The influence of Space Syntax on Cycling Movement. Proceedings of the 12 Space Syntax Symposium

227

THE INFLUENCE OF SPACE SYNTAX ON CYCLING MOVEMENT IN MANTA, ECUADOR

DANIEL ORELLANA[1,2]; MARÍA LAURA GUERRERO[1]

[1] LlactaLAB – Ciudades Sustentables, Departamento de Espacio y Población, Universidad de Cuenca

[2] Facultad de Ciencias Agropecuarias, Universidad de Cuenca

ABSTRACT

This study explores the effect of street network structure on cycling movement patterns in the city of Manta, Ecuador at street segment level, using crowdsourced information from STRAVA Metro and OpenStreetMap. Multiple linear regression models were used to determine the influence of the network structure on cycling activity when controlling for other variables such as household density, socio-economic status, cycling infrastructure and topography. Also, the variation between weekday and weekend activities, and between commuting and sport cycling was explored. Results show positive significant correlations between STRAVA cycling activity volumes and Normalised Angular Choice weighted by segment length (sl_NACH) in the study area. The final regression model including the influence of the network structure for the overall number of cyclists explained 4.6% of the log-likelihood of the intercept model, which represents an important improvement compared with the base model (2.7%). The Incidence Rate Ratio for Normalised Angular Choice weighted by segment length at 4500m radii was 2.75, implying that for an increase of one standard deviation of (sl_NACH) above the mean, there is an expected increase of about 275% in cycling activity. These results are useful for modelling scenarios where cycling infrastructure use can be potentiated using the increase this Space Syntax measure as design guidance.

KEYWORDS

Urban morphology, cycling movement patterns, Space Syntax, Active Mobility

1. INTRODUCTION

Promoting cycling as an active transportation mode, is on the urban mobility management agendas in many cities. Active travel contributes to sustainable urban development, has positive impact on health, causes little to no pollution, requires no fuel apart from normal nutrition, generates few accidents and doesn't require expensive infrastructure, contributing to solve some of the fundamental problems of car-dependant urban environments (Manum & Nordstrom, 2013).

Endeavours for promoting urban cycling have principally concentrated on building and improving infrastructure including dedicated lanes, cycle tracks and parking spaces. However, little attention has been paid to the interaction between cycling patterns and the urban grid (Law et al., 2014). The reasoning behind it is knowing what makes people cycle, and where the will cycle, will help to plan and design cities that promote active mobility. The cycling-built environment relation has been addressed before, finding that connectivity and land use diversity (Sae lens et al., 2003; Winters et al., 2013) and the existence of exclusive bicycle infrastructure (Chang & Chang, 2009; Fagnant & Kockelman, 2016; Winters et al., 2013) are key aspects. However, how the general structure of the network influences movement has been less addressed, and above all, it has been based on somehow narrow datasets (e.g. counts in certain sectors of the urban grid, limited times of day, or phone surveys susceptible to memory and perception bias of the surveyed subject). Therefore, few studies have objective count data with enough time aggregation and space coverage to draw a general pattern of cycle movement. Some authors highlight the importance of using data that accounts for a long period of time to have a comprehensive approach to bicycle research (Heinen et al., 2010).

Latin American transport research on cycling has been focused on health (Hoehner et al., 2008; Sadarangani et al., 2018) or general commuting where cycling is a component (Guerra et al., 2018). Some studies explore the relationship between cycling and the built environment more comprehensively, but limited to areas where specific initiatives allowed gathering data at a fine scale (Ziff et al., 2018). Consequently, detailed studies on the relationship between the urban structure and bicycle movement Latin America are missing.

Crowdsourcing platforms are a viable resource for studies requiring fine-scale data, especially in some regions with challenging economies, like Latin America. In this contribution, we examine hoe using large volunteered and crowdsourced geo-information datasets (OpenStreetMap and STRAVA Metro) contributes to explore the relationship between the urban network and cycling movement behaviour in Manta.

The following section briefly introduces the theoretical background related to space syntax and cycling behaviour relevant for our study. The next section exposes the methodology, including datasets, the mapping process, and statistical analyses. Section four presents and discusses the main results of the study, and finally conclusions are presented in Section Five.

Space Syntax and Bicycle Movement

The relationship between human behaviour and urban morphology is the nucleus of Space Syntax, a theory of space that extracts spatial properties and relates them to social performance (Hillier and Hanson, 1984). Hillier (2007) states that for cognitive reasons, the relation configuration-form deeply affects our spatial behaviour. Space syntax draws patterns from quantifying our responses to 'clues' given by inherent properties of built space (configurations and structure). The form of structures has proven to be very related to human conduct in different areas: natural movement and movement economies (Hillier, 2007), crime (Chiaradia et al., 2009; Tarkhanyan, 2015), poverty (Bolton et al., 2017), and others.

Cyclist's movement in urban environments goes hand in hand with the properties of urban structure. Raford, Chiaradia, & Gil (2007) found that accumulated cyclists' trips, regardless of origin and destination, seem to follow a compelling spatial logic well described by mean angular depth. In the study, this variable was the most robust explanatory variable for cyclist's counts variation at urban level, probably due that streets with lower angular mean depth are probabilistically "shallower" to more origin – destination pairs. Law et al. (2014) showed that Normalised angular choice at radii N had stronger explanatory power on cyclist movement than the London Cycle Superhighway presence (built after 2008), statistically wise, with data between 2003 and 2012. This meant that accessibility of the route was more influential than infrastructure to describe movement. Rybarczyk & Wu (2014) used Space Syntax's Visual Graph Analysis to analyse the link between cycling activities and the built environment, finding that utilitarian visual mean depth related positively with bicycle activity, because of the perceived simplicity of reaching the closest traversable space with visual mean depth. The same authors also found that disorder (visual entropy) near the trip origin reduced the likelihood of bicycling. All these studies show the importance of inherent properties of the built for determining or influencing route choice. They also show the relation between route choice and cognition or how we interpret or read space. Previously, research showed disaggregate modelling techniques and their adequacy for exploring cycling route choice and how it relates to an ample range of intervening effects, including urban form, socio-economic features and cycler's attributes (Parking et al., 2008).

Manum, Nordström, Gil, Nilsson, & Marcus (2017) explored Origin-Destination Betweenness (OD Betweenness), consists on scoring each segment for being on the shortest routes between a set of origins and destinations, instead of routes between all nodes of the network. The authors combined number of residents (origin weight) with the normalised weight of the destination (i.e. dividing the destination weight by the sum of all destination weights). They also advice adding speed and traffic safety as weights in the betweenness measure, and make time work as impedance factor to improve model's accuracy, based on 'convenience', aside from configuration.

