Biodiversity and healthy cities: examining the associations of geographical characteristics with cycling behaviours using Strava Metro data in Newcastle, UK

Home, P.[11], Boyd, D.[2] Li, S.[22]

[1]Nottingham Geospatial Institute, University of Nottingham, Jubilee Campus, Lenton, Nottingham NG7 2TU
[2]School of Geography, University of Nottingham, University Park, Nottingham, NG7 2RD

GISRUK 2023

Summary
Sustainable urban cities contribute positively to the promotion of active travel. Studies have shown the increasing access to greenspace improves health, however little is known about the impacts of biodiversity on active travel. We hypothesise an increased degree of biodiversity will be positively associated with increased cycling volume. We combine several environmental, ecological, and urban form datasets to understand the association between environmental characteristics including biodiversity and recreational cycling in Newcastle. Recreational cycling data was acquired from the Strava Metro dataset in 2018. After data processing, a linear mixed-effects model will be used to determine significant associations between cycling volume and independent variables. This research will contribute to our understanding of the contribution of nature and biodiversity towards public health, pertinent due to the ongoing biodiversity crisis and growing urban populations.

KEYWORDS: cycling, greenspace, Strava, crowdsourced geographic information, biodiversity

1. Background

Through the lens of designing sustainable and healthy cities, the promotion of active travel both occupationally and recreationally is a high priority for urban planners. The physical health benefits of cycling are numerous, reducing risk of coronary heart disease, high blood pressure and stroke (Götschi et al., 2015). Other benefits include reduced air pollution and improved traffic congestion (Taddei et al., 2015). Therefore, designing urban form that promotes active travel usage and elicits positive behaviour change is of importance to several stakeholders. Associations between cycling behaviour are very useful in contextual city planning and evidence-based decision making. Previous studies have shown both positive and negative bicycle usage associations with population density, street connectivity and motorized traffic volume (Sun et al., 2017; Griffin & Jiao, 2015; Zhao, 2014).

Greenspace in urban areas offer a myriad of ecosystem services including promoting recreational behaviours and active travel; and facilitate biodiversity (Figure 1). The Convention on Biological Diversity (CBD) and the World Health Organization (WHO) are collaborating to promote awareness of the influence of biodiversity on human health. Active travel mobility associations with some environmental characteristics including biodiversity are not yet fully understood, particularly in the UK. This case study will explore how social, environmental and biodiversity characteristics influence cycling behaviour in Newcastle.

[1] Philip.home@nottingham.ac.uk
[2] sabrina.li@nottingham.ac.uk

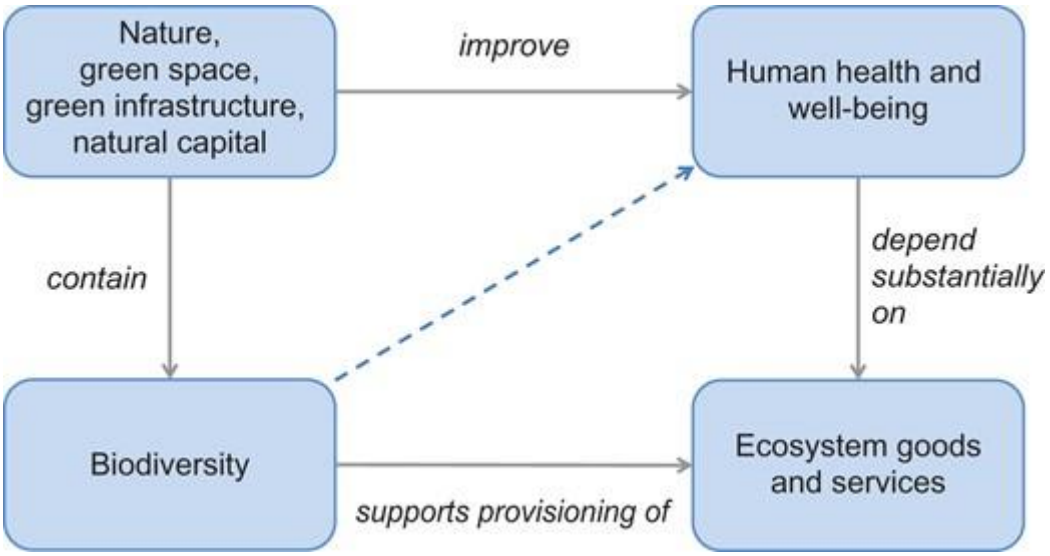

Figure 1: The linkages between human health, biodiversity and greenspace (Aerts et al., 2018).

