Incorporating human mobility to enhance epidemic response and estimate real-time reproduction numbers

Mousumi Roy[1*], Hannah E. Clapham[1], Swapnil Mishra[1, 2, 3†*]

1 Saw Swee Hock School of public health, National University of Singapore, Singapore
2 Institute of Data Science, National University of Singapore, Singapore
3 Department of Statistics and Data Science, National University of Singapore, Singapore

†On behalf of MLGH Network
* {mou.roy, hannah.clapham, swapnil.mishra}@nus.edu.sg

Abstract

Human mobility plays a critical role in the transmission dynamics of infectious diseases, influencing both their spread and the effectiveness of control measures. In the process of quantifying the real-time situation of an epidemic, the instantaneous reproduction number R_t appears to be one of the useful metrics widely used by public health researchers, officials, and policy makers. Since individuals can contract infections both within their region of origin and in other regions they visit, ignoring human mobility in the estimation process overlooks its impact on transmission dynamics and can lead to biased estimates of R_t, potentially misrepresenting the true epidemic situation. This study explicitly integrates human mobility data into a disease transmission model based on the renewal equation to estimate instantaneous reproduction numbers that account for spatial connectivity. Furthermore, we incorporate pathogen-specific generation time distribution, observational delay, and latent period, making the framework epidemiologically informed and flexible to a wide range of diseases. The framework is primarily validated using a simulated dataset and applies it to two different mobility settings at two different spatial scales as case studies: England and the north-east region of England. This analysis also investigates the spatial scalability of the framework, indicating that lower spatial resolution can diminish the effect of inter-regional mobility on the disease transmission, and we conclude that utilizing a finer spatial scale is advantageous on the basis of data availability and computational resources to obtain a better picture of the detailed transmission dynamics. This framework uses non-identifiable, publicly available datasets that are widely accessible, which proves the practical applicability and offers an alternative approach of individual-level modeling. These findings emphasize the necessity of incorporating human mobility into infectious disease models to enhance real-time estimations of effective reproduction numbers. Furthermore, a clear grasp of the transmission scenario also supports better-informed and more targeted public health measures.

Author summary

The real-time or instantaneous reproduction number R_t is a key metric for assessing the state of an epidemic at any given time. When estimating these numbers across multiple connected regions, human mobility plays a crucial role, as movement patterns

significantly influence disease transmission. Traditional epidemic models often assume homogeneous mixing, which does not reflect real-world interactions. On the other hand, individual-based models incorporate heterogeneous mixing at individual level but demands an extremely refined data and substantial computational support. To address these challenges, we employ a renewal equation-based transmission framework, particularly useful for its effectiveness in real-time epidemic analysis, by incorporating heterogeneous mobility flows at a chosen spatial resolution. This yields the estimates of spatially connected instantaneous reproduction number for each region. This improved understanding enables better assessment of the impact of mobility on disease transmission spread, and provides valuable insights for designing targeted epidemic control and intervention strategies.

1 Introduction

Human mobility significantly influences the transmission dynamics of infectious diseases [1, 2]. Hence, its been a constant effort from the researchers to incorporate human mobility within a transmission model in the process of analyzing disease spread [3, 4]. Human mobility pattern also changes with the time course of an outbreak [5]. A clear example was observed during the SARS-CoV2 pandemic, while lockdowns and other NPIs led to substantial reductions in mobility. These shifts were also closely linked to changes in transmission patterns, highlighting the need to account for mobility in the real-time estimation of reproduction numbers. Therefore, ignoring mobility in transmission models can lead to biased parameter estimates and compromise the accuracy of future predictions. In this study, we use *human mobility* to describe the movement of people between regions, which creates *spatial connectivity* and leads to a *mobile population* that affects how diseases spread. These three terms *human mobility*, *spatial connectivity*, and *mobile population* are used interchangeably throughout the paper to convey a similar meaning.

Various modeling strategies have been implemented to study infectious diseases at the population level, each presenting unique benefits and drawbacks. Compartmental mechanistic model [6], for example, in the absence of a spatial component often leads to biased estimation due to its intrinsic assumption of homogeneous mixing, a condition that rarely reflects the real-world mixing pattern [7, 8]. This leads to questions about the reliability of the model within a heterogeneous setting. A network-based model [9, 10], which overlays a network on top of compartmental mechanistic models, can incorporate information about heterogeneous mixing. However, this modeling approach is commonly used to understand the mechanism behind an outbreak, rather than to evaluate the dynamics of disease transmission in real time. On the other hand, individual-level models [11, 12] capture the nuanced complexities of human mobility, but are computationally intensive and require granular data for validation, which is often not easily accessible. An alternative is the approach based on the renewal equation [13, 14], which effectively models infection at the population level and is particularly valuable for real-time analysis of disease transmission. Incorporating a spatial component into the renewal equation-based framework could greatly enhance the estimation of spatially connected parameters, thereby improving the understanding of real-time transmission dynamics in the context of a population with heterogeneous mobility.

In the process of modeling and understanding a real time epidemic situation, the instantaneous or real-time reproduction number (R_t) serves as a metric that quantifies the average number of secondary cases generated from a primary infection at any particular time point. This metric is crucial for assessing the current state of the epidemic and evaluating the effectiveness of any intervention strategies implemented throughout its course [15, 16]. Inaccurate estimation of R_t can lead to a misleading

understanding of the epidemic situation, potentially resulting in misallocation of resources and improper implementation of control policies. To estimate the reproduction number R_t, various forms of observational data, such as, cases, hospitalizations, deaths, etc. can be utilized [14, 17]. Several estimation methods are available, each with its distinct advantages and methodological limitations [13, 17, 18]. When building the modeling framework, it is essential to consider the delay between the actual infection and the reporting of observational data that are used for estimation, particularly towards the end of the timeline of the datasets [19]. When daily observational data are lacking, a temporally aggregated dataset can also be used for estimation purposes [20, 21]. In spite of initiatives aimed at integrating heterogeneity within the renewal equation-based framework for estimating R_t over time [22–24], spatial expansion remains an underexplored direction with the goal of reducing bias within the estimates.

A significant challenge when working with mobility-induced transmission models is the limited availability of human mobility data. Various datasets serve as proxy data for human mobility, such as mobile phone data and public transport data, each with its advantages and limitations [25, 26], however, access to these datasets is often restricted. In that situation, well-established human mobility models can effectively simulate movement between regions, depending on factors like population, distance, and geographical significance [27–29]. In the absence of real-world mobility data, models like the gravity model, visitation model, etc. become particularly valuable tools for generating synthetic mobility patterns and exploring their implications for disease transmission dynamics.

Despite significant efforts by infectious disease modelers to integrate human mobility into transmission models in order to understand its influence on transmission pattern, this area still requires further investigation. In contrast to previous literature with renewal-equation based framework where in most of the studies mobility reduction appears as an aspect of behavioral changes [14, 30], we use explicit human mobility flow within the model to account inter-regional mobility effect over the transmission dynamics at any region. We apply this framework to the initial phase of the SARS-CoV-2 pandemic in England as a case study to illustrate its real-world applicability. Consequently, this study systematically investigates the influence of commuting populations and the epidemic conditions of interconnected regions on disease transmission dynamics, using both simulated data and a case study with necessary adaptations to the model to reflect the specific context. Furthermore, we observe the spatial scale itself is a vital factor, as working with larger spatial scales may diminish the observable effects of inter-regional mobility.

2 Materials and methods

In this study, we consider interconnected regions characterized by commuting populations, reflecting daily movement patterns from home to various destinations. We use a discrete-time renewal equation based approach [13], integrated with a spatial component, to assess the real-time impact of mobility on transmission scenario. This approach helps us understand how heterogeneous commuting patterns across different spatial scales influence transmission dynamics. New infections in any region arise not only from the infected population within the region but also from the mobile population between connected regions. It provides a thorough evaluation of how regional connectivity affects epidemic dynamics within a specific region. The estimated instantaneous reproduction number using this framework accounts both within region transmission and the effect of epidemic status of neighboring regions with the corresponding inter-regional mobility.

