Estimation the state of the Covid-19 epidemic curve in Mayotte

Solym MANOU-ABI* [1,2,4], Yousri SLAOUI[2], Julien BALICCHI[3]

December 31, 2021

[1] Institut Montpelliérain Alexander Grothendieck
UMR CNRS 5149, Université de Montpellier

[2] Laboratoire de Mathématiques et Applications
UMR CNRS 7031,Futuroscope,Université de Poitiers

[3] Agence Régionale de Santé de Mayotte, France
Centre Kinga- route Nationale, Kaweni, Mamoudzou

[4] Centre Universitaire de Formation et de Recherche
Rue de l'université, Dembeni, Mayotte

Abstract

We study in this work some statistical methods to estimate the parameters resulting from the use of an age-structured contact mathematical epidemic model in order to analyze the evolution of the epidemic curve of Covid-19 in the French overseas department Mayotte from march 13, 2020 to february 26,2021. Using several statistic methods based on time dependent method, maximum likelihood, mixture method, we fit the probability distribution which underlines the serial interval distribution and we give an adapted version of the generation time distribution from Package R0. The best-fit model of the serial interval was given by a mixture of Weibull distribution. Furthermore this estimation allows to obtain the evolution of the time varying effective reproduction number and hence the temporal transmission rates. Finally based on others known estimates parameters we incorporate the estimated parameters in the model in order to give an approximation of the epidemic curve in Mayotte under the conditions of the model. We also discuss the limit of our study and the conclusion concerned a probable impact of non pharmacological interventions of the Covid-19 in Mayotte such us the re-infection cases and the introduction of the variants which probably affect the estimates.

*Corresponding author

Key words and phrases: Covid-19, Mathematical model, Estimation, Reproductive number, Epidemiology, Parameter estimation.

1 Introduction

Following the emergence of the Covid-19 and its spread outside of China, Europe and now the all world are experiencing the pandemic. The COVID-19 virus has caused a great disruption to the human health, social life, developments, and economics. In response, several countries have implemented unprecedented non-pharmaceutical interventions including case isolation of symptomatic individuals and their contacts, the closure of schools and universities, banning of mass gatherings and some public events, and widescale social distancing including local and national lockdowns of populations with all but essential internal travel banned. Also in response to the rising numbers of cases and deaths, and to maintain the capacity of health systems to treat as many severe cases as possible, France like European countries and other continents, have implemented some process measures to control the epidemic including his overseas department outside the metropolitan France namely in Mayotte island. To understand trends in the development of the epidemic in Mayotte from march 13,2020, to february 26, 2021 we explore the estimation of important parameters in a justified mathematical model. To this end, we describe the evolution of an individual through several possible states: susceptible, exposed, infected including symptomatics and asymptomatics cases and removed or isolated. Mathematical models can be defined as a method of emulating real life situations with mathematical equations to expect their future behavior. In epidemiology, mathematical models are relevant tools in analyzing the spread and control of infectious diseases. Many mathematical models of the COVID-19 coronavirus epidemic have been developed, and some of these are listed in the following papers [9, 12, 13, 5, 6]. Based on the development of epidemiological characteristics of COVID-19 infection, a model of type SEIR can be appropriate to study the dynamic of this disease including some statistical method to estimate the associate parameters.
In this paper, we first introduce the observed data in Mayotte and the mathematical tools that justified the modeling procedure in section2. In section 3 we estimate the serial interval distribution by several methods and the best-fit is given by a mixture model. We also derive an estimation of the generation time distribution. Section 4 provide numerical simulation of the effective reproduction number and Section 5 deal with the numerical approximation of the epidemic curve under some conditions. The results are presented and a discussion on the limit of our study in the last section.

2 Data and tools for the mathematical analysis

This work is concerned with some statistical analysis of the dynamic of the Covid-19 epidemic curve in Mayotte over a given period. Note that the ob-

served epidemic in Mayotte started at march 13,2020 where the first detected and reported case travelled from the metropolitan France three days ago. Some others imported cases were confirmed later and lead to the actual state of the spread of the virus. In this paper we use a subset of the database from march 13,2020 to february 26,2021 for simplicity. This horizon time is chosen in such a way to justify the following fact: given that the starting of the vaccination around January, 25 and since it took on average of 3 or 4 weeks for a complete vaccination, this corresponds approximately at this end of study date. In addition, it coincided with the start of the second lockdown in Mayotte.
Note that data relating to covid-19 can be downloaded from `https://www.data.gouv.fr/fr/pages/donnees-sante/`. The following pictures illustrate the daily and cumulative reported number of confirmed infected cases as well as confirmed hospitalization, intensive care and death cases in Mayotte the above interval date.

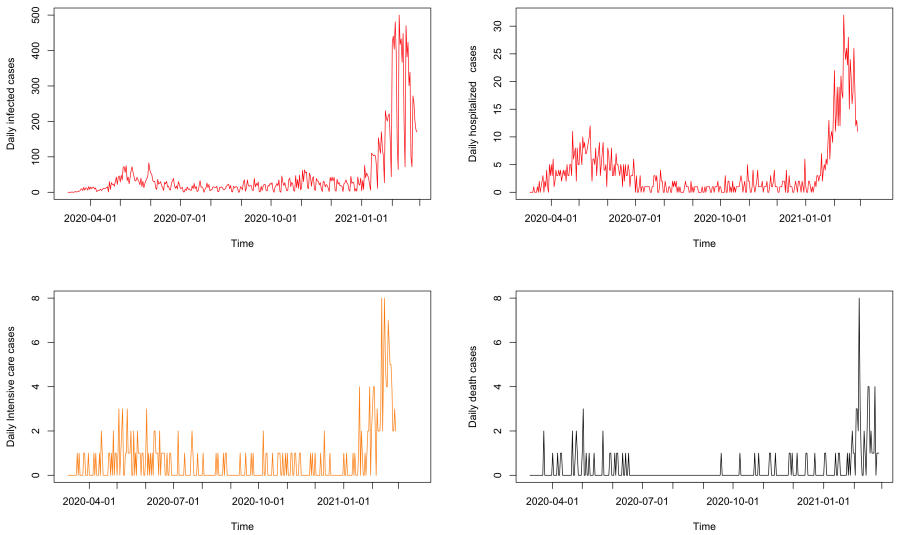


Figure 1: Reported daily cases of Covid-19 epidemic in Mayotte from march 13, 2021 to february 26, 2021

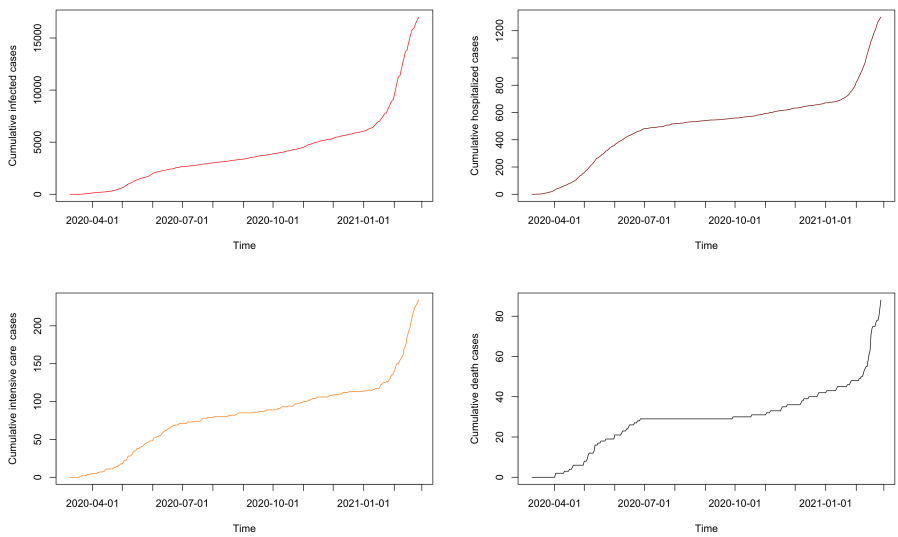


Figure 2: Cumulative cases of Covid-19 epidemic in Mayotte from march 13, 2021 to february 26, 2021

2.1 Statistical analysis by regression

Let $H(t)$, $ICU(t)$ and $D(t)$ denote respectively the number of Hospitalized, Intensive care unit and Death cases at time t. Regression is a predictive statistical approach for modeling relationship between a dependent variable say Y with a given set of observed variables say $X = (X_1, X_2, ..., X_n)$ in the form

$$Y = f(X, \eta) + e,$$

where η is unknown (vector) parameter and e representing an additive error that may stand for a white noise.
Now, let us analyse in the following lines the performance of some relevant regression linear model

- 90.84% of the variations in the cumulative intensive care $ICU(t)$ are explained by the variations of t the cumulative infected cases $I(t)$ with a coefficient around 0.013 and CI equal to $[0.012, 0.013]$ with a p-value $< 2 * 10^{-16}$ and the constant around 2.78.

- 88.89% of the variations in the cumulative death $D(t)$ are explained by the variations of the explanatory variable $I(t)$ with a coefficient around $4.42.10^{-3}$ and CI equal to $[0.0042, 0.0045]$ with a p-value $< 2.2 * 10^{-16}$ and a negligible effect of the constant.

- 86.88% of the variations in the cumulative hospitalized cases $H(t)$ are explained by the variations of the cumulative infected cases $I(t)$ with a coefficient around 0.07 with a CI equal to $[0.068, 0.07]$ with a p-value $< 2.2 * 10^{-16}$ and an important effect of the constant (intercept).

Based on the above analysis, we summarize the obtained predictions in the following pictures.

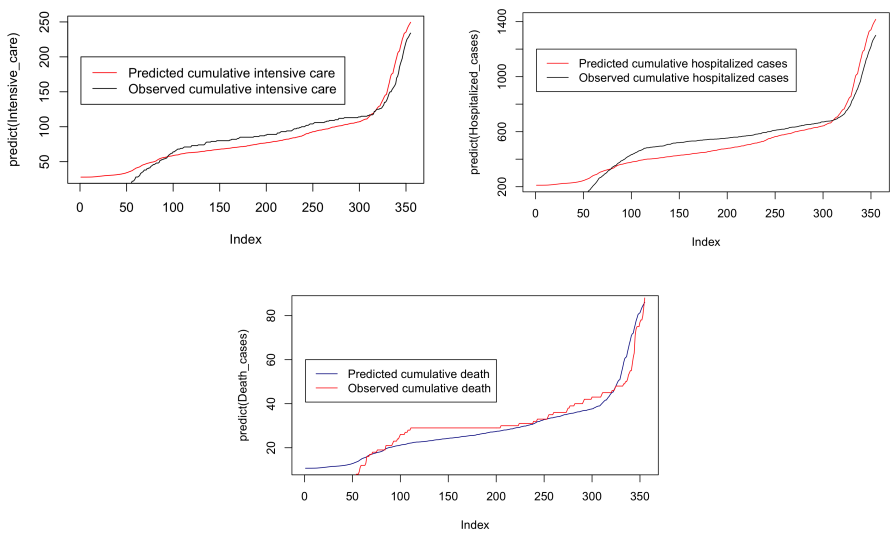


Figure 3: Predicted cumulative cases by linear regression

One can observed in the above pictures that the predicted cumulative cases based on the linear model is not better at certain points in time evolution. This may be due to some non-linear variations. In order to do better, let us deal with a fit based on Generalized Additive Model (GAM) which allows the to use of mix nonlinear, linear functions and categorial effects to data without overfitting. With a GAM, however, we can fit data with smooths, or splines, which are functions that can take on a wide variety of shapes. We fit a GAM using a polynomial transformation and predict the above desired cumulative cases as follows.

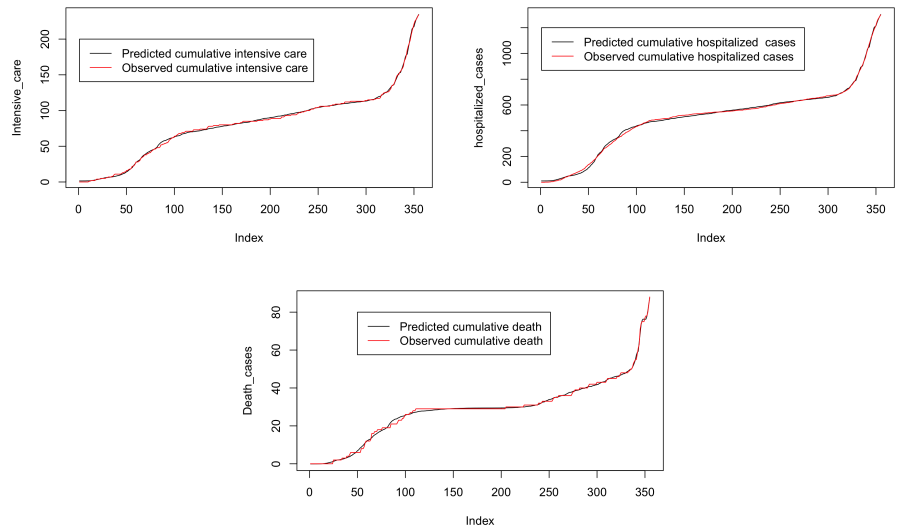


Figure 4: Predicted cumulative cases by GAM

2.2 Some definitions

Now let us specify some definitions concerning the infectivity process. Consider an infector i and infectee j.

Definition 1 *The generation time interval $\mathcal{G}_{i,j}$ is the time interval from the infection of i to infection of j.*

The generation time distribution $\mathcal{G}$ describes infectiousness with respect to the point of infection.

Definition 2 *The serial interval time $\mathcal{S}_{i,j}$ is the time interval between symptom onset of the infector i and symptom onset of the infectee j.*

Definition 3 *The incubation time interval $\mathcal{E}$ is the time interval between symptom onset and the infection.*

The infectious profile denoted by $\mathcal{P}_{i,j}$ is the time interval from the symptom onset of i to infection of j. It describes infectiousness relative to symptom onset. The above distributions are typically derived from contact tracing and can be captured by the following equations:

$$\mathcal{S}_{i,j} = \mathcal{P}_{i,j} + \mathcal{E}_j \tag{1}$$

$$\mathcal{G}_{i,j} = \mathcal{P}_{i,j} + \mathcal{E}_i \tag{2}$$

One can assume that $\mathcal{P}_{i,j}$ and $\mathcal{E}_j$ are plausibly independent as well as $\mathcal{E}_j$ independent of $\mathcal{E}_i$ so that $\mathcal{S} = \mathcal{P} * \mathcal{E}$. Note that the incubation and serial interval distributions can be estimated without further assumptions but not the generation time and the infectious profile. We refer the reader to [10, 20] for further explanations as concerns the way to derive the generation time distribution from the serial interval.

2.3 Infection model from the Markovian SIR and SEIR

To model the number of infections over time, we need to understand how the above transmission distributions appeared in a mathematical modeling. In the context of population dynamics, let us recall the following formation of Lotka and Euler in [14, 4]:

$$\Gamma(t) = \int_0^{+\infty} \Gamma(t-a)p(a)\nu(a)da,$$

where $\Gamma(t)$ is the number of births at time t, $p(a)$ is the probability of survival to age a and $\nu(a)$ is the fertility at age a. The above renewal equation was adapted to epidemics in the seminal work of Kermack and Mc Kendrick in [8]. To understand the above renewal equation in an epidemiological context, let us consider firstly the following simple Markovian SIR model in a closed population with size N partionned into Susceptible individuals S, Infectious individuals I and Recovered individuals R. We consider only transmission from person-to-person and we assume that infectious individuals become permanently immune after their infectious period during our horizon time study T. Of course in the case of Covid-19 we do not have permanent immunity.

$$\begin{aligned} dS(t) &= -\beta(t)S(t)\,\frac{I(t)}{N}dt, \\ dI(t) &= \beta(t)S(t)\,\frac{I(t)}{N}dt - h(t)I(t)dt, \\ dR(t) &= h(t)I(t)dt \end{aligned}$$

where $h(t)$ is the hazard rate function associated to the survival function $G(t)$ of the time infectivity. The time varying transmission rate is denoted by $\beta(t)$. For simple computations, we assume that $h(t)$ is constant and equal to γ the mean rate of the infectivity (being exponentially distributed with parameter γ) so that $G(t) = e^{-\gamma t}$. The second equation of the above system tell us that

$$\Gamma(t) = \frac{\beta(t)S(t)I(t)}{N} = \frac{\beta(t)S(t)}{\gamma N}\int_0^{t-t_0} g(\tau)\Gamma(t-\tau)d\tau$$

where t_0 is the starting date of the epidemic and $g(\tau) = \gamma e^{-\gamma\tau}$ is the density of the probability distribution function of the secondary infections according

to primary-infection: the so called generation time distribution (assumed to be exponentially distributed). The above formula means that new susceptible individuals are reported at time t according to the contact rate $\beta(t)\gamma^{-1}$ if they were infected before time t and remains infectious for the duration τ until date t with probability $g(\tau)$. The incidence increases exponentially quickly if

$$\frac{\beta(t)S(t)}{\gamma N} > 1.$$

In particular, at time $t = 0$, one have the following formula for the basic reproduction number at the begining of the outbreak:

$$R_0 = \beta\gamma^{-1}, \quad \beta = \beta(0).$$

$$R_0(t) = \frac{\beta(t)S(t)}{\gamma N}$$

and

$$\beta(t) = R_0(t)\frac{\gamma N}{S(t)} \tag{3}$$

Thus

$$R_0(t) = \frac{\Gamma(t)}{\int_0^{t-t_0} g(\tau)\Gamma(t-\tau)d\tau}$$

Therefore in a discrete time setting we have the following algorithmic formula for all $t > t_0$:

$$R_0(t) = \frac{\Gamma(t)}{g(0)\Gamma(t) + g(1)\Gamma(t-1) + ... + g(t-t_0)\Gamma(t_0)}$$

The following fact is easy to verify

Proposition 1 *In the above Markovian SIR model if $\gamma \in [\gamma_{min}, \gamma_{max}]$ then in a discrete time setting we have*

$$e^{-\gamma_{max}(t-t_0)}I(t_0) + \gamma_{max}^{-1}\sum_{t_0\leq\tau\leq t} g(t-\tau)\Gamma(\tau)$$

$$\leq I(t) \leq e^{-\gamma_{min}(t-t_0)}I(t_0) + \gamma_{min}^{-1}\sum_{t_0\leq\tau\leq t} g(t-\tau)\Gamma(\tau),$$

for all $t \geq t_0 + 1$.

We illustrate the above proposition as follow with $\gamma^{-1} \in [7, 21]$ days.
The above Markovian SIR model under learn the estimation of $I(t)$, this was predictable since it does not take into account the delay in the infectivity process due to the incubation.

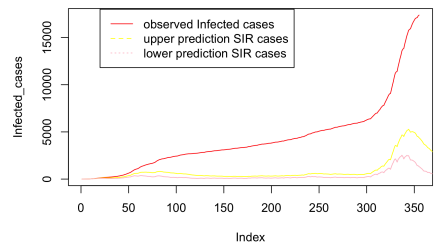


Figure 5: Estimated (lower and upper) cumulative infected cases with a markovian SIR model

Let is us consider a Markovian SEIR model where the population is partitionned into five categories : Susceptibles S; Exposed E infected but not yet infectious in a latent period with mean λ^{-1}; Infectious I being reported or unreported with an infectious period with mean γ^{-1}, Dead D and Recovered R.
The model meets the following assumption.

- The natural average mortality rate μ and a recruitment rate Λ can be considered but we shall assume that the recruitment rate Λ compensates succeptibles individuals who died:

  $$\Lambda(t) = \mu S(t).$$

  Given initial conditions $D(0) = D_0$ and

  $$S(0) = N - E(0) - I(0) - R(0) - D(0),\ E(0) = E_0, \quad I(0) = I_0, \quad R(0) = R_0,$$

  where N is the total fixed population size, the model consists of the following system of differential equations

  $$\begin{cases} dS(t) = \Lambda - \mu S(t) - \beta(t)\frac{S(t)I(t)}{N}dt, \\ dE(t) = \beta(t)\frac{S(t)I(t)}{N}dt - \lambda E(t)dt, \\ dI(t) = \lambda E(t)dt - \gamma I(t)dt \\ dR(t) = (\gamma - \alpha)I(t)dt \\ dD(t) = \alpha I(t)dt \end{cases} \tag{4}$$

  One can look by this model to keep track of the cumulative cases from the time of onset of symptoms $C(t)$ which is not a compartment and defined as $dC(t) = \lambda E(t)$.
  Assume that the probability of a susceptible to be alive at time $t > 0$ given that we was alive at time $t = 0$ follows an exponential distribution of parameter μ. Assume also that the incubation and the infectiousness are independent and follows respectively an exponential distribution with

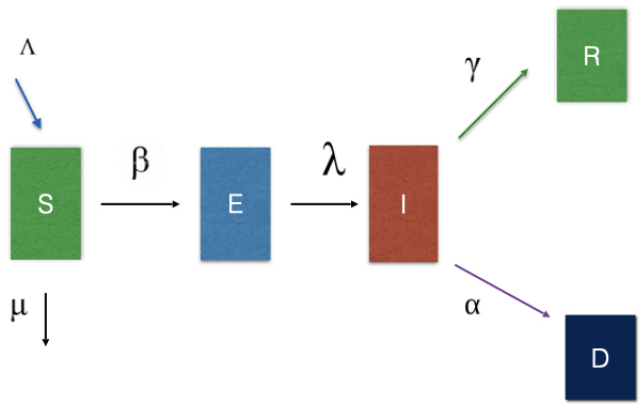


Figure 6: Compartments of the SEIRD model

parameter λ and γ and let G_λ, G_γ and G_μ their associate survival functions.

The following proposition follows from the disease equations.

Proposition 2 *Under the above assumption, we have in the Markovian SEIR model :*

$$E(t) = e^{-\lambda(t-t_0)}E(t_0) + \lambda^{-1}\sum_{u=t_0}^{t} g_\lambda(t-u)\Gamma(u)$$

and

$$I(t) = e^{-\gamma(t-t_0)}I(t_0) + \lambda\gamma^{-1}\sum_{v=t_0}^{t} g_\gamma(t-v)E(v)$$

so that if $\gamma \neq \lambda$ we have

$$I(t) = \gamma^{-1} g_\gamma(t-t_0) I(t_0) + \gamma^{-1} f_{\lambda,\gamma}(t-t_0) E(t_0) + \gamma^{-1} \sum_{t_0 \tau \leq t} f_{\lambda,\gamma}(t-\tau)\Gamma(\tau),$$

where $f_{\lambda,\gamma}(x) = \frac{\lambda\gamma}{\gamma-\lambda}(e^{-\lambda x} - e^{-\gamma x})$ is the density distribution of the two parameter (λ, γ) hypo-exponential distribution.

We illustrate the above proposition as follow with $\gamma^{-1} \in [7, 21]$ days and $\lambda^{-1} \in [1.5, 7]$ days

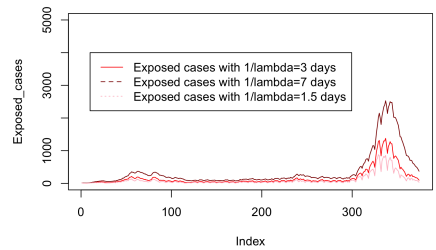


Figure 7: Estimated cumulative exposed cases with $E(0) = 15$ in a markovian SEIR model with various λ

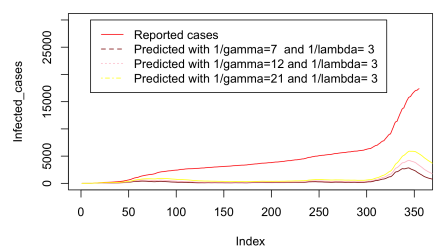


Figure 8: Infected cumulative cases with $E(0) = 15$ in a markovian SEIR model with various γ and fixed λ

Here again the Markovian SEIR model under learn the estimation of $I(t)$. This is due to the fact that the generation times are not exponential as we will see in the next section. For a deeply study of non-Markonvian SEIR model we refers the reader to the work of [5, 6].

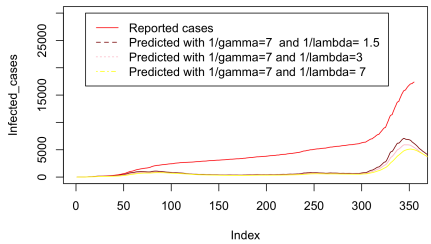


Figure 9: Infected cumulative cases with $E(0) = 15$ in a markovian SEIR model with various λ and fixed γ

3 Estimation of the serial interval and generation time distributions

Some major studies have been made on the determination of generation times through serial interval distributions and incubation time from real data collected, see for instance [7, 16]. For the rest of this paper, we will focus on the data collected in [7] where the authors report temporal patterns of viral shedding in 94 patients with laboratory-confirmed Covid-19 and modeled Covid-19 infectiousness profiles from a separate sample of 77 infector–infectee transmission pairs. The variation between individuals (infector/infectee) and the transmission process is summarized respectively by the distribution of incubation periods and the distribution of serial intervals. The database used contains the number of infected cases (in red), the lower limit of infectors (in black) and the upper limit of infectors (in green). The variation between individuals and chains of transmission is summarized respectively by the distribution of incubation periods and the distribution of serial intervals.

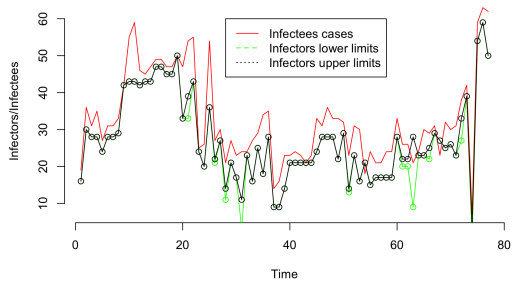


Figure 10: Evolution of the infector/ infectee pairs

The main purpose in this section is to study the above data of transmission pairs in order to be able to fit the best distribution of the serial interval which is computed as follow like in[7]:

$$serial\_interval = y - (\frac{x.ub + x.lb}{2})$$

with

- $x.lb$: Lower limit dates of onset of symptoms of infectors
- $x.ub$: Upper limit dates of onset of symptoms of infectors
- y: Dates of onset of symptoms of the infectee.

The incubation period is calculated by the formula

$$y - \text{ probable date of infection by the infector.}$$

The following picture show the histogram of the serial interval from the underline database.

Now, we will attempt to fit a theoretical distribution to the above histogram of the observed empirical distribution function of the serial distribution among a set of distributions that we will choose. A rapid analysis of the histogram seems to show that the serial interval can be fit by a family of exponential distribution or a mixture of them because of the appearance of several peaks. There are many methods for generating estimates of parameters for standard probability distributions such as the method of moments and the Maximum Likelihood.The use of others methods was generally confined to analytical work on mixture densities.

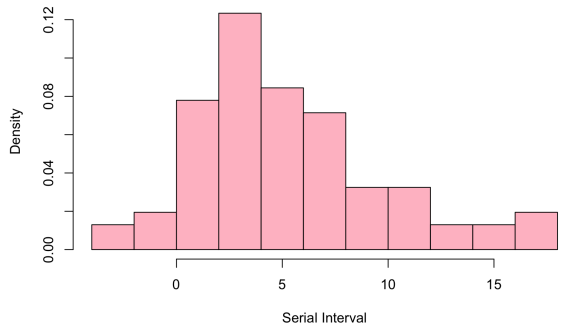


Figure 11: Histogram of the serial interval

Estimation of the serial interval distribution by the method of moments

In this section we use the method of moments amoung a set of classical laws. The method of moments consists in estimating the parameters of the distributions from the theoretical finite moments with their empirical estimates. The results of this method can be seen in the following picture.

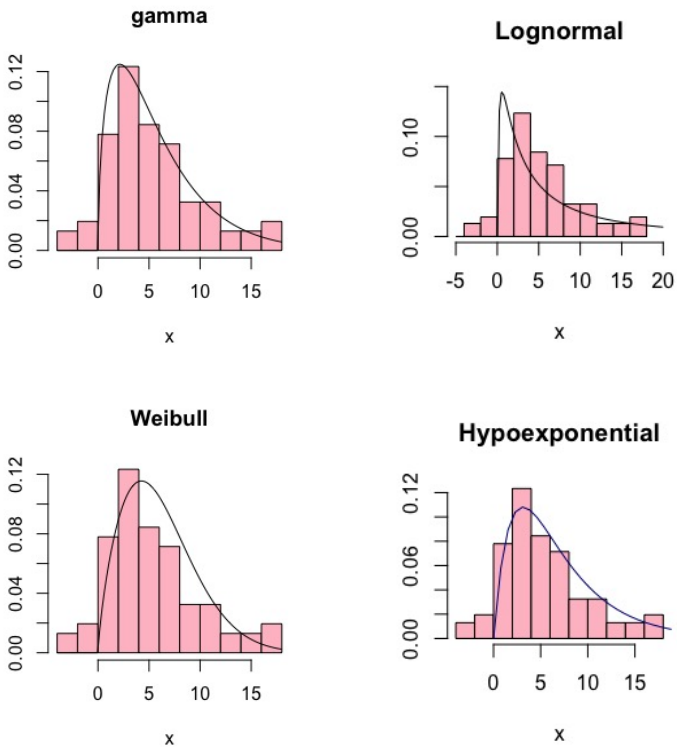


Figure 12: Fitted distributions with the method of moments

The above pictures mentioned namely fitted:

- Gamma distribution with parameters 1.584523 and 0.2735611
- Lognormal distribution with parameters 1.756514 and 1.526372
- Weilbull distribution with parameters 1.756514 and 6.87615

- Two parameter Hypo-exponential (denoted by *exp_conv*) distribution with parameters 1/4.9 and 1/2.1

Note that the standard exponential distribution (here with parameter 0.1726457) does give a good fit, as it can be seen in the following picture.

As we can see, the gamma, weibull seems to matche if we do not take into account the left tails of the distributions. Using the Kolmogorov-Smirnov non parametric test with a risk $\alpha = 0.05$ gives an p-value around 0.5229,

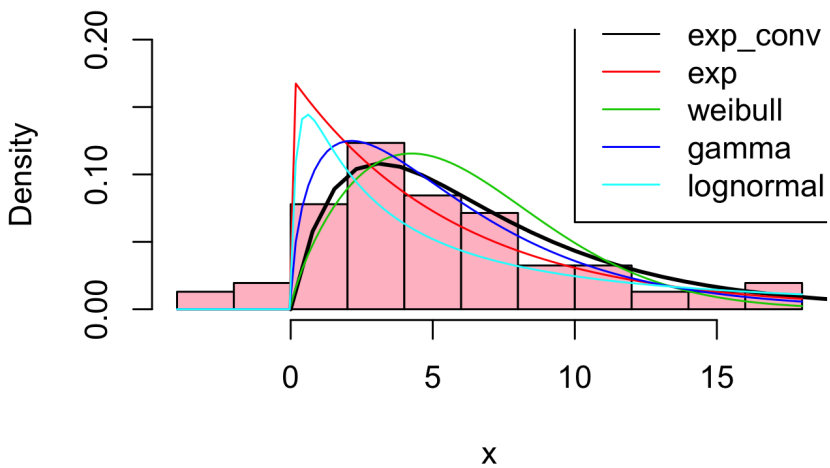


Figure 13: Summary (moment method) of fitted distributions.

$0.5237 * 10^{-9}$ and $6.312572 * 10^{-4}$ for Gamma, Weibull and Lognormal distribution respectively.

Note that the method of moments does not give best fit and this can be explained by the fact that this basic method does not use any optimization techniques as the following Likelihood method.

3.1 Estimation of the serial interval distribution by the Likelihood method

Here we consider the maximum likelihood estimator which is a statistical method used to estimate the parameters of the probability law of a given observation $x = (x_1, .., x_n)$ with probability densities $f(x_i|\theta)$ by determining the values of the parameters maximizing the log-likelihood function given by :

$$L(\theta) = \log(\prod_{i=1}^{n} f(x_i|\theta)) = \sum_{i=1}^{n} \log(f(x_i|\theta))$$

where θ is the (vector) parameter to be estimate. The likelihood function then represents the joint density of the individual observations x_i for any given level of the distribution parameters. The maximum likelihood estimate is the distribution parameter values, which maximize the likelihood function since these same values of the estimators also maximize the log-likelihood function. The *nlm* function of R-Package in [17] makes it

possible to optimize the log likelihood functions of the different laws that we pass as an argument in addition to the initialization of the parameters to be optimized and of the serial interval that we seek to determine its better distribution. The following picture represents several fitted densities estimated by this method.

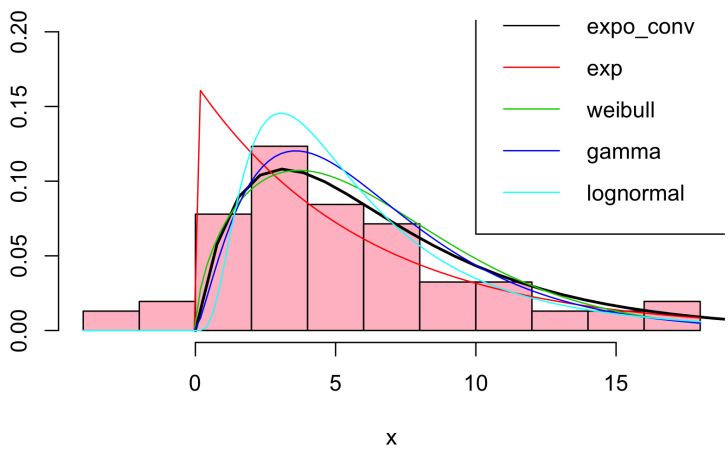


Figure 14: Estimation of distributions by the maximum likelihood method

Table 1: Log-likelihood, AIC and BIC values

Law	Log-likelihood	AIC	BIC
Gamma	192.83	387.66	390.01
Log-normal	192.51	421.76	424.10
Weibull	194.30	390.60	392.94
Hypo-exponential	179.76	Inf	Inf

We can see that compared to the method of moments, the Maximum Likelihood improve in some way the estimation. The log-likelihood values and widely applicable information AIC and BIC criteria is summarise in the following table

We note that that the gamma, log-normal and weibull models are the best to fit data. In addition one can see that the two parameter hypo- exponential distribution does not converge, which explains why the estimated parameters are negative.

```
                                  gamma          exp                      lognorm
minimum                        192.8342     209.8803                     192.5157
estimate          2.3157219, 0.3680613    0.1655624          1.608040, 0.702471
gradient   -9.941429e-07, 2.529532e-06 -1.98952e-06 -1.558913e-05, 6.264145e-05
code                                1            1                            1
iterations                         18            4                            8
                               weibull                exp_conv
minimum                        194.3016           1.797693e+308
estimate           1.568012, 7.047558 -0.5877867, -1.1314021
gradient   1.486328e-06, 1.935766e-07                   0, 0
code                                1                      1
iterations                         20                      0
```

Figure 15: Maximum likelihood method results

Note also that the left tails of the distributions don't matches and since the underline left tail data represents 7% of the data we neglect this part of the data. Another upcoming work will deal with the theoretical aspects of estimating heavy tails of distribution that may be appropriate for doing cross mixtures of heavy tails and exponential family laws.

3.2 Estimation of the serial interval distribution by mixture methods

Now let us deal with mixture modeling with exponential families. Mixture models are a powerful and flexible tools (see [15, 1]) to model an known smooth probability density function as a weighted sum of parametric density functions $f_j(x|\theta_j)$ in a possibly multivariate independent observations $x = (x_1, .., x_n)$ drawn from k several densities :

$$f(x) = \sum_{j=1}^{k} p_j f_j(x|\theta_j)$$

where k represents the number of classes or groups of the mixing, $p_j > 0$ the relative proportion of observations from each class j such that $\sum_{j=1}^{n} p_j = 1$. The mixture density will allow to find the parameter $\Theta = (p_j, \theta_j)$ of the model where θ_j is the (vector) parameter from class j.
The log-likelihood of the mixing density is defined by:

$$L(\Theta) = \log(\prod_{i=1}^{n} f(x_i|\Theta)) = \sum_{i=1}^{n} \log(\sum_{j=1}^{k} p_j f_j(x_j|\theta_j)).$$

The maximum likelihood estimates are:

$$\max \sum_{i=1}^{n} \left\{ \log \Big(\sum_{j=1}^{k-1} (p_j f_j(x_i|\theta_j)) + (1 - \sum_{j=1}^{k-1} p_j) f_k(x_i|\theta_k) \Big) \right\}$$

subject to

$$\sum_{j=1}^{k-1} p_j \leq 1.$$

For $k > 1$, the sum of term appearing inside a logarithm makes this optimization quite difficult. Solution of this non-linear optimization problem has long been a difficult task for applied researchers. There are, of course, many general iterative procedures which are suitable for finding an approximate solution of the likelihood equations and which have been honed to a high degree of sophistication within the optimization community.
For the data, a two-component mixture model is clearly a reasonable model based on the bi-modality evident in the histogram.
Using the R software function **optim**, it possible to minimize the log likelihood functions of the different laws taken as arguments of this function in addition to the initialization of the parameters to be optimized and the serial interval distributions we want the best-fit. The figures below summarize the estimation results of the mixture models. Gamma mixing, lognormal and weibull mixing seem to give better results compared to the others.

