MAPPING SPATIAL AND TEMPORAL VARIABILITY IN MOBILITY PATTERNS IN CALIFORNIA during 2020-2022

Evgeny Noi[a]*, Alexander Rudolph[b] and Somayeh Dodge[a]

[a] Department of Geography, UCSB, Santa Barbara, USA
[b] Department of Statistics, UCSB, Santa Barbara, USA
* noi@ucsb.edu

Keywords: COVID-19, local Moran's I, mobility, spatial autocorrelation, visual analytics.

Introduction

The COVID-19 pandemic has had a powerful and lasting impact on human mobility globally and across American communities. Previous geospatial research has investigated the effect of socio-demographic and economic determinants on COVID-19 and human mobility (Franch-Pardo et al., 2020). Most of the research efforts were aimed at analysis of the coronavirus spread (Huang et al., 2021; Sugg et al., 2021), decease modelling (Agbehadji et al., 2020) and assessment of non-pharmaceutical intervention effectiveness (Askitas et al., 2021). While important, these investigations had limitations. First, they relied heavily on just one data set to make inferences. And secondly, they assessed only a very brief observation period (typically, a few months before and after the start of the pandemic).

The data sets utilized in geographical analysis of movement and COVID-19 are complex and heterogeneous. They come from a variety of sources, including Global Positioning Systems (GPS), call detailed records (CDR), Bluetooth, internet-of-things sensors, smartphone applications and other location-aware technologies (Buckee et al., 2020). Because the data is usually processed by different companies that rely on privacy guarding algorithms, raw measurements are aggregated which may result in biases originating from modifiable areal unit problem (Fotheringham and Wong, 1991), scaling problem and spatial misalignment (Gotway and Young, 2002), and ecological fallacy (Piantadosi et al., 1988). Since mobility data and metrics come in a variety of forms and units, they tend to provide only an approximation to the actual human mobility and are hard to compare to one another. The assessment of such complex multi-source data is problematic, therefore research frameworks of mobility must provide a more balanced analytical approach to account for discrepancies and deficiencies of data. In this short paper, we motivate this necessity with an illustrative example from California that assesses mobility indices obtained from two different sources and demonstrate spatial and temporal variability in captured patterns.

Two widely used mobility data sets, Apple Mobility Trend Reports (AMTR) and Google Community Mobility Reports (GCMR), collected over two and a half years between 2020 and 2022, are assessed and compared to one another in terms of their spatial and temporal coverage and distribution. These two data sources have been used in multiple geographical analyses that investigated mobility during the pandemic, both separately and in combination with one another (Sulyok and Walker, 2020; Cot et al., 2021; Hadjidemetriou et al., 2020). Thus, characterizing variability in these mobility data sets would highlight the importance of multi-source flexible visual analytics for contextualizing human mobility during the pandemic and through different waves of COVID-19.

Method

Exploratory Visual Analysis

In contrast to some earlier research, this short paper assesses two and a half years of mobility data (January/February 2020 - April/May 2022) from Apple and Google. This allows to ascertain regional mobility trends in California (as a case study) in relation to the spread of different virus variants (e.g. Alpha, Delta, Omicron) and the associated shifts in mobility at the county level. We assess the data using two visual displays presented in Noi et al. (2021): recency and consistency map and line-path scatter plot. Both visual displays utilize Local Moran's I (Anselin, 1995) to capture local clusters (hotspots and coldspots) in California.

Recency and consistency map provides a spatial view of the study area, where the recency of a spatial cluster is mapped onto color intensity, such that low intensity red/blue corresponds to the less recent hotspot/coldspot behavior and high intensity red/blue corresponds to the most recent hotspot/coldspot behavior. The consistency of the hotspot/coldspot is mapped onto the size of the centroid marker and is measured as the number of times (weeks) the county was classified as statistically significant hotspot/coldspot. The recency and consistency map provides a quick at-glance overview of the spatial and temporal patterns captured in data. In particular, four types of behavior are easily discovered: recent and consistent hotspots/coldspots, recent and inconsistent hotspots/coldspots, non-recent and inconsistent hotspots/coldspots, and non-recent and consistent hotspots/coldspots. These four types of behavior are various extremes of the spectra, there are many different combinations in between as we will demonstrate.

