UNIVERSITY OF BIRMINGHAM
University of Birmingham Research at Birmingham

Using Twitter data in urban green space research

Roberts, Helen Victoria

DOI:
10.1016/j.apgeog.2017.02.008

License:
Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Roberts, HV 2017, 'Using Twitter data in urban green space research: A case study and critical evaluation', *Applied Geography*, vol. 81, pp. 13-20. https://doi.org/10.1016/j.apgeog.2017.02.008

Link to publication on Research at Birmingham portal

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.
•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
•User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 24. Apr. 2024

Applied Geography 81 (2017) 13–20

Contents lists available at ScienceDirect

Applied Geography

journal homepage: www.elsevier.com/locate/apgeog

Using Twitter data in urban green space research: A case study and critical evaluation

Helen Victoria Roberts

University of Birmingham, School of Geography Earth and Environmental Science, Birmingham B15 2TT, United Kingdom

ARTICLE INFO

Article history:
Received 30 June 2016
Received in revised form
7 February 2017
Accepted 8 February 2017
Available online 20 February 2017

1. Introduction

1.1. Green space

Urban green spaces are gaining an increasing amount of attention in both academic and policy arenas. The potential benefits they offer to human populations are increasingly significant in an urbanised society given their potential to facilitate improved human and societal well-being (Chiesura, 2004). The benefits that nature and green space can provide to human populations are broadly termed ecosystem services; encompassing the regulating, supporting, provisioning and cultural benefits that ecosystems provide (Costanza et al., 1997; Daily, 1997; Ehrlich & Ehrlich, 1981; MEA, 2005).

Urban green spaces are beneficial for human populations in a myriad of ways (Keniger, Gaston, Irvine, & Fuller, 2013) as is clear from the wide range of ecosystem services that have been identified (Costanza et al., 1997; Daily, 1997; Ehrlich & Ehrlich, 1981; MEA, 2005). The benefit of relevance to this case study is the interactions they facilitate between nature and human individuals in the surrounding communities. Provision and engagement with activity in such spaces is thought to facilitate social cohesion (Ewert & Heywood, 1991; Groenewegen, Berg, Vries, & Verheij, 2006), foster social empowerment (Westphal, 2003) and mitigate exclusion and isolation of certain groups (Seeland, Dübendorfer, & Hansmann, 2009; Shinew, Glover, & Parry, 2004). It has also been found that green spaces have a significant role in developing a sense of place, community identity and ownership of space (Kim & Kaplan, 2004). Urban green spaces are also important in facilitating interactions between people and nature. Such interactions have been found to be beneficial to human wellbeing; reducing stress and encouraging pro-social and sustainable behaviours (Maller, Townsend, Pryor, Brown, & St Leger, 2006).

While these ecosystem services have been defined, there is a significant lack of understanding as to how these benefits are transferred to and received by urban populations and the circumstances under which this can happen most effectively.

Insight into the interactions between urban populations and green spaces can provide direction for planners to instigate schemes to improve the quality of human life which is imperative given the increasing urbanisation occurring across the globe. Such endeavours to manage urban green spaces most effectively require research founded on an evidence based approach. Similarly, such information can influence local authority decision making and encourage the incorporation of functional and usable green spaces into the urban environment for the benefit of its urban citizens.

Understanding how green spaces are used is a fundamental starting point in improving insights about their significance for human and societal well-being. Therefore, knowing when and how people are engaging with urban green spaces is important, as well as an awareness of the potential barriers that may be preventing their use. Often a space can have multiple, simultaneous functions for different groups of users and understanding these can help in the development of sustainable land use management strategies (Peng, Chen, Liu, Lu, & Hu, 2016).

Previous attempts to examine how and when human populations make use of urban green spaces have followed two methodological approaches. The first of these relies on a qualitative, report based approach in which surveys, interviews and focus groups are used to gain information from participants on a range of topics. These have included the effects of deprivation on green space access (Jones, Hillsdon, & Coombes, 2009), features of green space which may promote increased use (Schipperijn, Sttigsdotter, Randrup, & Troelsen, 2010), personal motivations and barriers to using green spaces (Gidlow & Ellis, 2011), attempts to assess the non-market economic value of green space (Lockwood & Tracey, 1995; Del Saz Salazar & Garcia Menendez, 2007) and assessing the benefits people feel they receive from urban green space in times of heat stress (Lafortezza, Carrus, Sanesi, & Davies, 2009).

E-mail address: hxr133@bham.ac.uk.

http://dx.doi.org/10.1016/j.apgeog.2017.02.008
0143-6228/Crown Copyright © 2017 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Self-reporting techniques have also been employed to assess the impact of park improvements on use and activity (Cohen et al., 2009b). Whilst employed extensively in researching urban green space use, report based methods have a number of limitations. For example, as it is inherently reliant upon participant responses there may be issues with recall and social desirability biases (Evenson, Wen, Hillier, & Cohen, 2014) and it is difficult for the information received from participants to be independently validated. Indeed, where validations of questionnaire responses have been undertaken disagreements have been found. For example, Evenson, Wen, Golinelli, Rodrigues, and Cohen (2012) found only an acceptable agreement (67–82% percent agreement) between the actual and reported park visits of participants, using GPS monitors to validate responses.

The second way in which urban green spaces have been examined utilises an observational approach, treating the urban green space as a study location while researchers record the visitors and activities on-going within it. While used less extensively than report-based methods, such approaches have been employed in a range of contexts related to urban green spaces; including investigations into the types of activities that occur in green spaces (Tzoulas & James, 2010), features of a park associated with physical activity (Kaczynski, Potwarka, & Saelens, 2008) and the influence of meteorological variables on green space use (Thorsson, Lindqvist, & Lindqvist, 2004). Studies using this method have also investigated the effect of race, age and gender (Cohen et al., 2007; West, 1989) on the use of neighbourhood parks and their significance as a location for physical activity. Specific protocols such as the SOPARC (System for Observing Play and Recreation in Communities) have also been developed in an attempt to produce a standardised approach to observational methods of park use (McKenzie, Cohen, Sehgal, Williamson, & Golinelli, 2006). SOPARC has been used in investigations of the differences between rural and urban park visits (Shores and West, 2010) and how the installation of fitness zones affects physical activity engagement in parks (Cohen, Marsh, Williamson, Golinelli, & McKenzie, 2010). Observational methods have a significant limitation, requiring multiple observations over different days and seasons to ensure reliability (Cohen et al., 2009a) as park use patterns between specific observation times cannot be reliably estimated. Significant time is therefore required and as a result studies utilising observations tend to lack longitudinal depth.

