Research paper

Corresponding author:

Katarzyna Sila-Nowicka,

Urban Big Data Centre, 6-7 Lilybank Gardens, University of Glasgow, Glasgow, UK

email: Katarzyna.Sila-Nowicka@glasgow.ac.uk

Calibrating spatial interaction models from GPS tracking data: An example of retail behaviour.

Katarzyna Sila-Nowicka[1]

[1]Urban Big Data Centre, School of Political and Social Sciences, University of Glasgow, Lilybank Gardens 7, Glasgow, UK, G12 8RZ

Email: Katarzyna.sila-nowicka@glasgow.ac.uk

A. Stewart Fotheringham[2]

[2]School of Geographical Sciences and Urban Planning, Arizona State University

Coor Hall, 5th floor, 975 S. Myrtle Ave. Tempe, AZ 85287

Email: Stewart.Fotheringham@asu.edu

1. Acknowledgements

The authors would like to thank all the survey participants. Furthermore, we acknowledge the valuable assistance of Jan Vandrol and Taylor Oshan with the processing of the GPS logs and with coding.

2. Funding statement

This work was supported by the EU FP7 Marie Curie ITN GEOCROWD grant (FP7-PEOPLE-2010-ITN-264994).

Abstract:

Global Positioning System (GPS) technology has changed the world. We now depend on it for navigating vehicles, for route finding and we use it in our everyday lives to extract information about our locations and to track our movements. The latter use offers a potential alternative to more traditional sources of movement data through the construction of trip trajectories and, ultimately, the construction of origin-destination flow matrices. The advantage of being able to use GPS-derived movement data is that such data are potentially much richer than traditional sources of movement data both temporally and spatially. GPS-derived movement data potentially allow the calibration of spatial interaction models specific to very short time intervals, such as daily or even hourly, and for user-specified origins and destinations. Ultimately, it should be possible to calibrate continuously updated models in near real-time. However, the processing of GPS data into trajectories and then origin-destination flow matrices is not straightforward and is not well understood. This paper describes the process of transferring GPS tracking data into matrices that can be used to calibrate spatial interaction models. An example is given using retail behaviour in two towns in Scotland with an origin-constrained spatial interaction model calibrated for each day of the week and under different weather conditions (normal, rainy, windy). Although the study is small in terms of individuals and spatial context, it serves to demonstrate a future for spatial interaction modelling free from the tyranny of temporally static and spatially predefined data sets.

Keywords:

GPS movement data, spatial interaction modelling, retail analysis

3. Introduction

The measurement and recording of human mobility is vital for understanding many important elements of society such as the demand for transportation services, the optimal location of facilities and the redistribution of population. Until recently, exploring human mobility in detail was challenging because personal trip data collection methods consisted of expensive and time-consuming paper-and-pencil interviews, computer-assisted telephone interviews and computer-assisted self-interviews (Wolf, Guensler, & Bachman, 2001). As well as being expensive to collect, these data are also typically limited in terms of their spatial and temporal resolution. The development of sensors such as GPS trackers that capture movement data in real-time and at detailed spatial and temporal scales has transformed our ability to collect mobility data (M.-P. Kwan & Neutens, 2014). However, even though GPS trackers record an individual's location and movement very accurately, they do not record essential characteristics of travel behaviour such as travel mode or trip purpose (Shen & Stopher, 2014). To overcome this problem, various attempts have been made to infer individual behavioural from GPS trajectories (*inter alia* Wolf et al. 2001; Patterson et al. 2003;Di Lorenzo et al. 2012; Gong et al. 2012; Sila-Nowicka et al. 2016; and Xiao et al. 2016). However, the overwhelming majority of studies using GPS data simply report visual descriptions of movement patterns rather than exploring the deeper understanding of what factors might have been responsible for these patterns.

Over the past decade there have been many attempts to adapt other technologies such as Radio Frequency Identification (RFID), WiFi, Bluetooth, smart cards or GSM to study aspects of human mobility. The tracking applications of RFID technology related to mobility have been reported in transportation and logistics (*inter* alia Eckfeldt & Bruce 2005; Zuo et al. 2010 and Zacharewicz et al. 2011) and in the tracking of patients in hospitals (Cangialosi, Monaly Jr., & Yang, 2007). Most research relating to the use of location information obtained via Bluetooth and WiFi technologies within mobile phones focusses on predicting movement patterns by asking a few fundamental questions about future location, time spent there and social interactions during time spent in location (*inter alia* Anastasi & Borgia 2004: Vu et al. 2011). Yet another modern method of capturing travel behaviour is via smart cards which are used in most of the world's major cities to automatically pay for travel fares. Data collected from these cards provide an opportunity to study human mobility patterns, as well as the efficiency and other aspects of transportation services, but are necessarily limited to the on and off points of the public transportation service and do not necessarily capture the real origins and destinations of the movement (*inter alia* Long & Thill 2015; Zhong et al. 2015; Tonnelier et al. 2016). In recent years increasing effort has been put into the analysis of mobile phone data (recording movements between GSM towers) showing the potential of these data in identifying fine-grained variations in urban flows over time, for estimating movements in urban spaces, and identifying potential social interactions and significant places for individuals (*inter alia* Ratti et al. 2006; Kwan 2007; Ratti et al. 2010; Calabrese et al. 2013; Calabrese et al. 2014; Ahas et al. 2015; Behadili 2016).

Currently, however, these advances in movement data collection technologies are well ahead of the existing methods for extracting meaningful information from such data (Laube, Dennis, Forer, & Walker, 2007; J. Long & Nelson, 2012). Furthermore, there have been very few studies that have tried to analyse decision-making processes related to mobility using data from emerging technologies. There is a need therefore to determine if new forms of movement data can be translated into new insights about mobility behaviour. We do this through an examination of the calibration of spatial interaction with GPS data.

To crystalize the rationale for this paper, we turn to a quote by Golledge & Stimson (1997, p.5) about an earlier era of geography as the quantitative revolution was dawning: “geographers became experts on describing `what' was there and are now seeking to explain `why' or `how' things were there". This sentiment is pertinent today with a new wave of descriptive analysis breaking over the geographical shores propagated by emerging technologies that generate huge quantities of spatial data. As yet, these data have yet to yield much insight with the bulk of research limited to a description of patterns rather than an analysis of human behaviour. Our goal therefore is to move beyond description and to present a demonstration of the potential inherent in GPS-derived data for analysing and understanding human behaviour. We do this by focussing on two specific questions:

(i) What has to be done in order to transform GPS tracking data into origin-destination matrices that can be used for the calibration of spatial interaction models?

(ii) Is it possible to draw meaningful insights into mobility decision-making from the calibration of spatial interaction models with GPS-derived data? In particular we will investigate the possibility of calibrating spatial interaction models for different days of the week and for different weather conditions.

In order to answer the questions posed above, two preparatory steps need to be undertaken and which have been described elsewhere (Authors, 2016). These involve the initial collection of the GPS traces and the classification of these traces into semantically enriched trajectories. Here we concentrate on the transformation of the GPS movement data into origin-destination matrices and on the use of these matrices to calibrate interaction models of shopping behaviour for different days of the week and under different weather conditions.

4. Spatial Interaction Models in Retailing

Spatial interaction refers to movement or communication over space that results from a decision-making process (Batten & Boyce, 1987; Fotheringham & O'Kelly, 1989; A. Wilson, 1967, 1970). It can be defined in terms of the movement of people, goods or information and it covers behaviours such as migration, commuting, shopping, recreation, trips for educational purposes, airline passenger movement, the choice of health care services and patterns of telephone calls (more examples of spatial interactions are given by Haynes & Fotheringham (1984)). These behaviours are characterised by a common and fundamental principal whereby individuals trade off the benefit of interaction with the cost of overcoming the distance (separation) to a destination (Fischer, 2002). This trade-off is at the heart of all spatial interaction models. For instance, the most common form of spatial interaction model employed in retail analyses is often referred to as an origin-constrained spatial interaction model and has the form:

$$T_{ij} = \frac{O_i\, w_j^{\alpha} d_{ij}^{\beta}}{\sum_j \, w_j^{\alpha} d_{ij}^{\beta}}$$

or, equivalently,

$$T_{ij} = A_i O_i \, w_j^{\alpha} d_{ij}^{\beta}$$

where

$$A_i = \frac{1}{\sum_j \, w_j^{\alpha} d_{ij}^{\beta}}$$

and where T_{ij} represents the number of retail trips from origin i to outlet j, O_i is the total number of trips originating at *i*, A_i is a balancing factor which ensures that the total number of predicted trips from *i* is equal to O_i, w_j represents the attractiveness of outlet j which can be measured by a number of variables but is often measured by size which reflects the range of goods available and sometimes price levels, d_{ij} is the network distance between *i* and *j*, β indicates the sensitivity of the number of trips between *i* and *j* to the distance between them, and α is a parameter reflecting consumers' sensitivity to variations in store sizes. Examples of the use of this model to understand consumer spatial choice include Lakshmanan & Hansen, 1965; Fotheringham & Trew 1993; Clarke et al. 1998; Bhat et al. 2004; Rodriguez & Joo 2004;

Preston & McLafferty 2016; de Vries et al. 2009; Dolega et al. 2016; Nakaya et al. 2007; and Merino & Ramirez-Nafarrate 2015.

Common to most applications of spatial interaction models in retailing, however, is the dearth of appropriate trip data with which to calibrate the models. Quite often the data are just not available and so models cannot be calibrated. The parameters in the model are then guessed at or borrowed from other studies to allow the models to be used to estimate flows from residential areas to a set of stores under varying conditions to examine questions such as *"Where is the optimal location for a new store?* or "*If I locate a store here, how much custom will it cannibalise from my exiting stores?"* Where flow data are available it is then possible to estimate the model's parameters which will yield more accurate predictions of flows and will also yield behavioural information on consumer spatial choice. Estimates of α describe consumers' utility from selecting larger stores with a greater variety of products and possibly lower prices while estimates of β reflect consumers' sensitivities to distance as a deterrence in selecting a store. For example if β were zero then consumers would not be constrained by distance at all in their selection of a store to patronise and increasingly negative values of β reflect greater deterrence in overcoming longer distances.

