Spatial assessments of visitation and discourse about national parks using social media

Author

Norman, Patrick

Published

2020-04-03

Thesis Type

Thesis (PhD Doctorate)

School

School of Environment and Sc

DOI

10.25904/1912/1267

Spatial assessments of visitation and discourse about national parks using social media

Patrick Norman

B.Sc (Hons)

Griffith School of Environment and Science
Environmental Futures Research Institute
Griffith Sciences

Submitted in the fulfilment of the requirements of the degree of Doctor of Philosophy

January 2020

Statement of Originality

This work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

(Signed)_____

Patrick Norman

Abstract

Protected areas such as national parks provide many cultural ecosystem services, including opportunities for tourism and recreation. Understanding patterns of visitation to national parks is important socially, environmentally and economically, as is public discourse about the parks. With the rapid increase in the use of social media there is an enormous volume of publicly available spatial data about national parks posted online, but what are the benefits and limitations of this still novel source of information for research and park management? Publicly available spatial data includes global positioning system (GPS) route data posted on social media platforms by mountain bikers, walkers and runners, as well as more general geographic information embedded in texts on microblogging platforms such as Twitter. This thesis examines how both of these types of spatial social media data can be used for monitoring visitation and discourse about national parks at a range of spatial scales from single parks to global assessments.

Firstly, the thesis compared how walkers, runners and mountain bikers use a series of connected urban national parks and reserves close to the city of Brisbane, Australia, using GPS route data from the social media platform MapMyFitness (**Chapter 2**). Route data was correlated with trail counter data, indicating that it could be used as a surrogate for monitoring within and among parks, and may, in some cases, give more reliable data particularly for mountain biking. Clear differences were found in how the parks and reserves were used for the three activities with mountain bikers travelling further and using more of the parks than walkers and runners. Also, mountain bikers and walkers preferred to visit on weekends, while runners use the

parks more consistently throughout the week. The results highlighted how route data complements and expands on other visitor data including providing important spatial and temporal data to assist in managing issues such as conflict and unauthorised use.

Next GPS route data for walking and running was compared among different types of parks (urban to remote) and among social media platforms in south-east Queensland, Australia (**Chapter 3**). The amount of route data differed among the platforms MapMyFitness, GPSies and Wikiloc. Data from the fitness focused platform MapMyFitness was the most useful for visitor monitoring in the urban park, due to large number of routes posted. However, data from the platform only showed visitation on the formal trail network. In contrast, those posting routes for the remote park, particularly on Wikiloc, often went off the formal trail network. Therefore, the best social media platform to use for monitoring depends on the questions being asked and the type of park, with the popularity of platforms also varying among countries and over time.

To see if social media platforms could be used to assess the relative popularity of parks for walking, running and mountain biking, GPS route data was collected from Strava, MapMyFitness and Wikiloc for 40 national parks in south-eastern Queensland, Australia (**Chapter 4**). Although Strava was very popular with over 430,000 routes, the other platforms provided useful information for a wider range of activities (MapMyFitness) or a different group of walkers (Wikiloc). Distance to urban areas, and to a lesser extent, the types of trails and permitted activities, best explained the popularity of parks based on data from the two fitness-oriented

platforms, Strava and MapMyFitness. For Wikiloc, however, the elevation range of the park was the best predictor of popularity with more adventurous walkers wanting to access rugged and remote areas. Understanding what makes a park popular for each activity is important when managing social and environmental issues particularly in popular urban parks.

For the last results chapter, the focus changed from GPS route data within a single park or region, to using text from the microblogging platform Twitter at a global scale (**Chapter 5**). Twitter data collected over a 6-month period was used to assess the scale of the discourse about national parks globally and to assess spatial information in the texts, including identifying which parks are talked about and by whom. The discourse was massive, in terms of number of tweets (>2 million), number of tweeters (~750,000) and covered a wide range of parks, with 264 national parks talked about at least 100 times. The size of the discourse about parks varied among countries, but predominantly it was about North American parks, and often sent by accounts in the USA. The number of tweets was correlated with the relative popularity of parks with visitors for the 40 most talked about parks on Twitter, particularly for the most popular parks in the USA. Twitter is therefore a useful source of data for park management, particularly in countries where the platform is popular and/or there are globally renowned national parks.

Overall, social media data including GPS routes, but also Twitter text, are an increasingly useful source of spatial information about parks including when assessing the movements of visitors in parks, factors influencing park popularity and

what makes them noteworthy at a range of spatial scales from local urban parks to global networks. The individual results chapters, and overall results illustrated important benefits but also limitations with social media data including the range and scale of different data types as well as the ephemeral nature of social media platforms and the availability of data.

Table of Contents

List of Figures

Acknowledgements

Firstly, I would like to express my sincere gratitude to my primary supervisor Professor Catherine Pickering for the continuous support throughout my PhD study and related research, in particular her patience, motivation and knowledge.

I would also like to thank Dr Guy Castley. His wisdom constantly helped me throughout my PhD and he was always great person to approach when in need of some inspiration. My sincere thanks also go to Dr Alienor Chauvenet, Dr Eloise Skinner, Dr Jan Warnken, Dr Ryan Pearson, Luke Carpenter and Dr Clare Morrison for all helping me throughout this time. Without their precious support and knowledge, it would not be possible to conduct this research. I would also like to thank my family for all of their support in particular my mother Katherine Norman who kindly helped to proof read my final thesis.

Lastly, I would like to thank my wife Jenna Wraith. The PhD journey has been a wonderful time and that is largely due to sharing the experience with her.

Glossary

Geocode: The process of converting words, text, and numbers into a geographic location, such as an address into latitude and longitude (Cayo and Talbot, 2003).

Geographic information system (GIS): A system to store, analyse, manage and present geographic data (Antenucci et al., 1991).

Geolocation: The identification or estimation of a location, often to geographic coordinates such as a longitude and latitude (Djuknic and Richton, 2001).

Geotag: A location identifier included in the metadata of other types of media, such as an image or video (Torniai et al., 2009).

Global Positioning System (GPS): A navigation system where a location is calculated by orbiting satellites (Parkinson et al., 1996).

Metadata: The data which provides information about another piece of data, such as the time stamp when a social media post was produced (Skågeby, 2009).

Peri-urban: An area on the fringe of a city or town (Simon, 2008).

Route data: Spatial data about an entire journey, often being continuously recorded in small time intervals (Chen et al., 2011).

Point data: Spatial data about a single location, also known as point location data (Guo et al., 2012).

Public Participation Geographic Information System (PPGIS): The public involvement in policy making and planning, through sharing geographical information and knowledge (Sieber, 2008).

Social media platform: Interactive web-based application where users share information and connect with others (Obar and Wildman, 2015).

User generated content: A form of content created and shared by users, posted onto web-based platforms (Van Dijck, 2009).

Volunteered Geographic Information (VGI): User generated spatial information posted online (Coleman et al., 2009).

References

Antenucci, J.C., Brown, K., Croswell, P.L., Kevany, M.J. and Archer, H., 1991. *Geographic Information Systems: A guide to the technology*. Springer: New York City.

Cayo, M.R. and Talbot, T.O., 2003. Positional error in automated geocoding of residential addresses. *International Journal of Health Geographics*, *2*(1), pp.10.

Chen, L., Lv, M., Ye, Q., Chen, G. and Woodward, J., 2011. A personal route prediction system based on trajectory data mining. *Information Sciences*, *181*(7), pp.1264-1284.

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

Djuknic, G.M. and Richton, R.E., 2001. Geolocation and assisted GPS. *Computer*, *34*(2), pp.123-125.

Guo, D., Zhu, X., Jin, H., Gao, P. and Andris, C., 2012. Discovering spatial patterns in origin-destination mobility data. *Transactions in GIS*, *16*(3), pp.411-429.

Obar, J.A. and Wildman, S.S., 2015. Social media definition and the governance challenge-an introduction to the special issue. *Telecommunications Policy*, *39*(9), pp.745-750.

Parkinson, B.W., Enge, P., Axelrad, P. and Spilker Jr, J.J. eds., 1996. *Global positioning system: Theory and applications, volume II*. American Institute of Aeronautics and Astronautics. AIAA: Washington.

Sieber, R., 2006. Public participation geographic information systems: A literature review and framework. *Annals of the association of American Geographers*, *96*(3), pp.491-507.

Simon, D., 2008. Urban environments: issues on the peri-urban fringe. *Annual Review of Environment and Resources*, *33*, pp 167-185

Skågeby, J., 2009. Exploring qualitative sharing practices of social metadata: Expanding the attention economy. *The Information Society*, *25*(1), pp.60-72.

Torniai C., Battle S., Cayzer S., 2009. Sharing, Discovering and Browsing Geotagged Pictures on the World Wide Web. In: Scharl A., Tochtermann K. (eds) *The geospatial web. Advanced information and knowledge processing.* Springer: London.

Van Dijck, J., 2009. Users like you? Theorizing agency in user-generated content. *Media, Culture & Society*, *31*(1), pp.41-58.

Published papers included in the thesis

Parts of this thesis have already been published in refereed academic journals including Chapters 2, 3 and 4, which were co-authored with my supervisors. My contribution to these papers is outlined at the start of each of these chapters. The detailed bibliographic information for the papers is as follows:

Chapter 2: Norman, P., Pickering, C.M. and Castley, G., 2019. What can volunteered geographic information tell us about the different ways mountain bikers, runners and walkers use urban reserves? *Landscape and Urban Planning, 185*, pp.180-190.

Chapter 3: Norman, P. and Pickering, C.M., 2017. Using volunteered geographic information to assess park visitation: Comparing three on-line platforms. *Applied Geography, 89*, pp.163-172.

Chapter 4: Norman, P. and Pickering, C.M., 2019. Factors influencing park popularity for mountain bikers, walkers and runners as indicated by social media route data. *Journal of Environmental Management, 249*, pp.109413.

Appropriate acknowledgements of those contributing to the research but did not qualify as authors are also included in each paper.

(Signed):

(Date):

Patrick Norman

(Counter signed):

(Date):

Professor Catherine Marina Pickering

Chapter 1. Introduction

1.1. Protected areas

Protected areas are the major mechanism for the conservation of biodiversity globally, helping protect many species, ecosystems and landscapes as well as providing a wide range of other benefits to humans (Boyd et al., 2008; Gaston et al., 2008; Gaveau et al., 2012; Gray et al., 2016; Ma et al., 2013; Venter et al., 2014; Watson et al., 2011). Although there is a long history in different cultures of setting aside areas for conservation, the modern system of protected areas is considered to have started in the 1870s in the United States with the establishment of Yellowstone National Park (Tyrell, 2012). Over the next 150 years this model has been used in most countries, with the first national parks established in Australia in 1879, Canada in 1885 and New Zealand in 1887 (Boyd and Butler, 2009; Ramp et al., 2006; Tyrell, 2012). Now there are protected areas in nearly every country (United Nations Environment Programme, 2018), and cover almost 15% of the earths land surface (United Nations Environment Programme, 2018; Worboys et al., 2015). Once established, protected areas provide a safeguard helping to prevent areas from being developed or degraded by a range of human activities (Newsome et al.,2012; Struebig et al., 2015; Worboys et al., 2015).

The level of protection, types of human use, management strategies, terminology and status of protected areas can vary within and among countries (IUCN, 2019). Increasingly a global standard system is used to classify protected areas, formulated by the International Union for the Conservation of Nature which classifies global protected areas into seven categories (IUCN, 2019). These range from strict nature

reserves (Category Ia) where human impacts are kept to an absolute minimum, to areas with sustainable use of natural resources (Category VI) allowing for non-commercial natural harvesting (IUCN, 2019). One of the most common and well recognised types of protected areas under the IUCN system are category II: National Parks (Hall and Frost, 2009; Worboys et al., 2015; IUCN, 2019), which are

> *"Large natural or near natural areas set aside to protect large-scale ecological processes, along with the complement of species and ecosystems characteristic of the area, which also provide a foundation for environmentally and culturally compatible spiritual, scientific, educational, recreational and visitor opportunities"* (IUCN, 2019).

These protected areas are primarily set aside to conserve biodiversity and to promote education and recreation. There are globally renowned national parks in Europe, Africa, North and South America, Asia and Oceania, with the term national park used to refer to Category II protected areas in more than 86 countries (United Nations Environment World Conservation Monitoring Centre, 2019). With large areas protected in national parks globally, they provide a wide range of benefits to both people and the environment.

1.2. Importance of national parks

Ecosystems services, including those provided by national parks are critical in maintaining human populations and have an estimated global worth of over US$125 trillion annually (Costanza et al., 2014). National parks and other types of protected areas provide ecosystem services in many regions including clean air and water, agricultural soils, carbon sequestration, food and medicinal resources for millions of people as well as a range of cultural ecosystem services (FAO, 2019). The ability of

natural ecosystems to provide services is degraded when they are damaged or modified, thus conserving natural ecosystems in national parks ensures the continuity of their manifold benefits (Chazdon, 2008; Lecina-Diaz et al., 2019; Mukul et al., 2017; Palomo et al, 2013; Siikamäki et al., 2015). Among the many benefits provided by many national parks are cultural ecosystem services (Paracchini et al., 2014).

National parks provide a range of cultural ecosystem services, including spiritual services, education, cultural heritage and values, aesthetics, a sense of place as well as tourism and recreation opportunities (Plieninger et al., 2013). The resulting health benefits of parks is immense, estimated at US$6 trillion per year (Buckley et al., 2019). Ensuring areas important for providing cultural ecosystem services are included in national parks, can result in these areas being popular for tourism and recreation (Paracchini et al., 2014; Plieninger et al., 2013; Rossi et al., 2019).

1.3. Tourism and recreation in national parks

Globally, parks and other protected areas are important for tourism and recreation with over 8.3 billion visits annually, generating US$600 in economic benefits (Balmford et al., 2015). In many countries, national parks are the cornerstone of tourism industries and are often popular attractions, making them a critical part of the economy (Driml and Common, 1995; Hunt et al., 2015). Visitation to national parks is expected to grow rapidly over the coming decades, with much of this growth concentrated in particular parks (Karanth and DeFries, 2011; Leung et al., 2018; Monz et al., 2010) including those close to cities (Neuvonen et al., 2010). With visitation to national parks important and increasing, understanding the factors that influence visitation is vital for park management (Fletcher, 2011; Leung et al., 2018).

A wide range of tourism and recreation activities occur in national parks including sightseeing, camping, picnicking, cultural and educational activities, viewing wildlife, fishing as well as a number of trail-based activities such as walking, running and mountain biking (Figure 1.1.) (Eagles et al., 2002; Manning and Anderson, 2017; Newsome et al., 2012). The relative popularity of specific activities and total visitation varies amongst parks, countries and regions. Factors influencing visitation include proximity to cities (Kovacs-Györi et al., 2018; Neuvonen et al., 2010; Rossi et al., 2019; Schipperijn et al., 2010), climate (Hadwen et al., 2011), conservation status such as world heritage (Caust and Vecco, 2017), types of tourism and recreation infrastructure (Neuvonen et al., 2010), marketing and the regulation of activities by park agencies (Newsome et al., 2012). Park managers need to ensure there is adequate planning and research so parks are inclusive for a range of appropriate tourism and recreation activities whilst also minimising negative social and environmental impacts (Newsome et al., 2012). This is particularly important with rapid increases in visitation to many parks, overall and for specific activities (Newsome and Davies, 2009; Pickering et al., 2010a). Therefore, it is important to understand who, where and when people visit national parks and for what types of activities.

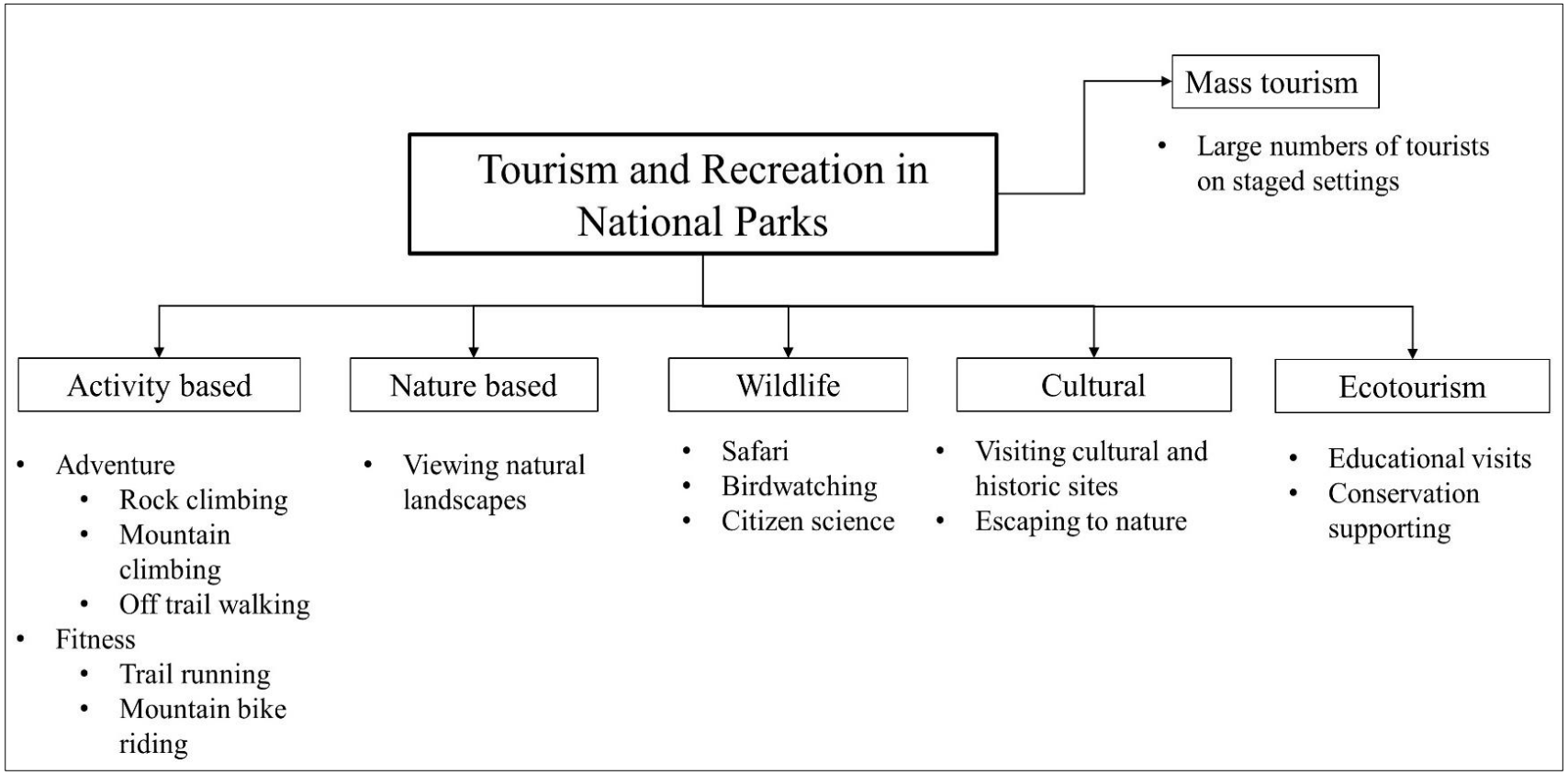

Figure 1.1. Some of the different types of tourism and recreation activities in national parks. Adapted from Newsome et al. (2012).

1.4. Monitoring visitation to parks

Understanding park visitation, including temporal and spatial patterns is important for managers, particularly when allocating resources such as staff and infrastructure (Worboys et al., 2015). Infrastructure for tourism and recreation range from general facilities such as visitor centres, entry gates, roads, carparks, mown areas with tables and toilets, campgrounds and walking and mountain biking trails among others (Worboys et al., 2015; Newsome and Davies, 2009). With increasing visitation there can be greater environmental impacts, requiring response such as limiting use and/or further hardening of facilities such as trails (Ballantyne and Pickering, 2015; Pickering and Norman, 2017). Visitation to national parks can also lead to a range of environmental impacts (Monz et al., 2013; Pickering and Hill, 2007; Wolf et al., 2019).

A wide range of environmental impacts occur from excessive visitation including damage to vegetation, reductions in animal abundance, pollution to waterways, soil erosion and increases in feral animals and plants (Pickering et al., 2010b; Pickering and Hill, 2007; Wolf et al., 2019). In areas of high visitation there can be significant damage with long lasting environmental impacts (Monz et al., 2013; Wolf et al., 2019). Where activities occur away from hardened infrastructure, such as riding and hiking off formal trails, there can be additional damage (Ballantyne and Pickering, 2015; Monz et al., 2013), including threatening rare plants and animals with restricted distributions (Hockett et al., 2010; Lacki, 2000; Wraith and Pickering, 2017). Understanding temporal and spatial patterns of visitation in and among parks can

identify areas were damage occurring already or likely to occur in the future (Lime and Stankey, 1971; Sayan and Atik, 2011; Symmonds et al., 2000).

Social issues can also arise in popular parks, as well as on and around popular park infrastructure such as trails (Arnberger et al., 2012; Sayan and Atik, 2011; Zeppel, 2010). This includes when the use of facilities exceeds carrying capacity or when facilities, such as trails, are used for different activities which can result in conflict among visitors (Cessford, 2002; Manning, 2002). For example, on mixed use trails conflict can occur between walkers, runners and mountain bikers due to differences in perceptions about appropriateness of specific activities in parks and the risk of injuries due to differences in speed (Mann and Absher, 2008; Pickering and Rossi, 2016; Spencer, 2012). To assess intensity and types of visitation, data on temporal and spatial patterns of visitation among and within parks is important, although this is often limited for many parks (Newsome et al., 2012).

1.5. Methods for spatially monitoring visitation to national parks

Traditionally there were a number of ways to assess spatial and temporal patterns of visitation in and among parks, including both on and off formal trail networks (Newsome et al, 2013; Rossi et al., 2015). These include trail counters, entrance and camping fees, video and aerial imagery as well as questionnaires and surveys about where people go and what they do in parks (Eagles, 2014; Newsome et al., 2012). This data is used to understand aspects of park visitation including; which parks people visit, use of formal tail networks vs off trail use, and to identify areas within parks that are particularly popular for different activities (Eagles, 2014; Newsome et al, 2013; Worboys et al., 2015). There are, however, important considerations and

constraints with visitor data, and as it is often time consuming and expensive to collect, potentially resulting in small sample sizes and across short time periods (Ankre et al., 2016; Newsome et al., 2012; Tarrant and Smith, 2002; Wolf et al., 2012).

Recently a range of types of volunteered geographic information has been used to assess park visitation (Levin et al., 2017; Wolf et al., 2015). This includes providing visitors with global positioning system (GPS) tracking devise to map movements within parks as well as harnessing online surveys including using apps and other systems where visitors can indicate where they went and what they prefer in parks (Brown and Fagerholm, 2015; Engen et al., 2018; Korpilo et al., 2017; Wolf et al., 2015). These methods, often referred to as public participation geographic information systems (PPGIS), can be very useful for assessing finer scale visitor movements in parks including mapping the use of trails for different activities (Brown, 2012; Wolf et al., 2012; Wolf et al., 2015). Another novel source of public spatial data that is starting to be used for monitoring visitors within national parks is social media.

