<u>Determining Cycle Mode Choice for Commuting using New Route Analysis Methods</u>

(Richard Thomas, University of Leeds GIS MSc 2013-14)

Introduction

Previous research has shown that the likelihood of choosing to commute by bicycle can be very dependent on the availability of acceptable cycle routes of reasonable distance, avoiding hills and with a minimum of cycling in motorized traffic. Commuting by bicycle is a good proxy for general utility cycling with the UK National Travel Survey (NTS) showing a significant correlation, with a ratio of 1:0.77 averaged over 2002-2010 (Goodman, 2013). Considering cycle ***routing*** specifically is important: the location of cycling infrastructure can be key to whether it is effective in encouraging cycle commuting. Although various analyses take into account general measures of infrastructure provision (for example Parkin et al, 2008; Schoner and Levinson, 2013), it has been recommended that specific commuting routes should be analysed as characteristics of routes have a strong influence on cycle commuting uptake (Schoner, 2013, p48). Knowledge of existing routes can inform policy on locating new cycle infrastructure, but it can also highlight areas where infrastructure is adequate but cycle uptake is poor - indicating a policy need to address other determinants. Transport simulation models have historically focused on motorized traffic and thus often do not have enough detail at a small street level to be representative for modelling non-motorized traffic.

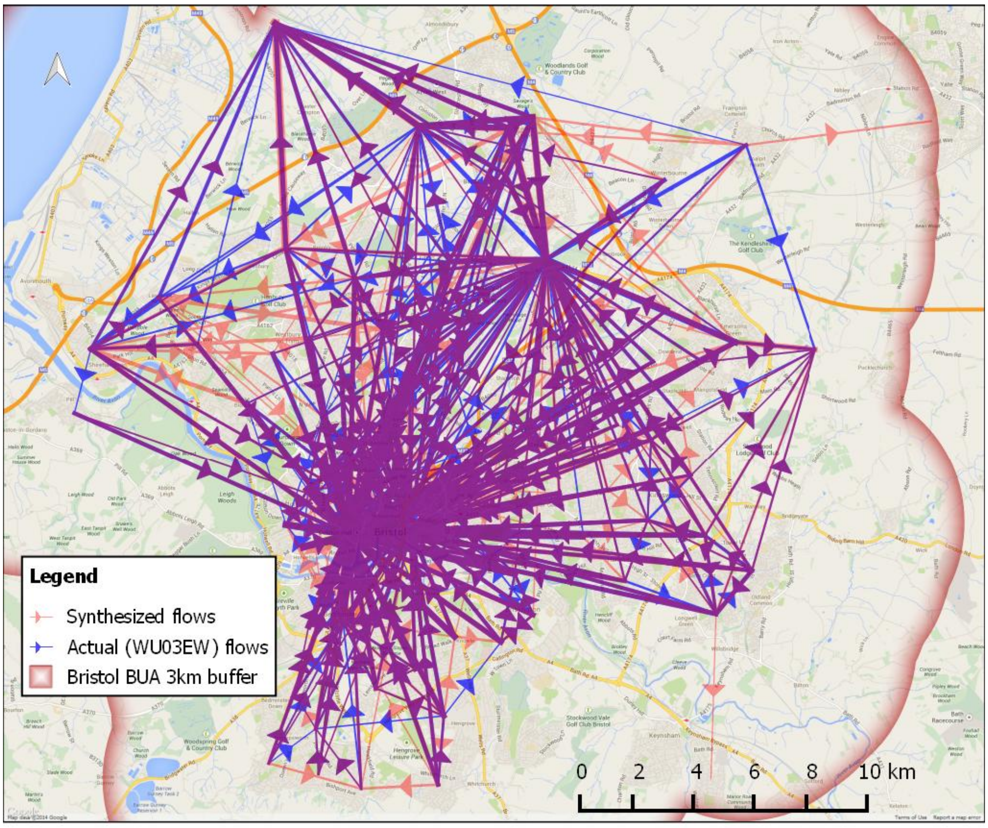
Taking the UK city of Bristol as a basis for analysis, this study used a third party cycle-routing engine (CycleStreets.net based on the OpenStreetMap data) to generate representative commuting routes and from there to generate 'cost functions' that characterise those routes. Routing was done between Population Weighted Centroids (PWCs) of the lowest census levels of geography: between Output Areas (OA) origins and Workplace Zone (WZ) destinations. MSOA-level 2011 census origin-destination commuting data (all modes of transport) became available in July 2014, but initial analysis done with this data suggested that MSOA PWCs were too coarse a spatial resolution to give representative routes. Unfortunately, 2011 OA-level commuting origin-destination census flow data will not become available until the end of 2014 and will even then only be accessible within a 'secure' environment, thus restricting its use. Thus the other new method of this study was to synthesize representative OA-level commuting flows using the constraints of other available OA-level census data.

