



Mobile Device & Big Data Workshop 1

November 28, 2023

Drs. Ashley D'Antonio, Troy Hall, and Madeline Aberg
Oregon State University





Background

- Broader project context
 - Big Basin Redwoods State Park
 - Statewide trainings
- What is visitor use management?
- Why focus a training on “big data”?
 - Statewide survey → interest and questions

Overview

- What is “Big Data” in VUM Context
 - Social Media Data
 - Volunteered Geographic Information (VGI)
 - Mobile Device Data (MDD)
- Opportunities
- Limitations
- Case Study Examples
- Discussion & Questions

What is “Big Data” in a VUM Context?

Big Data Definition :

- 1 Large datasets often from secondary data sources
- 2 “Volume, velocity, and variety”

Dagan & Wilkins (2023), Journal of Outdoor Recreation and Tourism



Why Use “Big Data” in VUM?

- 1 Volume
- 2 Velocity
- 3 Variety
- 4 “Passive” Participation

Examples of Big Data in VUM

1 Social Media

2 Crowdsourced and/or Volunteered Geographic Information (VGI)

3 Mobile Device (Cellphone or Human Mobility) Data

4 Reservation system (i.e., rec.gov), connected vehicles, community science

Social Media Data

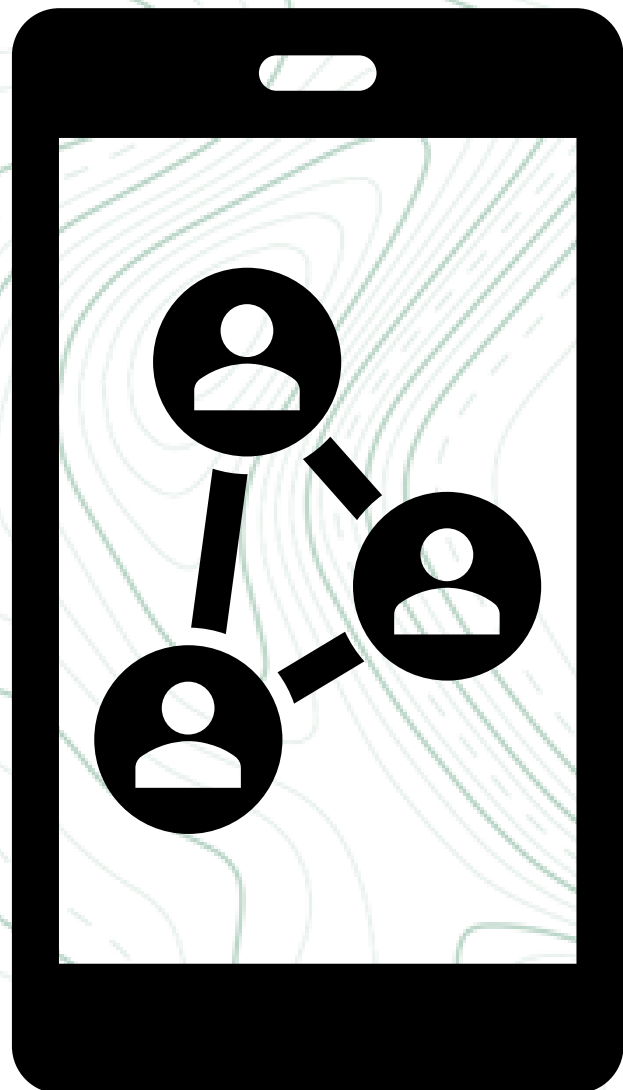




Wikimedia Commons, Luisbrunda

What is social media data?

- User-generated content that is publicly shared on digital platforms
- Social media may have volunteered geographic information (VGI)
- Examples:
 - Flickr
 - Twitter/X
 - Instagram



Opportunities

- Reduces time & resources in the field
- Can contain spatial information
- Contextual information
- Engagement between users
- Publicly available (using API) & low cost



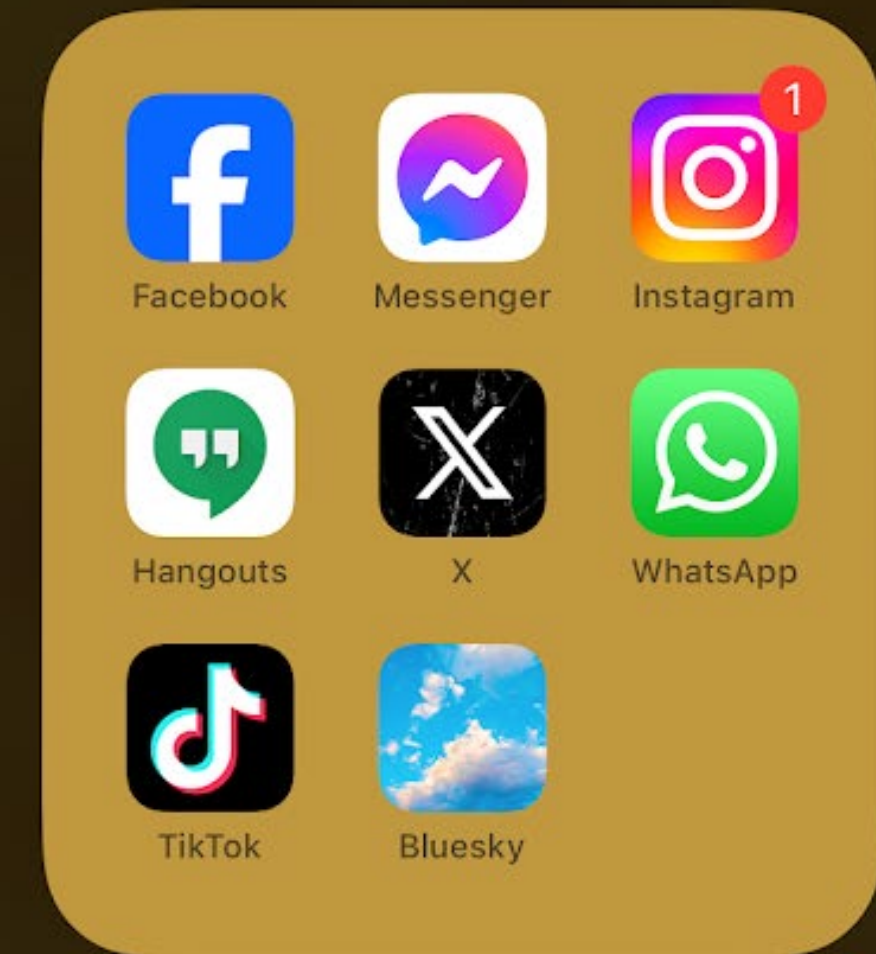
Strengths

- Correlated with visitor use
 - Relationships are variable
- Broad temporal scale predictions
 - i.e. visitor use estimations for an entire week vs. daily
- Text and/or photo context



