AGILE: GIScience Series, 3, 15, 2022. https://doi.org/10.5194/agile-giss-3-15-2022
Proceedings of the 25th AGILE Conference on Geographic Information Science, 2022.
Editors: E. Parseliunas, A. Mansourian, P. Partsinevelos, and J. Suziedelyte-Visockiene.
This contribution underwent peer review based on a full paper submission.

The Impact of Built Environment on Bike Commuting: Utilising Strava Bike Data and Geographically Weighted Models

Hyesop Shin[1], Costanza Cagnina[1], and Anahid Basiri[1]

[1]School of Geographical and Earth Sciences, University of Glasgow, G12 8QQ, UK

Correspondence: Hyesop Shin (Hyesop.Shin@glasgow.ac.uk)

Abstract.

Active travel provides significant public health benefits including improving physical and mental health and air quality. Given the geography of congested roads, availability of required infrastructure and cost of transportation in cities, promoting active travel, including cycling, can be a good solution for commuting within built environments. Having a better understanding of the key drivers that may influence bike ridership can help with designing cities that accommodate cyclists' needs for healthier citizens. This paper examines the built environment features that may affect commuting cyclists. We respectively employ Ordinary Linear Square (OLS) regression and Geographically Weighted Regression (GWR) for 136 Intermediate Zones of the city of Glasgow, UK. The results of GWR show that the significant local variation in green areas suggests that even though the global regression showed a negative association between the greenness and commute cycling, over half of the IZ areas had a strong positive association with the green areas. Building height and Public Transport Availability Index show geographic patterns where the residuals are fairly stationary across the study area with some clusters of high residuals. Performance wise, the results from GWR provided an R^2 of 0.73 which was higher than OLS at 0.3. Our results can provide insights into how to use crowdsourced cycling data when there are spatially and temporally limited resources available.

Keywords. Active Travel, Geographic Weighted Regression (GWR), Strava Data, Built Environment, Crowdsourced Data

1 Introduction

Active travel provides important benefits to public health by increasing physical activity (Doorley et al., 2015). According to Obesity Action Scotland, 28% of the adult population in Scotland is obese, 65% overweight or obese, and 16% of children aged 2-15 are at risk of obesity. As such, physical activity has been highly recommended as a factor of positive weight management, along with providing established benefits on mental health (Je et al., 2009).

One of the first steps to promote active travel is to understand the behaviours of individuals and patterns of their travels. Studies have shown that active travellers, especially (commuting) cyclists, were more motivated to commute where the built environments, such as wider bike lanes (Sun et al., 2017b), more green areas (Winters et al., 2010), convenience to workplaces (Zhao, 2014; Sarjala, 2019; Sun et al., 2017b), and connections to public transport (Zhao, 2014), were more advanced. Considering many cities have attempted to improve their bikeability, it is fundamental to conduct case studies to develop an understanding of the most influential factors in each area, as to develop tools to aid decision making and the evaluation of the implementation of investment schemes (Aldred, 2019).

Many studies which investigated the influence of the built environment on citizens' choice of transport rely on surveys and questionnaire data or cycling counts at fixed spots (Song et al., 2013; Aldred et al., 2019; De Vos et al., 2019; Echiburú et al., 2021). However, these studies have limitations for their small sample sizes as well as a short period of examination to sufficiently understand the spatial and temporal trends of cycling. Some studies have used smartphone devices to capture their route choices via GNSS data (Chen et al., 2018; Khatri et al., 2016; Lu et al., 2018; Scott et al., 2021; Ta et al., 2016). However, the platform for data gathering was computationally expensive for users to understand the logic and upload their trajectories, which, for these studies, resulted in a smaller sample size.

The popularity of activity tracking apps has increased to the point that people voluntarily share data on their recreational and utilitarian movements. Crowdsourced data such as this, also called volunteered geographic information (VGI), is often more spatially and temporally comprehensive than traditionally acquired data and encompasses a much larger user group (Ferster et al., 2018). The fitness tracking app Strava is one of the most popular in its genre, and it is becoming increasingly popular amongst researchers and urban planners as well due to the Strava Metro service, a data service that aims to enable the harnessing of Strava data for urban planning improvement (Strava, 2021).

This paper examines the built environment factors that affect commuting cyclists. This paper will achieve two objectives: i) to explore the travel patterns of Strava data between 2017 and 2018 with respect to the built environment features in Glasgow; ii) to compute the strength of the built environmental characteristics against cycling.

2 Related Works

The positive impacts of active travel (AT) for the community, environmental health and individual health are well recognised by governing and public health bodies (Doorley et al., 2015). An increase in active travel encouragement has led to growing yearly investments - the latest reaching £338 million[1]. According to the government report, around 34% of car journeys in Scotland are less than 4km long and could be done by walking or cycling instead. Active Travel Scotland published in 2014 a plan named 'A long-term vision for active travel in Scotland' [2] envisioning what an active travel optimised country looks like, being infrastructure one of the key features.

