

# Supplementary Information

Pieter Libin<sup>1,2</sup>, Arno Moonens<sup>1</sup>, Timothy Verstraeten<sup>1</sup>, Fabian Perez-Sanjines<sup>1</sup>, Niel Hens<sup>3</sup>, Philippe Lemey<sup>2</sup>, and Ann Nowé<sup>1</sup>

<sup>1</sup>Artificial Intelligence lab, Department of computer science, Vrije Universiteit Brussel, Brussels, Belgium

<sup>2</sup>KU Leuven - University of Leuven, Rega Institute for Medical Research, Clinical and Epidemiological Virology, Leuven, Belgium

<sup>3</sup>Centrum voor Statistiek, Hasselt University, Hasselt, Belgium

February 21, 2020

## 1 Introduction

This Supplementary Information document accompanies the paper titled “Multi-agent deep reinforcement learning for epidemic control”.

## 2 Census data

Each compartment model is representative of one of the districts defined in the main manuscript, and as such the compartment model is parametrised with the census data of the respective district, i.e., population counts stratified by age groups. We use the 2011 United Kingdom census data made available by NOMIS<sup>1</sup>. This dataset contains census data for all of the considered districts for the following age groups: 0-4, 5-7, 8-9, 10-14, 15, 16-17, 18-19, 20-24, 25-29, 30-44, 45-59, 60-64, 65-74, 75-84, 85-89, 90-90+.

To be compatible with our model, we need to map this census data to the age structure imposed by the Eames contact matrix: i.e., 0-4 years (children), 5-18 years (adolescents), 19-64 years (adults), 65-90+ years (elderly). To define this mapping, we will refer to the number of individuals with the symbol  $N$ , subscripted with the dataset type (i.e., NOMIS or Eames) and the age group.

For the age group 0-4 and 65-90+ we have a direct mapping:

$$\begin{aligned} N_{\text{Eames}, \text{Children}} &= N_{\text{NOMIS}, 0-4} \\ N_{\text{Eames}, \text{Elderly}} &= N_{\text{NOMIS}, 65-90+} \end{aligned} \tag{1}$$

However, as for the contact matrix the adolescents and adults are split between age 18 and 19, and for the census data these 2 age groups are aggregated, we need to make

---

<sup>1</sup><https://www.nomisweb.co.uk>

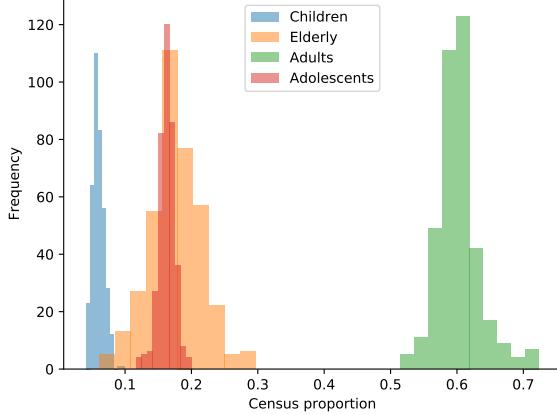


Figure 1: Histograms of the census proportions in the districts of Great Britain, according to Eames' age structure.

a custom mapping. We will aggregate all shared age groups and divide the common age group in two:

$$\begin{aligned} N_{\text{Eames,Adolescents}} &= N_{\text{NOMIS},5-17} + \left\lceil \frac{N_{\text{NOMIS},18-19}}{2} \right\rceil \\ N_{\text{Eames,Adults}} &= \left\lceil \frac{N_{\text{NOMIS},18-19}}{2} \right\rceil + N_{\text{NOMIS},20-64} \end{aligned} \quad (2)$$

When restructuring the census data according to the Eames age groups, we observe clear trends over the districts with respect to the proportion of children, adolescents, adults and elderly, as shown in Figure 1. However, the histograms in Figure 1 only show the marginalized distribution per age group. To reason about the distribution over all age groups, consider that we have a proportion of each of the age groups, and we thus can represent this data as a positive simplex [1], as defined in Definition 1.

**Definition 1** (Unit simplex). *A unit simplex [4], with  $D$  components, corresponds to the set:*

$$\mathbb{S}^D = \{ \langle x_1, \dots, x_D \rangle \mid \forall x_i : x_i > 0, \sum_{i=1}^D x_i = 1 \}. \quad (3)$$

This representation enables us to reason about this data in a statistical framework, and to visualize the four-dimensional data in a three-dimensional space by using the Barycentric coordinate system, as shown in Figure 2. Figure 2 shows that the census distribution exhibits a dense region with a limited number of outliers.

Note that we use the 2011 census dataset, rather than the more recent 2018 census dataset, to be fully compatible with the mobility dataset used to inform our between-patch transition model (see Section 3).

For each district, the base contact matrix is corrected to make it reciprocal, using that district's census data.

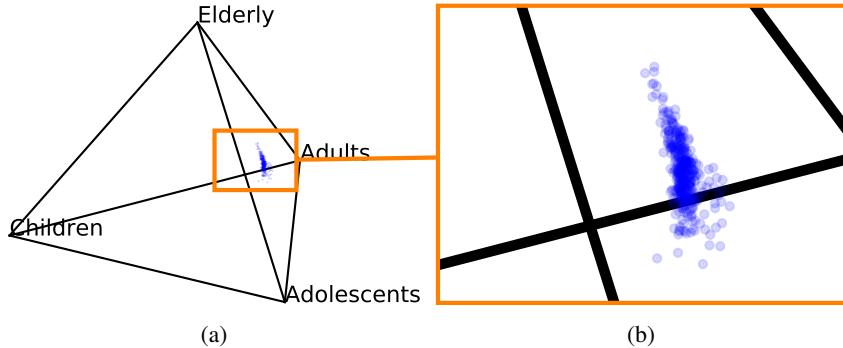


Figure 2: Barycentric projection of the census proportions in the districts of Great Britain, according to Eames’ age structure. Each scatter point corresponds to one district, and each axis corresponds to the proportion of the age groups it connects. The left panel shows the original census pyramid, and the right panel zooms in on the point cloud.

### 3 Between-patch model

Our model, that is comprised of a set of connected SEIR patches, is inspired by the recent BBC pandemic model [19]. The BBC pandemic model was in its turn motivated by the model presented in [16].

At each time step, our model decides whether a patch  $p$  becomes infected. This is modulated by the patch’s force of infection, which combines the potential of the infected patches in the system, weighted by a mobility model:

$$\dot{\phi}_p(t) = \sum_{p' \in \mathcal{P}} \mathcal{M}_{p'p} \cdot \beta \cdot (S_p^A(t))^\mu \cdot \mathcal{I}_{p'}(t), \quad (4)$$

where  $\mathcal{P}$  is the set of patches in the model,  $\mathcal{M}_{p'p}$  is the mobility flux between patch  $p'$  and  $p$ ,  $\beta$  is the probability of transmission on a contact,  $S_p^A(t)$  is the susceptible population of adults in patch  $p$  at time  $t$  and its contribution is modulated by parameter  $\mu$ , and  $\mathcal{I}_{p'}(t)$  is the infectious potential of patch  $p'$  at time  $t$ . We define this infectious potential as,

$$\mathcal{I}_{p'}(t) = I_{p'}^A(t) \cdot M_{AA}, \quad (5)$$

where  $I_{p'}^A(t)$  is the size at time  $t$  of the infectious adult population and  $M_{AA}$  is the average number of contacts between adults.