Consequently, using current methods of observation and subjective reporting the measurement of human interactions with urban green spaces is challenging in terms of achieving consistent results. Methodological progression and new approaches are required to overcome the limitations currently faced (Orr, Paskins, & Chaytor, 2014).

1.2. Crowdsourcing and social networks data

Various emerging technologies now have the potential to advance assessment techniques of human interaction with urban green spaces: how, when and why people use them, what activities occur within them, and how people feel while using them. Social networks and social media systems enable anyone connected to the internet to provide information about their current location, feelings and activities. As such they provide a source of sensing and information that can be used to understand motivational factors behind the habits of populations (Silva, Vaz de Melo, Almeida, Salles, & Loureiro, 2013). This is an example of crowdsourcing which, in its simplest form, refers to a group of people producing data that can be used by third parties to solve a problem (Estellés-Arolas & González-Ladrón-de-Guevara, 2012). In the context of this paper, the crowd is comprised of users of smart technology devices (Kleeman, Voss, & Rieder, 2008) who share information on social media platforms via the internet. The crowdsourcing of information is becoming increasingly utilised as individuals become progressively connected and accessible in the information age (Brand, 2012).

It is clear that the recent proliferation of mobiles devices are key to crowdsourcing (Kanhere, 2011), especially when obtaining crowdsourced information from social media platforms – a mobile device enables anyone connected to the internet to share their information at any time. These social networks provide a platform to create a human powered participatory sensing network (Demirbas, Bayir, Akcora, Yilmaz, & Ferhatosmanoglu, 2010) in which the mobile devices carried by the users become the nodes of the network, connected to provide continuous information to a server. A number of social networks are increasingly present in the day to day lives of millions of people around the world and have already been employed in an academic context (Su, Wan, Hu, & Cai, 2016).

Created and launched in 2006, Twitter is a free microblogging service which enables users to communicate through short statuses and messages of up to 140 characters in length. Anyone connected to the internet via a smart device or computer, and with a Twitter account has high speed access with the ability to receive and share information. This connectivity along with the large number of users makes Twitter a highly influential player in the distribution of information and opinion (Mathioudakis & Koudas, 2010). Its popularity is credited to the ability of users to gain insight into other users without having a connection with them (Russell, 2013; Suh, Hong, Pirolli, & Chi, 2010). Following a person on Twitter affords a user instant access to another person's profile without the need for the other to give permission, follow them back, or even be aware of them (Weng, Lim, Jiang, & He, 2010). Indeed many pages have open access status and do not require any sign up to Twitter, unlike other networks such as Facebook.

Information obtained from Twitter has already been used successfully in urban research. Tweet information has aided land use classifications of urban areas (Noulas, Scellato, Mascolo, & Pontil, 2011; Frias-Martinez & Frias-Martinez, 2014; Zhan, Ukkusuri, & Zhul, 2014) and has been used to investigate the emotional responses of people to urban spaces (Bifet & Frank, 2010; Hauthal & Burghardt, 2013; Klettner, Huang, Schmidt, & Gartner, 2013). It has also been shown to be useful in following how information spreads through urban areas (Malleson & Andresen, 2015; Yardi & Boyd, 2010) and extension apps can be used to monitor a range of environmental variables (Demirbas et al., 2010).

This paper draws on the successes of such studies and sees them as justification for the inclusion of Twitter data in urban research. While Twitter data has been used to investigate cityscapes in general, it has not yet been applied to the study of urban green spaces despite their significance as components of the urban landscape. This paper provides a first introduction of the utility of Twitter data in the study of urban green space and the potential of crowdsourced information in improving understandings of such spaces and their importance for human populations. Twitter has been selected over other social networks due to the ease of accessing public data as well as the large numbers of users on the network generating this information.

2. Methodology, dataset and thematic analysis

The method described herein explores the potential of Twitter data as a source of information about human interactions with urban green spaces. To gain access to this publicly available data, it is necessary to connect to the Twitter API. R Studio was used as the interface through which connection to the Twitter API was made. The 'twitteR' package is designed specifically for working with the

Twitter API and the necessary coding functions are already in place. Crucially this method makes use of the OAuth protocol, a method which enables third party researchers to access user data without gaining access to their password and other private information (Hawker, 2010; Russell, 2013). Access through OAuth grants a third party user an access token and an access token secret which act as their credentials to access the user data. Using this package a range of metadata is returned alongside the tweet text, as shown in Fig. 1.

To obtain the tweets for study, a search was made of Twitters REST API using the park name as the search query (e.g. "Aston Park"). This was the only search term used in order to obtain the full range of tweets related to each park and prevent restriction or biasing of the tweet responses to certain types of activity. Tweets were then manually screened to ensure those included in the sample were relevant and reported an interaction with the specified urban green space.

A basic example is now presented demonstrating the utility of such Twitter data in the assessment of human use of urban green space. The study sites are located in Birmingham, the second largest city in the United Kingdom with an estimated population of 1.1 million (ONS, 2014). Within the metropolitan area, there are nearly 600 parks, public open spaces and nature reserves (BCC, 2016), the most of any European city. They provide an important resource for the surrounding populations in terms of their contribution to cultural ecosystem service provision.

Using Twitter data collected from tweets concerning urban green spaces in Birmingham, the range of information that can be obtained using crowdsourcing and social network data is illustrated. A case in point is the variety of organised activities found to occur in these green spaces and their importance for social interactions, economic opportunity and community identity is subsequently discussed. A thematic approach was taken in the subsequent analysis bringing together themes from a number of literature typically disengaged from the ecosystem services debate, including economics, social policy and cultural studies.

