

Analysis of Uncertainty in Performance Indicators of a Complex Activity-Based Model: The Case of the Albatross Model System

Soora Rasouli, Theo Arentze and Harry Timmermans

*Eindhoven University of Technology
Urban Planning Group
P.O. Box 513
5600 MB Eindhoven
The Netherlands
Email: s.rasouli@tue.nl*

Abstract. This paper reports the main results of an uncertainty analysis conducted for the Albatross model in the Rotterdam region. The analysis focuses on the number of runs required to achieve stable coefficients of variation for different mobility indicators and the associated confidence intervals. Results indicate that few runs are required to obtain stable results. To achieve conventional confidence intervals substantially higher numbers of runs seem required.

1. MOTIVATION

In reporting the results of models of transportation demand, the assessment of uncertainty has received less attention than the consideration of goodness-of-fit and the interpretation of the sign and significance of estimated parameters. It is to be expected, however, that uncertainty analysis will play a more important role in the near future due to some current trends. First, the increased costs of data collection and reducing response rates in travel surveys have stimulated research on data imputation, increasing input uncertainty. Second, increasing political polarization that can be witnessed in several Western societies will also affect discussions on transportation policies. The value (and uncertainty) of travel forecasts has been increasingly scrutinized in this process and this trend will likely continue. Third, the shift from four-step models to activity-based models has implied a shift from deterministic to inherently probabilistic models, demanding an analysis of model uncertainty to differentiate policy effects from stochastic variability of the model itself.

Uncertainty in forecasting error can be attributed to two basic sources: input uncertainty and model uncertainty. Input uncertainty is concerned with the effects of changes in uncertain input data, due to measurement error or to scenario uncertainty, on the ultimate forecasts of the model. In contrast, model uncertainty consists of two types of error: specification error and estimation error. Specification errors result from a failure of the researcher to identify the true model, a simplification of the model or the statistical distribution of random components. Estimation error involves error in estimating the values of various constants and parameters in the model structure

This paper is concerned with model uncertainty. The aim of the analyses is to assess aspects of model uncertainty associated with the Albatross model system. In the present paper, the discussion is confined to the analysis of uncertainty as it relates to a set of aggregate performance indicators, generated by the Albatross model system. The following research questions guided the analyses: (i) How does model uncertainty vary with an increasing number of model runs?; (ii) How many runs are required to achieve the commonly used 95% confidence levels of certain performance indicators at the level of the synthetic population?

The remainder of this paper is organized as follows. First, in the next section, we will briefly summarize previous research on uncertainty analysis. This is followed by a summary of the Albatross model. The next two sections describe the analysis and results of the uncertainty analysis respectively. Finally, concluding remarks are given.

2. LITERATURE REVIEW

An examination of the rather scarce literature suggests that uncertainty analysis has been mainly restricted to classic 4 step models and discrete choice models. Researchers such as Ben-Akiva and Lerman (1985), De Jong *et al.* (1998) and Yang and Chen (2011) developed an analytic expression for the variance of outputs resulting from the variance of parameters. In case of simple models, these variances can be calculated in a straightforward manner. If the model is more complicated, t-ratios and standard errors of model parameters can be obtained by using Jackknifing or bootstrapping methods (e.g., De Jong *et al.* 1998, Armoogum 2003, Beser Hugosson 2005) or by drawing randomly from input and/or model distributions (e.g., Zhao and Kockelman 2001, Boyce and Bright 2003).

Castiglione (2003) examined the impact of the number of runs and other operational decisions on the convergence of model output statistics such as the mean number of trips. Similarly, Ziemis *et al.* (2011) ran 20 times two trip-based sub-models (trip routing and traffic simulation) of Transims and investigated differences between predicted traffic volumes and the final mean traffic volume on two roads. These studies found that the stochastic variability is quite small and that a relatively small number of runs is required to get stable average performance indicators. However, they suggested that further research should attempt to examine stochastic variability of alternative micro-simulation model systems as it is not immediately clear whether their results generalize to such model systems.

This paper picks up this suggestion by examining uncertainty in a computational process model of activity-travel behavior. It replicates and extends work by Cools *et al.* (2011), who analyzed the effects of some socio-economic variables on uncertainty of two travel indices (distance traveled and number of trips), generated by the Feathers model: the Flanders equivalent of Albatross (Janssen *et al.*, 2007). Their results indicated that micro-simulation error increases with model complexity, expressed as the degree of disaggregation of the model.

