Evaluating the reliability of the Utility-Based Dynamic OD Estimation on Large Networks

Guido Cantelmo[1], Francesco Viti[1]*

[1] Faculty of Science, Technology and Communication University of Luxembourg, Luxembourg
[*]Corresponding author: guido.cantelmo@uni.lu

Extended abstract submitted for presentation at the
7th International Conference on Transport Network Reliability (Sydney, 17-19 January, 2018)

1. Introduction

Simulation of traffic conditions requires accurate knowledge of the travel demand. In a dynamic context, this entails estimating time-dependent demand matrices, which are a discretised representation of the dynamic origin-destination (OD) flows. This problem, referred to as Dynamic OD Estimation (DODE) in literature, seeks for the best possible approximation of OD flows, which minimises the error between simulated and available traffic data. Traditional DODE models solve two optimisation problems, according to a bi-level formulation: the upper level updates the time-dependent OD flows, while in the lower level a dynamic traffic assignment model ensures consistency between demand and supply models.

Since DODE problems are usually underdetermined because of the high number of unknown variables [1], researchers have dealt with the critical issue of decreasing the number of decision variables in order to (i) obtain a smooth approximation of objective function [2] and (ii) to reduce the overall computational time [3]. Additionally, issues have been addressed, among others, to the nonlinear relation between link and demand flows [4], pointing out how having a reliable a-priori knowledge of the demand (seed matrix) is of paramount relevance in order to achieve a satisfactory outcome. Zhou and Mahmassani [5] highlight that, in order to provide a robust and reliable estimation, the demand should be considered as a convolution of three functional components: the "regular pattern", the "structural deviation" and the "random fluctuation". The regular pattern can be considered as the systematic component of the demand, the structural deviation is the influence of the specific conditions for which we are estimating/updating the OD matrix (weather conditions, road works,…) and, finally, the random component takes into account the random fluctuations of the demand. Since having a reliable knowledge of the seed matrix is equivalent to know the systematic component of the demand – or regular pattern - we argue that the overall reliability of the DODE depends on how accurate the knowledge of this component is.

The contribution of this paper is twofold. First, we assess the reliability of a new methodology, which aims at estimating the regular pattern within the OD matrix. In fact, the authors believe that, in order to have a reliable approximation of the systematic component of the demand, trip purpose needs to be explicitly included in the DODE model. Based on the empirical findings presented in [6], [7], the authors already formulated a Utility-Based Dynamic OD Estimation (UB-DODE) model, which has been presented in [8]. Previous studies evaluated mathematical and numerical properties for the model. In this work, the new methodology is tested on the real network of Luxembourg City, generalizing those findings. Second, we propose a modification of the well-established SPSA. By imposing a soft constraint to the research space of the model, we systematically increase the results reliability in terms of how likely we are to estimate the "regular pattern" of the OD matrix.

2. Methodology

While for a detailed overview of the model the interested reader can refer to [8], in this section we provide a general overview. The main difference with respect to the standard DODE formulation is in the lower level. We include within lower level DTA procedure a Departure Time Choice (DTC) model that performs the equilibrium through the utility maximisation theory, as proposed in [9]. The advantage of using this approach is twofold.

