School of Transportation Sciences

Master of Transportation Sciences

Master's thesis

Accessibility Measure and its Role in Disaggregate Activity based Model System

Mahmoud Abdelrahman

Thesis presented in fulfillment of the requirements for the degree of Master of Transportation Sciences, specialization Transport Policy and Planning

SUPERVISOR :
Prof. dr. ir. Tom BELLEMANS

CO-SUPERVISOR :
dr. Muhammad ADNAN

2021
2022

UHASSELT
KNOWLEDGE IN ACTION

www.uhasselt.be
Universiteit Hasselt
Campus Hasselt:
Martelarenlaan 42 | 3500 Hasselt
Campus Diepenbeek:
Agoralaan Gebouw D | 3590 Diepenbeek

School of Transportation Sciences

Master of Transportation Sciences

Master's thesis

Accessibility Measure and its Role in Disaggregate Activity based Model System

Mahmoud Abdelrahman

Thesis presented in fulfillment of the requirements for the degree of Master of Transportation Sciences, specialization Transport Policy and Planning

SUPERVISOR :

Prof. dr. ir. Tom BELLEMANS

CO-SUPERVISOR :

dr. Muhammad ADNAN

Table of Contents

List of Figures

List of Tables

Acknowledgment

This research would not have been possible without the support of many people. Many thanks to my co-supervisor, Prof. Dr. Muhammad Adnan, who read my revisions and helped make some sense of the confusion. Also thanks to my supervisor, Prof. Dr. Tom Bellemans, who offered guidance and support.

Thanks to the researchers of Transportation Research Institute (IMOB) of Hasselt University who have created and are developing FEATHERS, the target model in this research. Thanks to my family, and friends who offered support and love.

Finally, thanks to Al Alfi Foundation for granting me a merit-based scholarship to pursue this master's program. Special gratitude to MEP Dorien Rookmaker who gave me the opportunity to widen my horizons by working with her at the European Parliament.

Preface

Robust transport demand models are needed for formulating sensible transport policies. Traditionally, trip-based models were the dominant transport models which consider the personal trip as the forecasting element regardless of its purpose. Eventually, the Activity-Based Models (ABMs) emerge because they deal with travelling as a sequence of activities that are affected by the traveller’s needs and attributes. Accessibility refers to how attractive or easy to reach is the place. As it affects the travelling choices, incorporating its parameters in the ABMs is essential. However, accessibility is superficially defined by focusing merely on travel distance and time.

Hence, this research aims to incorporate new accessibility parameters into FEATHERS, an ABM developed for Flanders, Belgium. These parameters enrich the accessibility definition by reflecting various dimensions, i.e. Land-use (Opportunities), Transport (Network Impedance), Temporal (Services Availability) and Individual (Needs/Capabilities). They are incorporated as variables in the utility functions which predict the choice probability between alternatives following a discrete choice model. Their coefficients are estimated using Biogeme via the maximum likelihood technique.

The research proves the effectiveness of these parameters on the daily activities, location and transport mode choices of the traveller. For example, the individual who lives in areas with a low density of shops aggregates the shopping activities due to the long home-shop distances. Moreover, it is found that the relevance of the accessibility parameters varies across the models. The network impedance parameters fit in the Location Choice Model while the land use pattern is suitable for the Daily Activity Pattern Model. Within the Location Choice Model, the research has incorporated new parameters only in the education activities. Further activities can be tackled in future research. This research improves the FEATHER's forecasting accuracy by adding new accessibility dimensions. Consequently, other ABMs can be enhanced by practitioners using a similar methodology.

1. Introduction

This thesis will discuss the topic of accessibility measure and its role in disaggregate activity-based model. It goes across the topic from its background to the relevant literature review and the objective of conducting the research along with the methodology deployed to reach this objective.

After then, the case study examined by the research is explained in further details through illustrating its transport demand model. This includes portraying the various model's components and the accompanying input data which can help in understanding the flow of the rest of the thesis.

Consequently, the outcome of applying the aforementioned methodology of incorporating new accessibility parameters into FEATHERS model is presented in a separate section. Finally, the conclusion and lesson learned from the research are described in the last section.

In light of the previous overall description, the thesis comprises the following sections:

- Section 2: Research Background
- Section 3: Literature Review
- Section 4: Research Objective
- Section 5: Research Methodology
- Section 6: FEATHERS Model Components
- Section 7: New Parameters Incorporation
- Section 8: Conclusion and Discussion

2. Research Background

Accessibility term, especially in the transport modelling field, is an ambiguous term which can be defined and estimated in various methods as it is affected by physical and cultural barriers where the latter is more difficult to model (Gould, 1969). The traditional definition of accessibility is that it is the ability of a place to be arrived from, or to be arrived by, different places. Therefore, the characteristics and the arrangement of the transport system are main factors in the measurement of accessibility (Rodrigue, 2020).

(Litman, 2013) refers to the evolution of the transport-system performance evaluation from the traffic-based assessment (focus on motorized modes' travel speeds and costs) to mobility-based assessment (focus on passenger and good travel speeds and costs) then to accessibility-based assessment (focus on the capacity of individuals and businesses to get to needed facilities).

Given that reaching the destination is the eventual goal of mobility, accessibility-based assessment is the recommended way whereas it generates the most optimal and effective answers to the common transport issues. Hence, measuring accessibility is of utmost importance in the evaluation of any transport system or network.

Travel demand models are used to predict how changes in the transportation system's characteristics and the people who use it will react to changes in transportation demand. Travel demand models, in particular, are used to forecast travel behaviour and transportation service use within various socioeconomic conditions and land-use arrangements (Bhat et al, 2003).

According to (Chu et al., 2012), Travel Demand Models fall mainly under one of two categories, i.e. Trip-based Travel Models and Activity-based Travel Models. The first kind of travel model, also known as the four-step model, was initially established in the 50s in the last century. It considers the personal trips as the element of analysis and comprises four consecutive steps, i.e. Trip Generation, Traffic Distribution, Mode Choice, and Route Choice.

Later, the Tour-based Travel Models came as a more developed form of Trip-based Travel Models whereas the trips are aggregated in trip chains forming the so-called Tour. (Ferdous et al, 2012) found that the Tour-based Model had, generally, a better performance than the Trip-based Model in respect to the studied four aspects, i.e. car ownership, average work-trip travel time, work-flow distribution and work-flow distribution (over time-of-day), in the case study of Columbus Metropolitan Area in Ohio, USA.

On the other hand, the activity-based travel models were developed in the 70s. These models see travel as a resulting demand from the necessity to follow activity spread in a spatial context (Axhausen & Garling, 1992). In other words, its main concept is grounded on the assumption that travelling to specific places, at specific times of day and by specific transport modes is caused by the demand activity participation.

(Ben Akiva, 2008) mentions the main advantages and limitations of each travel model type as listed in **Table 2-1**. Moreover, **Figure 2-1** illustrates how each model type deals with the travel components where the schedule represents the Activity-based approach.

Table 2-1: Evaluation of Travel Model Approaches

Travel Model Type	Advantages	Limitations
Trip-based	1- Requires less computation 2- Well-established Approach	1- Spatial and Demographic Aggregation (Less accurate) 2- Travel Behavior does not change (No induced travel)
Tour-based	1- Links the trips in chains (See the bigger picture) 2- Widely Used Approach	1- No Individual Schedule considered 2- Temporal aspect is disregarded
Activity-based	1- Considers the schedule and temporal dimension in travel behavior 2- More accuracy in modelling travel behavior	1- Requires extensive computation 2- Not as well-established as Trip-/Tour-based models

Source: (Ben Akiva, 2008)

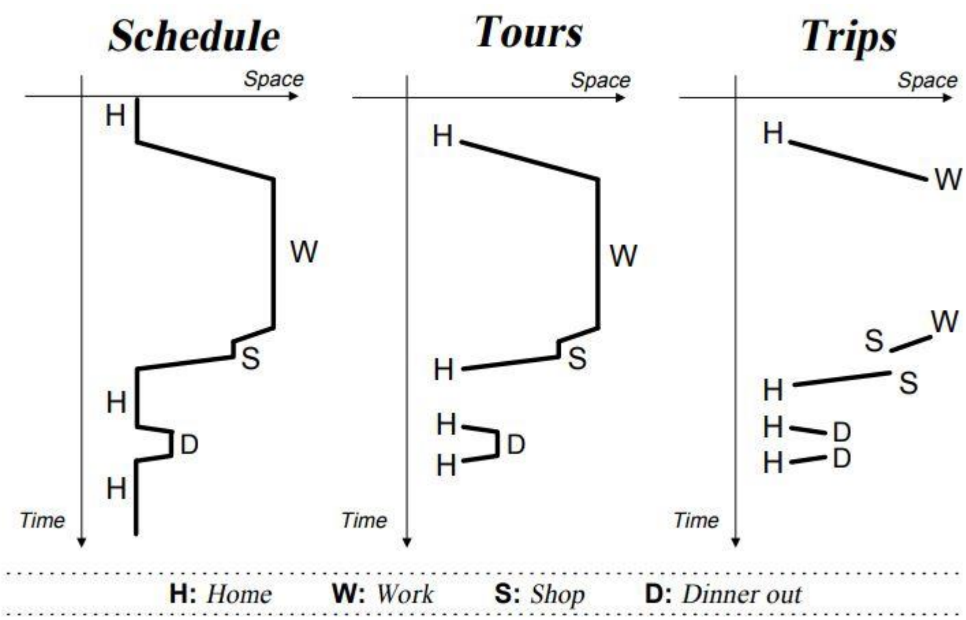
Furthermore, the guidelines titled Activity-Based Travel Demand Models: A Primer could summarise and compare the main travel/commute questions targeted to be tackled by each component of a Trip-Based Model versus an Activity-Based Model as shown in **Table 2-2** (National Academies of Sciences, 2014).

In this research project, the activity-based transport model will be adopted as the basis of the investigation. More details about the model used and the case study of this research will be clarified in the upcoming sections.

Table 2-2: Main Travel Questions of Different Models and their Answers

Main Travel Question	Trip-Based Model Components	Activity-Based Model Components
What activities do people want to participate in?	Trip generation	Activity generation and scheduling
Where are these activities?	Trip distribution	Tour and trip destination choice
When are these activities?	None	Tour and trip time of day
What travel mode is used?	Trip mode choice	Tour and trip mode choice
What route is used?	Network assignment	Network assignment

Source: (Activity-Based Travel Demand Models: A Primer, 2014)



Source: (Ben Akiva, 2008)

Figure 2-1: Comparison between Different Types of Travel Models

3. Literature Review

Several papers and publications are issued searching for the most suitable method and indicators in order to measure the accessibility in the most accurate possible way. Accessibility measures are of utmost importance to guarantee reasonable policy sensitivity at the different stages of the model development to alterations in infrastructure or land use patterns or both at the time.

(Berechman, 1981) referred to the extreme complexity of measuring accessibility which requires a lot of essential assumptions related to the human behaviour, social inclinations and spatial configuration of the potential activities. Ignoring one of these elements may cause a misleading outcome which affects on the transport condition understanding and selecting the proper policies as a reaction.

In order to understand the concept of measuring accessibility defined by (Castiglione, 2015), we need to understand the log-sum term which is broadly used in the literature. The log-sum is the summation of the denominator of the logit mode choice model which represents the anticipated maximum utility value related to the set of alternatives available for the traveller (Geurs et al., 2012).

According to (Castiglione, 2015), there are four sorts of accessibility variables incorporated in the travel models which are as follows:

1- Direct magnitudes of journey times, travel distances, and monetary costs from the transport network routes;
2- Comprehensive (Detailed) log-sums computed over available choices in models that comprise direct magnitudes;
3- Combined (Aggregate/approximate) log-sums computed over available choices in models that comprise direct magnitudes (suitable for day-level models); and
4- Buffer measures embodying the activity options and the urban scheme surrounding each parcel or micro-zone.

The direct measures are deployed, whenever applicable, in all transport-mode, travel-destination, and time-of-day choice models. Yet, the model sequence can make using the direct measures unmanageable as they are dependent on the results of later stages in the model. For instance, the travel time comes as an outcome after executing the trip distribution, mode choice and time-of-day choice models. Hence, the travel time cannot be estimated before conducting the aforementioned models/steps.

Here comes, in the activity-based models, the detailed log-sum approach to overcome this limitation where the restrictive sequence of the trip-based models is absent. For example, the detailed log-sums of the tour mode choice model can be computed and utilized in high-level choice models, e.g. tour-destination and time-of-day.

However, the detailed log-sums can be non-feasible due to their intensive computation process that needs enormous storage, long time and advanced hardware. Alternatively, the aggregate log-sums emerged to solve this challenge by making combinations from important features such as Origin Zone, Travel Purpose and Household Income Category.

On the other side, the buffer measures simulate the accessibility to/from attractive activities/destinations, e.g. Jobs, Schools and Transit Stops, in the vicinity which can be reached by walking, biking or short-automobile-trip. Thus, the buffer measures are beneficial when the model's spatial units are smaller than the buffer range which means that the spatial context should be defined in terms of parcels.

According to (National Academies of Sciences, 2014), one of the methods to get the buffer measures more precise and significant is to use on-street, shortest-route distance to determine the distance to the edge of the buffer area, instead of utilizing a straight line in the crow-flies form. Moreover, deploying distance-decay functions can diminish the sudden changes and gaps observed between the ranges of the buffer areas or what is called the cliff effects which are occasionally linked to the applying simple buffer measures.

(Hwang et al., 2017) conducted a study to investigate the level of accuracy of the outcome of activity-based travel models compared to collected travel data from navigation and smartcards using the big-data technology. Although the activity-based models are more comprehensive, as mentioned in the previous section, and consider several socio-economic parameters and factors, it is found that there is low accuracy when it comes to the start time of the trips and the non-work trips within the studied schedule.

Moreover, the degree of discrepancy between the model and the collected data is affected by land use and network attributes of the Traffic Analysis Zones (TAZs). This inclination is more notable in the case of public transport, because travel by a private car appears less influenced by transport network provisions and land use types. The results indicate that accuracy of simulating trips using the public modes should be enhanced.

