Received: 22 October 2024 | Revised: 6 May 2025 | Accepted: 12 June 2025

DOI: 10.1111/geoj.70033

THE GJ Royal Geographical Society with IBG Advancing geography and geographical learning WILEY

ARTICLE

Understanding place-to-place interactions using flow patterns derived from in-app mobile phone location data

Mikaella Mavrogeni | **Justin van Dijk** | **Paul Longley**

Department of Geography, University College London, London, UK

Correspondence
Mikaella Mavrogeni, Department of Geography, UCL, North West Wing, Gower Street, London WC1E 6BT, UK.
Email: mikaella.mavrogeni.19@ucl.ac.uk

Funding information
Economic and Social Research Council, Grant/Award Number: ES/Z504464/1

Abstract

Functional roles of neighbourhoods change throughout the day, as both a cause and consequence of human mobility fluctuations. Here we review how neighbourhoods can be characterised by origin–destination flows derived from individual level GPS-enabled in-app data. These are used to track individual trajectories from start to end points prior to aggregation. We leverage securely held individual level in-app mobile phone location data that preserve spatial and temporal flexibility in representing place-to-place interactions. The data are available at the individual level and are aggregated for reporting of origin–destination analysis at the Middle layer Super Output Area (MSOA) level to accommodate disclosure control and positional uncertainty. We show how in-app mobile phone location data for Greater London enhance our understanding of the relationships between places, and demonstrate how these relationships may change over the course of the day. Finally, we discuss how such analysis can inform transport policy and the contribution of our approach to extending geodemographic research.

KEYWORDS

geodemographics, geo-temporal analytics, in-app location data, interactional geographies, mobility patterns, origin–destination analysis

1 | BACKGROUND

Human geographers have long studied the forms and functions of settlement systems across a range of spatial scales, with a particular focus on intra-urban residential structure. This has been formalised as the study of geodemographics, which characterises populations based on where they live (Harris et al., 2005; Longley, 2015). At its core, geodemographics rely on data reduction techniques, such as k-means clustering, to summarise data collected by national statistical organisations and reveal social patterning at small spatial scales (Gale et al., 2016; Wyszomierski et al., 2024). In the UK, neighbourhood-scale census 'state-istics' (Dorling & Simpson, 1998; Louckx & Vanderstraeten, 2015) are shaped by government priorities, with the content of modern UK censuses tracing back to the 1920 Census Act

The information, practices and views in this article are those of the author(s) and do not necessarily reflect the opinion of the Royal Geographical Society (with IBG).

(Dewdney & Rhind, 2011). These classifications also reflect data collection constraints, such as the longstanding omission of income due to feasibility and ethical concerns.

Over time, geodemographic classifications have evolved in several ways. First, they have expanded beyond residential data to include workplace geographies, supplementing focus on where people live with where they work (Cockings et al., 2020). Second, they have been used to enable small-area estimates using survey data, such as estimates of fear of crime or travel behaviour (Ashby, 2005; Batey & Brown, 1995; Liu & Cheng, 2020). Third, they have been supplemented with commercial lifestyle data, extending their application to areas like education, policing and healthcare (Singleton, 2010). However, while these classifications highlight self-organising neighbourhood structures, they do not account for how neighbourhood functions shift throughout the day as indicated by human mobility, and remain limited in capturing temporally granular flows.

The increasing availability of Global Positioning Systems (GPS) in-app mobile phone location data allow for detailed analysis of complex place-to-place interactions of individuals beyond home-to-work trajectories. They can provide a more comprehensive view of daily activity patterns across different times of the day, week and year than traditional census and survey approaches (Arribas-Bel & Tranos, 2018; Gibbs et al., 2024). For instance, GPS data have been used to estimate hourly population density and temporally granular population counts to generate ambient population datasets (Deville et al., 2014; Liu et al., 2018), as well as to infer key activity locations, such as home and workplace. This enables a deeper understanding of individual activity locations (Alexander et al., 2015; Jiang et al., 2017).

GPS data sources not only allow for mapping of individuals' locations, but also enable the exploration of origin–destination (OD) flows between places and the temporal rhythms that structure them. OD analysis can be used to characterise human interactions, over the short, medium or long term, in order to facilitate the understanding of connections between places, and has gained significant traction with the advent of GPS location data (Guo et al., 2012; Schneider et al., 2013; Van Dijk et al., 2021; Wang et al., 2019). For instance, Yang et al. (2016) use in-app mobile phone location data for Shenzhen in China to classify areas into clusters exhibiting different characteristics of inbound and outbound human mobility flows throughout the day. Similarly, Xu (2022) applies in-app mobile phone location data to understand how exposure to racial diversity varies throughout the day, revealing that individuals often experience more diverse social environments than residential data alone suggest, and that such exposure is shaped by temporally dynamic activity spaces. Over the longer term, Ge and Fukuda (2016) use aggregated mobile phone data to estimate work-related travel in Tokyo over a five-year period, identifying consistent spatial patterns in commuting. Calafiore et al. (2021) use in-app mobile phone location data with a spatially weighted community detection approach to derive functional neighbourhoods in New York, illustrating how mobility-based methods can reveal dynamic urban boundaries shaped by actual movement patterns. Despite the increasing use of mobile phone data for OD analysis, the primary methodological focus of past research has been on developing and predicting temporally aggregated OD flows, neglecting the spatial dynamics of diurnal variations (Calabrese et al., 2011; Demissie & Kattan, 2022; Graser et al., 2019). This overlooks the finer spatial and temporal area-specific insights that can be obtained from mapping OD flows to leverage their full analytical potential.

Building on these developments, we propose to shift focus from residential geodemographics to interactional geographies. Unlike traditional geodemographic studies, this approach does not seek to classify populations based on residence alone, but rather on understanding place-to-place interactions as they unfold throughout the day. This makes it possible to examine the relationship between human activity and urban spaces (Kempinska et al., 2018; Vich et al., 2017). We therefore use individual level place-to-place interactions from in-app mobile phone location data to effectively visualise OD flows at a spatially and temporally granular scale, with a particular focus on the dominant patterns of interaction flows in three areas of Greater London: Canary Wharf, London Bridge (South Bank) and Kilburn. Through this analysis, we demonstrate how GPS location data can enhance spatial and temporal analyses of urban connectivity, establishing a proof of concept for interactional geographies as an extension to geodemographic classifications of residences or workplaces. The empirical objectives of this study are thus twofold:

1. To demonstrate how temporally and spatially granular OD matrices from in-app data go beyond traditional home-to-work commuting data to enable more comprehensive understanding of interactions between neighbourhoods through the course of the day.
2. To identify how interaction flow indicators can be used to capture an area's connectivity, extent of connections and volume of inbound and outbound flows by time of day, and how these can be integrated into geodemographics.

2 | DATA AND METHODS

2.1 | Study area, data and validation

The individual level GPS in-app location dataset available to us, acquired from a consolidator by the Consumer Data Research Centre (Longley et al., 2018), records use of 700+ apps by approximately 120,000 individuals in 2019 in London. Data are collected when the individual user accesses an app, when the app is running in the background, or both, depending on the specific user settings and privacy controls used by each application. This results in larger volumes of location impressions recorded from devices where the app is running in the background, thus revealing more activity insights. Unlike the pre-aggregated data used in most other studies, having access to individual level in-app impressions gives us the flexibility in aggregating and analysing the data without being constrained by the assumptions built into industry data products. It should also be noted that for commercial and disclosure control reasons, the functions of the applications are not known.

While the full dataset has UK-wide coverage, London is selected as the study area because it contributes 54% of all in-app mobile phone locations in the dataset. Moreover, London's complex and dynamic characteristics also make it a good candidate for empirical analysis (cf. Singleton & Longley, 2015; Wyszomierski et al., 2024). The data were collected between June 2016 and October 2020, but we limit our analysis to 2019 for reasons of computation management as well as to avoid the impact that the COVID-19 pandemic had on mobility patterns. The data include device IDs, application IDs, level of GPS accuracy determined through GPS receivers, longitude, latitude and timestamp of the local date and time of the event. The application ID attribute is used to restrict analysis to the 25 most heavily used applications, which together comprise 99.6% of all impressions recorded in London in 2019. The analysis is further limited to triangulated GPS points with at least 100 metre accuracy, which is common practice given that geolocation accuracy greatly improves with triangulation of position using two or more GPS satellites (Kumar & Dutt, 2020; Wang et al., 2019).

