EJTIR

Stability analysis of activity-based models: case study of the Tel Aviv transportation model

Shlomo Bekhor[1]

Faculty of Civil and Environmental Engineering, Technion – Israel Institute of Technology, Israel

Leonid Kheifits[2]

Independent Consultant, Israel

Michael Sorani[3]

A.B. Plan Ltd., Israel

The Tel Aviv activity based model structure is similar to other activity based models described in the literature. The model run is supposed to converge to the equilibrium between generated tours and corresponding level of service (LOS) data. However, individual tour generation uses random draws for various choices (activity, time of day, destination, and mode). This introduces simulation errors, which combined with population sampling and limited precision of static traffic assignments, prevents the convergence of the model results.

This paper analyses the above uncertainty sources on the basis of multiple model runs conducted for this study. Three averaging procedures are investigated and compared. Practical considerations regarding setting up the averaging procedures required for obtaining stable model results are discussed.

Keywords: activity-based models, population generator, stability analysis, traffic assignment.

1. Introduction

Activity-Based Models (ABM) have been implemented in various metropolitan areas. Although these models differ in their structure, input data and aggregation level, they share a general common structure: ABM are disaggregate models which simulate the decision processes of individuals listed in randomly generated synthetic population as random draws from choice sets (activity choice, destination choice, mode choice, and so on). Therefore, the outcome of a model run always includes a random component, and special measures are required to mitigate the random effects and to assure reproducibility of the model results.

Despite the practical importance of the subject, the available literature on stability of ABM results is far from being conclusive. Recently, Rasouli and Timmermans (2012) presented a review of uncertainty in travel demand forecasting models. They note that the literature on this topic is relatively scarce and there is a lack of systematic research. Most of the literature on the subject is related to traditional models. For example, Zhao and Kockelman (2002) investigated the stability of a traditional four-step travel demand model by simulating the propagation of uncertainty in a 25-zone network. The results indicated that the average uncertainty increases in the first three steps of the forecasting process (trip generation, trip distribution and mode choice), while the final traffic assignment step decreases average uncertainty.

[1] Technion City, Haifa, 32000, Israel T: +972 4 829 2460 F: +972 4 822 5716 E: sbekhor@technion.ac.il
[2] Gabizon 7, Tel Aviv, 67296, Israel T: +972 3 631 6067 E: kheifits@netvision.net.il
[3] Hamesila 22, Nesher, Israel T: +972 4 820 2055, F: +972 4 820 2053 E: abplan1@netvision.net.il

Specifically with respect to activity-based models, there are some attempts to quantify the uncertainty related to both simulation error and sampling error. Veldhuisen et al. (2000) examined the effect of Monte Carlo draws on regional aggregate activity patterns using the Ramblas micro-simulation model. The results of analyses showed that the effects of Monte Carlo error are negligible which can be expected given the level of aggregation of the results considered.

Castiglione et al. (2003) investigated the variability of the forecasts due to random simulation error, using the San Francisco ABM. The model system was run 100 times, each time changing only the sequence of random numbers used to simulate individual choices from the logit model probabilities. The paper showed that the mean values of key output variables obtained as an average for multiple simulation runs converge toward a stable value as the number of runs increases.

Bowman et al. (2006) described the Sacramento activity based model (SACSIM) and techniques to establish convergence. The technique is based on a modification of the classical method of successive averages (MSA) as follows. At each iteration, the ABM builds a partial matrix based on a fraction of the population. This matrix is then expanded and assigned to the network. Link volumes from the previous iteration and the current assignment are then averaged and travel times are updated. In the next iteration, another fraction of the population is used to generate a new partial demand matrix. The paper analyses the implementation of this method for different values of the procedure parameters.

Vovsha et al. (2008) present a discussion of practical ways to reach equilibrium within an activity-based model, using the New York City ABM. The paper described different methods for stabilizing the ABM results, including several averaging techniques.

Cools et al. (2011) investigated uncertainty related to the statistical distributions of random components, i.e. micro-simulation error. The assessment of the error was performed by systematically running the FEATHERS activity-based model using 10% population samples. Their conclusion is that the customary use of 10% sample for speeding the model runs may be sufficient for overall policy recommendations, but not for more refined decisions.

The studies discussed above considered different aspects of the ABM stability and demonstrated quite different approaches. However, practical aspects of the effects of population sampling and averaging procedures are still not fully explored. The present study is intended to add additional insights to these themes, related to practical needs that were revealed after two years of operating the Tel Aviv ABM.

Similar to other activity-based models, The Tel Aviv ABM system comprises a hierarchy of logit and nested logit models for the main stop of the tour - namely, activity type, time of day, destination, and mode (Cambridge Systematics, 2008). In most activity-based models, the "raw" output of the demand sub-model is the individual's list of activities and trips that include detailed information about departure time, destination and mode for each trip. This detailed output is aggregated into origin-destination (OD) matrices needed for the highway and transit assignments. The assignment outputs provide level of service (LOS) data, which in turn are used as inputs for the next iteration of the ABM run.

Experience of working with Tel Aviv ABM has revealed the need of a comprehensive analysis of uncertainty sources and in developing procedures for monitoring model stability. Practical importance of the above goals relates to the fact that the use of ABM for evaluation of development projects is ultimately based on the comparison of scenarios (e.g. "with" versus "without" project); and any random component of the difference between model results for alternative scenarios should be controlled.

This paper investigates the stability of activity-based model using the results of numerous runs of Tel Aviv ABM, thus continuing the line of research presented in Castiglione et al. (2003),

Bowman et al. (2006), Vovsha et al. (2008), Cools et al. (2011), and outlined in Rasouli and Timmermans (2012). The focus of the paper is on the random components associated with the model run itself, i.e., input errors are not discussed in this paper.

The rest of this paper is arranged as follows: the next section briefly describes the Tel Aviv Activity Based Model. Then, primary sources of uncertainty in the model results are identified. The subsequent sections discuss each of these sources, and the final section summarizes major findings and outlines directions for further research.

2. The Tel Aviv Model

The Tel Aviv metropolitan area (about 1,500 square kilometres) is the largest urban concentration in Israel, with a population of 3.3 million habitants in 2009. The area is divided into 1,219 traffic analysis zones (TAZ), and the planning network contains about 10,000 links and 1,000 transit lines. Figure 1 shows a map of the metropolitan area and the main cities.
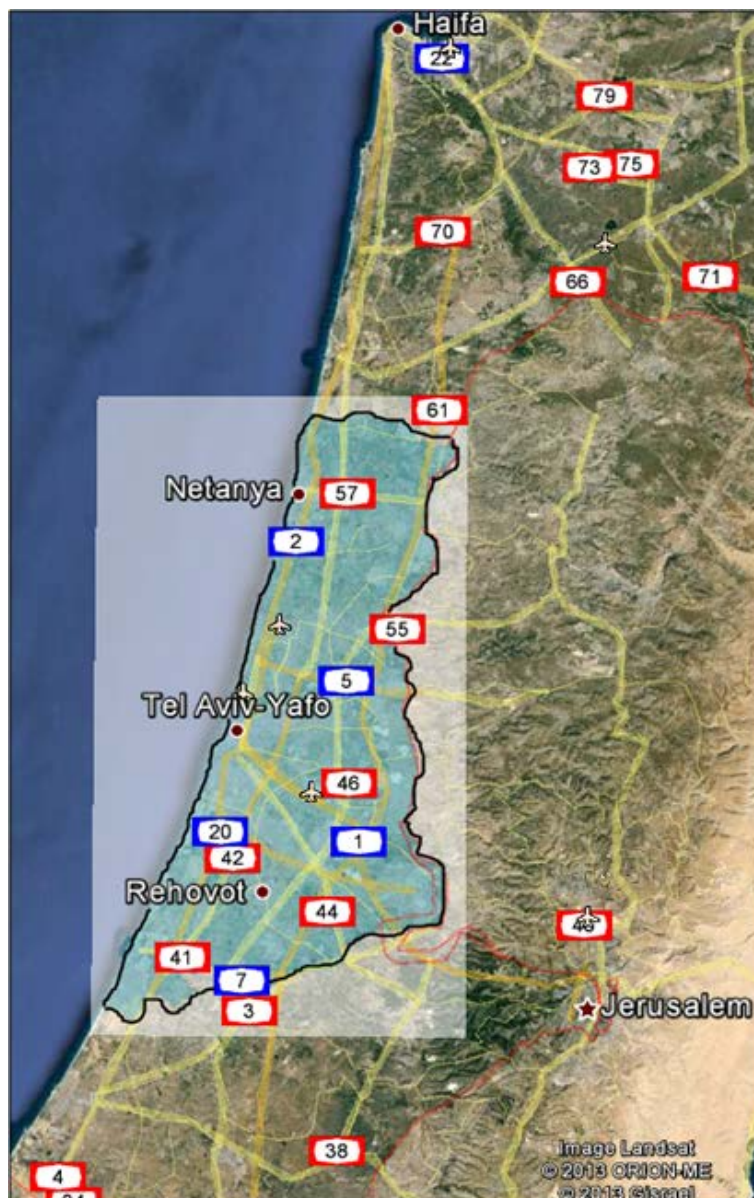

Figure 1.   Tel Aviv Metropolitan Area

The Tel Aviv ABM system comprises a hierarchy of logit and nested logit models for the main stop of the tour - namely, activity type, time of day, destination, and mode. The intermediate stops are then modelled conditional on the main stop models. The Israel National Travel Habits Survey (NTHS) conducted in 1996 is the primary data source for model development. The Tel Aviv model accounts for up to two tours for each person: the most important tour of the day, referred to as the primary tour, and the second most important tour of the day, or the secondary tour.

The structure of the activity schedule model system for the primary tour is shown in Figure 2. At the highest level of the model system the auto availability model predicts the probability of the number of autos available to the household. Following this model, the primary activity model determines a person's primary activity. The alternatives include work, education, shopping, other types of activity out of home, and staying at home. For activities outside the home, the model then determines the destination of the primary activity and the main mode of the tour. Destination choice models are estimated using the full set of 1,219 traffic analysis zones as choice alternatives. The main mode of the tour is determined by a combined revealed-preference and stated-preference mode choice model.