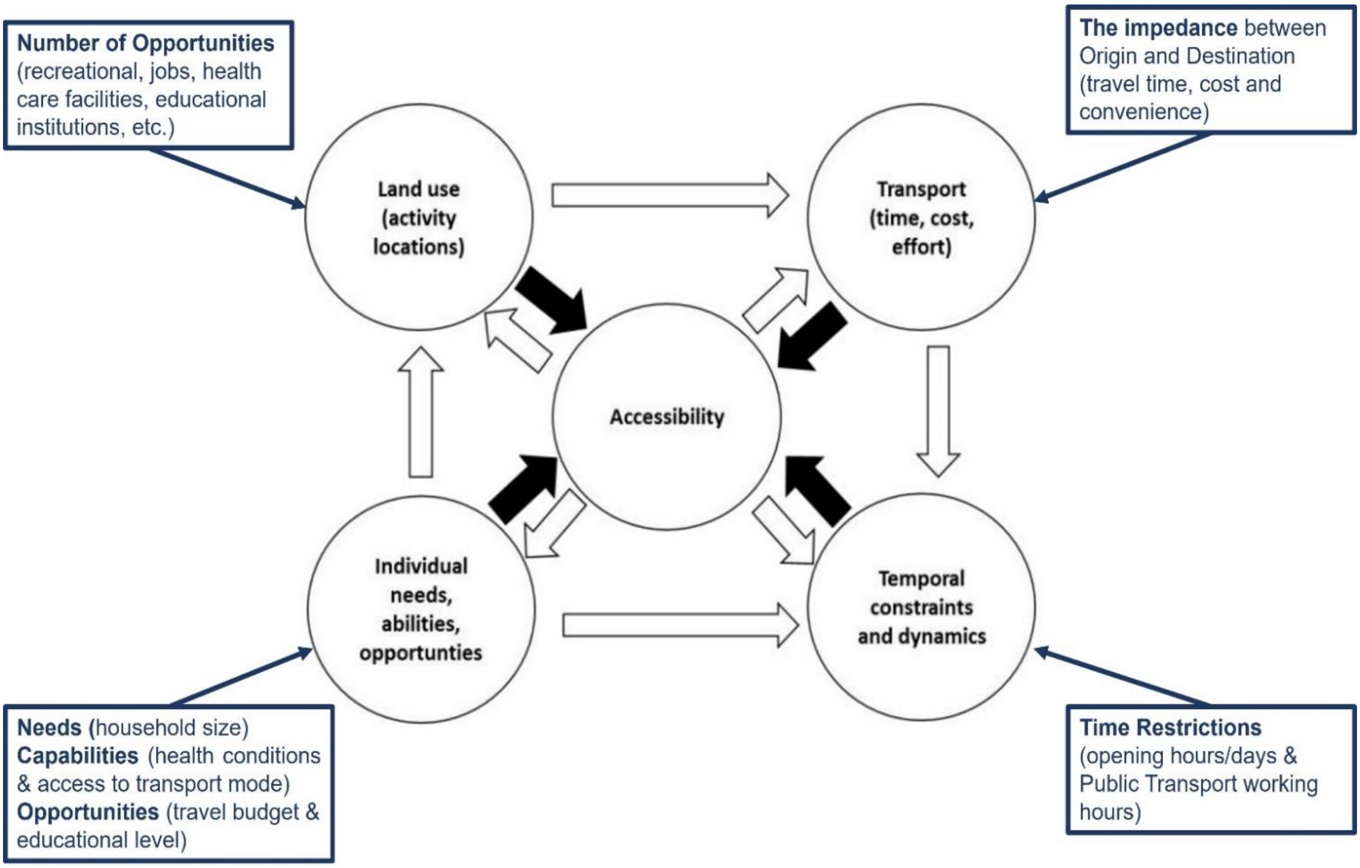
Based on these findings, such low accuracy has a potential impact on measuring accessibility. This impact is more significant during following the direct-magnitude approach while it is less in the case of the log-sum approaches where the latter combines different direct measures, e.g. travel time and distance, instead of relying solely on a specific direct measure which its accuracy is questionable depending on the transport mode used.

(Geurs & Van Wee, 2012) addresses the accessibility from a different perspective by defining its main four components, i.e. Land-use, Transport, Temporal and Individual, along with the interface between them which can be described as follows:

1- Land-use Component symbolizes the land-use parameters such as (a) the amount of opportunities (recreational, jobs, etc.) enclosed in the destination of the traveller, (b) the attractiveness of these opportunities for the traveller at his/her origin, (c) the demand-supply conflict due to the limited capacity of certain opportunities (health care facilities, educational institutions, etc.)
2- Transport Component represents the transport system in terms of the impedance elements (e.g. travel time, cost and convenience) of the trip conducted from an origin to a destination
3- Temporal Component considers the dynamic nature of the time-dimension of the travel impedance (e.g. peak hour) or the time restriction of the desired activities (e.g. opening hours/days)

4- Individual Component embodies the needs (correlated with age, household conditions, etc.), capabilities (correlated with health conditions, access to transport mode, income, etc.) and opportunities (correlated with the travel budget, educational level, etc.) of the passenger to move and reach the destination

Figure 3-1 depicts a summary visualization for the main components of accessibility and their interface.



Adopted from: (Geurs & Van Wee, 2012)

Figure 3-1: Main Components of Accessibility and their Interface

4. Research Objective

The transport model under study in this research project is the Forecasting Evolutionary Activity-Travel of Households and their Environmental RepercussionS (FEATHERS). FEATHERS is an activity-based micro-simulation model for travel demand forecasting. It was created by the Transportation Research Institute (IMOB) of Hasselt University (Bao et al, 2018).

The framework of FEATHERS adopts the ALBATROSS tool which is a Learning-based Transportation Oriented Simulation System. ALBATROSS is a household activity-based modelling structure. It is created through a scheduling module and models which are linked by conditional rules. ALBATROSS framework follows the below sequence:

1- The model is fed by an outline of a schedule with fixed/set-up activities to be conducted alongside a group of different sorts of constraints.
2- The scheduling procedure will add flexible/variable activities.
3- The outcome of the scheduling module will be a group of unsorted fixed/set-up and flexible/variable activities.
4- In the next step, the activities will be sequenced and mode, destination and time-of-day will be modelled using decision trees/nested form which is developed using collected data after taking all kinds of constraints into consideration.

In short, ALBATROSS is developed to formulate travel-decision trees depending on activity diary data. ALBATROSS anticipates what kind of activities is carried out along with its time, location, duration, with and the travel mode used. Furthermore, different contextual, spatial-temporal and organizational restrictions are merged in the model (Arentze, 2004).

ALBATROSS was originally built empirically for the Dutch Ministry of Transportation, Public Works and Water Management, trying to discover the prospect of the rule-based approach and create a travel demand model for the purposes of assessing proposed policies in the future. A case study was conducted for Rotterdam, verifying the potentiality of the rule-based approach. Since then, ALBATROSS has been advancing and enhancing. In line with the enhancement, ALBATROSS has been incorporated into an integrated activity-based framework FEATHERS (Bellemans et al., 2010), which has been executed for the Flanders region in Belgium and used to estimate vehicle emissions, air pollution parameters as the case of (Beckx et al., 2009) and energy consumption as the case of (Yang et al., 2010).

For each passenger with own specific characteristics, FEATHERS predicts if an activity, e.g. shopping, business, recreational, will be conducted or not. Consequently, the main questions related to the activity, i.e. to-where, which-transport-mode, for-how-long, are answered considering the characteristics of the traveller, e.g. age group and gender, of the household, e.g. car ownership, and of the spatial area, e.g. population, the number of stores.

As a result, individual schedules or travel patterns are anticipated which can be used to obtain the forecasted travel demand. These estimated travel patterns can be later deployed to deal with several policy challenges and their effect, for instance, the introduction of a new transport mode in the system (Bao et al, 2014).

The key objective of this study will be to answer the main research questions which are as follows:

1- How to incorporate the accessibility features in the activity-based model?
2- What is the most suitable way to measure the accessibility in the activity-based model?
   a) Can the location-based measures be the most compatible way for the activity-based model?

5. Research Methodology

As discussed in the Literature Review section, the log-sum approach has been deployed in many studies and research projects as a tool for measuring accessibility within the activity-based travel demand models. However, (Geurs & Van Wee, 2013) refer to the shortcomings associated with the log-sum measure whereas the main ones are the deficiencies linked to the traveller's willingness to pay (WTP) and ignoring the fact that the utility at the decisive moment of the choice may change from the utility during experiencing the choice selected itself. The latter limitation is confirmed by (Chorus et al, 2011) highlighting the inconsistency between what log-sum approach of accessibility intends to measure, i.e. experienced-utility, and what it in fact measures, i.e. decision-utility.

Accessibility is usually assessed by deploying either a travel model or a land-use model separately or sequentially lacking a closed-loop of feedback. This leads to ignoring the interface impacts that occur between both of the models and the derived competition effects within the opportunities caused by the accessibility enhancements. Eliminating the Transport-Land Use interactive relation and opportunities for competition from consideration can cause a significant underestimation of the policy outcomes. Recently, the development of Land-Use and Transport Interaction (LUTI) models increase the interest in utilizing such models to satisfactorily measure accessibility and appraise its influence (Wang et al, 2013).

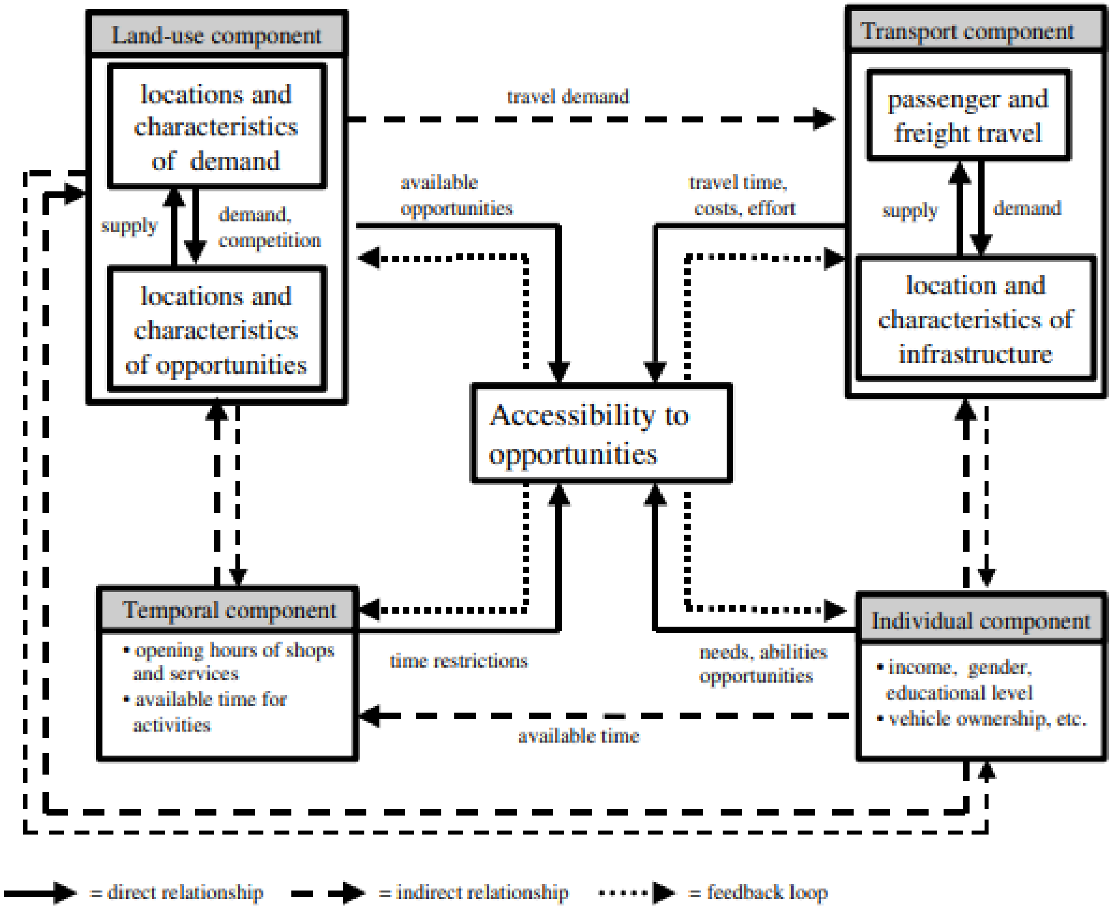
Transportation Research Institute (IMOB) has already an on-going research project (R-9900) that aims to improve FEATHERS by integrating a land-use model in order to give a more accurate analysis of tested scenarios of policies. Given the recommendation of (Wang et al, 2013) and the on-going land-use model project, the land-use component is the answer to the main research question whereas it is the most feasible path to take in order to incorporate the accessibility measures in FEATHERS.

Regarding the sub-question of the second research question, location-based accessibility measures will be investigated to see to what extent it is valid and accurate for the accessibility evaluation. (Miller, 2018) refers to the downside of considering purely location-based accessibility measures because it does not reflect the transport mode and the time-of-day. He admits that it makes the process of estimating accessibility more complicated.

According to (Geurs, 2018), the limitation of the potential accessibility measure (i.e. opportunities competition effect) as part of the location-based measure can be diminished by using modified approaches.

In this spatial interactive model, the balancing factors help to guarantee that the amount of traffic flow, e.g. trips, produced from and attracted by each Traffic Analysis Zone (TAZ) matches the correct number for that TAZ, e.g. population or job opportunities. Because they are inter-dependently associated, it is necessary to estimate them in an iterative way. Here comes (Miller, 2018) to highlight the convolution of measuring the attractiveness of zones. For instance, the shopping trip covers so wide range of possibilities where it can be going to buying something as simple as a milk bottle or as advanced as purchasing a new car.

Therefore, the sub-question is put on the table in order to go deeper and explore more this non-settled topic. However, the recommendation of (Geurs et al., 2004) for a hybrid approach could be adopted as a starting point. He recommends measuring the accessibility, firstly, based on macroscopic perspective/aggregate level through computing location- or utility-based accessibility. After then, the outcome of this aggregate-level approach can be fine-tuned on disaggregated level considering the individual-specific features and temporal aspect. **Figure 5-1** shows different components of the accessibility measures and explains the interaction between them.



Source: (Geurs et al., 2004)

Figure 5-1: The interaction between different accessibility components

6. FEATHERS Model Components

a) Structure of FEATHERS Model

As previously mentioned, FEATHERS is an activity-based micro-simulation model for travel demand forecasting. In simple terms, it predicts the daily activity schedule, as an output, by processing groups of data, as an input, collected via various surveys, i.e. synthetic population, household travel survey, land use parameters and transport network skim matrices.