3. ALBATROSS

Albatross is a rule-based system of activity-travel behavior, developed for the Dutch Ministry of Transportation (Arentze and Timmermans, 2004). It is based on the theory that human decision making, in case of repetitive behavior and large solution spaces, relies on heuristics rather than on exhaustive evaluation of solutions. As Figure 1 shows, Albatross uses a priority-based scheduling process in which mandatory activities are scheduled first and discretionary activities next. The rules underlying the model are empirically derived from activity-travel diary data using a CHAID-based induction method.

An important feature of the model is that generated schedules fully meet space-time constraints. For each scheduling decision, Albatross delineates the choice set/range, given all previous decisions, travel times and opening hours/availability of facilities. Consistently, the model determines the maximally available time window for an activity/trip across possible next choices. Having derived a decision tree for each choice facet, Albatross uses these trees for prediction. Consider a response variable that has Q levels and a decision tree with K leaf nodes.

In the prediction stage, the tree is used to classify new cases to one of the K leaf nodes based on attributes of the case. To that end, a response-assignment rule is specified that defines a response (decision) for each classified case. In many applications, deterministic rules such as a plurality rule are used. Such deterministic rules may yield the best predictions at the individual level, but fail to reproduce residual variance (if any) at leaf nodes in predictions. Therefore, a probabilistic assignment rule is used in Albatross. According to this rule, the probability of selecting the q -th response for each new case assigned to the k -th node is simply given by:

$$P_{kq} = \frac{f_{kq}}{N_k} \quad (1)$$

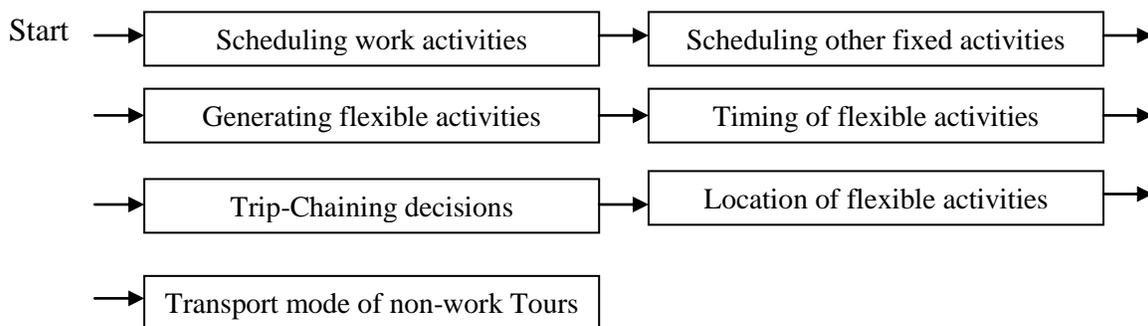


Figure 1: Main steps in the Albatross scheduling process

where f_{kq} is the number of cases of category q at leaf node k and N_k is the total number of cases at that node. This rule is sensitive to residual variance, but fails to take scheduling constraints into account. If such constraints are represented in the decision tree, the probabilistic rule would assign zero probability to infeasible categories. Therefore, to cover the general case, Albatross uses a refinement of the previous rule:

$$P_{kq} = \begin{cases} 0 & \text{if } q \text{ is infeasible} \\ f_{kq} / \sum_{q'} f_{kq'} & \text{otherwise} \end{cases} \quad (2)$$

where, q' is an index of feasible alternatives for the decision at hand.

4. ANALYSIS

The analysis of uncertainty of the Albatross model system was conducted for the city of Rotterdam, the second largest city in the Netherlands with a population of approximately 600,000 inhabitants in 2005. The analysis was conducted according to the following steps. First, a synthetic population of the city was created, using the approach described in Arentze and Timmermans (2008). Like many other population synthesizers, this approach uses iterative proportional fitting, but differs in that the consistency between individual and households is maintained by using relation matrices. Next, a 10% fraction of the synthetic population, consisting of 41,668 persons and 27,961 households, was randomly selected. To rule out the possibility that results are influenced by the sampled fraction, it was kept constant for all analyses. Then, for each sampled individual of this fraction, the Albatross model was run up to a

maximum of 100 times. In each run, the action state of each decision tree of the model was selected with a certain probability using Monte Carlo draws. These runs result in a probability distribution of each facet of the simulated activity-travel patterns and the associated performance indicators. Running the model multiple times allows one to analyze the effects of the number of runs on the uncertainty of the performance indicators. Uncertainty was measured in terms of the coefficient of variation:

$$CV = \frac{Sd}{Mean} \times 100 \quad (3)$$

The coefficients of variation were calculated for the following performance indicators (i) distance travelled per day per person for various activities, (ii) distance travelled per day per person for different transport modes and (iii) distance travelled per day per person for different start times. Weighted results are reported.