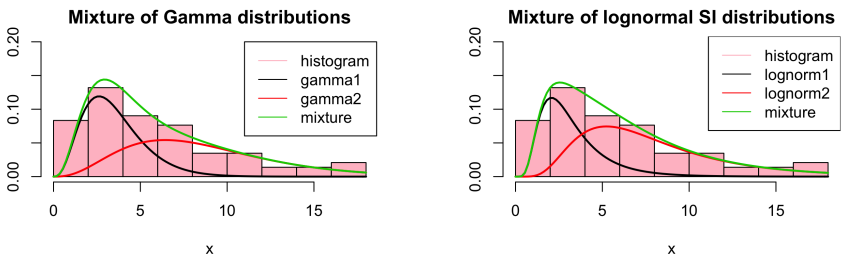


Figure 16: Mixture distributions by the Optim method

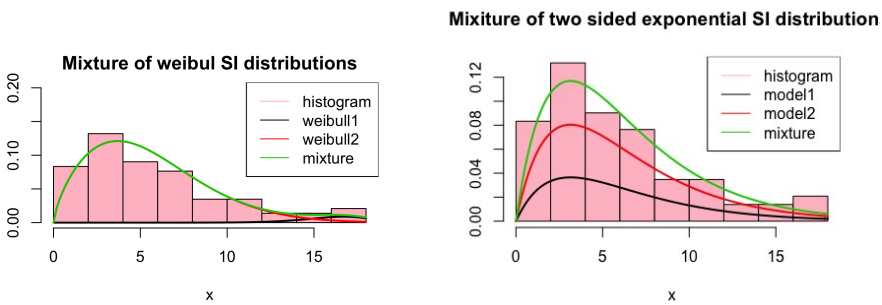


Figure 17: Mixture distributions by the Optim method

The Kolmogorov-Smirnov test, log-likehood, AIC and BIC values is summarised in the following table We conclude that the best-fit mixture is given by $p = 0.0406$ and two Weibull distributions with parameters (9.96140450, 16.90693027) and (1.70447135, 6.20140484).
We can see the visualization of the results of the mixtures of each law as follows.Note that one can still go further in the estimation by a cross-mixture among the best three-fit distributions and the a heavy tail distribution by the use of the EM algorithm. This will be done in another forthcoming theoretical and practical work.

Table 2: Kolmogorov-Smirnov test, log-likehood, AIC and BIC values

Law	P-value	alpha	Log-likehood	AIC	BIC
Gamma	0.686	0.05	192.06	-374.12	-362.40
Log-normal	$6.65*10^{-4}$	0.05	191.90	-373.81	-362.0935
Weibull	$1.16*10^{-16}$	0.05	191.60	-373.21	-361.49

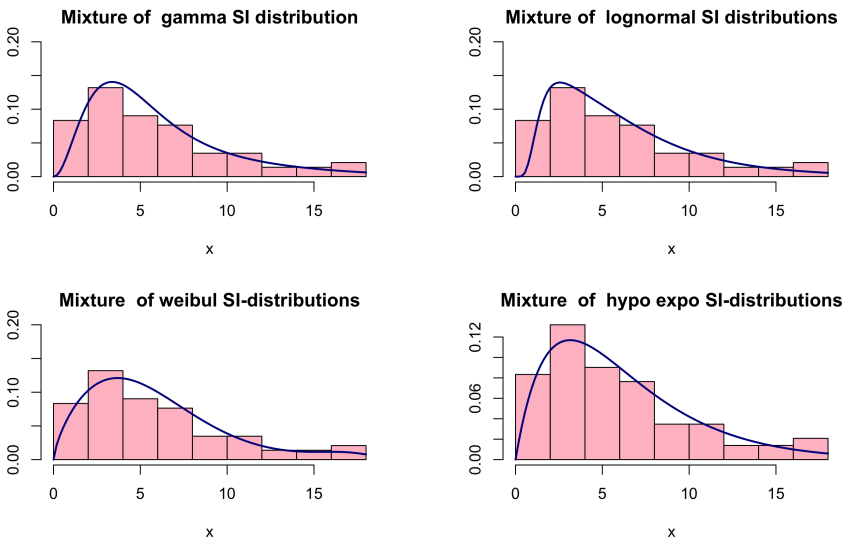


Figure 18: Result of each mixture density distribution

3.3 Results of the estimation on the incubation and infectious time distribution

Let us summarise here some well known works about the estimation of the infectiousness and incubation period of [11]. In this work, the authors established that the best fit bincubation period distribution is of mean 5.2 days with 95% CI equal to $[4.1, 7]$. Using statistical estimation method, the authors also show that the best-fit distribution of this incubation rate is managed by a log normal distribution of parameters 1.434065 and 0.6612.

Their analysis also suggests that viral shedding may begin 5 to 6 days before the appearance of the first symptoms and inferred that infectiousness started from 12.3 days (95% CI equal to $[5.9, 17]$ days before symptom onset and peaked at symptom onset (95% CI equal to $[-0.9, 0.9]$ days). The authors pointed out that with sensitivity analysis, the infectiousness start from 5, 8 and 11 days before symptom onset and infectiousness was shown to peak at 2 days before to 1 day after symptom onset, and the proportion of pre-symptomatic transmission ranged from 37% to 48%.

Estimation of the generation time distribution

Our aim here is the estimation of the generation time distribution by the use of the above best-mixture distribution. Let us mentionned that in the package R_0 (see [18, 2]), the best generation time distribution is fitted by

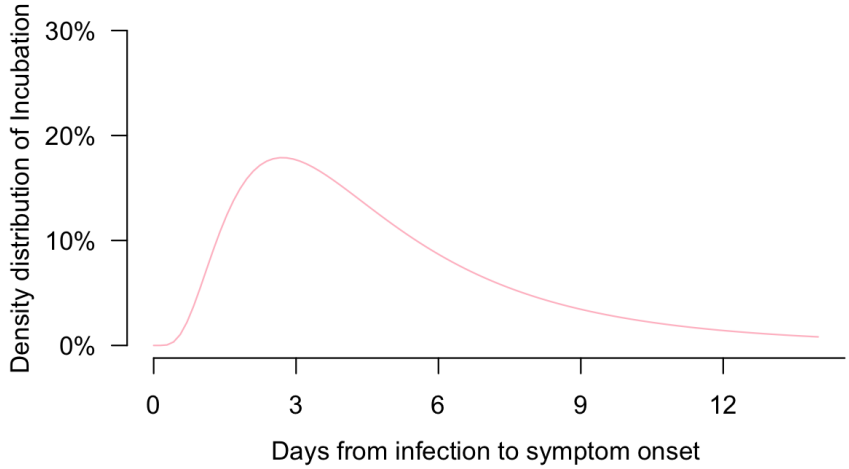


Figure 19: Density distribution of incubation by a best-fit log-normal distribution with parameters 1.4340 and 0.6612

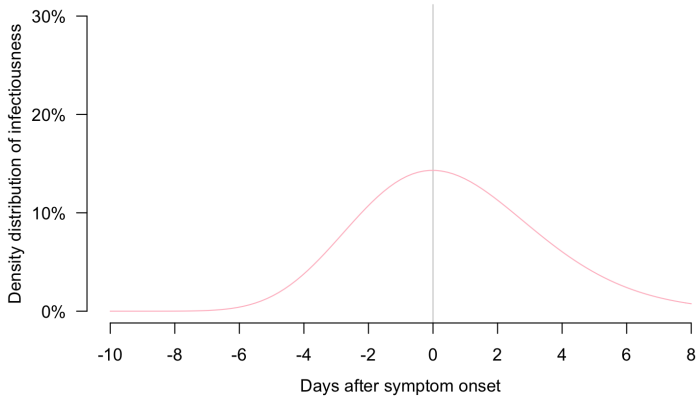


Figure 20: Density distribution of the infectious time distribution

the *generation.time* function, based on the epidemic incidence curve and the best-fit serial interval distribution choosen among "gamma", "lognormal" or "weibull" distributions using the Newton-Raphson optimization algorithm. In our case, we use the above Optim function to fit the best

serial interval distribution, which is given for instance by the Gamma distribution (since it turns out to be the best to match the data in mixture model). Thus, using our best-fit mixture model together with the Mayotte epidemic curve, we obtain a modified adapted generation time distribution of the Package $R0$. The result can be seen in the following picture.

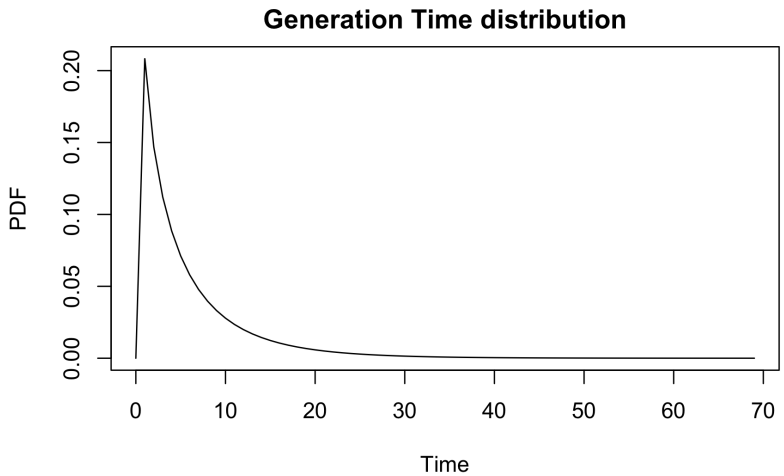


Figure 21: Generation time distribution by the best-fit mixture of the SI distribution (Weibull) by the optim method associated to package R0

4 Numerical estimates of the effective reproduction number

Understanding the development of an epidemic is important in order to control its spread. For this, various models point out the importance of the the basic reproduction number R_0 at the start of an epidemic. It is historically defined as the average number of new cases of infection generated by an individual during a period of infectivity (see [3]). As stated above, if this number is less than 1 then the epidemic will tend to die out. There are several methods for computing the parameter R_0 as well as its time dependent version as explained above. Let us mentioned some others parametric and non-parametric approaches. In the parametric approach, we will approach the estimation of R_0 by two method based on the attack rate of an epidemic and an exponential growth rate. The non-parametric approach will also be described from several angles.

4.1 Non parametric estimation based on exponential growth rate

The exponential growth rate r is the change per capita in the number of new cases per unit time during the early phase of an outbreak. It is an important measure for the speed of spread of an infectious disease. It being zero is, like the basic reproduction number $R_0 = 1$, a disease threshold. The disease can invade a population if the growth rate is positive and cannot invade (with a few initially individuals) if it is negative. In fact, it can be used to infer R_0. We describe here a non parametric method following J.Wallinga and M. Lipsitch in [21, 22]
Let η be the probability that an individual will remain infectious over a unit of time after being infected (i.e. age of infection) and β be the transmission rate at age of infection a. Then $\tau(a) = \eta(a)\beta(a)$ is the transmissibility of an infectious individual at the age of infection a, assuming that the entire population is susceptible. Hence

$$R_0 = \int_{a=0}^{\infty} \tau(a)\,\mathrm{d}a.$$

We can normalize $\tau(a)$ to be probability density function:

$$\begin{aligned} w(a) &= \frac{\tau(a)}{\displaystyle\int_{a=0}^{\infty} \tau(s)\mathrm{d}s} \\ &= \frac{\tau(a)}{R_0}. \end{aligned}$$

so that

$$\tau(a) = R_0 w(a).$$

Let $\Gamma(t)$ be the number of new infections during the time interval $]t; t+dt[$. Note that the new infections at time t are the sum of all infections caused by infectious individuals infected at a unit of time (i.e. at time $t = a$) if they remain infectious at time t (with an infectious age a) and their contact is sensitive. In other words

$$\begin{aligned} \Gamma(t) &= \int_0^{\infty} \Gamma(t-a)\tau(a)s(t)\mathrm{d}a = s(t)\int_0^{\infty} \Gamma(t-a)\tau(a)\mathrm{d}a \\ &= s(t)\int_0^{\infty} \Gamma(t-a)w(a)R_0\mathrm{d}a. \end{aligned}$$

At the start of an epidemic ($t = 0$), where the epidemic grows exponentially (with an exponential growth rate r). $s(t) \sim 1$ and $\Gamma(t) = c_0 e^{rt}$ where c_0 is the initial number of cases at time $t = 0$. Thus

$$c_0 e^{\lambda t} = R_0 \int_0^{\infty} c_0 e^{\lambda(t-a)} w(a)\mathrm{d}a$$

so that

$$R_0 = \frac{1}{\int_0^\infty e^{-ar} w(a)\mathrm{d}a} = \frac{1}{M(-r)} \tag{5}$$

where

$$M(r) = E(e^{-rA}) = \int_0^\infty e^{-ar} w(a)\mathrm{d}a$$

is the the moment generating function of the generation time distribution.For $t \neq 0$,

$$R_0(t) = R_0 * s(t) = \frac{\Gamma(t)}{\int_0^\infty \Gamma(t-a) w(a)\mathrm{da}}$$

4.2 Estimation based on attack rate

Let us consider again the elementary Markovian SIR model. Recall that the basic reproductive number and the time varying reproductive number are defined by the relation

$$R_0 = \frac{\beta}{\gamma} \quad \text{and} \quad R_0(t) = \frac{S(t)}{N}\frac{\beta}{\gamma} = \frac{S(t)}{N}\frac{\beta}{h(t)}.$$

In epidemiology, the attack rate AR is the percentage of an at-risk population that contracts the disease during a specified time interval. The term attack rate is sometimes used interchangeably with the term incidence proportion.The attack rate is calculated as the number of people who became ill divided by the number of people at risk for the illness. if we denote by

$$i(t) = \frac{I(t)}{N}, \quad s(t) = \frac{S(t)}{N},$$

it is straighfoward to see that

$$i(t) = i(0) + s(0) - s(t) + \frac{1}{R_0}\ln(\frac{s(t)}{s(0)}).$$

If at time $t = 0$,we assume that no one is retired then $i(0) + s(0) = 1$ and

$$i(t) = 1 - s(t) + \frac{1}{R_0}\ln(\frac{s(t)}{s(0)}).$$

By taking the limit and setting $s_\infty = 1 - AR$ we will have:

$$R_0 = \frac{-\ln(\frac{(1-AR)}{S0})}{AR - (1 - S_0)}. \tag{6}$$

where S_0 represents the percentage of the initial homogeneous and closed population assuming no intervention during the epidemic. From the relation

$$R_0(t) = R_0 s(t).$$

The varying attack rate is then defined by

$$AR(t) = \frac{S(t)}{R_0(t) - S(t)}.$$

The confidence interval with respect to the rate attack is given by

$$IC\ \ (AR)\ \ = AR \pm 1.96\sqrt{AR\frac{(1-AR)}{n}}$$

Note that this approach assumes that the transmission rate β is known which is not always true. This proves the limits of the parametric method.

4.3 Non parametric estimation based on Maximum likehood

This estimation method is proposed by White and Pagano in [23] and is based on the assumption that the number of secondary cases caused by an index case follows a poison distribution with the expected value R. Let $(N0, N1, ..., NT)$ be the incident case as a function of time with a time horizon T and w a distribution of the generation time. We estimate R by maximizing the log-likelihood

$$LL(R) = \sum_{t=1}^{T} \log(\frac{e^{-\mu_t}{\mu_t}^{N_t}}{N_t!})$$

where

$$\mu_t = R\sum_{i=1}^{k} N_{t-i}w_i$$

From the first derivative of the log-likelihood define by

$$\frac{\delta LL(R)}{\delta R} = \frac{\delta \sum_{t=1}^{T} log(\frac{e^{-\mu_t}{\mu_t}^{N_t}}{N_t!})}{\delta R} = \sum_{t=1}^{T}(-\sum_{i=1}^{k} N_{t-i}w_i + N_t\frac{1}{R}).$$

we can estimate R by solving the equation:

$$\frac{\delta LL(R)}{\delta R} = 0$$

$$\Longrightarrow \sum_{t=1}^{T}\sum_{i=1}^{k} N_{t-i}w_i = \sum_{t=1}^{T} N_t \frac{1}{R}$$

and

$$\hat{R} = \frac{\sum_{t=1}^{T} N_t}{\sum_{t=1}^{T}\sum_{i=1}^{k} N_{t-i}w_i} \tag{7}$$

4.4 Numerical results of the estimation

In this section we make numerical simulation based on the observed data in Mayotte island to estimate the time varying reproduction number using not only the non parametric method based on the generation time distribution fitted by the daily observed infected cases and the best mixture of the serial interval distribution given by the Weibull distribution.
The following picture represents the evolution of the time varying reproduction number $R_0(t)$

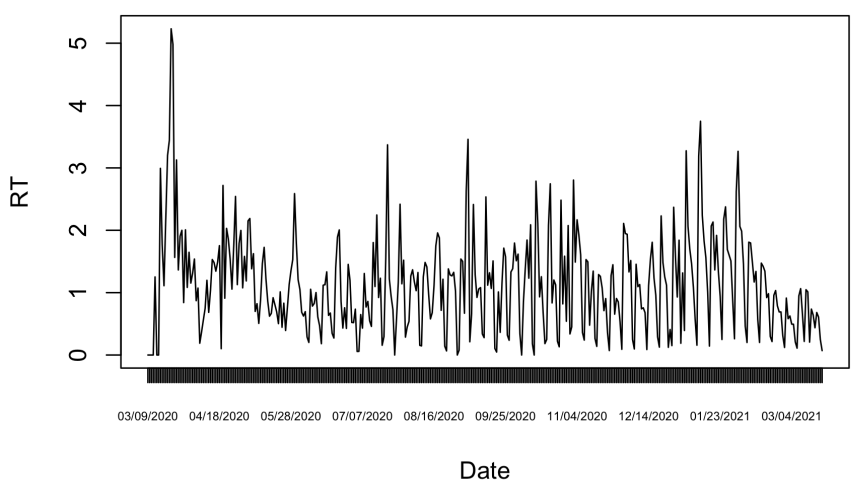


Figure 22: Evolution of the time varying reproductive number in Mayotte from march 13, 2020 to february 26, 2021

We can see that at the start of the epidemic the basic reproductive number is around 3, which is consistent with the litterature. In the following picture, we make the curve smooth using estimates values.

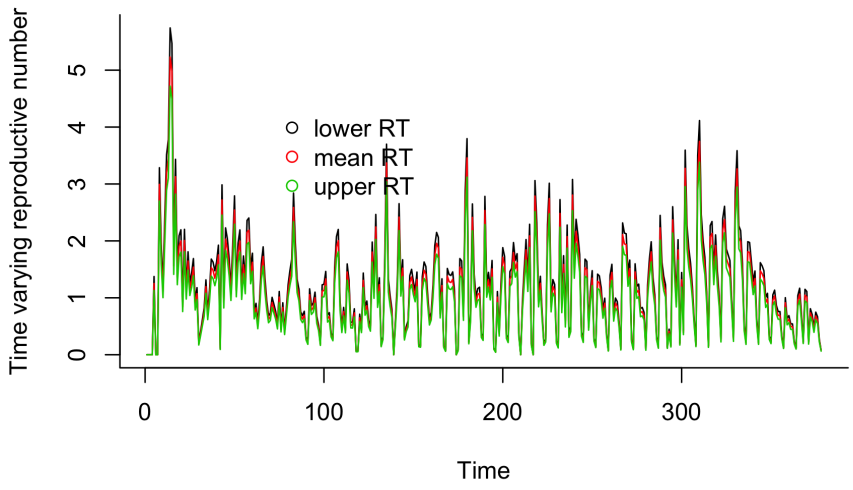


Figure 23: Time varying reproductive number in Mayotte with confidence interval of 95 per cent

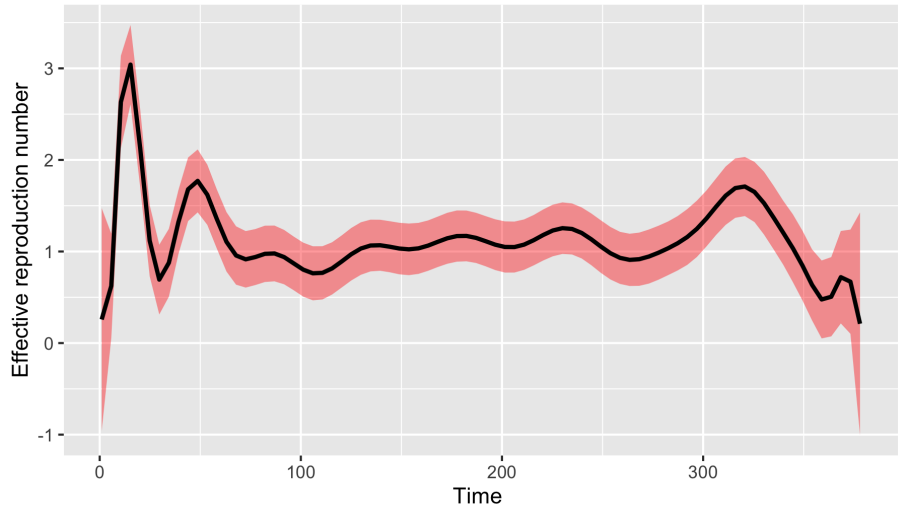


Figure 24: Evolution of the time varying reproductive number in Mayotte

Estimating of the number of reproduction number by the maximum likelihood method during a time horizon $T = 35$ provide an estimation of $R0$ around 1.018325 as present in the following

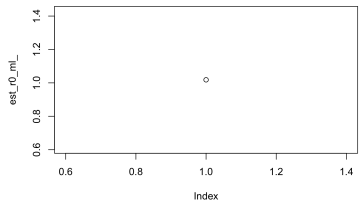


Figure 25: Basic reproductive number by maximum likelihood method

Estimation of the reproduction number by attack rate during a time horizon $T = 35$ giving an estimation of R_0 around 1.022491.

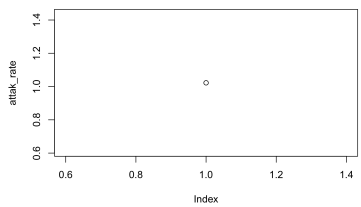


Figure 26: Basic reproductive by attack rate

Estimation of the reproduction number by exponential growth rate during a time horizon $T = 35$ giving an estimation of R_0 around 1.056758.

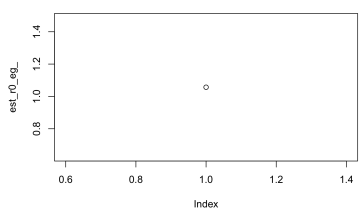


Figure 27: Basic reproductive number by exponential growth rate

5 Numerical approximation of the Covid-19 epidemic curve in Mayotte

We set up here the numerical simulations by incorporating the above statistical estimates. The model equations are an age structured contact model. Such model allow the heterogeneous structure of the population since there is variation by age in the contact. The inclusion of the age and the contact improves the realism of a model. Consider a population aggregated by age into n groups labelled by $i = 1, 2, ..., n$. The population within age group i is partitioned into Susceptibles S_i, Exposed E_i, infected but not yet infectious, in a latent period; Infected and Infectious I_i being reported or unreported; Recovered R_i. Let N_i be the total population of the age group i. We have

$$N_i = S_i + E_i + I_i + I_i + R_i.$$

If we assume that there is no change in the age group over time, we have

$$N = \sum_{i}^{n} N_i.$$

Contact matrix
Denote by $C_{i,j}$ the number of contacts of a susceptible individual in age group i with a susceptible individual in age group j;

Definition 4 *The matrix $(C_{i,j})_{1\leq i,j\leq n}$ is called the contact matrix. It represents the average number of daily contacts of an individual of class i with an individual of class j.*

Clearly, the total number of contacts between group i to group j must be equal to the total number of contacts between group j to group i. Thus for closed population, we have the following reciprocity relation

$$C_{i,j}N_i = C_{j,i}N_j,$$

where N_i is the population in group i. By dividing the places of contact we obtain the following model of social contact

$$C_{i,j} = C^H_{i,j} + C^S_{i,j} + C^W_{i,j} + C^{SC}_{i,j} + C^O_{i,j}$$

where $C^L_{i,j}$ is the number of contacts of a susceptible individual in age group i with a susceptible individual in age group j in a place L (Home, School, Work,Shopping Center, Other).
A database of contact matrix from eight European countries was the subject of a study projected in many other countries, see [19] for more details.

Unfortunately Mayotte island is not referenced in this study which provides the contact matrix for several countries. But it is possible to use the contact matrix of an island called Sao Tome and Principe, an African island country located in the Gulf of Guinea which structural population model is closed to the french ovearseas island Mayotte as we can see in the following population demographic.

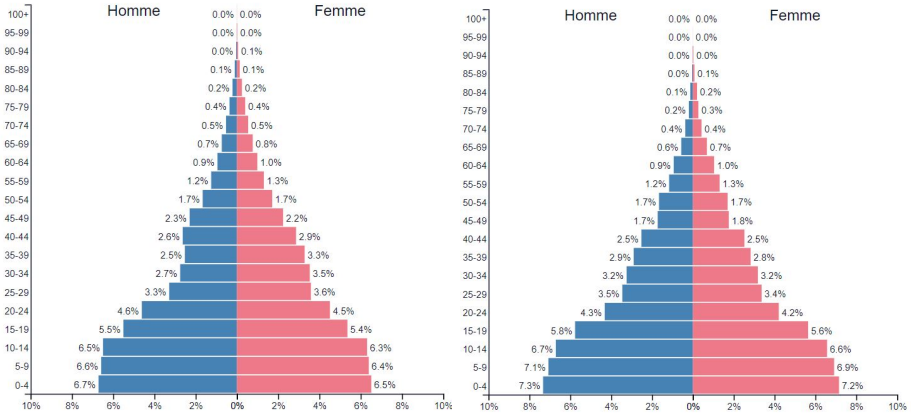


Figure 28: Age pyramid of Mayotte (left) and Sao Tome and Principe (right) in 2020. Source (Wikipedia)

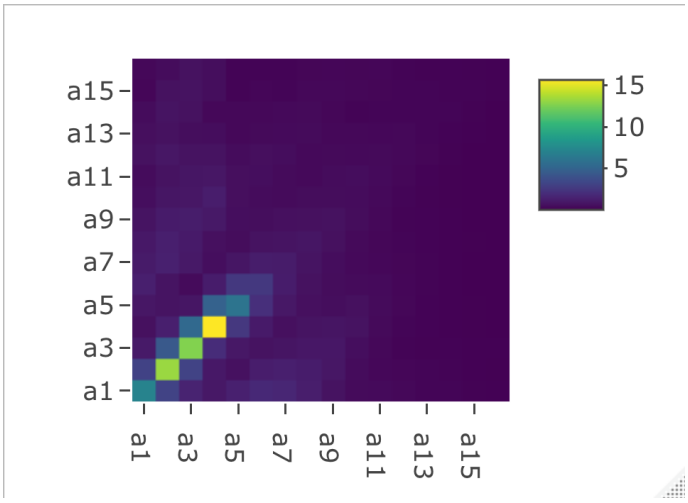


Figure 29: Contact matrix of Sao Tome and Principe

5.1 Simulation based on a simple age-structured contact SEIRD model

The differential equations that govern the trajectories of our model are formulated as follow:

$$\begin{cases} dS_i(t) = \mu N_i - \beta(t) S_i(t) \sum_{j=1}^{n} C_{i,j} \, \frac{I_j}{N_j} dt - \mu S_i(t) \\ dE_i(t) = \beta(t) S_i(t) \sum_{j=1}^{n} C_{i,j} \, \frac{I_j}{N_j} dt - (\lambda + \mu) E_i(t) dt, \\ dI_i(t) = \lambda E_i(t) dt - (\gamma_j + \alpha + \mu) I_i(t) dt \\ dR_i(t) = \gamma_j I_i(t) - \mu R_i(t) dt \\ dD_i(t) = (\alpha_j + \mu) I_i(t) dt \end{cases} \tag{8}$$

with

$$S(t) = \sum_{i=1}^{n} S_i(t), \quad I(t) = \sum_{i=1}^{n} I_i(t), \quad E(t) = \sum_{i=1}^{n} E_i(t), \quad R(t) = \sum_{i=1}^{n} R_i(t)$$

$$D(t) = \sum_{i=1}^{n} D_i(t),$$

β is the rate of infection on contact assumed intrinsic to the pathogen. We assume that all groups will experience background death from other causes at a rate μ. Since we consider closed populations, we also assume that background deaths are balanced by births into S at a rate μN. The induced mortality rate α_j depend on the age as well as the infectious rate γ_j. For simulation purpose we choose $n = 16$ to correspond to the 16 age group into which the contact matrix data is partitioned. Then the $5 * n = 80$ ordinary differential equations are then numerically integrated using R-software. The following table lists the parameters of the model we'll use for numerical simulation. These estimates are often wide due to wide variations from one infection to another. We also give the initial values of the model.

Table 3: List of the parameters and initial values of the model

Description	Value
Incubation rate λ	section 3.3 [11]
Infectious period by age : $(\gamma_1^{-1}, ..., \gamma_{16}^{-1})$	$(10, 10, 10, 10, 10, 10, 10, 14, 14, 14, 14, 21, 21, 28, 28, 28)$
$\beta(t)$: temporal transmission rates	Estimated by formula (3)
Induced mortality rate by age $(\alpha_1, ..., \alpha_{16})$	$(0, 0, 0, 0, 0, 0, 10^{-3}, 2.10^{-3}, 2.10^{-3}, 2.10^{-3}, 3.10^{-3}, 3.10^{-3}, 4.10^{-3}, 4.10^{-3}, 5.10^{-3}, 5.10^{-3})$
$\mu = 0.003$: natural mortality rate in 2019	`https://www.insee.fr/fr/statistiques/5424581`
C : contact matrix	Sao-Tome-Principe [19]
$E(0)$: average of all contact cases at the begining 15	ARS-Mayotte 2020
$(N_1, ...N_{16})$	`https://www.insee.fr/fr/statistiques/1893198`
$(S_1(0), ..., S_1(0))$	$(N_1 - I_1(0) - E_1(0), ..., N_{16} - I_{16}(0) - E_{16}(0))$
$(I_1(0), ..., I_{16}(0))$	$(0, 0, 0, 0, 0, 1, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0)$
$(E_1(0), ..., E_{16}(0))$	$(0, 0, 0, 1, 1, 1, 2, 2, 2, 3, 1, 1, 0, 1, 0, 0)$
$(R_1(0), ..., R_{16}(0))$	$(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$
$(D_1(0), ..., D_{16}(0))$	$(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$

The following pictures show by simulation the progress of the epidemic by age. By observing the curve of infected by age group, we observe that the epidemic peak is greater in the age group 1 to 4 corresponding to the age class $[0 - 19]$, followed by age group 5 to 9 ($[20 - 44]$), then age group 10 to 12 ($[45 - 59]$) and then age group 13 to 16 ($[60+]$).In this last class the age group $[60 - 64]$ is more affected than the class 65+. This shows that there is a sensitivity according to the age group. The age group $[0 - 19]$ is more exposed which means that it is a population to be protected by various means such as vaccination or isolation.

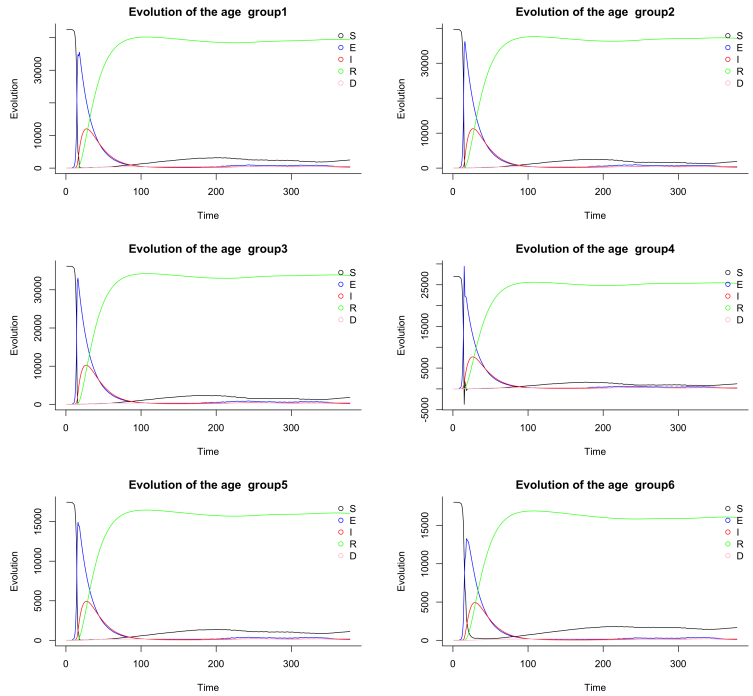


Figure 30: Simulating the age-structured contact model

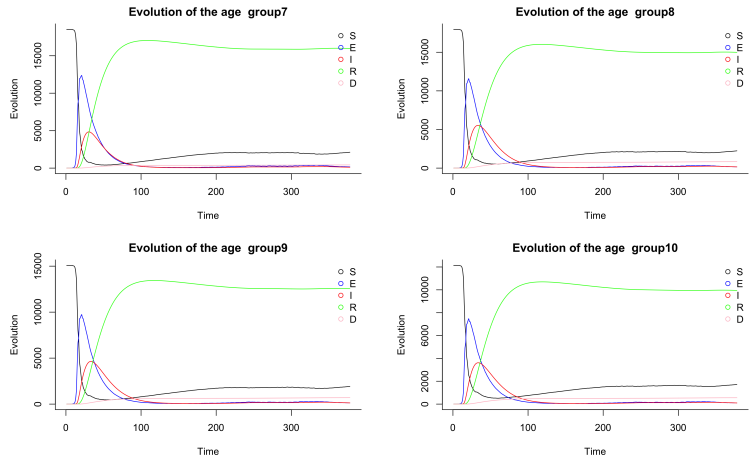


Figure 31: Simulating the age-structured contact model

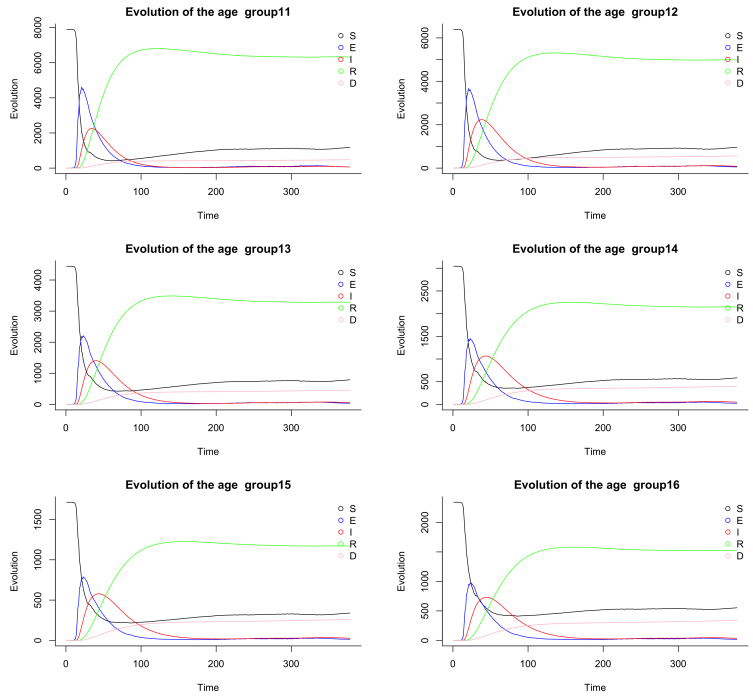


Figure 32: Simulating the age-structured contact model

6 Discussion and conclusion

In the above simulation, we did not integrate the vaccination because it was assumed that during the study interval of the epidemic curve, vaccination in Mayotte had no effect because it started in February 2021, see `https://www.mayotte.ars.sante.fr`. Given the current state of the epidemic in Mayotte, the above numerical simulation of the epidemic curve show that, new factors have contributed to modify the dynamics of the epidemic. It is important to note that the simulation is only theoretical and restricted by given conditions. A conclusion that can be drawn subject to assumptions is that mitigation and non-pharmacological interventions of the Covid-19 of any kind can possibly be considered in the model such as the factors of reinfection, immunity and the appearance of variants. Note also that some conditions can be dramatically different in reality, for instance the infectious and the virus mortality rates depending strongly on the age in this work. It should also be noted a bias in the approximate use of the contact matrix.