3. Results

Taking the summer months of 2015 (June, July August) as the study period, 24 out of 46 urban parks and green spaces were identified as hosting one or more organised event/s. The locations of these parks are given in Fig. 2. Parks were chosen to create a sample that reflects the variety of parks in Birmingham. Parks of varying characteristics were selected based on their size, the presence of woodland and water bodies, the presence of a number of different amenities, and their status as Green Flag parks, Nature Reserves and Active Parks locations.

From a total of 2847 tweets received over the study period 793 tweets relating to 61 separate events were identified, shown in Fig. 3. Tweets were categorised manually based on an assessment of their text and image content into 11 categories, one of which encompassed organised events. Other categories included physical activity, non-physical activity, nature related activity, charitable activity, economic activity, volunteering, political and religious focused tweets and information based tweets. The organised events category has been utilised herein to ensure a robust number of tweets for analysis.

The number of events recorded in each park ranged from 1 to 17 (Fig. 4). The highest number of tweets relating to a single event occurred at the Fusion Music Festival (Cofton Park) with a total of 335 tweets recorded. This is unsurprising given the size and popularity of the event with both locals and those from further afield who travel to the event, as was made clear in the tweets received.

The events identified provide opportunities for individuals to engage at a range of scales, from global events such as the Rugby World Cup (Eastside City Park), national events such as World Food Day (Cannon Hill Park) and the Summer Solstice (Lickey Hills), regional events such as the Big Hoot and Eid celebrations (Small Heath Park) and finally local events such as Edgbaston Regatta (Edgbaston Reservoir) and Acocks Green Carnival (Fox Hollies).

The overwhelming majority (80%) of the events identified were local events (Table 1), reflecting the important role of urban green spaces in providing a location where members of the local community can come together and socialise (Low, Taplin, & Scheld, 2009). Previous research has identified the role that urban green spaces play in developing an individual's sense of identity and feeling of connectedness with others (Kim & Kaplan, 2004). Such events may help to achieve this in the local communities in Birmingham. They may also be particularly important for older adults with limited mobility as being able to meet people in their local area is important for them to maintain social ties and a sense of connectedness to the community (Kweon, Sullivan, & Wiley, 1998).

From the tweets received it is also possible to identify the role of urban green spaces as places of engagement with a range of social, political and religious ideologies. For example, the Eid festival (Small Heath Park), Refugee week events, Vegan Picnic and Pankhurst Picnic (Cannon Hill Park) bring likeminded people together through a shared faith, or perspective. This again links to the importance of urban green spaces in the development of individual and community identity. In agreement with Mitchell (1995), urban green space is shown to be an important space into which different religious, social and political perspectives are brought and celebrated. Other cultural events were shown to occur, with music events such as the Birmingham Mela (Cannon Hill Park) providing the chance for communities to participate in cultural activities including music, dance and art.

The urban green spaces sampled also provided space for a range

	A
1	text
2	Loads of FREE #ParkLives fun at Aston Park (Fri 5pm) http://t.co/jWYCXmUbXy http://t.co/VdxZ1P7Iz4
3	Aston Hall and Park #birmingham #8000acres photo courtesy of http://t.co/vtUIYEdfMr http://t.co/bzMBMgkvx2
4	Another of Aston Park #birmingham #parklife #B6 Photo courtesy of http://t.co/DH6rne2MNJ http://t.co/LMf3RhaqFB
5	Great football, great drama , great sportmanship , great day at the Aston Park Rangers Football Tournament today http://t.co/4gfcUayGYx

B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
favorited	favoriteCount	replyToSN	created	truncated	replyToSID	id	replyToUID	statusSource	screenName	retweetCount	isRetweet	retweeted	longitude	latitude
FALSE	0	NA	19/05/2015 10:22	FALSE	NA	6.01E+17	NA	<a href="http:		0	FALSE	FALSE	NA	NA
FALSE	7	NA	19/05/2015 09:13	FALSE	NA	6.01E+17	NA	<a href="http:		10	FALSE	FALSE	-1.90145	52.452
FALSE	3	NA	18/05/2015 19:46	FALSE	NA	6E+17	NA	<a href="http:		0	FALSE	FALSE	NA	NA
FALSE	1	NA	18/05/2015 17:16	FALSE	NA	6E+17	NA	<a href="http:		2	FALSE	FALSE	NA	NA

Fig. 1. Example.csv file containing text and metadata information returned from the Twitter API.