Therefore, these predicted activity schedules can be assigned, in terms of traffic volumes, to the transport network of the study area as a standalone and final outcome. This loaded network with different types of movements can be deployed to forecast the effect of planned policy measures or develop what-if scenarios based on different decision-making options. Alternatively, these schedules can be used as input for other specialised models, such as air pollution estimation models and traffic noise exposure models. **Figure 6-1** illustrates how FEATHERS work along with the potential applications of its output. **Figure 6-2** shows an example of the output generated by other specialised models using FEATHERS day activity schedules. **Figure 6-3** depicts the interaction between the input, output and FEATHERS components and sub-models.

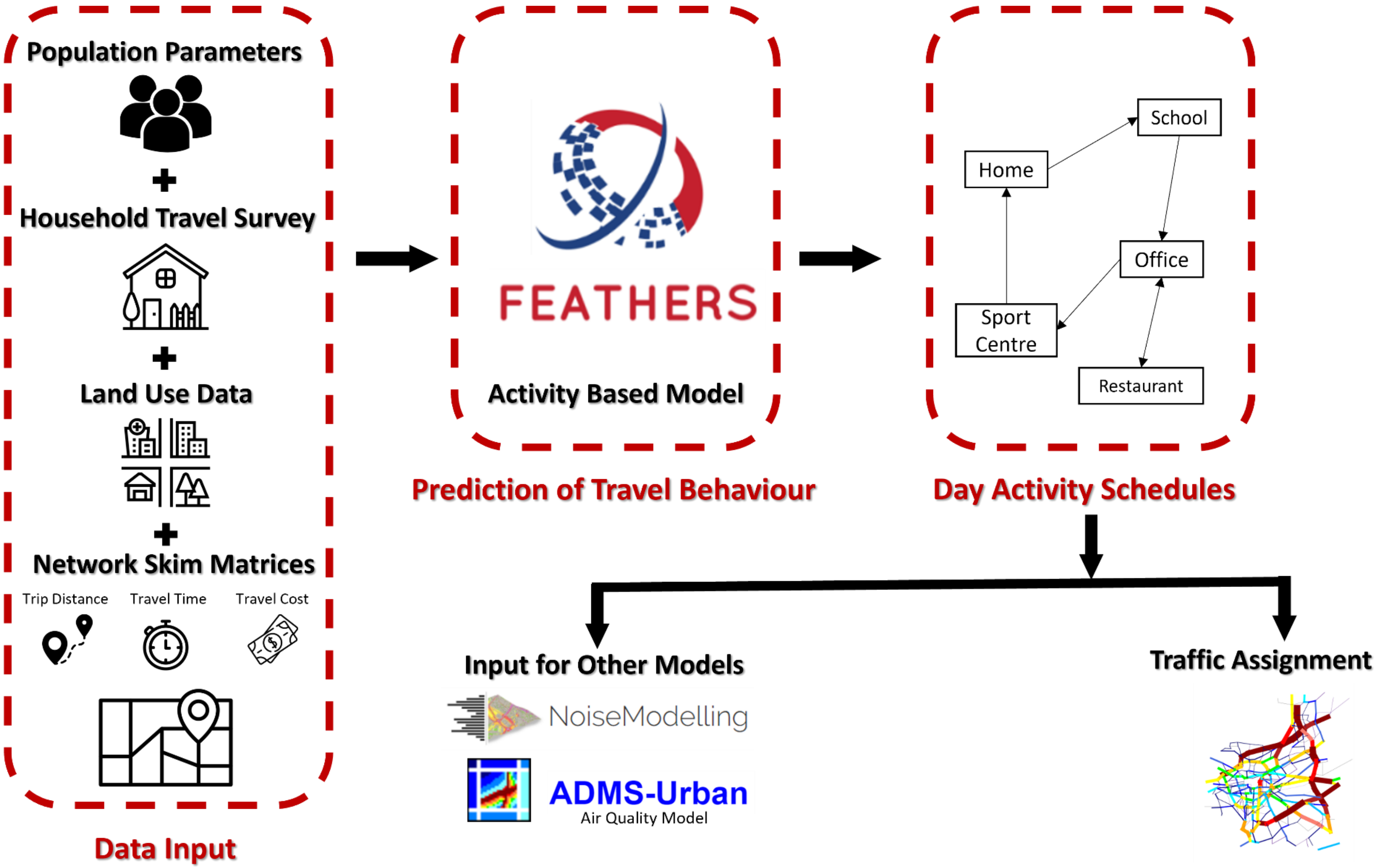


Figure 6-1: Input and Output of FEATHERS Model

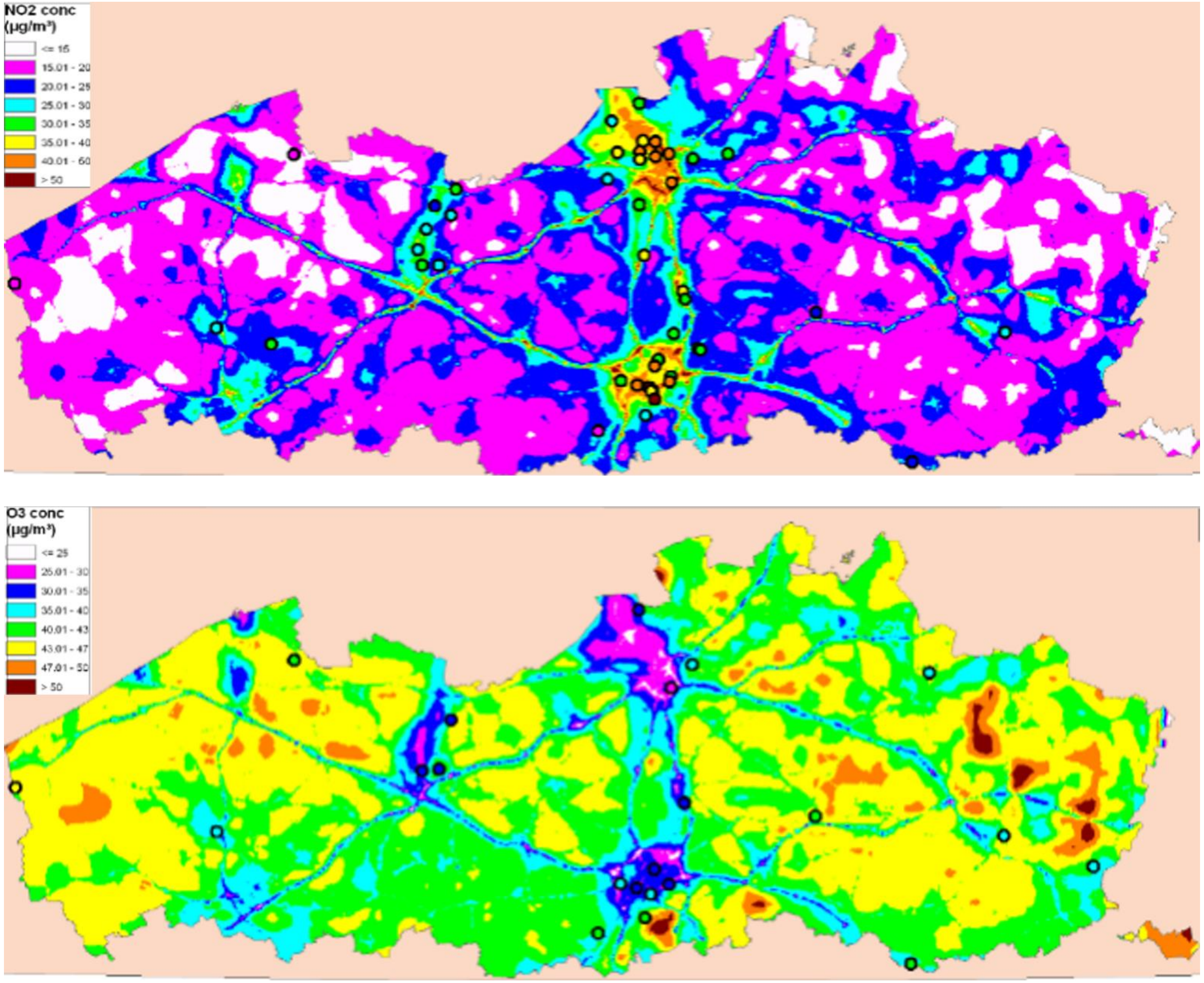


Figure 6-2: Examples of FEATHERS Output Applications (VITO Air Quality Model)

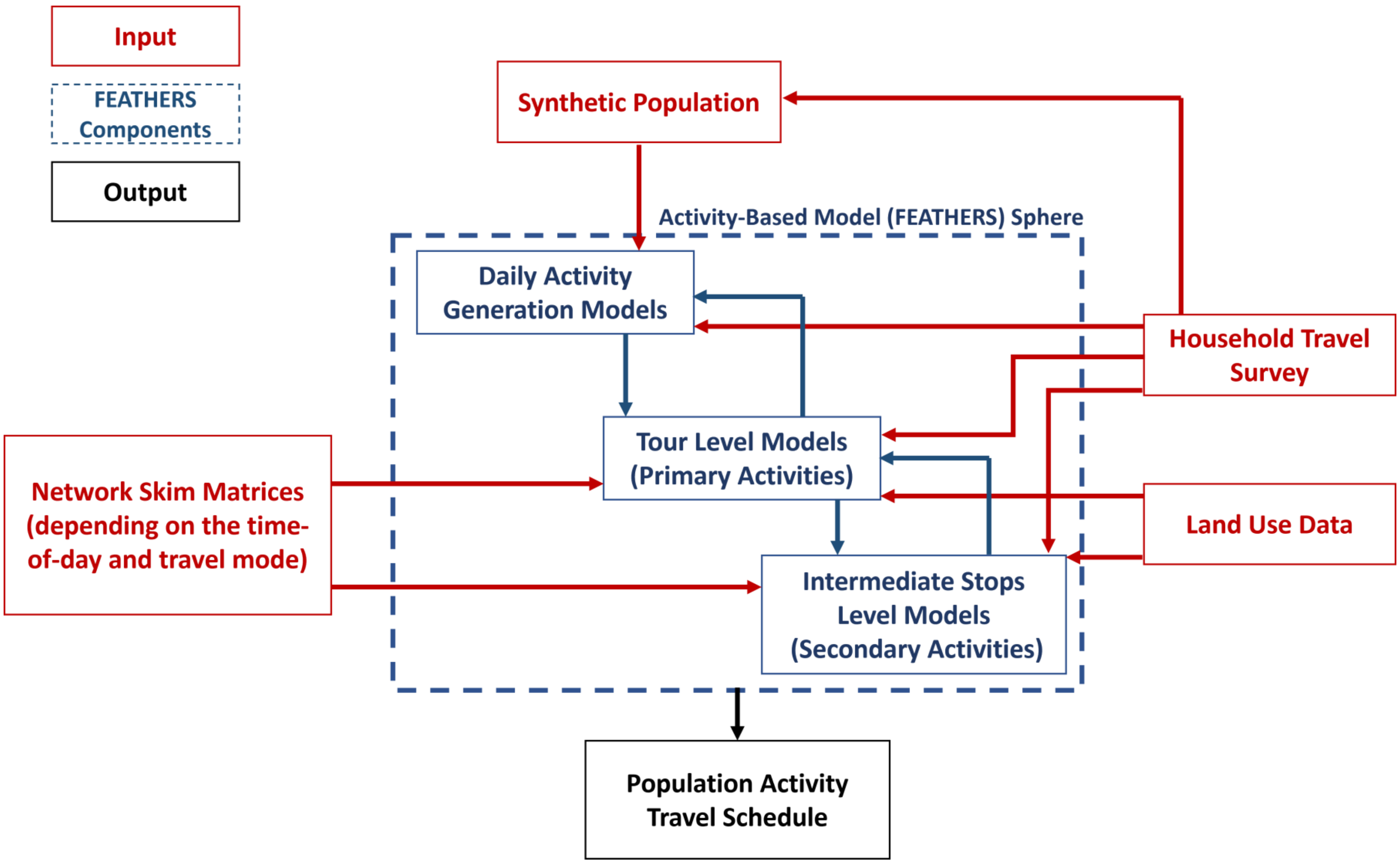


Figure 6-3: The Interaction of FEATHERS Model Components

b) Input data of FEATHERS Model

As previously explained, the input data used to feed FEATHERS model comprises four categories obtained from different sources or derived via various methodologies. They can be listed as follows:

i. Household Travel Survey

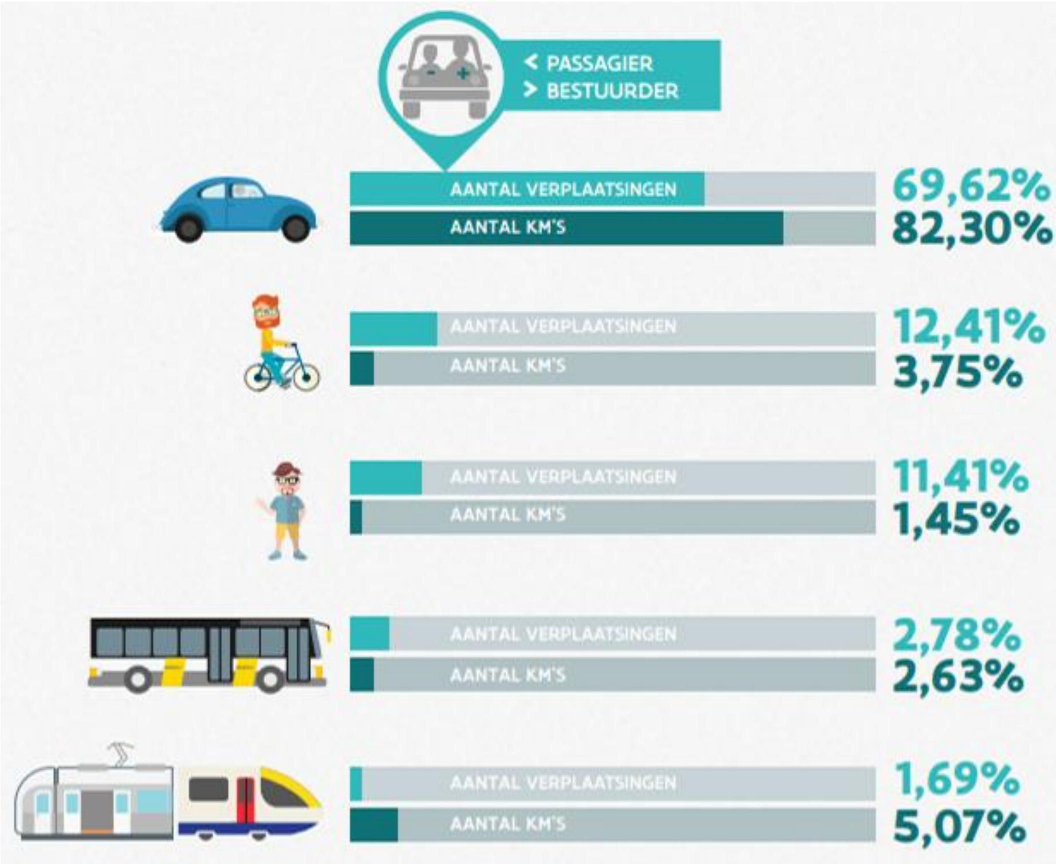
Since 1994 in the Flanders region of Belgium, the Flemish government has annually investigated the travel behavior of the region’s residents. For this purpose, Flemish residents aged 6 years and older, which form the target group of the survey, are asked about their travel and mobility behaviour.