2.1 Modeling infection within a region

To model the infections within a region, we use the renewal-equation based framework [13, 14, 18], where new infections at time t are described as,

$$I_t = (S_t/N)R_t \sum_{\tau=0}^{t-1} I_\tau g_{t-\tau} \tag{1}$$

where, $S_t = N - (\sum_{s=1}^{(t-1)} I_s)$ represents the susceptible population at time t, and N denotes the total population. Eq.(1) illustrates that the number of new infections at any given time depends on the prior infections, with their contributions weighted by the generation distribution, g, which captures time interval between primary and secondary infections. Here g is derived by discretizing the pathogen-specific density function $g(t)$ as $g_s = \int_{s-0.5}^{s+0.5} g(t)dt$ for $2, 3, 4, ...$ and $g_1 = \int_0^{1.5} g(t)dt$. The instantaneous reproduction number R_t quantifies the average number of secondary infections caused by a single primary infection at time t. Eq.(1) provides a robust foundation for understanding the transmission dynamics within a single region, which we refer to as the national model.

2.2 Extension of the framework with a spatial component

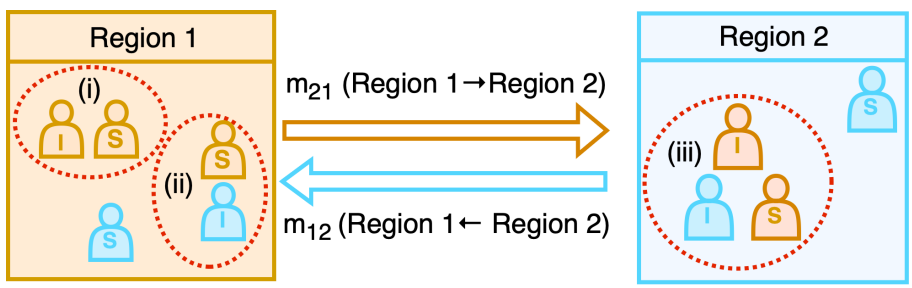

Fig 1. Schematic representation of disease transmission dynamics with mobility. The diagram illustrates two interconnected regions (Region 1 and Region 2), where individuals are classified as susceptible (S) and infected (I) can move between regions. The colors indicate the region of origin for each individual. The arrows represent mobility flows: m_{21} (movement from Region 1 to Region 2) and m_{12} (movement from Region 2 to Region 1). The diagram describes the disease transmission pathways for region 1, highlighted by red ovals: (i) local transmission due to interactions between susceptible and infected individuals of region 1 within region 1, (ii) infected individuals of region 2 commute to region 1, can transmit infection to the susceptible population of region 1, (iii) susceptible individuals from Region 1 who travel to Region 2, acquire the infection in region 2, and return as infected individuals. This framework captures the role of human mobility in shaping spatial disease transmission patterns.

Central to this study is the advancement of the existing epidemiological framework through the integration of a spatially explicit component, which systematically incorporates human mobility data to elucidate the mechanistic role of population movement in shaping the spatiotemporally heterogeneous disease transmission dynamics. A schematic representation of the extended framework is shown in Fig 1. For simplicity, we consider two regions: region 1 and region 2 represented by two different colors. The population in each region comprises susceptible and infected individuals, with a portion of the population commuting between regions. Additionally, we assume that recovery from the disease confers lifetime immunity, as recovered individuals cannot be reinfected. The mobility

between two regions $m_{ij}(j \to i)$, represents the proportion of the population of region j that commutes to region i. Hence, in Fig 1, m_{12} denotes the fraction of the population of region 2 that commutes to region 1, and similarly, m_{21} represents the reverse flow. The presence of mobile population implies that the infection in any region are influenced not only by local transmission but also by interactions with infected individuals from other connected regions. To illustrate the transmission dynamics in more detail, we focus on infection pathways affecting the population of region 1. As illustrated in Fig 1, the red ovals indicate three different routes through which individuals in region 1 can acquire infection. Susceptibles in region 1 can be exposed to infected individuals within their own region or can acquire infection by visiting region 2 and contract infection there. In summary, infection within any region can occur through three primary pathways: (i) local transmission within the region by its own infected population, (ii) infection through incoming infected commuters from other regions, and (iii) acquisition of infection during visit to another regions. These mobility-driven transmission mechanisms underscore the complexity of disease spread in interconnected populations.

Drawing from these observations, we develop a modeling framework that explicitly incorporates human mobility into disease transmission dynamics. Using Eq.(1), as the foundational structure, we extend the model by integrating a spatial component to account for mobility flows between regions. For the modeling framework, we consider a total of M inter-connected regions, and we model the infection dynamics in a region denoted as i, with the number of infections at time t represented as $I_{i,t}$, while allowing for population movement across M regions. At time t, the susceptible individuals from region i commuting to another region j are represented as $m_{ji}(t)S_{t,i}$. These susceptible individuals are exposed to all the infected population in region j at that time, given by $\sum_{k=1}^{M} m_{jk}(t) \sum_{\tau=0}^{t-1} I_{\tau,k} g_{t-\tau}$. This interaction between susceptible individuals from region i and infected population in region j results in new infection of region i, albeit occurring in region j. Since individuals from region i can commute to any regions, summing over all regions j yields the total infections in region i that is given by,

$$I_{t,i} = \sum_{j=1}^{M} \frac{m_{ji}(t)S_{t,i}}{\sum_{l=1}^{M} m_{jl}(t)N_l} R_{t,j} \sum_{k=1}^{M} m_{jk}(t) \sum_{\tau=0}^{t-1} I_{\tau,k} g_{t-\tau}, i = 1, 2, 3, ..., M. \tag{2}$$

Here, the denominator serves as a normalization term, describing the effective population in region j at that time. $R_{t,j}$ is the reproduction number of region j, indicates the average number of secondary infections occurring from a primary infection in that region. g is the generation distribution, and N_i is the total population of region i, while $S_{t,i}$ represents the number of susceptible individuals left in region i at time t. In the special case of completely isolated regions, the mobility matrix becomes an identity matrix (unity on the diagonal and with zero elsewhere). With this simplification, in the absence of inter-regional mobility, Eq.(2) reduces to Eq.(1) for each region.

As the next step, for each region i, the model is fitted to available observational data, such as infection data, case counts, or death data, to estimate the region-specific instantaneous reproduction number R_t. The model can be adjusted based on the nature of the available datasets and the specific context. We demonstrate the utility of our model through two distinct applications. First, we employ simulated daily incidence data as observations to validate the model in a controlled setting. Second, we apply the model to real-world data by analyzing the early spread of SARS-CoV-2 in England, using weekly mortality data as observations. These examples illustrate the flexibility of the framework in accommodating both different types of data and their aggregated nature. This spatially extended framework offers a comprehensive understanding of the impact of human mobility on disease transmission dynamics across interconnected regions. By explicitly incorporating mobility flows into the infection process, the model captures both local and travel-mediated transmission, providing a more realistic representation of

pathogen spread in spatially structured populations. Here we provide a detail description of three sources of infection.