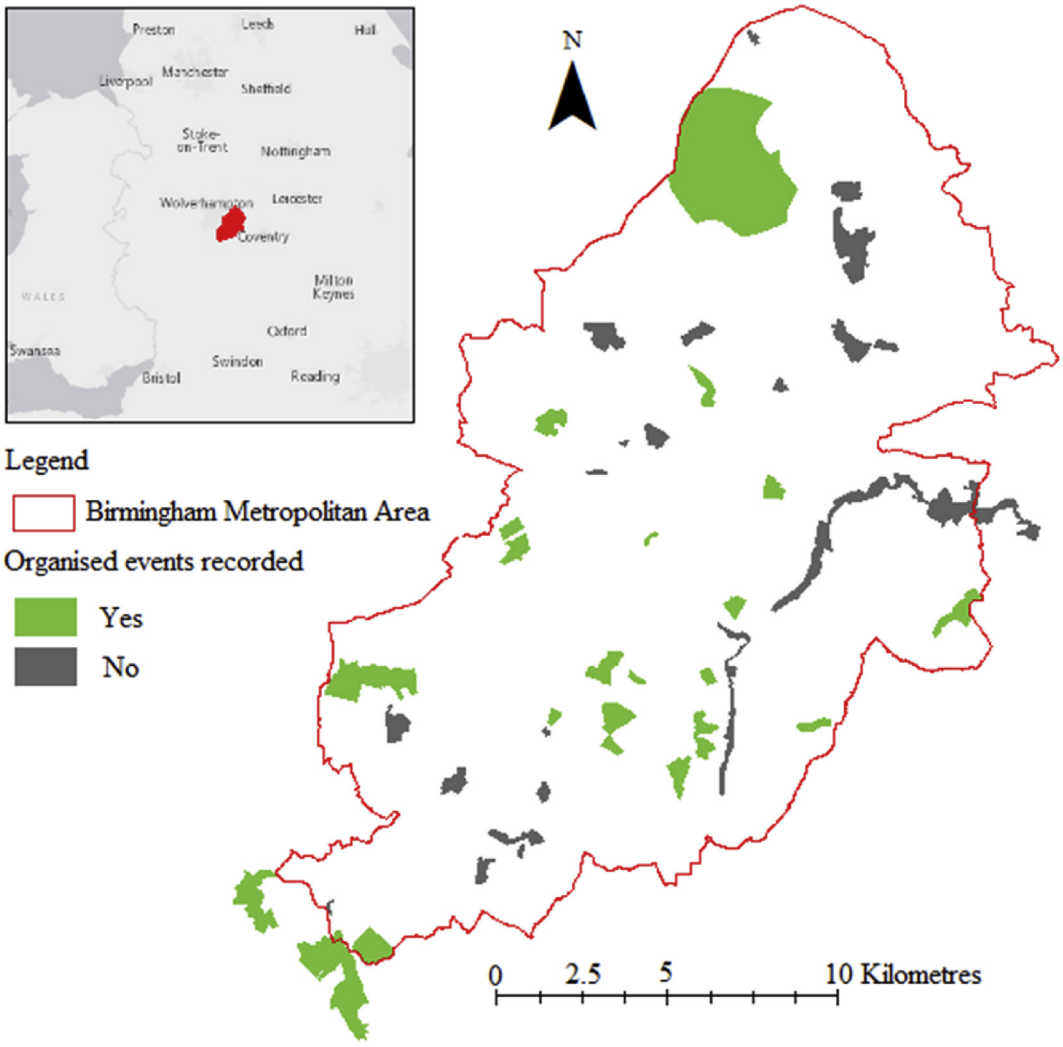


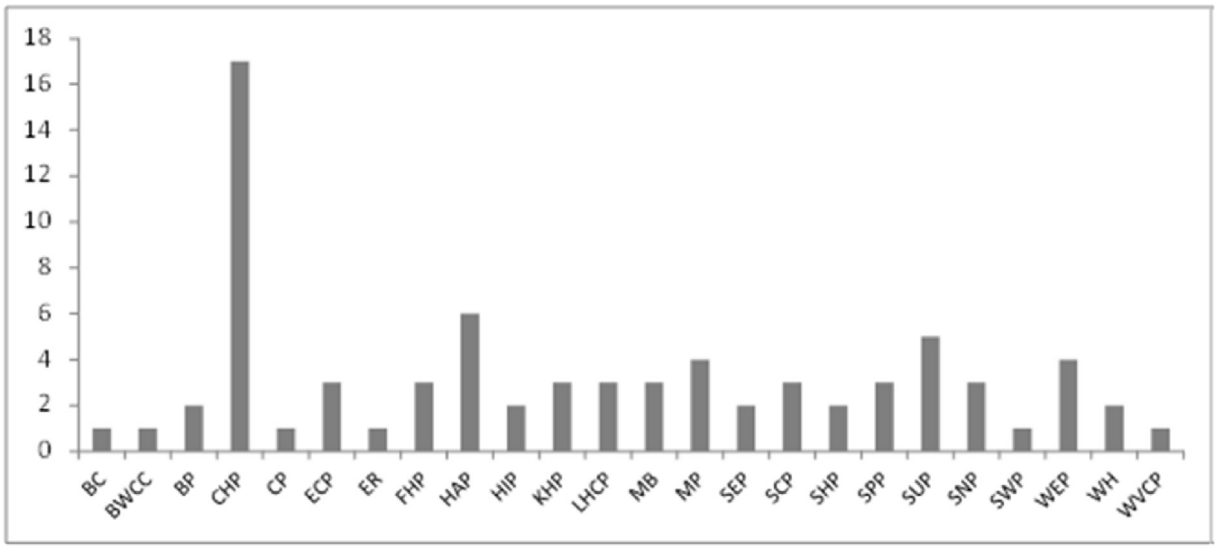
Fig. 2. The locations of parks in Birmingham hosting one or more organised event(s) during the summer months of 2015 where green indicates the occurrence of organised events and black indicates the remainder of the sampled parks. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 3. A word cloud showing the events captured using Twitter. Word size relates to the number of tweets received for each event, with larger words signifying more received tweets.

of activities aimed at facilitating social inclusion and empowerment for groups facing social isolation or other difficulties. Youth projects such as the Girls Youth Hub (Cannon Hill Park) and Sparkhill Youth Project (Sparkhill Park) have previously been identified as having an important role in social inclusion and integration of young people from a range of cultural backgrounds (Seeland et al., 2009). The opportunity that some events create for an individual to meet with others in a similar position to their own can also be beneficial. For example, Brummy Mummy Meetups enable new mothers to meet, socialise and discuss issues they may be facing.

The facilitation of interactions between humans and nature is an important role provided by urban green spaces (Maller et al., 2006). More natural environments have been found to have a restorative effect on cognition, providing an environment with less stressor and a variety of intriguing stimuli (Berman, Jonides, & Kaplan, 2008; Kaplan, 2001). Events which focus on bringing people into contact with nature such as Bioblitz, the Big Bog Lunch and flower planting help to facilitate these interactions and bring about improved mental well-being.

