

The Interaction Between Open Spaces and Mental Wellbeing :

A Study In Greater London

Vikki Houlden

Abstract—With increased migration of rural dwellers to the world's cities, it has never been more crucial for urban developers to consider the effects these environments may have on residents. As such, this paper seeks to understand how the distribution of open spaces may influence the mental wellbeing of the population within each London Ward. Weak correlations are found between different types of spaces and wellbeing, emphasising the complex relationships at play. A classification tool is implemented to discover how these environmental characteristics may be used as a predictor for likely levels of wellbeing. These findings demonstrate how people interact with and are affected by their home and work environments, and highlight the importance of considering the user's needs in design. While more research is required in this area, in order to include a greater range of built environment variables and draw more quantitative conclusions, it is clear that developers have a responsibility to ensure cities are beneficial to the population's mental health and wellbeing.

I. INTRODUCTION

SIGNIFICANT research has suggested that there exists a relationship between the built environment and both the physical and mental health of its residents. While it is evident that these interactions exist, it remains to be seen which attributes of a person's locality are of most importance. The purpose of this study is therefore to examine the relationship between wellbeing and open spaces in Greater London, to examine which types of spaces may be most beneficial, and consider how this may come to be. Existing research shall firstly be considered, to understand the current situation in the fields of mental health, wellbeing and greenspace benefits. From looking at this work, it is anticipated that larger and more natural areas may have the greatest benefits to the population. Therefore, different public spaces in each London Ward and the relevant probability of positive mental wellbeing are correlated, to observe the importance of different open space factors. In order to test the hypothesis that these characteristics may as such be predictive of better mental wellbeing, a classification model is then built. The strength and reliability of these

observations are then discussed, followed by recommendations for further work. The application of this knowledge is then emphasised, with recommendations for urban design being made.

10th January 2015

II. BACKGROUND

A. Greenspaces

It has been shown that those living in more natural environments are likely to have improved overall health, with greenery having been linked to a variety of health factors, ranging from levels of stress to recovery times following surgery [1]. Notably, Alcock et al recently demonstrated how those moving to greener environments benefited from increased mental health, which may in turn increase general health and happiness [2]. Fig.1 represents the benefits observed for those now residing in more natural areas.

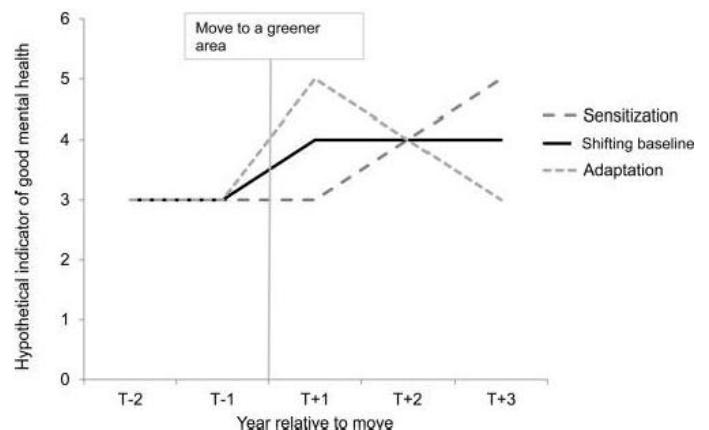


Fig. 1. "Graph Showing Hypothetical Temporal Patterns of Moving to a Greener Area" [2]

However, there are many other factors to consider, which may include the types, quantities and accessibility of greenspaces. Aside from suggestions that abundance of nature may increase physical activity and some investigations into stress hormones, limited research has examined the causal relationships in this field.

B. Wellbeing

The idea of wellbeing is still emerging, gaining in prominence and importance in policy. Rather than simply an absence of mental difficulty, the concept of wellbeing is based on a state of positive mental health, encompassing not just satisfaction but longer term happiness, enjoyment and ambition. The government currently defines it as a positive physical, social and mental state over a prolonged time, while others seek to more physically characterise key components [3]. Of perhaps the greatest political importance, the UK Government now measures the countries prosperity by the overall wellbeing, in addition to the traditional GDP [3]. Seligman describes a five-part Wellbeing model as comprising Positive Emotion, Engagement, Relationships, Meaning, and Accomplishment (PERMA) [4]. The idea may be represented simply by Dodge et al [5], as a balance between personal and external factors.