1.6. Use of social media to assess park visitation

With the recent advent of smartphones, the use of social media is rapidly increasing (Primack et al., 2017). As many smartphones have accurate GPS locating capabilities, people often post location data to social media along with other information, such as where they have been, what they have been doing and what they have seen (Furini and Tamanini, 2015; Munar and Jacobsen; 2014; Schwartz and Halegoua, 2015). Such highly accurate user generated spatial content is emerging as a valuable tool for

assessing park visitation at a range of spatial scales (Ghermandi and Sinclair 2019; Hausmann et al., 2018). Social media spatial data has been used by researchers to assess different aspects of visitation to parks including; the relative popularity of different parks (Hausmann et al., 2018), patterns of visitation within parks (Santos et al, 2016), preferences for different types of infrastructure (Orsi and Geneletti, 2013), temporal differences in visitation (Walden-Schreiner et al., 2018), and differences in the patterns of use of trails by runners and mountain bikers (Campelo and Mendes, 2016). However, the use of social media data for assessing national park visitation is dependent on access to, the volume of and the types of information available.

1.7. Availability and types of spatial social media data

The amount, type, availability of spatial data changes over time and among social media platforms (Toivonen et al., 2019). A number of different social media platforms are emerging as useful sources of data for assessing visitation to parks, including Twitter, Flickr, Weibo and YouTube (Ghermandi and Sinclair, 2019; Toivonen et al., 2019). Other large and well-known platforms are increasingly restricting access to data for research, either by not providing access/banning the use of public data (Facebook, Instagram and TripAdvisor) or by charging for detailed data (Strava) (Facebook, 2019; Instagram, 2019; Strava, 2019; TripAdvisor, 2019; Toivonen et al., 2019). Often data restrictions have been due to privacy concerns with the misuse of social media data, which was the case for Facebook (Schneble et al., 2018). Platforms designed to share information publicly, rather than within friend groups, are more likely to allow access to data for research, such as Flickr and Twitter (Flickr, 2019; Twitter, 2019). They along with some other platforms provide an

automated programming interface (API), which can be used to access and download data for analysis.

In addition to differences in access, the types of data available about parks varies among social media platforms (Toivonen et al., 2019). Available data can include basic metadata, such as user identification and the time and date of the post, multimedia data such as text, images and videos, through to spatial information such as geolocation data or GPS routes (Campelo and Mendes, 2016; Ghermandi and Sinclair, 2019; Toivonen et al., 2019). With this data being generated without the input of researchers, social media is now emerging as a valuable source of visitor information for assessing visitors to national parks.

Some social media platforms are focused on sharing spatial information about where people go for specific activities such as mountain biking, running and walking among others. This spatial data posted online is an increasingly common example of volunteered geographic information (VGI) (Flanagin and Metzger, 2008). Examples include adding check-ins to a location (Foursquare), finding hidden 'treasures' (Geocaching) and geolocated images (Flickr) (Figure 1.2.) (Flickr, 2019; Foursquare, 2019; Geocaching, 2019). In recent years, the popularity of platforms sharing spatial data has increased to such an extent, that they can provide large amounts of data about visitors in parks (Campelo and Mendes, 2016, Monteiro, 2016; Walden-Schreiner et al., 2018).

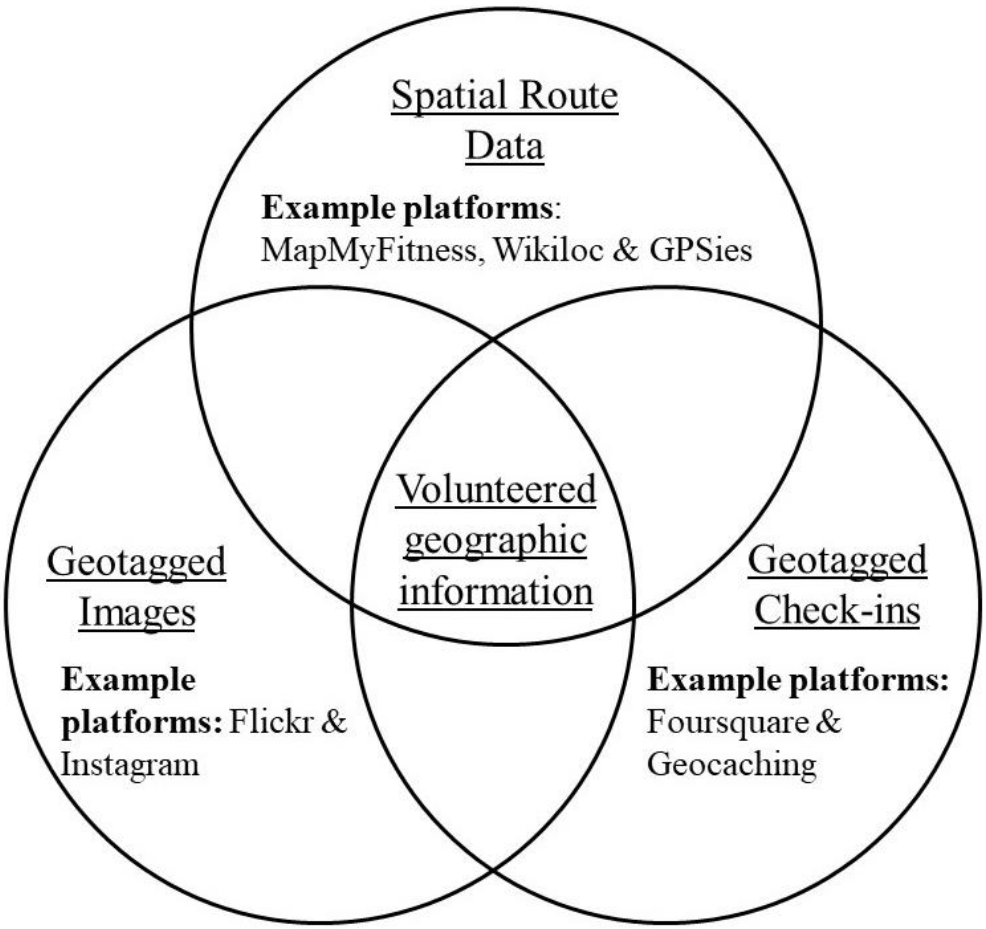

Figure 1.2. Some common types of volunteered geographic information, including examples of the types of social media platforms.

One potentially rich source of spatial data for assessing park visitation is GPS route data (Campelo and Mendes, 2016). Route data is often available from social media platforms targeting people who want to share information about their activities online including fitness platforms such as Strava, GPSies, Endomondo, Nike Run Club and MapMyFitness (Endomondo, 2019; MapMyFitness, 2019; Nike Run Club, 2019; Strava, 2019). Researchers have started using this source of spatial data to assess patterns of visitation including the use of trails (Campelo and Mendes, 2016; Jurado Rota et al., 2019; Santos et al, 2016). There are however still a range of questions about this data including, are there differences in the ways different visitors including

those engaging in different activities move throughout parks and can it be used to detect the popularity of different parks for different activities. As GPS route sharing platforms are less popular than many larger social media platforms, they are currently suited for assessing visitor movement at small spatial scales and for specific activities (Jurado Rota et al., 2019; Santos et al, 2016).

1.8. Social media discourse about national parks

Increasingly people are using social media to share their opinions about a wide range of issues including about national parks. Due to the scale of data available from Twitter, it is emerging as one of the most useful platforms for social media research about parks (Bassolas et al., 2016; Ghermandi and Sinclair, 2019). Due to privacy concerns, changes were made to the platform from 2014, limiting the automatic collection and display of spatial data about where posts were sent from (Leetaru, 2019). However, the texts of tweets can be used to assess some geographical patterns in park visitation and can provide insights into public discourse about national parks including public responses to controversial events (MacDonald et al., 2016), sentiments (Becken et al., 2017) and issues around natural heritage conservation (Halpenny and Blye, 2017). Information from the text includes which parks are discussed and by whom (Becken et al., 2017; Halpenny and Blye, 2017).

Assessing text data from Twitter presents a number of challenges that must be accounted for when using this platform for research. This includes language and character constraints as well as the peculiar writing style arising from the short length of tweets (Inkpen et al., 2017; Kern et al., 2016; Singh and Kumari, 2016). Even with these constraints, the size of the discourse and availability of Twitter data means that

text from the platform will continue to be a valuable source of social media data about parks. Important questions that need to be addressed regarding Twitter include, what is the scale of discourse on Twitter about parks, does it reflect the popularity of parks with visitors and how does the discourse vary among regions, countries and parks.

1.9. Comparisons with on ground data

A critical aspect of social media data for assessing park visitation, including for both text and GPS routes, is establishing how accurately it reflects visitation against other types of monitoring. As a result, some studies have compared visitation patterns from social media point location data with 'on-ground' visitation data (Balmford et al., 2015; Tenkanen et al., 2017; Wood et al., 2013). They found that point location data from Flickr, Twitter and Instagram, can provide similar information on relative park popularity as on-ground visitation data (Balmford et al., 2015; Tenkanen et al., 2017). However, geolocation data about where tweets were created, which was used for assessing a range of issues relating to parks (Ghermandi and Sinclair, 2019; Tenkanen et al., 2017) is no longer available from Twitter (Leetaru, 2019).

1.10. Aims

With vast amounts of publicly available data on social media including about national parks, it is important to assess how this data can be used in research and management. This includes evaluating how spatial data, such as GPS route data post to social media could be used to assess visitation within and among parks. Also, there is interest in seeing if the text in social media posts such as those on Twitter can be used for spatial analyses, now that detailed geolocation data about where tweets were sent from is no longer available. Therefore, the aim of this thesis was to assess how some types of

social media data can be used for spatial monitoring of visitation and discourse about national parks, including at different spatial scales from local urban parks to a global scale. It specifically addressed the following questions:

1. What are the benefits and limitations of different social media platforms as sources of GPS route data for visitor monitoring in and among parks (Chapter 3 and 4)?
2. How can GPS route data from social media platforms be used to assess spatial and temporal patterns of visitation within parks for walking, running and mountain biking (Chapters 2 and 3)?
3. What makes parks popular for walkers, runners and mountain bikers (Chapter 4)?
4. What is the scale of the discourse about national parks on Twitter, and how does this vary among countries (Chapter 5)?
5. Can social media data be used to predict park popularity (Chapters 3, 4 and 5)?

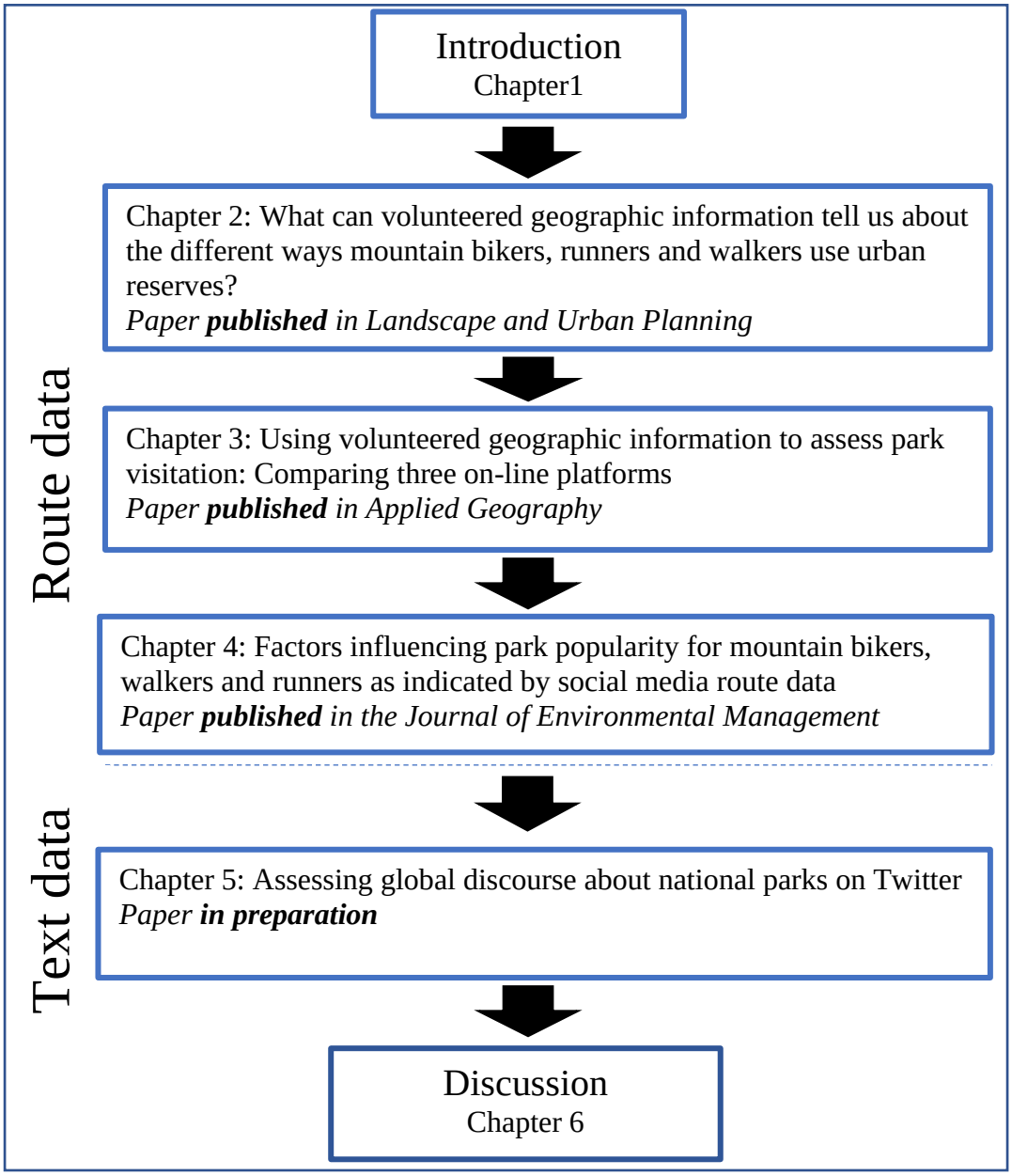

Figure 1.3. A schematic diagram of the thesis including the six chapters and outlining the predominant data type being assessed in each of the results chapters.

1.11. Structure of the thesis

This thesis has been prepared as a series of published and unpublished papers in accordance with Griffith University's Policy (Appendix A). The thesis consists of six

chapters including; an overall introduction (Chapter 1), the first results chapter/paper that assessed how walkers, runners and mountain bikers use trail networks within linked urban national parks and protected areas using GPS route data from MapMyFitness (Chapter 2), a chapter/paper assessing differences in route data posted to different social media platforms across three different parks (Chapter 3), an assessment of the factors influencing visitation to 40 national parks using GPS route data from three platforms (Chapter 4), a global scale assessment of the discourse on national parks using text data from Twitter (Chapter 5) and a general discussion of the implications of the results in the thesis (Chapter 6).

The four results chapters (Chapters 2-5) are in the form of manuscripts formatted to meet the requirements of the peer-reviewed academic journals where they have been published/prepared for. As a result, there is some repetition among the chapters including among the introductions, methods, results, discussions and reference lists and some differences in terminology (eg. protected areas/reserves/national parks, park visitors/users etc). Tables and figures in the individual results chapters were also formatted in accordance with the requirements of the individual journals as were references. The details of the papers, authors contributions, and publication status are also given at the beginning of the four results chapters. At the time of thesis submission, Chapters 2, 3 and 4 consist of published peer-reviewed academic journal articles, while Chapter 5 is in preparation for submission to a journal. Also, some of the research was presented at national and international conferences.

1.12. Published papers included as results chapters in the thesis

Chapter 2: Norman, P., Pickering, C.M. and Castley, G., 2019. What can volunteered geographic information tell us about the different ways mountain bikers, runners and walkers use urban reserves? *Landscape and Urban Planning, 185*, pp.180-190.

Chapter 3: Norman, P. and Pickering, C.M., 2017. Using volunteered geographic information to assess park visitation: Comparing three on-line platforms. *Applied Geography, 89*, pp.163-172.

Chapter 4: Norman, P. and Pickering, C.M., 2019. Factors influencing park popularity for mountain bikers, walkers and runners as indicated by social media route data. *Journal of Environmental Management, 249*, pp.109413.

1.13. Other publications completed during candidature but not included in this thesis:

1.13.1. Related published papers

Pickering, C.M. and **Norman, P.** (2017). Comparing impacts between hardened trails and informal recreational trails. *Journal of Environmental Management, 193*, pp.270-279.

Wraith, J., **Norman, P**. and Pickering, C., 2020. Orchid conservation and research: An analysis of gaps and priorities for globally Red Listed species. *Ambio*, pp.1-11.

Pickering, C. and **Norman, P.** (In review). Assessing discourses about controversial environmental management issues on social media: Tweeting about wild horses in a national park. *Journal of Environmental Management.*

1.13.2. Conference and symposium presentations

Norman, P. and Pickering, C.M. 2018. Using Volunteered Geographic Information (VGI) to assess the visitor use of parks. International Conference on Monitoring and Management of Visitors in Recreational and Protected Areas (MMV). Bordeaux, France. 28th-31st August 2018.

Norman, P. and Pickering, C.M. 2018. Using Volunteered Geographic Information (VGI) to assess the visitor use of parks. Using Social Media in Environmental Science Symposium. Gold Coast, Australia. 31st October 2018.

Norman, P. and Pickering, C.M. 2019. Assessing the spatial patterns of tweets about national parks: a global social media analysis. Annual Conference of the Institute of Australian Geographers. Hobart, Australia. 9th – 13th July 2019.

1.13.3. Government reports

Stevens, T., Warnken, J., Walters, K., Towner, S., **Norman, P.** and Piccolo. R., 2018. *Assessment of congestion and conflicting use management for the Gold Coast waterways, volume 1: Technical report.* Gold Coast Waterways Authority. Accessible from: https://gcwa.qld.gov.au/wp-content/uploads/2019/09/GCWA-Assessment-of-Congestion-and-Conflicting-Use-Management-for-the-Gold-Coast-Waterways.-Volume-1-Technical-Report.pdf

Stevens, T., Warnken, J., Walters, K., Towner, S., **Norman, P.** and Piccolo. R., 2018. *Assessment of congestion and conflicting use management for the Gold Coast waterways, volume 2: Discussion paper*. Gold Coast Waterways Authority. Accessible from: https://gcwa.qld.gov.au/wp-content/uploads/2019/09/GCWA-Assessment-of-Congestion-and-Conflicting-Use-Management-for-the-Gold-Coast-Waterways.-Volume-2-Discussion-Paper.pdf

References

Ankre, R., Fredman, P. and Lindhagen, A., 2016. Managers' experiences of visitor monitoring in Swedish outdoor recreational areas. *Journal of Outdoor Recreation and Tourism, 14*, pp.35-40.

Arnberger, A., Eder, R., Allex, B., Sterl, P. and Burns, R.C., 2012. Relationships between national-park affinity and attitudes towards protected area management of visitors to the Gesaeuse National Park, Austria. *Forest Policy and Economics, 19*, pp.48-55.

Ballantyne, M. and Pickering, C.M., 2015. The impacts of trail infrastructure on vegetation and soils: current literature and future directions. *Journal of Environmental Management, 164*, pp.53-64.

Balmford, A., Green, J.M., Anderson, M., Beresford, J., Huang, C., Naidoo, R., Walpole, M. and Manica, A., 2015. Walk on the wild side: estimating the global magnitude of visits to protected areas. *PLoS Biology, 13*(2), pp.e1002074.

Bassolas, A., Lenormand, M., Tugores, A., Gonçalves, B. and Ramasco, J.J., 2016. Touristic site attractiveness seen through Twitter. *EPJ Data Science, 5*(1), pp.12.

Becken, S., Stantic, B., Chen, J., Alaei, A.R. and Connolly, R.M., 2017. Monitoring the environment and human sentiment on the Great Barrier Reef: assessing the potential of collective sensing. *Journal of Environmental Management, 203*, pp.87-97.

Boyd, C., Brooks, T.M., Butchart, S.H., Edgar, G.J., Da Fonseca, G.A., Hawkins, F., Hoffmann, M., Sechrest, W., Stuart, S.N. and Van Dijk, P.P., 2008. Spatial scale and the conservation of threatened species. *Conservation Letters, 1*(1), pp.37-43.

Boyd, S.W. and Butler, R.W., 2009. *Tourism and the Canadian national park system. Tourism and national parks: International perspectives on development, histories, and change*, Routledge: New York City.

Brown, G., 2012. An empirical evaluation of the spatial accuracy of public participation GIS (PPGIS) data. *Applied Geography, 34*, pp.289-294.

Brown, G. and Fagerholm, N., 2015. Empirical PPGIS/PGIS mapping of ecosystem services: A review and evaluation. *Ecosystem Services, 13*, pp.119-133.

Buckley, R., Brough, P., Hague, L., Chauvenet, A., Fleming, C., Roche, E., Sofija, E. and Harris, N., 2019. Economic value of protected areas via visitor mental health. *Nature Communications, 10*(1), pp.1-10.

Campelo, M.B. and Mendes, R.M.N., 2016. Comparing webshare services to assess mountain bike use in protected areas. *Journal of Outdoor Recreation and Tourism, 15*, pp.82-88.

Caust, J. and Vecco, M., 2017. Is UNESCO World Heritage recognition a blessing or burden? Evidence from developing Asian countries. *Journal of Cultural Heritage, 27*, pp.1-9.

Cessford, G., 2003. Perception and reality of conflict: walkers and mountain bikes on the Queen Charlotte Track in New Zealand. *Journal for Nature Conservation, 11*(4), pp.310-316.

Chazdon, R.L., 2008. Beyond deforestation: restoring forests and ecosystem services on degraded lands. S*cience, 320*(5882), pp.1458-1460.

Costanza, R., de Groot, R., Sutton, P., Van der Ploeg, S., Anderson, S.J., Kubiszewski, I., Farber, S. and Turner, R.K., 2014. Changes in the global value of ecosystem services. *Global Environmental Change, 26*, pp.152-158.

Driml, S. and Common, M., 1995. Economic and financial benefits of tourism in major protected areas. *Australian Journal of Environmental Management, 2*(1), pp.19-29.

Eagles, P.F., 2014. Research priorities in park tourism. *Journal of Sustainable Tourism, 22*(4), pp.528-549.

Eagles, P.F., McCool, S.F., Haynes, C.D. and Phillips, A., 2002. *Sustainable tourism in protected areas: Guidelines for planning and management* (Vol. 8). IUCN: Gland.

Endomondo, 2019. Endomondo, viewed 20 December 2019. Retrieved from: https://www.endomondo.com/challenges/

Engen, S., Runge, C., Brown, G., Fauchald, P., Nilsen, L. and Hausner, V., 2018. Assessing local acceptance of protected area management using Public Participation GIS (PPGIS). *Journal for Nature Conservation, 43*, pp.27-34.

Facebook, 2019. Facebook, viewed 5 December 2019. Retrieved from: https://www.facebook.com/

FAO., 2019. Protecting ecosystem services and biodiversity: FAO's mission and solutions, viewed 10 September 2019. Retrieved from: http://www.fao.org/ecosystem-services-biodiversity/background/en/

Flanagin, A.J. and Metzger, M.J., 2008. The credibility of volunteered geographic information. *GeoJournal, 72*(3-4), pp.137-148.

Fletcher, R., 2011. Sustaining tourism, sustaining capitalism? The tourism industry's role in global capitalist expansion. *Tourism Geographies, 13*(3), pp.443-461.