While the economic potential of urban green space has been considered extensively from the perspective of the whole city in terms of increasing land value, storm water management and non-market assets (Smardon, 1988; Del Saz Salazar & Garcia Menendez, 2007; Millward & Sabir, 2011), few studies have accounted for their role as discrete spaces for economic activity to take place. 40 of the



BC Billesley Common **BWCC** Birmingham Wildlife Conservation Centre **BP** Brookvale Park
CHP Cannon Hill Park **CP** Cofton Park **ECP** Eastside City Park **ER** Edgbaston Reservoir
FHP Fox Hollies Park **HAP** Handsworth Park **HIP** Highbury Park **KHP** Kings Heath Park
LHCP - Lickey Hills Country Park **MB** Moseley Bog **MP** Moseley Park **SEP** Selly Park
SCP Sheldon Country Park **SHP** Small Heath Park **SPP** Sparkhill Park **SUP** Summerfield Park
SNP Sutton Park **SWP** Swanhurst Park **WEP** Ward End Park **WH** West Heath Park
WVCP Woodgate Valley Country Park

Fig. 4. The total number of events recorded per park, obtained from captured tweet information.

Table 1
The scale of the events identified.

Scale of event	No. of events of each scale
Global	3
National	3
Regional/City wide	6
Local	49

events identified (66%) had an economic element, providing local businesses, charities or larger organisations with the chance to increase brand exposure and make financial gains through participating in them. Food festivals and summer fetes are a particular example where local businesses set up stalls and are provided with an opportunity for engagement with the local community. On a larger scale, music festivals such as Fusion Festival (Cofton Park) exemplify the opportunities provided to a range of sectors from entertainment, food, security and logistics. Charity events such as those taking place in association with Refugee Week (Cannon Hill Park) provide fundraising opportunities for charities as well as promoting pro-social behaviour and improving individual well-being (Thoits & Hewitt, 2001).

From the three month study period it was possible to elicit a large amount of information from Twitter as to the events occurring in the sampled urban green spaces. The implications these have for individual and societal well-being are discussed and bring together themes from a number of disciplines.

4. Critique

This discussion has explored the potential of using crowd-sourced data from Twitter in explorations of human interactions with urban green spaces. This method was proposed following the identification of a number of limitations with the previous observational and self-report based methodologies employed; and the successful use of Twitter data in a range of urban related research. It is important to evaluate this method and highlight its utility compared to previously utilised methods, as well as identify any limitations which have become apparent in this use of Twitter as a data source to inform human use of urban green spaces.

4.1. Advantages of data collection using Twitter

A large criticism of observational approaches is that to be reliable, they require extensive repeat measurements at the same location (Cohen et al., 2009a), incurring large time and cost expenditures. Twitter data do not have this limitation; indeed tweets can be captured with ease as frequently as necessary providing an opportunity for more measurements to be taken at no extra time or financial cost, achieving greater longitudinal depth as a result. Compared to self-report based methods, tweets are often posted with a photograph giving visual evidence for validation purposes.

The method also provides an unobtrusive method of non-participation which is easy to replicate, improving the potential for a standardised approach to be developed. Being free, publicly available and instantly accessible, the data collection method using Twitter incurs no financial cost and takes significantly less time compared to previously used methods. The consistency of Twitter data in terms of the 140 character limit on tweets means that analysis is more straightforward than other mixed media posts of varying lengths (Highfield & Leaver, 2014). These attributes mean Twitter data is well placed to provide information on the well-being, behaviours and activities occurring within communities (Nguyen et al., 2016).

4.2. Issues identified with using Twitter

Despite these benefits, the use of this method also raises some issues which should be taken into consideration in order for it to be utilised most effectively. The first of these is that crowdsourcing information via social networks limits the base population to which one is investigating and there is a need to discuss the inherent biases in these datasets (Hannay & Baatard, 2011). Crucially, those members of urban populations who do not own a smart device are excluded from the sample population. This can have implications for examining the use of space by various sectors of the population such as older people (aged 75+) who show disproportionate levels of non-engagement with these forms of technology (Zickuhr & Madden, 2012). Various spaces throughout the cityscape are supposed to be spaces where all members of the community can come together, but the users of social media do not reflect this diversity (Schwartz & Hochman, 2014) and are therefore not a truly representative cross section of the population. It should also be noted that very limited demographic information is known about the population from which tweets are received. Information about age, occupation or ethnicity is not available through the method described herein which may limit the type of investigation which can be carried out using this method. The metadata provided with the downloaded tweets can go some way to addressing this problem. For example it is possible to ascertain the gender of Twitter users through a search of their profile name on Twitter.

Numerous studies have been undertaken to try to determine the types of people who engage actively with social media (Bendler, Wagner, Brandt, & Neumann, 2014; Coleman, Georgiadou, & Labonte, 2009) as a means to assess source credibility. As a general rule, extroverts tend to be more frequent users of social media (Correa, Hinsley, & De Zuniga, 2010) with adults (aged 18–49) making up an increasingly large proportion of those actively engaging through posts (Lenhart, Purcell, Smith, & Zickuhr, 2010). Subsections of the population may be missing in the received dataset due to the inherent biases in using this type of technology; however because no demographic information is available it is difficult to know the direction in which the sample is non-representative.

Methods based on crowdsourcing make use of mobile devices through which people communicate and create a network. This creates issues associated not with the data collection method itself but those that need to be taken into account to understand what limitations there may be on the data available for capture by the method. The quality of internet connection can vary substantially between mobile networks and signal may be intermittent in some areas. An area with limited or no connection to the internet may lead to areas with no recorded use which may not necessarily be an accurate reflection of reality (Chatzimilioudis, Konstantinidis, Laoudias, & Zeinalipour-Yazti, 2012). While this can limit the production of data, in urban areas such as Birmingham, poor internet connection and mobile phone coverage are unlikely to be problematic as much of the city has 4G coverage. Appropriate selection of where this method is employed can overcome this obstacle to effective use.

User privacy and the ethics of obtaining data in the ways described herein is an important area of consideration when engaging with crowdsourcing through social media (Ma, Wei, Chai, & Xie, 2008; Burghardt, Buchmann, Müller, & Böhm, 2009; Vicente, Freni, Bettini, & Jensen, 2011). Being able to access the necessary information without compromising the privacy of the user is extremely important to users and researchers alike. This is not a significant issue using the method described in this paper as only public Twitter accounts are used to provide information, i.e. those who have enabled anyone to view their profile. The use of the OAuth process also addresses the need for privacy and data protection ensuring no personal account details are accessible.