There are certain limitations on the studies mentioned in this section, related to data acquisition and geographic scale. Rybarczyk & Wu (2014) used phone surveyed data, while Raford, Chiaradia, & Gil (2007) used hand-drawn traces. Law et al. (2014) used gate counts for 50 locations in 2003 and 22 locations in 2012 in Elephant and Castle, London.

For analysing cycling behaviour, datasets expensive in logistic and economic terms. Moreover, most related studies have been made in developed countries such as UK, the US, Sweden, and not much is known about cycling in developing countries.

Our study contributes by introducing a methodological approach to explore the relationship between urban network structure and cycling movement using crowdsourced data and space syntax. This perspective can be cost-effective, especially in places where data is scarce.

2. DATASETS AND METHODS

Our approach consisted on fitting multiple regression models with different sets of control variables related to physical and structural characteristics of the street network, socio-economic factors and cycling infrastructure. Then, a variable representing the network structure was added to the base model and both models were compared to understand the relative influence of the street network structure on the volume of cycling activity.

Study Area

Manta is a coastal city in the province of Manabí, at the west of Ecuador, with an urban population of 217 553 inhabitants and an urban surface of 61.4Km2 (INEC, 2010). The total length of roads in the city is 706.63km (OpenStreetMap, 2016). It is city by the sea, limiting with the Pacific Ocean to the north. The mean altitude is 6 m.a.s.l. and mean temperature varies from 21 to 28ºC, with a dry weather in comparison to the rest of the Ecuadorian coast. The topography is variable: there is a plateau with relatively high slope changes that make neighbourhoods or household settlements to form in the portions of land where the slope is nearly flat (Figure 1). The local economy is based on industry and commerce, and its beaches are an important destination for domestic tourism. It has one university and a process of socio-spatial segregation is evident, as shown in Figure 2, where the quartiles of population with better living conditions are accumulated on the north west, by the sea shore and to the centre of the city, while those with worse living conditions accumulate to the south, over the slope of the city. The province of Manabí has the fifth highest crime rate in the country (Ministerio del Interior, 2019).

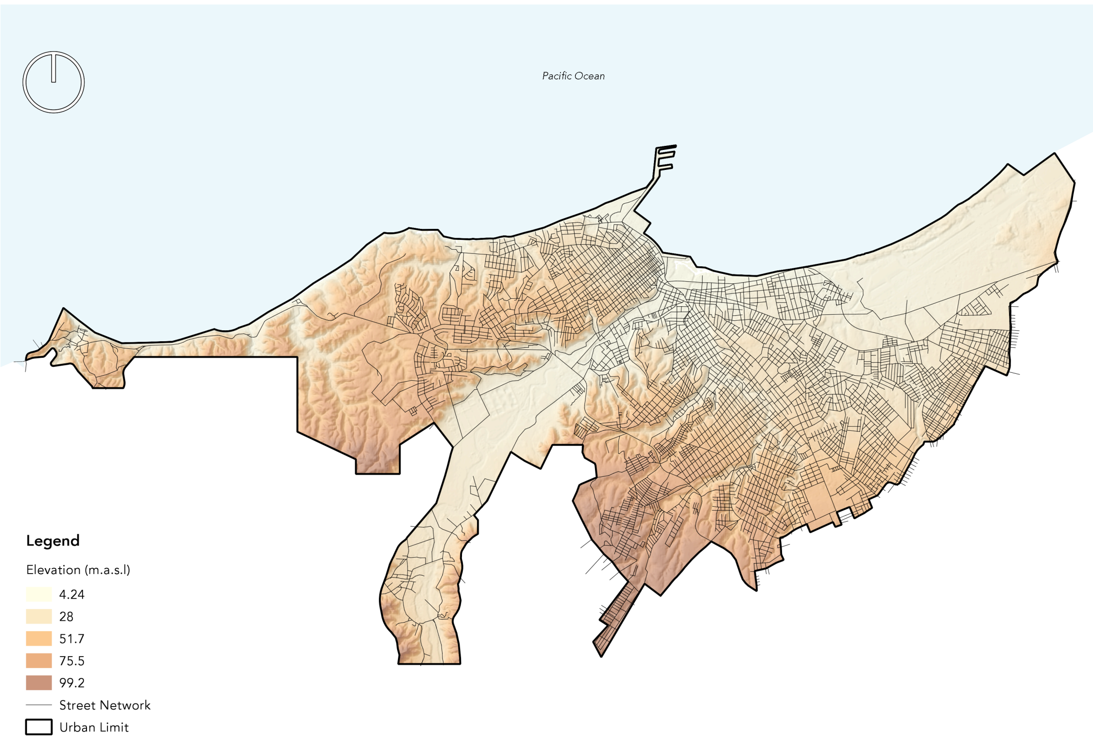

Figure 1. Study area: Manta (Ecuador).

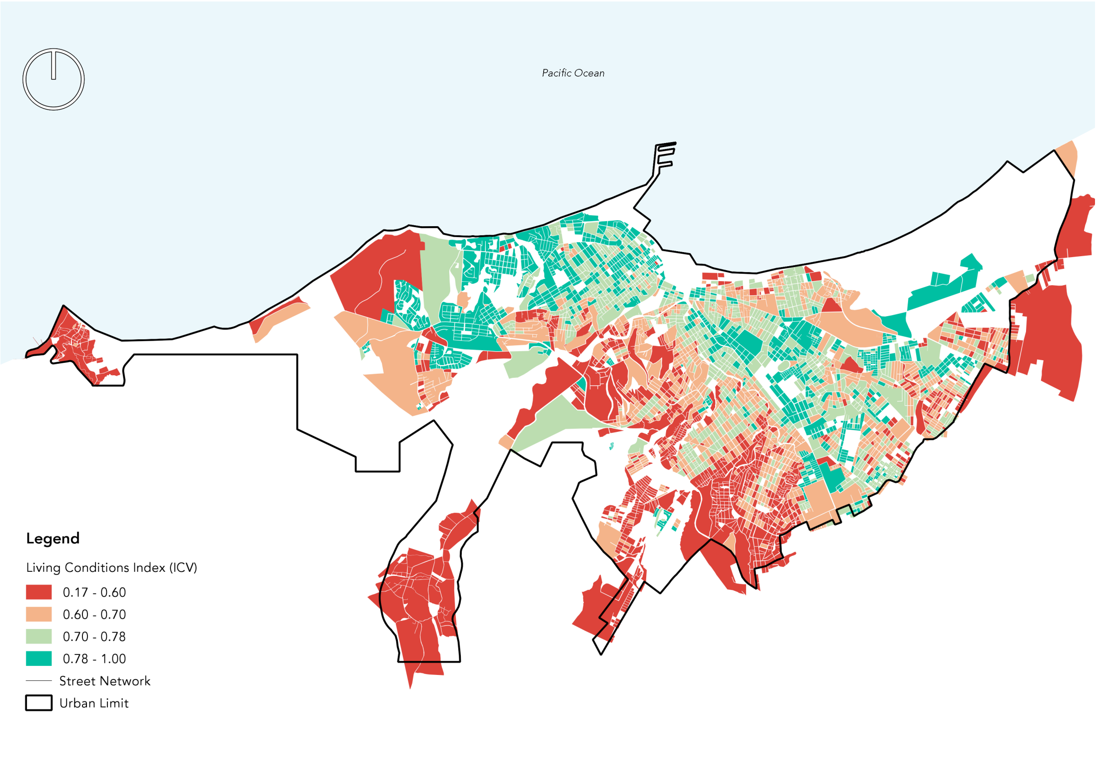

Figure 2. Socio-spatial segregation Manta (Ecuador).