Routing results were subsequently aggregated based on known commuting origins to give area-based measures which could then be used to build more traditional area-based aggregate regression models which would address the likelihood of whether for each area, local workers are likely to commute by bicycle, given typical locations commuted to from their area

Synthetic Origin-Destination Flow Data Generation

A microsimulation approach was considered, but the characteristics to be matched were entirely spatial (location and distance) rather than the (typically) socio-economic ones used in microsimulation so it seemed a less than ideal fit. However, some of the techniques of using randomised iteration not always restricted to lower cost solutions were adopted from that field. Unlike typical synthesis methods used for motorized traffic, there was little requirement for modelling dynamic network loading in response to traffic counts. Therefore a much simpler origin-destination synthesis could be used; a methodology similar to the SHAPE-2 gravity model of Barbour and Fricker (1994) was adopted. The resulting new Java software 'GenSynthFlow' used constraints data found in the 'Distance travelled to work ' 2011 census tables QS702EW (by residential location) and WP702EW (by workplace), with population and route synthesis done for one distance interval at a time (e.g. 2-5km). A gravity model was incorporated based on the assumption that the number of people travelling to a workplace will reduce with increased distance.

Because commuting would occur across the boundaries of any area modelled, it would not be expected to get a perfect fit to all the commuter counts. However, by modelling a larger area (4 complete Unitary Authorities covering most of the 20km maximum range required) than that of interest (Bristol Built Up Area), these errors should be minimized in the final analysis.



Finding a suitable statistical measure for overall validation of the synthesis against census MSOA origin-destination data proved difficult. However, comparison with actual MSOA-level data from 2011 census table WU03EW (figures 1 and 2) shows that synthesis should be reasonably representative for the routing requirements of this study (purple routes indicate where synthesized and actual flows coincide).

Figure 1: MSOA-level commuter flows larger than 100 within 3km buffer of Bristol Built Up Area (BUA)

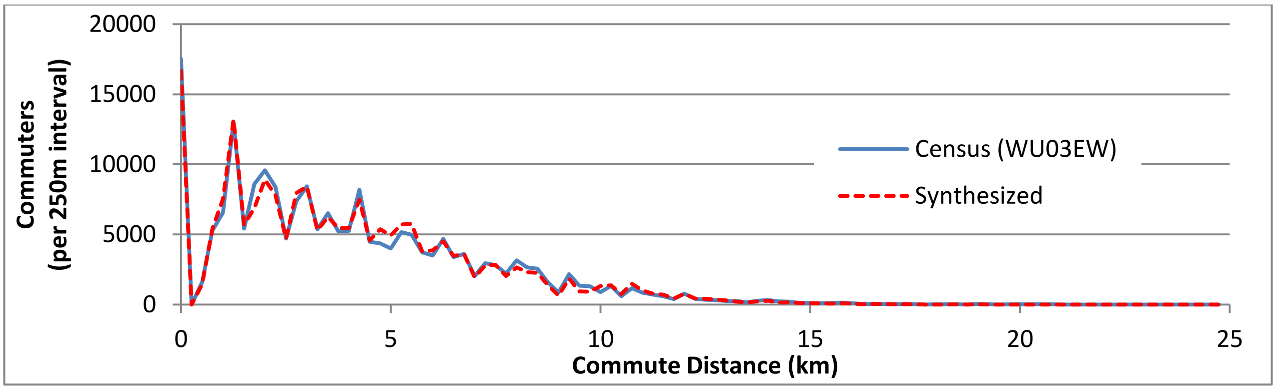


Figure 2: Distance distribution of synthesized and actual MSOA flow data within 3km buffer of BUA

Routing Cost Functions

Cycle-specific routing was performed using the CycleStreets.net (2014) 'balanced' routing algorithm to produce routes that would take advantage of cycle infrastructure and avoid busy roads, but only where it did not introduce too much of a detour. Given the large number of OAs within the BUA and the number of commute routes from each, it was decided to limit routing to the four most popular commutes (by any means) for each origin area. Routes were also limited to those with destinations within a 3km buffer around the BUA to avoid excluding popular cycling routes on the fringes of the BUA but ensure that cost functions largely reflect characteristics of the BUA. A minimum route distance of 1 km was used to exclude routes more likely to be walked.