2.3 Sources of infections:

Infections within a region can originate from three different sources. Using Eq.(2), we segregate the infected population based on the three different sources of transmission:

1. **Infections driven by own infected population within the region:** These infections occur within the region when susceptible individuals contract the disease from the infected population of their own region. This category captures purely local transmission dynamics and can be expressed as:

$$I_{t,i}^{in,loc} = \frac{m_{ii}(t)S_{t,i}}{\sum_{l=1}^{M} m_{il}N_l} R_{t,i} \left[m_{ii}(t) \sum_{\tau=0}^{t-1} I_{\tau,i} g_{t-\tau} \right], i = 1, 2, 3, ..., M. \quad (3)$$

2. **Mobility induced infections within the region:** These infections are caused by infected individuals from other regions visiting the region and transmitting the disease to the local susceptible population. This category is defined as infections due to mobility inside the region and can be expressed as,

$$I_{t,i}^{in,mob} = \frac{m_{ii}(t)S_{t,i}}{\sum_{l=1}^{M} m_{il}N_l} R_{t,i} \left[\sum_{k=1,k\neq i,} m_{ik}(t) \sum_{\tau=0}^{t-1} I_{\tau,k} g_{t-\tau} \right], i = 1, 2, 3, ..., M. \quad (4)$$

3. **Mobility induced infections outside the region:** These infections occur when individuals from a region commute to other regions, become exposed there, and get infected. This category is defined as infections due to mobility outside the region and can be expressed as :

$$I_{t,i}^{out,mob} = \sum_{j=1,j\neq i}^{M} \frac{m_{ji}(t)S_{t,i}}{\sum_{l=1}^{M} m_{jl}N_l} R_{t,j} \left[\sum_{k=1}^{M} m_{jk}(t) \sum_{\tau=0}^{t-1} I_{\tau,k} g_{t-\tau} \right], i = 1, 2, 3, ..., M. \quad (5)$$

2.4 Simulated data generation

We use a simulated infection dataset to validate the framework. The simulated dataset assumes three regions with total populations of 20000000, 10000000, and 15000000, respectively. These regions are interconnected through a mobility matrix (see Fig 2(b)), where $m_{ij}(j \to i)$ represents the fraction of the population of jth region that commutes to region i. Therefore, in the mobility matrix, each column corresponds to the outflow of individuals from a specific region (i.e., the jth column represents outflows from region j), while each row represents the inflows into a region (i.e., the ith row represents inflows into region i). Consequently, the sum of each column equals unity. The diagonal elements of the mobility matrix indicate the fraction of the population that does not commute, staying within their respective regions. According to the intrinsic characteristic of the framework, infections are assumed to occur homogeneously within each region governed by the corresponding reproduction numbers, $R_{t,i}, i = 1, 2, 3.$ (see Fig 2(c)). Based on the reproduction numbers, mobility matrix, and using Eq.(2), simulated data are generated as shown in Fig 2(d). The colors of the reproduction number and incidence data correspond to their respective regions. The generation distribution is obtained by discretizing the density function $g \sim Gamma(6.5, 0.62)$ [14,31], a SARS-CoV-2 specific generation distribution. Here, we consider the gamma distribution with mean 6.5 and

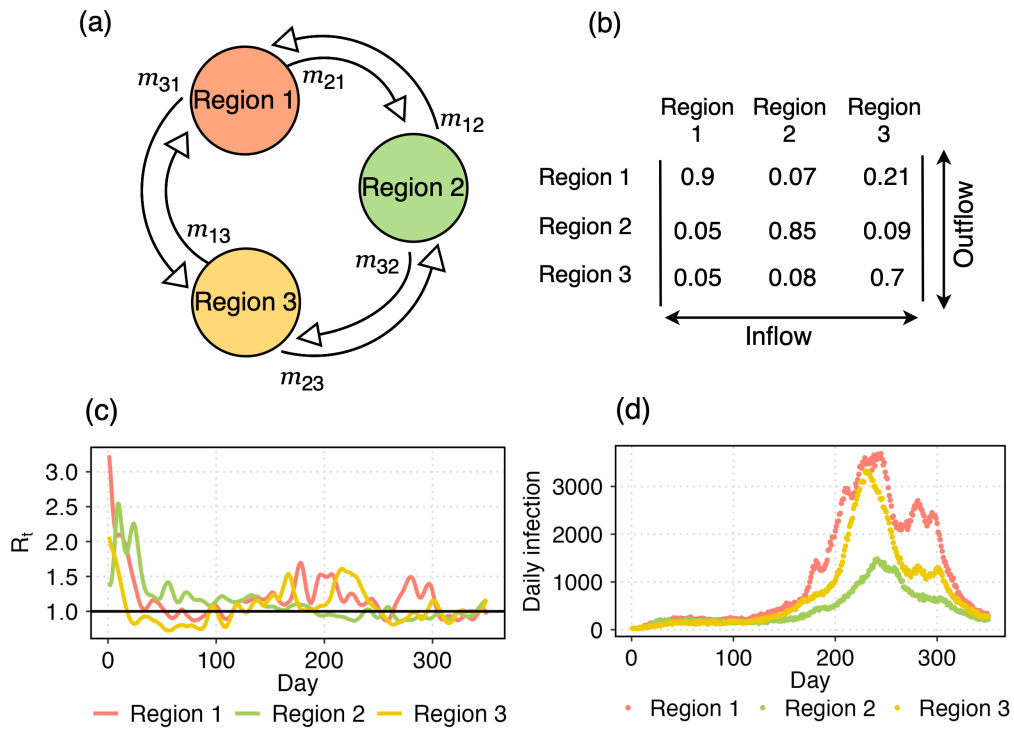

Fig 2. Simulated dataset (a) Three regions with the associated mobility flow. (b) The mobility matrix where each element m_{ij} denotes the fraction of population of region i commutes to region j. Therefore, the columns show the outflows from the respective region and each row represents the inflow to that specific region. (c) Instantaneous reproduction numbers R_t for the three regions. (d) With the mobility matrix and reproduction numbers, we simulate daily infection data using Eq.(2).

CV 0.62. However, in real-world scenarios, human mobility, total population, and observational datasets are typically the only available data sources. Using these, the goal is to estimate the instantaneous reproduction numbers R_t, for all regions to understand the epidemic dynamics in the presence of human mobility.

2.5 Model adjustments and assumptions for estimating R_t with the simulated data

R_t is modeled as geometric random walk, where at any time t, $log(R_t)$ is updated according to a normal distribution with the mean $log(R_{t-1})$ and a standard deviation σ as follows:

$$\begin{aligned} log(R_t) &\sim \mathcal{N}(log(R_{t-1}), \sigma) \\ log(R_0) &\sim \mathcal{N}(-1, 0.5) \\ \sigma &\sim \mathcal{N}^+(0, 0.5) \end{aligned} \tag{6}$$

Here, R_0 is assigned a weak prior to consider a wide range of values, and σ follows a truncated normal distribution (constrained to positive values). For connected R_t estimation, we use the mobility matrix as defined in Fig 2(b). In the disconnected case, we consider the matrix is a diagonal matrix with ones on the diagonal and zeros elsewhere, representing a scenario in which the entire population of each region does not commute to any other region.

2.6 Model adjustments and assumptions for case study: SARS-CoV-2 in England and North East region of England (2020)

We apply the framework to analyze the initial phases of the SARS-CoV-2 pandemic in England during 2020 as a case study. To address the challenge of low reporting rates in the early stages, we use weekly death data to estimate R_t for the nine regions of England: North East, North West, Yorkshire and the Humber, East Midlands, West Midlands, East, London, South East and South West. The mobility matrix (see Fig B(a) in S1 Text) is constructed over the nine regions. Weekly death data for each region are shown in Fig C in S1 Text. The framework is adapted for estimating reproduction numbers, modified the base model(2) as outlined below:

We use the same generation distribution, $g \sim Gamma(6.5, 0.62)$, as in the previous section to simulate infections. For death modeling, we consider the infection-to-death delay distribution $\pi \sim Gamma(5.1, 0.86) + Gamma(17.8, 0.45)$, which combines the infection-to-onset and onset-to-death [14, 30, 32] delays. Here, the gamma distributions are parametrized as Gamma(mean, CV).

The weekly death data $D_{w,i}$ for region i in week w is modeled using a negative binomial distribution:

$$D_{w,i} \sim \text{Negative binomial}\left(d_{w,i}, d_{w,i} + \frac{d_{w,i}^2}{\psi}\right)$$

$$\psi \sim \mathcal{N}^+(0, 5)$$

Here, $d_{w,i}$ is the expected number of weekly deaths, calculated as the sum of daily deaths over seven days. Daily deaths $d_{t,i}$ are derived from daily infections $I_{t,i}$ using the discrete sum,

$$d_{t,i} = ifr_i^* \sum_{\tau=0}^{t-1} I_{\tau,i}\pi(t-\tau),$$

where $ifr_i^* = ifr \cdot N(1, 0.1)$ represent the infection fatality ratio [14] of England with a noise around it corresponding to region i.