Datasets

We've used two main data sources. The first is OpenStreetMap (OpenStreetMap, 2016), a free volunteered geographic information (VGI) online mapping platform, which was used to create a road centreline map of street segments. The other dataset is a product from STRAVA Metro (Strava Metro, 2015), a commercial system, which employs information generated by Strava App users and collects GPS data (with mobile devices or otherwise), creating fine grained, aggregated, and anonymised datasets. This dataset contains bicycle trip counts per street segment between September 1, 2014 and September 30, 2015 for all streets in Cuenca's urban area. Both datasets allow having up-to-date large-scale data at fine geographical and temporal levels for analysing the link between cycling patterns and the structure of the urban grid, in detail. Also, the datasets world-wide coverage implies that our approach could be replicated and scaled.

This study used additional datasets for obtaining other variables, like the National Population and Household census of 2010 (INEC, 2010), used for household density and socio economic status. A Digital Terrain Model (DTM) at resolution of 3m from 2011 provided by the SIGTierras program (MAGAP, 2012), used to compute the slope of each street segment.

Variables

For the dependent variable we've used the total number of STRAVA cycling trips (activities) for one year counted on each street segment (*total activities*). For analysing temporal and user type variations we've added four additional dependent variables, obtained from the original STRAVA datasets: activity counts on weekdays (*total activities weekdays*), activity counts on weekends (*total activities weekends*), sports activities counts on weekdays (*sports activities weekdays*), commuting activities counts on weekdays (*commuting activities weekdays*).

We also considered a number of control variables that might influence cyclists' behaviour, based on literature: Heinen et al. (2010) and Parking et al. (2008) suggest socio-economic status, for which we've used *household density* and *living conditions index* (ICV) (Orellana and Osorio, 2014). Chang and Chang, (2009), Fagnant and Kockelman (2016), Winters et al. (2013) highlight infrastructure, that in our study is represented as *road hierarchy*, *existence of bikeways* and *number of connections.* Finally, physical variables, suggested by Saelens et al. (2003), Winters et al (2013), represented as *slope* of street segments.

For explaining the network structure, we used Space Syntax's Integration and Choice. For an appropriate assessment of cycling patterns, we computed Integration and Choice for a set of radii from local to global (400m, 800m, 1200m, 1600m, 2000m, 2500m, 3000m, 3500m, 4000m, 4500, 5000m, 6000m, 7000m, 8000m, 10000m, 15000m and global Rn), to identify the one that related the most with the dependent variable. Literature indicated that normalised angular versions these variables describe better the incidence of network configuration (Law et al., 2014; Turner, 2007). For that reason, we also computed *NACH* (normalised choice) and *NAIN* (normalised integration) for the corresponding radii.

Data preparation and mapping process

Space Syntax variables demands a simplified geometry of street network for computing its analysis (Kolovou et al., 2017). OpenStreetMap often represents space in more detail, like double lines for separate lanes, complex roundabouts, etc.; therefore, pre-processing the original OSM dataset was necessary for producing a simplified network.

The pre-processing consisted of five steps. First, we used QGIS, generalizing the original OSM geometries with the Douglas-Peucker algorithm, for reducing the number of vertices per line. Then, roundabouts were transformed to simple crossings one at a time, and multi-lane streets were simplified by generating 20m buffers for deriving a simple centreline using PostGIS. Third, the simplified network was cleaned with segmentation and snapping algorithms from GRASS. The fourth step consisted in importing unlinks from the original dataset, representing tunnels and overpasses. For the final step, we tested the validity of the resulting network with using the Space Syntax Toolkit Plugin.

For computing Space Syntax's' NACH and NAIN on the simplified road network we used the angular segment analysis algorithm form the Space Syntax Toolkit for QGIS. This process also allowed to obtain the number of connections. We applied the smallest geographical disaggregation level of the 2010 national census for Living conditions index and household density, calculating them for street blocks. Then these values were transferred to street segments, joining them by location and averaging values. Slope was obtained with every segment's first and last points altitude and its length. Finally, we added road hierarchy and cycleways and to segments using the original OSM classification by spatial match.

For transferring the cycling activity counts from the original network dataset, we used spatial join, with tolerance values of 50m for avenues and 7m for streets.

Statistical Analysis

There were three steps for the statistical analysis. First, individual regression models between each independent variable and *total activities* were fitted and evaluated. This step had double purpose: exploring the non-controlled influence of each variable, and selecting the right radii of the spatial variable for the following phases. Second, multiple regression "base" models were computed with the selected variables from the previous step. Variable selection process was based on each variable's individual influence, obtained on the first phase, and stepwise variable selection based on BIC (Bayesian Information Criterion) comparisons and Vuong's test (Vuong, 1989). Third, full models including the selected spatial variable were fitted for each type of cycling counts.

A full and a base model were fitted for each version of the dependent variable. Then the models were compared using the Mc Fadden's pseudo R2 for the LR statistic of the intercept-only model to obtain network structure's explicative potential on cycling activity. Given the logarithmic nature of the link function in negative binomial models, coefficients of a model can be interpreted more easily by exponentiating them. The resulting exponentiated coefficients are known as Incidence Rate Ratios (IRRs) and can be interpreted as the expected variation of the dependent variable for each unit of variation of the predictor, keeping all other predictors' values the same. Living Condition Index, Household density, and space syntax variables were transformed to Z scores for comparing and interpreting regression models' coefficients, because Z scores allow to read IRRs variations in terms of standard deviations from the mean, regardless of the variable's scale. All statistical analyses were performed in R (R Core Team, 2016).

3. RESULTS

The analysis included a total of 9847, after deleting null values and invalid geometries. Visual comparison showed relative correspondence between the streets with the highest values of NACH and streets with the highest values of cyclists' activities counts, as shown in Figures 3 and 4.

