Strengths and Limitations of Using UberMedia Smartphone Data To Augment Visitation Studies at OSMP

Brian Anacker* (OSMP; anackerb@bouldercolorado.gov), Colin Leslie* (OSMP; lesliec@bouldercolorado.gov), David M. Theobald* (Conservation Planning Technologies; dmt@davidmtheobald.com)

*Listed alphabetically; All authors contributed equally; contact any author for more information

Draft report: 12/3/21

Final report: 2/22/22

Executive Summary

We analyzed mobility data from UberMedia for two geographic areas ("Chautauqua" and "Bobolink") over 20-months (Jan 2020 – Aug 2021); the dataset included 2,975,539 geographic "pings". Our results are grouped into four areas: (1) Correlations: The number of devices per day were significantly related to daily visits ($P < 0.001$; $r = 0.59$ for Bobolink; $r = 0.85$ Chautauqua). However, the noise around the fit was substantial and concerning. More worrisome was the drastic change-in-over time in the amount of UberMedia data, which would be confounded with any "real" signal in increasing visitation. (2) Along-trail travel (heat map): Patterns of visitation along trails matched our expectations, both in space and time. For example, visitation was busier on the Chautauqua trail and the Royal Arch trail than surrounding trails, and visitation was higher in summer months than winter months. (3) Visitor origin: The percentage of "local" devices at Bobolink was 76%, versus 39% at Chautauqua, matching expectations of relatively lower rates of tourism at Bobolink. (4) Off-trail travel: Within Chautauqua, three possible hot spots of off-trail travel were identified. Moreover, some known undesignated trails were observable in the heat map. While our investigations of on-trail travel, device origin, and off-trail travel revealed patterns that made sense to us, we lacked the ability to correlate these patterns with known estimates, and thus have little confidence in using this information to guide any management choices. Further research and development (more sites, more time series, more calibration) is required before tapping this data to produce a credible, production-level source of information about mobility data (i.e., a "visitor" explorer tool).

Abstract

There is a critical need for data on visitation attributes to manage outdoor recreation but collecting data on the recreational use of parks is resource intensive and hence spatial and temporally limited. Here, we processed mobility data from UberMedia for two geographic areas ("Chautauqua" and "Bobolink") over 20-months (Jan 2020 – Aug 2021); the dataset included 2,975,539 geographic "pings". After a significant amount of processing, we were able to show that the mobile data: (1) describe day-to-day visitation levels, but not perfectly (r = 0.59 for Bobolink; r = 0.85 for Chautauqua); (2) map relative visitation levels among and along trails (i.e., heatmap), not just at access points; (3) show differences in which trails locals and non-locals visit (Percentage of visitors that lived w/in 22 miles of the trailhead was 76% and 39% for Bobolink and Chautauqua, respectively); and (4) locate off trail travel, albeit in a very limited way. We also found some major limitations: (1) the data were too sparse to examine visitation at a particular hour at a specific spot, meaning the data cannot reveal real-time patterns, and (2) year-over-year differences in ping density were large, reflecting changes in cell phone operating systems over time, and likely confounding any real "signal" of annual changes in visitation, and (3) the data could not easily separate use types (e.g., hiking vs. biking), without differentiating by speed of travel. Given these limitations, we suggest that further research and development (more sites, more time series, more calibration) is required before moving this data forward into a credible, production-level mobility data explorer. Meanwhile, Strava Metro introduced a new dashboard tool that, despite its own limitations, may be the more cost-effective direction for OSMP staff than UberMedia.

Keywords: Visitation, big data, mobility data, heat map, outdoor recreation

Introduction

Outdoor recreation is increasing in Colorado. For example, visitation to Rocky Mountain National Park increased from 2.9 to 4.4 million visits over the last 20 years. Municipal open space systems have experienced even greater increases in visitation, due to their proximity to neighborhoods. For example, recreational visits to the City of Boulder Open Space and Mountain Parks (OSMP) lands increased from 3 to 6.25 mission visits over the last 20 years. The trend of increasing visitation has many benefits. Outdoor recreation promotes visitors' mental and physical health (Outdoor Rx Collaborative), provides economic benefits in the hundreds of billions of dollars each year (BEA; Neher 2013; SCORP), and contributes to an ethic of nature appreciation and a community's quality of life (National Geographic). But, increasing outdoor recreation stresses systems that were designed for far fewer visitors. Moreover, increased trail use can facilitate the movement of weeds and pests, cause erosion, damage vegetation, impact water quality, and displace wildlife.

There is a critical need for data on visitation attributes. Moreover, increasing visitation affects areas across a landscape very differently, with some sites being more crowded than others, some sites attracting more tourists or residents than others, and some sites favoring more "nature play" than others. Basic visitation information is the foundation for reliable social and resource indicators. Unfortunately, few land managers measure visitation at the scales that are relevant to adaptive management. Crowdsourced location data may come to our collective rescue.

Crowdsourced location data is information collected from a group of visitors when users turn on location services in smart phone apps, like Google Maps, or track their activities using fitness apps, like Strava. The major advantages of using crowdsourced location data are that it is cheap, it is available for virtually anywhere with few sampling biases. Traditionally, collecting data on the recreational use of parks is resource intensive and hence spatial and temporally limited. Yet, truly massive datasets of location data have been collected from smartphones over the last 20 years. Here, we harness the power of this underused, crowdsourced data gold mine to provide the missing information on visitation that land managers sorely need.

The emergence of “big data” offers a promising additional source of data to generate information about visitor use patterns. The raw data collected by mobile devices (especially cell phones and smart mobile devices) have location, timestamp, and a unique (anonymous) identifier associated with each data point. Mobile data offer a number of advantages to understanding visitation (Wood et al. 2020), including: spatially and temporally comprehensive (not just at trailhead, trail counter, or on-trail, and 24/7/365); consistent information across all areas (not just the limited number of places sampled through traditional means), is conducted in a non-obtrusive manner, and after the initial analytical work, is cost-efficient.

Crowdsourced location data can be used for managing increasing visitation, but its potential is almost completely unrealized. Yet, to move beyond simple graphical displays (such as the Strava Global Heatmap), additional research and development must be conducted to produce credible information to decision makers based on rigorous information. A number of emerging scientific analyses (e.g., Grantz et al. 2020) and applications (e.g., COVID mobility reports) have tapped “big data" from mobile devices to understand visitor use patterns. Much of this work typically tends to be in urban areas and/or along transportation routes, especially roads, and make assumptions that might not readily apply to open space lands (e.g., QA/QC procedures from many data/application vendors assume this; Monz et al. 2020). Moreover, commercial products and vendors typically use proprietary, undisclosed statistical modeling and processing steps (e.g., Monz et al. 2020). Finally, our own preliminary investigations at OSMP have shown mobile data can be used to recover, and expand upon, insights around the "surge" in visitation associated with the stay-at-home order of March 2019 (Fig. 1).

Methods to process and analyze these data need to be appropriate for dealing with locational data that are subject to additional uncertainties that affect their quality, such as poor or uneven reception in rugged terrain and potentially uneven use by mobile phone users. Some emerging efforts are exploring use and comparison of statistical models to estimate use in more remote areas such as parks (Merrill et al. 2020). In addition to the scientific issues, there are pragmatic considerations as well, as these "big" data are indeed big. The volume of data can make it challenging to process with typical computational approaches. For example, analyzing and visualizing data using common desktop GIS or statistical software on a typical stand-alone workstation tends to be constrained to a few million features, whereas typical data volumes of the mobile data are in the billion to trillion range of features. Larger cloud-based solutions may decrease the computational overhead.

The primary goals of the research presented here are to test, validate, and estimate visitor use patterns by analyzing and mapping mobile data. By mapping and analyzing "big data" from mobile devices, managers and decision makers will be able to better understand where visitors go (both on and off-trail usage) and when they do so (seasonal, weekly, and daily usage patterns). This information will inform decision making around resource management, protection priorities, and collaboration with land management partners.

Methods

Data sources

We used data provided by UberMedia using their Vista report generator interface (https://vista.um.co/). After a sales call, UberMedia provided 10 free credits for us to use for free. Future credits were quoted at $75 per credit. The geographic and temporal limitation on the query size allowed per credit were untested, but UberMedia said such limits existed. Note, Vista has a toggle to allow queries to include more than 2 million ft^2 (~46 acres), which we turned on, but it is unclear if geographic extents above this size, or some other threshold, would be more expensive. Using Vista, you enter the parameters to query and the system returns a report in a matter of minutes, at the cost of one credit per query/report. We used these credits to download data from 2020 and 2021 for two different geographic extents (Map 1). These polygons were drawn by hand, on the fly, in the Vista software. We later created an additional shape in ArcGIS to “mask” the Chautauqua area more finely, largely to exclude baseline road and the parking lot (more below). As part of our report configuration, we selected the Common Daytime Location/Common Evening Location (CDL/CEL) add on, which provides the coordinates were each unique mobile device rests during the majority of the day (CDL; i.e., where the device works) and where it rests in the evening (CEL; i.e., where the devices lives/sleeps). We ran the queries over several months as we iterated on our R&D. The results that are presented in this report represent a combined ping dataset of 2,975,539 points. We also pulled a query of earlier years (starting Jan 2017) to see how data availability changes over time for Chautauqua; this query included 3,640,497 pings.

Data wrangling

An initial evaluation of the data was completed in Google Earth engine (Appendix 1). Inferences on the spatial accuracy and temporal resolution of the ping data were taken forward into a future work session completed in R, and described below.

As part of our pre-analysis planning, we decided that we wanted all R scripts to be self-contained and standalone from all other analysis script. To support this, we performed some initial data wrangling in order to establish a common, pre-formatted dataset, from which all the subsequent analyses scripts would begin. The scripts and data for this project are stored here: https://cityofboulder.sharepoint.com/sites/OSMPMobileDataProject/Shared%20Documents/Forms/AllItems.aspx (accessible only to City of Boulder staff).

The ping data from the UberMedia Vista system was the largest in terms of number of records and were exported from the Vista system as four separate tab separated data files (compressed as a Zip archive). Using R, we read in the four data files and then appended them into a single dataset. After a few adjustments to set data types, arrange the data, and filter to January 1, 2020 and later, we saved the data as an RData object, allowing us to read in a preformatted data object at the start of each analysis script. The CEL/CDL and Demographics datasets were exported from Vista as single tab separated datasets and did not need pre-processing before being used in subsequent analysis scripts.