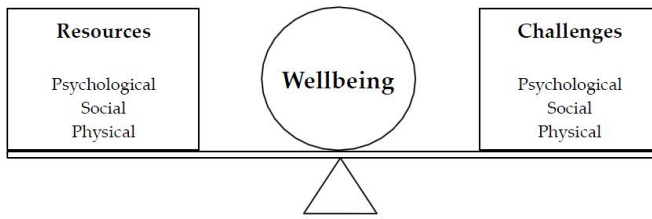


Fig. 2. "Definition of Wellbeing" [5]

As such, it is evident that an individuals Wellbeing may be dependent upon many variables, some of which are affected by their environment.

C. Wellbeing in the Built Environment

With an increasing number of people living in Urban environments, and an expected 50% of the global population living in cities by 2020, the built environment is an ever more important factor in the mental health and wellbeing of its residents [6]. Given that people will spend the majority of their time in a built environment- whether their home, route to work at the office or socialising- designers have a responsibility to ensure that places are beneficial to health. This is particularly applicable with regard to urban planning, where open spaces may be incorporated into the street network to improve the area. The types of spaces may also be important, with sports facilities and playing fields offering opportunities for physical activities, parks designed for socialising and natural environments ideal for exploring. While clearly offering a range of benefits, it is as yet unclear which types of spaces may be most critical in improving wellbeing in a region. However,

this is a complex field to investigate, given the subjective nature of ones experience of their environment, and the difficulties of obtaining reliable data.

D. Measuring Wellbeing

As such a complex idea, researchers differ over what constitutes an absolute measurement of wellbeing. From an objective view, an analysis may constitute external factors, such as family situation, living environment and income. More subjectively, it may consider a persons individual feelings over a period of time. One of the most prominent, the Warwick-Edinburgh-Mental-Wellbeing- Scale (WEMWBS), implements a 14-point self-report questionnaire looking at optimism, interest and relationships, with a 5-level answer scale [7]. This therefore gives a temporal insight into a range of internal factors. For a holistic view, as used in this report, the idea of wellbeing shall encompass both objective (external) and subjective (internal) factor; as discussed in more detail in the Wellbeing Data section .

E. Open Spaces in Greater London

As the largest Urban area in Britain, London has over 60,000 hectares of open space, ranging from waterways such as the Thames, to Royal and local square parks and green corridors, and with a population exceeding 9million, these spaces have the potential to influence the quality of life of both residents and workers in the city [8]. It is estimated that wellbeing may be improved in areas which are more open and offer potential for more natural, greener environments. Therefore, this offers perhaps the greatest scope for investigation, given the diverse range of environments and characteristics.

III. THE DATA

To observe the relationship between greenspace and wellbeing, two primary sources of data were used: wellbeing and population data from the London Data Store and open space data from the Greenspace Information for Greater London records.

A. A London Dataset

The London Data Store is an online depository for all types of open-source data relating to the Capital. The 2013 London Ward Profiles database includes almost 1Mb of data pertaining to each borough and ward. Of particular interest were population, area and geographical descriptors, in addition to built-up and open space statistics, which were all then extracted for this work. As all of this data was already processed to give an average for each Ward and Borough, this meant it was immediately accessible and available for implementation.

B. Wellbeing Data

The London Wards Wellbeing Probability Scores file contains a user-interactive database, where a range of different wellbeing attributes may be weighted by importance according to the purposes of the investigation; this allowed for a holistic measure, accounting for both personal and external factors. 2013 source data is also included for reference and extraction. To define a holistic measure of wellbeing, the following attributes were included and weighted equally: life expectancy, incapacity benefits claimants rate, unemployment rate, income support claimant rate, crime rate, deliberate fires, GCSE point scores, unauthorised pupil absences, children in out-of-work households, public transport accessibility scores and self-reported subjective wellbeing scores. The final value calculated indicates the probability that the population of each ward experiences positive wellbeing (Fig 3.).

