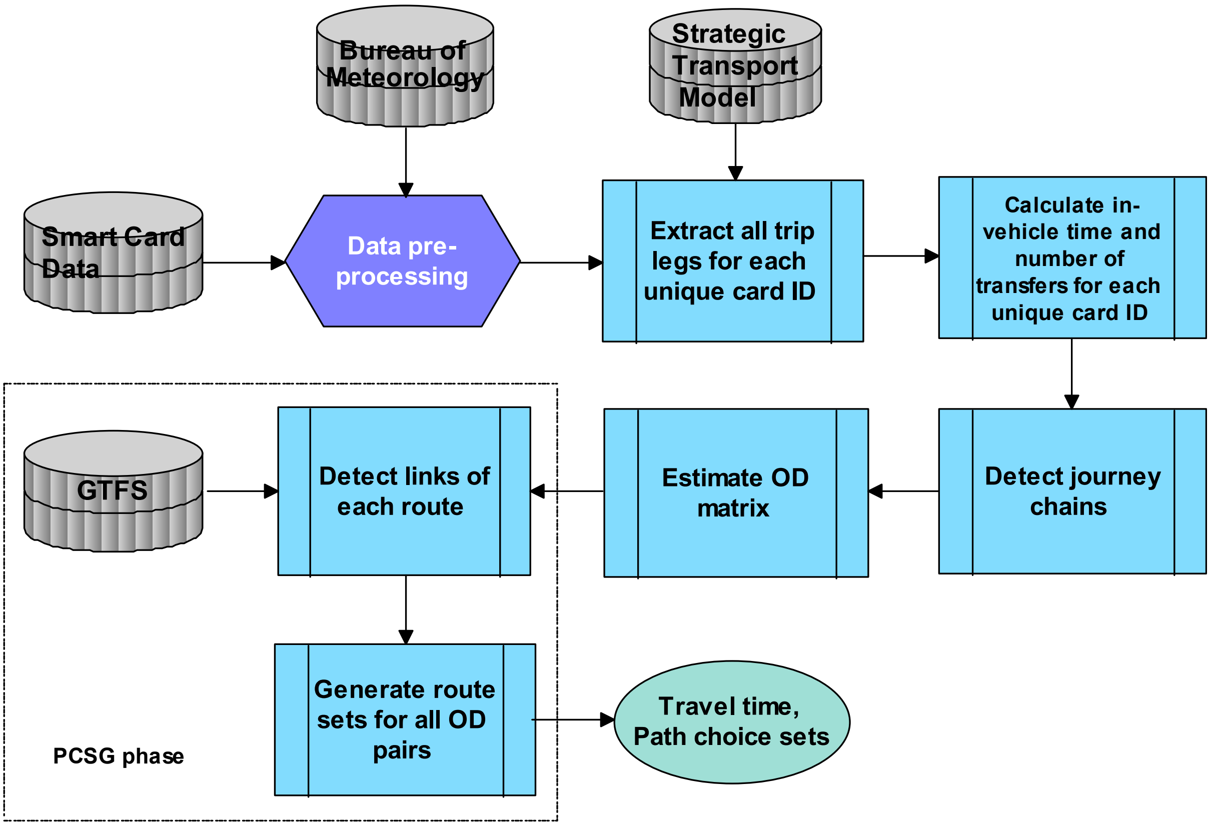


Table 2: Data characteristics

Row	Characteristics	Value
1	Number of OD pairs	24,878
2	Number of trip legs	95,658
3	Number of journeys	80,767
4	Number of paths in each path set	[1,19]
5	Number of passengers on each path	[1,272]
6	Number of passengers on each OD pair	[1,560]
7	Number of used links	12,255

Table 3 shows the number of OD pairs, journeys, and used links at different sample sizes. For each sample size, the OD pairs are selected randomly for five iterations.