Flickr, 2019. Flickr, viewed 5 December 2019. Retrieved from: https://www.flickr.com/

Foursquare, 2019. Foursquare, viewed 1 December 2019. Retrieved from: https://foursquare.com/

Furini, M. and Tamanini, V., 2015. Location privacy and public metadata in social media platforms: attitudes, behaviors and opinions. *Multimedia Tools and Applications, 74*(21), pp.9795-9825.

Gaston, K.J., Jackson, S.F., Cantu-Salazar, L. and Cruz-Piñón, G., 2008. The ecological performance of protected areas. *Annual Review of Ecology, Evolution, and Systematics, 39*, pp.93-113.

Gaveau, D.L.A., Curran, L.M., Paoli, G.D., Carlson, K.M., Wells, P., Besse-Rimba, A., Ratnasari, D. and Leader-Williams, N., 2012. Examining protected area effectiveness in Sumatra: importance of regulations governing unprotected lands. *Conservation Letters, 5*(2), pp.142-148.

Ghermandi, A. and Sinclair, M., 2019. Passive crowdsourcing of social media in environmental research: A systematic map. *Global Environmental Change, 55*, pp.36-47.

Geocaching, 2019. Geocaching, viewed 5 December 2019. Retrieved from: https://www.geocaching.com/play

GPSies, 2019a. GPSies, viewed 3 December 2019. Retrieved from: https://www.gpsies.com/

Gray, C.L., Hill, S.L., Newbold, T., Hudson, L.N., Börger, L., Contu, S., Hoskins, A.J., Ferrier, S., Purvis, A. and Scharlemann, J.P., 2016. Local biodiversity is higher inside than outside terrestrial protected areas worldwide. *Nature Communications, 7*, pp.12306.

Hadwen, W.L., Arthington, A.H., Boon, P.I., Taylor, B. and Fellows, C.S., 2011. Do climatic or institutional factors drive seasonal patterns of tourism visitation to protected areas across diverse climate zones in eastern Australia? *Tourism Geographies, 13*(2), pp.187-208.

Hall, C.M. and Frost, W., 2009. 21 The future of the national park concept. Tourism and National Parks: *International Perspectives on Development, Histories, and Change, 14*, pp.301.

Halpenny, E.A. and Blye, C.J., 2017. North American park agencies' evolving use of Twitter: A content analysis of 2014 and 2017 Tweets. *Parks, 23.2*, pp.79.

Hausmann, A., Toivonen, T., Slotow, R., Tenkanen, H., Moilanen, A., Heikinheimo, V. and Di Minin, E., 2018. Social media data can be used to understand tourists' preferences for nature-based experiences in protected areas. *Conservation Letters, 11*(1), pp.e12343.

Hockett, K.A., Clark, Y.F., Leung, J.L. and Park, L., 2010. *Deterring off-trail hiking in protected natural areas: Evaluating options with surveys and unobtrusive observation*. Virginia Tech College of Natural Resources, Forestry/Recreation Resources Management: Blacksburg.

Hunt, C.A., Durham, W.H., Driscoll, L. and Honey, M., 2015. Can ecotourism deliver real economic, social, and environmental benefits? A study of the Osa Peninsula, Costa Rica. *Journal of Sustainable Tourism, 23*(3), pp.339-357.

Inkpen, D., Liu, J., Farzindar, A., Kazemi, F. and Ghazi, D., 2017. Location detection and disambiguation from Twitter messages. *Journal of Intelligent Information Systems, 49*(2), pp.237-253.

Instagram, 2019. Instagram, viewed 5 December 2019. Retrieved from: https://www.instagram.com/

IUCN., 2019. Protected Area Categories, viewed 10 September 2019. Retrieved from: https://www.iucn.org/theme/protected-areas/about/protected-area-categories

Jurado Rota, J., Pérez Albert, M.Y. and Serrano Giné, D., 2019. Visitor monitoring in protected areas: An approach to Natura 2000 sites using Volunteered Geographic Information (VGI). *Geografisk Tidsskrift-Danish Journal of Geography, 119*(1), pp.69-83.

Karanth, K.K. and DeFries, R., 2011. Nature-based tourism in Indian protected areas: New challenges for park management. *Conservation Letters, 4*(2), pp.137-149.

Kern, M.L., Park, G., Eichstaedt, J.C., Schwartz, H.A., Sap, M., Smith, L.K. and Ungar, L.H., 2016. Gaining insights from social media language: Methodologies and challenges. *Psychological Methods, 21*(4), pp.507.

Korpilo, S., Virtanen, T. and Lehvävirta, S., 2017. Smartphone GPS tracking—Inexpensive and efficient data collection on recreational movement. *Landscape and Urban Planning, 157*, pp.608-617.

Kovacs-Györi, A., Ristea, A., Kolcsar, R., Resch, B., Crivellari, A. and Blaschke, T., 2018. Beyond spatial proximity—classifying parks and their visitors in London based on spatiotemporal and sentiment analysis of Twitter data. *ISPRS International Journal of Geo-Information, 7*(9), pp.378.

Lacki, M.J., 2000. Effect of trail users at a maternity roost of Rafinesque's big-eared bats. *Journal of Cave and Karst Studies, 62*(3), pp.163-168.

Lecina-Diaz, J., Alvarez, A., De Cáceres, M., Herrando, S., Vayreda, J. and Retana, J., 2019. Are protected areas preserving ecosystem services and biodiversity? Insights from Mediterranean forests and shrublands. *Landscape Ecology, 34*(10), pp.2307-2321.

Leetaru, K. 2019. Forbes: Visualizing Seven Years of Twitter's Evolution: 2012-2018, viewed 10 October 2019. Retrieved from https://www.forbes.com/sites/kalevleetaru/2019/03/04/visualizing-seven-years-of-twitters-evolution-2012-2018/#66c249b47ccf

Leung, Y.F., Spenceley, A., Hvenegaard, G. and Buckley, R., 2018. *Tourism and visitor management in protected areas: Guidelines for sustainability*. IUCN. Gland.

Levin, N., Lechner, A.M. and Brown, G., 2017. An evaluation of crowdsourced information for assessing the visitation and perceived importance of protected areas. *Applied Geography*, *79*, pp.115-126.

Lime, D.W. and Stankey, G.H., 1971. Carrying capacity: maintaining outdoor recreation quality. In: *The Forest Recreation Symposium. State University of New York College of Forestry*; 1971 October 12-14: US Department of Agriculture, Forest Service, Northeastern Forest Experiment Station. pp. 174-184.

Ma, Y., Chen, G., Grumbine, R.E., Dao, Z., Sun, W. and Guo, H., 2013. Conserving plant species with extremely small populations (PSESP) in China. *Biodiversity and Conservation*, *22*(3), pp.803-809.

Macdonald, D., Jacobsen, K., Burnham, D., Johnson, P. and Loveridge, A., 2016. Cecil: a moment or a movement? Analysis of media coverage of the death of a lion, *Panthera leo*. *Animals*, *6*(5), p.26.

Mann, C. and Absher, J.D., 2008. Recreation conflict potential and management implications in the northern/central Black Forest Nature Park. *Journal of Environmental Planning and Management*, *51*(3), pp.363-380.

Manning, R.E., 2002, January. How much is too much? Carrying capacity of national parks and protected areas. In Monitoring and management of visitor flows in recreational and protected areas. Conference proceedings. pp. 306-313.

Manning, R.E., Anderson, L.E. and Pettengill, P., 2017. *Managing outdoor recreation: Case studies in the national parks*. CABI. Boston.

MapMyFitness. (2019), MapMyFitness, viewed 5 December 2019. Retrieved from: http://www.mapmyfitness.com/

Monteiro, L., 2016. Using GPS data from Web 2.0 platforms to assess informal trail network and its impacts in protected areas. In Monitoring and management of visitor flows in recreational and protected areas. Conference proceedings. pp.270-272.

Monz, C.A., Cole, D.N., Leung, Y.F. and Marion, J.L., 2010. Sustaining visitor use in protected areas: future opportunities in recreation ecology research based on the USA experience. *Environmental Management*, *45*(3), pp.551-562.

Monz, C.A., Pickering, C.M. and Hadwen, W.L., 2013. Recent advances in recreation ecology and the implications of different relationships between recreation use and ecological impacts. *Frontiers in Ecology and the Environment*, *11*(8), pp.441-446.

Mukul, S.A., Sohel, M.S.I., Herbohn, J., Inostroza, L. and König, H., 2017. Integrating ecosystem services supply potential from future land-use scenarios in protected area management: A Bangladesh case study. *Ecosystem Services*, 26, pp.355-364.

Munar, A.M. and Jacobsen, J.K.S., 2014. Motivations for sharing tourism experiences through social media. *Tourism Management*, *43*, pp.46-54.

Neuvonen, M., Pouta, E., Puustinen, J. and Sievänen, T., 2010. Visits to national parks: Effects of park characteristics and spatial demand. *Journal for Nature Conservation*, *18*(3), pp.224-229.

Newsome, D. and Davies, C., 2009. A case study in estimating the area of informal trail development and associated impacts caused by mountain bike activity in John Forrest National Park, Western Australia. *Journal of Ecotourism*, *8*(3), pp.237-253.

Newsome, D., Moore, S.A. and Dowling, R.K., 2012. *Natural area tourism: Ecology, impacts and management* (Vol. 58). Channel view publications: Bristol.

Nike Run Club, 2019. Nike Run Club, viewed 3 December 2019. Retrieved from: https://www.nike.com/au/en_gb/c/nike-plus/running-app-gps

Orsi, F. and Geneletti, D., 2013. Using geotagged photographs and GIS analysis to estimate visitor flows in natural areas. *Journal for Nature Conservation*, *21*(5), pp.359-368.

Palomo, I., Martín-López, B., Potschin, M., Haines-Young, R. and Montes, C., 2013. National Parks, buffer zones and surrounding lands: mapping ecosystem service flows. *Ecosystem Services*, *4*, pp.104-116.

Paracchini, M.L., Zulian, G., Kopperoinen, L., Maes, J., Schägner, J.P., Termansen, M., Zandersen, M., Perez-Soba, M., Scholefield, P.A. and Bidoglio, G., 2014. Mapping cultural ecosystem services: A framework to assess the potential for outdoor recreation across the EU. *Ecological Indicators*, *45*, pp.371-385.

Pickering, C., Castley, J.G., Hill, W. and Newsome, D., 2010a. Environmental, safety and management issues of unauthorised trail technical features for mountain bicycling. *Landscape and Urban Planning*, *97*(1), pp.58-67.

Pickering, C.M., Hill, W., Newsome, D. and Leung, Y.F., 2010b. Comparing hiking, mountain biking and horse riding impacts on vegetation and soils in Australia and the United States of America. *Journal of Environmental Management*, *91*(3), pp.551-562.

Pickering, C.M. and Hill, W., 2007. Impacts of recreation and tourism on plant biodiversity and vegetation in protected areas in Australia. *Journal of Environmental Management*, *85*(4), pp.791-800.

Pickering, C.M. and Norman, P., 2017. Comparing impacts between formal and informal recreational trails. *Journal of Environmental Management*, *193*, pp.270-279.

Pickering, C.M. and Rossi, S., 2016. Mountain biking in peri-urban parks: Social factors influencing perceptions of conflicts in three popular National Parks in Australia. *Journal of Outdoor Recreation and Tourism*, *15*, pp.71-81.

Plieninger, T., Dijks, S., Oteros-Rozas, E. and Bieling, C., 2013. Assessing, mapping, and quantifying cultural ecosystem services at community level. *Land Use Policy, 33*, pp.118-129.

Primack, B.A., Shensa, A., Sidani, J.E., Whaite, E.O., Lin, L., Rosen, D., Colditz, J.B., Radovic, A. and Miller, E., 2017. Social media use and perceived social isolation among young adults in the US. *American Journal of Preventive Medicine, 53*(1), pp.1-8.

Ramp, D., Wilson, V.K. and Croft, D.B., 2006. Assessing the impacts of roads in peri-urban reserves: road-based fatalities and road usage by wildlife in the Royal National Park, New South Wales, Australia. *Biological Conservation, 129*(3), pp.348-359.

Rossi, S.D., Barros, A., Walden-Schreiner, C. and Pickering, C., 2019. Using social media images to assess ecosystem services in a remote protected area in the Argentinean Andes. *Ambio*, pp.1-15.

Rossi, S.D., Byrne, J.A. and Pickering, C.M., 2015. The role of distance in peri-urban national park use: Who visits them and how far do they travel? *Applied Geography, 63*, pp.77-88.

Santos, T., Mendes, R.N. and Vasco, A., 2016. Recreational activities in urban parks: Spatial interactions among users. *Journal of Outdoor Recreation and Tourism, 15*, pp.1-9.

Sayan, M.S. and Atik, M., 2011. Recreation carrying capacity estimates for protected areas: A study of Termessos National Park. *Ekoloji, 20*(78), pp.66-74.

Schipperijn, J., Ekholm, O., Stigsdotter, U.K., Toftager, M., Bentsen, P., Kamper-Jørgensen, F. and Randrup, T.B., 2010. Factors influencing the use of green space: Results from a Danish national representative survey. *Landscape and Urban Planning, 95*(3), pp.130-137.

Schneble, C.O., Elger, B.S. and Shaw, D., 2018. The Cambridge Analytica affair and internet-mediated research. *EMBO Reports, 19*(8).

Schwartz, R. and Halegoua, G.R., 2015. The spatial self: Location-based identity performance on social media. *New Media and Society, 17*(10), pp.1643-1660.

Siikamäki, P., Kangas, K., Paasivaara, A. and Schroderus, S., 2015. Biodiversity attracts visitors to national parks. *Biodiversity and Conservation, 24*(10), pp.2521-2534.

Singh, T. and Kumari, M., 2016. Role of text pre-processing in Twitter sentiment analysis. *Procedia Computer Science, 89*, pp.549-554.

Spencer, D.M., 2012. Understanding recreation activity participants as clientele groups of multi-use recreation areas. *Annals of Leisure Research, 15*(3), pp.227-242.

Strava, 2019. Strava, viewed 5 December 2019. Retrieved from: https://www.strava.com/

Struebig, M.J., Wilting, A., Gaveau, D.L., Meijaard, E., Smith, R.J., Abdullah, T., Abram, N., Alfred, R., Ancrenaz, M., Augeri, D.M. and Belant, J.L., 2015. Targeted conservation to safeguard a biodiversity hotspot from climate and land-cover change. *Current Biology, 25*(3), pp.372-378.

Symmonds, M.C., Hammitt, W.E. and Quisenberry, V.L., 2000. Managing recreational trail environments for mountain bike user preferences. *Environmental Management, 25*(5), pp.549-564.

Tarrant, M.A. and Smith, E.K., 2002. The use of a modified importance-performance framework to examine visitor satisfaction with attributes of outdoor recreation settings. *Managing Leisure*, *7*(2), pp.69-82.

Tenkanen, H., Di Minin, E., Heikinheimo, V., Hausmann, A., Herbst, M., Kajala, L. and Toivonen, T., 2017. Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas. *Scientific Reports, 7*(1), pp.17615.

Toivonen, T., Heikinheimo, V., Fink, C., Hausmann, A., Hiippala, T., Järv, O., Tenkanen, H. and Di Minin, E., 2019. Social media data for conservation science: a methodological overview. *Biological Conservation*, *233*, pp.298-315.

TripAdvisor, 2019. TripAdvisor, viewed 3 December 2019. Retrieved from: https://www.tripadvisor.com.au/

Tyrrell, I., 2012. America's National Parks: the transnational creation of national space in the progressive era. *Journal of American Studies, 46*(1), pp.1-21.

Twitter, 2019. Twitter, viewed 5 December 2019. Retrieved from: https://twitter.com/

United Nations Environment World Conservation Monitoring Centre., 2019. Protected Planet. Available from: https://www.protectedplanet.net/

United National Environment Programme., 2018. *Protected Planet Report 2018.* Accessible from: https://livereport.protectedplanet.net/pdf/Protected_Planet_Report_2018.pdf

Venter, O., Fuller, R.A., Segan, D.B., Carwardine, J., Brooks, T., Butchart, S.H., Di Marco, M., Iwamura, T., Joseph, L., O'Grady, D. and Possingham, H.P., 2014. Targeting global protected area expansion for imperiled biodiversity. *PLoS Biology, 12*(6), pp.e1001891.

Walden-Schreiner, C., Rossi, S.D., Barros, A., Pickering, C. and Leung, Y.F., 2018. Using crowd-sourced photos to assess seasonal patterns of visitor use in mountain-protected areas. *Ambio, 47*(7), pp.781-793.

Watson, J.E., Evans, M.C., Carwardine, J., Fuller, R.A., Joseph, L.N., Segan, D.B., Taylor, M.F., Fensham, R.J. and Possingham, H.P., 2011. The capacity of Australia's protected area system to represent threatened species. *Conservation Biology, 25*(2), pp.324-332.

Wolf, I.D., Croft, D.B. and Green, R.J., 2019. Nature conservation and nature-based tourism: A paradox? *Environments, 6*(9), pp.104.

Wolf, I.D., Hagenloh, G. and Croft, D.B., 2012. Visitor monitoring along roads and hiking trails: How to determine usage levels in tourist sites. *Tourism Management, 33*(1), pp.16-28.

Wolf, I.D., Wohlfart, T., Brown, G. and Lasa, A.B., 2015. The use of public participation GIS (PPGIS) for park visitor management: A case study of mountain biking. *Tourism Management, 51*, pp.112-130.

Worboys, G.L., Lockwood, M., Kothari, A., Feary, S. and Pulsford, I. eds., 2015. *Protected area governance and management*. ANU Press. Canberra.

Wraith, J. and Pickering, C., 2017. Tourism and recreation a global threat to orchids. *Biodiversity and Conservation, 26*(14), pp.3407-3420.

Zeppel, H., 2010. Managing cultural values in sustainable tourism: Conflicts in protected areas. *Tourism and Hospitality Research, 10*(2), pp.93-115.

Chapter 2. What can volunteered geographic information tell us about the different ways mountain bikers, runners and walkers use urban reserves?

In the previous introductory chapter, the background, general literature, aims and structure of the thesis were outlined. The following four chapters present the results of the thesis in the form of papers with their own aims, introduction, methods, results, discussions and reference lists. The first results chapter 'What can volunteered geographic information tell us about the different ways mountain bikers, runners and walkers use urban reserves?' assesses differences in park visitation for the three activities using GPS route data from the social media platform, MapMyFitness. This includes comparing spatial and temporal patterns in the use of trail networks in a series of national parks and local reserves in the city of Brisbane, Australia. Results from the route data were also compared with that off trail counters to assess validity.

This chapter consists of the published version of a paper co-authored with my supervisors. The citation of this paper is **Norman, P.,** Pickering, C.M., and Castley, G. (2019) What can volunteered geographic information tell us about the different ways mountain bikers, runners and walkers use urban reserves? *Landscape and Urban Planning, 185,* 180-190. My contribution of the paper includes: design, methods, data collection and analysis, mapping analyses, all figures, drafting the chapter and submitting it to the journal.

(Signed):

(Date):

Patrick Norman

(Counter signed):

(Date):

Doctor Guy Castley

(Counter signed):

(Date):

Professor Catherine Marina Pickering

Landscape and Urban Planning 185 (2019) 180–190

Contents lists available at ScienceDirect

Landscape and Urban Planning

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

Research Paper

What can volunteered geographic information tell us about the different ways mountain bikers, runners and walkers use urban reserves?

Patrick Norman*, Catherine Marina Pickering, Guy Castley

Environment Futures Research Institute, Griffith University, Gold Coast 4222, Australia

ARTICLE INFO

Keywords:
Visitor monitoring
Volunteered geographic information
Urban reserves
MapMyFitness
Park management
Trail use

ABSTRACT

Urban nature reserves are increasingly popular for recreational activities but who uses them, for what activities, and when and where do visitors go? Volunteered geographic information has only recently started to be used by managers to examine patterns of trail use in reserves. As yet, little is known about the benefits and limitations of this type of data. Therefore, we compared how mountain bikers, runners and walkers use popular reserves (Daisy Hill Conservation Park, Venman Bushland National Park and Bayview Conservation Area) close to Brisbane, Australia, based on extracted route data from the fitness tracking application MapMyFitness. Routes for all three activity types were displayed and analysed using ArcGIS. Mountain biking was more popular (49% of the 948 routes), than walking (27%) or running (24%) across the three reserves. Route data was effective at predicting the relative popularity of specific trails with significant linear regressions when compared with fixed-point trail counts ($R^2 = 0.681$). Bikers went further, used a greater range of trail combinations and used more reserves per trip than the other two activities. Weekends were popular for all three activities, but more so for biking (57%), than walking (44%), or running (43%). The results highlight how volunteered geographic information complements other trail data allowing urban planners and managers to better assess visitor movements in reserves. A number of specific management challenges were identified, such as bikers using more areas within and across the reserves, and in some cases, the use of private land and unauthorised access points.

A full copy of this paper is available at the publisher's website via the following link:

https://www.sciencedirect.com/science/article/pii/S0169204618304730

Chapter 3. Using volunteered geographic information to assess park visitation: Comparing three on-line platforms

The previous chapter assessed how mountain bikers, walkers and runners used trails within parks and reserves using data from a single fitness platform. It demonstrated that data from MapMyFitness could provide valuable data about the different way's walkers, runners and mountain bikers use urban parks. In the next chapter, differences in patterns of visitation were assessed for walkers and runners among social media platforms but for three different types of parks, varying from a peri-urban to a remote park. Specifically, GPS route data from MapMyFitness, Wikiloc and GPSies was used to map visitation on and off the formal trail networks. The benefits and limitations of the three platforms for researchers and managers was also assessed.

This chapter consists of the published version of a paper co-authored with my principle supervisor. The citation of this paper is **Norman, P.,** and Pickering, C.M. (2017). Using volunteered geographic information to assess park visitation: Comparing three on-line platforms. *Applied Geography, 89,* 163-172. My contribution of the paper includes: design, methods, data collection and analysis, liaising with park managers, mapping, all figures, drafting the chapter and submitting to the journal.

(Signed):

(Date):

Patrick Norman

(Counter signed):

(Date):

Professor Catherine Marina Pickering

Applied Geography 91 (2018) 168

Contents lists available at ScienceDirect

Applied Geography

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

Corrigendum to "Using volunteered geographic information to assess park visitation: Comparing three on-line platforms" [Applied Geography 89C (2017) 163-72]

Patrick Norman*, Catherine Marina Pickering

Griffith University, Australia

The following abstract was missing from the original version of the above-mentioned article:

"The emergence of web platforms where people share geographic data provides a new way to monitor how people use landscapes. This includes assessing park use, including for different recreational activities such as walking and running: but which platforms are best to use and why? To start to address these questions we compared geographic route data for park visitors from three large web share platforms: GPSies, MapMyFitness and Wikiloc. We obtained route data from all three platforms for an urban park (Daisy Hill), a peri-urban park (Lamington) and a more remote park (Mount Barney) in Australia. The largest number of user posted routes for the three parks was from MapMyFitness (1041 routes), including information on how walkers (572) and runners (469) use trails, particularly in Daisy Hill (705). However, the popularity of this platform is declining. In contrast, Wikiloc provided information on off track use, particularly in the more remote Mount Barney National Park. There were few trails on GPSies from the three parks, despite the platform's popularity in Europe. The results indicate that the best platform to use will depend on the questions asked and the popularity of the platform with local users. Volunteered geographic data can be a cost and time efficient way to collect visitor data compared to some other methods, such as trail counters and observations, but it only reflects particular users, and is geographically and temporally variable as the popularity of platforms change among users."