To estimate the instantaneous reproduction number $R_{t,i}$ for every region, we introduce the human behavioral changes through five different types of information from Google mobility data: retail and recreation, groceries and pharmacies, parks, transit stations, workplaces. Here , we define R_t as follows:

$$R_{t,i} = R_{0,i} \cdot f\left(-X_i P_t \alpha_{t,i} - \epsilon_{w_{t,i}}\right),$$

where, $f(x) = 2exp(x)/(1+exp(x))$, is twice the inverse logit function. $R_{0,i} \sim \mathcal{N}(3.28, \kappa)$ with $\kappa \sim \mathcal{N}^+(0, 0.5)$ for all regions. $\epsilon_{w_{t,i}}$ is the weekly effect that follows a random walk centered at 0, with the week index $w_{t,i} = [\frac{t-t_0}{7}] + 1$, where t_0 is the first day of seeding.

X_i is the region-specific effect for covariates $\alpha_{t,i}$, which are the averages of five different dimensions observed from Google mobility data. P_t is an index parameter capturing behavioral effects related to different lockdown periods, with distinct values x, y and z for the period till the end of first lockdown, from the end of first lockdown until the end of second lockdown, and post-second lockdown, respectively. We assume that initial seeding begins six weeks prior to the date on which each region in England reaches 10 cumulative reported deaths. From that date, we seed equal initial infections for six consecutive days at each region, modeled as:

$$I_{1,i} = I_{2,i} = I_{3,i} = I_{4,i} = I_{5,i} = I_{6,i} \sim \text{Exponential}(\frac{1}{\tau}),$$

$$\tau \sim \text{ Exponential}(0.01), \quad \text{i} = 1, 2, 3, ..., \text{M}.$$

By tailoring the model to account for spatial and temporal variations in epidemic trends, the framework provides a robust tool for understanding and managing regional dynamics of SARS-CoV-2 in England in the year 2020. The model fitting is shown in Fig D in S1 Text for nine regions of England.

We then apply the framework to a finer spatial scale, focusing shifts to the North East region, which is selected for detailed analysis because nearly 99% of its population does not commute to other regions. This characteristic makes it an ideal area for studying the relationship between human mobility and disease transmission on a more localized scale. The mobility matrix (see Fig B(b) in S1 Text) is constructed for the twelve lower-tier local authorities (LTLAs) within the North East region, ignoring mobility to regions outside of it. Notably, while 99% of the population does not commute to other regions, the mobility matrix for North East region reveals that a significant portion of the population commutes within the region, i.e., moving between LTLAs. We use the same model described in Eq.(2) and Sec.2.6, with data specific to the North East region (see Fig G in S1 Text). The model fitting across the 12 LTLAs are illustrated in Fig H inS1 Text.

3 Results

3.1 Model validation with simulated dataset

To validate the proposed framework (described in Sec.2), we compare the estimated R_ts with the ground truth values (see Fig 2(c)) used in the simulation. This validation ensures that the framework reliably captures the transmission dynamics under controlled conditions before applying it to real-world scenarios. In Fig 3((a)-(c)), we present the results on the simulated dataset to estimate the reproduction numbers R_t for three regions. The red curves represent the disconnected R_t, estimated under the assumption that no commuting population exists between the regions. The green curves represent the connected R_t estimated while accounting for mobility across the regions. The 95% credible intervals representing the uncertainty, are shown as the shaded bands around the mean estimates. The blue lines depict the original R_t as shown in Fig 2(c). The black horizontal dashed line indicates the $R_t = 1$ threshold, which differentiates epidemic growth ($R_t > 1$) or decline ($R_t < 1$).

Fig 3((a)-(c)) reveal clear discrepancies between the disconnected and connected estimates of R_ts. The red curves, representing R_t estimates that do not account for inter-regional mobility, frequently deviate from the actual R_t. These deviations are particularly pronounced in Region 2, where the disconnected R_t often provides misleading insights into the epidemic dynamics. Since the disconnected model does not address mobility-driven infections or inter-regional transmission, it intrinsically assumes that individuals remain within their respective regions and that infections occur exclusively within regional boundaries. For instance, between days 150 and 250 (Fig 3(b)), the disconnected R_t remains above the epidemic threshold of one, whereas the actual R_t (blue curve) exhibits a downward trend, indicating a declining epidemic. This discrepancy can be further understood by examining Fig 3(e), which shows that although the total infections in Region 2 increased during this period, a closer examination, as described in detail in the next paragraph, reveals the majority of new infections originated from mobility-driven transmission. Another notable discrepancy occurs between days 50 and 150 (Fig 3(b)), where the disconnected R_t underestimates the actual R_t, suggesting a lower epidemic severity than observed. However, the actual R_t follows a distinct trajectory during this period, highlighting the limitations of the disconnected model in capturing inter-regional transmission. The disconnected model shows a direct correspondence between the trend of R_t with the increase or decrease in the total number of infections within a region,

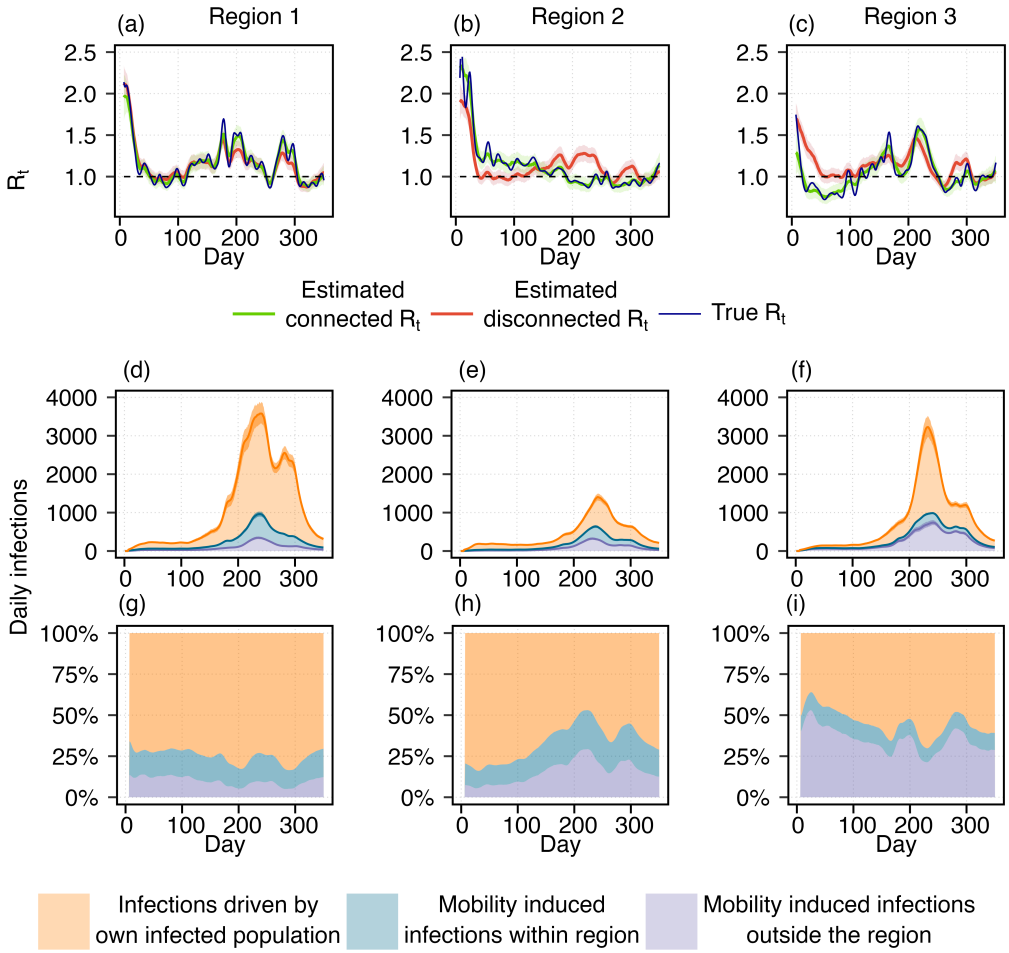

Fig 3. Estimation of time-varying reproduction number (R_t) using simulated data and contribution of mobility-driven infections. Top row (a–c): Blue curves represent the true R_t trajectories for three regions, with horizontal dotted lines indicating the epidemic threshold ($R_t = 1$). Green and red curves show the estimated R_t values with and without incorporating human mobility, respectively. Shaded areas around the curves denote the 95% credible intervals. The deviations between the two estimates highlights the importance of accounting for human mobility in transmission modeling. Bottom row (d–f): Stacked plots illustrate the three sources of daily infections. Orange shows the infections caused by the infected population of that same region. Blue shows the infections that generated within the region by infected incoming population form the other regions. Purple reflects mobility-induced infections acquired outside the region by the outgoing susceptibles. (g)-(i): The proportion of infections from three different sources. The colors are considered same as defined for the Figs.(d)-(f). These figures represent the considerable number of mobility driven infections in the disease transmission dynamics.