The current version of the model includes 5 time periods as follows: early morning (MO) from 03:00 to 06:30, morning peak (AM) from 06:30 to 08:30, mid-day (MD) from 08:30 to 15:00, afternoon peak (PM) from 15:00 to 20:00, and evening / late night (EV), from 20:00 to 03:00. Because the time period models predict the joint choice of departure times from home and main destination, there are 15 feasible choice alternatives.

Subsequent models determine if there is an intermediate stop on the way from home to the primary destination and/or from the primary destination back home. For each intermediate stop, the model determines the destination of the stop by taking into account the additional disutility caused by adding a stop to the tour. Once all the details of the primary tour, which includes the main activity of the day are determined, the model estimates if there is an additional secondary tour for which a similar set of models is estimated.

The hierarchy of the models is fixed and the general model structure is similar to the Bowman and Ben-Akiva (2001) approach. The results from these models are translated into O-D travel matrices that are assigned to the network. For details on the model application, see Bekhor et al. (2011) and Shiftan et al. (2003).
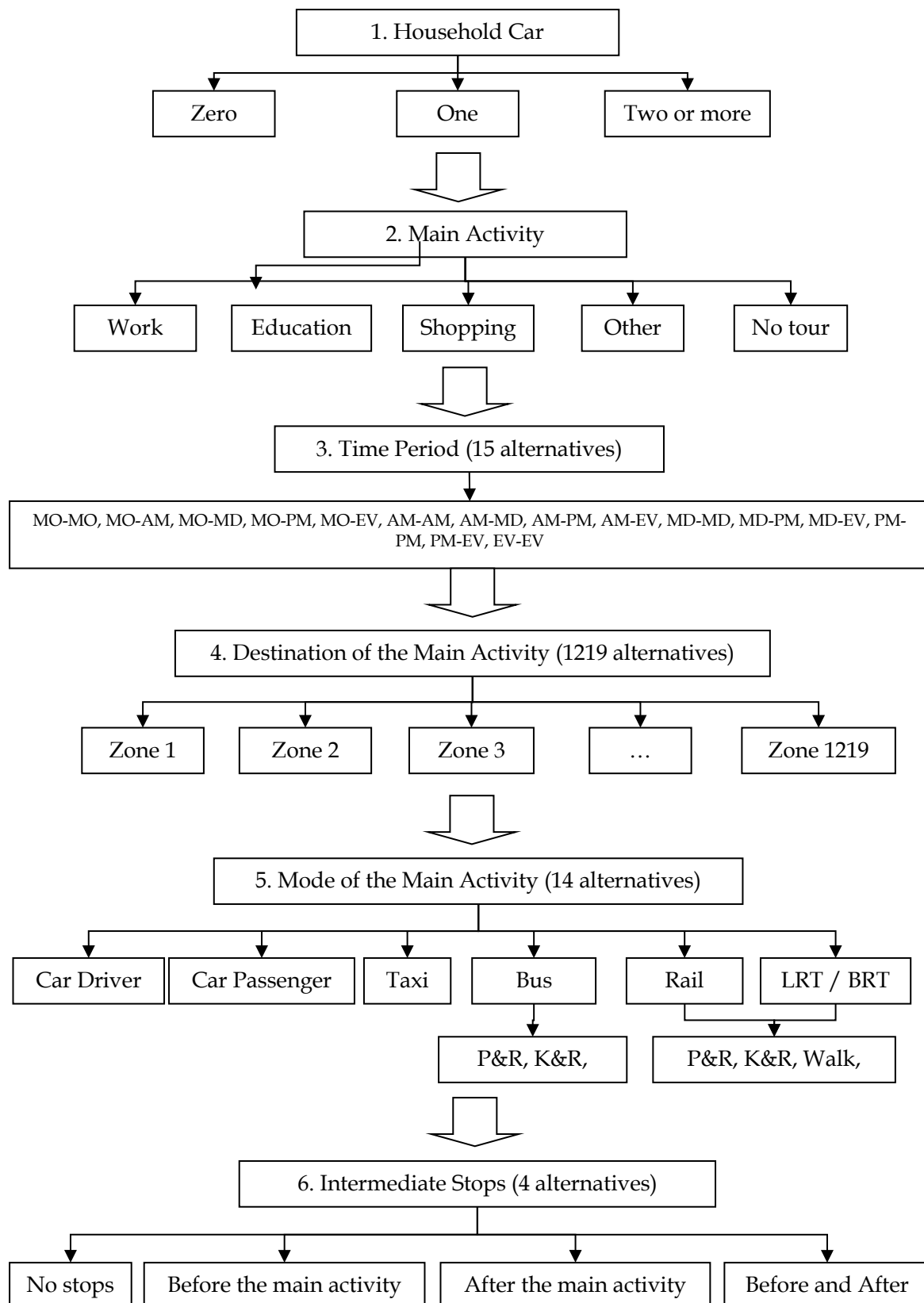
```
                    ┌─────────────────────┐
                    │  1. Household Car    │
                    └─────────────────────┘
           ┌────────────┬──────────────┬──────────────┐
      ┌─────────┐  ┌─────────┐   ┌──────────────┐
      │  Zero   │  │   One   │   │ Two or more  │
      └─────────┘  └─────────┘   └──────────────┘
```

| Zero | One | Two or more |

2. Main Activity

| Work | Education | Shopping | Other | No tour |

3. Time Period (15 alternatives)

MO-MO, MO-AM, MO-MD, MO-PM, MO-EV, AM-AM, AM-MD, AM-PM, AM-EV, MD-MD, MD-PM, MD-EV, PM-PM, PM-EV, EV-EV

4. Destination of the Main Activity (1219 alternatives)

| Zone 1 | Zone 2 | Zone 3 | … | Zone 1219 |

5. Mode of the Main Activity (14 alternatives)

| Car Driver | Car Passenger | Taxi | Bus | Rail | LRT / BRT |

| P&R, K&R, | P&R, K&R, Walk, |

6. Intermediate Stops (4 alternatives)

| No stops | Before the main activity | After the main activity | Before and After |

Figure 2.   Tel Aviv ABM structure (primary tour)

Figure 3 presents the Tel Aviv model application structure, and the interaction of the three main functional units: Population Generator (PG), Tour Generator (TG), and Trip Assignment Unit. The Population Generation unit is written in Visual FoxPro; the Tour Generator is a C# console

application; and the Assignment Unit is composed of a system of macro procedures for the EMME modelling package. Following is a brief description of the functionality of each unit.
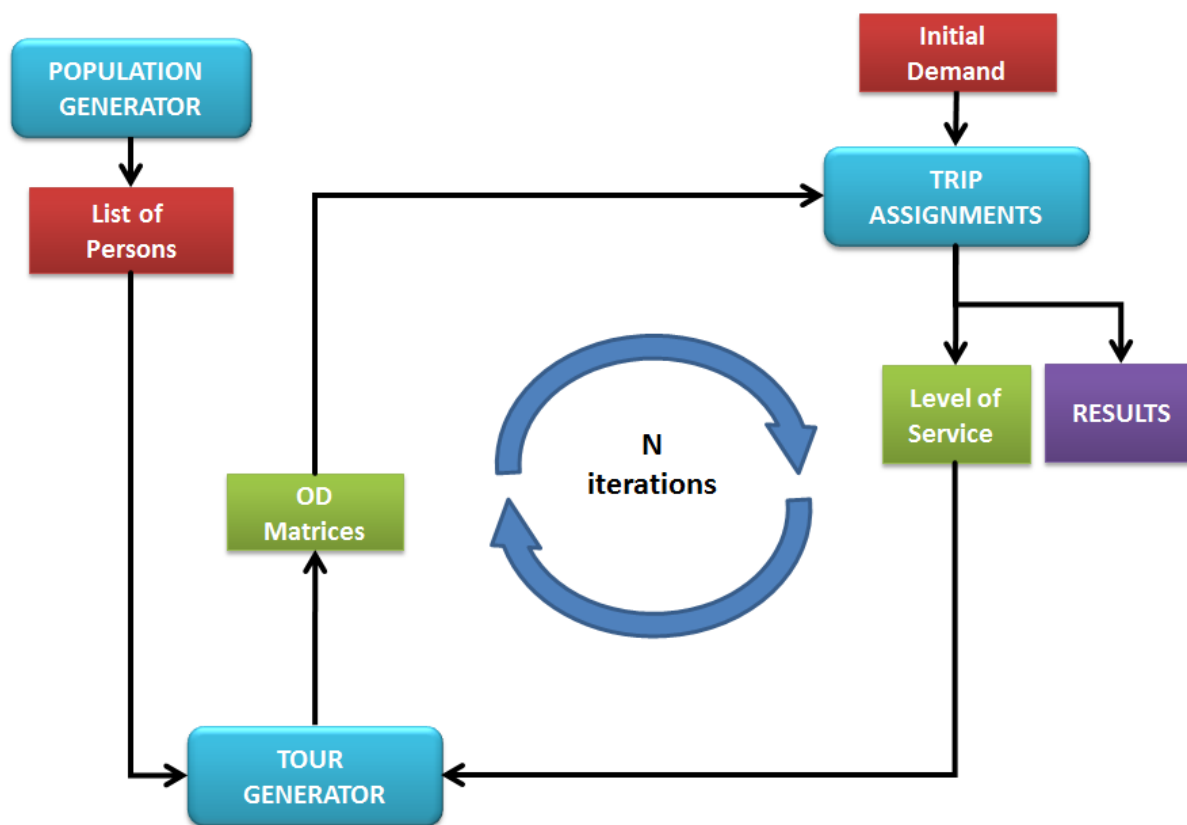

Figure 3.   Tel Aviv Model Application Structure

The stand-alone PG produces a list of individuals, which represents the population of the study area for a given year, and for a given sampling rate. It may be the full population, or a random sample often used to shorten run times. The list contains all the relevant attributes of the individuals needed for the Tour Generator unit.