Table 3: The associated data to the selected sample size percentages

Sample size (%)	No. of OD pairs	No. of journeys	No. of used links
1	250	840	2,344
5	1,250	3,620	5,907
10	2,490	7,270	8,037
30	7,464	25,578	10,454
40	9,952	31,413	11,152
50	12,450	39,270	11,390
70	17,415	57,283	11,933
100	24,878	80,767	12,255

3.2. Prior information

In the Bayesian inference framework, the prior distribution will be specified from researchers' subjective perspectives, independent of the data. Some previous researchers such as (He et al., 2002, Hofleitner et al., 2012, Li et al., 2013, Rahbar et al., 2016) assumed that link travel time follows a normal or lognormal distribution. In this paper, the authors assume that the link travel follows a normal distribution (Equation 5) with mean μ and standard deviation σ.

$$c \sim N(\mu,\sigma^2) \tag{5}$$

To select the parameters of the prior (μ and σ), three approaches are examined in this study: 1) uninformative, 2) informative, and 3) overly informative. In the first approach, it is assumed that there is no information on link travel time. We call this prior an uninformative prior with μ=5 and σ^2=5 for all links of the network. In the second approach, the authors use an estimated speed of public transport vehicles. The estimated speed, by considering the dwell time to allow passengers boarding and alighting, is 20 (km/h). After calculating the distance between each two successive stops, a unique μ is provided for each link. For example, for one link with 200 (m) length, the estimated μ would be 0.6 minute. The second approach's prior is considered as an informative prior. Finally, in the third approach, GTFS data are used to provide link travel times as prior information (so-called overly informative prior). Based on the GTFS data, the link travel times across the network vary from 1 to 37 minutes. Using the model-based clustering approach, links are categorized into seven groups, with μ = 1, 5, 9, 15, 23, 29, and 37 minutes. The model-based clustering approach is based on probability models wherein objects are assumed to follow a finite mixture of probability distributions, such that each component distribution represents a cluster (Oh and Raftery, 2012). The variance values (σ^2) for all three different prior distributions are 5 (min^2) to make sure that the selected distribution covers a greater distance from the mean value.

4. Results and discussion

The Bayesian model (Equation 2) is implemented in the R programming language (R Core Team, 2016) with these three different priors and is applied on different sample sizes using the smart card data. The model has been run for 108 different scenarios (3 priors × 7 sample sizes × 5 iterations, and with 3 priors × 100% data) on high performance computers of the University of Queensland. The estimated total travel times (in-vehicle and transfer time) are compared with actual data and GTFS. To show the associated errors, the percentage root mean square error is used. The lower the value of %RMSE, the lower the differences between the predicted values and the actual (Go Card) data and GTFS data. The %RMSE formula is presented as follows:

$$RMSE=\sqrt{\frac{\sum_{i=1}^{N}(Model_i-Actual_i)^2}{N}} \tag{6}$$

$$\%RMSE=\frac{RMSE}{\left(\frac{\sum_{i=1}^{N}Actual_i}{N}\right)} \times 100 \tag{7}$$

where $Actual_i$ is the observed value from GoCard data or GTFS, $Model_i$ is the predicted value, and N is the number of predictions.

