Journal of Park and Recreation Administration
https://doi.org/10.18666/JPRA-2020-10541
Volume 39, Issue 1, Spring 2021, pp. 123–130

Research Note

Mobile Device Data Analysis to Determine the Demographics of Park Visitors

Christopher Monz[a], Noah Creany[a], Jordan Nesbitt[a], and Milan Mitrovich[b]

[a] Department of Environment and Society, Utah State University, Logan, Utah
[b] Natural Communities Coalition, Irvine, California

Please send correspondence to Christopher Monz, chris.monz@usu.edu

Abstract

The systematic gathering of data on visitors to parks and protected areas (PPAs) is an important aspect of adaptive management, but also presents an ongoing challenge to conduct. Visitor questionnaires, often administered at entrance facilities, trailheads and parking areas, are the most common approach, but require significant staff time for field work and analysis and may not be inclusive of all visitors. In this paper, we present a novel, mobile device-based assessment strategy that can determine basic visitor demographic attributes without the need for field-based assessments. We compared results from this strategy to a common, trailhead questionnaire approach in three urban-proximate park locations and generally found similar results, except in a situation where a significant number of visitors may not have entered via the main location, and thus were missed by the survey. Overall, the mobile device strategy is likely more accurate in situations where visitors enter parks from multiple, sometimes informal entry locations, but is also somewhat limited in the types of data available.

Keywords

Visitor demographics, park and protected area management, big data, mobile device data

Introduction

The sustainable management of visitors to parks and protected areas (PPAs) continues to be complicated by rapidly changing social and technological phenomena that influence visitors' behavior, attitudes and perceptions regarding recreation and tourism activities (Hammit et al., 2015). Rapid changes in use levels and patterns are becoming increasingly apparent at both a worldwide scale (Balmford et al., 2015) and at the individual park and system level as well, for example, the dramatic increases in use occurring in many U.S. National Parks (NPS, 2020). Managers of national parks, wildlands and protected areas are often required to serve legal mandates to accommodate this

visitation while maintaining a high degree of resource protection. Thus, knowledge of visitor use, current visitation trends and visitor characteristics is constantly needed to balance visitor use management and resource protection.

Fundamental to visitor management in PPAs is an understanding of the basic characteristics of visitors, the numbers of visitors, where they enter and exit, and the spatial attributes of their visit (English & Bowker, 2018). Moreover, understanding the social correlates to outdoor recreation participation (i.e., the basic demographic, socio-economic characteristics, motivations, etc.), has also been important to PPA managers (Manning, 2011). Understanding these social drivers of recreation participation may be even more relevant today in an effort to understand the causal factors influencing some of the rapid changes in use levels being observed in many protected areas. In addition, PPAs, particularly in urban proximate locations, are now more than ever viewed as locations that are relevant to human health and well-being. As the USA transitions to an even more urbanized and ethnically diverse population, understanding use demographics provides initial information on equity of access and potential constraints and barriers to visitation—a step to allowing managers to develop strategies to make PPAs and their benefits more available to the overall populace.

Methodologically, visitor questionnaires, trip diaries and interview techniques have typically provided demographic information regarding PPA visitation, but these approaches require significant time investment from both visitors and researchers (Manning, 2011). These methods are also prone to issues of reporting accuracy and sampling challenges (Hallo et al., 2005). Today, mobile device use is prevalent among the general public for navigation, travel, and communication, and the associated data being captured by mobile service providers and mobile analytic firms yields information on recreation and tourism behaviors, home locations, and a range of demographic attributes useful to PPA managers (e.g., Xu et al., 2020). Consequently, a growing number of studies have applied this technology to PPA visitation (e.g., Korpilo et al., 2017; Monz et al., 2019; Rice et al., 2019) as mobile device approaches reduce the need for field data collection since the data sources exist and often can be analyzed from recent past history to examine trends. An important distinction is there are also methodologies using mobile device app data, such as personal fitness and tracking applications (e.g., Strava), and the associated data to understand PPA recreation use, but these approaches require the active participation by the visitor (Korpilo et al., 2018; Muñoz et al., 2019; Norman & Pickering, 2017). The mobile device approach and data discussed in this paper are passively collected and derived from mobile devices so long as common location-service settings are enabled and the visitor carries the device during their visit—which is becoming increasingly the norm in many PPAs.