The Tour Generator applies the activity-based model of travel behaviour to each person included in the generated population sample. As a result, description of each person's daily travels is obtained. The description includes types and time schedule of tours, attributes of intermediate stops, destinations for each activity, and transportation modes used. Then, tours are split into individual trips that are summarized into mode-specific demand matrices for different periods of day. These matrices are then used as input to the Trip Assignment Unit, which performs all necessary auto and transit assignments. The current version of the model uses static traffic assignment with link-based travel time functions.

After preparation of the list of individuals by PG, a model run is started with assignment of predefined initial demand by Trip Assignment Unit. The Level-of-Service (LOS) indices are obtained from assignments, including auto travel time, transit travel components like in-vehicle time, access walk time, waiting time, number of transfers, etc. These indices are passed to TG module where they affect the generation of individual travel tours, which in turn are transformed into demand matrices. Then, the assignments are made again with newly obtained demand. The base model loop is repeated several times in order to eliminate the influence of the initial demand.

Note that in general terms the iterative procedure of the ABM is quite similar to classic aggregate (deterministic) models: an initial demand is assigned to the networks, various LOS indices are

calculated, and then an updated demand is calculated for the assignment, calculation of new LOS indices and so on. The major difference between the models is in the way the demand matrices are calculated: in the aggregate model the OD matrices are obtained by superposing the distributions of partial mode-specific OD matrices for different market segments and/or trip types. In contrast, the OD matrices in the disaggregate ABM are obtained as a superposition of individual movements. These movements are based on instances of complex random choices (activity, time of day, destination, mode, etc.) made by each individual. As a result, the OD matrices in each iteration of the ABM run are random.

In deterministic (aggregate) trip-based models, the iterative run process converges to the equilibrium between demand and supply. This can be formally proved, assuming certain characteristics for the supply and demand functions. For a detailed discussion, see Patriksson (1994), which includes an extensive list of network equilibrium formulations and algorithms.

Considering the disaggregate ABM, it may be expected that the influence of initial demand will be decreased when number of iterations is growing. However, the randomness of results will remain, and special measures are required to obtain stable results. This will be discussed in depth in following sections.

3. Uncertainty sources

This section describes the uncertainty sources in the Tel Aviv ABM. The starting point of this section is the application structure described in Figure 3.

As mentioned in the previous section, the random nature of tour generating in the ABM brings uncertainty to the model's results. This "TG-related" randomness is the central source of uncertainty in ABM. In addition, there are two other elements of the ABM that can affect the level of uncertainty: population sampling and assignment method.

Population sampling is often used in runs of disaggregate models in order to shorten model run time (Rasouli and Timmermans 2012). For example, the ABM can be applied to a randomly chosen 10% of the individuals, and the resulting demand matrices formed by the superposition of the individual trips are then multiplied by 10, thus assuring the correct total demand.

The assignment method used in the model system is another uncertainty factor, since it may introduce an additional random noise into the model loop. This element of uncertainty is well-known in the literature (e.g. Horowitz, 1984; Parthasarathi and Levinson, 2010). This element is considered in the present paper for the sake of completeness, in order to bring together all relevant quantitative results related to the specific Tel Aviv ABM implementation.

The analysis of uncertainties will be conducted as follows: first, the assignment method will be presented. Then, the primary TG-related uncertainty will be considered, and averaging methods will be discussed. Finally, the effect of population sampling will be analysed.

The metric that will be used in this paper for the analysis of uncertainty, is the vehicle-hours travelled (VHT). This variable is one of major performance indices of the transportation system; it is widely used for evaluating benefits of transportation projects. This metric was also used in other studies (e.g. Bowman et al., 2006; Vovsha et al., 2008).

4. Assignment method

The traffic assignments of the Tel Aviv ABM are implemented using the EMME software. Accordingly, three deterministic traffic assignment algorithms available in EMME package were considered: the standard Frank-Wolfe algorithm (FW), the Frank-Wolfe algorithm with parallel computation (FWP), and path-based traffic assignment (PBTA).

Figure 4 compares the distribution of the vehicles hour travelled (VHT) according to the FW and PBTA assignments for AM peak hour car demand. The x-axis of the graph represents the iterations and the y-axis represents the distribution of link VHT differences for both assignment types, computed as the difference between the VHT of the current iteration and the VHT at equilibrium. Red (dark) colour indicates high percentage of links, and green (light) colour indicates low percentage of links.
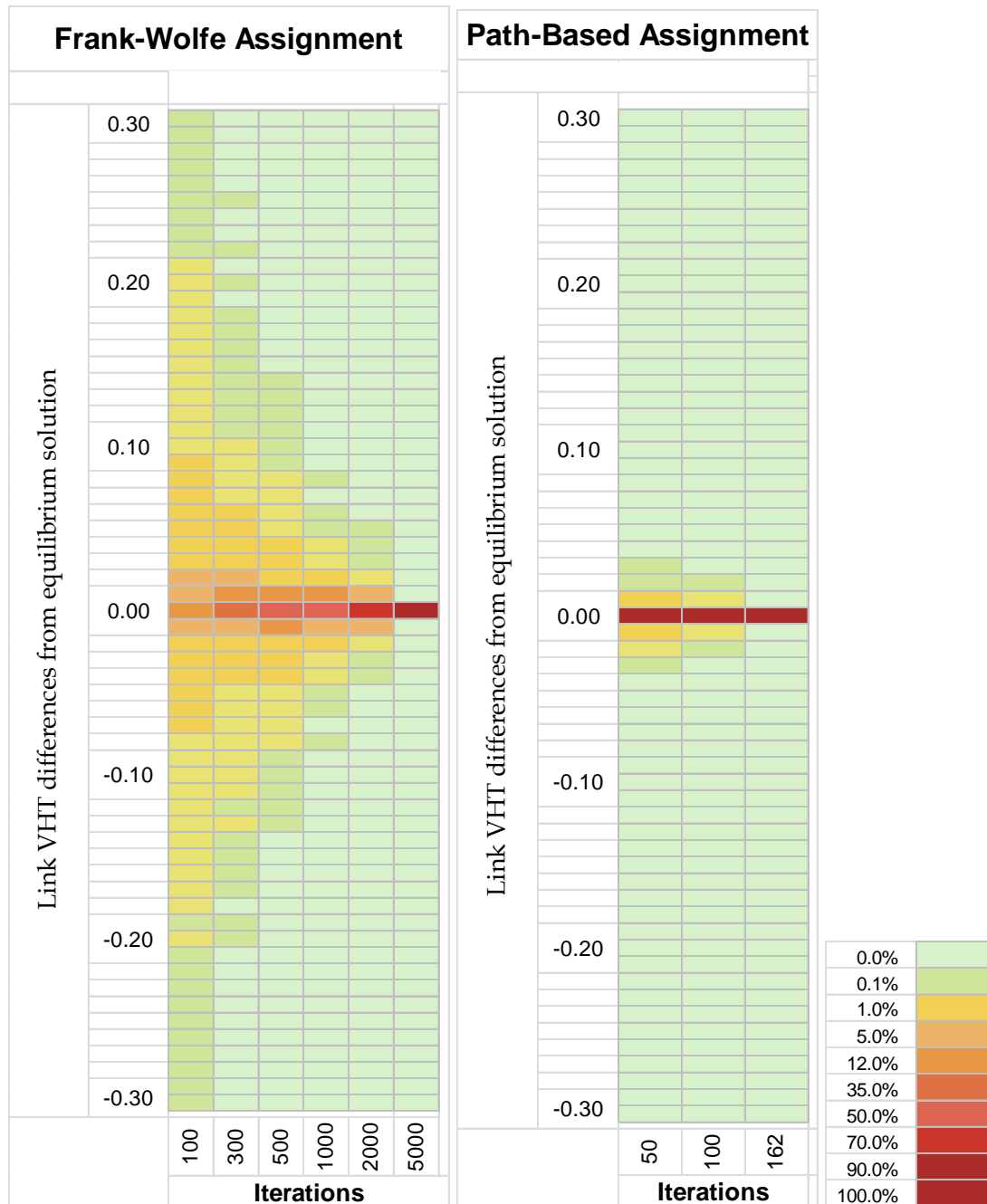

Figure 4.   Distribution of link VHT differences for FW and PBTA assignments

The results presented in Figure 4 shows that the FW algorithm needs many more iterations to reach the same convergence level of the PBTA. This result is not surprising because it is well known that path based algorithms converge faster compared to link based algorithms (Jayakrishnan et al., 1994).

Figure 5 presents dependence of link VHT standard deviation on run time of traffic assignment procedures. The standard deviation is calculated for VHT differences between current and equilibrium states in all links.
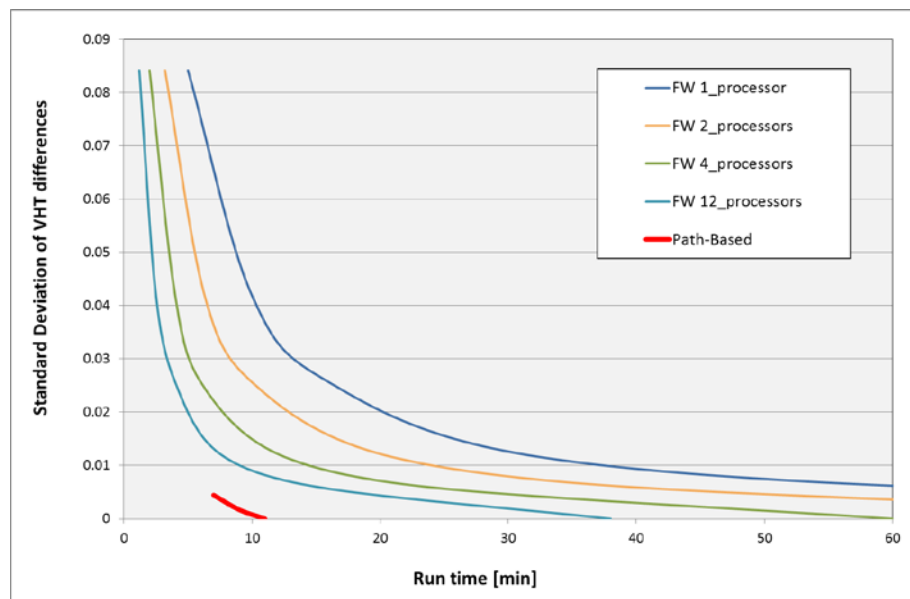

Figure 5.   Convergence of FW and path-based assignments

It can be seen, that the PBTA algorithm converges much faster, also when applying the FW algorithm with multiple processors. The analysis of the assignment results has shown that usage of PBTA algorithm may practically eliminate the assignment error. Therefore, in the subsequent sections the results do not consider the assignment method issue.