Limitations

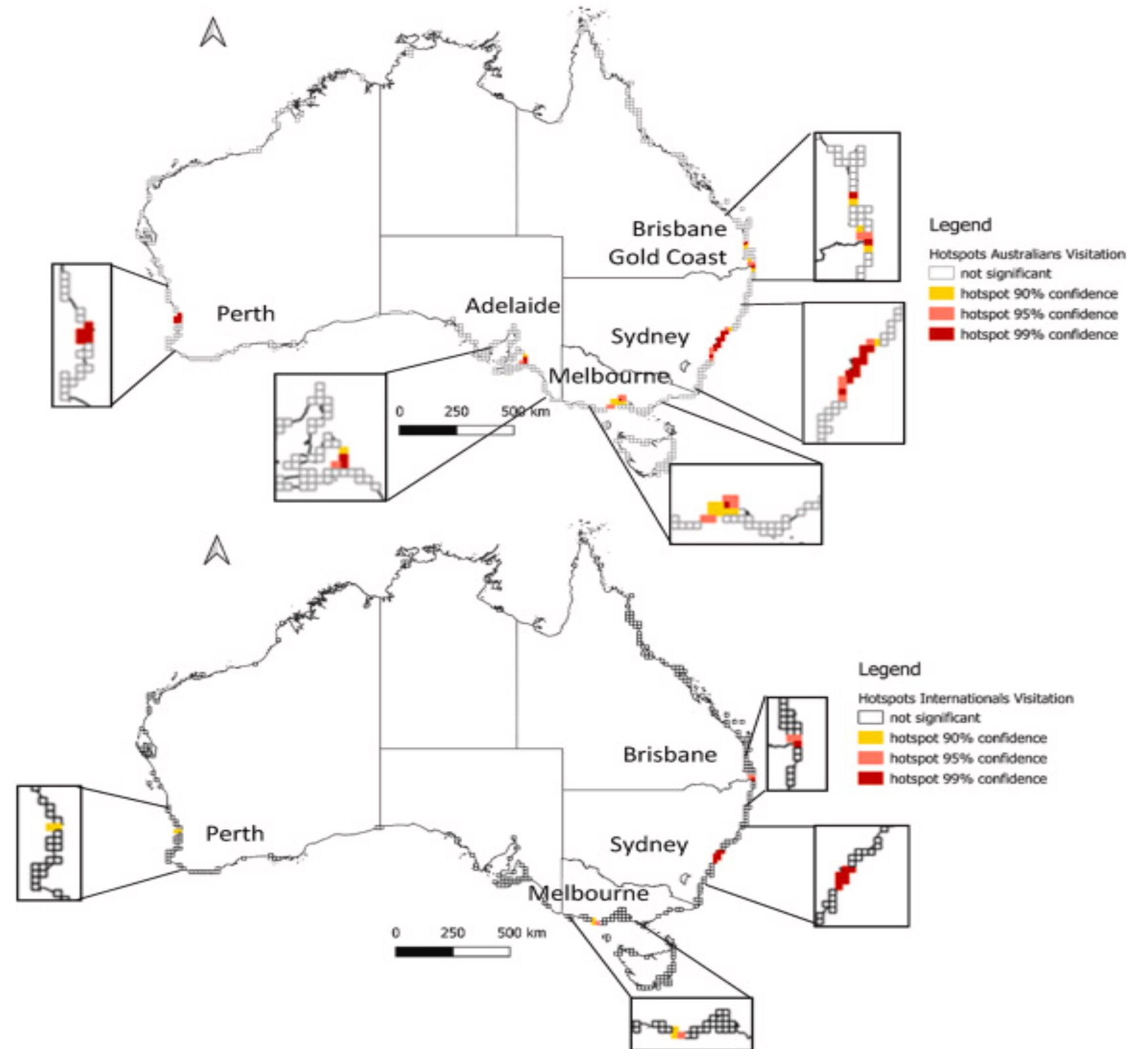
- Bias
 - Not representative of park users
 - Postings
- Changing privacy regulations & data availability
- Choosing relevant platform & validation



Social Media in VUM Example:

Popularity of Australian beaches:
Insights from social media images
for coastal management

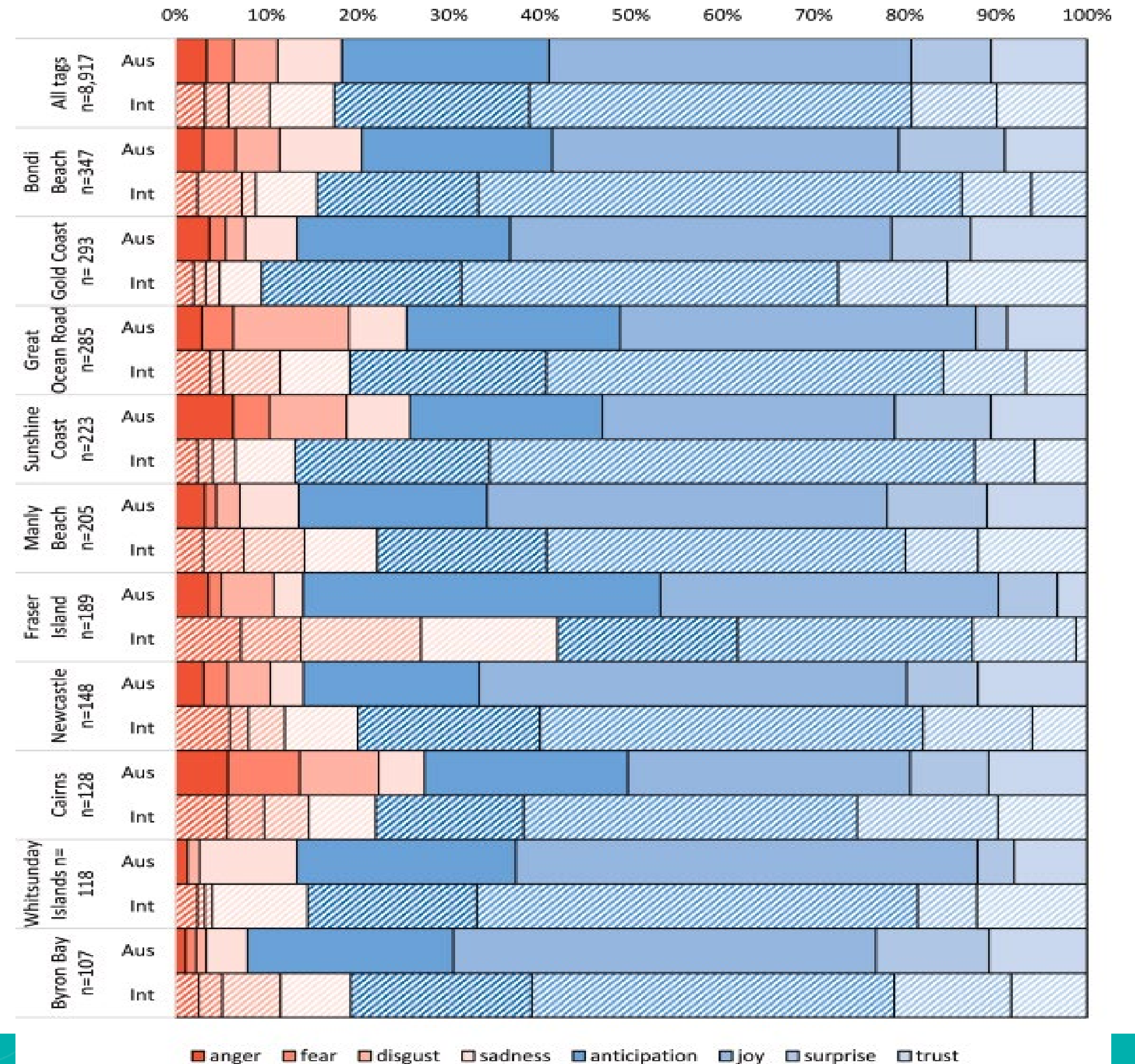
Mota & Pickering (2022) Ocean &
Coastal Management



Social Media in VUM Example:

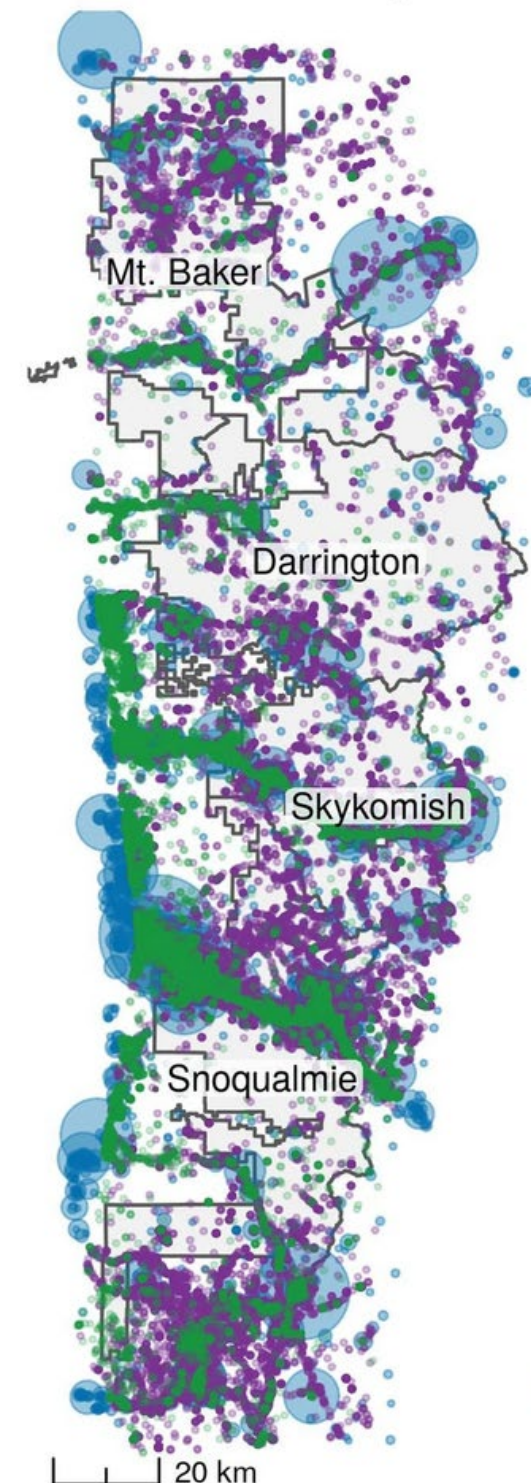
Popularity of Australian beaches:
Insights from social media images
for coastal management

