Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 52 (2021) 684–691

Transportation Research Procedia

www.elsevier.com/locate/procedia

23rd EURO Working Group on Transportation Meeting, EWGT 2020, 16-18 September 2020, Paphos, Cyprus

Trip chaining impact on within-day mode choice dynamics: Evidences from a multi-day travel survey

Ariane Scheffer[a]*, Richard Connors[a,b], Francesco Viti[a]

[a]Université du Luxembourg, 2, avenue de l'Université, L-4365 Esch-sur-Alzette, Luxembourg
[b]Institute for Transport Studies, 34-40 University Rd, LS2 9JT, Leeds, United Kingdom

Abstract

Mode choice is influenced by a large variety of factors, as for example users' socio-economic attributes or level of service for the different alternatives. In order to understand better what leads to temporal and spatial variations of modal split, we propose in this paper an analysis of a multi-day *travel* survey, with a series of descriptive statistics as well as inferential analysis on the correlation between mode choice and tour-specific attributes at both spatial and temporal levels. This paper discusses the importance of considering tour-based mode choice not only because it brings consistency between successive mode choices but also allows the inclusion of relevant tours' characteristics such as activity types, distances, time of the day, and previous mode choices. A total of 5848 home-based tours done in 2008 are studied in the area of Ghent, Belgium. Identified patterns show the importance of modelling dynamic mode choice with trip chaining and time of the day. The modal share of *car drivers* differs of more than 40% between hours of the day and about 30% between different activities. Furthermore, the definition of activity spaces by principal mode choice and home-work locations introduces the calibration of probabilistic aggregate Gaussian fit to visited points.

Keywords: Mobility; Mode choice dynamics; Multimodality; Trip chaining; Travel survey; Activity space

1. Introduction

Recent developments in terms of mobility and transport systems create major challenges for the future planning of networks. While in the past travel demand was mostly captured by single modes users, the development of new systems such as Mobility as a Service, as well as policies aiming to reduce car use and ownership, enhance the multimodal behaviour of travellers. In order to more reliably assess future planning and management of networks we believe it is necessary to estimate and predict temporal and spatial distributions of travel demand in transportation systems. Activity-based models already ensure consistency of successive trips which are triggered for undertaking activities (Ben-Akiva and Bowman, 1998). In this work, we aim to evaluate to what extent we can observe this possible cause-effects relationship at an aggregate level and apply it for modelling accurately dynamic modal splits. Traditional models capture such relationship by using a step-by-step approach, like the "four-step model" (Cascetta, 2009), which can be inconvenient for observing realistic behaviour.

10.1016/j.trpro.2021.01.082

Mode choice is a complex decision that involves many determinants at various levels. Obviously, the level of service and trips characteristics impact the choice set and relative attractiveness of options, but the decision involves socio-demographic characteristics and personal preferences or habits as well (Tyrinopoulos and Antoniou, 2013). (De Witte et al., 2013) undertook a comprehensive review and highlighted that, among a large number of interacting parameters, departure time and even more so, trip chaining are too often ignored, e.g. trip chaining is considered meaningful for 80% of the cases but included in only 20% of the papers. Household travel surveys are commonly used for studying single modal choices. (Pucher and Renne, 2003) show regional variations but also significant variations in modal split by trip purpose, nonmotorized modes being less represented for going to work. In the case of Melbourne, (Currie and Delbosc, 2011) observe that complexity of tours is lower for car users and underline the need of adding a spatial perspective. On the contrary, (Hensher and Reyes, 2000) rely on time budget and value of time and model mode choice using different logit models in order to demonstrate to what extent trip chaining is a barrier to public transport use. In the same vein (Krygsman et al., 2007) show that public transport hinders the inclusion of secondary activities in work tours, notably because of mode (un)availability at the workplace. Furthermore, they use a co-evolutionary approach to conclude that the intermediary activity decision is made most of the times before the mode decision. Similarly, (Ye et al., 2007) use econometric modelling to explore the directionality of the relationship between mode choice and complexity of trip chaining patterns using micro census travel survey and show that trip chain complexity precedes mode choice. To understand better the impact of the mode availability constraint, we concentrate our analysis on the impact of using owned resources on trip chaining, and focus on the home-work (HW) tour as they include trips which have the strongest impact on transport networks. HW tour is used to describe tours (also named 'trip chains') starting at home and finishing at home which include the activity work.