The line-path scatter plot provides an aspatial representation for individual counties over time. And in particular, it maps the weeks (temporal aspect) on the x-axis, and the consistency is mapped onto y-axis. Thus, the succession of cumulative consistency is converted into line segments and mapped into individual county paths. This allows to easily ascertain the temporal trends in spatial autocorrelation within data. Specifically, if the line-path slopes are close to one, the county behavior is most consistent (i.e. the county was classified as statistically significant hotspot/coldspot each week through the observation period). The counters located in the top left corner of the plots indicate the number of derived and statistically significant hotspots, coldspots, and flips. The flipping behavior is observed when the county shifts from being a hotspot to being a coldspot (or the other way).

Data

Apple published Mobility Trend Reports from January 13, 2020 until April 14, 2022. The data is available for each country, where Apple Maps is used. For the United States, the data is provided daily at the county level and is calculated based on the number of requests for Apple Maps navigation. To preserve privacy of its users, Apple uses rotating identifiers and thresholding algorithms. The mobility index is provided relative to the baseline volume of navigation requests on January 13, 2020. The temporal coverage of data (Figure 1a) appears to have several gaps in May 2020 (Monday, Tuesday), March 2021 (Friday) and March 2022 (Monday). For spatial coverage (Figure 1a) the lowest number of complete records is in the following counties: Modoc, Trinity, Plumas, Sierra and Alpine. These are the least populous counties in CA with population well under 20,000, so the gaps in data are likely attributed to privacy-thresholding algorithms.

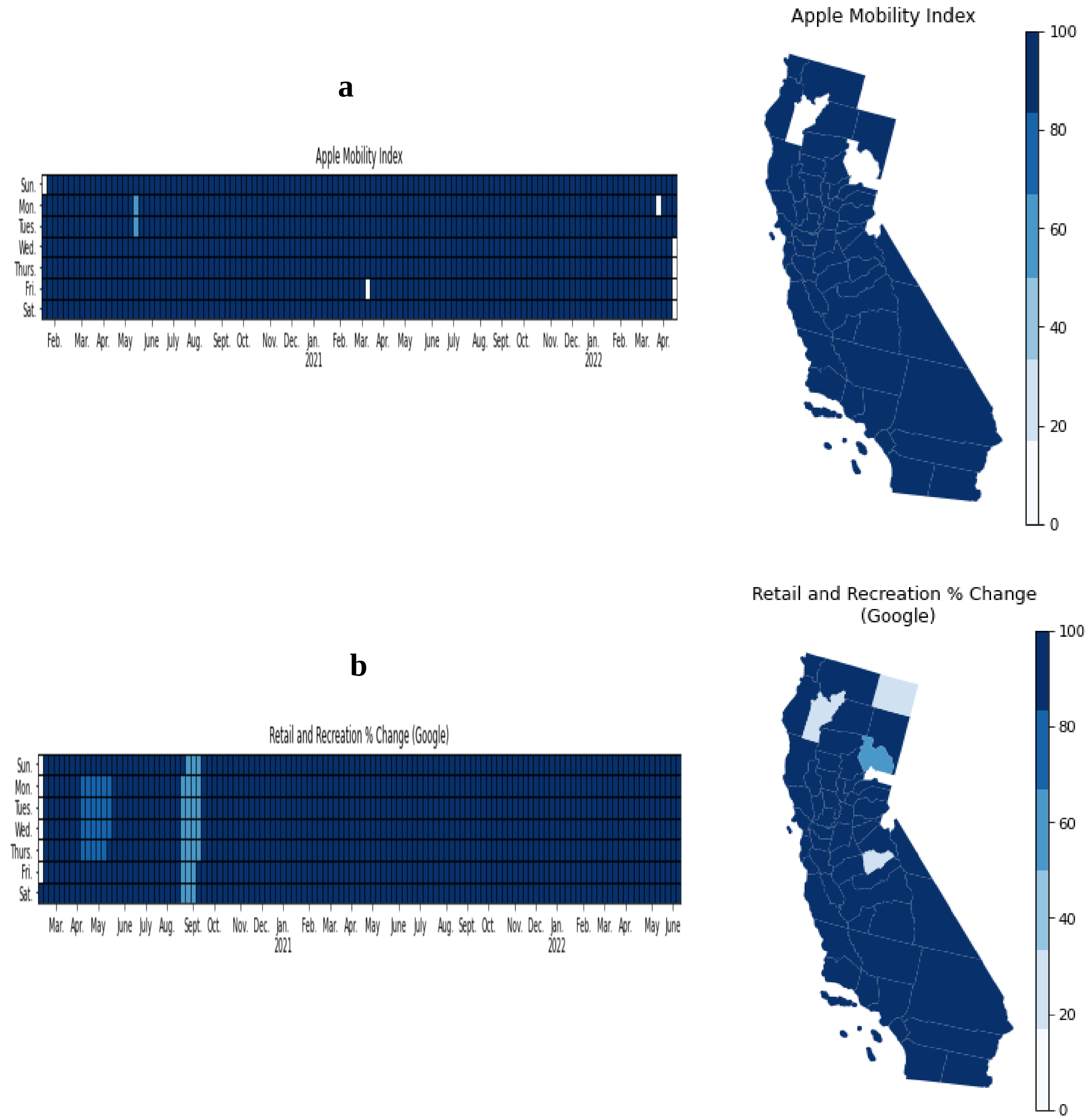
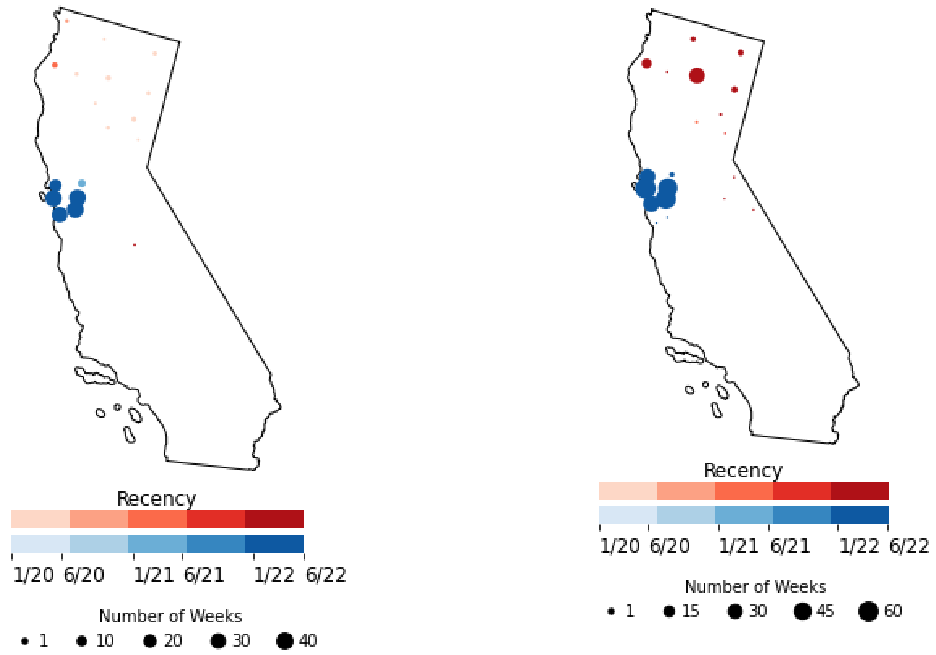