7 Supporting information

Additional supporting information of the original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s. Other data about the population used in this study are openly available from the French INSEE web site. Note that data relating to covid-19 can be downloaded from `https://www.data.gouv.fr/fr/pages/donnees-sante/`. As concerned the data on contact matrix, they are available in the supplementary material of [19].

References

[1] Tatiana Benaglia, Didier Chauveau, David R Hunter, and Derek S Young. mixtools: an r package for analyzing mixture models. *Journal of statistical software*, 32:1–29, 2010.

[2] Pierre-Yves Boelle, Thomas Obadia, and Maintainer Thomas Obadia. Package 'r0'. 2015.

[3] Odo Diekmann, Johan Andre Peter Heesterbeek, and Johan AJ Metz. On the definition and the computation of the basic reproduction ratio r 0 in models for infectious diseases in heterogeneous populations. *Journal of mathematical biology*, 28(4):365–382, 1990.

[4] Leonhard Euler. Recherches générales sur la mortalité et la multiplication du genre humain. *Mémoires de l'académie des sciences de Berlin*, pages 144–164, 1767.

[5] Raphaël Forien, Guodong Pang, and Étienne Pardoux. Epidemic models with varying infectivity. *SIAM Journal on Applied Mathematics*, 81(5):1893–1930, 2021.

[6] Raphaël Forien, Guodong Pang, and Étienne Pardoux. Estimating the state of the covid-19 epidemic in france using a model with memory. *Royal Society open science*, 8(3):202327, 2021.

[7] Xi He, Eric HY Lau, Peng Wu, Xilong Deng, Jian Wang, Xinxin Hao, Yiu Chung Lau, Jessica Y Wong, Yujuan Guan, Xinghua Tan, et al. Temporal dynamics in viral shedding and transmissibility of covid-19. *Nature medicine*, 26(5):672–675, 2020.

[8] William Ogilvy Kermack and Anderson G McKendrick. 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*, 115(772):700–721, 1927.

[9] Adam J Kucharski, Timothy W Russell, Charlie Diamond, Yang Liu, John Edmunds, Sebastian Funk, Rosalind M Eggo, Fiona Sun, Mark Jit, James D Munday, et al. Early dynamics of transmission and control of covid-19: a mathematical modelling study. *The lancet infectious diseases*, 20(5):553–558, 2020.

[10] Sonja Lehtinen, Peter Ashcroft, and Sebastian Bonhoeffer. On the relationship between serial interval, infectiousness profile and generation time. *Journal of the Royal Society Interface*, 18(174):20200756, 2021.

[11] Qun Li, Xuhua Guan, Peng Wu, Xiaoye Wang, Lei Zhou, Yeqing Tong, Ruiqi Ren, Kathy SM Leung, Eric HY Lau, Jessica Y Wong, et al. Early transmission dynamics in wuhan, china, of novel coronavirus–infected pneumonia. *New England journal of medicine*, 2020.

[12] Zhihua Liu, Pierre Magal, Ousmane Seydi, and Glenn Webb. A covid-19 epidemic model with latency period. *Infectious Disease Modelling*, 5:323–337, 2020.

[13] Zhihua Liu, Pierre Magal, Ousmane Seydi, and Glenn Webb. Predicting the cumulative number of cases for the covid-19 epidemic in china from early data. *arXiv preprint arXiv:2002.12298*, 2020.

[14] Alfred J Lotka. Relation between birth rates and death rates. *Science*, 26(653):21–22, 1907.

[15] Geoffrey J McLachlan, Sharon X Lee, and Suren I Rathnayake. Finite mixture models. *Annual review of statistics and its application*, 6:355–378, 2019.

[16] Hiroshi Nishiura, Natalie M Linton, and Andrei R Akhmetzhanov. Serial interval of novel coronavirus (covid-19) infections. *International journal of infectious diseases*, 93:284–286, 2020.

[17] Thomas Obadia, Romana Haneef, and Pierre-Yves Boëlle. The r0 package: a toolbox to estimate reproduction numbers for epidemic outbreaks. *BMC medical informatics and decision making*, 12(1):1–9, 2012.

[18] Thomas Obadia, Romana Haneef, and Pierre-Yves Boëlle. The r0 package: a toolbox to estimate reproduction numbers for epidemic outbreaks. *BMC medical informatics and decision making*, 12(1):1–9, 2012.

[19] Kiesha Prem, Kevin van Zandvoort, Petra Klepac, Rosalind M Eggo, Nicholas G Davies, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Alex R Cook, and Mark

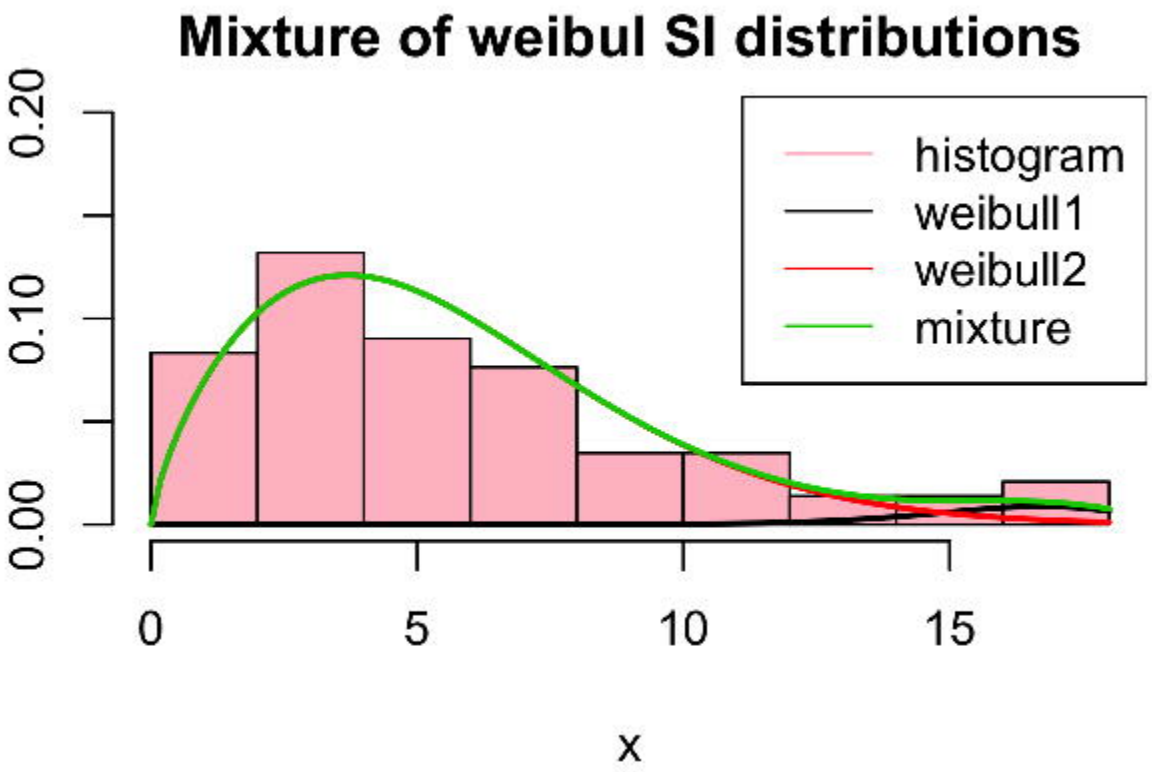
Jit. Projecting contact matrices in 177 geographical regions: an update and comparison with empirical data for the covid-19 era. *PLoS computational biology*, 17(7):e1009098, 2021.

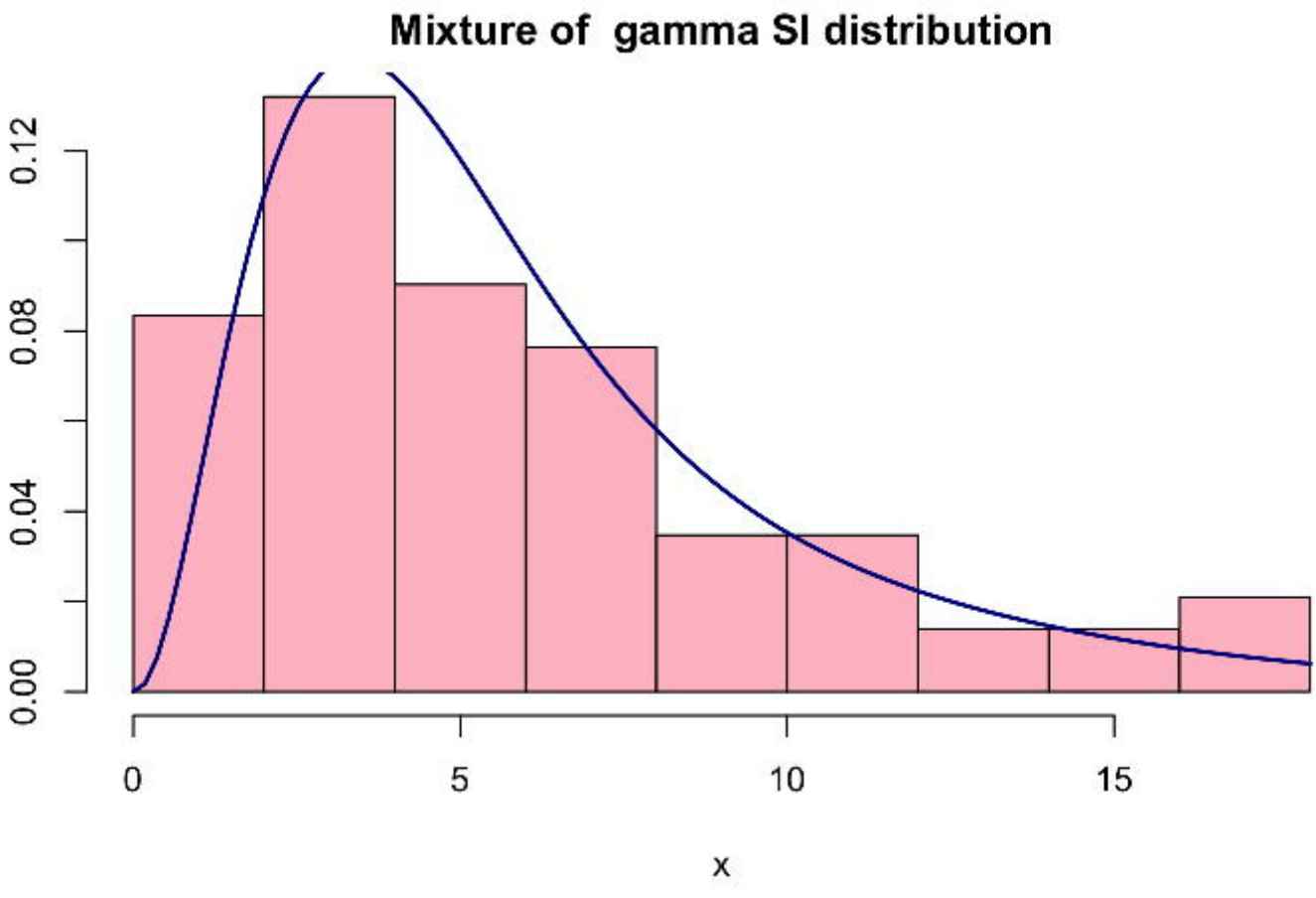
[20] Åke Svensson. A note on generation times in epidemic models. *Mathematical biosciences*, 208(1):300–311, 2007.

[21] Jacco Wallinga and Marc Lipsitch. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological Sciences*, 274(1609):599–604, 2007.

[22] Jacco Wallinga and Marc Lipsitch. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological Sciences*, 274(1609):599–604, 2007.

[23] Laura Forsberg White, Jacco Wallinga, Lyn Finelli, Carrie Reed, Steven Riley, Marc Lipsitch, and Marcello Pagano. Estimation of the reproductive number and the serial interval in early phase of the 2009 influenza a/h1n1 pandemic in the usa. *Influenza and other respiratory viruses*, 3(6):267–276, 2009.

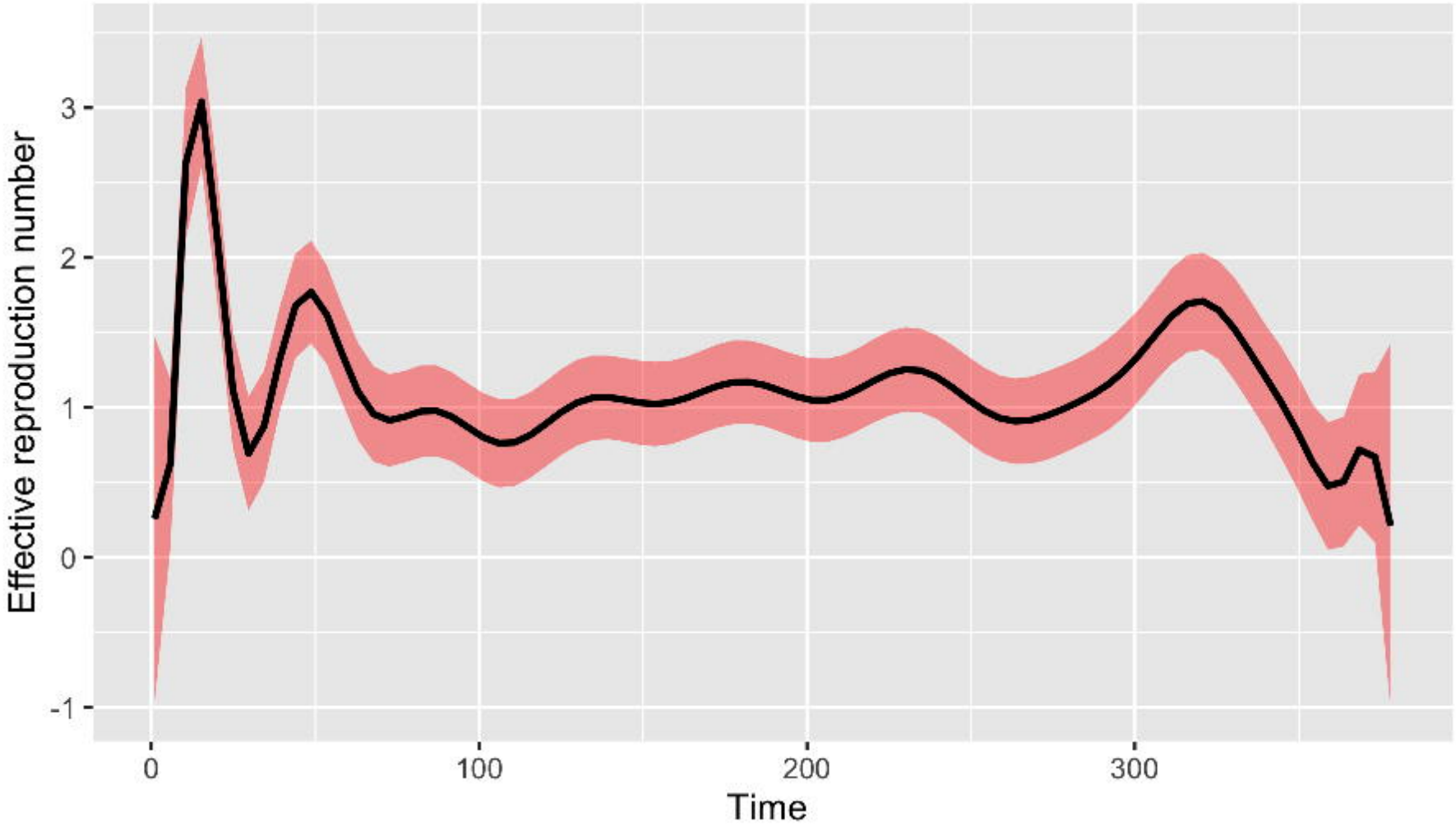

Mixture of weibul SI distributions
0.20
0.10
0.00
histogram
weibull1
weibull2
mixture
0
5
10
15
x

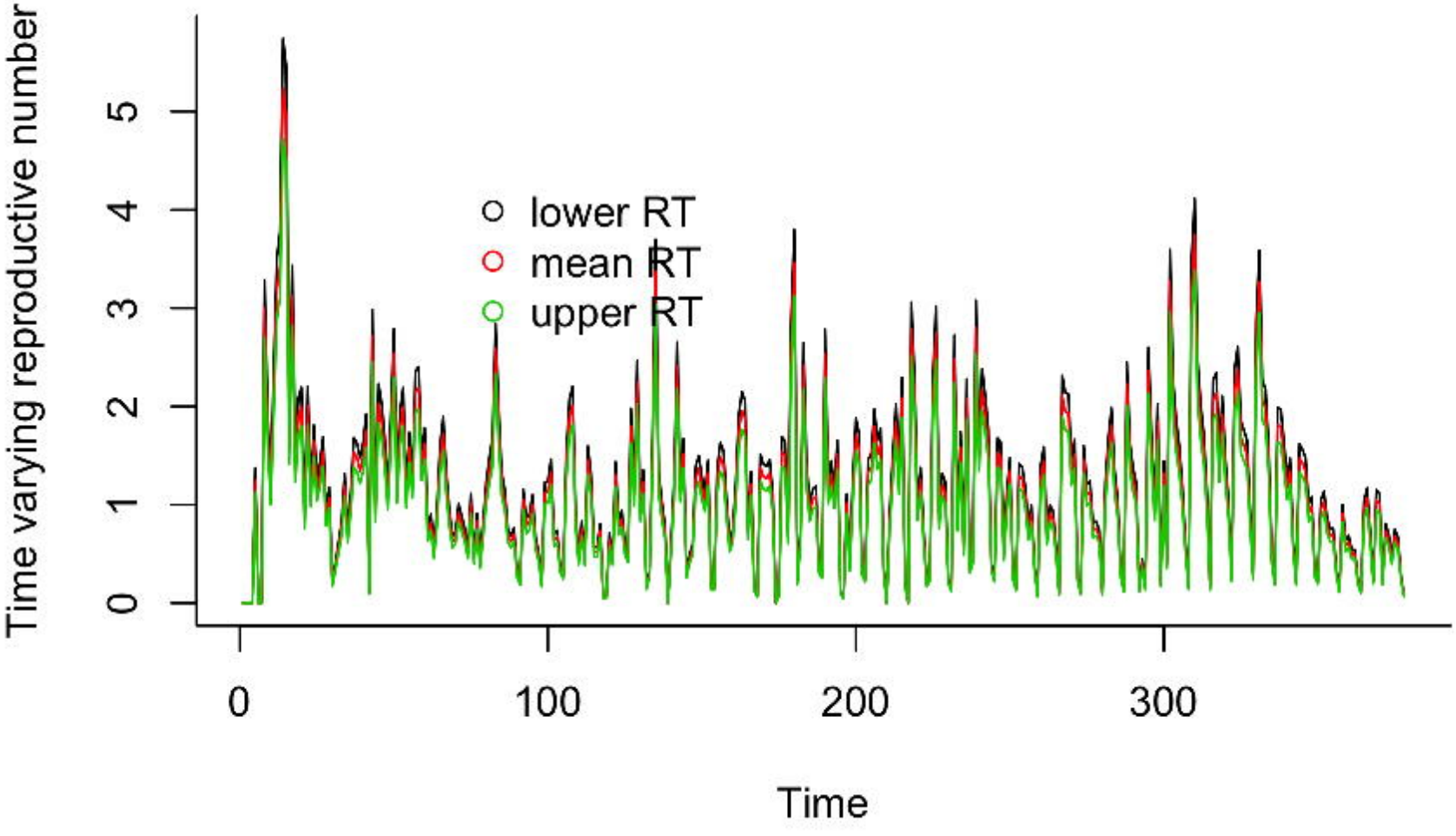

Mixture of gamma SI distribution
0.00
0.04
0.08
0.12
0
5
10
15
x

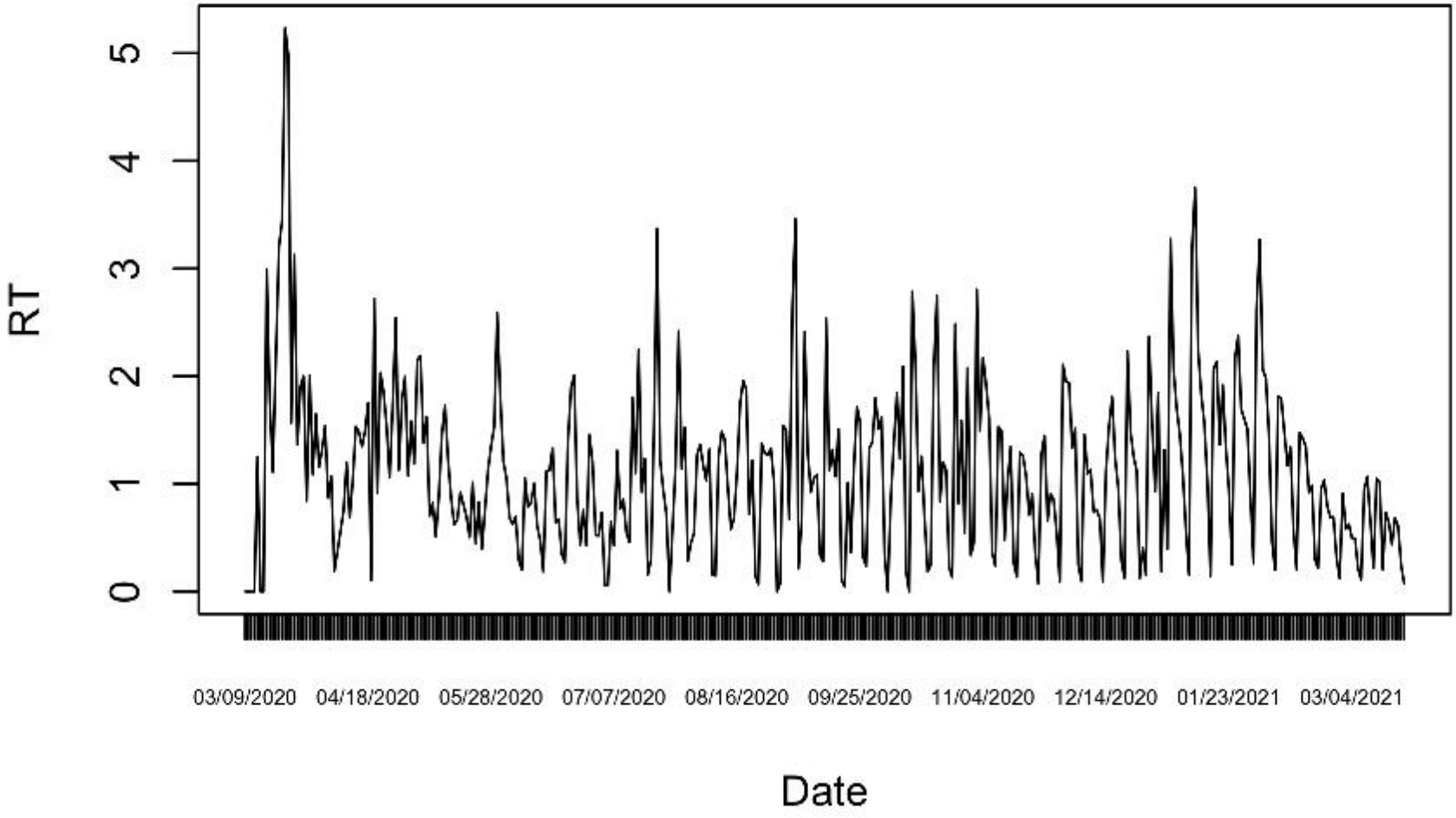
```
                                                        est_gamma
par         0.4668812, 3.9708818, 1.1336771, 3.8530206, 0.4436036
value                                                    191.5604
counts                                                    502, NA
convergence                                                     1
message                                                      NULL
                                        est_exp
par         0.4000000, 0.1589404, 0.1589404
value                                  204.4243
counts                                   40, NA
convergence                                   0
message                                     NULL
                                                      est_lognorm
par         0.3862606, 1.0274504, 0.5477013, 1.9496451, 0.5422352
value                                                    191.9062
counts                                                    502, NA
convergence                                                     1
message                                                      NULL
                                                         est_weib
par         0.04066447, 9.96140450, 16.90693027, 1.70447135, 6.20140484
value                                                      191.6061
counts                                                      502, NA
convergence                                                       1
message                                                        NULL
                                                     est_exp_conv
par         0.3121223, 0.3181467, 0.3177062, 0.3180879, 0.3176683
value                                                    193.2564
counts                                                    134, NA
convergence                                                     0
message                                                      NULL
```

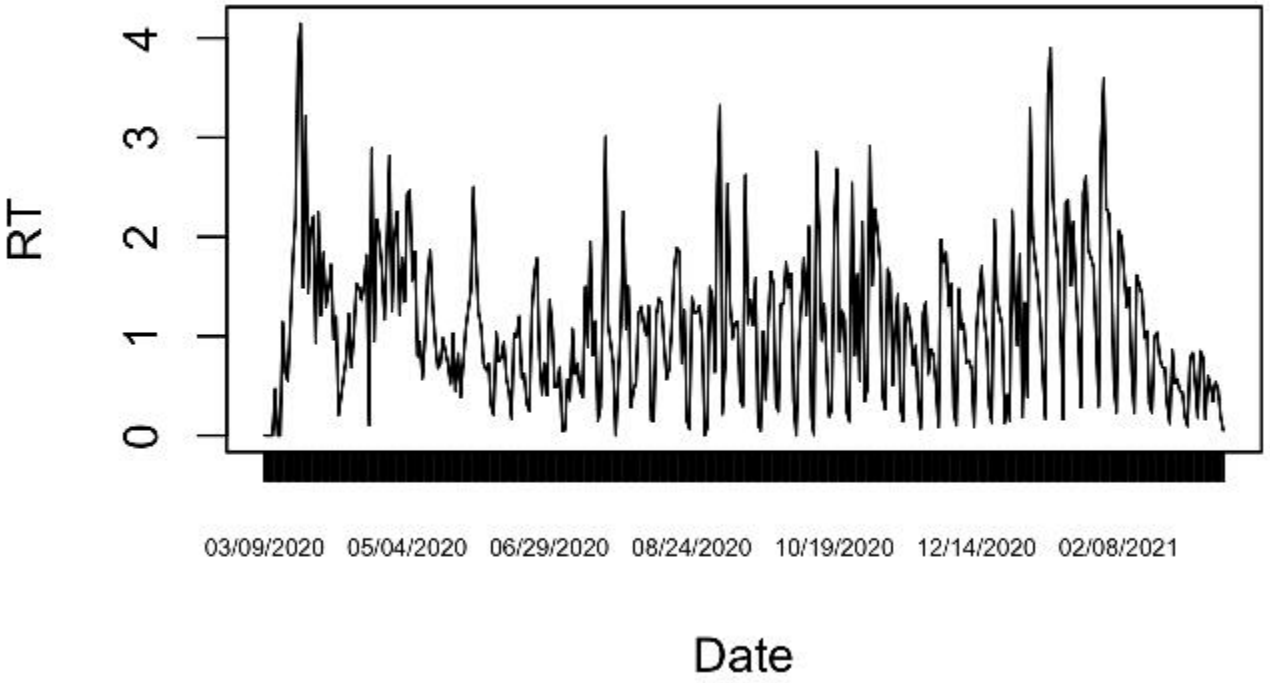
	gamma	exp	lognorm
minimum	192.8342	209.8803	192.5157
estimate	2.3157219, 0.3680613	0.1655624	1.608040, 0.702471
gradient	-9.941429e-07, 2.529532e-06	-1.98952e-06	-1.558913e-05, 6.264145e-05
code	1	1	1
iterations	18	4	8

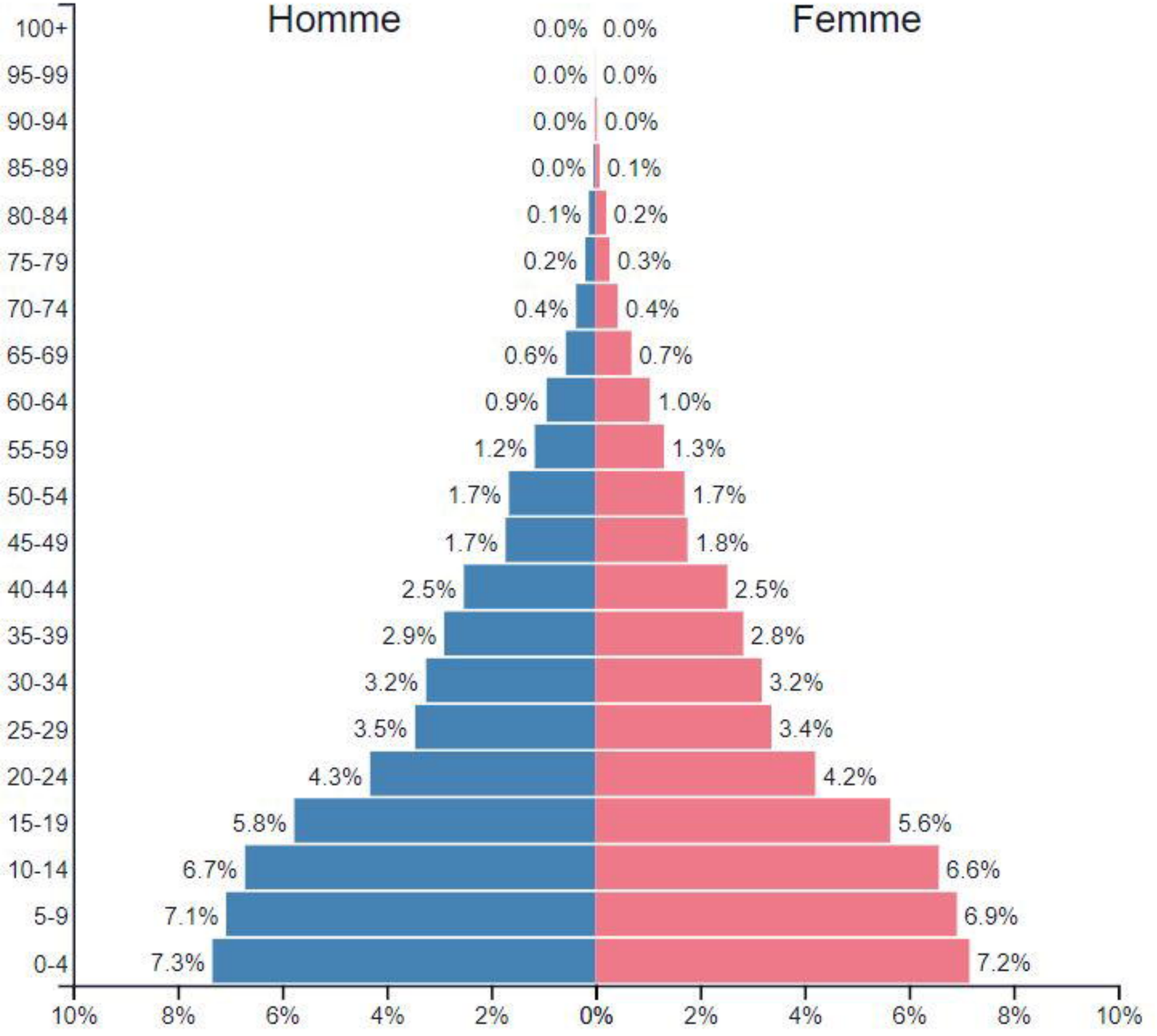
	weibull	exp_conv
minimum	194.3016	1.797693e+308
estimate	1.568012, 7.047558	-0.5877867, -1.1314021
gradient	1.486328e-06, 1.935766e-07	0, 0
code	1	1
iterations	20	0

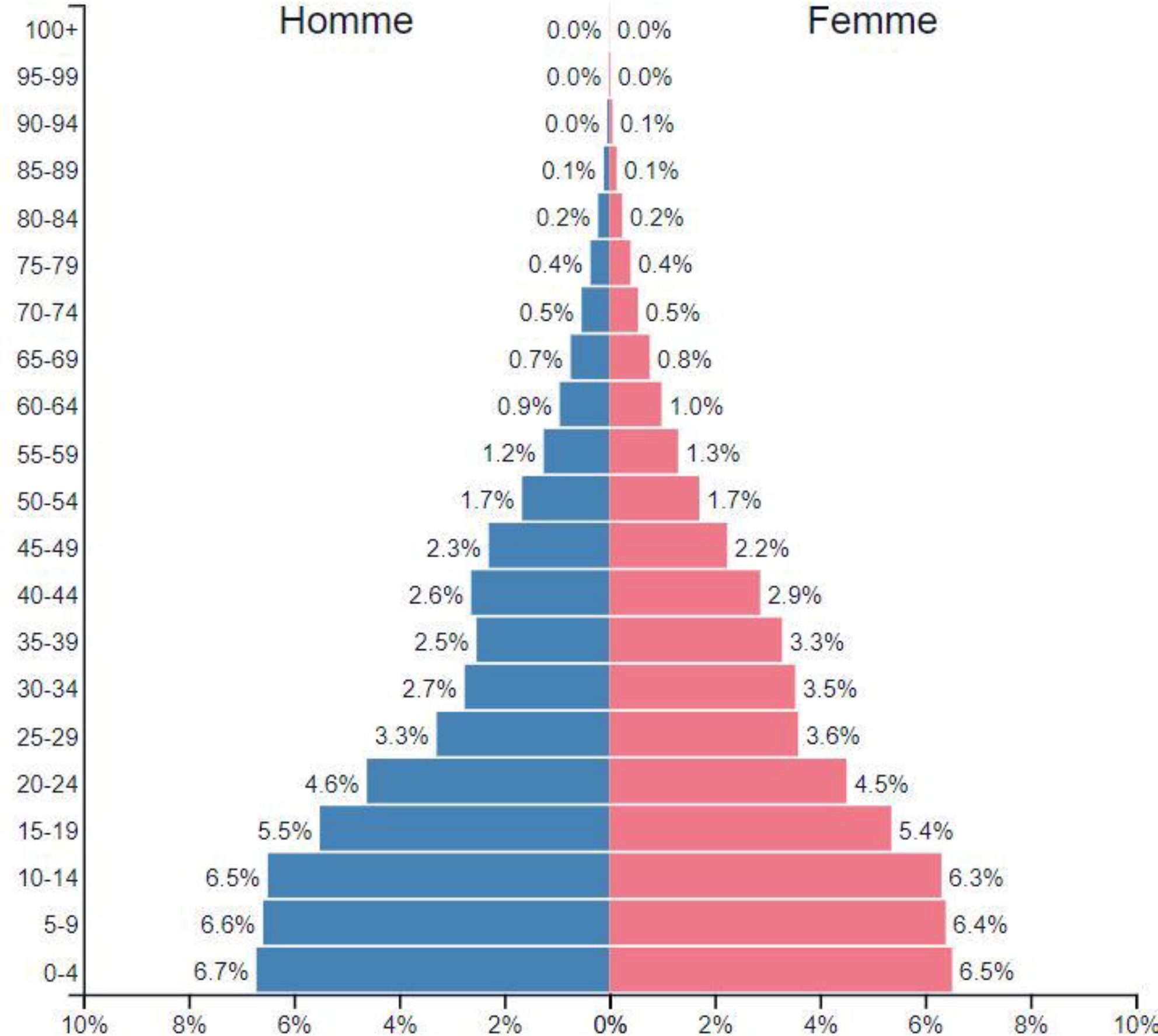

Effective reproduction number
3
2
1
0
-1
0
100
200
300
Time

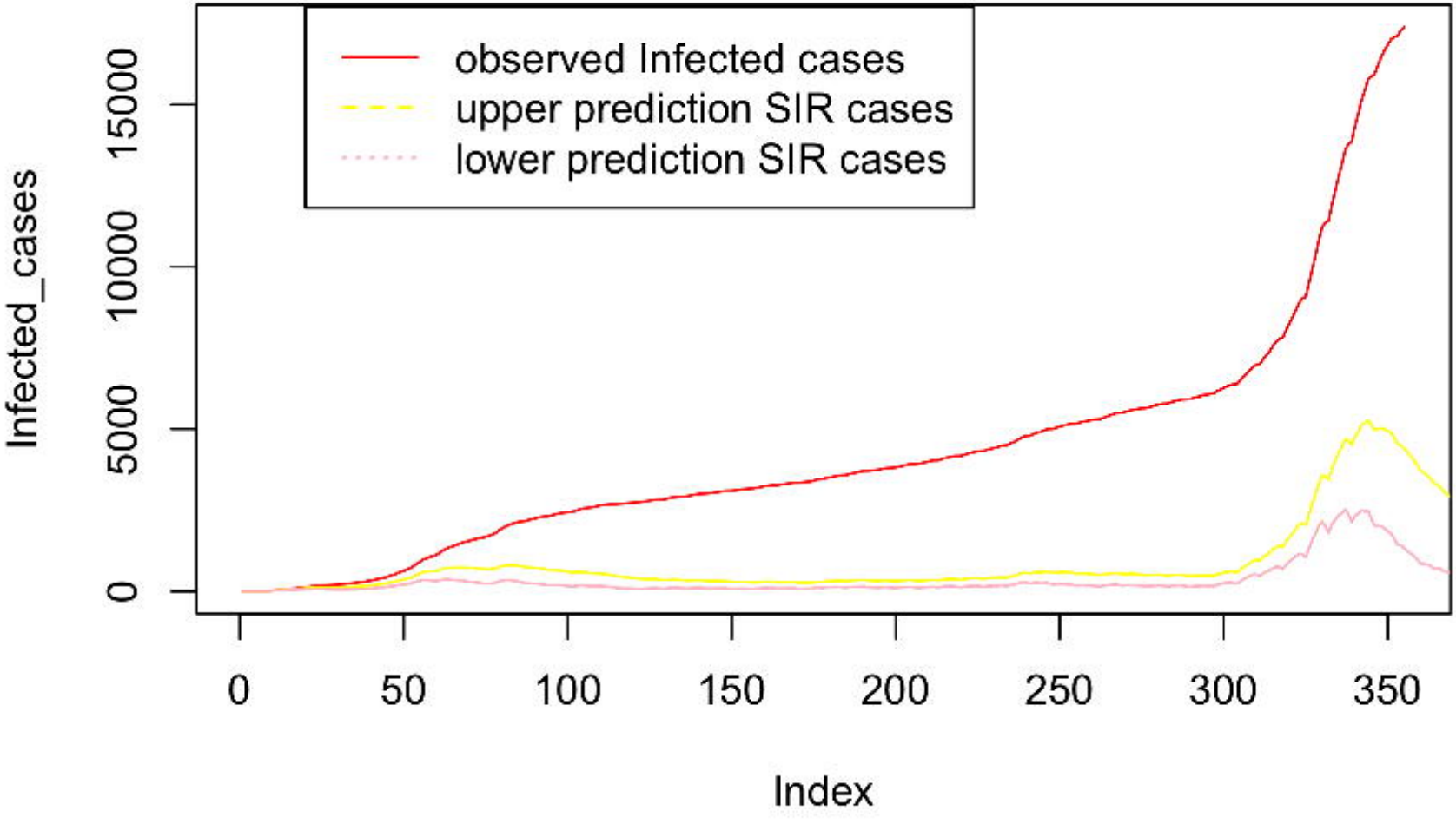

Time varying reproductive number
0
1
2
3
4
5
lower RT
mean RT
upper RT
0
100
200
300
Time