The complexity of trip chaining and its impact on mode choice has also been linked to the concept of activity space (AS). The notion of time-space prisms, introduced by (Hägerstrand, 1970) illustrates for example the necessity of carrying out "[roles] within a given duration, at given times and places" and constrains them with "geometrical shapes in terms of location in space, areal extension, and duration in time" whose parameters depend on the available modes of transport. This concept, refined by (Lenntorp, 1976) has been widely used notably in the activity-based approach to travel demand modelling (Bowman and Ben-Akiva, 2001). The AS geographical aspect has been described in many ways and applied for various purposes. (Patterson and Farber, 2015) reference 66 applications of AS and potential path areas in diverse fields and highlight 4main methods to estimate AS: ellipses and circles; network-based approaches; kernel density approaches; minimum convex-hull polygons. Some authors compared different models based on a series of criteria, for example (Schönfelder and Axhausen, 2003) regret the rigid assumptions of the confidence ellipse that makes the AS's size too high. However, the versatility of the AS concept prevents reaching a consensus on a single set of criteria appropriate for all use cases. In order to represent better AS, new geometries are still recently proposed (Li and Tong, 2016) and notably for integration in MATSim (Rai et al., 2007). To reflect better AS, (Perchoux et al., 2014) combines different methods to create a set of indicators in order to qualify individual space-time patterns. This study also concludes that active modes are used in small AS centred around home and that larger AS correspond to the use of motorized modes. While most of explorations on AS focus on individuals, (Harding et al., 2013) assess the relationship between mode choice and AS at the city level. They selected convex hull for describing AS of sampled individuals and use 'area to compactness ratio' for comparing transit, car and active modes. While they mostly have expected outcomes, transit users are associated to compactness ratios closest to 1.

In this work we propose a new definition of AS and propose to link it with mode choice through common factors for groups of users. Through analysis of multiday survey data we quantify how the sequence of modes is impacted by the sequence of activities generating the travel need. The "Behaviour and Mobility within the week" (BMW) database used in this work is a multiday travel survey collected in Ghent in 2008 (Castaigne, 2009). The opportunity to use the BMW database is twofold, not only the size of the database is large with representative modal split but also the duration of the survey allows to define AS for each individual in a robust way. Few similar databases exist and for example the *"Mobidrive"* study (Axhausen et al., 2002) constitutes an excellent reference in such analysis, as six weeks-diaries for 317 people were recorded and it was used for analysing both activity scheduling (Cirillo and Axhausen, 2010) and mode choice of complex tours (Cirillo and Axhausen, 2002). However, the Mobidrive database contains a share of 86,1% unimodal tours, while the BMW database reaches 33% multimodal tours. This is an opportunity to study more in detail the correlations and interconnection between successive mode choices, time of the day and activity characteristics.

2. Methodology

In order to test the following hypotheses, we analyse a travel diary collected for 707 individuals in the city of Ghent (Belgium) in 2008 (the city was about 237.250 inhabitants at that time). A description of the database and analysis of the variability of daily activity-travel pattern is available in (Raux et al., 2016). The goal here is to observe emerging behaviours from these respondents in order to detect quantitative characteristics to be applied in future aggregated dynamic mode choice models. The tested hypotheses are the following:

- Modal split changes over the day and is statistically correlated with activity choice dynamics
- The mode chosen for a trip strongly depends on the mode chosen at an earlier time of the day
- Owned resources such as bike and car increase this dependency
- AS varies with the most frequently mode used and can be described by knowing home and work locations

2.1 Data

We focus on the HW trip chain of workers from a spatial and temporal dimension. 404 respondents were selected for analysis who described at least one trip going to work, resulting in an average of 4 working days per person and 2.8 trips within the HW tour. 3543 unique points addresses of trip destinations were translated to GPS coordinates. Each trip is described by its origin, destination, starting, ending and travel time, modes, activity at origin and destination and their duration. Distances are the direct ‘crow flies’ distances. Home/work locations were identified for each worker as the most visited locations they referred to as *home*/*work*. The data includes 7 different modes of transport and their combination: car (or motorbike) driver, car passenger, train, bus (and tram), bike, walk and other and 12 activity types could be recorded in this multiday survey.