Mota & Pickering (2022) Ocean & Coastal Management



Social Media in VUM Example:

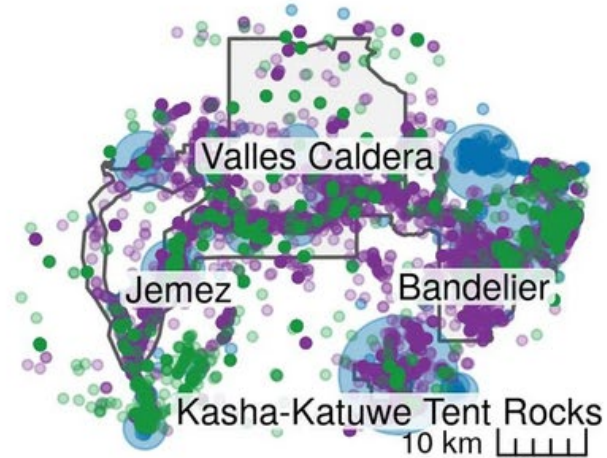
Western Washington



Western USA

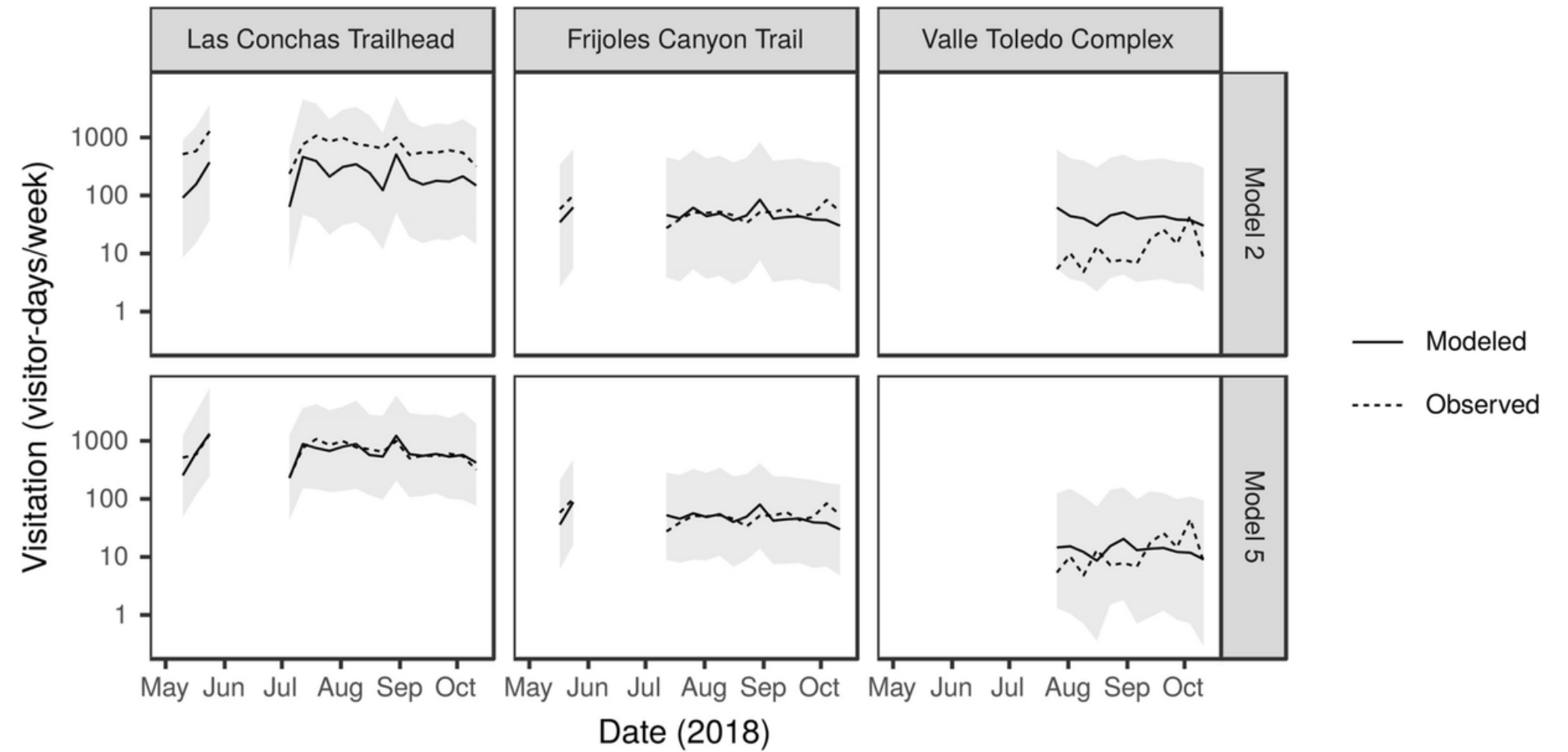


Northern New Mexico



Social Media

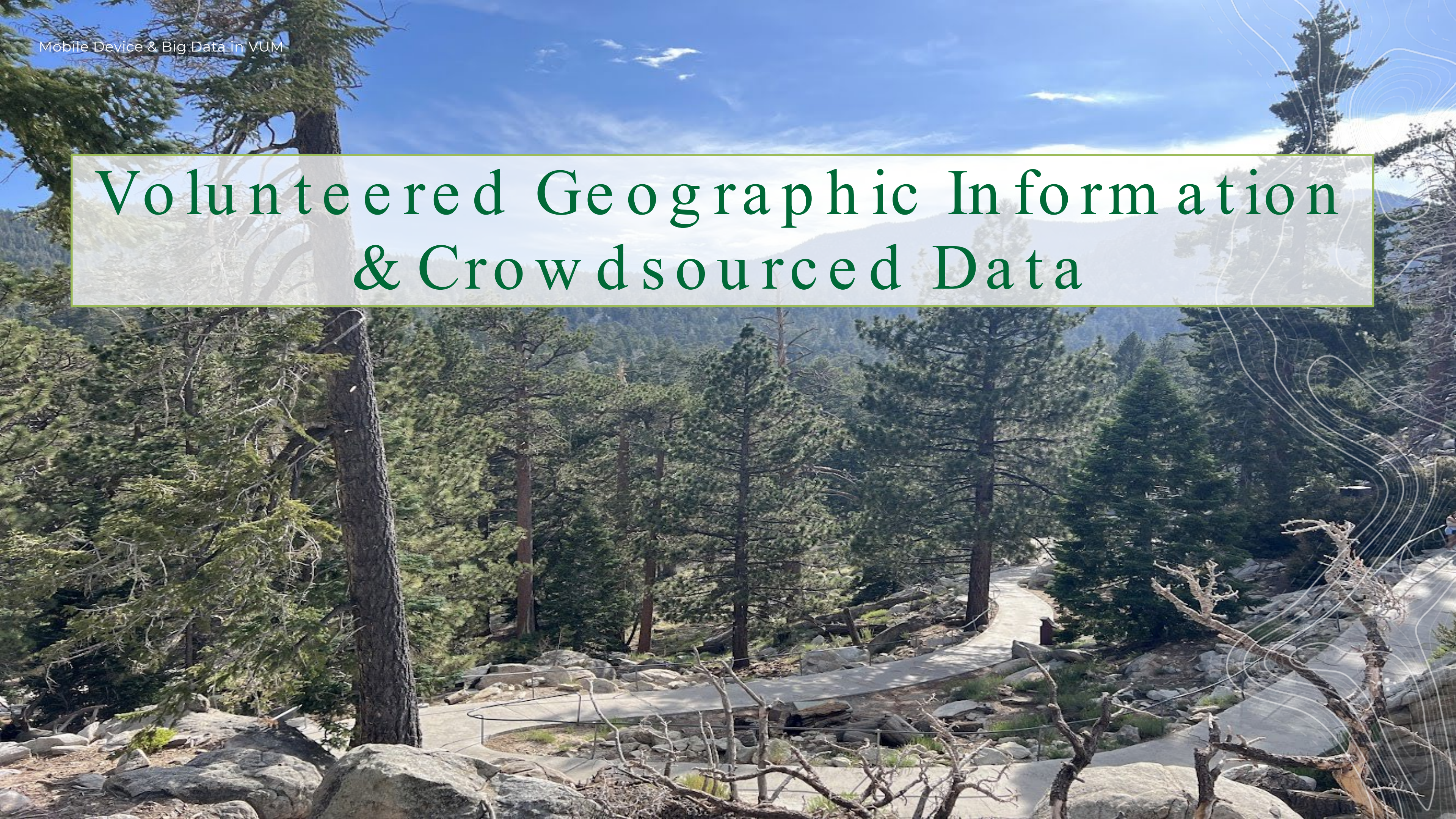
- Flickr
- Twitter
- Instagram

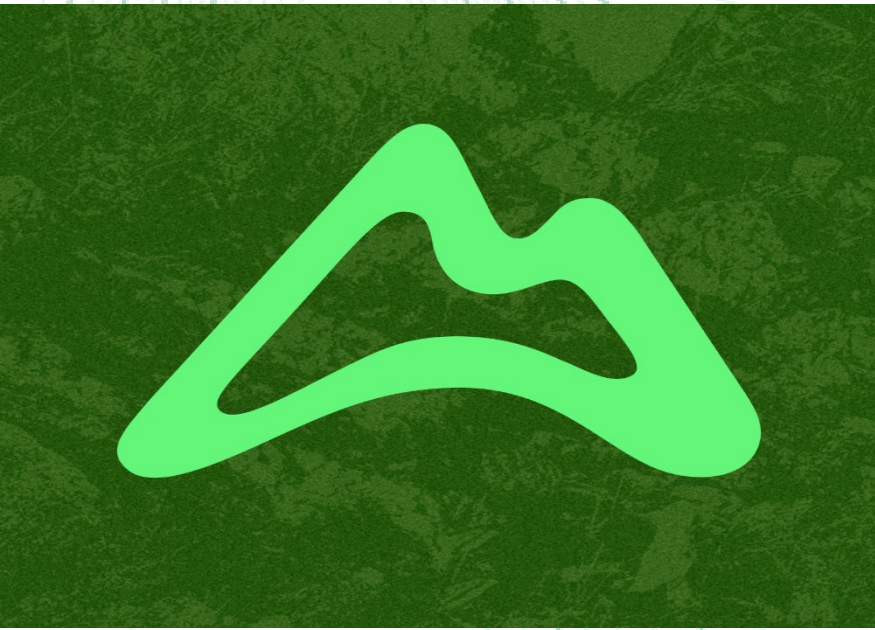


Next-generation visitation models using social media to estimate recreation on public lands

Wood et al., (2020) Scientific Reports

Volunteered Geographic Information & Crowdsourced Data





TRAILFORKS

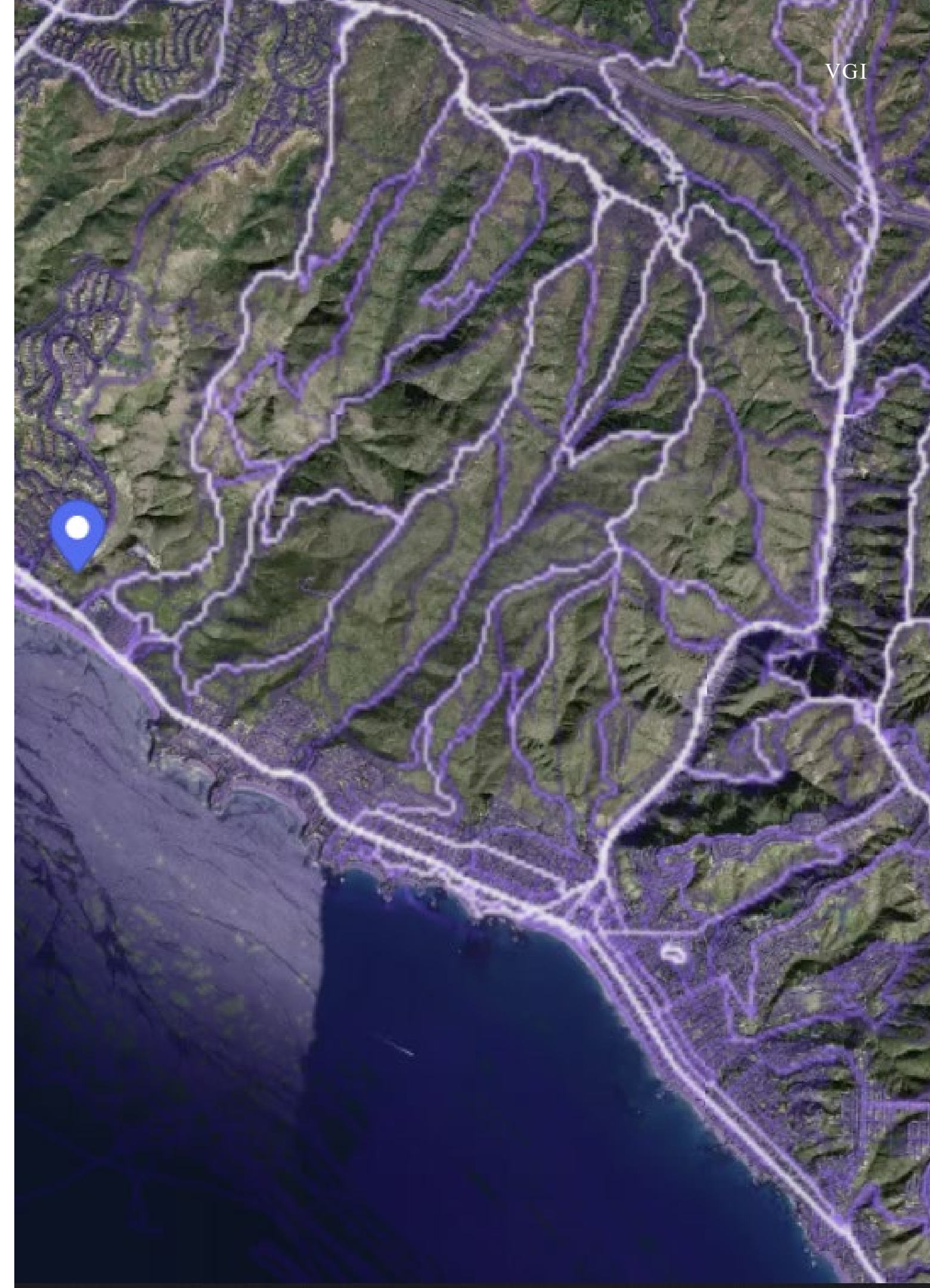


What is VGI data?

- User-generated content with geographic (spatial) information
- Social media data can have geographic info
- Examples (most fitness-tracking apps):
 - Strava
 - AllTrails
 - Trailforks

Opportunities

- Reduces time and resources in the field
- Publicly available & low cost
- Spatial and temporal precision
 - Broader scales
- Relative use levels & behavior



Strengths

- Best for destinations with a fitness or activity focus
- Comparing relative use levels or patterns



Limitations

- Representativeness
- Best if validated or calibrated
- Less precise at smaller spatial and temporal scales
- Changing popularity



VGI in VUM Example:

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

Norman & Pickering (2017). Applied Geography

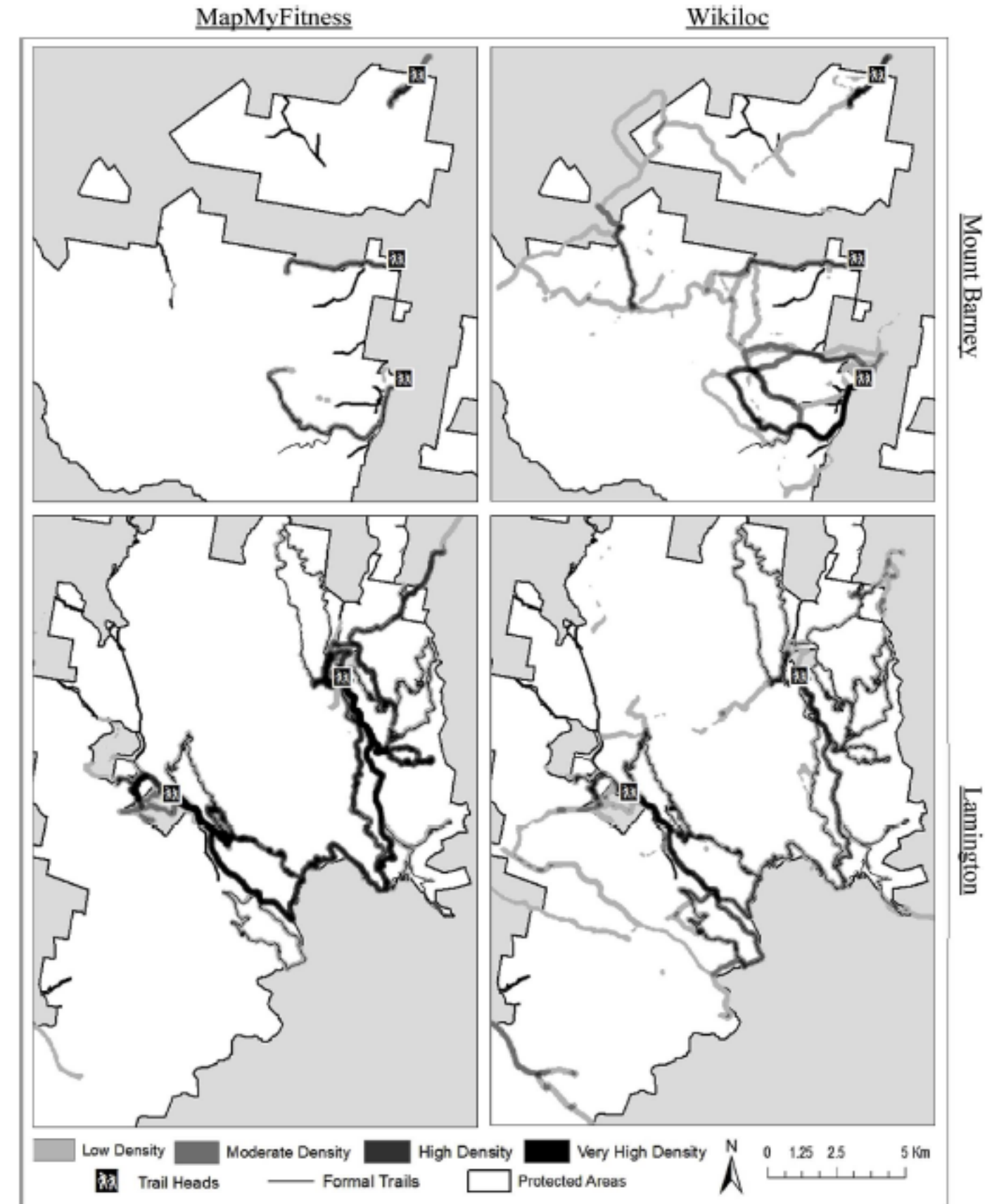
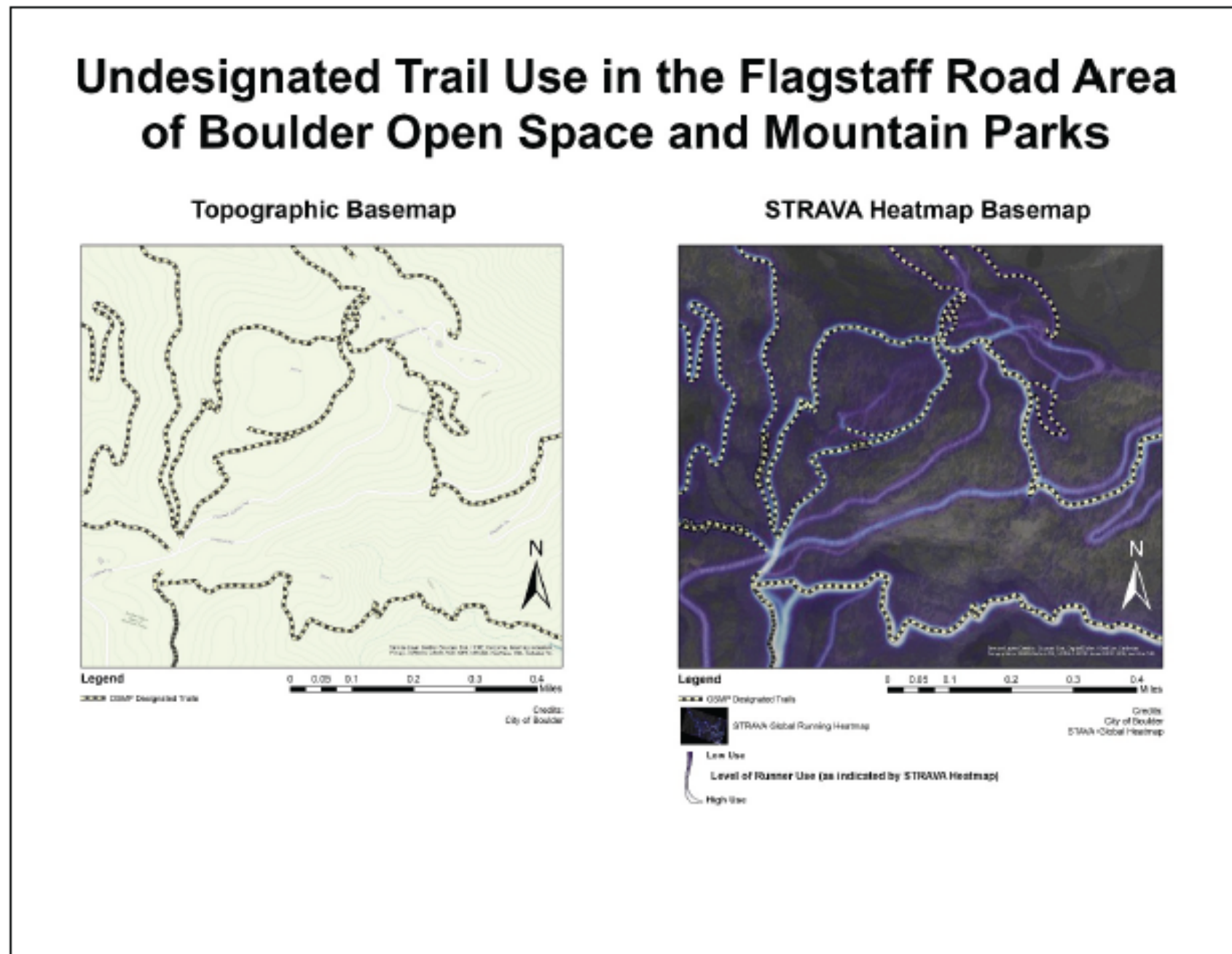


Fig. 5. Density maps of routes using data from MapMyFitness (Left) and Wikiloc (Right) for Mount Barney (Top) and Lamington National Park (Bottom).

VGI in VUM Example:



Detailing an Approach for Cost -
Effective Visitor-Use Monitoring
Using Crowdsourced Activity Data

Rice et al., (2019)
Journal of Parks and Recreation Administration

Figure 2. Maps of the Flagstaff Road portion of the Boulder Open Space and Mountain Parks

Mobile Device Data (Cellphone or Human Mobility Data)





