Ref. Ares(2019)1367833 - 28/02/2019

Behavioural Simulations Design Aspect Report

D4.3

February/2019

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 689954.

Project Acronym and Name	iSCAPE - Improving the Smart Control of Air Pollution in Europe	
Grant Agreement Number	689954	
Document Type	Technical Report	
Document version & WP No.	V. 01	WP 4
Document Title	Behavioural simulations design aspect report	
Main authors	Muhammad Adnan	
Partner in charge	University of Hasselt (UH)	
Contributing partners		
Release date	February, 2019	

Document Control Page	
Short Description	*The focus of this deliverable is mainly to describe the scope of the behavioural simulation and provide the details of different design aspects of the simulation framework. Much of the discussion regarding FEATHERS and MATSIM, along with an illustration of the implementation and execution of the policy scenarios are already provided in deliverable 4.2. Therefore, in order to avoid duplication, in this deliverable we focused more on aspects showing the use of different processes to obtain the required input data, estimation of models and calibration framework details of the simulator. The actual simulation execution and obtained results for the three cities in the base case and in a few policy scenarios will be reported in the deliverable 4.4.*

Review status	**Action**	**Person**	**Date**
	Quality Check	*Coordination Team*	28/02/2019
	Internal Review	*Erika Brattich (UNIBO)* *Athanasios Votsis (FMI)*	14/02/2019
Distribution	Public		

Revision history			
Version	**Date**	**Modified by**	**Comments**
V0.1	14/02/2019	Muhammad Adnan	First Draft for internal review
V0.2	27/02/2019	Muhammad Adnan	Final Draft addressing comments from the internal reviewers (submitted for Quality Check)

Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Table of Contents

Table of Contents

List of Figures

List of abbreviations

CHAID	Chi-square Automatic Interaction Detection
CRAB	Centraal ReferentieAdressenBestand
DOW	Description of Work
FEATHERS	Forecasting Evolutionary Activity-Travel of Housedhold and their Environmental RepurcussionS
FMI	Flemish Meteorological Institute
GIS	Geographical Information System
GPS	Global Positioning System
GTFS	General Transit Feed Specification
IMOB	Instituut Voor Mobiliteit (Transportation Research Instiute)
IPU	Iterative Proportional Updating
MATSIM	Multi-agent Transport Simulation
OD	Origin Destination
OLS	Ordinary Least Square
OSM	OpenStreetMap
OVG	Onderzoek VerplaatsingsGedrag
POI	Point of Interest
SPSA	Simultaneous Perturbation Stochastic Approximation
SRMSE	Standardized Root Mean Square Error
TAZ	Traffic Analysis Zones
UNIBO	University of Bologna
WP	Work Package

1 Executive Summary

This deliverable reports the work carried out as part of WP 4 under task 4.2.2, with two major aims. Firstly, it presents the scope of the behavioural simulation process and selected operational tools to develop a simulation framework. This is done to test different mobility related interventions and assess their impacts on improving air quality and individual exposure in other tasks of the same WP. In particular, Section 3 of this Deliverable provides some discussion on this line of action in continuation with the content already described in deliverable 4.2 of WP 4.

As a second major aim, the deliverable presents two behavioural simulation frameworks; the first one is more comprehensive and employs a complete activity-based model FEATHERS to represent behavioural notions of individual decision making with great detail, while the second framework employs a light activity-based model mainly developed to provide inputs for the MATSIM framework (mainly used as supply-side model within the behavioural simulator). In the first case, the model estimation process and the development of required inputs are time-consuming, whereas the light activity-based model has low data requirements and uses a simplified algorithm which can be used in absence of required datasets and also when resources are limited. Since the behavioural simulation framework is designed in such a manner that it uses openly available dataset, its use can be easily extended to other cities also not part of the iSCAPE project. Furthermore, the use of MATSIM also provides flexibility to assess individual exposure in a more appropriate manner as it keeps intact the movement of individuals from one point to another. This is important for task 7.3 of WP 7, whose aim is to estimate exposure based on different individual socio-economic classes and to further expand this to estimate health impacts for development of general guidelines.

The deliverable provides the basis for executing simulations for the base case and also for some policy scenarios for selected iSCAPE cities, a work which is carried out as part of task 4.2.3 of WP 4 and presented in deliverable 4.4. The important aspects described in this deliverable are the development process of synthetic population using IPU algorithm, estimation process of models within FEATHERS framework, light activity-based model development and calibration framework utilizing SPSA algorithm.

2 Introduction

Activity-based (AB) model in the transport modelling literature is described as state-of-the-art modelling approach to predict travel behaviour of the individual in a more detailed manner [*Beckx et al., 2009, Hao et al., 2010, Dons et al., 2011, Rasouli and Timmermans, 2014, Adnan et al., 2016, Adnan and Ahmed, 2017*]. Compared to the traditional trip-based approach that predicts travel pattern of a region in an aggregate manner e.g. the number of trips originated for a particular purpose, at a particular time of day and using a particular travel mode from small spatial units (often described as travel analysis zones (TAZ)) in a region. The unit of analysis in the trip-based approach is a TAZ [*Castiglione et al., 2015*]. Compared to the trip-based approach, AB modelling approach is disaggregated and the unit of analysis is an individual rather than a spatial unit. The approach provides flexibility to test a large variety of emerging policies and then able to provide the impacts of these policies on a subsequent travel pattern/routine of an individual in a day [*Axhausen and Gärling, 1992, Siyu, 2015*]. For example, increase in bus frequency of a particular route may encourage an individual to take bus for the morning commute to work, however, this may affect individual routine in a way that he may undertake another car tour in the evening for shopping after reaching home, however, before that intervention he may be used to perform this shopping activity on a return journey to home. Trip-based approach cannot capture such effect. Another promising feature of the activity-based approach is to integrate it with dynamic traffic simulation model in a way that individual identity would be kept intact while execution of the travel between the two activities [*Adnan et al., 2016*]. In this way, it is also known which individual is travelling on which link of the road network. Based on FEATHERS and MATSIM integration this can be achieved at the demand side as well as supply side. FEATHERS [*Bao et al., 2016*] is a demand-side activity-based model, while MATSIM [*Horni et al., 2016b*] in this integration will work as a supply-side tool that executes individual plans on the road network. This is important for iSCAPE because the goal of WP 4 is to test such mobility-oriented policies that can help improve air quality and reduce personal exposure. The approach followed by using an integration of FEATHERS and MATSIM provides such an output that can be utilised not only for estimation of emission and pollutant concentration but also help finding out individual exposure to pollutants in a variety of policy scenarios. Using such an approach exposure to vulnerable individuals can also be obtained, as individual identity is known at all stages of movement in the simulation process [*Shabanpour et al., 2017*]. Based on the above background, this deliverable reports the design aspects of such simulator in detail.

This section further describes the contextual details of this deliverable. Furthermore, it describes the scope under which this deliverable is prepared keeping in view the DOW mentioned for the WP 4 task 4.2.2.

2.1 Scope of the deliverable

This deliverable is prepared as an output of task 4.2.2 of WP 4, that aims to develop a behavioural simulation model that uses activity-based approach. This simulator is then used for exploring different mobility-based policy scenarios for the environmental analysis and also exposure assessments. The major purpose of this deliverable is to present the complete design of the mobility simulator, in relation to the required inputs, highlighting the estimation process of models with the available data and algorithms used to process inputs and outputs. The following points mention the complete aims of this deliverable, in-line with the DOW:

1) DOW for task 4.2.2 focuses more on the advantages of using activity-based approach for behavioural simulations as it provides a more appropriate analysis of exposure assessment by considering individual movement in space and time as opposed to a static approach. Furthermore, it also provides a more enriched approach for analysing impacts of a variety of mobility interventions. Based on this, the scope of this deliverable is to explain with more details the activity-based approach and how it will be used for simulating iSCAPE cities.
2) The design of the simulator is discussed in relation with the structure of the model system, available data and its pre-processing to get the required input for the simulator. Furthermore, this deliverable also highlights the relevant parameters which are helpful to calibrate the system. It is important to note that this deliverable only provide procedural details of various mechanism/processes used within the behavioural simulator. The actual results can be seen in deliverable 4.4.
3) Because of the data-driven nature of the activity-based model and the limited data availability from the iSCAPE cities, this deliverable also discusses the approach used for developing light activity-based model for the selected iSCAPE cities with appropriate justifications that fulfils the purpose stipulated in the DOW.
4) The deliverable also discusses the methodology adopted for exposure assessment as an outcome Tasks which aim to explore the impact of policy options in reducing air pollution; these results will form the basis to fulfil the commitments made in DOW under WP 7 task 7.3 which aims to assess exposure in relation to their socio-economic classes and further expand that estimate to ascertain health impacts.

