

# Contents

<b>1 SIR and SIR-D models</b>	<b>1</b>
1.1 SIR model . . . . .	1
1.2 SIR-D model . . . . .	2
<b>2 Markov Chain Monte Carlo approach</b>	<b>4</b>
2.1 MCMC . . . . .	4
2.2 MCMC steps . . . . .	6
2.3 Parameters estimation . . . . .	7
2.4 Likelihood . . . . .	9

## Abstract

Owing to the pandemic scenario of COVID-19 disease cases all over the world, the outbreak prediction has become extremely complex for emerging scientific research. Several epidemiological mathematical models of spread are increasing daily to forecast the predictions appropriately. In this study, the classical susceptible-infected-recovered (SIR) modeling approach was employed to study the different parameters of this model for Germany. The purpose is to have insights into political measures on the pandemic spread. In this respect, we have calculated parameters of SIR model in time, for genders and different age classes and parameters changes are studied alongside political measures like the number of testing, social distancing, imposed lockdowns, etc. The Monte Carlo method is applied to estimate parameters which beside better estimation helps us to have an insight of parameter reliability.

## 1 SIR and SIR-D models

In this section, we will create a customized mathematical model derived from SIR model.

### 1.1 SIR model

To understand the trend of infection, we will use mathematical epidemic model. Let's start discussion using the simplest model named SIR.

What is SIR model?

In 1927, Kermack and McKendrick [1] first proposed a class of compartmental models that simplified the mathematical modeling of infectious disease transmission. The SIR model is a set of general equations that explain the dynamics of an infectious disease spreading through a susceptible population.

Essentially, the standard SIR model is a set of differential equations that can be categorized as susceptible (if previously unexposed to the pandemic disease), infected (if currently colonized by the pandemic disease), and removed (either by death or recovery) as follows:

SIR model is a simple mathematical model to understand outbreak of infectious diseases.

(Susceptible  $\rightarrow$  Infectible  $\rightarrow$  Recovery/Removed) SIR model.

- S : Susceptible (= Population - Confirmed)
- I : Infected (=Confirmed - Recovered - Fatal)
- R : Recovered or Fatal (= Recovered + Fatal)

Model:

$$S \xrightarrow{\beta} I \xrightarrow{\gamma} R \quad (1)$$

$\beta$  : Effective contact rate [1/min]

$\gamma$  : Recovery(+Mortality) rate [1/min]

Ordinary Differential Equation (ODE):

$$\frac{dS}{dt} = -\frac{\beta}{N} SI \quad (2)$$

$$\frac{dI}{dt} = \frac{\beta}{N} SI - \gamma I \quad (3)$$

$$\frac{dR}{dt} = \gamma I \quad (4)$$

Where  $N = S + I + R$  is the total population,  $t$  is the elapsed time from the start date.

The basic reproduction number,  $R_0$ , is the ratio of transmission and recovery rates.

$$R_0 = \frac{\beta}{\gamma}. \quad (5)$$

## 1.2 SIR-D model

Because we are measuring the number of fatal cases and recovered cases separately, we can use two variables ("Recovered" and "Deaths") instead of "Recovered + Deaths" in the mathematical model.

What is SIR-D model?

- S : Susceptible (= Population - Confirmed)
- I : Infected (=Confirmed - Recovered - Fatal)
- R : Recovered
- D : Fatal

Model:

$$\begin{array}{c} S \xrightarrow{\beta} I \xrightarrow{\gamma} R \\ I \xrightarrow{\alpha} D \end{array} \quad (6)$$

$\alpha$  : Mortality rate [1/min]

$\beta$  : Effective contact rate [1/min]

$\gamma$  : Recovery rate [1/min]

Ordinary Differential Equation (ODE):