5. Tour Generator random component

This section analyses the simulation error produced by the TG component of the ABM system. The starting point of the analysis is the demand matrices that result from the TG component. According to the flowchart presented in Figure 3, the individual random choices are aggregated to form the demand matrices. There are over 30 demand matrices generated by the model for different modes and periods of day. To analyse the TG randomness effects, the paper considers the car demand matrix for the AM period.

Since the car demand matrix is formed by the individual trips, the cells of the matrix contain integer numbers. Figure 6 shows the distribution of cells (that is, OD pairs) by number of OD-trips for several consecutive iterations of an ABM run. Note that the total number of OD pairs is 1219 * 1219 = 1,485,961 cells. It can be seen that the distribution is very stable, and can be approximated with high accuracy by a power function; this fact may be a subject of separate analysis, far beyond the scope of the paper.
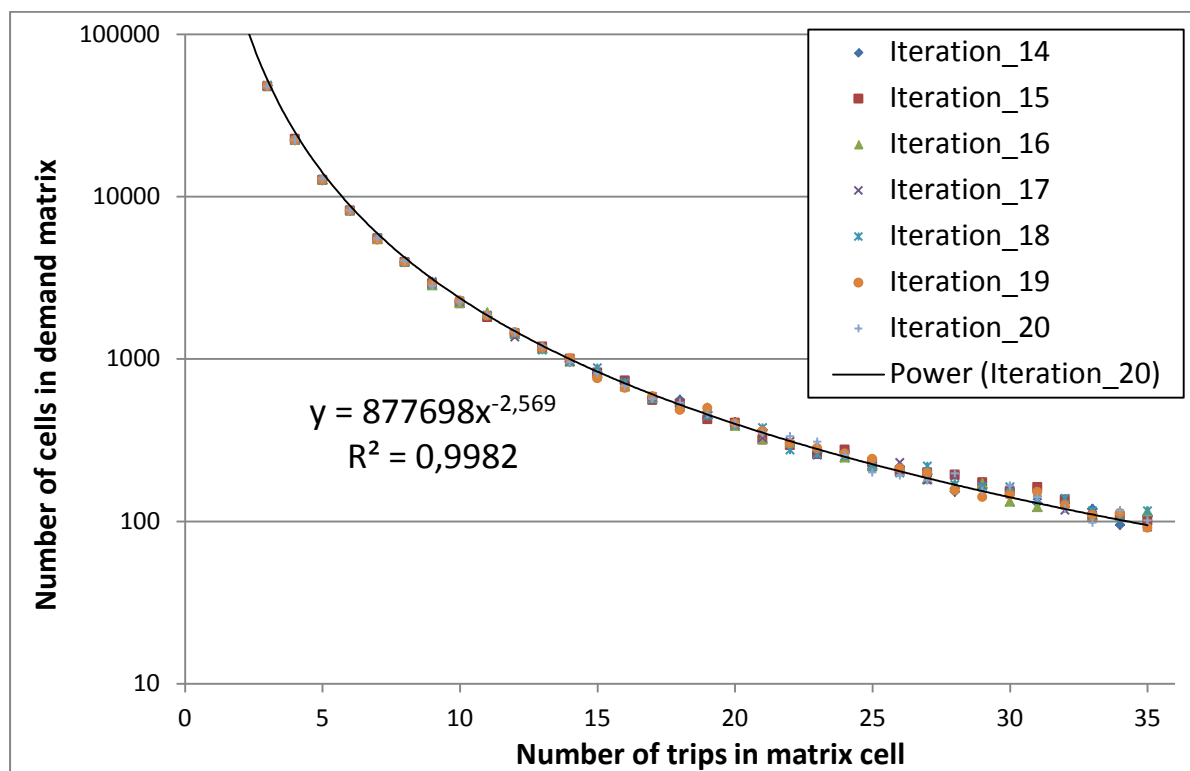
Figure 6.   Distribution of trips in the car demand matrix for different iterations

Although the overall distribution of trips in the matrix is quite stable, there are considerable changes at the cell level. In a given iteration of the model run, the number of non-empty cells is about 270,000, of which about 150,000 cells (55%) contain one trip. In the next iteration, about 117,000 non-empty cells (42%) transforms into empty cells, of them 93,000 (34% of non-empty cells, or 80% of transformed non-empty cells) were cells with one trip in the previous iteration. Similarly, there are changes in the opposite direction: about 117,000 empty cells transforms into non-empty cells in the next iteration. This "flip-flop" situation along the iterations is quite stable: Table 1 shows average values and standard deviation of above numbers of transformations for all different pairs of matrices for 10 iterations of the model.

Table 1.   Sparseness of demand matrices and pairwise comparison of matrices from different iterations

Parameter	Average	Standard Deviation
Total non-empty matrix cells	270,101	403
Total matrix cells with one trip	150,362	382
Total non-empty cells in iteration (n), corresponding to empty cells in iteration (n+1)	117,068	364
Total cells with one trip in iteration (n), corresponding to empty cells in iteration (n+1)	93,152	346

To conclude the analysis of demand matrix randomness, the overall coefficient of variation (CV) of the matrix was calculated (excluding cells, which are empty in both matrices): the average number of trips per cell is about 1.01, and standard deviation of cell changes is about 2.1, giving a relatively high CV value of about 2.

The next question for discussion is: how the randomness in the demand matrix is translated into the randomness of model results?

The issue of uncertainty propagation in transportation models was thoroughly investigated (see, for example, Zhao, Kockelman, 2002). The present analysis differs from previous studies due to the source of uncertainty: in most studies, the source of the error is in the input variables, whereas in the current study the uncertainty is produced intrinsically by the Tour Generator. Thus, no assumptions about type or intensity of the uncertainty are required.

To analyse the uncertainty of the model results, two elements are considered. The first one is the travel time matrix for the AM period auto assignment. This matrix is one of the several LOS matrices that affect the individual choices in the next iteration of the model. The second element is the total VHT by car for the AM peak period. This variable is one of the central performance indices used for comparison of different scenarios, and therefore may be considered as a representative of the overall model results.

Figure 7 shows the distributions of OD time differences, for several iterations of a model run. The differences are calculated between successive iterations at the cell level.
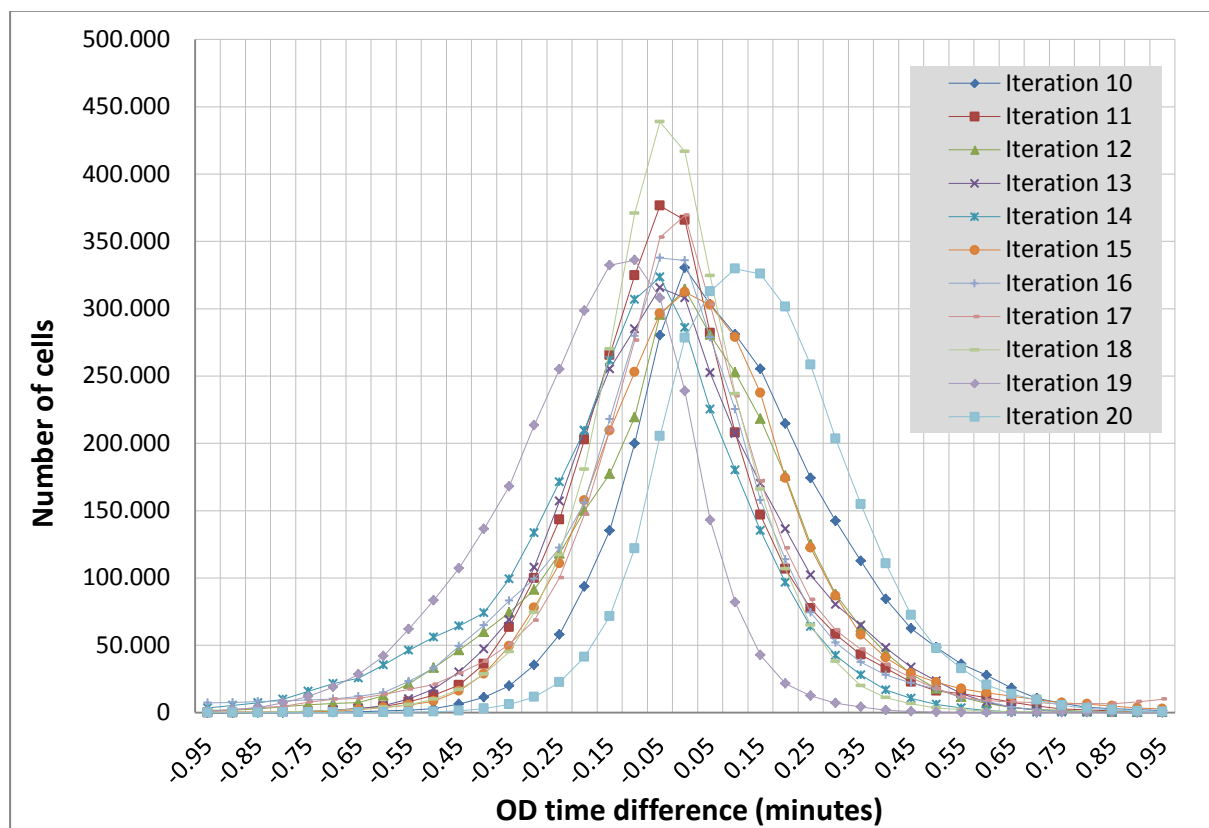

Figure 7.   Distribution of OD time differences for various iterations

The average difference varies in the range [-0.19, 0.15], and standard deviation varies in the range [0.16, 0.25]. Since the average OD time in the matrix (not weighted by demand) is about 41 minutes, the correspondent CV of time differences is about 0.007, which is significantly lower in comparison to the car demand matrix. There are two reasons that may explain this low CV: first, the well-known high smoothing power of the assignment procedure, which is related to its aggregative nature: the trips from about 270,000 cells (OD pairs) are assigned to the network with about 10,000 links, of which many are interdependent. Second, in the ABM the LOS values calculated from the assignment results in one iteration affect the formation of demand matrix for

the next iteration. Therefore, the demand randomness is not fully "white" and this also has an impact on the low CV value.