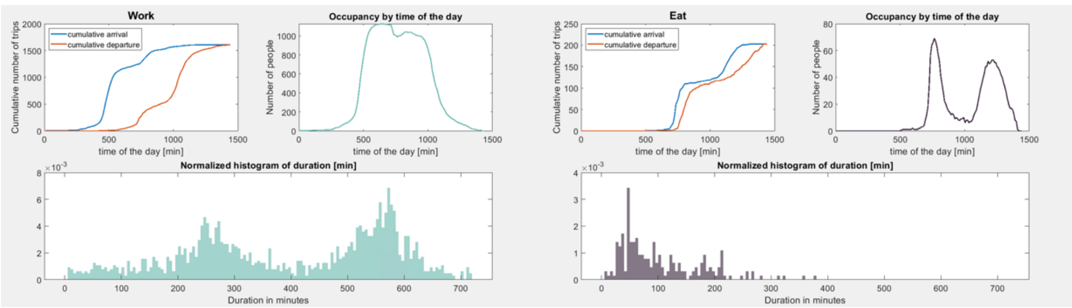

Figure 1 clustered activities characteristics: the example of (a) work and (b) eat

The focus not being on activity choice, a simplification has been made by clustering the activity types based on the following criteria: *occupancy*, *starting time* profile, *weekly frequency* (distribution from one to seven days in which the activity is done) and distribution of the *duration*. These four criteria have been normalized to one and grouped using k-means clustering into five groups of activities (home, work, daily tasks, personal business and eating). The *occupancy* is the difference between cumulative people starting and people ending the activity. This indicator represents, at each time of day, the potential number of people who may start a trip at later time, in order to start a subsequent activity. The profiles of different activities suggest that the negative correlations and offset between these curves translate activity chaining. Figure (1), shows the clustering criteria for an example of two activities.

K-means clustering was also applied to the 404 workers, to group them according to their mode use and their home and work locations. Because of the limited number of respondents belonging to each sub-group, geographical and behavioral attributes resulted in two independent groupings, applied for distinct parameters comparison. Users are firstly separated into groups based on the proportion of their trips done by each of the seven modes, resulting in the following five groups: car users, soft mode users, public transport users, train users and multimodal users (Figure 2a). Secondly, users were first classified by municipality. For those living and working in Ghent, home and work GPS coordinates and distance between the two points has been used. Using the Calinski-Harabasz criterion, suggested 5 groups. An example of resulting clusters is shown on Figure 2b.

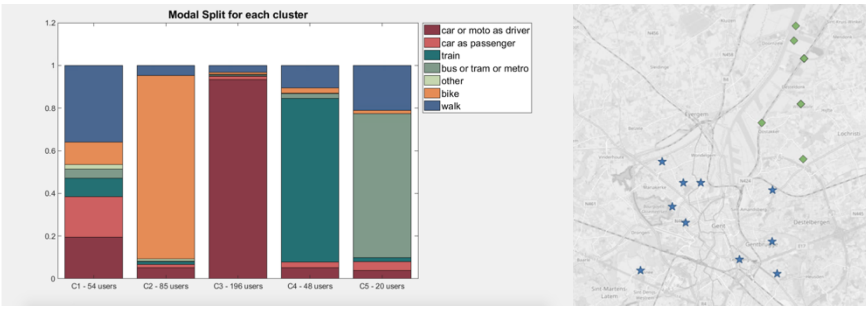

Figure 2 (a) modal split of 5 mode-based clusters (b) example of a Ghent HW locations-based cluster