2.1 Environment and Socioeconomic Characteristics and Cycling

The urgency to understand how cycling infrastructure affects behaviour has been motivated by the growing awareness of active travel benefits. Neighbourhood characteristics positively influence the tendency to cycle(Song et al., 2013; Macmillen and Stead, 2014; Aldred, 2019). De Vos et al. (2019) found the presence of appropriate infrastructure to drastically increase travel satisfaction in active travellers. Montreal saw a significant increase in cycling likelihood after years of safe-road enhancing networks (Zahabi et al., 2016). However, a study led over three selected sites in the UK, found that infrastructure alone may not be sufficient (Song et al., 2017). The cycling network is used as a variable in this analysis in two forms: segregated paths and paths shared with traffic. Road class was found to be strongly influential in cyclists' route choices in (Sun et al., 2017b) and (Le et al., 2021), therefore a bespoke classification of the road network is also included in the models. The relationship between cycling routes and land use has been observed to lean in different directions. Generally, the proximity to greenery is a key factor in studies that look at active travel mode share. Park and Akar (2019) found that their participants preferred routes surrounded by a low grade of the land-use mix; Zhao (2014) found that a mixed environment is preferred in Beijing, while Mäki-Opas et al. (2016) concluded that a high proportion of recreational green space contributed negatively to the levels of active commuting, whereas the presence of vegetation was the most important environmental characteristic in Van Holle et al. (2014). The aesthetic experience of the route was decisive for the participants in Stefansdottir (2014), especially in terms of greenery and distance from motorized traffic. In this study, in addition to the surrounding green spaces, it was decided to include the building height variable, attempting to consider the aesthetic quality of the most travelled routes. As per the socioeconomic aspect, the percentage of the unemployed population per output area is considered. People in employment have the necessity to commute, which paired with stable income often leads to higher usage of motorized vehicles (Brand et al., 2014). More recent research indicated that people who received higher education and are in employment develop 'greener' habits and are starting to move away from car use (Dadashova et al., 2020; Le et al., 2021; Song et al., 2017; Sun et al., 2017b). In other instances, the employment rate seems to have a non-zero correlation with cyclists' chosen routes (Sun et al., 2017b). Some studies looked at the correlation with public transport availability and revealed that high levels of public transport accessibility are associated with higher bicycle usage presumably due to high population density (Winters et al., 2010, 2011). These differences in findings can be motivated by the variety in study areas, cultural differences and other non-controlled factors. Few of these studies considered all variables concurrently, therefore the observed associations may have been influenced by unforeseen variables.

2.2 Cycling data

Many studies have analysed the influence of infrastructure and/or the built environment on cyclists' route choice through questionnaires, surveys or census data (Mertens et al., 2016; Echiburú et al., 2021; Ek et al., 2021). Lee and Moudon (2008) analysed respondent survey data and their results highlighted good street lighting to be the most important factor for choosing cycling, and traffic volume to be the most significant obstacle. Mertens et al. (2016)'s respondents identified cycle path type as the most important environmental factor. However, GPS-based solutions have become increasingly applied, with some mixed-approach exceptions. Krenn et al. (2014) investigated the association between the built environment characteristics and bicycling for transportation using GPS data of 70 participants and digitized routes acquired from a survey. Broach et al. (2012) provided the participants with handheld GPS devices to track their utilitarian trips and focused on bicyclists' preferences for facility types, e.g. route distance, turn frequency, intersection control, road type. Sarjala (2019) explored the relationship between the built environment and active travel in Tampere, Finland, with GPS

[1]https://www.gov.uk/government/news/338-million-package-to-further-fuel-active-travel-boom

[2]https://www.transport.gov.scot/publication/scottish-government-s-long-term-vision-for-active-travel-in-scotland/

enabled smartphones. Many authors employed data from the fitness app Strava, which raised the question of whether Strava data represents the demographic and route characteristics of the total cycling population. Strava data is crowdsourced data, and as such offers large volumes of data with the potential to be collected at the time and location that the apps are used; however, this type of data is limited by the target market of the provider company. For example, Griffin and Jiao (2019) stated that 3%–9% of bicycle trips counted on trails in Austin used Strava at the time of the count. However, this percentage can change based on the study area, road network availability, socioeconomic, seasonal, and demographic factors (Conrow et al., 2018; Jestico et al., 2016). Previous studies have explored the potential of crowdsourced data to predict real cycling volumes or in relation to other influential factors. Jestico et al. (2016) employed linear regression to quantify the strength of the association between crowdsourced data and manual cycling counts in Victoria, British Columbia, concluding that while representing a small user group, Strava data has the potential to predict actual cycling counts, especially in urban areas. Kwayu et al. (2021) found that this type of fitness crowdsourced data represents a biased set. Lin and Fan (2020) quantified the relationship between manual cycling count, Strava count and other relevant variables and proved Strava's quality assurance.

3 Methods

3.1 Study Area

The city of Glasgow has been chosen as a study area as the city has experienced a significant increase in cycling up to 11,000 cyclists per day, which is a 111% rise over 9 years (Sustrans, 2018). In addition, Glasgow has actively collected cycling information in tribute to the city council investments over the past 15 years, which resulted in a variety of findings related to weather, distance, recreational cycling, and built environment (Hong et al., 2020a, b; McArthur and Hong, 2019; Sun et al., 2017b).

3.2 Data and Software Availability

Cycling: The University of Glasgow's Urban Big Data Centre provided Strava data for 2017 and 2018 (https://www.ubdc.ac.uk). The data contains three products: Nodes, Edges, Origin/Destination (OD). The Edges product, based on the road network extracted from OpenStreetMap, contains each street segment on which activity was recorded with hourly measures. Nodes and OD were not used. The Strava app lets the user classify the activity as 'recreational' or 'commute', which then appears as a field in the products. In this analysis, the total activity count is used for the correlation with authoritative data, whereas the commute count (hereon referred to as CTCNT) is employed for the regression analysis.

For comparison with on-site cycling counts, The UK Department of Transport's annual average daily flow (AADF) product was used, which contains the cycling count at 200 locations across the city for 2017 and 2018.

Table 1. List of Variables

Type	Variable
Boundary	Intermediate Zone (IZ) [3]
Cycling	Strava 2017 Strava 2018
Environmental Variables	Building Height % of Green Space Public Transport Availability Index

Building Height: Building height reports the absolute and relative heights of buildings in the selected area. We used the maximum relative height, which is "the *maximum absolute height for building* minus *the absolute height of ground*".

