

RESEARCH ARTICLE

Mobility data resolution needed to inform predictive models of spatial epidemic spread from mobile phone data

Giulia Pullano¹, Shweta Bansal¹, Stefania Rubrichi², Vittoria Colizza^{3*}

1 Department of Biology, Georgetown University, Washington, District of Columbia, United States of America, **2** Orange Research - SENSE, Châtillon, France, **3** Sorbonne Université, INSERM, Sorbonne Université Modeling Outbreaks Center (SUMOC), Institut Pierre Louis d'Epidémiologie et de Santé Publique, Paris, France

* vittoria.colizza@inserm.fr



OPEN ACCESS

Citation: Pullano G, Bansal S, Rubrichi S, Colizza V (2026) Mobility data resolution needed to inform predictive models of spatial epidemic spread from mobile phone data. PLoS Comput Biol 22(7): e1014427. <https://doi.org/10.1371/journal.pcbi.1014427>

Editor: Shi Zhao, Tianjin Medical University, CHINA

Received: November 3, 2025

Accepted: June 9, 2026

Published: July 2, 2026

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pcbi.1014427>

Copyright: © 2026 Pullano et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution,

Abstract

Human mobility fundamentally shapes the spatial spread of infectious diseases, yet the level of detail required from mobility data to accurately inform epidemic models remains unclear. Mobile phone records offer unprecedented resolution on population movements, but little attention has been devoted however to determining (i) which aspects of mobility are epidemiologically relevant and (ii) what level of data resolution is necessary to capture spatial invasion dynamics. Using mobile phone records from 9.5 million users in Senegal (approximately 80% of the population), we systematically compare three approaches to aggregating mobility data for epidemic modeling. These approaches span a range of resolutions: high-resolution tracking of all individual displacements between consecutive visited locations (HR), medium-resolution accounting for time spent in all visited locations (MR), and low-resolution identification of the most-visited location (LR). We incorporate these mobility representations into a metapopulation epidemic model that explicitly accounts for transmission from residents, visitors, and returning travelers, and simulate diseases with varying transmissibility corresponding to controlled epidemic conditions, seasonal influenza-like transmission, and highly transmissible pathogens. We find that preserving all observed displacements in individual trajectories does not necessarily improve the epidemiological relevance of mobility in pathogens spatial transmission. Instead, displacement-based networks fragment long-range trips and underestimate key spatial connections relevant for disease spread. In contrast, approaches that capture where individuals spend most of their time (such as home, work, or school) more accurately reproduce spatial invasion patterns. Accounting for additional daily activities beyond these primary locations provides little additional epidemiological information. Our results suggest that lower-resolution mobility indicators capturing time spent at key locations are sufficient to inform predictive epidemic models. These findings have important implications for both epidemic modeling and data governance,

and reproduction in any medium, provided the original author and source are credited.

Data availability statement: The code and data for all analyses are publicly available at <https://github.com/EPIcx-lab/Senegal-coupling>. This repository includes monthly coupling matrices HR, MR, and LR at municipality level, and the metapopulation model code, with census data used to design the model.

Funding: This study was partially funded by: Agence Nationale de la Recherche project DATAREDUX (ANR-19-CE46-0008-03 to VC); Horizon Europe grants ESCAPE (101095619 to VC) and VERDI (101045989 to VC); EU Horizon 2020 grant MOOD (H2020-874850 to VC). GP's salary was partially supported by the Fritz-Family fellowship program from Georgetown University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The contents of this publication are the sole responsibility of the authors and don't necessarily reflect the views of the European Commission.

Competing interests: The authors have declared that no competing interests exist.

indicating that mobile phone data can be aggregated to reduce privacy issues while still providing the essential information needed to model spatial disease transmission.

Author summary

Human mobility patterns play a crucial role in the spatial spread of infectious diseases, as pathogens move between communities through the movements of infected individuals. During the COVID-19 pandemic, mobile phone companies shared unprecedented amounts of mobility data to support epidemic modeling. However, important questions remain: which aspects of human mobility are truly necessary to capture spatial disease transmission, and how detailed must mobility data be to inform epidemic models? Using mobile phone data from 9.5 million users in Senegal, we systematically compared three ways of summarizing individual mobility patterns for epidemic modeling. These approaches ranged from high-resolution tracking of every individual displacement to simpler representations that identify only the locations where people spend most of their time, such as home, work, or school. We found that tracking every movement does not improve epidemic predictions. Instead, the most informative mobility patterns for modeling spatial disease spread are the ones connecting locations where individuals spend sustained periods of time. Approaches that capture the connections between these key locations reproduce epidemic invasion patterns more accurately than those that record every displacement. These results suggest that highly detailed mobility data may not be necessary for epidemic modeling. Instead, aggregated mobility indicators capturing time spent at major activity locations can provide sufficient information to model spatial transmission. This finding has important implications for both epidemic modeling and data sharing practices, as it indicates that useful mobility data can be shared at lower resolution, reducing potential privacy risks while maintaining scientific value.

Introduction

Human mobility plays a fundamental role in the spatial spread of infectious diseases [1–3]. While many factors may contribute to epidemic dynamics – including environmental conditions, population susceptibility, pathogen transmissibility, and the built environment – human movement provides the primary mechanism through which pathogens propagate across geographic space. This mechanism explains the spatial invasion patterns observed across diseases such as influenza [4–6], COVID-19 [7–9], and SARS [10–12]. From daily commuting to international travel, the connections people create through movement provide the pathways along which infections spread between communities [13–15], offering novel non-local opportunities for exposure and transmission. Understanding these mobility patterns is therefore essential for anticipating epidemic trajectories and designing effective predictive models and interventions.

For several decades, epidemiological models have relied on traditional sources of mobility information [16], including census records [17], transportation statistics [5], commuting flows [13], and international air travel data [18]. However, these traditional sources often report data at coarse spatial and temporal resolutions and only in specific settings, limiting our ability to detect rapid behavioral shifts during public health emergencies [19].

The proliferation of mobile phones in the last decades has fundamentally transformed the observation of human mobility. By 2025, approximately 5.8 billion people, 71% of the global population, carried mobile devices that continuously generate geolocated data [20], by providing unprecedented insight into human behavior at population scale [16,21,22]. Unlike traditional data sources that capture only specific movement types, mobile phone data comprehensively record all trips, regardless of purpose, distance, or transportation mode, spanning spatial scales from local neighborhood movements to long-distance domestic travel [23]. This detection capability, combined with continuous tracking and high spatial-temporal resolution, has made mobile phones an extraordinarily rich source of mobility information and have opened new possibilities for studying human behavior and modeling disease spread.

Mobile phone data have indeed been widely adopted in spatial epidemic modeling. Theoretical and methodological studies have demonstrated their usefulness for quantifying mobility networks and connectivity patterns [24–27], and they have been applied to a variety of infectious diseases including malaria [25,28], cholera [29,30], schistosomiasis [31], dengue [32], HIV [33,34], 2013–2016 Ebola outbreak [35,36]. The COVID-19 pandemic marked a major turning point in the use of mobility data for public health [37,38]. Facing an unprecedented global emergency, telecommunications companies and technology platforms shared mobility data at unprecedented scales, enabling researchers to track population movements and support public health responses in near real-time [8,39–44].

Despite these advances, two key questions remain unresolved: (i) which characteristics of the mobility process are epidemiologically relevant for spatial disease spread, and (ii) what level of data resolution is actually necessary to capture invasion dynamics? Addressing these questions is essential for developing accurate epidemic models capable of informing public health decision-making. In the post-COVID-19 landscape, access to individual-level trajectories has become increasingly restricted, while aggregated mobility indicators remain the standard approach adopted by telecommunications. Identifying the minimum mobility information required for accurate epidemic prediction would allow data providers to generate indicators that are scientifically meaningful while avoiding the need to share highly detailed individual trajectories.

Integrating mobile phone data into predictive epidemic models requires translating individual-level trajectories into location-level coupling forces, quantitative measures representing the probability of disease transmission between connected places. This process involves two key choices: which locations should be connected (network structure), and how strongly they should be coupled (edge weights). Different aggregation methods encode different assumptions about which aspects of mobility are epidemiologically relevant. Some preserve complete movement sequences, tracking the order of displacements between locations; others focus on time allocation, recording how long individuals spend in different places; and others identify only primary destinations such as home, work, school as the main sites of transmission risk. Despite the widespread adoption of mobile phone data in epidemic modeling, these aggregation choices have not been systematically assessed because mobility data are typically pre-aggregated by data providers and researchers have rarely access to individual trajectories. Determining which mobility features drive disease spread, and at what resolution they must be measured, is therefore both a scientific and an ethical challenge. If highly detailed mobility information provides little improvement over simpler aggregations, unnecessary privacy risks may be introduced without substantial epidemiological benefit. Conversely, overly coarse aggregation may obscure key transmission pathways and reduce predictive accuracy.

Here, we investigate how the level of detail used to describe human mobility influences predictions of spatial epidemic spread. We compare alternative ways of summarizing population movement that capture different degrees of behavioral detail and evaluate how these representations affect predictions of when and where epidemics spread across space. By identifying which aspects of mobility information are essential for capturing spatial disease dynamics, our study clarifies

what type of mobility data is most informative for epidemic modeling and highlights the trade-offs between epidemiological accuracy and potential privacy risks.