$$\frac{dS}{dt} = -\frac{\beta}{N} SI \quad (7)$$

$$\frac{dI}{dt} = \frac{\beta}{N} SI - (\gamma + \alpha)I \quad (8)$$

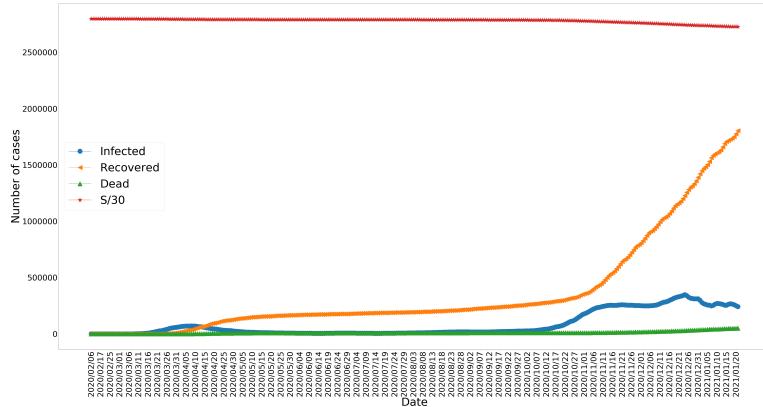
$$\frac{dR}{dt} = \gamma I \quad (9)$$

$$\frac{dD}{dt} = \alpha I \quad (10)$$

Where  $N = S + I + R + D$  is the total population,  $t$  is the elapsed time from the start date.

The basic reproduction number,  $R_0$ , is the ratio of transmission and recovery plus fatal rates.

$$R_0 = \frac{\beta}{\gamma + \alpha}. \quad (11)$$



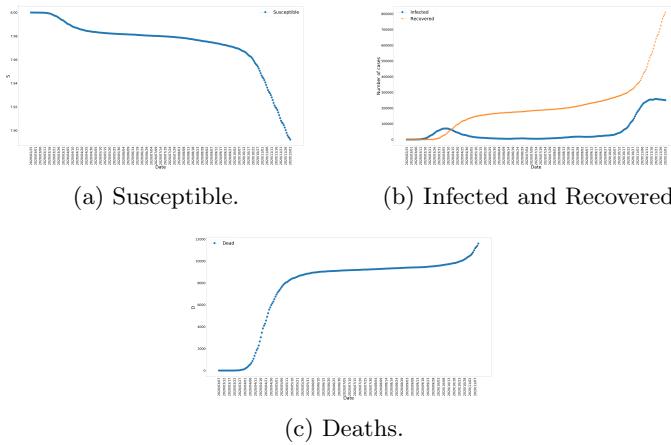
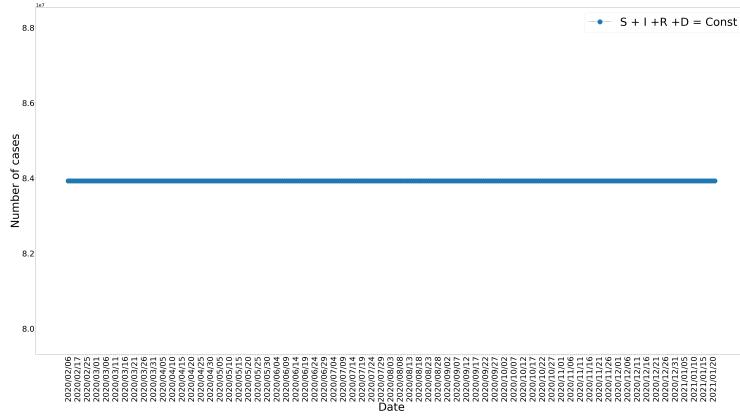


Figure 1: S, I, R, and D.

## 2 Markov Chain Monte Carlo approach

### 2.1 MCMC

Markov Chain Monte Carlo (MCMC) approaches provide a solution to this problem, offering a straightforward procedure for generating random draws from very complex distributions. Simulating random vectors drawn from some complex target distribution can be a very difficult task. The idea behind MCMC approaches is to successively draw samples from a far simpler distributions in such a way that distribution of the samples converge to the target distribution. MCMC approaches are so-named because one uses the previous sample value to randomly generate the next sample value, generating a Markov chain. While

there are wide range of MCMC methods, here we will apply Metropolis-Hastings algorithm [2]. In the case of Metropolis-Hastings method, one simulates draw from a complex target distribution by first drawing a random variable from a specified distribution and then using a second probability distribution to decide whether to accept or reject the realization.