2. Methods

2.1 Data

Cycling behaviour will be characterised by Strava Metro data, collected by the mobile fitness tracking platform, which records GPS trajectories and coordinates through time. The platform allows classification of activity including cycling, running and walking, which can be labelled as a recreational or commuting activity. Strava Metro is a cleaned and aggregated dataset that produces mobility data into three units: Street segments (edges), Intersections (nodes) and polygons of trip destination and origin (OD). Streets and Nodes data are created from the road network from OpenStreetMap. Strava Metro makes this dataset available to urban planners and active travel advocates to improve infrastructure for cyclists and pedestrians. Strava Metro data can accurately describe cycling behaviour. Firstly, Strava data is almost gapless over both time and space as long as users exist (Lee & Sener, 2021). Second, several studies have shown Strava data to be an indicator of the behaviour of the wider cycling population through high correlations with traditional cycling count data (Lee & Sener, 2021; Sun et al., 2017). Third, compared with typical count data, Strava has high spatial granularity and spatial/temporal coverage.

Strava Metro data characterising street-level cycling behaviour in Newcastle in 2018 has been acquired from the Urban Big Data Centre based at the University of Glasgow. Table 1 lists the attributes of the Streets file including volume at each edge (street) at a time. Figure 2 illustrates the pervasiveness of Strava usage across the North East of England.

Table 1: Fields in Streets file (Source: Sun et al., 2017)

Field	Description
Edge_id	Unique and permanent street identification number
Year	Numerical year format (yyyy)
Day	Numerical day format (1-365)
Hour	Numerical hour format (0-24)
Minute	Numerical minute format (0-59)
Count_Ride	Count of all purpose cycling trips (regardless of unique cyclists on section of street for the day, hour and minute.
Commute_Count_Ride	Count of commuting cycling trips (regardless of unique cyclists on section of street for the day, hour and minute.

Recreation_Count_Ride	Count of recreational cycling trips (regardless of unique cyclists on section of street for the day, hour and minute.

Figure : Strava Metro Global Heatmap of Newcastle Upon Tyne. Source: (Strava Metro, 2023)

The UK Department for Transport (DFT) collects annual average daily flow (AADF) data covering several roads in Newcastle upon Tyne. To validate Strava Metro bicycle data for acceptable representation of the wider cycling population, the Pearson correlation between bicycle counts on matched streets within both datasets will be calculated (Sun et al., 2017; Lee & Sener, 2021).

Auxiliary datasets describing environmental characteristics have been chosen based on literature and past studies including socio-economic data extracted from census collections. Landuse/greenspace and biodiversity metrics are still under consideration. Finally, traffic factors including motorised vehicle volume and accident density are integrated. Table 2 describes independent variables and their sources used in this study.

Table 2: Independent variables considered in this study.

Independent Variable	Variable Category	Source
Road Class	Road factors	OpenStreetMap
Road Length (km)		
Road Connectivity		
Population density	Socio-economic factors	Office for National Statistics
Employment density		
Distance to city centre	Urban fabric factors	
Distance to nearest bus stop		
Land use mix	Land use and greenspace factors	UK Centre for Ecology and Hydrology
Contiguity with greenspace		
Dominant Land use		
Biodiversity		National Biodiversity Network Natural History Museum
Traffic accident density	Traffic factors	Department for Transport
Volume of motor vehicles		

2.1 Characterising recreational cycling behaviour

To characterise recreational cycling behaviour and trips at the street level, the recreational cycling rate index (RCR) as devised by Sun et al. (2017) is defined as:

$$RCR(s,t) = \frac{trip\_cnt^{R}(s,t)}{trip\_cnt^{R}(s,t)trip\_cnt^{C}(s,t)} \quad (1)$$

Where trip_cnt^{R}(s, t)and trip_cnt^{C}(s,t) represent the respective number of recreational and commuting trips on streets during the time slot on all workdays. Independent variable data manipulation in GIS software is required to coerce data into useable format or preparing for linear model from sourced data.

2.2 Linear mixed effects model

To examine the association of environmental and biodiversity characteristics with cycling data, a linear mixed effects model taking into account both random and fixed effects will be used in this study. Predictor variables are shown in Table 2. For the sake of simplicity within the scope of this project, only one random effect is accounted for. i.e., the intercept, using the street as the group. This was selected as it is assumed some characteristics may influence Strava users that are invisible to available data, such as contextual characteristic and traffic calming measures. Figure 2 illustrates the designated workflow for the study.

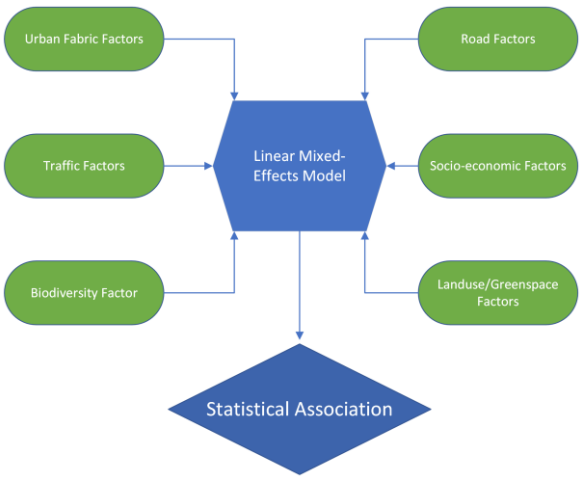

Figure 2: Association study workflow.