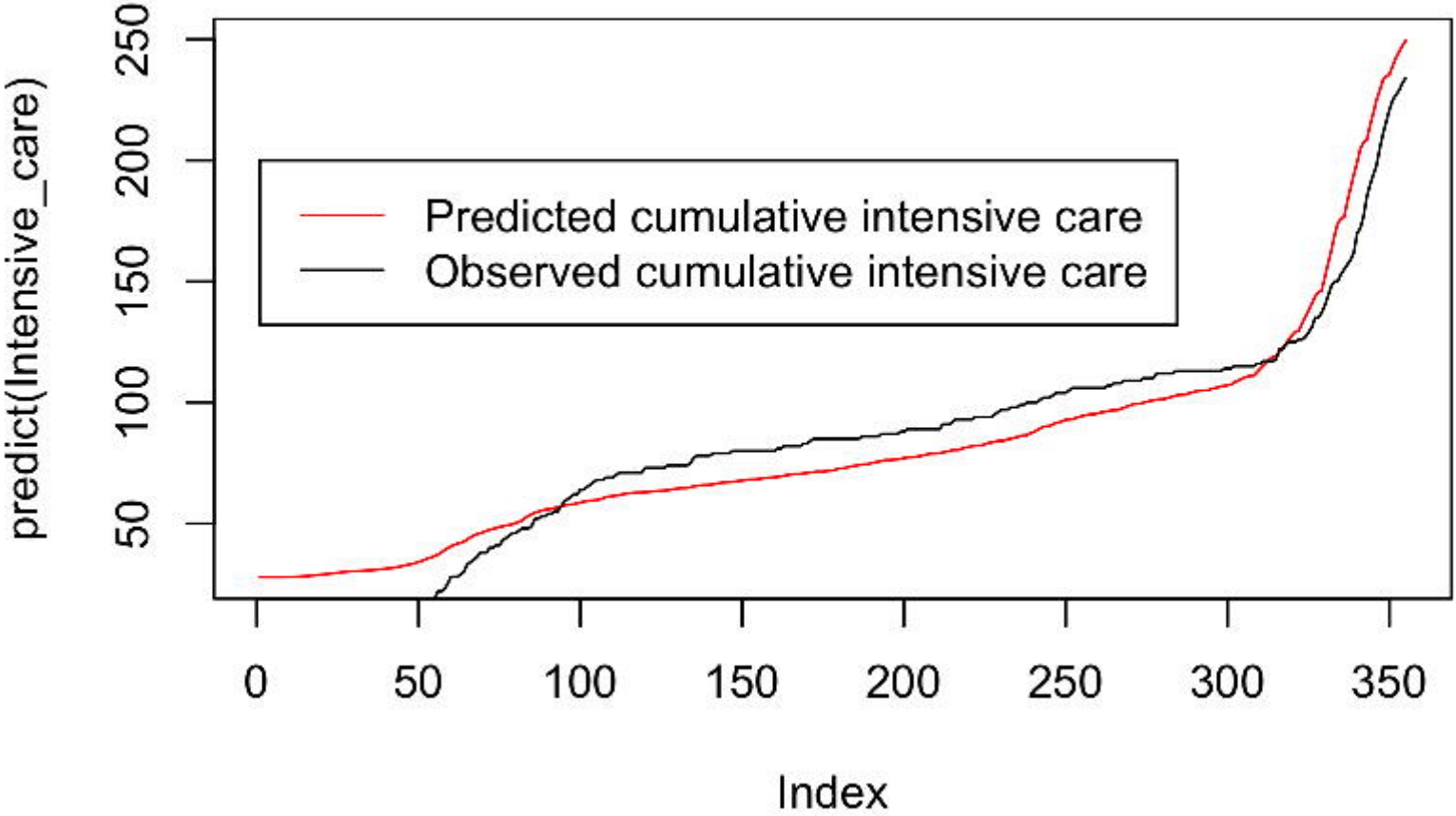

RT
0
1
2
3
4
5
03/09/2020
04/18/2020
05/28/2020
07/07/2020
08/16/2020
09/25/2020
11/04/2020
12/14/2020
01/23/2021
03/04/2021
Date

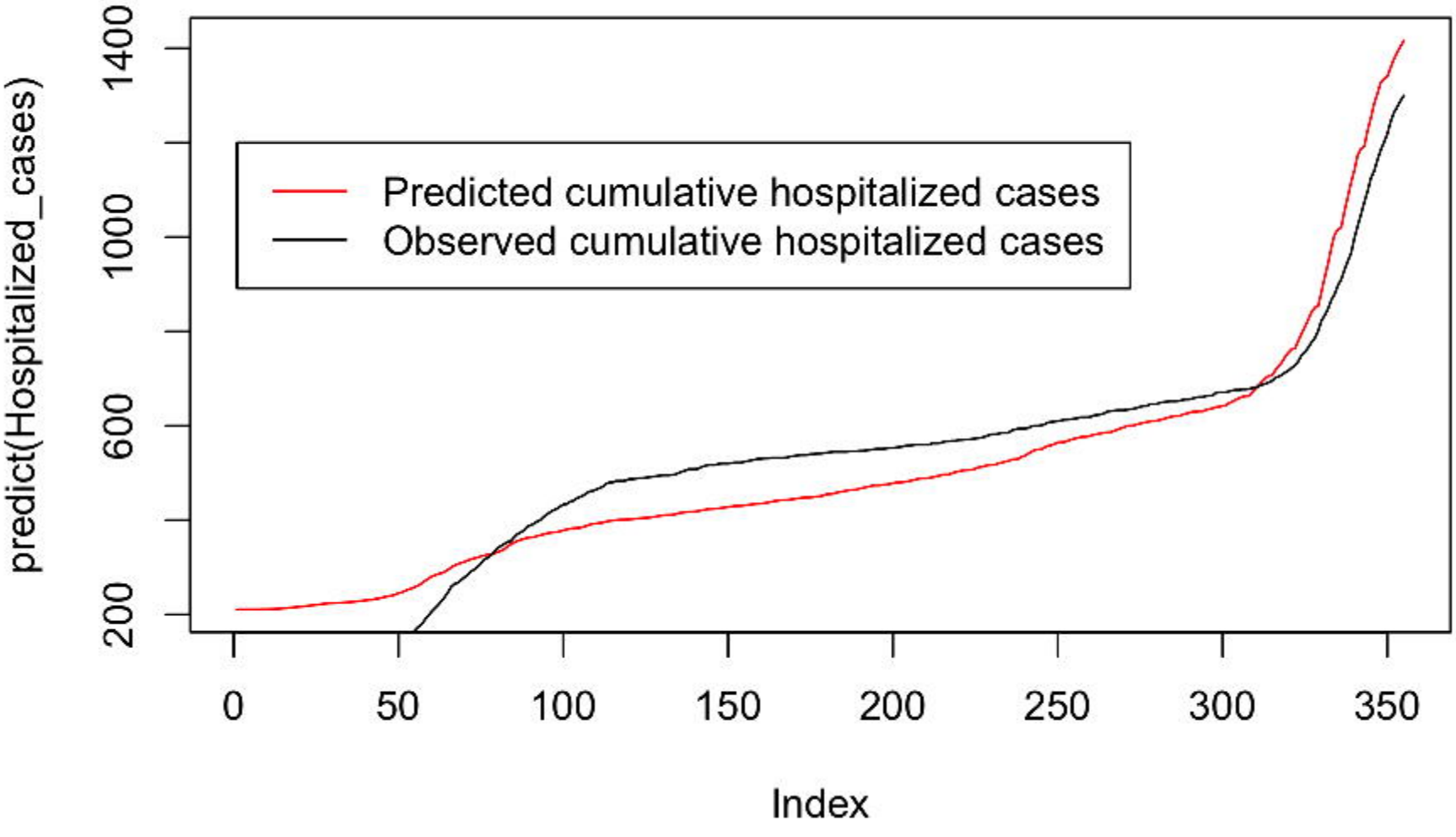

RT
0
1
2
3
4
03/09/2020
05/04/2020
06/29/2020
08/24/2020
10/19/2020
12/14/2020
02/08/2021
Date

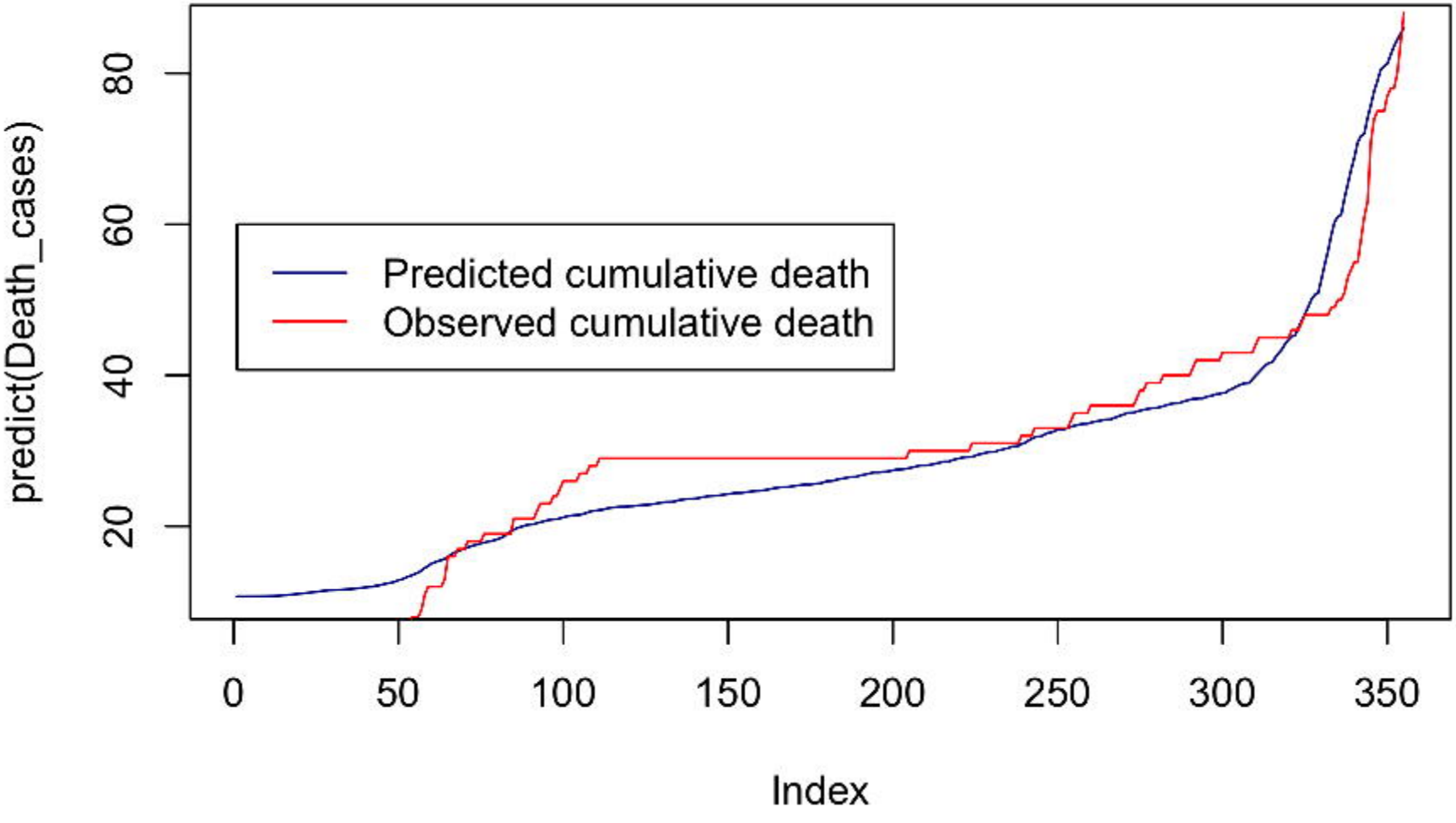

Homme
Femme
100+ 0.0% 0.0%
95-99 0.0% 0.0%
90-94 0.0% 0.0%
85-89 0.0% 0.1%
80-84 0.1% 0.2%
75-79 0.2% 0.3%
70-74 0.4% 0.4%
65-69 0.6% 0.7%
60-64 0.9% 1.0%
55-59 1.2% 1.3%
50-54 1.7% 1.7%
45-49 1.7% 1.8%
40-44 2.5% 2.5%
35-39 2.9% 2.8%
30-34 3.2% 3.2%
25-29 3.5% 3.4%
20-24 4.3% 4.2%
15-19 5.8% 5.6%
10-14 6.7% 6.6%
5-9 7.1% 6.9%
0-4 7.3% 7.2%
10% 8% 6% 4% 2% 0% 2% 4% 6% 8% 10%

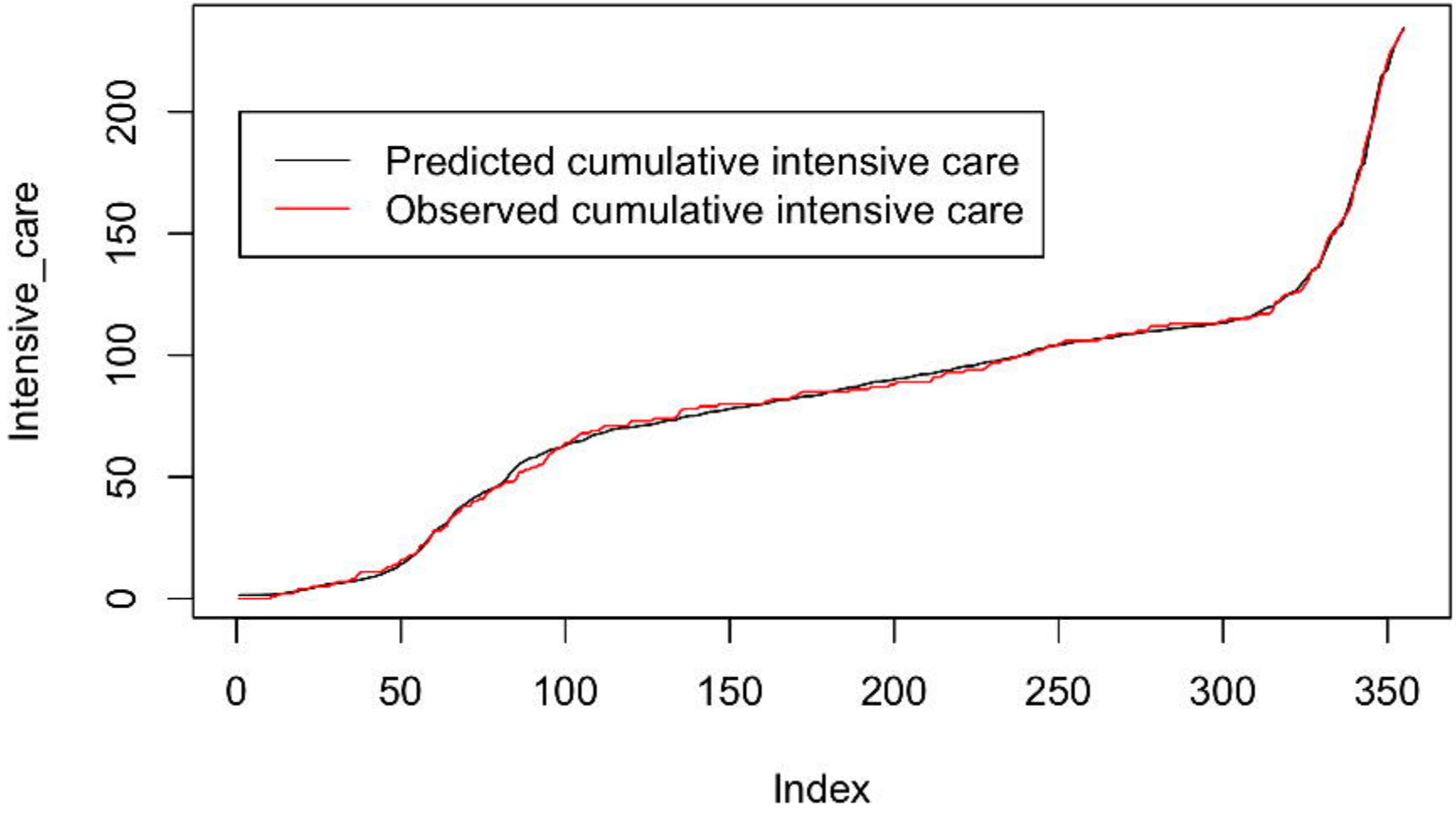

Homme
Femme
100+
95-99
90-94
85-89
80-84
75-79
70-74
65-69
60-64
55-59
50-54
45-49
40-44
35-39
30-34
25-29
20-24
15-19
10-14
5-9
0-4
0.0% 0.0%
0.0% 0.0%
0.0% 0.1%
0.1% 0.1%
0.2% 0.2%
0.4% 0.4%
0.5% 0.5%
0.7% 0.8%
0.9% 1.0%
1.2% 1.3%
1.7% 1.7%
2.3% 2.2%
2.6% 2.9%
2.5% 3.3%
2.7% 3.5%
3.3% 3.6%
4.6% 4.5%
5.5% 5.4%
6.5% 6.3%
6.6% 6.4%
6.7% 6.5%
10% 8% 6% 4% 2% 0% 2% 4% 6% 8% 10%

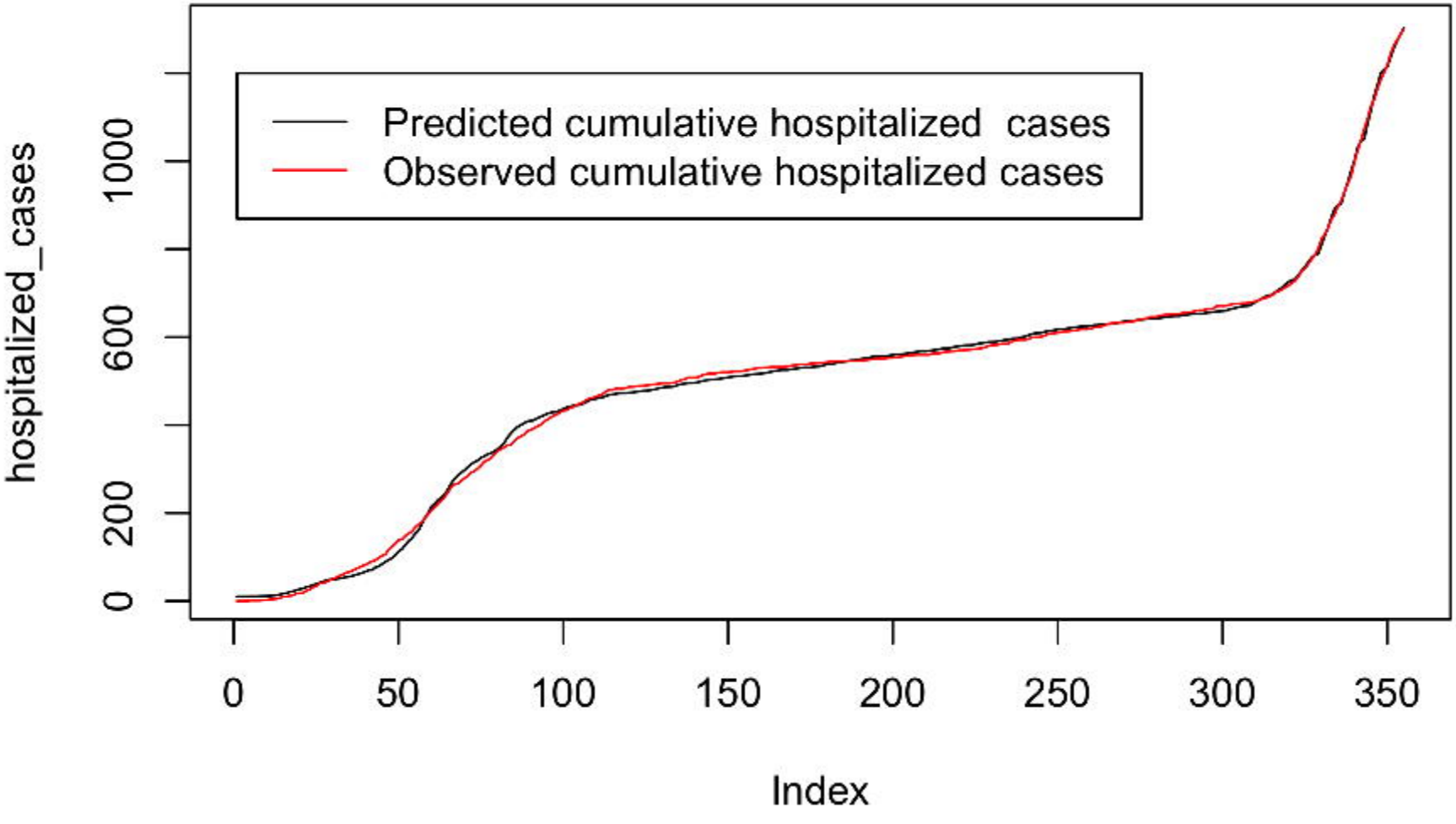

observed Infected cases
upper prediction SIR cases
lower prediction SIR cases
Infected_cases
0
5000
10000
15000
0
50
100
150
200
250
300
350
Index

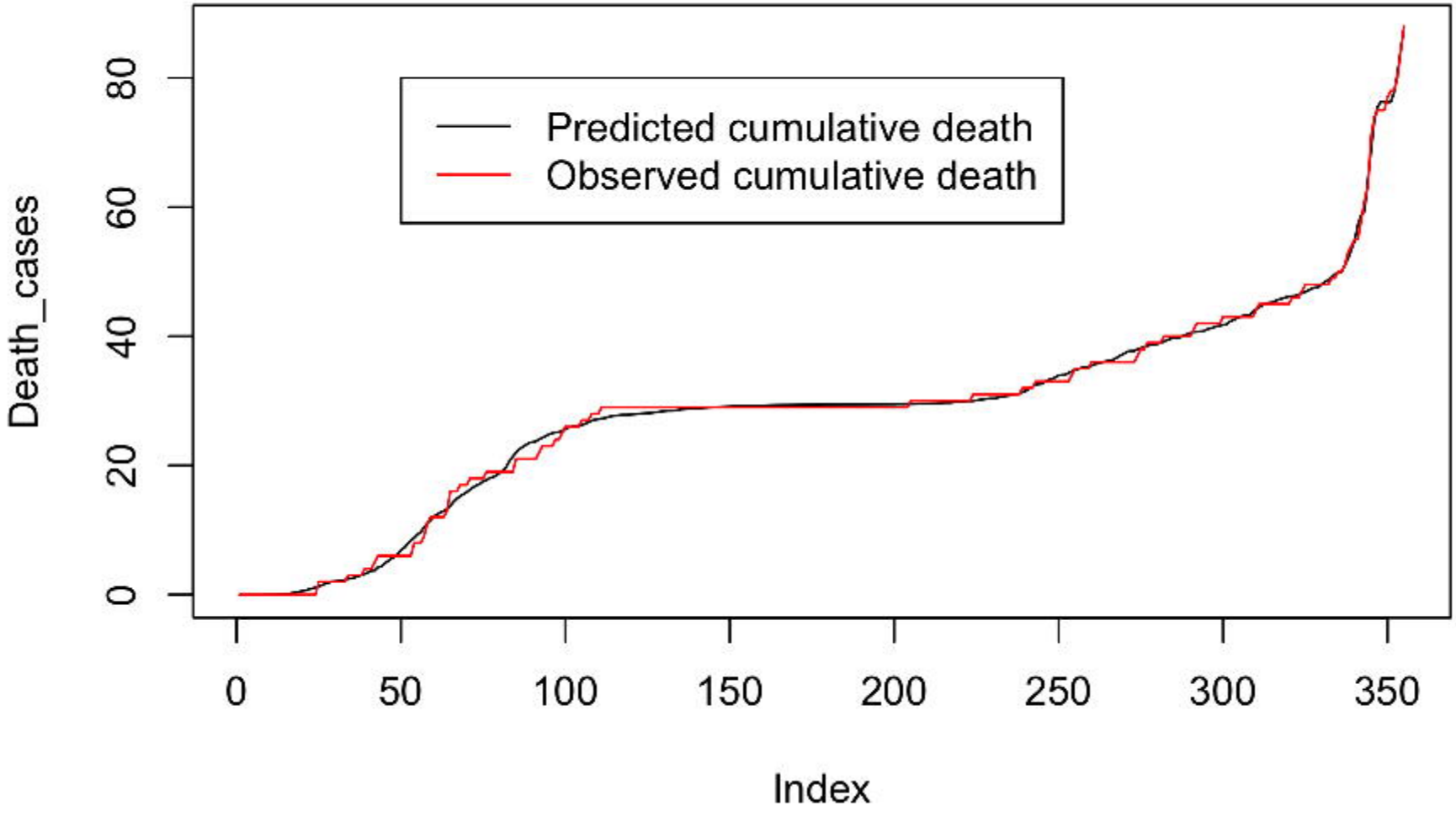

Predicted cumulative intensive care
Observed cumulative intensive care
predict(Intensive_care)
Index
0
50
100
150
200
250
300
350
50
100
150
200
250

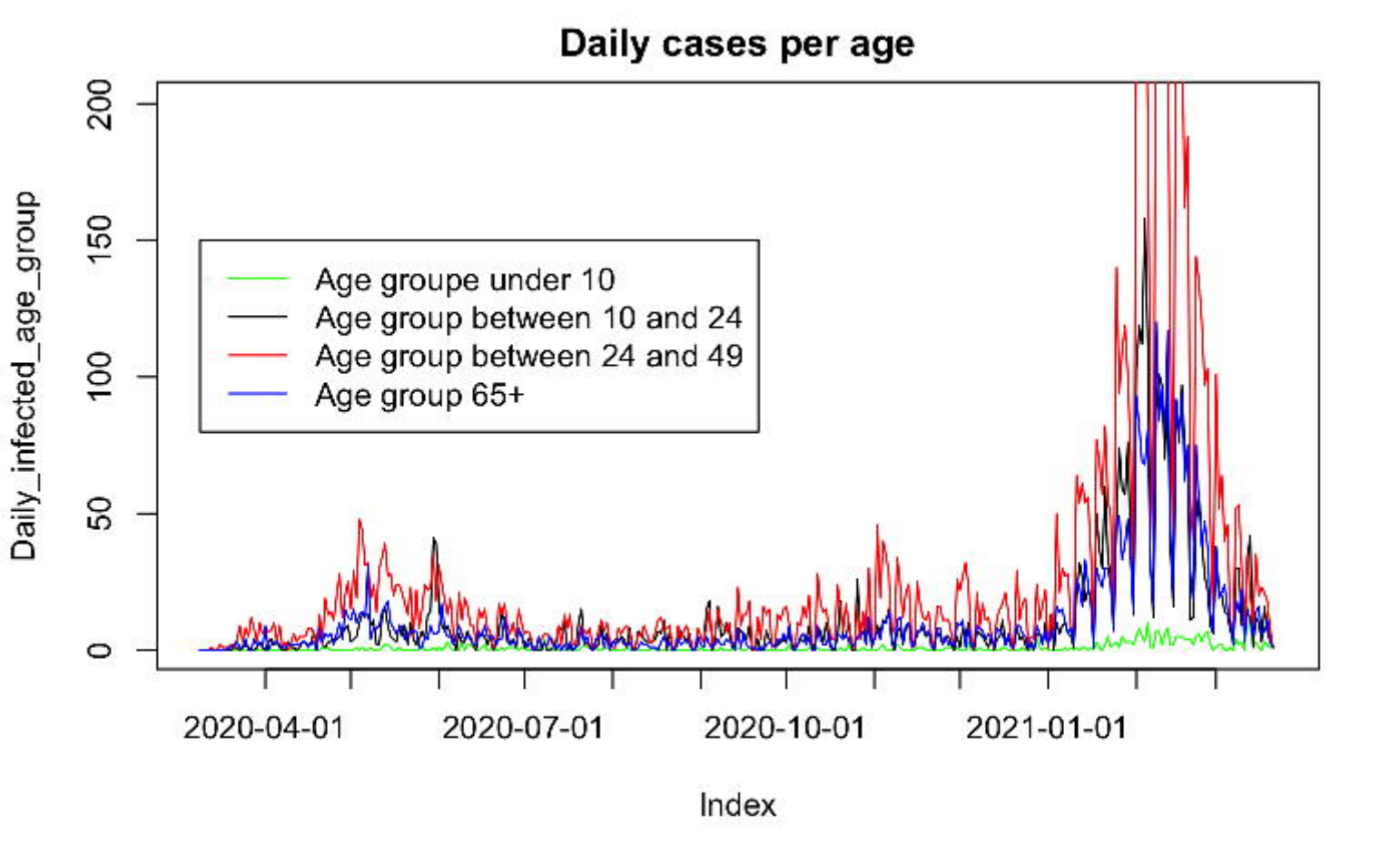

Predicted cumulative hospitalized cases
Observed cumulative hospitalized cases
predict(Hospitalized_cases)
200
600
1000
1400
0
50
100
150
200
250
300
350
Index

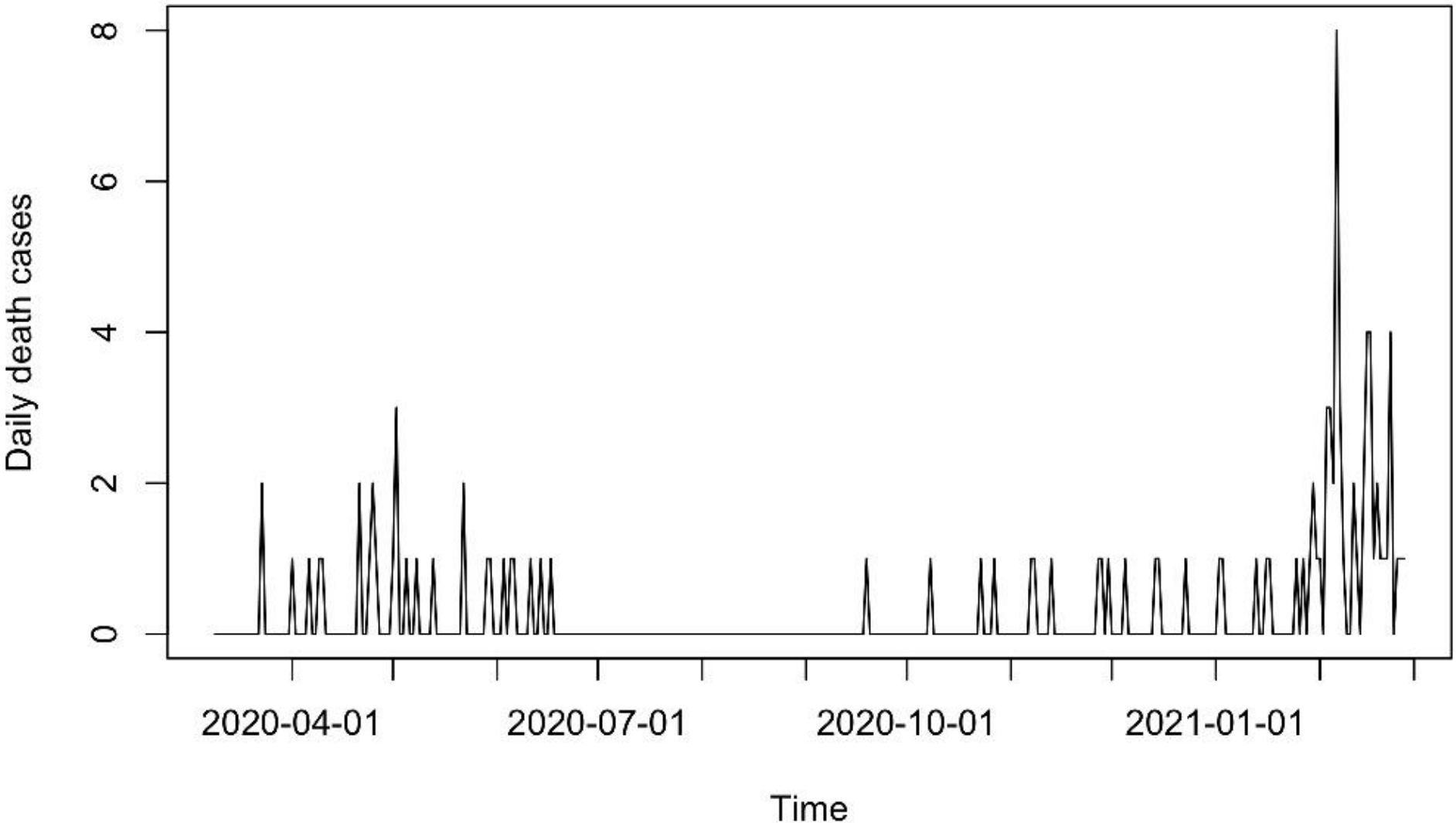

Predicted cumulative death
Observed cumulative death
predict(Death_cases)
20
40
60
80
0
50
100
150
200
250
300
350
Index