In this research note we describe a novel approach that uses existing mobile device data sources and available analysis tools to answer some fundamental demographic questions about PPA visitation. The overall goal of this work is to present a new method for gathering demographic data on PPA visitors. Our approach used data purchased from Streetlight Data, Inc., a transportation data analysis provider, and their associated web-based analysis tool and compared this directly to data we collected via a traditional trailhead questionnaire approach. Using Streetlight Data allows for the understanding of some visitor demographic attributes without the need to deploy field personnel to administer surveys. We present examples of our analysis findings at select parks as an illustration of an emerging method rather than a full examination of demographic trends at our study location. This paper reports on an expansion of the work con-

ducted in a previous study we published in this journal (Monz et al., 2019) that used Streetlight Data to examine overall use levels in park locations. We refer the reader to this previous paper and the references therein for a more complete discussion of the emergent work on this overall topic and for additional details on our mobile device data sources, methodological approaches and analysis procedures. We present these new findings in a brief format here.

Methods

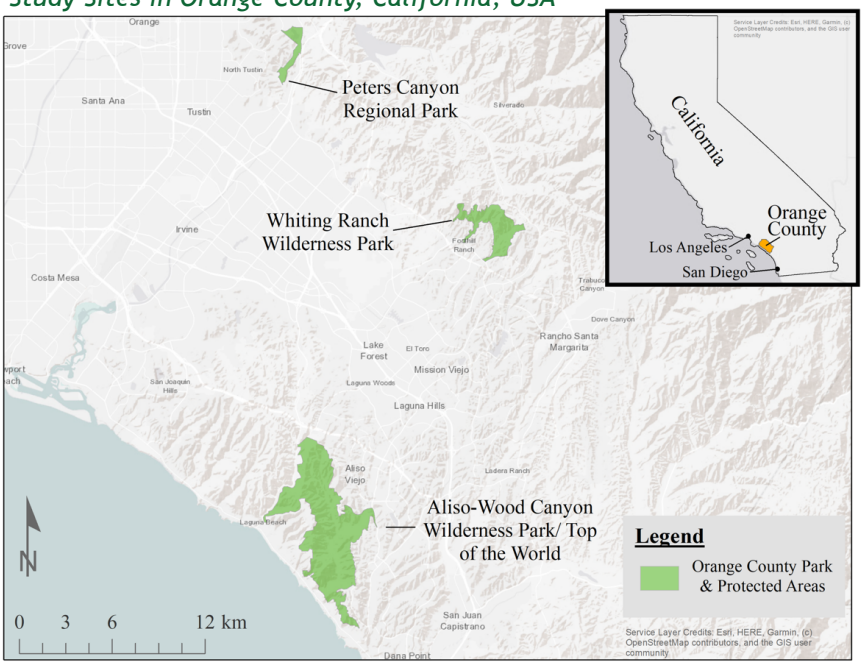
Study Site

The Nature Reserve of Orange County, California, USA (referred to throughout this paper as "the Reserve") consists of a total of 22 park units under various management designations. The Natural Communities Coalition (NCC; occonservation.org) is a nonprofit organization that coordinates conservation and management efforts across the Reserve, and thus while individual park unit differences exist, many programs and initiatives span jurisdictional boundaries. The Reserve is part of the California Chaparral and Woodlands Ecoregion with primary vegetation types of coastal sage scrub, oak woodland, native grassland, chaparral, Tecate cypress and riparian communities. The Reserve system parks provide outdoor recreation opportunities, such as hiking, running, mountain biking, beach recreation and nature appreciation in an urban-proximate setting to the over 3.2 million residents of Orange County (Center for Demographic Research, 2020).

Three Reserve parks were selected for this analysis: Peters Canyon Regional Park (PECA), Whiting Ranch Wilderness Park (WHRA), and Top of the World/ Aliso and Wood Canyons Wilderness Park (ALWO) (Figure 1). In May 2017, researchers administered a questionnaire with a basic demographic profile to a systematic random sample of visitors with an average response rate of 75% stratified across the main entrance locations of 10 parks in the Reserve system. For a more detailed description of the study site, questionnaire methods and data sources refer to Monz et al. (2019) and Sisneros-Kidd et al. (2019). This survey work suggested that approximately 90% of reserve visitors are carrying a mobile device during their visit. A comparison of the resulting survey dataset and the analysis tools available on the Streetlight Insight platform yielded three demographic variables that could be directly compared between the two approaches: race/ethnicity, average household income, and level of education.