2.2 Temporal analysis

Intuitively, different modes have different typical usage patterns, in particular based on the land-use and supply characteristics. In order to observe how they combine and how usage differs by time of day, we calculated the usage profiles for those modes and compared these patterns. The results are matched to individual observations through the transition probabilities. The transition matrix was calculated for all successive trips, for only commuting trips and finally for only intermediate trips, in order to estimate to what extent sets of modes complement one another and which are the most binding for round trips. To include all sizes of tours in the transition calculation, a notional mode corresponding to ending the trip chain was added. Additionally, we show that mode choice varies by trip purpose and secondly that the full activity sequence of workers impacts choices on the set of modes used. The way activities have been clustered distinguishes naturally systematic and non-systematic activities as can be seen in the weekly frequency. We believe that this difference impacts the travel time budget, due notably to the location that is more or less free to choose and leads to additional travel time for non-systematic activities higher than for systematic activities. This can explain, based on geographical dispersion, the relationship between the previously highlighted factors. In order to quantify the spatial aspect of mode choice dynamics, the AS of those workers has been analysed.

2.3 Spatial analysis

To show the link between temporal and spatial dynamics, the dispersion of each user's AS has been estimated using the method of 95% confidence ellipse. Such ellipses have few parameters to estimate and are commonly used for estimating AS. The 95% confidence ellipse ensures that number of visited points covered by the surface is high enough and so that it describes better the AS. The centre of the ellipse is the centre of mass of all visited points within the HW tour for the complete study period, weighted by the number of visits. The ellipse's axes are the square root of the eigenvalues of the covariance matrix, and the ellipse is then scaled to get the desired confidence interval, based on the number of degrees of freedom and the Chi-Square distribution. The resulting ellipses are compared based on: orientation angle, minor axis, major axis, aspect ratio and centre.

The hypothesis is that the ellipse is strongly defined by the home and work location, considered as *anchor points* for the user. This can be observed through the position of the centre of the ellipse and its orientation.

Because the multiday travel survey contains on average four working days for each worker, we first applied an outlier detection method, such that non-recurrent behaviours do not unduly impact the estimation of AS. This detection is based on the Orthogonalized Gnanadesikan-Kettenring robust estimator (Maronna and Zamar, 2002) using the Mahalanobis distance: $d_{(\mu,\Sigma)}(x_i)^2 = (x_i - \mu)'\Sigma^{-1}(x_i - \mu)$ with respect to location μ and covariance Σ. Detected outliers are not always deleted. Firstly, home and work locations are never discarded. Secondly, points closer than a given threshold to home/work are kept. If the number of visits to an outlier point is two or more, it is kept. Finally, for each user we try to retain enough points to compute the AS. Finally, 181 points have been removed: 26% of the *personal business* locations, 16% of *daily tasks* locations, 12,5% of *eat* locations and 8% of *work* locations.

Properties of individual ellipses can reveal emerging trends for groups of individuals. However, the application to aggregate location choice models is limited with such strongly constrained geometrical units. In order to apply the observations to groups of users and to use soft constraints, a gaussian fitting has been chosen to make ellipses probabilistic. The multivariate gaussian describes plausible visited areas given home and work regions. It is again parametrized by the centre of mass of the visited points and the model is estimated by the maximum likelihood, using the expectation-maximization algorithm.

3. Results

In this section, a selection of the most significant results is presented in order to support the validation of hypotheses noted above. The section is separated into temporal and spatial analysis. Firstly, indicative results about the correlation between these two aspects are described in order support the link between them (Figure 3).

First of all, the travel time inside the HW tour has been separated into two components: HW direct trip and extra travel time. For each of the five groups of users (according to their mode choice), the results indicate that there exists an upper bound, which differs with respect to the chosen mode. Public transport users reserve a large part to the HW travel time itself, in opposition to bike and car users who have an upper bound of respectively 80 and 100 minutes.