4.3. Improving the robustness of a Twitter captured dataset

It should also be noted that the data gleaned from social networks are rarely produced with the aim of it being utilised in scientific research. There may be inaccuracies in their narrative that seem inconsequential to the user but may have significant implications for the research output if utilised by the researcher (Flanagin & Metzger, 2008). While this paper has shown the potential of Twitter in generating a dataset suitable for investigating human interaction with urban green spaces, there are a number of ways in which the robustness of such a dataset can be improved for research.

Improvements could be made to the resultant dataset by actively engaging and encouraging people to tweet about a specific subject or location of interest. A tried and tested way in which this is achieved is to create a hashtag unique to the study which individuals could be encouraged to add to their tweets. This hashtag could then be inputted into the search query to pull out relevant tweets. This is already being utilised by political campaigns and commercial companies to enable the tracking of Twitter responses to their product or ideas and the creation of a cyber-community who interact together through the use of specific hashtags. While not a traditional approach to data collection, this inductive approach could be employed in the research community with project specific hashtags or accounts affording new opportunities and the creation of a more robust dataset. Success has already been seen to this end with the use of the @ecorecordings account encouraging citizen science engagement with nature sightings.

With respect to the method described in this paper which connects to the Twitter API, it is important to note that an exhaustive source of tweets is not returned. Any search made to the API provides information of tweets produced in the last 10 days or so. One way to overcome this limitation is to make use of the Firehose API, a feed provided by Twitter that allows access to all public tweets. A significant problem to the use of the Firehose data however is the restrictive cost, as well as the amount of resources required to retain the Firehose data (servers, network availability, and disk space). To ensure maximum possible capture of tweets when using the Twitter API it is advisable to make regular searches to the API approximately every 10 days improving the completeness of the dataset received. Examination of the metadata downloaded with the tweet text, shown in Fig. 1 can help to provide information about the tweets received. Information such as the time and date of creation and name of the creator provides context to the dataset and improves robustness.

5. Conclusion

This paper has presented a method to investigate human use and interaction with urban green spaces following a critique of the current approaches employed to this end. Twitter is presented as a source of data which can be gathered through crowdsourcing. An example case study of 46 locations in Birmingham, UK has shown the potential of this new approach over a three month period. Twitter data was found to be successful in providing information about the range of organised events occurring in these urban green spaces, indicating the diversity in how urban populations make use of them. A high prevalence of local events was identified along with the provision of opportunity for engagement with regional and national events. The study sites were found to host a range of activities facilitating community engagement with social, cultural, political, religious and nature based events, while also providing

space for a range of economic activities.

In comparison to observational and subjective reporting methods which have been used previously in this area of research, the method presented herein offers a number of benefits. These include the free, publically available and immediately accessible nature of Twitter data, improved longitudinal depth that this method affords and potential to produce a standardised procedure to investigations. It also addresses the time and cost constraints identified with previous methods.

While this method has been demonstrated successfully herein, there is a need to identify a number of confines which much be addressed in order to utilise it most appropriately and effectively. These include privacy issues, biases in the received datasets and a lack of demographic information about the individuals included in the dataset.

References

Bendler, J., Wagner, S., Brandt, T., & Neumann, D. (2014). Taming uncertainty in big data evidence from social media in urban areas. *Business and Information Systems Engineering, 6*(5), 279–288.

Berman, M., Jonides, J., & Kaplan, S. (2008). The cognitive benefits of interacting with nature. *Psychological Science, 19*(12), 1207–1212.

Bifet, A., & Frank, E. (2010). Sentiment knowledge discovery in twitter streaming data. In *DS proceedings of the 13th international conference on discovery science* (pp. 1–15). Berlin, Heidelberg: Springer-Verlag.

Birmingham City Council (BCC). (2016). *Parks and nature conservation*. http://www.birmingham.gov.uk/parks Accessed 4 May 2016.

Brand, J. (2012). *The crowdsourcing handbook – everything you need to know about crowdsourcing*. Brisbane: Emereo Publishing.

Burghardt, T., Buchmann, E., Müller, J., & Böhm, K. (2009). Understanding user preferences and awareness: Privacy mechanisms in location-based services. In *On the move to meaningful internet systems: OTM* (Vol. 2009, pp. 304–321). Berlin Heidelberg: Springer.

Chatzimilioudis, G., Konstantinidis, A., Laoudias, C., & Zeinalipour-Yazti, D. (2012). Crowdsourcing with smartphones. *Internet Computing - IEEE, 16*(5), 36–44.

Chiesura, A. (2004). The role of urban parks for the sustainable city. *Landscape and Urban Planning, 68*, 129–138.

Cohen, D., Golinelli, D., Williamson, S., Sehgal, A., Marsh, T., & McKenzie, T. (2009b). Effects of park improvements on park use and physical activity: Policy and programming implications. *American Journal of Preventative Medicine, 37*(6), 475–480.

Cohen, D., Marsh, T., Williamson, S., Golinelli, D., & McKenzie, T. (2010). Impact and cost-effectiveness of family fitness zones: A natural experiment in urban public parks. *Health & Place, 18*(1), 39–45.

Cohen, D., McKenzie, T., Sehgal, A., Williamson, S., Golinelli, D., & Lurie, N. (2007). Contribution of public parks to physical activity. *American Journal of Public Health, 97*(3), 509–514.

Cohen, D., Sehgal, A., Williamson, S., Marsh, T., Golinelli, D., & McKenzie, T. (2009a). New recreational facilities for the young and old: Policy and programming implications. *Journal of Public Health Policy, 30*(1), S248–S263.

Coleman, D., Georgiadou, Y., & Labonte, J. (2009). Volunteered geographic information: The nature and motivation of producers. *International Journal of Spatial Data Infrastructures Research, 4*(1), 332–358.