Several ‘cost functions’ were derived for each route and the results aggregated by origin area to build a cross-sectional linear regression model for cycle mode choice. Of these cost functions, three were selected (in addition to route distance) based on lack of collinearity and potential interest:

- **‘Directness’** (ratio of Euclidean to Routed commute distance) made a notable contribution to the final regression model and could possibly be further improved by replacing the Euclidean distance measure with a ‘fastest’ routed distance.
- **‘EffortRatio’** (estimated energy exerted divided by routed distance) intuitively should be a better determinant of cycle mode choice than just a measure of hilliness (a documented cycle commuting detractor) as it also incorporates energy expended at likely stop/start locations. However, it did not make a useful contribution to the regression model - possibly because stop/start behaviour might be more of a hindrance to car drivers (certainly in its impact on speed). As the routing engine used returns elevation measures at many points along the route, reworking with a derived direct measure of hilliness might be more useful.
- **‘Traffic’** (likelihood of cycling in motor traffic – see figure 3) proved representative of the level of cycle infrastructure, but was not helpful in cycle mode choice regression. This could indicate that infrastructure has little impact on cycle mode choice, but the measure gives little indication of how much the infrastructure is actually required at each location (it is always easier to install infrastructure in less congested streets). Thus, although the measure includes a rating for general types of road, it might be improved by combining it with estimated motor traffic counts from each road taken from established transport models based on actual traffic count data.