3. Evaluation

Data collation and curation are ongoing. The first stage has been to assess literature to identify environmental, biodiversity and social predictor variables. Simultaneously, various biodiversity data and metrics are being compared and evaluated for appropriateness to this study including biodiversity intactness, Shannon and Simpson Indexes. Once identified, GIS manipulation and linear modelling will be performed to identify associations and significance.

4. Conclusions and Future Research

- This is part of an ongoing research project, where outputs could contribute towards the wider evidence base of the impacts of greenspace and biodiversity on public health. Implications of this work include furthering past works on cycling associations, the importance of environmental characteristics in effective active travel design. It is important for local policy

makers to promote cycling in urban areas, reducing traffic congestion and encouraging physical activity.

- The project uses largely open datasets, meaning techniques can be applied to other areas of interest for future comparative studies. Comparison in this way will help indicate the contextual nature of aggregated cycling behaviour in different urban areas to best arm decision making in local context urban design to promote cycling behaviour.
- Presenting at GISRUK would enable further scrutiny and feedback, enabling an early career researcher such as myself to develop knowledge and skills in urban analytics, which would help to improve this project.

5. Acknowledgements

This research was funded by EPSRC Geospatial Systems Centre for Doctoral Training. This project has also benefited from data obtained from the Urban Big Data Centre. Road and urban form data taken from OpenStreetMap: Map data copyrighted OpenStreetMap contributors and available from https://www.openstreetmap.org.

References

Aerts, R., Honnay, O. & van Nieuwenhuyse, A. (2018) 'Biodiversity and human health: mechanisms and evidence of the positive health effects of diversity in nature and green spaces', *British Medical Bulletin*, 127(1), pp. 5–22.

Götschi, T., Garrard, J. & Giles-Corti, B. (2015) 'Cycling as a Part of Daily Life: A Review of Health Perspectives', *https://doi.org/10.1080/01441647.2015.1057877*, 36(1), pp. 45–71.

Griffin, G.P. & Jiao, J. (2015) 'Where does bicycling for health happen? Analysing volunteered geographic information through place and plexus', *Journal of Transport & Health*, 2(2), pp. 238–247.

Lee, K. & Sener, I.N. (2021) 'Strava Metro data for bicycle monitoring: a literature review', *Transport Reviews*, 41(1), pp. 27–47.

Strava Metro (2023) *Strava Metro | Map*. [Online] [online]. Available from: https://metroview.strava.com/map/demo#eyJtIjoic3RyZWV0cyJ9 (Accessed 17 January 2023).

Sun, Y., Du, Y., Wang, Y. & Zhuang, L. (2017) 'Examining Associations of Environmental Characteristics with Recreational Cycling Behaviour by Street-Level Strava Data', *International Journal of Environmental Research and Public Health 2017, Vol. 14, Page 644*, 14(6), p. 644.

Taddei, C., Gnesotto, R., Forni, S., Bonaccorsi, G., Vannucci, A. & Garofalo, G. (2015) 'Cycling Promotion and Non-Communicable Disease Prevention: Health Impact Assessment and Economic Evaluation of Cycling to Work or School in Florence', *PLoS ONE*, 10(4), .

Zhao, P. (2014) 'The Impact of the Built Environment on Bicycle Commuting: Evidence from Beijing', *Urban Studies*, 51(5), pp. 1019–1037.

Biographies

Philip Home is a first year PhD student at the University of Nottingham and Newcastle as part of the EPSRC Geospatial Systems CDT. Philip has a BSc in Environmental Science, an MSc in Risk and Disaster Science from the University College London and an MRes in Geospatial Data Science from the University of Nottingham. His research interests include data visualisation/storytelling and using spatial data to inform effective decision making.

Dr. Sabrina Li is Assistant Professor in Quantitative Geography at the University of Nottingham. Sabrina's research investigates the interactions between human health and the physical, social, and built environments. In particular, she is interested in understanding why certain populations are more susceptible to diseases and ill health than others, and the assessment of differential health impacts experienced by disadvantaged populations. Most of her work focuses on the social and environmental determinants of health.

Prof. Doreen Boyd is Professor of Earth Observation with research interests is in using cutting-edge geospatial data science on standard and novel data streams. The principal data stream is that captured by remote sensing systems. Applications cut across a range of UN SDGs, attracting over £10 million in research funding.