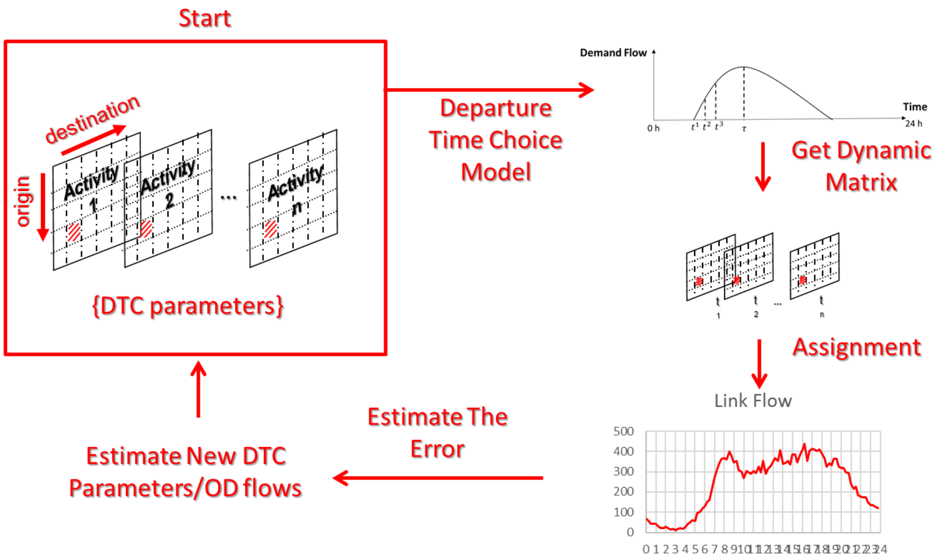

Figure 1: Illustrative representation of the UB-DODE model;

Firstly, this formulation systematically decreases the number of decision variables. The reason is that the parameters of the departure time choice model become the decision variables of the model. As a consequence, the UB-DODE becomes a parametric approach in which, for each OD pair, the model estimates the average departure time and its variance. The second advantage is that the DTC model can include different parameters for different activities, thus it explicitly accounts for the trip purpose within the DODE. Lastly, since each parameter directly affects a large number of time-dependent OD flows, the locality of the optimization problem strongly decreases. Figure 1 shows the main steps for the proposed UB-DODE model.

The proposed methodology can be implemented with most of the existing solution algorithms, including the well-established SPSA. In this paper we proposed to use the C-SPSA (Cluster-SPSA[10]), as it is intuitive to create different clusters for different type of variables (OD flows, preferred departure time,…) when calculating the gradient. Unfortunately, many of the desirable properties of convergence of the SPSA derive from the assumptions that the variables are independent. Clearly, this assumption does not hold for the UB-DODE, as OD and DTC parameters are highly correlated. Therefore, the SPSA has the tendency of exploring unrealistic solutions during the optimization. To avoid this behaviour, we proposed the follow equation to update the solution at each iteration:

$$\boldsymbol{X}_{i+1} = \boldsymbol{X}_i + \alpha \cdot \boldsymbol{G}_i \cdot P(\boldsymbol{X}_i + \alpha \cdot \boldsymbol{G}_i) \quad (1)$$

Where $\boldsymbol{X}_i$ is the vector of the variables to be updated at iteration i, α is the step size and $\boldsymbol{G}_i$ is the gradient. $P(\boldsymbol{X}_i + \alpha \cdot \boldsymbol{G}_i)$ represents the probability that a certain value is realistic for a certain variable. If for instance we are estimating the value for the preferred departure time for commuting to work in the morning, $P(\boldsymbol{X}_i + \alpha \cdot \boldsymbol{G}_i)$ will have a very high value between 6 and 9 am, while will be close to zero for unrealistic values, such as 1 am or 17 pm. This probability, whose parameters are an input for the optimization, acts like a constraint during the optimization, reducing drastically the number of unrealistic solutions.

3. Case study and results

In order to assess the reliability of the UB-DODE, we applied the procedure to the network of Luxembourg City. This network consists of more than 3400 links and 1400 nodes and represents the

typical mid-sized European city. Moreover, Luxembourg City has the typical structure of a metropolitan area, composed of a city centre, ring, and suburb areas.
Lastly, to support the claim that the model is ready for practical implementations, the simulation environment employed is PTV Visum, which is one of the most widely adopted software packages for traffic analysis.

After generating a realistic starting matrix through the conventional 4-Step model, employing available socio-demographic data, we performed three different experiments, using the UB-DODE to estimate the systematic component of the demand. Figure (2) reports some of the results. The first test – Experiment 1 - exploits the standard C-SPSA to estimate the purpose-dependent demand. As shown in Figure (2a), the estimated OD demand does not provide an adequate approximation of the real demand. In the second test – Experiment 2 - we investigated the possibility of using a different set of parameters, which were more likely to provide a realistic result. Figure (2b) shows how this new set of parameters (step size, perturbation, weights within the goal function) leads to an extremely accurate estimation of the demand.