Figure 1: Spatial and temporal data coverage in CA in 2020-2022. (a) – Apple Mobility Reports. (b) – Google Mobility Reports. The color values correspond to the percentage of complete records per day/county. Darker colors denote higher coverage (more complete records), while lighter colors denote lower coverage (less complete records).

Google started publishing data on February 15, 2020 and the data is still being published. The data is broken down by location and is available across the globe. For the United States, the data set is available at the county level and provides daily mobility values. As opposed to the AMTR, Google does not publish a single index. Instead, the indices are broken down by the type of destination (e.g. retail and recreation; grocery and pharmacy; parks; transit stations; workplaces; residential). The mobility indices are calculated as percentage change in the number of visits from the day of week during the baseline period (January 3 - February 6, 2020). While all these indices have various spatial and temporal distributions as outlined in Noi et al. (2022), we selected retail and recreation mobility as it more closely aligns with the AMTR. The temporal coverage for Google mobility trends (Figure 1b) appears similar to Apple, albeit with a higher degree of complete records. For the spatial coverage (Figure 1b), the number of missing values is lowest in Sierra and Alpine counties, followed by Modoc, Trinity and Mariposa (second lowest) and Plumas county (third lowest). Once again, this grouping of counties corresponds to the population residing in these counties, so the gaps can be attributed to privacy-thresholding algorithms (Aktay et al., 2020).