Results

To assess how mobility data aggregation shapes epidemic predictions, we constructed three mobility networks from individual mobile phone trajectories of 9.5 million users in Senegal, each encoding a decreasing level of movement detail: a high-resolution network (HR), a medium-resolution network (MR), and a low-resolution network (LR). These networks were integrated into a metapopulation epidemic modeling framework and used to simulate outbreaks across a range of transmission scenarios. We then compared the resulting epidemic arrival times and transmission pathways to evaluate which aspects of mobility are epidemiologically informative.

To evaluate how different levels of mobility data aggregation influence epidemic predictions, we compare three coupling representations derived from mobile phone data (Fig 1): high-resolution tracking of all individual displacements and connecting consecutive visited locations (HR) [45], medium-resolution coupling linking home to all visited locations and weighted by the time individuals spend in these locations (MR) [29], and low-resolution coupling linking home to the

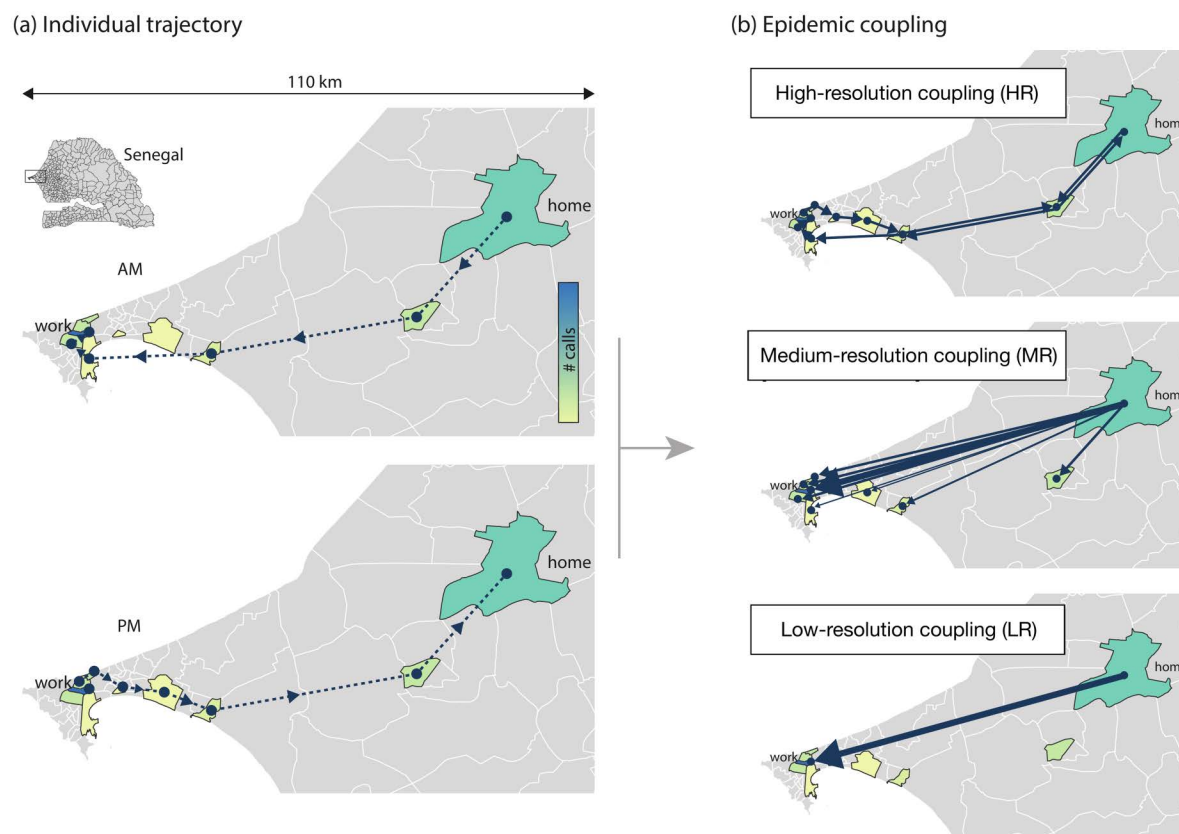


Fig 1. Mobility resolutions. (a) Individual trajectory showing mobile phone user movements throughout the day from home to work in the morning (top) and from work to home in the afternoon (bottom), with color indicating number of activities (calls, messages) in each municipality. (b) Resulting epidemiological coupling networks for three aggregation methods: HR) High-resolution method linking all consecutive call locations; MR) Medium-resolution method linking home to all visited locations weighted by time spent; LR) Low-resolution method linking home only to the most-visited location. The maps were generated in Python using administrative boundary shapefiles from the Global Administrative Areas database (GADM), available at <https://gadm.org>.

<https://doi.org/10.1371/journal.pcbi.1014427.g001>

most-visited location only (LR) [25]. We first examine the structural differences between the resulting mobility networks, focusing on coupling probabilities, connectivity patterns, and the presence of long-range links. We then quantify how these differences affect simulated epidemic dynamics using a metapopulation model under three transmissibility regimes: low transmission ($R_0 = 1.1$), representing controlled epidemic conditions such as SARS-CoV-2 under strong interventions; moderate transmission ($R_0 = 1.5$), consistent with seasonal influenza-like dynamics; and high transmission ($R_0 = 3.0$), representing highly transmissible pathogens such as early SARS-CoV-2 or Ebola during intense community transmission. Epidemic outcomes are compared in terms of spatial invasion patterns, arrival times across municipalities, and transmission pathways, allowing us to assess how mobility data resolution shapes predicted epidemic spread.

Location and commuting-based matrices better represent epidemic-relevant mobility

The three aggregation methods at different spatio-temporal resolutions (HR, MR, LR) yielded 12 daily average directed networks per method, one for each month. The resulting networks exhibit different topologies depending on the aggregation process. Multiple Regression Quadratic Assignment Procedure (MRQAP) revealed high correlations among all methods (regression coefficient $r_{LR,MR} = 1$, $r_{HR,MR} = 0.99$, $r_{HR,LR} = 0.99$). Despite these high correlations, the median coupling probability in HR is approximately one order of magnitude lower than in MR and LR (1.5×10^{-4} in LR, 1.0×10^{-4} in MR, 8.4×10^{-6} in HR, Fig 2a). Largest differences are on links that connect Urban/Rural municipalities (Fig 2c).

Conversely, outgoing probabilities are higher in HR compared with MR and LR (0.86 in HR, 0.62 in MR, 0.69 in LR; Fig 2b). Differences are mainly driven by rural areas in the southwest of Senegal (Fig 2d). Because HR is constructed

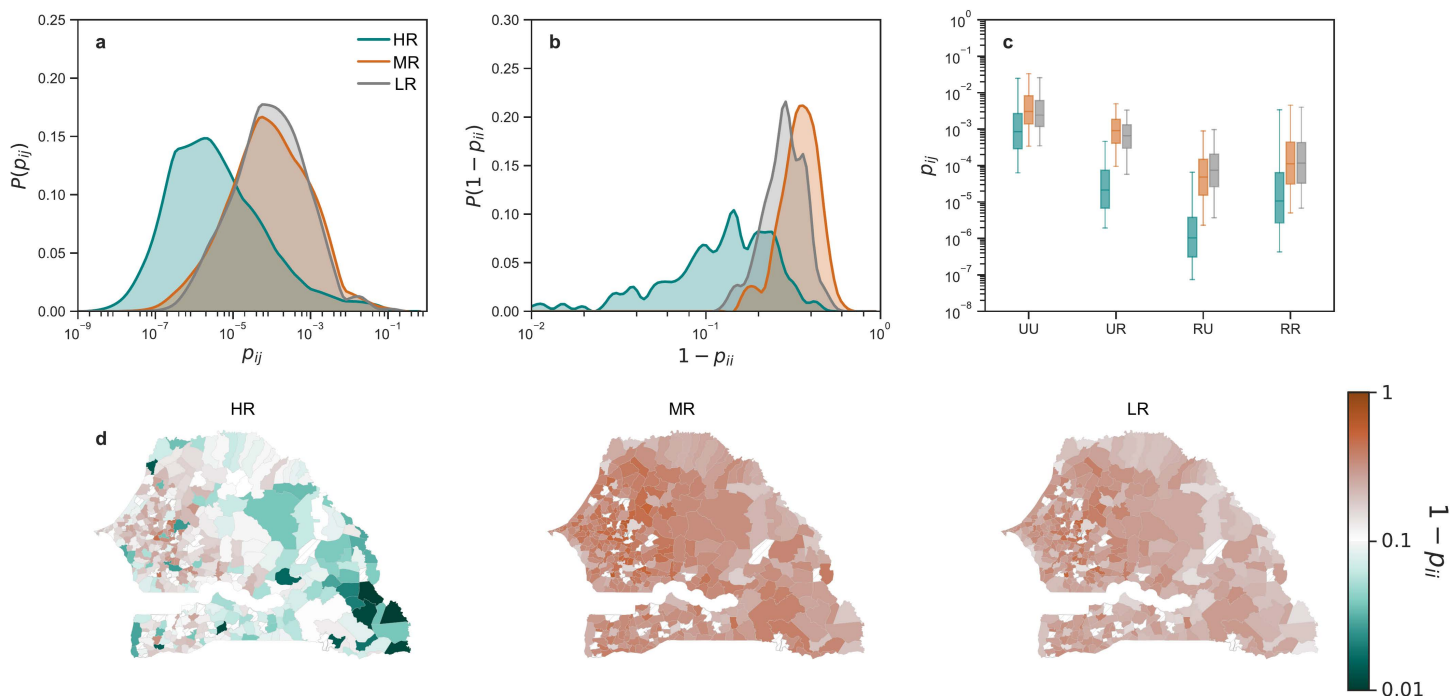


Fig 2. Coupling forces. **a)** Coupling probability distribution in HR, MR, and LR. Distribution includes matrices for all months in 2013. **b)** Outgoing probability ($1 - p_{ij}$) distribution in HR, MR, and LR. Distribution includes matrices for all months in 2013. **c)** Coupling probability stratified by municipality type. Box plots show the 95% reference range across all months for links connecting Urban-Urban (UU), Urban-Rural (UR), Rural-Urban (RU), and Rural-Rural (RR) municipalities. **d)** Spatial distribution of outgoing probability. Maps show the outgoing probability for each municipality in HR, MR, and LR for January 2013. The maps were generated in Python using administrative boundary shapefiles from the Global Administrative Areas database (GADM), available at <https://gadm.org>.

