

## Research Article

# Investigating the Minimum Size of Study Area for an Activity-Based Travel Demand Forecasting Model

**Qiong Bao, Yongjun Shen, Lieve Creemers, Bruno Kochan,  
Tom Bellemans, Davy Janssens, and Geert Wets**

*Transportation Research Institute (IMOB), Hasselt University, Wetenschapspark 5, Bus 6, 3590 Diepenbeek, Belgium*

Correspondence should be addressed to Yongjun Shen; [yongjun.shen@uhasselt.be](mailto:yongjun.shen@uhasselt.be)

Received 28 July 2014; Accepted 16 October 2014

Academic Editor: Wuhong Wang

Copyright © 2015 Qiong Bao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Nowadays, considerable attention has been paid to the activity-based approach for transportation planning and forecasting by both researchers and practitioners. However, one of the practical limitations of applying most of the currently available activity-based models is their computation time, especially when large amount of population and detailed geographical unit level are taken into account. In this research, we investigated the possibility of restraining the size of the study area in order to reduce the computation time when applying an activity-based model, as it is often the case that only a small territory rather than the whole region is the focus of a specific study. By introducing an accuracy level of the model, we proposed in this research an iteration approach to determine the minimum size of the study area required for a target territory. In the application, we investigated the required minimum size of the study area surrounding each of the 327 municipalities in Flanders, Belgium, with regard to two different transport modes, that is, car as driver and public transport. Afterwards, a validation analysis and a case study were conducted. All the experiments were carried out by using the FEATHERS, an activity-based microsimulation modeling framework currently implemented for the Flanders region of Belgium.

## 1. Introduction

As an alternative to the traditional four-step model of travel demand, the activity-based approach has currently been given more and more attention by transportation researchers and has resulted in the development of a number of practical models, such as ALBATROSS [1], RAMBLAS [2], CEMDAP [3], and FAMOS [4]. Since the approach focuses on the complete activity behavior patterns and adopts a holistic framework considering the individual interactions and spatiotemporal constraints, it explicitly reveals the inability of the conventional trip-based approach and can be used to address many policy issues and their impact, such as land use, energy consumption, emission, safety, and congestion pricing [5–7].

Although its usefulness in transportation planning and forecasting has been widely recognized, one of the practical limitations of applying most of the currently available activity-based models is their computation time, especially when large amount of population and detailed geographical unit level are taken into account. For instance, in

the FEATHERS (Forecasting Evolutionary Activity-Travel of Households and Their Environmental RepercussionS) [8], an activity-based microsimulation modeling framework currently implemented for the Flanders region of Belgium, it takes approximately 16 hours for a single model run based on the 10% of the full population of Flanders at the Building block level (currently the most disaggregated geographical level of detail for Belgium; see also Section 2) [9]. If the population fraction increases to 50%, which is also a frequently used fraction for the model operation, the FEATHERS framework will then take almost two days to complete the model execution. Moreover, if multiple model runs are required due to the consideration of stochastic error derived from the microsimulation approach [9], the computation time will be magnified dramatically, which makes the real-time application of the model particularly difficult or even impossible to realize.

In order to reduce the computation time when applying an activity-based model, several tradeoffs can be made in actual applications, one of which is to restrain the size of

the study area and to conduct the computation only for the selected region [10]. The other way round, it is also often the case that merely a small territory (e.g., a municipality) rather than the whole region or country is the focus of a specific study. Therefore, a relatively small study area surrounding the target territory is needed for investigation rather than taking the whole region into account. In this way, the computation time of the model could be saved effectively. The question then becomes, What should be the minimum size of the study area surrounding the target territory and how to determine it? Based on authors' current knowledge, however, not much research effort has been paid to this subject as yet. In most of the current studies using activity-based models, the size of the study area is chosen mainly via domain knowledge of researchers or practitioners themselves, which is actually rather arbitrary. In this study, by defining an accuracy level of the model, we propose an iteration approach to determine the minimum size of the study area required for a target territory when performing travel demand forecasting. More specifically, by adding a small zone to the target territory constituting a new study area each time, the accuracy of the model—defined as the difference between the occurrence of both the departing and the arriving trips derived based on this study area and that based on the whole region—is calculated. Such a procedure is repeated until the predefined accuracy level is satisfied. In the application, we investigate the required minimum size of the study area surrounding each of the 327 municipalities in Flanders, Belgium, with regard to two different transport modes, that is, car as driver and public transport. Afterwards, a validation analysis based on four extreme municipalities is conducted and a case study using the identified minimum study area for the city of Leuven is provided. All the experiments are carried out by using the FEATHERS framework.

The rest of this paper is structured as follows. In Section 2, we briefly introduce the FEATHERS framework and the levels of geographic detail of Flanders. The methodology proposed in this research to determine the minimum size of the study area is elaborated in Section 3, followed by a detailed demonstration of the experiment execution. In Section 4, the results are presented, validated, and further applied to a practical project. The paper ends with conclusions and future research in Section 5.