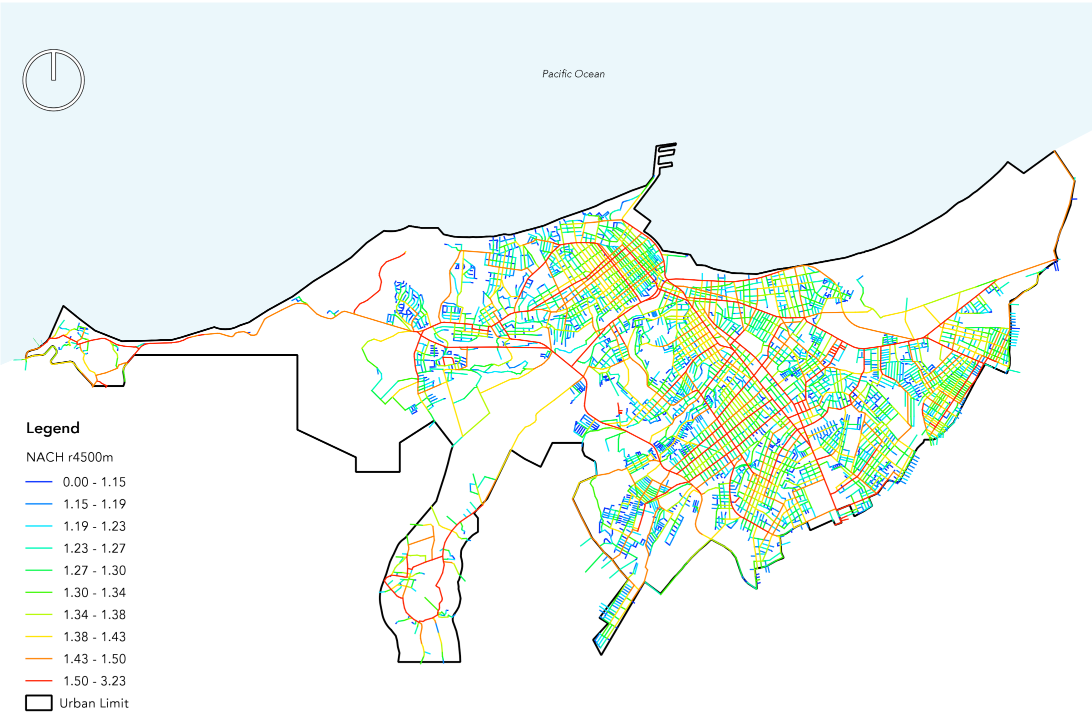

Figure 3. Deciles of NACH

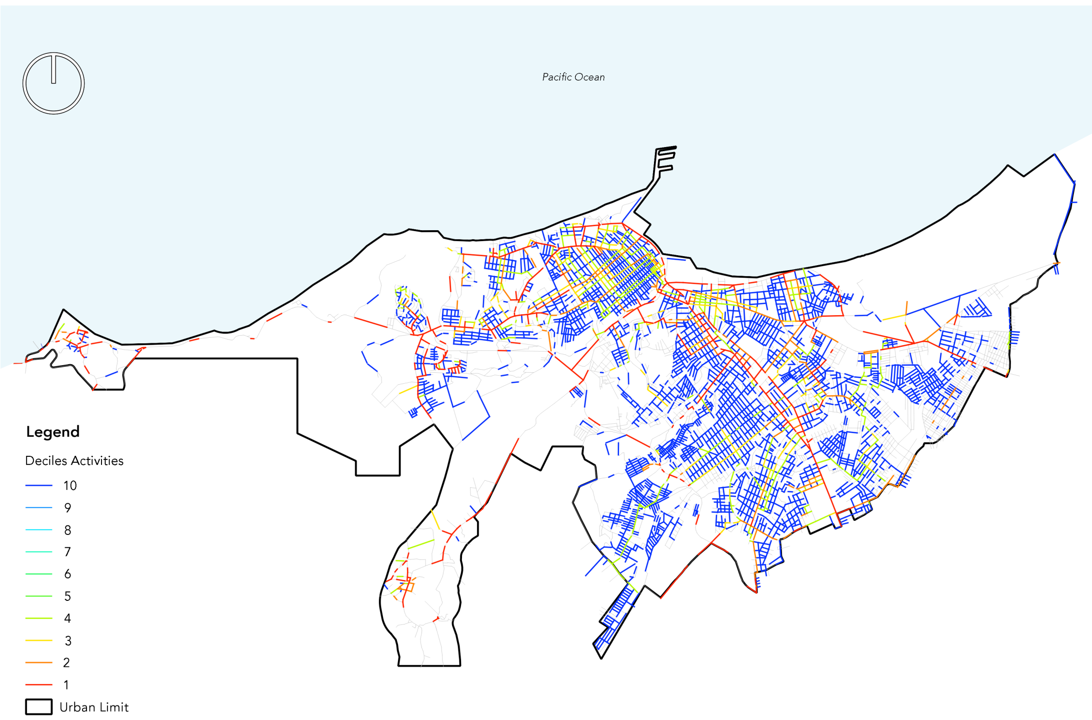

Figure 4. Deciles of total cycling activities counts

Table 1 shows descriptive statistics of the variables. As it was expected, several streets had 0 counts of cycling activity. The difference between the mean and the variance in each dependent variable shows evidence of over dispersion. These characteristics suggest that negative binomial models distributions are appropriate. Finally, there is a higher number of sports activities than commuting activities, as it was expected when using a sports app.

Table 1. Descriptive statistics of variables

	Mean	Median	Variance	Kurtosis	Asymmetry coefficient	Minimum	Maximum	Total
total activities	32.551	0	18762.12	115.52	8.526	0	3423	176037
total activities weekdays	20.105	0	7176.19	136.39	9.049	0	2257	108729
total activities weekends	12.446	0	2806.82	89.56	7.864	0	1166	67308
commuting activities weekdays	4.025	0	248.26	84.12	7.588	0	333	21769
sports activities weekdays	16.080	0	4838.89	150.96	9.422	0	1924	86960
z household density	0.007	-0.125	0.995	0.44	0.743	-1.501	5.556	-
z living conditions index	0.003	0.067	0.995	-0.02	-0.152	-4.744	2.457	-
road hierarchy	0.050	0	0.048	14.88	4.108	0	1	-
existence of bikeways	0.001	0	0.001	671.62	25.949	0	1	-
number of connections		5	1.430	-0.22	-0.614	1	8	-
Slope	0.027	0.017	0.001	11.82	2.842	0.000	0.326	-
z NACH_sl_4500	0.004	-0.065	0.925	0.45	0.379	-2.650	7.735	-

Individual associations

Table 2 shows the coefficients, Incidence Rate Ratios (IRR) and significance values for individual negative binomial models. All predictors, except Z *Living Conditions Index*, showed highly significant association with the cycling activity counts. We decided to keep all variables for the next steps (model fitting).

