**City form and well-being: What makes London neighborhoods good places to live?**

# City Form and Well-being: What makes London neighborhoods good places to live?

Alessandro Venerandi
University College London
Gower Street
London, UK
alessandro.venerandi.12@ucl.ac.uk

Giovanni Quattrone
University College London
Gower Street
London, UK
g.quattrone@cs.ucl.ac.uk

Licia Capra
University College London
Gower Street
London, UK
l.capra@ucl.ac.uk

## ABSTRACT

What is the relationship between urban form and citizens' well-being? In this paper, we propose a quantitative approach to help answer this question, inspired by theories developed within the fields of architecture and population health. The method extracts a rich set of metrics of urban form and well-being from openly accessible datasets. Using linear regression analysis, we identify a model which can explain 30% of the variance of well-being when applied to Greater London, UK. Outcomes of this research can inform the discussion on how to design cities which foster the well-being of their residents.

## CCS Concepts

•**Social and professional topics** → *Computer supported cooperative work;*

## Keywords

Urban Form; Well-being; Quantitative Analysis; Open Data

## 1. INTRODUCTION

The search for an urban form which fosters people's well-being has long been the research topic of many architects and scholars. Urban form refers to the physical and configurational features of the built environment, such as: built density, street connectivity, the amount of green areas. Probably, the first work ever carried out in this field dates as far as back as the first century BC when the Roman architect Vitruvius wrote the architectural treatise *De Architecura* and identified the method for good city planning in the grid street pattern. Since then, many others followed. On one side, urban designers and architects mostly based their theories on personal observations, thus undermining the generalizability of the results. On the other, researchers who implemented quantitative approaches focused on very few aspects of the built environment (e.g., presence of certain amenities) [9] or modeled single aspects of well-being (e.g., crime) [10]. To remedy this, we propose a quantitative method to analyze the relationship between urban form and well-being by considering multiple features of the urban environment and apply it to Greater London. The method comprises of two main steps: (i) extracting metrics of urban form and well-being from openly accessible datasets (e.g., OpenStreetMap) and (ii) performing linear regression with the metrics of urban form as independent variables and well-being index as dependent one. The outcome of this research is a set of variables of urban form which are related to different levels of well-being. Method and outcome can help urban planners and administrators in shaping cities which foster the well-being of their residents.

ACM ISBN 978-1-4503-2138-9.
DOI: 10.1145/1235

## 2. RELATED WORK

Different approaches have been used to study urban form and well-being. On the qualitative side, Jane Jacobs strongly supported the traditional compact city form, characterized by medium to high densities, highly walkable short urban blocks, and mixed use [11]. Le Corbusier was against this and favored bigger blocks, the use of cars, and the segregation of functions in specialized sectors [5]. Since these works are based on personal views, they are limited as they are hard to replicate or validate. Other researchers implemented quantitative methods. Vaughan *et al.*, for example, found that places with poor accessibility (e.g., back streets) were associated with disadvantaged classes, while places with a higher accessibility (e.g., main streets) were related to more well-off residents [17]. Other scholars studied the relationship between urban form and crime. Hillier [10] and Budd [3] studied dwelling typologies and came to the conclusion that the flat is the safest house type. Hillier also analyzed high densities in relation to crime and found that were overall beneficial against it [10]. Other researchers focused on social aspects of well-being (e.g., place attachment). Some academics reported that high urban densities enhanced social interactions [6], while low ones reduced them and also led to more car-dependent behaviors [4]. The quantitative studies above mentioned have two main limitations: (i) they focused on single aspects of the urban environment (e.g., accessibility, cul-de-sac) thus providing a limited understanding of the relationship between urban form and well-being; (ii) they modeled specific aspects of well-being (e.g., crime, place attachment) rather than more complex indexes. With the recent data revolution, large geodatasets have become easily available and computational social scientists started to investigate urban form. Quercia *et al.*, for example, quantitatively studied the relationship between people's visual perceptions and the happiness of places [14]. Naik *et al.* identified what city areas were related to different socioeconomic aspects such as wealth, safety, and uniqueness [13]. Others analyzed geographic variations of Twitter content to predict happiness [7] and well-being [16]. Inspired by this line of research, we propose a quantitative method to study the relationship between urban form and well-being in Greater London by extracting multiple metrics from openly accessible datasets. The approach then relies on linear regression to identify what specific set of metrics is related to well-being. We present the datasets used next.