During exploratory analysis, we discovered that major roadways adjacent to our study area (e.g., Baseline road next to Chautauqua) captured large numbers of pings from devices that were not associated with the recreational area of interest (i.e. vehicles moving on the roadways) t. To minimize the effects of pings from travel corridors such as this, we drew several analysis polygons by heads up digitizing them in ArcGIS. For the Chautauqua study area, our largest polygon excluded all roadways but included the Ranger Parking Lot. We also created two additional polygons for Chautauqua, one that excluded the Ranger Lot (i.e. only included trails) and one that applied a 100m inward buffer (to filter out potential edge effects). The geographic extent for Bobolink was drawn carefully to avoid roads and parking areas (Map 1), and thus did not need any geographic filtering after download.

We then used these analysis polygons to clip the data, which improved both analysis processing time and subsequent visualization by only retaining data with the area of interest.

Data analysis

1. Correlations

The Chautauqua and Bobolink area have permanently installed trail counters which count pedestrians in both directions in 15 minute increments. One of our initial research questions was to see how well the UberMedia ping data and trail counter data correlated. We explored correlations at multiple temporal and spatial scales and ultimately decided to analyze counts at the daily level.

The first step was to clip the ping data to the study area. For the Chautauqua area, we clipped the ping data using the most extensive analysis polygon, which included the Ranger Lot but excluded roadways. Devices can ping anywhere from one to hundreds of times within the study area on any given day. To get a better representation of unique visits (more in line with what the trail counters collect), we grouped the ping data by day and then counted the distinct device IDs present in the study. We built a linear model using the estimate of daily visits from the trail counter as the dependent variable and number of distinct device ids as the independent variable (Fig 2b).

The resulting terms from the linear model were then used to weight the daily count of device ids to produce a continuous time series of estimated visits at the trail counter. Trail counter data and the weighted device id counts were then overplotted to visualize daily and seasonal patterns (Fig 2a).

2. Along-trail travel (heat map)

Travel statistics along trail corridors were determined by rasterizing the ping data, with cell values representing a count of the number of pings that fell within the cell. Because different devices ping at different rates, with even the highest frequency devices pinging 10s to 100s of meters apart, we were limited in how fine of a temporal resolution we could extract from the data. Travel corridors were easily revealed (and confirmed by overlaying trails from OSMPs GIS) when analyzing ping data from the entire year.

One important consideration we had to deal with when rasterizing the data was not just the spatial distribution but how that interacted as a function of temporal distribution. Clear spatial patterns remained at the monthly level, but daily analysis proved to be too sparse to reveal all but the most heavily used trails. However, some finer temporal resolution analysis, such as hourly patterns, were possible so long as the hour of the day was aggregated across many weeks or months.

To create the raster layers, we defined an empty raster grid using the extent of all ping data. We tested multiple pixel resolutions but found that about 8 meters (square) yielded a good balance between being large enough to summarize ping densities consistently along a trail segment (as we expected values to remain relatively consistent along a segment) but small enough to reveal individual trail corridors.

The last step for most raster analyses was to apply a raster filter to remove any cells with values below a certain threshold. From our exploratory analysis, we determined that ping points could occasionally have high spatial error. Removing cells with counts below a certain threshold was one of the simpler methods to remove potential effects from erroneous ping data, while also improving readability of the final raster data when plotted visually.

3. Visitor origin

In addition to the point data provided by the ping dataset, UberMedia was also able to provide related data, summarized by device id. One such dataset was the common evening location, common daytime location (CELCDL) dataset. Based on where the device is commonly located at night, UberMedia can derive a number of additional variables including the general residence location (town, city, county) and associated distance of the common evening location to the analysis area. Using the device id as the matching variable, we joined these data to the ping dataset, resulting in a new ping dataset that included additional information about the device owner (aka visitor).

As we described above, by calculating the density of ping points across the analysis area, we were able to reveal travel corridors based on contiguous patterning of high value cells. Sub-setting the ping data based on newly joined variables, such as for a particular residency class or other category allowed us to create separate raster layers based on these variables. We then used these sub-set raster layers, in conjunction with the initial raster of all ping data to perform raster calculations. For example, by creating one raster of ping counts from all devices, and another raster of ping counts from only devices with a common evening location within a certain proximity of the analysis area, we were able to divide the “proximate” raster by the “all” raster to derive a proportional raster representing the percent of visits to an area from within a certain travel distance.

4. Off-trail travel

We analyzed the ping data for the presence of undesignated or off trail travel use. When contiguous lines of higher density raster cells were present, but designated trail features did not intersect with the raster lines, this suggested the presence of undesignated trail routes.

Furthermore, if after filter the final raster output for cells with counts that fell below a certain threshold (typically 2 or 3 ping counts) cells or groups of cells remained that were not associated with a travel corridor, this suggests potential off-trail destinations.

Results

1. Correlations

The number of devices per day were significantly related to daily visits ($P < 0.001$; Fig. 1, 2).

2. Along-trail travel (heat map)

Patterns of visitation along trails matched our expectations, both in space and time. For example, visitation was busier on the Chautauqua trail and the Royal Arch trail than surrounding trails (Fig. 4), and visitation was higher in summer months than winter months (Fig. 5). We don't, however, have any data to validate the along-trail travel patterns against.

3. Visitor origin

At Chautauqua, "Local" devices tended to be more prevalent away from the main Chautauqua trail and Royal Arch trail, where "Regional" and "Other" devices were more common (Fig. 7); likewise, Local devices were relatively common in mornings and evenings and winter months, while Regional and Other devices were relatively common in midday times and summer months. While these results are qualitatively in line with expectations from previous visitor surveys, it is difficult to know how to evaluate the statistical significance of strength of these results on visitor origin. These analyses were not repeated for Bobolink.

4. Off-trail travel

Within Chautauqua, three possible hot spots of off-trail travel were identified. Moreover, some known undesignated trails were observable in the heat map (Fig. 4). These analyses were not repeated for Bobolink.

Discussion

The UberMedia data was significantly related to visitation counts, but the noise around the fit (despite our best analytical efforts) was substantial and concerning. More worrisome was the drastic change-in-over time in the amount of UberMedia data (Fig. 9), which would be confounded with any “real” signal in increasing visitation, in both an absolute and relative sense. Likewise, our investigation of device origin and off-trail travel, although cursory at best, revealed patterns that made sense, but we lacked the ability to correlate these patterns with known estimates, and thus have little confidence in using this information to guide any management choices.

All four of our investigative areas (correlation with trail counts, heat maps, visitor origin, off trail travel) beg for a continued and substantial research and development investment by OSMP to help sharpen the inferences and move towards a credible, production-level, web-based mobility data explorer. At the minimum, OMSP would need to evaluate more sites and more time windows to better understand the signal in the mobility data and build area- and time-specific calibrations.

At this time, however, we recommend that OSMP staff and partners do not make this investment for a couple of reasons. First, the mobility data market appears to be highly volatile. Cell phone location services are changing as new laws and software are put in place, drastically changing the availability of device location data. Another volatility is that the cottage industry of location-services companies is changing rapidly as companies are bought and sold. Thus, any products OSMP would make today may not work in the future (i.e., data set dries up or the company folds). If OSMP wants to continue with R&D, it may be better to wait say 3-5 years for the market to settle down. Second, a huge number of analytical choices need to be made in the processing of the mobility data, and this complexity may be best left to the companies dedicated to this endeavor, assuming they are transparent about their choices. Third, the inference gains that we are able to make seem marginal, and/or easily recovered using other 3rd party heat maps.

A solid, cost-effective alternative may be the recently released Strava Metro. Our initial exploration suggests that this new dashboard tool shows reasonable signal in visitation patterns (i.e., heat map) along OSMP trails, but has a lot of other nice features: (1) distinguishes among hikers vs bikers, (2) is rendered for the entire county, allowing comparison with BCPOS and USFS trails, (3) can separate locals from tourists (because it knowns where each device sleeps at night), and (4) has the ability to look at specific time windows. Moreover, this dashboard appears to be free. Of course, a major limitation of Strava is that its location data comes from Strava users, which may skew heavily towards fitness-focused men in their 30s (personal observation). Thus, Strava Metro may be the more cost-effective direction for OSMP staff than UberMedia, but biases in the population being sampled should be evaluated. We suspect that the Strava Metro team and/or researchers in this field have already done this.

Literature cited

Grantz, K.H., Meredith, H.R., Cummings, D.A.T. et al. The use of mobile phone data to inform analysis of COVID-19 pandemic epidemiology. Nat Commun 11, 4961 (2020). https://doi.org/10.1038/s41467-020-18190-5

Neher C, Duffield J, Patterson D. Valuation of national park system visitation: the efficient use of count data models, meta-analysis, and secondary visitor survey data. Environ Manage. 2013 Sep;52(3):683-98. doi: 10.1007/s00267-013-0080-2. Epub 2013 May 29. PMID: 23716008.

Merrill NH, Atkinson SF, Mulvaney KK, Mazzotta MJ, Bousquin J (2020) Using data derived from cellular phone locations to estimate visitation to natural areas: An application to water recreation in New England, USA. PLoS ONE 15(4): e0231863. https://doi.org/10.1371/journal.pone.0231863

Monz, C., Mitrovich, M., D'Antonio, & 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. doi:10.18666/JPRA-2019-9899.

Wood, S.A., Winder, S.G., Lia, E.H., White, E.M., Crowley, C.S.L., Milnor, A.A. 2020. Next-generation visitation models using social media to estimate recreation on public lands. Scientific Reports 10:15419.