Even when data are available on consumers' shopping patterns and when spatial interaction models can be calibrated, the models typically yield limited information on consumer behaviour. This is because the data on individuals' movements over space are traditionally based on travel diaries or questionnaires which, besides being expensive to conduct, provide only a very limited snapshot of people's behaviour. They typically only represent behaviour over a broad period of time and often only for prescribed destination sets which are defined for the purposes of the survey. Recent years have brought new perspectives to spatial interaction modelling showing that using data from loyalty cards from major shopping retailers can improve forecasts concerning store patronage and store revenues

(Newing, Clarke, & Clarke, 2015). Nevertheless, until very recently it has not been possible to provide, for example, information on consumer behaviour at different times of the day or on different days of the week or during different weather conditions. Do consumers make different choices and exhibit different spatial behaviour, for example, during the week compared to the weekend, during the morning compared to the afternoon, or on days when it is raining compared to when it is dry? Traditional consumer surveys very rarely yield the data necessary to answer these questions.

However, the recent technological advances in recording the locations of individuals through their phones or with dedicated GPS trackers has the potential to radically change the spatial interaction modelling landscape by allowing the calibration of models for fine time intervals and for multitudes of different types of movement. These new forms of movement data have already begun to be employed in retailing. For instance, Yue et al. (2012) use GPS trajectories of taxi flows to compute trading areas around shopping centres in China; Lovelace et al.(2016) present a comparison of estimating shopping flows from a major mobile phone service provider, a commercial consumer survey and geotagged Twitter messages. Most recently, Lloyd & Cheshire (2017) investigate the feasibility of using geo-tagged Twitter data to define catchment areas for retail centres in part of the United Kingdom However, to date, there has been very limited discussion of the use of GPS trajectory data to calibrate spatial interaction models to better understand the dynamics of consumer spatial behaviour. This paper fills this gap in the literature and heralds a new era of spatial interaction modelling by showing the potential to calibrate models with new forms of geocoded data which allow variations in behaviour to be modelled over very short time intervals allowing us to better understand the dynamics of consumer behaviour.

5. Preliminaries

In order to investigate the feasibility of using GPS movement data to calibrate spatial interaction models, a sample of 150 individuals in two towns in Fife, Scotland were asked to carry GPS tracking devices (I-Blue 747 ProS) for a period of seven consecutive days (further details about the data collection methodology, ethical approval and data processing can be found in Authors' previous publication (2016). This generated 2,863,410 raw GPS points, with each location record containing participant ID, latitude, longitude, elevation, date and time. The data collection took place over a 4-month period in 2013 (September - December) and the movement data of 91 individuals in Dunfermline and 59 individuals in Glenrothes were tracked.

To extract information from these GPS traces, we filtered, pre-processed, segmented, classified and contextually enriched the data using a framework for mobility patterns analysis (travel mode and activity places) from a combination of GPS movement data and contextual information. The data processing schema (Figure 1) was designed following the methodologies of Yan et al. (2013) and Spaccapietra (2009).

Figure 1 somewhere here

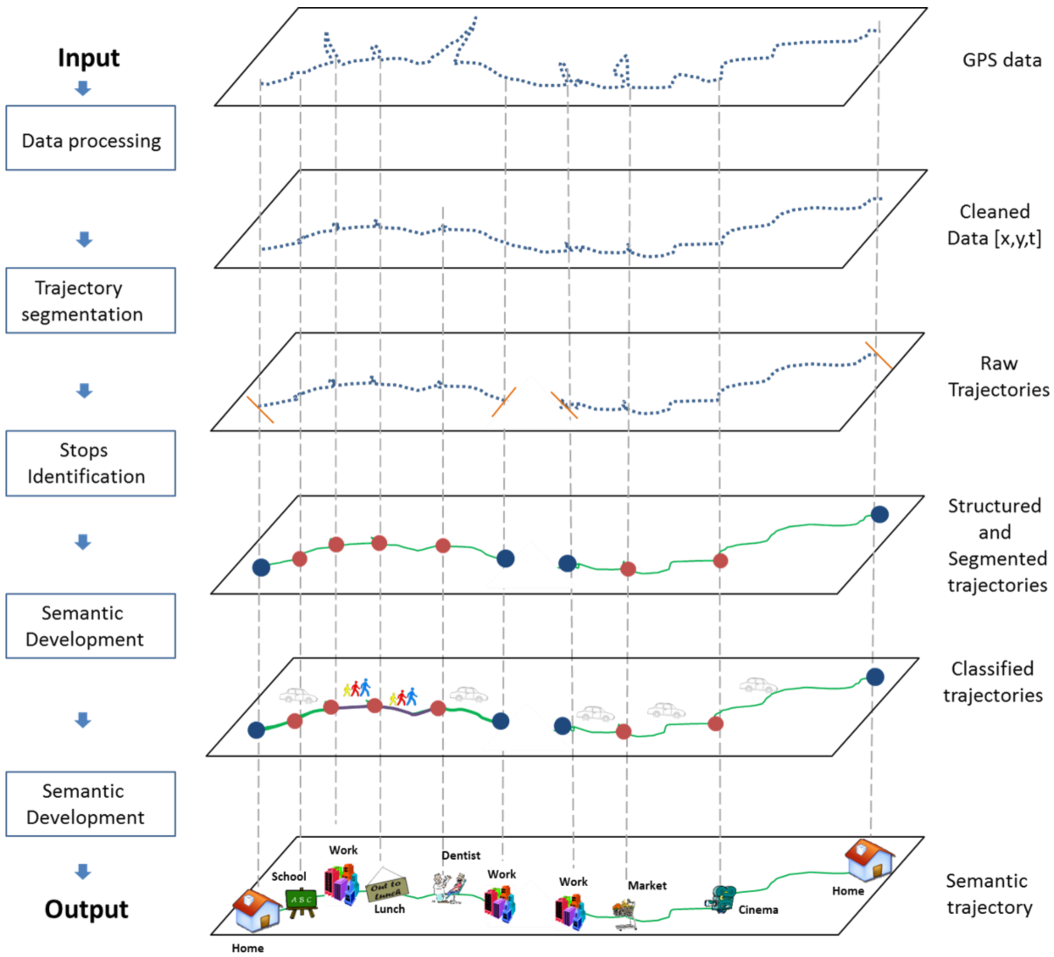


Figure 1. Visual flowchart for GPS data processing. The idea for visualisation is obtained from Yan et al. (2013) and Spaccapietra (2009) with steps of data processing modified in order to contextually enrich GPS movement data for this study.

The collected GPS movement data were first cleaned and filtered to minimise the number of erroneous records (those with low precision caused by the satellites' geometry).Then we segmented trajectories into homogeneous sub-trajectories using a procedure based on a new statistical measure implemented into a machine learning algorithm – a Spatio-Temporal Kernel Window developed by Authors (2016). Subsequently we applied a two-step feedforward neural network with a general backpropagation algorithm for segment classification; first to distinguish movement from non-movement segments and then to classify movement segments into specific travel modes (driving, walking, bus and train). The non-movement segments were classified based on their importance to a user into “home” and set of significant locations such as “work”, ”school”, “third place” and others which were compared to a Points/Places of Interest dataset (a combination of Ordnance Survey, OSM and self-created POI dataset) in order to contextually enrich them with functions such as: shopping, leisure, school/health or transport related.

From the semantically enriched trajectories we created individual trip chains for each participant which involved linking spatially and temporally interrelated trips (Zhao, Chua, & Zhao, 2012, p. 2). Each segment in the GPS dataset is labelled with either travel mode, possible activity or as an unidentified stop. By running a set of SQL queries, the travel chains can be retrieved. Using the trip chain structures, individual-thematic trips, such as commuting, shopping or leisure trips can be extracted. An example of the resulting database is given in Table 1.

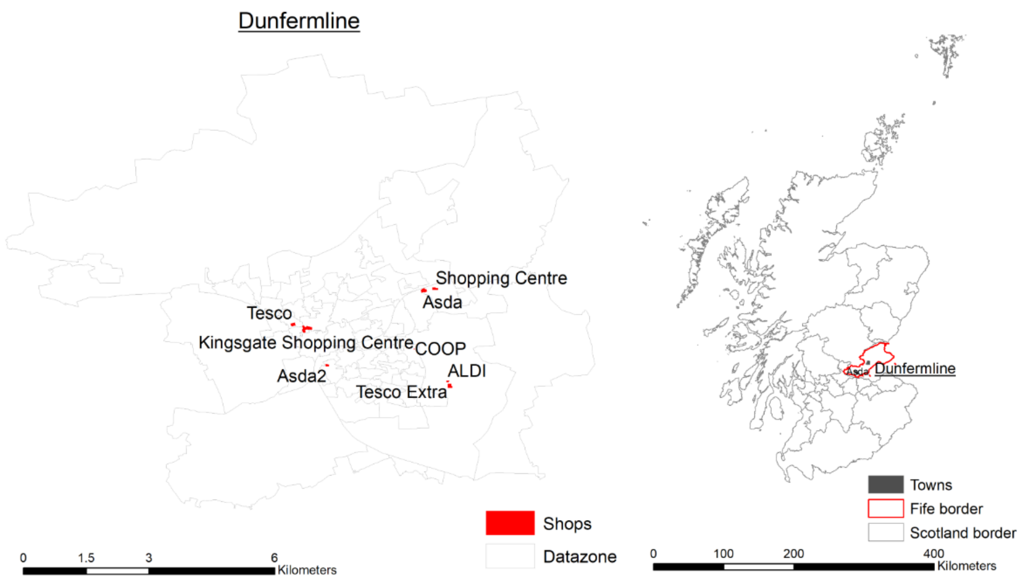
Table 1. An example of a trip chain derived from GPS trajectories

Participant_id	Start tstamp	Stop tstamp	Time spent [seconds]	Mode/purpose	Geographic unit
8	06:22	06:39	995	Home	Datazone A
8	06:58	15:25	30421	Work	Datazone B
8	15:42	15:55	780	Shopping	Datazone C
8	16:12	17:55	6169	Home	Datazone A
8	17:55	18:32	2236	Walk	Datazone A
8	18:32	20:21	6551	Home	Datazone A
8	20:25	20:29	210	Shopping	Datazone B
8	20:34	21:26	3133	Home	Datazone A

Places of residence were then aggregated to administrative units (datazones) to prevent privacy problems and so that it was possible to link the flow data with census variables. Finally, in order to calibrate a retail spatial interaction model, data are needed on distances between consumers and stores and on the size of stores or shopping centres. Road distances between datazone centroids and either stores or shopping centres were obtained from the OpenStreetMap (OSM) road network in both Dunfermline and Glenrothes. A set of possible shopping alternatives in both towns was created from a self-created POI dataset which combined three different POI datasets: the Ordnance Survey POI dataset; the Google Maps POI set; and the OSM POI dataset. From this amalgamated POI database, we identified the main supermarkets and shopping centres in both towns. We created these retail locations as polygons rather than points in order to decrease incorrect linking of trajectories with a shopping destination. The distributions of the datazones and the retail stores for both towns are shown in Figure 2.