without focusing on the mobility driven infections. Unlike the disconnected model, the connected framework accounts for these mobility-induced infections, resulting in a spatially connected estimation of R_t. As illustrated in Fig 3, the connected model (green curves in the upper row) provides a more detailed representation of transmission dynamics

by incorporating the effects of mobility between regions. These results highlight the importance of incorporating mobility into the model during the estimation of reproduction number. Relying on disconnected R_t can lead to severe misleading understanding of the epidemic situation, as seen in Region 2. This enhanced understanding is crucial for real-world applications, particularly in policy-making and resource allocation, where epidemiological estimates are essential for effective decision-making.

Fig 3((d)-(i)) illustrate the absolute number and proportion of infections arising from three distinct sources. The orange regions represent the infections occurring within the same region by its own infected population. The blue color indicates mobility-induced infections within the region, which occur when infected individuals from other regions visit and cause transmission. Finally, the purple regions denote the infections acquired outside the region when individuals travel to other regions and get infected there. Therefore, combining the above description, the blue and purple highlight infections driven by mobility. In this Fig 3((d)-(f)), the lines are the mean of the infection from different sources and the shaded bands indicate the corresponding 95% credible intervals. It is observed that a significant portion of infections is influenced by mobility, and this becomes especially important when R_t varies across regions and the epidemic behaves differently from one region to another. For instance, mobility can amplify infections when susceptible individuals travel to regions with increasing epidemic trends. These infections can subsequently trigger outbreaks in regions where the epidemic is otherwise stable or declining. This phenomenon is clearly observed in Fig 3(f) and Fig 3(i) up to day 200. During this period, region 3 experiences a substantial number of infections acquired outside as people travel to regions 1 and 2, which have higher R_t values. Despite region 3 having a lower local R_t, the mobility-driven infections inflate its overall incidence. This mobility driven effect cannot be captured by the disconnected model, which overlooks inter-regional interactions. As a result, the disconnected model overestimates R_t in region 3 during this period. In contrast, the connected model, which incorporates mobility, accurately captures this dynamics and aligns with the true epidemic trends (see Fig 3(c)). A detailed discussion in this context is provided in Fig A in S1 Text.

These results validate the effectiveness and necessity of incorporating human mobility into the model to accurately estimate reproduction numbers and better understand the epidemic situation. This approach is essentially critical in scenarios involving spatial heterogeneity in transmission dynamics.

3.2 Case study: Instantaneous reproduction number estimates over the regions of England

In Fig 4, we present the estimates of the instantaneous reproduction number (R_t) for the nine regions of England, comparing results with and without account for mobility. In Fig 4, the green curves represent the connected R_t estimates, which incorporate mobility flows between regions, while the red curves represent the disconnected R_ts, which disregards inter-regional mobility. The corresponding 95% uncertainty bands are shown in the same colors. The vertical dotted lines indicate the start and end of the first and second lockdowns, while the black horizontal line at $R_t = 1$ indicates the epidemic threshold, distinguishing growth ($R_t > 1$) from mitigation ($R_t < 1$). From Fig 4, we observe that the connected and disconnected estimated R_ts are nearly identical across all regions. This suggests that mobility does not have a significant impact on infections on the regional scale. Upon examining the mobility matrix (see Fig B(a) in S1 Text), it is observed that the fraction of the population commuting between regions is very small relative to the total population of each region. This makes the regions of England treatable as disjoint areas, where the disconnected model is sufficient for R_t estimation at this spatial scale. Transmission at the regional level can be treated as homogeneous

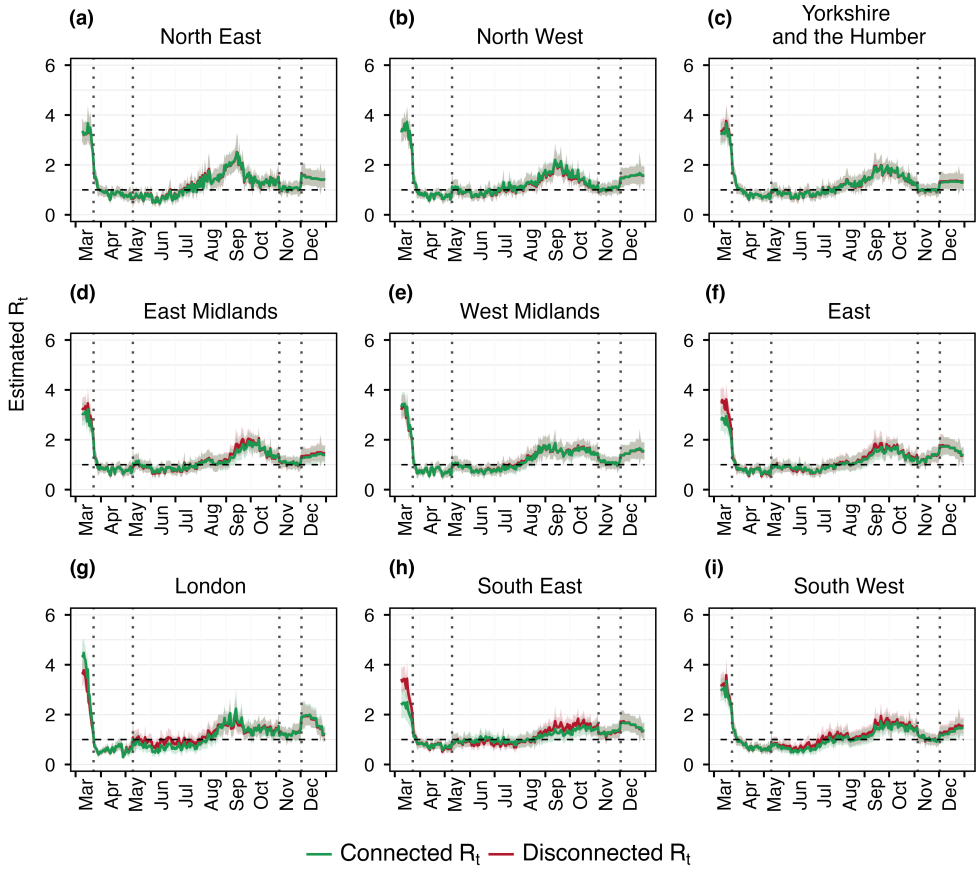

Fig 4. Estimated R_t over different regions of England. The red curve shows the disconnected R_t which is the reproduction number estimated separately for each region without considering mobility across them. Green curve is the connected R_t that is the estimated reproduction number with the considered mobility matrix showed in Fig B(a) in S1 Text. The vertical dotted lines show the first and second lockdown period, and the horizontal black solid line is the threshold for reproduction number that is $R_t = 1$.

spreading within regions, with minimal influence from inter-regional mobility. Therefore, at large spatial scales where inter-regional mobility accounts for only a negligible fraction of the total population, its impact on disease transmission becomes insignificant and does not substantially affect the R_t estimates. This suggests that, when the disease is already established at all regions, mobility can be reasonably excluded from models at such spatial scales.