The authors would like to apologize for any inconvenience that this may have caused to the readers.

A full copy of this paper is available at the publisher's website via the following link:

https://www.sciencedirect.com/science/article/abs/pii/S0143622817306409

Chapter 4. Factors influencing park popularity for mountain bikers, walkers and runners as indicated by social media route data

In the previous results chapters, route data was used to compare the use of trails for different activities within parks. The next chapter takes a broader scale approach, and assesses what makes parks popular for walking, running and mountain biking and why. Specifically, the number of routes posted about 40 different parks in south-eastern Queensland, from three platforms for the three activities was compared. A range of factors including geographic, social and park facilities, were then modelled to determine which best predicted the popularity of the parks for each platform and activity using data from each of the three platforms.

This chapter consists of the published version of a paper co-authored with my principle supervisor. The citation of this paper is **Norman, P.,** and Pickering, C.M. (2019). Factors influencing park popularity for mountain bikers, walkers and runners as indicated by social media route data. *Journal of Environmental Management, 249*, 109413. My contribution of the paper includes: design, methods, data collection, data analysis, mapping analyses, all figures, drafting the chapter and submitting to the journal.

(Signed):

(Date):

Patrick Norman

(Counter signed):

(Date):

Professor Catherine Marina Pickering

Journal of Environmental Management 249 (2019) 109413

Contents lists available at ScienceDirect

Journal of Environmental Management

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

Research article

Factors influencing park popularity for mountain bikers, walkers and runners as indicated by social media route data

Patrick Norman*, Catherine Marina Pickering

Griffith University, Environment Futures Research Institute, Griffith University, Gold Coast, 4222, Australia

ARTICLE INFO

Keywords:
Volunteered geographic information
Visitor monitoring
Geographic information systems
National parks
Recreation
Distance decay models

ABSTRACT

For most protected areas visitor data is limited, even for popular activities such as mountain biking, walking and running. With the exponential increase in volunteered geographic information, it is now possible to assess visitation to parks by people using social media. The number of routes posted to MapMyFitness, Strava and Wikiloc were used to assess the relative popularity of 40 national parks in South East Queensland, Australia for mountain biking, walking and running. Then, using generalised linear models, the topographic, geographic, management and environmental variables that best predicted park popularity among users were identified. There were more than 450,000 visits to the parks, with 98% of routes posted on Strava. Distance (road and direct distance) from large urban areas best predicted visitation using routes from MapMyFitness and Strava with three parks within 2 km of urban areas experiencing over 70,000 visits (Strava). For urban parks, recreational trail length best predicted usage, while for more remote parks, the direct distance to urban areas remained the most important factor. In contrast, people using the adventure platform Wikiloc preferred more remote parks with rugged terrain. The results highlight factors affecting park popularity including distance. With the expansion and densification of cities, including in Australia, urban and peri-urban parks are likely to experience increasing levels of use for fitness-based activities while more remote parks may remain attractive among people focused on adventure.

A full copy of this paper is available at the publisher's website via the following link:

https://www.sciencedirect.com/science/article/pii/S0301479719311314

Chapter 5. Assessing global discourse about national parks on Twitter

In the previous three results chapters, data from GPS route sharing social media platforms was used to assess visitation within and among national parks in one region: south-east Queensland, Australia. The following chapter takes a broader scale approach and uses a different type of data: text from the platform Twitter, to determine which parks are talked about publicly online and by whom. As a vast number of posts about parks only contain text from the platform, this chapter assesses the scale of discourse about National Parks and if the number of posts on social media reflect the number of visits to park.

This chapter consists of the unpublished version of a paper co-authored with my principle supervisor. The bibliographical information the chapter is **Norman, P.,** & Pickering, C. M. (in preparation) Assessing the discourse of national parks; a global Twitter analysis. My contribution of the paper includes: design, methods, data collection, data analysis, mapping, all figures and drafting the chapter.

(Signed):

(Date):

Patrick Norman

(Counter signed):

(Date):

Professor Catherine Marina Pickering

Assessing the discourse of national parks: A global Twitter analysis

Abstract

With the spread of social media, there is increasing discussion about national parks online globally, but which parks are talked about, by whom and how does this relate to visitation and park size? More than two million tweets were posted about national parks in six months on Twitter. Overall, 264 parks were tweeted about more than 100 times and of these, parks in the USA dominate the discourse (62.5%, 144 860 tweets). More tweets were sent by people from North America (59.4% of all tweets), compared to Europe (18.0%), Asia (9.1%), Africa (8.1%), Oceania (4.2%) and South America (1.5%). The rank order popularity of parks on Twitter was similar to actual visitation for the 40 most popular parks on Twitter, but not park size. Twitter can be a valuable source of data about parks particularly for assessing public views on management but, as with all social media, there are both benefits and limitations.

Keywords

nature-based tourism, park management, protected areas, public discourse, social media, volunteered geographic information

Introduction

Protected areas are critical for the conservation of species and ecosystems (Worboys et al., 2015). They provide a wide range of ecosystem services worth over US$600

billion per year, which underpins the wellbeing of billions of people (Worboys et al., 2015) and are popular destinations for tourism and recreation with an estimated 8 billion visits annually globally (Balmford et al., 2015). Although there are different types of protected areas, many are termed ‘national parks’ including high-profile protected areas with large visitor numbers (Newsome et al., 2012; Worboys et al., 2015). For example, national parks in the USA receive 85.4 million visits per year (National Park Service, 2017), and the term ‘national park’ is now used for protected areas in more than 85 countries (United Nations Environment World Conservation Monitoring Centre, 2019). Most of these areas are IUCN Category II protected areas, set aside to conserve natural processes with minimal impacts from humans, as well as providing recreational, scientific and educational benefits and opportunities (Eagles et al., 2002; Worboys et al., 2015).

There is considerable interest in national parks including how they are managed, and this is reflected in public debates about management decisions, particularly where there are conflicting views among stakeholders (Burns et al., 2011; Chamberlain, 2012; Marijnen and Verweijen, 2016; Pettebone et al., 2013). With the principles of public participation in decision making increasingly seen as fundamental to park management, agencies need to know more about people’s views on parks (Worboys et al., 2015). Traditional methods used to assess public opinions about parks include monitoring news sources, soliciting input from the public and community groups into management process via submissions, as well as conducting focus groups and surveys, both within parks and more broadly (Newsome et al., 2012; Worboys et al., 2015, Veal 2017).

With the advent of web 2.0, people are increasingly sharing information on social media, including where they have been, what they have done and their opinions, including on national parks (Ghermandi and Sinclair, 2019; Norman et al., 2019). Research quantifying the scale of discourse on issues, known as culturomics (Ladle et al., 2016), is starting to be used to better understand people's relationships with parks (Ghermandi and Sinclair, 2019; Tenkanen et al., 2017). There are a range of social media platforms now being used to improve the management of parks, (Heikinheimo et al., 2017; Norman and Pickering, 2017; Tenkanen et al., 2017; Walden-Schreider et al., 2018; Wood et al., 2016), with text based platform Twitter, considered to be a valuable source of public data about discourses on parks (Ghermandi and Sinclair, 2019; Nara et al., 2018).

Twitter is a micro-blogging platform where users (tweeters) can send brief (up to 280 characters) lines of text (tweets) on the platform that can be read by others (Twitter, 2019). It has a global reach with over 500 million tweets posted per day (Twitter Business, 2019). Its impact is increasing with Twitter used to create and respond to news (Orellana-Rodriguez and Keane, 2018), shape politics and policy (Tumasjan et al., 2011), affect human health (Mowery et al., 2017) and influence decisions made by a wide range of government agencies (Kim et al., 2015). Twitter data is relatively easy to access and as a result has become one of the most popular platforms for assessing public discourse and salience on a range of issues relevant to national parks (Ghermandi and Sinclair, 2019) and has been used to assess the relative popularity of parks in South Africa and Finland (Tenkanen et al., 2017).

To assist managers and researchers in assessing the usefulness of this novel source of data, we ask the following questions: (1) what is the global scope of the discourse on

Twitter regarding national parks, including (2) which parks were discussed (3) who talked about them and where were they from, (4) did visitation or park size relate to the number of tweets about a park, and (5) what is the source platform of the tweets, and (6) can geolocation data for tweets be used to assess within-park usage? Answers to these questions can assist in gauging both the benefits and limitations of Twitter for park management.

Methods

To provide a measure of the global scale of discourse relating to global national parks on Twitter, data for all tweets that included the term 'national park' was collected over a 6-month period. One challenge when searching Twitter for specific content, is how to select terms to extract data from the over 180 billion tweets sent per year (Twitter Business, 2019). For the current study a range of terms including 'protected area', 'conservation area' and 'state park' were tested with the phrase 'national park' the most effective, in part because it forms part of the name of a large number of protected areas globally, particularly IUCN Protected Area Category II protected areas (Eagles et al., 2002; Worboys et al., 2015). Tweet metadata and texts were collected using the archiving Google sheet, TAGS for Twitter (Hawksey, 2018) which provides an automated way to retrieve large numbers of tweets. All tweets which included the term 'national park' and associated metadata were collected in an online archiving Google sheet (Google, 2018) from 10 July 2018 to 9 January 2019 (183 days), with returns of up to 18 000 tweets per hour. Metadata for the tweets included tweet text, user (tweeter) identification number, user location, time and date of the post (given in GMT), and information on the platform used to post the tweet. Due to

the large number of tweets, processing of the data used the computing language R through R studio (R Core Team, 2019; RStudio Team, 2019).

Firstly, tweets were filtered to remove retweets, so the data better reflected distinct/original views. The original tweets were then searched to establish which park(s) were discussed. To do this, a series of n-grams were performed using the R package Tidytext to extract the most common three, four, five and six-word sequences within each tweet (Silge and Robinson, 2016). From these, text sequences each ending in the word ‘park’ was collected and assessed to establish if the sequence was in fact a park name. Tweets for all parks mentioned in more than 100 tweets were used in further analysis. Where different parks had the same name, such as Glacier National Park in Canada and the USA and Pigeon Island National Park in St Lucia and Sri Lanka, tweets were assessed to see if it would be possible to separate the parks by screening the text for the name of the country, but most did not include this information in the tweet. Therefore, parks with the same name had to be removed from the study due to uncertainties in which park was referenced to in the tweet.

For the remaining parks with over 100 tweets, additional information about each park was collected from the Protected Planet database on protected areas (United Nations Environment World Conservation Monitoring Centre, 2019) (Appendix 1). This included assigning each park to a country and continent. Also, where tweets referred to multiple parks, the first park mentioned was nominated as the one referred to for that tweet, which resulted in some parks having less than 100 tweets. Finally, each tweet was attributed to one of the commonly tweeted parks.

To determine the origin of tweeters, location data provided by tweeters was used to geo-code their origin using the R function geocodeOSM from the tmaptools package (Tennekes, 2019). This tool geocodes text using OpenStreetMap (OpenStreetMap, 2019) and results in a single set of coordinates per tweeter. After the coordinate data was cleaned to remove any anomalous locations, it was then used to determine if a tweet was sent by someone from the same country/continent (domestic) or if they were from another country or region (international) to the park mentioned in the tweet. This does not indicate where the tweet was sent from and relies on the accuracy of the location data provided by the tweeter. The location where the tweeter was based was then mapped using QGIS and the R package 'sp' (Pebesma and Bivand, 2005; QGIS Development Team, 2019; R Core Team, 2019).

In order to assess if popularity on Twitter related to park characteristics, data on visitor numbers, size of the parks and IUCN categories was obtained for each of the 40 most commonly tweeted parks (Appendix 2). Data on park size, country and IUCN category was obtained from the Protected Planet database (United Nations Environment World Conservation Monitoring Centre, 2019), while data on visitor numbers was obtained for each park from the most recent available data (Appendix 1). To assess if park size and visitation were correlated with the number of tweets and tweeters, Spearman's rank correlation coefficients were used, as values were not normally distributed. A Spearman's rank correlation coefficient was also conducted to determine if the total number of tweets about national parks per country was correlated with the popularity of Twitter (Hootsuite and we are social, 2018). This was assessed for the 14 countries with at least one park with over 100 tweets and were in the top 20 highest number of users on Twitter globally (Hootsuite and we are social, 2018).

Results

What was the scale of the discourse on national parks?

Globally there were 2.24 million tweets and retweets posted by 749 000 different tweeters about national parks over the six months, with an average of 370 000 per month (Table 1). Of these, 437 789 were original tweets (not retweeted), sent from 194 467 twitter accounts. Although 1 784 tweets about parks were sent from a single account, most Tweeters (72%) only posted once about national parks over the six months. Tweets were posted by users of 45 different languages, but nearly all were posted in English (92%), with Spanish and Japanese Tweets accounting for another 2.1% and 0.85% respectively.

Table 1. Details on the number of tweets that included the term 'national park' posted over the six-month sample period.

Steps in data collection	**# tweets**	**# tweeters**
All tweets posted including the term 'national park'	2 078 220	749 500
Only original tweets (once retweets removed)	437 789	194 464
Original tweets referring to common parks (i.e. the 264 parks with more than 100 tweets each)	231 671	93 107
Original tweets with location data relating to a country	150 932	51 801
Original tweets with geolocation data	36 921	19 593

Which parks were discussed on Twitter?

Some national parks were talked about a lot, some a moderate amount, but most were either uncommon in the discourse or, if they were tweeted about, it was not using the

term ‘national park’ (Table 1, Figure 1). The discourse on Twitter encompassed a wide range of parks, with 264 parks from 55 countries and all continents apart from Antarctica mentioned 100 or more times on Twitter (Figure 1). The discourse was dominated by parks in one country, with 66 parks in the USA tweeted about more than 100 times, and of the top 40 parks discussed on Twitter, 27 were from the USA. Just three of these parks accounted for 21.8% of all the tweets about the 264 parks, with Yosemite mentioned in 23 926 tweets, Yellowstone in 14 170 tweets and Rocky Mountain National Park in 9 568 tweets (Table 1, Figure 1). Outside the USA, the most commonly discussed national parks were Banff in Canada (7 631 tweets), and Kruger in South Africa (6 419). For the majority of the parks, there were far fewer tweets, with 230 national parks with between 100 and 1000 tweets.

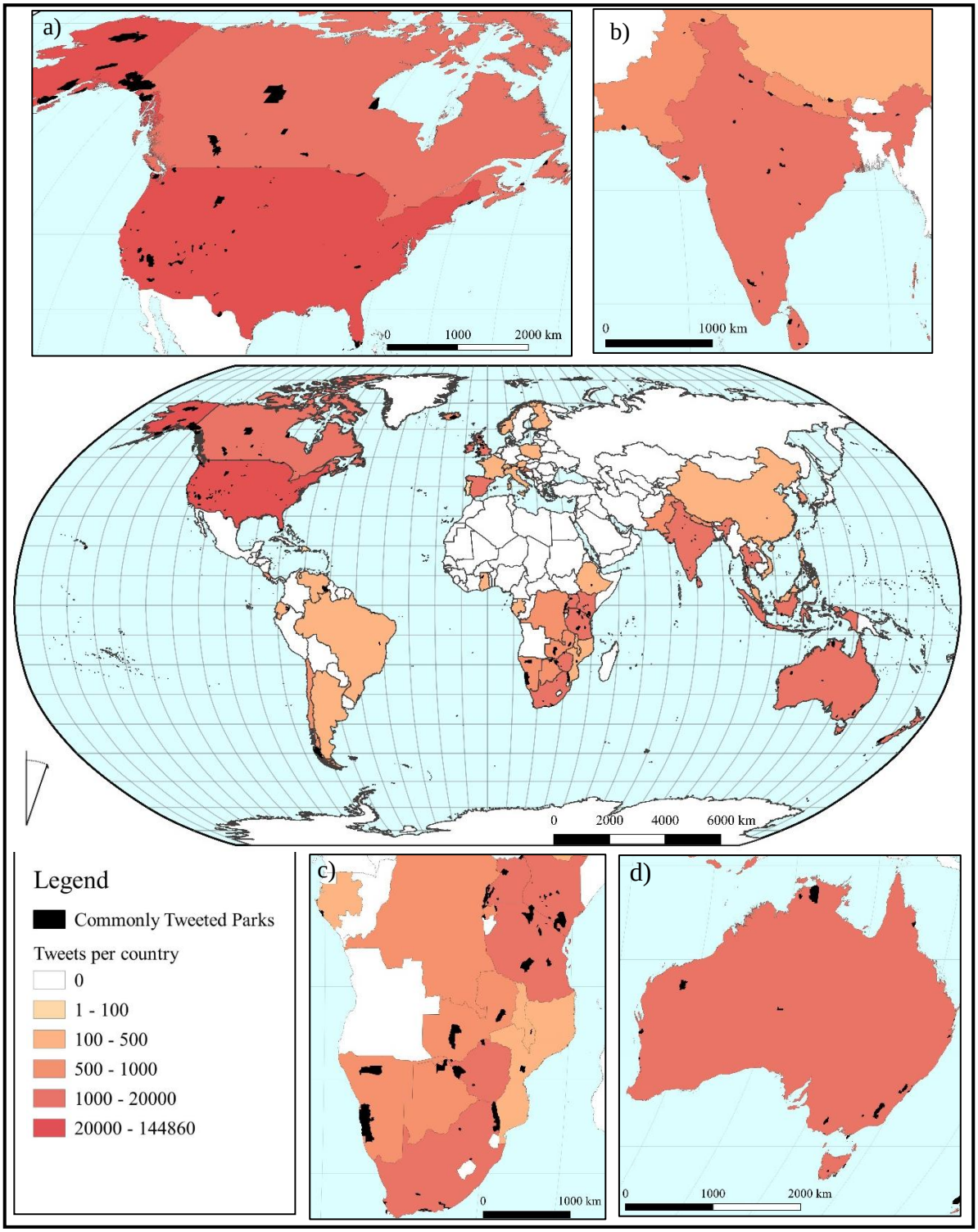

Figure 1. The number of tweets per country based on data for the 264 national parks with over 100 tweets. Inset maps show regions with many parks discussed on Twitter including: a) North America, b) South Asia, c) Southern Africa and d) Australia. The USA had the largest number of tweets about national parks (144 860 tweets).

Who talked about national parks and where were they from?

Of the 231 653 original tweets covering 264 parks, 150 932 contained information about where tweeters were from that could be geocoded to a country. Based on this information, tweets were sent by people from 187 countries but predominantly from the USA (76 754 tweets), United Kingdom (14 901) and Canada (11 920) (Figure 2). In contrast, there were 51 countries with less than 10 tweeters, with an overall median of 41 tweeters per country.

When assessed at a continental scale differences were apparent in who tweeted about where. Nearly all tweets sent by North Americans (59.4% of total), were about North American parks (90.6%), highlighting the dominance of people and parks from North America in the tweets (Figure 2). African tweeters were also focused on their own parks, with 92.1% of the tweets sent by people from Africa concerning African parks. In contrast, tweets sent by people from Europe (27 190 tweets) were more likely to talk about North American parks (50.7%), followed by parks in Europe (28.9% of their tweets). Tweets sent by people from Asia were mainly either about parks in Asia (45.4%) or North America (41.0%), with few about parks in Africa (6.3%). The few tweets sent by people from South America were mainly about North American parks (61.1%), with only 1.7% about parks in South America.

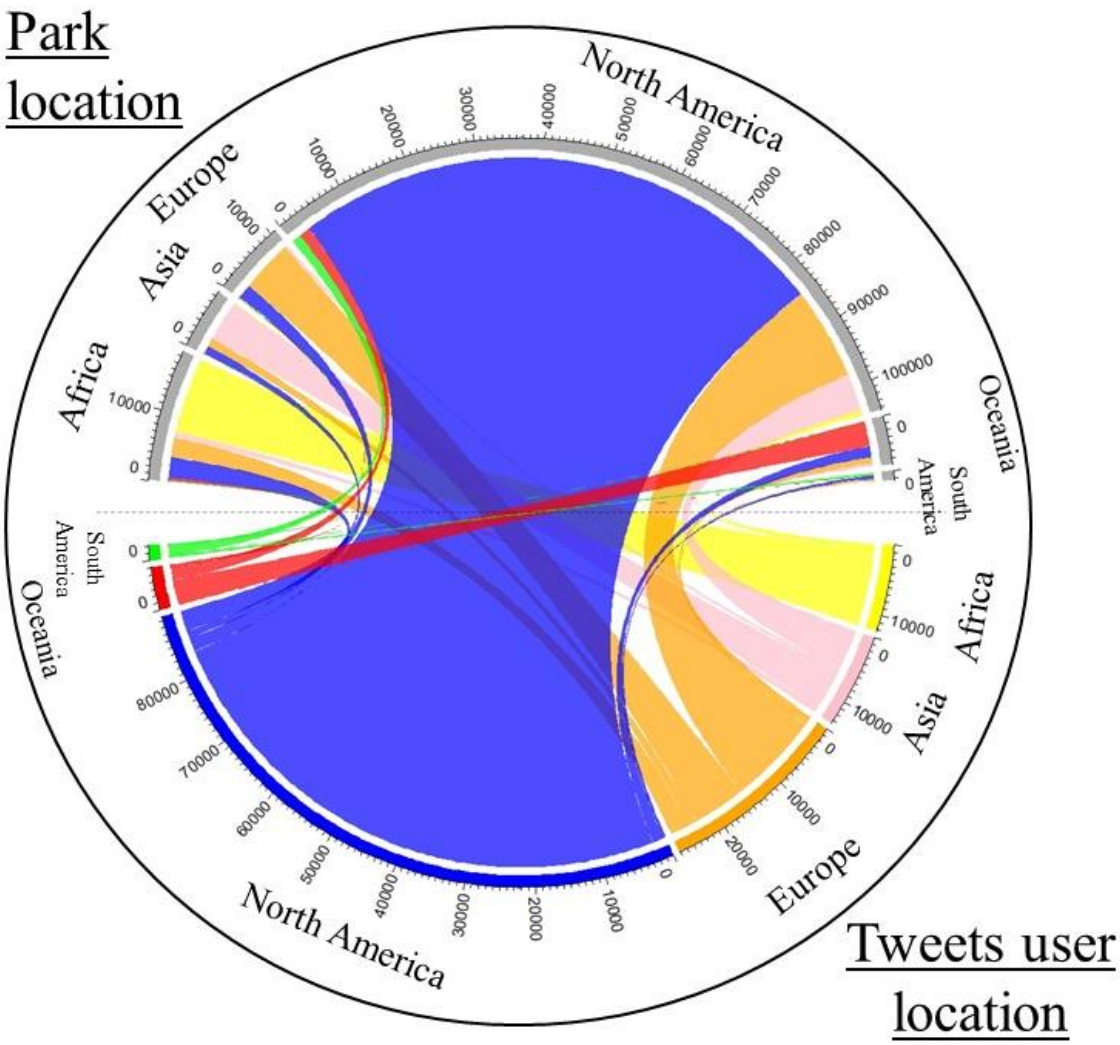

Figure 2: The continental distribution of tweets about national parks, showing where the tweeter said they were from (lower half) and the location of the park discussed within the tweet (upper half) for 150 932 tweets.