In the following a model means a new parameter, so  $m$  can be one of our parameters.  $m = (\beta, \gamma, \alpha)$ . As discussed the first step in Metropolis-Hastings method is to draw a new model  $m'$ , as a random variate from a probability density (proposal or jump probability distribution)  $Q(m'|m)$  which is conditional only on the current model  $m$ . A simple example for proposal density function could be Gaussian distribution with zero mean and diagonal covariance matrix  $C = \text{diag}\{\sigma_{m_1}^2, \sigma_{m_2}^2, \dots\}$ ,

$$Q(m'|m) \sim \exp\left(-\frac{1}{2}(m - m')^T C^{-1}(m - m')\right).$$

One can in practice generate a new model  $m'$  from the current model  $m$  by perturbing the  $i$ th component of a random variable  $u$  from a normal distribution  $N(0, 1)$  and set

$$m' = m + u\sigma_{m_i}$$

where  $\sigma_{m_i}$  the variance of the proposal. This type of proposal density function is common in Metropolis-hastings methods. The key character of this proposal distribution is its symmetry, namely the probability of generating a perturbed model  $m'$  from starting model  $m$  is the same as generating the reverse proposal distribution, ie probability of starting from  $m'$  and generating a sample at  $m$ . The second step will be acceptance or rejection of new model. Once the new model is generated by using the proposal distribution  $Q(m'|m)$ , the acceptance rate will be calculated  $\alpha(m'|m)$ , then a uniform random variable  $U$  between 0 and 1 will be generated. If  $U$  the new model is accepted and the current model  $m$  will be replaced with the new model  $m'$  and the chain jump to next step. If  $U > \alpha$  the new model is rejected and Markov chain will stay in current step and the next of process will be repeated. The acceptance rate is a key issue to ensure that the new model will be generated according to target density (here the posterior distribution). Gilks et al. [1996] showed that the chain will converge to posterior distribution  $p(m|d)$ , if

$$\alpha(m'|m) = \min\left(1, \frac{p(m'|d)}{p(m|d)} \frac{Q(m|m')}{Q(m'|m)}\right).$$

A key issue in successful implement of Metropolis-Hastings method or any other MCMC method is the number of runs until the chain converges the stationary process (the length of burn-in period). Typically the first 1000 to 5000 values of the chain are thrown out, and then various convergence tests are used to asses to see the chain is converged or not. A poor choice of starting values or proposal distribution can greatly increase the required burn-in time. One suggestion for starting point could be center of distribution, for example taking a value close to the distribution mode. A chain is said to be poor mixed if it

stays in small area of parameter space for long periods of time. A poorly mixing chain can arise because the target distribution is multimodal and our starting point can trap us near one of the nodes (such multimodal posterior distribution can arise if we have a strong prior in conflict with the observed data). One straightforward and common approach is to use multiple highly dispersed starting values to start several different chains [Gelman and Rubin, 1992]. In our case we can independently estimate parameters ( $\beta$ ,  $\gamma$  and  $\alpha$ ) from equation 6, 8, 9 respectively. Starting implementation of our algorithm with initial values for parameters the algorithm proceeds iteratively. The updating algorithm will be as follow:

- At every iteration we update the parameter of interest according to a Gaussian distribution, for example  $Q(m'|m)$  centered at current value of parameter  $m$ . The variance of Gaussian distribution is a tuning parameter which should be chosen.
- For updated parameters (new model), the forward model will be solved, namely the corresponding value of parameter (in the case of  $\beta$  this value is susceptibility  $S$ ) for the next time step. Then the likelihood and posterior of new model  $p(m'|d)$  are calculated.
- Now we should decide to accept or reject the new model, hence the acceptance rate  $\alpha(m'|m)$  should be calculated by Eq. (2.16). Since the proposal distribution is symmetric as discussed already  $Q(m'|m) = Q(m|m')$  so we have

$$\alpha(m'|m) = \min\left(1, \frac{p(m'|d)}{p(m|d)}\right).$$

If the posterior has a higher posterior value, it is always accepted and becomes the current model.

After we collected enough samples, we can get the mean value of desired parameter  $m$ , with following equation:

$$E(m) = \frac{1}{L} \int_{l=1}^L m_l$$

where  $L$  is the number of models collected in post burn-in period. Actually we estimate expected value of parameters, namely mean value of samples.