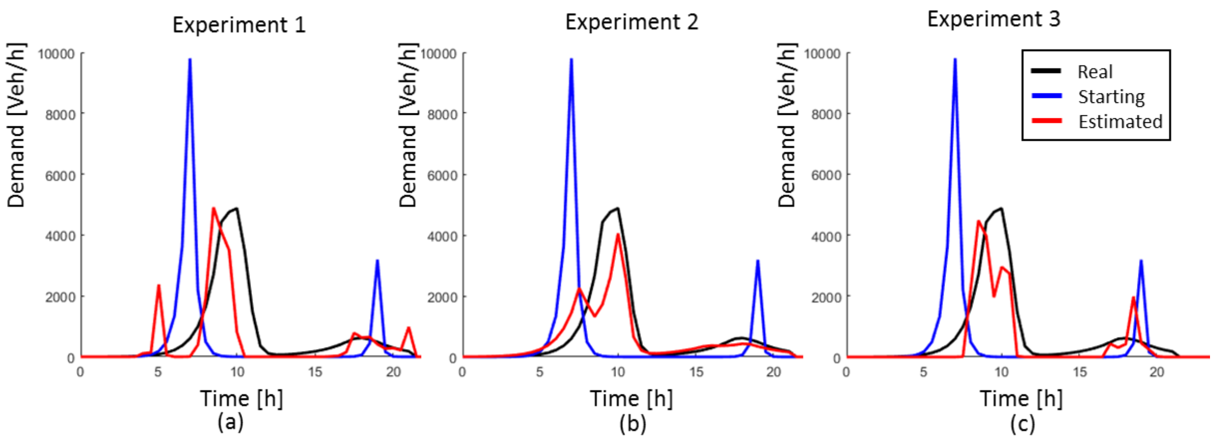

Figure 2: Demand generated from France to Luxembourg according to Experiment 1 (a),2 (b) and 3(c)

However, in this case study the real-demand is known, so it is relatively easy to properly calibrate the parameters of the model in order to improve the performances. Unfortunately, this is not the case for real applications. As a third option, in Experiment 3 we applied the C-SPSA combined with the soft constraint reported in Equation 1. Results show that the constrained C-SPSA achieves a satisfactory estimation of the demand, even with sub optimal set of parameters.

Figure (2) reports an intuitive representation of how accurate the model is in reproducing a realistic demand profile. However, we are interested in the behaviour of the model at a network level. While Experiment 2 seems to outperform its constrained counterpart for that specific traffic zone, we need to analyse if this observation holds at network level. Hence, we calculated the relative improvement in terms of RMSE (Root Mean Squared Error) for the Generated Demand Flows for each traffic zone as:

$$\Delta\_RMSE_{zone} = RMSE_{Zone}^{Starting} - RMSE_{Zone}^{Estimated} \quad (2)$$

The Δ_RMSE_{zone} term represents how close we are to reality in terms of temporal distribution. If Δ_RMSE is negative, it means that the error is higher for the estimated matrix than for the starting demand, if it is equal to zero there is no improvement, while if $\Delta\_RMSE > 0$ we improved the situation with respect to the initial situation. The larger Δ_RMSE is, the bigger is the improvement in the Estimated matrix. Figure (3) report the probability (3a) and cumulative probability (3b) of having a certain Δ_RMSE value for a generic traffic zone according to the three Experiments. Figure 3 clearly shows that, at network level, if the constraint formulation is applied, we systematically improve the quality of the estimated OD matrix with respect to the unconstrained scenario. Even when we adopt a good set of parameters – Experiment 2 – at network level the model is not as good as the constrained one, meaning that for each traffic zone properly calibrated – such as the one in Figure 2a – there is another one with a larger error. By contrast, the constrained formulation provides a more reliable estimation at network level.