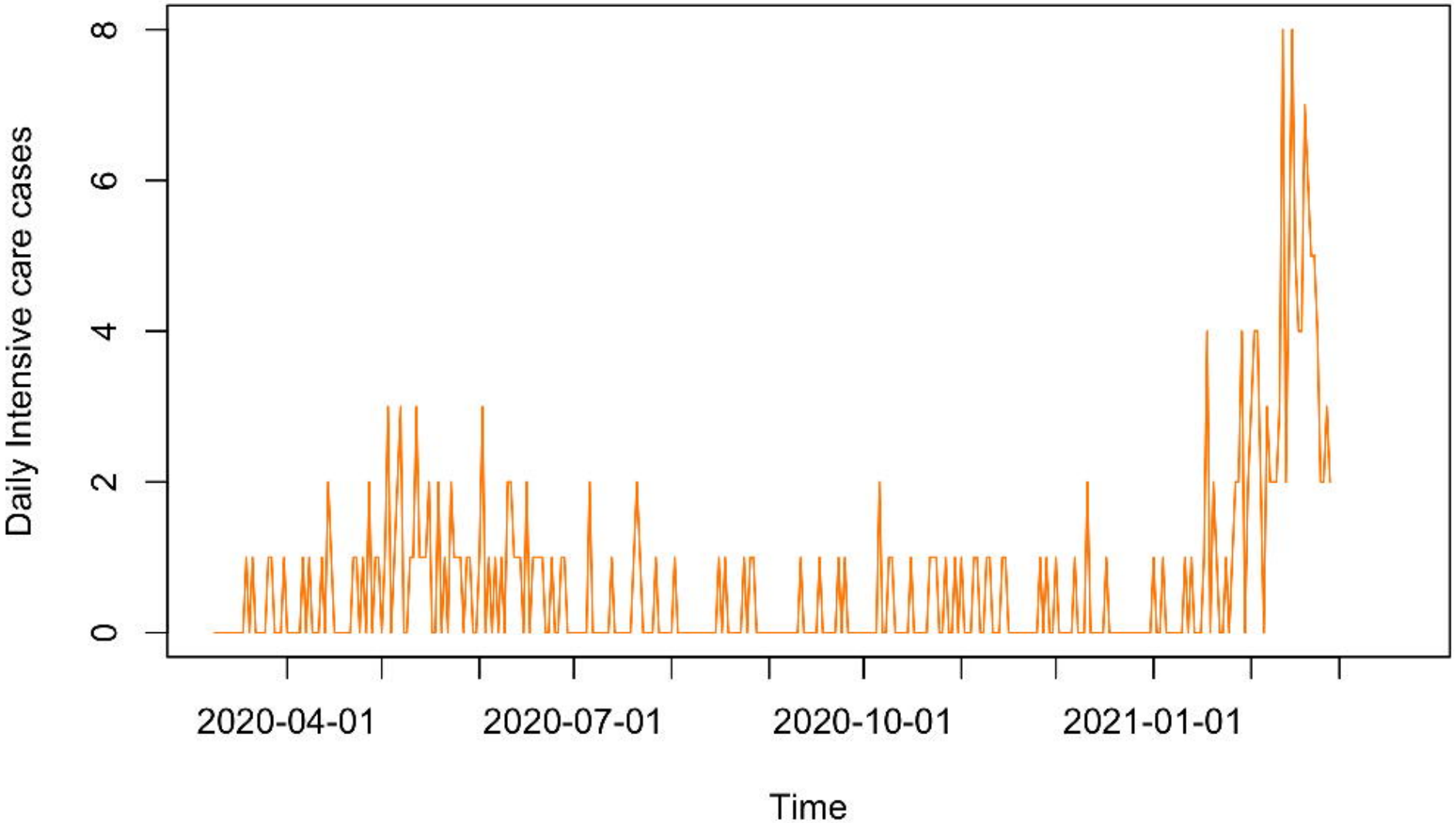

Predicted cumulative intensive care
Observed cumulative intensive care
Intensive_care
0
50
100
150
200
0
50
100
150
200
250
300
350
Index

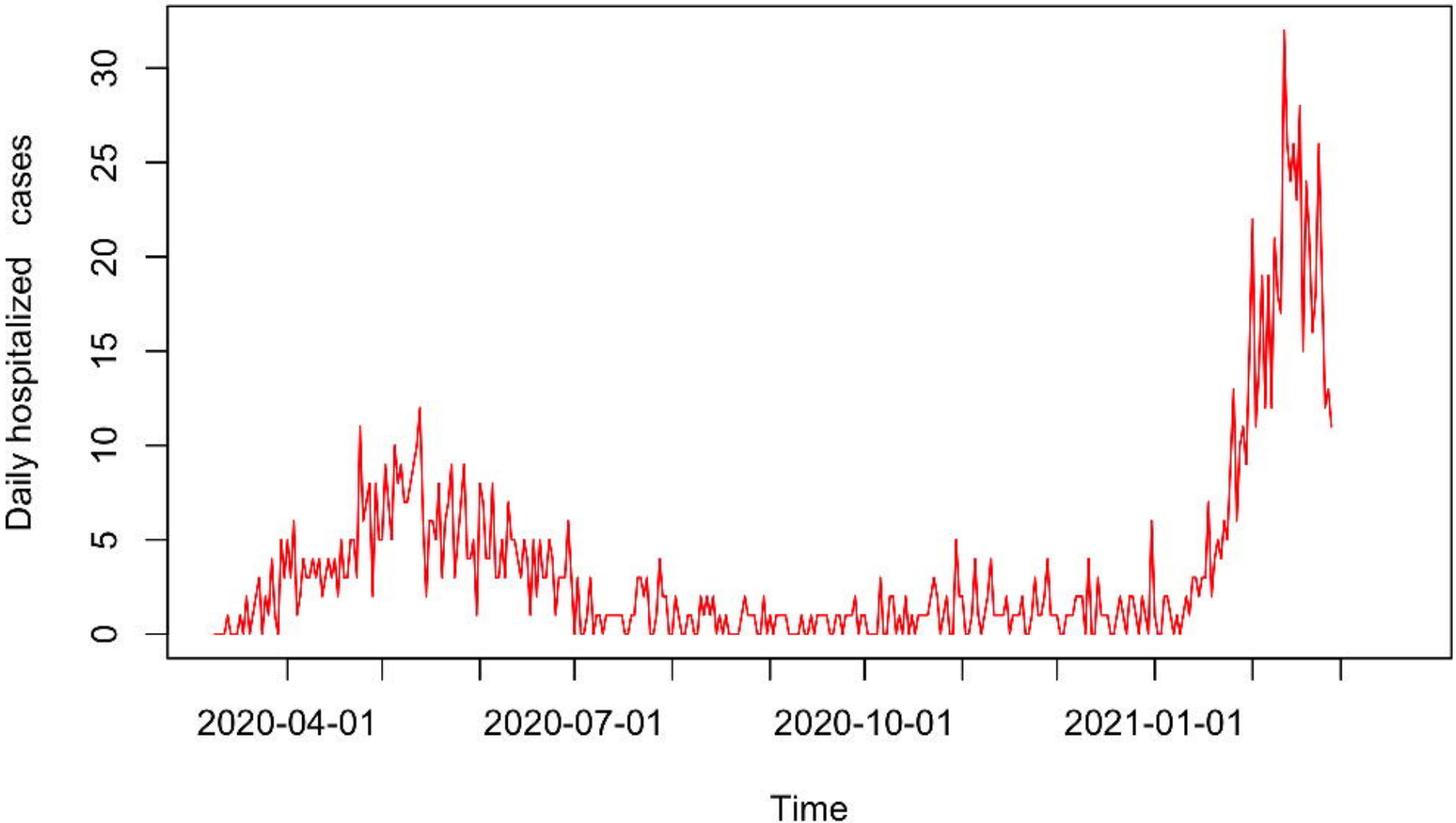

Predicted cumulative hospitalized cases
Observed cumulative hospitalized cases
hospitalized_cases
0
200
600
1000
0
50
100
150
200
250
300
350
Index

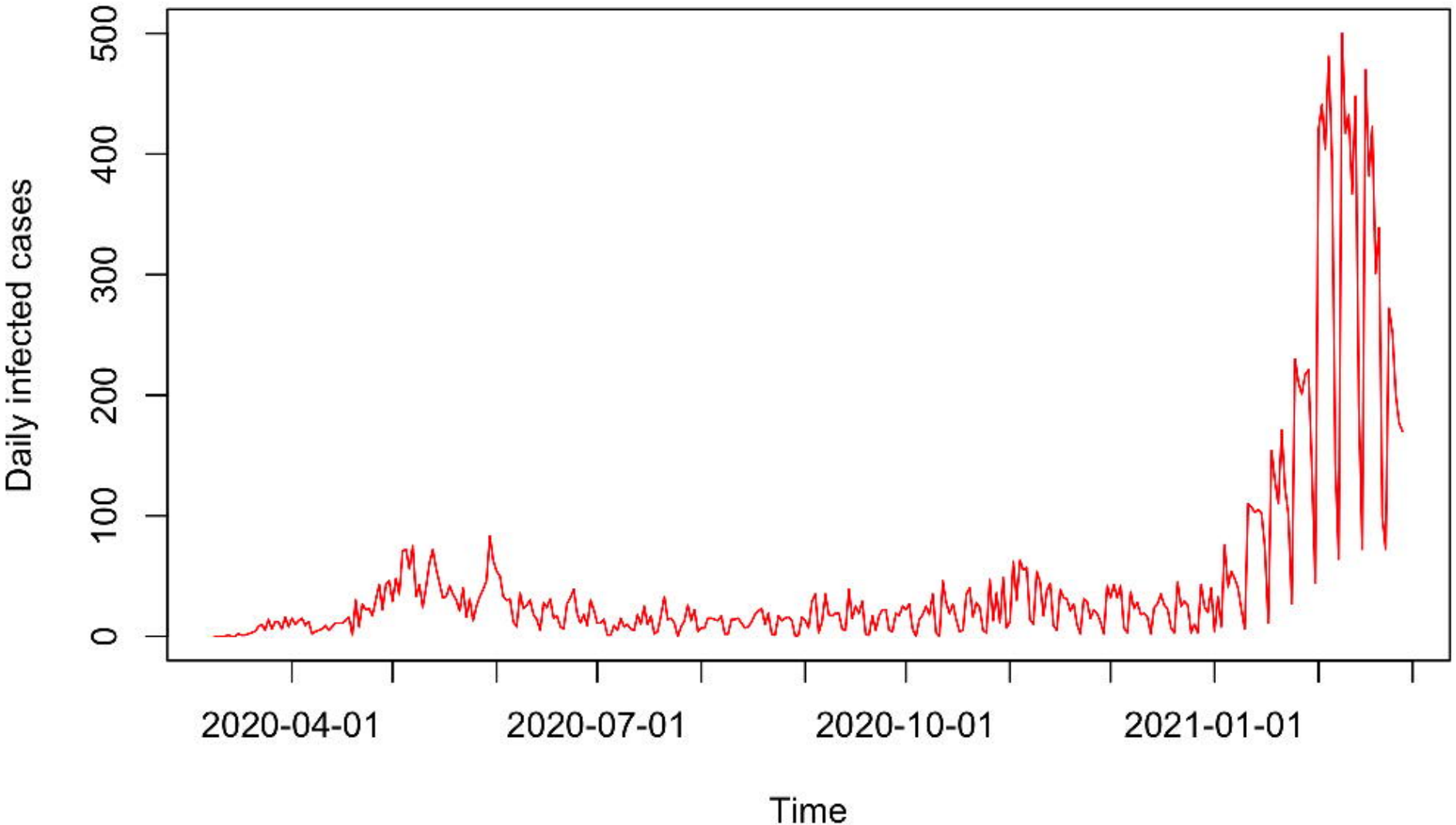

Predicted cumulative death
Observed cumulative death
Death_cases
0
20
40
60
80
0
50
100
150
200
250
300
350
Index

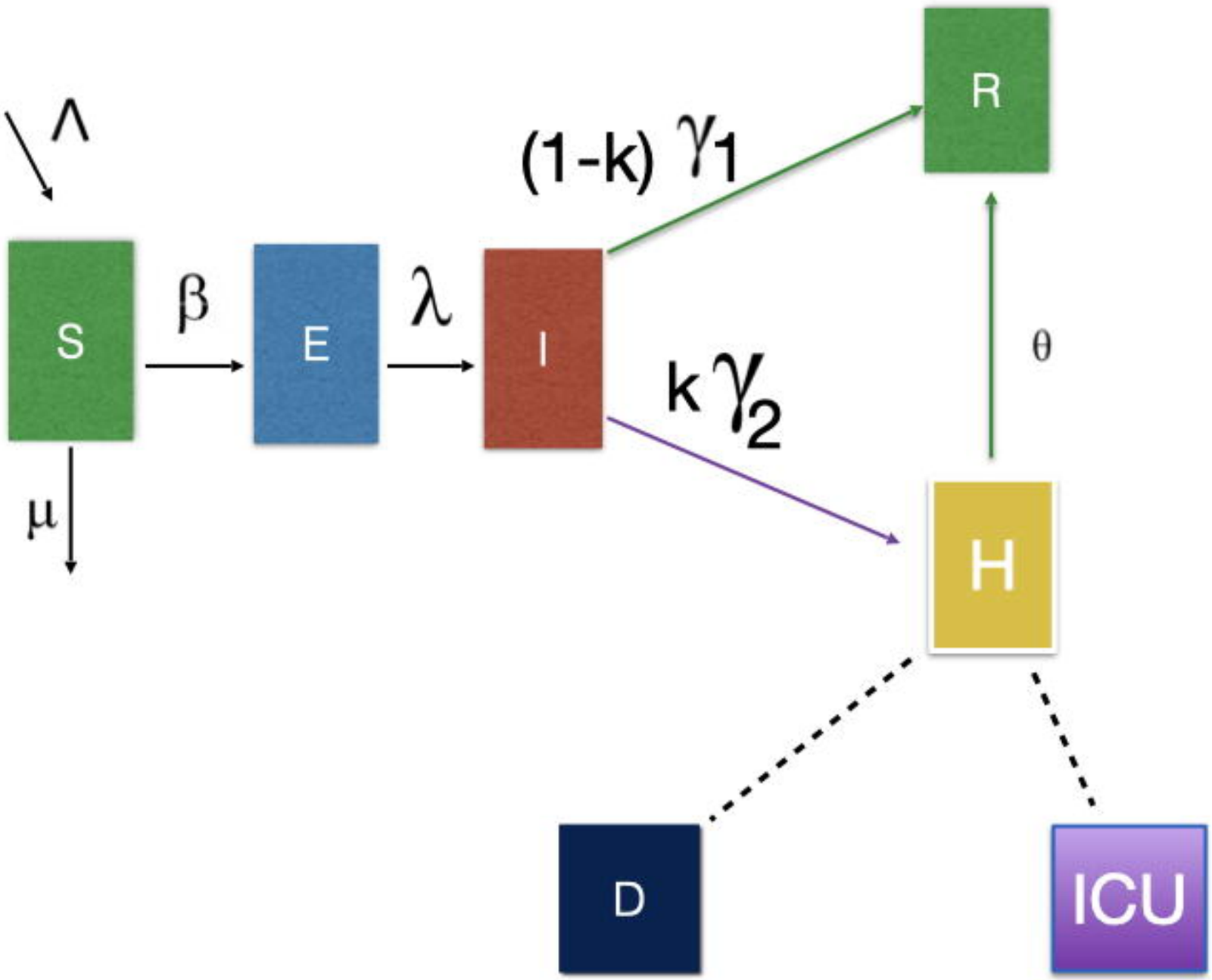

Daily cases per age
Daily_infected_age_group
0
50
100
150
200
Age groupe under 10
Age group between 10 and 24
Age group between 24 and 49
Age group 65+
2020-04-01
2020-07-01
2020-10-01
2021-01-01
Index

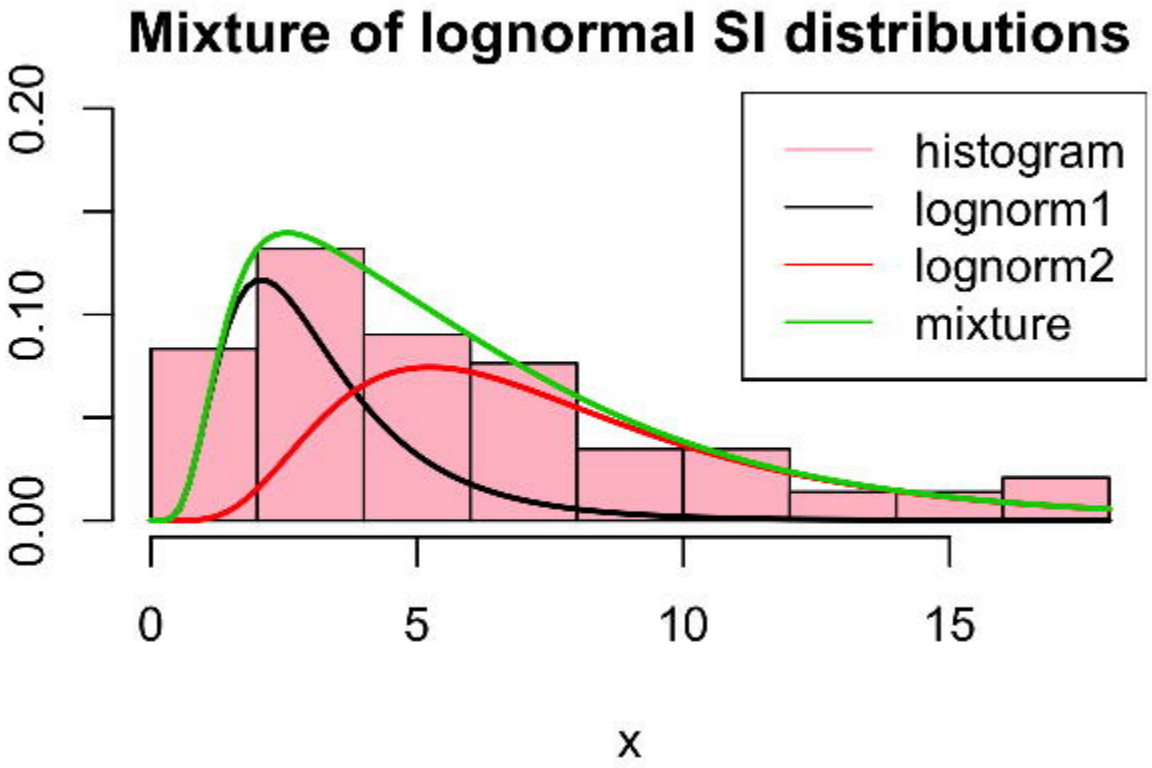

Daily death cases
0
2
4
6
8
2020-04-01
2020-07-01
2020-10-01
2021-01-01
Time

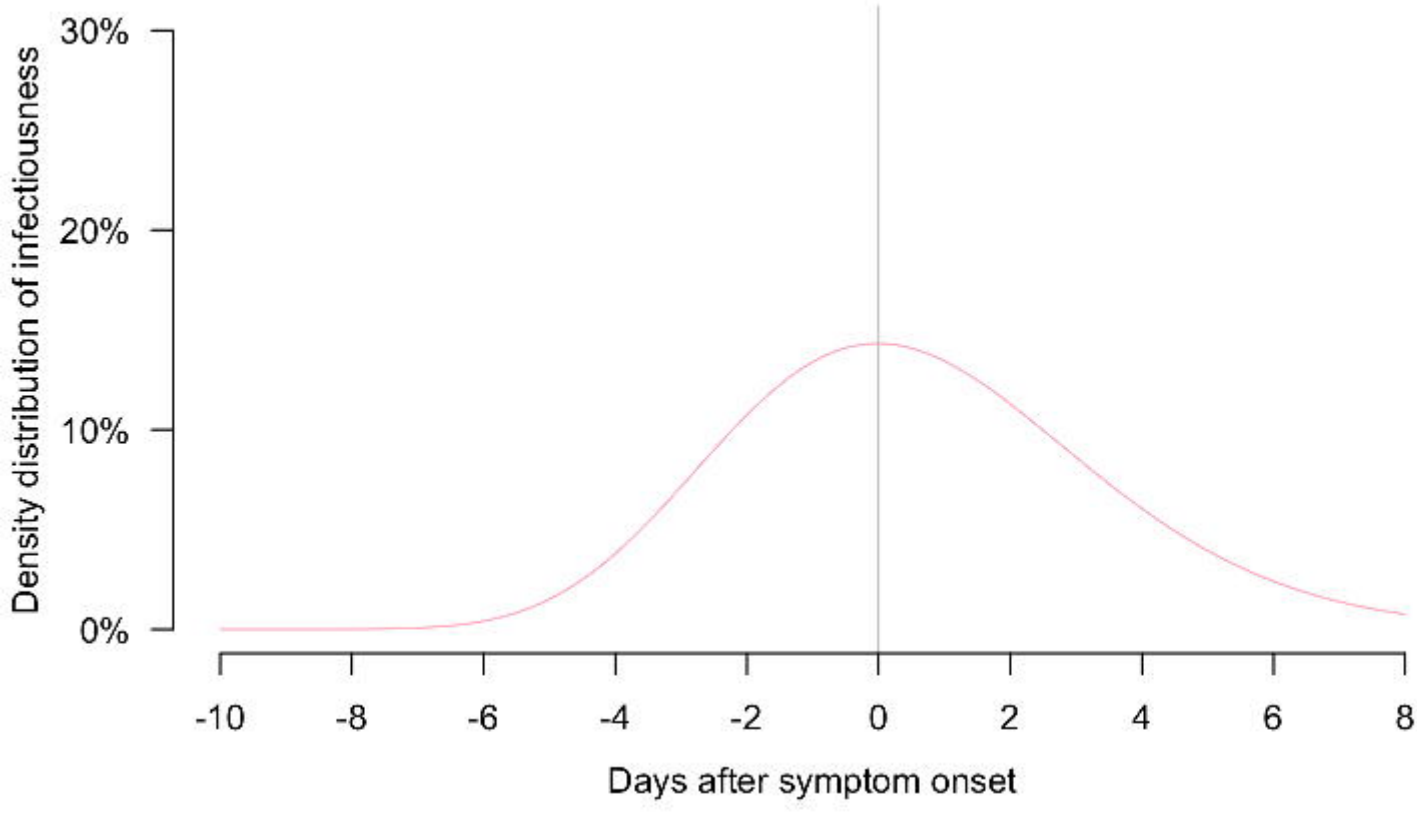

Daily Intensive care cases
0
2
4
6
8
2020-04-01
2020-07-01
2020-10-01
2021-01-01
Time

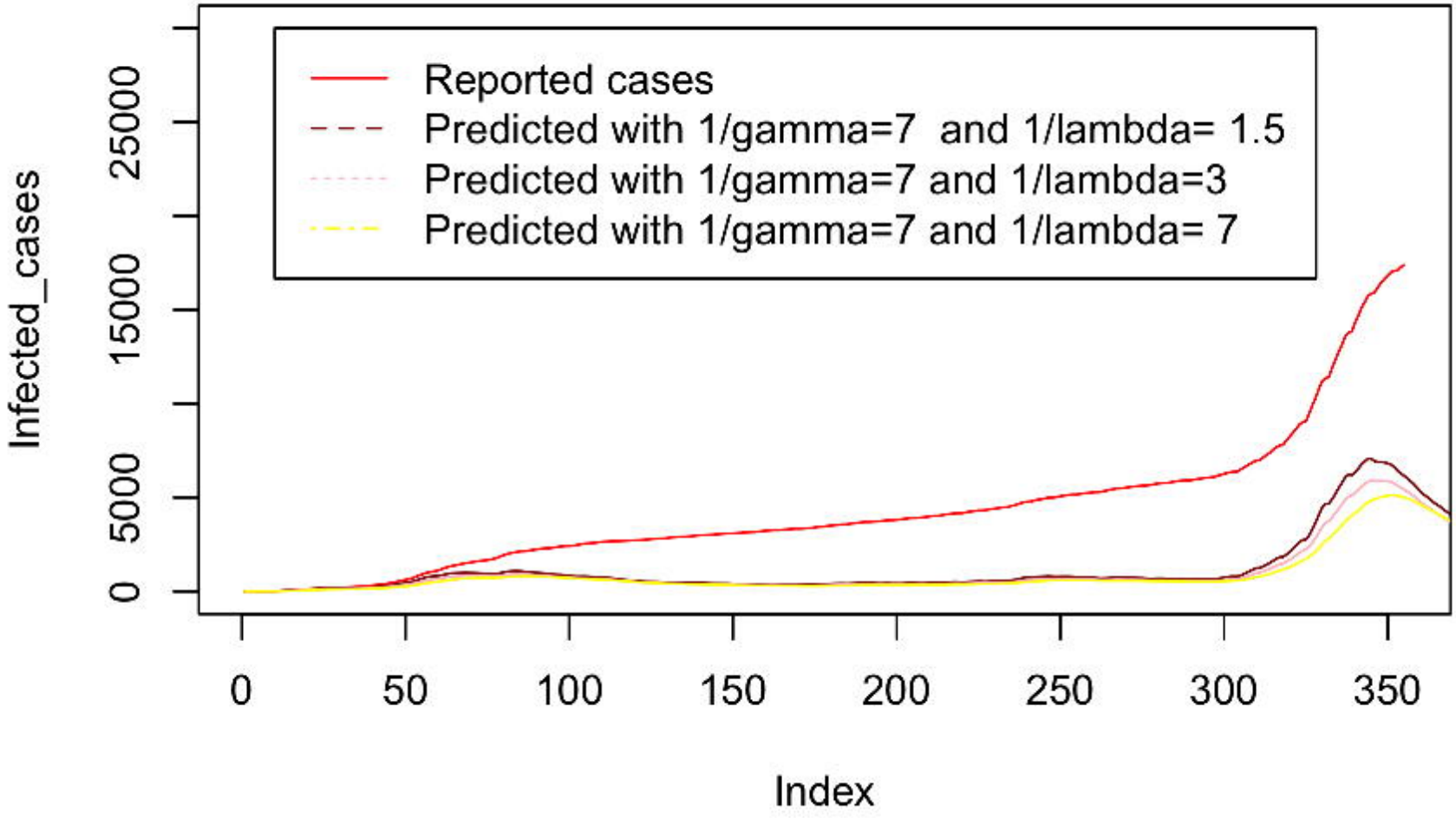

Daily hospitalized cases
0
5
10
15
20
25
30
2020-04-01
2020-07-01
2020-10-01
2021-01-01
Time

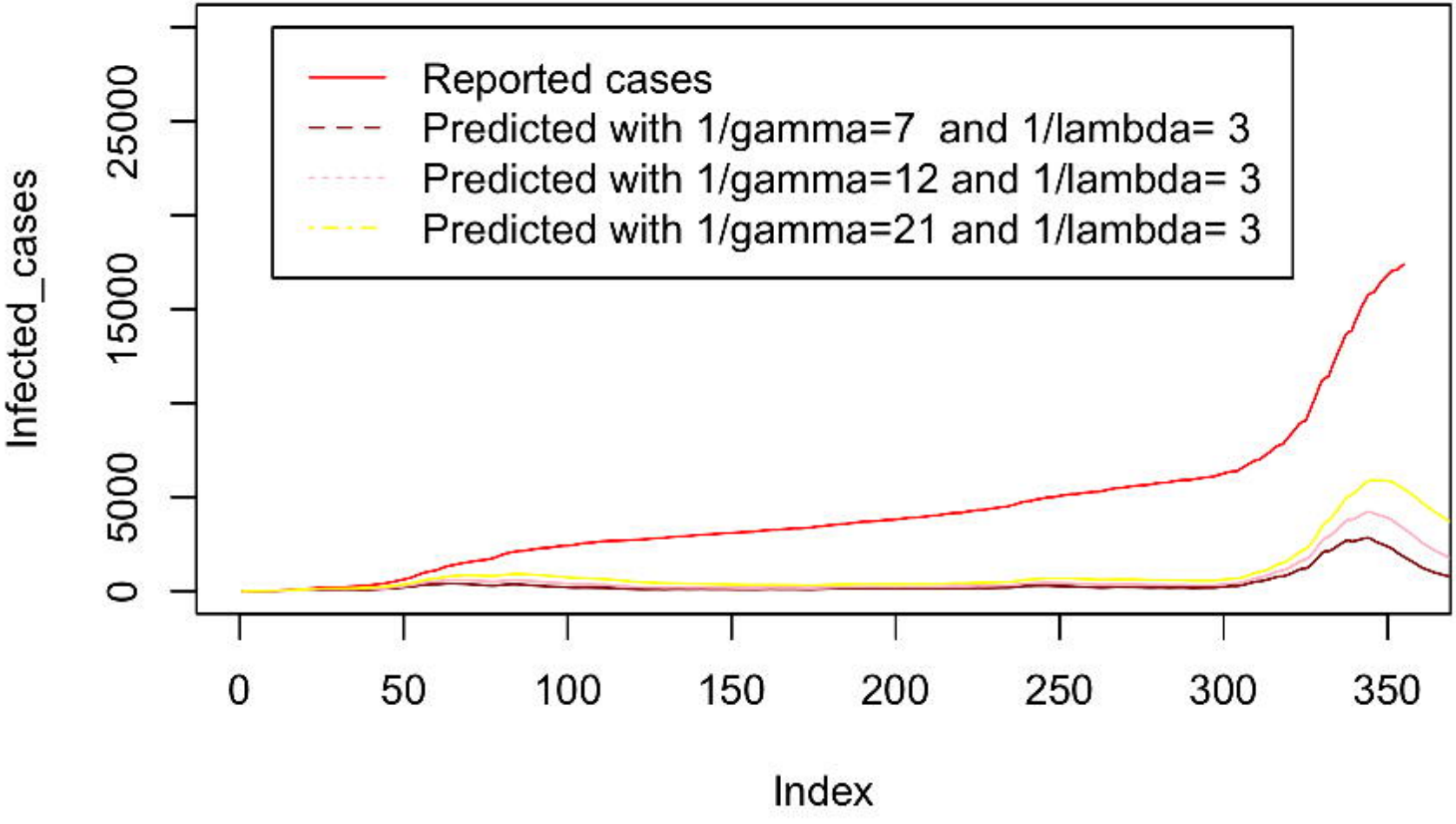

Daily infected cases
0
100
200
300
400
500
2020-04-01
2020-07-01
2020-10-01
2021-01-01
Time

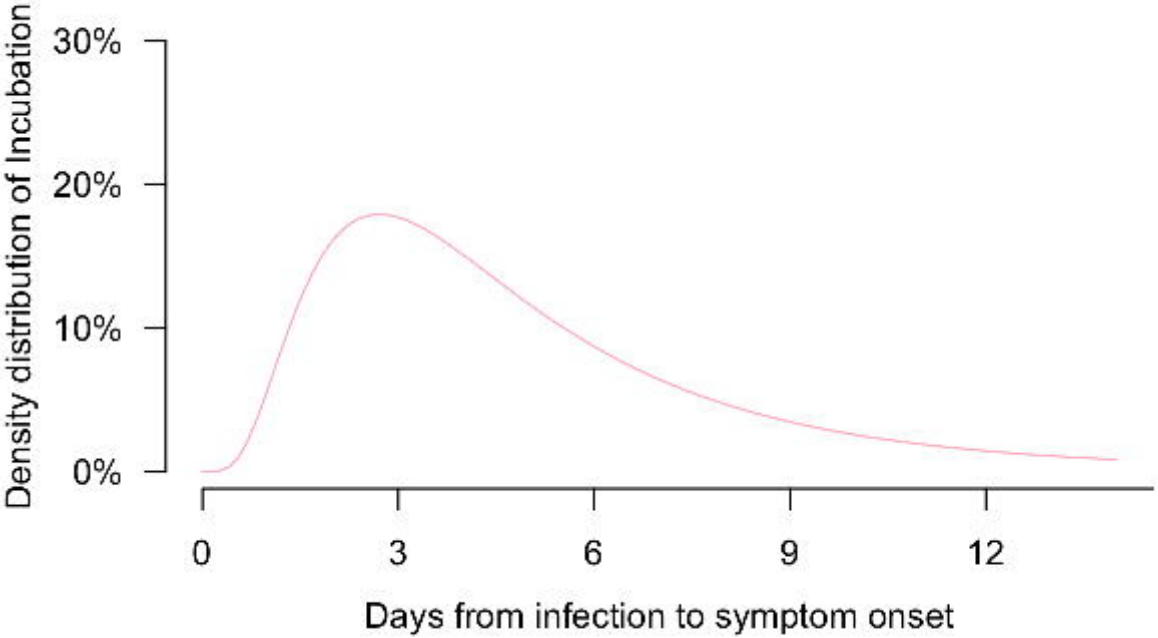

Λ
S
β
E
λ
I
(1-k) γ1
R
k γ2
θ
μ
H
D
ICU

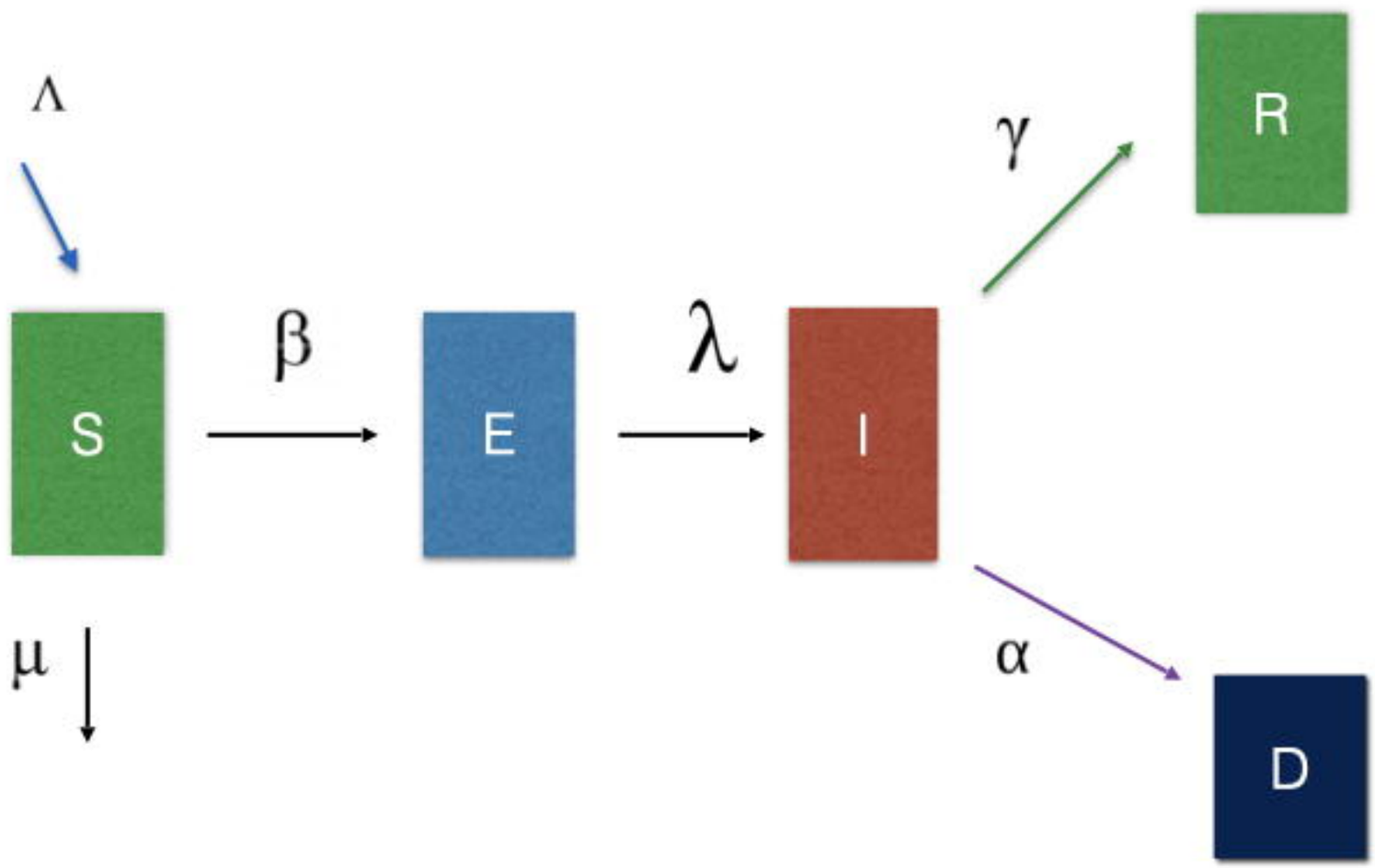

Mixture of lognormal SI distributions
0.20
0.10
0.00
histogram
lognorm1
lognorm2
mixture
0
5
10
15
x