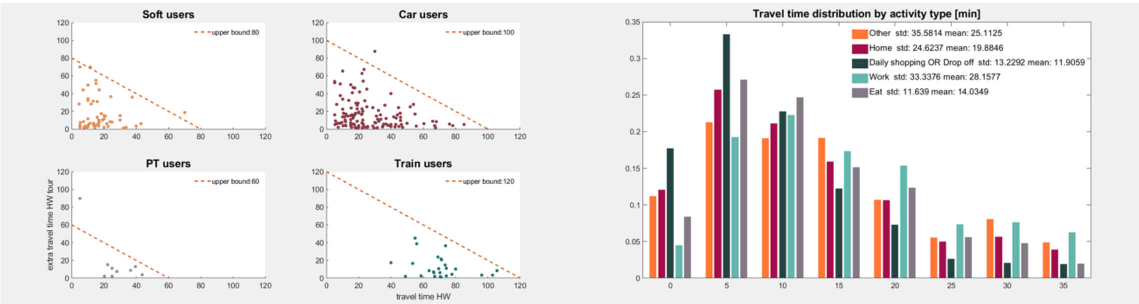

Figure 3 (a) HW travel time vs. additional travel time for four mode-based clusters (b) Travel time distribution by activity type

Results indicate also that the travel time is also following different distributions with respect to trip purpose. For eating or daily tasks, the travel time is on average less than fifteen minutes while the average travel time for work-related trips approaches half an hour.

3.1 Emerging behaviors at temporal levels

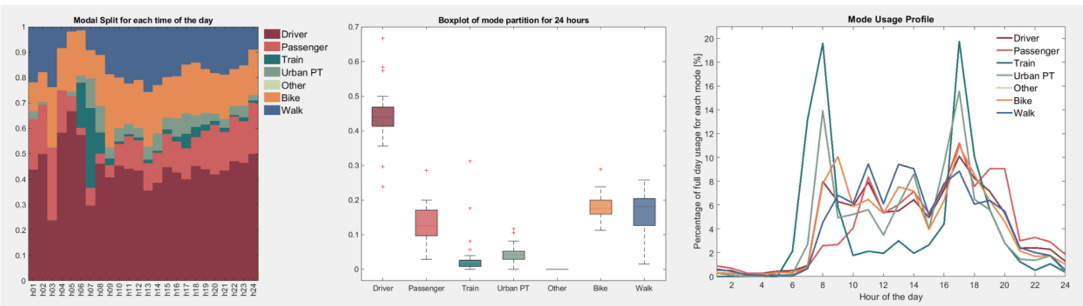

Figure 4 (a) Modal split by time of the day (b) Boxplot of distribution of modal share (c) Usage profile of different modes

Modal split significantly varies along the day (Figure 4a,b) and this can be explained by different usage types of modes (Figure 4c). As an example, car use as driver varies between 23.8% and 66.7%. The profile of each mode shows that they are used at different times of day. Public transport is used for commuting, owned vehicles are more similar to the general demand profile. Walking seems to be used as complementing mode during off-peak hours.

The difference between theoretical modal split by activity is indeed particularly significant for the following combination: train to go to work and walk to eat (Figure 5a). In terms of mode choice, this can be observed through the sequence of trip modes within tours. Figure 5c shows the transition matrices for choosing a mode after having used any other mode. The highest probabilities are walking after using public transport and for the rest is to use the same mode for successive trips. When looking at the two commuting trips, the probability to choose the same mode is always higher than 50% and higher than 90% for owned resources (reaching 95% for car users). Figure 5b shows the example of car commuters. When car is chosen to leave home the probability to choose the car later on the trip chain is much higher than other modes. Furthermore, when another mode is introduced in the chain, the probability of returning home is lower than to ultimately go back to car before ending the chain.

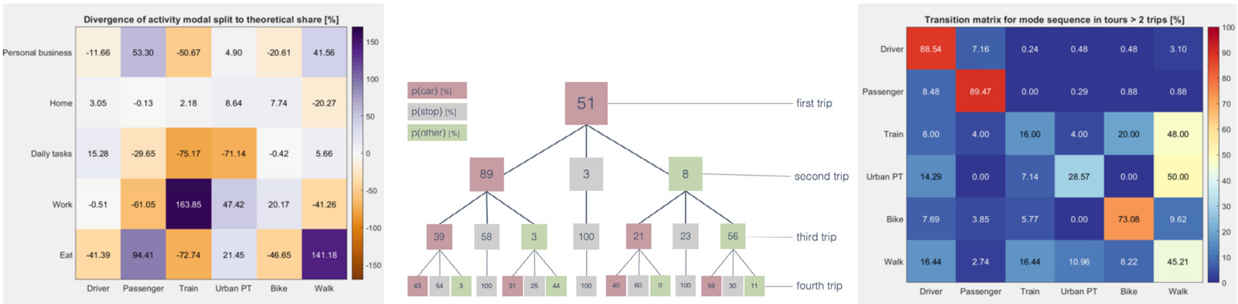

Figure 5 (a) Divergence of activity modal split to theoretical share (b) Car use probability tree (c) Transition matrix for mode sequence