Z Household Density showed strong negative association for all types of activities. Z *Living Conditions Index* had some significant correlation with weekday commute activities, and no significance for sports activities neither on weekdays nor on the weekends. *Road hierarchy*, *existence of bikeways* and *number of connections* all had strong association with all types of cycling activities. *Road hierarch*y and *existence of cycle paths* both have positive influence with a strong incidence ratio, while the amount of connections has negative influence on the activities count. *Slope* had strong negative association with the five types of counts, which is consistent with previous studies and with the general behaviour of cyclists who avoid steep slopes.

Finally, an assessment of NACH and NAIN variables at different radii indicated that Normalised Angular Choice weighted by segment length at 4500m radius (z *NACH_sl_4500*) had the strongest association with cyclist counts in the study area. Therefore, this variable was selected to represent the influence of the network structure for further analysis. There is some indication that the influence of the network is stronger for utilitarian activities that are closer to natural movement theory.

Table 2. Results of individual regressions

Dependent Variables	Independent Variables	coeff	IRR	p	
Total Trips	z Household Density	-0.76	0.47	<0.001	***
	z Living conditions index (ICV)	0.04	1.04	0.304	
	Connections	-0.36	0.697	<0.001	***
	Road Hierarchy	2.77	15.94	<0.001	***
	Bikeways	3.23	25.37	<0.001	***
	Slope	-8.84	1.45E-04	<0.001	***
	Z NACH sl4500	0.85	2.34	<0.001	***
Total trips weekdays	z Household Density	-0.74	0.48	<0.001	***
	z Living conditions index (ICV)	0.04	1.04	0.329	
	Connections	-0.356	0.7	<0.001	***
	Road Hierarchy	2.74	15.43	<0.001	***
	Bikeways	3.29	26.87	<0.001	***
	Slope	-7.91	3.66E-04	<0.001	***
	Z NACH sl4500	0.83	2.3	<0.001	***
Total trips weekends	z Household Density	-0.79	0.45	<0.001	***
	z Living conditions index (ICV)	0.04	1.04	0.335	
	Connections	-0.365	0.693	<0.001	***
	Road Hierarchy	2.82	16.82	<0.001	***
	Bikeways	3.13	22.95	<0.001	***
	Slope	-10.62	2.43E-04	<0.001	***
	Z NACH sl4500	0.89	2.44	<0.001	***
	z Household Density	-0.691	0.501	<0.001	***

Commuting trips weekdays	z Living conditions index (ICV)	0.112	1.118	0.004	**
	Connections	-0.328	0.72	<0.001	***
	Road Hierarchy	2.577	13.152	<0.001	***
	Bikeways	3.166	23.72	<0.001	***
	Slope	-8.032	3.30E-04	<0.001	***
	Z NACH sl4500	0.879	2.408	<0.001	***
Sports trips weekdays	z Household Density	-0.761	0.467	<0.001	***
	z Living conditions index (ICV)	0.019	1.019	0.635	
	Connections	-0.361	0.695	<0.001	***
	Road Hierarchy	2.775	16.046	<0.001	***
	Bikeways	3.32	27.667	<0.001	***
	Slope	-7.893	3.73E-04	<0.001	***
	Z NACH sl4500	0.83	2.292	<0.001	***

Multiple regression models

Different base models were fitted to decide a suitable set of control socio-economic, physical, and infrastructure variables using Likelihood Ratio (LR stats) and ANOVA tests for comparison. Although *living conditions index* had small contribution to the variation of the cyclist's counts on the individual regression, it was kept for the base model, since it became significant during the fitting process. On the contrary, *slope* and *existence of bikeways* became not significant when controlled for other variables, and Voong's tests suggested that their exclusion produced a parsimonious model without important loss of explanatory power. The characteristics of the city may support the decision of excluding these variables. There is only one dedicated bikeway in Manta, on a highway towards the exit of the city, and therefore its importance is rather limited when other variables, such as *road hierarchy* are considered. Slope, on the other hand, lost significance when controlled for the other variables: The denser neighbourhoods are located on the flatter areas. The final set of control variables is therefore composed by *z Household Density, z Living Conditions Index, Number of Connections,* and *Road Hierarchy.*

The base model for *total activities* had a Likelihood Ratio LR=787.3242 and a pseudo R2 of 0.027, which means that explained 2.7% of the log-likelihood of the intercept-only model. Once the network infrastructure is included in the model trough *z NACH_sl_4500*, the full model had a LR of 546.3554 and a pseudo R2 of 0.046. This suggests that the full model explained a 70% more than the base model. The influence of network's structure is therefore able to explain almost as much as all the control variables together.

The models for the different versions of the dependent variable exhibited a similar behaviour: In all cases, the inclusion of *z NACH_sl_4500* improved the model's prediction. For commuting activities on weekdays, the full model improved an 86% compared to the base model, explaining a 6.7% of the intercept-only model. For sport activities on weekdays, the improvement of the full model was somehow lower (a 66%) and explained a 4.8% of the intercept only model. This result suggest that the Normalised Angular Choice has a stronger effect for commuting cyclists than for sport cyclists. The difference between weekdays and weekends was unexpected, since the inclusion of *z NACH_sl_4500* had a stronger effect for predicting weekend activities than for weekday activities. One possible explanation is the higher variability on weekday activities.

Vuong's test Z values for Bayesian Information Criterion (BIC) revealed that full models were closer to the true models than the base model even when accounting for complexity. All values were significant with p-values < 0.001. In other words, the value of the information provided by *z NACH_sl_4500* is well worth. Table 3 presents a summary of each model. Figure 5 illustrates the improvement of each full model compared to the corresponding base model.