Theme	Health	Economic security	Safety	Children	Access	Happiness
Indicator name	Life Expectancy	Unemployment rate	Crime rate	Unauthorised Pupil Absence	Public Transport Accessibility Scores	Subjective well-being average score
Importance	10	10	10	10	10	10
10 = very important and 0 = not important at all						

Fig. 3. "Interactive Wellbeing Dataset"

Geographical data is also included for visual representation of wellbeing distribution, highlighting clusters of higher and lower probability positive mental wellbeing scores (Fig. 4). It is evident from this that the probability of positive mental wellbeing varies considerably across London, giving great scope for investigation into the causes of these differences.

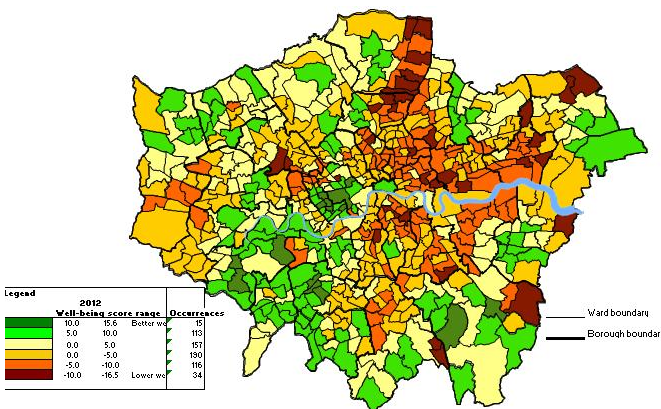


Fig. 4. "Wellbeing in Greater London"

C. GiGL Open Space Data

Two up to date 2014 data sets, of 3Mb combined, were obtained with permission of GiGL. The first contained the details of each individual open space across London, classified by borough; the second contained relevant information pertaining to each Ward. While a complicated and very time consuming task, combining these data sets allowed space descriptors, such as type, use, access and geographical location, to be localised by both ward and borough, for identification; this will be discussed in greater detail in section V-A.

IV. A HYPOTHESIS

With the knowledge that greenspaces may be beneficial to mental health, it follows that there could exist a positive relationship between open spaces and wellbeing. As such, the subsequent hypotheses have been drawn:

- 1) Areas with higher levels of open space may have an increased possibility of higher wellbeing
- 2) Wards with more natural, green environments may be most highly correlated with positive wellbeing
- 3) Levels of greenspace may be used as a predictor for better mental wellbeing.

V. DATA CLEANING

Once an agreement between the author and GiGL had been signed, access to all of the required data was granted. However, because of the different classifications used between the two data sets, extensive initial processing was required to combine the data sets.

A. Combining Data Sets

The first data set contained the dimensions of open spaces within each ward and borough, while the second comprised the name, borough, size and PPG17 (Planning Policy Descriptor) classification. In order for the two sets to be effectively combined, those open spaces which lay in multiple wards were assigned to their majority location, while the total area was still kept the same. All spaces which fell outside the Greater London area were dismissed as irrelevant to the study. These were then matched between ward and borough to give an appropriate level of geographical information. Using the PPG17 descriptors, the open spaces were then sorted according to type, and manually summed to give a total area of each type of space for each ward. The types included were: allotments, community gardens and city farms, amenity, cemeteries and churchyards, spaces for children and teenagers, civic spaces, green corridors, natural and semi-natural urban greenspace,

other, other urban fringe, outdoor sports facilities and parks and gardens. While this constituted a significant data reduction, it was necessary so that it could be combined with the wellbeing data. All spaces without a classification were assigned to other, although these were then removed from the set, along with other urban fringe, as it was decided that these did not contribute as qualitative descriptors of the spaces. This data was then in a state to be combined with the London datasets; it should be noted that total open space statistics for each ward was checked between the GiGL and London data before being included. As data describing the City of London was only presented by the Data store as a whole entity, the greenspace for each ward was combined into one. In order to make the quantities of each type of space comparable, population data was used to calculate area per capita for each type.

B. Missing Values

As there were several wards which did not include quantitative information for at least one type of open space, either because they had no classification or were not collected, it was assumed that there was no open space of this type in this class, and so all missing values were set to 0.