## 2. FEATHERS Framework for Flanders

FEATHERS (Forecasting Evolutionary Activity-Travel of Households and Their Environmental RepercussionS) [8] is a microsimulation framework particularly developed to facilitate the implementation of activity-based models for transport demand forecasting. Currently, an activity-based model similar to the ALBATROSS model [11] is embedded, in which a sequence of 26 decision trees, derived by means of the chi-squared automatic interaction detector (CHAID) algorithm, is used in the scheduling process and decisions are based on a number of attributes of the individual (e.g., age, gender), of the household (e.g., number of cars), and of the geographical zone (e.g., population density, number of shops). For each individual person with its specific attributes,

the model simulates whether an activity (e.g., shopping, working, leisure activity, etc.) is going to be carried out or not. Subsequently, the location, transport mode, and duration of the activity are determined, taking into account the attributes of the individual. Based on the estimated schedules or activity-travel patterns, travel demand can then be extracted and assigned to the transportation network.

Currently, the FEATHERS framework has been implemented for the Flanders region of Belgium (e.g., [12–14]) and is fully operational at six levels of geographic detail of Flanders, that is, Building block (BB) level, Subzone level, Zone level, Superzone level, Province level, and the whole Flanders level. Figure 1 illustrates the hierarchy of the geographical layers with different granularities.

In practice, to predict the travel demand or the total number of trips happening within a specific zone of Flanders, also named target territory, we normally have to calculate both the departing trips (i.e., the trips from this target territory to the whole Flanders region) and the arriving trips (i.e., the trips from the whole Flanders region to this target territory) (see Figure 2(a)). As a consequence, the more the detailed geographical unit level is considered, the longer the computation time is needed. For instance, to run FEATHERS at the Subzone level, approximately 16 hours is needed based on the 50% of the full population of Flanders. If the most disaggregated geographical level of detail, that is, the BB level is under consideration, the FEATHERS framework will then take almost two days to complete the model execution. Therefore, how to effectively reduce the model computation time is a practical issue of applying this framework.

## 3. Methodology and Experiment

In this study, we aim to find an effective solution to the computation time problem of activity-based models in general and the FEATHERS in particular. As described in the above section, to estimate the total number of trips happening within a target territory, whole Flanders is normally used as the study area to calculate both the departing and arriving trips of this territory. However, if we can find a relatively small study area surrounding this target territory within which most of the departing and arriving trips are generated, it is then not necessary to take the whole Flanders region into account (see Figure 2(b)). In this way, the computation time of the model could be saved effectively. The question then becomes, What should be the minimum size of the study area surrounding the target territory and how to determine it? In this research, we propose an iteration approach based on which we investigate the minimum size of the study area needed for each of the 327 municipalities (i.e., the Superzone level) in Flanders, Belgium. The whole procedure is illustrated in Figure 3.

More specifically, by generating the basic prediction dataset from the activity-based model inside FEATHERS, we obtain the whole activity-travel pattern or schedule information for each individual in Flanders, based on which the origin and destination (OD) matrices can be derived.

Next, for each particular Superzone  $i$ , one more zone  $j$  (which could be a Superzone, a Subzone, or a Building block)

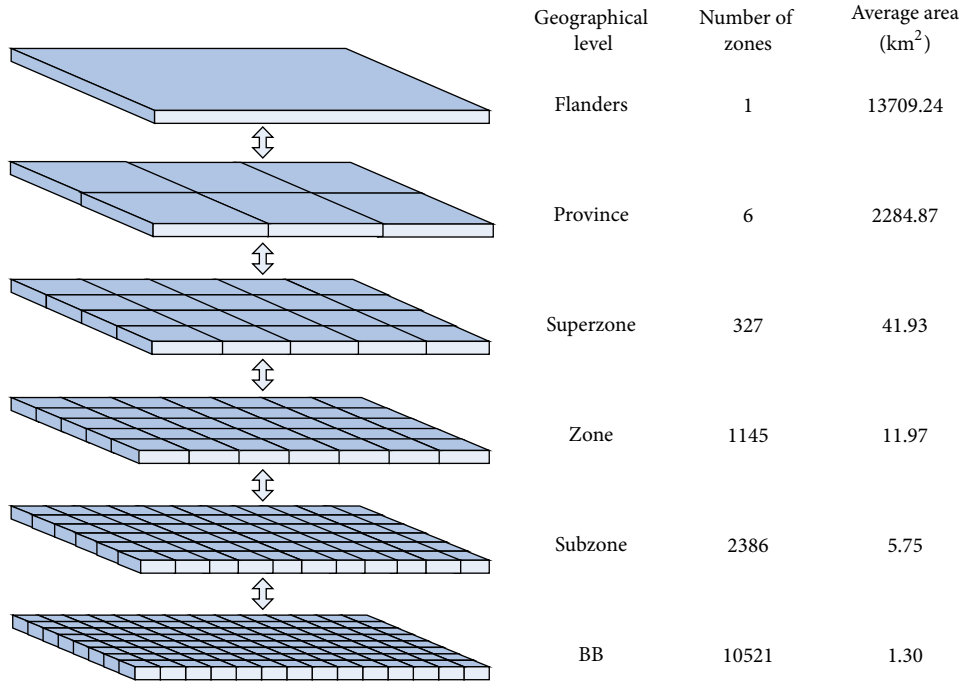


FIGURE 1: Six levels of geographic detail of Flanders used in the FEATHERS.