3.3 Findings on higher spatial resolution: North East region

To assess whether the estimates using the model provide any additional or different insights when applied at a finer spatial scale, we apply the framework at higher resolution. The high proportion of mobile populations between the LTLAs of North East region has a notable impact on disease transmission, as evidenced by Fig 5. Significant differences in the estimated reproduction numbers are observed in LTLAs such as Hartlepool, Middlesbrough, and North Tyneside. These variations highlight that a considerable number of infections arise from movement between LTLAs, as shown in Fig 6, and the estimated disconnected R_ts fail to accurately capture these transmission dynamics,

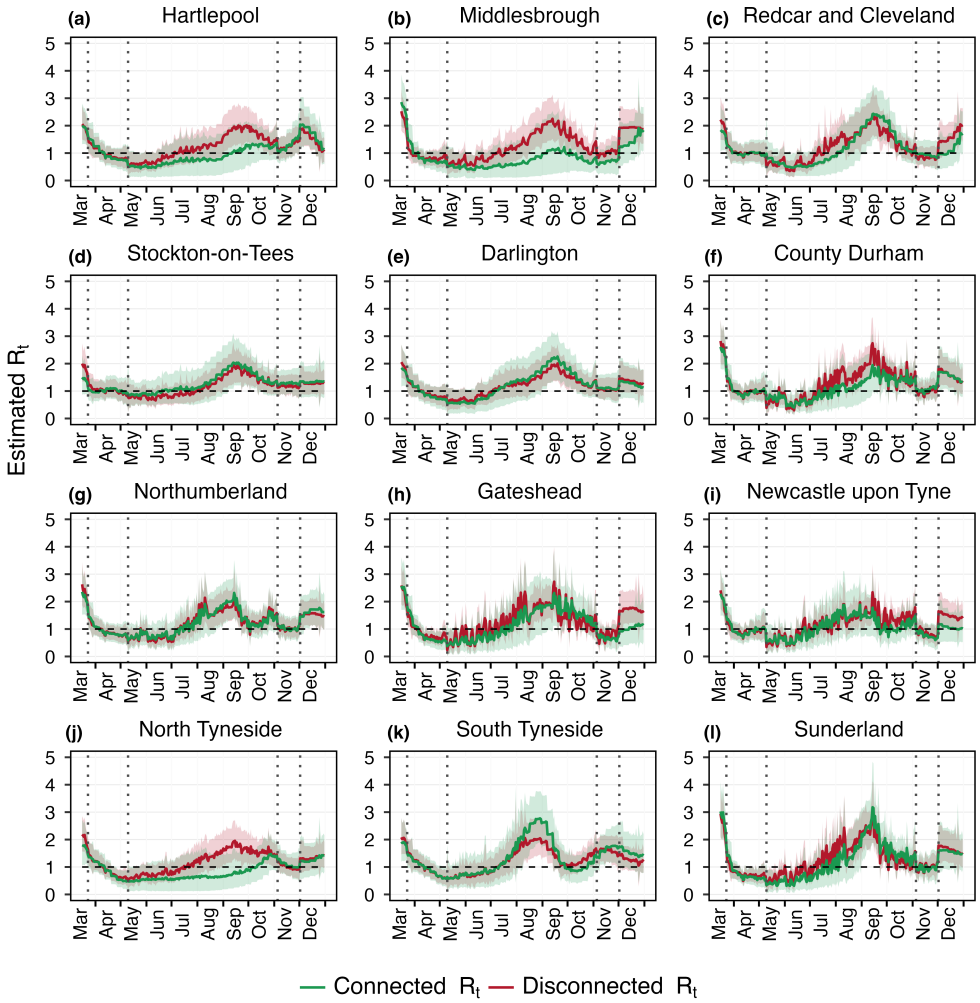

Fig 5. Estimated R_t over the LTLAs of North East region. The green and red curve shows the estimation of real-time reproduction number with and without considering mobility respectively. The dotted vertical lines represents the lockdown period and black horizontal line shows the threshold for reproduction number at $R_t = 1$. Red and green band show the 95% uncertainty interval corresponding to the disconnected and connected R_t estimates. The mobility matrix is shown in Fig B(b) in S1 Text.

often resulting in underestimation or overestimation of the disease transmission. This underscores the importance of incorporating intra-regional mobility into models to better understand localized disease transmission patterns.

In Fig 6 and Fig 7, we examine the absolute number and the proportion of daily infections originating from three distinct sources across the 12 LTLAs. The infections from different sources are depicted by the colors orange, blue and purple. Orange indicates infections resulting from local transmission by own infected population within the LTLA, while blue and purple represent mobility-driven infections occurring inside and outside the LTLA, respectively. While Fig 5 highlights the mobility-driven effects on the R_t estimates, Fig 6 and Fig 7 provide direct evidence of a substantial portion of infections arising from mobility. In particular, areas such as Hartlepool, Middlesbrough, Newcastle upon Tyne, and North Tyneside show mobility-driven infections contributing

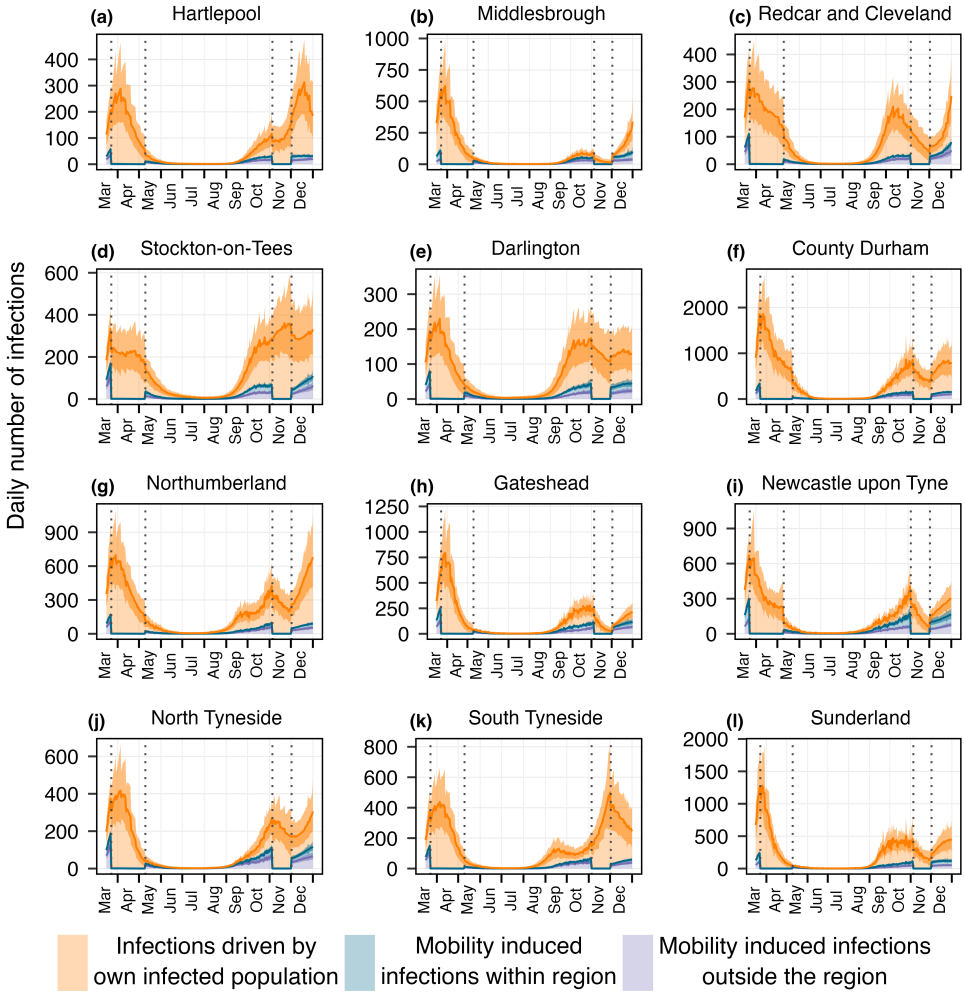

Fig 6. Daily infections from three different sources over the LTLAs of North East region. Orange color represents the infections occur within the LTLA due to the infected population of that same LTLA. Blue and purple depict the mobility induced infections happen within and outside the LTLA. The dotted vertical lines indicate two lockdown periods while mobility driven infections go down and most of the infections are generated by the local transmission. Here, the lines represent the mean of the estimates and band shows the 95% uncertainty around it.

to approximately 50% of the total cases, with the proportion rising to as high as 75% in North Tyneside. During the lockdown periods, the reduction in mobility led to a marked decrease in mobility-driven infections, resulting in transmission dynamics being predominantly driven by local spread. These findings underline the crucial role of inter-LTLAs mobility in shaping transmission dynamics and emphasize the importance of accounting for such mobility in public health interventions and resource allocation strategies.