C. Outliers

To ensure correlations could correctly be drawn from this data, it was decided to alter outlying points which lay more than 3 standard deviations from the mean each outlier was replaced with the mean value for that attribute. This process was manually applied to each open space type, although not if there were more than 3, as this may have statistical significance. As the attributes civic spaces and community gardens were applicable to only 10% of the wards, and did not correlate with a notable wellbeing score, these were also removed from the set.

D. Groups

To enable data classification, wellbeing scores were assigned into groups as follows

Range	Classification
<-10	Very low
<-2	Low
<2	Average
<10	Good
<20	Very good

Fig. 5. "Arbitrary Wellbeing Classification Table"

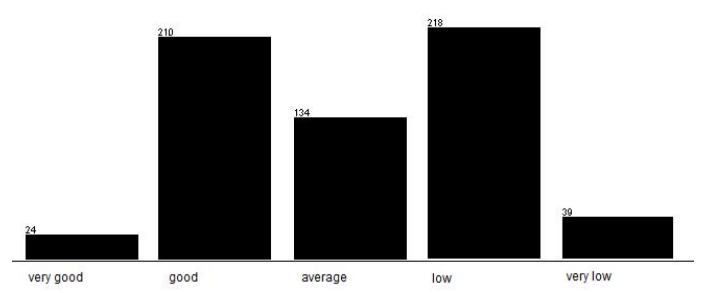


Fig. 6. "Visual Distribution of Classes"

VI. DATA PROCESSING

A. Correlations

To begin observing patterns within the data, wellbeing probability scores were plotted against each open space characteristic in turn. The linear equation of each correlation was calculated, as well as the coefficient of determination. The r^2 value is used to examine how the variations in wellbeing may be explained by the different open space factors; this was used to establish the strength of the relationship between wellbeing and each descriptor.

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad (1)$$

$$SS_{(tot)} = \sum_i (y_i - \bar{y})^2 \quad (2)$$

$$SS_{(res)} = \sum_i (y_i - f_i)^2 \quad (3)$$

$$r^2 = 1 - \frac{SS_{(res)}}{SS_{(tot)}} \quad (4)$$

Where \bar{y} is the mean value of y , which in this case is the Open Space Area of each type, and f_i refers to the estimated value of y at each point i , that of the Line of Best Fit. This equation therefore compares the variance at each point $SS_{(tot)}$ to the residual sum of squares.

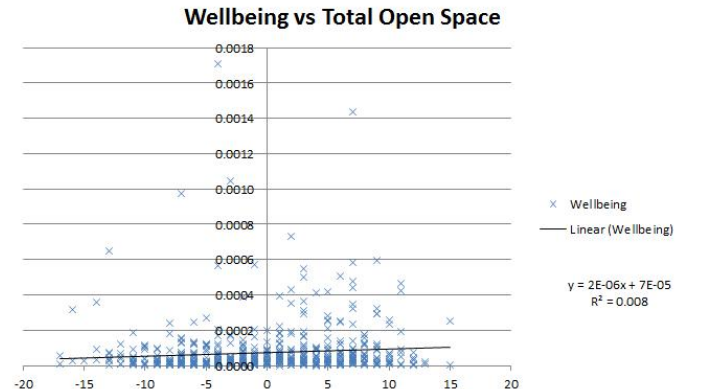


Fig. 7. "Wellbeing and Total Open Space Per Population"

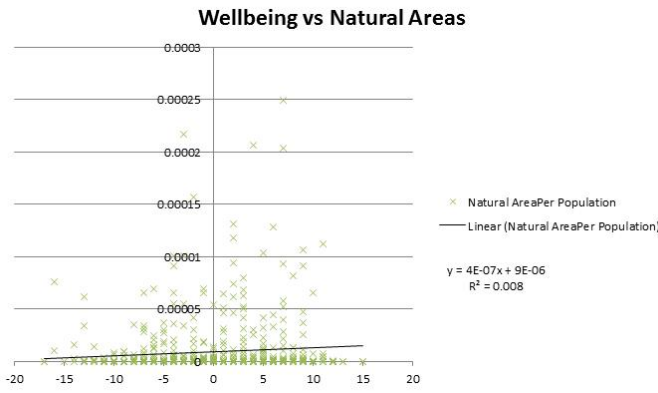


Fig. 8. "Wellbeing and Total Natural Spaces Per Population"