This survey is conducted by filling in two questionnaires and keeping track of the respondent’s movements in a travel booklet for one day (24 hours). One of the questionnaires is about the mobility behavior of the family (family questionnaire) as a whole and the other one is concerning the individual's own mobility behavior (person questionnaire). An extract of data points related to around 8800 individuals has been selected to train the FEATHERS model and its sub-models.

The aim of this survey project is to discover how Flanders and Brussels residents move by inquiring about a randomly chosen sample of Flemings and Brussels residents about their movements spread over a full year. The research is a valuable source of data used for formulating regional and local policy decisions.

These decisions seek to enhance traffic safety of the road network, optimising public transport infrastructure, improving the environmental conditions and combating traffic congestion. **Figure 6-4** illustrates the modal split in Flanders as an example of the output of the survey analysis report of the year 2015-2016 in terms of the total number of movements “Aantal Verplaatsingen” and total number of kilometers travelled “Aantal Km’s” for each transport mode (Department Mobiliteit & Openbare Werken, 2016).

Research of the Movement Behaviour in Flanders or Onderzoek VerplaatsingsGedrag (OVG) Vlaanderen, as the survey is called in Dutch, follows the provisions of the General Data Protection Regulation (GDPR) whereas all the data collected are associated with a key code to keep the individuals anonymous for whoever carrying out analysis or developing research publications.



Source: (Department Mobiliteit & Openbare Werken, 2016)

Figure 6-4: Modal Split in Flanders Region (2015-2016)

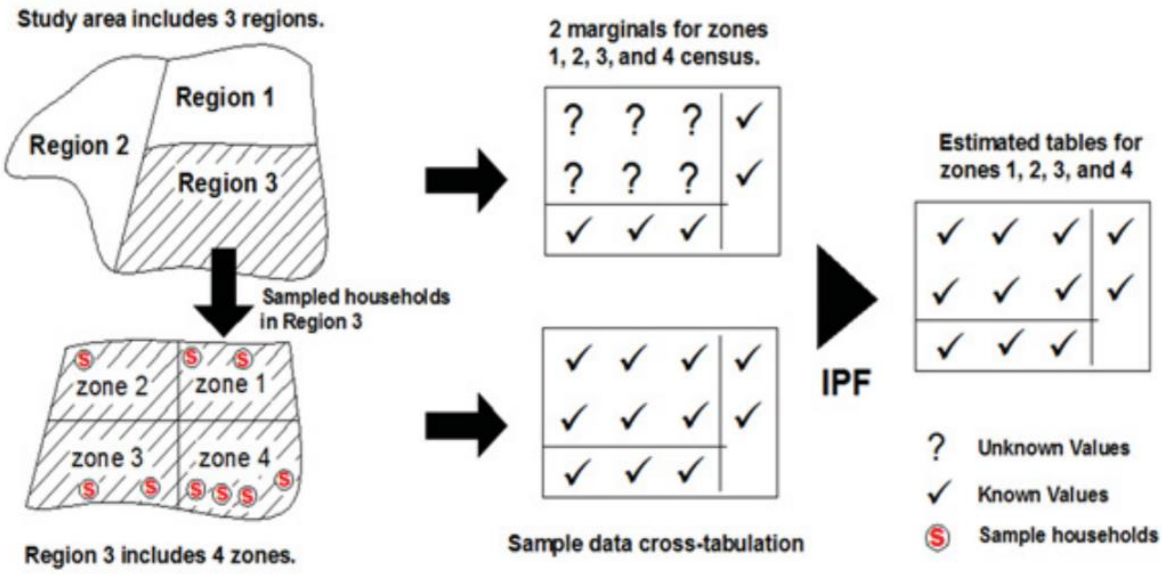
ii. Synthetic Population

The complete socioeconomic data of the research region population is required for activity-based travel demand modeling. Because collecting such precise data for the entire population is too costly, population synthesis has been presented as a way to forecast the data and synthesize it based on a sample.

Iterative Proportional Fitting (IPF) is a significantly cheaper technique for forecasting population features that have been the subject of extensive research due to its many benefits (Choupani et al, 2016).

Population synthesis applying the Iterative Proportional Fitting (IPF) approach contains two steps, i.e. fitting and allocation. In the fitting stage, some cross-tabulation of (k) attributes of households sampled from a zone is obtainable. These available attributes are known also as seed or reference tables.

However, these types of cross-tabulations are not available for zone populations. As solely marginals of the k variables are obtainable for zones, (Beckman et al., 1996) mention that comparable cross-tabulations can be fine-tuned for zones using IPF estimator and reference tables as shown in **Figure 6-5**.



Source: (Choupani et al, 2016)

Figure 6-5: Iterative Proportional Fitting (IPF) Concept

The zoning system adopted in FEATHERS comprises a total number of super-zones and sub-zones equal to 327 and 2,386 respectively. The super-zones correspond to the number of municipalities located in the Flanders region while the sub-zones reflect the smaller geographic units, e.g. districts and neighbourhoods. This zoning system is utilized in both Network Skim Matrices and Land Use Data as will be explained further in the upcoming sub-sections.

Figure 6-6 illustrates the example of Hasselt Municipality (Super-zone) which is depicted in a polygon filled in orange colour while the Hasselt Centre district (Sub-zone) is coloured yellow. By adopting two different levels of zoning details FEATHERS model can fit various applications.

For instance, the level of granularity in the case of a sub-zone system is necessary for the accuracy of the model's output required for the local policy decision-making. On the other hand, the level of granularity in the case of the super-zone system is sufficient for observing and analysing the movements between the cities and provinces enclosed within the Flanders region boundary.

The synthetic population of FEATHERS model includes 6,294,756 persons spread over the Flanders region. Each person in this population is labeled with a unique attribute value, i.e. person_id, and assigned to a specific zone among the aforementioned zoning system, i.e. location_id.

Each row in the synthetic population database represents a single individual whereas the gender is expressed in a binary form, i.e. 1 for male and 2 for female, and the age is stated in the form of categorical data.

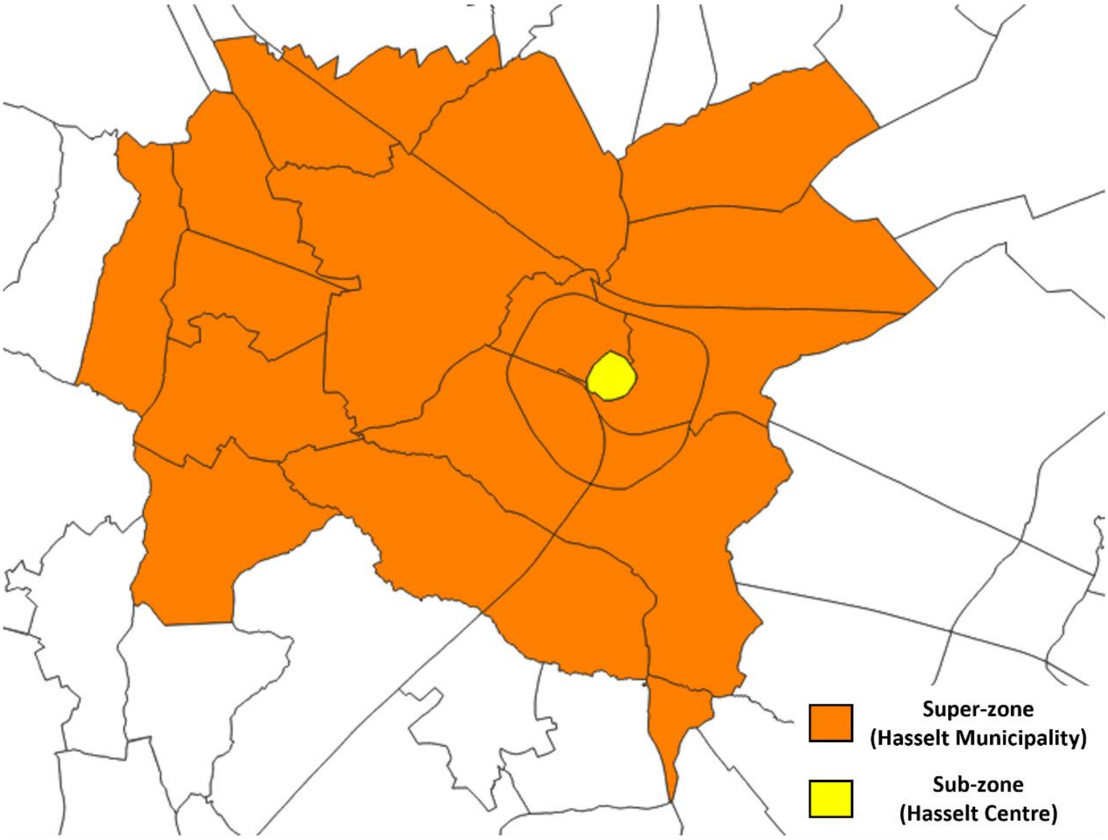


Figure 6-6: Zoning System Structure in FEATHERS Model (Hasselt City Example)

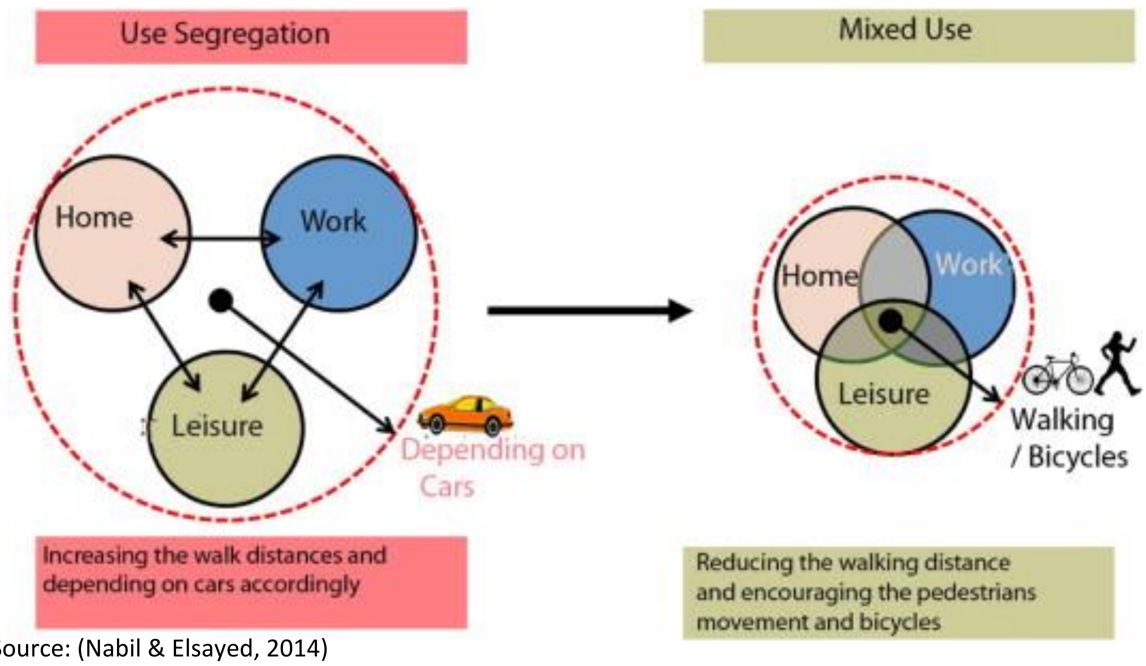
iii. Land Use Data

In transport demand modelling regardless of the model type whether it is a four-step (traditional transport model), tour-based or activity-based model, land use parameters of the geographical scope is essential, along with the socioeconomic population data, for developing the model.

For instance, land use data has a direct impact on trip generation step, i.e. trip production and trip attraction, within the traditional transport models framework. On the other hand, land use data have an effect on the type and amount of activities conducted by the individuals within the activity-based transport models. Moreover, it plays a vital role in the mode choice and the time-of-day when the trip/activity is carried out.

(Litman, 2006) referred to the observed correlation between the vehicle-mile travelled and the mixed land use, to what extent related land uses (housing, commercial, institutional) are mixed in the same spatial unit. It is found that a better land use mix leads to a reduction in the vehicle travel distance per person, and boosts the use of alternative transport modes, especially walking for doing some basic daily errands.

The districts that contain a decent level of land use mixture normally have between 5% and 15% lower vehicle-miles travelled. **Figure 6-7** visualises the impact of adopting mixed land use zones on travel behaviour.



Source: (Nabil & Elsayed, 2014)

Figure 6-7: The relationship between the existence of Mixed Use and the Travel Behaviour

Likewise, (Litman, 2006) mentioned the role of the land use density factor, which is defined as the number of persons or jobs enclosed per unit of land area, e.g. square km and hectare, in influencing the travel demand and movement patterns. It is discovered that enhanced density causes a reduction in vehicle travel per person. The relation is each 10% rise in land use densities usually diminishes the vehicle-miles travelled per person by 2.5% on average.