The data originate from self-selecting user bases, with further bias likely arising from user consents to share app use data, and socio-demographic selectivity potentially frustrating population-wide inferences (Lovelace et al., 2016). Elsewhere, research has tried to assess the over- or underrepresentation of the underlying population by using demographic data. Sinclair et al. (2023), for instance, use these mobile phone location data for Glasgow across a period of 3 years to benchmark the assigned home locations against the Acorn geodemographic classification types (see CACI, 2025). Their findings showed that despite the fact that the sample covers only a small percentage of the population, the dataset shows a good socio-demographic representation compared with the population at large. Because our focus is on examining interaction flows and relations between places, we compare our data with footfall estimates from an aggregated call detail record mobile phone location dataset collected by BT (2024), provider to 30% of mobile phone users in the UK under the EE brand (Statista, 2022).

The BT data, referred to as 'BT footfall' data, were obtained from the Greater London Authority (GLA) high street data service, with aggregations made available for a tessellation of 350-metre hexagons. For comparison, we aggregated our in-app mobile phone location data to the same 350-metre hexagonal grid. The results indicate that, in most grid cells, our in-app mobile phone location data consistently capture between 0.00001% and 3% of footfall activity data. Whereas this seems little in comparison, what is important for our current purposes is that our data provide a consistent spatial representation of users across most areas in London. A larger sample than the BT data occurs in less than 1.2% of the hexagonal grids, and this is the case only in parks, such as Bushy Park and Richmond Park, and Heathrow Airport. Together, we believe this to be an acceptable representation of the patterns captured in the much larger footfall dataset. Moreover, where the aggregated BT footfall data obscure OD patterns, our mobile phone in-app dataset preserves individual movements.

2.2 | Method

The first step in deriving place-to-place interactions from the data is to convert the timestamped individual location impressions of each device into activity locations with a start and end time (see Table 1 for raw data structure). We do this by employing spatial and temporal clustering, using agglomerative clustering with the Ward method (Gibbs et al., 2023; Pedregosa et al., 2011). This approach combines raw, timestamped in-app mobile phone location data points that are spatially and temporally close, grouping them into a single visit of known duration for each device. Data points recorded at the same location share (approximate) coordinates, but this proximity may be influenced by GPS-induced imprecision.

When clustering these raw data points into visits (see l_0, l_1 and l_2 in Figure 1), the spatial threshold used for clustering can vary depending on the degree of imprecision in the data. For locations that show greater imprecision, such as those resembling l_0, a smaller threshold is applied, while for locations with less imprecision, like l_2, a larger threshold is used. Clusters of points that are spatially close are then defined as a visit, with the arrival and departure times determined by when the device first and last appears at that specific location.

To define activity locations, we first specify a minimum visit duration based on the distribution of apparent arrivals and departures. A key data cleaning step involves removing implausibly short visits to screen out any non-static impressions that could otherwise introduce noise into the analysis. For this reason, we remove apparent visits of less than 5 min to filter intermediate locations from activity destinations, following accepted practices (Fang et al., 2018; Transportation Research Board and National Academies of Sciences, Engineering, and Medicine, 2018; Xu, 2022). A visit therefore begins when a device remains present at a location for at least 5 min, and the first recorded appearance at a location marks the start of the visit.

To derive flows between different activity locations, we first clean the data and organise observations by device and time of occurrence. This sorting allows us to link a location visited by a device to the next location it moves to, facilitating the formation of interaction flows. After linking origins to destinations, we calculate the time elapsed between leaving one location and arriving at the next. The final cleaned dataset contains almost 17 million activity locations, generating around 10 million flows between locations. In a final step, we remove any rows where the time between visits exceeds 2 hour. We set this two-hour threshold because even the most distant locations in Greater London can be reached within this time. Figure 2 presents the workflow used to obtain individual OD interaction data.

To move from individual OD flows to place-to-place interactions, we aggregate the data to the Middle layer Super Output Area (MSOA) level. A further hourly breakdown is used to understand how interactions vary according to the time of day. We choose to conduct our analysis at this level to accommodate likely issues of positional uncertainty. MSOAs are administrative geographies used in England and Wales, and defined as part of the decennial census to facilitate the reporting of data, and typically consist of 2000–6000 households. While the lower granularity level may limit the depth of analysis, this guarantees adherence to minimum statistical disclosure requirements and ensures that interaction origins have multiple destinations. To protect locational privacy, a threshold of 10 observations is applied within each MSOA, resulting in minimal data loss (Welpton, 2019). The choice of administrative geographies further ensures that census data can be brought into the analysis at a later stage of the research, and facilitates profiling of city-wide patterns consistent with previous research (Friedrich et al., 2010; Schlaich et al., 2010). Crucially, it also provides a foundation for our ambition to extend geodemographic analysis with place-to-place interactions.

To visualise relationships between neighbourhoods, we use the centroids of MSOAs to create Euclidean distance lines connecting origin and destination areas. The centroids serve as reference points for drawing the flow lines, providing a consistent and spatially accurate way to represent the connection between two MSOAs. Each line is then assigned an attribute indicating the total number of flows it represents, providing a clear visualisation of the intensity of interactions between areas. We apply edge bundling in some of the outputs to minimise visual clutter, making overlapping OD flows easier to interpret (see Graser et al., 2019).

TABLE 1 The structure of location visits, in which the median latitude and longitude of raw data points within a cluster are used to define activity locations (data in the table are synthetic).

App ID	User ID	Start datetime	Leaving datetime	Cluster	Visit longitude	Visit latitude
111	12,344	2019-01-01 00:00:05	2019-01-01 12:40:45	1	51.5244	−0.1342

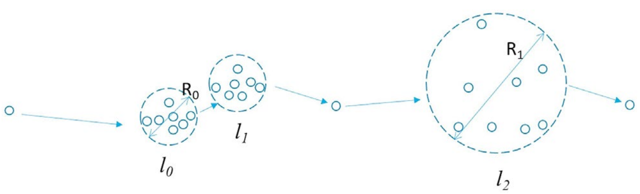

FIGURE 1 Trajectories of three illustrative visits (after Wang et al., 2019).