Table 3. Model summary (comparison between base and full models)

Dependent Variable	Model	2 x log-likelihood	McFadden's R2†	Vuong's Z value for BIC††
Total Activities	Intercept	-28896.21		
	Base	-28108.89	0.027	
	Full	-27562.53	0.046	8.132 (BIC)***
Total activities weekdays	Intercept	-25228.73		
	Base	-24530.26	0.028	
	Full	-24042.98	0.047	7.886 (BIC)***

Total activities weekends	Intercept	-21673.10		
	Base	-20952.77	0.033	
	Full	-20446.01	0.057	8.192 (BIC)***
Commuting activities weekdays	Intercept	-16782.41		
	Base	-16183.57	0.036	
	Full	-15653.12	0.067	8.585 (BIC)***
Sports activities weekdays	Intercept	-23015.25		
	Base	-22345.69	0.029	
	Full	-21904.31	0.048	7.508 (BIC)***

† Compared to the intercept-only model
†† Compared to the base model. Significance codes: *:<0.05, **:<0.01, ***:<0.001,

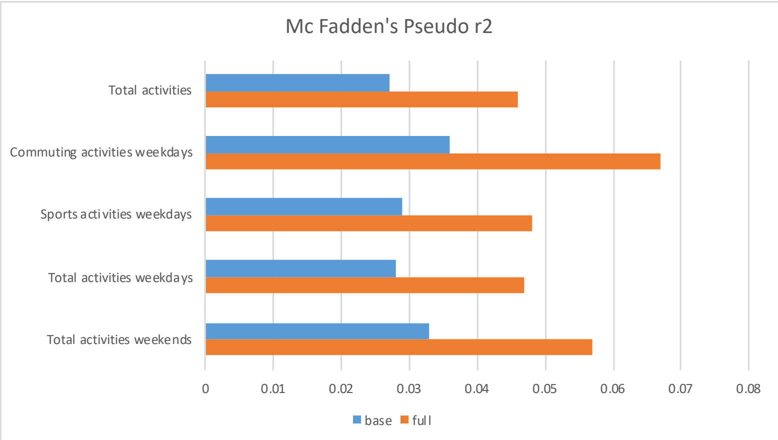

Figure 5. Model improvement compared to base model

The influence of each variable in the full-generalized linear models is summarized in Table 4. All predictors were significant at 99.9% (p<0.001).

Road Hierarchy had an IRR of 6.73 (CI: 4.92 - 9.51) for *total activities*. This means that primary roads had, in average between about five and nine more cycling activities that other roads keeping all the other predictors the same. *Number of connections* had a negative influence: each additional connection will diminish the cycling activity in the segment to the half (IRR=0.56, CI: 0.52–0.60). This means that cyclists seem to avoid segments with complex intersections where they might find potential conflict or insecurity. Household density also had a negative influence on cycling activity, since an increment of 1SD on density will diminish the cycling activity on between 30% and 40% (IRR=0.67, CI:0.625–0.715). IRR for *z Living Conditions Index* was 1.25 (CI:1.17–1.32), implying that cycling activity is between 17% and 32% higher in streets with 1 SD above the mean of living conditions. Although a higher influence of living conditions might be expected for data collected with a smartphone app, it seems that STRAVA cyclists don't have a strong preference for cycling in high-living conditions areas for both, sport and commuting activities. Finally, *z NACH_sl_4500* had an important influence: an increment of cycling activity between 2.5 and 3 times is expected for 1 SD of increment of Normalised Angular Choice.

Figure 6 illustrates the IRRs and 2.5% - 97.5% confidence intervals of each predictor for comparability of different outcome variables. First, CI for different dependent variables largely overlap, which indicates that predictors' influence is somehow uniform for weekdays and weekends and for sports and commuting cycling. Only for *commuting activities weekdays* the influence of *road hierarchy* was visibly lower, whereas for *z connections* and *z household density,* was slightly lower (closer to 1) and for *z living conditions index* slightly higher. *Z NACH_sl_4500* behaved similar for all the dependent variables. On the other hand, *Road hierarchy* had a somehow stronger influence for activities on the weekends compared to weekdays.

For commuting activities during weekdays, z *household density* and *z living conditions index* had IRRs of 0.703 and 1.27 respectively. This means that increasing one standard deviation (SD) above the mean of household density implies a reduction of about 30% in cycling activities counts and one SD increase in the living conditions index would be associated with an increase of 27% in the cycling activities keeping all the other variables the same. The relation with the amount of connections is negative (IRR

0.53), implying a reduction of 47% of cycling activity. Contrarily, road hierarchy is associated with an increase of 470% cycling activities. Normalised Angular Choice (*z NACH_sl_4500*) has an IRR of 2.75, implying that for an increase of one SD in the value of *z NACH_sl_4500*, there is an expected increase of 275% in the cycling activity.

The model for sports activities during weekdays also showed strong correlations for *household density* and *living conditions index*. The IRRs showed that an increase in 1SD of household density would decrease the expected cyclist counts in 34%, compared to the 29.7% reduction for all cycling activity in the previous model. This means that during the week utilitarian cyclists tend to move through denser areas than sports cyclists. *Living conditions index* has an IRR of 1.216, implying that there is an increase of 21% in the expected cycling activity for 1SD of increment of *living conditions index*, which mean that commuters during weekdays tend to move slightly more through areas with better living conditions. The number of connections have a negative effect of reducing 52% of cycling activities; compared to the 47% reduction for commuters, it would mean that sports cyclists tend to move through less connected segments. Finally, for the spatial variable, an increase in one SD of *z NACH_sl_4500* has an expected increase of 269% in cycling activity. This means that the spatial structure has a higher influence on commuting activities, comparing it to the 275% increase on commuting activities.

When comparing the models for total activities during weekdays and total activities during weekends, there are slight differences between the IRR values, but those differences are smaller than between the other two models, and since the confidence intervals overlap, there is not enough evidence to reject the null hypothesis that the models are different. To illustrate this, Figure 6 represents the confidence intervals for the variables and their means. Despite of it, it is worth highlighting an important pattern in these models. Road hierarchy has an influence ratio of 630% for weekday activities and 744% for weekend activities. Also, the influence of the spatial variable, *z NACH_sl_4500* has an IRR of 271% and 278% in total activities during weekdays and weekends respectively.