Therefore, the FEATHERS model framework incorporates fundamental land use data attributes associated with the aforementioned sub-zones. The involved attributes along with their description can be summarized in **Table 6-1**.

Table 6-1: Land use data incorporated in FEATHERS model

Attribute	Value Type	Description
SUNZONE0	Integer (0,2385)	Unique value for each sub-zone in the Flanders region
POP_2	Integer	Population per sub-zone
DAG_4	Integer	Number of day-shops located in the sub-zone
LLSCH_5	Integer	Total number of students residing in the sub-zone
URB_6	Integer	Categorical data indicating the urbanisation degree of the sub-zone
NONDAG_7	Integer	Number of non-day(night)-shops located in the sub-zone
BANK_8	Integer	Number of different bank branches/outlets located in the sub-zone
ARBTOT_9	Integer	Total number of workers/employees residing in the sub-zone
N_HORECA	Integer	Number of hospitality places located in the subzone
AREA	Integer	The spatial area of the sub-zone

iv. Network Skim Matrices

The impedance is a measure of how difficult it is to go from one location to another. It is usually measured as a mix of trip time, travel distance, and travel expenditures, or as a combination of all three. Parking, gasoline, maintenance, and transit charges are all possible travel expenses. Access time, egress time, wait time, and the number of transfers is frequently included in a generic expenses term in transportation, commonly called "Level of Service-LoS" variables.

(Wang & Vom Hofe, 2007) mentioned the importance of the concept of impedance and its vital role in forecasting the volumes of the vehicular traffic assignments on the road network. The concept is characterized by an inverse relationship between the trip costs and the attractiveness of selecting the route. In other words, the higher the trip cost on a certain route is, the greater the impedance of this route will be. Bearing in mind, the high impedance means less attractiveness of the route.

That is why, according to (Gazis, 2002), the majority of the traffic assignment models follow one of the two main diminishing impedance theories. It occurs either through:

1) the individual traveller selects the route which has the minimum value of the impedance (also known as the user-optimal solution), or
2) the model assigns the travellers on various routes so the average trip impedance is minimized for all the road users (also known as the system-optimal solution)

The trip impedance skim matrix is among the most important intermediate output inside any model of travel demand forecasting and is an essential input for activity-based travel demand forecasting models. It indicates some attributes such as the travel time between different zones, the degree of travel time reliability, the associated costs to the trip conducted, etc. These attributes vary depending on the time of day when the trip is conducted. The traditional method to obtain skim matrices is to execute multiple times of time-dependent shortest path calculations (Tian et al., 2020).

In the case of the FEATHERS model, the Network Skim/LoS matrices include the travel time, travel distance and travel cost for each available transport mode between all the possible Origin-Destination pairs.

7. New Parameters Incorporation

After explaining the structure of FEATHERS and showing the taxonomy of the input data, this section sheds light on the selected sub-models within FEATHERS to be tested for incorporating the new accessibility parameters and reflecting them in these sub-models utility functions.

Before explaining the process of incorporation of new parameters in the aforementioned sub-models, it is necessary to clarify how these models are developed and operated. All these sub-models are operated under the Pandas package which is part of the Python Data Analysis Library. They are developed using the Biogeme package which is created to estimate the parameters of different models using the concept of maximum likelihood estimation. Biogeme is especially intended for the discrete choice type of models according to (Bierlaire, M., 2020).

The Maximum Likelihood Estimation (MLE) is a statistical technique utilised to estimate the parameters within a presumed probability distribution considering set of available data. This is accomplished by maximizing a likelihood function to the extent that the available data is most probable under the presumed statistical model. The scaling step within the parameter area which maximizes the value of the likelihood function is known as the maximum likelihood estimate (Rossi, 2018).

In simple terms, some of the input data mentioned in the previous section are used to formulate utility functions that can be used to predict future choices and scenarios. These utility functions consist of variables/parameters which are associated with coefficients, referred to in this context as betas. The various choice models of the FEATHERS include already utility functions defined by a set of parameters.

In order to achieve the objective of this research, it is necessary to incorporate new parameters in these utility functions representing the accessibility of the sub-zones/Traffic Analysis Zones (TAZs) enclosed in the study area of the FEATHERS model. By incorporating these new parameters, new dimensions will be merged to the process of predicting the daily activity schedules through FEATHERS.

In each sub-section, the utility functions contained in the as-received version of FEATHERS are displayed along with the relevant parameters/variables. After then, the new parameters proposed to be incorporated in the as-received utility functions are shown along with testing results of incorporating them. The criteria used to assess the statistical significance of the new parameter is the p-value indicator. The acceptable range of p-value is (p-value ≤ 0.1).

Figure 7-1 illustrates in detail the mechanism of FEATHERS and the outcome of each step in the process of predicting the day activity schedule starting from the population data. Out of the sub-models depicted in the flow chart, 6 sub-models have been selected which are enclosed in the Day-Level and Primary-Activity-Level boxes. These sub-models can be categorized based on 3 main models of the FEATHERS framework as listed in **Table 7-1**.

Table 7-1: The Selected Sub-models for the New Parameters Incorporation

Model Group Level	Model Group	Sub-model Name	Activity Type
Day	**Daily Activity Pattern**	choose__DAP__1_1	All
Primary Activity	**Location Choice**	choose_primary_activity_location_2_1_activity_type_5	Education
		choose_primary_activity_location_2_1_activity_type_6	Shopping
		choose_primary_activity_location_2_1_activity_type_23	Work + Business
		choose_primary_activity_location_2_1_activity_type_47	Drop-off/Pick-up + Other
	Mode Choice	choose_primary_activity_transport_mode_2_2	All

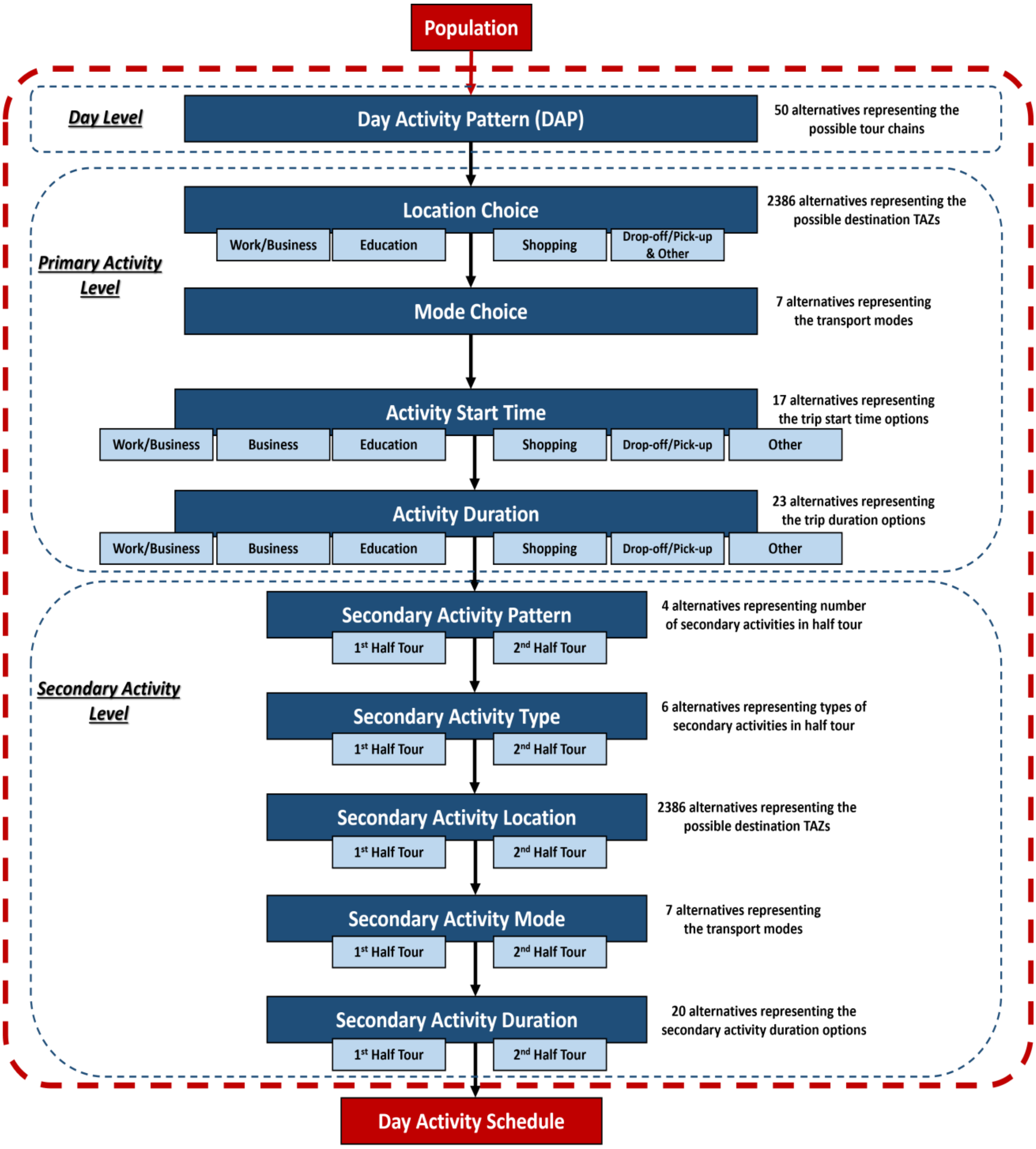


Figure 7-1: Holistic Flow Chart of FEATHERS Framework

a) Group 1: Daily Activity Pattern

The Daily Activity Pattern (DAP) choice model predicts the trip chain that will be conducted by the individual during the day (day level model). In order to predict this daily schedule, all the possible alternatives and their combination should be considered. **Table 7-2** lists the activity type alternatives included in DAP. By combining different activity types, the total number of possible trip chains is 50 alternatives. **Figure 7-2** shows an example of combining a trip chain.

Table 7-2: Activity Types as defined in FEATHERS

Activity Type	Activity Code
Home	0
Work	2
Business	3
Drop-off/Pick-up	4
Education	5
Shopping	6
Other	7

Figure 7-2: Combining Different Activity Types to form a Trip Chain

Each of the 50 alternatives represents a choice that has a unique utility function. Such utility function is defined in terms of the relevant parameters are as follows:

- ✓ Alternative Specific Constant (***b_cte_n***) is a constant value, i.e. the intercept, of the function reflecting the special characteristics of the choice which is not represented by the parameters of the utility function. Therefore, there are 50 values of b_cte to represent all the possible choices, i.e. trip chains.
- ✓ Gender (***b_gender_man***) is the coefficient associated with the fact whether the individual is male or not.

- ✓ Driving License (***b_driving_license***) is the coefficient associated with the fact whether the individual has a driving license or not.
- ✓ Area of the Origin TAZ (***b_area***) is the coefficient that represents the area of the Traffic Analysis Zone where the origin of the trip chain is located.

By combining the above-mentioned parameters (typed in blue) along with their coefficient (typed in green), the utility function formula becomes as below:

$$VString[n] = \begin{cases} b\_cte\_n \\ + b\_gender\_man * (gender == 1) \\ + b\_driving\_license * (driving\_license == 1) \\ + b\_area * AREA \end{cases}$$

<u>Where:</u>

n is the Trip Chain Code, for example (267) – Work-Shopping-Other

VString[n] is the probability of choosing trip chain (n)

(gender == 1) is a binary value (0 for Female, 1 for Male)

(driving_license == 1) is a binary value (0 for Without License, 1 for With License)

(AREA) is an integer value in square kilometres

As can be noted from the considered parameters, the land use and the population data are not properly reflected. That is why, the proposed new parameters are as follows:

- ➢ Day Shops (***b_day_shops***) is the coefficient which is related to the number of shops open during the day (10 AM – 6 PM) located in the origin TAZ
- ➢ Night Shops (***b_night_shops***) is the coefficient which is related to the number of shops open during the night (6 PM – 2 AM) located in the origin TAZ
- ➢ Students (***b_students***) is the coefficient which is related to the number of students residing in the origin TAZ
- ➢ Banks (***b_banks***) is the coefficient which is related to the number of bank branches or outlets located in the origin TAZ
- ➢ Workers (***b_workers***) is the coefficient which is related to the number of workers residing in the origin TAZ
- ➢ HORECA (***b_horeca***) is the coefficient associated with the number of hospitality businesses located in the origin TAZ. HORECA is an acronym used for any Food and Beverage Services (F&B) industry sector. The acronym HORECA is a combination of the first 2 letters of (HOtel, REstaurant, CAfé). It is more commonly used in the so-called Benelux region, i.e. Belgium, Netherlands and Luxembourg.

It is worth mentioning that the decision of which new parameters should be incorporated in each utility function is made depending on the activities involved in the trip chain. For instance, the utility function formula for the trip chain (267) – (Work-Shopping-Other) incorporates ***b_workers, b_day_shops, b_night_shops*** and ***b_horeca***. The activity type (Other) is assumed to be going for recreational purposes such as drinking coffee or having lunch. That is why ***b_horeca*** is incorporated in this case.

Regarding the new parameters, their decimal logarithmic values are calculated before involving in the utility function. Moreover, these parameters are raised to the power (m) whereas (m) is the number of repeating the same activity type within the predicted trip chain.

As such, an example of new utility function formula for the trip chain code (244) – (Work-Education-Education) after incorporating the new parameters and the corresponding coefficients becomes as below:

$$\textbf{VString[244]} = \left\{ \begin{array}{l} b\_cte\_244 \\ + b\_gender\_man * (gender == 1) \\ + b\_driving\_license * (driving\_license == 1) \\ + b\_area * AREA \\ + b\_workers * Log\ (ARBTOT\_9) \\ + b\_students * (Log\ (LLSCH\_5))^2 \end{array} \right.$$

<u>Where:</u>

VString[244] is the probability of choosing trip chain (244) – (Work-Education-Education)

(gender == 1) is a binary value (0: Female, 1: Male)

(driving_license == 1) is a binary value (0: Without License, 1: With License)

(AREA) is an integer value in square kilometres

(ARBTOT_9) is an integer value of the total number of workers in the origin TAZ

(LLSCH_5) is an integer value of the total number of students in the origin TAZ

As can be seen in the example, the logarithmic value is squared, i.e. raised to the power of 2, due to the repetition of conducting the activity type of Education (4) twice in the trip chain (244). **ANNEX A**: Utility Functions of the Daily Activity Pattern (DAP) lists all the 50 utility functions after incorporating the new parameters and coefficients.