$\mathcal{M}$  is a matrix based on the mobility dataset provided by NOMIS<sup>2</sup>. This dataset describes the amount of commuting between the districts in Great-Britain.

In general, this between-patch model is constructed from first principles i.e., census data, a mobility model, the number of infected individuals and the transmission potential of the virus. However, for the parameter  $\mu$  that modulates the contribution of the susceptibles in the receptive patch, while it is commonly used in literature [16, 13, 18], no such intuition is readily available. Therefore, this parameter is fitted to match the properties of the epidemic that is under investigation [16, 13, 18].

To validate our model, we conduct two experiments. Firstly, we compare our model to the original compartment model and perform a sensitivity analysis with respect to

<sup>2</sup>We use the NOMIS WU03UK dataset that was released in 2011.

parameter  $\mu$ . Secondly, we show that our model fits the recent influenza pandemic of 2009, by choosing an appropriate value for all model parameters.

Given this time-dependent force of infection, we model the event that a patch becomes infected with a non-homogeneous Poisson process [28, 26, 7]. A Poisson process can be used to model the occurrence of events with a given intensity (see Definition 3), and non-homogeneous Poisson processes generalize this concept to time-dependent intensities (see Definition 4). As the process' intensity depends on how the model (i.e., the set of all patches) evolves, we cannot sample the time at which a patch becomes infected a priori. Therefore, we determine this time of infection using the time scale transformation algorithm [10]. Firstly, we explain the generic time scale transformation algorithm (Section 3.1). Secondly, we adjust the algorithm to our setting (Section 3.2).

### 3.1 Time scale transformation algorithm

The time scale transformation algorithm enables us to determine the time at which an event, modelled by a non-homogeneous Poisson process, will take place [10].

We will start by formally defining the homogeneous and non-homogeneous Poisson process. A Poisson process is a counting arrival process, defined on a sample space  $\Omega$  with probability measure  $P$ .

**Definition 2** (Arrival process). *An arrival process is a stochastic process [10],*

$$\mathcal{N} = \{\mathcal{N}_t; t \geq 0\}, \quad (6)$$

*such that for any  $\omega \in \Omega$ , the mapping  $t \rightarrow \mathcal{N}_t(\omega)$ , has  $\mathcal{N}_0 = 0$ , is non-decreasing, increases only by integer jumps and is right continuous.*

**Definition 3** (Homogeneous Poisson process). *A homogeneous Poisson process is an arrival process [10, 9],*

$$\mathcal{P} = \{\mathcal{P}_t; t \geq 0\}, \quad (7)$$

*for which these axioms hold:*

1. *for almost all  $\omega \in \Omega$ ,  $t \rightarrow \mathcal{P}_t(\omega)$  jumps in steps of size 1*
2. *the number of arrivals within any interval  $[t..t+s]$ , is independent of the history of arrivals prior to  $t$*
3. *the process is time-homogeneous*

From this definition, we can show that for each homogeneous Poisson process  $\mathcal{P}$ :

$$\forall t \geq 0 : P(\mathcal{P}_t = k) = \frac{e^{-\lambda t} (\lambda t)^k}{k!}, \quad (8)$$

for some constant  $\lambda \geq 0$ , where  $\lambda$  signifies the intensity (i.e., rate) of the process.

The concept of a Poisson process can be generalized to a *non-homogeneous* Poisson process by removing the time-homogeneity requirement:

**Definition 4** (Non-homogeneous Poisson process). *A non-homogeneous Poisson process is an arrival process [10],*

$$\mathcal{P}^{\lambda(t)} = \{\mathcal{P}_t^{\lambda(t)}; t \geq 0\}, \quad (9)$$

*for which these axioms hold:*

1. for almost all  $\omega \in \Omega$ ,  $t \rightarrow \mathcal{P}_t(\omega)$  jumps in steps of size 1
2. the number of arrivals within any interval  $[t, t+s]$ , is independent of the history of arrivals prior to  $t$

$\mathcal{P}_t^{\lambda(t)}$  has a time-dependent rate that is specified by the intensity function  $\lambda(t)$ , where  $\lambda(t) \geq 0$ .

We define the process' cumulative intensity function:

**Definition 5** (Cumulative intensity function). A non-homogeneous Poisson process  $\mathcal{P}^{\lambda(t)}$  with intensity function  $\lambda(t)$  has a cumulative intensity function:

$$\Lambda(t) = \int_0^t \lambda(s)ds \quad (10)$$

Furthermore, we can show that [25]:

$$\mathbb{E}[\mathcal{P}_{t+h}^{\lambda(t)} - \mathcal{P}_t^{\lambda(t)}] = \int_t^{t+h} \lambda(s)ds, \quad (11)$$

and thus we have that  $\Lambda(t)$  is the expectation function of  $\mathcal{P}_t^{\lambda(t)}$ :

$$\Lambda(t) = \mathbb{E}[\mathcal{P}_t^{\lambda(t)}] \quad (12)$$

From Definition 5, it is clear that  $\Lambda(t)$  will be a non-decreasing function and at least right-continuous.

The crucial theorem that underlies the time scale transformation algorithm denotes that the arrival times in a non-homogeneous Poisson process can be mapped to a homogeneous Poisson process with rate 1 (Theorem 1). We present an example that demonstrates this theorem in Figure 3.

**Theorem 1** (Mapping non-homogeneous Poisson processes). Let  $\Lambda$  be a continuous non-decreasing cumulative intensity function. Then,

$$T_1, T_2, \dots \quad (13)$$

are the arrival times in a non-homogeneous Poisson process if and only if

$$\Lambda(T_1), \Lambda(T_2), \dots \quad (14)$$

are the arrival times in a homogeneous Poisson process with rate 1 [10].

The time scale transformation algorithm uses the relation in Theorem 1 to transform a homogeneous Poisson process with  $\lambda = 1$  into a non-homogeneous Poisson process with expectation function  $\Lambda$ . The homogeneous process is formed by sampling from an exponential probability distribution with  $\lambda = 1$ . To make this transformation possible, a time inverse function of  $\Lambda$  is required:

**Definition 6** (Time inverse of  $\Lambda(t)$ ). The time inverse  $\tau$  of an expectation function  $\Lambda(t)$  for a non-homogeneous Poisson process  $\mathcal{P}_t^{\lambda(t)}$ :

$$\tau(s) = \inf\{t : \Lambda(t) > s\} \quad (15)$$

In Algorithm 1, we formalize this procedure. At each step  $i$ , we obtain a sample  $X_i$  from an exponential probability distribution with rate  $\lambda = 1$ , which is added to the set of samples  $\mathcal{X}_i$ . The sum of the elements in  $\mathcal{X}_i$  represents the  $i^{\text{th}}$  arrival in the homogeneous Poisson process, and is transformed into the  $i^{\text{th}}$  arrival in the non-homogeneous Poisson process using the inverse time function  $\tau(\cdot)$ .