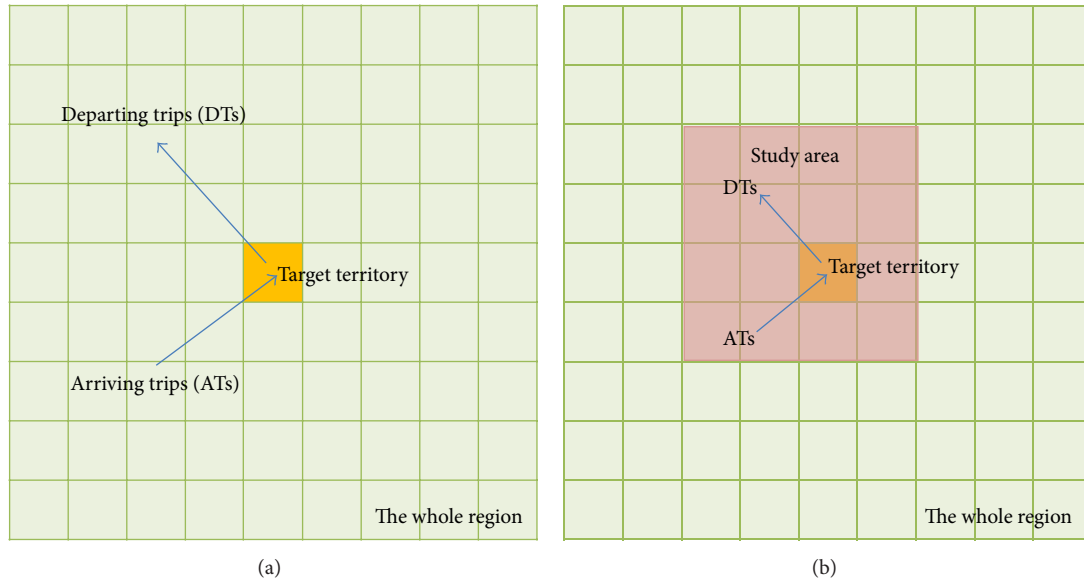


FIGURE 2: The schematic diagrams of travel demand calculation for a target territory.

with the shortest centroid distance to the target Superzone  $i$  is added constituting a new study area (SA). Then, the travel demand (i.e., the number of trips) of both departing mode and arriving mode within this new study area is computed,

and the difference ( $D_{ij}$ ) with that based on whole Flanders can be calculated by (1), which can be further used to estimate the accuracy rate ( $A_{ij}$ ) of this study area by (2):

$$D_{ij} = \frac{\sqrt{(\#DTs \text{ in } SA_{ij} - \#DTs \text{ in Flanders})^2 + (\#ATs \text{ in } SA_{ij} - \#ATs \text{ in Flanders})^2}}{\sqrt{(\#DTs \text{ in Flanders})^2 + (\#ATs \text{ in Flanders})^2}} \quad (1)$$