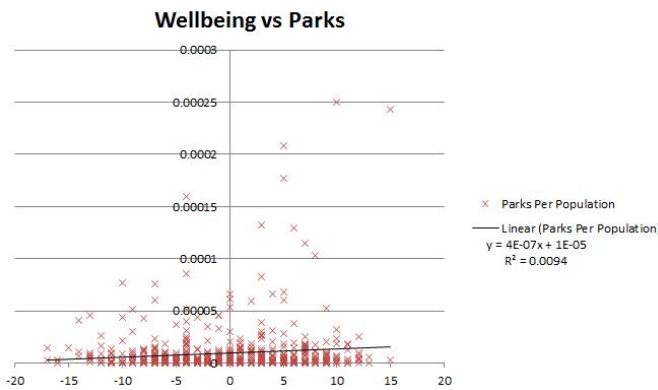


Fig. 9. "Wellbeing and Total Park and Playground Space Per Population"

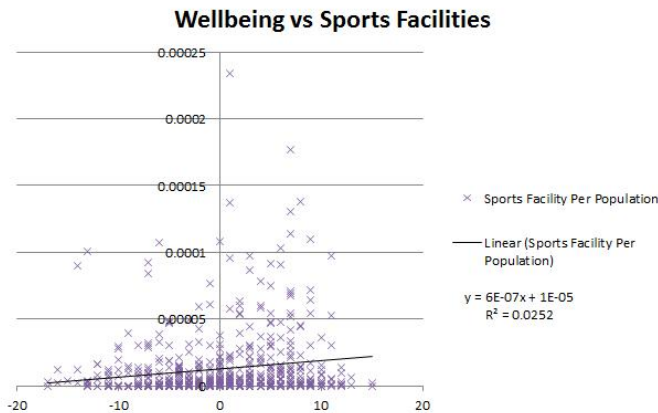


Fig. 10. "Wellbeing and Total Area of Sports Facilities Per Population"

As shown, it was found that total open space, parks, natural areas and sports facilities were correlated positively with wellbeing. The coefficients of determination were much closer to 0 than 1, meaning that the correlation was weak. By examining the plots, however, it may be inferred that there is a greater distribution of area

where wellbeing was higher, suggesting that areas with greater quantities of these types of open space are more likely to have a positive probability of wellbeing than those with lower areas of each open space.

B. Building a Classifier

WEKA software was used to investigate whether a level of wellbeing may be determined from open space characteristics. Wellbeing classes, as described above, were included, along with 7 nominal attributes: cemeteries, areas for children and teenagers, green corridors, natural areas, sports facilities and park areas per population in each ward; this totalled 626 data points. A K-nearest neighbours approach was used with K=2 to create a wellbeing class descriptor; a summary of the results is shown below.

Correctly Classified Instances	390	62.4 %
Incorrectly Classified Instances	235	37.6 %
Kappa statistic	0.4905	
Mean absolute error	0.146	
Root mean squared error	0.2695	
Relative absolute error	51.0557 %	
Root relative squared error	71.3171 %	
Total Number of Instances	625	

Fig. 11. "Summary Statistics from Weka K-Nearest-Neighbour"

With over 60% of the instances correctly classified, this represents a reasonable fit, implying that comparable levels of wellbeing are associated with similar open space attributes. The confusion matrix can be examined to discover which classifications and misclassifications have been made.

== Confusion Matrix ==

a	b	c	d	e	<-- classified as
24	0	0	0	0	a = very good
3	131	0	0	0	b = average
10	48	160	0	0	c = low
0	13	17	9	0	d = very low
9	50	71	14	66	e = good

Fig. 12. "Confusion Matrix from Weka K-Nearest-Neighbour"

The Kappa statistic and the Receiver Operating Characteristic (ROC) should also be considered to determine the robustness of the classifier, as follows:

$$\kappa = \frac{Pr(a) - Pr(e)}{1 - Pr(e)} \quad (5)$$

Kappa compares the number of correctly classified instances $Pr(a)$ to the probability that there is random agreement between the attributes $Pr(e)$, where a value of 0 infers that the data matches are entirely random and 1, that the variables are in complete agreement. As such, a mid-range value such as this signifies that the matches are not coincidental, while the agreement is not strong. Therefore further clarification is required.