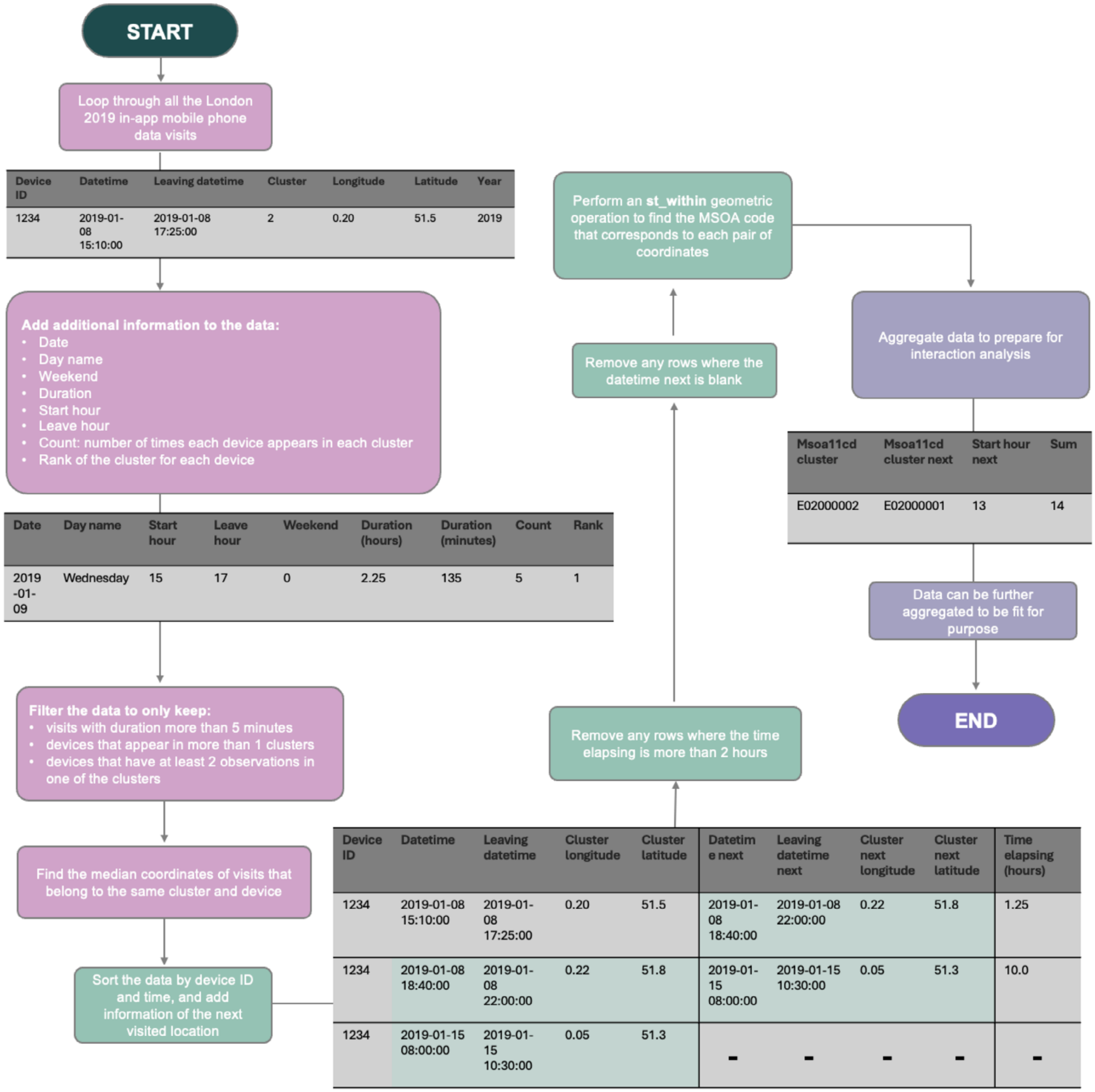

FIGURE 2 Workflow for obtaining interaction flows from GPS-derived activity locations.

3 | RESULTS

3.1 | Visualising place-to-place interactions

Figure 3 illustrates the place-to place-interactions derived from the individual OD mobile phone location data. The most connected MSOAs are London's West End and South Bank, and, while some other parts of Greater London have strong connectivity with these areas, there are also many flows between town centres and MSOAs across the area. For example, Croydon is well connected to central London MSOAs as well as its adjacent MSOAs, while Eltham in south-eastern London shows strong connections with neighbouring areas and Canary Wharf. These observed patterns may manifest good transport links, that make it possible to cover longer distances over any given time window.

FIGURE 3 Edge-bundled non-directional origin–destination (OD) flows between each Middle layer Super Output Area (MSOA) in Greater London, 2019 (minimum threshold = 600 interactions).

The interaction flows can further be broken down temporarily and directionally. Figure 4 shows the intensity of these interactions by time of day, standardised by the area of each borough. Flows may either begin or end in the boroughs shown, highlighting how different parts of the city are engaged at various times.

The most prominent pattern is the high volume of interactions observed between 8 am and 7 pm across the capital. Boroughs such as Tower Hamlets, Southwark, Newham, Lambeth, Islington, Hammersmith & Fulham, and the City of London show clear peaks in activity during the morning and afternoon. The City of London, Camden and Westminster have the most intense activity patterns, while the lowest levels of activity are found in Richmond-upon-Thames, Redbridge, Hillingdon, Havering, Enfield and Bromley. While Figure 4 provides a useful overview of borough-level patterns throughout the day, it offers only limited insights into the multi-dimensional nature of hourly interaction flows. To explore these dynamics in greater detail, Figure 5 maps the origins and destinations of flows during four selected one-hour time windows: midnight, 6 am, 12 noon and 6 pm.

Commencing at 12 midnight, interactions are very low, and the only areas with more than 60 interactions are (a) Oxford Street and Covent Garden and (b) Heathrow and Whitton (see Figure 5a). By 6 am interactions are becoming

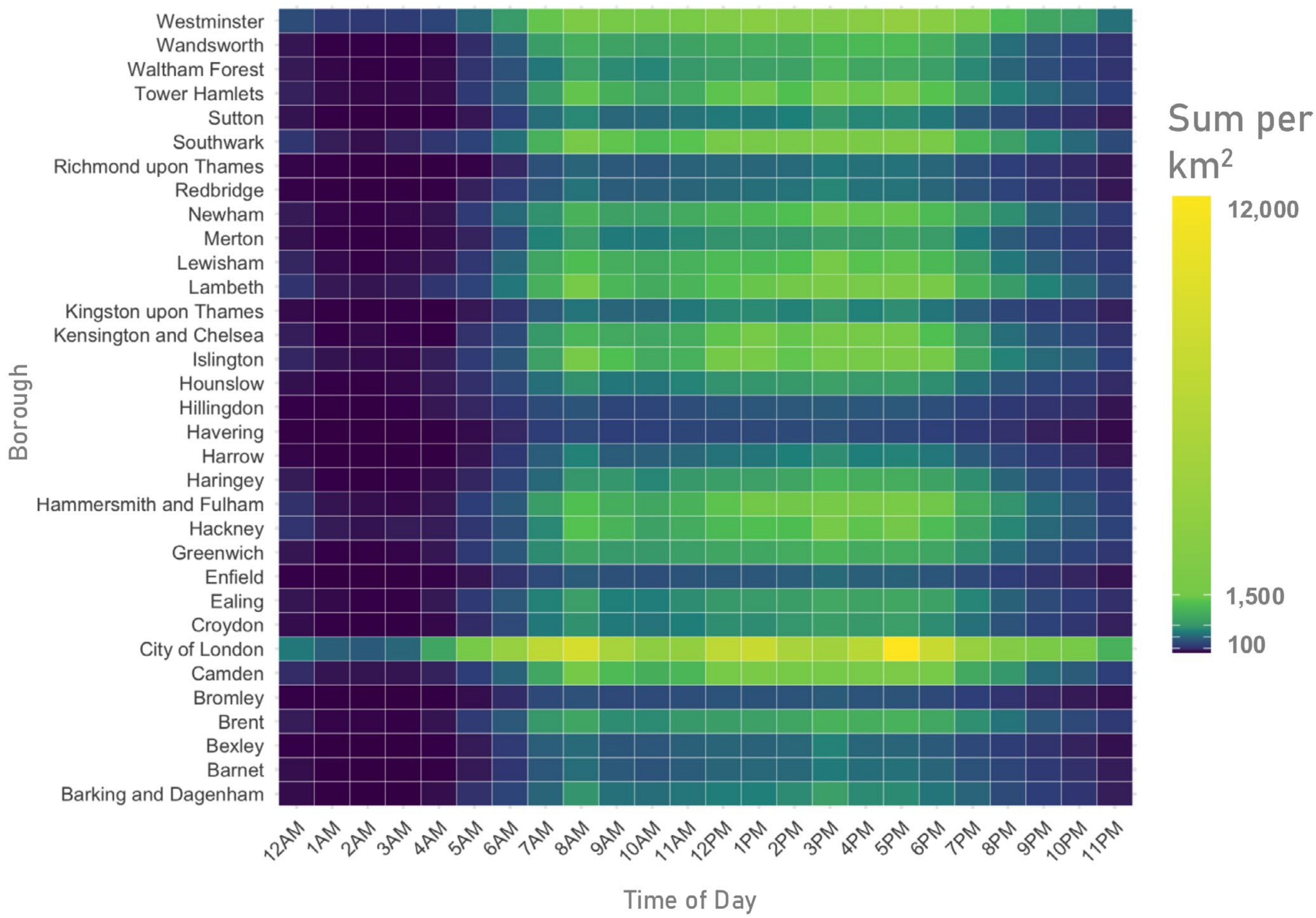

FIGURE 4 Interaction flows per km^2, by borough and time of day.

much higher, manifesting commuting to work, especially between suburban MSOAs and central London MSOAs. Midday interactions are shorter distance, manifesting loci of daytime activities (see Figure 5b,c). By 6 pm, such interactions are supplemented by return commuting (see Figure 5d), indicating the presence of both commuting patterns and other localised flows.

3.2 | Visualising shifting neighbourhood functions

As outlined in the introduction, the main aim of this work is to use detailed individual-level data to better understand how interactions between places unfold over the course of the day. This serves as a proof of concept for how such data can be used to generate insights into place-to-place interaction and the shifting roles of neighbourhoods over time. In this final part of the analysis, we demonstrate how focusing on specific areas can help reveal these changing functions, using directional movement patterns for three case study areas: Canary Wharf (Figure 6), London Bridge (South Bank) (Figure 7) and Kilburn (Figure 8). This analysis assumes that the 8 am and 6 pm flows primarily reflect commuting patterns, and 1 pm and 11 pm flows represent local daily activity and the night-time economy, respectively. The Canary Wharf and London Bridge MSOAs, which are predominantly commercial, experience a net inflow at 8 am and net outflow at 6 pm, while the Kilburn area, which is mostly residential, experiences the opposite pattern.