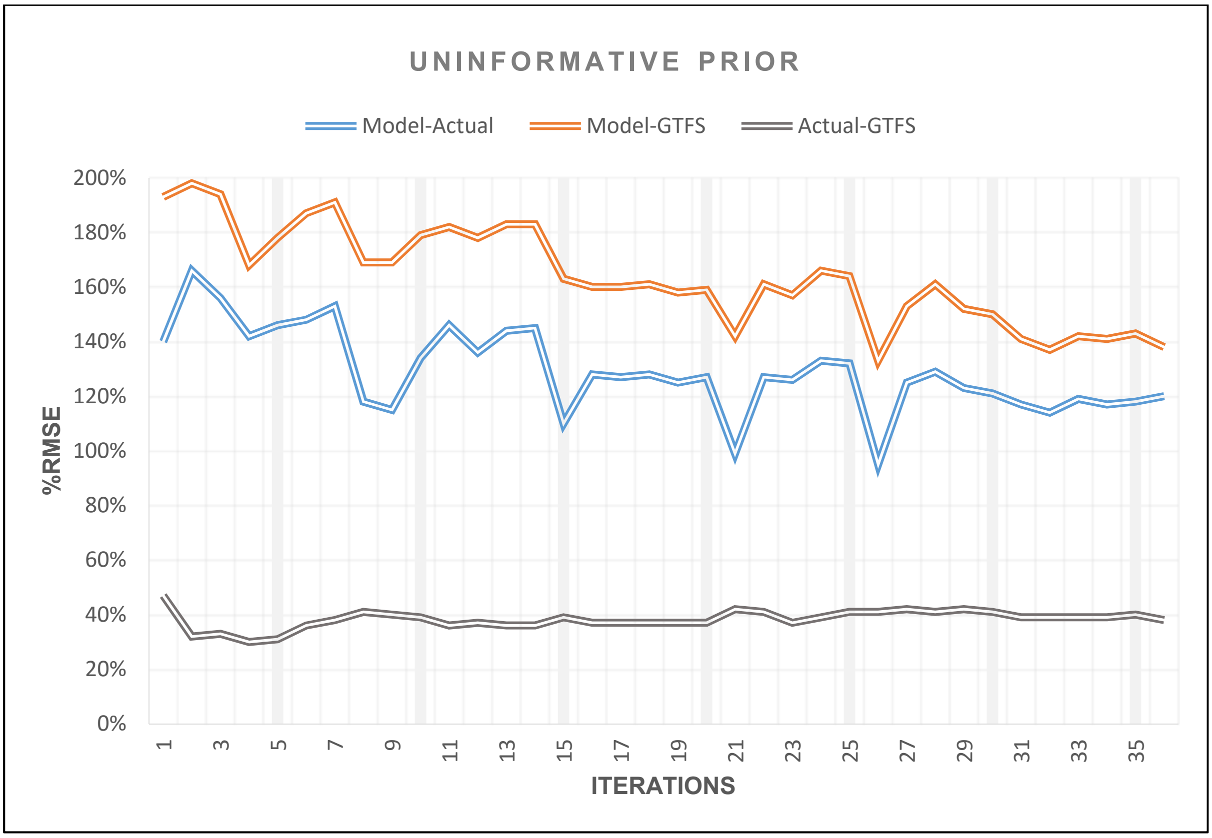
Table 4 presents the %RMSE between the estimated travel time from the Bayesian model and the actual data and GTFS data, for different sample sizes as well as different priors. For example, iterations number six to ten belong to 5% sample size. As can be seen, the best results belong to the model with the overly informative prior. In this case, the sample size does not have a strong influence on the posterior information. On the other hand, the worst results belong to the model with the uninformative prior. However, by increasing the sample size, the %RMSE gets smaller, although this decrease is not significant. For the model with informative prior, both the prior and the sample size have a significant effect on the posterior. This means that both the data and the prior inform us what values of estimated parameters are more plausible. The final column shows the %RMSE between the actual data and GTFS data. The presented values in the last column provide a base for comparing the results of three approaches.

Table 4: The %RMSE between the estimated travel time, actual data, and GTFS data

Sample size (%)	Iteration	Uninformative		Informative		Overly informative		Actual & GTFS
		Actual	GTFS	Actual	GTFS	Actual	GTFS	
1	1	140	193	66	101	35	26	47
	2	166	198	78	96	25	15	32
	3	156	194	65	88	29	33	33
	4	142	168	74	97	27	30	30
	5	146	178	77	104	24	30	31
5	6	148	187	62	87	28	40	36
	7	153	191	61	86	30	34	38
	8	118	169	62	87	30	38	41
	9	115	169	63	98	28	39	40
	10	134	179	63	98	29	38	39
10	11	146	182	60	90	27	31	36
	12	136	178	60	91	26	33	37
	13	144	183	59	94	26	33	36
	14	145	183	61	85	26	31	36
	15	110	163	59	94	25	37	39
30	16	128	160	38	46	26	30	37
	17	127	160	41	53	26	31	37
	18	128	161	41	53	26	30	37
	19	125	158	42	54	27	31	37
	20	127	159	39	46	26	30	37
40	21	99	142	41	53	29	37	42
	22	127	161	39	49	29	33	41
	23	126	157	41	54	26	31	37
	24	133	166	40	50	28	33	39
	25	132	164	38	47	25	30	41
50	26	95	133	38	47	29	36	41
	27	125	153	38	47	30	33	42
	28	129	161	37	46	29	33	41
	29	123	152	40	49	30	35	42
	30	121	150	37	47	29	33	41
70	31	117	141	35	44	29	32	39
	32	114	137	37	46	29	32	39
	33	119	142	36	44	29	31	39
	34	117	141	35	43	29	32	39
	35	118	143	36	45	29	33	40
100	36	120	138	32	39	24	34	38

The errors for the sample sizes, three priors, and five iterations are also presented in Figures 3-6. All iterations of each sample size are separated from other sample sizes by a solid gray vertical line. As can be seen in Figure 3, although the error between the model and actual data decreases with an increase in the sample size, even by using 100% sample size, the uninformative prior is misleading the posterior since the gap between %RMSE of 'Model-Actual' and 'Actual-GTFS' is very large. The prior represents information on the link travel time, while the total travel time, the most important attribute in path choice (Jánošíková et al., 2014), is built from that. It may be concluded that, in the presence of an uninformative prior, even a large set of data does not guarantee a reasonable posterior for the principal variables of this model. Therefore, ESS identification, using this amount of data associated with the uninformative prior, is not possible.