# 3. DATASETS

To carry out this work, we had to access datasets with the following information: configuration of the urban environment, age of the housing stock, and well-being scores for Greater London. We extracted the information concerning the first two points from OpenStreetMap, Foursquare and OS VectoMap District. Well-being scores were extracted from the London Datastore.

**OpenStreetMap (OSM)** is an openly accessible and editable map of the world. OSM consists of three types of spatial entities: *nodes* which represent amenities, *ways* which represent roads, and *relations* which are used to group together other spatial objects (e.g., bus routes). For the purpose of this study, we focused on ways which are the basic components of the street network to later extract variables descriptive of urban form for Greater London.

**Foursquare** is a social media platform where users can share their whereabouts with friends by checking-in into places (i.e., amenities). A Foursquare place is defined by a pair of geographic coordinates (i.e., latitude, longitude), a name, and a category (e.g., theater, Japanese restaurant, pawn shop). For the aim of this study, we crawled the Foursquare places of Greater London through the official Foursquare API between the 01/04/2014 and the 06/04/2014.

**OS VectorMap District** is an accurate map of the UK in vector format. It is released by the Ordnance Survey (i.e., the official UK mapping agency) and it is freely accessible. OS VectorMap District contains information about several geographic entities (e.g., roads, buildings, bodies of water, stations). For the purpose of this work, we selected the information relative to the buildings (i.e., building footprints) of Greater London.

**London Datastore** is an official web portal with statistical data for Greater London. Information on well-being for the areas under study was extracted from this repository. The well-being score is computed as the weighted mean of 12 different non-economic indicators such as life expectancy and crime rate. In this work, we used the latest available score which dates back to 2013. We expected well-being to be related to socioeconomic factors. A wealth of studies – see for example Blanchflower and Oswlad [2] – has, in fact, reported that better-off people also tend to experience greater wellness. We thus tested this assumption by correlating well-being with the official deprivation index for the UK, the Index of Multiple Deprivation (IMD). Given the spatial nature of the data analyzed, we used a correlation technique which controlled for spatial autocorrelation (an issue which can diminish the robustness of findings). By applying this technique, we found that well-being and deprivation were correlated, with no sign of spatial autocorrelation (i.e., $r =$ -0.85, $p-value = 0$). Socioeconomic factors seemed thus to be highly explanatory of well-being; however, they seemed not to fully explain it. We argue that urban form also plays a role in defining well-being. We next present a method to quantify to what extent urban form can serve to indicate well-being.

# 4. METHOD

The method comprises of two steps: computation of metrics of urban form per areal unit of analysis and regression analysis with well-being as dependent variable.

## 4.1 Unit of analysis

The areal unit of analysis chosen for this study was the *ward*, which is the areal unit for which well-being data is available. Wards are UK official administrative boundaries and represent electoral areas as well as ceremonial entities. Greater London is subdivided in 625 wards with an average size of 255 hectares. In the remainder of this paper, we will refer to wards as "neighborhoods".

## 4.2 Metrics

After having obtained the datasets presented in the previous section, we extracted a set of descriptors of urban form. The chosen metrics can be broadly grouped in 3 main categories: Street Network Configuration, Amenities' Offering, and Neighborhood Age.

*Street Network Configuration (SNC).* We identified a total of 7 metrics related to the configuration of the street network. Three are proposed by the authors of this paper and were extracted from OSM. These are:

- Dead-end density ($deden$) is calculated as the ratio between the number of dead-ends and the areal unit. The hypothesis behind this metric is that dead-end roads might be detrimental to people's well-being as they promote a life style based on the use of cars, which diminishes the number of travels made by foot. This decreases the amount of physical activity undertaken by residents and might consequently lead to less well-being.