To obtain the demographic data from Streetlight, a Zone Activity Analysis (ZAA) was conducted on the available web browser-based interface (Streetlight InSight) by delimiting polygons in the parking lots proximate to exact locations where visitors were intercepted for the 2017 survey and restricting the analysis to the same dates in May of 2017 when sampling occurred in the three park study areas. The ZAA tool provides an index of visitation (Streetlight Index) and estimates the demographic information of mobile devices that pass through these delineated spatial and temporal zones. The Streetlight data analysis process will account for the number of unique devices in a delineated zone during a particular time, and thus analysis results will typically have varying totals depending on the number of individuals present. The Streetlight algorithm determines the device's likely home location, via the approximate location a mobile device remains between late night and early morning. In order to maintain privacy, Streetlight aggregates device locations to a one-kilometer grid and matches these to the appropriate U.S. Census Block (U.S. Census Bureau, 2020a) to estimate demo-

Figure 1
Study Sites in Orange County, California, USA



graphic data based on 2010 U.S. Census data. The 2017 questionnaire dataset required preparation for analysis by collapsing or merging subcategories of race/ethnicity and education to match the Streetlight /U.S. Census data. Finally, the survey and Streetlight datasets were entered into SPSS (Version 26, IBM Corp., 2017) for statistical analysis. We performed a chi-square test of homogeneity to compare the proportions average household income, and highest level of education between the 2017 survey dataset and the Streetlight output samples for visitors to all three park locations. A Fischer's exact test was performed on race/ethnicity for visitors to all the three park locations due to an inadequate sample size for the Chi-Squared test of homogeneity.

Results

A common finding across the three areas was the significant difference in proportions of the highest level of education between datasets, with the lowest level (high school or less) being markedly different between the two approaches (Table 1). In general, the survey approach returned a lower proportion of visitors in this category compared to the Streetlight demographic analysis. Consequently, the higher education levels tended to be similar or higher on the survey results as compared to Streetlight. In PECA a significant difference was found in comparisons of the proportions of visitor's ethnicity, with the most marked difference being in the proportion of visitors identifying as Latino, with 42.1% in the survey results and 21.1% in the Streetlight approach. No significant differences were found with other comparisons across the three parks in the analysis

Table 1

A Comparison of Visitors' Demographic Information for the Three Parks on Orange County, CA

	Peters Canyon Regional Park			Whiting Ranch Wilderness Park			Aliso and Wood Canyons Wilderness Park		
	2017 Survey	Streetlight		2017 Survey	Streetlight		2017 Survey	Streetlight	
Race or Ethnicity									
Hispanic or Latino	45 (42.1%)	86 (21.1%)		8 (11.6%)	563 (16.2%)		7 (13.2%)	228 (17.9%)	
American Indian or Alaskan Native	2 (1.9%)	2 (0.5%)		1 (1.4%)	12 (0.3%)		1 (1.9%)	6 (0.5%)	
Asian	5 (4.7%)	59 (14.5%)	$x^2_{5,1} = 32.45$, p<.001	8 (11.6%)	447 (12.9%)	$x^2_{5,1} = 6.98$, p=.190	7 (13.2%)	132 (10.4%)	$x^2_{5,1} = 4.86$, p=.383
Black/African	3 (2.8%)	12 (2.9%)		0 (0.0 %)	84 (2.4%)		0 (0.0%)	35 (2.7%)	
White	51 (47.7%)	21 (51.8%)		58 (39.2%)	946 (32.6%)		26 (29.9%)	261 (25.2%)	
Other/Prefer not to respond	1 (2.6%)	37 (9.1%)		4 (5.8%)	345 (10.0%)		4 (7.5%)	132 (10.4%)	
Education									
HS Graduate or less	9(8.0%)	114 (33.5%)		15 (10.1%)	627 (21.6%)		8 (9.2%)	316 (30.5%)	
Some College /Associate Degree	43 (38.4%)	102 (30.0%)	$x^2_{3,5} = 28.30$, p<.001	37 (25.0%)	865 (29.8%)	$x^2_{3,1} = 19.79$, p<.001	22 (25.3%)	287 (27.5%)	$x^2_{3,1} = 29.57$, p<.001
B.A./B.S.	39 (34.8%)	76 (22.4%)		58 (39.2%)	946 (32.6%)		26 (29.9%)	261 (25.2%)	
M.S./Ph.D./J.D./M.D.	21 (18.8%)	48 (14.1%)		38 (25.7%)	464 (16.0%)		31 (35.6%)	172 (16.2%)	
Income									
$35,000 or less	23 (23.0%)	70 (20.6%)		17 (12.1%)	386 (13.3%)		18 (21.7%)	180 (17.4%)	
$35,000-$50,000	11 (11.0%)	33 (9.7%)		10 (7.1%)	229 (7.9%)		6 (7.2%)	97 (9.4%)	
$50,000-$75,000	18 (18.0%)	69 (20.3%)	$x^2_{6,1} = 1.65$, p=.949	17 (12.1%)	497 (16.5%)	$x^2_{6,1} = 8.20$, p=.224	7 (8.4%)	180 (17.4%)	$x^2_{6,1} = 7.46$, p=.280
$75,000-$100,000	14 (14.0%)	39 (11.5%)		19 (13.6%)	406 (14.0%)		16 (19.3%)	136 (13.1%)	
$100,000-$150,000	19 (19.0%)	65 (19.1%)		29 (20.7%)	703 (24.2%)		14 (16.9%)	136 (13.1%)	
$150,000-$200,000	6 (6.0%)	28 (8.2%)		23 (16.4%)	345 (11.9%)		6 (7.2%)	79 (7.6%)	
$200,000 or more	9 (9.0%)	36 (10.6%)		25 (17.9%)	354 (12.2%)		16 (19.3%)	172 (16.6%)	
Total (n)	115	340		151	2903		96	1037	