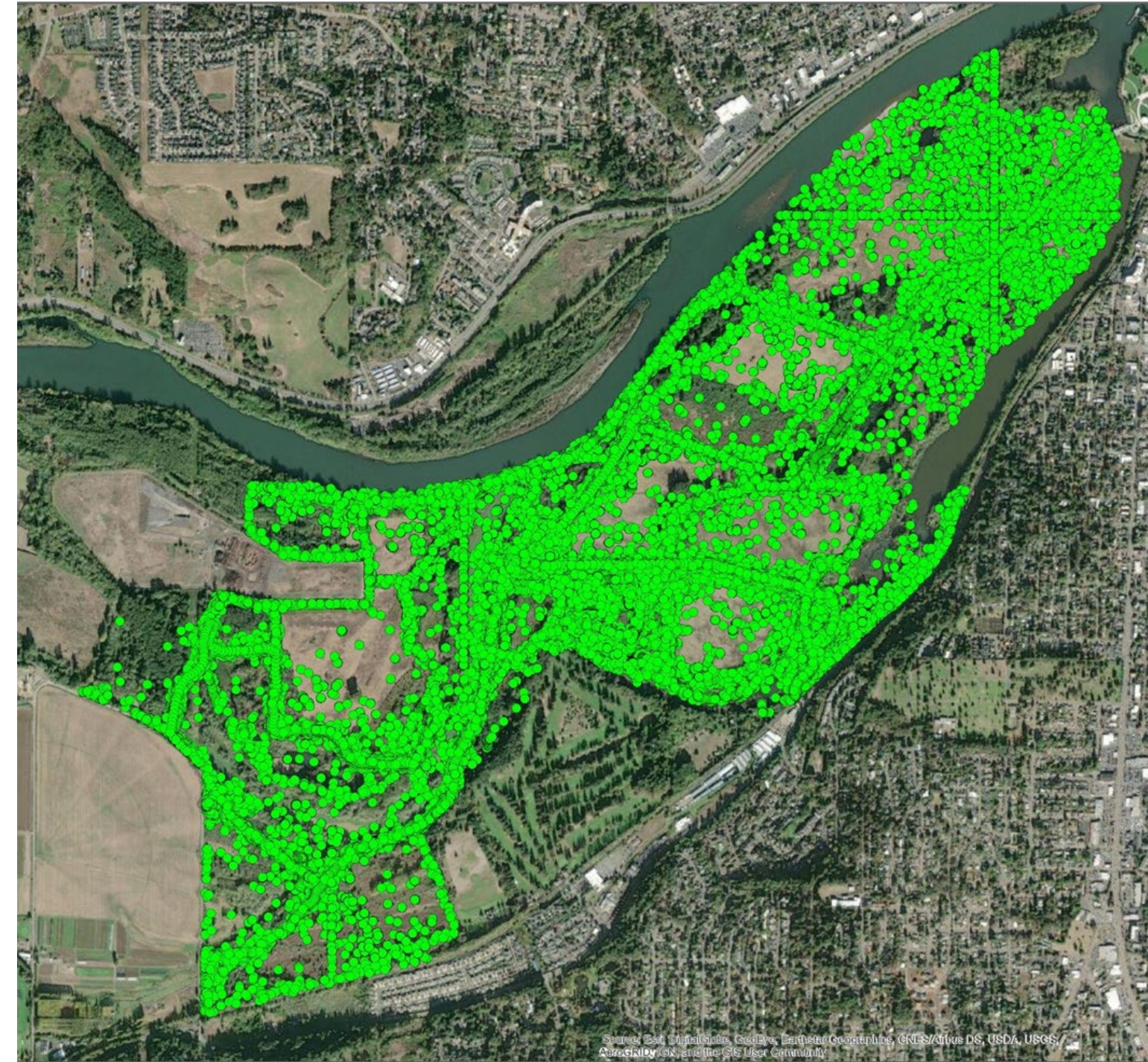
What is Mobile Device Data?

- Geographic data associated with movement of mobile device
- Where is the data coming from?
 - Device's internal GPS
 - Location-based Services (LBS)
- Often acquired through a vendor
 - Many vendors
 - Rapidly changing landscape



Opportunities

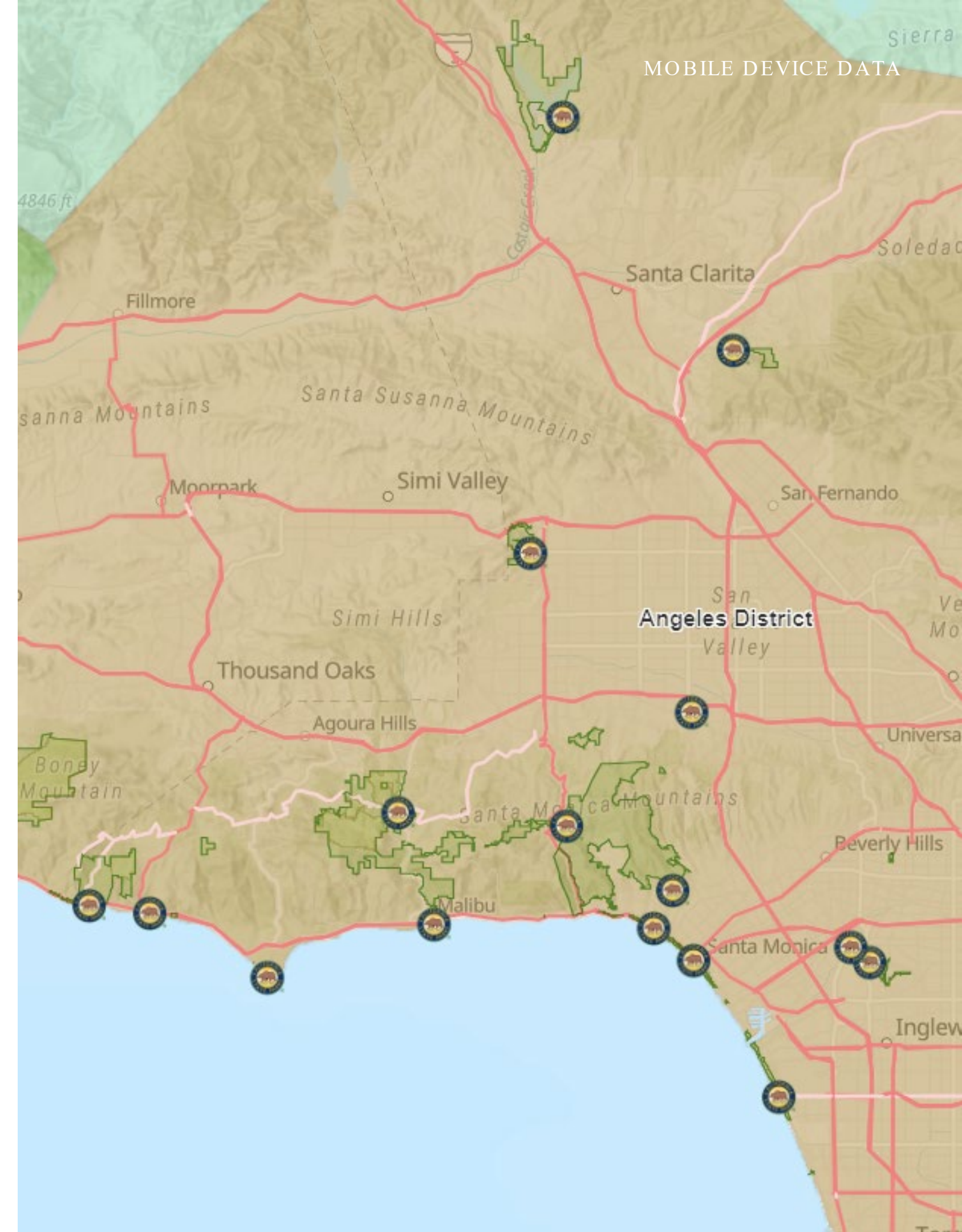
- High precision
- Variety of temporal scales & broad spatial availability
- Reduces time & resources in the field
- Additional info:
 - Demographics (some limitations)
 - Home location



Raw mobile device data (3 months in 2021) for an urban park in Salem, Oregon

Strengths

- Estimating visitation at broad spatial scales (i.e., entire park)
- Parks with porous boundaries
- Broad spatial patterns
- Network-level VUM questions