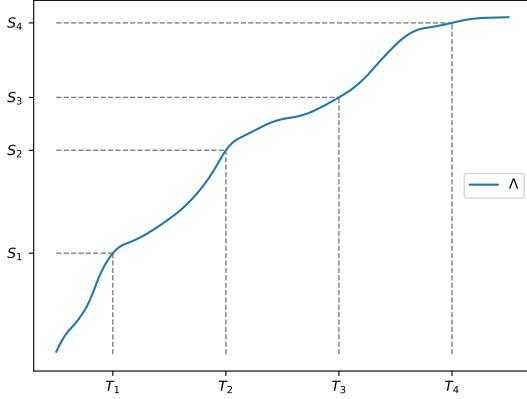


Figure 3: A visual example of Theorem 1:  $T_1, T_2, \dots$  form a non-homogeneous Poisson process with expectation function  $\Lambda$  if and only if  $S_1, S_2, \dots$  form a homogeneous Poisson process with rate 1.

```

 $\mathcal{X}_0 = \emptyset$ 
for  $i = 1, \dots$  do
|    $X_i \sim \text{Exp}(\lambda = 1)$ 
|    $\mathcal{X}_i = \mathcal{X}_{i-1} \cup \{X_i\}$ 
|    $t_i = \tau \left( \sum_{x \in \mathcal{X}_i} x \right)$ 
end

```

**Algorithm 1:** Time scale transformation algorithm

### 3.2 Time scale transformation algorithm to model the infection of patches

In order to use the time scale transformation algorithm (Algorithm 1) in our epidemiological model, note that the patches' internal state is updated in a discrete number of time steps. We determine a patch's intensity  $\phi_p(t)$  (Equation 4) at the end of each day. This results in a sequence of intensities between which we linearly interpolate to obtain a piecewise linear intensity function:

$$\lambda(t) = \text{line}(t, \lceil t \rceil - 1, \lceil t \rceil, \phi_p(t)), \quad (16)$$

where

$$\text{line}(x, x_1, x_2, f) = f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1) \quad (17)$$

interpolates linearly between  $(x_1, f(x_1))$  and  $(x_2, f(x_2))$ .

This piecewise linear intensity function  $\lambda(t)$  is continuous and thus its cumulative counterpart  $\Lambda(t)$  is continuous as well. Furthermore, as  $\phi_p(t) \geq 0$  for all  $t \geq 0$ ,  $\Lambda(t)$  is non-decreasing.

As  $\phi_p(t)$  depends on the simulation state at time  $t$ , it is clear that we cannot evaluate this function beyond the current simulator time step. However, the definition of

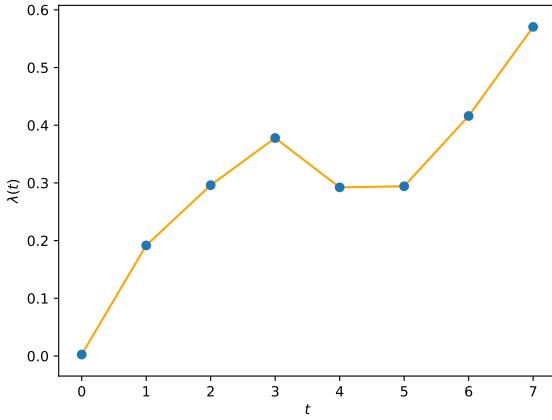


Figure 4: An example of a piecewise linear intensity function for a patch in our model (see Equation 16). The blue scatter points represent the evaluation of  $\phi(t)$  (Equation 4) at discrete time steps (i.e., the end of each day). The orange connecting lines represent the linear interpolation between  $\phi(i - 1)$  and  $\phi(i)$ .

the time inverse  $\tau$  (Definition 6) shows that we can use the arrival time in the homogeneous Poisson process as a threshold for the arrival time in the non-homogeneous Poisson process. We formalize this threshold-based time scale transformation algorithm in Algorithm 2. Note that this algorithm approximates the original algorithm as we check whether the threshold is surpassed at discrete time steps.

```

 $X \sim \text{Exp}(\lambda = 1)$ 
for  $t = 1, \dots$  do
  if  $\Lambda(t) \geq X$  then
    | Trigger event
    |  $X^{(t)} \sim \text{Exp}(\lambda = 1)$ 
    |  $X = X + X^{(t)}$ 
  end
end

```

**Algorithm 2:** Time scale transformation algorithm using discrete time steps

## 4 Model validation

Our objective is to construct a model that is representative for contemporary Great Britain with respect to population census and mobility trends. This model is to be used to study school closure intervention strategies for future influenza pandemics. While in many studies [18, 16, 13], a model is created specifically to fit one epidemic case, we aim for a model that is robust with respect to different epidemic parameters, most importantly  $R_0$ , the basic reproduction number.

To validate our model according to these goals, we conduct two experiments. In the first experiment, we compare our patch model to a SEIR compartment model that

uses the same contact matrix and age structure. While we do not expect our model to behave exactly like the compartment model, as the patches and the mobility network that connects them induces a different dynamic, we do expect to see similar trends with respect to the epidemic curve and peak day. In the second experiment we show that our model is able to reproduce the trends that were observed during the 2009 influenza pandemic, commonly known as the swine-origin influenza pandemic, that originated in Mexico. The 2009 influenza pandemic in Great Britain is an interesting case to validate our model for three main reasons. Firstly, the pandemic occurred quite recently and thus our model’s census and mobility scheme should be a good fit, as both the datasets on which we base our census and mobility model were released in 2011. Secondly, due to the time when the virus entered Great Britain, the summer holiday started 11 weeks after the emergence of the epidemic. The timing of the holidays had a severe impact on the progress of the epidemic and resulted in a epidemic curve with two peaks. This characteristic epidemic curve enables us to demonstrate the predictive power of our age-dependent contact model with support for school closures. Thirdly, the number of symptomatic cases that occurred in Great Britain during the 2009 pandemic was recorded meticulously and is publicly available [20].

#### 4.1 Comparison to the Eames SEIR compartment model

In this experiment, we compare our patch model to a simple SEIR model that encompasses the same age structure and contact matrix [12], to which we will refer as Eames-SEIR from this point onwards. We consider a stochastic implementation of the Eames-SEIR. This experimental setting will be central to the reinforcement learning experiments, related to finding optimal school closure policies, that we present in the main manuscript.

Following [12] and [5], we use a latent period of one day ( $\zeta = \frac{1}{1}$ ) and an infectious period of 1.8 days ( $\gamma = \frac{1}{1.8}$ ). We perform our experiment for a set of  $R_0$  values within the range of 1.4 to 2.4, in steps of 0.2. This range is considered representative for the epidemic potential of influenza pandemics [8, 21].