<https://doi.org/10.1371/journal.pcbi.1014427.g002>

from consecutive mobile phone events, periods in which individuals remain in the same location generate self-loops, substantially reducing the probability of moving. Trips are partitioned into multiple short consecutive segments, causing the corresponding movement probability to be distributed across several intermediate links rather than concentrated on a single origin–destination connection, as occurs in MR and LR. Because of these two effects, coupling probabilities in HR are smaller and diagonal elements in the HR matrices are larger, whereas non-diagonal elements are smaller, producing the observed differences between HR and the matrices MR and LR.

We then examined differences between coupling matrices by analyzing (i) links shared across aggregation methods and (ii) links present in one aggregation but absent in the others (Fig 3). For links shared between HR and MR, coupling probabilities in HR are systematically lower, with differences increasing with geographical distance (Fig 3a–3c). When considering links with relative variation above a threshold ($\Delta^{HR,MR}(p_{ij}) > x$), the most discrepant links show coupling probabilities in HR up to 1,000 times lower than in MR (Fig 3a). These strongly discrepant links predominantly connect rural and urban municipalities at long distances (Fig 3a–3c). The comparison between HR and MR revealed similar patterns. In contrast, links shared between MR and LR show much smaller differences. Coupling probabilities in LR differ by at most one order of magnitude from MR, and

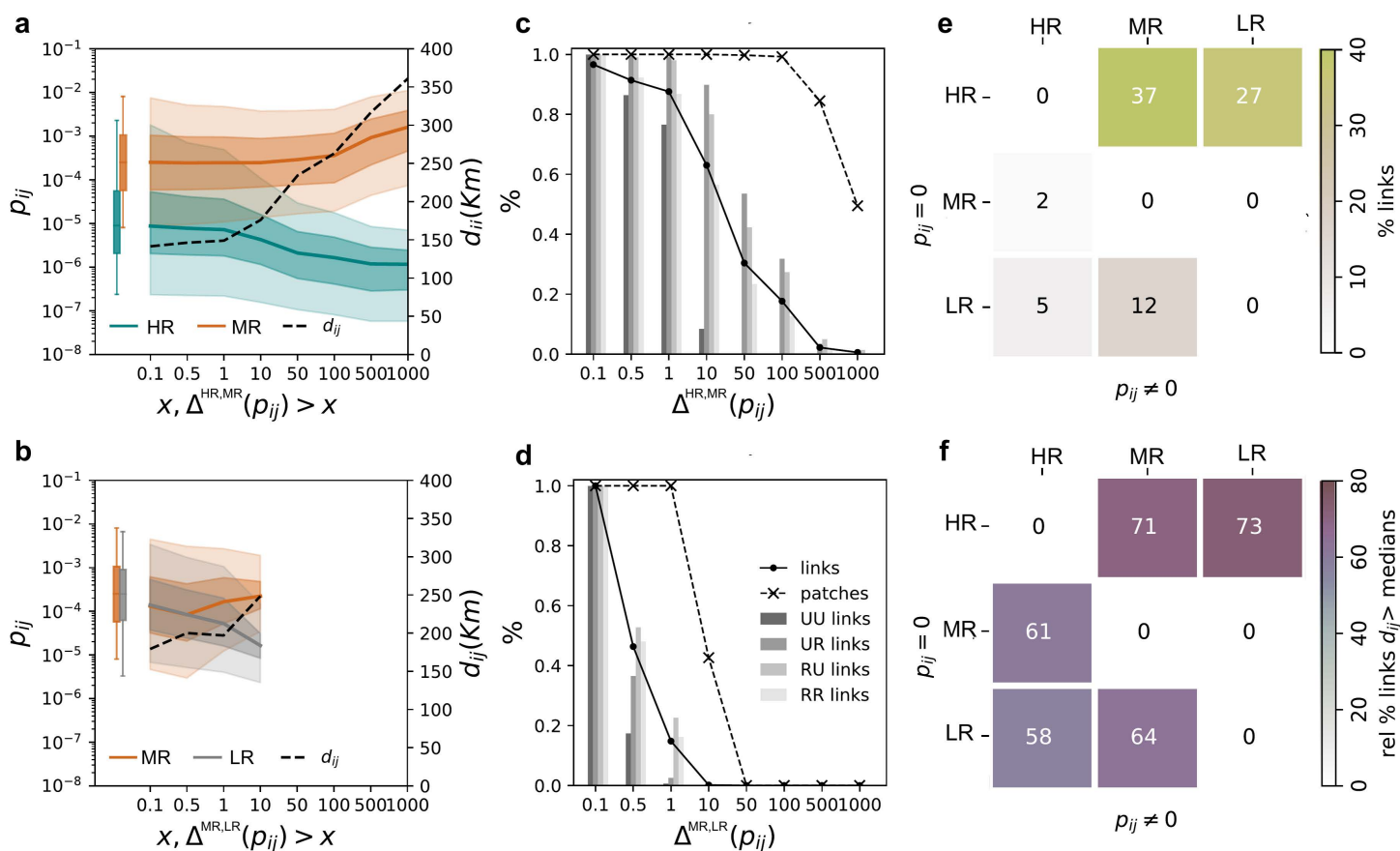


Fig 3. Differences between HR, MR and LR. **a, b** Coupling probability distributions for common links in all 12 matrices as a function of relative variation threshold (x). Subsets include only links where the relative variation Δ between methods exceeds x . Lines show coupling probabilities for HR and MR (**a**) or MR and LR (**b**); dashed lines indicate median geographical distance of links for each subset. **c, d** Characteristics of link subsets as a function of relative variation threshold (x) for all 12 matrices. Solid line: percentage of links in each subset. Dotted line: percentage of locations (nodes) represented in each subset. Stacked bars: distribution of links by municipality type (UU: Urban-Urban, UR: Urban-Rural, RU: Rural-Urban, RR: Rural-Rural). **e-f** Matrices showing unique links unique to each method for all 12 matrices. Yellow matrix: percentage of links present in the column method but absent in the row method. Violet matrix: among unique links, percentage connecting municipalities at distances greater than the median.

<https://doi.org/10.1371/journal.pcbi.1014427.g003>

these differences remain relatively stable across geographical distances (Fig 3b–3d). Also in this case, the largest differences involve links connecting rural areas (Fig 3d). We then examined links present in only one aggregation. 37% of links in MR are absent from HR (Fig 3e, yellow matrix) and 71% of these unique MR links connect municipalities separated by distances greater than the median (Fig 3f, violet matrix). Similarly, 27% of links in LR are absent from HR, with 73% corresponding to long-range connections. Conversely, only 12% of links in MR are absent from LR (Fig 3e). These MR-only links represent connections to secondary activity locations with lower mobility flows. This small difference indicates that the additional destinations captured by MR, beyond the primary workplace/school in LR, contribute little to the overall mobility network structure.

Together, these findings demonstrate that by tracking every displacement, HR fragments trips and misses the long-distance connections critical for spatial disease invasion. In contrast, methods that aggregate mobility based on where individuals spend time (MR and LR) naturally capture the long-range locations where sustained exposure may occur.

Displacement-based coupling both slows and constrains simulated epidemic spread

The lower coupling probabilities observed in HR translated directly into delayed epidemic arrival times. In HR, infections reached municipalities between a few weeks and 410 days after seeding, whereas in MR and LR, arrival times did not exceed 250 days (Fig 4a). The median relative variation in arrival times between MR and LR was near 40 across all R_0 values (Fig 4a), whereas comparisons between HR and either MR or LR showed median variations around 100%.