All these new elements and their definitions were merged in the Biogeme discrete choice model in order to estimate the coefficient values of the new incorporated parameters. The detailed output of running the Biogeme model is produced in the form of HTML file. The statistical results of the estimation process related to the new coefficients are summarised in **Table 7-3** below.

According to the depicted results, ***b_day_shops*** and ***b_night_shops*** are statistically significant because their p-value is within the acceptable range (p-value ≤ 0.1). On the other hand, the rest of the coefficients fail to meet this criterion. Hence, they are statistically insignificant and should not be considered in the new utility function formula as there is no strong evidence of the correlation between their parameters and the individual's choice decisions within the daily activity pattern.

Table 7-3: Statistical Results of Biogeme Model - Daily Activity Pattern (DAP)

Coefficient	Estimated Value	Standard Error	t-test	p-value
b_banks	0.1500	0.1922	0.7805	0.4351
b_day_shops	-0.5331	0.2921	-1.8253	**0.0680**
b_horeca	-0.0686	0.0695	-0.9873	0.3235
b_night_shops	0.5389	0.2975	1.8113	**0.0701**
b_students	-0.0401	0.0266	-1.5060	0.1321
b_workers	-0.0325	0.1259	-0.2585	0.7960

b) Group 2: Location Choice

After choosing the trip chains through the Daily Activity Pattern, the Location Choice Model comes to predict the destination zone selected by the traveller to carry on the activities planned. The concept of the Location Choice Model is similar to the trip distribution step in the sequence of the traditional travel demand models, i.e. four-step models. As the outcome of this model is specifying the destination zone, the number of alternatives in the choice set is equal to the total number of the sub-zones or Traffic Analysis Zones (TAZs) in FEATHERS, i.e. 2,386 sub-zones.

The utility functions used in the destination prediction are defined in terms of some attributes/parameters which vary across all the sub-zones of the Flanders region. These parameters are listed as follows:

- ✓ (***b_distance***) is the coefficient associated with the travel distances between all the sub-zones
- ✓ (***b_radius_intrazonal***) is the coefficient representing the geographical size of the origin sub-zone
- ✓ (***b_area)*** is the coefficient representing the area of the possible destination sub-zones
- ✓ (***b_size)*** is the coefficient representing the population size of the possible destination sub-zones

- ✓ ***(b_business)*** is the coefficient associated with the activity type whether it is business or not.
- ✓ ***(b_gender_man)*** is the coefficient associated with gender whether the individual is male or not.
- ✓ ***(b_bring_get)*** is the coefficient associated with the activity type whether it is pick-up/drop-off or not.

By combining the above-mentioned parameters (typed in blue) along with their coefficients (typed in green), the utility function formulas for the targeted activity types are as follows:

<u>For Activity Type 5 – Education</u>

```
V[i] =   b['b_distance’] * var['distance_'+str(i)]) * var['interzonal_'+str(i)]
       + b['b_radius_intrazonal’] * var['radius_'+str(i)] * (1 - var['interzonal_'+str(i)])
       + b['b_size'] * log(var['D_POP_2_'+str(i)] + b['b_area’] * var['D_AREA_'+str(i)] + 0.1)
```

<u>For Activity Type 6 – Shopping</u>

```
V[i] =   b['b_distance’] * var['distance_'+str(i)]) * var['interzonal_'+str(i)]
       + b['b_radius_intrazonal’] * var['radius_'+str(i)] * (1 - var['interzonal_'+str(i)])
       + b['b_size’] * log(var['D_POP_2_'+str(i)] + 0.1)
```

<u>For Activity Type 23 – Work + Business</u>

```
V[i] =   b['b_distance’] * var['distance_'+str(i)]) * var['interzonal_'+str(i)]
       + b['b_business’] * (var['activity_type']==3) * var['interzonal_'+str(i)]
       + b['b_gender_man’] * (var['gender']==1) * var['interzonal_'+str(i)]
       + b['b_radius_intrazonal’] * var['radius_'+str(i)] * (1 - var['interzonal_'+str(i)])
       + b['b_size’] * log(var['D_POP_2_'+str(i)] + b['b_area']*var['D_AREA_'+str(i)] + 0.1)
```

<u>For Activity Type 47 – Pick-up/Drop-off + Other</u>

```
V[i] =   b['b_distance’] * var['distance_'+str(i)]) * var['interzonal_'+str(i)]
       + b['b_radius_intrazonal’] * var['radius_'+str(i)] * (1 - var['interzonal_'+str(i)])
       + b['b_bring_get’] * (var['activity_type']==4) * var['interzonal_'+str(i)]
       + b['b_size’] * log(var['D_POP_2_'+str(i)] + 0.1)
```

Where:

V[i] is the probability of choosing the sub-zone number i as a destination

var['distance_'+str(i)]) is a vector of distance values between the sub-zone i and the rest of sub-zones in Flanders region

var['interzonal_'+str(i)]) is a vector of binary values (0: intrazonal activity, 1: interzonal activity)

var['radius_'+str(i)]) is a vector of values of the radius of the circle derived from converting the sub-zone area to an equivalent circle shape

var['D_AREA_'+str(i)]) is a vector of values of the areas of the sub-zones

var['D_POP_2_'+str(i)]) is a vector of values of the population of the sub-zones

var['activity_type']==3 is a binary value (0: Non-business activity, 1: Business activity)

var['activity_type']==4 is a binary value (0: Non-pick-up/-Drop-off activity, 1: Pick-up/Drop-off activity)

var['gender']==1 is a binary value (0: Female, 1: Male)

This research focuses on incorporating new parameters in the utility function of Activity Type 5 – Education. The rationale behind concentrating on Education activities is that it is relatively more probable travelling between the Flanders provinces/regions to follow study programs at Universities and High Schools (Hogescholen). On the other side, it is less probable to do the same, for instance, in the case of shopping activities. However, studying incorporating new parameters in the utility functions of other activities can be the objective of future research papers about FEATHERS.

As can be noted from the considered parameters, the utility function do not consider any accessibility measures that represent to what extent the destination sub-zone is easy/affordable to reach in terms of travel distance, travel time and travel cost. Therefore, the proposed new parameters to fill these gaps are as follows:

- Travel Distance (***b_access_a***) is the coefficient associated with the mean of travel distances using the available transport modes from sub-zone (i) to the rest sub-zones of the model.

- Travel Cost (***b_access_k***) is the coefficient associated with the mean of travel costs using the available transport modes from sub-zone (i) to the rest sub-zones of the model.

- Travel Time (***b_access_t***) is the coefficient associated with the mean of travel times using the available transport modes from sub-zone (i) to the rest sub-zones of the model.

As such, the new utility function formula of Activity Type 5 – Education of the Location Choice Model after incorporating the new parameters and the corresponding coefficients becomes as below:

$$V[i] = \begin{cases} (b[\text{'b\_distance'}] * var[\text{'distance\_'}+str(i)]) * var[\text{'interzonal\_'}+str(i)] \\ + (b[\text{'b\_radius\_intrazonal'}] * var[\text{'radius\_'}+str(i)] + b[\text{'b\_area'}] * var[\text{'D\_AREA\_'}+str(i)]) * (1 - var[\text{'interzonal\_'}+str(i)]) \\ + log((b[\text{'b\_access\_t'}] * var[\text{'GC\_t'}] + (b[\text{'b\_access\_a'}]-1) * var[\text{'GC\_a'}] * var[\text{'distance\_'}+str(i)] + b[\text{'b\_access\_k'}] * var[\text{'GC\_k'}])^2) \\ + b[\text{'b\_size'}] * log(var[\text{'D\_POP\_2\_'}+str(i)] + 0.1) \end{cases}$$

<u>Where:</u>

V[i] is the probability of choosing sub-zone (i) as a destination

var['distance_'+str(i)]) is vectors of distance values between the sub-zone i and the rest of sub-zones in Flanders region

var['interzonal_'+str(i)]) is a binary value (0: intrazonal activity, 1: interzonal activity)

var['radius_'+str(i)]) is a vector of values of the radius of the circle derived from converting the sub-zone area to an equivalent circle shape

var['D_POP_2_'+str(i)]) is a vector of values of the population of the sub-zones

var['GC_k']) the mean of travel monetary costs using the available transport modes from sub-zone (i) to the rest sub-zones of the model

var['GC_t']) the mean of travel times using the available transport modes from sub-zone (i) to the rest sub-zones of the model

var['GC_a'] the mean of travel distances using the available transport from sub-zone (i) to the rest sub-zones of the model

GC_sum is obtained by combining these 3 new parameters (var['GC_k']), var['GC_t']) and var['GC_a']). It is an origin-property indicator that represents the extent to which a certain sub-zone as an origin of a trip is accessible from the rest of the sub-zones. **Figure 7-3** depicts the accessibility indicator in a form of a heatmap where the high values, coloured in red, refer to the high generalised trip cost, i.e. less accessible sub-zone. It is worth mentioning that these values are average of all the possible destinations from each origin sub-zone.

The accessibility indicator values in the map are converted through the max-min normalisation formula (scale 0-100) to make it easier to compare between the sub-zones. As shown in the heatmap, the centric sub-zones have the highest accessibility values (lowest generalised costs). This accessibility diminishes by going further either to the far East (Remersdaal City) or the far West (Veurne City).

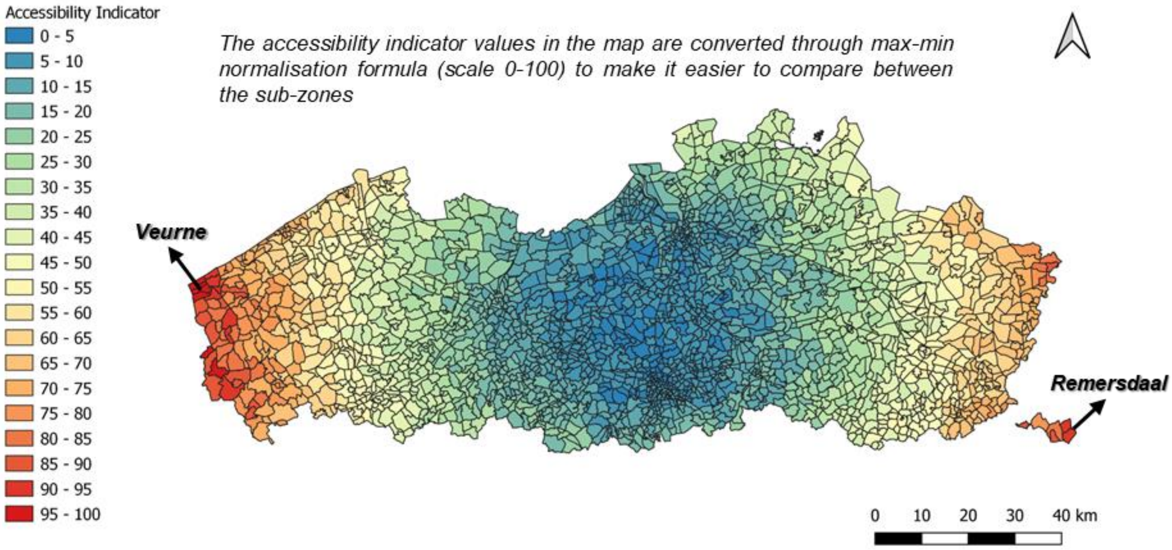


Figure 7-3: Accessibility Indicator (GC_sum) Heatmap (Normalised Values)

The statistical results of the estimation process related to the new coefficients for Activity Type 5 are summarised in **Table 7-4** below. According to the shown results, ***b_access_k*** and ***b_access_t*** are statistically significant because their p-value is within the acceptable range (p-value ≤ 0.1). On the other hand, ***b_access_a*** fails to meet this criterion. Thus, ***b_access_a*** should not be considered in the new utility function formula.

Table 7-4: Statistical Results of Biogeme Model – Location Choice Model (Activity Type 5)

Coefficient	Estimated Value	Standard Error	t-test	p-value
b_access_a	0	1.8e+308	0.000	**1.000**
b_access_k	-11,792	1.2e+03	-9.475	**0.000**
b_access_t	-6,671	6.9e+02	-9.620	**0.000**

c) Group 3: Mode Choice

The Mode Choice Model comes in the sequence of FEATHERS framework after the Location Choice Model. It predicts the transport mode which will be utilised by the traveller to carry on the trip chain selected within the Daily Activity Pattern (DAP) Model in order to commute to the destination zones predicted within the Location Choice Model. FEATHERS framework comprises 7 different transport mode types as listed in **Table 7-5** below.

Table 7-5: Transport Modes as defined in FEATHERS

Transport Mode Type	Mode Code (n)
Micro 5 km/h (e.g. walk)	1
Micro 15 km/h (e.g. bike, step-scooter)	2
Micro 25 km/h (e.g. e-bike, motor-scooter)	3
Private (e.g. car driver)	4
Shared private (e.g. car passenger)	5
Shared on-demand (e.g. minibus, shared taxi)	6
Shared traditional (e.g. public transport: bus, tram, metro, train)	7

Therefore, the choice set of this model includes 7 alternatives. Each of these alternatives represents a choice that has a unique utility function. Such utility function is defined in terms of the relevant parameters are as follows:

- ✓ Alternative Specific Constant (***b_cte_n***) is a constant value, i.e. the intercept, of the function reflecting the special characteristics of the choice which is not represented by the parameters of the utility function. Therefore, there are 7 values of b_cte_n to represent all the possible choices, i.e. transport modes.
- ✓ Age Category (***b_age_person_cat_m_n***) is the coefficient associated with the age category of the individual.
- ✓ Gender (***b_gender_man_n***) is the coefficient associated with the gender whether the individual is male or not.
- ✓ Driving License (***b_driving_license_n***) is the coefficient associated with the fact of whether the individual has a driving license or not.