Table 4. Final Regression Models

Dependent Variable	Independent Variable	coefficient	IRR	2.50%	97.50%	p	sig
Total Activities	(Intercept)	2.336	10.335	9.611	11.133	<0.001	***
	z Household Density	-0.403	0.668	0.625	0.715	<0.001	***
	z Living conditions index (ICV)	0.217	1.242	1.166	1.324	<0.001	***
	Connections	-0.580	0.560	0.522	0.600	<0.001	***
	Road Hierarchy	1.907	6.733	4.916	9.507	<0.001	***
	z NACH sl4500	1.012	2.752	2.529	2.998	<0.001	***
Total activities weekdays	(Intercept)	1.894	6.647	6.162	7.183	<0.001	***
	z Household Density	-0.389	0.678	0.632	0.727	<0.001	***
	z Living conditions index (ICV)	0.207	1.230	1.152	1.314	<0.001	***
	Connections	-0.584	0.558	0.519	0.599	<0.001	***
	Road Hierarchy	1.841	6.303	4.545	9.030	<0.001	***
	z NACH sl4500	1.000	2.718	2.487	2.974	<0.001	***
Total activities weekends	(Intercept)	1.308	3.698	3.423	4.002	<0.001	***
	z Household Density	-0.436	0.647	0.601	0.696	<0.001	***
	z Living conditions index (ICV)	0.229	1.258	1.174	1.347	<0.001	***
	z Connections	-0.559	0.572	0.531	0.615	<0.001	***
	Road Hierarchy	2.007	7.444	5.364	10.678	<0.001	***
	z NACH sl4500	1.024	2.784	2.549	3.043	<0.001	***
Commuting activities weekdays	(Intercept)	0.373	1.452	1.347	1.567	<0.001	***
	z Household Density	-0.352	0.703	0.655	0.756	<0.001	***
	z Living conditions index (ICV)	0.245	1.278	1.197	1.364	<0.001	***
	z Connections	-0.528	0.590	0.549	0.632	<0.001	***
	Road Hierarchy	1.548	4.701	3.442	6.596	<0.001	***
	z NACH sl4500	1.013	2.755	2.530	3.003	<0.001	***
Sports activities weekdays	(Intercept)	1.649	5.199	4.805	5.637	<0.001	***
	z Household Density	-0.402	0.669	0.622	0.720	<0.001	***
	z Living conditions index (ICV)	0.196	1.216	1.137	1.302	<0.001	***
	z Connections	-0.594	0.552	0.512	0.596	<0.001	***
	Road Hierarchy	1.905	6.722	4.789	9.773	<0.001	***
	z NACH sl4500	0.992	2.697	2.457	2.963	<0.001	***

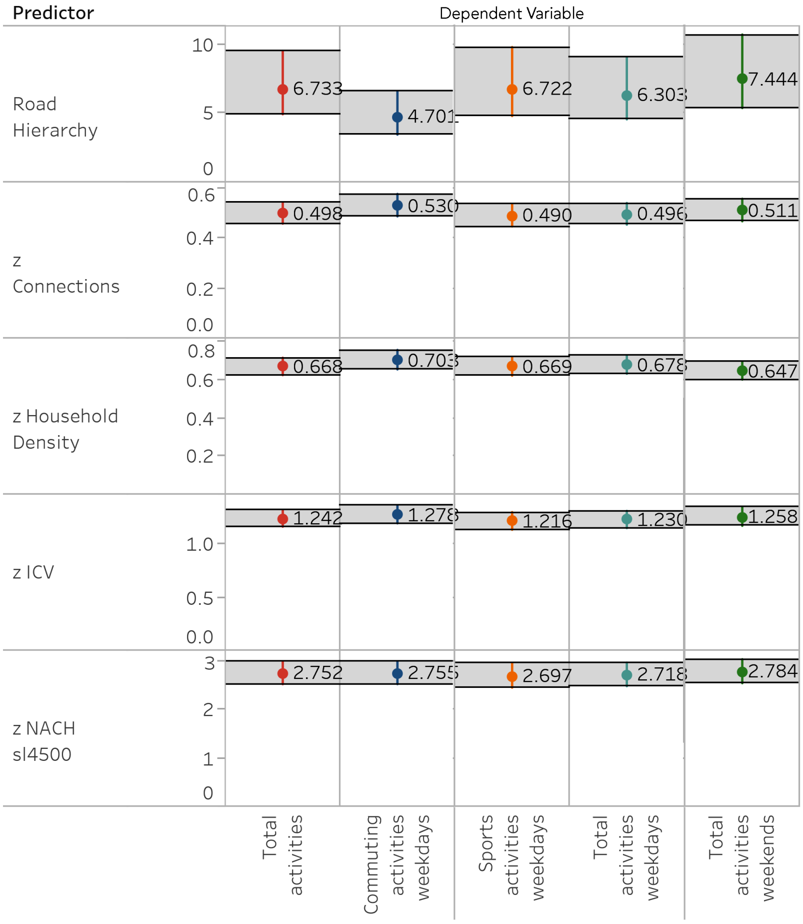

Figure 6. IRR and 2.5% – 97.5% Confidence Intervals of each predictor for full models

4. CONCLUSIONS

In this contribution, we studied the bond between urban's street network structure and cyclists' movement behaviour in a Latin American city. We used multiple regression models to understand better the importance of urban form with Space Syntax measures when controlling for other variables such as living conditions, infrastructure, and household density. We used crowdsourced data for our study including open (OpenStreetMap) and proprietary (STRAVA Metro) sources.

Spatially speaking, the results in this study evidence that street network's structure has an important influence on cycling activity, captured by the STRAVA app. When the regression models included Normalised Angular Choice, the prediction power almost doubled in relation to base models. This contributes to the evidence supporting theory of natural movement for cycling behaviour in Latin American cities. Also, we found that the inclusion of NACH in models predicting commuting cycling had a larger improvement than in those for sport cycling. Commuting cycling behaviour, therefore, seems to be closer to natural movement theory than cycling behaviour in our dataset. In morphological terms, Space Syntax's Normalised Angular Choice weighted by segment length at radius 4500m seems to be appropriate to describe the pattern of relations between elements within the urban grid and cyclists' movement in our study area, which also means that it can potentially be used as a tool for determining the layout of cycling infrastructure in the city. From the temporal scope, there was not enough evidence to show that weekday and weekend models are different.

This research revealed other important outcomes. The results for the lack of incidence of slope, contrary to theory (Saelens et al., 2003; Winters et al., 2013), indicates that although there might be some universal trends, models should adapt local or regional conditions of functioning.

Our analysis can be replicated anywhere data is available. Geographic information platforms, whether they are commercial or free, such as STRAVA Metro and OpenStreetMap provide wide global coverage and might be a viable alternative for exploring cycling patterns and how they relate to the built environment. A limitation in this study is the potential socio-economic bias from STRAVA data, that might be limited to high economic status users who possess smart devices capable of supporting the platform, and probably mobile data plans. Consequently, if there is such bias, STRAVA's universe might not represent the whole population's spectrum.

The results are relevant for cycling infrastructure design: In order to boost the use of cycleways, bicycle parking, and other facilities, urban planners and designer should select locations with high values of normalised choice. Upon further research in other urban areas, the inclusion of space syntax measures might be recommended as a design guideline for cycling infrastructure.

Future research could improve the regression model, as well comparing different cities to test the nature of results. For example, a deeper, more detailed investigation on the cyclists themselves might help illustrating why there is no apparent difference of behaviour between weekends and weekdays, and to test if movement responds to patterns of segregation and crime. The inexistence of a differentiated road network for cyclists in Manta is a big window of opportunity for informed urban design using tools and methods like the ones we displayed here, to test scenarios. On-site information should be taken, to validate the results obtained with this dataset. Finally, regression models should be included in the process of designing a network of cycling infrastructure, as an evidence-generator tool.

REFERENCES

Al-Sayed K, Turner A, Hillier B, et al. (2014) Space Syntax Methodology. Bartlett School of Architecture, UCL.