Overall people were more likely to tweet about local parks with 60.6% of all the tweets sent by people about parks in their own country. This varied among parks and countries, with tweets about Kanger Valley National Park in India (148 tweets) being entirely sent by Tweeters from India, while for Hohe Tauern National Park in Austria, only one of the 215 tweets, was sent by someone from Austria. For the top 40 parks,

Rocky Mountain National Park in the USA had the highest proportion of domestic tweets (88.7%), while Krka National Park in Croatia had the lowest number of domestic tweets (3.9%).

Among countries, there were also differences in the numbers of domestic and international tweets. Of the 55 countries with a commonly tweeted park, domestic tweeters posted 30.9% of the tweets, highlighting how there is often a strong international focus on some parks. This included seven from Africa and five from Asia. Gabonese parks were most likely to be discussed by locals (91.5%), as were those in Great Britain (80.7% locals) and Pakistan (79.5% locals). In contrast, parks in the Dominican Republic, Austria, Argentina and Switzerland where rarely tweeted about by locals with less 5% of the tweets sent by domestic tweeters. Although many (76 754) tweets about national parks were sent by tweeters from the USA, United Kingdom (14 901) and Canada (11 920), use of twitter was not correlated with the size of the discourse on parks among countries (Spearman's rho = 0.341, P = 0.233, Appendix 2).

Is park visitation or size related to salience?

The relative popularity of parks on ground was similar to their popularity on Twitter (Table 2, Figure 3). This can be seen in the significant correlations between the rank order of number of visits and tweets (Spearman's rho = 0.578), and tweeters (0.580) for the 40 most popular parks on Twitter, with even stronger relationships when data for tweeters from the same country as the park was used (0.594). The relationship was slightly tighter for the USA parks, both between number of visits and number of tweets (Spearman's rho = 0.863), and the number of tweeters (0.769). However, for

individual parks the relationship between visitor numbers and tweets and tweeters varied dramatically, limiting the capacity to use tweets as an actual estimate of visitation within a park (Table 2). For example, although there was an average of 0.77 tweets per thousand visits for the most talked about 40 parks on Twitter, values ranged from 0.02 tweets per thousand visits for North Cascades National Park in the USA, to 4.33 tweets per thousand visits for Snowdonia National Park in the United Kingdom (Figure 3, Appendix 1).

Table 2. Results of Spearman's rank correlation coefficients between number of visits to a park and number of tweets and tweeters for the top 40 Parks on Twitter globally (top triangle) and the top 27 parks from the USA on Twitter (bottom triangle), with significant correlations in bold. Where ϱ is the Spearman's rank correlation coefficient and P is the probability value.

	Park Size	#Visits	#Tweets	#Tweeters	#Domestic Tweeters
Park Size		ϱ = 0.060, P > 0.7	ϱ = 0.227, P > 0.1	ϱ = 0.205, P > 0.2	ϱ = 0.2, P > 0.2
#Visits	ϱ = - 0.060, P > 0.7		**ϱ = 0.578, P < 0.001**	**ϱ = 0.580, P < 0.001**	**ϱ = 0.594, P < 0.001**
#Tweets	ϱ = 0.099, P > 0.6	**ϱ = 0.863, P < 0.001**		**ϱ = 0.958, P < 0.001**	**ϱ = 0.884, P < 0.001**
#Tweeters	ϱ = 0.121, P > 0.5	**ϱ = 0.769, P < 0.001**	**ϱ = 0.962 P < 0.001**		**ϱ = 0.934, P < 0.001**
#Domestic Tweeters	ϱ = 0.052 P > 0.7	**ϱ = 0.784 P < 0.001**	**ϱ = 0.952 P < 0.001**	**ϱ = 0.981 P < 0.001**	

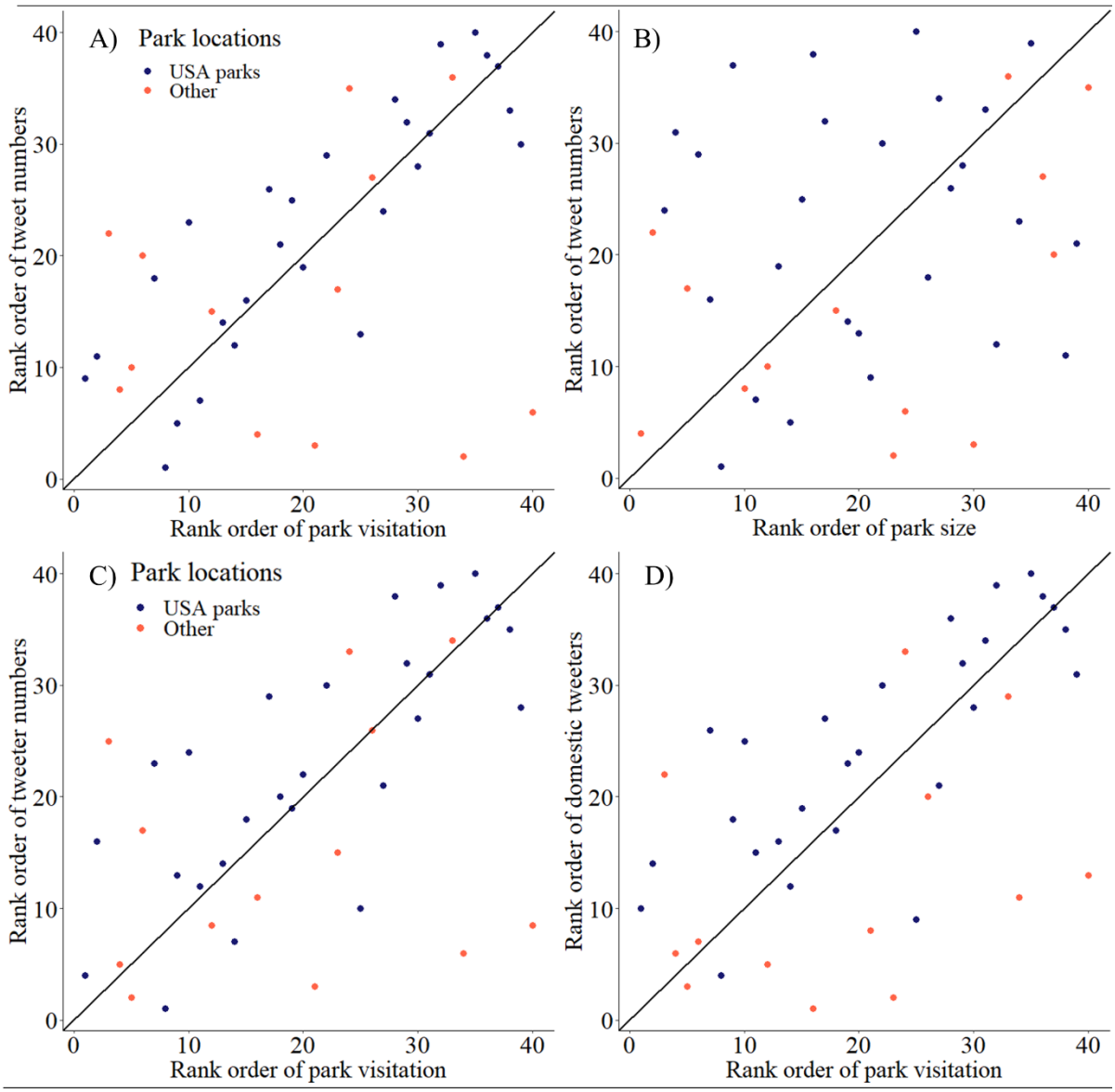

Figure 3: Correlations between the rank order of tweets and visitation numbers (A), or park size (B), as well as correlations between the number of tweeters (C) and domestic tweeters (D) against park visitation. Estimates of visits per park were sourced from the most recent available data. Due to the dominance of USA parks (27 of 40), these have been coloured separated (dark blue) from parks from other countries (light red).

Which platforms were originally used to post tweets and where were tweets sent from?

Reflecting people's increasing use of multiple social media platforms, original tweets about national parks on Twitter were posted from 3 054 different sources. The most common source of tweets were direct posts on Twitter itself or one of its derivative platforms (42.3%), although posts to Twitter from other popular social media platforms were also common, particularly from Instagram (12.4%) and Facebook (7.2%). Also, 12.3% of the tweets were from IFTTT, a platform that allows users to simultaneously post to several social media accounts, although it is unknown which social media platform was the original target.

Of the 437 789 original tweets, 36 921 (8.4%) contained geolocation data. Although this may appear to provide useful data about where people tweeted from, as opposed to where the tweeter was from, there appear to be major issues with the accuracy of the data. After screening the tweets, it was found that almost all (86.5%) were from Instagram with just 0.014% of the geolocated tweets posted directly onto Twitter. As a result, the geolocation data for most tweets relate to a very few locations within a park, and these are often the nominated places with locations created by users (Figure 4). For example, although there were 101 geolocated tweets for Tiede National Park, on the Canary Islands, Spain, they only refer to four locations within the Park, and three were the summit of the volcano or the nominated Tiede National Park default location.

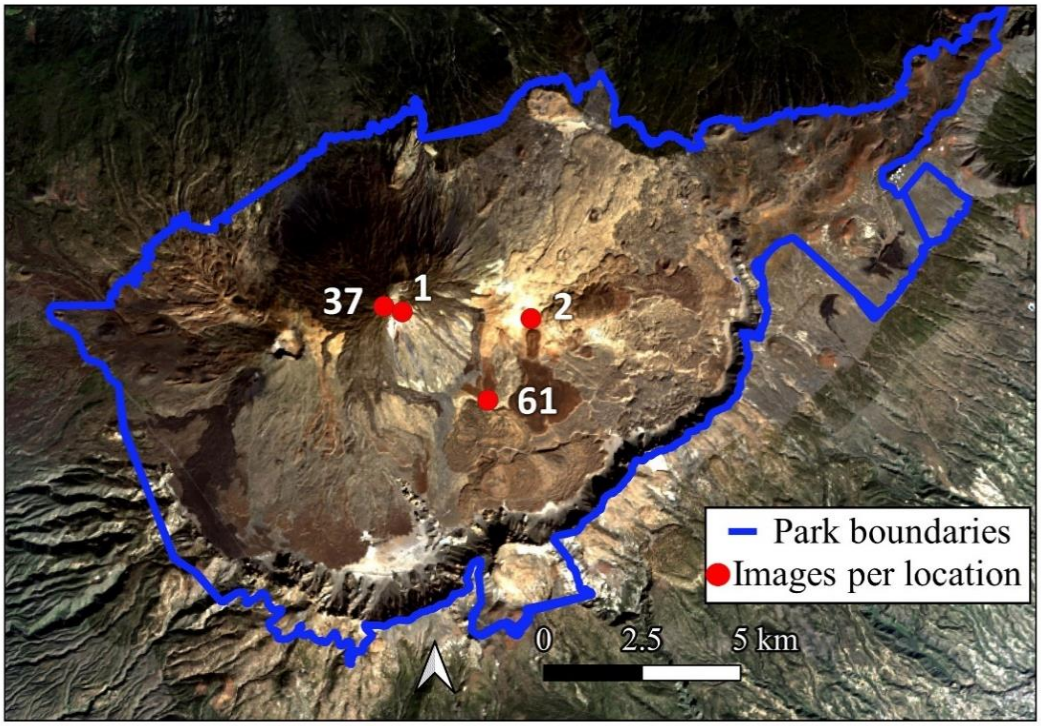

Figure 4: Geolocated tweets were almost entirely originally posted to Instagram, which resulted in stacked locations (red dots), seen here from 101 geolocated tweets from four locations within the Teide National Park boundary (blue outline) (Landsat 8 image sourced via Google Earth Engine, 2019).

Discussion

The scale and extent of the results indicate that Twitter can provide valuable information about the public discourse for a wide range of national parks globally and that large numbers of users of the platform talk about parks. The results also highlight that salience of tweets can reflect relative popularity among parks, but they no longer provide reliable estimates of either the number of visitors to specific parks or where people go within parks. These results add to the still limited research on how Twitter data can be used to assess public discourse on park management at a range of scales,

from a single trail network (Wilson et al., 2019), across urban and peri-urban parks (Hamstead et al., 2018) to entire park systems (Tenkanen et al., 2017).

The study found that there was a relationship between the relative popularity of a park on Twitter and actual visitation across 40 parks globally, and within USA parks, providing similar results to those found among parks in Finland and slightly stronger than that for South African parks where the relationship was weaker (Tenkanen et al., 2017). However, some parks were more popular relatively on Twitter than on ground with order of magnitudes differences in the ratio of tweets or tweeters to park visits. This is because people can send tweets about parks anywhere in the world, and discourse on Twitter, including about parks, is often driven by specific events in the news (Macdonald et al. 2016). Data from other social media sites including Flickr and Instagram appear to be a more reliable measure of relative and actual visitation and in the case of Flickr, can provide high resolution geolocation data both among and within parks (Heikinheimo et al., 2017; Walden-Schreider et al., 2018; Wood et al., 2016).

Studies globally have found that social media platforms vary in popularity across countries and regions (Sonter et al., 2016; Wood et al., 2013) and this was evident in tweets about national parks. Even in countries with very high use of social media, data from social media only represent a subsample of visitors to parks (Ghermandi and Sinclair, 2019; Tenkanen et al., 2017).). How well social media platforms represent actual visitation data depends on the specific park/parks and the platform (Tenkanen et al., 2017). Surveys of visitors' to parks use of social media found that 28% of respondents have, or were likely to, post their trip onto a social media platform,

including 7% to Twitter for visitors to Finnish National Parks Heikinheimo et al. (2017), where for Crater Lake National Park in the USA, it was found that 76% of visitors surveyed are active social media users with 28% using Twitter (Wilkins et al., 2018). The variability in the use of social media is an important consideration for researchers and managers when selecting social platforms to address specific questions.

Tweets about national parks were dominated by North Americans and they mostly discussed North American national parks. In part this reflects the dominance of Twitter by people from North America, predominantly from the USA, who are the largest single users of Twitter (Hootsuite and we are social, 2018). The dominance of North America in the discourse on Twitter about parks is also likely to reflect a strong interest in national parks in the region and that the term national park occurs in the title of many famous protected areas in the USA and Canada. The use of Twitter and other social media platforms is affected by a range of factors including motivations of the users, features of platforms, but also government regulation, language and access to the internet among others (Mao and Qian, 2015; Norman and Pickering, 2017; Stockmann and Luo, 2017; Woods et al. 2013). As a result of these and other factors, there was no correlation between tweets about national parks and popularity of Twitter per country.

Globally most tweets were sent by people from the same country as the park. Although the views of people from other countries can be important for the management of parks, including international tourists, agencies and governments are often more interested in the views of their own citizens, including local visitors to

parks, and for many parks, most visitors live close to the park (León et al., 2015; Muñoz et al., 2019; Rossi et al., 2015). Understanding locals' views is particularly beneficial when there are changes in management, for example in access, fees, infrastructure and regulations (Garbuzov et al., 2015; León et al., 2015). Twitter is also increasingly used as a rapid communication outlet to deal with unplanned events that affect parks, including managing fires, floods, hurricanes, rioting and political unrest (Garbuzov et al., 2015; León et al., 2015). As Twitter is used by billions of people to discuss important issues and opinions, it will be increasingly useful as a source of data to better understand international and domestic discourse on parks, complementing and expanding on other sources of data such as surveys (Acar and Muraki, 2011; Thelwall et al., 2011).

Twitter data can be used by researchers and managers to access data from other social media platforms such as Facebook and Instagram, with users posting the same content to multiple outlets. However, care must be taken as inaccuracies in data from one platform are also transferred over to others, such as geotagged tweets. Although 8.4% of tweets on national parks were geolocated, nearly all (86.5%) of these came from Instagram, and hence there was limited capacity to use the data to understand within park visitation patterns. On Instagram, users specify a location via names, then choose the place or nearest place from a list of pre-existing locations from the platform. Nominating a location often results in many geolocated tweets being associated with a few 'named' locations within a park. Prior to 2014, each Twitter post had associated co-ordinates as default although, after changing this to an op-in feature, it has become very unpopular (Leetaru, 2019). As a result, other social media platforms such as Flickr and fitness route sharing platforms can provide better insights into smaller scale

visitor movements within parks due to the higher accuracy of their spatial data, even though they are less popular than Twitter (Ghermandi and Sinclair, 2019; Walden-Schreider et al., 2018; Norman and Pickering, 2017).

There are important logistical issues when using Twitter as a data source for research and management. An important consideration is that Twitter imposes time limits on data via its automated programming interface (API). The standard API allows free access for up to a maximum of 18 000 tweets from the last nine days. Accessing older tweets requires use of Twitter's Premium API giving access to 30 days of data but it is limited to 100 tweets per request (Twitter, 2019). As a result of this limited capacity to easily access past tweets combined with the stochastic nature of tweets, with sudden but ephemeral engagement topics on Twitter, it may not be possible to use Twitter to address some issues (Ghermandi and Sinclair, 2019). Similarly, there are challenges with terminology when searching for tweets. These are particularly problematic when there is no single term or sets of terms that can be used to find all the required tweets on a specific topic. We found that 'national park' provided large numbers of tweets, but it excluded the discourse on many popular and important protected areas including World Heritage Areas, and Conservation Areas and others where national parks did not form part of the name. It also excluded much of the discourse in other languages.

There are well recognised biases in who posts to social media, with certain demographics heavily influencing results (Hargittai, 2018; Sorokowska et al., 2016; Tenkanen et al., 2017). Twitter is no exception to this. However, as it is one of the most common platforms globally, it can be a useful source of data, provided the

biases in user groups are understood and accounted for. These limitations apply to all social media data, requiring results to be interpreted in according to patterns in who uses platforms and why (Ghermandi and Sinclair, 2019). Such considerations also arise with other sources of data used to assess public discourse on parks, where certain views and people can disproportionately dominate responses, including in surveys and stakeholder groups (Armstrong, and Overton, 1977; D'Antonio and Monz, 2016; East, 2017; Veal, 2017, Worboys et al. 2015).

Conclusion

The public discourse about national parks globally on Twitter is massive, both in terms of number of tweets and number of tweeters. The number of tweets was related to the relative popularity of parks, but a poor predictor of actual visitation with the ratio of tweets to visitors varying by orders of magnitude among parks. The size of the discourses on parks varied amongst countries with many of those tweeting in the USA and many of the tweets about USA parks. People were also more likely to tweet about parks in their own countries, but those in Europe talk a lot about USA parks. More detailed analysis of the content and sentiments associated with tweets about national parks including for particular parks, regions and topics are likely to provide important insights into the emotions, issues and opinions of people about these important natural areas, but there are limitations with this source of data, several of which were highlighted here.

References

Acar, A., and Y, Muraki. 2011. Twitter for crisis communication: lessons learned from Japan's tsunami disaster. *International Journal of Web Based Communities*, *7*(3), 392-402.

Armstrong, J.S., and T.S. Overton. 1977. Estimating nonresponse bias in mail surveys. *Journal of Marketing Research*, *14*(3), 396-402.

Association of National Park Authorities. 2011. Estimated Visitor Figures/Spend in UK National Parks. Cardiff, United Kingdom. file:///C:/Users/s2679022/Downloads/445%20Estimated%20Visitor%20Spend%20figures%20for%20UK%20National%20Parks%20(May%2011).pdf

Balmford, A., J.M. Green, M. Anderson, J. Beresford, C. Huang, R. Naidoo, M. Walpole and A. Manica. 2015. Walk on the wild side: Estimating the global magnitude of visits to protected areas. *PLoS biology*, *13*(2), p.e1002074.

Brett, M.R. 2018. Tourism in the Kruger National Park: Past development, present determinants and future constraints. *African Journal of Hospitality, Tourism and Leisure*, *7*(2).

Burns, G. L., J. Macbeth. and S. Moore. 2011. Should dingoes die? Principles for engaging ecocentric ethics in wildlife tourism management. *Journal of Ecotourism*, *10*(3), 179-196.

Calcagni, F., A.T.A. Maia, J.J.T. Connolly and J. Langemeyer. 2019. Digital co-construction of relational values: Understanding the role of social media for sustainability. *Sustainability Science.* https://doi.org/10.1007/s11625-019-00672-1

Canadian Government. 2019, Parks Canada Attendance 2017-18. Retrieved 14 March, 2019, from https://www.pc.gc.ca/en/docs/pc/attend/table3

Chamberlain, E.C., M.B. Rutherford and M.L. Gibeau. 2012. Human perspectives and conservation of grizzly bears in Banff National Park, Canada. *Conservation Biology*, *26*(3), 420-431.

Croatian Ministry of Tourism. 2017. Tourism in Figures 2017. Zagreb, Croatia. https://mint.gov.hr/UserDocsImages/AA_2018_c-dokumenti/180608_HTZTUBENG_2017.PDF

Daley, J. 2019. Indonesia Considers Closing Komodo Island Because Poachers Keep Stealing the Dragons. Retrieved 31 July, 2019, from https://www.smithsonianmag.com/smart-news/there-be-no-dragons-indonesia-considers-closing-komodo-island-tourists-180971877/

D'Antonio, A., and C. Monz. 2016. The influence of visitor use levels on visitor spatial behavior in off-trail areas of dispersed recreation use. *Journal of Environmental Management*, *170*, 79-87.

Eagles, P.F., S.F. McCool, C.D Haynes and A. Phillips. 2002. Sustainable Tourism in Protected Areas: Guidelines for Planning and Management IUCN, Cambridge: IUCN Publications.

East, D., P. Osborne, S Kemp and T. Woodfine. 2017. Combining GPS and survey data improves understanding of visitor behaviour. *Tourism Management*, *61*, 307-320.

Espenshade, J.L., J.D. Murdoch, T.M. Donovan, R.E. Manning, C.A. Bettigole and J. Austin. 2018. Public acceptability of development in the Northern Forest of Vermont, USA -The influence of wildlife information, recreation involvement, and demographic characteristics. *PloS One*, *13*(12), e0203515.

Hargittai, E. 2018. Potential biases in big data: Omitted voices on social media. *Social Science Computer Review*, 0894439318788322.

Garbuzov, M., K.A. Fensome and F.L. Ratnieks. 2015. Public approval plus more wildlife: twin benefits of reduced mowing of amenity grass in a suburban public park in Saltdean, UK. *Insect Conservation and Diversity*, *8*(2), 107-119.

Ghermandi, A., and M. Sinclair. 2019. Passive crowdsourcing of social media in environmental research: A systematic map. *Global Environmental Change*, *55*, 36-47.

Google. 2018. About Google docs. Retrieved 15 June, 2019, from https://www.google.com/docs/about/

Google Earth Engine. 2019. Google Earth Engine: A planetary-scale platform for Earth science data & analysis. Retrieved 29 July, 2019, from https://earthengine.google.com/

Harbrow, M. 2013. Visitor Trends Report. Wellington, New Zealand. https://www.doc.govt.nz/Documents/about-doc/role/visitor-research/twt-visitor-trends-report.pdf

Hamstead, Z.A., D. Fisher, R.T. Ilieva, S.A. Wood, T. McPhearson and P. Kremer. 2018. Geolocated social media as a rapid indicator of park visitation and equitable park access. *Computers, Environment and Urban Systems*, *72*, 38-50.