Figure 6 shows that Canary Wharf experiences a large influx of people at 8 am, principally from Tower Hamlets, Newham and Lambeth. At 6 pm there is a corresponding outflow of individuals to some adjacent MSOAs as well as to further flung suburban locations. The flows at 1 pm and 11 pm are much smaller. Canary Wharf is principally a commercial area but also houses many workers, so while there is an evident net inflow of commuters, there is also some local commuter traffic. Comparing Figures 6 and 7, we observe similarities in the net inflow at 8 am and outflow at 6 pm between Canary Wharf and London Bridge, reflecting their shared role as commercial hubs serviced by daily commuters.

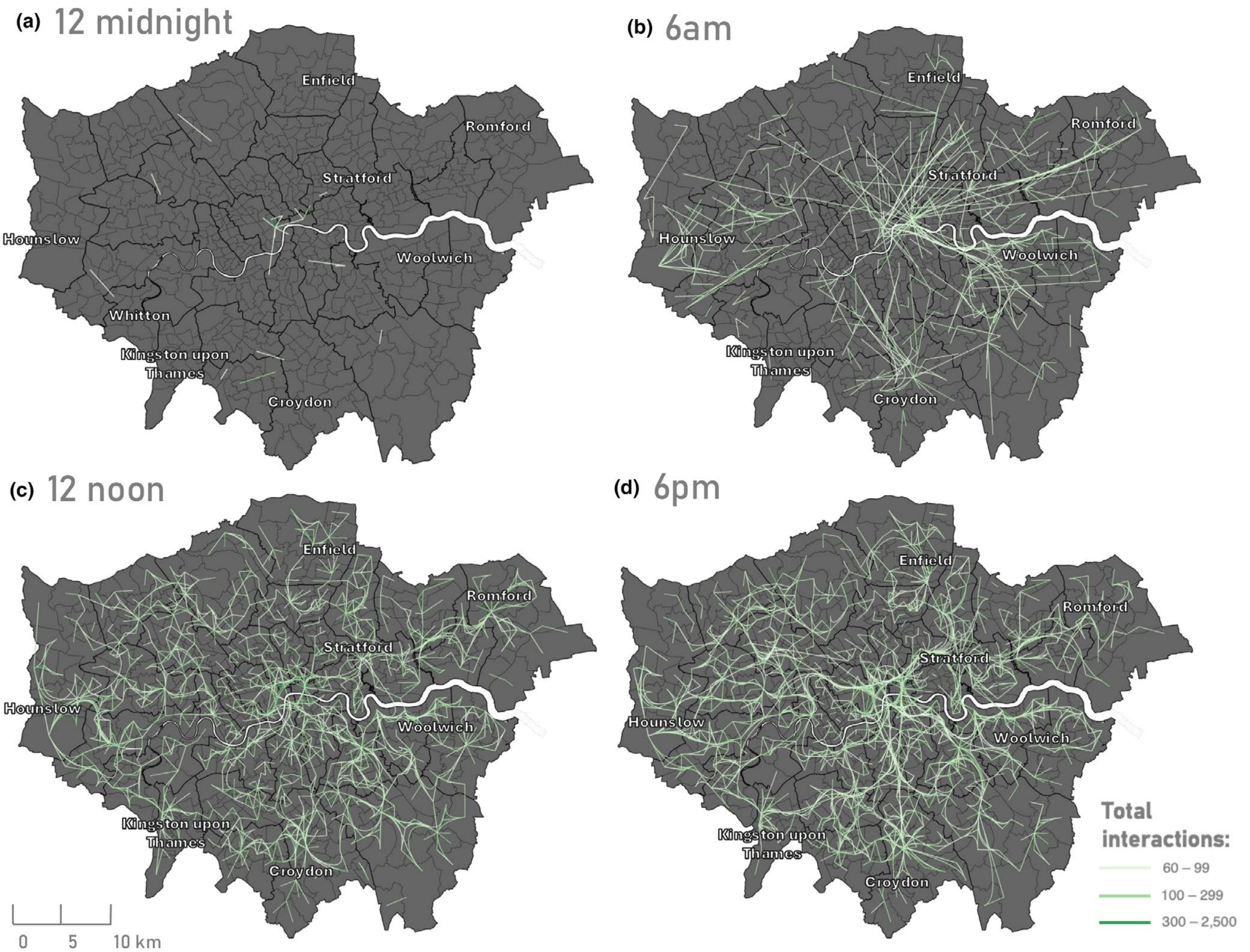

FIGURE 5 Edge-bundled hourly non-directional origin–destination (OD) flows between Middle layer Super Output Areas (MSOAs) in Greater London in 2019 commencing at (a) 12 midnight, (b) 6 am, (c) 12 noon and (d) 6 pm (minimum threshold = 60 interactions).

The key difference between the two is that London Bridge's flows are largely directed towards South London, with strong links to Croydon, whilst Canary Wharf's flows are most pronounced with northern and central London neighbourhoods. At 6 pm, London Bridge's outflows demonstrate a strong connection with east and south-east London, whereas its 8 am inflows are more widely distributed across commuting areas. Many commuters to London Bridge come from Newham or Tower Hamlets in east London, or Mill Hill, East Barnet, Harrow, Woodford and Newbury in north London. Canary Wharf and London Bridge show similar inflow and outflow volumes during peak commute hours. What distinguishes the two areas is the extent and location of the other areas to which they are connected.

For Kilburn (Figure 8), inflows and outflows are shown simultaneously rather than separately. We focus on a residential MSOA in Kilburn to examine net flows, where negative values indicate net outflows and positive values indicate net inflows. This approach allows for a straightforward assessment of the overall balance of movement into and out of the area at different times of day, providing a clearer picture of its changing function within the wider urban context. Net flows in Kilburn are relatively small and largely confined to the same or neighbouring boroughs. The only notable exception occurs around 8 am, when there is a net outflow to more distant MSOAs such as Marylebone, Soho, Whitechapel, Bayswater, Edgware and Northwood. In contrast, inflows at that time originate from nearby areas including Willesden, Wembley and Marble Arch, with no significant outflows to those locations. By 1 pm, net inflows are observed from Hammersmith and Kensington, while net outflows are directed towards Wembley and Brent Cross. At 6 pm, inflows exceed outflows, and by 11 pm, all recorded flows are net inflows. Given that any flows between MSOAs of less than 10 are excluded from the analysis due to disclosure control, Kilburn appears to be locally embedded within its immediate surroundings but has limited connectivity to more destinations. Overall, Kilburn's flow volumes are evidently much lower than those of Canary Wharf and London Bridge.