Furthermore, we need to choose a value for the parameter in the between-patch model, i.e.,  $\mu$ , that modulates the contribution of susceptible adults in the receiving patch (see Section 3). This parameter is typically fitted towards data, however, in this experiment and in the reinforcement learning experiments in the main manuscript, we consider a model to investigate future epidemics. Our goal is to calibrate our model such that it produces peak days that are similar to the peak days in Eames-SEIR [12], which is a prominent model for pandemic influenza that moreover generates peak days that are in agreement with earlier work [14]. Therefore, we investigate the effect of  $\mu$  in this setting, through a sensitivity analysis. We consider  $\mu$  in the interval  $[0, 1]$ , where the left end of the interval (i.e.,  $\mu = 0$ ) signifies that the contribution of susceptible adults is ignored and the right end of the interval (i.e.,  $\mu = 1$ ) signifies that the contribution of adults is not modulated. In Figure 5, we show the results for the sensitivity analysis for  $\mu \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 1\}$ , together with the peak days for the Eames-SEIR model. From these results, it is clear that the different values for  $\mu$  form a gradient within the interval  $[0, 1]$ .

However, no value of  $\mu$  provides a good fit for all of the considered  $R_0$ ’s, when comparing the peak days to the Eames-SEIR model. Rather, we can discern a log-relationship between  $\mu$  and the best fit for the different  $R_0$ ’s. Based on this observation, we propose to define:

$$\mu = \log(R_0) \cdot s, \quad (18)$$

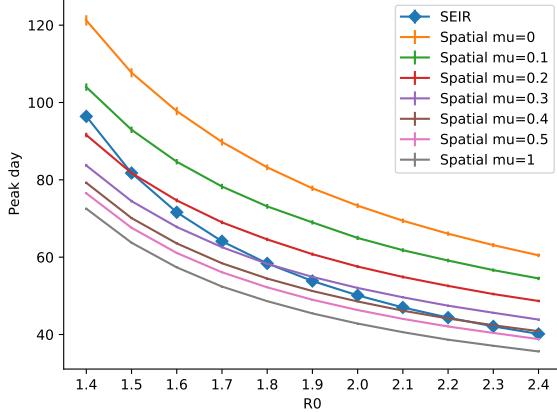


Figure 5: Time of peak day (y-axis) for  $R_0 \in \{1.4, 1.6, 1.8, 2.0, 2.2, 2.4\}$  (x-axis). A curve is shown for different values of  $\mu$  (plain curve) and the peak days as produced by the SEIR-Eames model (curve with diamond scatter points). For each  $R_0$ , 100 stochastic trajectories were sampled and the bound signifies the 95% confidence interval of the sample.

where  $s$  is a scaling factor. For this experimental setting, we find that  $s = .6$  provides a good fit for all of the considered  $R_0$ 's, which we show in Figure 6.

Provided this choice of  $\mu$ , when we compare the epidemic trajectories of our spatial model with the SEIR-Eames model in Figure 7, we observe similar trends with respect to the shape of the trajectory distributions. The main difference is that the epidemic curves grow slower in our spatial model than in the Eames-SEIR model and also achieve a lower peak incidence. This is expected, as we constrain mixing in our spatial model within the districts, and thus increase the resolution of our model, which has been shown to more accurately predict peak incidence [23].

Furthermore, in Figure 8, we show the number of districts that get infected over time for different  $R_0$  values. This shows that all districts get infected, and the time it takes for all districts to reach this point depends mainly on the transmission-ability of the virus strain.

We expect the attack rate to be similar for the SEIR-Eames and spatial model. When all districts get infected, the attack rate in the spatial model is the sum of the attack rates of a set of SEIR-Eames models (i.e., one Eames-SEIR model per district). We verified that the attack rates are indeed nearly identical, as shown in Figure 9, with little variance for either of the models.

## 4.2 2009 influenza pandemic in Great Britain

The virus responsible for the 2009 influenza pandemic arrived in Great Britain during the first week of May 2009 (week 19). Following this introduction, the epidemic grew for 11 weeks until the summer school holidays started, after which the epidemic showed its first peak. After the school holidays, the epidemic was rekindled and grew to a second peak. In Figure 10 we show the weekly case count, as recorded by the British Health Protection Agency (HPA) and the time at which the school holidays

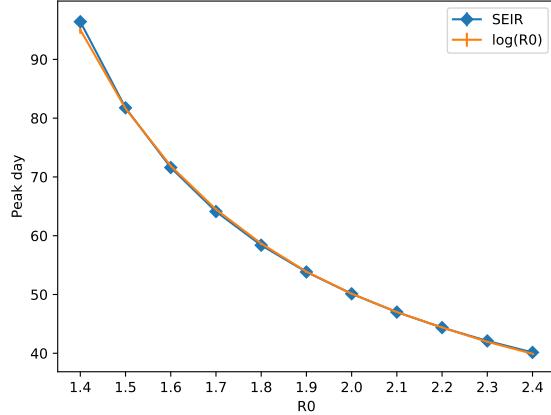


Figure 6: Number of peak days (y-axis) for  $R_0 \in \{1.4, 1.6, 1.8, 2.0, 2.2, 2.4\}$  (x-axis). A curve is shown for  $\mu = \log(R_0) \cdot 0.6$  (orange curve) and the peak days as produced by the SEIR-Eames model (blue curve with diamond scatter points). For each  $R_0$ , 100 stochastic trajectories were sampled and the bound signifies the 95% confidence interval of the sample.

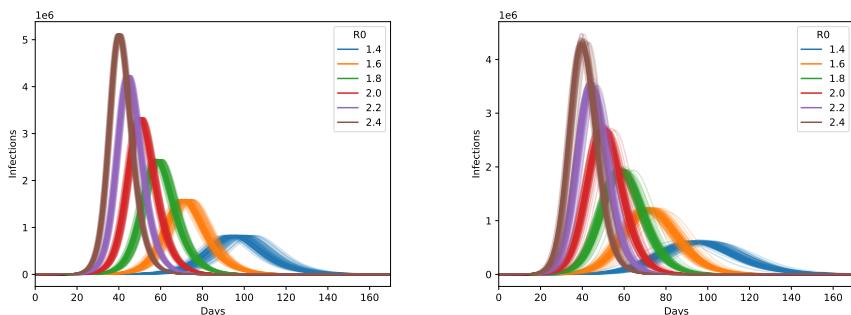


Figure 7: Epidemic trajectories for the Eames-SEIR model (left panel) and the spatial model (right panel). One epidemic trajectory encodes the number of infections per day. Trajectory distributions are shown for  $R_0 \in \{1.4, 1.6, 1.8, 2.0, 2.2, 2.4\}$ , with a different colour per reproductive number. For each  $R_0$ , the distribution consists out of 100 trajectory samples.