Figure 2 somewhere here

A



B

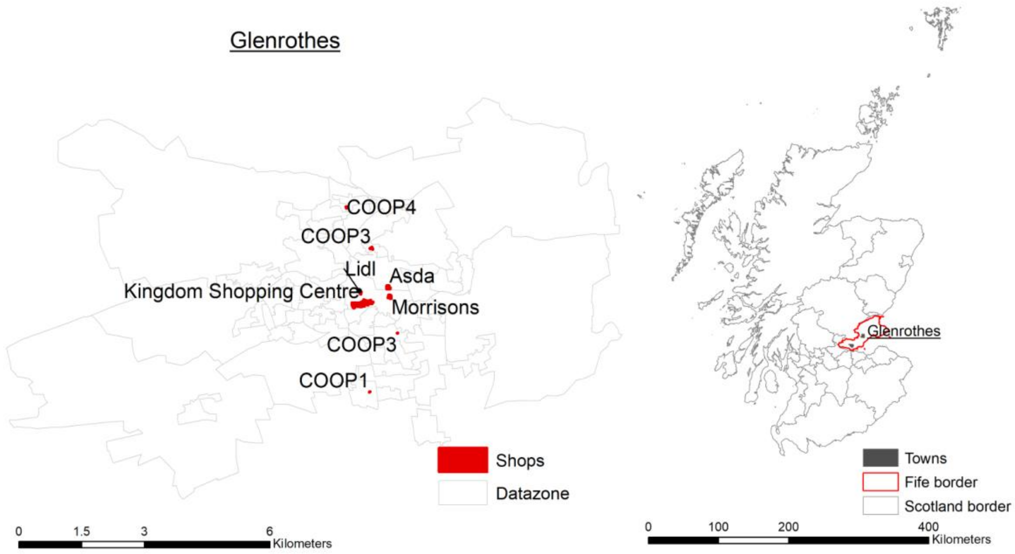


Figure 2. Retail stores in the two towns: A- Dunfermline, B- Glenrothes.

The size of each retailing opportunity is used as a measure of store attractiveness in the model and is created by obtaining a building floor area from the digitised building layer from OpenStreetMap as a proxy of retail area. The sites were verified with Google Street view to confirm retail activity and to identify whether retail area occupied more than one floor. Furthermore, to identify only the actual shopping trips we used opening times for shops to filter out retail-related trips from outside the time range.

Because the GPS data are time and date stamped, this provides the opportunity to calibrate models separately by time of the day, day of the week and for different weather conditions. In order to identify the weather conditions on different days for which the GPS traces were collected, we referred to the local weather conditions for Dunfermline and Glenrothes given by the website Weatherunderground (www.wunderground.com). This contains data on date, time, temperature, humidity, pressure, visibility, wind direction, wind speed and occurrences of rain at locations with meteorological stations nearest Dunfermline and Glenrothes (Edinburgh airport and Leuchars, respectively)

6. Origin-destination matrices

Without further processing trip chains from GPS traces are useful and can provide valuable information on people's activity patterns. However this information is largely limited to visualisations representing spatial patterns of activities along with some descriptive statistics. To be of more use, the data need to be transformed into origin-destination matrices which then can form the basis of calibrating spatial interaction models. Figure 3 represents an origin-destination matrix for an interaction system with *m* origins and *n* destinations. The elements, T_{ij}, of this (m x n) matrix indicate the number of flows between origin *i* and destination *j*. Each row of the matrix represents flows from origin *i* and the columns represent flows into destination *j*. The total number of flows from origin *i* and the total flows into destination *j* are given by the marginal totals, O_i and D_j respectively and the sum of all flows in the system is given by T.

Figure 3 somewhere here

For the calibration of models of retail behaviour, we use only the home-based shopping trips derived from the trip-chaining individual datasets in accord with usual practice (Newing et al., 2015). The GPS traces yielded 280 and 290 individual home-based shopping trips in Dunfermline and Glenrothes respectively (Figure 4).

Origin \ Destination	1	2	3	. . . . . .	n	Total
1	T_{11}	T_{12}	.	. . . . . .	T_{1n}	O_1
2	T_{21}	T_{22}	.	. . . . . .	T_{2n}	O_2
3	.	.	.	. . . . . .	.	.
.	.	.	.	. . . . . .	.	.
.	.	.	.	. . . . . .	.	.
.	.	.	.	. . . . . .	.	.
.	.	.	.	. . . . . .	.	.
.	.	.	.	. . . . . .	.	.
m	T_{m1}	T_{m2}	.	. . . . . .	T_{mn}	O_m
Total	D_1	D_2	.	. . . . . .	D_n	T

Figure 3: An example of origin-destination matrix.

Figure 4 somewhere here

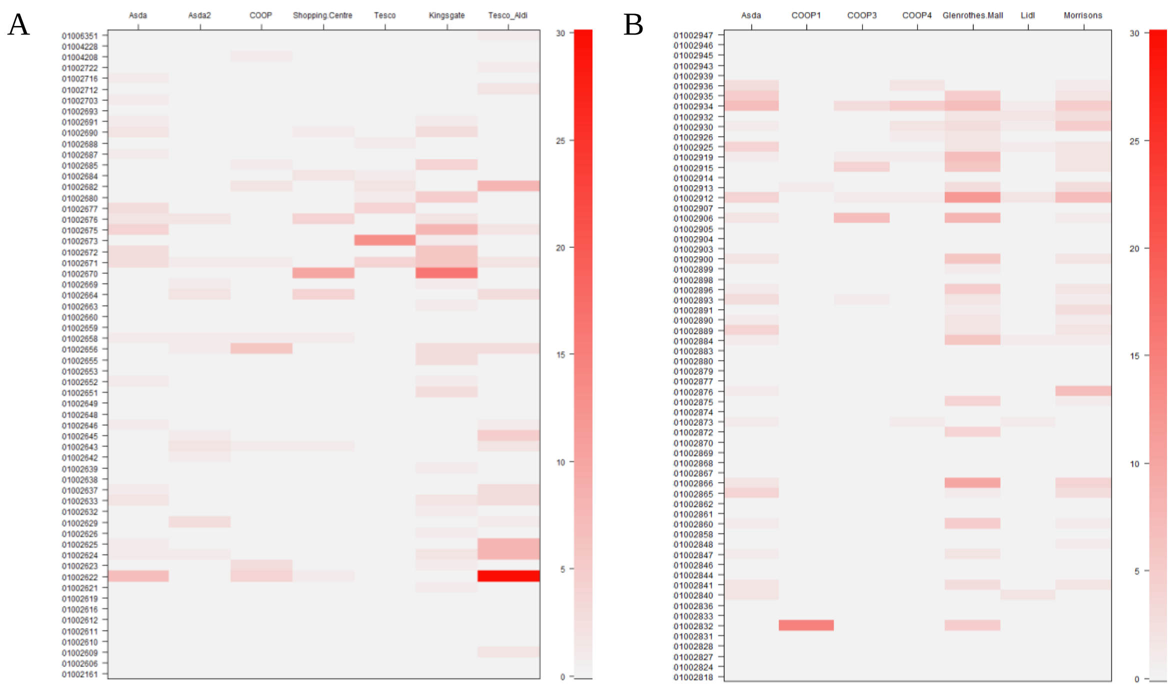


Figure 4. Origin-Destination matrices for A- Dunfermline, B- Glenrothes

7. Results

The primary goal of this paper is to provide evidence that GPS data can be used to calibrate spatial interaction models. In doing so we also highlight the potential for calibrating more temporally disaggregate models to produce new insights into spatial decision-making. To calibrate the models, we use the OD matrices derived from the GPS data of individual home-based shopping trips as described above and a python-based version of the SIMODEL-code (Williams & Fotheringham, 1984) called PySI. In order to compare the distance-decay parameter estimates between the two towns we used a power function of distance rather than an exponential form; the former allows consistent comparison of the estimates because they are elasticities and hence unaffected by scale.

We began by calibrating the origin-constrained spatial interaction model presented in equation (1) with shopping flow data to all stores in both towns. These results are shown in Table 2A which includes parameter estimates and standard errors from this calibration along with the r-squared value.

Table 2 A and 2 B somewhere here

Full set of trips								
	Dunfermline				Glenrothes			
Parameter	Est. value	Std error	t-value	p-value	Est. value	Std error	t-value	p-value
R^2	0.776				0.711			
α	0.635	0.074	8.542	0.000	0.514	0.040	12.383	0.000
β	-0.943	0.078	-12.038	0.000	-0.921	0.137	-6.731	0.000

α- trade area, β- distance decay parameter, *-insignificant

Reduced set of trips								
	Dunfermline				Glenrothes			
Parameter	Est. value	Std error	t-value	p-value	Est. value	Std error	t-value	p-value
R^2	0.817				0.708			
α	0.614	0.000	7.645	0.000	0.862	0.111	7.780	0.000
β	-1.023	0.093	-10.903	0.000	-1.322	0.193	-6.833	0.000

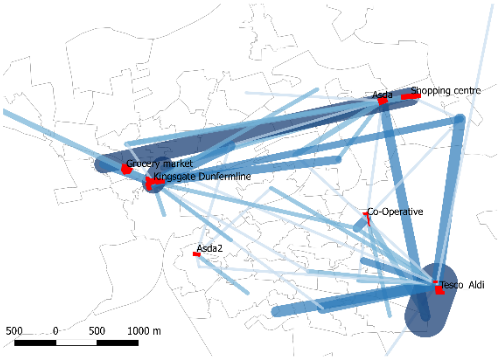
α- trade area, β- distance decay parameter, *-insignificant

Table 2: Origin-constrained spatial interaction model calibrated for shopping trips from GPS trajectories. A- model calibrated for all the trips; B- model calibrated for a reduced set of trips.