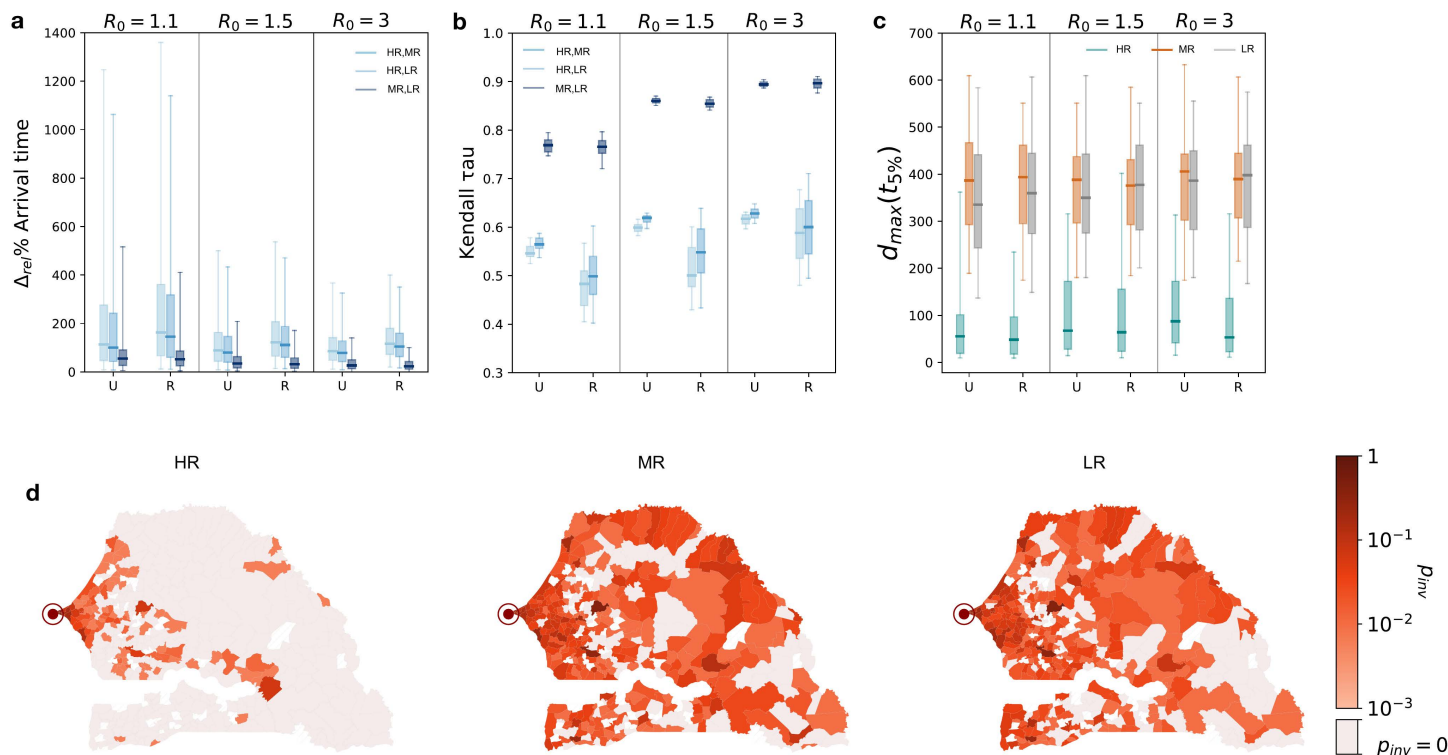


Fig 4. Differences on simulated epidemic outcomes. **a)** Relative variation in arrival times for each municipality between combinations of the three methods. **b)** Kendall tau rank correlation coefficients comparing arrival time orderings between HR vs. MR, HR vs. LR, and MR vs. LR. **c)** Maximum geographical distance reached from the epidemic seed at $t_{5\%}$. $t_{5\%}$ is defined as the time when the 5% of the municipalities have become infected ($t_{5\%}$). Boxplots in panels (a–c) indicate the 95% reference range. **d)** Spatial distribution in map of the invasion (p_{inv}) at $t_{5\%}$ for the three methods HR, MR, LR, for $R_0 = 1.1$, with a municipality of the capital Dakar as seed. The location with the red dot is the epidemic seed. The maps were generated in Python using administrative boundary shapefiles from the Global Administrative Areas database (GADM), available at <https://gadm.org>.

<https://doi.org/10.1371/journal.pcbi.1014427.g004>

MR and LR not only produced similar arrival time distributions but also predicted nearly identical spatial invasion sequences, with high rank correlation (Kendall $\tau=0.75-0.90$; [Fig 4b](#)). In contrast, HR showed substantially lower rank correlation with both MR and LR, indicating fundamentally different invasion pathways. The spatial extent of early spread also differed markedly across methods. At $t_5\%$ - the time when 5% of the municipalities have been reached by the transmission - the maximum geographical distance reached from the seed was substantially smaller in HR than in MR and LR ([Fig 4c](#)), reflecting HR systematic underestimation of long-range transmission. With $R_0 = 1.1$, the invasion probability of the top 5% infected municipalities in HR clustered tightly around the seed (e.g., near Dakar), whereas MR and LR produced more spatially heterogeneous invasion patterns with early spread to distant locations ([Fig 4e](#)).

Beyond arrival times, the spatial structure of epidemic invasion differed across aggregation methods. We analyzed invasion trees, representing the most likely transmission pathways through which the epidemic spreads from seed locations, for two contrasting scenarios: seeding from Dakar (the capital of Senegal) and from a rural municipality in Saraya (the rural and most distant location from Dakar, in the top 2% for variation in outgoing probability between HR and MR). MR and LR produced strikingly similar invasion trees, with betweenness centrality distances below 0.05, indicating nearly identical hierarchical spread patterns ([Fig 5a](#)). In contrast, HR differed substantially from both MR and LR (betweenness distances 0.05-0.25), revealing fundamentally different transmission pathways. In HR, the distribution exhibits a long tail extending to $d_{inv} = 15-20$, indicating that some municipalities are reached only through many sequential transmission steps from one location to another ([Fig 5b-5c](#)). In contrast, MR and LR show more compact distributions with maximum invasion distances of approximately 10 steps, reflecting spatial spread through long-range mobility connections.

To empirically ground our theoretical findings, we analyzed the invasion dynamics of the early COVID-19 pandemic in Senegal in 2020. The epidemic exhibited rapid long-range jumps between major urban hubs rather than gradual geographic diffusion ([Fig 6c](#)). The first confirmed cases were reported in Dakar, the capital and location of the international airport where the virus was first introduced. The virus was also detected in Touba, a major religious and economic hub located several hundred kilometers from Dakar. The rapid spread across distant locations highlights the importance of long-range mobility connections in shaping the early spatial spread of the epidemic. We compared the observed invasion order across departments with the arrival times predicted by the three mobility representations. Because early estimates of COVID-19 transmissibility varied widely (approximately $R_0 = 1.5-3.5$), we evaluated model predictions across this plausible range. The heterogeneous invasion pattern observed in Senegal is well reproduced by the MR and LR models but not by the HR model. For $R_0 = 1.5$, the predicted and observed arrival rankings are strongly correlated for MR and LR (Spearman rank correlations of $r=0.65$ and 0.77 , respectively), whereas HR shows weak and non-significant agreement ($r=0.24$; [Fig 6a](#)). Consistent results are obtained using Kendall's τ ($\tau=0.51$ and 0.61 for MR and LR respectively, versus non-significant association for HR). At the higher end of the estimated range ($R_0=3$), MR and LR continued to reproduce the observed invasion ordering (Spearman's $r=0.77$ for both models), while HR showed weaker agreement ($r=0.54$; [Fig 6b](#)). The consistent performance of MR and LR across this empirically plausible range of R_0 values indicates that mobility representations capturing where individuals spend time better reproduce the spatial ordering of epidemic invasion. These findings demonstrate that intermediate-range and long-range mobility coupling, rather than purely local displacement transmission, were the dominant drivers of the heterogeneous spatial spread observed during the early COVID-19 epidemic in Senegal.

Finally, to test the robustness of our findings, we performed a sensitivity analysis by introducing time-corrected versions of the three coupling matrices: HR', MR', and LR'. Those matrices are designed to reduce the bias introduced by heterogeneity in users' calling activity. HR' redefines displacement weight as the time elapsed between consecutive moves; MR' replaces call counts with the effective time spent in each visited location; and LR' applies the same correction to the most visited location matrix, ensuring coupling strengths reflect time allocation rather than call frequency (see [S1 Text](#) for details).