Figure 8 shows the variations of overall VHT by car for AM peak hour over iterations of a model run. The variations are presented for 3 independent model runs, which differ only by the random seed value. The independent model runs are respectively identified as "scen 2403", "scen 2422" and "scen 2425".
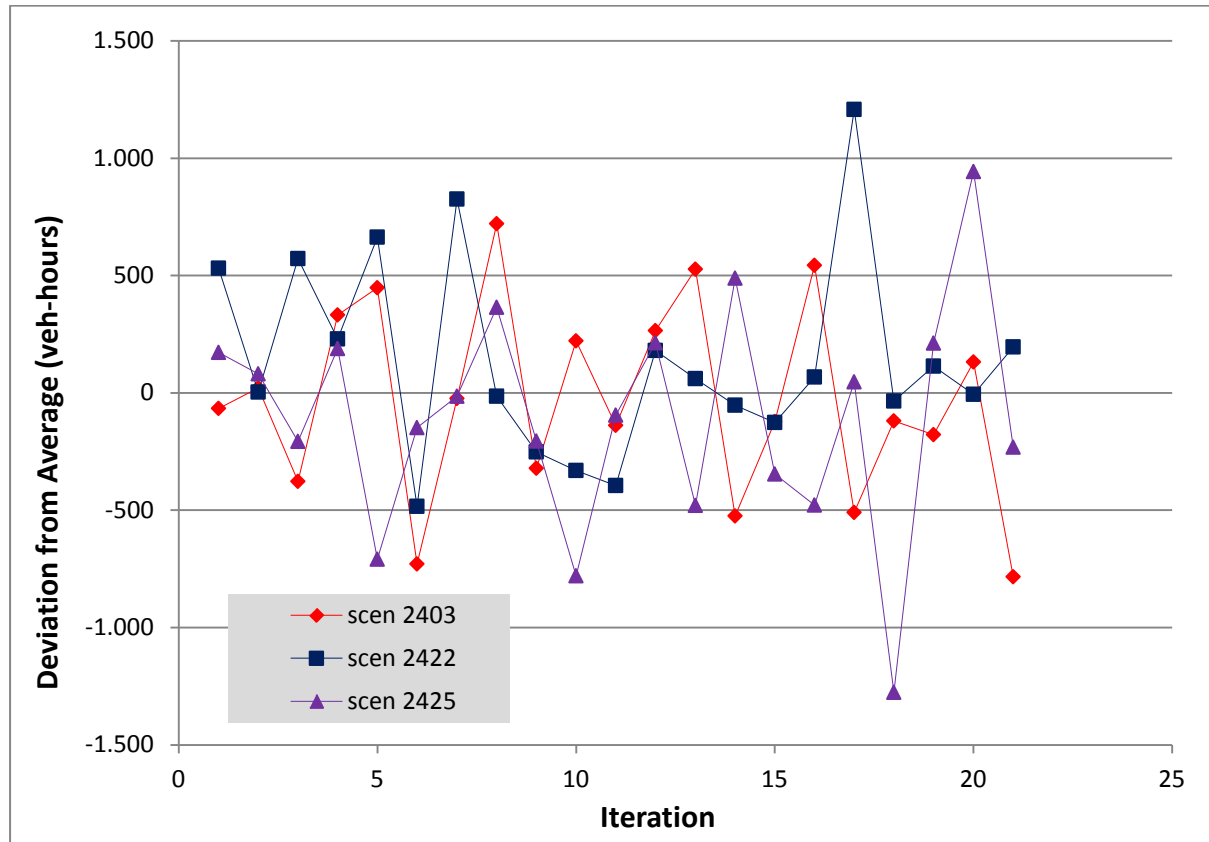

Figure 8.   Variations of VHT in ABM run

Statistical tests applied to the VHT values show that the values are random with 95% confidence. This means that after several iterations of a model run, the result (VHT value) is stabilized around the average, with a standard deviation of about 420 vehicle-hours. An important consequence can be derived here: the model loop [LOS->Tour Generator->Demand Matrices->LOS] suppresses considerably the random component, otherwise the deviation from the average VHT would have grown with iterations, since at each iteration a new portion of uncertainty is added. This consequence is quite natural and intuitive: for example, if the LOS is deteriorated for some OD pair, then the probability of performing an activity for this OD pair will decrease, which correspondently will reduce the number of trips for this OD pair, and vice versa.

The standard deviation of VHT obtained above is usually not acceptable for practical purposes. For example, small to medium projects may bring VHT savings of 100-200 vehicle-hours, and the random error in model results will not allow for correct project judging. Therefore, measures should be taken in order to stabilize model results. Possible methods to average results are discussed in the next section.

6. Averaging methods

Different averaging methods can be used to stabilize the model run results. For example, in Castiglione et al. (2003) and Cools et al. (2011) a simple average of model results was considered, whereas in Vovsha et al. (2008) averaging of demand matrices, link volumes, link times and LOS skims were considered.

In this paper three averaging methods are chosen for analysis, which are respectively presented in the following Figures 9, 10 and 11: "MSA-R", "MSA-M", and "Quasi-Aggregation" methods.
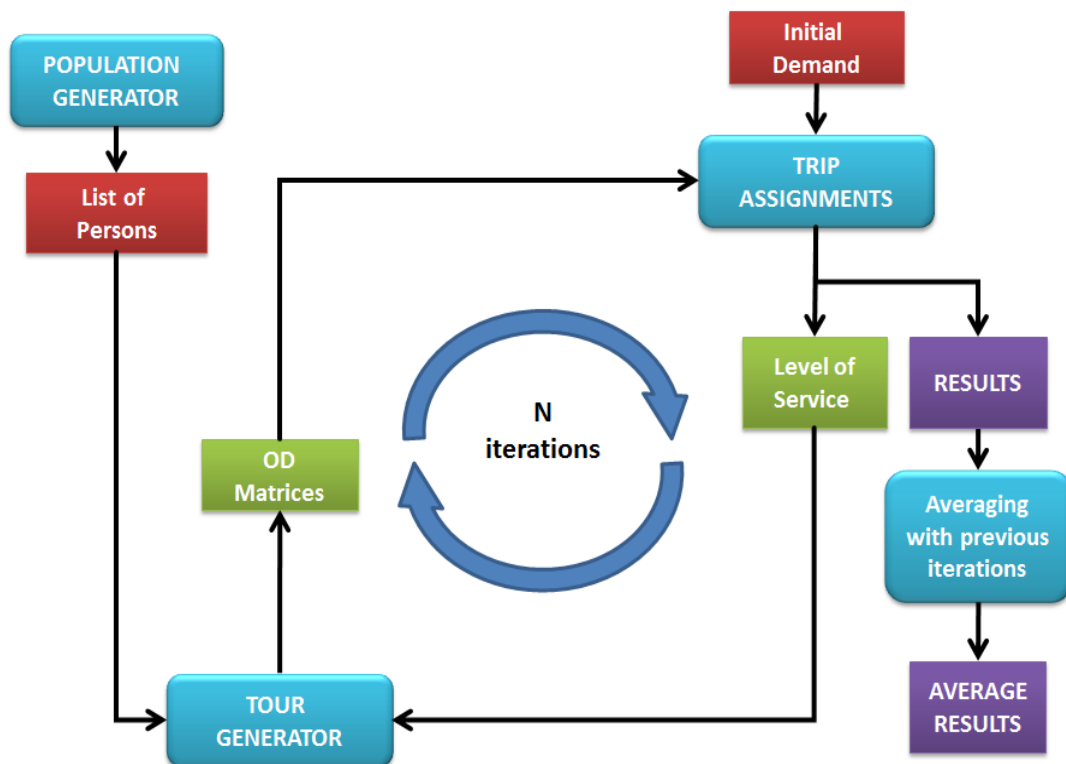
Figure 9.   MSA-R method

The "MSA-R" method is based on the fact that after several initial iterations, the randomness of results is stabilized. Therefore, the results obtained after each model iteration can be successively averaged.
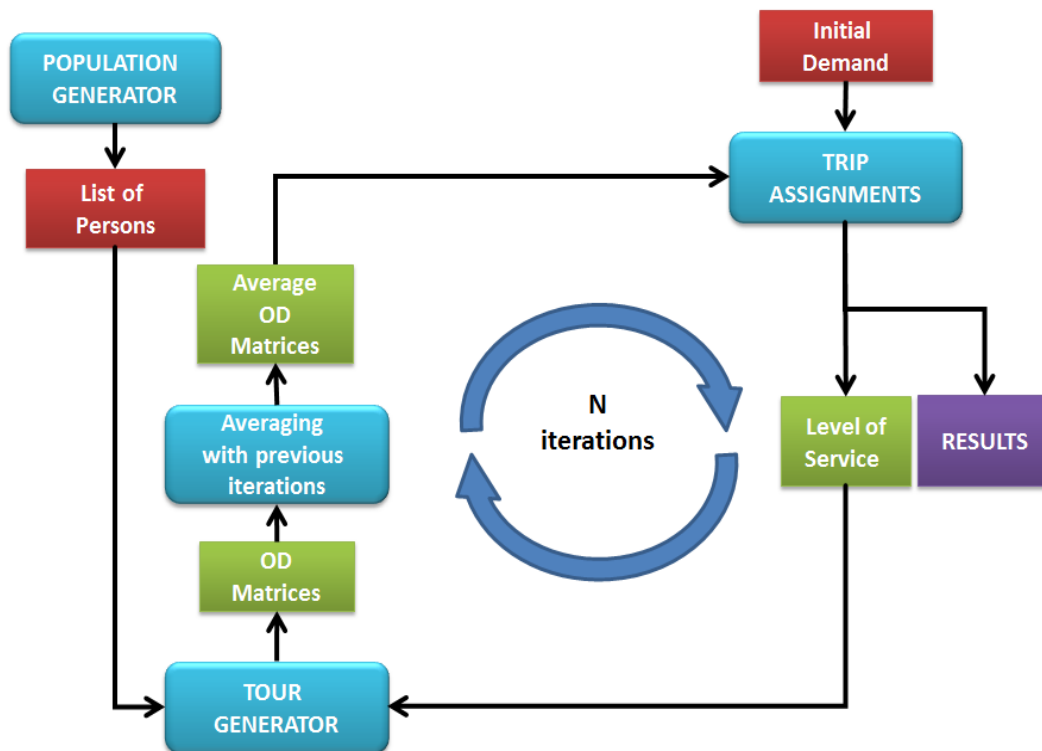
Figure 10.   MSA-M method