Hawksey, M. 2019. #TAGS. Retrieved 15 June, 2019, from https://tags.hawksey.info/

Heikinheimo, V., E.D. Minin, H. Tenkanen, A. Hausmann, J. Erkkonen and T. Toivonen. 2017. User-generated geographic information for visitor monitoring in a national park: a comparison of social media data and visitor survey. *ISPRS International Journal of Geo-Information*, *6*(3), 85.

Hootsuite and we are social. 2018. Global Digital Statshot Q2, Retrieved 26 September, 2019, from https://datareportal.com/reports/digital-2018-q2-global-digital-statshot

Kenyan Bureau of Statistics. 2018. Economic Survey 2018. Nairobi, Kenya. file:///C:/Users/s2679022/Downloads/Economic%20Survey%202018.pdf

Kim, S.K., M.J. Park and J.J. Rho. 2015. Effect of the government's use of social media on the reliability of the government: Focus on Twitter. *Public Management Review*, *17*(3), 328-355.

Ladle, R.J., R.A. Correia, Y. Do, G-J. Joo, C.C.M. Malhado, R. Proulx, J-M. Roberge and P. Jepson. 2016. Conservation culturomics. *Frontiers in Ecology and Environment*, *14*, 269-275.

Leetaru, K. 2019. Forbes: Visualizing Seven Years Of Twitter's Evolution: 2012-2018. Retrieved 21 October, 2019, from https://www.forbes.com/sites/kalevleetaru/2019/03/04/visualizing-seven-years-of-twitters-evolution-2012-2018/#66c249b47ccf

León, C.J., J. de León, J.E. Araña and M.M. González. 2015. Tourists' preferences for congestion, residents' welfare and the ecosystems in a national park. *Ecological Economics*, *118*, 21-29.

McCormick, T.H., H. Lee, N. Cesare, A. Shojaie and E.S. Spiro. 2017. Using Twitter for demographic and social science research: Tools for data collection and processing. *Sociological Methods and Research*, *46*(3), 390-421. Doi: 10.1177/0049124115605339

Mao, Y.Y., and Y.Y. Qian. 2015. Facebook use and acculturation: The case of overseas Chinese professionals in Western countries. *International Journal of Communication*, *9*(1), 2467-2486.

Macdonald, D.W., K.S. Jacobsen, D. Burnham, P.J. Johnson, and A.J. Loveridge, 2016. Cecil: a moment or a movement? Analysis of media coverage of the death of a lion, Panthera leo. *Animals, 6.* doi:10.3390/ani6050026

Marijnen, E., and J. Verweijen. 2016. Selling green militarization: the discursive (re) production of militarized conservation in the Virunga National Park, Democratic Republic of the Congo. *Geoforum*, *75*, 274-285.

Mowery, D., H. Smith, T. Cheney, G. Stoddard, G. Coppersmith, C. Bryan and M. Conway. 2017. Understanding depressive symptoms and psychosocial stressors on Twitter: a corpus-based study. *Journal of Medical Internet Research*, *19*(2), e48.

Muñoz, L., V. Hausner, G. Brown, C. Runge and P. Fauchald. 2019. Identifying spatial overlap in the values of locals, domestic-and international tourists to protected areas. *Tourism Management*, *71*, 259-271.

Nara, A., M.H. Tsou, J.A. Yang and C.C. Huang. 2018. The opportunities and challenges with social media and big data for research in human dynamics. In *Human Dynamics Research in Smart and Connected Communities* (pp. 223-234). New York City: Springer.

National Park Service. 2019. National Park Service Visitor Use Statistics. Retrieved 14 March, 2019, from https://irma.nps.gov/Stats/

Newsome, D., S.A. Moore and R.K. Dowling. 2012. Natural area tourism: Ecology, impacts and management. Bristol: Channel View Publications.

Norman, P., and C.M. Pickering. 2017. Using volunteered geographic information to assess park visitation: Comparing three on-line platforms. *Applied Geography*, *89*, 163-172.

Norman, P., C. Pickering and J.G Castley. 2019. What can volunteered geographic information tell us about the different ways mountain bikers, runners and walkers use urban reserves? *Landscape and Urban Planning*, *185*, 180-190. Doi: https://doi.org/10.1016/j.landurbplan.2019.02.015

OpenStreetMap. 2019. OpenStreetMap. Retrieved 29 July, 2019, from https://www.openstreetmap.org/

Orellana-Rodriguez, C., and M.T. Keane. 2018. Attention to news and its dissemination on Twitter: A survey. *Computer Science Review*, *29*, 74-94.

Pebesma, E.J., and R.S. Bivand., 2005. Classes and methods for spatial data in R. R News 5 (2), https://cran.r-project.org/doc/Rnews/.

Pettebone, D., B. Meldrum, C. Leslie, S.R. Lawson, P. Newman, N. Reigner and A. Gibson. 2013. A visitor use monitoring approach on the Half Dome cables to reduce crowding and inform park planning decisions in Yosemite National Park. *Landscape and Urban Planning*, *118*, 1-9.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing. http://www.R-project.org/.

RStudio Team. 2019. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

Rossi, S.D., J.A. Byrne, C.M. Pickering and J. Reser. 2015. 'Seeing red' in national parks: How visitors' values affect perceptions and park experiences. *Geoforum*, *66*, 41-52.

Silge, J., and D. Robinson. 2016. tidytext: Text Mining and Analysis Using Tidy Data Principles in R. *JOSS*, *1*(3). Doi: 10.21105/joss.00037

Sherren, K., M. Smit, M. Holmlund, J.R. Parkins and Y. Chen. 2017. Conservation culturomics should include images and a wider range of scholars. *Frontiers in Ecology and Environment. 14*, 289-290.

Sonter, L. J., K.B. Watson, S.A. Wood and T.H. Ricketts. 2016. Spatial and temporal dynamics and value of nature-based recreation, estimated via social media. *PLoS One*, *11*(9), e0162372.

Sorokowska, A., A. Oleszkiewicz, T. Frackowiak, K. Pisanski, A. Chmiel and P. Sorokowski. 2016. Selfies and personality: Who posts self-portrait photographs? *Personality and Individual Differences*, *90*, 119-123.

Stockmann, D., and T. Luo. 2017. Which social media facilitate online public opinion in China? *Problems of Post-communism*, *64*(3-4), 189-202.

United Nations Environment World Conservation Monitoring Centre. 2019. Protected Planet. Retrieved 11 March, 2019, from https://www.protectedplanet.net/

Tenkanen, H., E. Di Minin, V. Heikinheimo, A. Hausmann, M. Herbst, L. Kajala and T. Toivonen. 2017. Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas. *Scientific Reports*, *7*(1), 17615

Tennekes, M. 2019. tmaptools: Thematic Map Tools. R package version 2.0-2. https://cran.r-project.org/web/packages/tmaptools/index.html

Thelwall, M., K. Buckley and G. Paltoglou. 2011. Sentiment in Twitter events. *Journal of the American Society for Information Science and Technology*, *62*(2), 406-418.

Toivonen, T., V. Keikinheimo, C. Fink, A. Hausmann, T. Hippala, O. Jarv, H. Tenkanen and E. Di Minin. 2019. Social media data for conservation science: A methodological overview. *Biological Conservation*, *233*, 298-315.

Tumasjan, A., T.O. Sprenger, P.G. Sandner and I.M. Welpe. 2011. Election forecasts with Twitter: How 140 characters reflect the political landscape. *Social Science Computer Review*, *29*(4), 402-418.

Twitter. 2019. Getting Started. Retrieved 29 July, 2019, from https://help.twitter.com/en/twitter-guide

Twitter Business. 2019. Be what's happening. Retrieved 30 September, 2019, from https://business.twitter.com/

Veal, A.J. 2017. *Research methods for leisure and tourism.* London: Pearson.

Walden-Schreiner, C., S.D. Rossi, A. Barros, C. Pickering and Y.F. Leung. 2018. Using crowd-sourced photos to assess seasonal patterns of visitor use in mountain-protected areas. *Ambio*, *47*(7), 781-793.

Wilkins, E.J., J.W. Smith and R. Keane. 2018. Social media communication preferences of national park visitors. *Applied Environmental Education and Communication, 15*(1).

Wilson, T., R. Lovelace and A.J. Evans. 2019. A path toward the use of trail users' tweets to assess effectiveness of the environmental stewardship scheme: An Exploratory analysis of the Pennine Way National Trail. *Applied Spatial Analysis and Policy*, *12*(1), 71-99.

Wood, S.A., A. Guerry, J. Silver and M. Lacayo. 2013. Using social media to quantify nature-based tourism and recreation. *Scientific Reports,3*, 2976.

Worboys, G., M. Lockwood, A. Kothari, S. Feary and I. Pulsford. 2012. Protected area governance and management. Canberra: ANU Press.

World Bank Group. 2015. Tanzania's Tourism Futures. Washington, USA. http://documents.worldbank.org/curated/en/204341467992501917/pdf/96150-REVISED-PN-P150523-PUBLIC-Box393206B.pdf

Chapter 6: Discussion

6.1. Introduction

Understanding where, why and who visits parks as well as the size of public discourse about national parks is important for their effective management (Eagles et al., 2002; Worboys et al., 2015). It can assist managers in allocating resources, reducing environmental impacts, limiting conflict among visitors and assessing public concerns about parks (Newsome et al., 2012). In this thesis spatial social media data was used to evaluate public interactions with parks, including; how social media route data can be used to assess patterns of visitation and activities within national parks, to identify which factors influence park visitation and determine which parks are discussed online and by whom. Using spatial social media data, aspects of visitation were also assessed including differences in the temporal and spatial use of parks by walkers, runners and mountain bikers.

All results chapters are either published (Chapters 2, 3 and 4) or prepared for publication (Chapter 5) and include their own abstracts, introductions, methods, results, discussions and reference lists. This final chapter discusses the results across the whole thesis in relation to the broader literature, including how the research in this thesis advances the discipline, explores the potential of social media as an emerging data source and identifies future research directions. Within each of the results chapters are detailed discussions of the implications of the research, while this final discussion chapter takes a broader scale assessment of the significance and limitations of the research.

The results of the thesis highlight that there is a large volume of user generated spatial data available from some social media platforms that is relevant to park management including GPS route data. Short-term (daily, weekly, monthly and seasonal) temporal and spatial patterns of use of trail networks were identified from route data, although due to the ephemeral popularity of fitness platforms, the data is not suitable for assessing longer term trends (among years). The factors influencing the popularity of parks differed among user groups and activities, particularly between fitness trail based and adventure-based park visitors. Social media data was found to reflect on-ground use of trails as well as popularity among parks, with route data found to accurately represent trail and park usage. Text data from Twitter was also found to be valuable for assessing national park discourse at a global scale including identifying which parks were discussed and by people from which countries. Social media data was also able to provide answers for questions traditionally difficult to obtained data about, such assessing factors that influence visitation rates to parks for different activities, as well as differences in patterns of use among activities and to assess off trail use. The following sections summarise the key contributions of the thesis to knowledge about visitor monitoring using social media data as well as future research directions.

6.2. Aim 1: What are the benefits and limitations of different social media platforms as sources of GPS route data for visitor monitoring in and among parks (Chapters 3 and 4)?

Social media platforms where GPS route data are posted differ greatly in the amount and types of data available for research (Figure 6.1.) (Norman and Pickering, 2017; Norman and Pickering, 2019). In the thesis four different GPS route sharing platforms were assessed to see if they could provide useful data for national parks. The most

popular platform was Strava, with over 430,000 routes posted across 40 parks (Chapter 4). However, access to detailed route data from the platform was restricted in 2018 after massive controversy about unintentional use of the data (McKenna et al., 2019). At that time, a global heatmap of GPS route data highlighted the location and layout of classified USA army bases within conflict zones including in Somalia and Iraq (McKenna et al., 2019). This illustrates an ongoing issue with social media data, where it has been used for unintended purposes, with important privacy concerns. This has resulted in rapid changes in access to data affecting the suitability of specific platforms for park monitoring (Norman and Pickering, 2019).

Another important issue with social media data highlighted in the thesis was that social media platforms vary in popularity among countries and regions (Norman and Pickering, 2017). This was evident in Chapter 3 where the platform GPSies contained very few routes for the Australian parks being assessed. In contrast, Campelo and Mendes (2016) and Santos et al. (2016) both found the platform useful when assessing parks in Portugal with over 400 routes within the parks studied. Therefore, it is important when social media data is used for monitoring visitation to ensure there is both access and adequate amounts of data.

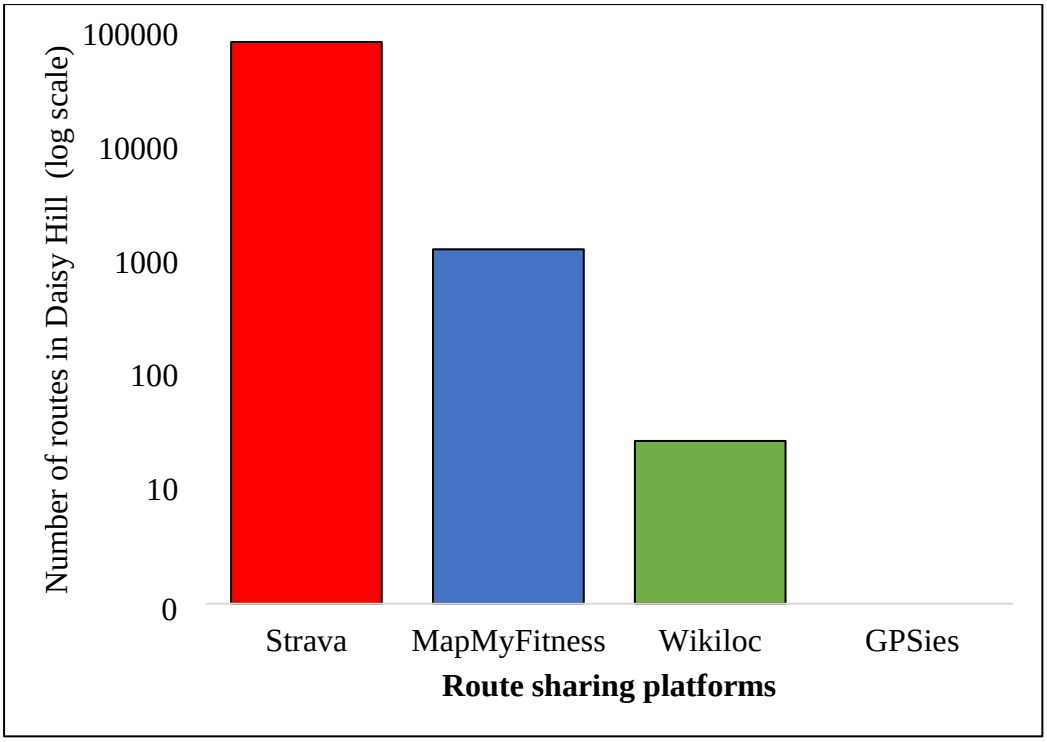

Figure 6.1. Variation in the number of GPS routes for a single park, Daisy Hill Conservation Park in Brisbane, Australia posted to different GPS route sharing platforms based on results from Chapters 2, 3 and 4 (Norman et al., 2019; Norman and Pickering, 2019; Norman and Pickering, 2017).

There were considerable differences in the types of data available from GPS route sharing platforms (Chapters 2, 3 and 4). For example, on Strava the data available when the research in the thesis was undertaken consisted of just the total number of passes per segment of trail. In contrast, data from GPSies, MapMyFitness and Wikiloc consisted of downloadable routes in the form of GPX files, which included user identifiers and timestamps of when the posts were made (GPSies, 2019a; MapMyFitness, 2019; Strava, 2019; Wikiloc, 2019). The scale and type of data available from different fitness platforms affects the types of analyses that can be performed and hence the types of questions that can be addressed (Norman and Pickering, 2019). When entire GPX files can be obtained, a wide range of questions

can be addressed by analysing patterns of trail use, including to determine route lengths and shapes as well as hot and cold spots of trail use (Jurado Rota et al., 2019). When platforms provide limited data such as Strava, researchers and managers cannot undertake detailed spatial analysis within parks (Norman and Pickering, 2019). However, due to the enormous popularity of Strava as seen in Chapter 4, this platform can still provide valuable data when comparing popularity among parks.

As already highlighted above, access to GPS route data, like other data varies among social media platforms (Table 6.1.) (Ghermandi and Sinclair, 2019; Norman and Pickering, 2019). Even when data is available to researchers, there are differences in how quickly data can be gathered or downloaded. When social media platforms provide an API for data collection, gathering data can be relatively easy and fast (Ghermandi and Sinclair, 2019; Toivonen et al., 2019). For instance, the platform GPSies provided access to GPS routes and metadata via an accessible API (GPSies, 2019b). It allowed GPX files to be downloaded on mass, providing researchers with a rapid tool to access large amounts of data (Campelo and Mendes, 2016). However, for some other platforms, APIs have not been created, and as a result, researchers must visit the website directly to gather data (Norman et al., 2019; Norman and Pickering, 2017). This was necessary for collecting data from Wikiloc and MapMyFitness used in Chapters 2, 3 and 4. Collecting every GPS route and subsequent metadata requires manual downloading each route, which was less efficient than using an API, potentially limiting the use of these platforms in future research, particularly for popular parks. Another option would be to purchase social media GPS route data directly from platforms.

Table 6.1. Some of the social media platforms used to assess visitation and discourse about national parks (Ghermandi and Sinclair, 2019; Toivonen et al., 2019); including the types of data available and if it was used in the thesis. Platforms with an automated programming interface (API) with public posts accessible for research is indicated.

Platform type	Name	Spatial data		Other types of data				Included in the thesis	Accessible API for public posts
		Route	Point	Temporal	Text	Image	Video		
Mainly route data	**MapMyFitness**	Yes	No	Yes	No	No	No	Yes	No
	Wikiloc	Yes	No	Yes	Yes	Yes	No	Yes	No
	Strava	Limited	No	No	No	Yes	No	Yes	Limited
	GPSies	Yes	No	Yes	No	No	No	Yes	Yes
Mainly image/video	**Instagram**	No	Limited	Yes	Yes	Yes	Yes	No	No
	Flickr	No	Yes	Yes	Yes	Yes	Yes	No	Yes
	Panoramio[!]	No	Yes	Yes	Yes	Yes	No	No	Closed
	Youtube	No	Limited	Yes	Yes	No	Yes	No	Yes
Multimedia and text	**Twitter**	No	Limited	Yes	Yes	Yes	Yes	Yes	Yes
	Facebook	No	Limited	Yes	Yes	Yes	Yes	No	No

[!] Closed platform

Many social media platforms once they become popular, monetize access to user generated data (Strava Metro, 2019). For example, Strava currently provides detailed spatial data through their Strava Metro program targeting local and regional governments, which dramatically increases the costs of obtaining detailed route data (Heesch et al., 2016; Strava Metro, 2019; Sun et al., 2017). Due to the popularity of the platform in some regions, purchasing data from Strava may be cost effective for some researchers and park agencies, but for many others, the costs will be prohibitive (Heesch et a l., 2016).

The highly ephemeral nature of social media including both the popularity of platforms and access to data are important considerations (Zhang and Gupta, 2018). In recent years, controversies have affected social media platforms, including the Cambridge Analytica scandal effecting Facebook (Schneble et al., 2018). Privacy concerns and issues arising from the misuse of personal data including from social media platforms resulted in the European Union initiating stronger privacy regulations, including the General Data Protection Regulation (GDPR) (European Commission, 2019). These regulations provide users of social media with greater control over their content including its use by third parties (Tikkinen-Piri et al., 2018). Since its implementation there has been large scale changes in the type and amount of information available from social media platforms including for research. A key recommendation of the regulations was that researchers, among others, have to ensure that people's data is carefully stored and not shared, and that API restrictions imposed by each platform are adhered to (Tikkinen-Piri et al., 2018). Although this will help prevent third party misuse of the data, it is likely that further restrictions will be introduced in the future. In any case researchers and managers should be cautious

when obtaining, storing and analysing social media data, particularly if it contains information which could be traced back to individuals.

Aim 2: How can GPS route data from social media platforms be used to assess spatial and temporal patterns of visitation within parks for walking, running and mountain biking (Chapters 2 and 3)?

Social media GPS route data is a rich source of information about visitation as it provides highly detailed spatial information about national parks (Chapters 2 and 3 Norman et al, 2019; Norman and Pickering, 2017). The greater detail about the movements of visitors within parks available from some social media platforms contrasts with that sourced using stationary monitoring techniques on ground such as trail counters (Byczek et al., 2018; Wolf et al., 2012). With detailed route data, it is possible to understand not only visitor use of specific locations, but also trail choices, directions, speeds, flows and directional patterns as well as map hot and cold spots of use of trail network (Byczek et al., 2018; Campelo and Mendes, 2016; Kidd et al., 2015; Santos et al., 2016; Wolf et al., 2015). Social media GPS route data also has many benefits compared to other sources of Volunteered Geographic Information, such as geolocated images for assessing visitor monitoring, as geolocated images only provide snapshots of visitor use rather than complete patterns of movement (Wei et al., 2015) (Figure 6.2.).

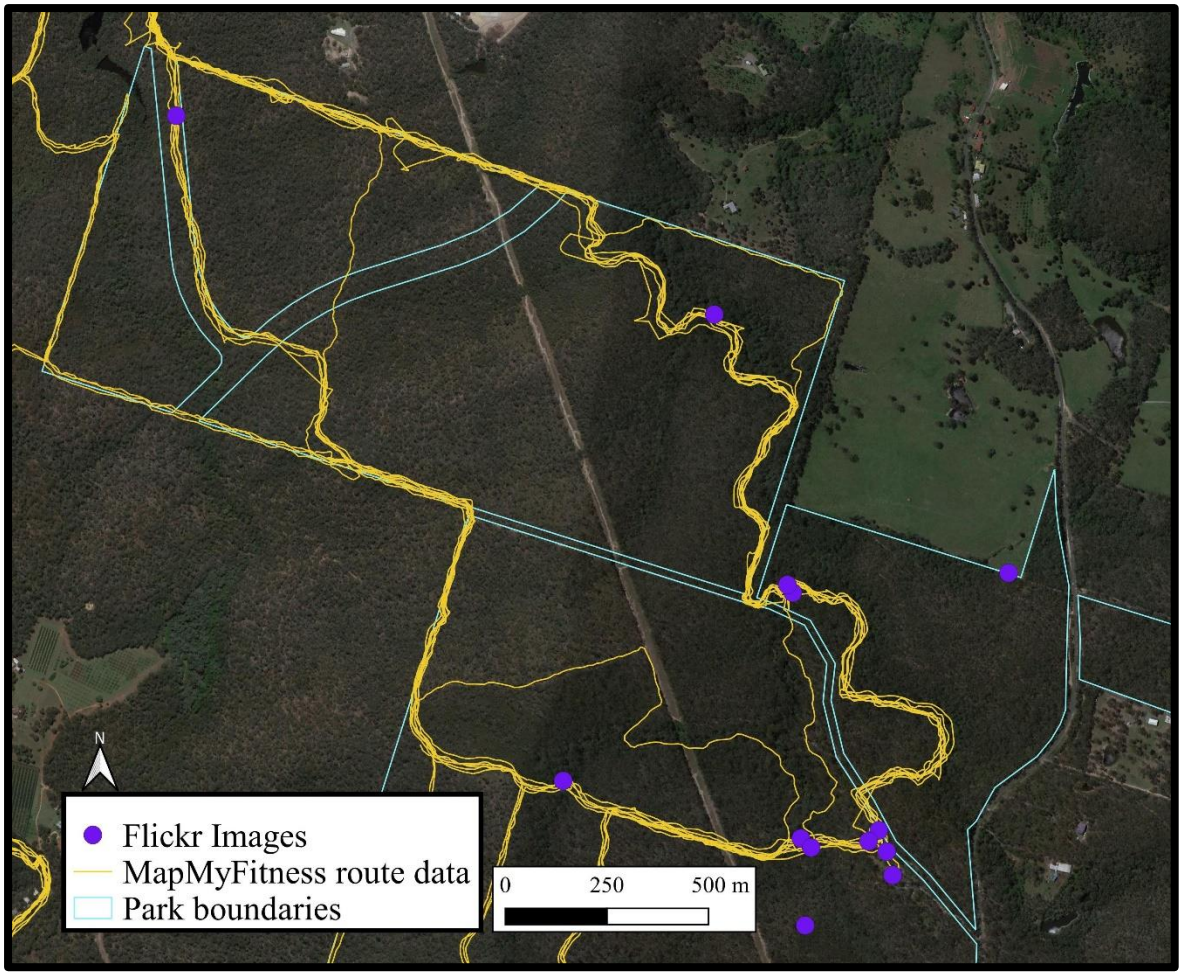

Figure 6.2. A map of all geolocated Flickr images (N=14) and the equivalent number of MapMyFitness routes for walking and running (N=14) within Venman National Park, Brisbane, Australia, illustrating how GPS route data can provide greater detail about visitation compared to point data from social media. Flickr data was sourced using the platforms API (Flickr, 2019).