Green Space (%): Provided by the OS Mastermap, the product of the green space categorises green areas based on their primary function, meaning it describes the purpose of the area (e.g., golf course, public park, cemetery, etc.).

Public Transport Availability Index (PTAI): PTAI takes account of both service frequency and service area and includes public transport on rail and wheels (i.e., coaches and buses) at LSOA level[4].

The data were pre-processed and implemented based on the R packages of tidyverse (Wickham et al., 2019), `sf` (Pebesma et al., 2018), `tmap` (Tennekes, 2018), `spgwr` (Bivand et al., 2020), and `spdep` (Bivand et al., 2005). For reproducibility, we stored the aggregated data and the execution code to our GitHub repository https://github.com/dataandcrowd/AGILE2022. Note that due to the data sharing policy, we are not allowed to share the raw Strava dataset to the public.

3.3 Geographically Weighted Regression

If the residuals of an Ordinary Least Squares (OLS) regression are distributed spatially, a Geographically Weighted Regression (GWR) may be used. This means the relationship between the dependant and independent variables may vary geographically (Brunsdon et al., 1998), i.e. GWR assumes one global model can not describe this relationship and finds the geographical changes in the model (Fotheringham et al., 2003).

$Y_i = a_0(lat_i, long_i) + \sum_1^k (lat_i, lon_i) x_{ik} + \epsilon_i,$

[4]LSOA: lower layer super output area is the second smallest geographic unit in England's demographic datasets

where *lat* and *long* identify the coordinates of location i where the value lies on, the parameter estimates are the continuous function $a_k(lat_i, lon_i)$ k number of independent variables x at location i, distinguishing the equation from OLS regression by this spatial reference. By defining a bandwidth for a grid that moves over data, GWR fits a model to the subset of data that falls into the grid, giving the most weight to the points that are closest to the one at the centre.

Given this paper works at administrative level data, we have taken an adaptive window size. In other words, we used the bandwidth that considers the immediate four neighbours (0.037×136 IZs ≈ 4 IZs). AICc was used to compare the performance of both global and local models (Feuillet et al., 2015).

4 Results

4.1 Spatial Distribution of Variables

The average number of Strava users reported in 2017 were 128,697 persons per IZ. This figure increased to 135,287 in 2018. Figure 1 shows that the distribution is positively skewed as many people have reported short-term journeys (Conrow et al., 2018). Regarding the spatial aspect, the users have frequently cycled in the city centre than the outskirts (see Figure 2).

Table 2. Descriptive Statistics

Variable	Min	Median	Mean	Max
Strava 2017	1,760	79,293	128,697	1,312,075
Strava 2018	2,275	81,615	135,287	1,421,505
Green	1.3	8.1	8.8	20.6
B.Height(m)	5.7	8.1	9.03	21.1
PTAI	245.8	808.5	1,022.2	4,990.8

Figure 3-Top shows that the per cent of green areas (by Intermediate Zones) gradually tends to decrease as it goes outside the city centre. The average is 8% across the whole area but the lowest is situated in the city centre as well as the city south.

The average building heights per IZs were 9 meters across Glasgow, but noticeably high rise buildings are concentrated in the city centre. The City Centre South was the highest at 21.1m followed by City Centre East and City Centre West (see Figure 3-Mid).

PTAI (Public Transport Availability Indicators) tend to be more clustered in the city centre (>3000) and around the major bus routes (>2000) while the north and the east are relatively lower (<1000) (see Figure 3-Bot).

Figure 1 depicts that the distributions of both the response and the predictor variables are highly skewed. Hence, a natural log was applied across all variables to closely normalise the distributions of all variables to meet the assumption of the OLS and GWR regressions.

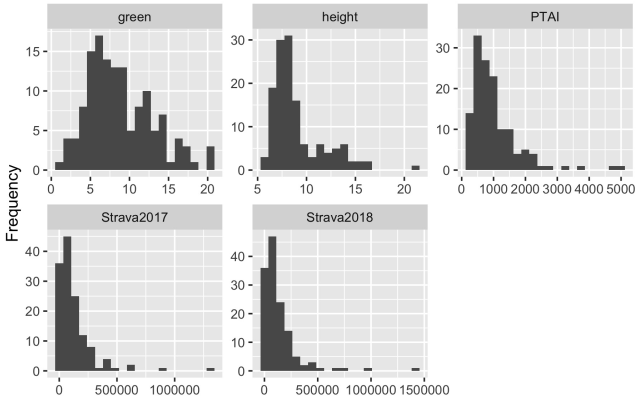

Figure 1. Histogram of the variables

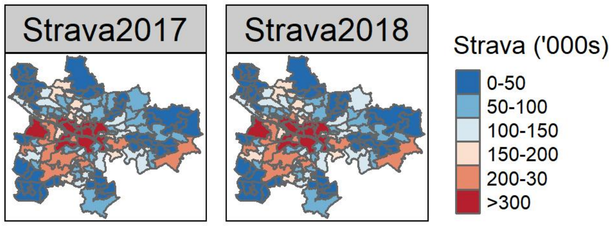

Figure 2. Distribution of Strava Counts

4.2 Ordinary Least Square Regression

The OLS results are shown in Table 3. Results show that the log-transformed green area (%) and building height variables are significant in the 2017 and 2018 models. The findings are consistent with the results of the following studies (Sun et al., 2017a; Heikinheimo et al., 2020; Hochmair et al., 2019; Munira and Sener, 2020). The logged PTAI was slightly over the p-value<0.05 threshold for the 2017 model and was on par with the 0.05 threshold for the 2018 model.