Tables

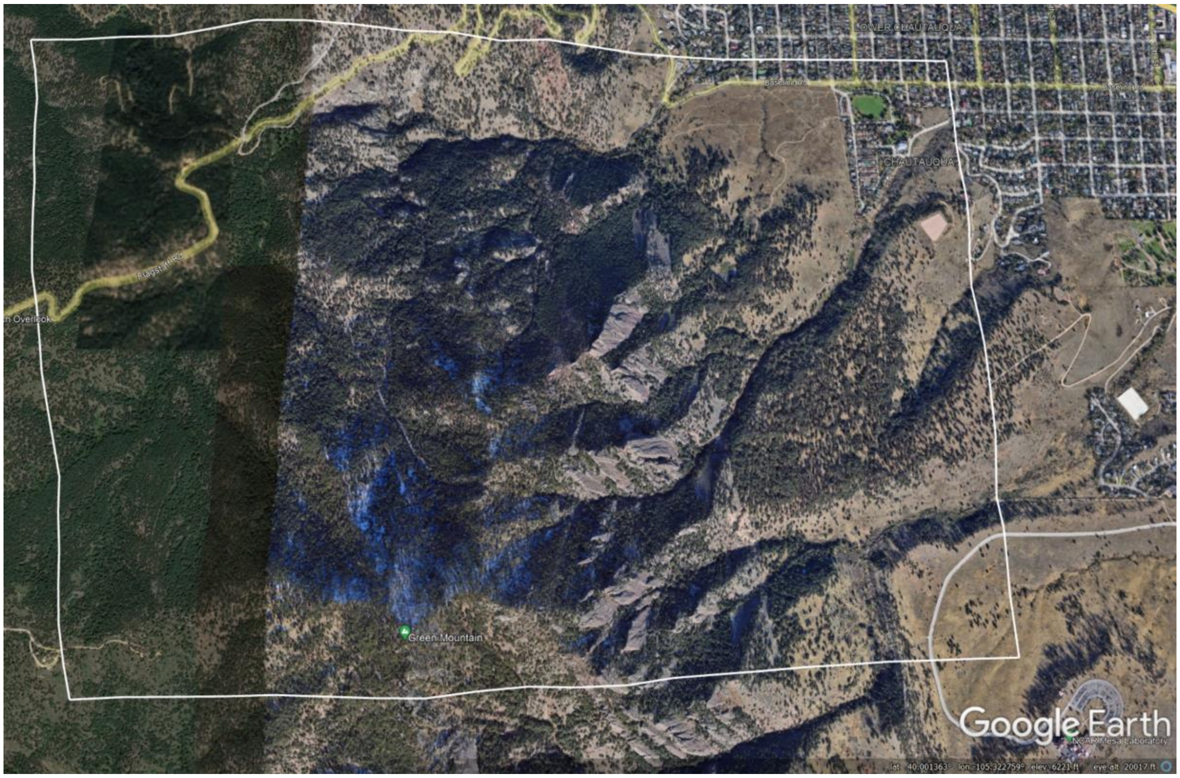
Table 1. Device origin (where the device regularly resides during their overnight hours) by area (Chautauqua and Bobolink). $X^2 = 28.0$, $P < 0.0001$

Group	Name	Chautauqua %of Devices	Bobolink % of Devices
<22 miles from home	Local	39%	76%
22-150 miles from home	Regional	22%	9%
>150 miles from home	Other	39%	15%

Table 2. Month-over-month patterns by device origin (where the device regularly resides at night) at Chautauqua. Numbers are proportions of devices.

	local	regional	other
Jan	0.4251572	0.2918239	0.2830189
Feb	0.4709091	0.2490909	0.2800000
Mar	0.4534161	0.3198758	0.2267081
Apr	0.5381295	0.3553957	0.1064748
May	0.3691932	0.4914425	0.1393643
Jun	0.3921035	0.3839346	0.2239619
Jul	0.3324675	0.3318182	0.3357143
Aug	0.3464516	0.3187097	0.3348387
Sep	0.4553506	0.2464945	0.2981550
Oct	0.4225466	0.2287105	0.3487429
Nov	0.4414020	0.3088719	0.2497262
Dec	0.4703947	0.2878289	0.2417763

Maps

a. 

b. 

Map 1. Two geographic extents examined in the study (white rectangles). Chautauqua (a), measuring 1,881 acres in area, and Bobolink (b), measuring 37 acres in area. Not presented at scale (i.e., Bobolink is relatively smaller than it appears).

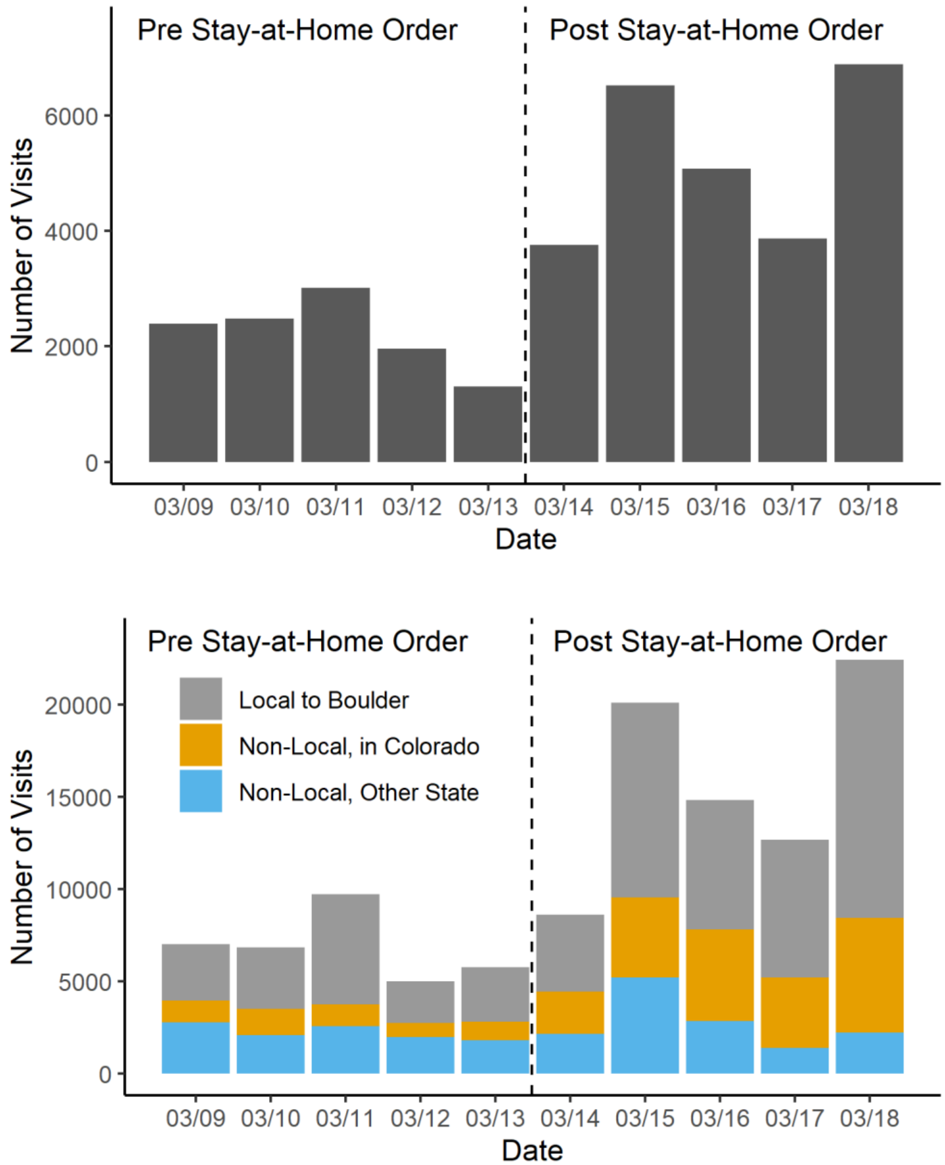
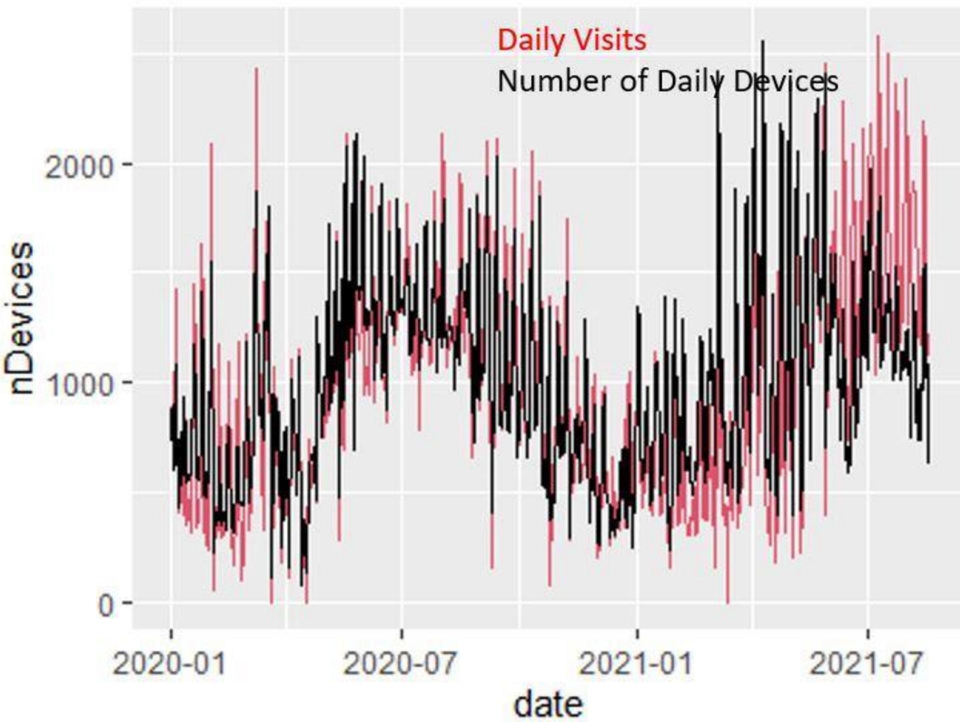


Fig. 1. Large increases in visitation Chautauqua driven by the Stay-at-Home order in Boulder, CO. Top panel shows data collected on-site using trail counters. Bottom panel shows mobile data. The mobile data recovers the same signal of increased visitation as the trail counters, and the mobile data adds a new insight on where visitors are from. Note, this data was sourced from Unacast for March of 2020; the remainder of this report is based on UberMedia (a competitor of

Unacast) from 2020 and 2021. This early result encouraged us to investigate mobile data in more depth.

a.



b.