This shows the impact of owned resources on the propensity to have a multimodal trip chain. When a person leaves home with a bike or car, (s)he will return home with this mode, and mostly use this mode for intermediate trips. However, there is not a big difference in the number of trips made with respect to that decision.

3.2 Emerging behaviors at the spatial level

AS were calculated only when there were enough unique points, i.e. 283 workers were used for ellipses parameter estimation. Figure 6a shows the orientation of the calculated AS in comparison with the orientation of the HW direct segment. The strong correlation of 0.8 supports the use of HW axis as an approximation of AS orientation. The correlation is even higher between the centre of the AS and the middle point of the HW segment.

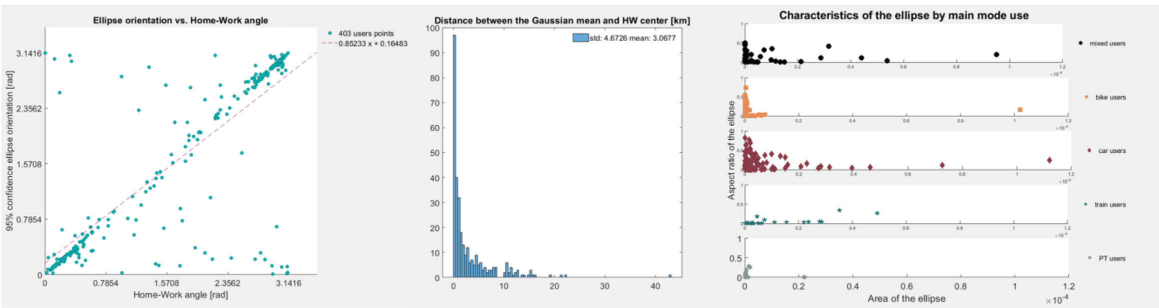

Figure 6 Characterization of ellipses (a) orientation vs HW axis (b) position of the centre vs HW centre (c) shape by mode

Figure 6b shows the distribution of the distance between these two points. The median is 1.1km, being higher for two types of users: the commuting distance is high (more than 100 km) or home and work are both located in peripheral areas. Figure 6c shows the area and aspect ratio of the ellipses for different user types. We clearly see that the aspect ratio is low for public transport (close to zero for all train users) and that the area is the smallest for urban public transport users. A high aspect ratio (eccentricity tending to zero and circular shape) reflects an AS less impacted by home and work locations with secondary activities potentially off the HW direct route (typically car and bike users). In opposition, if the AS has an eccentricity closer to 1, it means that the secondary activities are more likely to be located on the HW segment or close around those two focal points.

The clustered users reveal that the typical ellipses for the mode-specific groups have characteristic shapes: thin elongated ellipses for public transport users; larger for train and smaller for bus users. Ellipses are more rounded for private transport users: larger for car and smaller for bike users (Figure 7a). 30% of the travellers visited only two points in their HW tour, each day and resulted in the estimation of a segment; in order to relax the constraint of choosing an activity location strictly between the home and work location for these users and estimate AS even with low number of observations, we use the revealed regularities of similar users and proposed a probabilistic approach to the AS estimation. Figure 7b shows the output of aggregate probabilistic AS for the same group of users as seen on Figure 2b and for which only seven out of eleven users where subject to the single AS calculation.