To check the multicollinearity amongst the predictors, we used the variance inflation factor (VIF) to check whether the acceptable scores are usually below 10 (Yang et al., 2020). The variables showed 1.45, 1.98, and 2.08 for green area, PTAI, and building height respectively.

R^2 is around 0.3 for both years, meaning that the explanatory power is not too strong. In other words, the variables contain a higher amount of unexplainable variability.

The residuals are plotted in Figure 4. It is vital to examine the OLS results and the residuals to understand whether there exists a spatial autocorrelation. The Morans'I returned 0.36 for both 2017 and 2018 with p-values less than 0.01. This means that the residuals have a weak spatial autocorrelation. Strava 2017 and 2018 were separately fitted

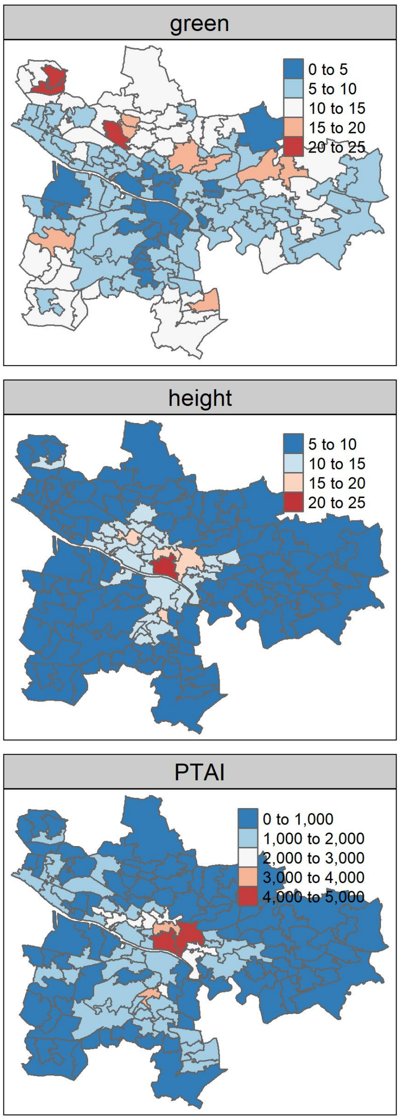

Figure 3. Distribution of Variables: %Green spaces, PTAI, and Average Building Heights

Table 3. Results of OLS Regression. All of the variables were log transformed.

2017	β	Stn.Err	VIF
Intercept	7.07***	1.40	
log(Green area)	-0.48*	0.20	1.45
log(PTAI)	0.37$^{\cdot}$	0.2	1.98
log(Building height)	1.18*	0.49	2.08
R^2	0.30		
2018	β	Stn.Err	VIF
Intercept	6.65***	1.40	
log(Green area)	-0.43*	0.20	1.45
log(PTAI)	0.39*	0.2	1.98
log(Building height)	1.27$^{\cdot}$	0.49	2.08
R^2	0.31		

to the predictor variables using GWR. As the Strava data were positively skewed, we normalised them before we implemented the models.

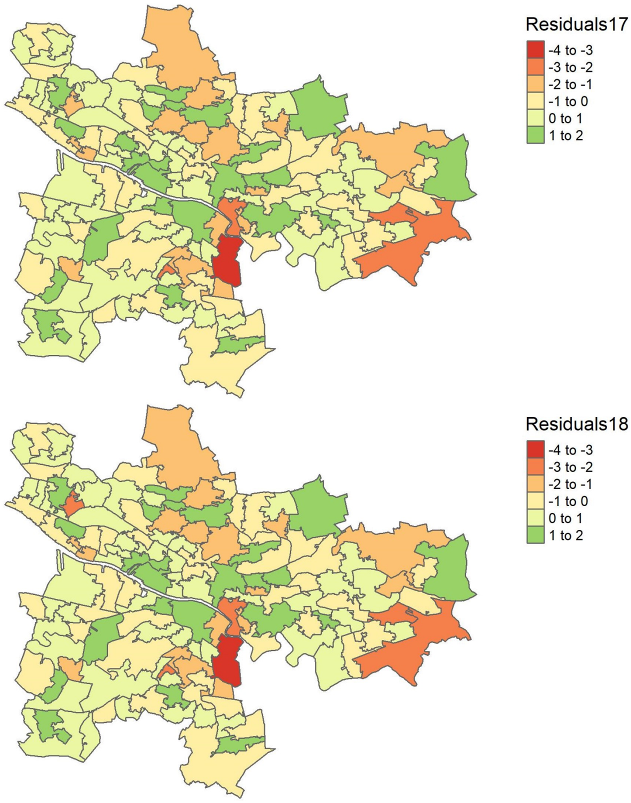

Figure 4. OLS residuals for the response variables. Morans'I for both residuals computed for 0.36 (weak clustering)

4.3 GWR: The Model Fit

The GWR model is presented together with the OLS results in Table 4. In the table, the R^2 were 0.735 and 0.733 for the GWR, which significantly outperformed the OLS results. Regarding the goodness-of-fit (AIC), the GWR results provide 300 and 303 for both years of Strava data, which, with the exponential transformation, imply a huge difference to the original fit.

Table 4. Comparison of Model Fit

Model	Response	R^2	AIC
OLS	Strava 2017	0.30	401
	Strava 2018	0.31	401
GWR	Strava 2017	0.735	300
	Strava 2018	0.733	303

Figures 5 and 6 exhibits the coefficients of the GWR results. For all predictors, a notable spatial variation of the local R^2 is seen across Glasgow ranging from 0.8 to around 0.07 in 2017 and 2018. The highest R^2s are identified in the city centre to the northern area of Glasgow, while the lowest R^2s are mostly situated in the eastern part of the city. However, despite the remarkable improvement on the overall R^2s from the GWR models, further investigation on the infrastructural aspects is needed in terms of the areas with lower predictive power e.g. East Glasgow is lower than the global R^2.