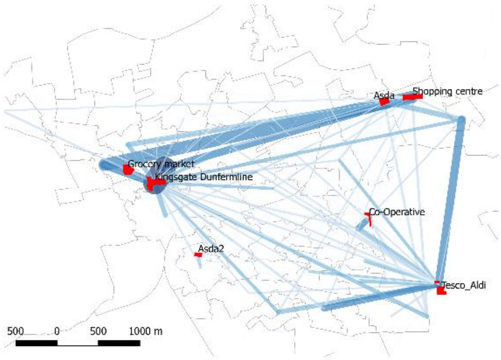
For both towns, the estimated parameters for store size and distance are significant with a p-value < 0.001. The estimated store size parameters for Dunfermline and Glenrothes are 0.635 and 0.514, respectively, indicating that a store's perceived attractiveness by consumers increases at a decreasing rate as size increases so there are diminishing returns to adding to a store's size. The estimated distance decay parameters are -0.943 for Dunfermline and -0.921 for Glenrothes indicating a reasonably strong degree of distance-deterrence in shopping behaviour. These values are in line with results from the calibration of retail shopping models based on traditional survey data (Dolega et al., 2016; Nakaya et al., 2007). The predictive power of the calibrated models, represented by R^2, is 0.78 for Dunfermline and 0.71 for Glenrothes indicating that the model fits the data reasonably accurately. A difference in means test indicates suggests that there is a significant difference (p<0.0001) between the two store size parameters but no significant difference between the estimated distance-decay parameters (p=0.019).

In both towns there is a one dominant retail complex which is multifunctional and contains not only food but also bookstores, boutiques, pharmacies and other possible stores and the inclusion of this multipurpose centre in modelling purely grocery shopping is therefore likely to bias the results. For this reason we excluded the trips to the multifunctional centre in both towns and recalibrated the model. The removal of the two shopping centres reduced the number of flows to 174 in Dunfermline and 212 in Glenrothes. Table 2B contains the results obtained from calibrating the model with this reduced data set.

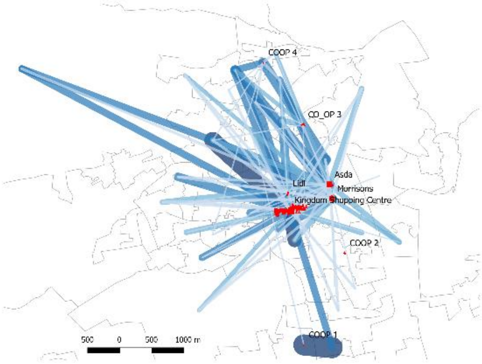
Again all the estimated parameters are significant with p-values < 0.001. When solely grocery trips are analysed, the size of the store becomes a more important factor for consumers in Glenrothes but not in Dunfermline. There is an increase in the strength of the distance-decay effect in both towns but more noticeably so in Glenrothes suggesting that trips to the large multipurpose centre in both towns are less constrained by distance than are pure grocery shopping trips. The R-squared value is relatively unchanged for Glenrothes but increases to 0.82 for Dunfermline suggesting that here the model provides a more accurate representation of grocery decision-making that for general shopping. A difference of means test on both the estimates of the store size and distance-decay parameters suggest there is a significant difference in shopping behaviour in the two towns (p<0.0001) The predicted and observed flow patterns for both data sets in both towns are shown in Figure 4.



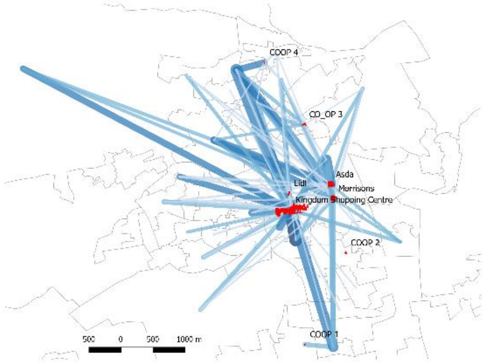
a) Observed shopping flows in Dunfermline



b) Predicted shopping flows in Dunfermline



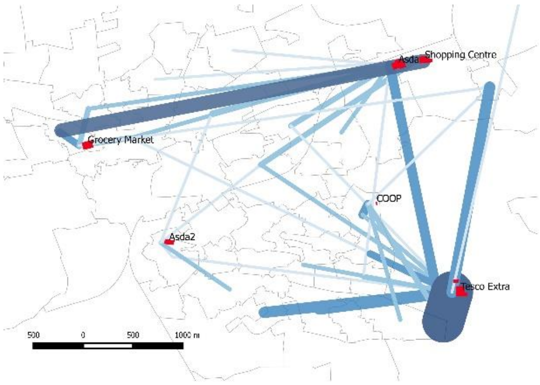
c) Observed shopping flows in Glenrothes



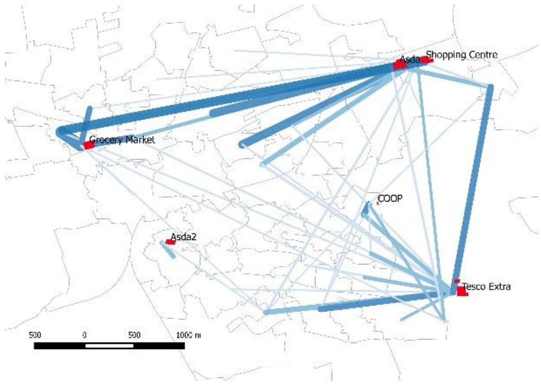
d) Predicted shopping flows in Glenrothes

Legend

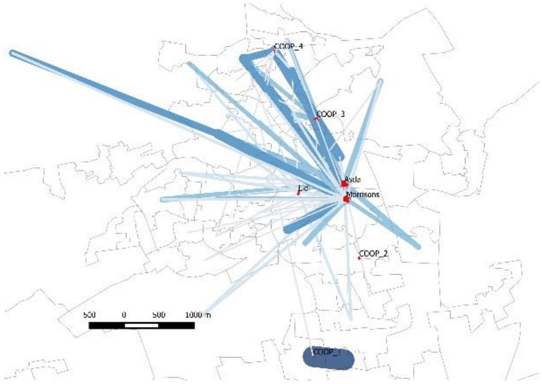
Figure 5: Observed and predicted flows from the initial models (all stores included).



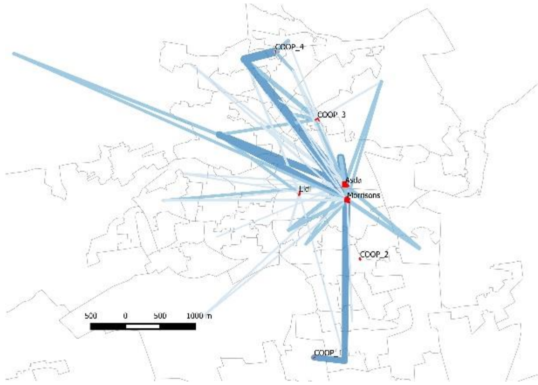
a) Observed shopping flows in Dunfermline



b) Predicted shopping flows in Dunfermline



c) Observed shopping flows in Glenrothes



d) Predicted shopping flows in Glenrothes

Legend

Figure 6: Observed and predicted flows from the initial models (reduced number of stores).

5.1. A comparison of retail behaviour at weekends compared to during the week

An important feature of using GPS traces to study retail behaviour is the ability to examine behaviour at different times of the day or on different days during the week. Here, because of the relatively small sample size, we demonstrate this feature by comparing shopping patterns during the week and on the weekend[1]. For both towns we calibrate the spatial interaction model separately for the two origin-destination matrices representing flows that take place Monday to Friday and those which take place on either Saturday or Sunday. In all cases we use the full

[1] In theory with GPS-derived data it is possible to calibrate spatial interaction models separately for each hour of the day or for periods such as rush hour and non-rush hour and also to disaggregate by consumer type.

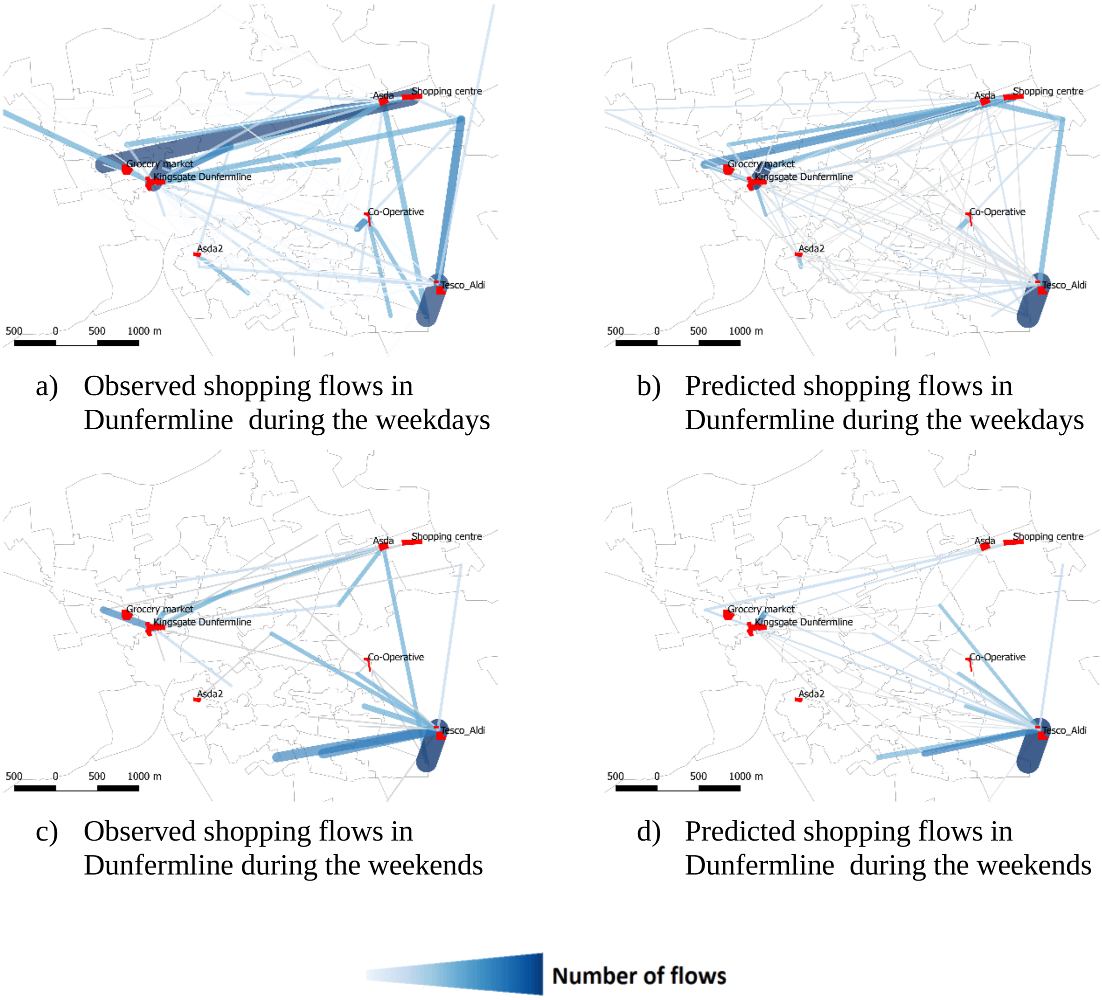
set of retail stores. The results are given in Table 3 and indicate some interesting differences in retail behaviour. In both towns the perceived attractiveness of large stores is much greater at the weekend than during the week (α increases from 0.40 to 0.85 in Dunfermline and from 0.41 to 0.90 in Glenrothes) suggesting that shopping trips on the weekend either have more of a social component to them whereby larger stores offer greater opportunities for diverse types of shopping or that the shopping trips are longer and more products are bought. In Dunfermline the perception of distance as a deterrent to shopping increases at the weekend (β decreases from -0.97 during the week to -1.22 at the weekend) whereas in Glenrothes there is relatively little distance deterrence at the weekends compared to during the week (β increases from -1.00 during the week to -0.12 at weekends. In all comparisons of parameter estimates between weekday shopping and weekend shopping, the differences are significant at $p<0.0001$. The ability of the spatial interaction to replicate flows is slightly better when those flows take place during the week compared to on the weekend. The patterns of both observed and predicted flows for the weekend and during the week are shown in Figures 7 and 8.