The "MSA_M" method is based on averaging the demand matrices after each iteration. As presented in Section 5, the demand matrices exhibit a high variability which ultimately affects the randomness of model results. Averaging the demand matrices will smooth the values in matrix cells and thus improve the model stability.
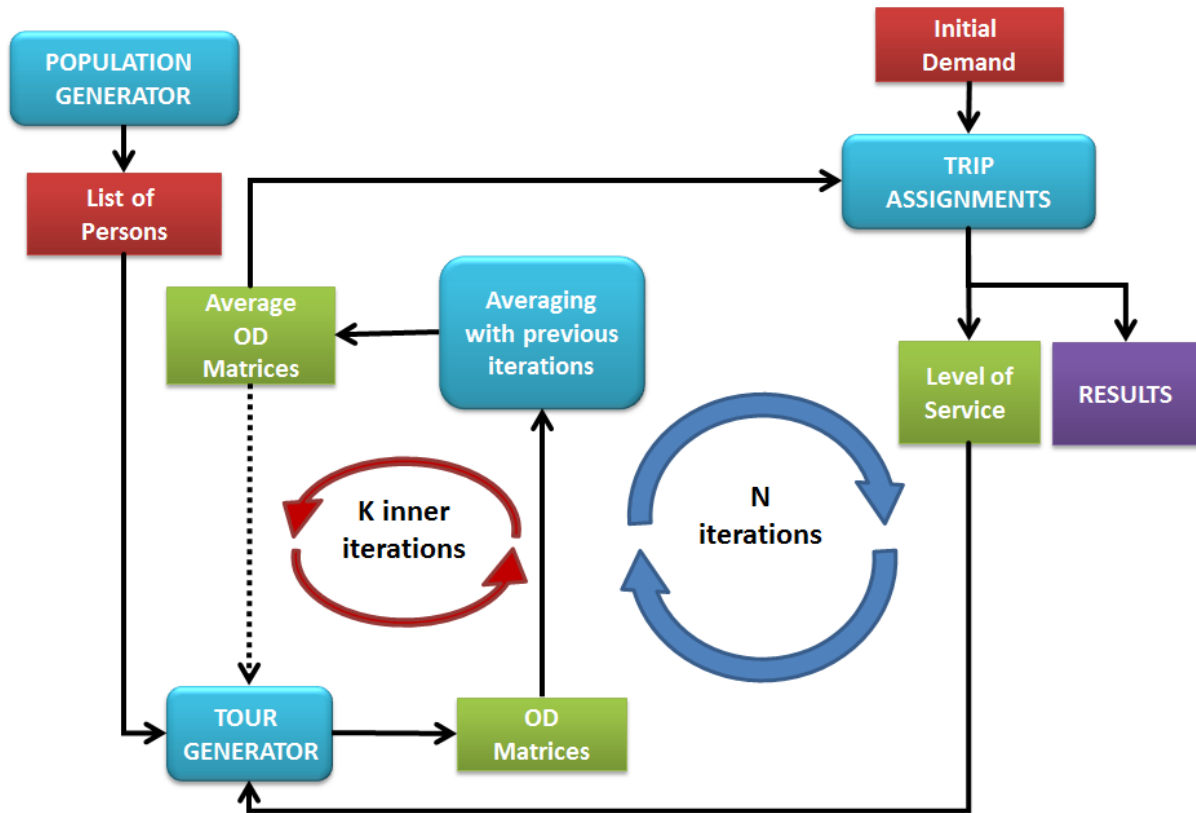
Figure 11.   "Quasi- aggregation" method

The "Quasi-aggregation" method is a resemblance of aggregate modelling. After making an iteration of model run and obtaining LOS matrices, the TG is activated several times using same population and same LOS, and demand matrices obtained each time are averaged. Then assignments are made, and new LOS matrices are obtained for the next iteration. The method is called "Quasi-aggregation" since asymptotically, when the number of TG activations (inner loop) increasing, the averaged demand matrices will represent the probability distributions of OD trips, just like in aggregate models. Of course, unrealistically large number of inner iterations would be required to reach the good approximation of probability distributions, however the smoothing of demand matrices will have the positive effect, and omitting assignments after each inner iteration will save run time.

The following equations describe the averaging procedures:

MSA – R: $$\widehat{VHT}_a^{(n+1)} = \frac{\overline{VHT}_a^{(n)}}{n+1} + \left(1 - \frac{1}{n+1}\right) VHT_a^{(n+1)} \tag{1}$$

MSA – M: $$\hat{T}_{OD}^{(n+1)} = \frac{\hat{T}_{OD}^{(n)}}{n+1} + \left(1 - \frac{1}{n+1}\right) T_{OD}^{(n+1)} \tag{2}$$

Quasi-Aggregated: $$\hat{T}_{OD}^{(n+1)} = \frac{\hat{T}_{OD}^{(n)}}{n+1} + \left(1 - \frac{1}{n+1}\right) \frac{1}{K} \sum_{k=1}^{K} T_{OD}^{(n+1,k)} \tag{3}$$

Where n indicates iteration number, k is inner iteration number for the "Quasi-aggregation" procedure; VHT_a is the vehicle hours travelled at link a; T_{od} is the demand for trips for the origin-destination pair OD. As can be seen in the equations, the averaging procedure is recursive; variables with cap (^) refers to the output of the averaging procedure, and symbols without caps are input variables.

For each averaging procedure three independent model runs were conducted, with different random seeds, indicated by "scen" numbers. Figures 12, 13, 14 respectively present, for each

averaging method, the deviation of the total VHT for AM peak hour as a function of iteration number.
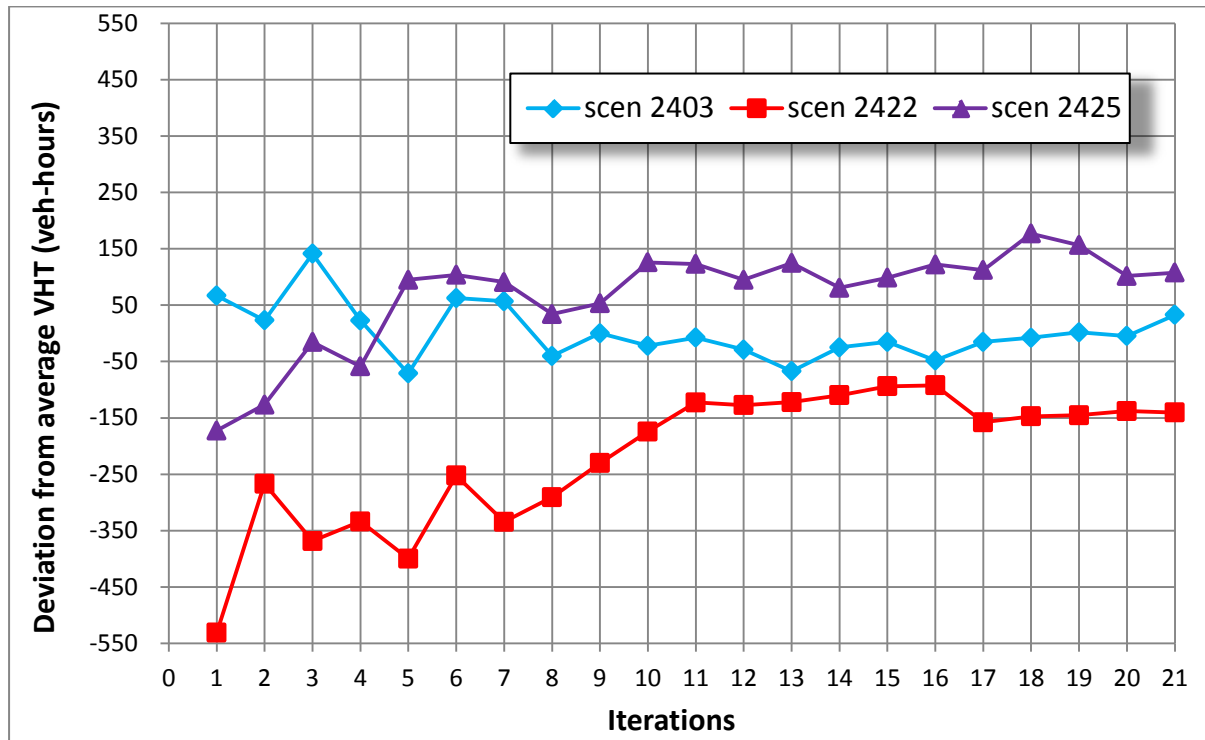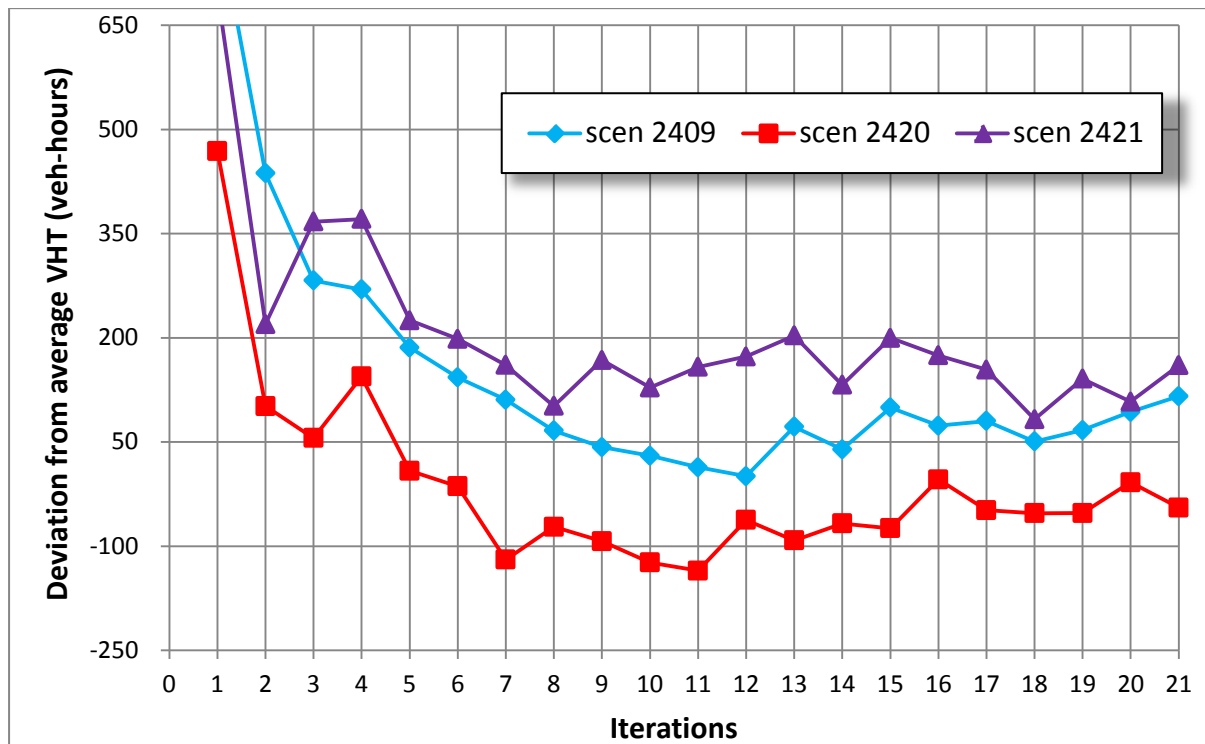