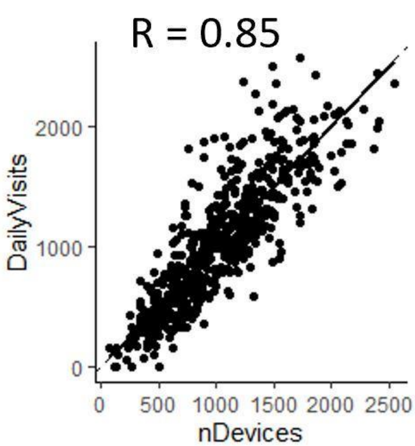
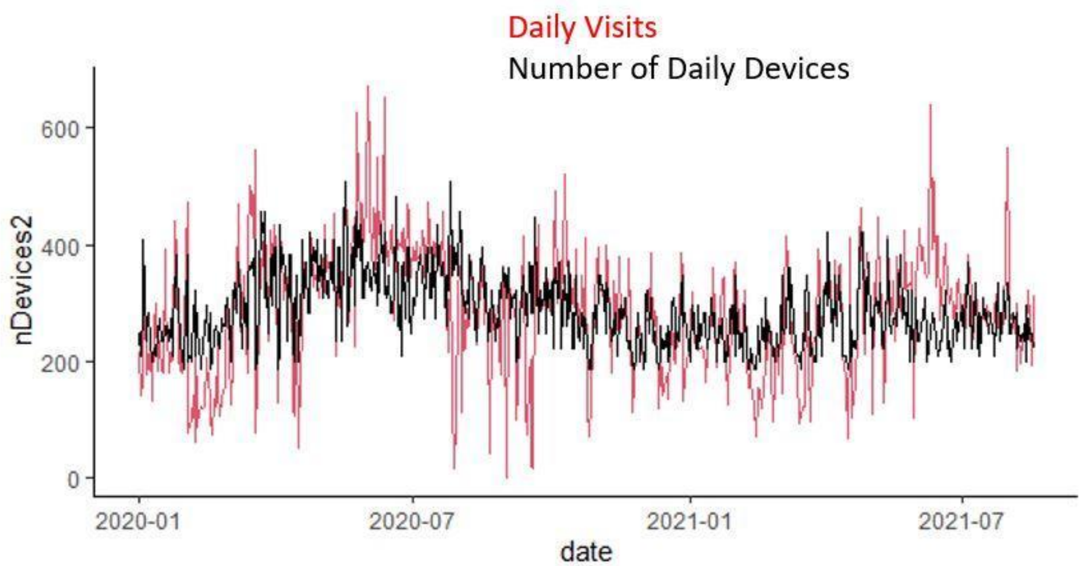


Fig. 2. Trend in the number of daily devices (black line, data from mobile phones) and daily visits (red line, data from trail counters) over a 20-month period (Jan 2020 – Aug 2021) at

Chautauqua (a). Correlation of the number of daily devices (y axis) and daily visits (x axis) for the same dataset (b).

a.



b.

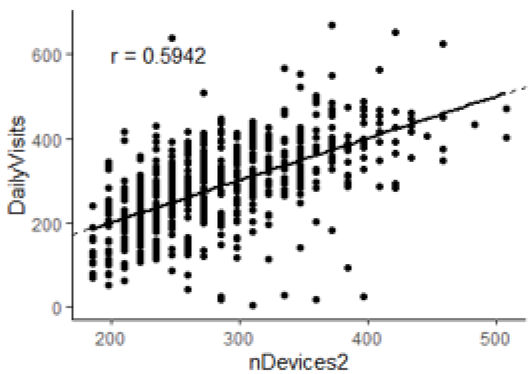
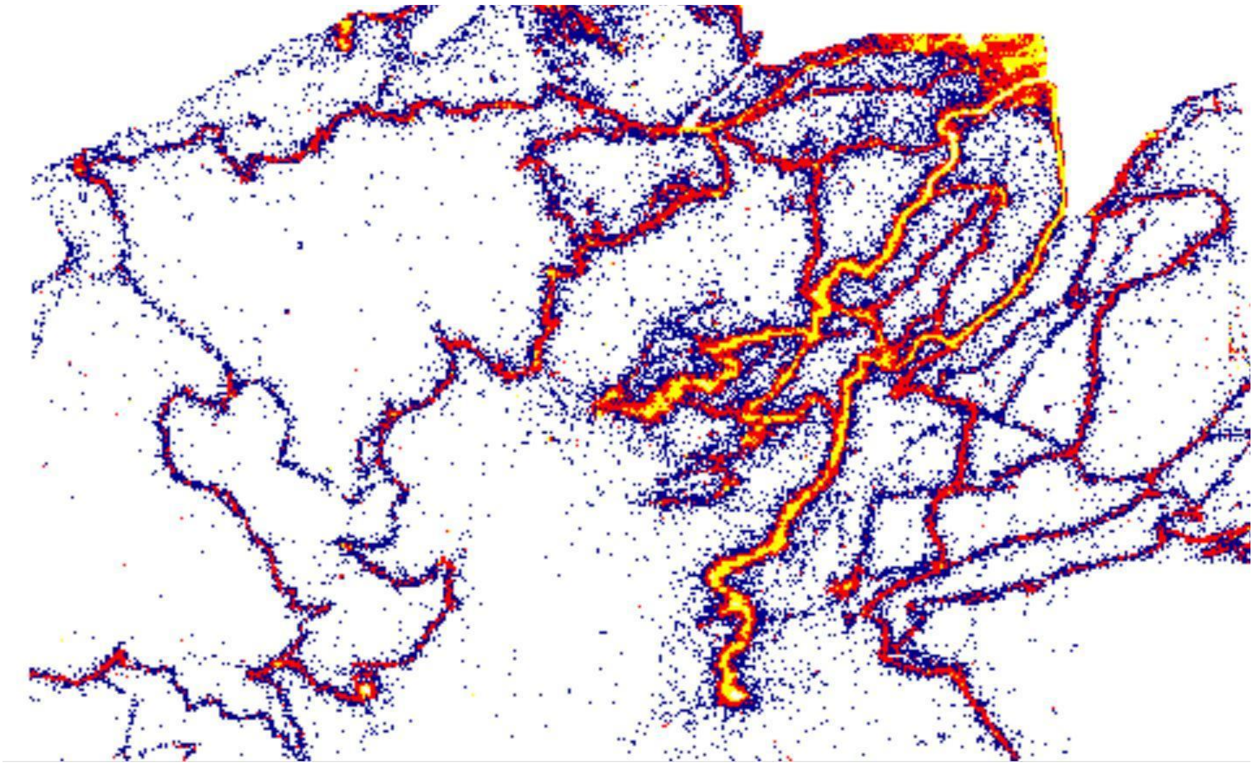


Fig. 3. Trend in the number of daily devices (black line, data from mobile phones) and daily visits (red line, data from trail counters) over a 20-month period (Jan 2020 – Aug 2021) at Bobolink (a). Correlation of the number of daily devices (y axis) and daily visits (x axis) for the same dataset (b).

a.



b.

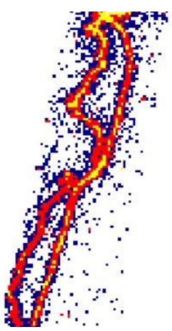


Fig. 4. Heat map of visitation for the Chautauqua (ab) and Bobolink (b) areas. Bright yellow areas are higher use, red medium, and blue low use. Raster cells are ca. 25 m^2 (5m x 5m)

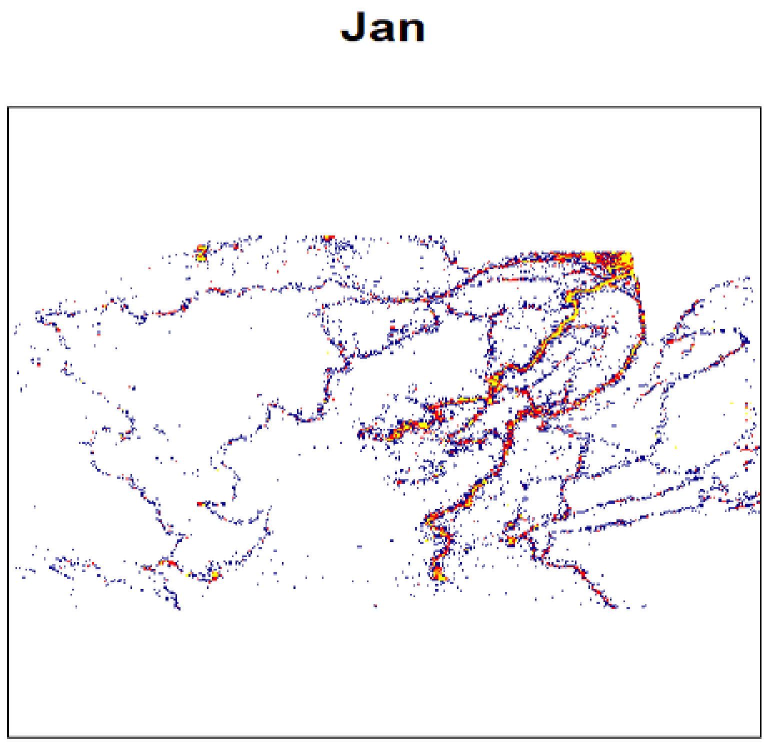


Fig. 5. Animated heat map of month-over-month visitation at Chautauqua.

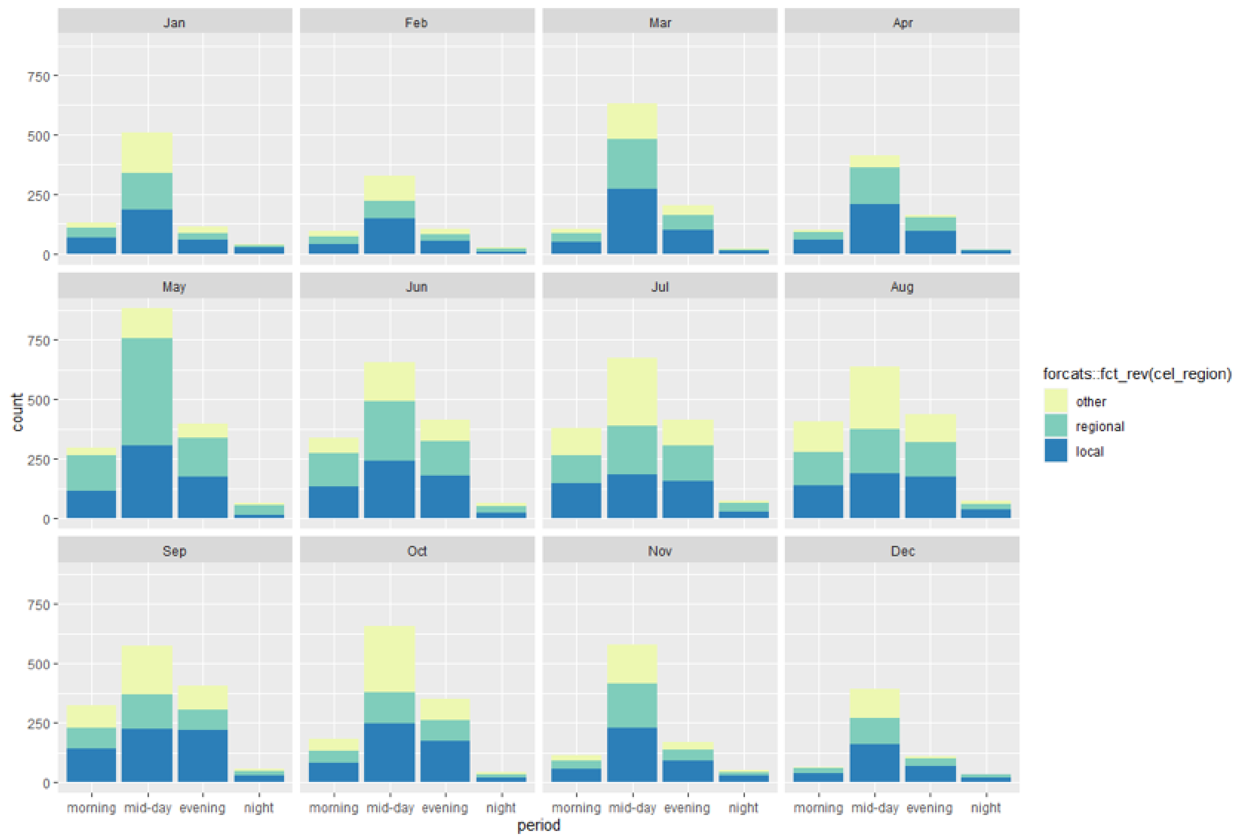


Fig. 6. Visitation (based on mobile data) by month (facets), time of day (x axis) and device origin (where the device regularly resides at night; colors) at Chautauqua.