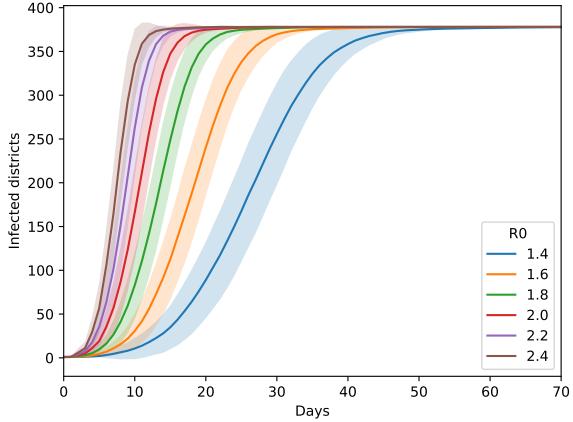


Figure 8: Number of infected districts (y-axis) per day (x-axis) for  $R_0 \in \{1.4, 1.6, 1.8, 2.0, 2.2, 2.4\}$ . For each  $R_0$ , 100 stochastic trajectories were sampled, of which the curve represents the mean, and the bound represents the standard deviation of the samples.

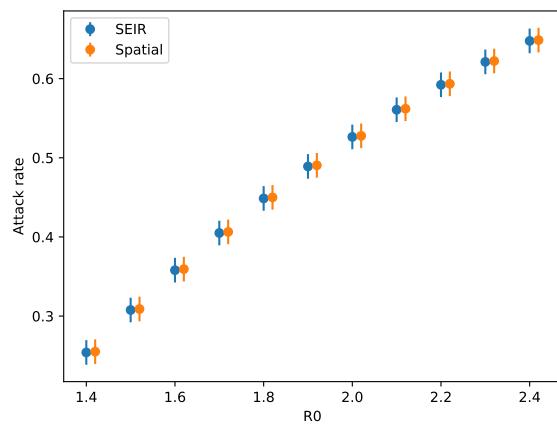


Figure 9: Attack rate (y-axis) for  $R_0 \in \{1.4, 1.6, 1.8, 2.0, 2.2, 2.4\}$  (x-axis). Results are shown for the Eames-SEIR model (blue scatter) and the spatial model (orange scatter). For each model, we depict the standard deviation as bars on top of the scatter points. For each  $R_0$ , 100 stochastic trajectories were obtained.

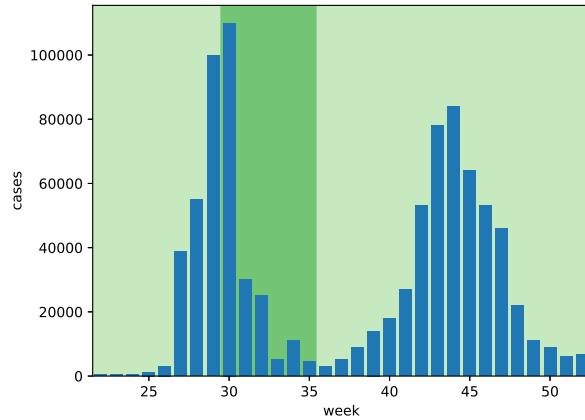


Figure 10: This figure shows the amount of cases that were recorded by the HPA on a weekly basis (blue bars). The background in this figure signifies the time of the summer holidays (dark green).

$R_0$	Source
1.22-1.58	[15]
1.3-1.7	[29]
1.21-1.35	[24]
1.75	[6]
1.87-2.07	[11]
1.31	[27]
1.16-1.59	[20]

Table 1: Overview of basic reproduction numbers from literature.

take place.

To reproduce this distinctive epidemic curve, we use our original model as it was described in the main manuscript. We consider two free parameters: the basic reproduction number and the time of the infectious period. The general consensus is that the basic reproduction number was moderate during the 2009 influenza pandemic, with estimates ranging from 1.16 to 2 [20, 27, 11, 15, 29, 24, 6]. We present a detailed overview of the reported basic reproduction number estimates in Table 1. For the period of infectiousness we found estimates of 1.8, 2.5 and 3.38 days [12, 6, 27]. We present a detailed overview of the infectious period estimates in Table 2.

Given these prior estimates, we parametrize our model with a basic reproduction

Infectious period	Source
1.8	[12]
2.5	[6]
3.38	[27]

Table 2: Overview of infectious periods from literature.

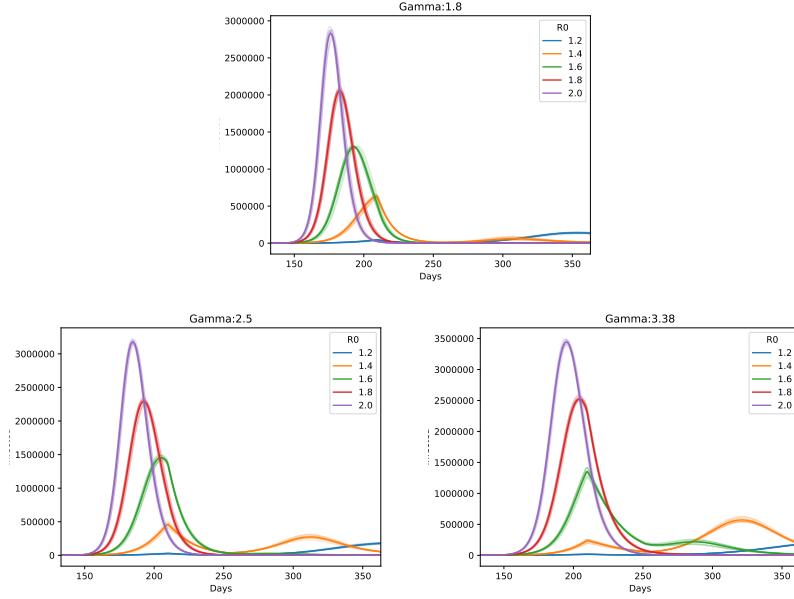


Figure 11: We demonstrate our model for  $R_0 \in \{1.2, 1.4, 1.6, 1.8, 2.0\}$  (enumerated in the legend) and an infectious period of 1.8 days (top panel), 2.5 days (bottom left panel) and 3.38 days (bottom right panel). For each parameter combination, we show a set of stochastic trajectories (light coloured lines) and the mean of these trajectories (dark coloured line). For clarity, we only show 10 stochastic trajectories in this Figure.

number that is in the range of 1.2 to 2.0 and consider a duration of infectiousness of respectively 1.8, 2.5 and 3.38 days.

For this experiment, we found,

$$\mu = \log(R_0) \cdot 2.74, \quad (19)$$

to be a good fit for the overall comparison. In Figure 11 we show the epidemic curve for our model with respect to these parameters. In general, the epidemic curves that result from using an infectious period of 1.8 days are insufficient to reproduce the trends of the 2009 pandemic. For the other infectious periods (i.e., 2.5 and 3.38), we show that for all but the highest reproductive numbers we observe an epidemic curve with 2 peaks. Furthermore, we observe a deeper trough in the epidemic curve when an infectious period of 2.5 days is chosen.