- Green areas ($ga$) is computed as the ratio between the surface covered by green areas and the areal unit. The hypothesis is that presence of green areas is beneficial for citizens' wellness, as it might reduce air pollutants, encourage more physical activity, and also have a therapeutic function against stress.

- Irregularity ($irr$) is calculated as the standard deviation of the node degrees per areal unit normalized on the average node degree. Irregularity measures whether a street network is more similar to a grid or to an organic structure. In this case, the hypothesis is that a more organic-shaped street network makes travels by car harder, thus increasing the share of trips by foot or bike. This, in turn, might raise the levels of well-being.

Four metrics were derived from previous works. The first three were extracted from OSM, while the last one was extracted from the OS VectorMap District dataset. These are:

- Connected node ratio (*cnr*) is computed as the ratio between the number of non dead-end intersections and the total number of intersections per areal unit [12]. It is a measure of network connectivity and walkability. These were considered necessary factors for thriving neighborhoods in Jane Jacobs' work [11].

- Intersection density (*iden*) is measured as ratio between the number of non dead-end intersections and the total number of intersections per areal unit [12]. It is a measure of street network density. Previous qualitative work [11] claimed that density is an important feature for a prosper urban environment.

- Betweenness (*bet*) measures the property of a place to lie in-between others and it has been shown to be strongly related to key dynamics in cities, for example street quality [15]. This is calculated at the street network level; however, for the aim of this work, it has been aggregated at the level of the areal unit of analysis by computing its maximum value.

- Percentage of open space (*osperc*) is calculated as percentage ratio between the surface of the areal unit minus its built up part, divided by the surface of the areal unit [1]. This measure describes a city area in terms of presence of open spaces. This directly relates to density, as fewer open spaces correspond to a more condensed urban fabric. The theoretical background for this metric is the same as the one presented for Intersection density [11].

*Amenities' Offering (AO).* In a previous work, the same authors of this paper found a relationship between the presence of specific amenities and socioeconomic deprivation [18]. In this work, we explored whether a similar relationship exists with respect to citizens' well-being. To do so, we used *Offering Advantage* (OA), a formula that captures whether a city area offers more of a certain amenity compared to the average offering of that amenity for the whole city under study. We applied this calculation to the 332 types of amenities present in the Foursquare dataset of Greater London. At this point, we checked what categories were related to well-being by means of correlation analysis with control for spatial autocorrelation. Since we were testing hundreds of values, the chance of obtaining false positive results increased. We thus applied a technique, called False Discovery Rate (FDR), to exclude those variables which were false positives (i.e., $q$-value>0.05). At the end of this process, we obtained 11 variables. For matter of brevity, we mention just 3: OA of fried chicken restaurants, OA of Caribbean restaurants, and OA of factories.

*Neighborhood Age (NA).* As we presented in section 2, building traditional neighborhoods or modernist ones was, and still is, a highly contested topic. To test this idea, we introduced a metric which measured whether a city area offered more historic properties compared to the average offering of Greater London. To this end, we chose 1928 as temporal threshold, the year of the Congres Internationaux de Architecture Moderne (CIAM). This represented a turning point in the planning practice worldwide: from that point

| *Topic* | *Ind. variable* | $p$-value | $\beta$ |
|---|---|---|---|
| | *IMD* | *** | -0.865 |
| | adj. $R^2 = 0.75$ | | |
| SNC | *deden* | * | -0.144 |
| | *ga* | * | 0.043 |
| | *irr* | * | -0.042 |
| | *cnr* | | -0.091 |
| | *iden* | * | 0.170 |
| | *bet* | *** | 0.146 |
| | *osperc* | | 0.032 |
| AO | *fried chicken* | | -0.002 |
| | *dentist* | | 0.024 |
| | *cricket* | | 0.016 |
| | *grocery store* | | -0.028 |
| | *factory* | | -0.009 |
| | *golf course* | | 0.030 |
| | *Italian rest.* | *** | 0.071 |
| | *African rest.* | * | -0.039 |
| | *Caribbean rest.* | * | -0.039 |
| | *wine shop* | *** | 0.072 |
| | *salon barbershops* | | -0.006 |
| NA | *preciam* | *** | 0.089 |
| | adj. $R^2 = 0.30$ | | |
| | Moran's index = 0.15 | | |