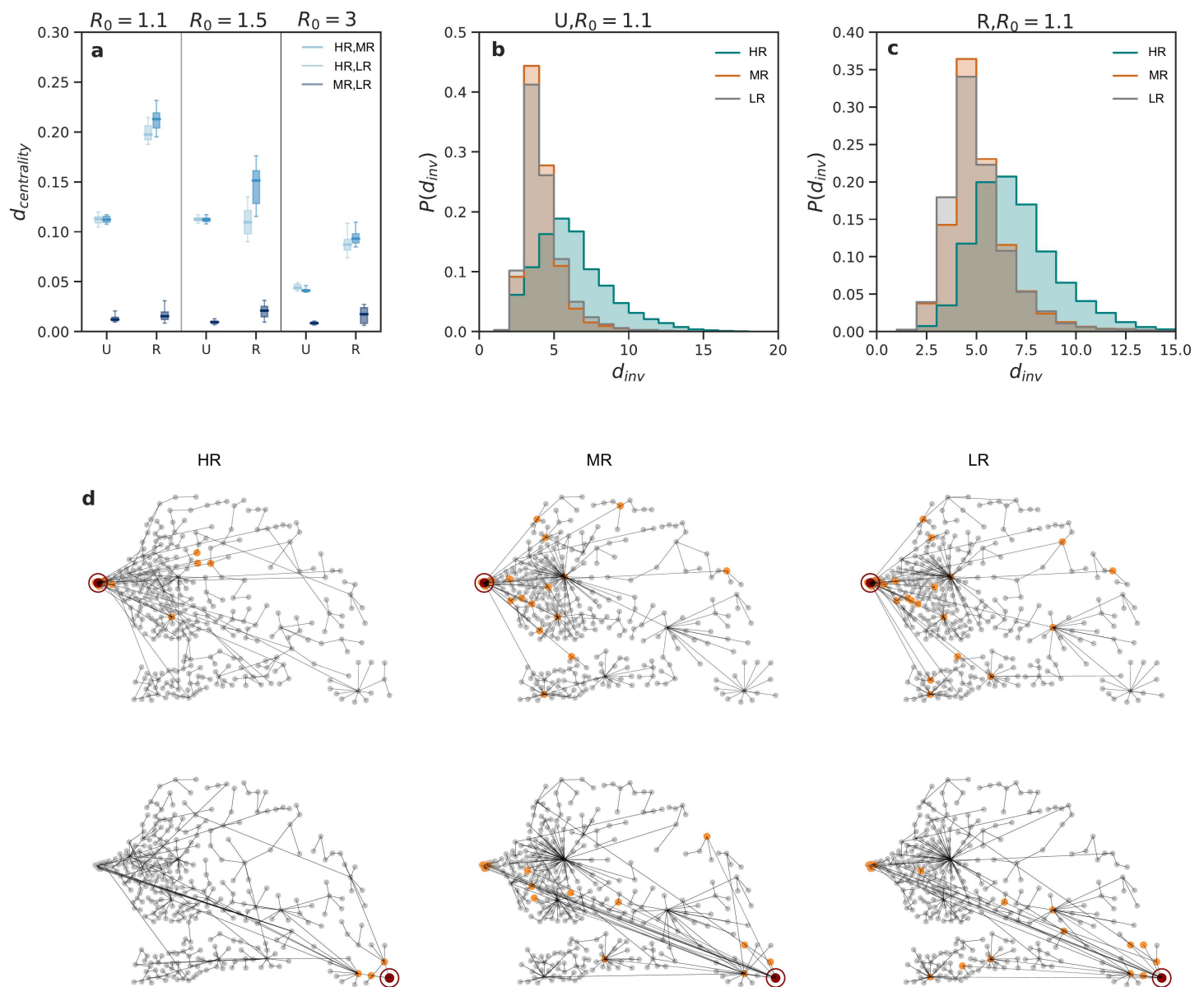


Fig 5. Epidemic invasion trees. Epidemic invasion trees were computed for $R_0 = 1.1, 1.5, 3$ considering both an urban (U) and a rural seed (R). **a)** Boxplots show the 95% reference range of the betweenness centrality distance index between invasion trees from each pair of methods. **b-c)** Distribution of the invasion distances across the three methods for $R_0 = 1.1$. **d)** The invasion trees for $R_0 = 1.1$. Epidemic seeds are shown in red. Locations directly infected by the seed are in orange, and locations infected by secondary spread are in gray. The maps were generated in Python using administrative boundary shapefiles from the Global Administrative Areas database (GADM), available at <https://gadm.org>.

<https://doi.org/10.1371/journal.pcbi.1014427.g005>

While MR' and LR' showed minimal differences from their original counterparts in terms of epidemic outcomes, HR 'demonstrated substantial improvements over HR, producing spatial diffusion patterns more similar to MR and LR (Fig O in S1 Text – Fig S in S1 Text). However, HR' still exhibited limitations, with invasion distance distributions remaining higher than in MR and LR due to approximately 40% of epidemiologically relevant links present in MR and LR being absent in

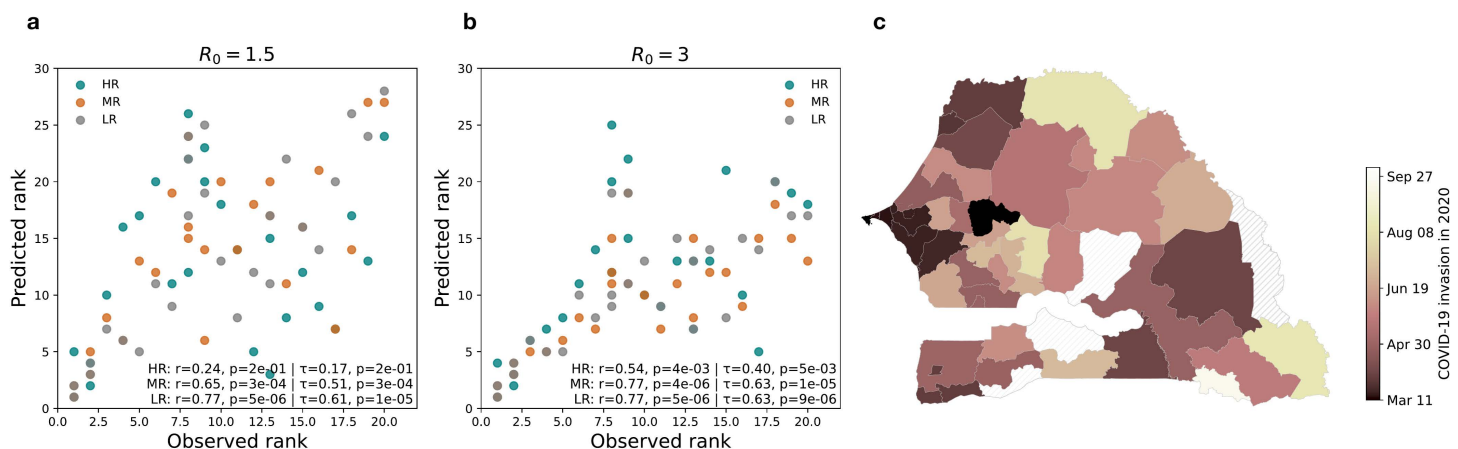


Fig 6. Comparison of observed and predicted COVID-19 invasion during the 2020 first wave. (a) Scatter plot comparing the ranking of observed early COVID-19 arrival times in 2020 with model-predicted arrival times across departments, for the HR, MR, and LR methods. Model predictions correspond to the scenario with $R_0 = 1.5$ and Dakar as the seed location. (b) Same as panel (a), but for the scenario with $R_0 = 3$ and Dakar as the seed location. Spearman's (r) and Kendall's (τ) correlation coefficients, together with their associated p -values, are reported in both panels. (c) Spatial distribution of the COVID-19 invasion in Senegal in 2020, shown as department-level arrival times. Dakar and Touba colored in black are the first ones to report confirmed COVID-19 cases on March, 11, 2020. The maps were generated in Python using administrative boundary shapefiles from the Global Administrative Areas database (GADM), available at <https://gadm.org>.

<https://doi.org/10.1371/journal.pcbi.1014427.g006>

HR' (Fig O in S1 Text). This suggests that even refined displacement-based approaches may not adequately capture epidemiological couplings between locations for diseases where daily mobility patterns drive transmission.

Discussion

We systematically evaluated how different levels of mobile phone data aggregation affects predictions of spatial epidemic spread. Using mobile phone data from Senegal, we compared three approaches to constructing coupling matrices from individual trajectories: high-resolution tracking of all displacements (HR) [45], medium-resolution representations based on time spent across visited locations (MR) [29], and low-resolution identification of the most-visited location (LR) [25]. Our results show that increasing mobility resolution does not necessarily improve epidemic predictions. Instead, approaches that capture where individuals spend most of their time produce more realistic spatial invasion patterns than those that track every displacement.

Two main findings emerge from our analysis. First, preserving maximum mobility detail by tracking all displacements (HR), without accounting for time spent in each place, produces biased epidemic predictions, including delayed arrival times, underestimated long-range transmission, and unrealistic radial diffusion patterns, compared with the other approaches (MR, LR) and previous works [27,46]. HR yields coupling probabilities approximately one order of magnitude lower than MR and LR and misses a substantial fraction of long-range connections. These results align with previous work showing that the duration of travel and time spent at destination are key drivers of transmission risk [47,48]. Second, MR and LR produce highly similar and more realistic epidemic patterns, indicating that identifying the locations where individuals spend sustained periods of time—such as homes, workplaces, and schools—is sufficient to capture the mobility structure relevant for spatial disease transmission. Tracking secondary activities beyond these most-visited location provides minimal additional epidemiological value for spatial disease spread. This is consistent with a substantial body of literature demonstrating that commuting fluxes between home and regularly visited locations are primary drivers of spatial epidemic spread, shaping geographic invasion [13,14].

The poor performance of HR stems from a mismatch between mobility representation and epidemiological processes driving spatial transmission. By linking consecutive mobile phone events, HR fragments continuous trips into multiple short segments and fails to identify the true origin–destination structure of mobility, connecting locations where sustained transmission occurs. This fragmentation has two critical consequences: (i) it substantially underestimates long-distance connectivity by breaking journeys into short hops, (ii) it assigns equal weight to each displacement regardless of time spent, failing to distinguish epidemiologically important locations (homes, workplaces, schools) from places individuals merely pass through.