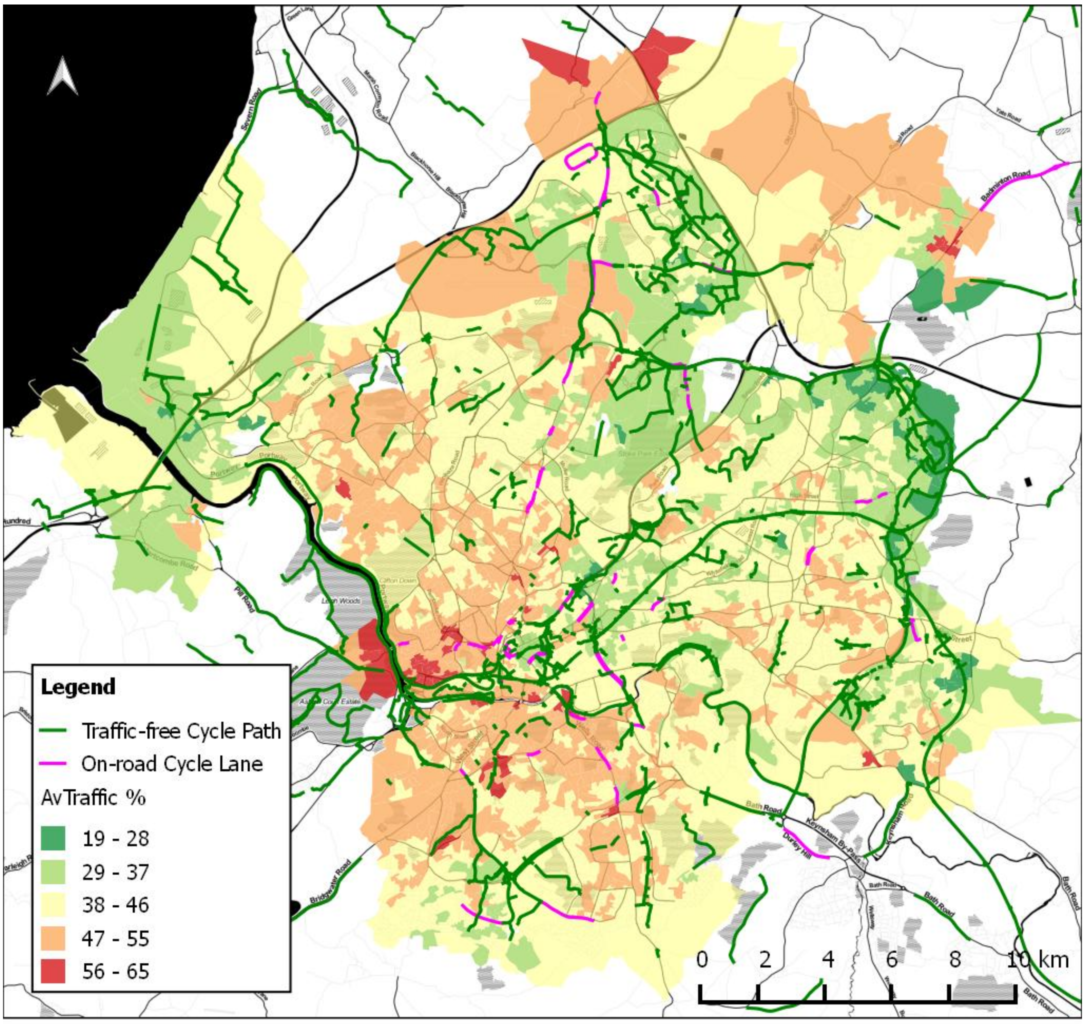


Figure 3: Mean traffic exposure measure by OA for commutes less than 20 km from Bristol BUA

Conclusions

Although all the routing cost functions were statistically significant, only 'Directness' made a useful contribution to the fit of the final OA-based regression model – one that was greater than car ownership or population density (both of which are noted in the literature as useful determinants).Results might be improved with a hybrid approach: calculating routing costs at OA level (using improvements to cost functions outlined above) perhaps on a large number of routes, then aggregating these results up to a higher (LSOA/MSOA) level in order to take advantage of the better overall model fit (with other data) achievable at those levels. Although this study used an area-based aggregate data model, the full potential power of routing-based cost functions might lie in non-aggregate modelling where each individual's commuting (or general utility) cycling origin-destination locations are known.

It is clear from this study that consideration of the effects of fine-grain route details tallied with realistic commuting flows is important. An additional output from this study's methods was the generation of maps of summed routes of simulated cycle commuter flows. These should give a good representation of potential cycle network hotspots where it would be worthwhile to target provision or upgrading of suitable cycle infrastructure. To relate such a function to the city of the analysis: in the recently published draft Bristol Cycle Strategy (Bristol City Council, 2014) there is a strong emphasis on the provision of a coherent network of ***direct*** cycle 'freeways' specifically aimed at efficient commuting and utility cycling. Knowing potential hotspots for cycle commuting based on likely commuting flows could thus be helpful in designing and staging the implementation of such a network.

References

Barbour, R. and Fricker, J.D. 1994. Estimating an origin-destination table using a method based on shortest augmenting paths. *Transportation Research Part B: Methodological.* **28**(2), pp.77-89.

Bristol City Council. 2014. *Bristol Cycling Strategy (Draft July 2014).* [Online]. [Accessed 25 August 2014]. Available from: https://www.citizenspace.com/bristol/city-transport/cycle-strategy

CycleStreets.net. 2014. *Journey Planner API (Application Programming Interface).* [Online]. [Accessed 25 August 2014]. Available from: http://www.cyclestreets.net/api/v1/journey/

Goodman, A. 2013. Walking, cycling and driving to work in the English and Welsh 2011 census: trends, socio-economic patterning and relevance to travel behaviour in general. *PloS one.* **8**(8), pe71790.

Parkin, J.Wardman, M. and Page, M. 2008. Estimation of the determinants of bicycle mode share for the journey to work using census data. *Transportation.* **35**(1), pp.93-109.

Schoner, J.E. 2013. *Catalysts and Magnets: built environment effects on bicycle commuting*. [Online]. Thesis, University of Minnesota. [Accessed 25 August 2014]. Available from: http://nexus.umn.edu/

Schoner, J.E. and Levinson, D.M. 2013. The Missing Link: Bicycle Infrastructure Networks and Ridership in 74 US Cities. In: *Transportation Research Board 92nd Annual Meeting*.

The full dissertation and associated Java software is available online at:
http://richard-thomas.github.io/GIS_MSc/