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	1	0.037	0.522	1	0.686	0.983	very good
	0.978	0.226	0.541	0.978	0.697	0.909	average
	0.734	0.216	0.645	0.734	0.687	0.882	low
	0.231	0.024	0.391	0.231	0.29	0.964	very low
	0.314	0	1	0.314	0.478	0.892	good
Weighted Avg.	0.624	0.127	0.722	0.624	0.594	0.9	

Fig. 13. "Detailed Accuracy from Weka K-Nearest-Neighbour"

$$TPR = \frac{TP}{P} = \frac{TP}{TP + FN} \quad (6)$$

$$FPR = \frac{FP}{N} = \frac{FP}{TP + FN} \quad (7)$$

$$ROC = \frac{TPR}{FPR} \quad (8)$$

Where $TP = TruePositives$, $FP = FalsePositives$, $TN = TrueNegatives$, $FN = FalseNegatives$, $P = TotalPositives$ and $N = TotalNegatives$.

The ROC value (Fig. 11) is then used to compare the True TPR and False Positive Rates FPR (Fig.10), and therefore ranges from 0.5 to 1, with an output of 0.5 showing that matches classes are all due to chance. With a value here of 0.9, it can be ascertained that there exists some positive correlation between wellbeing and types of open space. This can also be represented graphically, with the colour line indicating the threshold.

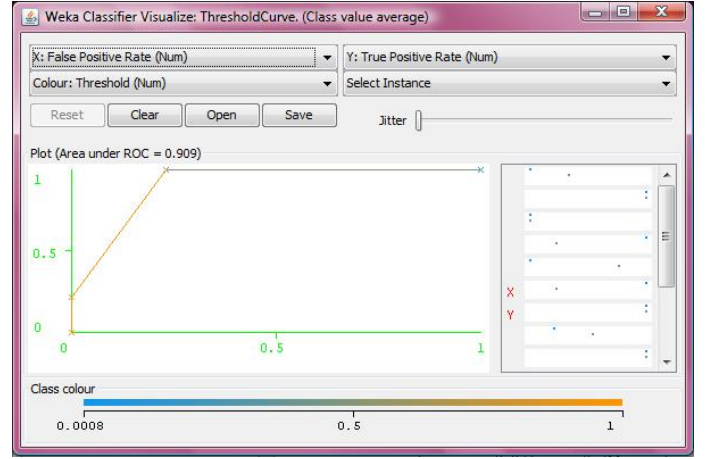


Fig. 14. "Receiver Operating Characteristic Total for Wellbeing Classes"

VII. ANALYSIS

As shown in section (VI-A), there exists a positive yet weak correlation between wellbeing and several greenspace characteristics, most notably quantity of sports facilities per population, but also amount of nature, park and total open spaces per capita in each ward. Based upon visual inspection, it appears that reasonably low coefficients of determination are due to the change in distribution as probability of positive wellbeing increases. In effect, as wellbeing rises, there seems to be a greater chance that there is a higher area of open space in that ward. While this is inferred, it implies that open space is but one of the factors which influence a persons wellbeing. It could be suggested that, as the strongest relationship was observed with sports facilities, that exercise could be encouraged in these wards and promotes both physical and mental health. This also emphasises the complexity of the variables here, and highlights the need for consideration of further built environment characteristics. When considering the K-Nearest-Neighbour Model, a significant percentage of data points were correctly assigned to their wellbeing class, a positive classification reinforced by both the Kappa and ROC Values, along with the confusion matrix, which all together showed that the findings were indeed significant. This implies that wellbeing may be a function of open space characteristics, although with wellbeing being a concept of some complexity, this model would need to take account of a much greater range of internal and external factors in order to have an improved representation and therefore accuracy. It is, however, evident that open spaces may play an important role in influencing the wellbeing of urban residents.