It is worth noting that the ages of the population are grouped into 7 age categories as listed in **Table 7-6.**

Table 7-6: Population Age Categories

Range (years)	Age Category (m)
Age ≤ 12	1
13 ≤ Age ≤ 17	2
18 ≤ Age ≤ 34	3
35 ≤ Age ≤ 54	4
55 ≤ Age ≤ 64	5
65 ≤ Age ≤ 74	6
Age ≥ 75	7

By combining the above-mentioned parameters (typed in blue) along with their coefficients (typed in green), the utility function formula becomes as below:

$$VString[n] = \begin{cases} b\_cte\_n \\ + b\_gender\_man\_n * (gender == 1) \\ + b\_driving\_license\_n * (driving\_license == 1) \\ + \sum_{m=2}^{7} [b\_age\_person\_cat\_m\_n * (age\_person\_cat == m)] \end{cases}$$

Where:

n is the Transport Mode Code within the range (1,7), for example (5) – Shared private (e.g. car passenger)

m is the age category within the range (1,7)

VString[n] is the probability of choosing Transport Mode Code (n)

(gender == 1) is a binary value (0: Female, 1: Male)

(driving_license == 1) is a binary value (0: Without License, 1: With License)

(age_person_cat==m) is a binary value (0: age in category m, 1: age out of category m)

As can be noted from the considered parameters, indirect accessibility measures reflecting the abilities and opportunities are not represented. Hence, the proposed new parameters to fill these gaps are as follows:

- Day Type (***b_day_cat***) is the coefficient associated with the fact whether the day is weekday or weekend when the transport mode is used.
- Education Level (***b_education***) is the coefficient associated with the educational level of the individual.

- Household Income (***b_hh_income***) is the coefficient associated with the economic stratum in terms of income/salary to which the individual belongs.

As such, the new utility function formula of the Mode Choice Model after incorporating the new parameters and the corresponding coefficients becomes as below:

$$VString[n] = \begin{cases} b\_cte\_n \\ + b\_gender\_man\_n * (gender == 1) \\ + b\_driving\_license\_n * (driving\_license == 1) \\ + b\_day\_cat * (day\_cat) \\ + b\_education * (education) \\ + b\_hh\_income * (hh\_income) \\ + \sum_{m=2}^{7} [b\_age\_person\_cat\_m\_n * (age\_person\_cat == m)] \end{cases}$$

Where:

(day_cat) is a binary value (0: Weekend, 1: Weekday)

(education) is the education category within the range (1,7)

(hh_income) is the economic category within the range (1,7)

ANNEX B: Utility Functions of the Mode Choice Model lists all the 7 utility functions after incorporating the new parameters and coefficients.

Similar to the process carried out in the case of the previous two models, the statistical results of the estimation process related to the new coefficients are summarised in **Table 7-7** below. According to the depicted results, ***b_day_cat*** and ***b_hh_income*** are statistically significant because their p-value is within the acceptable range (p-value ≤ 0.1). On the other hand, **b_education** fails to meet this criterion. Thus, **b_education** should not be considered in the new utility function formula.

Table 7-7: Statistical Results of Biogeme Model – Mode Choice Model

Coefficient	Estimated Value	Standard Error	t-test	p-value
b_day_cat	0.2756	0.1008	2.7346	**0.0062**
b_education	-0.0117	0.0293	-0.3973	0.6911
b_hh_income	0.0736	0.0276	2.6692	**0.0076**

8. Conclusion and Discussion

Regarding the Daily Activity Pattern (DAP), the choice decisions for the trip chains are limited in the as-received version of the FEATHERS model to some parameters. This group of parameters is correlated to characteristics of the traveller such as whether s/he has a driving license or geographical attribute of the origin sub-zone of the traveller. It neglects the rich picture which can be described by the land use and demographic data of each sub-zone.

Consequently, room for improvement was found to make the choice prediction more accurate. The enhancement of the Daily Activity Pattern (DAP) has been achieved by adding the land use and demographic data parameters that reflect the effect of the land use mix of the origin sub-zone on the activity pattern within the FEATHERS study region. The rationale behind it was based on the definition of the accessibility measures presented by (Geurs & Van Wee, 2012) as shown in **Figure 3-1** whereas:

- the land use component is reflected in terms of "*how many banks and shops are available nearby the household*"
- the individual conditions are reflected in terms of "*how many students/workers in the origin sub-zone need to go to the school/workplace*"

Among the new parameters proposed for DAP, b_banks, b_day_shops, b_horeca and b_night_shops fall under the "land use component" and the available options within the individual's sub-zone while b_students and b_workers are projecting the "individual conditions" to either study or work. The hypotheses of the correlation between the number of day and night shops available in the sub-zone and the DAP choices are approved.

As an example to clarify this approved correlation, the utility function of shopping is:

VString[6] = """b_cte_6 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1)   + b_area * AREA + b_day_shops* Log_DAG_4 + b_night_shops* Log_NONDAG_7

Assuming there is a **Male** with a **driving license** who lives in **Sub-zone No. 47** (Area = 268, Day-shops = 289, Night-shops = 289), the predicted utility/probability that he will plan for a shopping activity is:
V(Shopping) = 0.416806 + (-0.083514 * 1) + (-1.568903 * 1) + (0.000297 * 268) + (-0.5331 * log(289)) + (0.5389 * log(289)) = **-1.141**

If the **same person** lives in **Sub-zone No. 1630** (Area = 268, Day-shops = 71, Night-shops = 71), the predicted utility/probability that he will plan for a shopping activity is:
V(Shopping) = 0.416806 + (-0.083514 * 1) + (-1.568903 * 1) + (0.000297 * 268) + (-0.5331 * log(71)) + (0.5389 * log(71)) = **-1.145**

The takeaway from the statistical results of this example is that:

- the fewer shops are located in the origin zone, the fewer shopping activities are expected to be planned in the individual daily schedule.

This is due to the desire of the individual to aggregate the shopping activities as the (home-shop) distances are longer because of the low shops density in the sub-zone.

Regarding the Location Choice Model, the probability of choosing a certain sub-zone as a destination within the FEATHERS framework was obtained based on the interzonal distances between this sub-zone and the other sub-zones in addition to the area and the population size of this sub-zone. Such a notion neglects the overall accessibility, in terms of travel time, distance and monetary cost, of each sub-zone in the Flanders region.

Therefore, three new parameters have been introduced in this research representing the mean values of travel time, travel distance and travel cost to travel from the sub-zone to all the rest of the sub-zones. Thereafter, the mean of these values per origin sub-zone are calculated across the different transport modes to get mean travel time (GC_t), mean travel distance (GC_a) and mean travel cost (GC_k). Afterward, (GC_t), (GC_a) and (GC_k) are normalised in a 0-100 scale. By taking the sum of these normalised values, single accessibility (GC_sum) indicator can be derived for each origin sub-zone.

It is found that the central sub-zones, e.g. Brussels, have higher accessibility indicators than those on the edges of the Flanders region, e.g. Remersdaal City. The main takeaway from the new parameter incorporation is that:

- the inhabitants of centric sub-zones have an advantageous connectivity conditions to travel to the rest of sub-zones.

The negative values of the coefficients b_access_k and b_access_t prove that when the GC_k and GC_t values increase the utility value of travelling to whatever destination sub-zone decrease due to the higher travel costs/poorer accessibility.

It is worth mentioning that these results are applicable to only Activity Type 5 – Education. The rationale behind concentrating on Education activities is that it is relatively more probable travelling between the Flanders provinces/regions to follow study programs at Universities and High Schools (Hogescholen). On the other side, it is less probable to do the same, for instance, in the case of shopping activities. However, studying incorporating new parameters in the utility functions of other activities can be the objective of future research papers about FEATHERS..

Regarding the Mode Choice Model, the prediction of the chosen transport mode within the existing FEATHERS version focuses only on some socio-economic attributes of the individual such as age, gender and owning a driving license. On the other hand, it abandons some elements of the accessibility described in **Figure 3-1** namely the temporal constraints and the individual's socio-economic attributes other than the aforementioned ones.

Accordingly, the necessity for merging the new parameters appeared in order to augment the prediction precision. It has been accomplished by involving the day type on which the activity is conducted by the individual, i.e. b_day_cat, the education level of the individual, i.e. b_education, and the income of the household to which the individual belongs, i.e. b_hh_income.

The data included in the Level-of-Service (LoS) matrix, i.e. cost, time and distance, of each transport mode are not incorporated in this model because the destination sub-zone is not available in the mode choice data file. Therefore, the LoS values cannot be associated with the mode choice via the destination sub-zone number.

Between the added parameters to the Mode Choice Model, b_day_cat falls under the "temporal contains" while b_education and b_hh_income are projecting the "individual abilities and opportunities". The hypotheses of the correlation between the day type and the household income from a side and the Mode Choice prediction are approved. The findings from the results of the estimation of the coefficients can be interpreted as follows:

- On weekdays, the utilisation of the available transport modes is higher than their counterparts during the weekend.
- The higher household income is, the higher usage of the available transport modes is forecasted

Throughout this study, the targeted research questions, focusing on the Day Level and the Primary Activity Level, were answered as follows:

1- How to incorporate the accessibility features in the activity-based model?
   - ✓ In order to incorporate the accessibility aspect in an activity-based model, e.g. FEATHERS, diverse parameters should be considered in the utility functions. The approach adopted in this research is the summary statistics of the direct magnitudes of accessibility measures. Nevertheless, the logsum approaches, i.e. Detailed logsum or Aggregate logsum, can be investigate in future research on FEATHERS.

2- What is the most suitable way to measure accessibility in the activity-based model?
   - ✓ It is noted that there is no one-fit-for-all measure, For example, the location-based measure is compatible with the Location Choice Model while the land use pattern and the socio-economic characteristics are suitable for the Daily Activity Patter (DAP) and the Mode Choice Model respectively.

   a) Can the location-based measures be the most compatible way for the activity-based model?
      - ✓ The location-based measure is the simplest way and most relevant one in the case of the Location Choice Model. However, it was necessary to reflect the location by considering jointly the travel time, travel distance and monetary travel cost, not only the travel distance. They are the elements of the generalised cost which affect the destination/location choice.

REFERENCES

(Arentze, 2004) Arentze, Theo & Timmermans, H.J.P.. (2004). Albatross: A Learning Based Transportation Oriented Simulation System. Transportation Research Part B: Methodological. 38. 613-633. 10.1016/j.trb.2002.10.001.

(Axhausen & Garling, 1992) Axhausen, K., Garling, T. (1992). "Activity-based approaches to travel analysis: conceptual frameworks, models and research problems", Transport Reviews, 12, 324-341.

(Bao et al, 2014) Bao, Qiong & Kochan, Bruno & Bellemans, Tom & Janssens, Davy & Wets, Geert. (2014). Geographical Extension of the Activity-based Modeling Framework FEATHERS. Procedia Computer Science. 32. 774-779. 10.1016/j.procs.2014.05.490.

(Bao et al, 2018) Bao, Qiong & Kochan, Bruno & Shen, Yongjun & Creemers, Lieve & Bellemans, Tom & Janssens, Davy & Wets, Geert. (2018). Applying FEATHERS for travel demand analysis: Model considerations. Applied Sciences. 8. 211. 10.3390/app8020211.

(Beckman et al., 1996) Beckman, R.J., Baggerly, K.A., McKay, M.D . 1996 . Creating synthetic baseline populations. Transportation Research Part A: Policy and Practice 30, 415-429.

(Beckx et al., 2009) Beckx, C., Panis, L. I., 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. Environmental Impact Assessment Review 29 (3), 179–185.

(Ben Akiva, 2008) MIT OpenCourseWare. (2008). Transportation Systems Analysis: Demand and Economics. Retrieved from: https://ocw.mit.edu/courses/civil-and-environmental-engineering/1-201j-transportation-systems-analysis-demand-and-economics-fall-2008/lecture-notes/MIT1_201JF08_lec05.pdf

(Berechman, 1981) Berechman, J. (1981), Transportation, Temporal, and Spatial Components of Accessibility, by Lawrence D. Burns. Geographical Analysis, 13: 185-187. https://doi.org/10.1111/j.1538-4632.1981.tb00726.x

(Bhat et al, 2003) Bhat, R., Koppelman, S. (2003). "Activity-Based modeling of travel demand" Handbook of transportation Science. International Series in Operations Research & Management Science, Volume 56, Part 2, 39-65.

(Bierlaire, M., 2020) Bierlaire, M. (2020). A short introduction to PandasBiogeme. Technical report TRANSP-OR 200605. Transport and Mobility Laboratory, ENAC, EPFL.

(Castiglione, 2015) Castiglione, J., Bradley, M. A., & Gliebe, J. (2015). Activity-based Travel Demand Models. Transportation Research Board.

(Choupani et al, 2016) Choupani, Ahad & Mamdoohi, Amir Reza. (2016). Population Synthesis Using Iterative Proportional Fitting (IPF): A Review and Future Research. Transportation Research Procedia. 17. 10.1016/j.trpro.2016.11.078.

(Chorus et al, 2011) Chorus, Caspar & Jong, Gerard. (2011). Modeling experienced accessibility for utility-maximizers and regret-minimizers. Journal of Transport Geography - J TRANSP GEOGR. 19. 10.1016/j.jtrangeo.2011.02.009.

(Chu et al., 2012) Chu, Zhaoming & Cheng, Lin & Chen, Hui. (2012). A Review of Activity-Based Travel Demand Modeling. 48-59. 10.1061/9780784412442.006.

(Department Mobiliteit & Openbare Werken, 2016) Onderzoek verplaatsingsgedrag Vlaanderen (2015-2016) Analyserapport. Retrieved from: https://assets.vlaanderen.be/image/upload/v1597747051/samenvatting_zt1htk.pdf

(Ferdous et al, 2012) Ferdous, Nazneen & Bhat, Chandra & Vana, Lakshmi & Schmitt, David & Bowman, John & Bradley, Mark & Pendyala, Ram & Anderson, Rebekah & Giaimo, Gregory. (2012). A Comparison of the Four-Step Versus Tour-Based Models in the Context of Predicting Travel Behavior Before and After Transportation System Changes. Transportation Research Record Journal of the Transportation Research Board. 2303. 10.3141/2303-06.

(Gazis, 2002) Gazis, Denos. (2002). The Origins of Traffic Theory. Operations Research. 50. 69-77. 10.1287/opre.50.1.69.17776.

(Geurs et al., 2004) Geurs, Karst & Wee, Bert. (2004). Accessibility evaluation of land-use and transport strategies: Review and research directions. Journal of Transport Geography. 12. 127-140. 10.1016/j.jtrangeo.2003.10.005.

(Geurs & Van Wee, 2012) Geurs, K.T. and B. van Wee (2004), "Accessibility evaluation of land-use and transport strategies: Review and research directions". Journal of Transport Geography, Vol. 12/2, pp. 127-140.