Results

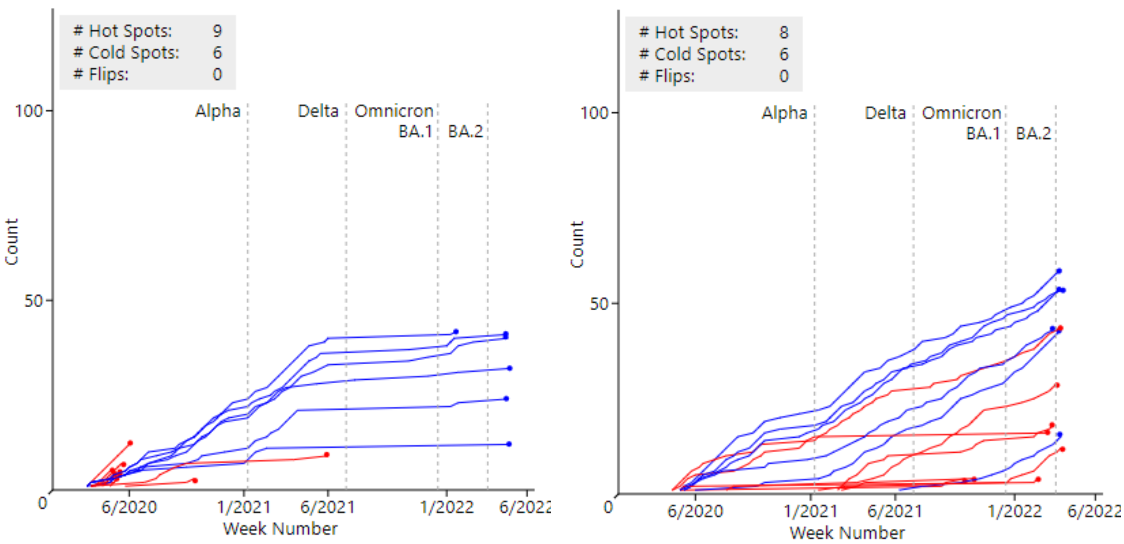
Since both Apple and Google reports denote percentage change in mobility from a set baseline period, hotspots would denote counties with associated increase in mobility located next to other counties with increases in mobility. On the other hand, coldspots would denote counties with associated decrease in mobility located next to other counties with decrease in mobility. Since local autocorrelation assesses individual county values in relation to the sample average, only state trends and patterns can be derived.



a **b**

Figure 2: Recency and consistency map for California in 2020-2022. (a) – Apple Mobility Reports. (b) – Google Mobility Reports.

The recency and consistency maps for California are provided in Figure 2. Both Apple and Google capture the coldspots in the Bay Area: in Alameda, San Francisco, Contra Costa, San Mateo, and Marin county. These areas have increased density of IT jobs, and they easily switched to working remotely starting from the first weeks of the COVID-19 pandemic. Working from home significantly reduced human mobility for retail and recreation in the Bay Area, thus prompting the counties to exhibit a coldspot behavior. The consistency of coldspots derived from Google mobility reports appears to be higher across the counties surrounding the Bay Area. The most drastic difference between Apple and Google mobility reports appears in Northern California, where Apple misses the recent hotspot (in Shasta, Humboldt, Lassen, Modoc, and Siskiyou counties). This can be potentially attributed to the large number of attractions in Northern California, such as wineries and camp sites in numerous national forests (Shasta-Trinity National Forest, Klamath National Forest), lakes (Lake Alhamanor, Eagle Lake, Honey Lake), which are not captured well by Apple navigational requests. This inconsistency points to selection bias in Apple mobility reports and must be acknowledged. Since the data assessed in this paper only looks at the hotspots/coldspots within the state of California, the changes in mobility in counties in Southern California are not captured. This is expected in assessing local indicators of spatial associations at the regional level: there are only 58 counties in California, and the South appears closer in values to sample average.



a **b**

Figure 3: Line-path scatter plot for California in 2020-2022. (a) – Apple Mobility Reports. (b) – Google Mobility Reports.

The line-path scatter plots are provided in Figure 3. The first thing we notice is that Apple mobility metrics peak at consistency of 40 weeks and plateau around week 70 (second half of May 2021). This indicates that Apple mobility reports fail to capture the changes in mobility signal for both Delta and Omicron virus waves. The hotspot/coldspot consistency, on the other hand, as was noted earlier, is much higher for local clusters derived from Google data, which captures the changes in mobility signal from COVID-19 waves associated with different variants. Furthermore, the line-path scatter plot supports our earlier findings on the hotspot behavior. The consistency of hotspots in Apple data is at or below 10, with the majority of counties stopping hotspot behavior around May 2021. For Google data, we observe a more consistent hotspot behavior that goes well into the 2021-2022.