Figure 3: The %RMSE for the model with uninformative prior



In Figure 4, the errors are higher at low sample size and reduce as the sample size increases in the both comparisons of the model with actual data and with GTFS. Comparing Figures 3 and Figure 4 demonstrates that the prior's role is substantial; however, the posterior is influenced by the prior as well as the data. Interestingly, the convergence in Figure 4 is happening with the 30% sample size (from iteration sixteen), and the %RMSE between the model and actual data is even lower than the %RMSE of the actual data and GTFS using a 40% sample size (from iteration twenty four). As can be seen from the graph, the ESS is about 40%-50% sample size.

Figure 4: The %RMSE for the model with informative prior

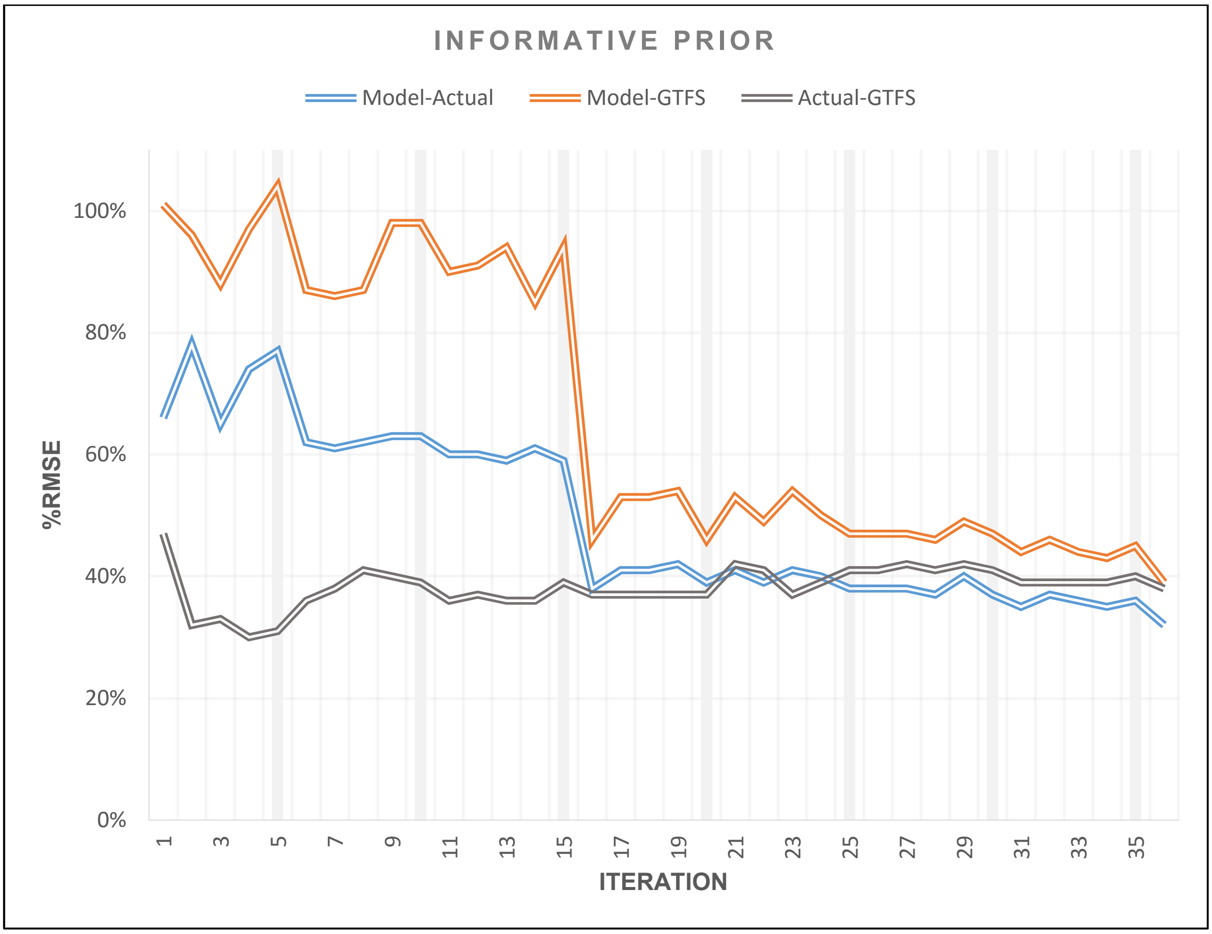


Figure 5, which shows the average of % RMSE of each sample size from Figure 4, confirms that the ESS is about 50% sample size. This figure illustrates that by using 50% sample size, the %RMSE of Model-Actual is less than those of Actual-GTFS and Model-GTFS.

