Model behaviour: GPS v GIS to examine our journey to work

Alice Dalton[12], Andrew Jones[23], David Ogilvie[24]

[1] School of Environmental Sciences, University of East Anglia, Norwich, NR4 7TJ, UK

Tel. +44 (0)1603 591361; Fax +44 (0)1603 591327

a.dalton@uea.ac.uk, www.alicedalton1.wordpress.com

[2]UKCRC Centre for Diet and Activity Research (CEDAR), Institute of Public Health, Cambridge, UK

[3]Norwich Medical School, University of East Anglia, Norwich, NR4 7TJ, UK

[4]Medical Research Council Epidemiology Unit, Institute of Public Health, Cambridge, UK

KEYWORDS: GPS, commute, transport, travel mode, physical activity and health

1. Introduction

In recent years, research has sought to understand how to encourage health promoting travel behaviours such as walking and cycling (Bauman et al. 2012). Some of this work has looked at which features of the environment on the travel route between home and work are predictors of such behaviours (Fraser and Lock 2010; Panter and Jones 2010). Often, objective measures of the characteristics of these routes are used, usually based on shortest distance or fastest travel time, modelled using geographic information system (GIS) software (Duncan et al. 2009). This is relatively straightforward and simple, requiring just postcodes or house locations as starting points and work locations for the destination of interest. However, whilst these GIS based measures are ubiquitous due to their methodological simplicity, they may not well reflect the actual route taken (Oliver et al. 2010), they cannot tell us much about where healthy behaviours occur (Dill 2009), and therefore are limited in their ability to help us learn where to target interventions (Maddison and Ni Mhurchu 2009).

Recently, studies have moved towards the use of global positioning system (GPS) devices to record locations visited by participants, providing new opportunities to measure the actual route taken, and thus validate, supplement or replace the GIS based measures. For example, research has looked into the child’s journey to school (Duncan and Mummery 2007) and cycling activity in an urban area (Dill 2009). The findings suggested that GIS routes are not representative of actual routes taken, but to our knowledge, there has been no study comparing the use of GIS and GPS to measure travel behaviour to work amongst adults. Therefore, we do not know how our conclusions on the role of the environment as a determinant of travel behaviour are influenced by the assumptions of GIS modelling. This study attempts to address this shortcoming using GPS amongst a sample of commuters in the city of Cambridge, UK.

1.1 Study background

This research utilised data obtained from a sample of commuters taking part in the Commuting and Health in Cambridge study in Cambridge, UK. The details of this study have been outlined previously (Ogilvie et al. 2010). Participants were aged 16 and over, working in Cambridge and living within 30km of the city, but not in the immediate area of where they worked. They were sampled using a workplace recruitment strategy that targeted a variety of workplaces and employers in a range of geographical locations across the city centre and urban fringe. The data used for this analysis were collected between May and November 2010 using postal questionnaires and GPS devices (Panter et al. 2011). Participants reported their travel to and from work, and a total of 197 respondents submitted at least two days of GPS data which recorded the route(s) taken for their journey to and from work. They also returned questionnaires with valid postcodes which could be used to identify their home and work locations. Ethical approval for the study was obtained from the Hertfordshire Research Ethics Committee and written informed consent was provided by each participant.

1.2 Research questions

The aim of this research was to compare the actual commuting routes with modelled routes created using GIS software. The following three questions were addressed:

1. How well do routes modelled in GIS software reflect actual routes recorded using GPS devices?
2. How strongly are differences associated with travel mode?
3. How does using GPS modify our conclusions on the contribution of travel to work on energy expenditure and carbon emissions?

2. Method

Actual commute routes for every participant journey to and from home and work were identified by extracting recorded GPS data points from randomly selected individual participant data files, based on work start/end times and travel mode (cycle; walk; bus; car; car and walk; car and cycle), as reported in the participant questionnaires. These were read into ESRI's® ArcMap™ 9.3 and projected onto Ordnance Survey (OS) National Grid. Redundant data points (recorded either before or after the participant left/arrived at home/work) and scatter (points recorded when the device was unable to establish a correct position, such as when inside buildings) were removed manually, then routes (linear features) were created for every individual journey using Hawths Analysis Tools for ArcGIS™ (Beyer 2004).