Note: Within each park and data source values are frequencies followed by percentages, with overall Chi-squared statistics immediately to the right of the two data columns examined.

Discussion

This analysis demonstrates an approach using mobile device data that can provide managers of PPAs demographic data for planning of park facilities, staffing, and long-term monitoring of visitors. Streetlight demographic estimates were, for the most part, consistent with the sample from the 2017 survey for race/ethnicity and income for two of the three parks in this analysis. However, visitors' level of education was significantly different between the survey and Streetlight dataset for the three parks. Visitor demographic data has typically been collected via in-field surveys which involve a considerable amount of time, planning, and funding to staff. The Streetlight platform provides the opportunity to capture this demographic information longitudinally, places less burden on visitors to participate, and provides managers of PPAs the ability to monitor and gather information from visitors who enter PPAs through informal or unofficial entrances (Monz et al., 2019).

Results from this analysis indicate some differences in multinomial proportions between the datasets which could be due to a range of factors, some of which are discussed below. First, the proportions of two of the three demographic variables in PECA were significantly different between the 2017 survey and Streetlight dataset. Sampling of visitors for the post-experience survey to these PPAs was conducted at formal parking entrances, which may have systematically biased our sample to exclude visitors who passed through the Streetlight zone but entered from an informal or non-designated entrance. Previous analyses we conducted using the Insight platform of the differences between visitors entering from the formal and informal at Peters Canyon indicates that visitors entering from the informal entrances and have different temporal use patterns, origin zip codes, and demographic profiles than visitors entering from the formal entrance (Mitrovich & Monz, 2018). This suggests a strength of Streetlight for visitor use monitoring in urban-proximate parks with porous boundaries and where sampling these visitors with traditional visitor intercept surveys is difficult to operationalize and may not gather data on important segments of the visitor population.

These results suggest several sources of error in both approaches. The method using Streetlight data to determine the demographic characteristics of visitors to these PPAs is based on a calculated home location of the mobile device, which is then aggregated to the 2010 Census Blocks. This could falsely identify an individual's home location if they do not work during traditional work hours, and Streetlight has noted this as a limitation of the platform. The demographic characteristics of these Census Blocks in Orange County, CA have likely shifted over time as the area has grown by an estimated 176,000 people since the 2010 Census (U.S. Census Bureau, 2020b). The 2010 Census data may be less representative of the population in areas of dynamic growth and change that has occurred over the decade since the last Census. The forthcoming 2020 Census data may reflect more up-to-date and representative demographic information when these data are integrated into the Insight platform, but the dissimilarities of education profiles observed between the survey and Streetlight could also be explained by visitors to these parks may be more highly educated than their non-visiting neighbors. However, a more systematic issue with Streetlight Data is whether unknown biases exist toward certain socio-economic, racial groups, or visitors' likelihood of carrying a mobile device. We are aware of at least one urban PPA system that is conducting similar analyses with the Streetlight platform and although unpublished, also observed some differences with people of color and lower-income groups in comparisons with

survey-based data (Huting et al., 2019). Regardless, the low proportions in both datasets of minority and less educated visitors underscores the importance of an awareness of the disparities in access and opportunities for recreation in PPAs among minority and less affluent groups. Finally, we acknowledge limitations of the analysis as a result of data preparation. In order to conduct this analysis race/ethnicity, highest level of education, and average household income required response categories be merged or collapsed to maintain consistency of measurement between the two datasets. For example, level of education started with seven groups from the 2017 survey but was reduced to four which resulted in less sensitivity of the test.