2.2 Layout of the report

The report is structured into various key sections. Section 3 presents the operation tools that will be utilized to develop the integrated simulator i.e. integration of FEATHERS behavioural simulation and MATSIM traffic simulation frameworks. This section also describes the kind of policies that are possible to examine using the simulator. Section 4 discusses details about the approach followed to develop the complete model for the city of Hasselt and also those to develop a light version of the activity-based model for other selected iSCAPE cities based on data availability. In addition to this, details are also provided for the tools used for pre-processing the data to convert them into the format appropriate as input for the simulator. Section 5 discusses the integration of the model output as input for the development of the emission inventory to be used to feed the successive dispersion model; additionally, this section also discusses a methodology for estimating the exposure based on the cumulative outputs from the integrated behavioural and environmental model. Section 6 draws the main conclusions of the deliverable and puts forward some recommendations for the whole iSCAPE project.

3 Behavioural Simulations- Operational Tools

This section focuses more on brief details of the operational tools used for the development of an integrated behavioural simulator. The scope of the behavioural simulator is to provide information on traffic to be used easily as input for the development of the emission inventory and the following dispersion model. Furthermore, the output can also be used for exposure analysis as it also contains information of the entire itinerary of each individual in the considered city (such that their participation in out-of-home activities and related travel). Individual-level based simulator is usually known as a micro simulator in transport modelling literature. The final output available for the development of the emission inventory and the dispersion model is in the form of traffic volume on the road network; however, the advantage of using activity-based approach is to use the results of the dispersion model and individuals' relevant position in the network to estimate their exposure to air pollutants more appropriately. This section briefly discusses the operational tools utilised to develop an integrated behavioural simulator.

3.1 Policies/Interventions for simulation and their scope

The major purpose for developing the integrated behavioural simulator is to test the impact of behavioural changes and traffic management policies firstly over the traffic activity data on the road network and secondly but more importantly over the urban air quality. In relation to this, deliverable 4.2 already explained the cost/utility functions incorporated in the simulator and already discussed a large variety of emerging mobility relevant policies that can be helpful to improve the air quality of the city. Deliverable 4.2 describes thoroughly these policies/interventions along with the necessary steps required for their implementation within the simulator. The readers are directed to deliverable 4.2 for a more comprehensive illustration of the policies; however, to ensure smooth readability, below we provide a list of these policies:

- Enhancement of public transport infrastructure
- Reduction in public transport fares
- Restricting car use in core city areas
- Road pricing
- Flexible working hours or telecommuting
- Opening/Closing times for Shopping malls/shops
- Parking regulations

The first order impact of the above-mentioned interventions/policies are on the individual activity-travel behaviour, e.g. a change in the choices of mode/ route/ departure times or in an activity sequence. Because of the changed behaviour of the individuals in an area, the second order impact of these interventions may be on the overall urban air quality. In turn, this may affect the exposure of inhabitants. The above discussion describes the scope under which relevant policies are to be tested; the behavioural simulator should be developed and designed according to these requirements. Traditional transport models which are widely used all over the world are aggregate in nature, following a notion of trip-based approach [*de Dios Ortuzar and Willumsen, 2011*]. These models predict the number of trips on the basis of defined spatial units (also called traffic analysis zones (TAZ)) of the study area and then distribute these trips by developing origin-destination matrices classified according to the vehicle types. The assignment model then takes this input to assign traffic on the network to ascertain volume (in units e.g. vehicles/hr) on the different links of the road network. The implementation of a particular policy/intervention leads to increase/decrease in the number of trip counts, their change in distribution within defined spatial units, or change in modes shares and links volumes. These models cannot provide impacts of policies onto changes in individual mobility behaviour due to their aggregate nature [*Ben-Akiva and Bowman, 1998*]. Activity-based microsimulation models are instead more detailed and keep track of individuals considering all aspects of an individual activity-travel. However, these models can only predict activity-travel schedules and require an assignment model for defining routes of travel. Integrating the activity-based model output with traditional traffic assignment model may cause loss of entire detailed information along with individual identity on the road network [*Lu et al., 2015b*]. More appropriate integration should allow to maintain this information when an individual activity-travel pattern is executed on the network. MATSIM [*Horni et al., 2016a*] and SimMoblity [*Adnan et al., 2016*] are two open source platforms that provide a framework within which individual identity can be retained at the supply side. The documentation available for the use of MATSIM is more detailed and is being used quite widely by the transportation researchers and planners, and therefore it is selected as part of developing a behavioural simulation platform.

3.2 FEATHERS-Activity-based Model Design

The structure of FEATHERS model is explained in detail in deliverable 4.2. It actually follows a rule-based approach and the model system is based on decision trees. According to [*Baqueri et al., 2019*], within FEATHERS in order to predict the individual travel routine of the day, the model structure is such that decisions are taken for the day pattern of an individual first e.g. decisions about the number of work episodes, home-based tour and their types, and intermediate stops are taken. These decisions for an individual are then further enriched by taking decisions about activity duration and location choice for the primary activity of the tour and then these decisions are made for the secondary activities in the tour. Decisions in relation to the activity start time and mode choices are then taken to complete all the scheduling dimension required to make the activity schedule of an individual. The process is well described through a framework presented in deliverable 4.2 (Please see figure 1, pg: 12). In relation to this deliverable, the major emphasis is to design the process through which the simulation framework can be used. The following points are important to describe the design:

1) Estimation of the models with the relatively new dataset to predict activity-schedules for the whole Flanders region that include the city of Hasselt.
2) Other primary settings of the simulator and the required data (e.g. land use data, zoning structure, travel skim matrices and their resolution)
3) Preparation of a study area population along with socio-economic characteristics

Section 4 of this report describe the above points with more details particularly in relation to the available datasets.

3.3 FEATHERS-MATSIM Integration

As described in section 3.1, the integration of the FEATHERS with MATSIM models is inevitable in order to execute the schedule onto the road network to determine traffic volumes on each link and to maintain all information generated from FEATHERS. MATSIM requires as input the FEATHERS schedule along with other input data; the model optimizes the schedules to provide traffic volumes on the road network. In the literature, MATSIM itself has been recognized as an activity-based model, however, the behavioural basis on which MATSIM optimizes schedule is weak compared to the standalone Activity-based model. MATSIM uses a score function to optimize schedule in relation with network situation, and through optimization it can change a variety of scheduling dimensions such as routes, transport mode, locations, activity duration, its time of day and also add and delete activities from the pattern. However, it can also act as an assignment model where only changes in route are allowable while other scheduling dimensions can be kept constant. A cyclic loop can be established between activity-based model and MATSIM, where changes of routes are allowed on MATSIM side (i.e. supply side) and changes in other dimensions can be allowed within an activity-based model (i.e. demand side). The system works under iterative framework to sought equilibrium between the demand and supply side. In the literature [*Adnan et al., 2016, Lu et al., 2015b, Auld and Mohammadian, 2012*], this is known as an integrated simulation framework.

In case of Hasselt city, the simulation will undergo this integrated framework. However, for other iSCAPE cities the area-specific inputs are provided to MATSIM, as it will simulate the cities standalone by allowing changes in all scheduling dimensions. This can be considered as a light activity-based model.

Apart from the point mentioned above, the following design aspects will be described in more detail in section 4.

1) Road network preparation for iSCAPE cities
2) Preparation of Public transport network and bus schedules and matching bus stops location onto the prepared road network for iSCAPE cities
3) In case of light activity-based model, preparation of schedules considering the available data from the iSCAPE cities.
4) Development of calibration mechanism

The approach followed in the development of the above process is to use as much open source available data as possible, so that the process is flexible to be implemented and data used for the simulator is in required standard formats.

4 Behavioural Simulation- System Design

Section 3 discussed the type of operational tools that will be employed for the behavioural simulations along with their role and scope. This section provides more details on different aspects of the design in relation to the city of Hasselt and to other iSCAPE cities.