Figure 5: The %RMSE average for the model with informative prior

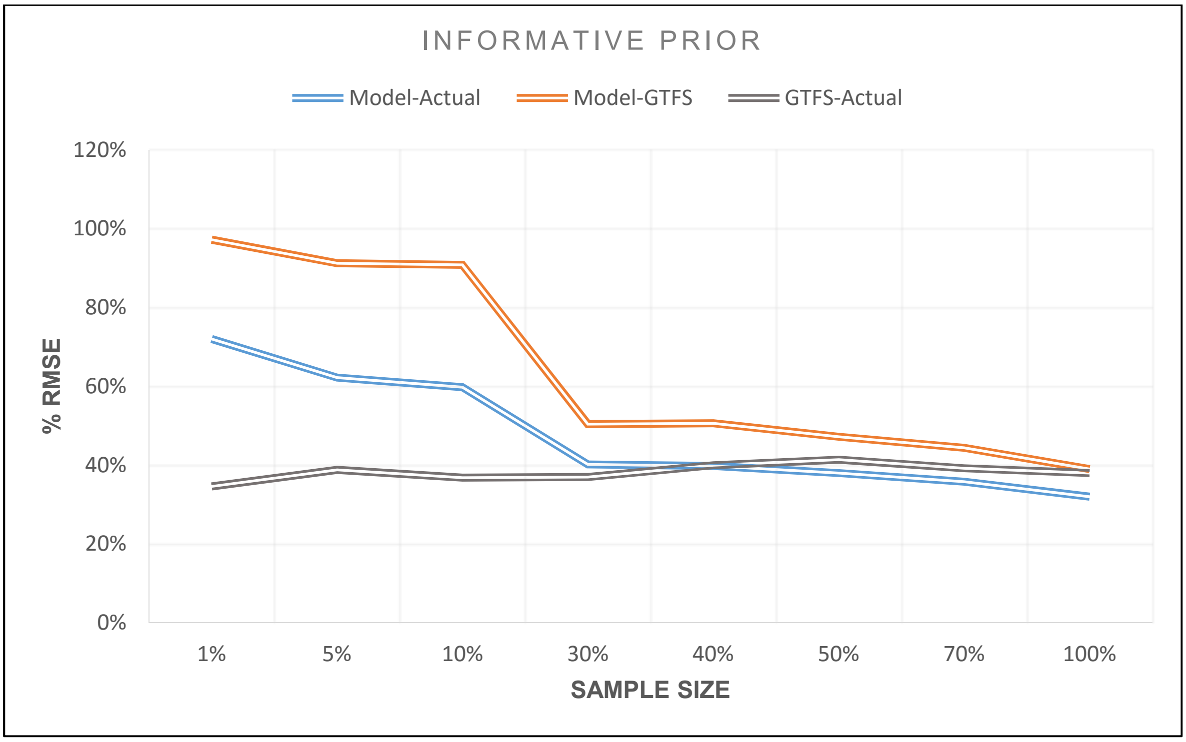
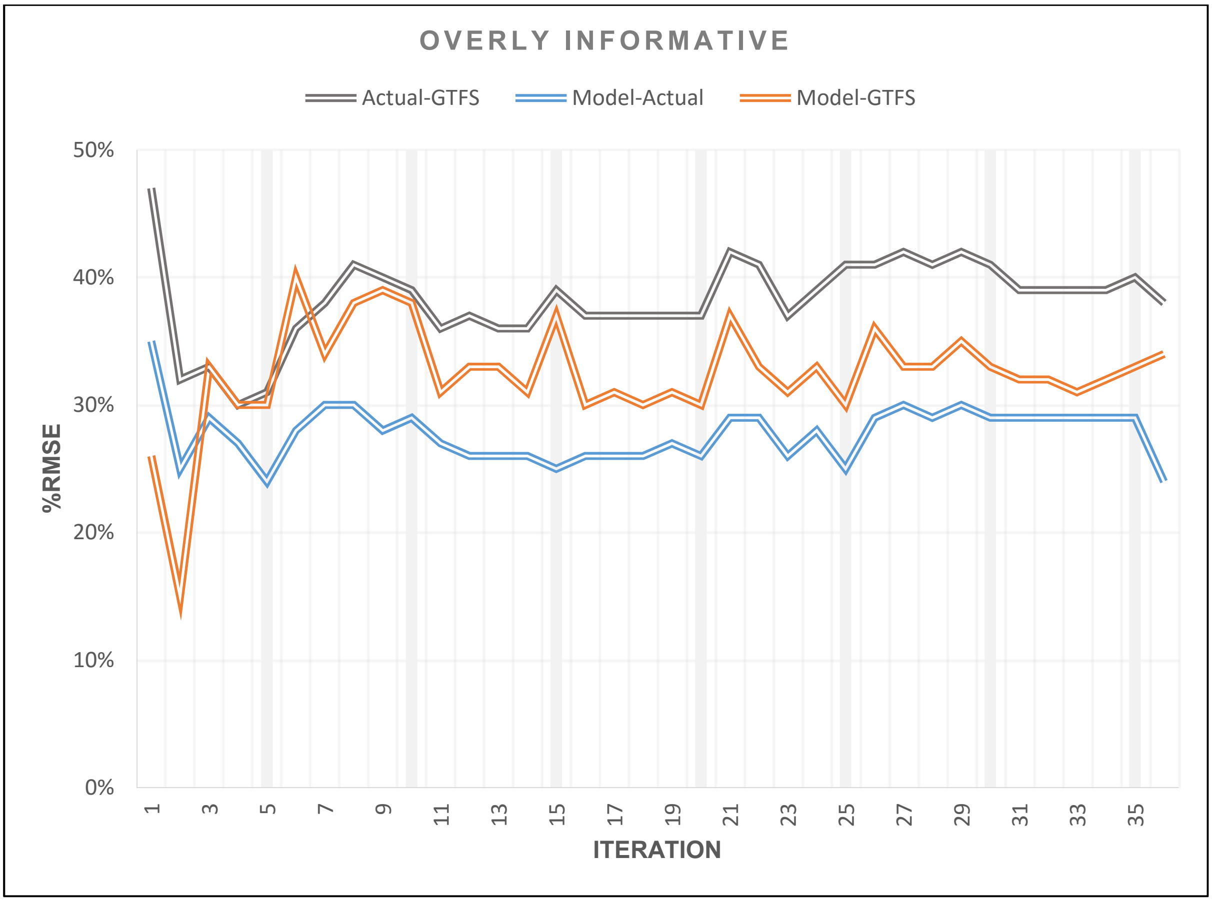


Figure 6 illustrates that the posterior is strongly influenced by the prior. In this case, even with the smallest sample size, the %RMSE of the estimated values are low. A possible explanation for this might be that the GTFS, as the utilized prior, is the transit schedule based on which the transit services should operate. Hence, the actual travel times are very close to this overly informed prior. In summary, comparing Figures 4 and Figure 5 demonstrates that the prior's role is substantial. In the case of overly informed prior, the posterior is mainly influenced by the prior and not much by the data. Stability is more visible by using 10% and larger sample sizes. Therefore, the ESS in this case would be a 10% sample size.