## 2.2 MCMC steps

$$P(\beta_i, \gamma_i, \alpha_i | \text{data}) = \frac{P(\text{data} | \beta_i, \gamma_i, \alpha_i) P(\beta_i, \gamma_i, \alpha_i)}{\sum_{j=1}^N P(\text{data} | \beta_j, \gamma_j, \alpha_j) P(\beta_j, \gamma_j, \alpha_j)}$$

Starting point

- Parameter space (covering all possible parameter values for all parameters in model)

- For each possible parameter value we can compute the likelihood =  $P(\text{data}|\text{parameter values})$
- For each parameter value we know the prior probability =  $P(\text{parameter values})$
- We can therefore compute prior  $\times$  likelihood for any given point in parameter space

Steps

- Start in random position on probability landscape ( $X$ ). Compute prior  $\times$  likelihood here. Let's call that  $p_x$ .
- Based on current position: attempt move to new position ( $Y$ ) by randomly drawing from "proposal distribution":  $q(Y|X)$
- (For example, the proposal distribution can be a normal distribution with mean  $X$  and standard deviation 1)
- Compute prior  $\times$  likelihood at new position. We'll call that  $p_Y$ .
  - (a) If move ends higher up, i.e.  $p_Y > p_X$ : accept move
  - (b) If move ends below: accept move with probability

$$p_{\text{accept}} = \frac{p_Y \times q(X|Y)}{p_X \times q(Y|X)}$$

- If  $q(X|Y) = q(Y|X)$ , i.e.,  $q$  is symmetric, this becomes:

$$p_{\text{accept}} = \frac{p_Y}{p_X}$$

- write parameter values for accepted moves in file (if proposed move is not accepted: write previous values again)
- After many, many repetitions points will be sampled in proportion to the height of the probability landscape: we therefore have an empirical approximation of the distribution.

### 2.3 Parameters estimation

In order to get time dependent parameters we needed to calculate them in a sliding window. So, we considered a window with length of 30 days (namely 30 points) and the for each window we estimate our parameters.

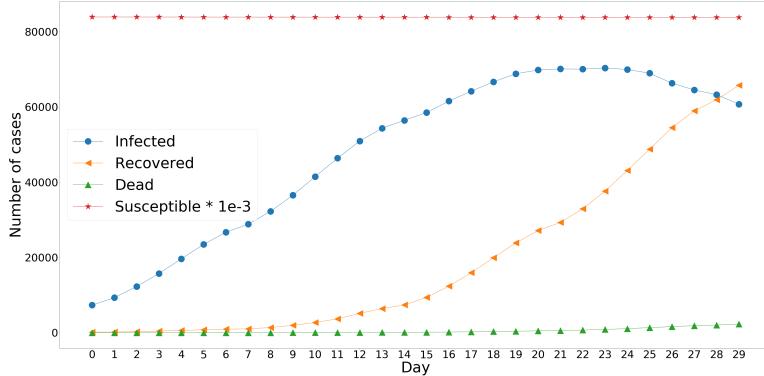
From equation 6 we estimate  $\beta$ , from equation 8 parameter  $\gamma$  and from 9,  $\alpha$  is calculated.

Now considering only a window, we have 30 values for  $S$  and  $I$ . As we can see in equation 21 we can get  $S(t+1)$  from  $S(t)$  and  $I(t)$ . Now, for a given  $\beta$

and  $S(t)$ ,  $I(t)$  we can calculate  $S(t+1)$ . So this is the basis of our method. It means we do a random walk for  $\beta$  and using current susceptibility and infection (namely  $S(t)$  and  $I(t)$ ) we predict a value for the next day, lets call it  $S_{pre}(t+1)$ .

Doing this for all points we can have an array of predicted  $S_{pre}^i(t+1)$  for  $i = 1..29$ . Therefor for each sliding window for any given  $\beta$  we can calculate  $S_{pre}(t+1)$ , and at the same time we have the true value for this time point  $S(t+1)$ . This is actually how sampling works. We do a random walk for  $\beta$  and accept or reject each new point with metropolis Hastings algorithm which is already explained.

As it can be seen in figure below we have plotted  $S$ ,  $I$ ,  $R$  and  $D$  for one of mentioned sliding windows.



Actually the data in this window (mentioned figure) is used to estimate all parameters. We here show only the results of sampling for  $\beta$ . For the rest we do the same procedure. In the following we can see the samples for  $\beta$  after burn-in period for mentioned window.