4.1 HASSELT Case

For the Hasselt case, as mentioned earlier, a detailed Activity-based model is developed which is integrated with MATSIM to obtain traffic (volume) on the road network. This is explained in the next few sub-sections.

4.1.1 Available data

A variety of datasets has been used to estimate sub-models within Activity-based model and also to calibrate the outputs derived from the FEATHERS-MATSIM Integration. Datasets used for the Hasselt case are as follows:

- Onderzoek VerplaatsingsGedrag (OVG) Vlaanderen dataset: This dataset contains details of the personalized mobility routine of sampled individuals in the Flemish region of Belgium. The survey includes Flemish residents aged 6 and older as study population. The survey contains information about household characteristics of sampled individual, their personal characteristics and at least one-day trip diary information [*Cools, 2013, Kochan et al., 2013*]. These surveys have been usually conducted periodically once a year since the year 1995. However, there are sometimes gaps of more than a year [*DMPW, 2019*]. The dataset of around 8800 individuals has been selected to train the sub-models of FEATHERS.
- GPS (Global Positioning System) based activity-travel routine data for Hasselt residents. This data was collected as part of task 4.1 of WP 4. More details can be obtained from deliverable 4.1. FEATHERS models are adjusted further with this data for appropriate prediction of schedules,
- Census data of 2015 conducted in Belgium, along with data that contains spatial unit-based (zone-based) marginal distribution of different population characteristics. Because of the privacy issues with complete census data, the marginal distribution data is used to develop a synthetic population of the study area. For this purpose, an iterative proportional updating algorithm [*Cho et al., 2014*] is being used to generate the synthetic population of Flanders.
- Detailed land use data in terms of residential units and firms, shop addresses within Flanders. This detailed land use data is helpful to assign exact locations within TAZ. This data is known as Centraal ReferentieAdressenBestand (CRAB) and is authentic source of addresses in Flanders [*CRAB, 2019*]. It contains all the official

addresses and their geographical locations and is administered by the Flemish towns and municipalities. It is available online and open to everyone[1].

- Skim Matrices (TAZ- based) containing travel time and cost for different modes of travel in an entire Flanders region. This data is developed using personalized travel survey [*Kochan et al., 2013*]. This is required for estimation mode choice, location choice and time-of-day models within FEATHERS.
- OpenStreetMap (OSM) data, which is open access data. This data is converted after necessary cleaning to provide a road network to MATSIM.
- General Transit Feed Specification (GTFS) data: this data contains information about public transport routes, stops and their schedule in a standard format. This data gives the necessary details to establish public transport network as an input to MATSIM after employing the necessary cleaning and integration process with OSM data.
- Traffic volume data on important links (Motorways, national roads) of Flanders for calibration of the final output from MATSIM.

4.1.2Data Processing Procedures and Model Development

A variety of data processing procedures has been used to develop input for the two tools (i.e. FEATHERS and MATSIM) and also to use and to calibrate these input models. Figure 1 illustrates this in a more comprehensive manner. Rectangular shapes in Figure 1 present procedures and oval shapes present inputs/outputs for a particular procedure. Arrows show the data flow from one procedure to another in the simulation framework. The next few sub-sections describe the important procedures/algorithms and models that are developed and used as part of developing a simulation model for the Hasselt case as mentioned in figure 1.

[1]https://overheid.vlaanderen.be/informatie-vlaanderen/producten-diensten/centraal-referentieadressenbestand-crab

Figure 1: Schematic design of behavioural simulation framework for Hasselt Case

4.1.2.1 Iterative Proportional Updating Algorithm

The Iterative Proportional Updating (IPU) algorithm has been developed to generate a synthetic population to be used within FEATHERS for simulation. The algorithm code has been developed at IMOB, Uhasselt, in Python language. This algorithm was first developed by [Ye et al., 2009]. to match both household and person marginals by updating sample household weights. In the fitting step, household and person type constraints are estimated using the fitting procedure, followed by the calculation of sample household weights by the IPU algorithm [*Cho et al., 2014*]. Based on the household selection probabilities, households are sampled and expanded and finally the population of household is generated. Sampled household weights generated in the fitting steps are used to generate household selection probabilities.

IPU algorithm requires two types of data: seed data and target marginal information. The seed data contain disaggregate population which normally describes enough details of the population (e.g. household and personal level information together with socio-economic characteristics), but there are only a small number of individual elements in the seed data. On the other hand, target marginal has only information about the sum of attributes in a spatial unit e.g. total number of males, total number of females, number of individuals in a particular age category, number of households with one car etc. The goal of IPU is to use the seed data in a way that it expands the sampled household and persons to fit into all considered target marginal attributes of a particular TAZ. The following simple process is used for zone-to-zone fitting.

1. Reading seed data in a study area of interest, on a TAZ level.
2. Reading target marginal in a corresponding area.
3. Fitting the seed data to the target marginal.
4. Drawing a synthetic population for the study area.

To check the accuracy of estimation, the synthetic population needs to be validated against real population based on some goodness-of-fit measure. We used standardized root mean square error (SRMSE) as a validation statistic presented as Eq.1.

$$SRMSE = \frac{\sqrt{\sum_i \sum_j \frac{\left(t_{ij} - \widehat{t_{ij}}\right)^2}{m.n}}}{\sum_i \sum_j \frac{t_{ij}}{m.n}} \qquad (1)$$

where $\widehat{t_{ij}}$ is the estimated number of population elements with attributes *i* and *j* and t_{ij} is an observed number of population, *m* and *n* are the number of attributes values for attributes *i* and *j* respectively. A zero value of SRMSE means a perfect match, while a value of ‘1’ indicates no matching between estimated and observed data.

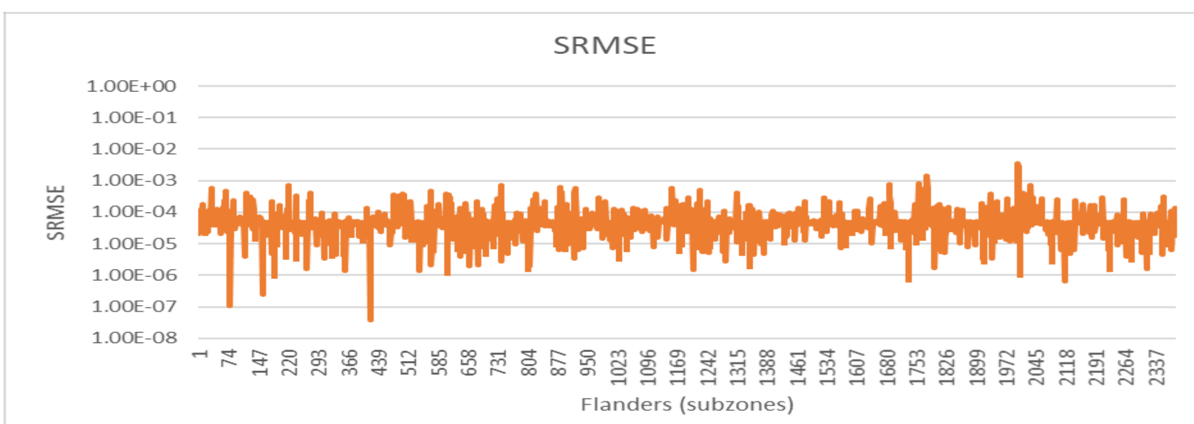

Figure 2: Goodness of fit (SRMSE) for each Zone

This algorithm is used to generate the synthetic population of entire Flanders. The highest SRMSE obtained is of about 0.0028. Figure 2 provide the plot for such values for all subzones of the Flanders that also include zones for the city of Hasselt. The output contains the Household and person .txt files of the study area. This is an important input for running a simulation within FEATHERS.

4.1.2.2 OSM and GTFS data Integration for MATSIM

This process is about the preparation of the road network using OSM data and the successive integration of OSM data with GTFS data to obtain the bus stops mapping on the road network. In this way, a road network is prepared where not only cars can travel but at the same time buses can travel according to their defined routes and schedules. This is an important input for running a MATSIM simulation for a study area. Both OSM and GTFS data are openly available and therefore the process developed and outlined here can be utilized for any city for which both datasets are available in the same formats. A comprehensive algorithm has been developed at IMOB, Uhasselt, in Java language. We describe here a few details; however, interested readers are directed to [*Vuurstaek et al., 2018*].