In contrast, MR and LR focus on time allocation, which better reflects exposure opportunities. Time spent in a location is strongly related to exposure opportunity and transmission risk [48], whereas brief passages through intermediate locations contribute little to disease spread. By preserving the directional relationship between residences and destinations, MR and LR also reconstruct the characteristic spatial pathways through which epidemics propagate between communities [10]. Consistent with our simulation results, comparison with the observed heterogeneous spatial invasion of COVID-19 in Senegal [49] shows that these representations better reproduce the ordering of epidemic arrival across locations.

Our findings are likely to generalize beyond the specific context analyzed here. Across the range of transmissibility explored ($R_0 = 1.1, 1.5, 3.0$), representing controlled, moderate, and high transmission scenarios, the relative performance of the mobility representations remained consistent. Differences between methods were most pronounced at low transmissibility ($R_0 = 1.1$), when spatial heterogeneity plays a greater role in shaping epidemic spread, and diminished at higher transmission levels. This robustness suggests that our conclusions reflect fundamental interactions between mobility structure and transmission processes rather than pathogen-specific characteristics. Although mobility patterns differ across socioeconomic settings, large-scale human movement is consistently characterized by structural regularities that can be captured by gravity- or radiation-type models across both high-income and low- and middle-income countries [14,16]. These regularities reflect common constraints on travel behavior, such as population size, distance, and activity patterns. While the magnitude of mobility flows and the relative importance of commuting versus long-distance travel may vary across contexts, the underlying origin–destination structure of human mobility tends to remain similar. Spatial scale may also influence the magnitude of differences between aggregation methods: at coarser spatial resolutions, averaging across locations may reduce the apparent discrepancies between mobility representations [27]. However, the central finding of our study remains robust: epidemiologically meaningful connectivity is better captured by linking residences to destinations where individuals spend time than by reconstructing sequences of consecutive displacements.

While this study focuses on Senegal, we are aware that mobility patterns differ across socioeconomic contexts, with implications for how aggregation methods may behave. In low-income settings, daily mobility often consists of fewer, shorter trips (mean ≈ 5 km), whereas in high-income settings mobility typically exhibits more frequent and longer commuting trips (mean ≈ 10 – 15 km). That said, a substantial body of literature has established that human mobility is well described by parsimonious generative models, gravity and radiation models, across both high-income and low- and middle-income country (LMIC) settings [16], and Balcan et al. [14] demonstrated that commuting flows can be reliably fitted at a global scale, suggesting that the fundamental structure of daily human circulation between home and regularly visited locations is a general feature of mobility that transcends development contexts. These structural differences, moreover, do not alter our methodological conclusions, because the failure of HR arises from trip fragmentation rather than from any property of trip length itself: by chaining sequentially visited locations, HR severs the epidemiologically meaningful connection between where people reside and where they spend significant time, misrepresenting the coupling forces that drive epidemic invasion regardless of characteristic trip distances. While the specific magnitude of differences between methods may vary with local mobility patterns, infrastructure, and socioeconomic conditions, the core distinction between origin–destination preservation and trip fragmentation should generalize across settings. The spatial scale of analysis may nonetheless affect the generalizability of our quantitative results, as using municipalities as spatial units allows us to capture substantial heterogeneity across locations, whereas coarser aggregations would smooth over these

differences and likely attenuate the observed magnitude of discrepancies between methods. Indeed, we expect differences between methods to diminish at coarser scales, where averaging obscures the heterogeneities that distinguish approaches, but empirical evaluation across scales remains an important direction for future work. Validation in diverse geographic and economic contexts would further strengthen the universality of these findings.

Our study has a number of limitations. First, although we compared simulated invasion patterns with observed COVID-19 spatial spread in Senegal, we did not perform a full calibration of the epidemic model to incidence time series or evaluate predictive accuracy across the entire epidemic wave. Our analysis instead focuses on the comparative impact of mobility representations on spatial transmission patterns. Second, mobile phone data capture user activity only when calls or messages occur, which may introduce biases in the reconstruction of mobility trajectories. Although our aggregation methods partially mitigate this issue by focusing on time allocation rather than individual events, variations in phone usage behavior may still influence estimated mobility patterns. The landscape of mobility data has evolved considerably, and a broader range of data sources, including XDR records, GPS traces, Wi-Fi access logs, and other high-resolution digital traces, now offer complementary approaches that may reduce reliance on event-driven CDR data and further improve the accuracy of mobility estimation. Third, our analysis focuses on mobility-driven transmission and does not explicitly incorporate other factors that may influence spatial spread, such as heterogeneities in contact behavior, local interventions, or environmental conditions. Despite these limitations, our results provide practical guidance for both epidemic modeling and mobility data sharing.

In the post-COVID-19 era, direct access to individual trajectories has become increasingly restricted, while aggregated mobility indicators are more widely available. Our findings challenge the assumption that more granular data yields better epidemic predictions. Instead, aggregated indicators describing time spent at major activity locations appear sufficient for predictive spatial models. This insight supports the development of privacy-preserving data-sharing frameworks that reduce the sensitivity of mobility data while preserving their epidemiological value.

Methods

Mobile phone dataset

We analyzed anonymized Call Detail Records (CDRs) provided by Orange Senegal covering January–December 2013. Each CDR corresponds to a call or text and includes a timestamp and the serving cell tower, allowing the user's location to be inferred at the antenna level. The dataset covers approximately 9.57 million users ($\approx 80\%$ of the national population), mapped to municipalities via 15,999 antennas distributed across the territory. To ensure spatial consistency, we retained only municipalities covered throughout 2013, resulting in 394 municipalities (46 urban and 348 rural). The 46 urban municipalities are located in the following cities: Dakar, Guediawaye, Pikine, Rufisque, Thiès. We kept users active for more than 30 days and excluded extreme activity profiles (>1000 events/week) to reduce biases associated with bots and call centers.

From individual trajectories to coupling matrices

Coupling matrices quantify the connectivity strength between spatial locations that enables pathogen transmission through human movement. Each element p_{ij} represents the monthly averaged probability that an individual residing in i travels to location j in a given day. Consequently, p_{ij} captures the opportunity for exposure of individuals residing in location i to individuals present in location j and potential transmission in j . Mobile phone communication events (calls and SMS) provide geo-location data at cell tower resolution. Each event marks the user's physical presence at a specific time and place, allowing us to track individual daily consecutive displacements and time spent in different locations as a proxy of the number of activities in a location. By aggregating individual mobility patterns across all users, we extracted three coupling matrices at different resolution levels, computing 12 daily average matrices for each month in 2013 at the municipality level for each method. The coupling matrices of the three methods are defined as follows:

High resolution coupling matrix (HR) [45]

$$p_{ij}^{HR} = \frac{\sum_{u_i} HR_{ij}^{u_i}}{\sum_k \sum_{u_i} HR_{ik}^{u_i}}$$

where $HR_{ij}^{u_i}$ is the number of times user u_i moves from location i to j on consecutive calls. This high-resolution method keeps the temporal sequence of individual displacements. The coupling represents the probability of moving from one place to another, without considering the time spent in each location. Each displacement has the same weight regardless of the time elapsed between consecutive calls.

Medium resolution coupling matrix (MR) [29]

$$p_{ij}^{MR} = \frac{\sum_{u_i} MR_{ij}^{u_i}}{\sum_k \sum_{u_i} MR_{ik}^{u_i}}$$

where $MR_{ij}^{u_i}$ is the number of calls made in j by user u_i living in i . Home locations are identified as the location where users made the most calls during nighttime (7 pm- 7 am), a well-established method used in literature [7]. This medium-resolution method loses information about the temporal sequence of displacements but captures the coupling between home and all visited locations. Under the assumption that the call frequency is proportional to the time spent in a location, this method measures probability of being in j given residence in i .

Low resolution coupling matrix (LR) [25]

$$p_{ij}^{LR} = \frac{\sum_{u_i} LR_{ij}^{u_i}}{\sum_k \sum_{u_i} LR_{ik}^{u_i}}$$

where $LR_{ij}^{u_i}$ is the amount of time user u_i spends at his most visited location daily in j . The most visited location was defined as where a user made the maximum number of calls during a 24-hour time period. This lowest-resolution method considers only locations where users spend most of their time (home, work or school), neglecting all other less-visited locations. The coupling probability represents the probability of spending most of the time in j given the residence in i . Since the most visited location often coincides with the workplace, this method approximates commuting fluxes [9].

We also tested refined versions of these matrices to correct for biases in user activity patterns and time representation (see [S1 Text](#) for details).

Epidemic metapopulation model

To quantify the impact of the coupling matrices on the modelled epidemic diffusion, we developed a stochastic, discrete-time and non-Markovian metapopulation model [4,5]. The model accounts for disease transmission arising from spatial co-location of individuals due to daily mobility patterns. The Senegalese population was spatially divided in 394 municipalities, with links between municipalities defined by the time-varying coupling matrices p_{ij} .