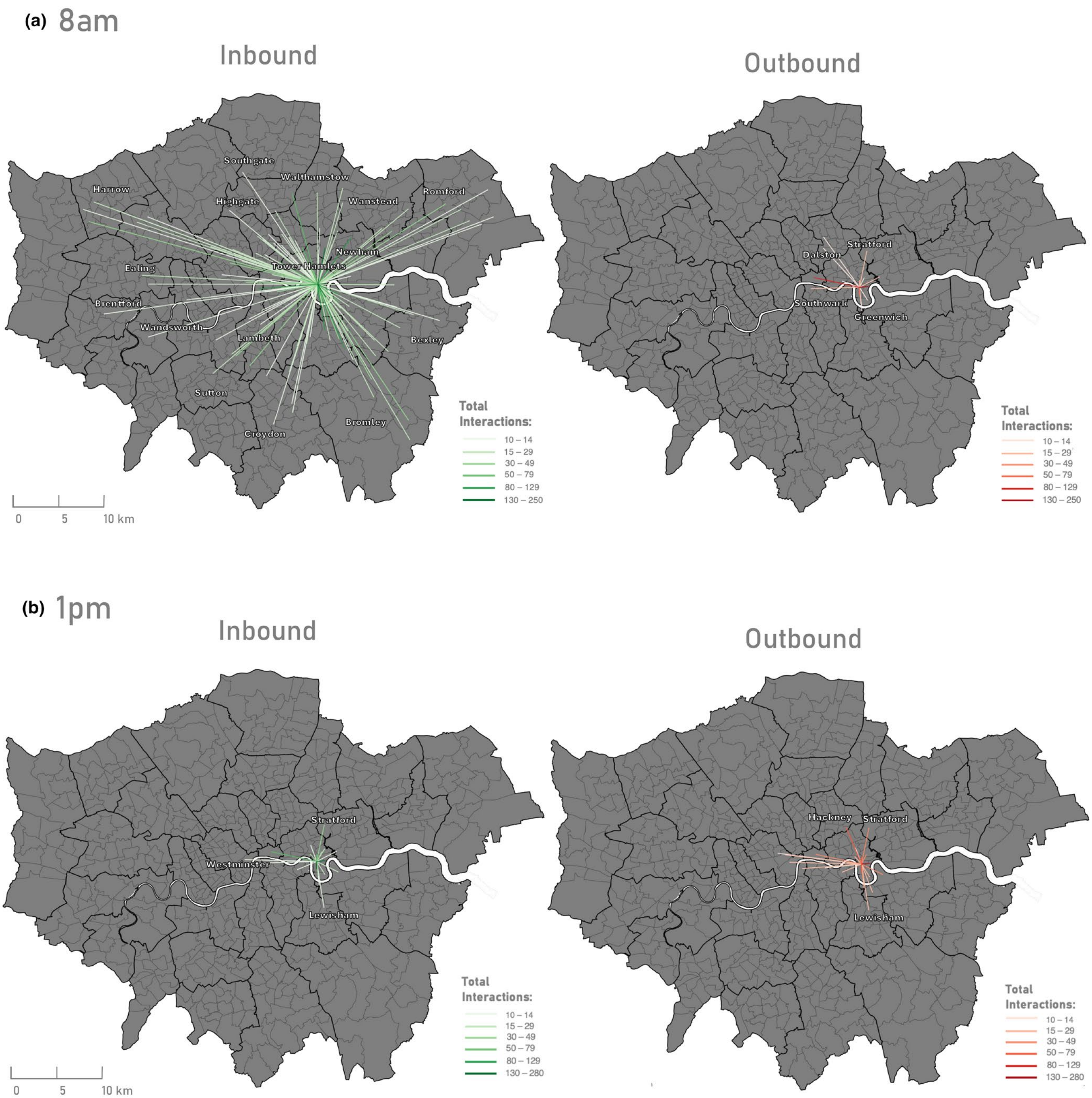

FIGURE 6 Hourly directional interactions terminating or originating in Canary Wharf and other MSOAs beginning at (a) 8 am, (b) 1 pm, (c) 6 pm and (d) 11 pm.

4 | DISCUSSION

The MSOA-interaction analysis demonstrates the useability of mobile phone location data for understanding connectivity, reach and links between locations throughout the day. This approach facilitates the exploration of how different areas assume varying functional roles as a result of fluctuations in human mobility. By providing a temporally granular breakdown of interaction flows, it offers insights into the dynamic aspects of these movements. As demonstrated in Section 3.2, this method allows us to focus on individual MSOAs, providing a nuanced view of how places function differently at different times of the day.

Unlike the Census, which captures only home-to-work trajectories without timestamps (Martin et al., 2018), GPS location data provide a more comprehensive view of daily activity patterns across different times of the day, week and

FIGURE 6 (Continued)

year. The finer temporal granularity of GPS location data can equip policymakers to understand how urban spaces function dynamically, informing more responsive urban planning, transport policies and service provision. Additionally, the 2021 Census data are heavily influenced by the COVID-19 pandemic, making it them representative of typical mobility patterns (Harrington & Hadjiconstantinou, 2022). By using 2019 GPS location data—the most recent pre-pandemic year available to us—we ensure that our analysis reflects more stable and generalisable trends.

The directional interaction flow analysis for Canary Wharf, London Bridge and Kilburn reveals significant variations in interaction patterns across the different areas, indicating the diverse functioning of labour markets and the integrity of neighbourhood structures. While Kilburn exhibits a limited number of interaction flows with neighbouring regions, Canary Wharf and London Bridge display a broader connectivity, reaching areas further away. Despite their extensive connections, these areas demonstrate distinct interaction profiles, bringing focus to urban dynamics and multifaceted urban landscapes. Even though the exact purpose of an interaction flow may be unknown, when the general function of

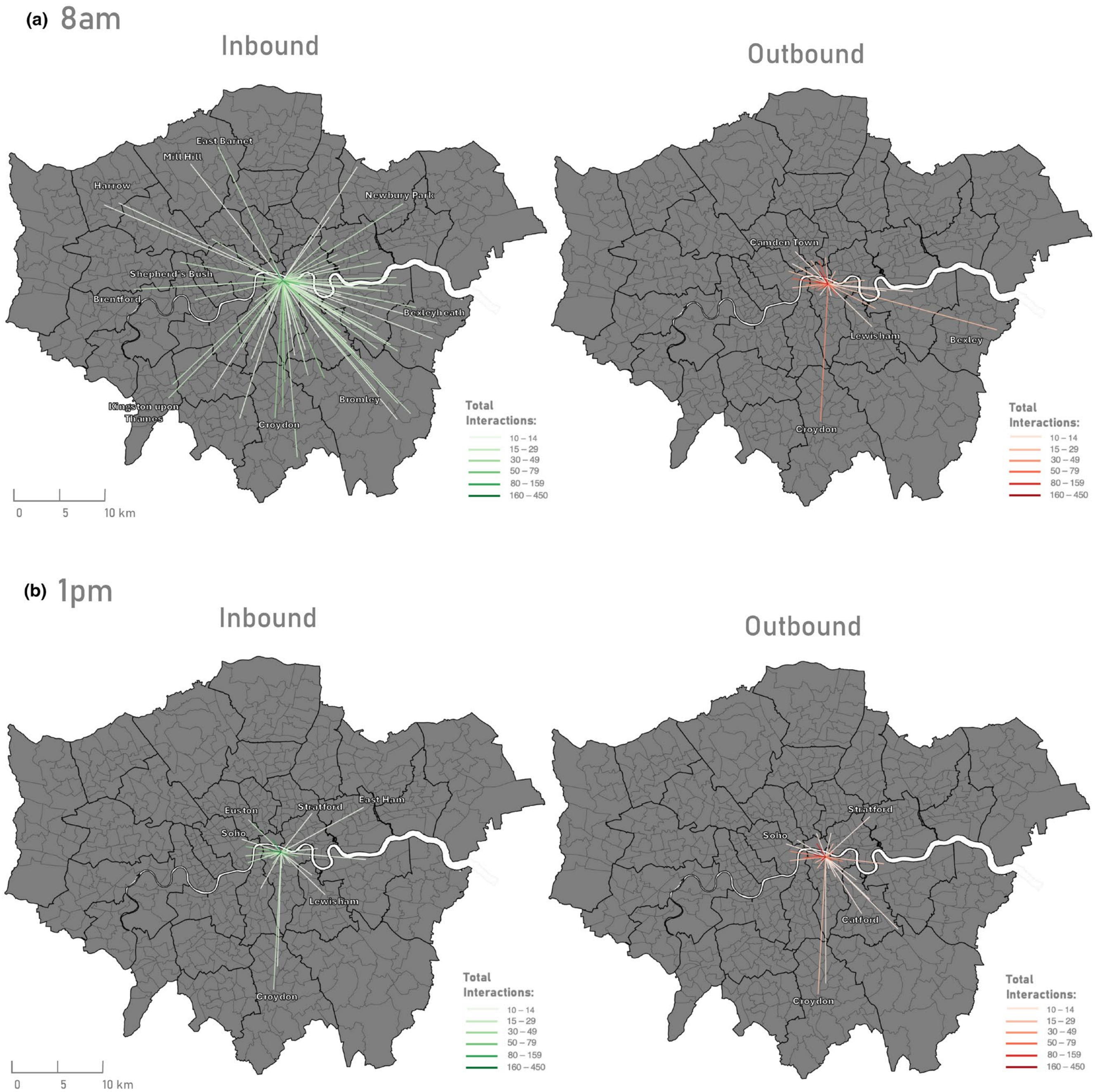

FIGURE 7 Hourly directional interactions between London Bridge and other MSOAs in Greater London in 2019 commencing at (a) 8 am, (b) 1 pm, (c) 6 pm and (d) 11 pm.

a location is coupled with time of day, it is possible to make inferences between different types of flows and to distinguish between a night-worker flow and a leisure flow, at 11 pm for example. This analysis can thus inform decision-making requiring identification of neighbourhoods with differing connectivity characteristics—for example, those with low activity and poor connectivity that might be upgraded to serve community needs.

Building directly upon this work, our future research will formalise interaction flows as summary indicators that will capture an area's connectivity, extent of connections and volume of inbound and outbound flows by time of day. This will allow us to extend conventional geodemographics by incorporating the variegated functional characteristics of any study area. Combined with socio-economic and demographic data, new geodemographic classifications will move beyond understanding place solely in terms of social similarities to also include the dynamic interaction flows that shape and define unique neighbourhood contexts.