The first major step of the algorithm is to prepare a clean and reduced version of the OSM data. Cleaning here means that the information available will be enriched, corrected and completed based on some straight forward rules. Reduction indicates the removal of not useful objects from the network with the goal to produce a strong connected graph. In addition to this, OSM network data are prepared so that roads with traffic flow in both directions are specified once as forward and once as backward to ensure that it connects well with the GTFS bus stop location data on both sides of the road separately.

In the second major step, bus stops are extracted from GTFS while removing anomalies, with the purpose to find a set of candidate links in the OSM data for each bus stops. This is done based on the concept of *projection of point on a curve* and *projection of point on network.* The bus stop mapping algorithm is based on the idea that public transport operator minimizes the total distance driven to complete the bus trips. Several numerical experiments were performed to see the performance of this algorithm. The validation results have shown good accuracy of the mapped bus stop with their actual location. Other details of the algorithm are also provided in supplementary publications [*Cich et al., 2016*] and [*Vuurstaek et al., 2017*].

The algorithm is prepared in a way that it can generate the input data for MATSIM simulation of all iSCAPE cities. In addition, it can be easily implemented in other cities provided that OSM and GTFS data is available.

4.1.2.3 FEATHERS Model Estimation their Adjustment

In general, the estimation of the activity-based model is done via three major datasets i.e. activity-travel diary along with socio-economic attributes, land use data and skim matrices based on a particular zoning system. The activity-based model itself includes a variety of sub-models, starting from an activity-pattern model, and then tour scheduling models for primary activity (such as mode, location and time-of-day) followed by scheduling models for intermediate stops in a tour. FEATHERS platform is developed in a way that all input data is provided in a particular standard format (e.g. trip/activity data along with socio-economic attributes, land use data and skim matrices along with zonal shape file), and the platform itself estimates decision trees for all sub-models in the system. As explained in deliverable 4.2, a decision tree is developed by recursively splitting a sample of observations into increasingly homogeneous groups in terms of a given response variable [*Kochan et al., 2013*]. FEATHERS utilizes Chi-square automatic interaction detection (CHAID) algorithm which evaluates splits based on a Chi-squared measure of significance of differences in response distribution between groups. Socio-economic attributes are more important for initial activity patterns decisions; however, land use and skim matrices are also important for decisions such as choices of transport mode, location, duration and time-of-day. In addition to this, sub-models in FEATHERS are arranged in a hierarchical order, and estimation of a lower order sub-model also uses variables that are derived from earlier models to represent the strong connectivity between a variety of scheduling decisions. For example, the location of activity is determined prior to the mode used to travel for that activity. Therefore, attributes defining the location (distance from the origin, location attractiveness measure etc.) are part of the variable through which decision tree for the mode choice is estimated. Estimation results and goodness-of-fit of the estimated decision tree can be examined via prediction accuracy from the contingency matrices.

The FEATHERS platform is developed at IMOB, UHasselt and based on the new datasets obtained for the Flanders region, all the sub-models (total of 28) were re-estimated. The complete list of the models is available in table 2 of deliverable 4.2. The available datasets (i.e. OVG, skim matrices and land use data) are processed in order to meet the format required by the platform. Goodness-of-fit of all estimated models are examined, and it is found that only a few models (three) have a prediction accuracy of less than 60%. These models are lower level models in the hierarchy and the primary reason for such low accuracy is the availability of enough number of observations for such decisions in the data sets (e.g. start time of secondary activity in the 2nd half of a home-based tour). All models related to the primary activity of any type of tour have prediction accuracy of more than 80%. The outputs obtained from the model are then used to compare with the test data set kept aside. The accuracy of prediction is found more than 90% for aggregate measures (such as number of tours, mode shares distribution) and for some disaggregate measures it is noted around 60-75%, this is considered reasonable for such an extensive data-driven models[*Petrik et al., 2018*].

The FEATHERS model is then simulated using a segmented synthetic population for Hasselt city and schedules of individuals are predicted. These schedules are then compared with previously collected GPS based activity-travel routine data to validate schedule prediction for the city of Hasselt city. This comparison is made via the distribution of several scheduling dimensions of their travel pattern such as primary activity duration distribution, transport mode distribution for primary activity etc. Relevant FEATHERS decisions trees are slightly adjusted in terms of the splitting value of the attributes in the decision trees to minimize the differences between predicted and observed distributions. The process is done by trial and error method, as running FEATHERS again and again for such a small segment of the population is less time-consuming.

4.1.2.4 MATSIM-FEATHER Integration

The first step towards the integration is to provide MATSIM with an input in the form of schedules of a population of the study area. MATSIM is based on a microsimulation framework, not only in terms of individuals but also for the very fine resolution of activity locations. While a schedule (also termed as a plan in MATSIM) is executed in the supply side of MATSIM for a particular individual, activity location (destination) at the end of the trip is considered as the starting node of the last link of the route from an origin location to destination. However, the output available from FEATHERS is in terms of zones (that contains many links of the road network). In order to define activity location in the given schedules from FEATHERS, there may be a variety of approaches that can be employed. The simplest is to use any random nodes of the MATSIM network within the zone to assign a given location. Another methodology could be to identify node based on a distribution of attraction measure (e.g. POI density of a region within the zone), and then using a Monte Carlo simulation to assign the region first and then randomly select a node within that region. Given the availability of detailed land use data in the form of CRAB dataset for the entire Flanders region, we then followed a procedure to assign to each nearest node in the MATSIM network an address according to its land use given in the CRAB. These nodes are then assigned to a particular type of activity location in the schedules randomly within the zone. For example, if for a given trip from a schedule of an individual the destination of the trip is work in a particular zone, then from the set of nodes within that zone that have been given a work type land use, a node is selected randomly to assign a destination of that trip. It is important to note that the CRAB data does not provide information about the number of people living or working at a particular address. Therefore, during the allocation of activity location at a particular node it may happen that a particular node has been allocated more or a smaller number of trips end of individuals as compared to the actual capacity of that node. Notwithstanding this limitation, this is the best we can achieve in terms of location accuracy of activities. With the use of this procedure for Flanders, case schedules from FEATHERS are changed into individual plans.

The second major step is to run MATSIM simulation based on all generated inputs. MATSIM itself follows an iterative algorithm, where not only route assignment process is employed but it calculates scores of alternative plans of an individual (the alternative plans are generated by changing activity start times, activity duration, modes for trip, activity sequence and routes based on an input plan of the individual). MATSIM employs a co-evolutionary algorithm to optimise the plan based on the network situation and attempt to establish an equilibrium. Usually, 250 or more simulations runs are recommended to achieve optimized results. It should be noted that MATSIM does not provide any visualization capabilities to examine the results and requires further processing of the results to get the desired outputs. At IMOB Uhasselt, we developed a macro to obtain 24-hour traffic counts (cars and buses) on each link of the road network. In addition to this, usually, a portion of the population (selected randomly, around 10% to 30% depending on population size) is simulated within MATSIM to allow for the lower run time. The results are then extrapolated further. Based on MATSIM output, new skim matrices are generated (travel times on zonal basis) and then used to run FEATHERS again. The FEATHERS-MATSIM cycle is then run a few times to attain global equilibrium. However, to achieve this faster in a few runs we employed a calibration process which is explained below.

4.1.2.5 Calibration Process

The calibration process is developed in a manner to change only a few important parameters within the MATSIM configuration file and road network parameters. The parameters used for optimization are related to mode utility score function (mode params) with the aim that travel mode use distribution in the study area is satisfied. Along with these parameters, other sets of parameters used for calibration are link capacities (given in the road network data) and parameters related to early departure and late arrival. The objective function is the ordinary least square (OLS) type that minimizes the difference between the observed traffic volume and predicted traffic volumes (only cars) on links where data is available. The algorithm used to find out the new set of parameters to be used for the next calibration iteration is SPSA (simultaneous perturbation stochastic algorithm). The algorithm is useful for optimization of complex problems involving range of processes to obtain final outcomes [*Lu et al., 2015a*]. It calculates the step size that is required to be added into the previous value of the parameters (to obtain the new one) by estimating an approximate difference of objective function by simultaneously perturbing all parameters. More details can be found in [*Lu et al., 2015a*], where they have demonstrated the use of this algorithm for calibrating large-scale traffic assignment model. It should be noted that within one calibration simulation there are 250 MATSIM simulation runs, and therefore, the process is quite time-consuming.