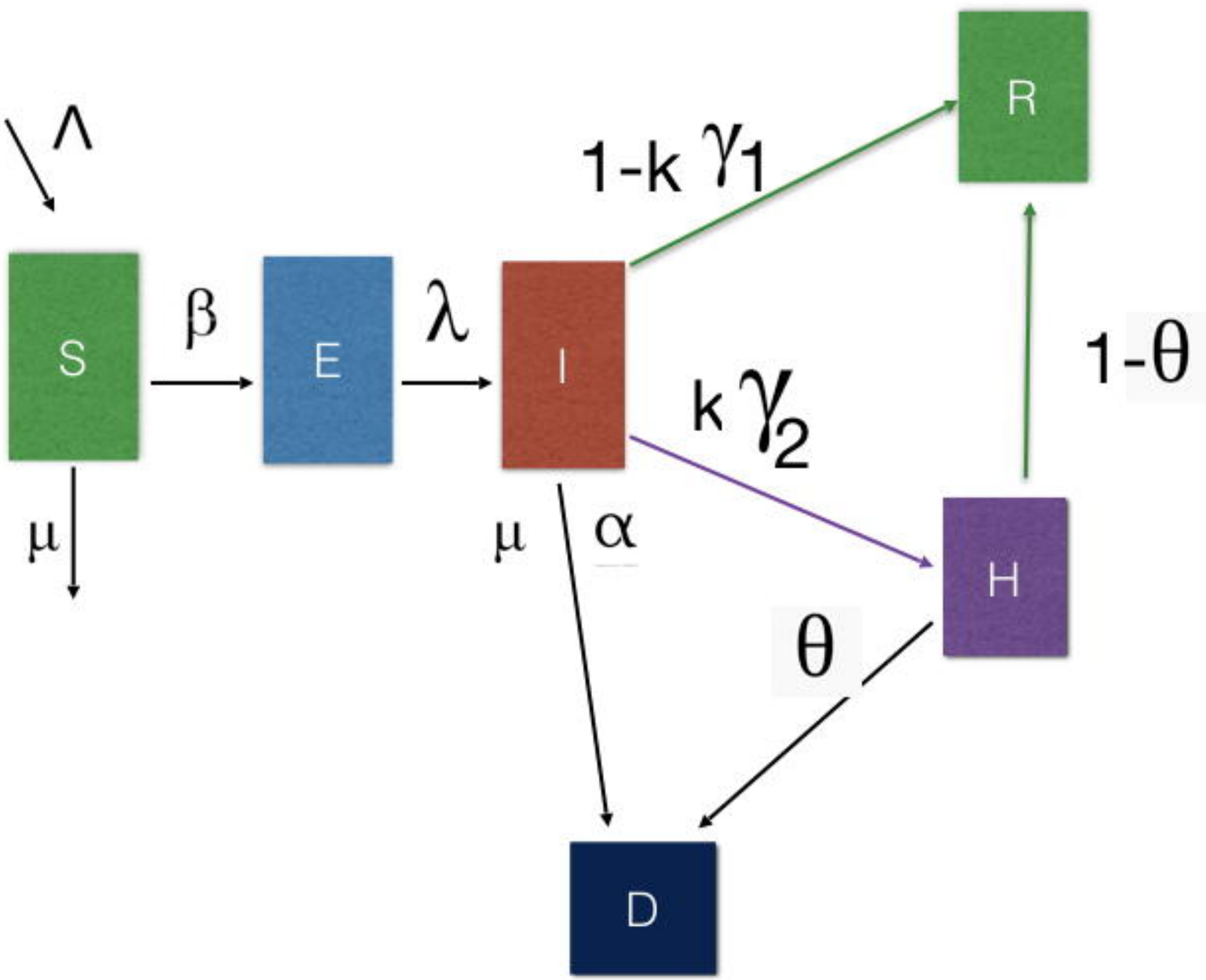

Density distribution of infectiousness
30%
20%
10%
0%
-10
-8
-6
-4
-2
0
2
4
6
8
Days after symptom onset

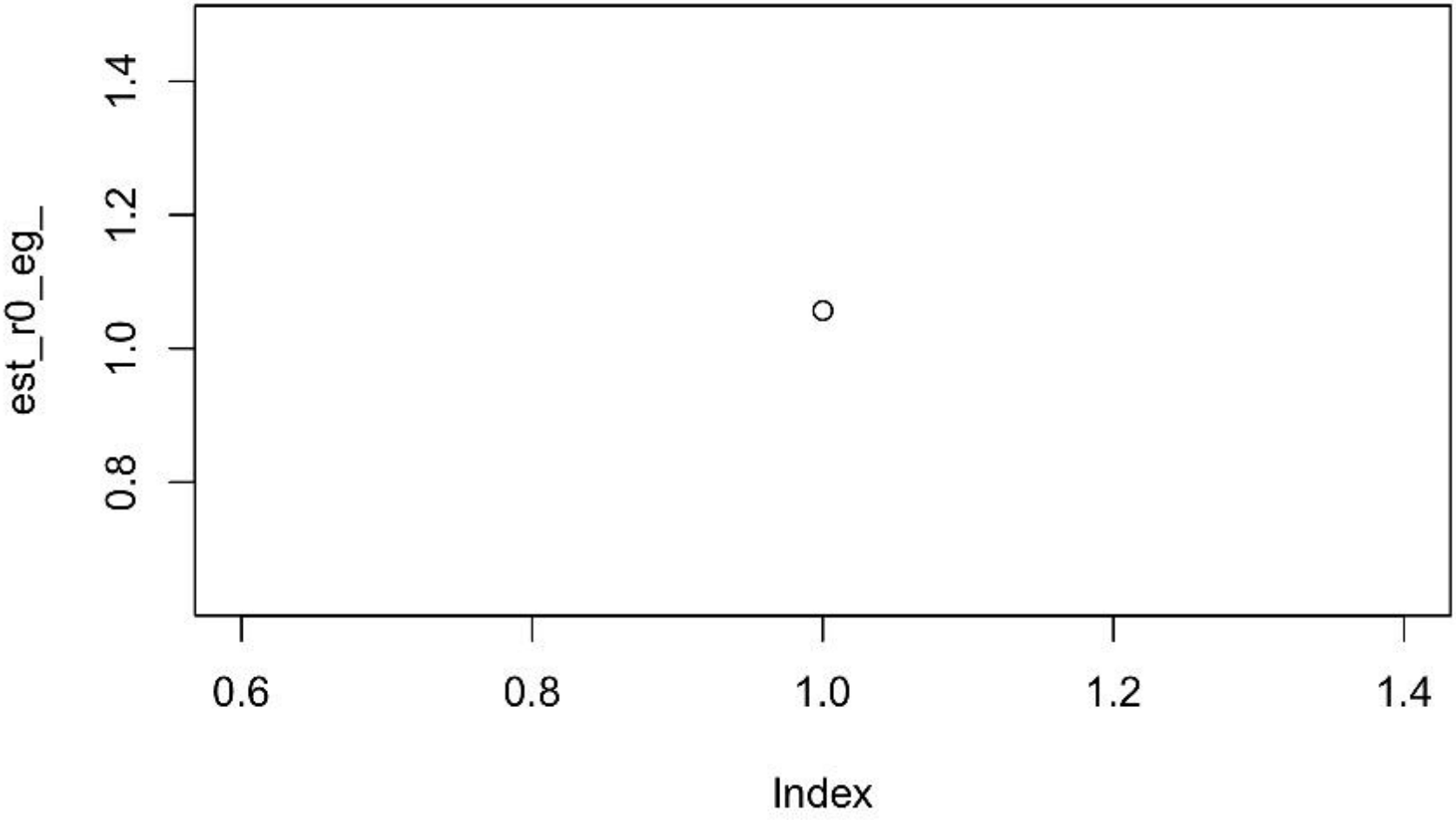

Reported cases
Predicted with 1/gamma=7 and 1/lambda= 1.5
Predicted with 1/gamma=7 and 1/lambda=3
Predicted with 1/gamma=7 and 1/lambda= 7
Infected_cases
0
5000
15000
25000
0
50
100
150
200
250
300
350
Index

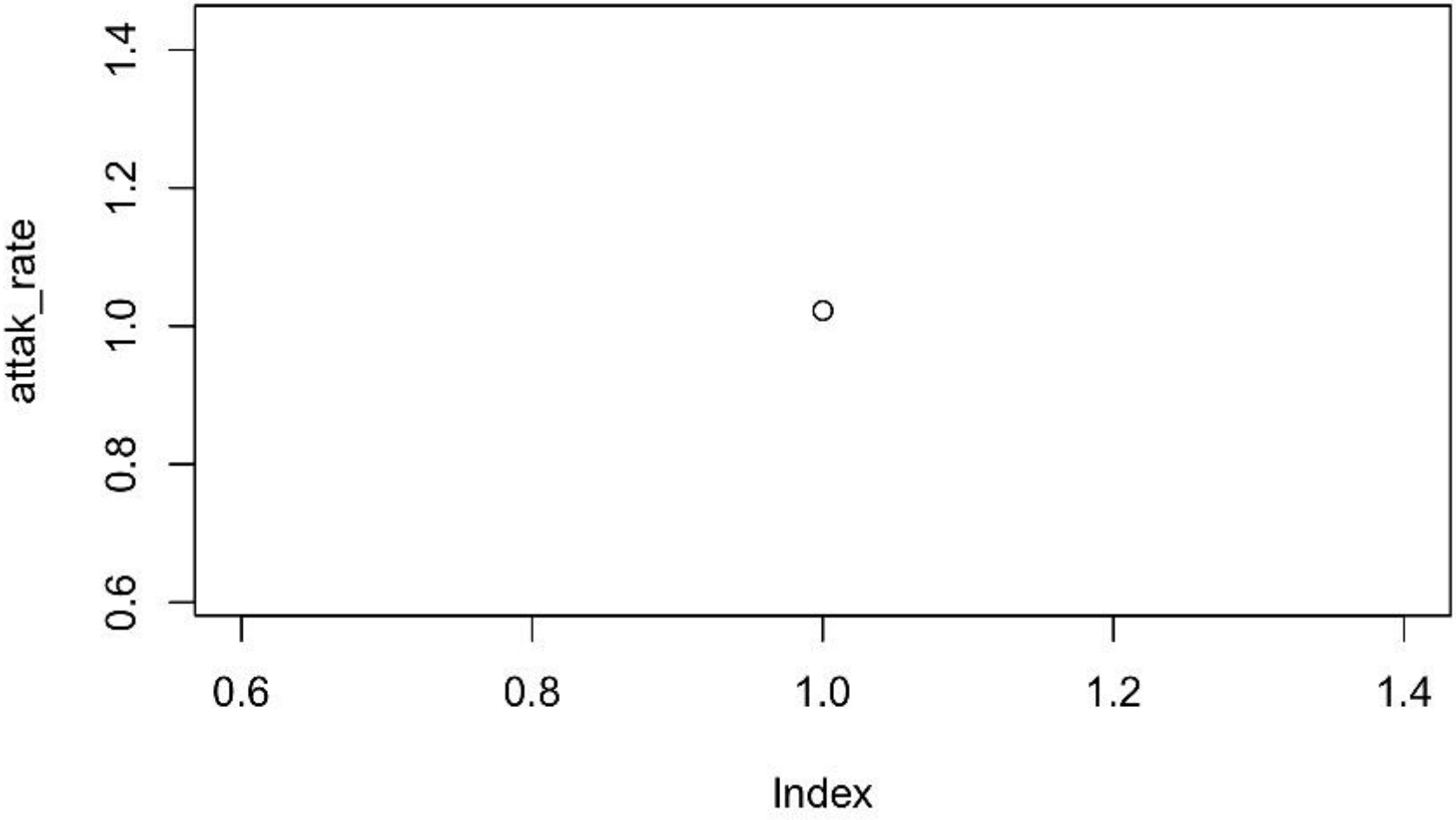

Reported cases
Predicted with 1/gamma=7 and 1/lambda= 3
Predicted with 1/gamma=12 and 1/lambda= 3
Predicted with 1/gamma=21 and 1/lambda= 3
Infected_cases
0
5000
15000
25000
0
50
100
150
200
250
300
350
Index

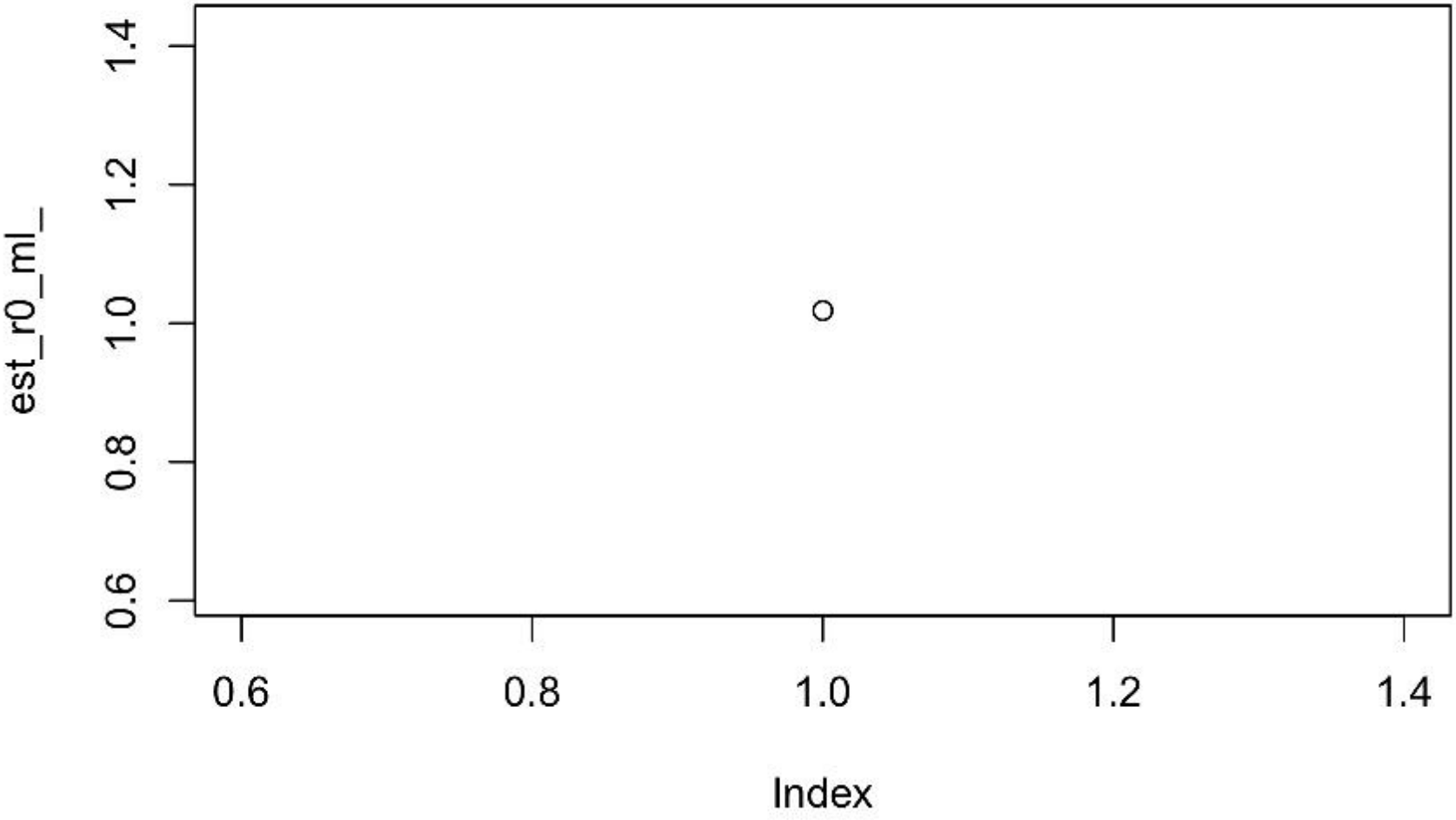

Density distribution of Incubation
30%
20%
10%
0%
0
3
6
9
12
Days from infection to symptom onset

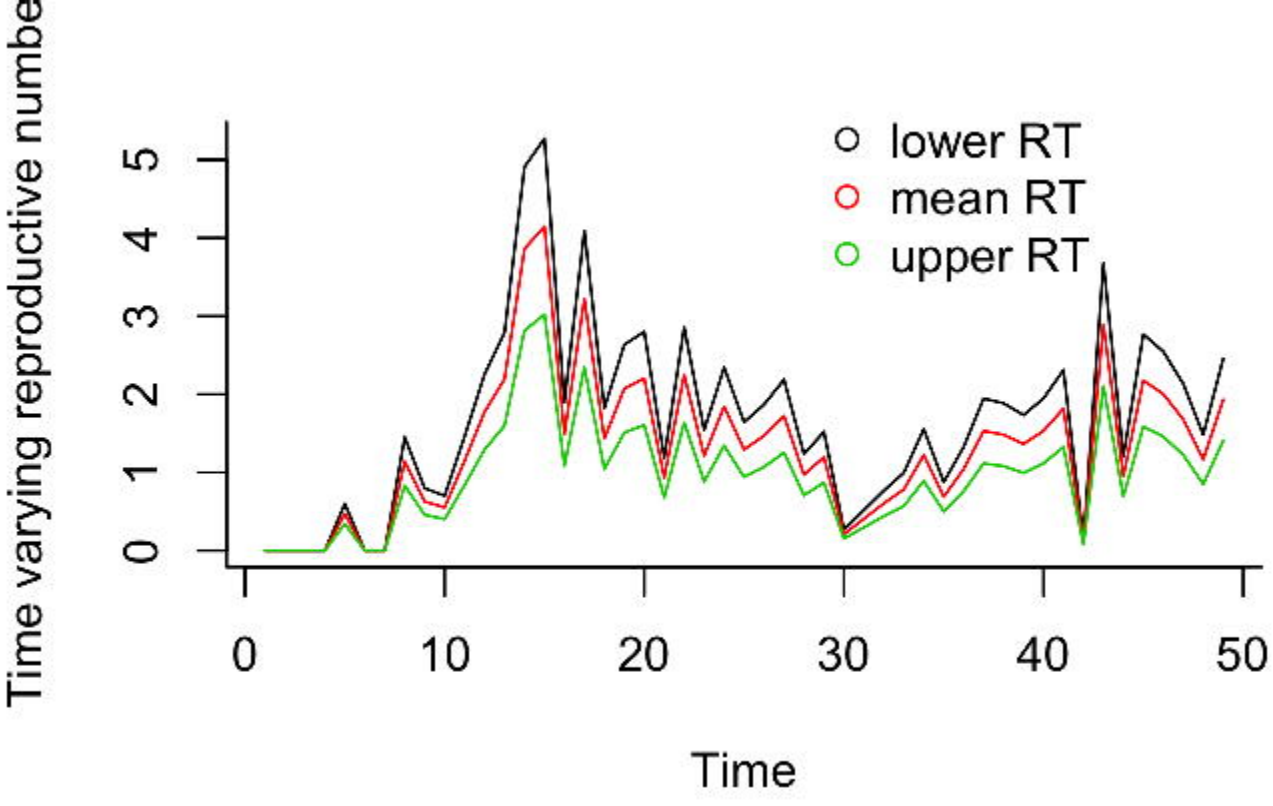

Λ
S
β
E
λ
I
γ
R
μ
α
D

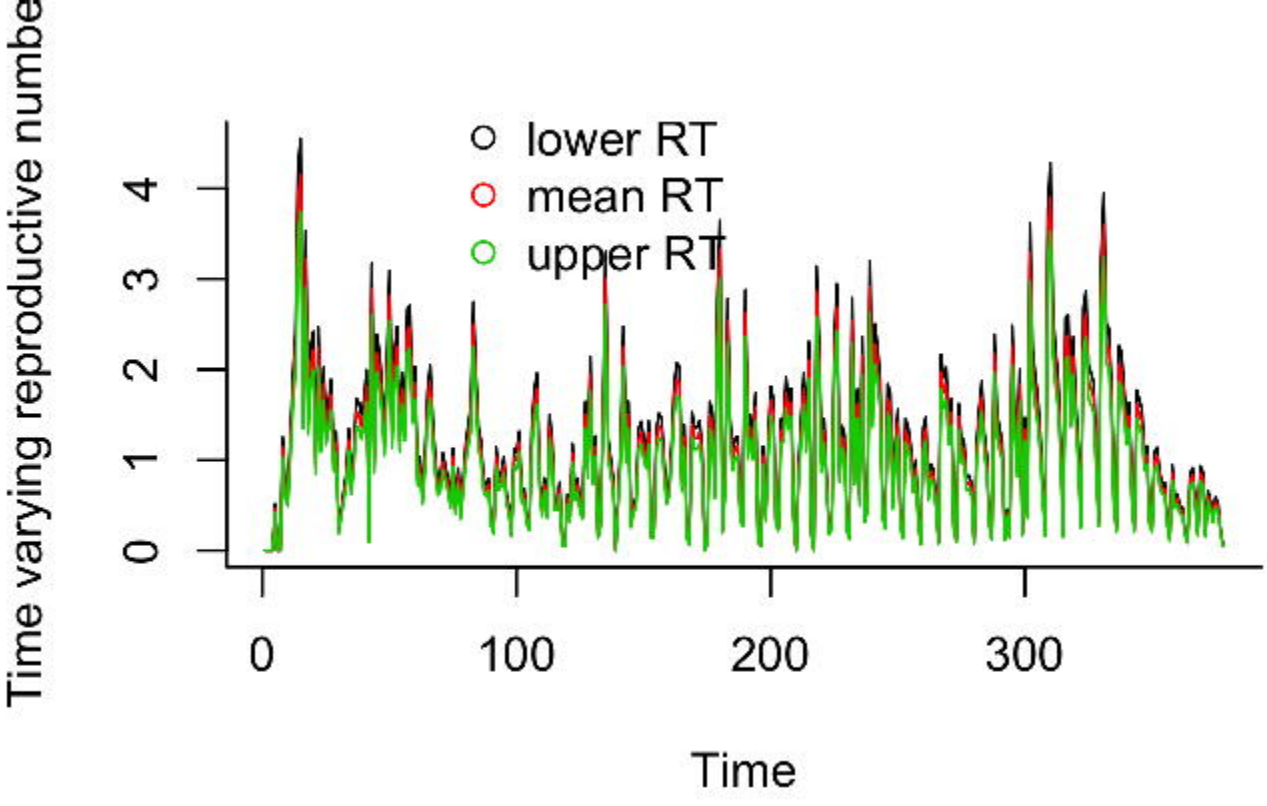

Λ
S
β
E
λ
I
1-k γ1
R
k γ2
1-θ
μ
μ
α
H
θ
D

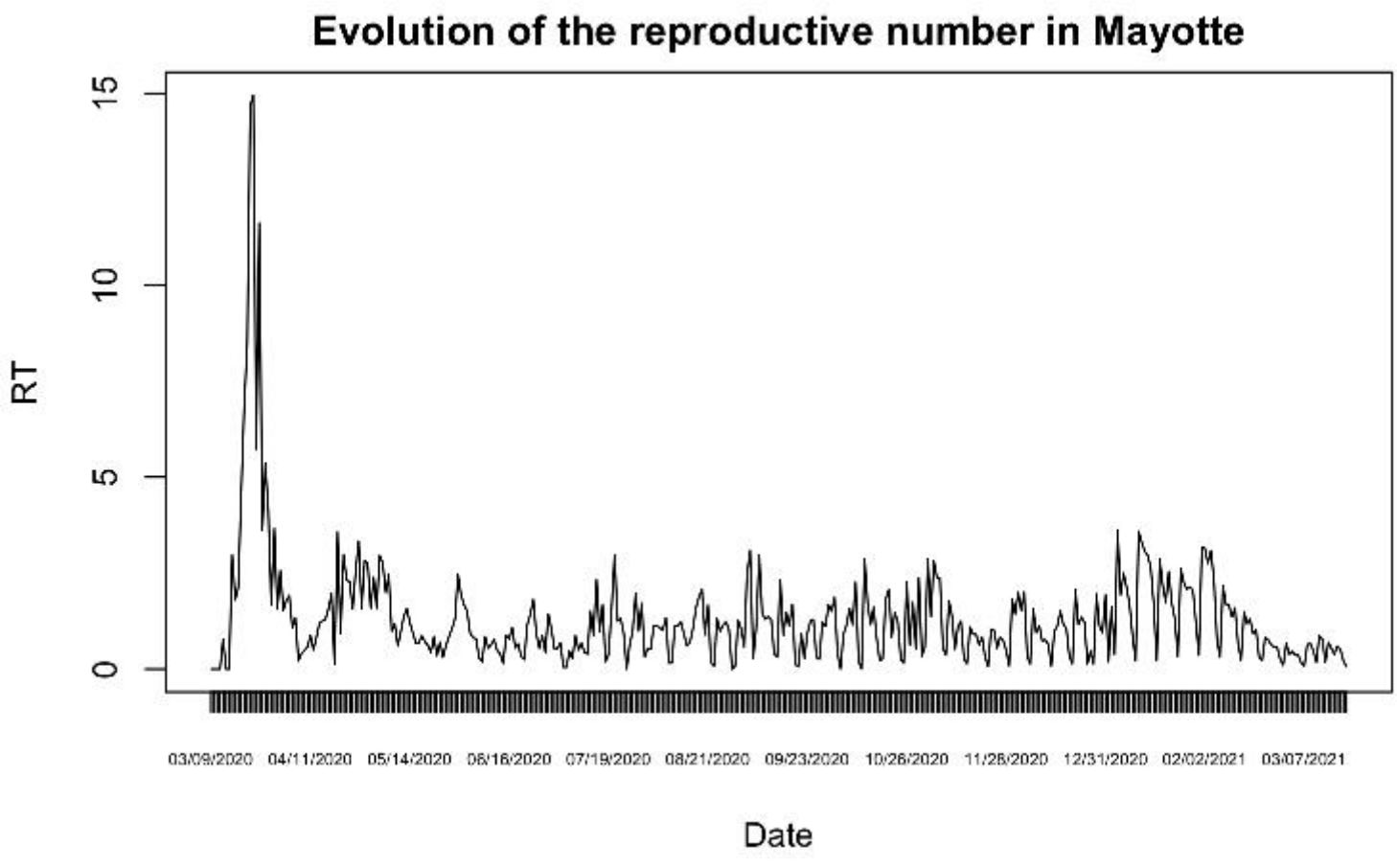

est_r0_eg_
0.8
1.0
1.2
1.4
0.6
0.8
1.0
1.2
1.4
Index

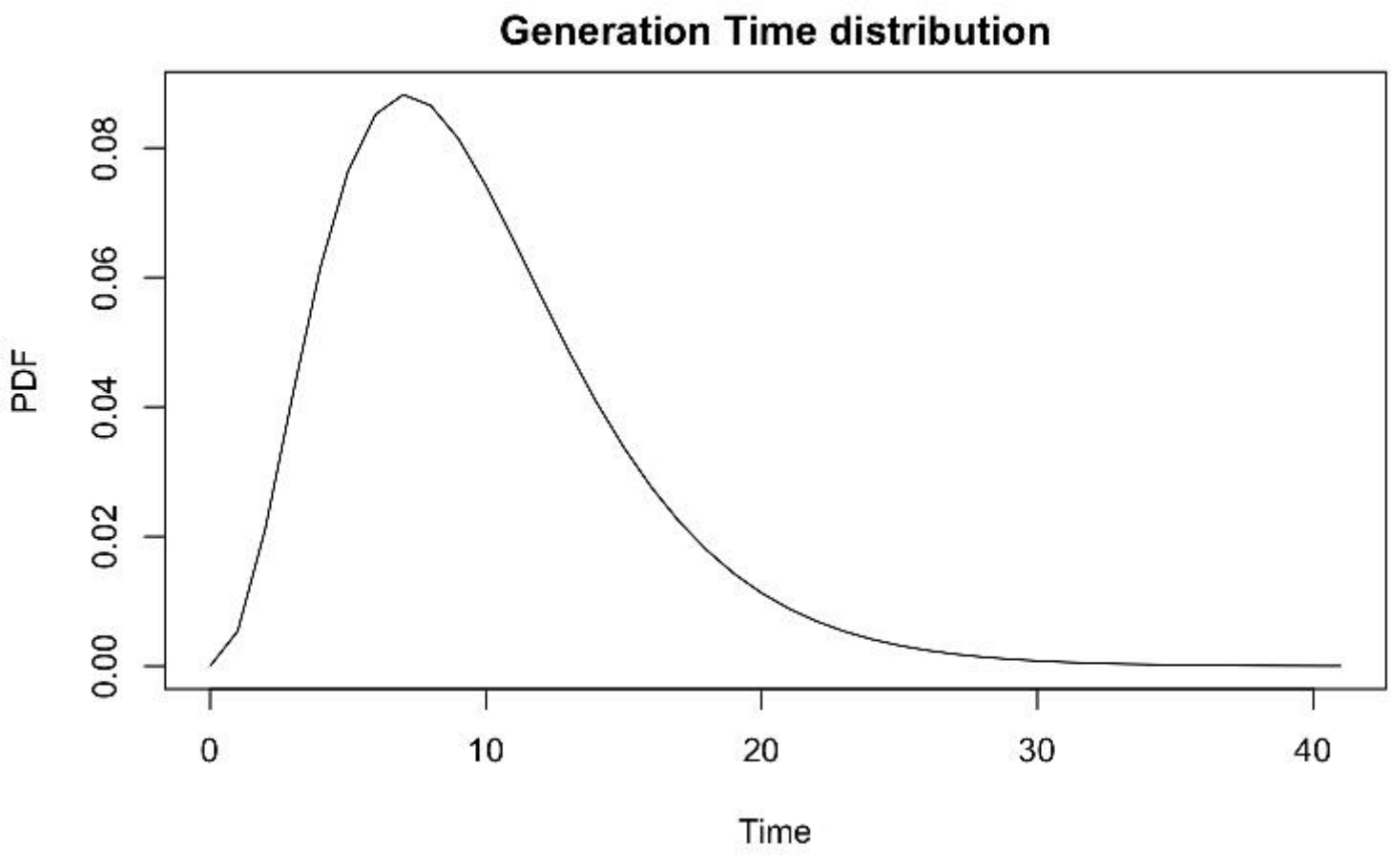

attak_rate
0.6
0.8
1.0
1.2
1.4
0.6
0.8
1.0
1.2
1.4
Index

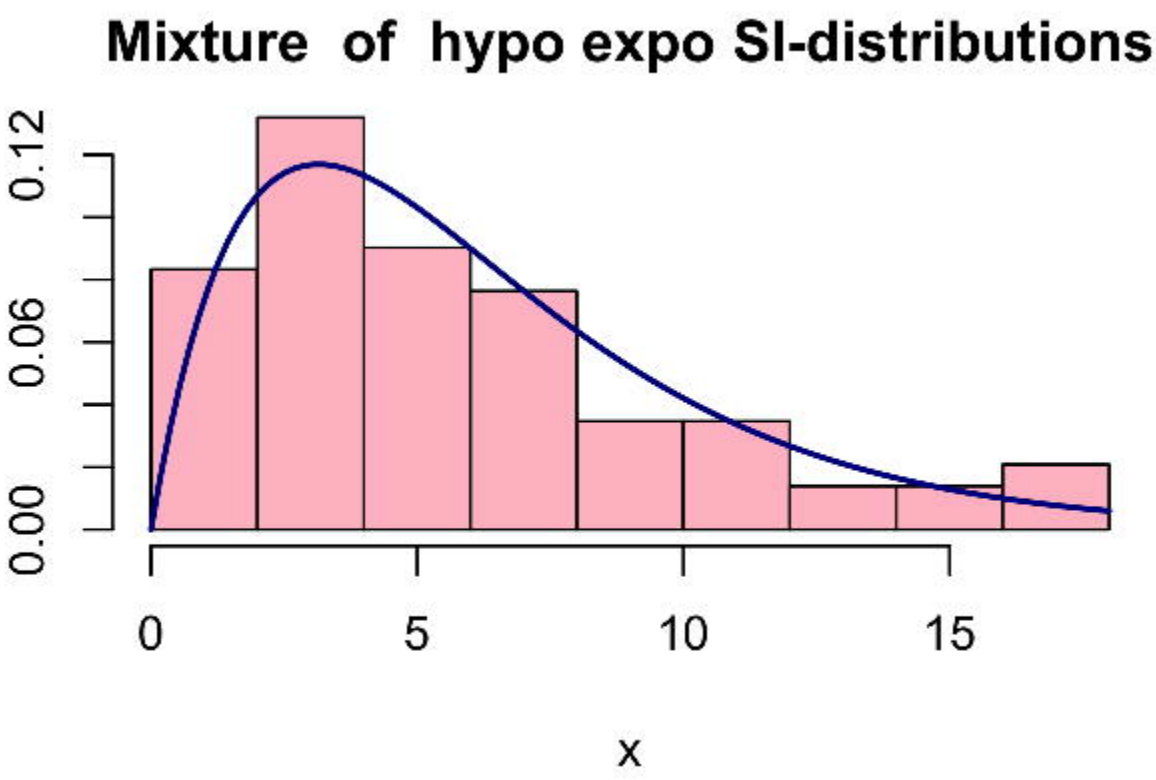

est_r0_ml_
1.4
1.2
1.0
0.8
0.6
0.6
0.8
1.0
1.2
1.4
Index

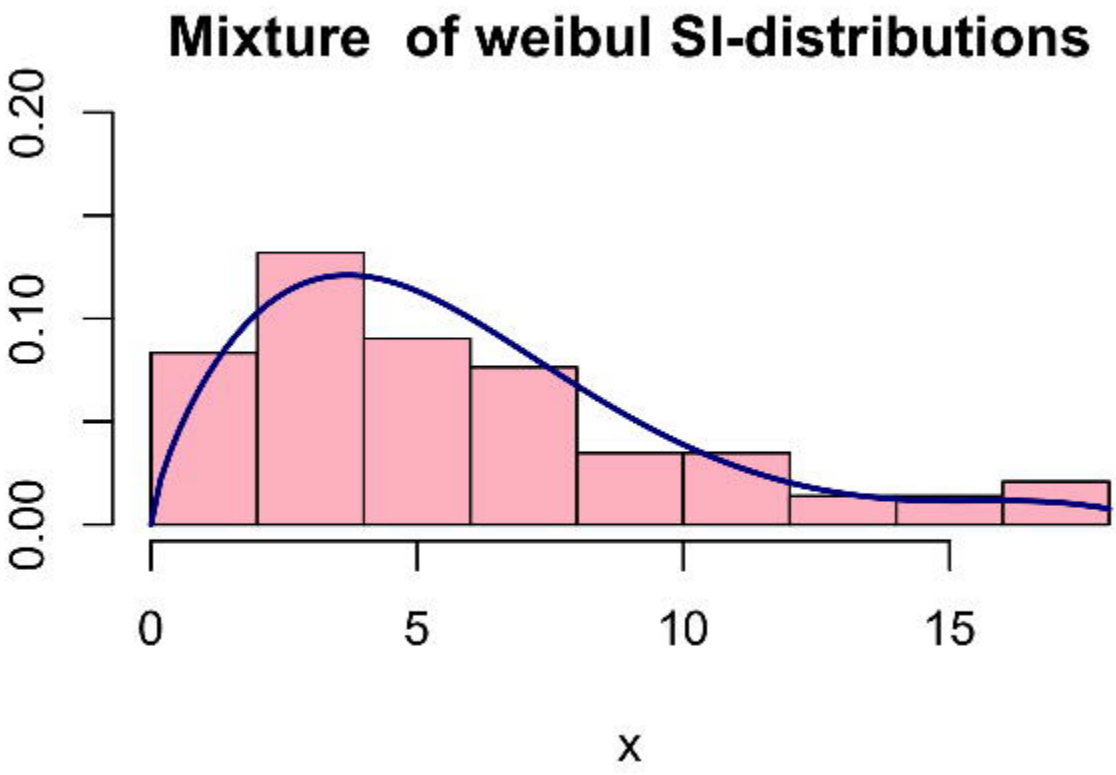

lower RT
mean RT
upper RT
Time varying reproductive number
0
1
2
3
4
5
0
10
20
30
40
50
Time