Bolton T, Francis N and Froy F (2017) The impact of Space Syntax on urban policy making. In: Proceedings of the 11th Space Syntax Symposium (eds T Heitor, M Serra, J Pinelo Silva, et al.), 2017, pp. 48.1–48.13. Instituto Superior Técnico.

Chang HL and Chang HW (2009) Exploring recreational cyclists' environmental preferences and satisfaction: Experimental study in Hsinchu technopolis. Environment and planning. B, Planning & design 36(2): 319–335.

Chiaradia A, Hillier B and Schwander C (2009) Spatial Economics of Crime. In: 7th International Space Syntax Symposium, Stockholm, Sweden, 2009. Available at: http://www.sss7.org/Proceedings/08%20Spatial%20Configuration%20and%20Social%20Structures/017_Chiaradia_Hillier_Schwander.pdf.

Fagnant DJ and Kockelman K (2016) A direct-demand model for bicycle counts: the impacts of level of service and other factors. Environment and planning. B, Planning & design 43(1): 93–107.

Guerra E, Caudillo C, Monkkonen P, et al. (2018) Urban form, transit supply, and travel behavior in Latin America: Evidence from Mexico's 100 largest urban areas. Transport Policy 69. Elsevier: 98–105.

Heinen E, van Wee B and Maat K (2010) Commuting by Bicycle: An Overview of the Literature. Transport Reviews 30(1). Routledge: 59–96.

Hillier B (2007) Space is the machine. Space Syntax.

Hillier B and Hanson J (1984) The Social Logic of Space. Cambridge: Cambridge University Press.

Hillier, B. & Iida, S. (2005), Network and psychological effects in urban movement, In: A.G. Cohn and D.M. Mark (Eds.): COSIT 2005, LNCS 3693, pp. 475

Hoehner CM, Soares J, Parra Perez D, et al. (2008) Physical activity interventions in Latin America: a systematic review. American journal of preventive medicine 34(3). Elsevier: 224–233.

INEC, (2018). INEC. [On-line]
Available at: http://www.ecuadorencifras.gob.ec/base-de-datos-censo-de-poblacion-y-vivienda/[Accessed: April 2018].

Kolovou I, Gil J, Karimi K, et al. (2017) Road Centre Line Simplification Principles for Angular Segment Analysis. In: Proceedings 11th INTERNATIONAL SPACE SYNTAX SYMPOSIUM, 2017, pp. 163.1–163.16. Instituto Superior Técnico.

Law S, Sakr FL and Martinez M (2014) Measuring the Changes in Aggregate Cycling Patterns between 2003 and 2012 from a Space Syntax Perspective. Behavioral Sciences (4): 278–300.

MAGAP, (2012). Programa SIGTierras. [On-line]
Available at: http://www.sigtierras.gob.ec/descargas/
[Accessed: May 2017].

Manum, B., & Nordstrom, T. (2013). INTEGRATING BICYCLE NETWORK ANALYSIS IN URBAN DESIGN: Improving bikeability in Trondheim by combining space syntax and GIS-methods using the place syntax tool. In Y. O. Kim, H. T. Park, & K. W. Seo (Eds.), Proceedings of the Ninth International Space Syntax Symposium (p. 028:1-028:14). Seoul. Retrieved from http://sss9sejong.or.kr/paperpdf/gusd/SSS9_2013_REF028_P.pdf

Manum, B., Nordström, T., Gil, J., Nilsson, L., & Marcus, L. (2017). Modelling bikeability; Space syntax based measures applied in examining speeds and flows of bicycling in Gothenburg. International Space Syntax Symposium, (July), 1379–1394. Retrieved from http://www.11ssslisbon.pt/docs/proceedings/papers/89.pdf%0A

Ministerio del Interior, 2019. Ministerio del Interior. [On line]
Available at: http://cifras.ministeriodelinterior.gob.ec/comisioncifras/inicio.php#
[accessed on: 12 January 2019].

Musakwa W and Selala KM (2016) Mapping cycling patterns and trends using Strava Metro data in the city of Johannesburg, South Africa. Data in brief 9: 898–905.

OpenStreetMap (2016) OpenStreetMap. Available at: http://www.openstreetmap.org/ (accessed 1 May 2017).

Orellana D and Osorio P (2014) Segregación socio-espacial urbana en Cuenca, Ecuador. Analitika, Journal of Statistical Analysis 8(2): 27–38.

Parking J, Warman M and Page M (2008) Estimation of the determinants of bicycle mode share for the journey to work using census data. Transportation 1(35): 93–109.

Raford N, Chiaradia A and Gil J (2007) Space Syntax: The Role of Urban Form in Cyclist Route Choice in Central London. Available at: http://escholarship.org/uc/item/8qz8m4fz (accessed 7 August 2018).

R Core Team (2016) R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Available at: http://www.r-project.org/.

Rybarczyk G and Wu C (2014) Examining the Impact of Urban Morphology on Bicycle Mode Choice. Environment and planning. B, Planning & design 41(2). SAGE Publications Ltd STM: 272–288.

Sadarangani KP, Von Oetinger A, Cristi-Montero C, et al. (2018) Beneficial association between active travel and metabolic syndrome in Latin-America: A cross-sectional analysis from the Chilean National Health Survey 2009--2010. Preventive medicine 107. Elsevier: 8–13.

Saelens BE, Sallis JF and Frank LD (2003) Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. Annals of behavioral medicine: a publication of the Society of Behavioral Medicine 25(2): 80–91.

Strava Metro, 2015. Strava Metro. [On line] Available at: https://metro.strava.com/

Tarkhanyan L (2015) Urban design and drug crime: uncovering the spatial logic of drug crime in relation to the urban street network and land use mosaic in London. Johnson S and Vaughan L (eds). Doctoral. UCL (University College London). Available at: http://discovery.ucl.ac.uk/1464459/ (accessed 7 August 2018).

Turner A (2007) From Axial to Road-Centre Lines: A New Representation for Space Syntax and a New Model of Route Choice for Transport Network Analysis. Environment and planning. B, Planning & design 34(3). SAGE Publications Ltd STM: 539–555.

Vuong QH (1989) Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. Econometrica: journal of the Econometric Society 57(2): 307.

Winters M, Brauer M, Setton EM, et al. (2013) Mapping bikeability: A spatial tool to support sustainable travel. Environment and planning. B, Planning & design 40(5): 865–883.

Zieff SG, Musselman EA, Sarmiento OL, et al. (2018) Talking the Walk: Perceptions of Neighborhood Characteristics from Users of Open Streets Programs in Latin America and the USA. Journal of urban health: bulletin of the New York Academy of Medicine. Springer. DOI: 10.1007/s11524-018-0262-6.