One aspect of park visitation that has been difficult to assess using traditional survey methods is the extent, location and nature of visitation beyond formal trails (Norman and Pickering, 2017). Social media GPS route data provided important data about off trail use in parks, including where people go and for what activities, with data often relatively easy to obtain (Norman et al., 2019). In Chapters 2 and 3 there was a high percentage of visitors sharing data about off trail use of parks, revealing areas of high

conservation value that were being damaged. In some situation visitors seem to be unaware that they were entering restricted areas, including locations that are sensitive for environmental and cultural reasons, as they went on to posted routes online for others to follow (Norman and Pickering, 2017). Several studies have assessed off trail use of parks either through direct observations (Walden-Schreiner and Leung, 2013) and/or through mapping the resulting damage including the formation of unauthorised trails (Barros et al., 2013; Newsome and Davies, 2009; Walden-Schreiner and Leung, 2013). Additional methods to monitor off trail use of parks is beneficial, particularly to detect such use before environmental damage becomes so severe it results in the loss of vegetation and the formation of new trails.

Social media platforms are often ephemeral in their popularity, with data such as the number of routes posted to platforms changing over time (Norman and Pickering, 2017). An example of this was the popularity of MapMyFitness which was used as a source of data in Chapters 2 and 3. Although data over multiple years was available from the platform, the popularity of the platform declined in the study region, and hence it provided more of a historical assessment of visitation rather than information on current visitation to parks. The change in the popularity of the platform may have been due to range of factors, including the rise of others such as Strava (Norman et al., 2019). Due to the ebbs and flows in the use of different social media platforms, they are not useful for longer term monitoring of parks (Norman and Pickering, 2017). Other popular social media platforms that were used for monitoring park visitation in the past, have also declined rapidly, with the highly popular image sharing platform, Panoramio closing entirely (Panoramio, 2019).

Social media GPS route data was useful for assessing temporal differences in park visitation among recreation activities, such as hiking, running and mountain biking. This type of temporal data is highly valuable for park management. In Chapter 2, differences in how the parks were used between weekends and weekdays was very important with runners visiting the parks most days of the week, while mountain bikers were more likely to visit on weekends, and walkers on Sundays (Norman et al., 2019). These temporal trends have important implications when allocating staff and ensuring the adequacy of other resources (bins, toilets etc), as well as minimising conflict among users at peak times of visitation (Eagles et al., 2002; Newsome et al., 2012; Worboys et al., 2015). Studies using other methods such as video monitoring and observational surveys also found differences in temporal patterns of visitation in parks, including the popularity of weekends for walkers and mountain bikers and hence the greater potential for conflict at these times (Arnberger, 2006; Porter and Wescott, 2004; Wolf et al., 2015).

The use of social media platforms varies depending on people's interests and activities and this was seen in Chapters 2, 3 and 4 for park visitation. The dominance of the fitness-based users on some route sharing platforms means that there is a bias towards these activities and visitors. For example, fitness focused park visitors mainly use trails while other visitors often use different facilities and locations such as picnickers, sightseers, campers, mountaineers and others (Norman and Pickering, 2017; Norman and Pickering, 2019; Walden-Schreiner et al., 2018). Results from Chapters 3 and 4 demonstrated that visitors posting GPS routes to fitness-based platforms differed in their preferences for trails and parks compared to those posting to adventure focused social media platforms. Depending on the questions being asked

by park managers and researchers, it is important to ensure that the best platform for the analysis is chosen, which in many cases could be from a range of sources (Figure 6.3.) (Norman and Pickering, 2017).

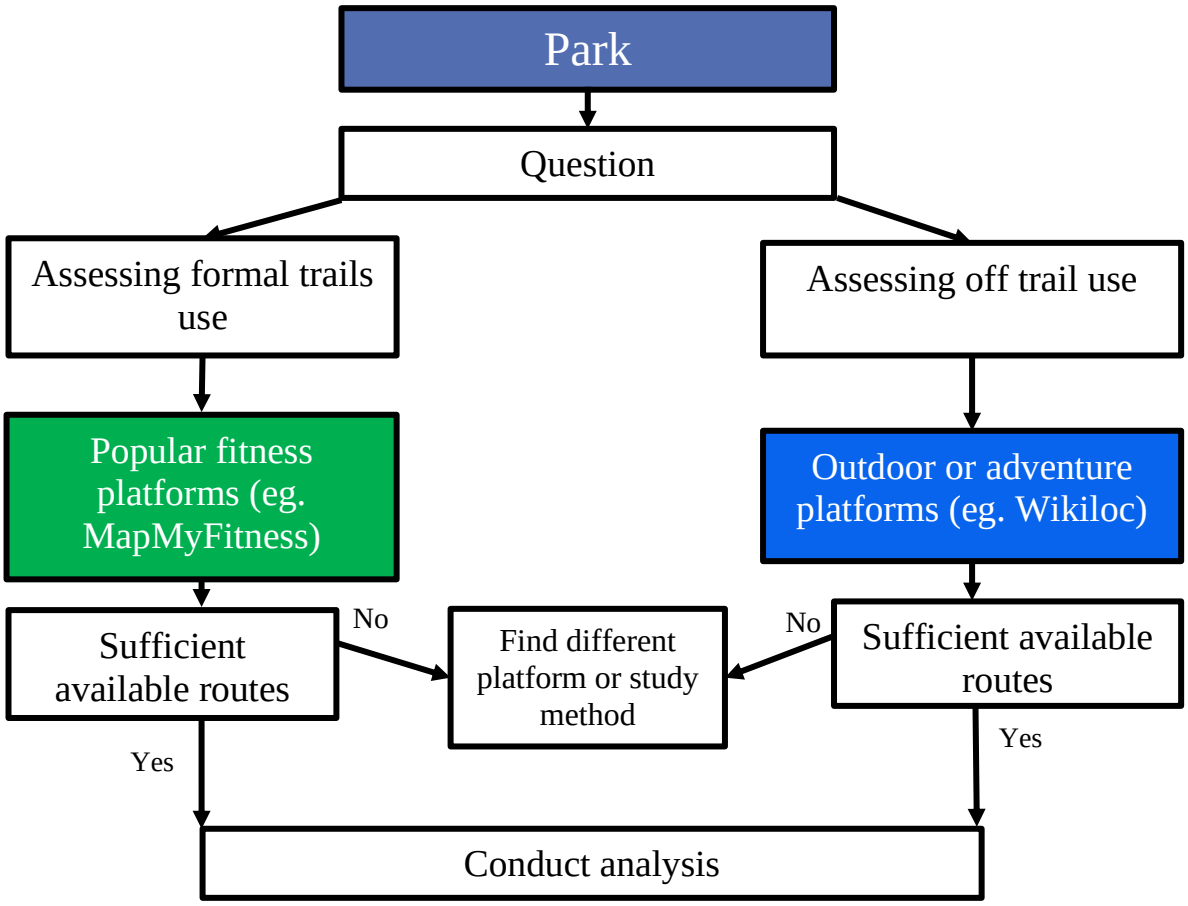

Figure 6.3. A simple decision tree for selecting the most appropriate social media platform for GPS route data for visitor monitoring in national parks.

6.3. Aim 3: What makes national parks popular for walkers, runners and mountain bikers (Chapter 4)?

Social media GPS route data can provide valuable information about the factors that influence the popularity of a park including for different types of recreation (Norman and Pickering, 2019). A number of factors influenced visitation to parks in south-

eastern Queensland for walking, running and mountain biking, many of which are likely to be important in other regions. The most important factor for mountain biking, running and walking on trails in south-eastern Queensland was distance to urban areas (Chapter 4). For Strava, there was a clear threshold with parks within 2 km of urban areas having 14.8 times more visitor that those further away (Figure 6.4.). Similar effects of distance on visitation to parks were found in other studies of park visitation using survey data (Schipperijn et al. 2010), including in the same region (Rossi et al. 2015). Studies using point data from social media also found that accessibility was very important, and it was more influential than charismatic animals in determining park visitation in sub-Saharan Africa (Hausmann et al., 2017). Understanding accessibility and the importance of distance is particularly important for parks close to cities as increasing urbanization is resulting in rapid increases in visitation to such parks (Levin et al., 2017). In contrast more remote parks can be popular where they provide specific activities and facilities, such as those found here for visitors focused more on adventure than fitness (Chapter 4).

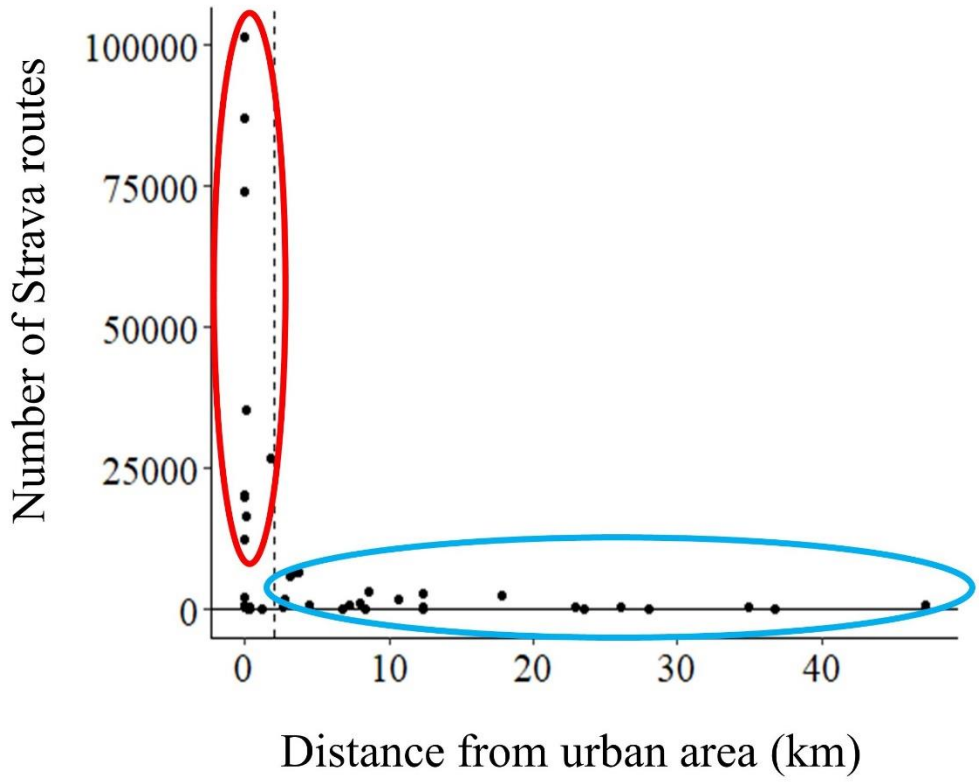

Figure 6.4. The relationship between distance to urban areas and visitation to national parks based on Strava route data, showing how parks within 2 km of urban areas (dashed line) had higher visitation (red oval) than those further away (blue oval). Adapted from a Figure in Chapter 4 (Norman and Pickering, 2019).

Another factor influencing the number of visits to parks is the presence of amenities and infrastructure, including recreational trails (Kaczynski et al. 2008, Norman and Pickering, 2019; Tenkanen et al., 2017). The existence of recreational trails allows visitors to access more of the parks including areas away from roads and entrance points. This can provide opportunities to escape crowds and busy areas allowing visitors to experience solitude and escape (Oishi, 2013). Careful consideration must be undertaken when planning and constructing infrastructure such as trails, as

increasing visitation including to otherwise remote areas, can lead to environmental and social impacts (Ballantyne and Pickering, 2015; Wolf et al., 2019).

The provision of single use trails can also increase the popularity of parks (Norman and Pickering, 2019). Results from Chapter 4 indicate that mountain bike specific trails were important factors in attracting riders, with parks with such facilities attracting more riders than similar parks with only multi-purpose management trails (Norman and Pickering, 2019). Other studies have found that dedicated trails and infrastructure are favoured by mountain bikers, and that they often prefer single use narrow trails with trail technical features (Goeft and Alder, 2001; Pickering et al., 2010). In and around cities unmet demand for riding experiences can result in mountain bikers constructing their own trails and trail technical features with important environmental and social impacts (Newsome and Davies, 2009; Pickering et al., 2010). Therefore, understanding the demand for specific infrastructure by some types of visitors is important, including identifying the most appropriate locations for trails while ensuring the overarching goals of nature conservation are still maintained.

6.4. Aim 4: What is the scale of the discourse about national parks on Twitter, and how does this vary among countries (Chapter 5)?

Twitter was the largest and most popular social platform used in the thesis, with over 330 million users globally (Sinnenberg et al., 2016; Hootsuite, 2018). Results from Chapter 5 highlight the enormity of the public debate about national parks, with over 2 million tweets during a 6-month period on Twitter. Limitations imposed by Twitter prevent easy access to past data, as the standard API only allows for the retrieval of tweets up to nine days prior to the request (Twitter, 2019). Even with this restriction,

the scale of content on Twitter means that it is a valuable source of data about parks (Halpenny and Blye, 2017; Tenkanen et al., 2017). With most tweets now lacking point geolocation data, text mining is required to obtain relevant spatial data such as which parks are discussed and by people from what countries.

There are important implications for park management from the global discourse about parks on Twitter. Firstly, tweets about specific parks were mostly sent by people from the same country with locals sending 60.6% of the tweets about a park on average (Chapter 5). Due to the platform used by news and media outlets, politicians and influencers as well as the wider public, it provides a useful source of data about a range of issues including those relevant to parks. This includes assessing how people feel when visiting parks and green spaces (Kovacs-Györi et al., 2018; Palomino et al., 2016). It can be used to understand public opinion about issues impacting national parks, such as pest management, poaching and attitudes to changes in access to parks (MacDonald et al., 2016; Wilkins et al., 2018).

The choice of search terms is important when accessing discourse on Twitter, as variations in language, characters and terms, can affect which tweets are identified using API (Lachlan et al., 2016). As ‘national park’ is a widespread term used to refer to Category II protected areas in over 85 countries (United Nations Environment World Conservation Monitoring Centre, 2019), the phrase allowed data about a large number of protected areas (264 parks) from many countries (55) to be identified from Twitter (Chapter 5). It did not cover some parks in some regions, such as Spanish speaking countries where many Category II protected areas are called ‘parque nacional’ (United Nations Environment World Conservation Monitoring Centre,

2019). Constraints around search terms is a common issue when using Twitter data, therefore it is important to test relevant search words and phrases (Lachlan et al., 2016; Sinnenberg et al., 2017).

When assessing large scale social media data sets, such as those obtained here from Twitter about parks, specific skills are required (Chapter 5). Such data sets rapidly exceed the capacity of some types of software. For example, the Twitter park dataset was too big for Microsoft Excel spreadsheets which can only accommodate a maximum of 1,048,576 rows (Microsoft Office, 2019). Therefore, for large datasets, researchers may have to rely on computing languages, such as Python and R to perform analyses (Python Software Foundation, 2019; R Core Team, 2019). Specialised skills, such as these are important considerations when assessing a large amounts of user generated data including from social media (Gandomi and Haider, 2015; Kern et al., 2016).

6.5. Aim 5: Can social media data be used to predict park popularity (Chapters 3, 4 and 5)?

The accuracy of social media data was assessed by comparing it with on-ground visitation data in Chapters 2, 3 and 5. At small spatial scales, social media GPS route data could accurately reflect relative trail usage within parks, although this does depend on the social media platform used and park in question (Norman et al., 2019; Norman and Pickering, 2017). For popular platforms such as MapMyFitness with data for a range of activities, route data can provide valuable insights into how parks are used, but as mentioned the popularity of platforms varies of time and among countries (Campelo and Mendes, 2016).

At a global scale, the size of the discourse about a park on Twitter was similar to its relative popularity with visitors, at least for the 40 most discussed parks (Chapter 5). When only the data for popular parks in the United States was assessed, there was an even stronger relationship suggesting that within country comparisons may provide more accurate assessments. However, this was a rank order analysis, with the ratio of tweets to visitors, orders of magnitude different among parks and hence, Twitter has limited capacity to accurately predict visitor numbers without other forms of visitor monitoring. The capacity of Twitter data for detecting visitation is also dependent on where the parks are located as the platform's popularity varies among countries and is restricted or banned in others, with the prominent example of China the platforms use (Stockmann and Luo., 2017). These factors make Twitter less useful for assessing park visitation compared to other sources of geolocated visitor data on social media, particularly at smaller scales or in certain countries (Folch et al., 2018; Manoruang and Asavasuthirakul, 2019).

Where there was spatial information included in tweets, they were nearly always for tweets that were originally posted to Instagram and then shared to Twitter and hence they relied on Instagram's geolocation methods. In order to share locations, Instagram users nominate a place using text and then selecting from a list of nominated locations rather than using GPS location data directly from a smartphone, limiting the quality of the spatial data (Instagram Help, 2019). Prior to Twitter changing its geolocation protocol in 2014, tweets were automatically assigned coordinates from the smartphones GPS, with varying accuracy depending on the settings and type of device used (Leetaru, 2019). Due to the vast amount of data created prior to 2014,

geolocation point data associated with tweets was utilised by researchers to assess the wide range of questions, including park visitation (Hamstead et al., 2018; Kovacs-Györi et al., 2018; Tenkanen et al., 2017). Searching for tweets using geolocation data is no longer possible, as accurate geolocation was removed from the platform for privacy and security reasons in 2014 (Leetaru, 2019; Twitter, 2019).

Although social media data can be useful for estimating the relative popularity of trails (Chapter 2) and parks (Chapter 5), social media data does not represent all visitors to parks (Hausmann et al., 2018; Wilkins et al., 2018). Surveys of park visitors in Kruger National Park and Crater Lake National Park found that 25% and 24% of visitors use Twitter respectively (Hausmann et al., 2018; Wilkins et al., 2018). The use of social media is common among park visitors, with 72.5% of nature-based visitors were likely to share their experiences online, though this varies among platforms (Hausmann et al., 2018). More generally research on social media platforms shows that people from certain regions, ages, and wealth are more likely to use social media, for example, Twitter is more popular in some countries for younger, wealthier people and those in politics and news organisations (Orellana-Rodrigues and Keane, 2018; Pew Research Centre, 2018). Similar biases occur with GPS route sharing platforms, with those posting routes to Strava often younger (between 25-54) with more men posting than women (Colorado Government, 2018; Heesch et al., 2016). These issues with who provides data on social media is an important issue not only when using data to better understand park visitation, but more broadly reflecting the increasing impact of the digital divide and differences between digital natives and digital immigrants (Ahn and Jung, 2016; Friemel, 2016; Lavery et al., 2018).

Even within users of social media there are further biases, with some individuals posting far more content than others (Panek et al., 2013; Wood et al., 2013). Methods used to help mitigate the emphasis on some visitors over others, include capping each visitor to a single post, converting data to visitor days or visitor park days (Pickering et al., 2020; Wood et al., 2013). The best methods to use to deal with the range of biases with social media, depend on the question being asked, the social media platform and the dominance of some visitors on social media (Ghermandi and Sinclair 2019; Tenkanen et al. 2017; Wood et al., 2013).

6.6. Contribution to knowledge

Collecting visitation data and public discourse about national parks has traditionally been time consuming and expensive (Ankre et al., 2016; Eagles et al., 2002; Newsome et al., 2012). With the rise of social media, large amounts of relevant data about park visitation is being produced and shared on social media without the involvement of researchers or managers (Ghermandi and Sinclair, 2019; Toivonen et al., 2019). This thesis has started to address important gaps in the research on social media including how data from different platforms can be used to assess spatial patterns in park visitation. Specifically, this thesis found:

1. Large amounts of data about the spatial movements of visitors within national parks is available on social media.
2. The type of data available from platforms differs.
3. Access to data varies among platforms, with some providing APIs that facilitate data collection, while others no longer provide access or provide limited data.

4. Route data is useful to assess spatial patterns of visitation in parks including informal trails and off-trails use.
5. Route data can highlight differences in visitation among recreation activities.
6. It can also provide useful information about temporal patterns of visitation within parks, including for different activities.
7. Social media platforms are used by a range of visitors, whom are attracted to parks for different reasons.
8. The distance to urban areas as well as the presence of trail infrastructure, effects the popularity of parks.
9. A vast amount of the public discourse about national parks on Twitter is produced by people in the same country as the park.
10. The amount of discourse posted about popular national parks on Twitter, reflects visitation numbers at a rank order scale.
11. Route data can reflect on-ground visitation for different activities in parks.

6.7. Research context and future research directions

The thesis examined the benefits of two sorts of social media data: GPS route data to examine adventure and fitness focused visitors going hiking, running or mountain biking, and Twitter text to assess the discourse on parks. There are a wide range of other platforms available that contain different types of data (Table 6.1), but also provide information about different types of park visitors to GPS route data. For example, data about park visitors interested in wildlife viewing, cultural tourism and those accessing remote areas for activities such as camping and hiking has been obtained from image sharing platforms such as Flickr and Panoramio (Ghermandi and Sinclair, 2019; Sessions et al., 2016; Tenkanen et al., 2017; Figueroa-Alfaro and

Tang, 2016; Pickering et al., 2020; Walden-Schreiner et al., 2018). There are, however, a number of visitor types sharing data online who could also provide important spatial information about visitation.