In Figure 12, we show a set of model realisations in conjunction with the symptomatic case data, which shows that we were able to closely match the epidemic trends observed during the British pandemic in 2009. This model was configured with a basic reproductive number of 1.4 and infectious period of 2.6. The reproductive number is in good concordance with the general consensus that the virus responsible for the 2009 pandemic exhibited a moderate infectiousness. While the infectious period slightly differs from the value reported by [6] (i.e., 2.5 days), it lies well within the confidence bounds reported in this study (confidence interval: 1.1-4.0 days). Note that our model reports the number of infections, while the HPA only recorded symptomatic cases. Therefore we scale the epidemic curve with a factor of  $\frac{1}{4}$ . While atypical, this large

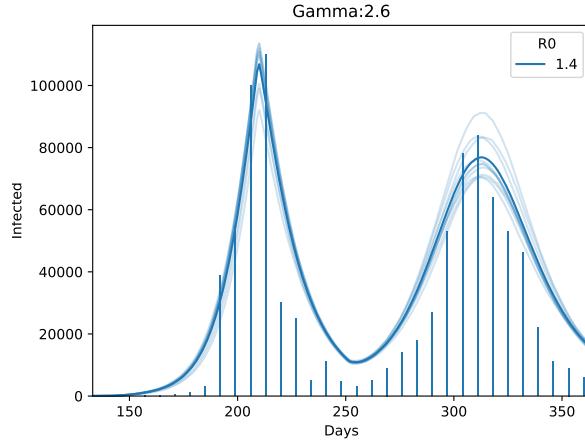


Figure 12: We show that our model, using a reproductive number of 1.4 and an average duration of infectiousness of 2.6 days is able to match the trends observed in the British pandemic of 2009. For clarity, we only show 10 stochastic trajectories in this Figure.

number of asymptomatic cases produced by our model is in line with earlier serological surveys [22] and with previous modelling studies [20].

## 5 Selecting districts to establish a ground truth

To establish a ground truth, we select 10 districts that are representative of the population heterogeneity in Great Britain. To this end, we remind the reader that in Section 2, we analysed the population heterogeneity by representing the population structure as a positive simplex. We select 10 districts: one district that is representative for the average of this distribution and a set of nine districts that is representative for the diversity in this distribution. To determine the average district, we consider the population heterogeneity distribution over all districts, and determine the Aitchison's mean (Definition 7) of this distribution [3]. We then select the district that is closest to the Aitchison's mean (Definition 7) according to the Aitchison distance (Definition 8), as shown in Figure 13.

**Definition 7** (Aitchison's mean). *Given a set of points from a unit simplex (Definition 1),*

$$P = \{p^{(i)} \mid p^{(i)} \in \mathbb{S}^D\}_{i=1}^N, \quad (20)$$

*the Aitchison's mean [4] is:*

$$C_A(P) = \frac{\langle h_1, \dots, h_D \rangle}{\sum_{d=1}^D h_d}, \quad (21)$$

*where,*

$$h_d = \left( \prod_{p^{(i)} \in P} p_d^{(i)} \right)^{(1/N)}, \quad (22)$$

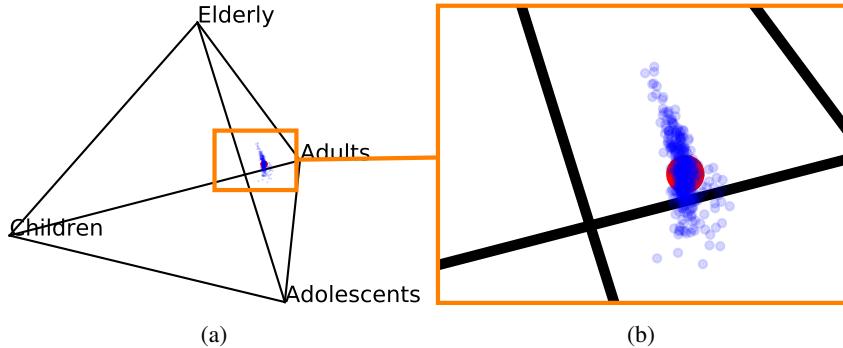


Figure 13: Barycentric projection of the census proportions in the districts of Great Britain (blue scatter points), according to Eames’ age structure. The geometric mean of this distribution is shown as a red scatter point. The left panel shows the original census pyramid, and the right panel zooms in on the point cloud.

is the geometric mean of the  $d$ -th component over all simplex points in  $P$ .

**Definition 8** (Aitchison distance). *Given two points from a unit simplex  $p, q \in \mathbb{S}^D$  (Definition 1), we define the Aitchison distance function [2]:*

$$d_A(p, q) = \left[ \sum_{d=1}^D \left( \log \frac{p_d}{h(p)} - \log \frac{q_d}{h(q)} \right) \right]^{1/2}, \quad (23)$$

where,

$$h(p) = \left( \prod_{d=1}^D p_d \right)^{(1/D)}, \quad (24)$$

denotes the geometric mean of  $p$ . This distance defines a metric on the simplex sample space.

Next, we determine the outer extreme points, as these represent the most diverse census points. To do this, we compute the convex hull of the point cloud (i.e., the smallest convex set of points that contains the point cloud), as shown in Figure 14.

We proceed by taking the points that belong to the surface of the convex hull, of which we make a sub-selection of 9 census points. As the convex hull consists out of 21 points, we consider all  $k$ -combinations (with  $k = 9$ ) and select the set of points that maximizes the minimum Aitchison distance between the selected points, as shown in Figure 15.

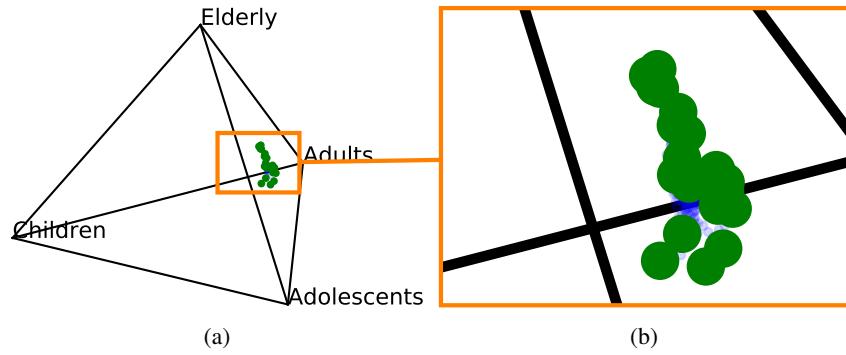


Figure 14: Barycentric projection of the census proportions in the districts of Great Britain (blue scatter points), according to Eames' age structure. The census points that are part of the convex hull are shown in green. The left panel shows the original census pyramid, and the right panel zooms in on the point cloud.