$$A_{ij} = (1 - D_{ij}) \times 100. \quad (2)$$

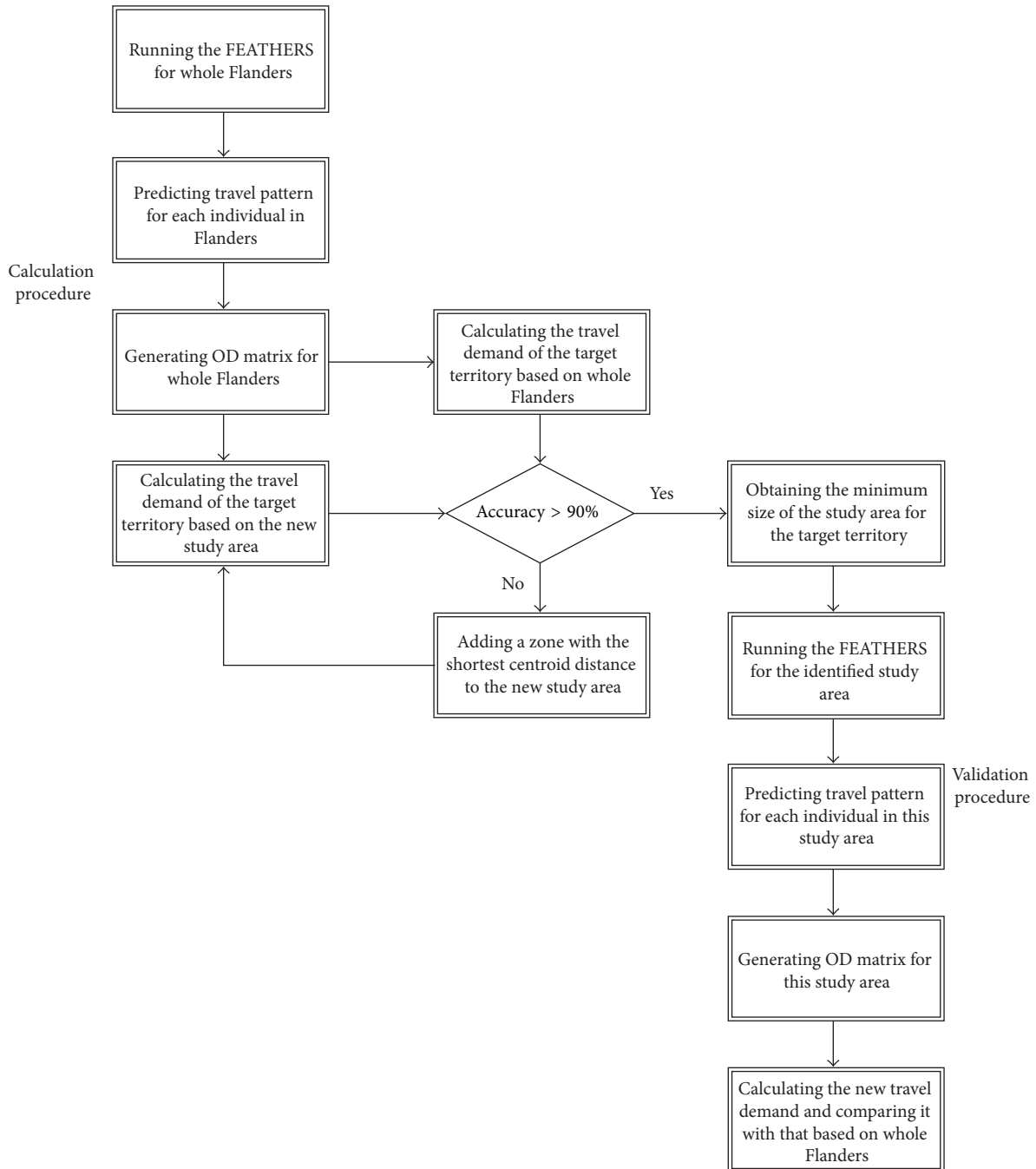


FIGURE 3: The flow diagram of the calculation and the validation procedure.

Such a procedure is repeated; that is, more zones are added from the close to the distant, until the predefined accuracy requirement is reached. The obtained study area is thus the minimum size needed for the travel demand prediction of the target municipality, and the centroid distance between the last zone that was added into the study area and the target municipality is defined as the minimum radius of the study area surrounding this municipality. Here, the radius is different from its conventional conception but refers in particular to the centroid distance between the added zone and the target municipality. Therefore, the radius for

each municipality increases discretely, and when the radius increases once, only one zone is counted into the study area.

In the experiments, the FEATHERS framework is executed at Subzone level for the 50% fraction of the full population, and 90% accuracy level is selected, based on which we investigate the minimum size of the study area needed for each of the 327 municipalities in Flanders, Belgium, with regard to two different transport modes, that is, car as driver and public transport, respectively. Moreover, as a validation procedure (see also Figure 3), we run the FEATHERS again based on the identified study area, respectively, for four

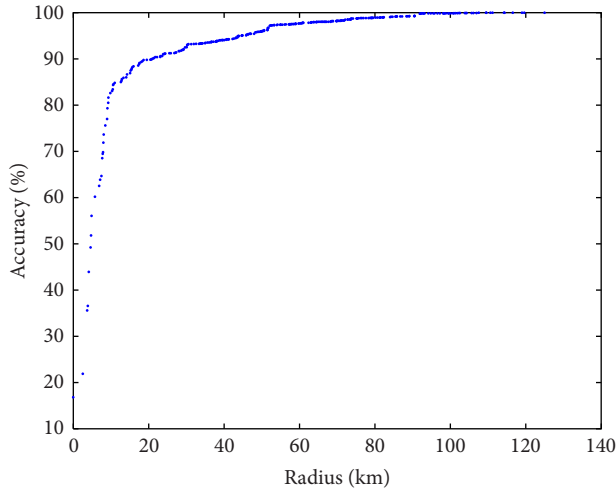


FIGURE 4: The relation between the accuracy level and the required radius for a randomly selected municipality.

extreme cases, that is, the municipality with the longest study area radius and the one with the shortest for each mode. The accuracy of the model in each case is examined, and the degree of the reduction in computation time is estimated. Finally, we apply our results to a practical project which investigates the potential impact of light rail initiatives on travel demand at a local network, and the minimum study area surrounding the city of Leuven is selected as a case study to perform prediction of the travel demand.

## 4. Results and Discussion

By applying the methodology described in Section 3, the corresponding results are presented and further discussed in the following sections.

**4.1. The Minimum Study Area Required for 327 Municipalities in Flanders.** In the experiment, for a particular municipality, by adding the surrounding municipalities one by one from the near to the distant, the radius of the study area is increasing each time. As a result, the difference of the travel demand regarding the departing mode and the arriving mode between the current study area and whole Flanders, according to (1) and (2), is expected to decrease gradually, while the accuracy level of the new study area is going to increase correspondingly. Taking a randomly selected municipality as an example, the relationship between the achieved accuracy and the required radius of the study area is presented in Figure 4.

To further obtain the minimum radius of the study area needed for each of the 327 municipalities in Flanders, a 90% accuracy level is selected in this study with respect to two different transport modes, that is, car as driver and public transport. The distribution of the results for all the 327 municipalities is illustrated in Figure 5.

The figure shows that the required radius for both the modes follows the normal distribution, and the average of the minimum radius needed for the car as driver mode is

39.70 km with the standard deviation of 6.74 km, while the average radius is 50.84 km for the public transport mode with the standard deviation of 6.33 km. In other words, to achieve the same 90% accuracy level for each municipality, the public transport mode needs in general a relatively larger study area compared to the car as driver mode. It can be partly explained by the fact that people in Flanders are more likely to choose the public transport mode (e.g., train) for a long distance trip, especially when the distance is larger than 50 km. Even so, given the fact that the area of whole Flanders is 13709.24 km<sup>2</sup>, the size of the study area needed for most of the municipalities is reduced to a great extent for both the transport modes, and therefore the computation time of the FEATHERS would be saved remarkably (see Section 4.2).

To give a more clear representation of the results, we visualize the calculated minimum radius of the study area for each of the 327 municipalities by using a color theme with 14 different colors for the car as driver mode and the public transport mode, respectively (see Figures 6 and 7), in which the displayed color is more green when a shorter radius is needed for a municipality, while the displayed color tends to be red when the radius needed is becoming larger.