The force of infection in location i at time t is given by:

$$\lambda_i = \lambda_{ii} + \sum_{j \neq i} \lambda_{ji}^v + \sum_{j \neq i} \lambda_{ij}^r$$

with

$$\lambda_{ii} = \beta p_{ii}^2 \frac{I_i}{N_i}; \lambda_{ji}^v = \beta p_{ii} p_{ji} \frac{I_j}{N_i}; \lambda_{ij}^r = \beta p_{ij} \frac{\hat{I}_j}{N_j}$$

p_{ij} is the coupling probability between location i and j at time t and β the transmissibility rate. The time index t is omitted from the formulas for clarity.

The three terms of the force of infections correspond to three distinct transmission contributions: (λ_{ii}) residents who remain in their home municipality and get infected locally; (λ_{ij}^v) infected visitors from other municipalities temporarily present in i ; and (λ_{ji}^r) residents of i who are visiting location j for the day, get infected in j and return to i [50]. All three pathways are implicitly accounted for through the coupling probabilities p_{ij} , consistently with the first-order equilibrium solution derived analytically in Balcan et al. [14].

The effective population at time t is:

$$\hat{N}_i = p_{ii}N_i + \sum_j p_{ji}N_j$$

and the effective number of infections in i at time t is:

$$\hat{I}_i = p_{ii}I_i + \sum_j p_{ji}I_j$$

$\hat{N}_i$ and $\hat{I}_i$ are time-dependent quantities recomputed at each discrete time step to reflect the instantaneous mixing configuration of the population across locations.

We explored three epidemic scenarios corresponding to different transmission intensities: (i) low transmissibility ($R_0 = 1.1$), representing controlled epidemic conditions, as observed for SARS-CoV-2 during periods of effective non-pharmaceutical interventions; (ii) medium transmissibility ($R_0 = 1.5$), representing moderate transmissibility, consistent with typical estimates reported for seasonal influenza and the 2009 H1N1 pandemic; (iii) high transmissibility ($R_0 = 3$), compatible with estimates reported for SARS-CoV-2 in susceptible populations and for Ebola virus disease during intense community transmission.

Comparison between coupling matrices

We compared the three coupling matrices through network analysis and epidemic outcomes. For the network analysis, we performed Multiple Regression Quadratic Assignment Procedure (MRQAP) to test correlation between coupling probabilities [51]. This method is specifically designed for network data where observations are not independent. We examined relative variations in coupling probabilities and outgoing probabilities, stratifying by urban/rural municipality pairs and geographical distance (computed using the Haversine formula). The outgoing probability from location i is defined as follows:

$$p_{out} = \sum_{i \neq j} p_{ij} = 1 - p_{ii}$$

where p_{ii} is the probability of not moving.

Comparison between simulated epidemic outcomes

For epidemic outcomes, we tracked two main observables (1) the arrival time t_a of the epidemic in each location, defined as the first time that 10 individuals become infected; (2) the epidemic invasion tree [14] which represents the most likely transmission path. For the invasion tree, whenever a disease-free location i experienced an infection, we tracked a directed link from the location of origin. For each scenario, for each run, we extracted the invasion path. By cumulating over all runs, we obtained an invasion path where link weights represent the percentage of runs in which that link appeared. The invasion tree was extracted as the maximum spanning tree of this weighted invasion path. To avoid the

stochastic fluctuations of the epidemic simulations, we computed invasion trees by doing 1000 runs, and selecting randomly for 50 times 400 runs.

We compared the invasion trees using (a) the betweenness centrality distance [52]; and (b) the invasion distance measured as the number of edges connecting each node with the seed location of the invasion tree.

Supporting information

S1 Text. Supplementary material. Extended methods and analyses.
(DOCX)

Author contributions

Conceptualization: Vittoria Colizza.

Data curation: Giulia Pullano, Stefania Rubrichi.

Formal analysis: Giulia Pullano.

Funding acquisition: Stefania Rubrichi, Vittoria Colizza.

Investigation: Giulia Pullano, Shweta Bansal, Stefania Rubrichi, Vittoria Colizza.

Methodology: Giulia Pullano, Shweta Bansal, Stefania Rubrichi, Vittoria Colizza.

Project administration: Vittoria Colizza.

Resources: Stefania Rubrichi, Vittoria Colizza.

Supervision: Stefania Rubrichi, Vittoria Colizza.

Visualization: Giulia Pullano.

Writing – original draft: Giulia Pullano.

Writing – review & editing: Giulia Pullano, Shweta Bansal, Stefania Rubrichi, Vittoria Colizza.

References

1. Lessani MN, Li Z, Jing F, Qiao S, Zhang J, Olatosi B, et al. Human mobility and the infectious disease transmission: a systematic review. *Geo Spat Inf Sci*. 2024;27(6):1824–51. <https://doi.org/10.1080/10095020.2023.2275619> PMID: 40046953
2. Findlater A, Bogoch II. Human mobility and the global spread of infectious diseases: a focus on air travel. *Trends Parasitol*. 2018;34(9):772–83. <https://doi.org/10.1016/j.pt.2018.07.004> PMID: 30049602
3. Castelli F. Human mobility and disease: a global challenge. *J Travel Med*. 2004;11(1):1–2. <https://doi.org/10.2310/7060.2004.13610> PMID: 14769279
4. Charu V, Zeger S, Gog J, Bjørnstad ON, Kissler S, Simonsen L, et al. Human mobility and the spatial transmission of influenza in the United States. *PLoS Comput Biol*. 2017;13(2):e1005382. <https://doi.org/10.1371/journal.pcbi.1005382> PMID: 28187123
5. Merler S, Ajelli M. The role of population heterogeneity and human mobility in the spread of pandemic influenza. *Proc Biol Sci*. 2010;277(1681):557–65. <https://doi.org/10.1098/rspb.2009.1605> PMID: 19864279
6. Bajardi P, Poletto C, Ramasco JJ, Tizzoni M, Colizza V, Vespignani A. Human mobility networks, travel restrictions, and the global spread of 2009 H1N1 pandemic. *PLoS One*. 2011;6(1):e16591. <https://doi.org/10.1371/journal.pone.0016591> PMID: 21304943
7. Zhang M, et al. Human mobility and COVID-19 transmission: a systematic review and future directions. *Ann GIS*. 2022;28:501–14.
8. Kraemer MUG, Yang C-H, Gutierrez B, Wu C-H, Klein B, Pigott DM, et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science*. 2020;368(6490):493–7. <https://doi.org/10.1126/science.abb4218> PMID: 32213647
9. Xiong C, Hu S, Yang M, Luo W, Zhang L. Mobile device data reveal the dynamics in a positive relationship between human mobility and COVID-19 infections. *Proc Natl Acad Sci U S A*. 2020;117(44):27087–9. <https://doi.org/10.1073/pnas.2010836117> PMID: 33060300
10. Colizza V, Barrat A, Barthélemy M, Vespignani A. Predictability and epidemic pathways in global outbreaks of infectious diseases: the SARS case study. *BMC Med*. 2007;5:34. <https://doi.org/10.1186/1741-7015-5-34> PMID: 18031574
11. Ruan S, Wang W, Levin SA. The effect of global travel on the spread of sars. *Math Biosci Eng*. 2006;3(1):205–18. <https://doi.org/10.3934/mbe.2006.3.205> PMID: 20361819