4.4 GWR: Model Coefficients

Tables 5 and 6 exhibit the coefficients from the GWR models to those of OLS models (global). Besides the intercepts, the per cent of Green areas range from -0.24 to 0.23 in 2017 and from -0.2 to 0.28 in 2018; PTAI range from -0.001 to 0.0039 in 2017 and from -0.002 to 0.003 in 2018, and Building height range from -0.44 to 0.35 in 2017 and from -0.37 to 0.4 in 2018. Based on the estimation of the GWR model, the spatial distribution of the coefficients can be seen in Figures 5 and 6. Each predictor shows a spatially diverse outcome regarding its coefficients. For the per cent of green areas and commute cycling, significant positive coefficients are distributed over the north (inc. Gilshochil, Parkhouse), east (inc. Haghill, Gallowgate), and the south-west (inc. Pollock). While the coefficient of PTAI is less sensitive across Glasgow, relatively high values of the coefficient include the western areas (e.g. Pollock and Drumoyne) and the northwestern areas (e.g. Drumchapel). The coefficient of building height shows a contrast between the high and low areas. The high areas are situated in the east (11 IZs), parts of north and west (4 IZs respectively), while significantly low areas are in the northwest (5 IZs).

Table 5. GWR Coeffencicts for Strava 2017 Users

Variable	Min	Median	Max	Global
Intercept	-8.73	5.95	20.48	7.07
log(Green)	-2.23	0.01	1.99	-0.48
log(PTAI)	-0.79	0.52	2.78	0.37
log(B.Height)	-4.06	0.45	3.48	1.18

5 Conclusion

This study respectively applied OLS and GWR regression to identify the global and local associations between key predictors and cycling using Strava Metro's cycling commute data for the period of 2017-18. Variables of of green areas, building height, and public transport availability were collected and cleaned based on the Intermediate Zones (IZs) of Scotland.

Table 6. GWR Coeffencicts for Strava 2018 Users

Variable	Min	Median	Max	Global
Intercept	-7.77	5.88	20.37	6.65
log(Green)	-1.96	0.05	1.99	-0.43
log(PTAI)	-0.91	0.51	2.64	0.39
log(B.Height)	-3.56	0.57	3.87	1.27

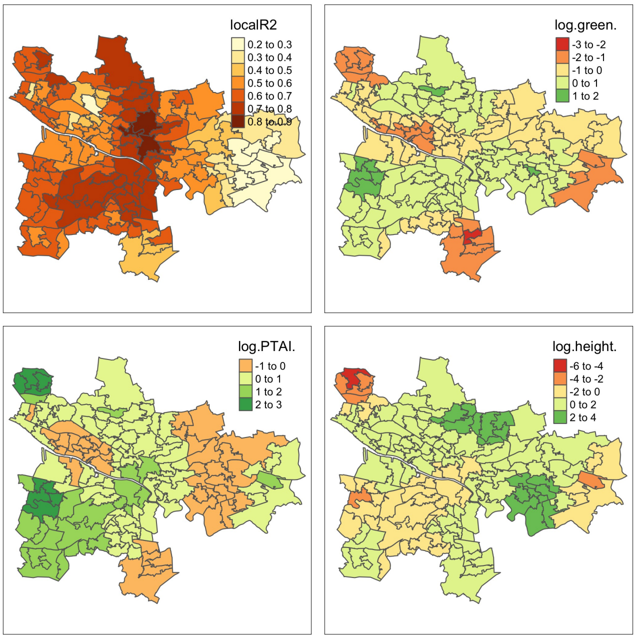

Figure 5. Estimated Coefficients of Strava Cycling (2017) and the Environmental Variables

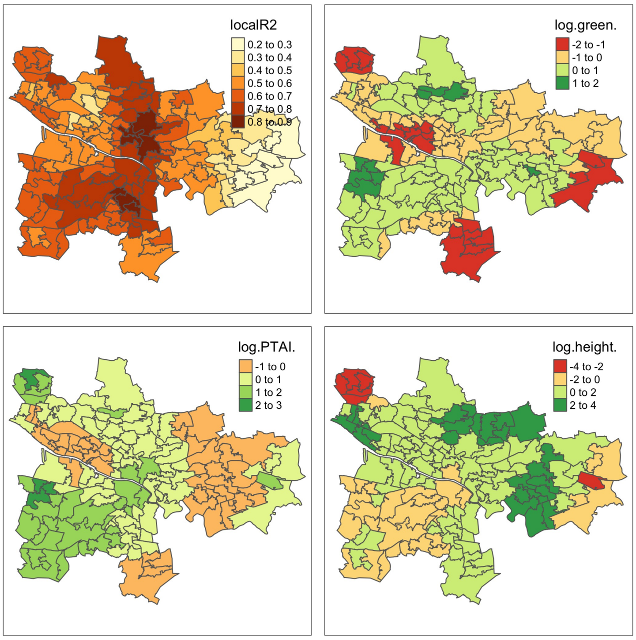

Figure 6. Estimated Coefficients of Strava Cycling (2018) and the Environmental Variables

Overall, the significant local variation in the effects of green areas (log.green) suggests that even though the global regression showed a negative association between the greenness and commute cycling, there were over half of the IZ areas that have a strong positive association with the green areas in both 2017 and 2018. The positive areas were mostly situated in the northwest (e.g. Netherton) or along the river Clyde where parks and cycling paths were constructed. This was something that the OLS regression (global) cannot reveal.