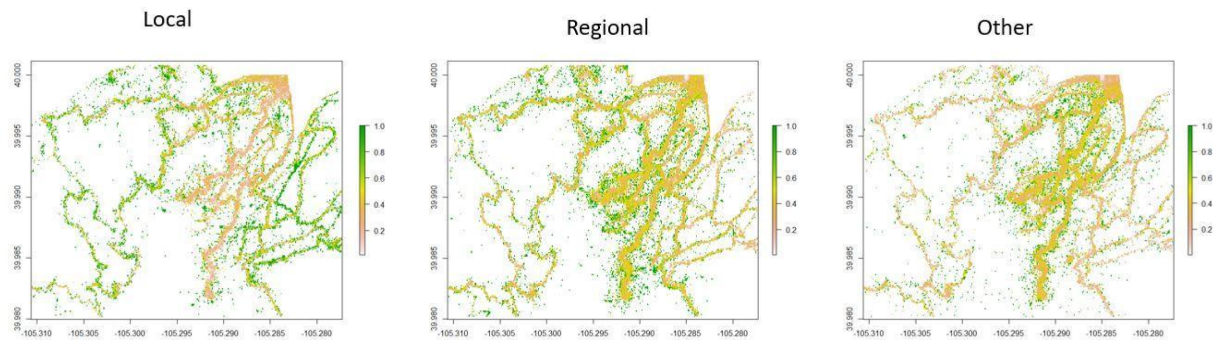


Fig. 7. Heat maps of visitation (based on mobile data) by device origin (where the device regularly resides at night).

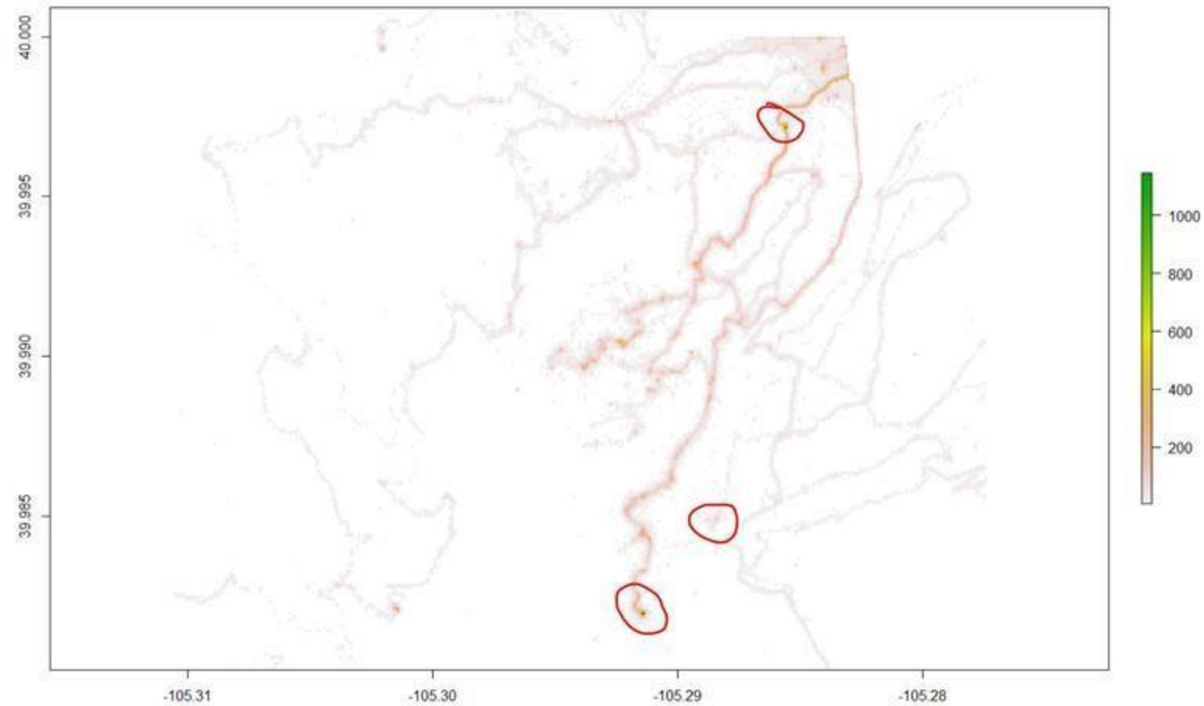


Fig. 8. Three possible locations of off-trail hotspots.

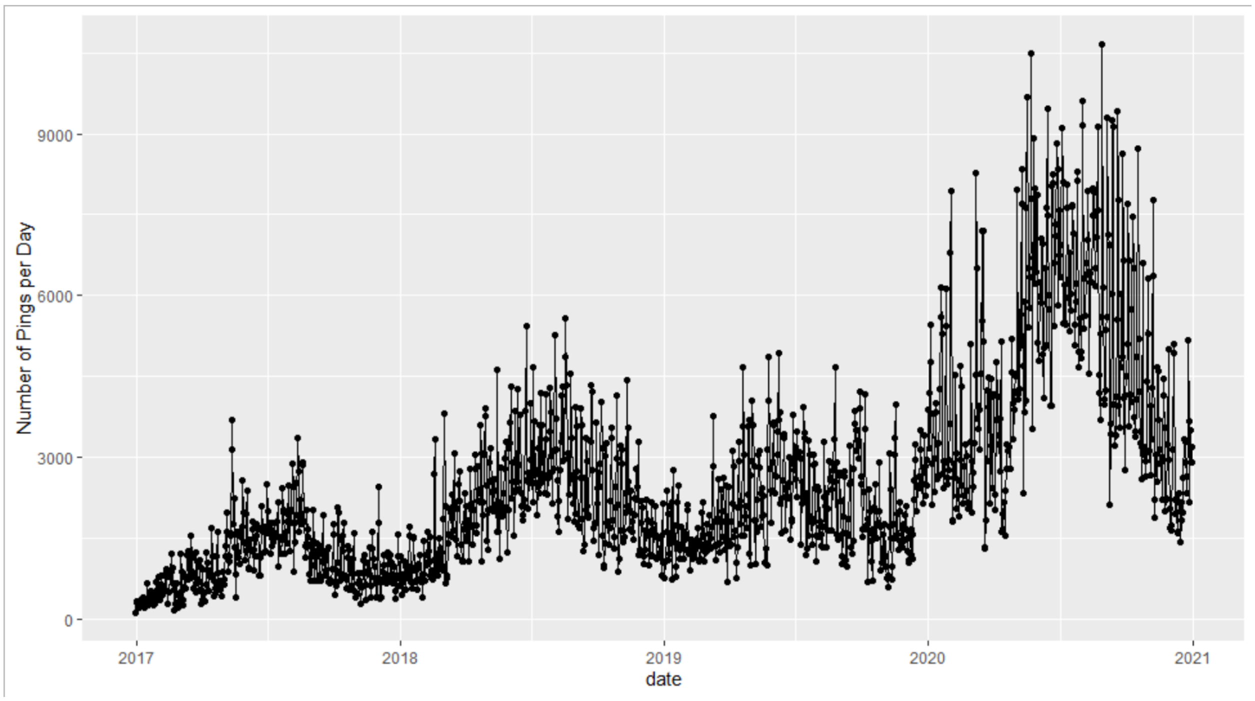


Fig. 9. Trend in the amount of data available from UberMedia at Chautauqua. Note, this date range (Jan 2017 – Dec 2020) differs from the date range used in the rest of the report (Jan 2020 - Aug 2021).

Appendices

Appendix 1. Initial data exploration in Google Earth Engine completed by Dave Theobald.