Table 1: Model's outcome. Symbols' interpretation: '.' $p$-value $< 0.1$, '*' $p$-value $< 0.05$, '**' $p$-value $< 0.01$, '***' $p$-value $< 0.001$.

on, the modernist approach (e.g., "tower in the park") became the norm for many new developments. We thus calculated the OA of those properties which were built before the 1928: OA of preciam properties (*preciam*).

### 4.3 Linear regression

After having defined the areal unit of analysis, we computed the morphological metrics presented in the previous section for the 625 London neighborhoods. At this point, we took two steps to make variables ready to be introduced in a linear regression model. To have more understandable and comparable regression coefficients, we first normalized the variables which were skewed through exponentiation and, second, we calculated their $z$ scores. In section 3, we showed that socioeconomic factors did not fully explain well-being. To quantify what urban form was able to capture of what the socioeconomic factors could not explain, we performed the following steps. Firstly, we performed a regression analysis with socioeconomic deprivation as independent variable and well-being as dependent one. Secondly, we used the residuals of this model as dependent variable in a second regression with the metrics of urban form as independent variables. The last step consisted in checking whether residuals showed or not spatial autocorrelation. To ascertain this point, we used a technique called Moran's test.

## 5. RESULTS AND DISCUSSION

The result of the correlation analysis presented in section 3 was confirmed by the outcomes of the regression analysis. IMD was able to explain 75% of the variance of well-being. To quantify the explanatory power of urban form, we input the residuals of this first model into a second one, with the 19 metrics descriptive of the urban environment as independent variables. We present a summary of the results below, for more details please refer to Table 1. The model showed an adjusted $R^2$ value of 0.30 with the most important and significant regression coefficients being Intersection density

($\beta$ = 0.17), Betweenness ($\beta$ = 0.15), and Dead-end density ($\beta$ = -0.14). The Moran's test for this model was significant with an observed index of 0.15, meaning that only small traces of spatial autocorrelation were present in the residuals. The results highlighted the existence of a relationship between urban form and well-being. By looking at signs and strengths of $\beta$ coefficients, we suggest that the typical London neighborhood with good levels of well-being has the following characteristics: it is well-connected and easily accessible (negative value of *deden* and positive value of *bet*), it is characterized by green areas and predominance of historic properties (positive values of *ga* and *preciam*), its street network is dense and tends to be grid-shaped (positive value of *iden* and negative value of *irr*). For what concerns the Foursquare venues analyzed, Italian restaurants and wine shops tend to be present in London neighborhoods with good well-being levels (positive values of *Italian rest.* and *wine shops*); conversely, African and Caribbean restaurants tend to be absent (negative values of *African rest.* and *Caribbean rest.*). Findings seem to be consistent with what Jane Jacobs identified as thriving neighborhoods, that is well-connected, dense, and walkalbe places characterized by buildings of different ages [11]. We argue that this kind of neighborhood is linked to well-being for the following reasons: a dense and well-connected street network might encourage more walking and less driving which, in turn, might positively affect well-being in two ways: more physical activity and less air and noise pollution. A dense urban environment can also enhance social interactions which might foster sense of belonging and perceived safety thus positively affecting the psychological well-being of a person. Furthermore, the presence of green areas might decrease levels of stress and air and noise pollution. Italian restaurants and wine shops with their offering of quality food might be linked to better eating habits, which, in turn, might positively impact people's well-being; conversely, African and Caribbean restaurants, which might sell less healthy food, might be associated with a worse diet and thus to lower levels of well-being. This work can have practical as well as theoretical implications. It can help urban planners and city administrators to design neighborhoods which foster the wellness of their residents. It can be used by researchers to test other urban theories and check whether results hold for other geographic contexts and for different time frames. This work has three main limitations. Firstly, it is confined to London and thus results only hold, at the moment, for this city. The second limitation concerns the direction of causality for the relationship between urban form and well-being. This study, in fact, does not clarify whether specific urban features attract people with higher well-being or whether the arrival or presence of residents with more wellness attract certain features of the built environment. We suppose that these two causes are combined thus making very hard for any study to disentangle them. The last limitation is linked to the explanatory power of our model. Although the model presented could explain the 30% of what the socioeconomic index is not able to capture in terms of well-being, a large part remains unexplained. It is plausible that other explanatory variables are missing (e.g., built density).