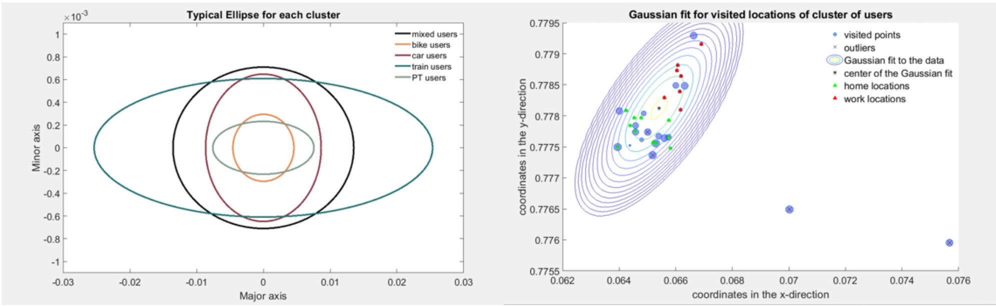

Figure 7 (a) Typical ellipse for 5 mode-based clusters (b) Gaussian fit for visited locations of a Ghent HW cluster of users

We can see that characteristics of the fitted gaussian after outlier detection and removal follows the same rules as what has been observed for individual workers. The centre is situated between the home and work areas and the contour lines of the distribution are oriented along this HW axis. Additionally, a single ellipse does not represent properly the option of 'crossing the line' and perform and activity outside the AS limits. That is why the estimation of distributions provides a more realistic perspective to be used in future research among other in order to estimate secondary activities locations.

4. Conclusion

The proposed work presented descriptive statistics and empirical analysis on mode choice in relationship with time of day and chains of activities as well as spatial distribution of visited locations in the HW tour. Starting hypotheses have been supported by a sample of Ghent population described in the BMW database and these empirical observations can be used as the basis for more accurate travel demand modelling. The modal split varies throughout the day and successive mode choices are strongly correlated to each other. This stands in particular for owned vehicles (bicycle or car), as there is a constraint of carrying the resource around. The usage profile of modes and the transition matrix show complementarity of modes, in particular walking which can be considered as a mainstay of many multimodal trip chains. Activities which are most frequent are usually located within a shorter distance from home or work, while longer distances are travelled for infrequent activities for which the destination cannot be substituted, like business trips or visits to family or friends. Parameters defining AS can be estimated by knowing the home and work locations of an individual which gives a good approximation of its center, orientation and one of the axes of the ellipse. Aspect ratio and area are governed instead by the commuting mode choice. For applying these observations to aggregated groups and using AS as a soft constraint for choice modelling, we fit, instead of a single ellipse, a bivariate normal distribution for groups of users. These findings can be used for estimating mode specific travel demand by time of day and regions. Future work will focus on exploiting those implications in dynamic demand modelling and include these revealed regularities both in mode choice and location choice models. Observed outcomes on the individual AS will be included in the calibration of potential areas for secondary activities locations using the gaussian fit. Furthermore, limitations have shown to be linked to the neglect of level-of-service and land use. They need to be included in the estimation of AS to make them more accurate and adaptable in order to potentially apply it to planning and new services implementation.

Acknowledgements

This research was funded by the EU-FEDER project MERLIN.

References

Axhausen, K., Zimmermann, A., Schönfelder, S., Rindsfüser, G., Haupt, T., 2002. Observing the rhythms of daily life: A six-week travel diary. Transportation 29, 95–124. https://doi.org/10.1023/A:1014247822322

Ben-Akiva, M.E., Bowman, J.L., 1998. Activity Based Travel Demand Model Systems, in: Marcotte, P., Nguyen, S. (Eds.), Equilibrium and Advanced Transportation Modelling. Springer US, Boston, MA, pp. 27–46. https://doi.org/10.1007/978-1-4615-5757-9_2

Bowman, J.L., Ben-Akiva, M.E., 2001. Activity-based disaggregate travel demand model system with activity schedules. Transportation Research Part A: Policy and Practice 35, 1–28. https://doi.org/10.1016/S0965-8564(99)00043-9

Cascetta, E., 2009. Transportation systems analysis: models and applications, 2nd ed. ed, Springer optimization and its applications. Springer, New York.

Castaigne, M., 2009. Behaviour and Mobility during the week "BMW."

Cirillo, C., Axhausen, K.W., 2010. Dynamic model of activity-type choice and scheduling. Transportation 37, 15–38. https://doi.org/10.1007/s11116-009-9218-8

Cirillo, C., Axhausen, K.W., 2002. MODE CHOICE OF COMPLEX TOURS, in: Publication of: Association for European Transport. Presented at the European Transport Conference 2002MVA, Limited; Association for European Transport.