For the Flanders case, we ran only a few calibration-based simulations, and once obtained reasonable results, new skim matrices are generated from MATSIM output and compared with the matrices used to run FEATHERS. We found that the cumulative difference between the two matrices is less than 20% and therefore, to reduce the computational effort, we assumed that the FEATHERS-MATSIM cycle has already achieved an equilibrium. This is because the observed traffic volume data used for calibration and zonal travel skim matrices of Flanders which is used to run FEATHERS are compatible with each other. Therefore, once obtained the calibrated network volumes, travel time matrices are very much equivalent to that previously used for running FEATHERS. Results of the calibration are illustrated in more details in deliverable 4.4. In this deliverable, we are only describing the procedural details. Figure 3 presents the base-case outputs in the form of rush hour traffic on the road network of the simulation after calibration for the city of Hasselt.

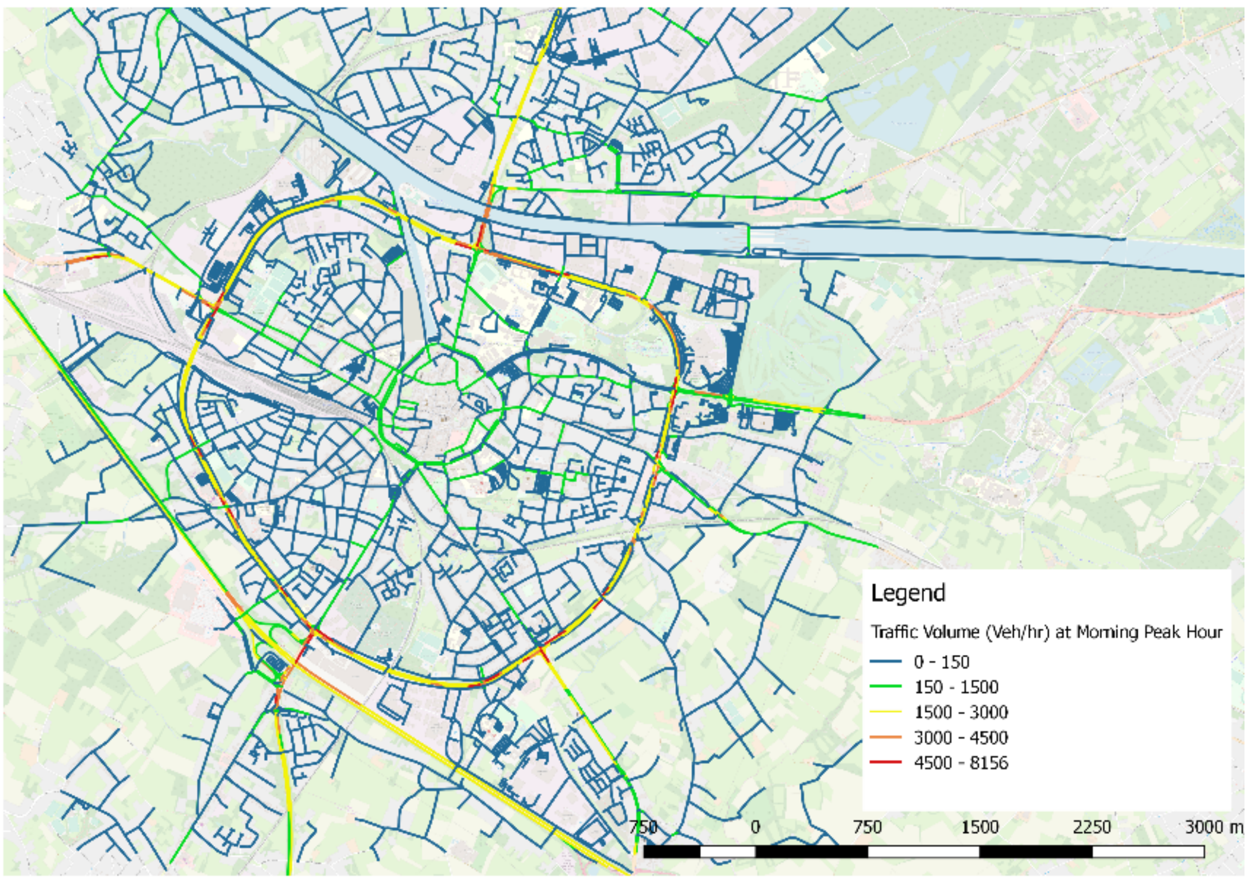

Figure 3: Morning peak hour traffic volume (vehicles/hr) on reduced Hasselt Network

4.2 Other iSCAPE Cities

Because of the unavailability of a detailed and complete data set as that available for the Hasselt case, a different procedure was adopted to perform behavioural simulation for other iSCAPE cities. Instead of using FEATHERS, for other iSCAPE cities, we mainly relied on MATSIM. In addition to this, we also relied mainly on open access data or data which is easily available from city agencies repositories. At the same time, we ensured to make the inputs for MATSIM as detailed as possible. As mentioned in the DOW, the behavioural simulations along with the development of emission inventories and dispersion models are limited to selected iSCAPE cities due to a variety of practical limitations. We have selected three iSCAPE cities to cover different environment and climate conditions by selecting one city from north, central and south part of Europe within the consortium. Specifically, we have selected Vantaa from Finland, Hasselt from Belgium and Bologna from Italy. For the purpose of this deliverable, the Hasselt case is described in great detail in section 4.1. In this section, the design aspects of the behavioural simulator for Bologna and Vantaa are described.

4.2.1 Available data

The available dataset from Bologna and Vantaa are described in detail here:

- *Census and target marginal data for administrative units:* In the case of Bologna, the target marginal data is available mainly for different attributes, such as Gender, Age and family size for around 2333 census sections. Furthermore, other population statistics are derived from various online resources [*Comune_di_Bologna, 2019*]. In the case of Vantaa, along with Gender and Age, other socio-economic attributes are available as target marginal for administrative units within the city (This dataset is available from respective city agency and supplied by the consortium partner FMI). Along with this, other population statistics were derived from online resources. These datasets helped to develop a synthetic population of the two cities.
- *Zones and administrative boundaries:* In the case of Bologna, the spatial units or zones used as TAZ are represented 2333 census sections. Furthermore, the shapefile of the census sections is also downloadable from the online resource [*Comune_di_Bologna, 2019*]. Similarly, in the case of Vantaa, the city administrative spatial units are much larger compared to Bologna. The city is divided into 60 zones (also called districts) and the shapefile is also received from the respective consortium partner. This is required for allocating activity locations in terms of zones for predicted schedules.
- Land use data for Bologna case is also available from [*Comune_di_Bologna, 2019*]. In addition to this, point of interest (POI) data using Google API for various land use were also collected for Bologna and Vantaa. This POI density data is used for assigning the location of activities (i.e. home, work and other activity locations) within various zones.
- GPS based activity-travel routine data for few Bologna residents is available for two weeks. This data was collected as part of task 4.1 of WP 4. More details can be obtained from deliverable 4.1 and also from [*Ahmed et al., 2018*]. Light activity-based model is then adjusted further with this data for the appropriate prediction of schedules for Bologna. In the case of Vantaa, this dataset is not available, however, for Helsinki region (a major city near Vantaa), a variety of travel and activity related research has been published. Schedules are adjusted in light of such published data.

- OpenStreetMap (OSM) data for both cities is downloadable. This data is then converted after necessary cleaning to provide a road network to MATSIM.
- General Transit Feed Specification (GTFS) data for both cities are available. This data contains information about public transport routes, stops and their schedule in a standard format. This data provides the necessary details to establish public transport network as an input to MATSIM after employing the necessary cleaning and integration process with OSM data.
- Traffic volume data on important links of Bologna city is available from consortium partner UNIBO. In a similar manner instead of traffic volume, network travel time for important links is available for Vantaa from FMI. These datasets are used for calibrating the final output from MATSIM.