(Geurs et al., 2012) Geurs, K.T., M. de Bok and B. Zondag (2012), "Accessibility benefits of integrated land use and public transport policy plans in the Netherlands", in Accessibility Analysis and Transport Planning: Challenges for Europe and North America, K. T. Geurs, K. Krizek and A. Reggiani (eds.), Edward Elgar, la, pp. 189-216. and research directions". Journal of Transport Geography, Vol. 12/2, pp. 127-140.

(Geurs & Van Wee, 2013) Geurs, K. T., van Wee, B., & Rietveld, P. (2013). Discussing the logsum as an accessibility indicator. 45-45. Abstract from XII NECTAR 2013 International Conference, Sao Miguel Island, Portugal.

(Geurs, 2018) Geurs, K. T. (2018). Transport Planning With Accessibility Indices in the Netherlands: Discussion Paper. OECD.

(Gould, 1969) Gould, P. (1969). Spatial Diffusion, Commission on College Geography, Association of American Geographers. Washington, DC.

(Hwang et al., 2017) Hwang, Jeong & Kim, Hyunmyung & Cho, Sungjin & Bellemans, Tom & Lee, Won Do & Choi, Keechoo & Cheon, Seung & Joh, Ch. (2017). An examination of the accuracy of an activity-based travel simulation against smartcard and navigation device data. Travel Behaviour and Society. 7. 10.1016/j.tbs.2017.01.001.

(Litman, 2006) Todd Litman (2006), Smart Growth Policy Reforms, Victoria Transport Policy Institute (www.vtpi.org); at www.vtpi.org/smart_growth_reforms.pdf

(Litman, 2013) Todd Litman (2013), "The New Transportation Planning Paradigm," ITE Journal (www.ite.org), Vol. 83, June, pp. 20-28; at www.vtpi.org/paradigm.pdf

(Miller, 2018) Eric J. Miller (2018) Accessibility: measurement and application in transportation planning, Transport Reviews, 38:5, 551-555, DOI: 10.1080/01441647.2018.1492778

(Nabil & Elsayed, 2014) Nabil, Noha & Elsayed, Gehan. (2014). Influence of mixed land-use on realizing the social capital. HBRC Journal. 52. 10.1016/j.hbrcj.2014.03.009.

(National Academies of Sciences, 2014) National Academies of Sciences, Engineering, and Medicine 2014. Activity-Based Travel Demand Models: A Primer. Washington, DC: The National Academies Press. https://doi.org/10.17226/22357

(Tian et al, 2020) Tian, Ye & Chiu, Yi-Chang & Sun, Jian & Chai, Chen. (2020). Sunsetting Skim Matrices: A Trajectory-Mining Approach to Derive Travel Time Skim Matrix in Dynamic Traffic Assignment for Activity-base Model Integration.

(Rodrigue, 2020) Rodrigue, J. (2020). The Geography of Transport Systems (5th ed.). Routledge.

(Rossi, 2018) Rossi, Richard J. (2018). Mathematical Statistics : An Introduction to Likelihood Based Inference. New York: John Wiley & Sons. p. 227. ISBN 978-1-118-77104-4.

(Wang et al, 2013) Wang, Yang and Monzón de Cáceres, Andrés and Di Ciommo, Floridea (2013). Assessing the impact of adaptive accessibility on the optimal transport policy implementation by using an integrated land-use/transport model for Madrid. In: "NECTAR 2013 International Conference: Dynamics of global and local networks", 16/06/2013 - 18/06/2013, São Miguel Island, Azores, Portugal. p. 46.

(Wang & Vom Hofe, 2007) Wang, Xinhao & Hofe, Rainer. (2007). Research Methods in Urban and Regional Planning. 10.1007/978-3-540-49658-8.

(Yang et al., 2010) Yang, D., Arentze, T., Timmermans, H., 2010. Primary and secondary effects of teleworking policies on household energy consumption. In: Proceedings of the 12th World Conference on Transport Research, Lisbon, Portugal.

ANNEX A: Utility Functions of the Daily Activity Pattern (DAP)

VString[2] ="""0"""

VString[0] = """b_cte_0 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA"""

VString[7] = """b_cte_7 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_horeca* Log_N_HORECA"""

VString[5] = """b_cte_5 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_students* Log_LLSCH_5"""

VString[6] = """b_cte_6 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_day_shops* Log_DAG_4 + b_night_shops* Log_NONDAG_7"""

VString[27] = """b_cte_27 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_workers* Log_ARBTOT_9 + b_horeca* Log_N_HORECA"""

VString[57] = """b_cte_57 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_students* Log_LLSCH_5 + b_horeca* Log_N_HORECA"""

VString[67] = """b_cte_67 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_day_shops* Log_DAG_4 + b_night_shops* Log_NONDAG_7 + b_horeca* Log_N_HORECA"""

VString[77] = """b_cte_77 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_horeca* pow(Log_N_HORECA, 2)"""

VString[26] = """b_cte_26 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_workers* Log_ARBTOT_9 + b_day_shops* Log_DAG_4"""

VString[55] = """b_cte_55 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_students* pow(Log_LLSCH_5, 2)"""

VString[4] = """b_cte_4 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_students* Log_LLSCH_5"""

VString[66] = """b_cte_66 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_day_shops* pow(Log_DAG_4, 2) + b_night_shops* pow(Log_NONDAG_7, 2)"""

VString[557] = """b_cte_557 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area* AREA + b_students* pow(Log_LLSCH_5, 2) + b_horeca* Log_N_HORECA"""

VString[56] = """b_cte_56 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_students* Log_LLSCH_5 + b_day_shops* Log_DAG_4 + b_night_shops* Log_NONDAG_7"""

VString[3] = """b_cte_3 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_banks* Log_BANK_8"""

VString[22] = """b_cte_22 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_workers* Log_ARBTOT_9"""

VString[677] = """b_cte_677 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_day_shops* Log_DAG_4 + b_night_shops* Log_NONDAG_7 + b_horeca* pow(Log_N_HORECA, 2)"""

VString[25] = """b_cte_25 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_workers* Log_ARBTOT_9 + b_students* Log_LLSCH_5"""

VString[24] = """b_cte_24 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_workers* Log_ARBTOT_9 + b_students* Log_LLSCH_5"""

VString[44] = """b_cte_44 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_students* Log_LLSCH_5"""

VString[46] = """b_cte_46 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_students* Log_LLSCH_5 + b_day_shops* Log_DAG_4 + b_night_shops* Log_NONDAG_7"""

VString[267] = """b_cte_267 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_workers* Log_ARBTOT_9 + b_day_shops* Log_DAG_4 + b_night_shops* Log_NONDAG_7 + b_horeca* Log_N_HORECA"""

VString[577] = """b_cte_577 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_students* Log_LLSCH_5 + b_horeca* pow(Log_N_HORECA, 2)"""

VString[47] = """b_cte_47 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_students* Log_LLSCH_5 + b_horeca* Log_N_HORECA"""

VString[667] = """b_cte_667 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_day_shops* pow(Log_DAG_4, 2) + b_night_shops* pow(Log_NONDAG_7, 2) + b_horeca* Log_N_HORECA"""

VString[777] = """b_cte_777 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_horeca* pow(Log_N_HORECA, 3)"""

VString[446] = """b_cte_446 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_students* pow(Log_LLSCH_5, 2) + b_day_shops* Log_DAG_4 + b_night_shops* Log_NONDAG_7"""

VString[277] = """b_cte_277 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_workers* Log_ARBTOT_9 + b_horeca* pow(Log_N_HORECA, 2)"""

VString[45] = """b_cte_45 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_students* Log_LLSCH_5"""

VString[567] = """b_cte_567 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_students* Log_LLSCH_5 + b_day_shops* Log_DAG_4 + b_night_shops* Log_NONDAG_7 + b_horeca* Log_N_HORECA"""

VString[467] = """b_cte_467 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_students* Log_LLSCH_5 + b_day_shops* Log_DAG_4 + b_night_shops* Log_NONDAG_7 + b_horeca* Log_N_HORECA"""

VString[227] = """b_cte_227 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_workers* Log_ARBTOT_9 + b_horeca* Log_N_HORECA"""

VString[447] = """b_cte_447 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_students* pow(Log_LLSCH_5, 2) + b_horeca* Log_N_HORECA"""

VString[556] = """b_cte_556 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_students* pow(Log_LLSCH_5, 2) + b_day_shops* Log_DAG_4 + b_night_shops* Log_NONDAG_7"""

VString[477] = """b_cte_477 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_students* Log_LLSCH_5 + b_horeca* pow(Log_N_HORECA, 2)"""

VString[666] = """b_cte_666 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_day_shops* pow(Log_DAG_4, 3) + b_night_shops* pow(Log_NONDAG_7, 3)"""

VString[37] = """b_cte_37 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_banks* Log_BANK_8 + b_horeca* Log_N_HORECA"""

VString[226] = """b_cte_226 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_workers* Log_ARBTOT_9 + b_day_shops* Log_DAG_4 + b_night_shops* Log_NONDAG_7"""

VString[23] = """b_cte_23 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_workers* Log_ARBTOT_9 + b_banks* Log_BANK_8"""

VString[444] = """b_cte_444 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_students* pow(Log_LLSCH_5, 3)"""

VString[244] = """b_cte_244 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_workers* Log_ARBTOT_9 + b_students* pow(Log_LLSCH_5, 2)"""

VString[247] = """b_cte_247 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_workers* Log_ARBTOT_9 + b_students* Log_LLSCH_5 + b_horeca* Log_N_HORECA"""

VString[5577] = """b_cte_5577 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1)+ b_area * AREA + b_students* Log_LLSCH_5 + b_horeca* pow(Log_N_HORECA, 2)"""

VString[257] = """b_cte_257 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_workers* Log_ARBTOT_9 + b_students* Log_LLSCH_5 + b_horeca* Log_N_HORECA"""

VString[266] = """b_cte_266 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_workers* Log_ARBTOT_9 + b_day_shops* pow(Log_DAG_4, 2) + b_night_shops* pow(Log_NONDAG_7, 2)"""

VString[4447] = """b_cte_4447 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1)+ b_area * AREA + b_students* pow(Log_LLSCH_5, 3) + b_horeca* Log_N_HORECA"""

VString[4444] = """b_cte_4444 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1)+ b_area * AREA + b_students* pow(Log_LLSCH_5, 4)"""

VString[36] = """b_cte_36 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1) + b_area * AREA + b_banks* Log_BANK_8 + b_day_shops* Log_DAG_4 + b_night_shops* Log_NONDAG_7"""

VString[6777] = """b_cte_6777 + b_gender_man* (gender==1) + b_driving_license* (driving_license==1)+ b_area * AREA + b_day_shops* Log_DAG_4 + b_night_shops* Log_NONDAG_7 + b_horeca* pow(Log_N_HORECA, 3)"""

ANNEX B: Utility Functions of the Mode Choice Model

VString[1] = """0"""

VString[2] = """b_cte_2 + b_age_person_cat_2_2*(age_person_cat==2) + b_age_person_cat_3_2*(age_person_cat==3)+b_age_person_cat_4_2*(age_person_cat==4)+b_age_person_cat_5_2*(age_person_cat==5)+ b_age_person_cat_6_2*(age_person_cat==6)+ b_age_person_cat_7_2*(age_person_cat==7)+b_gender_man_2*(gender==1)+b_driving_license_2*(driving_license==1) + b_day_cat * (day_cat) + b_education * (education) + b_hh_income * (hh_income)"""

VString[3] = """b_cte_3 + b_age_person_cat_2_3*(age_person_cat==2) + b_age_person_cat_3_3*(age_person_cat==3)+b_age_person_cat_4_3*(age_person_cat==4)+b_age_person_cat_5_3*(age_person_cat==5)+ b_age_person_cat_6_3*(age_person_cat==6) +b_age_person_cat_7_3*(age_person_cat==7)+b_gender_man_3*(gender==1)+b_driving_license_3*(driving_license==1) + b_day_cat * (day_cat) + b_education * (education) + b_hh_income * (hh_income)"""

VString[4] = """b_cte_4 + b_age_person_cat_2_4*(age_person_cat==2) + b_age_person_cat_3_4*(age_person_cat==3)+b_age_person_cat_4_4*(age_person_cat==4)+b_age_person_cat_5_4*(age_person_cat==5)+ b_age_person_cat_6_4*(age_person_cat==6) +b_age_person_cat_7_4*(age_person_cat==7)+b_gender_man_4*(gender==1)+b_driving_license_4*(driving_license==1) + b_day_cat * (day_cat) + b_education * (education) + b_hh_income * (hh_income)"""

VString[5] = """b_cte_5 + b_age_person_cat_2_5*(age_person_cat==2) + b_age_person_cat_3_5*(age_person_cat==3)+b_age_person_cat_4_5*(age_person_cat==4)+b_age_person_cat_5_5*(age_person_cat==5)+ b_age_person_cat_6_5*(age_person_cat==6) + b_age_person_cat_7_5*(age_person_cat==7)+b_gender_man_5*(gender==1)+b_driving_license_5*(driving_license==1) + b_day_cat * (day_cat) + b_education * (education) + b_hh_income * (hh_income)"""

VString[6] = """b_cte_6 + b_age_person_cat_2_6*(age_person_cat==2) + b_age_person_cat_3_6*(age_person_cat==3)+b_age_person_cat_4_6*(age_person_cat==4)+b_age_person_cat_5_6*(age_person_cat==5)+ b_age_person_cat_6_6*(age_person_cat==6) + b_age_person_cat_7_6*(age_person_cat==7)+b_gender_man_6*(gender==1)+b_driving_license_6*(driving_license==1) + b_day_cat * (day_cat) + b_education * (education) + b_hh_income * (hh_income)"""

VString[7] = """b_cte_7 + b_age_person_cat_2_7*(age_person_cat==2) + b_age_person_cat_3_7*(age_person_cat==3)+b_age_person_cat_4_7*(age_person_cat==4)+b_age_person_cat_5_7*(age_person_cat==5)+ b_age_person_cat_6_7*(age_person_cat==6) + b_age_person_cat_7_7*(age_person_cat==7)+b_gender_man_7*(gender==1)+b_driving_license_7*(driving_license==1) + b_day_cat * (day_cat) + b_education * (education) + b_hh_income * (hh_income)"""