Table 3 somewhere here

	Week				Weekend			
Parameter	Est. value	Std error	t-value	p-value	Est. value	Std error	t-value	p-value
	Dunfermline							
R^2	0.788				0.670			
α	0.394	0.095	9.449	0.000	0.850	0.156	5.447	0.000
β	-0.968	0.041	-3.013	0.000	-1.218	0.151	-8.083	0.000
	Glenrothes							
R^2	0.611				0.568			
α	0.406	0.046	0.308	0.000	0.900	0.095	9.449	0.000
β	-1.001	0.147	-10.523	0.003	-0.124	0.041	-3.013	0.000
α- trade area, β- distance decay parameter, *-insignificant								

Table 3: Weekend versus weekday shopping behaviour

Figures 7 and 8 about here



a) Observed shopping flows in Dunfermline during the weekdays

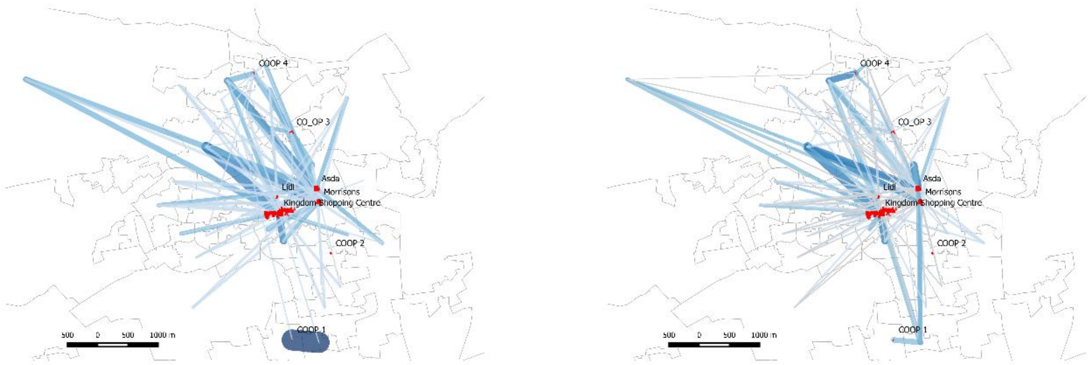
b) Predicted shopping flows in Dunfermline during the weekdays

c) Observed shopping flows in Dunfermline during the weekends

d) Predicted shopping flows in Dunfermline during the weekends

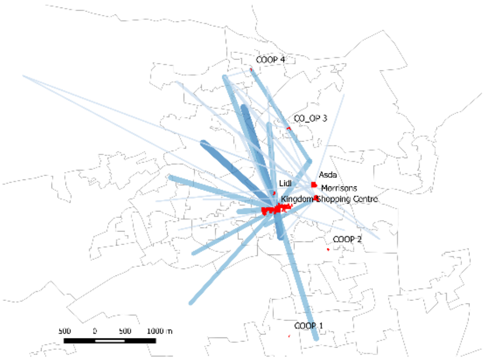
Legend

Figure 7: Observed and predicted patterns of shopping during the week and on weekends in Dunfermline.

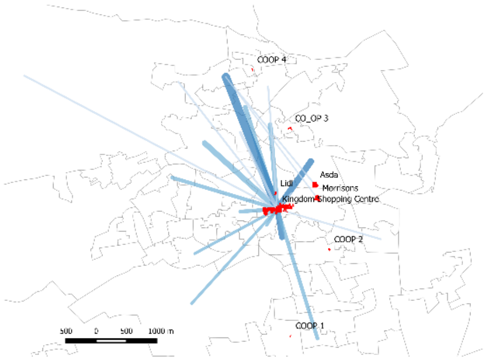


a) Observed shopping flows in Glenrothes during the weekdays

b) Predicted shopping flows in Glenrothes during the weekdays



c) Observed shopping flows in Glenrothes during the weekends



d) Predicted shopping flows in Glenrothes during the weekends

Legend

Figure 8: Observed and predicted patterns of shopping during the week and on weekends in Glenrothes

5.2. A comparison of shopping behaviour under different weather conditions

The effect of weather on consumer behaviour and spending has received only limited attention in the marketing literature (Bruyneel, Dewitte, Franses, & Dekimpe, 2005; Murray, Muro, Finn, & Leszczyc, 2010; Niemira, 2005) and to our knowledge has not received any attention when using GPS movement data in conjunction with spatial interaction models. Here we demonstrate how weather-specific spatial interaction models can be calibrated through the use of GPS-derived flow matrices. To do this we developed a methodology to assign weather conditions to each of the GPS points in the study area. Readings from meteorological station at Edinburgh Airport were used to annotate the GPS trajectories for Dunfermline and data from the meteorological station at the RAF base in Leuchars were used to assign weather conditions for each of the GPS trajectories in Glenrothes. These data were obtained through the wunderground.com website and the selection of these two meteorological stations was based on their proximity to the two towns. The weather data were collected in 30 minutes - 1 hour intervals, so we “interpolated" the values to make them match the trajectory points which were collected for much finer time intervals. Figure 9 highlights the method of assigning the weather data (i.e. rain occurrence and wind speed) to each of the trajectories. The process of transferring weather condition values to a GPS point (x_i; y_i; t_i) is based on the annotation of binary rain reading R (1 for the rain, 0 for no rain) and strength of the wind W. Having a GPS point x_1 from 17:59 which happens to be in between two weather readings from 17:50 (R=1,W=120km/h) and 18:20 (R=0, W=100km/h), we would assign the rain condition R_1 to the time of the nearest weather reading, therefore $R_{1=}$ 1 as $R_{17:50}$. Wind values represented by W are calculated as an average of the two nearest readings to the time of the GPS point so W_1 would be equal to an average between $W_{17:50}$ and $W_{18:20}$ which is (120+100)/2=110km/h.

Figure 9 somewhere here

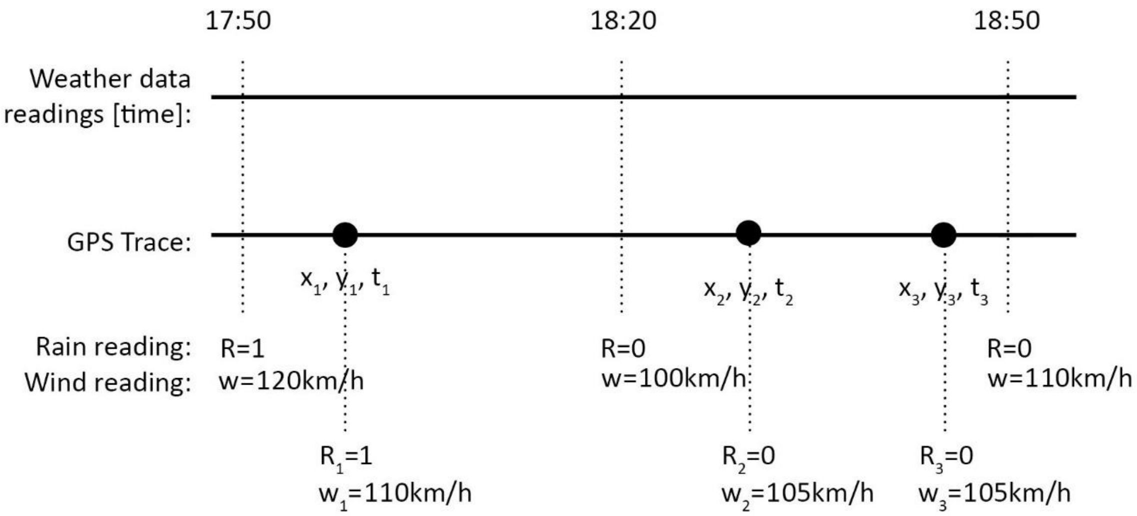


Figure 9: The process of assigning weather condition values to a GPS point (xi; yi; ti). R represents the binary rain reading (1 for the rain, 0 for no rain), W represents strength of the wind. Rain values are assigned to a GPS point based on the existence of rain in any of the two

nearest readings. Wind values represented by W are calculated as an average of the two nearest readings to the time of the GPS point (x_i; y_i).