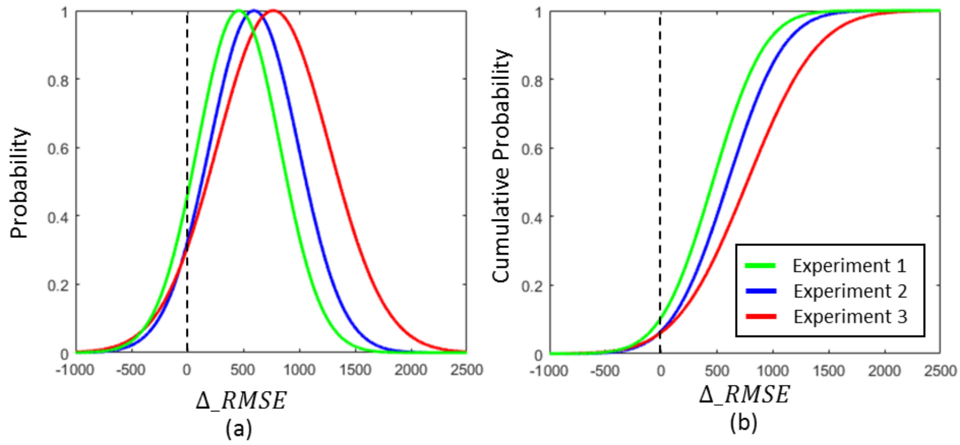

Figure 3: (a) Probability and (b) Cumulative Probability of improving the starting OD matrix for each Experiment

The final goal of this study is to replicate the same experiment under several conditions, in order to verify the robustness of the constrained/ unconstrained UB-DODE approach. This will furthermore allow to assess the conditions under which the proposed model is capable of estimating a reliable demand profile, compatible with existing on-line/off-line DODE models.

References

[1] V. Marzano, A. Papola, and F. Simonelli, “Limits and perspectives of effective O-D matrix correction using traffic counts,” *Transp. Res. Part C Emerg. Technol.*, vol. 17, no. 2, pp. 120–132, 2009.

[2] T. Djukic, G. Flötteröd, L. Van, and S. Hoogendoorn, “Efficient real time OD matrix estimation based on Principal Component Analysis,” presented at the IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, 2012, pp. 115–121.

[3] E. Cipriani, M. Florian, M. Mahut, and M. Nigro, “A gradient approximation approach for adjusting temporal origin-destination matrices,” *Transp. Res. Part C Emerg. Technol.*, vol. 19, no. 2, pp. 270–282, 2011.

[4] X. Zhou, C.-C. Lu, and K. Zhang, “Dynamic origin–destination demand flow estimation under congested traffic conditions,” in *Proceedings of the 91st Transportation Research Board Annual Meeting, Washington, DC*, 2012.

[5] X. Zhou and H. S. Mahmassani, “A structural state space model for real-time traffic origin–destination demand estimation and prediction in a day-to-day learning framework,” *Transp. Res. Part B Methodol.*, vol. 41, no. 8, pp. 823–840, Oct. 2007.

[6] G. Cantelmo and F. Viti, “Activity Demand: An empirical analysis on the influence of the activities on the traffic within-day demand profile,” in *4rd hEART Conference*, 2015.

[7] G. Cantelmo and F. Viti, “Effects of incorporating activity duration and scheduling utility on the equilibrium-based dynamic traffic assignment,” *DTA2016 proceedings (Submitted to Transp. Res. Part B).*, Jun. 2016.

[8] G. Cantelmo, F. Viti, E. Cipriani, and M. Nigro, “A Utility-based Dynamic Demand Estimation Model that Explicitly Accounts for Activity Scheduling and Duration.,” *ISTTT2017 proceedings (Accepted for Special IssueTransp. Res. Series.*, Jun. 2017.

[9] Z.-C. Li, W. H. K. Lam, and S. C. Wong, “Bottleneck model revisited: An activity-based perspective,” *Transp. Res. Part B Methodol.*, vol. 68, pp. 262–287, Oct. 2014.

[10] A. Tympakianaki, H. N. Koutsopoulos, and E. Jenelius, “c-SPSA: Cluster-wise simultaneous perturbation stochastic approximation algorithm and its application to dynamic origin–destination matrix estimation,” *Transp. Res. Part C Emerg. Technol.*, vol. 55, pp. 231–245, Jun. 2015.