Figure 12.   VHT convergence for MSA-R procedure

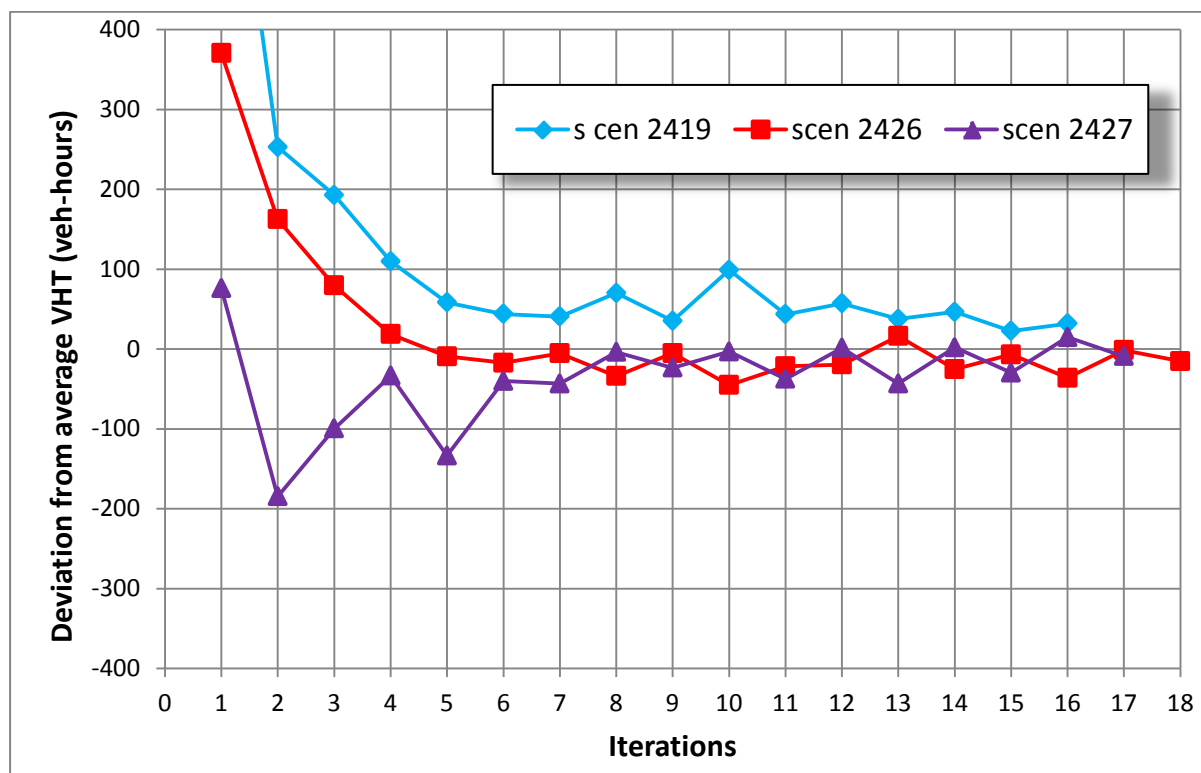*Figure 13.   VHT convergence for MSA-M procedure*

Figure 14.   VHT convergence for "Quasi-aggregation" procedure

The main results of the averaging procedures are summarized in Table 2.

Table 2.   **Convergence of averaging procedures**

Averaging procedure	Number of iterations	Standard Deviation of VHT at last iteration (computed for 3 runs)	Average deviation of last iteration VHT from global average
MSA-R	20	120	-33
MSA-M	20	140	-9
Quasi-Aggregation	16 (6 inner each)	35	43

The analysis of Figures 12 – 14 and Table 2 indicates that the convergence rate of all procedures follows, by the order of magnitude, the "$\sqrt{n}$" rule of thumb: for MSA-R and MSA-M the VHT standard deviation decreases from the original 420 to 120-140 after 20 iterations, that is comparable to the expected $420/\sqrt{20} = 93.9$; the resulting VHT standard deviation of 35 for Quasi-aggregate procedure after 16 iterations with 6 inner iterations each is even closer to $420 /\sqrt{16 \times 6} = 42.8$.

Another important observation from the averaging procedures for different initial seeds is that all of them converge to the same stationary point.

However, it should be noted that statistical data available from the model runs is not sufficient for accurate estimation of the convergence rate; at this stage of research the above statements can be considered as indications only. Nevertheless, these results allow adding a feature of accuracy monitoring to the ABM, in the sense that a desired variable estimate can be obtained together with the prediction of the number of iterations required for a given accuracy. Of course, the monitoring should refer to the performance indices specific to the study conducted.

Next, the comparison of the efficiency of the averaging procedures is made. The run time, and not number of iterations is used here in order to assure comparability between different averaging procedures. Figure 15 shows the convergence of VHT estimates as a function of model run times for all model runs analysed. A total of 9 runs were performed (3 averaging methods with 3 different initial seeds).

Note that the MSA-M procedure requires more time for assignments than MSA-R, since for path-based assignment used the run time depends on number of non-zero cells in demand matrix, and in MSA-M procedure the number of such cells increases with the iterations, whereas in MSA-R this number is almost constant. Further, Quasi-aggregate procedure has different proportion between number of TG runs and number of assignment runs, depending on amount of inner iterations.
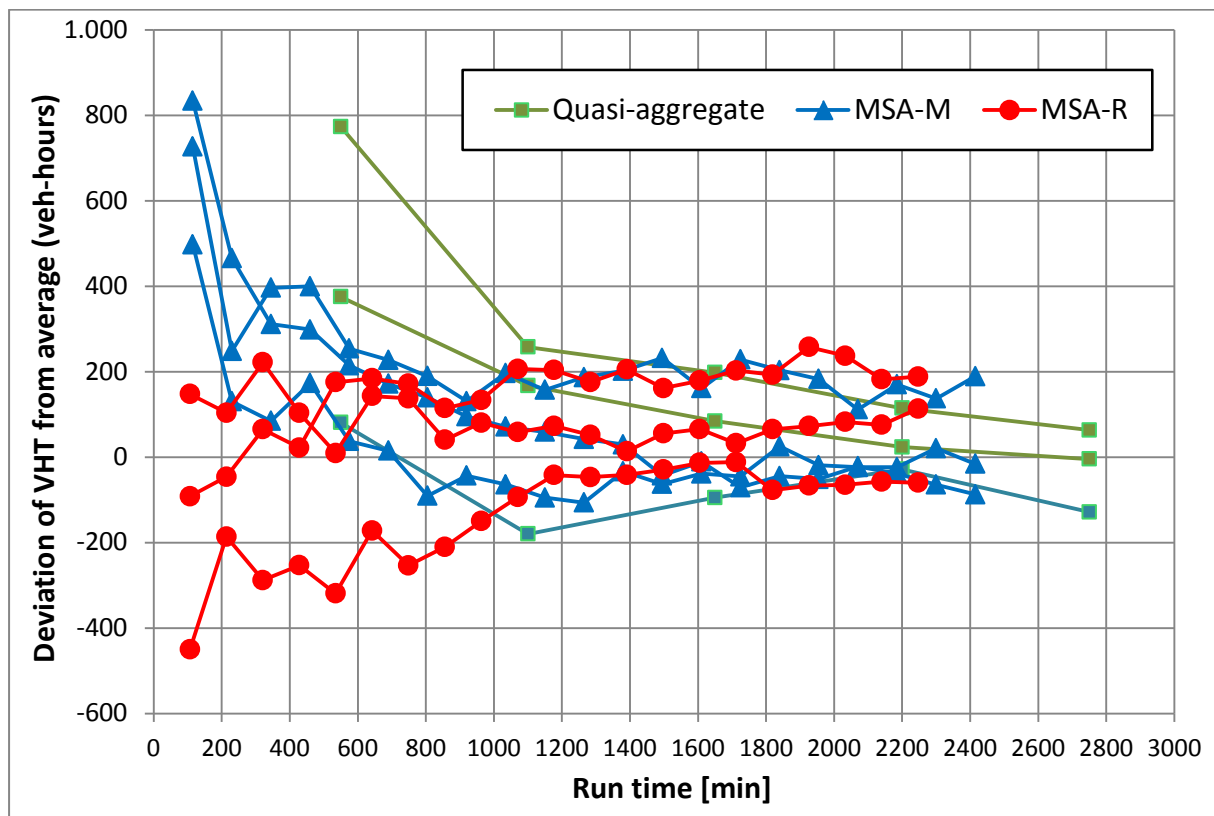

Figure 15.   Convergence of averaging procedures as a function of run time

Despite the above mentioned differences, it can be seen that the efficiency of all procedures is quite similar, and no one of them can be preferred based on available statistical data. Further analysis is required to refine the resulting conclusion. However, no significant changes in behaviour of averaging procedures could be expected, and therefore, the choice of procedure most probable should be based on other considerations, like convenience of implementation.