4.2.2Light Activity-based Model

In absence of detailed activity-travel diary data for both cities, we have developed a light activity-based model based on certain rules (described as under) for estimating schedules for the population. A simpler version of a similar algorithm was developed for estimating schedules for the student population in Flanders [*Adnan and Knapen, 2017*]. The heuristic algorithm is as follows:

- Schedules obtained for Flanders case of FEATHERS are used as Schedules for Bologna and Vantaa after making a variety of adjustments.
- FEATHERS schedules (in terms of activity sequence, such as home-work-home or home-work-shopping-home) frequencies are grouped into the type of Population characteristics of Flanders (Gender, age, Family size, etc) that are available for Bologna and Vantaa. These are developed as multivariate joint distribution. For example, for a specific value of gender, age and family size, relative frequencies are estimated for various type of activity sequence patterns.
- Using a Monte Carlo simulation method, for Bologna and Vantaa, a synthetic individual having a particular population characteristic has been assigned a particular activity sequence type.
- Once a sequence type is assigned, for each activity type distribution of duration, activity start time is derived from the collected GPS data from Bologna. Based on these distributions, again using a Monte Carlo simulation method, duration and activity start time are allocated to each activity in the activity sequence. Similarly, from the travel mode distribution for different activities, a travel mode is also assigned for a trip between two activities. In the case of Vantaa, a similar method is followed using published data from Helsinki region and overall Finland from these studies [*Pääkkönen, 2013, Karasmaa and Pursula, 1997, Helminen and Ristimäki, 2007, Karasmaa, 2007, Durand et al., 2018, Ng and Acker, 2018, Czepkiewicz et al., 2018*].
- The next major step to allocate activity location zones is based on the residential POI density of each zone, where all individuals are assigned a particular zone (for home location) in the respective city. A similar approach is followed for allocation of work activity location in terms of zones. For secondary activities of the tour (e.g.

shopping, leisure, recreational), again a similar method is followed to allocate zone from the set of candidate zones. The candidate zones, however, are ensuring that the zone location is within the threshold limit (i.e. 15 km for Vantaa and 10 km for Bologna) from the origin zone. This is to avoid giving farthest zone to the secondary activity of the tour.

- Some consistency checks are carried out to finalize the schedule. These are as follows:
  - Bicycle and walk trips are not more than 8km and 2 km respectively.
  - If the personal vehicle (i.e. car/bicycle) is selected for a particular tour, it is ensured that all other trips of the tour are made from the same mode.
  - It is also ensured that cumulative mode shares for these schedules followed other sources of data available from [*Comune_di_Bologna, 2019*].
  - Activity duration of the second and last activity in the pattern is adjusted as such that the last activity *home* can be started as late as 3:00 pm next day.

4.2.3Data Processing Procedures and Model Development

Figure 4 presents the behavioural simulation framework to simulate the other two iSCAPE cities. The framework follows the similar processes as those described for the Hasselt case. However, in this section, only major differences are highlighted.

Comparing the two frameworks adopted to carry on simulations in Hasselt and in other iSCAPE cities, it is clear that majority of steps are almost similar; however, instead of using FEATHERS, a light Activity-based (AB) model is used, as explained in section 4.2.1. Another major difference is the absence of integration link between the MATSIM and AB model. This is because the light version of the AB model will not enrich any behavioural notions in the predicted schedules once the new skim matrices are generated from MATSIM. Therefore, MATSIM here is a standalone important model where individuals' plans are optimized based on the network conditions. Another different characteristic here is that instead of using detailed land use data, we are using POI density of different land use within a particular zone to assign a network node for a particular activity location in the schedule. The population developed for the two cities are having SRMSE values of around 0.14 and 0.23 respectively. These are still good values given that required seed data was not appropriately available.

The outputs then provided to develop the emission inventory to fed in the dispersion model are in the form of a road network that contains traffic volumes of cars, buses and other heavy vehicles. Within the simulation process for the Hasselt case, an OD matrix was available for freight vehicles. This matrix was independently assigned to the same road network used for MATSIM. The result of this assignment was then later combined with the outputs of MATSIM. However, in the cases of Bologna and Vantaa, due to the absence of this data, we have assumed freight vehicles as some percentages (depending on the type of road) of car traffic. Furthermore, this volume is also calibrated with the available traffic volume data for Bologna. For Vantaa, we followed some available publications providing also some details of freight traffic in Finland.

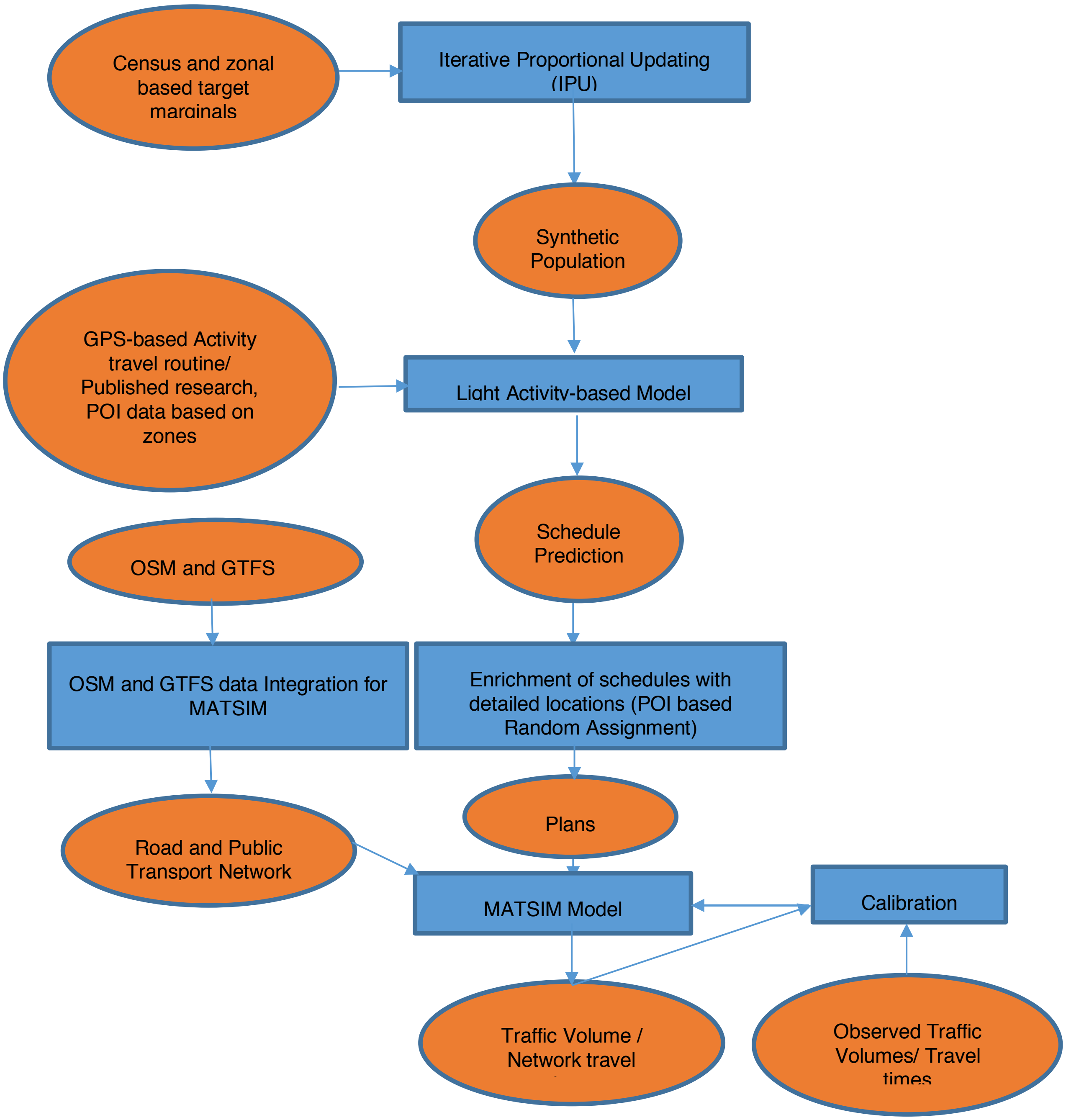

Figure 4: Schematic design of behavioural simulation framework for other iSCAPE cities

5 Exposure Assessment model

This section presents a few details about the following processes once the behavioural simulation outputs are available. The outputs from the behavioural simulation model are in the form of road network shapefiles, containing information on traffic for morning peak hour of the day. Furthermore, an hourly distribution of traffic normalized with respect to the peak hour traffic in the morning is also provided for the whole day. This data is critical for developing the emission inventory and estimating pollutant concentrations from dispersion simulations. Deliverable 4.5 provides a more comprehensive discussion of the use of the outputs of behavioural simulations to construct emission inventories to be used as input to dispersion models employed for the selected cities of iSCAPE to evaluate the exposure to air pollution within the base and also a few policy scenarios both in the present and in the future scenarios. The discussion here is based on the use of the outputs from the dispersion model to derive individual exposure. Here we are explaining only the algorithmic details, the results of this will be part of deliverable 7.3 of WP 7.