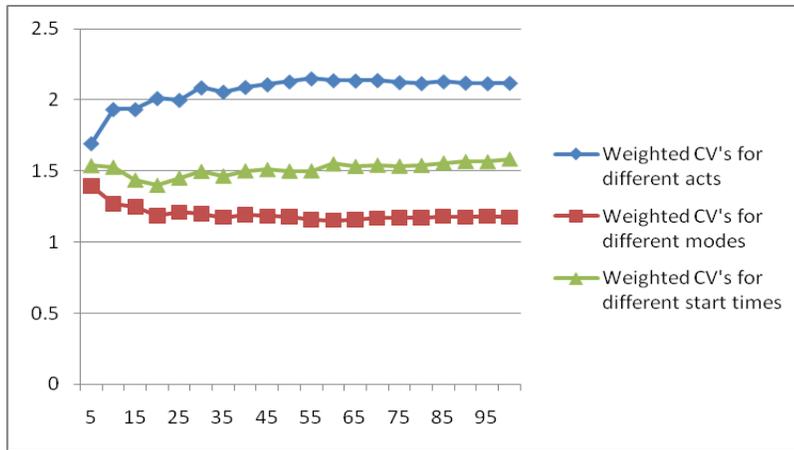


Figure 2: CV of sample vs. number of runs

5. RESULTS

5.1 Model uncertainty for increasing number of model runs

As shown in Figure 2, stochastic variability expressed in terms of CV is relatively small: less than 5% even in the worst case. Figure 2 also shows that the CV fluctuation decreases significantly after 25-30 runs for all performance indicators and further increasing the number of runs does not substantially change the resulting CV.

5.2 Confidence intervals for different number of runs

Table 1 lists the estimated confidence intervals for the coefficient of variation for a selected set of performance indicators, using McKay's (1932) and Miller's (1991) approximations. The confidence intervals are based on the common 95% probability level. It shows that the confidence intervals decrease with an increasing number of runs. They do not differ that much between the three performance indicators. At the maximum number of 100 runs, the confidence intervals are around 28 per cent. It implies that many more runs are required to achieve conventional levels of accuracy.

Table 1: Confidence intervals of a set of performance indicators for different number of runs

	No of runs	McKay				Miller			
		Lower boundary	Upper boundary	Diff lower from mean(%)	diff upper from mean(%)	Lower boundary	Upper boundary	diff lower from mean(%)	diff upper from mean(%)
Work	100	0.0138	0.0183	12.2060	16.1475	0.0135	0.0179	14.0967	13.7153
	50	0.0132	0.0197	16.4245	24.6368	0.0127	0.0189	19.9935	19.6144
	5	0.0116	0.0558	40.1238	187.5193	0.0059	0.0328	69.3383	69.2352
CarD	100	0.0092	0.0122	12.1997	16.1905	0.0090	0.0120	14.0818	14.0818
	50	0.0086	0.0128	16.4676	24.6170	0.0085	0.0125	17.8593	21.7428
	5	0.0094	0.0452	40.0892	187.7446	0.0048	0.0266	69.3135	69.3135
3am-10am	100	0.0122	0.0161	12.2022	16.1011	0.0119	0.0158	13.9691	13.8824
	50	0.0105	0.0157	16.6005	24.6227	0.0101	0.0151	19.8021	19.8021
	5	0.0068	0.0326	40.1762	187.2247	0.0035	0.0192	69.3054	69.3054

6. IMPLICATIONS FOR TRAVEL DEMAND FORECASTING PRACTICE

This paper has reported some main results of an uncertainty analysis of the Albatross model. The analyses specifically focused on (1) the relationship between uncertainty and the number of model runs and (ii) the confidence in the predicted uncertainty of selected performance indicators as a function of the number of model runs. To that effect, the Albatross model was run up to a maximum of 100 times

Results indicate that after 25-30 runs, the weighted coefficients of variation for distance traveled for activities, transportation mode and different start times converge to some stable value. In addition, after 100 runs confidence intervals of the coefficients of variation were still substantial.