Having assigned weather conditions to each trajectory and hence to each individual shopping trip, we were able to disaggregate the flow matrix in each town into four sub-matrices: one containing only those trips that took place when it was raining; one when it was dry; one containing only those trips when it was deemed very windy (wind speeds in excess of 35 km/h); and one containing trips taking place when the conditions were relatively still. A summary of the average distances in metres of shopping trips that took place under these four weather conditions is given in Table 4. In both towns shopping trips in both the rain and when it is windy are shorter on average than when it is not raining and not windy.

Table 4 about here

Table 4: Comparison of the mean observed distances [m] for the shopping trips during different weather conditions in the two towns.

	Mean observed distance [m]				
Town	All trips*	Trips in the rain	Trips with no rain	Trips when windy	Trips when no wind
Dunfermline	1641	1515	1924	1515	1746
Glenrothes	1600	1616	1832	1707	1784

*All home-based shopping trips within each city without disaggregating an OD matrix into sub-matrices based on weather condition.

The results of calibrating the retail spatial interaction model on each of the four origin-destination matrices including flows to the multipurpose shopping centre in both towns are given in Table 5. The results indicate that aspects of shopping behaviour do change according to weather conditions. For instance in both Dunfermline and Glenrothes shoppers are more attracted to larger stores and perceive distance to be more of deterrent when it is raining (all comparisons of parameter estimates are significant at least at p=0.0017). Windy conditions also have a significant impact on shopping behaviour. In both towns there is a significant increase ($p<0.0001$) in distance decay under windy conditions. However, the results of varying wind conditions on the attractiveness of large retail outlets is less convincing. Although in Dunfermline there is a significant increase in the attractiveness of large stores when it is windy, the reverse is the case in Glenrothes with a significant decrease in the attractiveness of large stores (in both tests, $p<0.0001$).

Table 5 somewhere here

Table 5: Calibration results for different weather condition - full choice set of stores.

	Full choice set of stores							
	Dunfermline							
Parameter	Est. value	Std error	t-value	p-value	Est. value	Std error	t-value	p-value
	Rain				No Rain			
R^2	0.693				0.746			
α	0.805	0.110	7.306	0.000	0.494	0.101	4.862	0.000
β	-0.972	0.110	-8.822	0.000	-0.930	0.112	-8.238	0.000
	Wind				No Wind			
R^2	0.760				0.683			
α	0.670	0.105	6.378	0.000	0.593	0.106	5.599	0.000
β	-1.000	0.105	-9.462	0.000	-0.867	0.117	-7.370	0.000
	Glenrothes							
	Rain				No Rain			
R^2	0.658				0.738			
α	0.472	0.065	7.243	0.000	0.544	0.151	10.679	0.000
β	-1.233	0.201	-6.142	0.000	-0.645	0.192	-3.352	0.001
	Wind				No Wind			
R^2	0.533				0.564			
α	0.483	0.058	8.347	0.006	0.538	0.055	9.680	0.000
β	-1.083	0.195	-5.554	0.000	-0.785	0.194	-4.040	0.000

α- trade area, β- distance decay parameter, *-insignificant

The above calibrations were repeated with the multipurpose shopping centre in each town removed from the analysis. The results are given in Table 6. These results reinforce those above. Under rainy and windy conditions, consumers tend to have a greater preference for larger stores and for stores in close proximity to their residences. This is most clearly seen in Glenrothes where the estimated distance-decay parameter is -0.86 in dry conditions and -2.06 in wet conditions. In still conditions, the estimated distance-decay parameter is -0.82 whereas in windy conditions it is -2.19. Similar, although less dramatic, effects are seen in Dunfermline. These results are important because they demonstrate the use of GPS-derived flow data to calibrate disaggregated spatial interaction models and that shopping behaviour varies according to weather conditions.

Table 6 about here

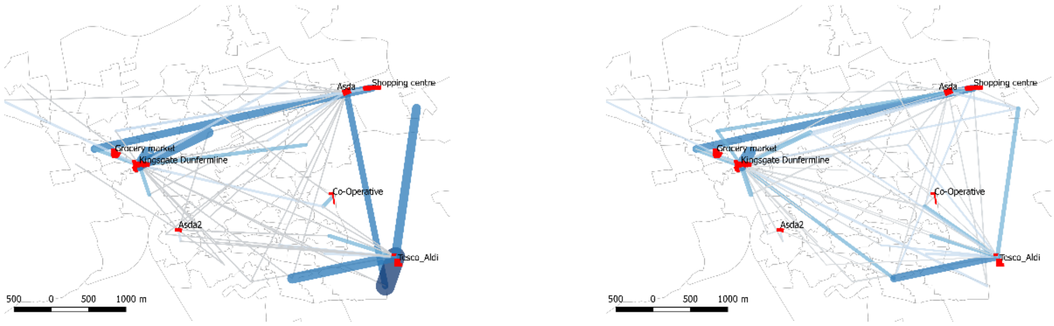
Table 6: Calibration results for different weather conditions - grocery only stores

	Reduced choice set of stores							
	Dunfermline							
Parameter	Est. value	Std error	t-value	p-value	Est. value	Std error	t-value	p-value
	Rain				No Rain			
R^2	0.776				0.759			
α	0.778	0.123	6.302	0.000	0.484	0.107	4.511	0.000
β	-1.042	0.138	-7.307	0.000	-1.017	0.129	-7.883	0.000
	Wind				No Wind			
R^2	0.788				0.739			
α	0.659	0.117	5.606	0.000	0.566	0.111	8.098	0.000
β	-1.092	0.126	-8.691	0.000	-0.923	0.142	-6.476	0.000
	Glenrothes							
	Rain				No Rain			
R^2	0.610				0.617			
α	1.132	0.222	5.098	0.000	0.776	0.127	4.511	0.000
β	-2.061	0.362	-5.697	0.000	-0.855	0.246	-7.883	0.000
	Wind				No Wind			
R^2	0.520				0.398			
α	1.307	0.217	6.016	0.000	0.626	0.131	4.755	0.000
β	-2.191	0.368	-5.948	0.000	-0.821	0.235	-3.501	0.001

α- trade area, β- distance decay parameter, *-insignificant

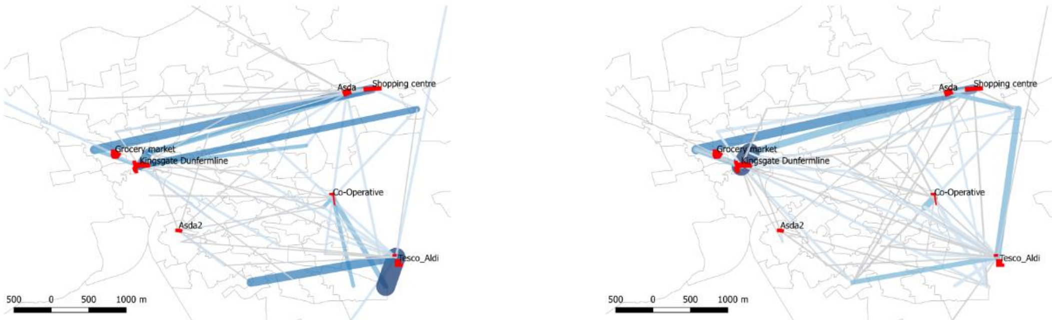
Comparing the above results with those from the set of destinations including the multipurpose centres, consumers in Glenrothes appear to be much more sensitive to weather conditions when it comes to deciding on choice of grocery store than consumers in Dunfermline. The estimated distance-decay parameters from Glenrothes become much more negative in rainy and windy conditions when the multipurpose shopping centre is removed from the analysis whereas the equivalent estimates for Dunfermline are much more stable. The observed and predicted flow patterns for Dunfermline shoppers under various weather conditions are shown in Figure 10 and the equivalent flows for Glenrothes are shown in Figure 11.

Figure 10 and 11 about here



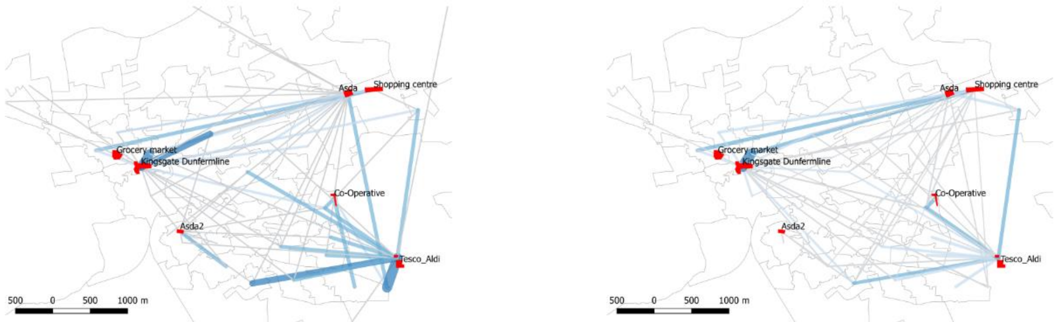
a) Observed shopping flows in Dunfermline on rainy days

b) Predicted shopping flows in Dunfermline on rainy days



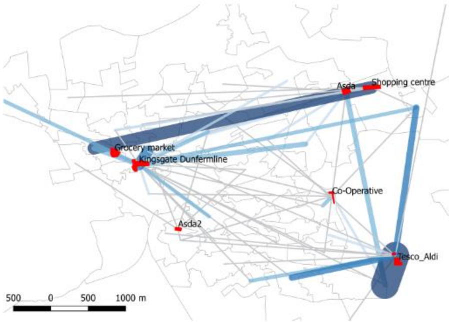
c) Observed shopping flows in Dunfermline on dry days

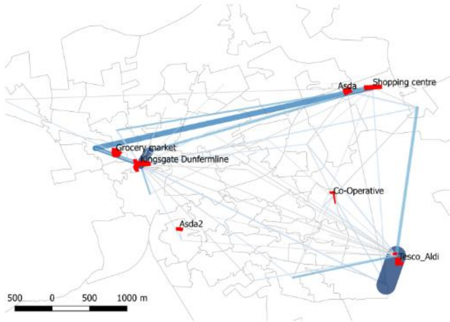
d) Predicted shopping flows in Dunfermline on dry days



e) Observed shopping flows in Dunfermline on windy days

g) Observed shopping flows in Dunfermline on windy days

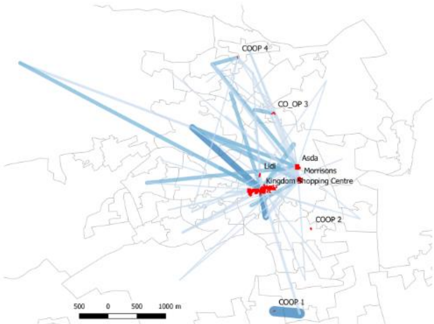




f) Observed shopping flows in Dunfermline on non-windy days

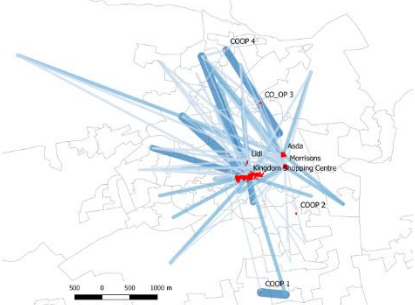
h) Observed shopping flows in Dunfermline on non-windy days