5.1 Exposure Assessment Framework

Individual exposure to pollutants is estimated based on a module that processes individual activity-travel routine data (MATSIM outputs) with the help of 3rd party data (outputs of dispersion model from UNIBO) related to ambient pollutant concentrations. An important feature why MATSIM was used is that along with traffic volume data, it also provides a complete activity-travel plan of each individual in a synthetic population. This output, which represents a complete itinerary of individual starting from home to coming back home at a later time period, can easily be displayed on a GIS platform. The information can be easily converted into a format similar to GPS-based trajectory data of an individual as shown in figure 5. On one hand, outputs of deliverable 4.5 (from UNIBO) provides maps of air pollutants concentrations at the urban scale. The critical process involved in this algorithm is matching pollutant concentration data with activity-travel plans to assess dynamic exposure of a person.

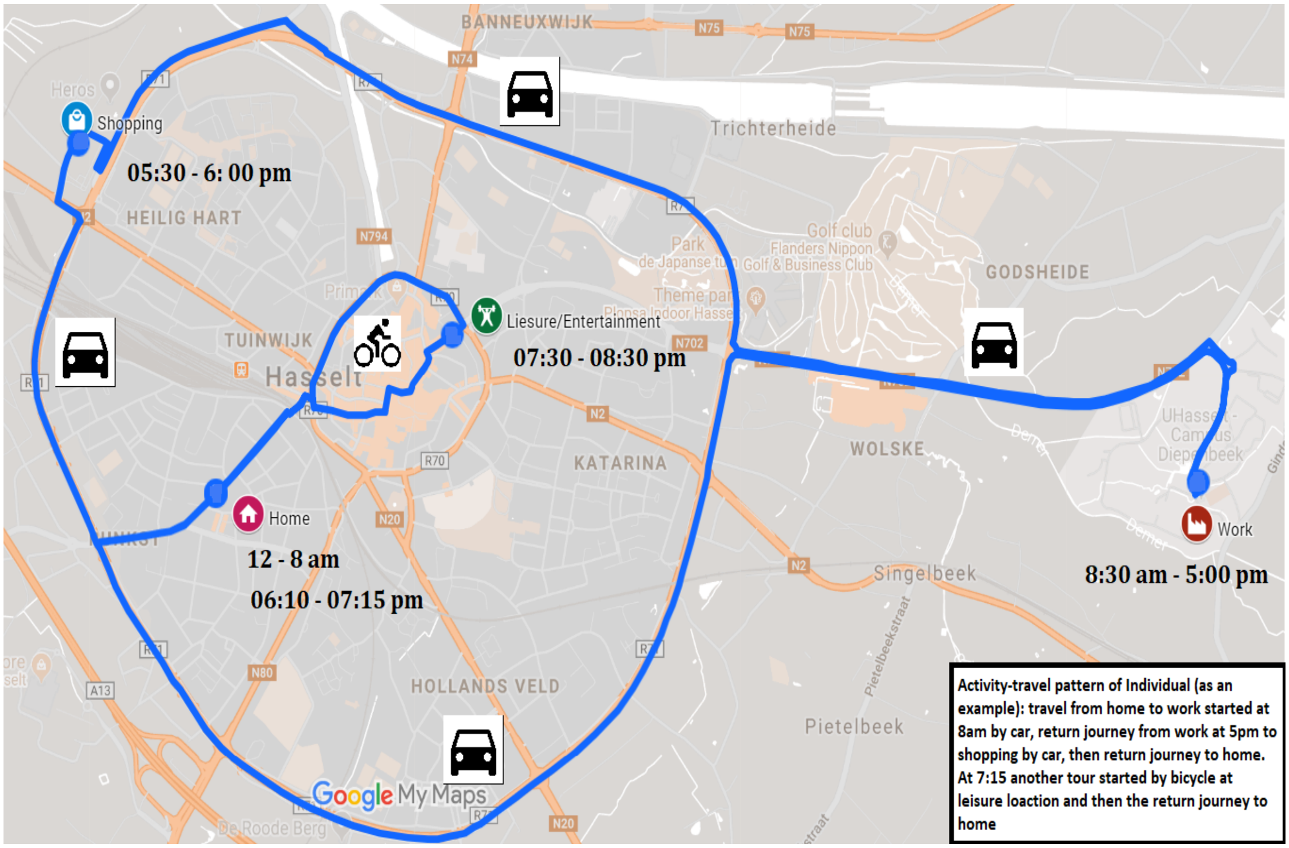

Figure 5: Activity-travel pattern (time-of-day, mode, route, and activity information)

The framework employed here is similar to that used in deliverable 4.1 and also reported in [*Ahmed et al., 2018*]. Specifically, the matching is carried out by assigning each x.y pair of the obtained trip trajectory to the respective grid-based pollutant concentration zones to obtain the exposure at that point w.r.t time. Further details are as follows:

- Exposure of a trip and activity is calculated by taking summation of the time spent in different pollutant concentration levels defined in deliverable 4.5.
- In case of an activity exposure, the spatial location from the predicted output along with the timestamp for each activity is available for each individual. For that particular location and duration of activity at a given time period, changes in the pollutant concentration categories are determined and time spent within each category is noted. The procedure is repeated for all activities individuals have performed in a given day and the results are aggregated for the time spent in each category of pollutant concentration.
- In relation to the trip, each trip is decomposed in several segments on the basis of their association with a particular grid of pollutant concentration (as available from pollutant concentration maps). For the duration of the trip in a given time period, pollutant concentration level belong to that grid is assigned. The process is repeated for all segments of the trip, and then for a particular trip, time duration

spent for each category of the pollutant concentration is determined. The process is repeated for all the trips that individuals have performed in a given day.
- Once duration distributions for trips and activities are available for each individual, it is then further aggregated to represent the overall exposure of activity-travel behaviour in a given day.

According to this algorithm, the appropriateness of the exposure estimation mainly depends on the granularity of the pollutant concentration maps (e.g. their space-time resolution). The higher the space-time resolution the more appropriate the exposure estimate of the individual in the synthetic population. Deliverable 4.5 discuss the basis in more details in relation to the space–time resolution adopted for the development of pollutant concentration maps.

6 Conclusion

The focus of this deliverable is mainly to describe the scope of the behavioural simulation and provide the details of different design aspects of the simulation framework. Much of the discussion regarding FEATHERS and MATSIM, along with an illustration of the implementation and execution of the policy scenarios are already provided in deliverable 4.2. Therefore, in order to avoid duplication, in this deliverable we focused more on aspects showing the use of different processes to obtain the required input data, estimation of models and calibration framework details of the simulator. The actual simulation execution and obtained results for the three cities in the base case and in a few policy scenarios will be reported in the deliverable 4.4.

The processes described in this deliverable are as follows:

1) Use of iterative proportional updating (IPU) algorithm to generate the synthetic population to provide an input for FEATHERS and MATSIM.
2) An algorithm to integrate GTFS and OSM data to obtain a mapping of bus stops to develop an input for MATSIM simulation (this is general algorithm and can be applied to all cities)
3) Development of calibration framework that employs SPSA algorithm
4) Development of a light activity-based model to develop schedules and integration of it with MATSIM simulation platform.
5) Exposure assessment framework based on the output of deliverable 4.4 and 4.5.

Hasselt case behavioural simulation framework is more detailed and comprehensive, but at the same time, the implementation of such a framework is time-consuming. Therefore, a light Activity-based model that uses simplified data is presented. The use of MATSIM in the simulation framework provides flexibility to employ a range of policy scenarios and at the same time, to obtain as detailed as possible results enabling the possibility to extract the impact of traffic management policies on a disaggregate level. This disaggregate level of output will be used to estimate exposure analysis as part of WP 7.