Correa, T., Hinsley, A., & De Zuniga, H. (2010). Who interacts on the web?: the intersection of users' personality and social media use. *Computers in Human Behaviour, 26*(2), 247–253.

Costanza, R., d'Arge, R., Groot, R., Farber, S., Grasso, M., Hannon, B., et al. (1997). The value of the world's ecosystem services and natural capital. *Nature, 387*, 253–260.

Daily, G. (Ed.). (1997). *Nature's services: Societal dependence on natural ecosystems*. Washington DC: Island Press.

Del Saz Salazar, S., & Garcia Menendez, L. (2007). Estimating the non-market benefits of an urban park: Does proximity matter? *Land Use Policy, 24*, 296–305.

Demirbas, M., Bayir, M., Akcora, G., Yilmaz, Y., & Ferhatosmanoglu, H. (2010). Crowd-sourced sensing and collaboration using twitter. In *World of wireless mobile and multimedia networks (WoWMoM), 2010 IEEE International Symposium on a 1–9*. IEEE.

Ehrlich, P., & Ehrlich, A. (1981). *Extinction: The causes and consequences of the of the disappearance of species*. New York: Random House.

Estellés-Arolas, E., & González-Ladrón-de-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science, 38*(2), 189–200.

Evenson, K., Wen, F., Golinelli, D., Rodrigues, D., & Cohen, D. (2012). Measurement properties of a park use questionnaire. *Environment and Behaviour*, 1–22.

Evenson, K., Wen, F., Hillier, A., & Cohen, D. (2014). Assessing the contribution of parks to physical activity using GPS and accelerometry. *Medicine and Science in Sports & Exercise, 45*(10), 1981–1987.

Ewert, A., & Heywood, J. (1991). Ground development in the natural environment. Expectations, outcomes and techniques. *Environment and Behaviour, 23*(5), 592–615.

Flanagin, A., & Metzger, M. (2008). The credibility of volunteered geographic information. *GeoJournal, 72*(3–4), 137–148.

Frias-Martinez, V., & Frias-Martinez, E. (2014). Specteral clustering for sensing urban land use using Twitter activity. *Engineering Applications of Artificial Intelligence, 35*, 237–245.

Gidlow, J., & Ellis, N. (2011). Neighbourhood green space in deprived urban communities: Issues and barriers to use. *Local Environment, 16*(10), 989–1002.

Groenewegen, P., Berg, A., Vries, S., & Verheij, R. (2006). Vitamin G: Effects of green space on health, well-being, and social safety. *BMC Public Health, 6*(1), 1.

Hannay, P., & Baatard, G. (2011). GeoIntelligence: Data mining locational social media content for profiling and information gathering. In *Proceedings of the 2nd international cyber resilience conference Australia*.

Hauthal, E., & Burghardt, D. (2013). Extraction of location-based emotions from photo platforms. In J. Krisp (Ed.), *Progress in location-based services*. Springer.

Hawker, M. (2010). *Developer's guide to social programming: Building social context using Facebook, google friend connect, and the twitter API*. UK: Pearson Education.

Highfield, T., & Leaver, T. (2014). A methodology for mapping Instagram hashtags. *First Monday, 20*(1–5).

Jones, A., Hillsdon, M., & Coombes, E. (2009). Greenspace access, use, and physical activity: Understanding the effects of area deprivation. *Preventative Medicine, 49*(6), 500–505.

Kaczynski, A., Potwarka, L., & Saelens, B. (2008). Association of park size, distance, and features with physical activity in neighborhood parks. *American Journal of Public Health, 98*(8), 1451–1456.

Kanhere, S. (2011). Participatory sensing: Crowdsourcing data from mobile smartphones in urban spaces. In *Mobile data management (MDM), 2011 12th IEEE international conference on 2*. IEEE.

Kaplan, R. (2001). The nature of the view from home: Psychological benefits. *Environment & Behavior, 33*, 507–542.

Keniger, L., Gaston, K., Irvine, K., & Fuller, R. (2013). What are the benefits of interacting with nature? *International Journal of Environmental Research and Public Health, 10*(3), 913–935.

Kim, J., & Kaplan, R. (2004). Physical and psychological factors in sense of community new urbanist Kentlands and nearby Orchard Village. *Environment and Behavior, 36*(3), 313–340.

Kleeman, F., Voss, G., & Rieder, K. (2008). Un(der)paid innovators: The commercial utilization of consumer work through crowdsourcing. *Science, Technology and Innovation Studies, 4*(1), 5–26.

Klettner, S., Huang, H., Schmidt, M., & Gartner, G. (2013). Crowdsourcing affective responses to space. *Kartographische Nachrichten, 2*(3), 66–72.

Kweon, B., Sullivan, W., & Wiley, A. (1998). Green common spaces and the social integration of inner-city older adults. *Environment and Behavior, 30*(6), 832–858.

Lafortezza, R., Carrus, G., Sanesi, G., & Davies, C. (2009). Benefits and well-being perceived by people visiting green spaces in periods of heat stress. *Urban Forestry & Urban Greening, 8*(2), 97–108.

Lenhart, A., Purcell, K., Smith, A., & Zickuhr, K. (3rd Feb 2010). Social media & mobile internet use among teens and young adults Millennials. Pew Internet & American Life Project http://pewinternet.org/Reports/2010/Social-Media-and-Young-Adults.aspx Accessed 9 October 2015.

Lockwood, M., & Tracey, K. (1995). Nonmarket economic valuation of an urban green park. *Journal of Leisure Research, 27*(2).

Low, S., Taplin, D., & Scheld, S. (2009). *Rethinking urban parks: Public space and cultural diversity*. Texas: University of Texas Press.

Maller, C., Townsend, M., Pryor, A., Brown, P., & St Leger, L. (2006). Healthy nature healthy people:'contact with nature'as an upstream health promotion intervention for populations. *Health Promotion International, 21*(1), 45–54.

Malleson, N., & Andresen, M. (2015). The impact of using social media data in crime rate calculations: Shifting hot spots and changing spatial patterns. *Cartography and Geographic Information Science, 42*(2), 112–121.

