

# UK Age and Space structured Covid-19 model

Chris Jewell, Barry Rowlingson, Jon Read

April 8, 2020

## 1 Concept

We wish to develop a model that will enable us to assess spatial spread of Covid-19 across the UK, respecting the availability of human mobility data as well as known contact behaviour between individuals of different ages.

A deterministic SEIR state transition model is posited in which individuals transition from Susceptible to Exposed (i.e. infected but not yet infectious) to Infectious to Removed (i.e. quarantined, got better, or died).

We model the infection rate (rate of  $S \rightarrow E$  transition) as a function of known age-structured contact from Polymod, known human mobility between MSOAs (Middle Super Output Area), and Census-derived age structured population density in regions across the UK.

Currently, this model is populated with data for England only, though we are in the process of extending this to Scotland, Wales, and Northern Ireland.

## 2 Data

### 2.1 Age-mixing

Standard Polymod social mixing data for the UK are used, with 17 5-year age groups  $[0-5), [5-10), \dots, [75-80), [80-\infty)$ . Estimated contact matrices for term-time  $M_{tt}$  and school-holidays  $M_{hh}$  were extracted of dimension  $n_m \times n_m$  where  $n_m = 17$ .

### 2.2 Human mobility

2011 Census data from ONS on daily mean numbers of commuters moving from each Residential MSOA to Workplace MSOA. MSOAs are aggregated to Local Authority Districts (LADs) for which we have age-structured population density. The resulting matrix  $C$  is of dimension  $n_c \times n_c$  where  $n_c = 152$ . Since this matrix is for Residence to Workplace movement only, we assume that the mean number of journeys between each LAD is given by

$$T = C + C^T$$

with 0 diagonal.

### 2.3 Population size

Age-structured population size within each LAD is taken from publicly available 2019 Local Authority data giving a vector  $N$  of length  $n_m n_c = 2584$ , i.e. population for each of  $n_m$  age groups and  $n_c$  LADs.

### 3 Model

#### 3.1 Connectivity matrix

We assemble a country-wide connectivity matrices as Kronecker products, such that

$$M^* = I_{n_c} \otimes M$$

and

$$C^* = C \otimes \mathbf{1}_{n_m \times n_c}$$

giving two matrices of dimension  $n_m n_c \times n_m n_c$ .  $M^*$  is block diagonal with Polymod mixing matrices.  $C^*$  expands the mobility matrix  $C$  such that a block structure of connectivity between LADs results.

#### 3.2 Disease progression model

We assume an SEIR model described as a system of ODEs. We denote the number of individual in each age-group-LAD combination at time  $t$  by the vectors  $\vec{S}(t), \vec{E}(t), \vec{I}(t), \vec{R}(t)$ . We therefore have

$$\begin{aligned} \frac{d\vec{S}(t)}{dt} &= -\epsilon \frac{\vec{S}(t)}{N} - \beta_t \left[ M^* \vec{I}(t) + \beta_2 w_t \bar{M} C^* \frac{\vec{I}(t)}{N} \right] \frac{\vec{S}(t)}{N} \\ \frac{d\vec{E}(t)}{dt} &= \epsilon \frac{\vec{S}(t)}{N} + \beta_t \left[ M^* \vec{I}(t) + \beta_2 w_t \bar{M} C^* \frac{\vec{I}(t)}{N} \right] \frac{\vec{S}(t)}{N} - \nu \vec{E}(t) \\ \frac{d\vec{I}(t)}{dt} &= \nu \vec{E}(t) - \gamma \vec{I}(t) \\ \frac{d\vec{R}(t)}{dt} &= \gamma \vec{I}(t) \end{aligned}$$

where  $\bar{M}$  is the global mean person-person contact rate, and  $w_t$  is the total rail ticket sales in the UK expressed as a fraction of the 2019 mean (a proxy for reduction in travel). Parameters are:

$$\epsilon_t = \begin{cases} \epsilon_0 & \text{if } t < T_b \\ 0 & \text{otherwise} \end{cases}$$

with  $T_b$  a time beyond which imports to the UK are considered to cease;

$$\beta_t = \begin{cases} \beta_1 & \text{if } t < T_L \\ \beta_1 \beta_3 & \text{otherwise} \end{cases}$$

with  $T_L$  the date of lock-down restrictions,  $\beta_1$  a baseline transmission rate, and  $\beta_3$  giving the ratio of post-lockdown to pre-lockdown transmission; commuting infection ratio  $\beta_2$ ; latent period  $\frac{1}{\nu} = 4$  days; and infectious period  $\frac{1}{\gamma}$ . Typically, we assume that contact with commuters is  $\beta_2 = \frac{1}{3}$  of that between members of the same age-LAD combination assuming an 8 hour working day.

#### 3.3 Noise model

Currently, and subject to discussion, we assume that all detected cases are synonymous with individuals transitioning  $I \rightarrow R$ . We assume the number of new cases in each age-LAD combination are given by

$$y_{ik}(t) \sim \text{Negative Binomial}(r, \phi(R_{ik}(t) - R_{ik}(t-1)))$$

where  $\phi$  is the case reporting fraction (i.e. proportion of infections that are eventually detected) and  $r$  is an overdispersion parameter.

### 3.4 Inference

We are interested in making inference on  $\epsilon_0, \beta_1$ , and  $\gamma$ . Priors are currently:

$$\begin{aligned}\epsilon_0 &\sim \text{Gamma}(1, 1) \\ \beta_1 &\sim \text{Gamma}(1, 1) \\ \beta_3 &\sim \text{Gamma}(20, 20) \\ \gamma &\sim \text{Gamma}(10, 40) \\ r &\sim \text{Gamma}(0.1, 0.1)\end{aligned}$$

specified to express relative ignorance about  $\epsilon_0$ ,  $\beta_1$ , and  $r$ , strong information on  $\beta_3$  and  $r$ , and a belief that the infectious period should be approximately 4 days.

### 3.5 Implementation

The model is currently implemented in Python3, using Tensorflow 2.1 with the RK5 differential equation solver implemented in the `DormandPrince` class provided by Tensorflow Probability 0.9. Code may be found at <http://fhm-chicas-code.lancs.ac.uk/jewell/covid19uk>.

The code is a work in progress – see README.

## 4 Results