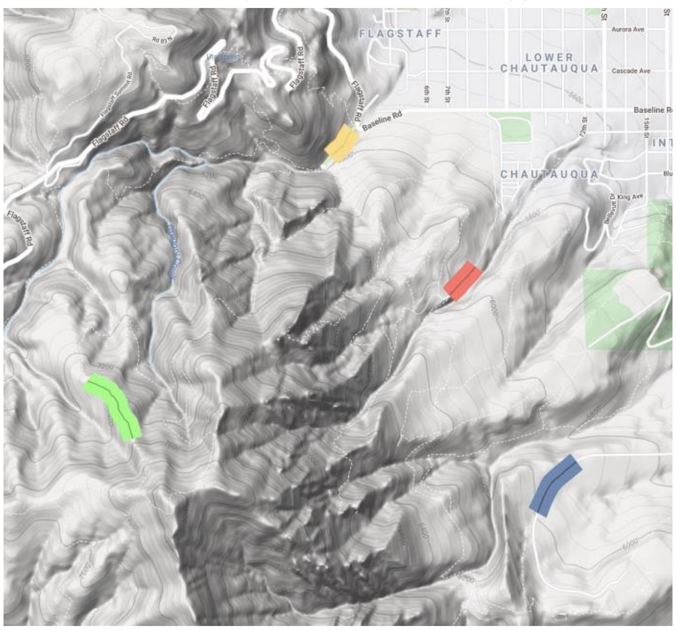
NOTES DMT 20210621

Filtering (DMT GEE script path; URL)

- Date: UberMedia data from 2018-06-01 to 2020-12-31
- Spatially - to the Chautauqua interior polygon

Estimating spatial precision of pings, particularly valuable for understanding reasonable pixel size for heat maps (GEE script path; URL)

- Identified 4 different locations that would be roughly representative of different
- Heads up digitizing of road/trail segment, buffered by 50 m
- terrain/vegetation contexts and cell tower reception
- Locations:
    - NCAR road: open terrain, close to towers, many pings
    - Gregory Canyon Road: canyon, sheltered, many pings
    - BluebellRoad:
    - Ranger Trail: sheltered, rugged, obscured(?) from cell towers, fewer points



- For each ping, get the distance (m) from the center-line of test area, within the 50 m buffer
    - Summary results, settled on using 2 SD as rough estimate:
    - Context,count,mean,SD,precision68,precision95,95th
    - NCAR,23201,3.77,5.6,9.37,14.97,11.36
    - Greg,11927,5.40,8.12,13.52,21.64,22.09
    - Bluebell,11863,6.51,7.69,10.11,19.15,22.35
    - Ranger,6.44,7.09,13.53,20.62,37.87

- Approximate median of the estimate at 95% is 21 m

Heat maps (GEE script path; URL)

- Each ping assigned “population” value of 1…
- Kernel density (gaussian) of 20 m radius was applied to set of pings
- Normalized by the number of days to enable rough comparisons to other time-slices

UserId-days

- Summary stats: 8465 unique id-days in TC buffer, 1.65 50%, 5 at 75%, 9.2 at 90%, 13.3 at 95%, 45.6 at 99%, and 643 at 100%
- Unique Id Days # and percentiles:
- 8465
- JSON
- Object (6 properties)
- p100:
- 643
- p50:
- 1.655577623303542
- p75:
- 5.008421052631579
- p90:
- 9.198966408268733
- p95:
- 13.345549738219896
- p99:
- 45.588235294117645
-

----------- summary stats of spatial precision mini analysis

NCAR sample

JSON

Object (13 properties)

max:

49.923804633259344

mean:

3.7691435082557825

min:

0.001894416205393969

sample_sd:

5.597131357335474

sample_var:

31.327879431268045

sum:

87447.8985350424

sum_sq:

1056410.481879384

total_count:

23201

total_sd:

5.597010733403963

total_var:

31.32652914983917

valid_count:

23201

weight_sum:

23201

weighted_sum:

87447.8985350424

JSON

Gregory Canyon sample

JSON

Object (13 properties)

max:

49.9462035903996

mean:

5.396822282792743

min:

0.002274814775111783

sample_sd:

8.121943900947024

sample_var:

65.96597273013056

sum:

64367.89936686905

sum_sq:

1134092.3043792169

total_count:

11927

total_sd:

8.12160340819315

total_var:

65.9604419199746

valid_count:

11927

weight_sum:

11927

weighted_sum:

64367.89936686905

JSON

Bluebell sample

JSON

Object (13 properties)

max:

49.589553491836696

mean:

6.513914174989794

min:

0.004989799965270549

sample_sd:

7.690482968225195

sample_var:

59.14352828456181

sum:

77274.56385790392

sum_sq:

1204920.4093916265

total_count:

11863

total_sd:

7.690158824035541

total_var:

59.1385427388917

valid_count:

11863

weight_sum:

11863

weighted_sum:

77274.56385790392

JSON

Ranger sample

JSON

Object (13 properties)

max:

49.09348430448336

mean:

6.441635985506967

min:

0.0040294711413162215

sample_sd:

7.092297402373241

sample_var:

50.30068244371022

sum:

17431.066976781854

sum_sq:

248347.9343136563

total_count:

2706

total_sd:

7.090986805041092

total_var:

50.28209386926687

valid_count:

2706

weight_sum:

2706

weighted_sum:

17431.066976781854

-------------------------- Quick summary stats for filtering ----------

Running from 2018-06-01 to 2020-12-31

JSON

all pings

3640497

JSON

Chat. interoir # pings and unique user IDs:

1980549

68483

JSON

pings and IDs interior:

1424532

47199

JSON

pings and IDs in TCChat segment:

39282

8465

JSON

pings and IDs TCChat hour: 0

42

22

JSON

pings and IDs TCChat hour: 1

60

22

JSON

pings and IDs TCChat hour: 2

64

34

JSON

pings and IDs TCChat hour: 3

79

46

JSON

pings and IDs TCChat hour: 4

184

77

JSON

pings and IDs TCChat hour: 5

554

178

JSON

pings and IDs TCChat hour: 6

999

248

JSON

pings and IDs TCChat hour: 7

957

340

JSON

pings and IDs TCChat hour: 8

1313

472

JSON

pings and IDs TCChat hour: 9

2037

764

JSON

pings and IDs TCChat hour: 10

2654

1057

JSON

pings and IDs TCChat hour: 11

4168

1311

JSON

pings and IDs TCChat hour: 12

4160

1291

JSON

pings and IDs TCChat hour: 13

4314

1252

JSON

pings and IDs TCChat hour: 14

4259

1214

JSON

pings and IDs TCChat hour: 15

3343

1004

JSON

pings and IDs TCChat hour: 16

3113

865

JSON

pings and IDs TCChat hour: 17

2610

629

JSON

pings and IDs TCChat hour: 18

1736

501

JSON

pings and IDs TCChat hour: 19

1036

335

JSON

pings and IDs TCChat hour: 20

738

197

JSON

pings and IDs TCChat hour: 21

435

95

JSON

pings and IDs TCChat hour: 22

112

39

JSON

pings and IDs TCChat hour: 23

300

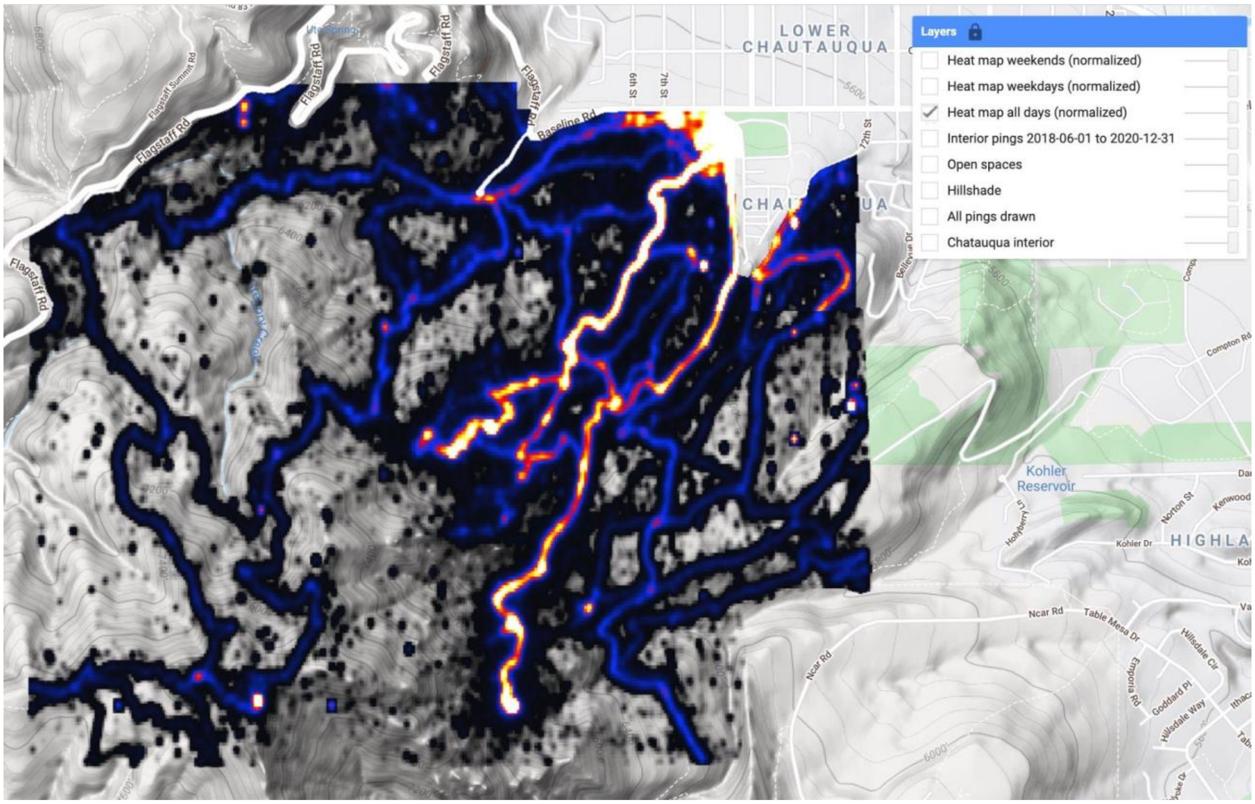
28

-- end summary stats

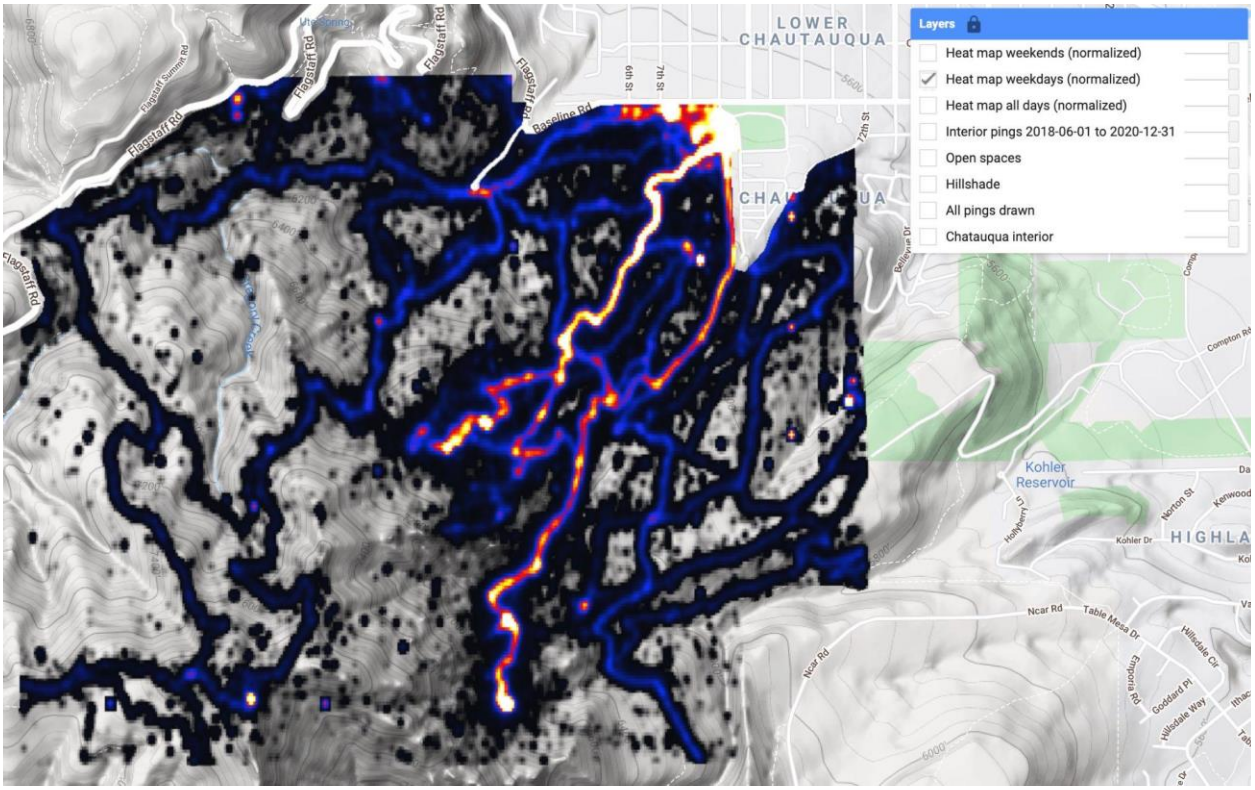
- Filter by Chaut. Trail counter buffer of 21 m
-

Some quick screen snapshots

All dates, all days

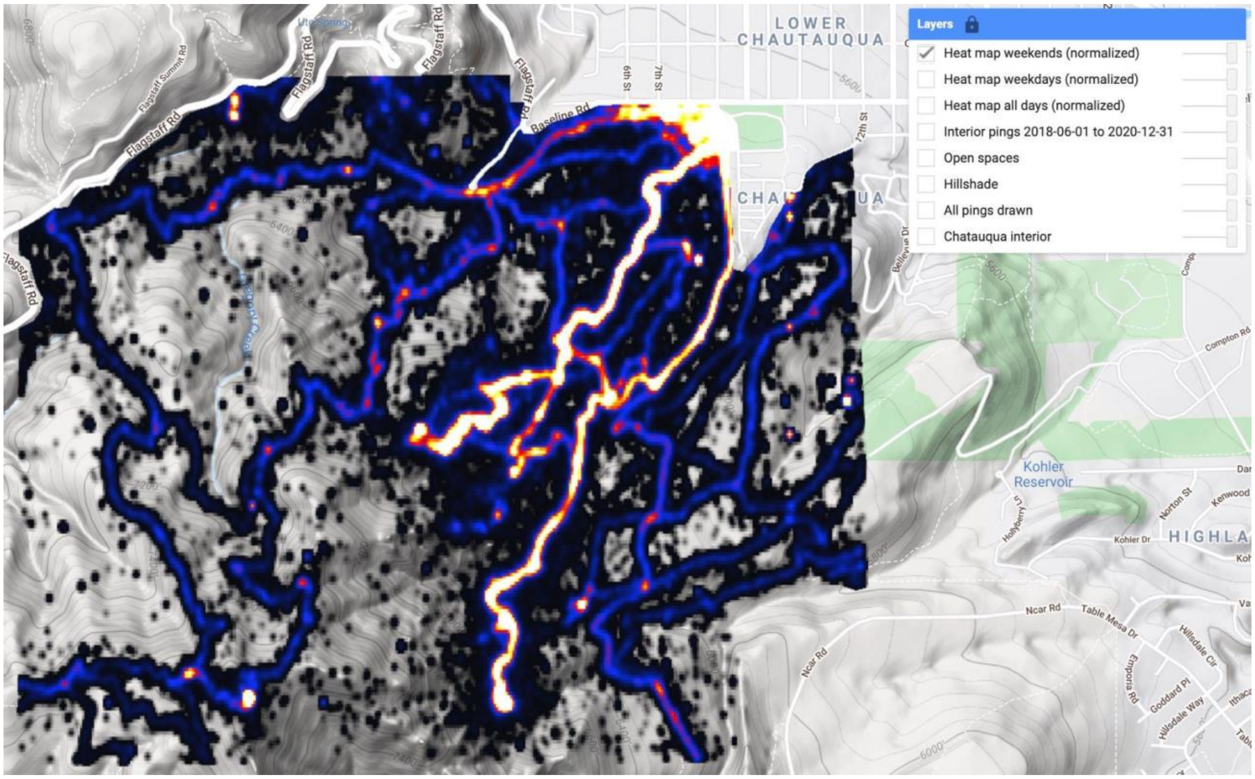


Weekdays all dates

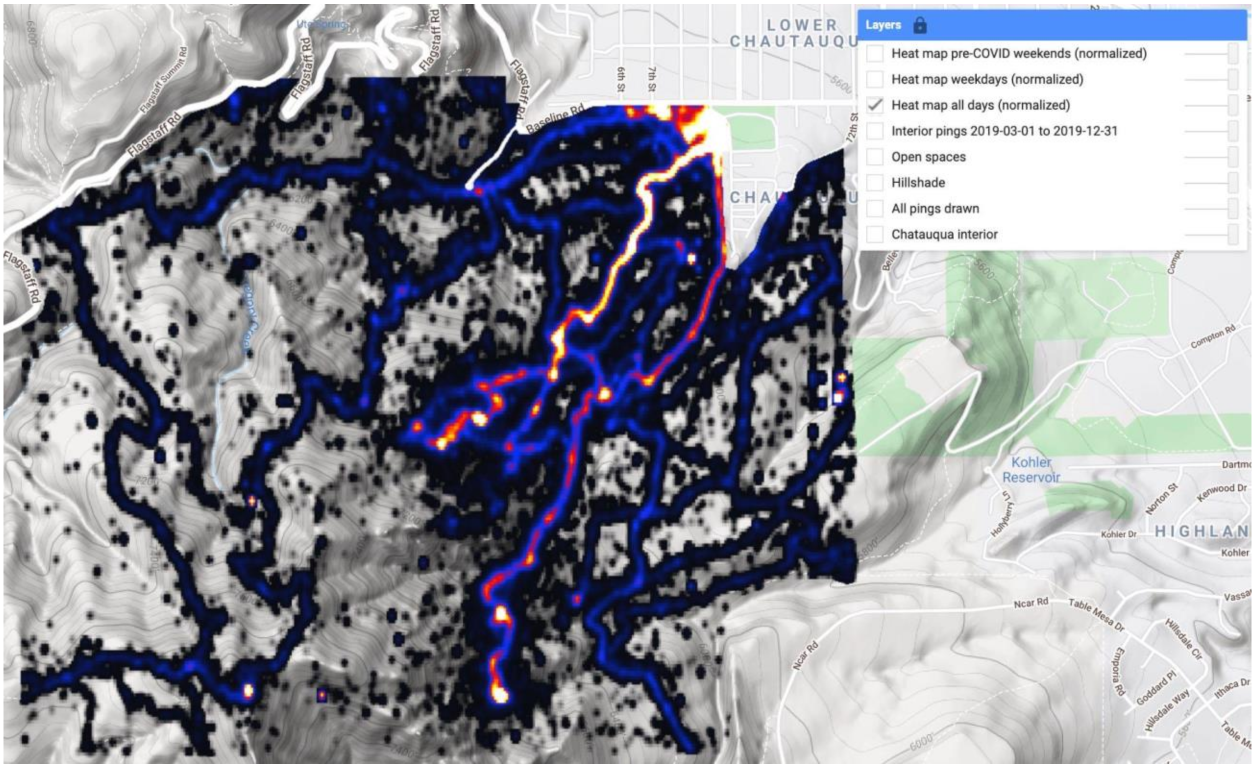


Layers
Heat map weekends (normalized)
Heat map weekdays (normalized)
Heat map all days (normalized)
Interior pings 2018-06-01 to 2020-12-31
Open spaces
Hillshade
All pings drawn
Chatauqua interior
LOWER CHAUTAUQUA
Flagstaff Rd
Baseline Rd
Kohler Reservoir
Ncar Rd
Table Mesa Dr