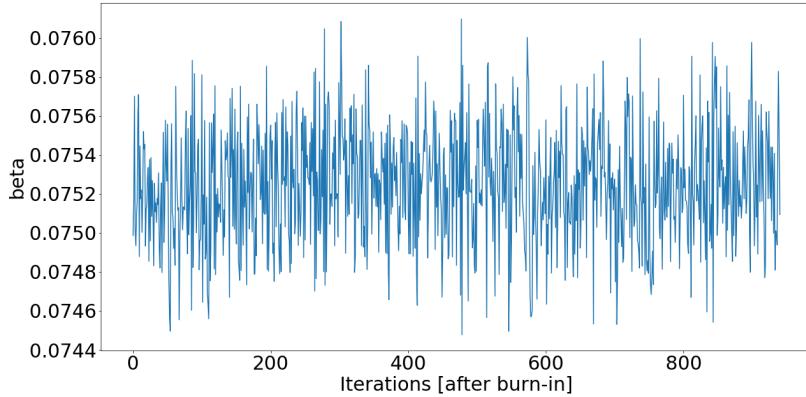
$\beta$  is actually and simply the mean value of samples shown in figure (Fig??). Again we slide the window and we estimate  $\beta$  for other time points. The same thing is implemented for  $\alpha$  and  $\gamma$ .

In our study, we estimate  $\beta$ , (Fig2), and  $\gamma$ ,(Fig3) , by time sampling for different [age groups](#) and [Gender](#).

## 2.4 Likelihood

from equation 6 we can have the stochastic model as

$$S(t+1) = (S(t) - \frac{\beta}{N} S(t) I(t)) dt + \epsilon = S(t+1)_{pre} + \epsilon \quad (12)$$



where  $\epsilon$  represent error vector and  $S(t + 1)_{pre}$  stands for predicted susceptibility by a given  $\beta$  and true  $S(t + 1)$  will be our data,  $d$ . The measurement error is assumed to be distributed normally and without correlation

$$p(d|\beta) \sim N(0, \sigma_\beta^2 \mathbb{1}) \sim \exp\left(\frac{-||S(t + 1) - S(t + 1)_{pre}||^2}{2\sigma_\beta^2}\right) \quad (13)$$

where  $\mathbb{1}$  is identity matrix. The likelihood can be calculated in the same manner for the other parameters.

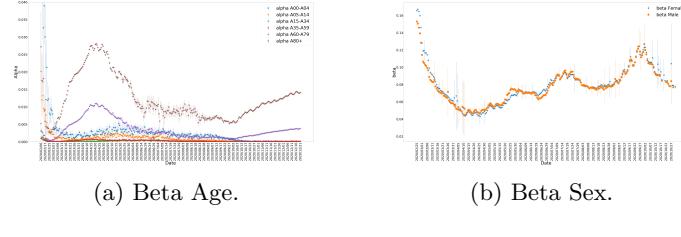


Figure 2: Effective contact rate.

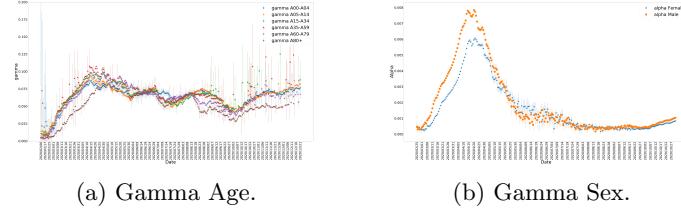


Figure 3: Recovery rate.

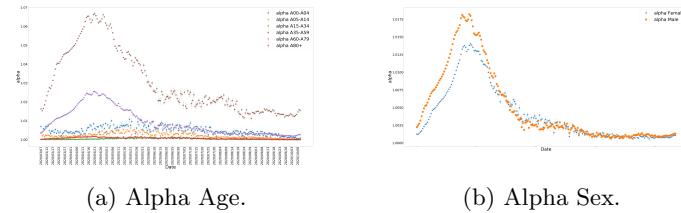


Figure 4: Mortality rate.

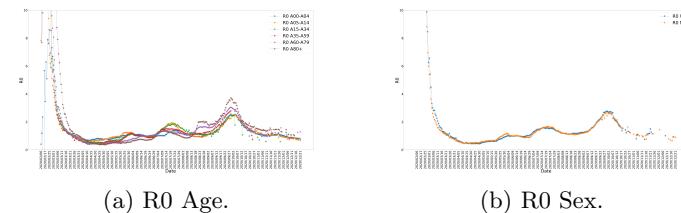


Figure 5: Reproduction number.

## References

- [1] Kermack W.O., McKendrick AG. Proceedings of the royal society of London series A, containing papers of a mathematical and physical character. Vol. 115. 1927. A contribution to the mathematical theory of epidemics; pp. 700721. [CrossRef] [Google Scholar]
- [2] Metropolis and Ulam, 1949, Metropolis et al., 1953, Hastings, 1970.