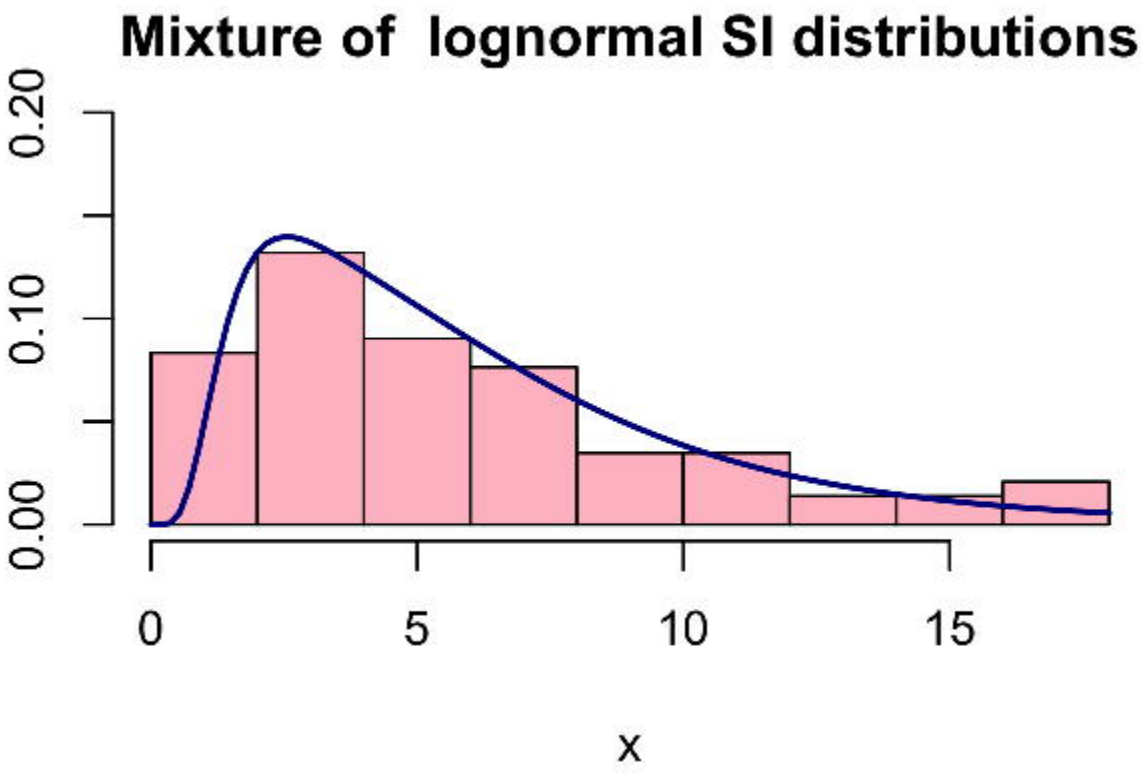

lower RT
mean RT
upper RT
Time varying reproductive number
0
1
2
3
4
0
100
200
300
Time

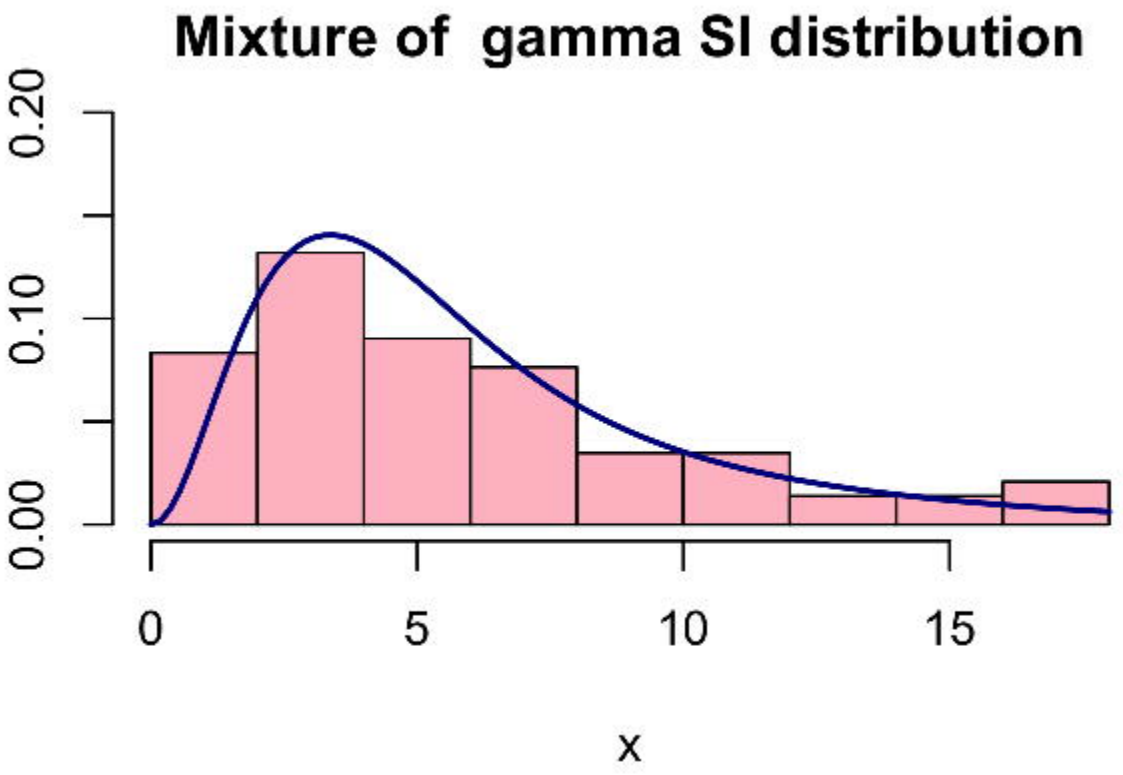

Evolution of the reproductive number in Mayotte
RT
0
5
10
15
03/09/2020
04/11/2020
05/14/2020
06/16/2020
07/19/2020
08/21/2020
09/23/2020
10/26/2020
11/28/2020
12/31/2020
02/02/2021
03/07/2021
Date

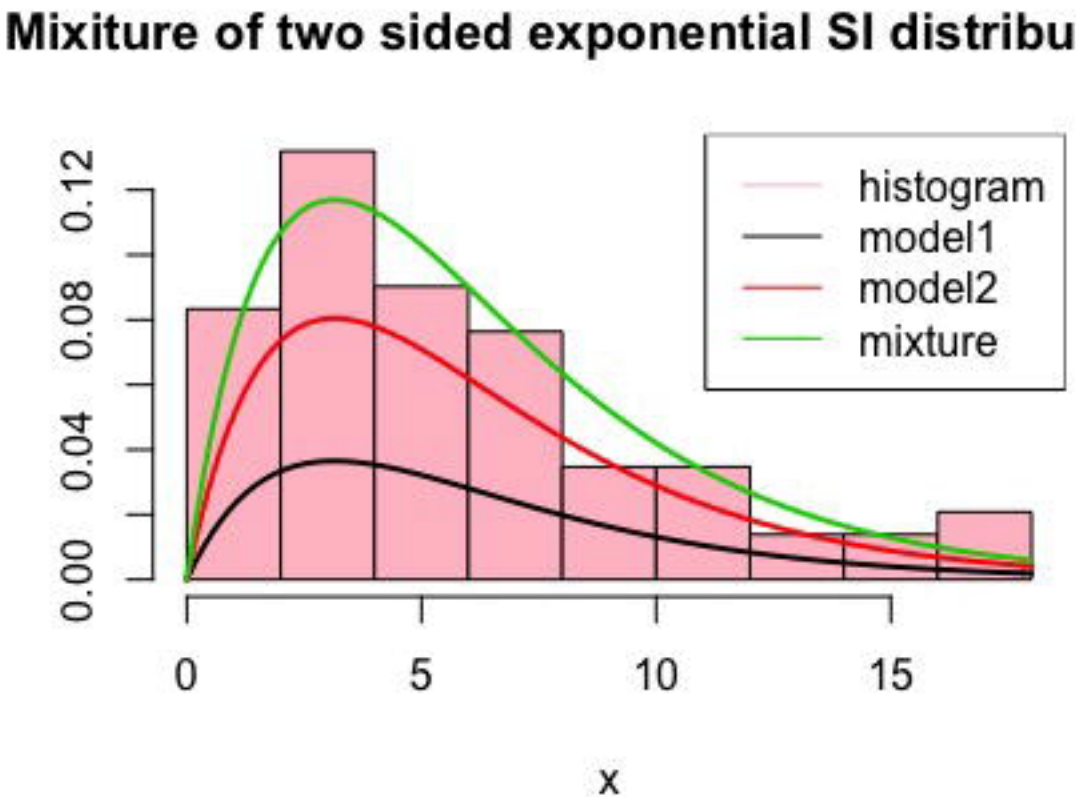

Generation Time distribution
PDF
0.00
0.02
0.04
0.06
0.08
0
10
20
30
40
Time

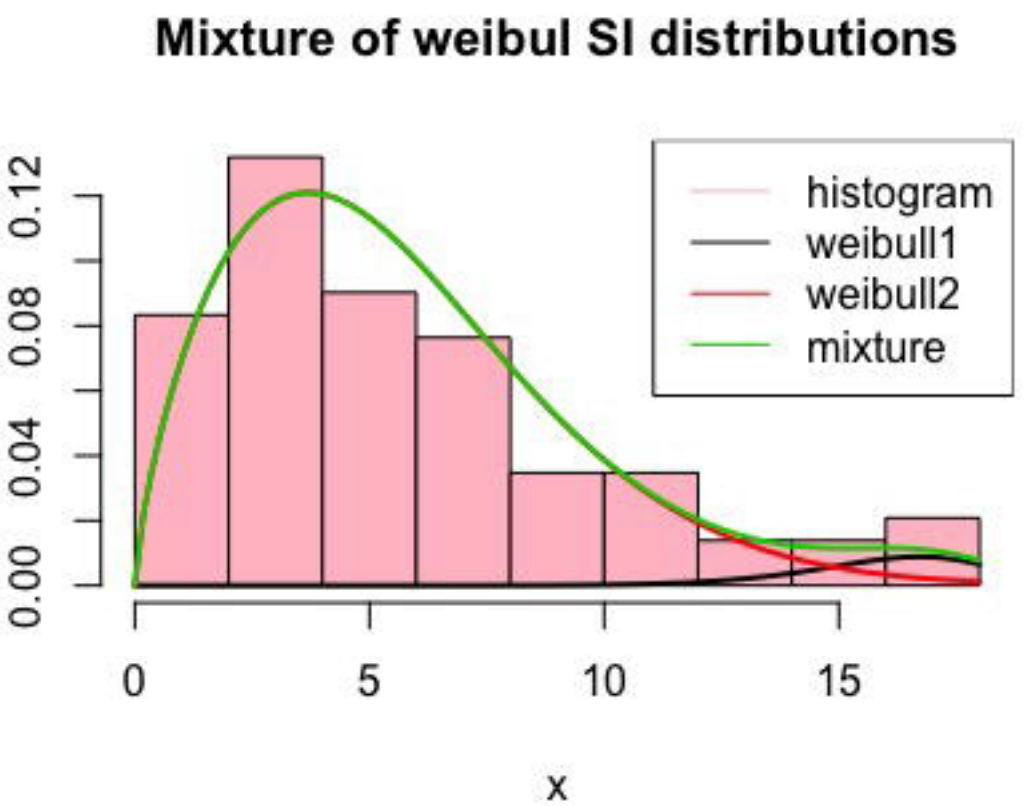

Mixture of hypo expo SI-distributions
0.00
0.06
0.12
0
5
10
15
x

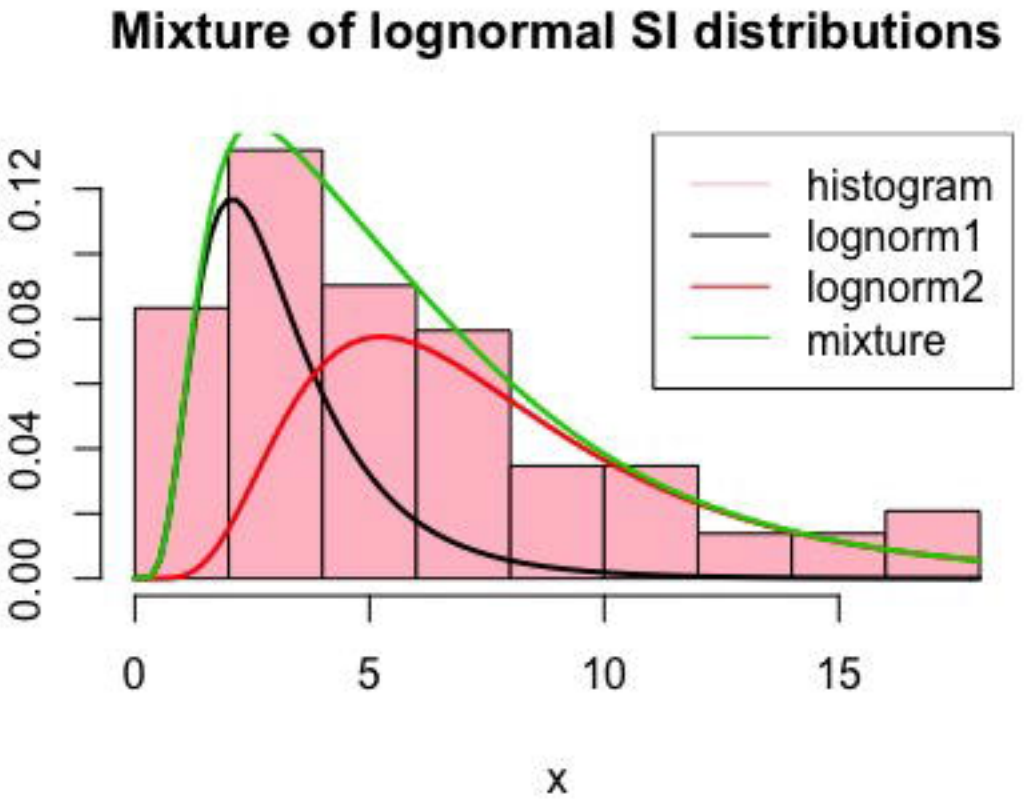

Mixture of weibul SI-distributions
0.20
0.10
0.00
0
5
10
15
x

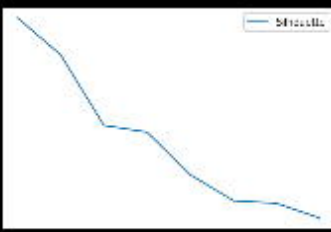

Mixture of lognormal SI distributions
0.20
0.10
0.00
0
5
10
15
x

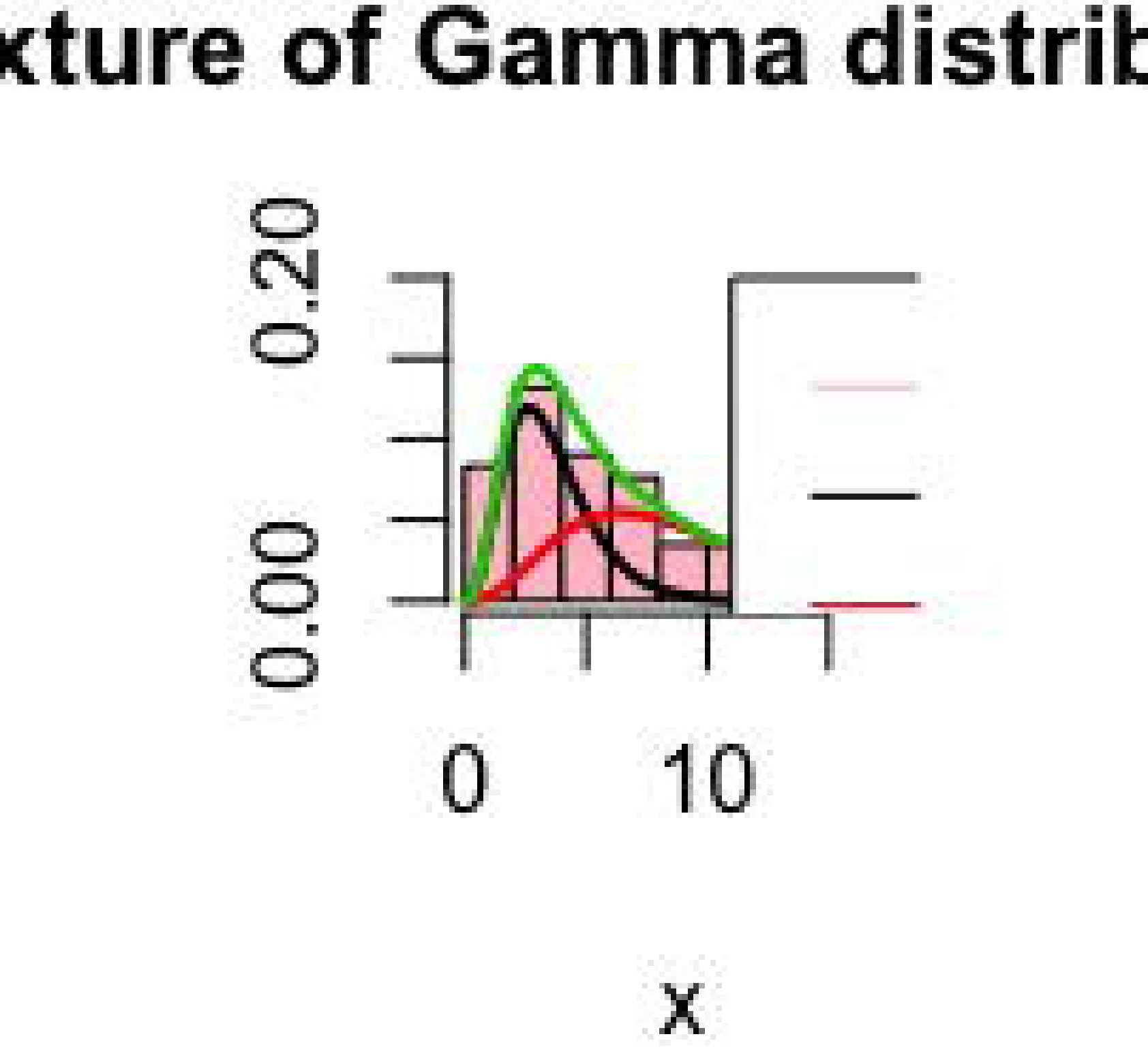

Mixture of gamma SI distribution
0.20
0.10
0.00
0
5
10
15
x

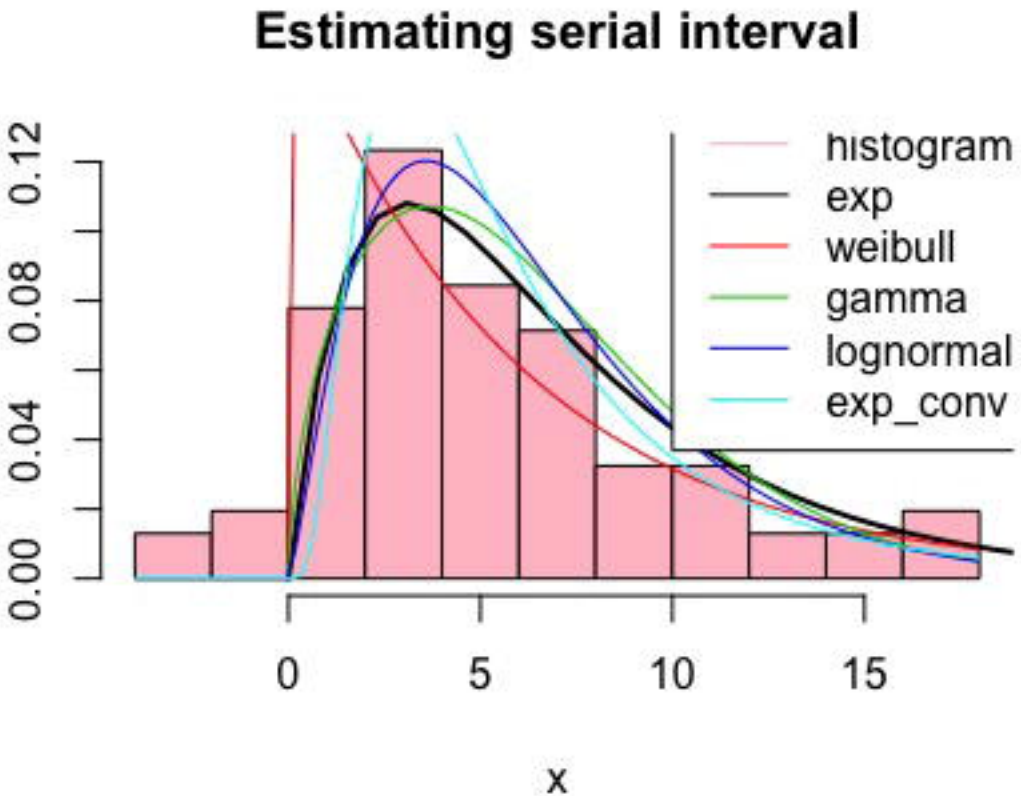

Mixiture of two sided exponential SI distribution
histogram
model1
model2
mixture
0.00
0.04
0.08
0.12
0
5
10
15
x

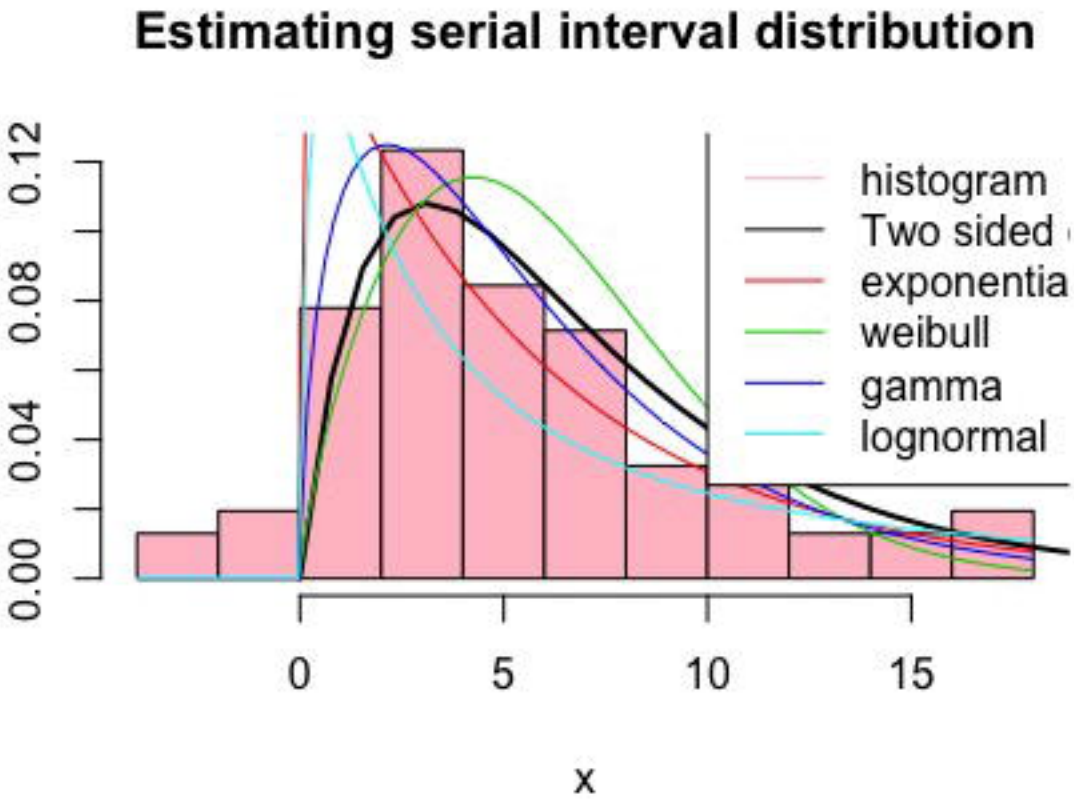

Mixture of weibul SI distributions
histogram
weibull1
weibull2
mixture
0.00
0.04
0.08
0.12
0
5
10
15
x

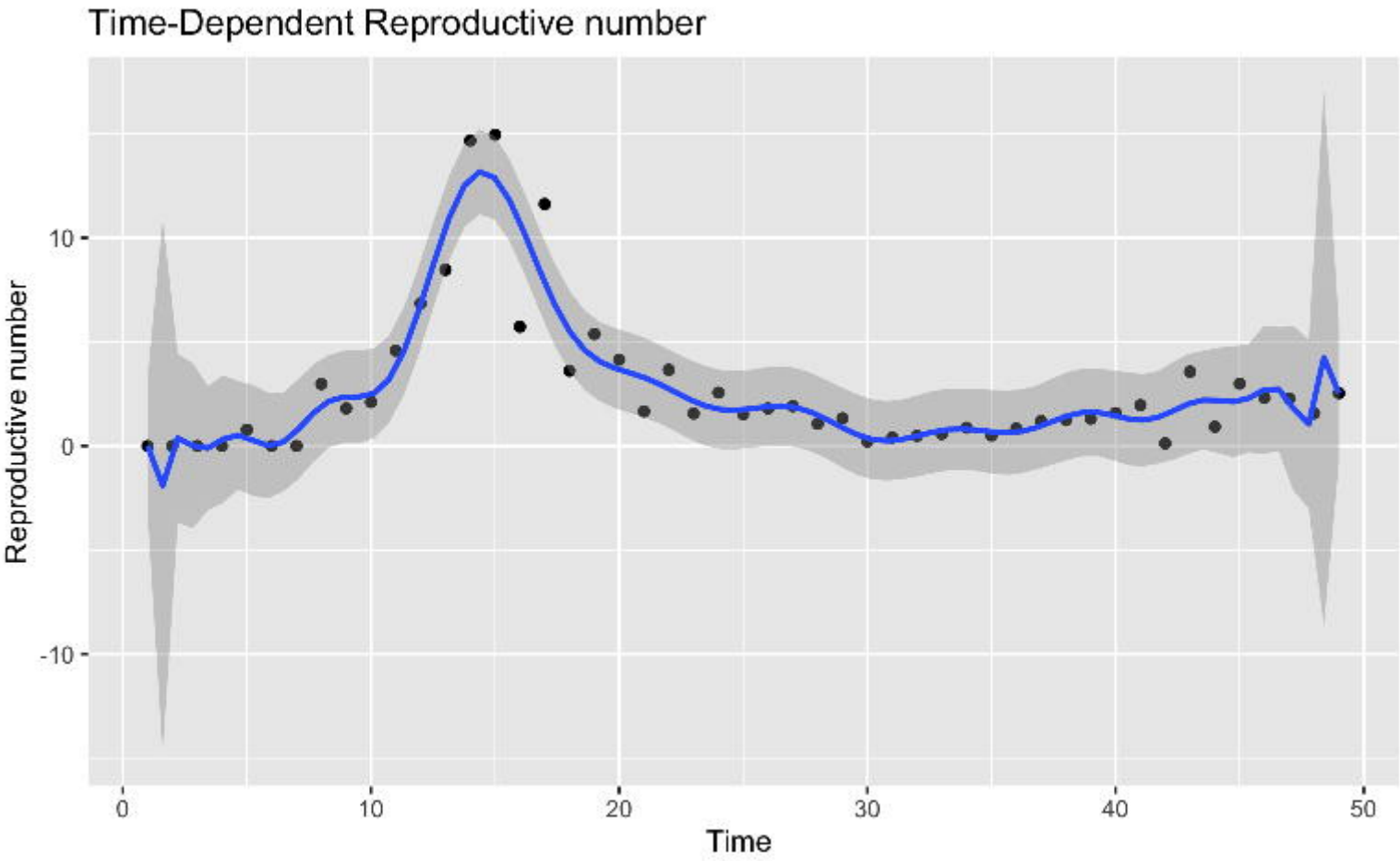

Mixture of lognormal SI distributions
histogram
lognorm1
lognorm2
mixture
0.00
0.04
0.08
0.12
0
5
10
15
x

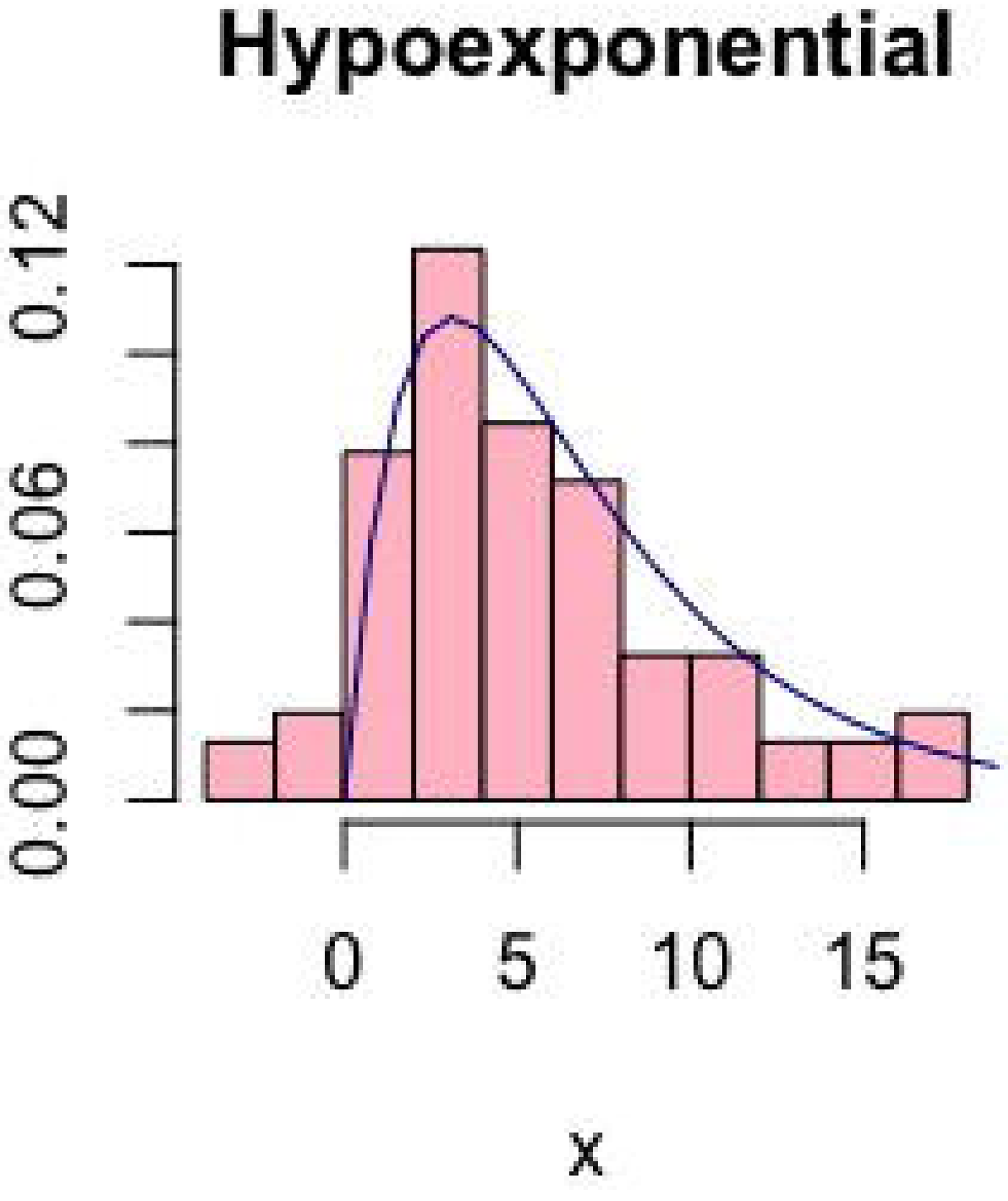


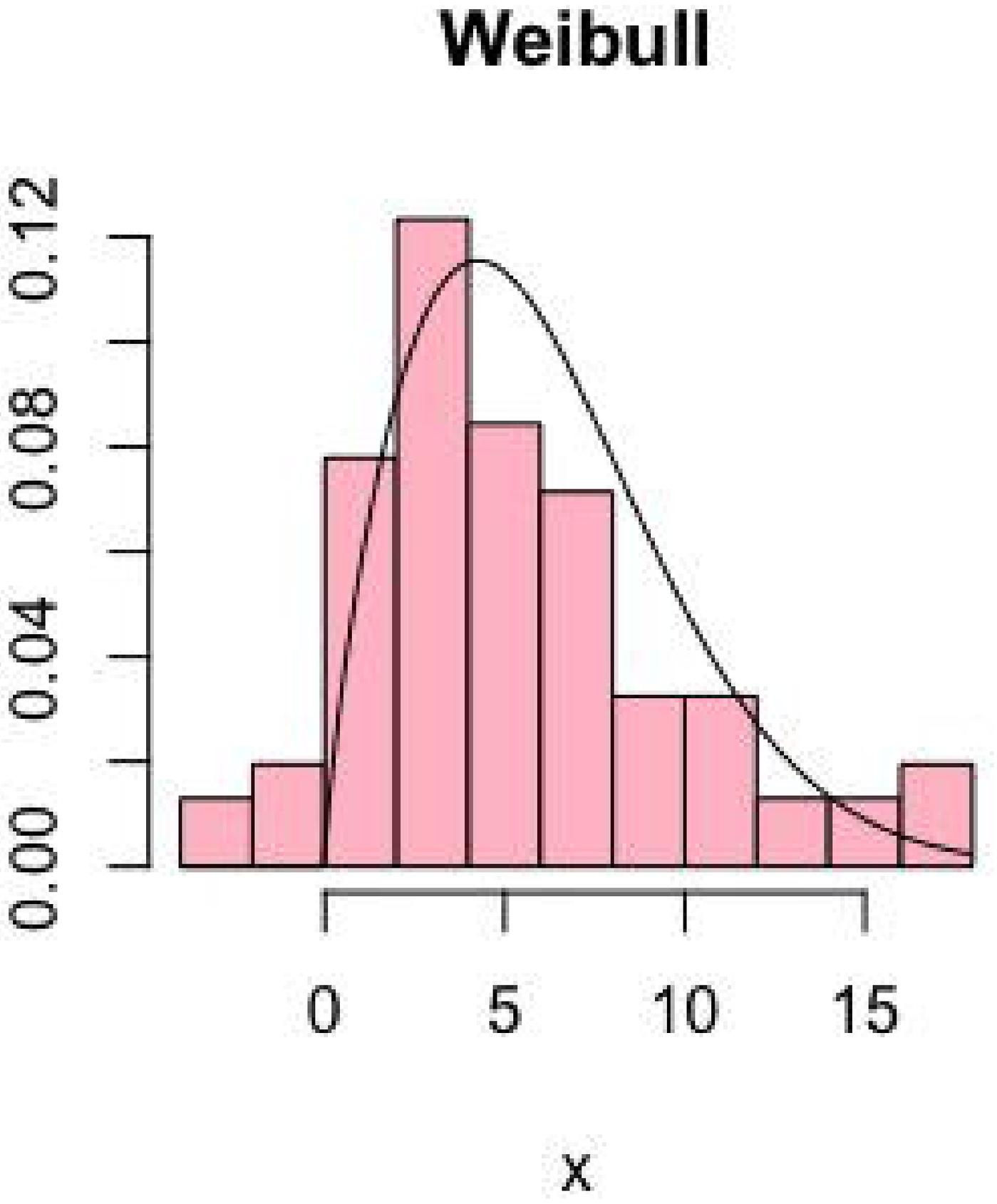

kture of Gamma distrib
0.00
0.20
0
10
x

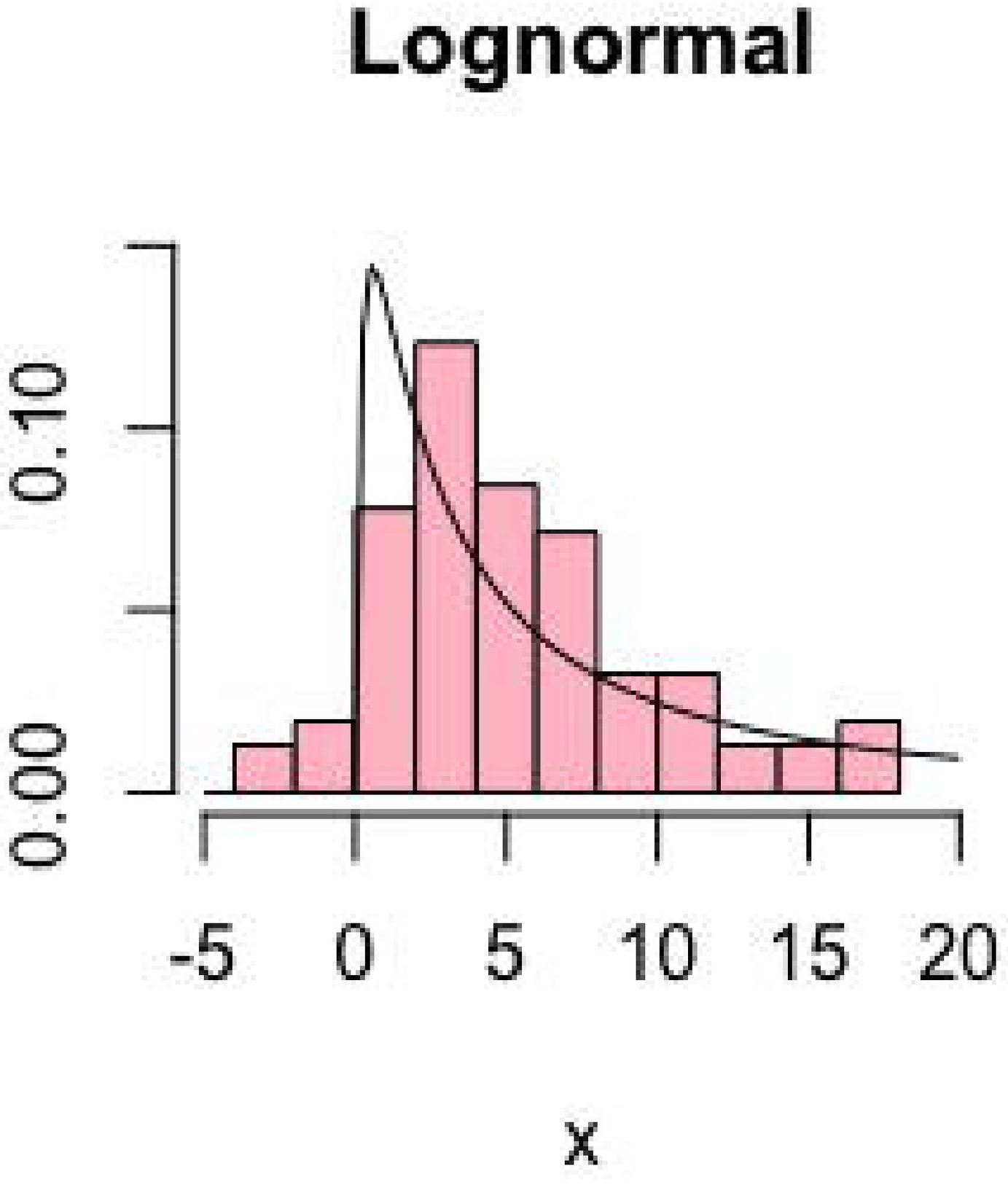

Estimating serial interval
histogram
exp
weibull
gamma
lognormal
exp_conv
0.00
0.04
0.08
0.12
0
5
10
15
x