Home and work postcodes taken from the questionnaires were georeferenced using the OS Address Layer 2® database. Modelled routes to work were created using GIS software, based on the shortest distances along the appropriate route network: OS MasterMap® Integrated Transport Network™ (ITN) for journeys made by car and public transport; ITN plus local authority rights-of-way data (public footpaths, bridleways and byways) for those made by cycle and foot.

Geographical exposure variables were quantified for both the modelled routes (GIS) and actual routes (GPS), including distance, route directness, proportion of route along A and B (i.e. busy) roads and the number of destinations (retail, food, leisure, education) within 100m of the route (PointX Ltd 2010), as shown in Figure 1. Differences between the modelled and actual routes were examined by travel mode. The degree of correspondence between modelled and actual routes was calculated by buffering the modelled routes by 15m (allowing for positional error and road width) and calculating the percentage of the actual route that was within this boundary. Energy expenditure in Kcal/hour on actual and modelled routes was estimated using the metabolic equivalent for each mode of travel (Ainsworth et al. 2011). Carbon Dioxide emissions were estimated for car journeys, using values for an average vehicle (Defra 2012).

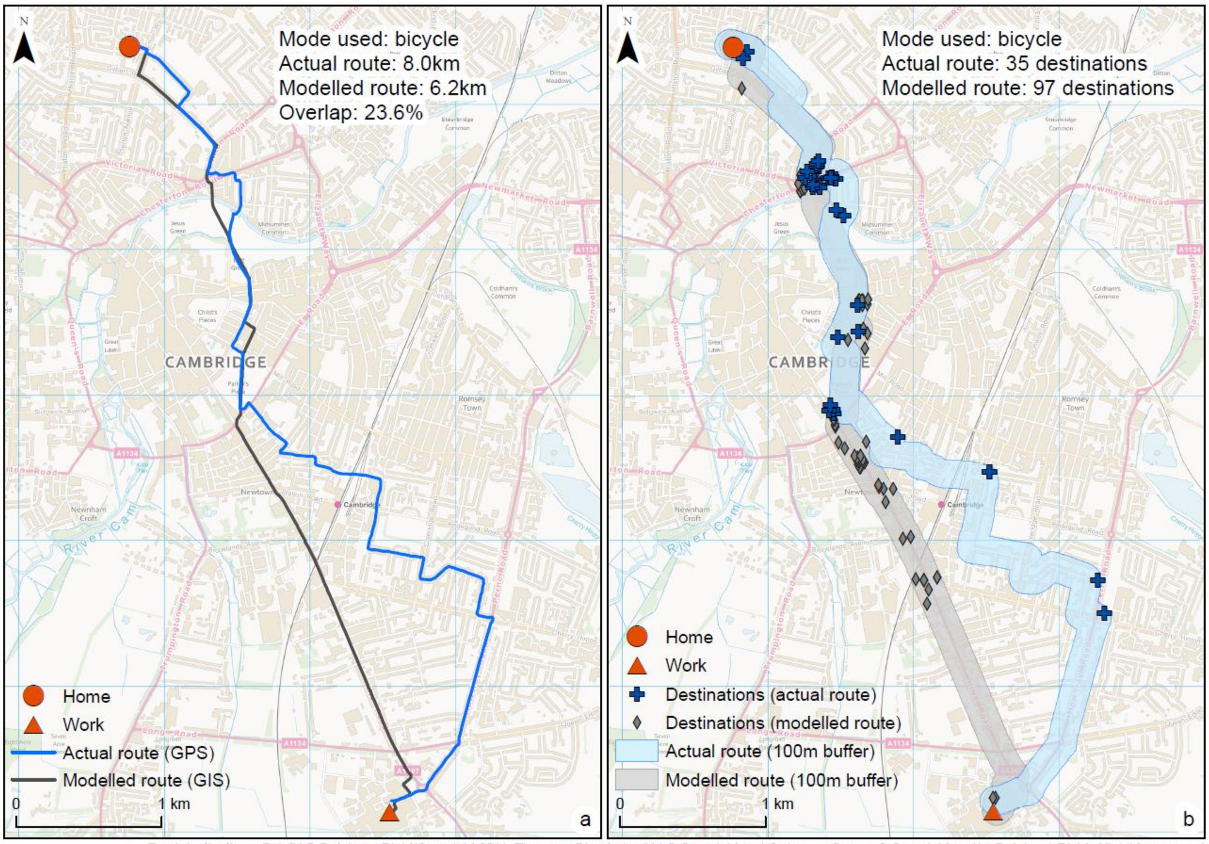

Figure 1. Example of (a) an actual and a modelled journey between home and work for one participant and (b) of the locations of all destinations within a 100m buffer of each route