Building height (log.height) shows geographic patterns where the residuals are fairly stationary across the study area with some clusters of high residuals in the north and the central east. We carefully speculate that low skylines and decent cycling infrastructures might have increased people to cycle to work more frequently (Teixeira et al., 2020). The Public Transport Availability Index i.e. PTAI, which was chosen as a proxy to associate public transport access points with bike commute, show a positive relationship overall, but the residual variation was low and homogenous throughout the study extent, meaning the index was neither more nor less effective.

In terms of the model performance, the outcome of GWR performed better than that of OLS with lower AIC values (i.e. 400 vs 300) and higher R^2 (i.e. 0.7 vs 0.3). This suggests that GWR is more suitable to detect the spatial variability between the predicting variables and commute cycling ridership. Further research is needed in terms of how to improve the goodness-of-fit of the model. One way for improvement is to include more variables, for example, length of separate bike paths or meteorology, to find out whether many combinations of data can help understand the commute ridership. Another way is to thoroughly check how different bandwidth of the neighbouring areas in GWR leads to different results (Comber et al., 2021).

This research had some limitations. First, the Strava data employed were aggregated at the year level: a finer temporal resolution would allow for the consideration of daily time slots and therefore lead to different findings. Additionally, this research made no distinction of demographic characteristics of the users, which would be a valuable addition in understanding the levels of over/under-representation of certain user groups.

6 Acknowledgement

The work presented in this paper has been funded by the UK Research and Innovation (UKRI) Future Leaders Fellowship MR/S01795X/2.

References

Aldred, R.: Built Environment Interventions to Increase Active Travel: a Critical Review and Discussion, Current Environmental Health Reports, 6, 309–315, https://doi.org/10.1007/s40572-019-00254-4, 2019.

Aldred, R., Croft, J., and Goodman, A.: Impacts of an active travel intervention with a cycling focus in a suburban context: One-year findings from an evaluation of London's in-progress mini-Hollands programme, Transportation Research Part A: Policy and Practice, 123, 147–169, https://doi.org/10.1016/j.tra.2018.05.018, 2019.

Bivand, R., Bernat, A., Carvalho, M., Chun, Y., Dormann, C., Dray, S., Halbersma, R., Lewin-Koh, N., Ma, J., Millo, G., et al.: The spdep package, Comprehensive R Archive Network, Version, pp. 05–83, 2005.

Bivand, R., Yu, D., Nakaya, T., Garcia-Lopez, M.-A., and Bivand, M. R.: Package 'spgwr', R software package, 2020.

Brand, C., Goodman, A., and Ogilvie, D.: Evaluating the impacts of new walking and cycling infrastructure on carbon dioxide emissions from motorized travel: A controlled longitudinal study, Applied Energy, 128, 284–295, https://doi.org/10.1016/j.apenergy.2014.04.072, 2014.

Broach, J., Dill, J., and Gliebe, J.: Where do cyclists ride? A route choice model developed with revealed preference GPS data, Transportation Research Part A: Policy and Practice, 46, 1730–1740, https://doi.org/10.1016/j.tra.2012.07.005, 2012.

Brunsdon, C., Fotheringham, S., and Charlton, M.: Geographically weighted regression, Journal of the Royal Statistical Society: Series D (The Statistician), 47, 431–443, 1998.

Chen, P., Shen, Q., and Childress, S.: A GPS data-based analysis of built environment influences on bicyclist route preferences, International Journal of Sustainable Transportation, 12, 218–231, https://doi.org/10.1080/15568318.2017.1349222, 2018.

Comber, A., Brunsdon, C., Charlton, M., Dong, G., Harris, R., Lu, B., Lü, Y., Murakami, D., Nakaya, T., Wang, Y., et al.: A route map for successful applications of Geographically Weighted Regression, Geographical Analysis, 2021.

Conrow, L., Wentz, E., Nelson, T., and Pettit, C.: Comparing spatial patterns of crowdsourced and conventional bicycling datasets, Applied geography, 92, 21–30, 2018.

Dadashova, B., Griffin, G. P., Das, S., Turner, S., and Sherman, B.: Estimation of Average Annual Daily Bicycle Counts using Crowdsourced Strava Data, Transportation Research Record, 2674, 390–402, https://doi.org/10.1177/0361198120946016, 2020.

De Vos, J., Schwanen, T., Van Acker, V., and Witlox, F.: Do satisfying walking and cycling trips result in more future trips with active travel modes? An exploratory study, International Journal of Sustainable Transportation, 13, 180–196, https://doi.org/10.1080/15568318.2018.1456580, 2019.

Doorley, R., Pakrashi, V., and Ghosh, B.: Quantifying the Health Impacts of Active Travel: Assessment of Methodologies, Transport Reviews, 35, 559–582, https://doi.org/10.1080/01441647.2015.1037378, 2015.

Echiburú, T., Hurtubia, R., and Muñoz, J. C.: The role of perceived satisfaction and the built environment on the frequency of cycle-commuting, Journal of Transport and Land Use, 14, 171–196–171–196, https://doi.org/10.5198/jtlu.2021.1826, 2021.

Ek, K., Wårell, L., and Andersson, L.: Motives for walking and cycling when commuting – differences in local contexts and attitudes, European Transport Research Review, 13, 46, https://doi.org/10.1186/s12544-021-00502-5, 2021.

Ferster, C. J., Nelson, T., Robertson, C., and Feick, R.: 1.04 - Current Themes in Volunteered Geographic Information, in: Comprehensive Geographic Information Systems, edited by Huang, B., pp. 26–41, Elsevier, Oxford, https://www.sciencedirect.com/science/article/pii/B9780124095489096202, 2018.