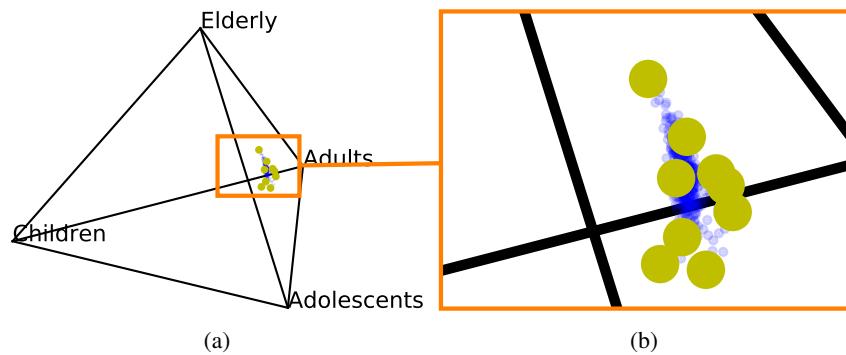


Figure 15: Barycentric projection of the census proportions in the districts of Great Britain (blue scatter points), according to Eames' age structure. The census point that were selected out of the convex hull are shown in yellow. The left panel shows the original census pyramid, and the right panel zooms in on the point cloud.

## 6 PPO learning curves ( $R_0 = 1.8$ )

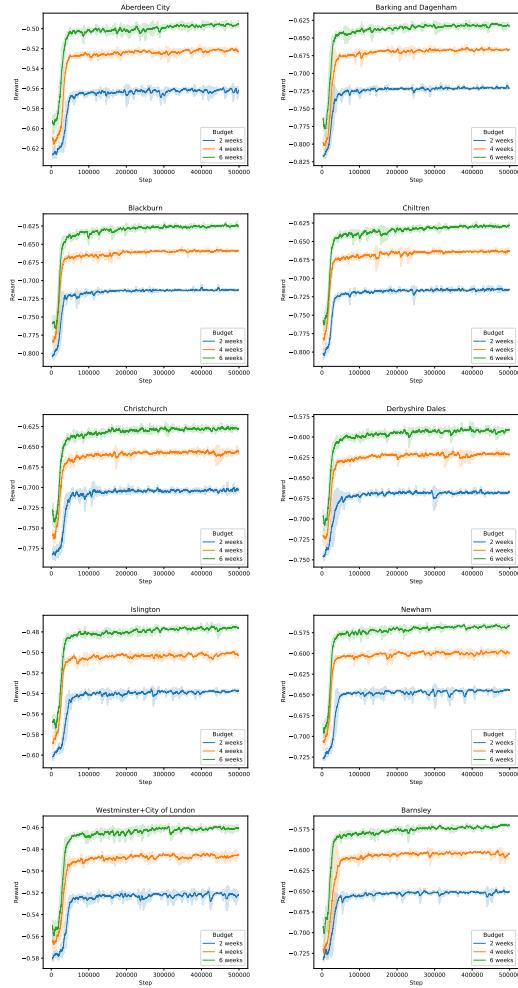


Figure 16: PPO learning curves for  $R_0 = 1.8$ .

## 7 PPO learning curves ( $R_0 = 2.4$ )

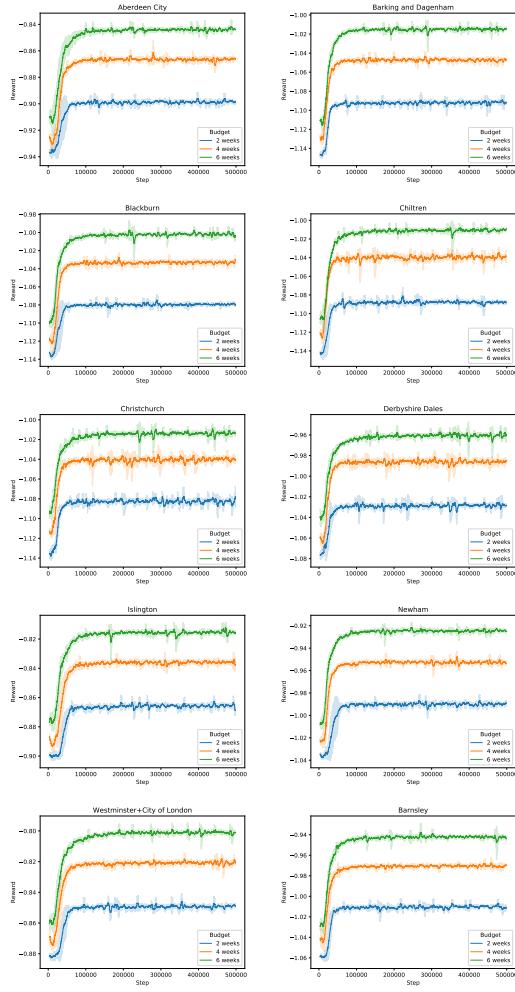


Figure 17: PPO learning curves for  $R_0 = 2.4$ .

## 8 QMIX reward curves

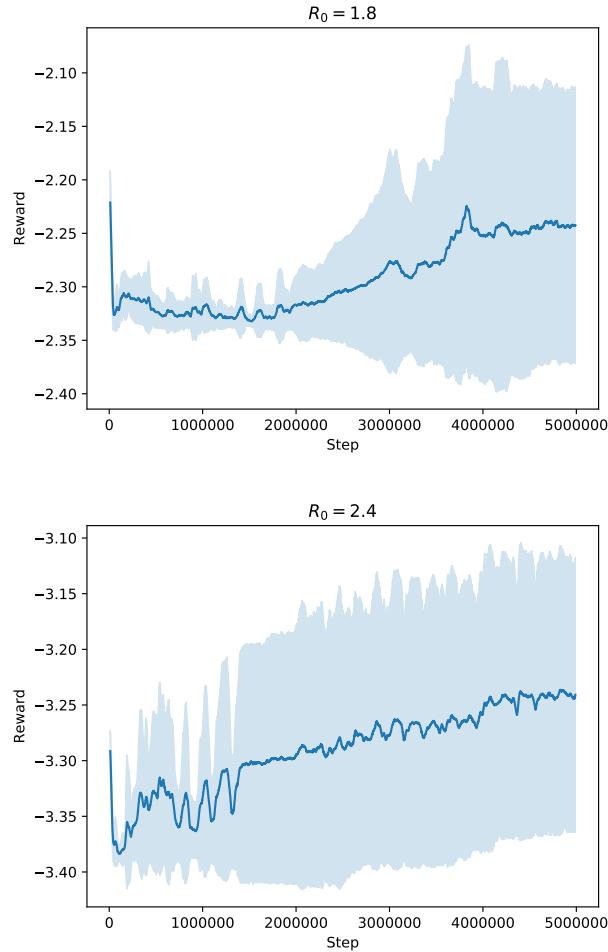


Figure 18: We show the reward curves for multi-district QMIX for  $R_0 = 1.8$  (top panel) and  $R_0 = 2.4$  (bottom panel). The shaded area shows the standard deviation of the reward signal, over 5 multi-district QMIX runs.