Figure 6: The %RMSE for the model with overly informative prior



5. Summary

Data collection is one essential prerequisite in analyzing the passengers' path choice behaviour. The main challenge faced by many researches is that how much data is required to get a certain level of accuracy? In this paper, we conducted an approach for computing the ESS in a variety of sample size and prior information levels. First, we selected the path choice model used by Rahbar et al. (2017). That study, taking advantage of high-quality travel data provided by smart card data, presented a transit assignment framework using a Bayesian inference approach. Second, we chose different sample sizes from one weekday data of the SEQ transit network including bus, train, and ferry modes. At each sample size, five iterations were generated to overcome the bias of the randomly selected sample size. Third, priors were generated using three definitions: 1) uninformative, 2) informative, and 3) overly informative. The combination of each sample size with each prior gave a particular scenario. Finally, the results of different scenarios are compared with the actual data and GTFS data. Detecting the ESS for the model associated with an uninformative prior is not possible. For the model with an informative prior, the ESS is a 50% sample (12,450 OD pairs and 39,270 journeys). Lastly, for the model with an overly informative prior, the ESS is a 10% sample (2,490 OD pairs and 7,270 journeys). The results highlight the effect of sample size and prior information on the accuracy level of the estimated path choice parameters (link travel time in this study).

Further research could explore the effect of sample size on the accuracy level of other parameters of the path choice model (Θ_1 and Θ_2). In addition, other distributions for priors, rather than a normal distribution, can be examined.

Acknowledgements

The authors would like to acknowledge TransLink (the public transport authority of Queensland, Australia) for providing the data for this research. Also, we would like to acknowledge the Australian Government Research Training Program for offering the scholarship.

References

BERGER, J. O., MORENO, E., PERICCHI, L. R., BAYARRI, M. J., BERNARDO, J. M., CANO, J. A., DE LA HORRA, J., MARTÍN, J., RÍOS-INSÚA, D. & BETRÒ, B. 1994. An overview of robust Bayesian analysis. *Test,* 3**,** 5-124.

BUREAU OF METEOROLOGY. 2017. *Climate Data Online. Bureau of Meteorology* [Online]. Available: http://www.bom.gov.au/qld/ [Accessed 9/20/2017].

CONGDON, P. 2001. *Bayesian statistical modelling,* Chichester;New York;, Wiley.

FU, Q., LIU, R. & HESS, S. A Bayesian modelling framework for individual passenger's probabilistic route choices: a case study on the London underground. Transportation Research Board 93rd Annual Meeting, 2014.

GALBRAITH, R. A. & HENSHER, D. A. 1982. Intra-metropolitan transferability of mode choice models. *Journal of Transport Economics and Policy***,** 7-29.

GOOGLE INC. 2013. General Transit Feed Specification.

HAZELTON, M. L. 2008. Statistical inference for time varying origin–destination matrices. *Transportation Research Part B,* 42**,** 542-552.

HAZELTON, M. L. 2010. Bayesian inference for network-based models with a linear inverse structure. *Transportation Research Part B,* 44**,** 674-685.

HE, R. R., LIU, H. X., KORNHAUSER, A. L. & RAN, B. 2002. *Temporal and spatial variability of travel time*, Center for Traffic Simulation Studies.

HOFLEITNER, A., HERRING, R. & BAYEN, A. Probability distributions of travel times on arterial networks: a traffic flow and horizontal queuing theory approach. 91st Transportation Research Board Annual Meeting, 2012.

JÁNOŠÍKOVÁ, Ľ., SLAVÍK, J. & KOHÁNI, M. 2014. Estimation of a route choice model for urban public transport using smart card data. *Transportation planning and technology,* 37**,** 638-648.

KARASMAA, N. 2007. Evaluation of transfer methods for spatial travel demand models. *Transportation Research Part A: Policy and Practice,* 41**,** 411-427.

KARASMAA, N. & PURSULA, M. 1997. Empirical studies of transferability of Helsinki metropolitan area travel forecasting models. *Transportation Research Record: Journal of the Transportation Research Board***,** 38-44.

LEE, M. & SOHN, K. 2015. Inferring the route-use patterns of metro passengers based only on travel-time data within a Bayesian framework using a reversible-jump Markov chain Monte Carlo (MCMC) simulation. *Transportation Research Part B: Methodological,* 81, Part 1**,** 1-17.

LI, B. 2005. Bayesian Inference for Origin-Destination Matrices of Transport Networks Using the EM Algorithm. *Technometrics,* 47**,** 399-408.

LI, B. 2009. Markov models for Bayesian analysis about transit route origin–destination matrices. *Transportation Research Part B: Methodological,* 43**,** 301-310.