3. Results

This is a work in progress. Statistical modelling is being carried out to investigate which modelled (GIS) route characteristics can predict goodness of fit of actual (GPS) routes.

Of the 283 trips studied, there was a generally even split between modes, as would be expected based on the selection of cases for this analysis (Table 1). The lower prevalence of walking trips reflects the fact that it was only possible to obtain 24 trips in total from the whole sample. As anticipated based on the study area, mean distances and durations of trips were around 20km and 36 minutes.

In terms of distance from home to work, cyclists, bus users and all car users all followed a route up to 35% longer than the GIS model would suggest, with walkers choosing a slightly shorter route (Table 2). For all except walkers, the actual routes were also less direct. Walkers and cyclists chose quieter routes, where the percentage of A and B roads was 52% and 35% less than on the modelled routes. In terms of destinations en route, all participants chose a route which had fewer retail, food, leisure and education establishments in total, with all car users showing the most pronounced difference.

The greatest similarity between modelled and actual routes, shown by the amount by which the routes overlap, was experienced with bus users, whereas the least similar routes were those using car plus walk (Figure 2). Overall, the actual routes were only found to show overlap with the modelled routes by an average of 33% of their length in our sample.

Energy expenditure would be underestimated for cyclists by an average of 7% if modelled routes were used. Carbon emissions from car journeys would appear to be underestimated by an average of 26% if one were to rely on a GIS estimate of the distance.

Table 1. General trip characteristics of actual routes per mode (n = 283)

Mode	n (trips)	n (people)	Distance (mean, km)	Route directness (mean, ratio)	Duration (mean, minutes)	Actual speed (km/hr per journey)			
						Mean	SE	Min	Max
Bike	50	10	7.5	0.72	25.2	18.3	0.6	11.2	35.1
Bus	52	13	21.0	0.70	46.0	25.7	1.4	9.6	50.9
Car	53	13	26.0	0.65	34.2	47.3	1.8	14.4	73.9
Car+bike	53	9	23.9	0.73	38.4	37.9	1.3	22.2	58.3
Car+walk	51	11	30.6	0.68	46.8	41.0	1.7	14.1	68.7
Walk	24	6	1.3	0.88	14.0	5.6	0.1	4.3	7.2

Table 2. Environmental characteristics of actual and modelled trips (averages, n = 283)

Characteristic	Mode	Actual (GPS)	Modelled (GIS)	Absolute difference	CI (95%)		Diff-erence	P
					Lower	Upper	%	
Route distance (mean, km)	Bike	7.5	6.6	0.90	-0.79	2.59	13.2	<0.001
	Bus	21.0	16.5	4.50	-0.09	9.09	27.1	
	Car	26.0	19.3	6.70	3.12	10.28	35.1	
	Car+bike	23.9	19.1	4.80	2.41	7.19	25.4	
	Car+walk	30.6	24.3	6.30	2.55	10.05	26.0	
	Walk	1.3	1.5	-0.20	-0.97	0.57	-11.3	
Route directness (ratio, 1 = most direct)	Bike	0.72	0.80	-0.08	-0.11	-0.05	-9.2	<0.001
	Bus	0.70	0.86	-0.16	-0.19	-0.13	-18.8	
	Car	0.65	0.85	-0.20	-0.25	-0.15	-22.8	
	Car+bike	0.73	0.90	-0.17	-0.20	-0.14	-18.5	
	Car+walk	0.68	0.86	-0.18	-0.21	-0.15	-21.1	
	Walk	0.88	0.66	0.22	0.14	0.30	33.2	
Busy roads (A&B) on route (mean, proportion)	Bike	0.09	0.13	-0.04	-0.07	-0.01	-34.6	0.014
	Bus	0.64	0.66	-0.02	-0.10	0.06	-3.0	
	Car	0.70	0.70	0.00	-0.09	0.09	0.0	
	Car+bike	0.74	0.76	-0.02	-0.07	0.03	-2.7	
	Car+walk	0.76	0.62	0.14	0.06	0.22	21.6	
	Walk	0.03	0.07	-0.04	-0.09	0.01	-51.5	
Destinations (retail, food, leisure, education) on route (mean, number)	Bike	93.6	101.6	-8.00	-48.06	32.06	-7.9	<0.001
	Bus	159.5	213.3	-53.80	-116.17	8.57	-25.2	
	Car	47.9	136.5	-88.60	-128.71	-48.49	-64.9	
	Car+bike	72.1	130.8	-58.70	-103.56	-13.84	-44.9	
	Car+walk	65.9	192.6	-126.70	-166.21	-87.19	-65.8	
	Walk	11.4	15.8	-4.40	-16.56	7.76	-28.0	