VIII. CONCLUSIONS

It can be concluded that, as hypothesised, there exists a relationship between mental wellbeing and the quantity of open space within Greater London. The correlations between higher wellbeing probability and the types of open space represent the importance of including public and greenspaces in urban environments. However, while it was anticipated that more natural environments could be more beneficial, it was in fact found that amount of sports facilities were most influential; this implies interplay between physical and mental health, with people perhaps more likely to take advantage of their local facilities, and be discouraged if they would be required to travel further. Despite this, it should not be discounted that natural and park areas were also correlated with increased wellbeing, and these along with total open space should be considered as contributing factors for improving the health of an area. In addition, a range of open space descriptors were able to be implemented as a tool for classifying wellbeing. While this was shown to be much more than coincidental, a lack of strength in the reliability tests highlight the need for further factors to be considered, and perhaps opens whole area of research for investigation.

A. Limitations

Because of the different data sources and associated compatibility issues with this work, extensive data cleaning was required, which reduced the number of data points from 12,000 to 600, converting the characteristics of each individual open space in Greater London into a collection of attributes for each ward. However, had a more localised wellbeing dataset been available, this could have improved the accuracy of these investigations. This would require access to the original wellbeing data set, as opposed to the generalised set which was openly available, although unfortunately personal data such as this is not accessible from the London Data Store. In addition, any data pertaining to mental health is closed-source, due to privacy issues, and as such could not be applied to describe wellbeing in more medical terms. Self-reported measures are also subjective, and as such wellbeing was combined with a number of environmental classifiers for a more holistic view, although ideally more thorough, personal and emotional descriptors could have been included; combining these with more focussed geographical locators would have been ideal. It should also be noted that, while the data retrieved from the London Data Store was collected for 2013, the open space data was more up to date, describing all open spaces in 2014. While this is not hugely problematic, it

is something to which the data and subsequent analyses may still have been sensitive to, so could perhaps explain more of the variation observed; in short, more accurate and representative results may be obtained if all data relates to exactly the same time period. Overall, utilising more thorough measures and increasing the number of data points may have increased the effectiveness of this work.

B. Recommendations

As the correlations observed were weak, this suggests that where higher levels of wellbeing are observed, there is more chance that an increased quantity of open space may be present, rather than there being a direct association between the two. As such, further research needs to be done to consider a greater number of environmental characteristics and investigate subsequent relationships. This could perhaps include more localised data, as mentioned above, and more thorough wellbeing and environmental descriptors. It is recommended that accessibility, facilities and safety issues be considered for each space, in addition to greenspace attributes, perhaps examination of the types of foliage and what sort of opportunities each space offers; for example, this could include whether the area is sheltered, seated, and more suited to socialising, walking or sports. It is estimated that, with a larger, more localised and descriptive data set may yield greater conclusion and find further application.

C. Applications

Although this is an initial study and more conclusive results would hold more weight, in line with the discussions above, it is evident that built environments, and in particular open spaces, are important for the wellbeing of residents. With ever greater numbers of people moving into cities, and London in particular, the effects of negative environments will only be exacerbated by increasing population density. As such, it is important that urban developers are able to make informed choices to create healthy, beneficial environments. In light of policy, the Government is also beginning to take notice of how wellbeing may affect the prosperity of the Nation, and so should encourage the creation of such environments, to perhaps improve the wellbeing of the population in general.

ACKNOWLEDGMENT

The author would like to thank Greenspace Information for Greater London for permission to work with Open Space Data.

REFERENCES

- [1] Agnes E. Van Den Berg, Terry Hartig, and Henk Staats. Preference for nature in urbanized societies: Stress, restoration, and the pursuit of sustainability. *Journal of Social Issues*, 63(1):79–96, 2007.
- [2] I. Alcock, M. White, Wheeler, L. Fleming, and M. Depledge. Longitudinal effects on mental health of moving to greener and less green urban areas. 2013.
- [3] Mental Health Foundation. Measuring wellbeing: An introductory briefing. 2011.
- [4] M. Seligman. *Flourish*. Free Press, New York, 2012.
- [5] R. Dodge, A. Daly, J. Huyton, and L. Sanders. The challenge of defining wellbeing. 2012.
- [6] World Health Organisation. Global health observatory, urban population growth, Accessed: 5th December 2014.
- [7] R. Tennant, L. Hiller, R. Fishwick, S. Platt, S. Joseph, S. Weich, J. Parkinson, J. Secker, and S. Stewart-Brown. The warwick-edinburgh mental well-being scale (wemwbs): Development and uk validation. 2007.
- [8] London Data Store.