## 6. REFERENCES


[1] D. Banister, S. Watson, and C. Wood. Sustainable cities: transport, energy, and urban form. *Environment and Planning B*, 24:125–144, 1997.

[2] D. G. Blanchflower and A. J. Oswald. Well-being over time in Britain and the USA. *Journal of public economics*, 88(7):1359–1386, 2004.

[3] T. Budd. Burglary of domestic dwellings: Findings from the British Crime Survey. 1999.

[4] R. W. Burchell, N. A. Shad, D. Listokin, H. Phillips, A. Downs, S. Seskin, J. S. Davis, T. Moore, D. Helton, and M. Gall. *The costs of sprawl-revisited*. Number Project H-10 FY'95. 1998.

[5] L. Corbusier. *Manière de penser l'urbanisme: urbanisme des CIAM*, volume 1. Éditions de l'architecture d'aujourd'hui, 1946.

[6] A. Duany, E. Plater-Zyberk, and J. Speck. *Suburban nation: The rise of sprawl and the decline of the American dream*. Macmillan, 2001.

[7] M. R. Frank, L. Mitchell, P. S. Dodds, and C. M. Danforth. Happiness and the patterns of life: A study of geolocated tweets. *Scientific reports*, 3, 2013.

[8] L. Freeman. The effects of sprawl on neighborhood social ties: an explanatory analysis. *Journal of the American Planning Association*, 67(1):69–77, 2001.

[9] B. Giles-Corti and R. Donovan. Socioeconomic status differences in recreational physical activity levels and real and perceived access to a supportive physical environment. *Preventive Medicine*, (35):601–611, 2002.

[10] B. Hillier and O. Sahbaz. An evidence based approach to crime and urban design, or, can we have vitality, sustainability and security all at once. *R. Cooper, G. Evans, C. Boyko, C.(eds.), Designing Sustainable Cities: Decision-making Tools and Resources for Design, Wiley Blackwell*, pages 163–186, 2009.

[11] J. Jacobs. *The life and death of great American cities.* New York: Random House, 1961.

[12] W. Marshall and N. Garrick. The shape of sustainable street networks for neighborhoods and cities. In *Congress for the New Urbanism XVII, Denver, CO*, 2009.

[13] N. Naik, J. Philipoom, R. Raskar, and C. Hidalgo. Streetscore–predicting the perceived safety of one million streetscapes. In *CVPRW 2014*.

[14] D. Quercia. Urban: Crowdsourcing for the good of London. In *WWW 2013*.

[15] A. M. Remali, S. Porta, and O. Romice. Correlating street quality, street life and street centrality in Tripoli, Libya. *The Past, Present and Future of High Streets*, 2014.

[16] H. A. Schwartz, J. C. Eichstaedt, M. L. Kern, L. Dziurzynski, R. E. Lucas, M. Agrawal, G. J. Park, S. K. Lakshmikanth, S. Jha, M. E. P. Seligman, and L. H. Ungar. Characterizing geographic variation in well-being using tweets. In *ICWSM 2013*.

[17] L. Vaughan, D. L. C. Clark, O. Sahbaz, and M. M. Haklay. Space and exclusion: Does urban morphology play a part in social deprivation? *Area*, 37(4):402–412, 2005.

[18] A. Venerandi, G. Quattrone, L. Capra, D. Quercia, and D. Saez-Trumper. Measuring urban deprivation from user generated content. In *CSCW 2015*.