While the effect of mobility may not be significant at broader regional scales, it becomes highly evident at smaller spatial scales, such as LTLAs. This highlights that the spatial scale plays a critical role in assessing the impact of mobility on disease transmission, with finer resolutions revealing dynamics that may otherwise remain obscured over bigger areas.

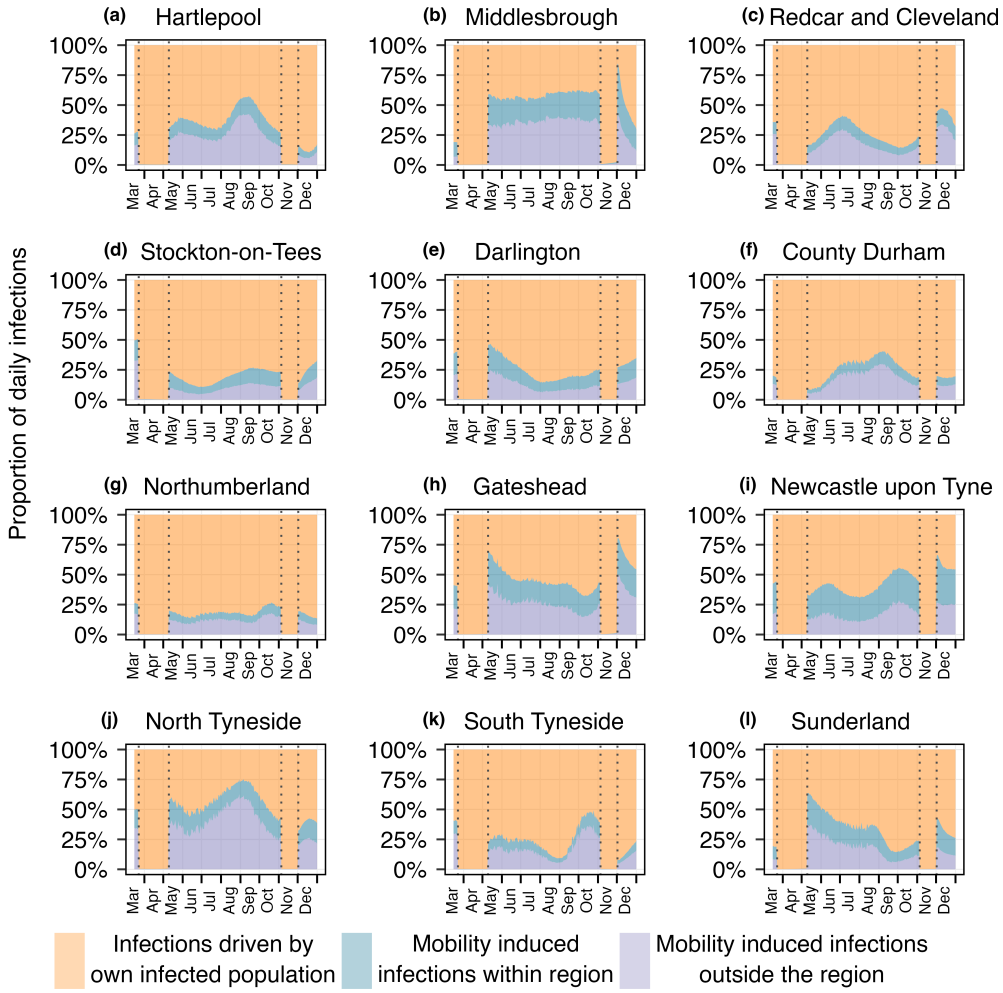

Fig 7. Proportion of daily infections from three different sources across the LTLAs of the North East region. Orange, blue, and purple indicate the proportions of daily infections attributed to three distinct sources. The dotted vertical lines represent two lockdown periods, during which nearly all infections resulted from local transmission.

4 Discussion

In this study, we developed a spatially extended epidemiological framework that improves the renewal equation based disease transmission model by incorporating human mobility. Unlike approaches that treat mobility as a reductive factor within a single region [13, 14], we explicitly accounted for incoming and outgoing infected populations, offering a more comprehensive representation of spatial transmission dynamics. The results highlight the substantial role of mobile population in shaping the dynamics of infectious disease transmission and in influencing the estimation of time-varying reproduction numbers (R_t). While some recent studies have attempted to integrate mobility flows into transmission models [24], further investigation is needed to address key aspects such as the contribution of different sources of infection, the role of spatial scale, etc. Through applications to both simulated data and real-world SARS CoV-2 death data in England, our study addresses these limitations and demonstrates that overlooking inter-regional mobility

can lead to significant biases in epidemic assessment and incorrect depiction of disease transmission trends.

Using simulated data, we validated our framework by comparing estimated R_t values with known ground truth trajectories. In scenarios with heterogeneous epidemic intensities across regions, we observed that disconnected models (which ignores mobility) often tend to overestimate or underestimate the true R_t, as they fail to account for mobility-driven infections that redistribute transmission risk across regions. When mobility is not accounted for, estimated reproduction numbers (R_t) of each region tend to closely follow the rise or fall in total infections in that region. In such cases, increases in infections simply lead to higher disconnected R_t estimates, and similarly, when infections decline, disconnected R_t estimates also decrease. It does not account for the infections that happen by the infected population from other regions due to mobility. In contrast, the connected model incorporated a mobility matrix that captured human mobility flow, hence the estimated R_t aligned closely with the ground truth and revealing the key contribution of mobility-driven infections to the observed epidemic patterns. The decomposition of infection sources into local, incoming, and outgoing mobility-driven transmissions provided more insights about the effect of mobility on the transmission process. These results proves the importance of accounting for human movement to improve the precision of epidemic models and real-time surveillance.

We applied this framework to SARS-CoV-2 weekly death data across nine regions of England, our framework revealed that mobility had a limited impact at coarse spatial resolution. At this spatial scale (regional level), the proportion of the population commuting between regions was minimal, and transmission occurred primarily within each region. Consequently, disconnected models provided comparable estimates of (R_t), suggesting that for large administrative units with limited inter-regional mobility, the added complexity of mobility integration may not be essential for real-time epidemic tracking. However, the situation changed markedly at a finer spatial resolution. Focusing on the 12 Lower Tier Local Authorities (LTLAs) within the North East region, we observed substantial inter-LTLA mobile population and, correspondingly, a clear difference between connected and disconnected (R_t) estimates. In certain LTLAs such as North Tyneside and Middlesbrough, mobility-driven infections accounted for up to 75% of total infections at some time points. During lockdown periods, reductions in mobility were reflected in a shift back toward local transmission dominance. These findings underscore that the impact of mobility on transmission and the utility of mobility-informed modeling depends strongly on the spatial scale of analysis.

Our framework provides a flexible and computationally efficient alternative to individual-based models that require extensive data and computational resources. By incorporating real-world human mobility flows, this model supports a more detailed understanding of how regional transmission trends interact and evolve over time. In the context of spatially heterogeneous epidemics, the impact of population movement on transmission dynamics becomes more pronounced, where differences in reproduction numbers across regions can trigger outbreaks in low-risk areas through population movement. This is especially relevant in real-world contexts, where such disparities are common during the early phases of an epidemic. As the epidemic progresses and spreads more uniformly across regions, local variations may still emerge due to differences in policy implementation, healthcare capacity, demographic profiles, or socio-economic factors. For public health policymakers, these insights are extremely useful. Accurate estimation of spatially connected R_t values can guide targeted interventions, such as localized lockdowns or vaccination prioritization, especially in regions where mobility plays a substantial role in disease spread. Failing to account for these dynamics may result in missed opportunities to suppress transmission or, conversely, in unnecessary restrictions in regions not driving epidemic growth.