## 9 PPO hyper-parameters

- Number of local steps: 1024
- Batch size: 128 (i.e., 8 minibatches)
- Clipped Surrogate Objective epsilon  $\epsilon$ : 0.2
- Number of epochs: 24
- Entropy coefficient: 0.0059
- $\gamma$ : 0.99
- Generalized advantage estimation  $\lambda$ : 0.95
- Neural network of actor and critic:
  - Number of hidden layers: 1
  - Number of units per hidden layer: 20
  - Non-linearity: hyperbolic tangent (tanh)
  - learning rate  $\alpha$ : 0.002
  - Optimizer: Adam [17]
  - Gradient norm clipping threshold: 1.0

## References

- [1] J. Aitchison. Principal component analysis of compositional data. *Biometrika*, 70(1):57–65, 1983.
- [2] J. Aitchison. On criteria for measures of compositional difference. *Mathematical Geology*, 24(4):365–379, 1992.
- [3] J. Aitchison. Principles of compositional data analysis. *Lecture Notes-Monograph Series*, pages 73–81, 1994.
- [4] J. Aitchison and V. Pawlowsky-Glahn. The one-hour course in compositional data analysis or compositional data analysis is simple. In *Proceedings of IAMG*, volume 97, pages 3–35, 1997.
- [5] M. Baguelin, A. J. Van Hoek, M. Jit, S. Flasche, P. J. White, and W. J. Edmunds. Vaccination against pandemic influenza a/h1n1v in england: a real-time economic evaluation. *Vaccine*, 28(12):2370–2384, 2010.
- [6] D. Balcan, H. Hu, B. Goncalves, P. Bajardi, C. Poletto, J. J. Ramasco, D. Paolotti, N. Perra, M. Tizzoni, W. Van den Broeck, et al. Seasonal transmission potential and activity peaks of the new influenza a (h1n1): a monte carlo likelihood analysis based on human mobility. *BMC medicine*, 7(1):45, 2009.
- [7] M. Barthélémy, C. Godreche, and J.-M. Luck. Fluctuation effects in metapopulation models: percolation and pandemic threshold. *Journal of theoretical biology*, 267(4):554–564, 2010.
- [8] N. E. Basta, D. L. Chao, M. E. Halloran, L. Matrajt, and I. M. Longini. Strategies for pandemic and seasonal influenza vaccination of schoolchildren in the United States. *American journal of epidemiology*, 170(6):679–686, 2009.
- [9] D. P. Bertsekas and J. N. Tsitsiklis. *Introduction to probability*, volume 1. Athena Scientific Belmont, MA, 2002.
- [10] E. Cinlar. *Introduction to stochastic processes*. Courier Corporation, 2013.
- [11] U. De Silva, J. Warachit, S. Waicharoen, and M. Chittaganpitch. A preliminary analysis of the epidemiology of influenza a (h1n1) v virus infection in thailand from early outbreak data, june-july 2009. *Eurosurveillance*, 14(31):19292, 2009.
- [12] K. T. Eames, N. L. Tilston, E. Brooks-Pollock, and W. J. Edmunds. Measured dynamic social contact patterns explain the spread of h1n1v influenza. *PLoS computational biology*, 8(3):e1002425, 2012.
- [13] R. M. Eggo, S. Cauchemez, and N. M. Ferguson. Spatial dynamics of the 1918 influenza pandemic in england, wales and the united states. *Journal of the Royal Society Interface*, 8(55):233–243, 2010.
- [14] N. M. Ferguson, D. A. Cummings, C. Fraser, J. C. Cajka, P. C. Cooley, and D. S. Burke. Strategies for mitigating an influenza pandemic. *Nature*, 442(7101):448, 2006.

- [15] C. Fraser, C. A. Donnelly, S. Cauchemez, W. P. Hanage, M. D. Van Kerkhove, T. D. Hollingsworth, J. Griffin, R. F. Baggaley, H. E. Jenkins, E. J. Lyons, et al. Pandemic potential of a strain of influenza a (h1n1): early findings. *science*, 324(5934):1557–1561, 2009.
- [16] J. R. Gog, S. Ballesteros, C. Viboud, L. Simonsen, O. N. Bjørnstad, J. Shaman, D. L. Chao, F. Khan, and B. T. Grenfell. Spatial transmission of 2009 pandemic influenza in the us. *PLoS computational biology*, 10(6):e1003635, 2014.
- [17] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [18] S. M. Kissler, J. R. Gog, C. Viboud, V. Charu, O. N. Bjørnstad, L. Simonsen, and B. T. Grenfell. Geographic transmission hubs of the 2009 influenza pandemic in the united states. *Epidemics*, 26:86–94, 2019.
- [19] P. Klepac, S. Kissler, and J. Gog. Contagion! the bbc four pandemic—the model behind the documentary. *Epidemics*, 2018.
- [20] R. J. Kubiak and A. R. McLean. Why was the 2009 influenza pandemic in england so small? *PloS one*, 7(2):e30223, 2012.
- [21] J. Medlock and A. P. Galvani. Optimizing influenza vaccine distribution. *Science*, 325(5948):1705–1708, 2009.
- [22] E. Miller, K. Hoschler, P. Hardelid, E. Stanford, N. Andrews, and M. Zambon. Incidence of 2009 pandemic influenza a h1n1 infection in england: a cross-sectional serological study. *The Lancet*, 375(9720):1100–1108, 2010.
- [23] H. L. Mills and S. Riley. The spatial resolution of epidemic peaks. *PLoS computational biology*, 10(4):e1003561, 2014.
- [24] H. Nishiura, G. Chowell, M. Safan, and C. Castillo-Chavez. Pros and cons of estimating the reproduction number from early epidemic growth rate of influenza a (h1n1) 2009. *Theoretical Biology and Medical Modelling*, 7(1):1, 2010.
- [25] S. Osaki. *Applied stochastic system modeling*. Springer Science & Business Media, 2012.
- [26] G. S. Tomba and J. Wallinga. A simple explanation for the low impact of border control as a countermeasure to the spread of an infectious disease. *Mathematical biosciences*, 214(1-2):70–72, 2008.
- [27] A. R. Tuite, A. L. Greer, M. Whelan, A.-L. Winter, B. Lee, P. Yan, J. Wu, S. Moghadas, D. Buckeridge, B. Pourbohloul, et al. Estimated epidemiologic parameters and morbidity associated with pandemic h1n1 influenza. *Can Med Assoc J*, 182(2):131–136, 2010.
- [28] L. Wang and J. T. Wu. Characterizing the dynamics underlying global spread of epidemics. *Nature communications*, 9(1):218, 2018.
- [29] Y. Yang, J. D. Sugimoto, M. E. Halloran, N. E. Basta, D. L. Chao, L. Matrajt, G. Potter, E. Kenah, and I. M. Longini. The transmissibility and control of pandemic influenza a (h1n1) virus. *Science*, 326(5953):729–733, 2009.