Based on these two figures, we can clearly see that, for most of the municipalities of Flanders, the required minimum radius (or study area) is relatively larger for the public transport mode than that for the car as driver mode. In other words, the computation time of the FEATHERS will be reduced more when the car as driver mode is under consideration. Moreover, for both the modes, especially the public transport mode, the municipalities lying on the border of Flanders generally need larger radius to reach the given accuracy level compared to those located in the relatively central position of Flanders. For instance, the city of Bruges—one of the leading seaside resorts in Belgium—requires the longest radius, no matter which transport mode is considered. In addition, the capital city Brussels, serving as a traffic hub of the region, also needs a larger study area for both the modes compared to its neighboring municipalities, since a great number of trips are happening every day between this municipality and all the others.

**4.2. Extreme Cases in Flanders Based on the Identified Minimum Study Area.** To verify the results we obtained, the validation procedure shown in Figure 3 is conducted by considering the four extreme cases, that is, the municipality with the longest study area radius and the one with the shortest radius for both the car as driver mode and the public transport mode (see Figures 8 and 9).

By running the FEATHERS again based on the identified study area for each extreme case, we can predict the travel pattern of each individual, based on which the origin and destination matrix in this territory can be derived. Then, by calculating the new travel demand of both departing mode and arriving mode within this study area and further comparing it with the one based on whole Flanders, we can examine the achieved accuracy level of the model. More importantly, by recording the computation time for each extreme case, the degree of the time reduction can be estimated. The results are shown in Table 1.

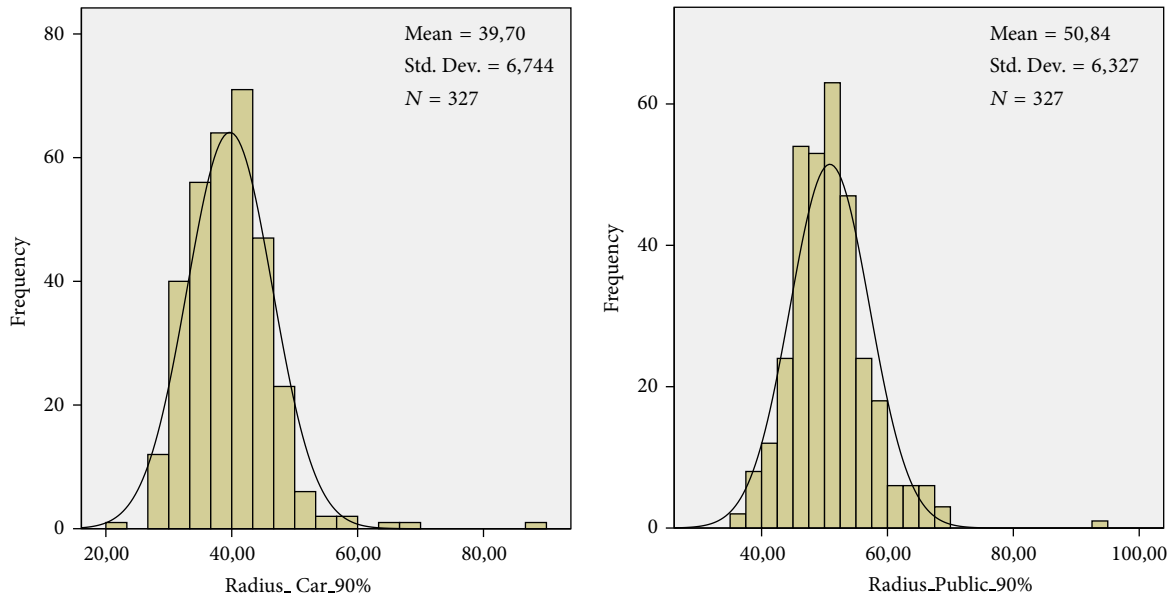


FIGURE 5: The distribution of the minimum radius needed for each of the 327 municipalities in Flanders with respect to car as driver mode and public transport mode.