Feuillet, T., Charreire, H., Menai, M., Salze, P., Simon, C., Dugas, J., Hercberg, S., Andreeva, V. A., Enaux, C., Weber, C., et al.: Spatial heterogeneity of the relationships between environmental characteristics and active commuting: towards a locally varying social ecological model, International Journal of Health Geographics, 14, 1–14, 2015.

Fotheringham, A. S., Brunsdon, C., and Charlton, M.: Geographically weighted regression: the analysis of spatially varying relationships, John Wiley & Sons, 2003.

Griffin, G. P. and Jiao, J.: Crowdsourcing Bicycle Volumes: Exploring the role of volunteered geographic information and established monitoring methods, Tech. rep., https://osf.io/preprints/socarxiv/e3hbc/, 2019.

Heikinheimo, V., Tenkanen, H., Bergroth, C., Järv, O., Hiippala, T., and Toivonen, T.: Understanding the use of urban green spaces from user-generated geographic information, Landscape and Urban Planning, 201, 103 845, https://doi.org/https://doi.org/10.1016/j.landurbplan.2020.103845, 2020.

Hochmair, H. H., Bardin, E., and Ahmouda, A.: Estimating bicycle trip volume for Miami-Dade county from Strava tracking data, Journal of transport geography, 75, 58–69, 2019.

Hong, J., McArthur, D. P., and Livingston, M.: The evaluation of large cycling infrastructure investments in Glasgow using crowdsourced cycle data, Transportation, 47, 2859–2872, https://doi.org/10.1007/s11116-019-09988-4, 2020a.

Hong, J., Philip McArthur, D., and Stewart, J. L.: Can providing safe cycling infrastructure encourage people to cycle more when it rains? The use of crowdsourced cycling data (Strava), Transportation Research Part A: Policy and Practice, 133, 109–121, https://doi.org/10.1016/j.tra.2020.01.008, 2020b.

Je, D., Sn, B., Jm, J., Mm, M., Jw, R., Bk, S., and Undefined: American College of Sports Medicine Position Stand. Appropriate physical activity intervention strategies for weight loss and prevention of weight regain for adults., Medicine and Science in Sports and Exercise, 41, 459–471, https://doi.org/10.1249/mss.0b013e3181949333, 2009.

Jestico, B., Nelson, T., and Winters, M.: Mapping ridership using crowdsourced cycling data, Journal of Transport Geography, 52, 90–97, https://doi.org/10.1016/j.jtrangeo.2016.03.006, 2016.

Khatri, R., Cherry, C. R., Nambisan, S. S., and Han, L. D.: Modeling Route Choice of Utilitarian Bikeshare Users with GPS Data, Transportation Research Record, 2587, 141–149, https://doi.org/10.3141/2587-17, 2016.

Krenn, P. J., Oja, P., and Titze, S.: Route choices of transport bicyclists: a comparison of actually used and shortest routes, International Journal of Behavioral Nutrition and Physical Activity, 11, 31, https://doi.org/10.1186/1479-5868-11-31, 2014.

Kwayu, K. M., Lyimo, S. M., and Kwigizile, V.: Characteristics of cyclists using fitness tracker apps and its implications for planning of bicycle transport systems, Case Studies on Transport Policy, 9, 1160–1166, https://doi.org/10.1016/j.cstp.2021.06.004, 2021.

Le, H. T. K., Buehler, R., and Hankey, S.: Correlates of the Built Environment and Active Travel: Evidence from 20 US Metropolitan Areas, Environmental Health Perspectives, 126, 77 011, https://doi.org/10.1289/EHP3389, 2021.

Lee, C. and Moudon, A. V.: Neighbourhood design and physical activity, Building Research Information, 36, 395–411, https://doi.org/10.1080/09613210802045547, 2008.

Lin, Z. and Fan, W. D.: Modeling bicycle volume using crowdsourced data from Strava smartphone application, International Journal of Transportation Science and Technology, 9, 334–343, https://doi.org/https://doi.org/10.1016/j.ijtst.2020.03.003, 2020.

Lu, W., Scott, D. M., and Dalumpines, R.: Understanding bike share cyclist route choice using GPS data: Comparing dominant routes and shortest paths, Journal of Transport Geography, 71, 172–181, https://doi.org/10.1016/j.jtrangeo.2018.07.012, 2018.

Macmillen, J. and Stead, D.: Learning heuristic or political rhetoric? Sustainable mobility and the functions of 'best practice', Transport Policy, 35, 79–87, https://doi.org/10.1016/j.tranpol.2014.05.017, 2014.

Mäki-Opas, T. E., Borodulin, K., Valkeinen, H., Stenholm, S., Kunst, A. E., Abel, T., Härkänen, T., Kopperoinen, L., Itkonen, P., Prättälä, R., Karvonen, S., and Koskinen, S.: The contribution of travel-related urban zones, cycling and pedestrian networks and green space to commuting physical activity among adults – a cross-sectional population-based study using geographical information systems, BMC Public Health, 16, 760, https://doi.org/10.1186/s12889-016-3264-x, 2016.

McArthur, D. P. and Hong, J.: Visualising where commuting cyclists travel using crowdsourced data, Journal of Transport Geography, 74, 233–241, https://doi.org/10.1016/j.jtrangeo.2018.11.018, 2019.

Mertens, L., Cauwenberg, J. V., Ghekiere, A., Bourdeaudhuij, I. D., Deforche, B., de Weghe, N. V., and Dyck, D. V.: Differences in environmental preferences towards cycling for transport among adults: a latent class analysis, BMC Public Health, 16, https://doi.org/10.1186/s12889-016-3471-5, 2016.