Mathioudakis, M., & Koudas, N. (2010). Twittermonitor: Trend detection over the twitter stream. In *Proceedings of the 2010 ACM SIGMOD international conference on management of data* (pp. 1155–1158). ACM.

Ma, X., Wei, Z., Chai, Y., & Xie, K. (2008). Intergrating map services and location-based services for georeferenced individual data collection. In *Geoscience and remote sensing symposium, 2008 IGARSS 2008 IEEE international 5 V-81*. IEEE.

McKenzie, T., Cohen, D., Sehgal, A., Williamson, S., & Golinelli, D. (2006). System for observing play and recreation in communities (SOPARC): Reliability and feasibility measures. *Journal of Physical Activity and Health, 6*(1), S208–S222.

MEA. (2005). *Ecosystems and human well-being* (Vol. 5). Washington DC: Island Press.

Millward, A., & Sabir, S. (2011). Benefits of a forested urban park: What is the value of Allan Gardens to the city of Toronto, Canada? *Landscape and Urban Planning, 100*(3), 177–188.

Mitchell, D. (1995). The end of public space? People's Park, definitions of the public, and democracy. *Annals of the Association of American Geographers, 85*(1), 108–133.

Nguyen, Q., Kath, S., Meng, H., Li, D., Smith, K., VanDerslice, J., et al. (2016). Leveraging geotagged Twitter data to examine neighbourhood happiness, diet and physical activity. *Applied Geography, 73*, 77–88.

Noulas, A., Scellato, S., Mascolo, C., & Pontil, M. (2011). Exploiting semantic annotations for clustering geographic areas and users in location-based social networks. In *Proceedings of the international conference on weblogs and social media*.

Office for National Statistics (ONS). (2014). *Mid-2014 Mid-year population estimates*.

BDB2015/04 Birmingham demographic briefing. BCC.
Orr, S., Paskins, J., & Chaytor, S. (2014). Valuing urban green Space: Challenges and opportunities. In *UCL policy briefing*. London: UCL Public Policy.
Peng, J., Chen, X., Liu, Y., Lu, H., & Hu, X. (2016). Spatial identification if multi-functional landscapes and associated influencing factors in the Beijing-Tianjin-Hebei region, China. *Applied Geography, 74*, 170–181.
Russell, M. (2013). *Mining the social web: Data mining Facebook, twitter, LinkedIn, Google+, GitHub, and more*. California: O'Reilly Media Inc.
Schipperijn, J., Sttigsdotter, U., Randrup, T., & Troelsen, J. (2010). Influences on the use of urban green space – a case study in Odense. *Denmark Urban Forestry & Urban Greening, 9*(1), 25–32.
Schwartz, R., & Hochman, N. (2014). *The social media life of public spaces: Reading places through the lens of geo-tagged data*. Routledge, New York: Locative Media.
Seeland, K., Dübendorfer, S., & Hansmann, R. (2009). Making friends in Zurich's urban forests and parks: The role of public green space for social inclusion of youths from different cultures. *Forest Policy and Economics, 11*(1), 10–17.
Shinew, K., Glover, T., & Parry, D. (2004). Leisure spaces as potential sites for interracial interaction: Community gardens in urban areas. *Journal or Leisure Research, 36*, 36–355.
Shores, K., & West, S. (2010). Rural and urban park visits and park-based physical activity. *Preventive Medicine, 50*, S13–S17.
Silva, T., Vaz de Melo, P., Almeida, J., Salles, J., & Loureiro, A. (2013). A comparison of Foursquare and Instagram to the study of city dynamics and urban social behaviour. In *Proceedings of the 2nd ACM SIGKDD international workshop on urban computing 4*. ACM.
Smardon, R. (1988). Perception and aesthetics of the urban environment: Review of the role of vegetation. *Landscape and Urban Planning, 15*(1), 85–106.
Suh, B., Hong, L., Pirolli, P., & Chi, E. (2010). Want to be retweeted? Large scale analytics on factors impacting retweet in twitter network. In *Social computing, 2010 IEEE second international conference* (pp. 177–184). IEEE.
Su, S., Wan, C., Hu, Y., & Cai, Z. (2016). *Characterizing geographical preferences of international tourists and the local influential factors in China using geo-tagged photos on*.
Thoits, P., & Hewitt, L. (2001). Volunteer work and well-being. *Journal of Health and Social Behaviour*, 115–131.
Thorsson, S., Lindqvist, M., & Lindqvist, S. (2004). Thermal bioclimatic conditions and patterns of behaviour in an urban park in Göteborg. *Sweden International Journal of Biometeorology, 48*(3), 149–156.
Tzoulas, K., & James, P. (2010). Peoples' use of and concerns about green space networks: A case study of Birchwood, Warrington new town, UK. *Urban Forestry and Urban Greening, 9*(2), 121–128.
Vicente, C., Freni, D., Bettini, C., & Jensen, C. (2011). Location-related privacy in geo-social networks. *Internet Computing - IEEE, 15*(3), 20–27.
Weng, J., Lim, E., Jiang, J., & He, Q. (2010). Twitterrank: Finding topic-sensitive influential twitterers. In *Proceedings of the third ACM international conference on Web search and data mining* (pp. 261–270). ACM.
West, P. (1989). Urban region parks and black minorities: Subculture, marginality, and interracial relations in park use in the Detroit metropolitan area. *Leisure Sciences, 11*(1), 11–28.
Westphal, L. (2003). Urban greening and social benefits: A study of empowerment outcomes. *Journal of Arboroculture, 29*, 137–147.
Yardi, S., & Boyd, D. (2010). Tweeting from the town square: Measuring geographic local networks. In *Proceedings of the international conference on weblogs and social media*. AAAI.
Zhan, X., Ukkusuri, S., & Zhul, F. (2014). Inferring urban land use using large-scale social media check-in data. *Networks and Spatial Economics, 14*(3–4), 647–667.
Zickuhr, K., & Madden, M. (2012). Older adults and internet use. *Pew Internet & American Life Project, 6*.