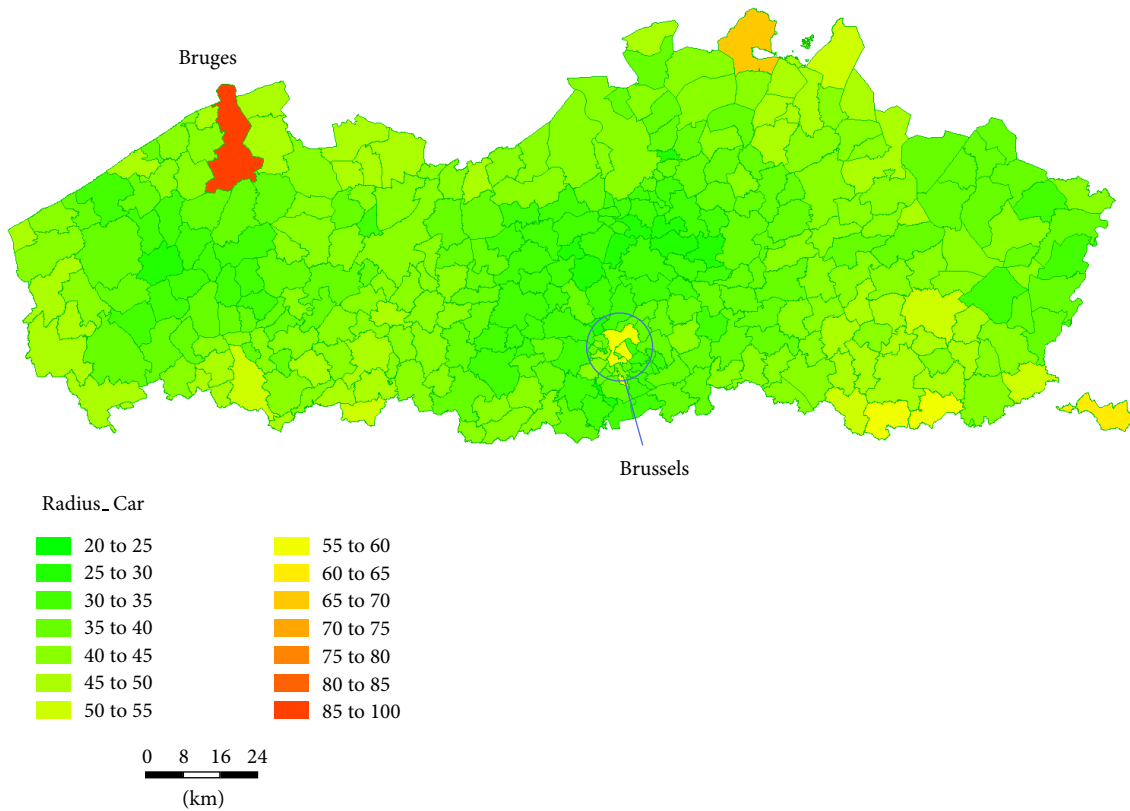


FIGURE 6: The color theme of the minimum radius needed for 327 municipalities of Flanders with respect to car as driver mode at 90% accuracy level.

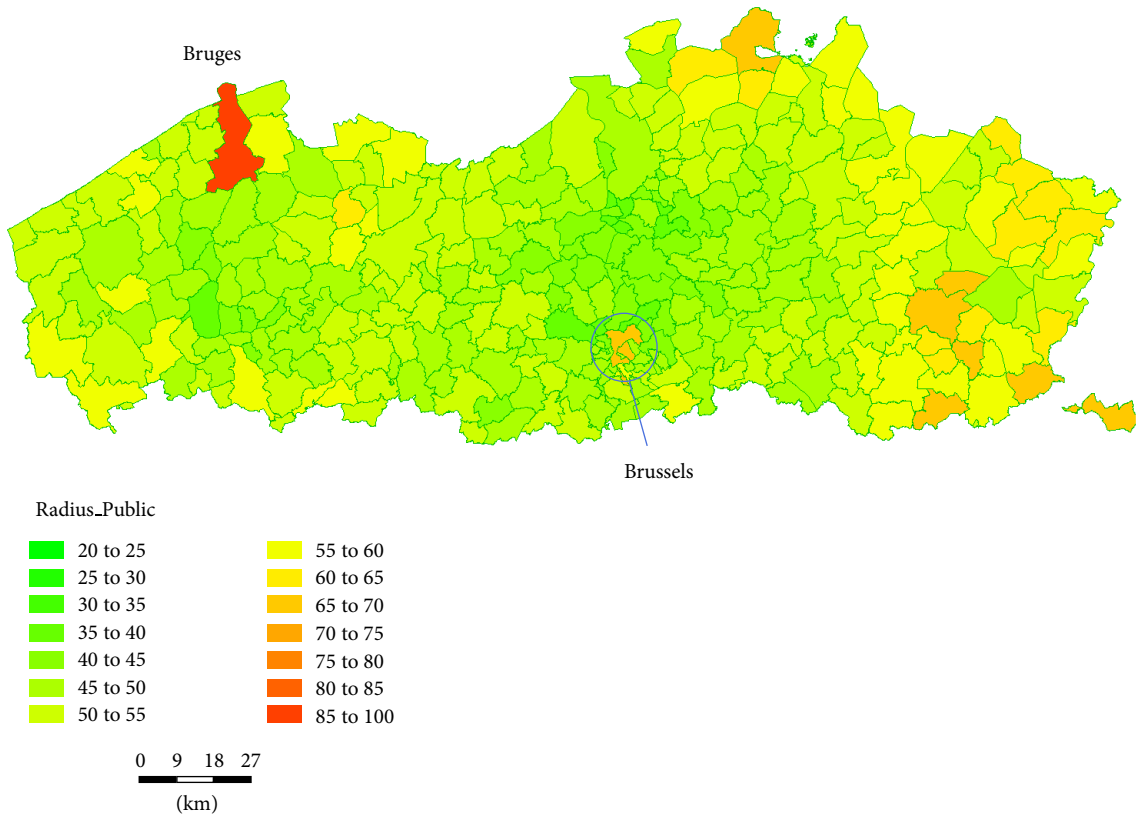


FIGURE 7: The color theme of the minimum radius needed for 327 municipalities of Flanders with respect to public transport mode at 90% accuracy level.

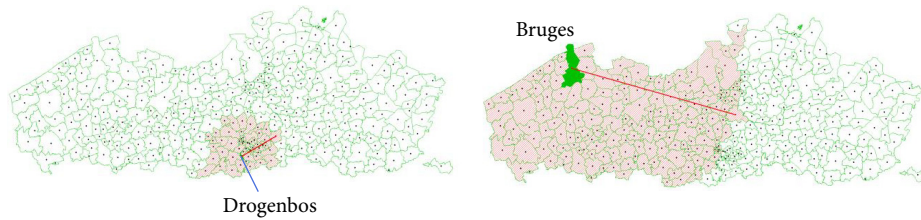


FIGURE 8: The municipality with the shortest and the longest required radius of study area with respect to car as driver mode at 90% accuracy level.