These results have some implications for travel demand forecasting practice. First, modelers should run the model several times to arrive at a stable value of the stochastic uncertainty of the model. In the present case, the number of runs is 25-35, but other models and/or other study areas may require a different number of runs. Second, approximately 500 runs are required to achieve conventional confidence intervals of the coefficients of variation.

It would be interesting to compare the results of the present study with those obtained for other types of models. It should be mentioned however that it is impractical at this stage since the limited number of relevant studies have mostly considered other performance indicators. For example, Ziems et al. (2011) investigated the effect of model uncertainty on volume and travel time through links, while Castiglione (2003) investigated uncertainty of the San Francisco model in terms of the number of tours. In addition, it should be realized that a valid comparison of different models ideally should involve the same study area. Future work should however address systematic comparisons of model uncertainty of the same model across different areas, or of different models in the same area. We plan to report the results of such comparisons in future publications.

Acknowledgement

This work has been funded by the European Commission, under the Seventh Framework Programme, by Contract no. 248488 within project The Uncertainty Enabled Model Web. The views expressed herein are those of the authors, and are not necessarily those of the European Commission.

References

- Arentze, T.A. and Timmermans, H.J.P. (2004), A learning-based transportation oriented simulation system, *Transportation Research B*, 38, 613-633.
- Arentze, T.A. and Timmermans, H.J.P. (2008), Albatross: Overview of the model, applications and experiences, *Proceedings Innovations in Travel Modeling Conference*, Portland, Oregon.
- Armoogum, J. (2003), Measuring the impact of uncertainty in travel demand modeling with a demographic approach, *European Transport Conference*, Strasbourg, France, 11 pp.
- Ben-Akiva, M.E. and Lerman, S. (1985), *Discrete Choice Analysis*. Cambridge: MIT Press.
- Beser Hugosson, M. (2005), Quantifying uncertainties in a national forecasting model, *Transportation Research A*, 39, 531-547.
- Boyce, A.M. and Bright M.J. (2003), Reducing or managing the forecasting risk in privately-financed projects, *European Transport Conference*, Strasbourg, France, 46 pp.
- Castiglione, J., Freedman, J. and Bradley, M. (2003), Systematic investigation of variability due to random simulation error in an activity-based micro-simulation forecasting model, *Transportation Research Record*, 1831, 76-88.
- Cools, M., Kochan, B., Bellemans, T., Janssens, D. and Wets, G. (2011), Assessment of the effect of micro-simulation error on key travel indices: Evidence from the activity-based model Feathers, to appear in *Transportation Research Record*, 1558.
- Jong, G. de., Ettema, D., Gunn, H., Cheung, F. and Kleijn, F. (1998), Preparatory research into updating values of travel time in The Netherlands, *Proceedings WCTR Conference* Antwerp, Belgium.
- Janssens, D., Wets, G., Timmermans, H.J.P. and Arentze, T.A. (2007), Modelling short-term dynamics in activity-travel patterns: Conceptual framework of the Feathers model, *Proceedings WCTR Conference*, Berkeley, USA.
- McKay, A. T. (1932), Distribution of the coefficient of variation and the extended 't' distribution, *Journal of the Royal Statistical Society*, 95, 695-698.
- Miller, E.G. (1991), Asymptotic test statistics for coefficient of variation, *Communications in Statistical Theory & Methods*, 20, 3351-3363.
- Yang, C. and Chen, A. (2011), Sensitivity-based uncertainty analysis of a combined travel demand model, *Transportation Research Record*, to appear in *Transportation Research Record*, 3560
- Zhao, Y., and Kockelman, K.M. (2001), The propagation of uncertainty through travel demand models: An exploratory analysis, *Annals of Regional Science*, 36, 909-921.
- Ziems, S., Sana, B., Plotz, J. and Pendyala, R.M. (2011), Stochastic variability in microsimulation modeling and convergence of corridor-level characteristics, to appear in *Transportation Research Record*, 3560.