7 References / Bibliography

Adnan, M. & Ahmed, S. 2017. Environmental Effects of Behavioural Actions. *Delivarable 4.1.* www.iscapeproject.eu: Transportation Research Institute (IMOB).

Adnan, M. & Knapen, L. 2017. Predictions of Daily Schedules for Students in Leuven for a FEATHERS-MATSIM Integration. *BIVEC-GIBET 2017 Transport Research Days* Leige, Belgium.

Adnan, M., Pereira, F. C., Azevedo, C. M. L., Basak, K., Lovric, M., Raveau, S., Zhu, Y., Ferreira, J., Zegras, C. & Ben-Akiva, M. SimMobility: A Multi-scale Integrated Agent-based Simulation Platform. 95th Annual Meeting of the Transportation Research Board Forthcoming in Transportation Research Record, 2016.

Ahmed, S., Adnan, M., Janssens, D., Brattich, E., Kumar, P., di Sabatino, S. & Shakshuki, E. M. 2018. Estimating pro-environmental potential for the development of mobility-based informational intervention: a data-driven algorithm. *Personal and Ubiquitous Computing*, 1-16.

Auld, J. & Mohammadian, A. K. 2012. Activity planning processes in the Agent-based Dynamic Activity Planning and Travel Scheduling (ADAPTS) model. *Transportation Research Part A: Policy and Practice,* 46, 1386-1403.

Axhausen, K. W. & Gärling, T. 1992. Activity-based approaches to travel analysis: conceptual frameworks, models, and research problems. *Transport reviews,* 12, 323-341.

Bao, Q., Kochan, B., Shen, Y., Bellemans, T., Janssens, D. & Wets, G. 2016. Activity-Based Travel Demand Modeling Framework FEATHERS. *Transportation Research Record: Journal of the Transportation Research Board,* 2564, 89-99.

Baqueri, S. F. A., Adnan, M., Kochan, B. & Bellemans, T. 2019. Activity-based model for medium-sized cities considering external activity-travel: Enhancing FEATHERS framework. *Future Generation Computer Systems*.

Beckx, C., Int Panis, L., Arentze, T., Janssens, D., Torfs, R., Broekx, S. & Wets, G. 2009. A dynamic activity-based population modelling approach to evaluate exposure to air pollution: methods and application to a Dutch urban area. *Environ Impact Assess Rev,* 29.

Ben-Akiva, M. E. & Bowman, J. L. 1998. Activity Based Travel Demand Model Systems. *In:* MARCOTTE, P. & NGUYEN, S. (eds.) *Equilibrium and Advanced Transportation Modelling.* Boston, MA: Springer US.

Castiglione, J., Bradley, M. & Gliebe, J. 2015. *Activity-based travel demand models: A primer*.

Cho, S., Bellemans, T., Creemers, L., Knapen, L., Janssens, D. & Wets, G. 2014. Synthetic population techniques in activity-based research. *Data Science and Simulation in Transportation Research.* IGI Global.

Cich, G., Vuurstaek, J., Knapen, L., Yasar, A.-U.-H., Bellemans, T. & Janssens, D. 2016. Data Preparation to Simulate Public Transport in Micro-Simulations Using OSM and GTFS. *Procedia Computer Science,* 83, 50-57.

Comune_di_Bologna. 2019. *Open Data Commune Di Bologna* [Online]. Available: http://dati.comune.bologna.it/ [Accessed].

Cools, M. Boosting Response Rates: Evidence from Flemish Travel Behavior Survey. Proceedings of the 92nd Annual Meeting of the Transportation Research Board (DVD-ROM), 2013. Transportation Research Board of the National Academies.

CRAB 2019. Centraal ReferentieAdressenBestand, Vlanderen.

Czepkiewicz, M., Ottelin, J., Ala-Mantila, S., Heinonen, J., Hasanzadeh, K. & Kyttä, M. 2018. Urban structural and socioeconomic effects on local, national and international travel patterns and greenhouse gas emissions of young adults. *Journal of Transport Geography,* 68**,** 130-141.

de Dios Ortuzar, J. & Willumsen, L. G. 2011. *Modelling transport*, John Wiley & Sons.

DMPW. 2019. *Mobiel Vlanderen* [Online]. Department of Mobility and Public Works, Flanders, Belgium. Available: https://www.mobielvlaanderen.be/index.php [Accessed 2019].

Dons, E., Beckx, C., Arentze, T., Wets, G. & Panis, L. 2011. Using an activity-based framework to determine effects of a policy measure on population exposure to nitrogen dioxide. *Transportation Research Record: Journal of the Transportation Research Board***,** 72-79.

Durand, A., Harms, L., Hoogendoorn-Lanser, S. & Zijlstra, T. 2018. Mobility-as-a-Service and changes in travel preferences and travel behaviour: a literature review. *The Hague: Netherlands Institute for Transport Policy Analysis*.

Hao, J., Hatzopoulou, M. & Miller, E. 2010. Integrating an activity-based travel demand model with dynamic traffic assignment and emission models: Implementation in the greater Toronto, Canada, area. *Transportation Research Record: Journal of the Transportation Research Board***,** 1-13.

Helminen, V. & Ristimäki, M. 2007. Relationships between commuting distance, frequency and telework in Finland. *Journal of Transport Geography,* 15**,** 331-342.

Horni, A., Nagel, K. & Axhausen, K. (eds.) 2016a. *Multi-Agent Transport Simulation MATSim,* London: Ubiquity Press.

Horni, A., Nagel, K. & Axhausen, K. W. 2016b. *The multi-agent transport simulation MATSim*, Ubiquity Press London.

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.

Kochan, B., Bellemans, T., Janssens, D. & Wets, G. 2013. Validation of an Activity-Based Traffic Demand Model for Flanders Implemented in the FEATHERS Simulation Platform. *In:* WANG, W. & WETS, G. (eds.) *Computational Intelligence for Traffic and Mobility.* Paris: Atlantis Press.

Lu, L., Xu, Y., Antoniou, C. & Ben-Akiva, M. 2015a. An enhanced SPSA algorithm for the calibration of Dynamic Traffic Assignment models. *Transportation Research Part C: Emerging Technologies,* 51**,** 149-166.

Lu, Y., Adnan, M., Basak, K., Pereira, F. C., Carrion, C., Saber, V. H., Loganathan, H. & Ben-Akiva, M. E. Simmobility mid-term simulator: A state of the art integrated agent based demand and supply model. 94th Annual Meeting of the Transportation Research Board, Washington, DC, 2015b.

Ng, W.-S. & Acker, A. 2018. Understanding Urban Travel Behavior by Gender.

Pääkkönen, H. 2013. Experiences with the light-version time diary in Finland.

Petrik, O., Adnan, M., Basak, K. & Ben-Akiva, M. 2018. Uncertainty analysis of an activity-based microsimulation model for Singapore. *Future Generation Computer Systems*.

Rasouli, S. & Timmermans, H. 2014. Activity-based models of travel demand: promises, progress and prospects. *International Journal of Urban Sciences,* 18**,** 31-60.

Shabanpour, R., Javanmardi, M., Fasihozaman, M., Miralinaghi, M. & Mohammadian, A. 2017. Investigating the applicability of ADAPTS activity-based model in air quality analysis. *Travel Behaviour and Society*.

Siyu, L. 2015. *Activity-based Travel Demand Model: Application and Innovation.*

Vuurstaek, J., Cich, G., Knapen, L., Ectors, W., Yasar, A.-U.-H., Bellemans, T. & Janssens, D. 2018. GTFS bus stop mapping to the OSM network. *Future Generation Computer Systems*.

Vuurstaek, J., Cich, G., Knapen, L., Yasar, A.-U.-H., Bellemans, T. & Janssens, D. 2017. GTFS Bus Stop Mapping to the OSM Network. *Procedia Computer Science,* 109**,** 50-58.

Ye, X., Konduri, K., Pendyala, R. M., Sana, B. & Waddell, P. A methodology to match distributions of both household and person attributes in the generation of synthetic populations. 88th Annual Meeting of the Transportation Research Board, Washington, DC, 2009.