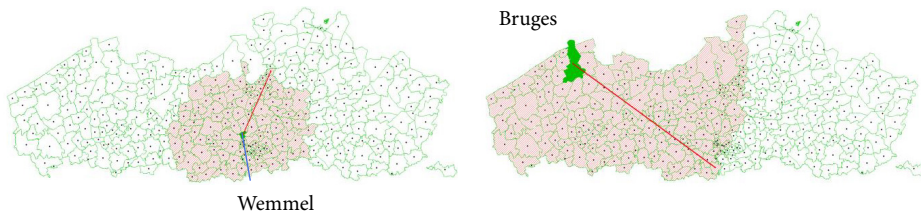


FIGURE 9: The municipality with the shortest and the longest required radius of study area with respect to public transport mode at 90% accuracy level.

TABLE 1: The validation results for the extreme cases with respect to car as driver mode and public transport mode.

Travel mode	Municipality	Radius	Accuracy level	Running time (hour)	Time saving
Car as driver	Drogenbos	Shortest: 21.2 km	86.9%	4	75%
	Bruges	Longest: 89.82 km	87.1%	9	43.8%
Public transport	Wemmel	Shortest: 36.17 km	86.3%	7.5	53.1%
	Bruges	Longest: 94.24 km	85.4%	10.5	34.4%

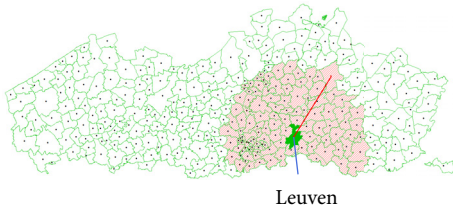


FIGURE 10: The study area for the case study of Leuven.

After performing the travel demand forecasting based on the rebuilt study area, we can see that all these four cases show a very high accuracy rate (all above 85%). Moreover, by restraining the size of the study area, the computation time of the model is reduced dramatically, from 16 hours in the original case to the best 4 hours (i.e., 75% time saving). We can therefore conclude that running the activity-based model inside FEATHERS using the rebuilt study area for each municipality of Flanders will improve the model's operational efficiency significantly.

**4.3. A Case Study Applying the Identified Minimum Study Area for Leuven.** Having successfully identified and validated the minimum size of the study area for each of the 327 municipalities in Flanders, Belgium, we now apply the results to a practical project which investigates the potential impact of light rail initiatives on travel demand at a local network in Flanders. In doing so, the city of Leuven is selected as a case study to perform prediction of the travel demand. The city owns quite large transport potential and is yet reasonably compact in size. Nevertheless, the city has no urban or regional light rail system so far.

Based on the identified study area for the city of Leuven, which is shown in Figure 10, the analysis is conducted by performing two scenarios in FEATHERS. Initially there is a null scenario that is limited to the situation where no light rail network is included. The public transport network contains only train lines (e.g., NMBS) and bus lines (e.g., De Lijn). In the second scenario, the light rail network is integrated with the current public transport network, which is therefore called the light rail scenario.

After running FEATHERS for these two scenarios, based on the same study area shown in Figure 10, the results are compared, and we find that the addition of the light rail network has a relatively positive impact on the public transport related trips in Leuven. The share of the public transport related trips increases by approximately 7% compared to the

null scenario. However, there is no significant change for other transport modes, such as car as driver (-0.22%) and car as passenger (-1.26%), and slow mode such as vulnerable road users (-0.20%). This result is in line with other international researches (e.g., [15]), indicating that apart from a reasonable increase in the public transport related trips, the implementation of a single light rail system has only limited effects on the overall modal split. Such consistent findings, from another viewpoint, verify the rationality of the selected study area. More importantly, by considering only the limited study area rather than whole Flanders, it turns out that the computation time is saved considerably for such an analysis. For a single model run, less than 6 hours is needed, which is only 38% of the computation time when whole Flanders is under consideration.

## 5. Conclusions and Future Research

The requirement of large computation time is currently one of the most important practical issues of applying activity-based models for travel demand forecasting. In this study, we investigated the possibility of restraining the size of the study area in order to reduce the computation time, as in many cases, only a small territory rather than the whole region is the focus of a specific study. By introducing an accuracy level of the model, which is defined as the difference between the occurrence of both the departing and the arriving trips derived based on the study area and that based on the whole region, we proposed in this research an iteration approach to determine the minimum size of the study area needed for a target territory, and both the calculation procedure and the validation procedure were designed. In the application, all of the 327 municipalities of Flanders, Belgium, were studied using the FEATHERS, an activity-based microsimulation modeling framework, and the minimum size of the study area needed for each of these municipalities was computed with regard to the car as driver mode and the public transport mode, respectively, given the accuracy level of 90%. The results indicated that the municipalities lying on the border of Flanders (e.g., the city of Bruges) as well as some traffic hub cities (e.g., the capital city Brussels) generally need longer radius to reach the given accuracy level. Meanwhile, for most of the municipalities of Flanders, the required minimum radius (or study area) is relatively larger for the public transport mode than that for the car as driver mode.