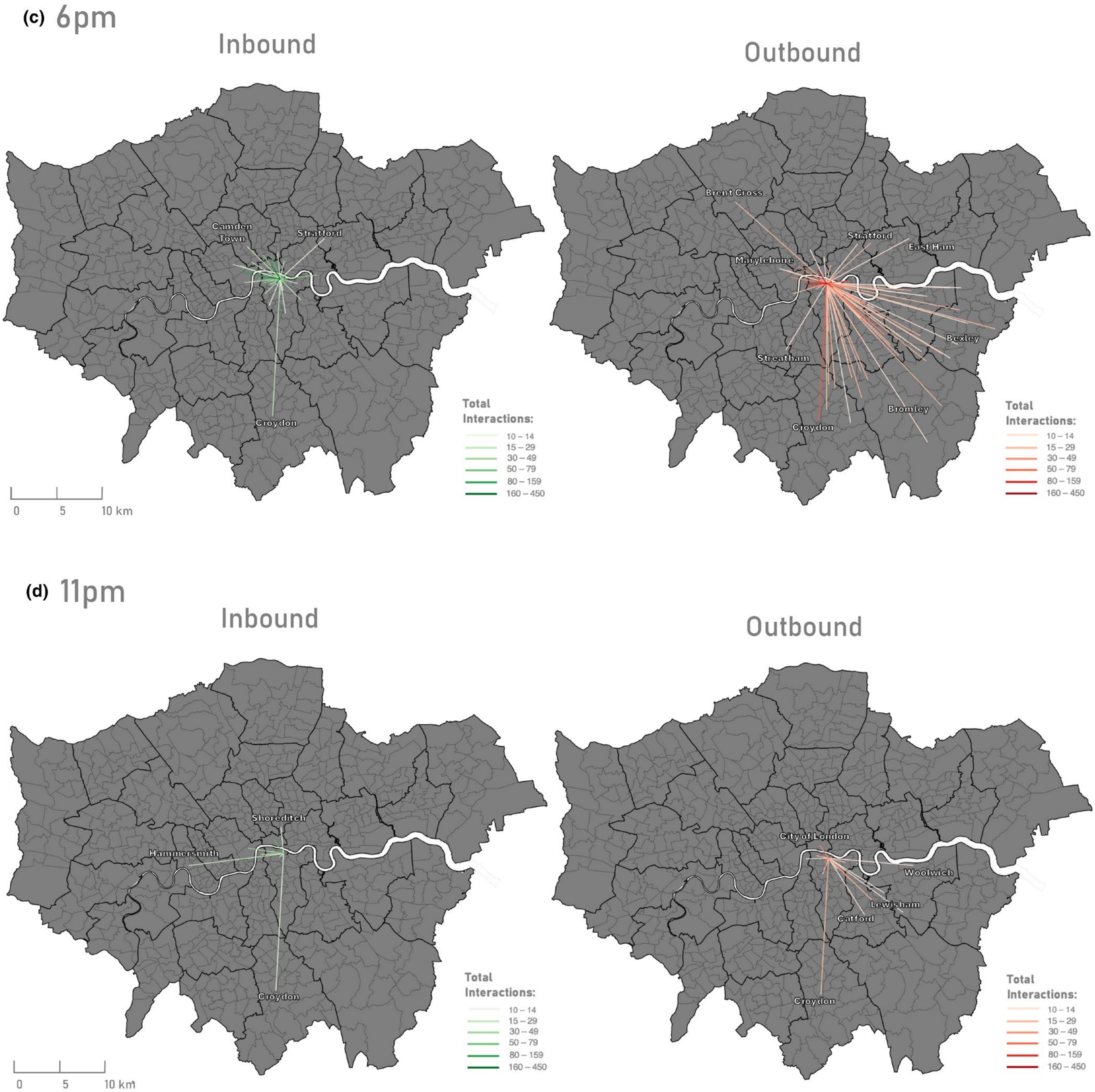

FIGURE 7 (Continued)

The underlying data enable calculation of transitions from origins to destinations at the MSOA level, which is not possible using pre-aggregated app or mobile phone data products. Unlike aggregated location data made available by data consolidators, OD patterns are not obscured, and spatial as well as temporal flexibility is preserved, enabling more bespoke analysis of how neighbourhoods function as part of a city system. By preserving granular movement data, we provide a richer understanding of urban connectivity, overcoming the limitations of commercial aggregation that often lacks methodological transparency. This shifts the focus from profiling the attributes of areas to profiling the functional interdependencies between them; while also acknowledging the temporal component of these place-to-place interactions. Through aggregation at the MSOA level, the study minimises the impact of GPS inaccuracies given that the coarser geographical scale reduces uncertainty and improves the useability of the dataset. Timeframes can similarly be scaled to intervals that are sensitive to context. This additionally addresses ethical considerations regarding the sensitive nature of the individual level data (cf. Sieg et al., 2024). A remaining limitation, however, is that adherence to minimum

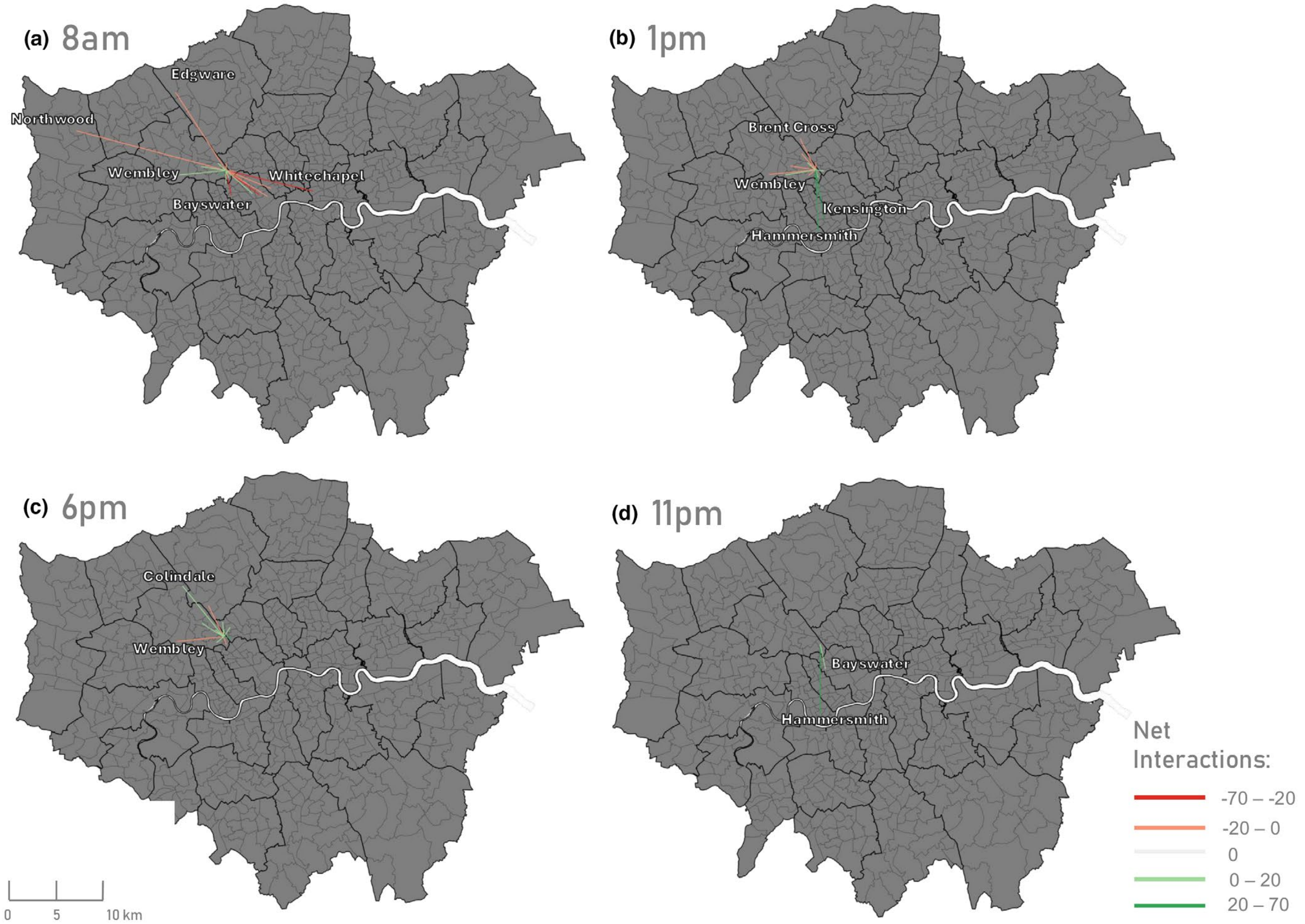

FIGURE 8 Hourly net origin–destination (OD) flows for the interactions between Kilburn and other London Middle layer Super Output Areas (MSOAs) in 2019, commencing at (a) 8 am, (b) 1 pm, (c) 6 pm and (d) 11 pm.

disclosure control requirements requires that insights of less prevalent interaction flows can only be revealed when using a larger dataset. Notwithstanding these challenges, urban planners and local policy makers can benefit from these types of analyses.