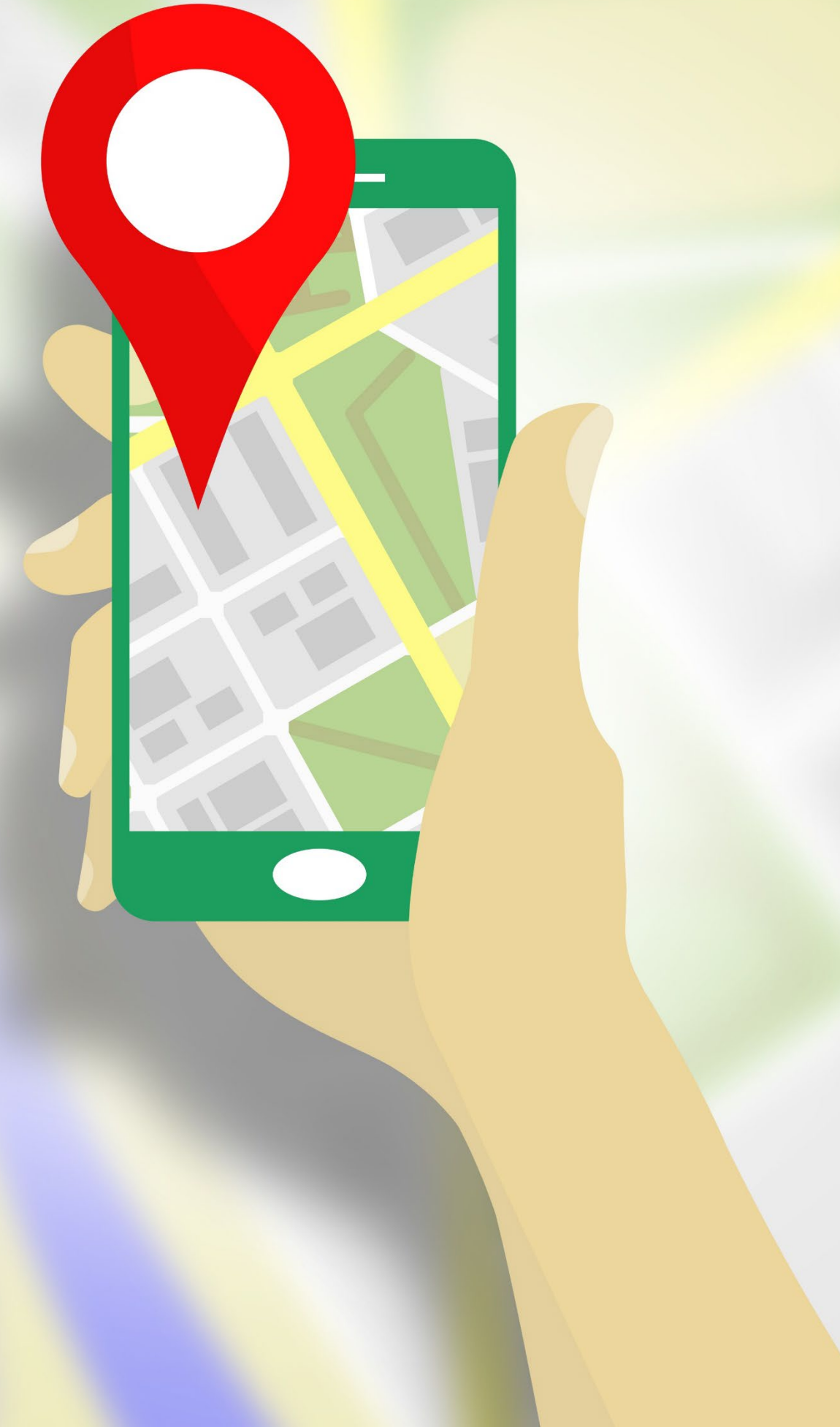
Limitations

- Not suitable for all VUM questions
- Can be expensive
- Transparency - “black box” of data processing & algorithms
- Calibration and validation still required



Limitations

- Demographic data
 - Park visitors don't match census blocks
 - Disparities in smartphone use/ownership
- Staff and volunteers overrepresented
- Spatial accuracy
- Behavioral aspect
 - Local vs. non-local, phone activity



Challenging Settings

- Parks with very close proximity to urban settings
- Parks with limited signal
- Small spatial scales
- Where calibration/validation data is not available



MDD in VUM Example :

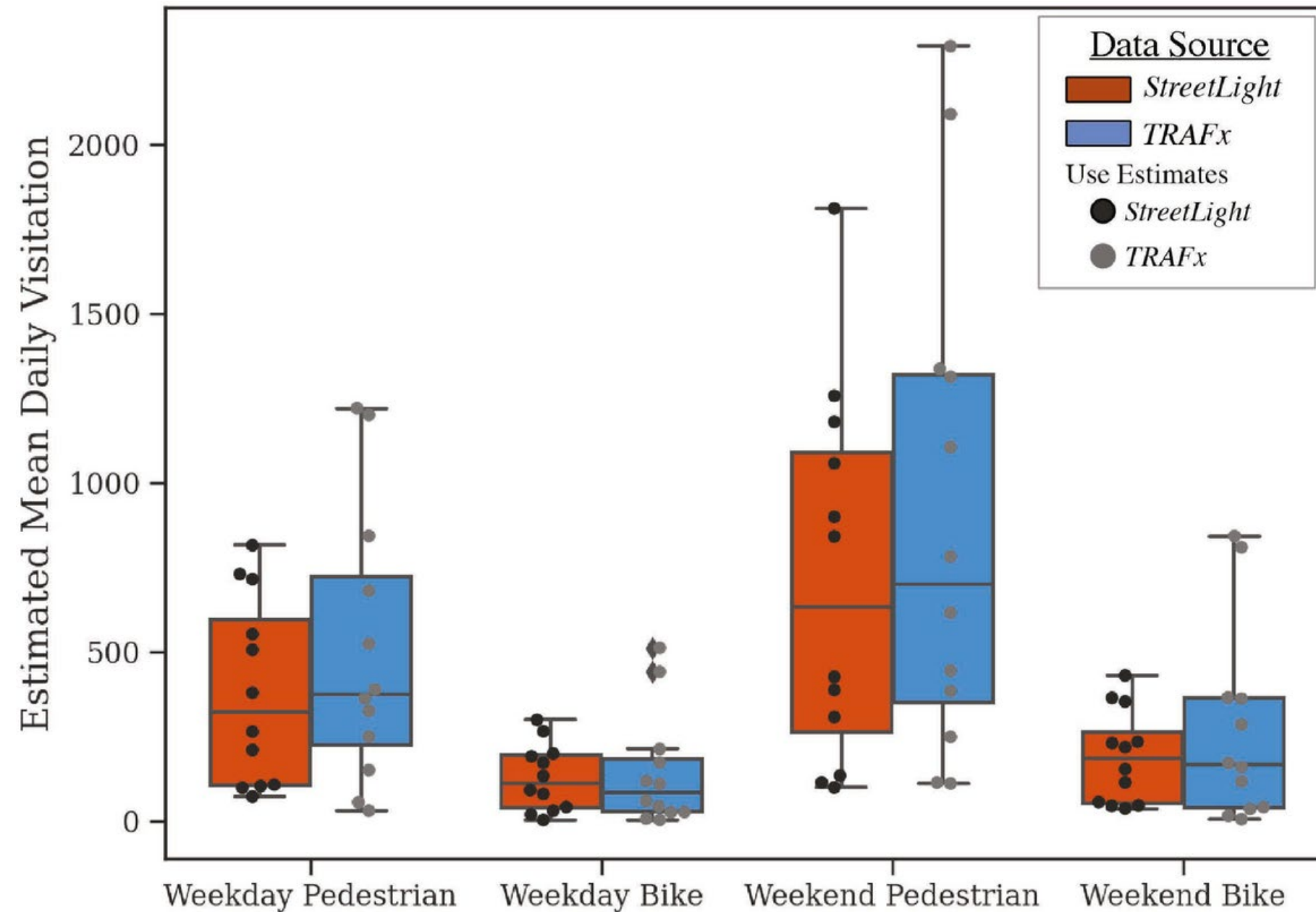


Fig. 2. *Streetlight* and *TRAFx* estimates of mean daily visitor use across the three parks in this analysis (n=12). Dots overlaid on the boxplots illustrate use estimate observations at four trailheads in three parks.

Estimating trail use and visitor spatial distribution using mobile device data: An example from the Nature Reserve of Orange County, California USA

Creany et al., (2021) Environmental Challenges

MDD in VUM Example:

Estimating trail use and visitor spatial distribution using mobile device data: An example from the Nature Reserve of Orange County, California USA