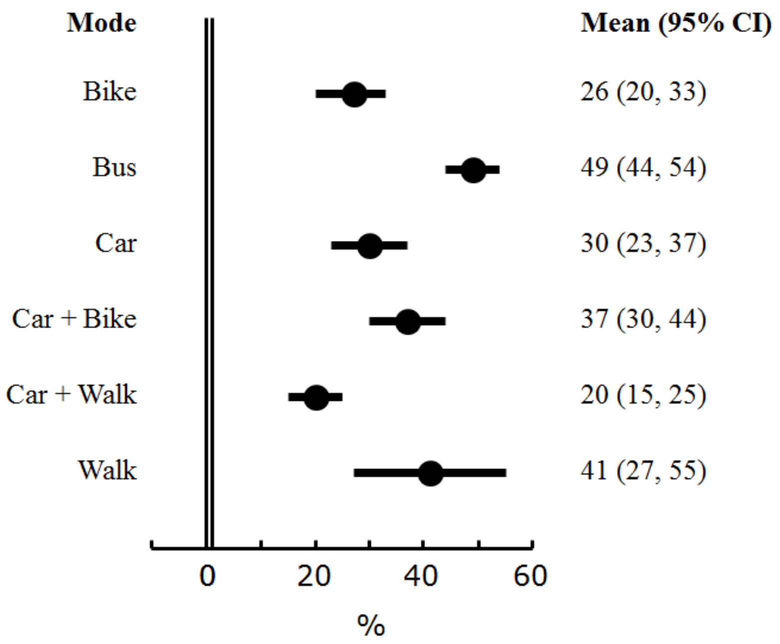

Figure 2. Route overlap between actual and modelled routes (n = 283)

4. Discussion

The results suggest that modelling routes based on the shortest distance between home and work will introduce considerable error into analyses as the modelled routes do not well reflect the characteristics of those actually taken. Generally, walkers may choose a route that is shorter, more direct and significantly less busy in terms of traffic, using cut-throughs and other informal or unmapped routes. Car users, by contrast, may choose routes that are significantly longer, with much fewer destinations than a model would indicate.

This study examined actual routes to work and benefitted from a large, well characterised sample residing in heterogeneous urban and rural environments. In addition, a high prevalence of cycling in the study area allowed us to investigate healthy travel behaviours in particular. However, there are potential limitations with using this type of data. Issues include processing time and data management complications due to file size, occasional lack of agreement between self-report and GPS data, missing data and incomplete (therefore invalid) journeys, and spatial error. In addition, there is a lack of consensus regarding methods of trip extraction from GPS data (Kerr et al. 2011). It is also possible that participants may have modified their usual travel behaviour during the period when they knew their route choices were being objectively monitored.

5. Conclusions

Whilst modelled routes may be appropriately applied in studies wishing to describe general travel characteristics, such as distance between home and work, our findings suggest that their use may not be recommended in those aiming to detail the association between environmental characteristics and the magnitude of participation in health promoting behaviours such as walking and cycling. Researchers should not rely on uncritical assumptions regarding the validity of GIS-modelled routes as a proxy for the actual routes followed by commuters. Where estimates of environmental exposures are required, the use of GPS is recommended despite the resource implications of their application.