7. Population Generator random component

Different types of PG-related sources of randomness exist, as described in Rasouli and Timmermans (2012). First, the PG generates list of persons with random characteristics based on forecast of aggregate control variables. The person characteristics are not uniquely defined by the aggregate control variables and hence may depend on the PG algorithm or its implementation.

Further, in order to accelerate the model's run, a sample from the full population is often used. A sample is taken randomly, and trips of each person in a sample take proper weight to assure correct total number of trips in a system.

The generation of full population in the study area also includes random variations to the model results, because the source of the synthetic population is based on aggregate measures such as total population and employment.

In this paper only the effects of sampling on the ABM results are considered.

Two population samples were prepared by random selection of 10% and 50% of persons from the full list formed by Population Generator. For each sample, three independent model runs were conducted, assigning weights of 10 and 2, respectively, to all trips generated in each iteration for the 10% and 50% samples. Figure 16 shows the results of the model runs for the different population samples and initial seeds (indicated by the "scen" number).
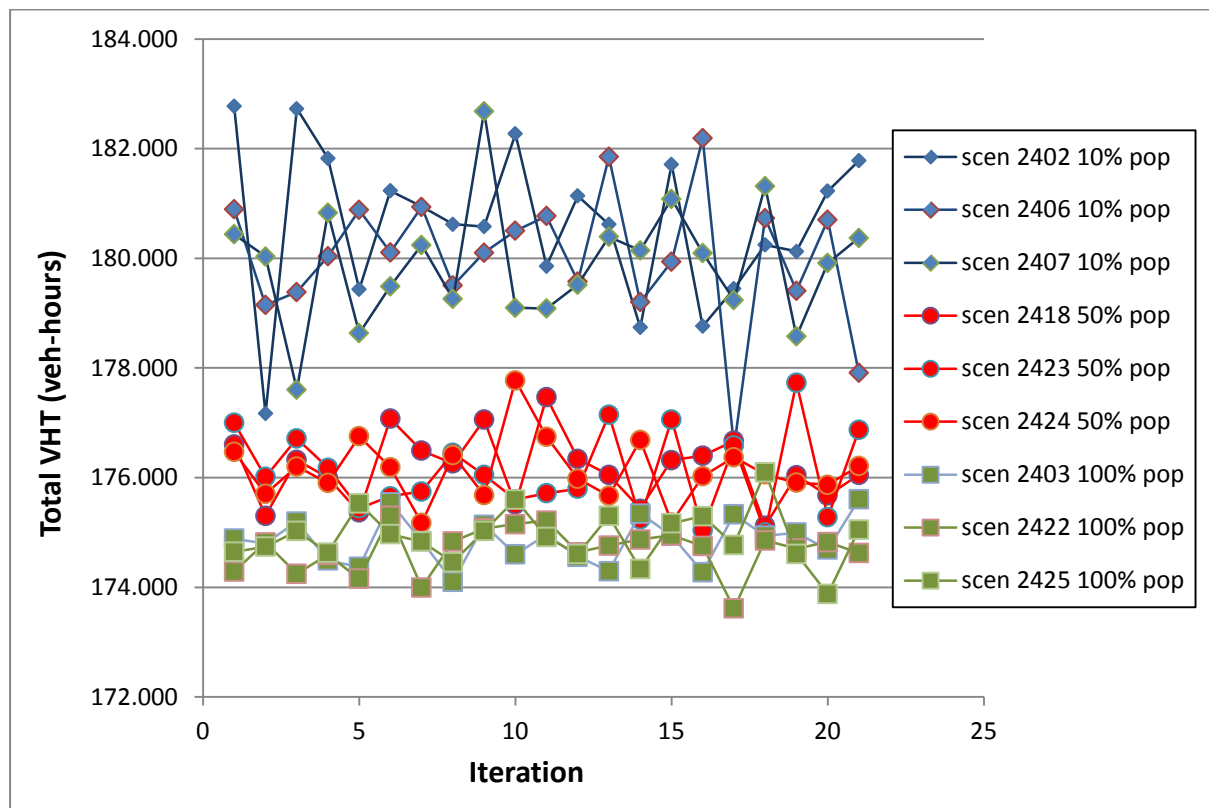

Figure 16.   Variations of VHT for different population sample size

The results presented above show, first of all, that the sample size affects the stationary point of ABM results. To our knowledge, this influence has never been reported.

The coarse structure of the demand elements might make the model less flexible for smaller sample sizes. Therefore, to assure the validity of results, an additional run of the ABM with 10% sampling was made, for different random choice of sample population. The results appeared to be statistically indistinguishable from those obtained for the original 10% sample, which means that the 10% sample is representative of the full population. Table 3 summarizes the convergence of the ABM run for different sample sizes.

Table 3.   **Convergence of ABM for different sample sizes**

Sample size s	Average VHT	Standard deviation of VHT variations		Standard deviation after 20 iterations	
		Observed	Theoretical $\sigma_{100\%}/\sqrt{s}$	Observed	Theoretical $\sigma/\sqrt{20}$
100%	174,700	420	420	120	92.9
50%	176,150	650	594	160	145.3
10%	180,320	1,250	1,330	250	279.5

The results indicate also, that the "$\sqrt{n}$" rule of thumb works both for the number of iterations and for the sampling rate. An important conclusion from this rule is that using the population sampling does not bring any significant savings in number of iterations, if the goal is to assure certain accuracy of ABM results. Moreover, if the total run time is considered, the run with 10% sample would require much more assignments, and therefore, may be comparable to the run with 100% sample only for Quasi-aggregation procedure, with proper increase of the number of inner iterations.

8. Summary and Outlook

Assuming that input variables are given and described by a fixed list of individuals, this paper analysed three sources of uncertainty in the model results: the assignment precision, the simulation error and the population sampling-related error. In line with previous studies, the uncertainty component related to the assignment procedures may practically be eliminated by using path based assignment algorithms.

The randomness of TG requires implementation of an averaging procedure to obtain stable results of a model run. This paper analysed three different averaging procedures, all three procedures appeared to show similar efficiency. The analysis performed allows introducing the run convergence monitoring and definition of stopping criteria based on the "$\sqrt{n}$" rule of thumb for the performance indices of interest. Future research may consider more "aggressive" averaging procedures, such as LOS averaging proposed in Vovsha et al. (2008).

As expected, the variability in model results was proportional to the inverse of the square root of the population sample size. In addition, the paper reveals that the sample size affects the stationary point of ABM results. It was shown also that the common practice of using 10% population sample does not save run time to obtain the same accuracy level of the results for the full population.

The influence of synthetic population and its randomness on the variability of model results requires further study: in addition to quantitative aspects analysed in the paper, the foundations of population synthesizing should be thoroughly examined.

Acknowledgments

References

Bekhor, S., Dobler, C., and Axhausen K.W. (2011). Integration of activity-based and agent-based models: case of Tel Aviv, Israel. *Transportation Research Record,* 2255, 38-47.

Bowman, J. L., and Ben-Akiva, M. E. (2001). Activity-based disaggregate travel demand model system with activity schedules. *Transportation Research Part A: Policy and Practice*, 35(1), 1-28.

Bowman, J. L., Bradley, M. A., and Gibb, J. (2006). The Sacramento activity-based travel demand model: estimation and validation results. Paper presented at the *European Transport Conference.*

Cambridge Systematics Inc. (2008). *Tel-Aviv Activity Schedule Travel Demand Model System: A Tour-Based Approach*. Final Report prepared for the Ministry of Transport.

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 microsimulation error on key travel indices: Evidence from the activity-based model Feathers. Paper presented at the *90th Annual Meeting of the Transportation Research Board,* Washington D.C.

Horowitz, J. L. (1984). The stability of stochastic equilibrium in a two-link transportation network. *Transportation Research Part B*, 18(1), 13–28.

Jayakrishnan, R., Tsai, W. T., Prashker, J. N., and Rajadhyaksha, S. (1994). A faster path-based algorithm for traffic assignment. *Transportation Research Record*, 1443, 75–83.

Patriksson, M. (1994). *The traffic assignment problem: models and methods*. VSPBV, The Netherlands.

Parthasarathi P., and Levinson, D. (2010). Post-construction evaluation of traffic forecast accuracy. *Transport Policy*, 17(6), 428–443.

Rasouli, S., and Timmermans, H.J.P. (2012). Uncertainty in travel demand forecasting models: literature review and research agenda. *Transportation Letters: The International Journal of Transportation Research*, 4, 55-73.

Shiftan, Y., Ben-Akiva, M.E., Proussaloglou, K., de Jong, G., Popuri, Y., Kasturirangan, K., and Bekhor, S. (2003). Activity-based modeling as a tool for better understanding travel behaviour. Paper presented at the *10th Conference of the International Association of Travel Behaviour Research*, Lucerne, Switzerland.

Veldhuisen, K.J., Timmermans, H.J.P., and L.L. Kapoen (2000). Microsimulation model of activity-travel patterns and traffic flows: Specification, validation tests and Monte Carlo error. *Transportation Research Record*, 1706, 126-135.

Vovsha, P., Donnelly, R., and Gupta, S. (2008). Network equilibrium with activity-based microsimulation models: the New York experience. *Transportation Research Record*, 2054(1), 102-109.

Zhao, Y., and Kockelman, K. (2002). The propagation of uncertainty through travel demand models: an exploratory analysis. *The Annals of Regional Science*, 36, 145–163.