Creany et al., (2021) Environmental Challenges

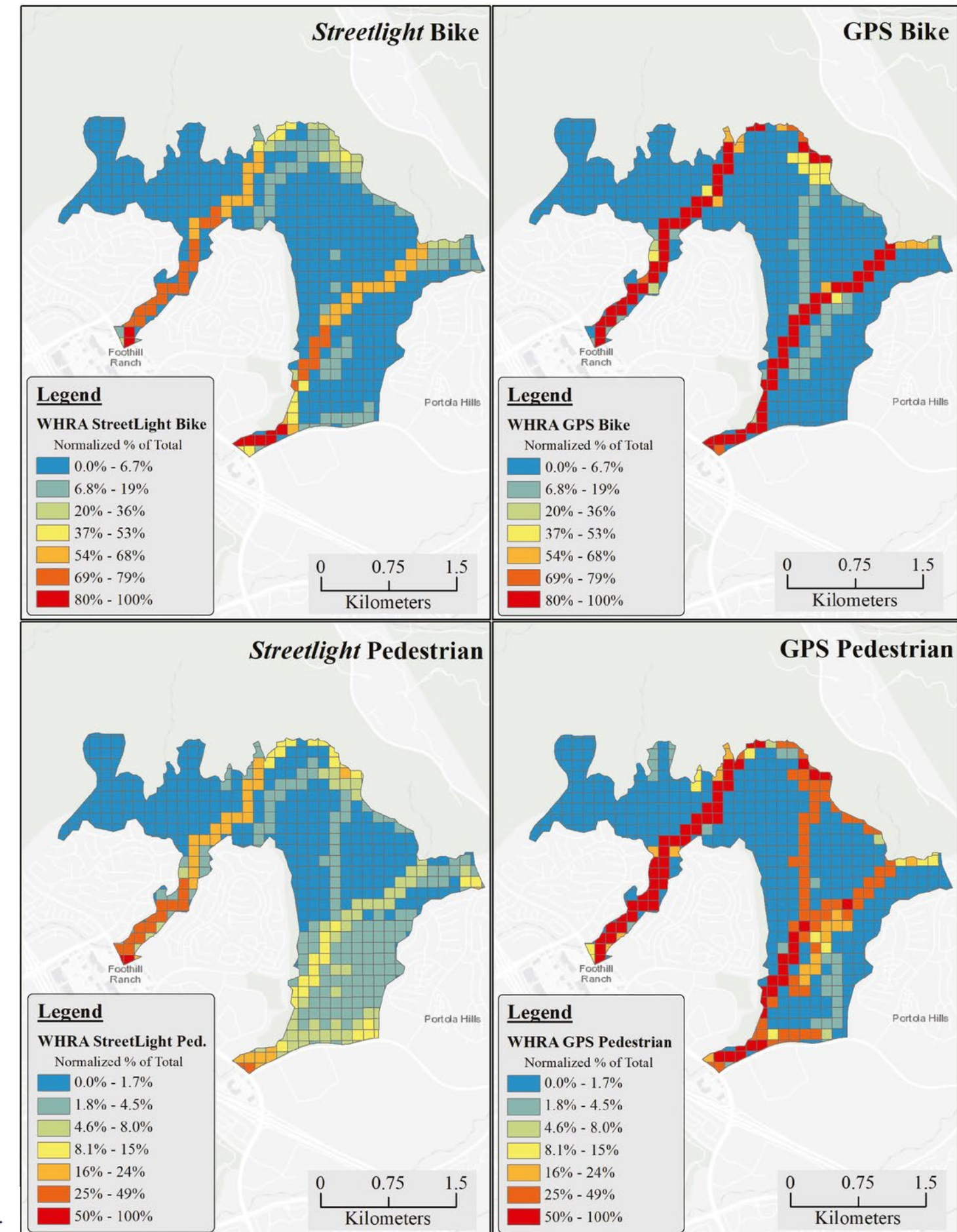


Fig. 5. Laguna Coast Wilderness and Crystal Cove State Park (LCW/CCSP) comparisons of Streetlight and GPS spatial distribution and density of use.

MDD in VUM Example :

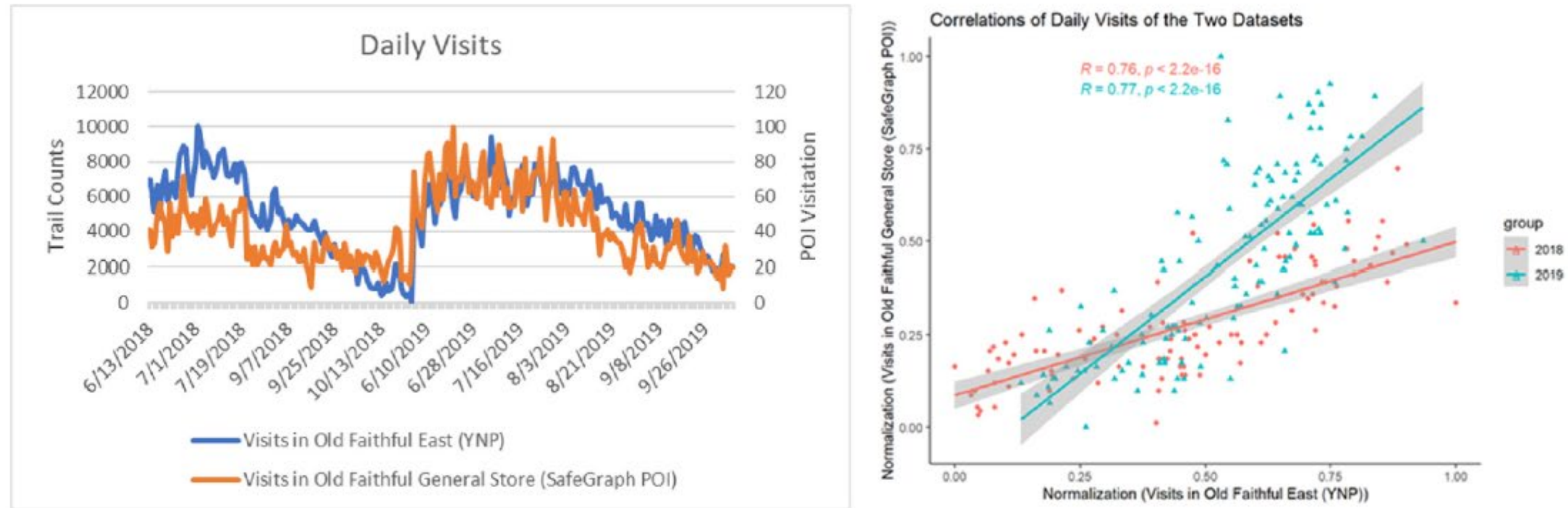


Fig. 4. Patterns (Left) and Correlations (Right) of Daily Visits between Trail Counters in Old Faithful East and SafeGraph POI in Old Faithful General Store. *Note.* The shaded areas (right) represent the 95% confidence interval.

Assessing the validity of mobile device data for estimating visitor demographics and visitation patterns in Yellowstone National Park



Erik Mclean (Pexels)



Other Sources of Big Data

- Reservation systems
- Active participation via apps
- Connected vehicles
- Community science apps (eBird, iNaturalist)

Ethical & Privacy Considerations





Take Home Messages

- Wealth of passive data available
- Many strengths and opportunities with big data!
- Important questions to consider:
 - Best data for the question/need?
 - Can you validate the data?
 - What does the data represent?
 - Ethical and privacy considerations

Selected References

Baird, T., Stinger, P., Cole, E., & Collins, R. (2022). Mobile Device Data for Parks and Public Lands Transportation Planning: A Framework for Evaluation and Applications. *Transportation Research Record*, 2676(8), 490–500. <https://doi.org/10.1177/03611981221083911>

Creany, N., Monz, C., D'Antonio, A., Sisneros-Kidd, A., Wilkins, E., Nesbitt, J., & Mitrovich, M. (2021). Estimating trail use and visitor spatial distribution using mobile device data: An example from the nature reserve of orange county, California USA. *Environmental Challenges*, 4, 100171. <https://doi.org/10.1016/j.envc.2021.100171>

Dagan, D. T., & Wilkins, E. J. (2023). What is “big data” and how should we use it? The role of large datasets, secondary data, and associated analysis techniques in outdoor recreation research. *Journal of Outdoor Recreation and Tourism*, 100668. <https://doi.org/10.1016/j.jort.2023.100668>

Liang, Y., Yin, J., Pan, B., Lin, M. S., Miller, L., Taff, B. D., & Chi, G. (2022). Assessing the validity of mobile device data for estimating visitor demographics and visitation patterns in Yellowstone National Park. *Journal of Environmental Management*, 317, 115410. <https://doi.org/10.1016/j.jenvman.2022.115410>

McKittrick, M. K., Schuurman, N., & Crooks, V. A. (2022). Collecting, analyzing, and visualizing location-based social media data: Review of methods in GIS-social media analysis. *GeoJournal*, 88(1), 1035–1057. <https://doi.org/10.1007/s10708-022-10584-w>

Rice, W., Mueller, J. T., Graefe, A., & DerrickTaff. (2019). Detailing an Approach for Cost-Effective Visitor-Use Monitoring Using Crowdsourced Activity Data. *Journal of Park and Recreation Administration*, 37(2), 144–155. <https://doi.org/10.18666/JPRA-2019-8998>

Schirck-Matthews, A., Hochmair, H., & Paulus, G. (2023). Comparison of reported outdoor activities in Florida State Parks among three fitness tracker apps. *Journal of Leisure Research*, 54(1), 46–71. <https://doi.org/10.1080/00222216.2022.2153097>

Whitney, P., Rice, W. L., Sage, J., Thomsen, J. M., Wheeler, I., Freimund, W., & Bigart, E. (2023). Developments in big data for park management: A review of mobile phone location data for visitor use management. *Landscape Research*, 1–19. <https://doi.org/10.1080/01426397.2023.2198762>

Venter, Z. S., Gundersen, V., Scott, S. L., & Barton, D. N. (2023). Bias and precision of crowdsourced recreational activity data from Strava. *Landscape and Urban Planning*, 232, 104686. <https://doi.org/10.1016/j.landurbplan.2023.104686>

Discussion:

What opportunities do you see for big data?



Questions?