LI, R., CHAI, H. & TANG, J. 2013. Empirical Study of Travel Time Estimation and Reliability. *Mathematical Problems in Engineering,* 2013.

MAHER, M. J. 1983. Inferences on Trip Matrices from Observations on Link Volumes: A Bayesian Statistical Approach. *Transportation Research,* 17**,** 435-447.

MEI, B., COONEY, T. A. & BOSTROM, N. R. 2005. Using Bayesian Updating to Enhance 2001 NHTS Kentucky Sample Data for Travel Demand Modeling. *Journal of Transportation and Statistics,* 8**,** 71.

MOLLA, M. M. I. 2017. *A stochastic Bayesian update and logistic growth mapping of travel-time flow relationship.* North Dakota State University.

MORITA, S., THALL, P. F. & MÜLLER, P. 2008. Determining the effective sample size of a parametric prior. *Biometrics,* 64**,** 595-602.

OH, M.-S. & RAFTERY, A. E. 2012. Model-based clustering with dissimilarities: A Bayesian approach. *Journal of Computational and Graphical Statistics*.

PERRAKIS, K., KARLIS, D., COOLS, M. & JANSSENS, D. 2015. Bayesian inference for transportation origin-destination matrices: the Poisson–inverse Gaussian and other Poisson mixtures. *Journal of the Royal Statistical Society: Series A (Statistics in Society),* 178**,** 271-296.

PERRAKIS, K., KARLIS, D., COOLS, M., JANSSENS, D., VANHOOF, K. & WETS, G. 2012. A Bayesian approach for modeling origin–destination matrices. *Transportation Research Part A: Policy and Practice,* 46**,** 200-212.

R CORE TEAM 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

RAHBAR, M., MESBAH, M., HICKMAN, M. & TAVASSOLI, A. 2016. Determining route choice behaviour of public transport passengers using Bayesian statistical inference. *Australasian Transport Research Forum (ATRF), 38th.* Victoria, Melbourne, Australia.

RAHBAR, M., MESBAH, M., HICKMAN, M. & TAVASSOLI, A. 2017. Determining route-choice behaviour of public transport passengers using Bayesian statistical inference. *Road & Transport Research: A Journal of Australian and New Zealand Research and Practice,* 26**,** 64-72.

RASHIDI, T., AULD, J. & MOHAMMADIAN, A. 2013. Effectiveness of bayesian updating attributes in data transferability applications. *Transportation Research Record: Journal of the Transportation Research Board***,** 1-9.

SANTOSO, D. & TSUNOKAWA, K. 2005. Spatial transferability and updating analysis of mode choice models in developing countries. *Transportation Planning and Technology* 28(5)**,** 341-355.

SANTOSO, D. S. & TSUNOKAWA, K. 2010. Comparison of updating techniques in transferability analysis of work trip mode choice models in developing countries. *Journal of Advanced Transportation,* 44**,** 89-102.

SUN, L. J., LU, Y., JIN, J. G., LEE, D. H. & AXHAUSEN, K. W. 2015. An integrated Bayesian approach for passenger flow assignment in metro networks. *Transportation Research Part C-Emerging Technologies,* 52**,** 116-131.

TALVITIE, A. & KIRSHNER, D. 1978. Specification, transferability and the effect of data outliers in modeling the choice of mode in urban travel. *Transportation,* 7**,** 311-331.

TAVASSOLI, A., ALSGER, A., HICKMAN, M. & MESBAH, M. How close the models are to the reality? Comparison of Transit Origin-Destination Estimates with Automatic Fare Collection Data. Australasian Transport Research Forum (ATRF), 38th, 2016, Melbourne, Victoria, Australia, 2016.

TRANSLINK. 2017. *About go card* [Online]. Department of Transport and Main Roads. Available: https://translink.com.au/tickets-and-fares/go-card/about-go-card [Accessed 9/20/2017].

WEI, C. & ASAKURA, Y. 2013. A Bayesian approach to traffic estimation in stochastic user equilibrium networks. *Transportation Research Part C-Emerging Technologies,* 36**,** 446-459.

YAMAMOTO, T., MIWA, T., TAKESHITA, T. & MORIKAWA, T. 2009. Updating dynamic origin-destination matrices using observed link travel speed by probe vehicles. *Transportation and Traffic Theory 2009: Golden Jubilee.* Springer.