Beyond social media there are other sources of data about park visitors online including citizen science websites. This includes eBird, iNaturalist or GBIF that contain large amounts of data that is often free and accessible for research (eBird, 2019; iNaturalist, 2019; GBIF, 2019). The data often relates to visits to national parks to see flora and fauna (Devictor et al., 2010; Kallimanis et al., 2017), and include observations and geolocation data (Conrad and Hilchey, 2011). The scale of the data is increasing with over 12 million bird sighting records on Ebird from over 7,500 users just in Australia (eBird, 2019). Although many of these platforms are not by definition, social media platforms due to the lack of ability to follow other users, the information posted can also be useful to assess park visitation. With many citizen science platforms increasing in popularity, exploring the use of such data to assess engagement is likely be an emerging area of research.

6.9. Conclusion

The thesis assessed the spatial aspects of visitation and discourse about national parks posted on social media, focusing primarily on GPS route and text data. Overall social media data provided valuable information about visitation and discourse about national parks, with route data providing insights into visitation by people engaged in different activities in parks as well as providing insights into which factors influence visitation. Likewise, data from Twitter provides information about the scale and diversity of the discourse about national parks globally, including which parks were

popular. Further research using social media for assessing park visitation and discourse is necessary including assessing its benefits and limitations including biases around user demographics. However, with the expanding popularity of social media it is likely that it will continue to be a valuable source of data about national parks into the future.

References

Ahn, J. and Jung, Y., 2016. The common sense of dependence on smartphone: A comparison between digital natives and digital immigrants. *New Media and Society, 18*(7), pp.1236-1256.

Ankre, R., Fredman, P. and Lindhagen, A., 2016. Managers' experiences of visitor monitoring in Swedish outdoor recreational areas. *Journal of Outdoor Recreation and Tourism*, *14*, pp.35-40.

Arnberger, A., 2006. Recreation use of urban forests: An inter-area comparison. *Urban Forestry and Urban Greening, 4*(3-4), pp.135-144.

Barros, A., Gonnet, J. and Pickering, C., 2013. Impacts of informal trails on vegetation and soils in the highest protected area in the Southern Hemisphere. *Journal of Environmental Management, 127*, pp.50-60.

Ballantyne, M. and Pickering, C.M., 2015. The impacts of trail infrastructure on vegetation and soils: current literature and future directions. *Journal of Environmental Management*, *164*, pp.53-64.

Byczek, C., Longaretti, P.Y., Renaud, J. and Lavorel, S., 2018. Benefits of crowd-sourced GPS information for modelling the recreation ecosystem service. *PloS One, 13*(10), pp.e0202645.

Campelo, M.B. and Mendes, R.M.N., 2016. Comparing webshare services to assess mountain bike use in protected areas. *Journal of Outdoor Recreation and Tourism, 15*, pp.82-88.

Colorado Government, 2018. *Strava MetroData analysis summary report*. Accessible from: https://www.codot.gov/admin/programs/bikeped/documents/strava-analysis-summary_06-25-18.pdf

Conrad, C.C. and Hilchey, K.G., 2011. A review of citizen science and community-based environmental monitoring: issues and opportunities. *Environmental Monitoring and Assessment*, *176*(1-4), pp.273-291.

Devictor, V., Whittaker, R.J. and Beltrame, C., 2010. Beyond scarcity: Citizen science programmes as useful tools for conservation biogeography. *Diversity and Distributions*, *16*(3), pp.354-362.

eBird, 2019. eBird, viewed 10 December 2019. Retrieved from: https://ebird.org/home

Eagles, P.F., McCool, S.F., Haynes, C.D. and Phillips, A., 2002. *Sustainable tourism in protected areas: Guidelines for planning and management* (Vol. 8). IUCN: Gland.

European Commission, 2019. EU data protection rules, viewed 30 December 2019. Retrieved from: https://ec.europa.eu/info/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules/eu-data-protection-rules_en

Figueroa-Alfaro, R.W. and Tang, Z., 2017. Evaluating the aesthetic value of cultural ecosystem services by mapping geo-tagged photographs from social media data on Panoramio and Flickr. *Journal of Environmental Planning and Management*, *60*(2), pp.266-281.

Flickr, 2019. Flickr, viewed 5 December 2019. Retrieved from: https://www.flickr.com/

Folch, D.C., Spielman, S.E. and Manduca, R., 2018. Fast food data: Where user-generated content works and where it does not. *Geographical Analysis*, *50*(2), pp.125-140.

Friemel, T.N., 2016. The digital divide has grown old: Determinants of a digital divide among seniors. *New Media and Society*, *18*(2), pp.313-331.

Gandomi, A. and Haider, M., 2015. Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), pp.137-144.

GBIF, 2019. GBIF, viewed 3 December 2019. Retrieved from: https://www.gbif.org/

Ghermandi, A. and Sinclair, M., 2019. Passive crowdsourcing of social media in environmental research: A systematic map. *Global Environmental Change*, *55*, pp.36-47.

Goeft, U. and Alder, J., 2001. Sustainable mountain biking: A case study from the southwest of Western Australia. *Journal of Sustainable Tourism*, *9*(3), pp.193-211.

GPSies, 2019a. GPSies, viewed 3 December 2019. Retrieved from: https://www.gpsies.com/

GPSies, 2019b. API Documentation. Available from: https://www.gpsies.com/api/GPSiesAPI_en.pdf

Halpenny, E.A. and Blye, C.J., 2017. North American park agencies' evolving use of Twitter: A content analysis of 2014 and 2017 Tweets. *Parks*, *23.2*, pp.79.

Hamstead, Z.A., Fisher, D., Ilieva, R.T., Wood, S.A., McPhearson, T. and Kremer, P., 2018. Geolocated social media as a rapid indicator of park visitation and equitable park access. *Computers, Environment and Urban Systems*, *72*, pp.38-50.

Hausmann, A., Toivonen, T., Heikinheimo, V., Tenkanen, H., Slotow, R. and Di Minin, E., 2017. Social media reveal that charismatic species are not the main attractor of ecotourists to sub-Saharan protected areas. *Scientific Reports*, *7*(1), pp.763.

Heesch, K.C., James, B., Washington, T.L., Zuniga, K. and Burke, M., 2016. Evaluation of the Veloway 1: A natural experiment of new bicycle infrastructure in Brisbane, Australia. *Journal of Transport and Health*, *3*(3), pp.366-376.
iNaturalist, 2019.

iNaturalist, viewed 10 December 2019. Retrieved from: https://www.inaturalist.org/

Hootsuite, 2018. Global Digital Statshot Q2, view 26 November, 2019. Retrieved from: https://datareportal.com/reports/digital-2018-q2-global-digital-statshot

Instagram Help, 2019. How do I add a location sharing my photo or video?, viewed 10 December 2019. Retrieved from: https://help.instagram.com/408972995943225

Jurado Rota, J., Pérez Albert, M.Y. and Serrano Giné, D., 2019. Visitor monitoring in protected areas: an approach to Natura 2000 sites using Volunteered Geographic Information (VGI). *Geografisk Tidsskrift-Danish Journal of Geography*, *119*(1), pp.69-83.

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

Kallimanis, A.S., Panitsa, M. and Dimopoulos, P., 2017. Quality of non-expert citizen science data collected for habitat type conservation status assessment in Natura 2000 protected areas. *Scientific Reports*, *7*(1), pp.8873.

Kern, M.L., Park, G., Eichstaedt, J.C., Schwartz, H.A., Sap, M., Smith, L.K. and Ungar, L.H., 2016. Gaining insights from social media language: Methodologies and challenges. *Psychological Methods*, *21*(4), pp.507.

Kidd, A.M., D'Antonio, A., Monz, C., Heaslip, K., Taff, D. and Newman, P., 2018. A GPS-based classification of visitors' vehicular behavior in a protected area setting. *Journal of Park and Recreation Administration*, *36*(1).

Kovacs-Györi, A., Ristea, A., Kolcsar, R., Resch, B., Crivellari, A. and Blaschke, T., 2018. Beyond spatial proximity—classifying parks and their visitors in London based on spatiotemporal and sentiment analysis of Twitter data. *ISPRS International Journal of Geo-Information*, *7*(9), pp.378.

Lachlan, K.A., Spence, P.R., Lin, X., Najarian, K. and Del Greco, M., 2016. Social media and crisis management: CERC, search strategies, and Twitter content. *Computers in Human Behavior*, *54*, pp.647-652.

Lavery, M.P., Abadi, M.M., Bauer, R., Brambilla, G., Cheng, L., Cox, M.A., Dudley, A., Ellis, A.D., Fontaine, N.K., Kelly, A.E. and Marquardt, C., 2018. Tackling Africa's digital divide. *Nature Photonics*, *12*(5), pp.249.

Levin, N., Lechner, A.M. and Brown, G., 2017. An evaluation of crowdsourced information for assessing the visitation and perceived importance of protected areas. *Applied Geography*, *79*, pp.115-126.

Leetaru, K. 2019. Forbes: Visualizing seven years of Twitter's evolution: 2012-2018, viewed 10 October 2019. Retrieved from: https://www.forbes.com/sites/kalevleetaru/2019/03/04/visualizing-seven-years-of-twitters-evolution-2012-2018/#66c249b47ccf

Macdonald, D., Jacobsen, K., Burnham, D., Johnson, P. and Loveridge, A., 2016. Cecil: a moment or a movement? Analysis of media coverage of the death of a lion, *Panthera leo*. *Animals*, *6*(5), pp.26.

Manoruang, D. and Asavasuthirakul, D., 2019. Quality analysis of online geocoding services for Thai text addresses. *Engineering and Applied Science Research*, *46*(2), pp.86-97.

MapMyFitness. (2019), MapMyFitness, viewed 5 December 2019. Retrieved from: http://www.mapmyfitness.com/

McKenna, A.T., Gaudion, A.C. and Evans, J.L., 2019. The role of satellites and smart devices: Data surprises and security, privacy, and regulatory challenges. *Penn State Law Review*, *123*(591).

Microsoft Office, 2019. Excel specifications and limits, viewed 15 November 2019. Retrieved from: https://support.office.com/en-us/article/excel-specifications-and-limits-1672b34d-7043-467e-8e27-269d656771c3

Newsome, D. and Davies, C., 2009. A case study in estimating the area of informal trail development and associated impacts caused by mountain bike activity in John Forrest National Park, Western Australia. *Journal of Ecotourism*, *8*(3), pp.237-253.

Newsome, D., Moore, S.A. and Dowling, R.K., 2012. *Natural area tourism: Ecology, impacts and management* (Vol. 58). Channel View publications: Bristol.

Norman, P. and Pickering, C.M., 2017. Using volunteered geographic information to assess park visitation: Comparing three on-line platforms. *Applied Geography*, *89*, pp.163-172.

Norman, P. and Pickering, C.M., 2019. Factors influencing park popularity for mountain bikers, walkers and runners as indicated by social media route data. *Journal of Environmental Management*, *249*, pp.109413.

Norman, P., Pickering, C.M. and Castley, G., 2019. What can volunteered geographic information tell us about the different ways mountain bikers, runners and walkers use urban reserves? *Landscape and Urban Planning, 185*, pp.180-190.

Oishi, Y., 2013. Toward the improvement of trail classification in national parks using the recreation opportunity spectrum approach. *Environmental Management*, *51*(6), pp.1126-1136.

Orellana-Rodriguez, C. and Keane, M.T., 2018. Attention to news and its dissemination on Twitter: A survey. *Computer Science Review*, *29*, pp.74-94.

Palomino, M., Taylor, T., Göker, A., Isaacs, J. and Warber, S., 2016. The online dissemination of nature–health concepts: Lessons from sentiment analysis of social media relating to "nature-deficit disorder". *International Journal of Environmental Research and Public Health*, *13*(1), pp.142.

Panek, E.T., Nardis, Y. and Konrath, S., 2013. Mirror or Megaphone?: How relationships between narcissism and social networking site use differ on Facebook and Twitter. *Computers in Human Behavior*, *29*(5), pp.2004-2012.

Panoramio, 2019. Panoramio has been discontinued, viewed 3 December 2019. Retrieved from: https://www.panoramio.com/

Pew Research Centre, 2018. Social Media Use in 2018, Appendix A: Detailed table, viewed 2 November 2019. Retrieved from: https://www.pewresearch.org/internet/2018/03/01/social-media-use-2018-appendix-a-detailed-table/

Pickering, C., Castley, J.G., Hill, W. and Newsome, D., 2010. Environmental, safety and management issues of unauthorised trail technical features for mountain bicycling. *Landscape and Urban Planning*, *97*(1), pp.58-67.

Pickering, C., Walden-Schreiner, C., Barros, A. and Rossi, S.D., 2020. Using social media images and text to examine how tourists view and value the highest mountain in Australia. *Journal of Outdoor Recreation and Tourism*, *29*, pp.100252.

Porter, C. and Wescott, G., 2004. Recreational use of a marine protected area: Point Lonsdale, Victoria. *Australasian Journal of Environmental Management*, *11*(3), pp.201-211.

Python Software Foundation, 2019. Python Language Reference, version 3.7. Available at http://www.python.org

R Core Team, 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

Rossi, S.D., Byrne, J.A. and Pickering, C.M., 2015. The role of distance in peri-urban national park use: Who visits them and how far do they travel? *Applied Geography, 63*, pp.77-88.

Santos, T., Mendes, R.N. and Vasco, A., 2016. Recreational activities in urban parks: Spatial interactions among users. *Journal of Outdoor Recreation and Tourism, 15*, pp.1-9.

Schipperijn, J., Ekholm, O., Stigsdotter, U.K., Toftager, M., Bentsen, P., Kamper-Jørgensen, F. and Randrup, T.B., 2010. Factors influencing the use of green space: Results from a Danish national representative survey. *Landscape and Urban Planning, 95*(3), pp.130-137.

Schneble, C.O., Elger, B.S. and Shaw, D., 2018. The Cambridge Analytica affair and Internet-mediated research. *EMBO Reports, 19*(8).

Sessions, C., Wood, S.A., Rabotyagov, S. and Fisher, D.M., 2016. Measuring recreational visitation at US National Parks with crowd-sourced photographs. *Journal of Environmental Management, 183*, pp.703-711.

Sinnenberg, L., Buttenheim, A.M., Padrez, K., Mancheno, C., Ungar, L. and Merchant, R.M., 2017. Twitter as a tool for health research: a systematic review. *American Journal of Public Health, 107*(1), pp.e1-e8.

Sinnenberg, L., DiSilvestro, C.L., Mancheno, C., Dailey, K., Tufts, C., Buttenheim, A.M., Barg, F., Ungar, L., Schwartz, H., Brown, D. and Asch, D.A., 2016. Twitter as a potential data source for cardiovascular disease research. *JAMA Cardiology, 1*(9), pp.1032-1036.

Strava, 2019. Strava, viewed 5 December 2019. Retrieved from: https://www.strava.com/

Strava Metro, 2019. Strava Metro, viewed 3 December 2019. Retrieved from: https://metro.strava.com/

Stockmann, D., and Luo, T., 2017. Which social media facilitate online public opinion in China? *Problems of Post-communism, 64*(3-4), pp.189-202.

Sun, Y. and Mobasheri, A., 2017. Utilizing Crowdsourced data for studies of cycling and air pollution exposure: A case study using Strava Data. *International Journal of Environmental Research and Public Health, 14*(3), pp.274.

Tenkanen, H., Di Minin, E., Heikinheimo, V., Hausmann, A., Herbst, M., Kajala, L. and Toivonen, T., 2017. Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas. *Scientific Reports, 7*(1), pp.17615.

Tikkinen-Piri, C., Rohunen, A. and Markkula, J., 2018. EU General Data Protection Regulation: Changes and implications for personal data collecting companies. *Computer Law and Security Review, 34*(1), pp.134-153.

Toivonen, T., Heikinheimo, V., Fink, C., Hausmann, A., Hiippala, T., Järv, O., Tenkanen, H. and Di Minin, E., 2019. Social media data for conservation science: a methodological overview. *Biological Conservation*, *233*, pp.298-315.

Twitter, 2019. Search Tweets, viewed 10 October 2019. Retrieved from: https://developer.twitter.com/en/docs/tweets/search/overview

United Nations Environment World Conservation Monitoring Centre., 2019. Protected Planet. Available from: https://www.protectedplanet.net/

Walden-Schreiner, C. and Leung, Y.F., 2013. Spatially characterizing visitor use and its association with informal trails in Yosemite Valley meadows. *Environmental Management*, *52*(1), pp.163-178.

Walden-Schreiner, C., Rossi, S.D., Barros, A., Pickering, C. and Leung, Y.F., 2018. Using crowd-sourced photos to assess seasonal patterns of visitor use in mountain-protected areas. *Ambio*, *47*(7), pp.781-793.

Wei LY., Zheng Y., Peng WC. (2015) Mining popular routes from social media. In: Baughman A., Gao J., Pan JY., Petrushin V. (eds). pp.93-116. *Multimedia Data Mining and Analytics*. Springer: Basel

Wikiloc, 2019. Wikiloc, viewed 3 December 2019. Retrieved from: https://www.wikiloc.com/

Wilkins, E.J., Smith, J.W. and Keane, R., 2018. Social media communication preferences of national park visitors. *Applied Environmental Education and Communication*, pp.1-15.

Wolf, I.D., Croft, D.B. and Green, R.J., 2019. Nature Conservation and Nature-Based Tourism: A Paradox? *Environments*, *6*(9), pp.104.

Wolf, I.D., Hagenloh, G. and Croft, D.B., 2012. Visitor monitoring along roads and hiking trails: How to determine usage levels in tourist sites. *Tourism Management*, *33*(1), pp.16-28.

Wolf, I.D., Wohlfart, T., Brown, G. and Lasa, A.B., 2015. The use of public participation GIS (PPGIS) for park visitor management: A case study of mountain biking. *Tourism Management*, *51*, pp.112-130.

Wood, S.A., Guerry, A.D., Silver, J.M. and Lacayo, M., 2013. Using social media to quantify nature-based tourism and recreation. *Scientific Reports*, *3*, pp.2976.

Worboys, G.L., Lockwood, M., Kothari, A., Feary, S. and Pulsford, I. eds., 2015. *Protected area governance and management*. ANU Press: Canberra.

Zhang, Z. and Gupta, B.B., 2018. Social media security and trustworthiness: overview and new direction. *Future Generation Computer Systems*, *86*, pp.914-925.

Appendices

Appendix 1

Griffith University policy on the inclusion of papers within the thesis. Accessed November 2019 https://intranet.secure.griffith.edu.au/research/griffith-graduate-research-school/preparing-your-thesis/inclusion-of-papers-within-the-thesis.

Inclusion of papers within the thesis

Higher degree by research is a program of independent supervised study that produces significant and original research outcomes, culminating in a thesis, exegesis or equivalent (refer to Higher Degree by Research Thesis). Inclusion of papers within a thesis is not a suitable thesis format for all research projects, for example: collaborative projects where there may be several co-authors for each paper which may make it difficult for the examiner to establish the independence of the candidates work; where primary data is not collected, or results obtained, until late in the candidature; or where the research will not produce a logical sequence of papers that are able to be presented as an integrated whole. Candidates should also take into account whether this thesis format is an accepted practice within their discipline and likely to be received well by the thesis examiners (refer also to the examination requirements below). Candidates are required to consult with their supervisor(s) early in their candidature to determine if this thesis format is appropriate. It is expected that candidates will identify as part of the confirmation of candidature milestone if their thesis is to be prepared in this format. Candidates should consult their Group specific guidelines in addition to the requirements detailed below. Candidates are also encouraged to attend the workshop: 'Inclusion of papers within a thesis' offered by the Griffith Graduate Research School.

Refer also to the Griffith University Code for the Responsible Conduct of Research, specifically the sections pertaining to publication ethics and the dissemination of research findings, and authorship.

Status of papers A thesis may include papers that have been submitted, accepted for publication, or published. Some disciplines may specify a variation to the status of papers requirement, refer to your Group specific guidelines.

Type of papers For the purpose of this requirement, papers are defined as a journal article, conference publication, book or book chapter. Papers which have been rejected by a publisher must not be included unless they have been substantially rewritten to address the reviewers' comments, or have since been accepted for publication. Some disciplines may specify a variation to the type of papers requirement, refer to your Group specific guidelines.

Number of papers A thesis may be entirely or partly comprised of papers. A paper maybe included as a single chapter if the paper contributes to the argument of the thesis, or several papers may form the core chapters of the theses where they present a cohesive argument. Where a thesis is entirely comprised of papers, there is no minimum requirement for the number of papers that must be included (except as noted below) and is a matter of professional judgment for the supervisor and the candidate.

Overall, the material presented for examination needs to reflect the research thesis standard required for the award of the degree. For example, PhD candidates, on the basis of a program of independent supervised study, must produce a thesis that makes a significant and original contribution to knowledge and understanding in the relevant field of study. This remains a matter of professional judgment for the supervisor and the candidate.

Where a thesis is entirely comprised of papers, some disciplines may specify a minimum number of papers to be included, refer to your Group specific guidelines.

Authorship The candidate should normally be principal author (that is, responsible for the intellectual content and the majority of writing of the text) of any work included in the body of the thesis. Where a paper has been co-authored, the candidate is required to have made a substantial contribution to the intellectual content and writing of the text, Co-authored work in which the candidate was a minor author can only be used and referenced in the way common to any other research publication cited in the thesis. A signature from the corresponding author is required in order to include co- authored material in the body of the thesis, refer to the declarations section below.

For co-authored papers, the attribution of authorship must be in accordance with the Griffith University Code for the Responsible Conduct of Research, which specifies that 'authorship must be based on substantial contributions in one or more of:

- conception and design of the research project
- analysis and interpretation of research data
- drafting or making significant parts of the creative or scholarly work or critically revising it so as to contribute significantly to the final output'.

Some disciplines may specify a variation to the authorship requirement, refer to your Group specific guidelines.

Quality of papers Candidates should endeavour to publish their research in high quality peer reviewed publications. Papers to be included in the body of the thesis should be published (or submitted for publication) in reputable outlets that are held in higher regard in the relevant field of research. Candidates should consult their supervisor(s) for advice on suitable publications specific to their research discipline. Some disciplines may specify quality standards that must be met for papers to be included, refer to your Group specific guidelines.

The library also provides support and advice to candidates on choosing a journal. Candidates are advised to note in particular advice in order to avoid 'predatory' publishers.

- Strategic Publishing Guide for Authors
- Publishing in Open Access journals

Copyright As copyright in an article is normally assigned to a publisher, the publisher must give permission to reproduce the work in the thesis and put a digital copy on the institutional repository. Information on how to seek permission is available at: Copyright and Articles in thesis. If permission cannot be obtained, students may still include the publication in the body of the thesis, however following examination the relevant chapter(s) will be redacted from the digital copy to be held by the Griffith University Library so that the copyright material is not made publicly available in the

institutional repository. Students are required to advise the copyright status of each publication included in the thesis via a declaration to be inserted in the thesis, as detailed below.

Students requiring further advice regarding copyright issues can contact the Information Policy Officer on (07) 3735 5695 or copyright@griffith.edu.au.

Group and discipline requirements

Some Groups or Elements may specify additional requirements for including papers within a thesis, refer below:

- Arts, Education and Law
- Griffith Business School (PDF 214k)
- Griffith Health
- Griffith Sciences (PDF 271k)

Appendix 2

Supplementary material from **Chapter 2**

Interactive online map which can be viewed at
https://www.sciencedirect.com/science/article/pii/S0169204618304730

Appendix 3

Supplementary material from **Chapter 3**
Interactive online map which can be viewed at
https://www.sciencedirect.com/science/article/pii/S0143622817306409?via%3Dihub