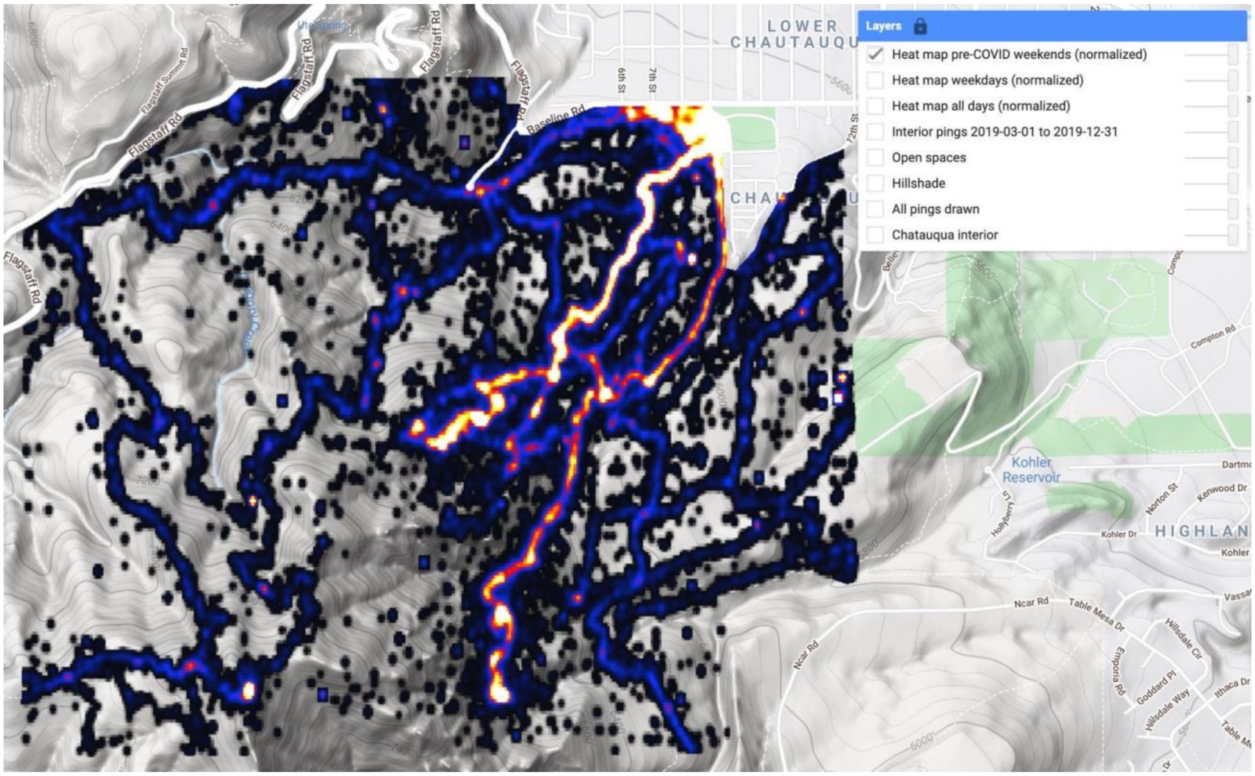
Weekends all



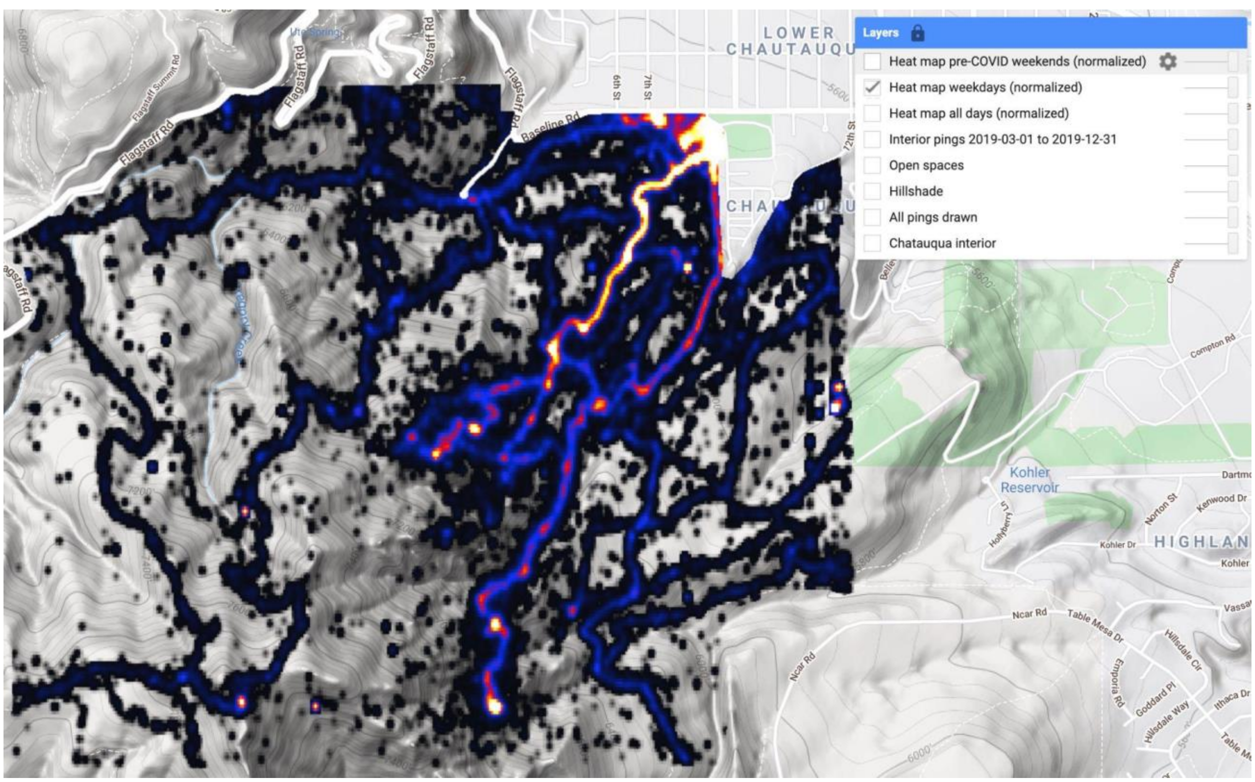
Pre-COVID all days



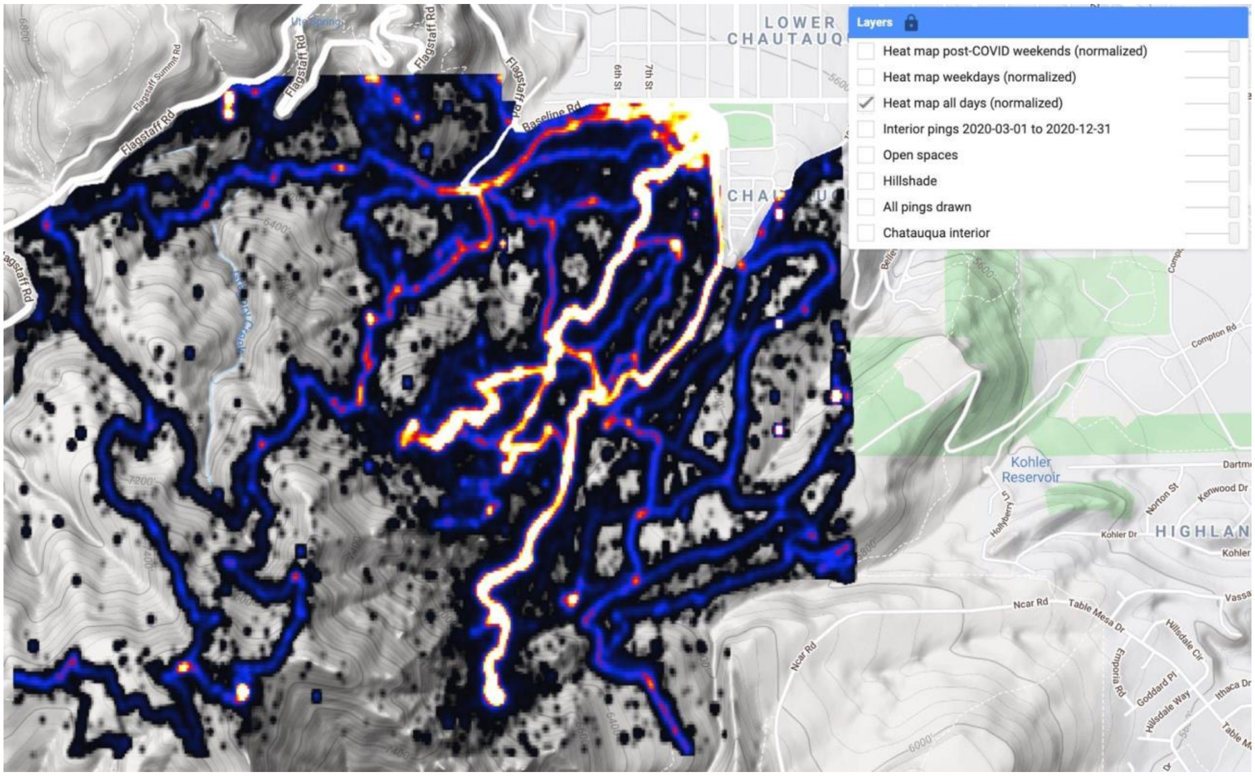
Weekends pre-COVID



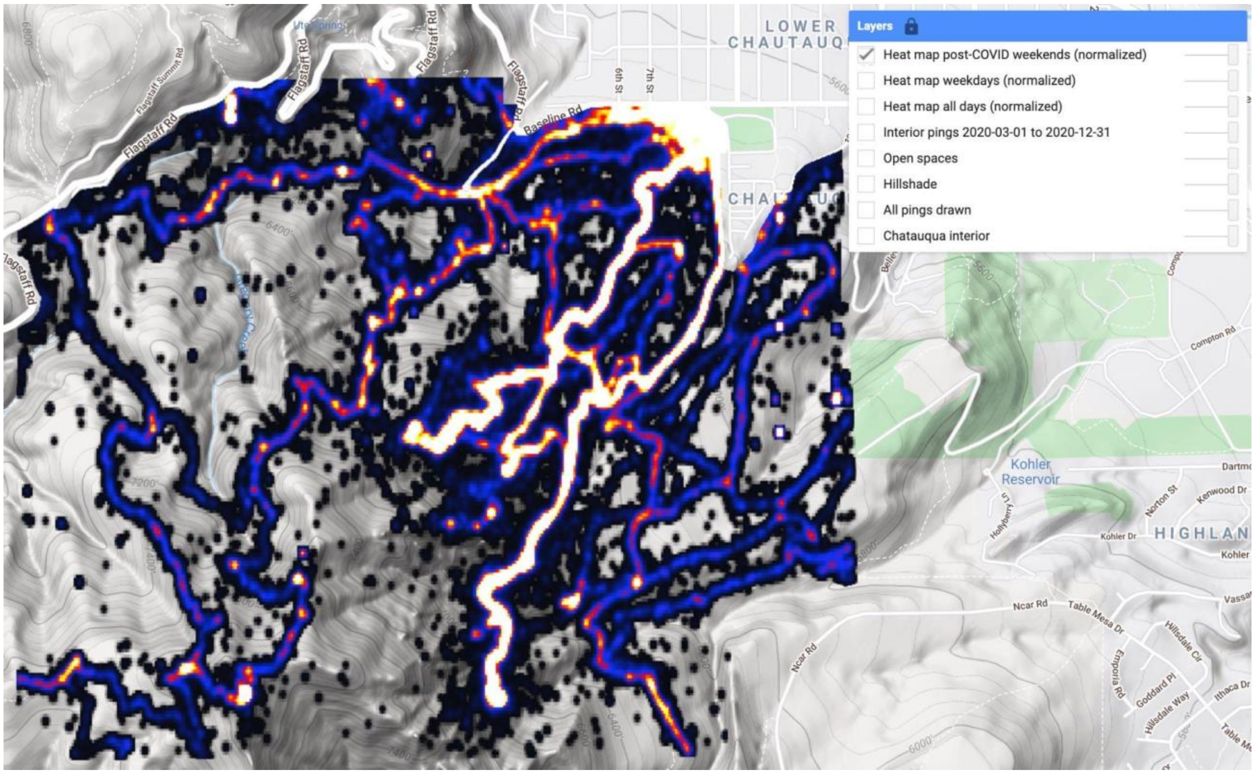
Weekdays pre-COVID



All post-covid



Post-covid weekends



Post-covid weekdays