Further research might expand the analysis to explore seasonal variations as well as weekday and weekend interaction flows to provide additional insights for urban planning and transportation management. The two-hour threshold used as a cut-off in the analysis for the time elapsing between leaving a location and first appearing at the next location could be increased to assess how it will affect the results. Our analysis is limited to OD flows within London, hence one direction for future research would be to extend the analysis to examine flows into and out of Greater London, which is particularly useful for understanding external flow patterns in terms of both volume and direction. Additionally, the disaggregate nature of the data means that information on location types might be used to filter analysis to only leisure (home to leisure) or commuting (home to office) interaction flows. Lastly, distance and speed thresholds could also be used to achieve the segmentation of OD flows by mode of transport, such as foot, car, bus and train.

5 | CONCLUSION

Geodemographic classifications have long supported general-purpose analysis of populations based on their residential locations, using variables from traditional censuses. This paper has explored how understanding spatial behaviour can be enriched by incorporating interaction flows, rendering geodemographics amenable to investigation of the ways in which areas interact through individual activity patterns. This marks a shift from the analysis of static urban form, to a more dynamic understanding of urban function. Our contribution is to demonstrate how in-app mobile phone location data can be used to understand the functional and interactional characteristics of an area throughout the day using OD analysis, something that will be incorporated in geodemographics in the next phase of this research. Through this study, we use interaction matrices as an analytical tool to explore urban flows dynamically across both time and space. Our results illustrate how different MSOAs, such as those in Canary Wharf, London Bridge and Kilburn, take on varying functional roles throughout the day as a result of human mobility fluctuations. This moves beyond the Census' home and work dichotomy to an area-based perspective of interaction flows.

Despite the self-selective nature of mobile phone location data, our findings show that OD analysis remains a valuable tool for distinguishing interaction patterns and understanding urban mobility dynamics. Given declining participation in travel surveys, mobile phone location data provide a promising way of analysing place-to-place interactions across continuous spatial and temporal scales despite limitations arising from sample size. We note that standardised, pre-aggregated data products risk obscuring important spatial and temporal variations, through their 'one size fits all' approach, and instead emphasise the benefits of using disaggregated data that allow flexibility in manipulation. By conducting analysis at the MSOA level, we mitigate issues of positional uncertainty arising from GPS data, meet disclosure control requirements, and facilitate linkage to demographic data. Ultimately, this study lays the groundwork for a new generation of geodemographic classifications, that incorporate mobility into holistic representation of the configuration as well as socioeconomic composition of neighbourhoods. The methodology of the study is also scalable and adaptable to other mobile phone location datasets where device IDs are known.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Dr Hamish Gibbs for assisting with the pre-processing of the data used in the current research using agglomerative clustering and scikit learn. The data for this research have been provided by the Geographic Data Service (GeoDS), an Economic and Social Research Council (ESRC) Smart Data Research Investment, under project ID ES/Z504464/1. We also disclose receipt of funding for the research from the ESRC UBEL Doctoral Training Partnership, and Didobi, through the Collaborative Studentships doctoral projects initiative (https://ubel-dtp.ac.uk/co-fundedandcollaborative/).

DATA AVAILABILITY STATEMENT

The dataset used in this study was collected in compliance with GDPR regulations and participants in the study have provided informed consent for the use of their data by partner organisations and for research purposes. Participants are smartphone users of mobile phone applications that have data sharing agreements with the data provider. Consent was obtained via agreement to the terms of service, and all participants are anonymised using hashed IDs. The data are securely held by the Geographic Data Service (GeoDS) and are accessible in an aggregated format through the Secure tier of its data service for public good research applications following successful application (see https://data.geods.ac.uk/dataset/huq-aggregated-in-app-location-dataset). The service's Trusted Research Environment, the Data Safe Haven (DSH), is certified to ISO 27001. Researchers can request access through the GeoDS under controlled conditions. Data processing is carried out under the public interest derogation for research (Article 89 of General Data Protection Regulation). In order to release outputs from the DSH, outputs are subject to Statistical Disclosure Control (SDC) checks by two independent GeoDS Data Scientists, to ensure no personal or identifiable information is disclosed. Rigorous procedures ensure that privacy and ethical considerations are addressed throughout, from data handling to the dissemination of research outputs.

ETHICS STATEMENT

Ethical guidelines have been followed and ethical approvals have been obtained through the University College London Research ethics service (ref: 24007/001).

ORCID

Mikaella Mavrogeni https://orcid.org/0009-0000-3204-0128
Justin van Dijk https://orcid.org/0000-0001-5496-425X
Paul Longley https://orcid.org/0000-0002-4727-6384

REFERENCES

Alexander, L., Jiang, S., Murga, M. & González, M.C. (2015) Origin–destination trips by purpose and time of day inferred from mobile phone data. *Transportation Research Part C: Emerging Technologies*, 58, 240–250. Available from: https://doi.org/10.1016/j.trc.2015.02.018

Arribas-Bel, D. & Tranos, E. (2018) Characterizing the spatial structure(s) of cities "on the fly": The space-time calendar. *Geographical Analysis*, 50, 162–181. Available from: https://doi.org/10.1111/gean.12137

Ashby, D.I. (2005) Policing neighbourhoods: Exploring the geographies of crime, policing and performance assessment. *Policing and Society*, 15, 413–447. Available from: https://doi.org/10.1080/10439460500310079

Batey, P.J. & Brown, P. (1995) From human ecology to customer targeting: The evolution of geodemographics. In: Longley, P.A. & Clarke, G.P. (Eds.) *GIS for business and service planning*. Cambridge, MA: GeoInformation International.

BT. (2024) Anonymised & aggregated by BT mobility. Available from: https://data.london.gov.uk/high-street-data-service/

CACI. (2025) Acorn: A geodemographic segmentation of the UK's population. Available from: https://acorn.caci.co.uk/

Calabrese, F., Di Lorenzo, G., Liu, L. & Ratti, C. (2011) Estimating origin-destination flows using Mobile phone location data. *IEEE Pervasive Computing*, 10, 36–44. Available from: https://doi.org/10.1109/MPRV.2011.41

Calafiore, A., Palmer, G., Comber, S., Arribas-Bel, D. & Singleton, A. (2021) A geographic data science framework for the functional and contextual analysis of human dynamics within global cities. *Computers, Environment and Urban Systems*, 85, 101539. Available from: https://doi.org/10.25436/E26K5F

Cockings, S., Martin, D. & Harfoot, A. (2020) Developing a National Geodemographic Classification of workplace zones. *Applied Spatial Analysis*, 13, 959–983. Available from: https://doi.org/10.1007/s12061-020-09337-4

Demissie, M.G. & Kattan, L. (2022) Estimation of truck origin-destination flows using GPS data. *Transportation Research Part E: Logistics and Transportation Review*, 159, 102621. Available from: https://doi.org/10.1016/j.tre.2022.102621

Deville, P., Linard, C., Martin, S., Gilbert, M., Stevens, F.R., Gaughan, A.E. et al. (2014) Dynamic population mapping using mobile phone data. *Proceedings. National Academy of Sciences. United States of America*, 111, 15888–15893. Available from: https://doi.org/10.1073/pnas.1408439111

Dewdney, J.C. & Rhind, D. (2011) The British and United States' censuses of population. In: Pacione, M. (Ed.) *Population geography: Progress and prospect*. Abingdon, Oxon: Routledge.

Dorling, D. & Simpson, S. (1998) Statistics in society: The arithmetic of politics. In: Dorling, D. & Simpson, S. (Eds.) *Statistics in society: The arithmetic of politics*. London: Arnold Publishers.