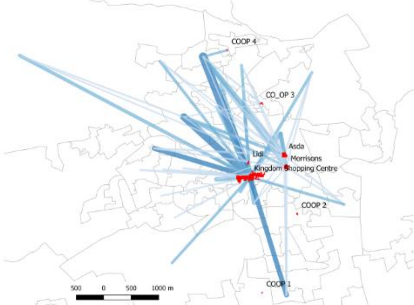
Figure 10: Observed and predicted patterns of shopping in different weather conditions in Dunfermline



a) Observed shopping flows in Glenrothes on rainy days

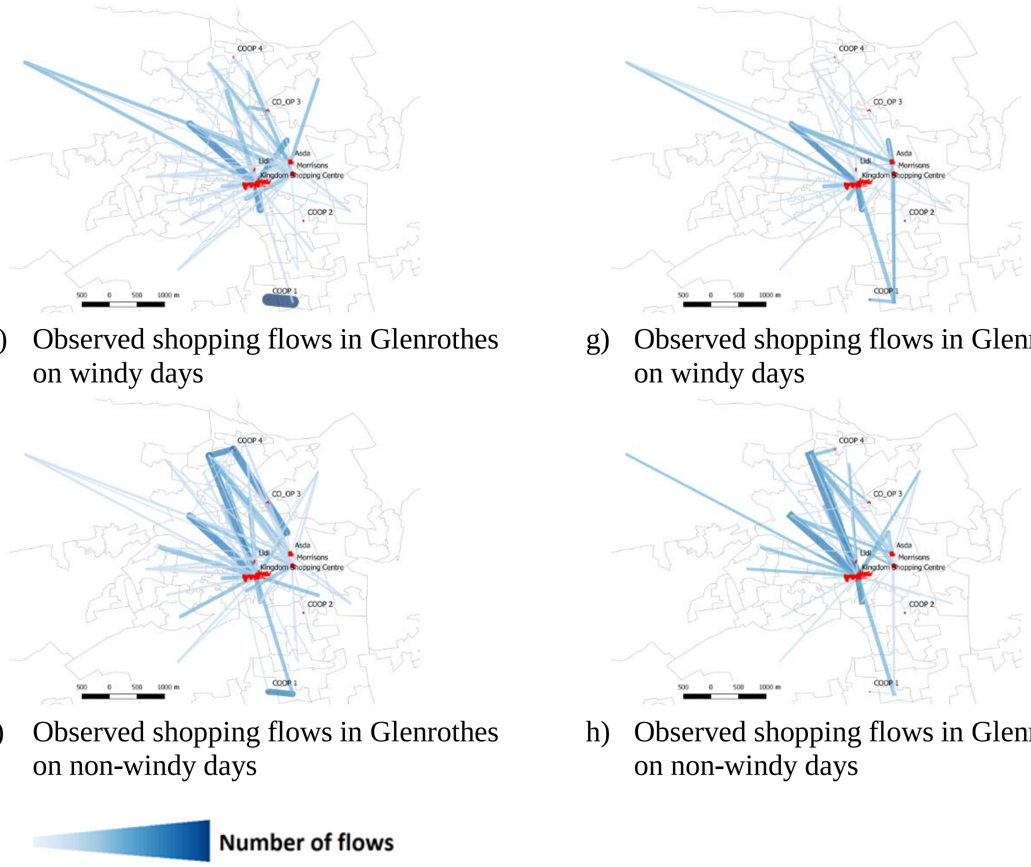
b) Predicted shopping flows in Glenrothes on rainy days





c) Observed shopping flows in Glenrothes on dry days

d) Predicted shopping flows in Glenrothes on dry days



e) Observed shopping flows in Glenrothes on windy days

g) Observed shopping flows in Glenrothes on windy days

f) Observed shopping flows in Glenrothes on non-windy days

h) Observed shopping flows in Glenrothes on non-windy days

Figure 11: Observed and predicted patterns of shopping in different weather conditions in Glenrothes

8. Discussion and conclusions

In this paper we introduce a framework for calibrating spatial interaction models using flows derived from GPS data. We focus on one type of model commonly employed in retailing – a production-constrained spatial interaction model– which we use to investigate shopping behaviour in two towns in Scotland, Dunfermline and Glenrothes. To demonstrate the potential of GPS traces for the calibration of spatial interaction models, we calibrate separate models for weekend shopping trips and weekday shopping trips and for shopping trips taking place in different weather conditions. For the latter, we designed a methodology to assign weather conditions to the GPS traces and then calibrated models for rainy versus dry conditions and for windy versus calm conditions. Significant differences in shopping behaviour were measured for both different periods of the week and under different weather conditions. To our knowledge, such differences have not been identified previously in the calibration of retail choice models because of a lack of suitable data. . This study takes advantage of increasingly available GPS trajectory data to produce origin-destination flow matrices which are used to calibrate spatial interaction models.

One of the recurring issues with the use of GPS trajectories for studies about spatial behaviour is the `noisiness' of the data caused by the unpredictability of how the trackers were used. In our study, participants were asked to carry their fully charged trackers with them at all times. In practice, trackers occasionally ran out of charge for various reasons, participants forgot to take them with them for certain trips, and the trackers occasionally lose the GPS signal connection. These issues need to be addressed to increase the utility of such data but it would seem inevitable that as GPS-based tracking becomes more reliable and the traces become more available, this form of data collection will replace conventional methods for understanding human spatial behaviour.. Because current GPS trackers have limitations regarding convenience and reliability, this study is only at the forefront of the use of such technology in the field of spatial interaction modelling and has clear limitations in terms of sample size and potential bias. However, as people become increasingly used to sharing their locational information and GPS trackers become more universal (such as through reporting apps on smart phones), these limitations will diminish in importance and the value added by having movement data which is time-stamped and spatially comprehensive will be increasingly recognised. GPS-based technology is changing how we are able to view and understand the world and how people interact with their environment. It is part of the broader concepts of ’Citizens as Sensors’, ‘Collective Sensing’ and ‘Citizen Science’ (Goodchild, 2007), in which “*people act as non-technical sensors with contextual intelligence and comprehensive knowledge*" (Resch, 2013, p. 393). GPS-based technology has already changed the world in major ways: we now depend on it for navigation and for finding out information on our surroundings. It is not difficult to imagine a world in which everyone is a sensor relating information about our movement patterns and our environment to central repositories. We are just at the beginning of such developments. Hence, this paper is very timely. There is a need to understand the necessary steps involved in transforming raw GPS data from individuals into usable trip trajectories and origin-destination matrices and to understand the limitations and potential uses of such data. Consequently, although the methods and results discussed in this paper are drawn from rather crude and relatively small samples in a limited spatial context, two relatively small towns in Scotland, they have the potential to guide future analysis of movement patterns and spatial behaviour using volunteered geographic information.

9. References

Ahas, R., Aasa, A., Yuan, Y., Raubal, M., Smoreda, Z., Liu, Y., … Zook, M. (2015). Everyday space–time geographies: using mobile phone-based sensor data to monitor urban activity in Harbin, Paris, and Tallinn. *International Journal of Geographical Information Science*, *29*(11), 2017–2039. http://doi.org/10.1080/13658816.2015.1063151

Anastasi, G., & Borgia, E. (2004). Wi-fi in ad hoc mode: a measurement study. *PerCom 2004.*

Batten, D., & Boyce, D. (1987). Spatial interaction, transportation, and interregional commodity flow models.

Bhat, C., Guo, J., Srinivasan, S., & Sivakumar, A. (2004). Comprehensive Econometric Microsimulator for Daily Activity-Travel Patterns. *Transportation Research Record*, *1894*(1), 57–66. http://doi.org/10.3141/1894-07

Bruyneel, S., Dewitte, S., Franses, P., & Dekimpe, M. (2005). *Why consumers buy lottery tickets when the sun goes down on them. The depleting nature of weather-induced bad moods*. *ERIM Report Series Research in Management ERS-2005-045-MKT*.

Calabrese, F., Diao, M., Di Lorenzo, G., Ferreira, J., & Ratti, C. (2013). Understanding individual mobility patterns from urban sensing data: A mobile phone trace example. *Transportation Research Part C: Emerging Technologies*, *26*, 301–313. http://doi.org/10.1016/j.trc.2012.09.009

Calabrese, F., Ferrari, L., & Blondel, V. D. (2014). Urban Sensing Using Mobile Phone Network Data: A Survey of Research. *ACM Computing Surveys*, *47*(2), 1–20. http://doi.org/10.1145/2655691

Cangialosi, A., Monaly Jr., J., & Yang, S. (2007). Leveraging RFID in hospitals: Patient life cycle and mobility perspectives. *IEEE Communications Magazine*, *45*(9), 18–23. http://doi.org/10.1109/MCOM.2007.4342874

Clarke, G., Langley, R., & Cardwell, W. (1998). Empirical applications of dynamic spatial interaction models. *Computers, Environment and Urban Systems*, *22*(2), 157–184. http://doi.org/10.1016/S0198-9715(98)00021-0

de Vries, J. J., Nijkamp, P., & Rietveld, P. (2009). Exponential or power distance-decay for commuting? An alternative specification. *Environment and Planning A*, *41*(2), 461–480. http://doi.org/10.1068/a39369

Di Lorenzo, G., Reades, J., Calabrese, F., & Ratti, C. (2012). Predicting personal mobility with individual and group travel histories. *Environment and Planning B: Planning and Design*, *39*(5), 838–857. http://doi.org/10.1068/b37147

Dolega, L., Pavlis, M., & Singleton, A. (2016). Estimating attractiveness , hierarchy and catchment area extents for a national set of retail centre agglomerations. *Journal of Retailing and Consumer Services*, *28,* 78–90. http://doi.org/10.1016/j.jretconser.2015.08.013

Eckfeldt, B., & Bruce. (2005). What does RFID do for the consumer? *Communications of the ACM*, *48*(9), 77. http://doi.org/10.1145/1081992.1082024

Fischer, M. M. (2002). Neural Spatial Interaction Modeling. In M. M. Fischer∗ & Y. Leung (Eds.), *Geocomputational modelling: Thechniques and appliactions* (Vol. 34, pp. 195–219). Berlin: Springer.