Acknowledgements

This project was funded by the National Institute for Health Research Public Health Research (NIHR PHR) programme (project number 09/3001/06). Visit the PHR programme website for more information. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the PHR programme, NIHR, NHS or the Department of Health.

Biography

Alice is a Senior Research Associate with the University of East Anglia. She is interested in the health implications of human interactions with natural/built environments. She has a PhD from the University of the West of England and an MSc from Birkbeck. She previously worked as a consultant for Halcrow.

References

Ainsworth B, Haskell W, Herrmann S, Meckes N, DR. BJ, Tudor-Locke C, Greer J, Vezina J, Whitt-Glover M and Leon A (2011). *Compendium of physical activities: activity categories* [Online]. Healthy Lifestyles Research Center, College of Nursing & Health Innovation, Arizona State University. Available: https://sites.google.com/site/compendiumofphysicalactivities/Activity-Categories [Accessed 23 October 2012].

Bauman AE, Reis RS, Sallis JF, Wells JC, Loos RJF and Martin BW (2012). Correlates of physical activity: why are some people physically active and others not? *The Lancet,* **380**(9838)**,** pp 258-271.

Beyer H (2004). *Hawth's Analysis Tools for ArcGIS.* [Online]. Available: http://www.spatialecology.com/htools [Accessed 21 June 2011].

Defra (2012). *2012 Guidelines to Defra / DECC's GHG Conversion Factors for Company Reporting.*

Dill J (2009). Bicycling for transportation and health: the role of infrastructure. *Journal of Public Health Policy,* **30**(Suppl 1)**,** pp S95-S110.

Duncan MJ, Badland HM and Mummery WK (2009). Applying GPS to enhance understanding of transport-related physical activity. *Journal of Science and Medicine in Sport,* **12**(5)**,** pp 549-556.

Duncan MJ and Mummery WK (2007). GIS or GPS? A Comparison of Two Methods For Assessing Route Taken During Active Transport. *American Journal of Preventive Medicine,* **33**(1)**,** pp 51-53.

Fraser SDS and Lock K (2010). Cycling for transport and public health: a systematic review of the effect of the environment on cycling. *The European Journal of Public Health,* **21**(6)**,** pp 738-743

Kerr J, Duncan S and Schipperijn J (2011). Using global positioning systems in health research: a practical approach to data collection and processing. *American Journal of Preventative Medicine,* **41**(5)**,** pp 532-540.

Maddison R and Ni Mhurchu C (2009). Global positioning system: a new opportunity in physical activity measurement. *International Journal of Behavioral Nutrition and Physical Activity* [Online], **6**,(73), doi:10.1186/1479-5868-6-73 [Online 4 November 2009].

Ogilvie D, Griffin S, Jones A, Mackett R, Guell C, Panter J, Jones N, Cohn S, Yang L and Chapman C (2010). Commuting and health in Cambridge: a study of a 'natural experiment' in the provision of new transport infrastructure. *BMC Public Health* [Online], **10**,(703), 10.1186/1471-2458-10-703 [Online 16 November 2010].

Oliver M, Badland H, Mavoa S, Duncan M and Duncan J (2010). Combining GPS, GIS, and Accelerometry: Methodological Issues in the Assessment of Location and Intensity of Travel Behaviors. *Journal of Physical Activity and Health,* **7**(1)**,** pp 102-108.

Panter J, Griffin S, Jones A, Mackett R and Ogilvie D (2011). Correlates of time spent walking and cycling to and from work: baseline results from the commuting and health in Cambridge study. *International Journal of Behavioral Nutrition and Physical Activity* [Online], **8**,(124), 10.1186/1479-5868-8-124 [Online 10 November 2011].

Panter JR and Jones AP (2010). Attitudes and the environment as determinants of active travel in adults: what do and don't we know? *J Phys Act Health* **7**(4)**,** pp 551–561.

PointX Ltd (2010). *Ordnance Survey Points of Interest* [Online]. Southampton: OS. Available: www.ordnancesurvey.co.uk/oswebsite/products/points-of-interest/index.html [Accessed].