Smartphones and the apps visitors use during outdoor recreation experiences have become the focus of a growing body of research to study their effect on visitor behavior and use-patterns. The approach we outlined contributes to this work by demonstrating a novel application of passive mobile device data acquisition and analysis. This process places minimal burden on the visitor, and helps managers of PPAs understand visitor characteristics and trends, and offers potential methodological advantages to traditional on-site visitor intercept surveys.

Disclosure Statement: The authors have no disclosures or competing interests to declare.

Funding: The authors thank the Natural Communities Coalition for providing funding support for this project especially James Sulentich and Danny Fry for extensive advice and expertise. Lead author (Monz) recognizes the Utah Agricultural Experiment Station, The Institute of Outdoor Recreation and Tourism and the USU Ecology Center for ongoing support and assistance.

References

Balmford, A., Green, J. M. H., Anderson, M., Beresford, J., Huang, C., Naidoo, R., Walpole, M., & Manica, A. (2015). Walk on the wild side: Estimating the global magnitude of visits to protected areas. *PLOS Biology, 13*(2), e1002074. https://doi.org/10.1371/journal.pbio.1002074

Center for Demographic Research, California State University, Fullerton. (2020, September). *Orange County jurisdiction demographics.* http://www.fullerton.edu/cdr/demographics/

English, D., & Bowker, J. M. (2018). Introduction to the special issue on visitor monitoring. *Journal of Park and Recreation Administration, 36*(1), ix–x.

Hallo, J. C., Manning, R. E., Valliere, W., & Budruk, M. (2005). A case study comparison of visitor self-reported and GPS recorded travel routes. In *Proceedings of the 2004 Northeastern Recreation Research Symposium: March 31-April 2, 2004* (Vol. 326, p. 172). US Department of Agriculture, Forest Service, Northeastern Research Station.

Huting, J., Roten, L., & Vandegrift, D., (2019, August). Using location-based services data to analyze visitation to regional parks. Presentation to the Metropolitan Parks and Open Space Commission, St. Paul, MN, United States.

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

Korpilo, S., Virtanen, T., Saukkonen, T., & Lehvävirta, S. (2018). More than A to B: Understanding and managing visitor spatial behaviour in urban forests using public participation GIS. *Journal of Environmental Management, 207*, 124–133.

Manning, R. E. (2011). *Studies in outdoor recreation: Search and research for satisfaction.* Oregon State University Press.

Mitrovich, M., & Monz, C. (2018, September 24-26). *A new way to measure park use with big data.* Paper presented at the StreetLight Data Customer Summit, Minneapolis, MN, United States.

Monz, C., Mitrovich, M., D'Antonio, A., & Sisneros-Kidd, A. (2019). Using mobile device data to estimate visitation in parks and protected areas: An example from the Nature Reserve of Orange County, California. *Journal of Park and Recreation Administration, 37*(4).

Muñoz, L., Hausner, V. H., & Monz, C. A. (2019). Advantages and Limitations of Using Mobile Apps for Protected Area Monitoring and Management. *Society & Natural Resources, 32*(4), 473–488.

National Park Service (NPS). (2020, February). *NPS Visitor use statistics.* https://irma.nps.gov/Stats/

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

Sisneros-Kidd, A., D'Antonio, A., Creany, N., Monz, C. & Schoenleber, C. (2019). *Recreation use and human valuation on the Nature Reserve or Orange County, CA*. Utah State University. https://occonservation.org/recreation-management/

U.S. Census Bureau. (2019). *2017 Data Profiles. Data Profiles for Orange County, CA*. https://www.census.gov/acs/www/data/data-tables-and-tools/data-profiles/

U.S. Census Bureau (2020a). *What are census blocks?* https://www.census.gov/newsroom/blogs/random-samplings/2011/07/what-are-census-blocks.html

U.S. Census Bureau. (2020b). *U.S. Census Bureau QuickFacts: Orange County; California; United States.* https://www.census.gov/quickfacts/fact/table/orangecountycalifornia,US/POP010210

Xu, F., Nash, N., & Whitmarsh, L. (2020). Big data or small data? A methodological review of sustainable tourism. *Journal of Sustainable Tourism, 28*(2), 144–163.