12. Bowen JT, Laroe C. Airline networks and the international diffusion of severe acute respiratory syndrome (SARS). *Geogr J*. 2006;172(2):130–44. <https://doi.org/10.1111/j.1475-4959.2006.00196.x> PMID: [32367889](#)
13. Viboud C, et al. Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science*. 2006;312:447–51.
14. Balcan D, Colizza V, Gonçalves B, Hu H, Ramasco JJ, Vespignani A. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc Natl Acad Sci U S A*. 2009;106(51):21484–9. <https://doi.org/10.1073/pnas.0906910106> PMID: [20018697](#)
15. Hufnagel L, Brockmann D, Geisel T. Forecast and control of epidemics in a globalized world. *Proc Natl Acad Sci U S A*. 2004;101(42):15124–9. <https://doi.org/10.1073/pnas.0308344101> PMID: [15477600](#)
16. Barbosa H, et al. Human mobility: models and applications. *Phys Rep*. 2017;734.
17. Eubank S, Guclu H, Kumar VSA, Marathe MV, Srinivasan A, Toroczkai Z, et al. Modelling disease outbreaks in realistic urban social networks. *Nature*. 2004;429(6988):180–4. <https://doi.org/10.1038/nature02541> PMID: [15141212](#)
18. Colizza V, Barrat A, Barthélemy M, Valleron A-J, Vespignani A. Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions. *PLoS Med*. 2007;4(1):e13. <https://doi.org/10.1371/journal.pmed.0040013> PMID: [17253899](#)
19. Wardle J, Bhatia S, Kraemer MUG, Nouvellet P, Cori A. Gaps in mobility data and implications for modelling epidemic spread: A scoping review and simulation study. *Epidemics*. 2023;42:100666. <https://doi.org/10.1016/j.epidem.2023.100666> PMID: [36689876](#)
20. The Mobile Economy. <https://www.gsma.com/mobileeconomy/>
21. Maxmen A. Can tracking people through phone-call data improve lives? *Nature*. 2019;569(7758):614–7. <https://doi.org/10.1038/d41586-019-01679-5> PMID: [31142864](#)
22. Blondel VD, Decuyper A, Krings G. A survey of results on mobile phone datasets analysis. *EPJ Data Sci*. 2015;4(1). <https://doi.org/10.1140/epjds/s13688-015-0046-0>
23. González MC, Hidalgo CA, Barabási A-L. Understanding individual human mobility patterns. *Nature*. 2008;453(7196):779–82. <https://doi.org/10.1038/nature06958> PMID: [18528393](#)
24. Lai S, Farnham A, Ruktanonchai NW, Tatem AJ. Measuring mobility, disease connectivity and individual risk: a review of using mobile phone data and mHealth for travel medicine. *J Travel Med*. 2019;26(3):taz019. <https://doi.org/10.1093/jtm/taz019> PMID: [30869148](#)
25. Wesolowski A, Eagle N, Tatem AJ, Smith DL, Noor AM, Snow RW, et al. Quantifying the impact of human mobility on malaria. *Science*. 2012;338(6104):267–70. <https://doi.org/10.1126/science.1223467> PMID: [23066082](#)
26. Pullano G, Alvarez-Zuzek LG, Colizza V, Bansal S. Characterizing US spatial connectivity and implications for geographical disease dynamics and metapopulation modeling: longitudinal observational study. *JMIR Public Health Surveill*. 2025;11:e64914. <https://doi.org/10.2196/64914> PMID: [39965190](#)
27. Tizzoni M, Bajardi P, Decuyper A, Kon Kam King G, Schneider CM, Blondel V, et al. On the use of human mobility proxies for modeling epidemics. *PLoS Comput Biol*. 2014;10(7):e1003716. <https://doi.org/10.1371/journal.pcbi.1003716> PMID: [25010676](#)
28. Tatem AJ, Huang Z, Narib C, Kumar U, Kandula D, Pindolia DK, et al. Integrating rapid risk mapping and mobile phone call record data for strategic malaria elimination planning. *Malar J*. 2014;13:52. <https://doi.org/10.1186/1475-2875-13-52> PMID: [24512144](#)
29. Finger F, Genolet T, Mari L, de Magny GC, Manga NM, Rinaldo A, et al. Mobile phone data highlights the role of mass gatherings in the spreading of cholera outbreaks. *Proc Natl Acad Sci U S A*. 2016;113(23):6421–6. <https://doi.org/10.1073/pnas.1522305113> PMID: [27217564](#)
30. Bengtsson L, Gaudart J, Lu X, Moore S, Wetter E, Sallah K, et al. Using mobile phone data to predict the spatial spread of cholera. *Sci Rep*. 2015;5:8923. <https://doi.org/10.1038/srep08923> PMID: [25747871](#)
31. Ciddio M, Mari L, Sokolow SH, De Leo GA, Casagrandi R, Gatto M. The spatial spread of schistosomiasis: a multidimensional network model applied to Saint-Louis region, Senegal. *Adv Water Resour*. 2017;108:406–15. <https://doi.org/10.1016/j.advwatres.2016.10.012> PMID: [29056816](#)
32. Wesolowski A, Qureshi T, Boni MF, Sundsøy PR, Johansson MA, Rasheed SB, et al. Impact of human mobility on the emergence of dengue epidemics in Pakistan. *Proc Natl Acad Sci U S A*. 2015;112(38):11887–92. <https://doi.org/10.1073/pnas.1504964112> PMID: [26351662](#)
33. Isdory A, Mureithi EW, Sumpter DJT. The impact of human mobility on HIV transmission in Kenya. *PLoS One*. 2015;10(11):e0142805. <https://doi.org/10.1371/journal.pone.0142805> PMID: [26599277](#)
34. Valdano E, Okano JT, Colizza V, Mitonga HK, Blower S. Using mobile phone data to reveal risk flow networks underlying the HIV epidemic in Namibia. *Nat Commun*. 2021;12(1):2837. <https://doi.org/10.1038/s41467-021-23051-w> PMID: [33990578](#)
35. Peak CM, Wesolowski A, Zu Erbach-Schoenberg E, Tatem AJ, Wetter E, Lu X, et al. Population mobility reductions associated with travel restrictions during the Ebola epidemic in Sierra Leone: use of mobile phone data. *Int J Epidemiol*. 2018;47(5):1562–70. <https://doi.org/10.1093/ije/dyy095> PMID: [29947788](#)
36. Wesolowski A, Buckee CO, Bengtsson L, Wetter E, Lu X, Tatem AJ. Commentary: containing the ebola outbreak - the potential and challenge of mobile network data. *PLoS Curr*. 2014;6. <https://doi.org/10.1371/currents.outbreaks.0177e7fcf52217b8b634376e2f3efc5e> PMID: [25642369](#)
37. Oliver N, Lepri B, Sterly H, Lambiotte R, Deletaille S, De Nadai M, et al. Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. *Sci Adv*. 2020;6(23):eabc0764. <https://doi.org/10.1126/sciadv.abc0764> PMID: [32548274](#)
38. Grantz KH, Meredith HR, Cummings DAT, Metcalf CJE, Grenfell BT, Giles JR, et al. The use of mobile phone data to inform analysis of COVID-19 pandemic epidemiology. *Nat Commun*. 2020;11(1):4961. <https://doi.org/10.1038/s41467-020-18190-5> PMID: [32999287](#)

39. Pullano G, Valdano E, Scarpa N, Rubrichi S, Colizza V. Evaluating the effect of demographic factors, socioeconomic factors, and risk aversion on mobility during the COVID-19 epidemic in France under lockdown: a population-based study. *Lancet Digit Health*. 2020;2(12):e638–49. [https://doi.org/10.1016/S2589-7500\(20\)30243-0](https://doi.org/10.1016/S2589-7500(20)30243-0) PMID: [33163951](#)
40. Chang S, Pierson E, Koh PW, Gerardin J, Redbird B, Grusky D, et al. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature*. 2021;589(7840):82–7. <https://doi.org/10.1038/s41586-020-2923-3> PMID: [33171481](#)
41. Badr HS, Du H, Marshall M, Dong E, Squire MM, Gardner LM. Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. *Lancet Infect Dis*. 2020;20(11):1247–54. [https://doi.org/10.1016/S1473-3099\(20\)30553-3](https://doi.org/10.1016/S1473-3099(20)30553-3) PMID: [32621869](#)
42. Gatto M, Bertuzzo E, Mari L, Miccoli S, Carraro L, Casagrandi R, et al. Spread and dynamics of the COVID-19 epidemic in Italy: effects of emergency containment measures. *Proc Natl Acad Sci U S A*. 2020;117(19):10484–91. <https://doi.org/10.1073/pnas.2004978117> PMID: [32327608](#)
43. Pullano G, Di Domenico L, Sabbatini CE, Valdano E, Turbelin C, Debin M, et al. Underdetection of cases of COVID-19 in France threatens epidemic control. *Nature*. 2021;590(7844):134–9. <https://doi.org/10.1038/s41586-020-03095-6> PMID: [33348340](#)
44. Cryptic transmission of SARS-CoV-2 and the first COVID-19 wave | *Nature*. <https://www.nature.com/articles/s41586-021-04130-w>
45. Lima A, Domenico MD, Pejovic V, Musolesi M. Disease containment strategies based on mobility and information dissemination. *Sci Rep*. 2015;5:10650.
46. Panigutti C, Tizzoni M, Bajardi P, Smoreda Z, Colizza V. Assessing the use of mobile phone data to describe recurrent mobility patterns in spatial epidemic models. *R Soc Open Sci*. 2017;4(5):160950. <https://doi.org/10.1098/rsos.160950> PMID: [28572990](#)
47. Giles JR, Zu Erbach-Schoenberg E, Tatem AJ, Gardner L, Bjørnstad ON, Metcalf CJE, et al. The duration of travel impacts the spatial dynamics of infectious diseases. *Proc Natl Acad Sci U S A*. 2020;117(36):22572–9. <https://doi.org/10.1073/pnas.1922663117> PMID: [32839329](#)
48. Poletto C, Tizzoni M, Colizza V. Human mobility and time spent at destination: impact on spatial epidemic spreading. *J Theor Biol*. 2013;338:41–58. <https://doi.org/10.1016/j.jtbi.2013.08.032> PMID: [24012488](#)
49. Petit V, Robin N, Martin N. Spatiality and temporality of the Covid-19 epidemic in Senegal. The production process of health data with regard to territorial discontinuities. *Rev Francoph Sur Santé Territ*. 2021;3. <https://doi.org/10.4000/rfst.1179>
50. Sabbatini CE, Pullano G, Di Domenico L, Rubrichi S, Bansal S, Colizza V. The impact of spatial connectivity on NPIs effectiveness. *BMC Infect Dis*. 2024;24(1):21. <https://doi.org/10.1186/s12879-023-08900-x> PMID: [38166649](#)
51. Dekker D, Krackhardt D, Snijders TAB. Sensitivity of MRQAP tests to collinearity and autocorrelation conditions. *Psychometrika*. 2007;72(4):563–81. <https://doi.org/10.1007/s11336-007-9016-1> PMID: [20084106](#)
52. Donnat C, Holmes S. Tracking network dynamics: a survey of distances and similarity metrics. *Phys Stat*. 2018. <https://doi.org/10.48550/arXiv.1801.07351>