Fotheringham, A. S. (1983). Some theoretical aspects of destination choice and their relevance to production-constrained gravity models. *Environment and Planning A*, *15*(8), 1121–1132. http://doi.org/10.1068/a151121

Fotheringham, A. S., & O’Kelly, M. E. (1989). *Spatial Interaction Models*. *International Encyclopedia of the Social & Behavioral Sciences*. Elsevier. http://doi.org/10.1016/B0-08-043076-7/02519-5

Fotheringham, A. S., & Trew, R. (1993). Chain image and store-choice modeling: the effects of income and race. *Environment and Planning A*, *25*(2), 179–196. http://doi.org/10.1068/a250179

Golledge, R. G., & Stimson, R. J. (1997). *Spatial behaviour: a geographic perspective*. *Spatial behaviour: a geographic perspective*. Guilford Press.

Gong, H., Chen, C., Bialostozky, E., & Lawson, C. T. (2012). A GPS/GIS method for travel mode detection in New York City. *Computers, Environment and Urban Systems*, *36*(2), 131–139. http://doi.org/10.1016/j.compenvurbsys.2011.05.003

Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal*, *69*(4), 211–221. http://doi.org/10.1007/s10708-007-9111-y

Haynes, K., & Fotheringham, A. S. (1984). *Gravity and Spatial Interaction Models*. Vol. 2. Beverly Hills: Sage publications.

Kwan, M.-P., & Neutens, T. (2014). Space-time research in GIScience. *International Journal of Geographical Information Science*, *28*(5), 851–854. http://doi.org/10.1080/13658816.2014.889300

Kwan, M. P. (2007). Mobile communications, social networks, and urban travel: Hypertext as a new metaphor for conceptualizing spatial interaction. *Professional Geographer*, *59*(4), 434–446. http://doi.org/10.1111/j.1467-9272.2007.00633.x

Lakshmanan, J. R., & Hansen, W. G. (1965). A retail market potential model. *Journal of the American Institute of Planners*, *31*(2), 134–143. http://doi.org/10.1080/01944366508978155

Laube, P., Dennis, T., Forer, P., & Walker, M. (2007). Movement beyond the snapshot - Dynamic analysis of geospatial lifelines. *Computers, Environment and Urban Systems*, *31*(5), 481–501. http://doi.org/10.1016/j.compenvurbsys.2007.08.002

Lerman, S., & Liu, T. (1984). Microlevel econometric analysis of retail closure. *London Papers of the Regional Science Association*, *14*.

Lloyd, A., & Cheshire, J. (2017). Deriving retail centre locations and catchments from geo-tagged Twitter data. *Computers, Environment and Urban Systems*, *61*, 108–118. http://doi.org/10.1016/j.compenvurbsys.2016.09.006

Long, J., & Nelson, T. (2012). A review of quantitative methods for movement data. *International Journal of Geographical Information Science*, *27*(2), 1–27. http://doi.org/10.1080/13658816.2012.682578

Long, Y., & Thill, J. C. (2015). Combining smart card data and household travel survey to analyze jobs-housing relationships in Beijing. *Computers, Environment and Urban Systems*, *53*, 19–35. http://doi.org/10.1016/j.compenvurbsys.2015.02.005

Lovelace, R., Birkin, M., Cross, P., & Clarke, M. (2016). From Big Noise to Big Data: Toward the Verification of Large Data sets for Understanding Regional Retail Flows. *Geographical Analysis*.

Merino, M., & Ramirez-Nafarrate, A. (2015). Estimation of retail sales under competitive location in Mexico. *Journal of Business Research*, 1–7. http://doi.org/10.1016/j.jbusres.2015.06.050

Murray, K., Muro, F. Di, Finn, A., & Leszczyc, P. (2010). The effect of weather on consumer spending. *Journal of Retailing and Consumer Services*, (17.6), 512–520.

Nakaya, T., Fotheringham, S. A., Hanaoka, K., Clarke, G., Ballas, D., & Yano, K. (2007). Combining microsimulation and spatial interaction models for retail location analysis. *Journal of Geographical Systems*, *9*(4), 345–369. http://doi.org/10.1007/s10109-007-0052-2

Newing, A., Clarke, G. P., & Clarke, M. (2015). Developing and Applying a Disaggregated Retail Location Model with Extended Retail Demand Estimations. *Geographical Analysis*, *47*(3), 219–239. http://doi.org/10.1111/gean.12052

Niemira, M. (2005). Weather matters. *The Impact of Climate, Weather and Seasons on …*.

Patterson, D. J., Lin, L. A., Fox, D., & Kautz, H. (2003). Inferring high-level behavior from low-level sensors. *Ubicomp 2003: Ubiquitous Computing*, *2864*, 73–89. http://doi.org/10.1007/b93949

Preston, V., & McLafferty, S. (2016). Revisiting Gender, Race, and Commuting in New York. *Annals of the American Association of Geographers*, 1–11. http://doi.org/10.1080/00045608.2015.1113118

Ratti, C., Frenchman, D., Pulselli, R. M., & Williams, S. (2006). Mobile Landscapes: Using Location Data from Cell Phones for Urban Analysis. *Environment and Planning B: Planning and Design*, *33*(5), 727–748. http://doi.org/10.1068/b32047

Ratti, C., Sobolevsky, S., Calabrese, F., Andris, C., Reades, J., Martino, M., … Strogatz, S. H. (2010). Redrawing the Map of Great Britain from a Network of Human Interactions. *PLoS One*, *5*(12), e14248. http://doi.org/10.1371/journal.pone.0014248

Resch, B. (2013). People as sensors and collective sensing-contextual observations complementing geo-sensor network measurements. In *Progress in Location-Based Services* (pp. 391–406). Springer Berlin Heidelber.

Rodriguez, D. A., & Joo, J. (2004). The relationship between non-motorized mode choice and the local physical environment. *Transportation Research Part D: Transport and Environment*, *9*(2), 151–173. http://doi.org/10.1016/j.trd.2003.11.001

Shen, L., & Stopher, P. R. (2014). Review of GPS Travel Survey and GPS Data-Processing Methods. *Transport Reviews*, *34*(3), 316–334. http://doi.org/10.1080/01441647.2014.903530

Siła-Nowicka, K., Vandrol, J., Oshan, T., Long, J. A., Demšar, U., & Fotheringham, A. S. (2016). Analysis of human mobility patterns from GPS trajectories and contextual information. *International Journal of Geographical Information Science*, *30*(5), 881–906. http://doi.org/10.1080/13658816.2015.1100731

Spaccapietra, S. (2009). Semantic Trajectories Stefano Spaccapietra SeCoGIS 2009, Gramado GeoPKDD. - ppt download.

Timmermans, H. (1984). Discrete Choice Models versus Decompositional Multiattribute Preference Models: A Comparative Analysis of Model Performance in the Context of Spatial Shopping-behaviour.

Tonnelier, E., Baskiotis, N., Guigue, V., & Gallinari, P. (2016). Smart card in public transportation: Designing a analysis system at the human scale. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)* (pp. 1336–1341). IEEE. http://doi.org/10.1109/ITSC.2016.7795730

Vu, L., Do, Q., & Nahrstedt, K. (2011). Jyotish: A novel framework for constructing predictive model of people movement from joint wifi/bluetooth trace. *Pervasive Computing*.

Williams, P., & Fotheringham, A. (1984). The calibration of spatial interaction models by maximum likelihood estimation with program SIMODEL.

Wilson, A. (1967). A statistical theory of spatial distribution models. *Transportation Research*, *1*(3), 253–269. http://doi.org/10.1016/0041-1647(67)90035-4

Wilson, A. (1970). *Entropy in Urban and Regional Modelling*. Pion Ltd.

Wilson, J. P., & Fotheringham, a S. (2008). *The Handbook of Geographic Information Science*. *The handbook of geographic*. John Wiley & Sons. http://doi.org/10.1002/9780470690819

Wolf, J., Guensler, R., & Bachman, W. (2001). Elimination of the travel diary: Experiment to derive trip purpose from global positioning system travel data. *Transportation Research Record*, *1768*(1), 125–134. http://doi.org/10.3141/1768-15

Xiao, G., Juan, Z., & Zhang, C. (2016). Detecting trip purposes from smartphone-based travel surveys with artificial neural networks and particle swarm optimization. *Transportation Research Part C: Emerging Technologies*, *71*, 447–463. http://doi.org/10.1016/j.trc.2016.08.008

Yan, Z., Chakraborty, D., & Parent, C. (2013). Semantic trajectories: Mobility data computation and annotation. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *4*(3), 49.

Yue, Y., Wang, H. dong, Hu, B., Li, Q. quan, Li, Y. guang, & Yeh, A. G. O. (2012). Exploratory calibration of a spatial interaction model using taxi GPS trajectories. *Computers, Environment and Urban Systems*, *36*(2), 140–153. http://doi.org/10.1016/j.compenvurbsys.2011.09.002

Zacharewicz, G., Deschamps, J.-C., & Francois, J. (2011). Distributed simulation platform to design advanced RFID based freight transportation systems. *Computers in Industry*, *62*(6), 597–612. http://doi.org/10.1016/j.compind.2011.04.009

Zhao, Z., Chua, G., & Zhao, J. (2012). Evolution of trip chaining patterns in London from 1991 to 2010. In *Innovations in Improving the Sensitivity of Models*.

Zhong, C., Manley, E., Müller Arisona, S., Batty, M., & Schmitt, G. (2015). Measuring variability of mobility patterns from multiday smart-card data. *Journal of Computational Science*, *9*, 125–130. http://doi.org/10.1016/j.jocs.2015.04.021

Zuo, Y., Pimple, M., & Lande, S. (2010). A Framework for RFID Survivability Requirement Analysis and Specification. *Innovations in Computing Sciences*, (5), 26–34.

List of tables

List of figures