To verify the results we obtained, a validation analysis was carried out by running the FEATHERS based on the identified study area for four extreme cases, that is, the



municipality with the longest study area radius and the one with the shortest radius for both the car as driver mode and the public transport mode. It turned out that, within the identified minimum size of each study area, the computation time of FEATHERS was reduced considerably (up to 75%), while the model still reached a very high accuracy rate. Moreover, a case study was also performed which investigated the potential impact of light rail initiatives on travel demand at a local network in Flanders, using the results derived for the city of Leuven. The findings confirmed that when only a particular territory is needed for consideration in a specific study, it is possible to rebuild a relatively small study area for investigation. Running the model in such a restrained study area will improve the model's operational efficiency significantly. Consequently, the results obtained in this paper can be consulted as a reference for those who plan to use the FEATHERS framework, while for the other activity-based models, the methodology proposed in this paper with respect to the calculation of minimum size of study area can also be repeated.

In the future, more aspects need to be investigated. First, other accuracy levels can be considered, and the best tradeoff between the accuracy rate and the computation time can be discussed. Moreover, apart from analyzing the results based on different transport modes, other valuable travel indices could be taken into account as well, such as activity types. In addition, exploration on detailed reasons behind the different size of the study area needed for each target territory is also worthwhile, which will in turn validate this modeling framework and facilitate its further development and dissemination.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## References

- [1] T. A. Arentze, F. Hofman, H. van Mourik, and H. J. P. Timmermans, "ALBATROSS: multiagent, rule-based model of activity pattern decisions," *Transportation Research Record*, no. 1706, pp. 136–144, 2000.
- [2] J. Veldhuisen, H. Timmermans, and L. Kapoen, "RAMBLAS: a regional planning model based on the microsimulation of daily activity travel patterns," *Environment and Planning A*, vol. 32, no. 3, pp. 427–443, 2000.
- [3] C. R. Bhat, J. Y. Guo, S. Srinivasan, and A. Sivakumar, "A comprehensive microsimulator for daily activity-travel patterns," in *Proceedings of the Conference on Progress in Activity-Based Models*, Maastricht, The Netherlands, 2004, CD-ROM.
- [4] R. M. Pendyala, R. Kitamura, and A. Kikuchi, "FAMOS: the Florida activity mobility simulator," in *Proceedings of the Conference on Progress in Activity-Based Models*, Maastricht, The Netherlands, 2004.
- [5] P. M. Jones, F. S. Koppelman, and J. P. Orfueil, "Activity analysis: state-of-the-art and future directions," in *New Developments in Dynamic and Activity-based Approaches to Travel Analysis*, P. M. Jones, Ed., pp. 34–55, Gower Publishing, Aldershot, UK, 1990.
- [6] M. G. McNally, "The activity-based approach," in *Handbook of Transport Modelling*, D. A. Hensher and K. J. Button, Eds., pp. 53–70, Elsevier Science, Oxford, UK, 2007.
- [7] M. Malayath and A. Verma, "Activity based travel demand models as a tool for evaluating sustainable transportation policies," *Research in Transportation Economics*, vol. 38, no. 1, pp. 45–66, 2013.
- [8] T. Bellemans, B. Kochan, D. Janssens, G. Wets, T. Arentze, and H. Timmermans, "Implementation framework and development trajectory of FEATHERS activity-based simulation platform," *Transportation Research Record*, no. 2175, pp. 111–119, 2010.
- [9] Q. Bao, B. Kochan, T. Bellemans, D. Janssens, and G. Wets, "Investigating micro-simulation error in activity-based travel demand forecasting using confidence intervals," in *Proceedings of the 92nd Annual Meeting of the Transportation Research Board*, CD-ROM, Washington, DC, USA, 2013.
- [10] Q. Bao, Y. Shen, B. Kochan, T. Bellemans, D. Janssens, and G. Wets, "Research on restrained study areas for effective activity-based travel demand forecasting," in *Proceedings of the 14th COTA International Conference of Transportation Professionals*, pp. 3267–3279, Changsha, China, July 2014.
- [11] T. A. Arentze and H. J. P. Timmermans, "A learning-based transportation oriented simulation system," *Transportation Research Part B: Methodological*, vol. 38, no. 7, pp. 613–633, 2004.
- [12] B. Kochan, T. Bellemans, D. Janssens, and G. Wets, "Assessing the impact of fuel cost on traffic demand in Flanders using activity-based models," in *Proceedings of the Travel Demand Management Symposium*, Vienna, Austria, 2008.
- [13] D. Kusumastuti, E. Hannes, D. Janssens, G. Wets, and B. G. C. Dellaert, "Scrutinizing individuals' leisure-shopping travel decisions to appraise activity-based models of travel demand," *Transportation*, vol. 37, no. 4, pp. 647–661, 2010.
- [14] L. Knapen, B. Kochan, T. Bellemans, D. Janssens, and G. Wets, "Activity-based modeling to predict spatial and temporal power demand of electric vehicles in Flanders, Belgium," *Transportation Research Record*, no. 2287, pp. 146–154, 2012.
- [15] S. E. Volosin, S. Paul, R. M. Pendyala, B. Grady, and B. Gardner, "The application of a microsimulation model system to the analysis of a light rail corridor: insights from a transims deployment," in *Proceedings of the 91th Annual Meeting of the Transportation Research Board*, Washington, DC, USA, 2012.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