Fang, Z., Jian-yu, L., Jin-jun, T., Xiao, W. & Fei, G. (2018) Identifying activities and trips with GPS data. *IET Intelligent Transport Systems*, 12, 884–890. Available from: https://doi.org/10.1049/iet-its.2017.0405

Friedrich, M., Immisch, K., Jehlicka, P., Otterstätter, T. & Schlaich, J. (2010) Generating Origin–Destination Matrices from Mobile Phone Trajectories. *Transportation Research Record: Journal of the Transportation Research Board*, 2196, 93–101. Available from: https://doi.org/10.3141/2196-10

Gale, C.G., Singleton, A.D., Bates, A.G. & Longley, P.A. (2016) Creating the 2011 area classification for output areas (2011 OAC). *Journal of Spatial Information Science*, 12(12), 1–27. Available from: https://doi.org/10.5311/JOSIS.2016.12.232

Ge, Q. & Fukuda, D. (2016) Updating origin–destination matrices with aggregated data of GPS traces. *Transportation Research Part C: Emerging Technologies*, 69, 291–312. Available from: https://doi.org/10.1016/j.trc.2016.06.002

Gibbs, H., Ballantyne, P., Cheshire, J., Singleton, A. & Green, M.A. (2024) Harnessing mobility data to capture changing work from home behaviours between censuses. *Geographical Journal*, 190, e12555. Available from: https://doi.org/10.1111/geoj.12555

Gibbs, H., Eggo, R.M. & Cheshire, J. (2023) Detecting behavioural bias in GPS location data collected by mobile applications. *medRxiv*. https://doi.org/10.1101/2023.11.06.23298140

Graser, A., Schmidt, J., Roth, F. & Brändle, N. (2019) Untangling origin-destination flows in geographic information systems. *Information Visualization*, 18, 153–172. Available from: https://doi.org/10.1177/1473871617738122

Guo, D., Zhu, X., Jin, H., Gao, P. & Andris, C. (2012) Discovering spatial patterns in origin-destination mobility data. *Transactions in GIS*, 16, 411–429. Available from: https://doi.org/10.1111/j.1467-9671.2012.01344.x

Harrington, D.M. & Hadjiconstantinou, M. (2022) Changes in commuting behaviours in response to the COVID-19 pandemic in the UK. *Journal of Transport & Health*, 24, 101313. Available from: https://doi.org/10.1016/j.jth.2021.101313

Harris, R., Sleight, P. & Webber, R. (2005) *Geodemographics, GIS and neighbourhood targeting, mastering GIS*. West Sussex, UK; Hoboken, N.J.: Wiley.

Jiang, S., Ferreira, J. & Gonzalez, M.C. (2017) Activity-based human mobility patterns inferred from mobile phone data: A case study of Singapore. *IEEE Transactions on Big Data*, 3, 208–219. Available from: https://doi.org/10.1109/TBDATA.2016.2631141

Kempinska, K., Longley, P. & Shawe-Taylor, J. (2018) Interactional regions in cities: Making sense of flows across networked systems. *International Journal of Geographical Information Science*, 32, 1348–1367. Available from: https://doi.org/10.1080/13658816.2017.1418878

Kumar, S. & Dutt, S. (2020) The global positioning system: Popular accuracy measures. *Materials Today Proceedings*, 33, 4797–4801. Available from: https://doi.org/10.1016/j.matpr.2020.08.380

Liu, Y. & Cheng, T. (2020) Understanding public transit patterns with open geodemographics to facilitate public transport planning. *Transportmetrica A: Transport Science*, 16(1), 76–103. Available from: https://doi.org/10.1080/23249935.2018.1493549

Liu, Z., Ma, T., Du, Y., Pei, T., Yi, J. & Peng, H. (2018) Mapping hourly dynamics of urban population using trajectories reconstructed from mobile phone records. *Transactions in GIS*, 22, 494–513. Available from: https://doi.org/10.1111/tgis.12323

Longley, P. (2015) *Geographic information science & systems, Fourth edition*. (Longley, P.). Hoboken, NJ: Wiley.

Longley, P.A., Cheshire, J. & Singleton, A. (Eds.). (2018) *Consumer data research*. London, UK: UCL Press. Available from: https://doi.org/10.14324/978-1-78735-388-6

Louckx, K. & Vanderstraeten, R. (2015) Household and state-istics: Cornerstones of Society in Population Censuses (Belgium, 1846–1947). *Social Science History*, 39, 201–215. Available from: https://doi.org/10.1017/ssh.2015.51

Lovelace, R., Birkin, M., Clarke, M. & Cross, P. (2016) From big noise to big data: Towards the verification of large datasets for understanding regional retail flows. *Geographical Analysis*, 48, 59–81. Available from: https://doi.org/10.1111/gean.12081

Martin, D., Gale, C., Cockings, S. & Harfoot, A. (2018) Origin-destination geodemographics for analysis of travel to work flows. *Computers, Environment and Urban Systems*, 67, 68–79. Available from: https://doi.org/10.1016/j.compenvurbsys.2017.09.002

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O. et al. (2011) Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12, 2825–2830.

Schlaich, J., Otterstätter, T. & Friedrich, M. (2010) Generating trajectories from Mobile phone data, presented at 89th annual meeting of the Transportation Research Board, Washington, D.C.

Schneider, C.M., Belik, V., Couronné, T., Smoreda, Z. & González, M.C. (2013) Unravelling daily human mobility motifs. *J. R. Soc. Interface*, 10, 20130246. Available from: https://doi.org/10.1098/rsif.2013.0246

Sieg, L., Gibbs, H., Gibin, M. & Cheshire, J. (2024) Ethical challenges arising from the mapping of mobile phone location data. *The Cartographic Journal*, , 61, 1–14. Available from: https://doi.org/10.1080/00087041.2024.2318055

Sinclair, M., Maadi, S., Zhao, Q., Hong, J., Ghermandi, A. & Bailey, N. (2023) Assessing the socio-demographic representativeness of mobile phone application data. *Applied Geography*, 158, 102997. Available from: https://doi.org/10.1016/j.apgeog.2023.102997

Singleton, A.D. (2010) The geodemographics of educational progression and their implications for widening participation in higher education. *Environment and Planning A*, 42(11), 2560–2580. Available from: https://doi.org/10.1068/a42394

Singleton, A.D. & Longley, P. (2015) The internal structure of greater London: A comparison of national and regional geodemographic models: The internal structure of greater London. *Geography and Environment*, 2, 69–87. Available from: https://doi.org/10.1002/geo2.7

Statista. (2022) Most used mobile network providers in the United Kingdom (UK) in 2022. Accessed 15th 8 2024.

Transportation Research Board & National Academies of Sciences, Engineering, and Medicine. (2018) *Cell phone location data for travel behavior analysis*. Washington, DC: The National Academies Press.

Van Dijk, J.T., Lansley, G. & Longley, P.A. (2021) Using linked consumer registers to estimate residential moves in the United Kingdom. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 184, 1452–1474. Available from: https://doi.org/10.1111/rssa.12713

Vich, G., Marquet, O. & Miralles-Guasch, C. (2017) Suburban commuting and activity spaces: Using smartphone tracking data to understand the spatial extent of travel behaviour. *Geographical Journal*, 183, 426–443. Available from: https://doi.org/10.1111/geoj.12220

Wang, F., Wang, J., Cao, J., Chen, C. & Ban, X. (2019) Extracting trips from multi-sourced data for mobility pattern analysis: An app-based data example. *Transportation Research Part C: Emerging Technologies*, 105, 183–202. Available from: https://doi.org/10.1016/j.trc.2019.05.028

Welpton, R. (2019) Statistical disclosure control handbook. https://doi.org/10.6084/M9.FIGSHARE.9958520

Wyszomierski, J., Longley, P.A., Singleton, A.D., Gale, C. & O'Brien, O. (2024) A neighbourhood output area classification from the 2021 and 2022 UK censuses. *Geographical Journal*, 190, e12550. Available from: https://doi.org/10.1111/geoj.12550

Xu, W. (2022) The contingency of neighbourhood diversity: Variation of social context using mobile phone application data. *Urban Studies*, 59, 851–869. Available from: https://doi.org/10.1177/00420980211019637

Yang, X., Zhao, Z. & Lu, S. (2016) Exploring Spatial-Temporal Patterns of Urban Human Mobility Hotspots. *Sustainability*, 8, 674. Available from: https://doi.org/10.3390/su8070674

How to cite this article: Mavrogeni, M., van Dijk, J. & Longley, P. (2025) Understanding place-to-place interactions using flow patterns derived from in-app mobile phone location data. *The Geographical Journal*, 00, e70033. Available from: https://doi.org/10.1111/geoj.70033