A key strength of this framework lies in its generalizability. It can be used for different diseases as the disease-specific epidemiological parameters such as serial interval, incubation period, infection-to-death delay are adapted within the framework. It can be applied to different mobility settings, as an example in this study we considered a high and a low inter-regional mobility settings. The ability to integrate various types of observational data like cases data, hospitalization data, mortality data and even with temporally aggregated data, makes this framework extremely valuable during the early phases of an epidemic or in settings with limited case reporting, and researchers need to rely on the alternative data sources. However, its accuracy is contingent upon the quality of the mobility data. In our case study, we relied on publicly available commuting data, which are often aggregated and may not reflect real-time changes in movement patterns, such as those induced by behavioral shifts or NPIs. Future work could enhance this model by integrating dynamic mobility data, such as mobile phone-based metrics, and extending it to capture other heterogeneities, such as age structure or vaccination coverage.

5 Conclusions

This study highlights the importance of incorporating human mobility explicitly within into models to better understand its impact on disease transmission dynamics. The regions with commuting populations, the incidence of infections does not solely depend on local outbreaks; instead, our framework demonstrates that commuting populations significantly contribute to the emergence of new infections in connected regions. Consequently, to gain a comprehensive understanding of the epidemic situation in any area, integrating human mobility into the model is essential to mitigate biases that may arise from mobility patterns.

Although this framework demands fewer computational resources compared to the individual-based model, the computational requirements may still increase substantially when dealing with a large number of regions, depending on the geographical context and specific research question at hand. Thus, a promising future direction for this study would be to focus on minimizing computational resource demands while maintaining accuracy, particularly for finer spatial scales.

Code availability

All the codes related to this study are available at `https://github.com/mlgh-sg/Spatial_transmission`

Supporting information

S1 Text This document includes all the other figures and datasets related to this study.

Acknowledgments

We are thankful to Zhi Ling for helping with computational understanding with STAN. Mousumi Roy would like to acknowledge National Research Foundation via The NRF Fellowship Class of 2023 (NRF-NRFF15-2023-0010) award (awarded to Swapnil Mishra) for financial support.

References

1. Belik V, Geisel T, Brockmann D. Natural human mobility patterns and spatial spread of infectious diseases. Physical Review X. 2011;1(1):011001.

2. Buckee C, Noor A, Sattenspiel L. Thinking clearly about social aspects of infectious disease transmission. Nature. 2021;595(7866):205–213.

3. Riley S. Large-scale spatial-transmission models of infectious disease. Science. 2007;316(5829):1298–1301.

4. Balcan D, Colizza V, Gonçalves B, Hu H, Ramasco JJ, Vespignani A. Multiscale mobility networks and the spatial spreading of infectious diseases. Proceedings of the national academy of sciences. 2009;106(51):21484–21489.

5. Zhang J, Litvinova M, Liang Y, Wang Y, Wang W, Zhao S, et al. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. Science. 2020;368(6498):1481–1486.

6. Kermack WO. A contribution to the mathematical theory of epidemics. Proceedings of the royal society of london Series A, Containing papers of a mathematical and physical character. 1927;115(772):700–721.

7. Barabási AL, Bonabeau E. Scale-free networks. Scientific american. 2003;288(5):50–9.

8. Gonzalez MC, Hidalgo CA, Barabasi AL. Understanding individual human mobility patterns. nature. 2008;453(7196):779–782.

9. Bansal S, Grenfell BT, Meyers LA. When individual behaviour matters: homogeneous and network models in epidemiology. Journal of the Royal Society Interface. 2007;4(16):879–891.

10. Keeling MJ, Eames KT. Networks and epidemic models. Journal of the royal society interface. 2005;2(4):295–307.

11. Allen LJ. An introduction to stochastic epidemic models. In: Mathematical epidemiology. Springer; 2008. p. 81–130.

12. DeAngelis DL. Individual-based models and approaches in ecology: populations, communities and ecosystems. CRC Press; 2018.

13. Bhatt S, Ferguson N, Flaxman S, Gandy A, Mishra S, Scott JA. Semi-mechanistic Bayesian modelling of COVID-19 with renewal processes. Journal of the Royal Statistical Society Series A: Statistics in Society. 2023;186(4):601–615.

14. Flaxman S, Mishra S, Gandy A, Unwin HJT, Mellan TA, Coupland H, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature. 2020;584(7820):257–261.

15. Parag KV, Thompson RN, Donnelly CA. Are epidemic growth rates more informative than reproduction numbers? Journal of the Royal Statistical Society Series A: Statistics in Society. 2022;185(Supplement_1):S5–S15.

16. Parag KV, Cori A, Obolski U. When is the R= 1 epidemic threshold meaningful? medRxiv. 2024; p. 2024–10.

17. Abbott S, Hellewell J, Thompson RN, Sherratt K, Gibbs HP, Bosse NI, et al. Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts. Wellcome Open Research. 2020;5(112):112.

18. Cori A, Ferguson NM, Fraser C, Cauchemez S. A new framework and software to estimate time-varying reproduction numbers during epidemics. American journal of epidemiology. 2013;178(9):1505–1512.

19. Gostic KM, McGough L, Baskerville EB, Abbott S, Joshi K, Tedijanto C, et al. Practical considerations for measuring the effective reproductive number, R t. PLoS computational biology. 2020;16(12):e1008409.

20. Nash RK, Bhatt S, Cori A, Nouvellet P. Estimating the epidemic reproduction number from temporally aggregated incidence data: A statistical modelling approach and software tool. PLOS Computational Biology. 2023;19(8):e1011439.

21. Ogi-Gittins I, Hart W, Song J, Nash R, Polonsky J, Cori A, et al. A simulation-based approach for estimating the time-dependent reproduction number from temporally aggregated disease incidence time series data. Epidemics. 2024;47:100773.

22. Gorji H, Stauffer N, Lunati I. Emergence of the reproduction matrix in epidemic forecasting. Journal of the Royal Society Interface. 2024;21(216):20240124.

23. Green WD, Ferguson NM, Cori A. Inferring the reproduction number using the renewal equation in heterogeneous epidemics. Journal of The Royal Society Interface. 2022;19(188):20210429.

24. Trevisin C, Bertuzzo E, Pasetto D, Mari L, Miccoli S, Casagrandi R, et al. Spatially explicit effective reproduction numbers from incidence and mobility data. Proceedings of the National Academy of Sciences. 2023;120(20):e2219816120.

25. Hu T, Wang S, She B, Zhang M, Huang X, Cui Y, et al. Human mobility data in the COVID-19 pandemic: characteristics, applications, and challenges. International Journal of Digital Earth. 2021;14(9):1126–1147.

26. Shi L, Chi G, Liu X, Liu Y. Human mobility patterns in different communities: a mobile phone data-based social network approach. Annals of GIS. 2015;21(1):15–26.

27. Barbosa H, Barthelemy M, Ghoshal G, James CR, Lenormand M, Louail T, et al. Human mobility: Models and applications. Physics Reports. 2018;734:1–74.

28. Simini F, González MC, Maritan A, Barabási AL. A universal model for mobility and migration patterns. Nature. 2012;484(7392):96–100.

29. Schläpfer M, Dong L, O'Keeffe K, Santi P, Szell M, Salat H, et al. The universal visitation law of human mobility. Nature. 2021;593(7860):522–527.

30. Unwin HJT, Mishra S, Bradley VC, Gandy A, Mellan TA, Coupland H, et al. State-level tracking of COVID-19 in the United States. Nature communications. 2020;11(1):6189.

31. Bi Q, Wu Y, Mei S, Ye C, Zou X, Zhang Z, et al. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. The Lancet infectious diseases. 2020;20(8):911–919.

32. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. The Lancet infectious diseases. 2020;20(6):669–677.