Munira, S. and Sener, I. N.: A geographically weighted regression model to examine the spatial variation of the socioeconomic and land-use factors associated with Strava bike activity in Austin, Texas, Journal of Transport Geography, 88, 102 865, https://doi.org/https://doi.org/10.1016/j.jtrangeo.2020.102865, 2020.

Park, Y. and Akar, G.: Why do bicyclists take detours? A multilevel regression model using smartphone GPS data, Journal of Transport Geography, 74, 191–200, https://doi.org/10.1016/j.jtrangeo.2018.11.013, 2019.

Pebesma, E. J. et al.: Simple features for R: standardized support for spatial vector data., R J., 10, 439, 2018.

Sarjala, S.: Built environment determinants of pedestrians' and bicyclists' route choices on commute trips: Applying a new grid-based method for measuring the built environment along the route, Journal of Transport Geography, 78, 56–69, https://doi.org/10.1016/j.jtrangeo.2019.05.004, 2019.

Scott, D. M., Lu, W., and Brown, M. J.: Route choice of bike share users: Leveraging GPS data to derive choice sets, Journal of Transport Geography, 90, 102 903, https://doi.org/10.1016/j.jtrangeo.2020.102903, 2021.

Song, Y., Preston, J. M., and Brand, C.: What Explains Active Travel Behaviour? Evidence from Case Studies in the UK, Environment and Planning A: Economy and Space, 45, 2980–2998, https://doi.org/10.1068/a4669, 2013.

Song, Y., Preston, J., and Ogilvie, D.: New walking and cycling infrastructure and modal shift in the UK: A quasi-experimental panel study, Transportation Research Part A: Policy and Practice, 95, 320–333, https://doi.org/10.1016/j.tra.2016.11.017, 2017.

Stefansdottir, H.: A Theoretical Perspective on How Bicycle Commuters Might Experience Aesthetic Features of Urban Space, Journal of Urban Design, 19, 496–510, https://doi.org/10.1080/13574809.2014.923746, 2014.

Strava: Strava Metro, https://metro.strava.com/, 2021.

Sun, Y., Du, Y., Wang, Y., and Zhuang, L.: Examining Associations of Environmental Characteristics with Recreational Cycling Behaviour by Street-Level Strava Data, International Journal of Environmental Research and Public Health, 14, https://doi.org/10.3390/ijerph14060644, 2017a.

Sun, Y., Mobasheri, A., Hu, X., and Wang, W.: Investigating Impacts of Environmental Factors on the Cycling Behavior of Bicycle-Sharing Users, Sustainability, 9, 1060, https://doi.org/10.3390/su9061060, 2017b.

Sustrans: Paths for Everyone: Sustrans' review of the National Cycle Network 2018, Tech. rep., Sustrans, https://www.sustrans.org.uk/media/2804/paths_for_everyone_ncn_review_report_2018.pdf, 2018.

Ta, N., Zhao, Y., and Chai, Y.: Built environment, peak hours and route choice efficiency: An investigation of commuting efficiency using GPS data, Journal of Transport Geography, 57, 161–170, https://doi.org/10.1016/j.jtrangeo.2016.10.005, 2016.

Teixeira, I. P., Rodrigues da Silva, A. N., Schwanen, T., Manzato, G. G., Dörrzapf, L., Zeile, P., Dekoninck, L., and Botteldooren, D.: Does cycling infrastructure reduce stress biomarkers in commuting cyclists? A comparison of five European cities, Journal of Transport Geography, 88, 102 830, https://doi.org/https://doi.org/10.1016/j.jtrangeo.2020.102830, 2020.

Tennekes, M.: tmap: Thematic Maps in R, Journal of Statistical Software, 84, 1–39, 2018.

Van Holle, V., Van Cauwenberg, J., Deforche, B., Goubert, L., Maes, L., Nasar, J., Van de Weghe, N., Salmon, J., and De Bourdeaudhuij, I.: Environmental invitingness for transport-related cycling in middle-aged adults: A proof of concept study using photographs, Transportation Research Part A: Policy and Practice, 69, 432–446, https://doi.org/10.1016/j.tra.2014.09.009, 2014.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., et al.: Welcome to the Tidyverse, Journal of open source software, 4, 1686, 2019.

Winters, M., Brauer, M., Setton, E. M., and Teschke, K.: Built Environment Influences on Healthy Transportation Choices: Bicycling versus Driving, Journal of Urban Health, 87, 969–993, https://doi.org/10.1007/s11524-010-9509-6, 2010.

Winters, M., Davidson, G., Kao, D., and Teschke, K.: Motivators and deterrents of bicycling: comparing influences on decisions to ride, Transportation, 38, 153–168, https://doi.org/10.1007/s11116-010-9284-y, 2011.

Yang, H., Zhang, Y., Zhong, L., Zhang, X., and Ling, Z.: Exploring spatial variation of bike sharing trip production and attraction: A study based on Chicago's Divvy system, Applied Geography, 115, 102 130, https://doi.org/https://doi.org/10.1016/j.apgeog.2019.102130, 2020.

Zahabi, S. A. H., Chang, A., Miranda-Moreno, L. F., and Patterson, Z.: Exploring the link between the neighborhood typologies, bicycle infrastructure and commuting cycling over time and the potential impact on commuter GHG emissions, Transportation Research Part D: Transport and Environment, 47, 89–103, https://doi.org/10.1016/j.trd.2016.05.008, 2016.

Zhao, P.: The Impact of the Built Environment on Bicycle Commuting: Evidence from Beijing, Urban Studies, 51, 1019–1037, https://doi.org/10.1177/0042098013494423, 2014.