Discussion and Conclusion

This short paper demonstrated that the spatial and temporal coverage across multi-source mobility data is not uniform. Therefore, researchers must be careful about making inferences on such data. While both Apple and Google mobility reports provide a relatively consistent and complete spatial and temporal coverage for their corresponding mobility indices, these data vary in a variety of ways. And specifically, local variation in Google data appears to be bicentric with two local clusters: a coldspot in Bay Area and a hotspot in Northern California. In contrast, Apple data appears to be monocentric only capturing the coldspot within the Bay Area.

The analysis presented here only considers one scale of analysis: state-level. Thus, only local variation is captured. It would be interesting to contrast and compare these indices in relation to the national average, which might indicate interesting changes in mobility in the Southern California, particularly for San Diego and Los Angeles metro areas as noted in Noi et al. (2021).

Acknowledgements:

The authors gratefully acknowledge the support from the National Science Foundation Award BCS #2043202, and the UCSB Vice Chancellor for Research COVID-19 Seed Program to conduct this research.

References

Agbehadji, I. E., Awuzie, B. O., Ngowi, A. B., & Millham, R. C. (2020). Review of big data analytics, artificial intelligence and nature-inspired computing models towards accurate detection of COVID-19 pandemic cases and contact tracing. *International journal of environmental research and public health*, *17*(15), 5330.

Aktay, A., Bavadekar, S., Cossoul, G., Davis, J., Desfontaines, D., Fabrikant, A., ... & Wilson, R. J. (2020). Google COVID-19 community mobility reports: anonymization process description (version 1.1). *arXiv preprint arXiv:2004.04145*.

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical analysis*, *27*(2), 93-115.

Askitas, N., Tatsiramos, K., & Verheyden, B. (2021). Estimating worldwide effects of non-pharmaceutical interventions on COVID-19 incidence and population mobility patterns using a multiple-event study. *Scientific reports*, *11*(1), 1-13.

Buckee, C. O., Balsari, S., Chan, J., Crosas, M., Dominici, F., Gasser, U., ... & Schroeder, A. (2020). Aggregated mobility data could help fight COVID-19. *Science*, *368*(6487), 145-146.

Cot, C., Cacciapaglia, G., & Sannino, F. (2021). Mining Google and Apple mobility data: Temporal anatomy for COVID-19 social distancing. *Scientific reports, 11*(1), 1-8.

Fotheringham, A. S., & Wong, D. W. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and planning A, 23*(7), 1025-1044.

Franch-Pardo, I., Napoletano, B. M., Rosete-Verges, F., & Billa, L. (2020). Spatial analysis and GIS in the study of COVID-19. A review. *Science of the total environment, 739*, 140033.

Gotway, C. A., & Young, L. J. (2002). Combining incompatible spatial data. *Journal of the American Statistical Association, 97*(458), 632-648.

Hadjidemetriou, G. M., Sasidharan, M., Kouyialis, G., & Parlikad, A. K. (2020). The impact of government measures and human mobility trend on COVID-19 related deaths in the UK. *Transportation research interdisciplinary perspectives, 6*, 100167.

Huang, J., Kwan, M. P., & Kan, Z. (2021). The superspreading places of COVID-19 and the associated built-environment and socio-demographic features: A study using a spatial network framework and individual-level activity data. *Health & Place, 72*, 102694.

Noi, E., Rudolph, A., & Dodge, S. (2021). A novel method for mapping spatiotemporal structure of mobility patterns during the COVID-19 pandemic. GIScience 2021 Short Paper Proceedings.

Noi, E., Rudolph, A., & Dodge, S. (2022). Assessing COVID-induced changes in spatiotemporal structure of mobility in the United States in 2020: a multi-source analytical framework. *International Journal of Geographical Information Science, 36*(3), 585-616.

Piantadosi, S., Byar, D. P., & Green, S. B. (1988). The ecological fallacy. *American journal of epidemiology, 127*(5), 893-904.

Sugg, M. M., Spaulding, T. J., Lane, S. J., Runkle, J. D., Harden, S. R., Hege, A., & Iyer, L. S. (2021). Mapping community-level determinants of COVID-19 transmission in nursing homes: A multi-scale approach. *Science of the Total Environment, 752*, 141946.

Sulyok, M., & Walker, M. (2020). Community movement and COVID-19: a global study using Google's Community Mobility Reports. *Epidemiology & Infection, 148*.