Currie, G., Delbosc, A., 2011. Exploring the trip chaining behaviour of public transport users in Melbourne. Transport Policy 18, 204–210. https://doi.org/10.1016/j.tranpol.2010.08.003

De Witte, A., Hollevoet, J., Dobruszkes, F., Hubert, M., Macharis, C., 2013. Linking modal choice to motility: A comprehensive review. Transportation Research Part A: Policy and Practice 49, 329–341. https://doi.org/10.1016/j.tra.2013.01.009

Hägerstrand, T., 1970. What about people in Regional Science? Papers of the Regional Science Association 24, 6–21. https://doi.org/10.1007/BF01936872

Harding, C., Patterson, Z., Miranda-Moreno, L., 2013. Activity Space Geometry and Its Effect on Mode Choice | Semantic Scholar. Presented at the TRB.

Hensher, D.A., Reyes, A.J., 2000. Trip Chaining as a barrier to the propensity to use public transport. Transportation 27, 341–361. https://doi.org/10.1023/A:1005246916731

Krygsman, S., Arentze, T., Timmermans, H., 2007. Capturing tour mode and activity choice interdependencies: A co-evolutionary logit modelling approach. Transportation Research Part A: Policy and Practice 41, 913–933. https://doi.org/10.1016/j.tra.2006.03.006

Lenntorp, B., 1976. Paths in Space-time Environments: A Time-geographic Study of Movement Possibilities of Individuals. Royal University of Lund, Department of Geography.

Li, R., Tong, D., 2016. Constructing human activity spaces: A new approach incorporating complex urban activity-travel. Journal of Transport Geography 56, 23–35. https://doi.org/10.1016/j.jtrangeo.2016.08.013

Maronna, R.A., Zamar, R.H., 2002. Robust Estimates of Location and Dispersion for High-Dimensional Datasets. Technometrics 44, 307–317.

Patterson, Z., Farber, S., 2015. Potential Path Areas and Activity Spaces in Application: A Review. Transport Reviews 35, 679–700. https://doi.org/10.1080/01441647.2015.1042944

Perchoux, C., Kestens, Y., Thomas, F., Hulst, A.V., Thierry, B., Chaix, B., 2014. Assessing patterns of spatial behavior in health studies: Their socio-demographic determinants and associations with transportation modes (the RECORD Cohort Study). Social Science & Medicine 119, 64–73. https://doi.org/10.1016/j.socscimed.2014.07.026

Pucher, J., Renne, J.L., 2003. Socioeconomics of Urban Travel: Evidence from the 2001 NHTS [WWW Document]. URL /paper/Socioeconomics-of-Urban-Travel%3A-Evidence-from-the-Pucher-Renne/3cd62e2331d2d65ba8318aae53fee401f6233af7 (accessed 5.27.20).

Rai, R.K., Balmer, M., Rieser, M., Vaze, V.S., Schönfelder, S., Axhausen, K.W., 2007. Capturing Human Activity Spaces: New Geometries. Transportation Research Record 2021, 70–80. https://doi.org/10.3141/2021-09

Raux, C., Ma, T.-Y., Cornelis, E., 2016. Variability in daily activity-travel patterns: the case of a one-week travel diary. Eur. Transp. Res. Rev. 8, 26. https://doi.org/10.1007/s12544-016-0213-9

Schönfelder, S., Axhausen, K.W., 2003. Activity spaces: measures of social exclusion? Transport Policy 10, 273–286. https://doi.org/10.1016/j.tranpol.2003.07.002

Tyrinopoulos, Y., Antoniou, C., 2013. Factors affecting modal choice in urban mobility. European Transport Research Review 5, 27–39. https://doi.org/10.1007/s12544-012-0088-3

Ye, X., Pendyala, R.M., Gottardi, G., 2007. An exploration of the relationship between mode choice and complexity of trip chaining patterns. Transportation Research Part B: Methodological 41, 96–113. https://doi.org/10.1016/j.trb.2006.03.004