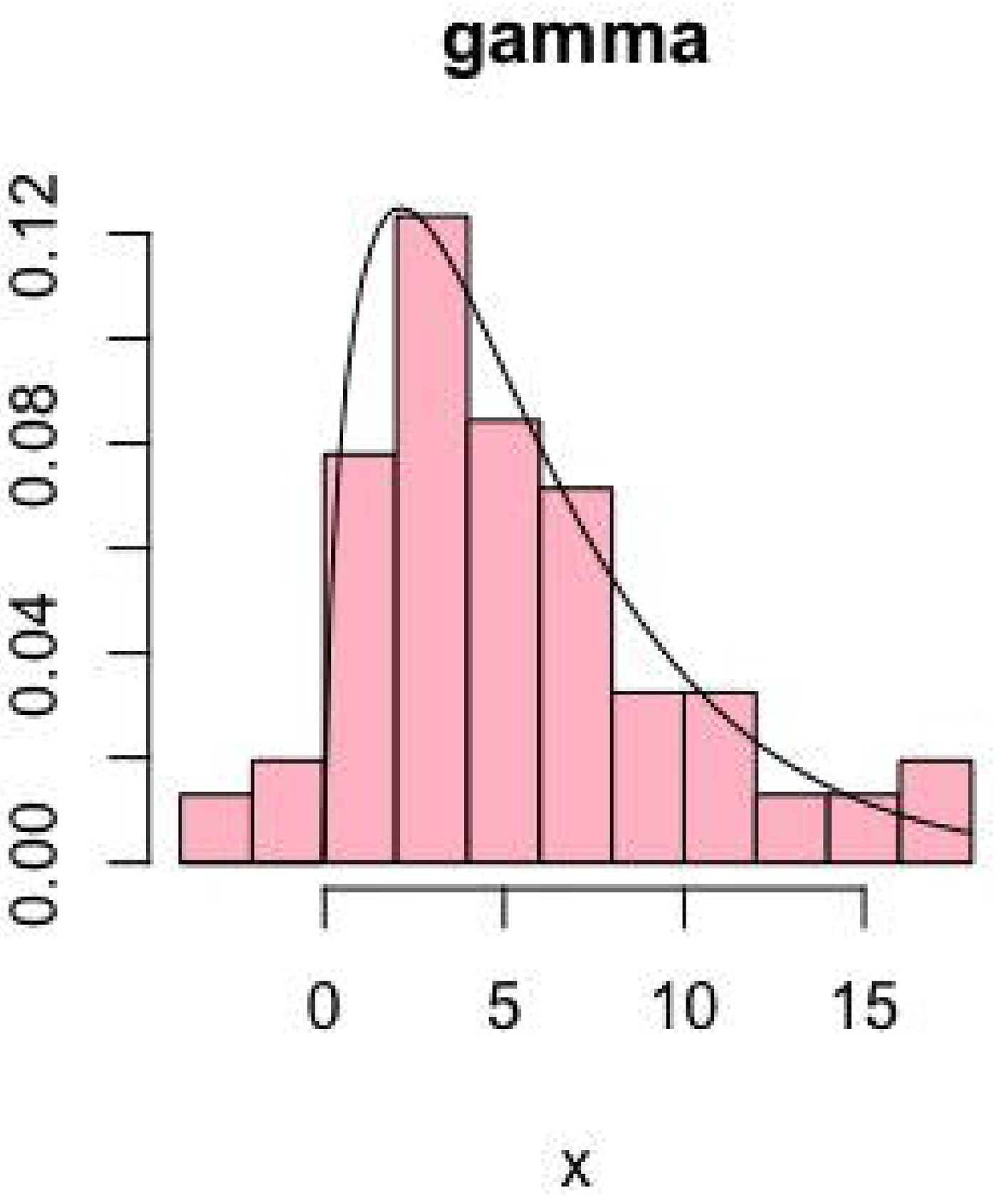

Estimating serial interval distribution
0.00
0.04
0.08
0.12
0
5
10
15
x
histogram
Two sided
exponentia
weibull
gamma
lognormal

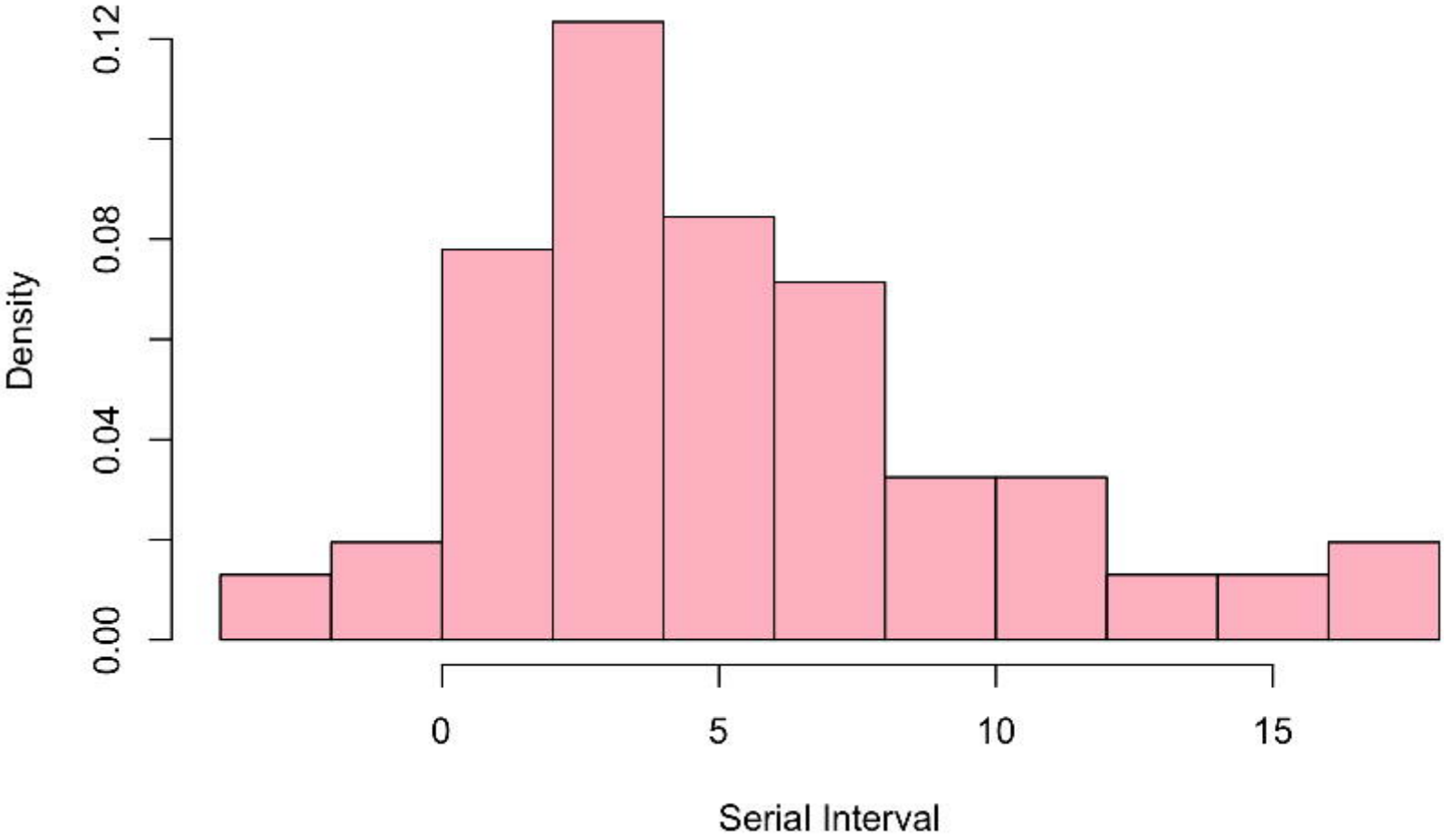

Time-Dependent Reproductive number
Reproductive number
10
0
-10
0
10
20
30
40
50
Time

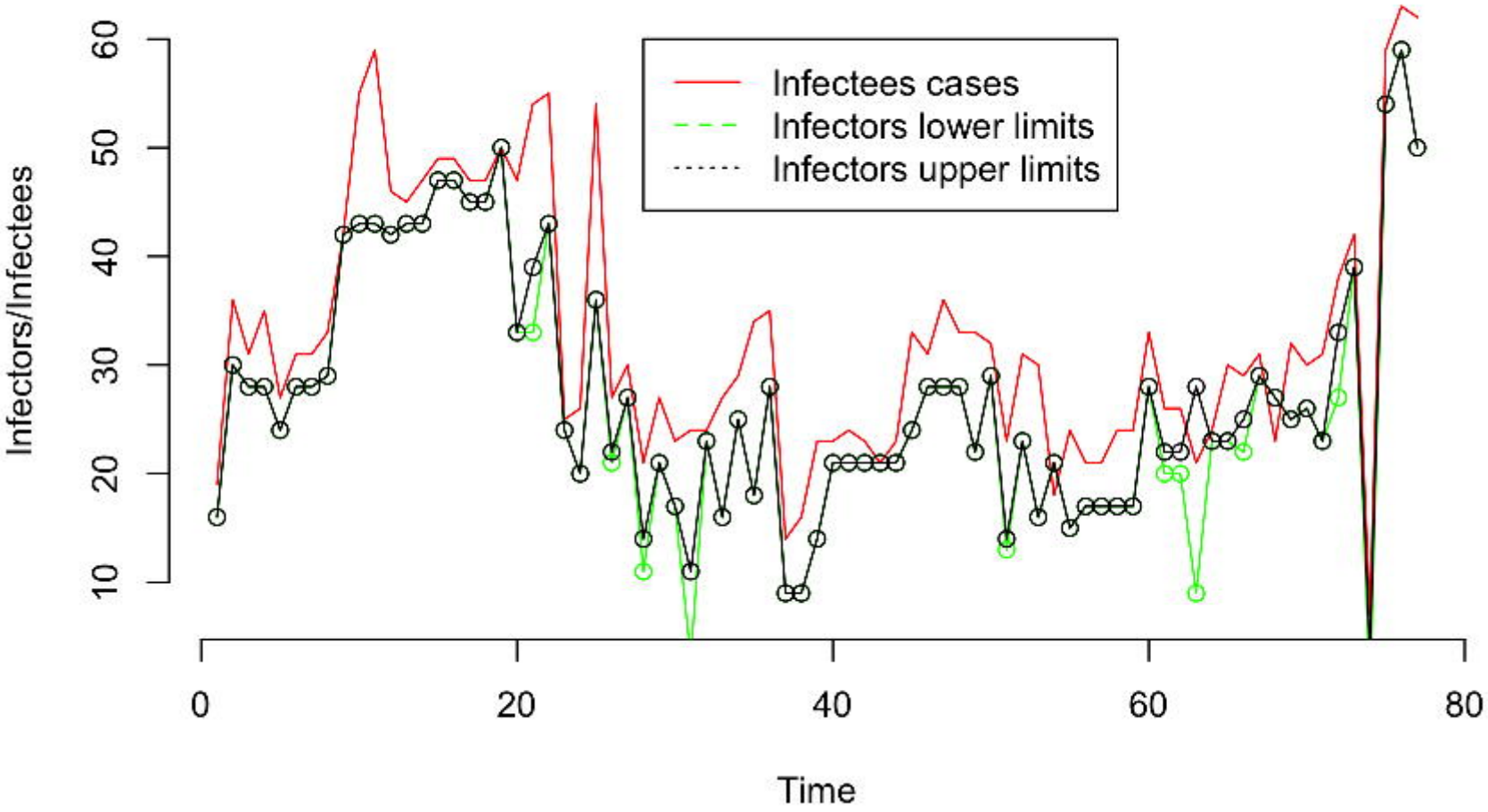

Hypoexponential
0.00
0.06
0.12
0
5
10
15
x

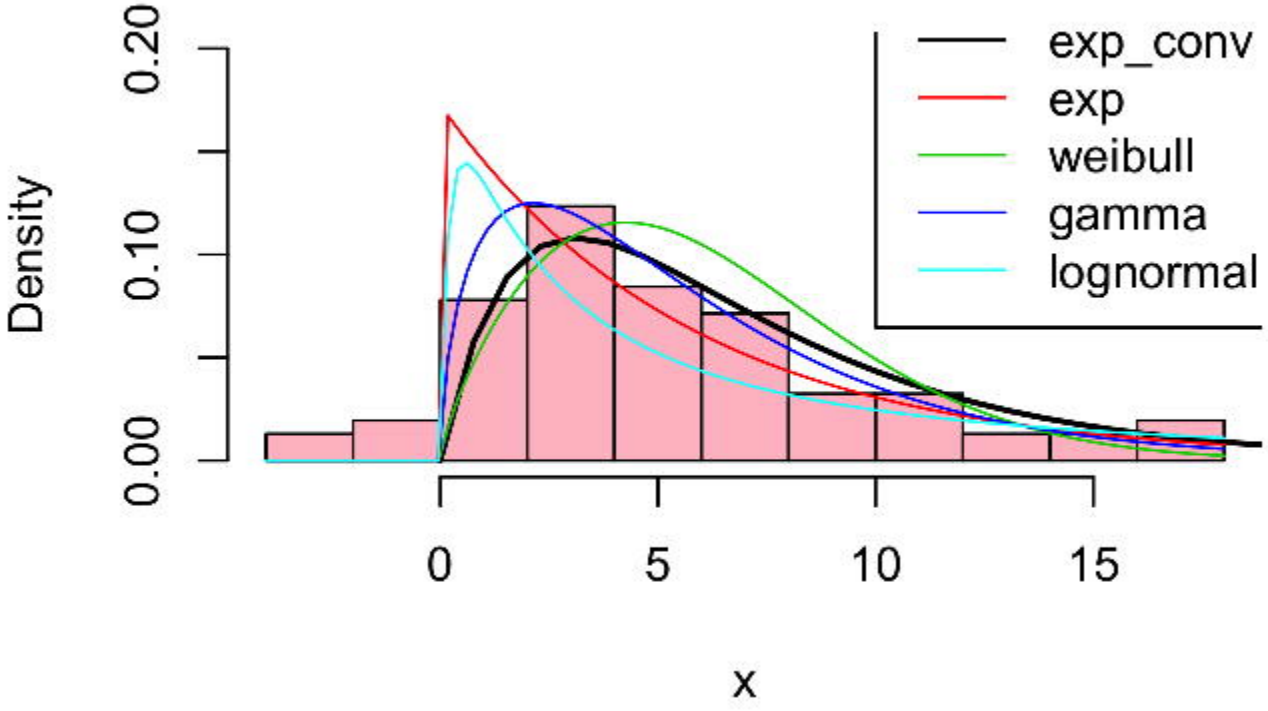

Weibull
0.00
0.04
0.08
0.12
0
5
10
15
x

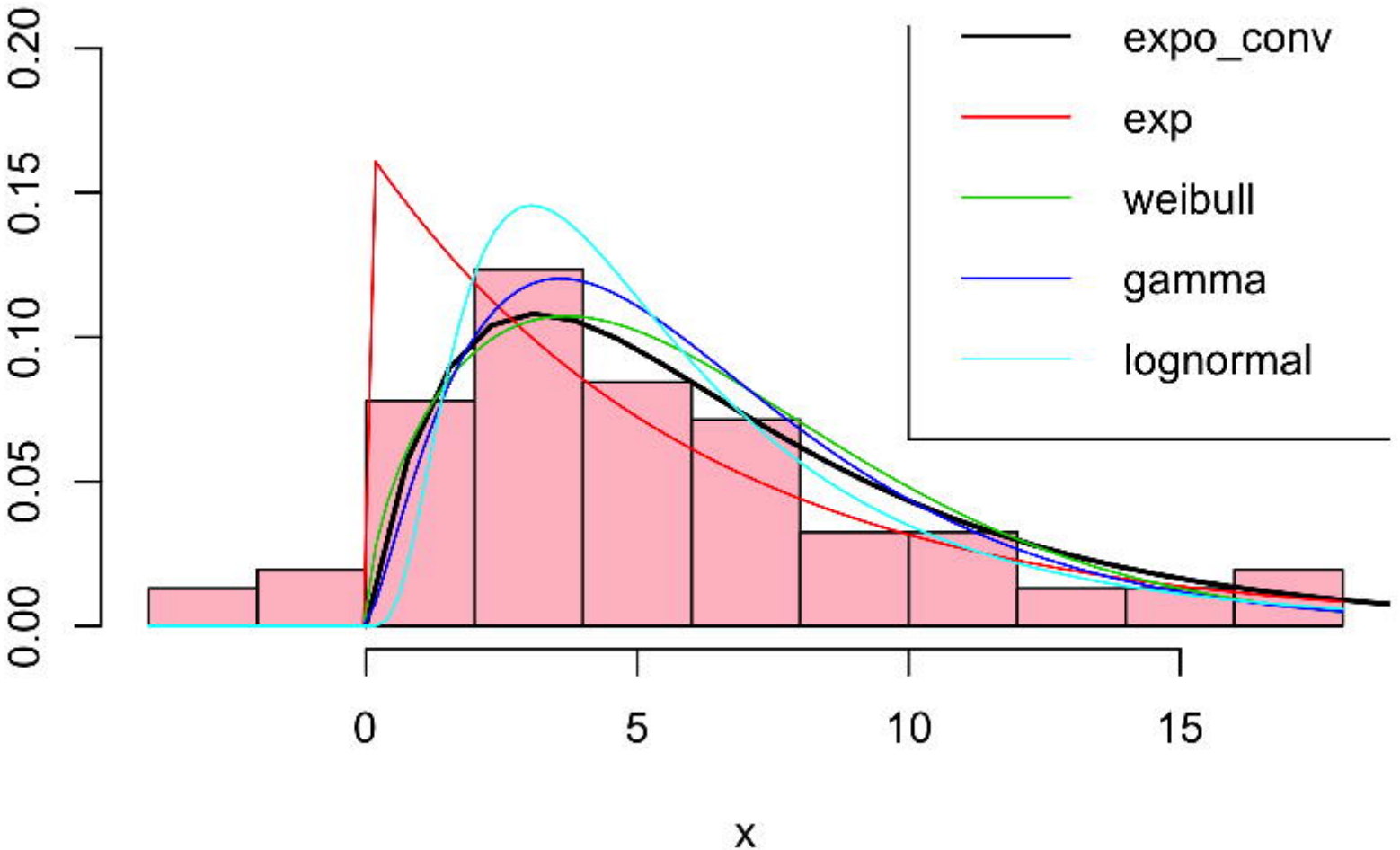

Lognormal
0.00
0.10
-5
0
5
10
15
20
x

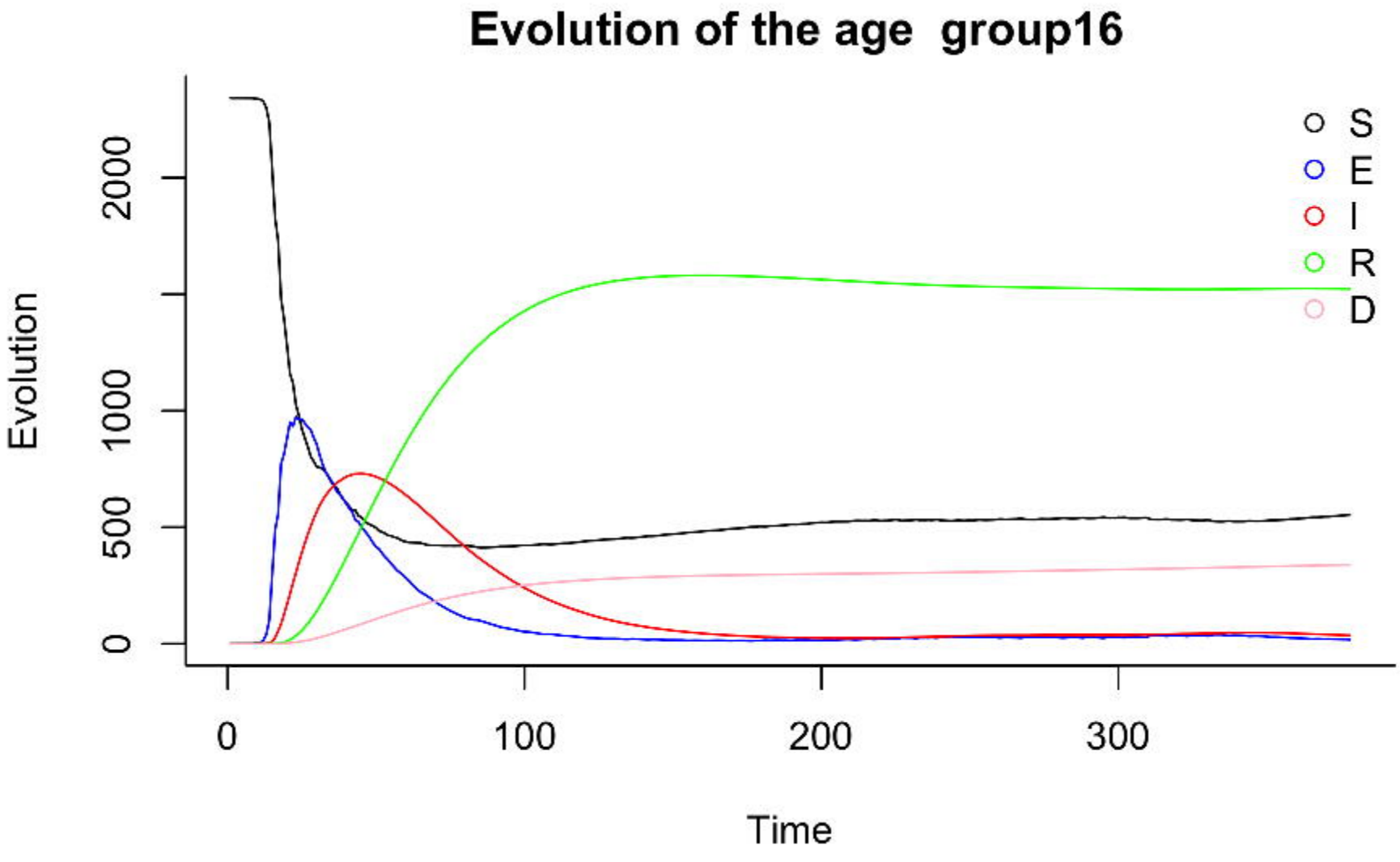

gamma
0.00
0.04
0.08
0.12
0
5
10
15
x

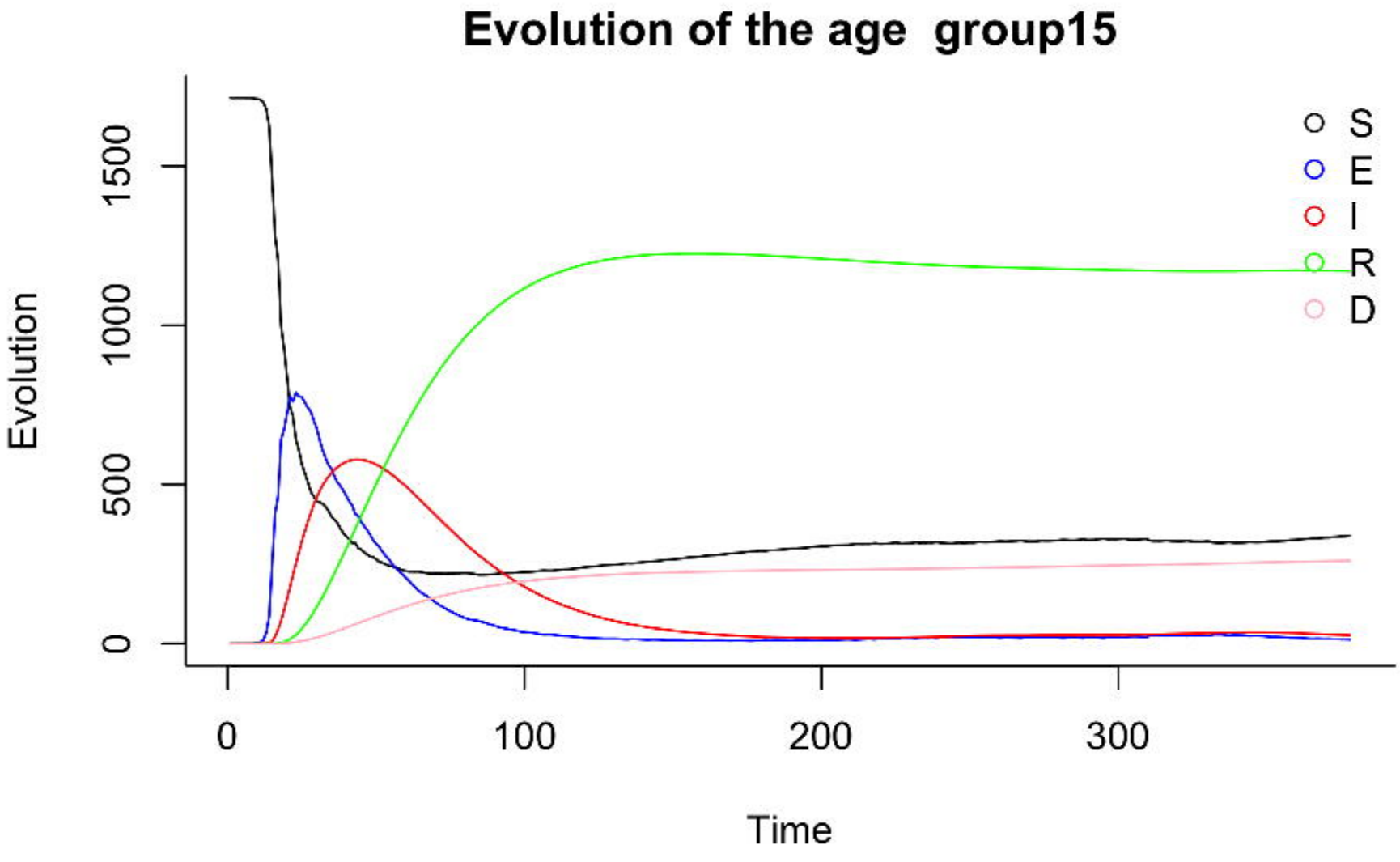

Density
0.00
0.04
0.08
0.12
0
5
10
15
Serial Interval

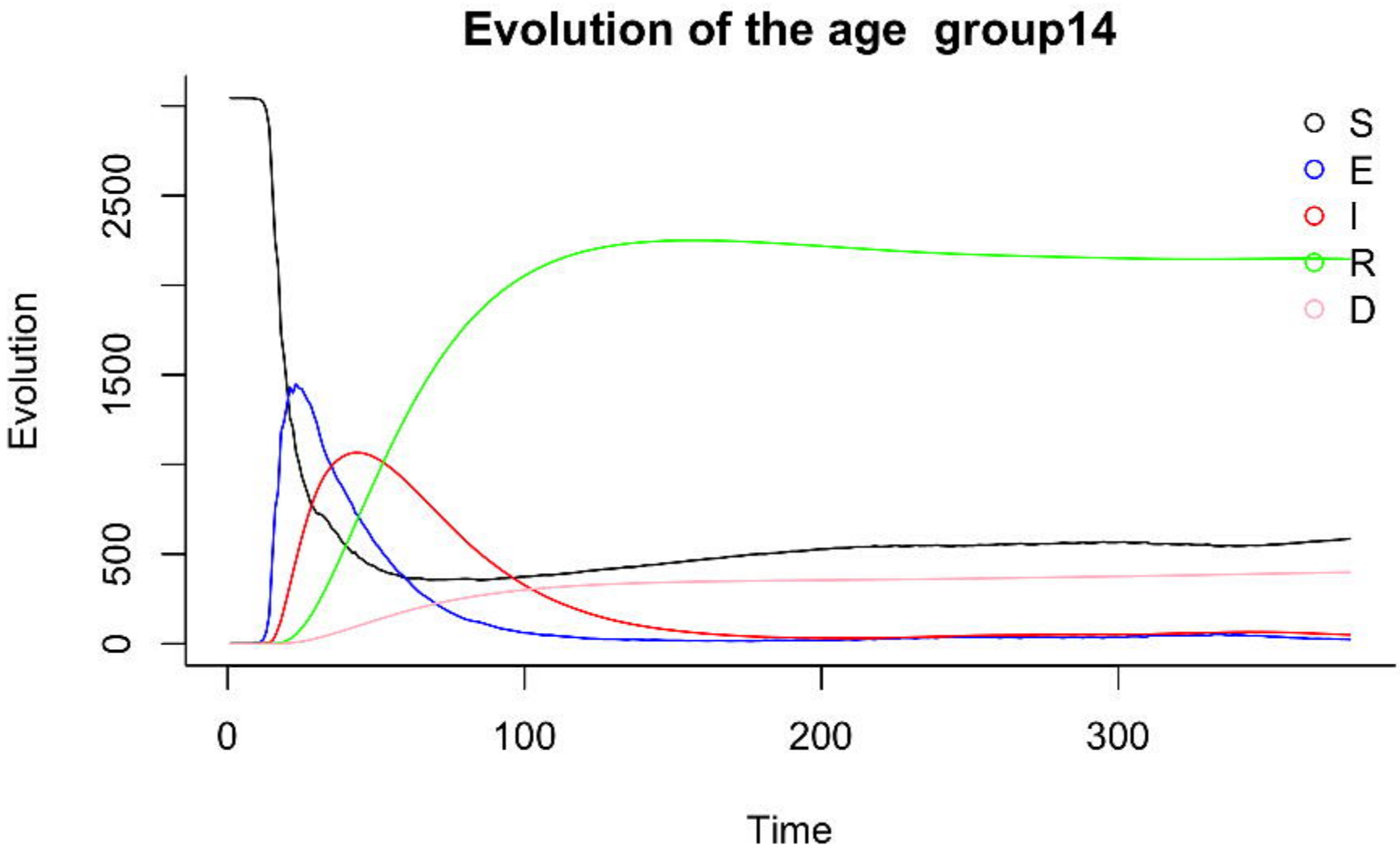

Infectees cases
Infectors lower limits
Infectors upper limits
Infectors/Infectees
10
20
30
40
50
60
0
20
40
60
80
Time

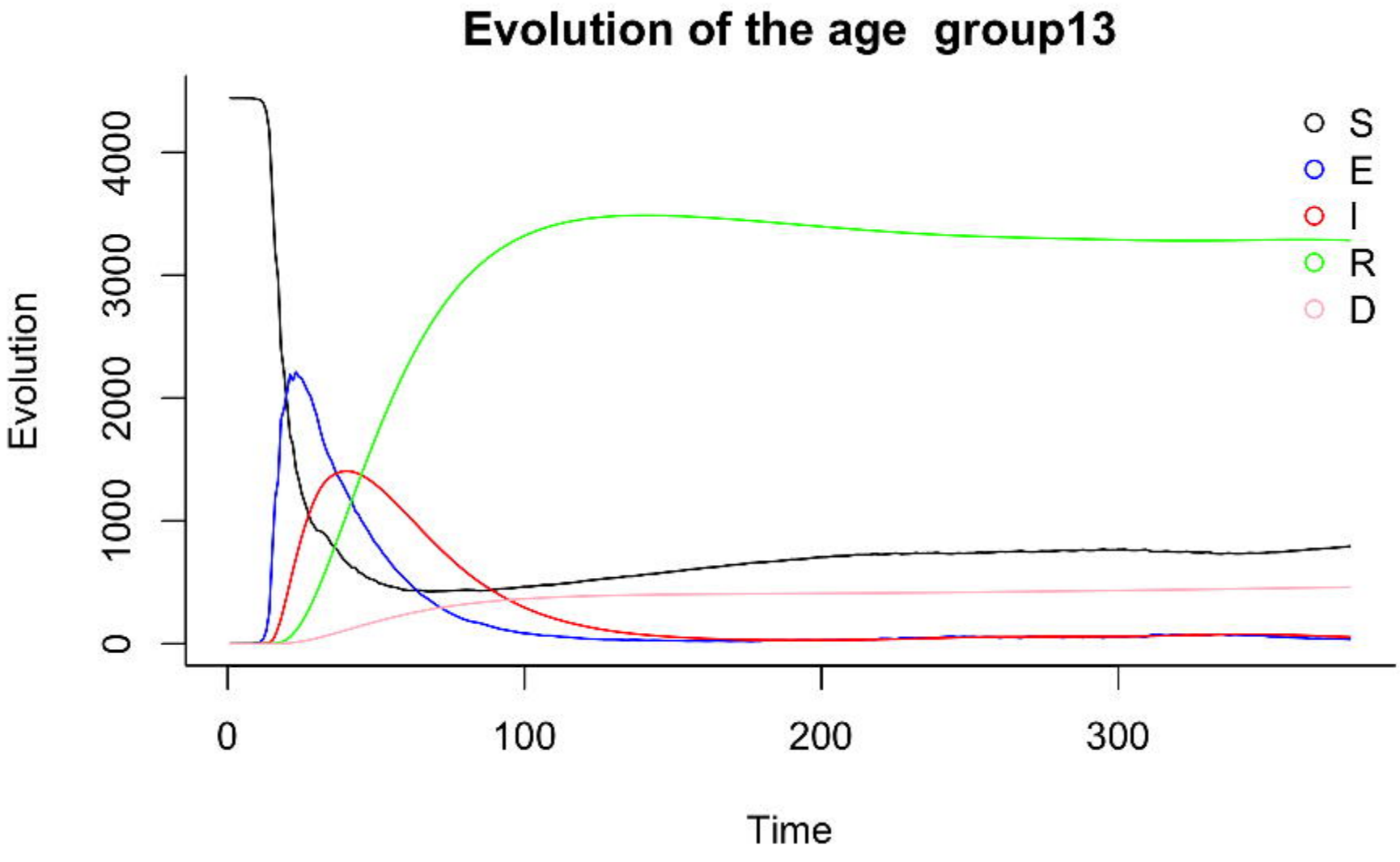

exp_conv
exp
weibull
gamma
lognormal
Density
0.00
0.10
0.20
0
5
10
15
x

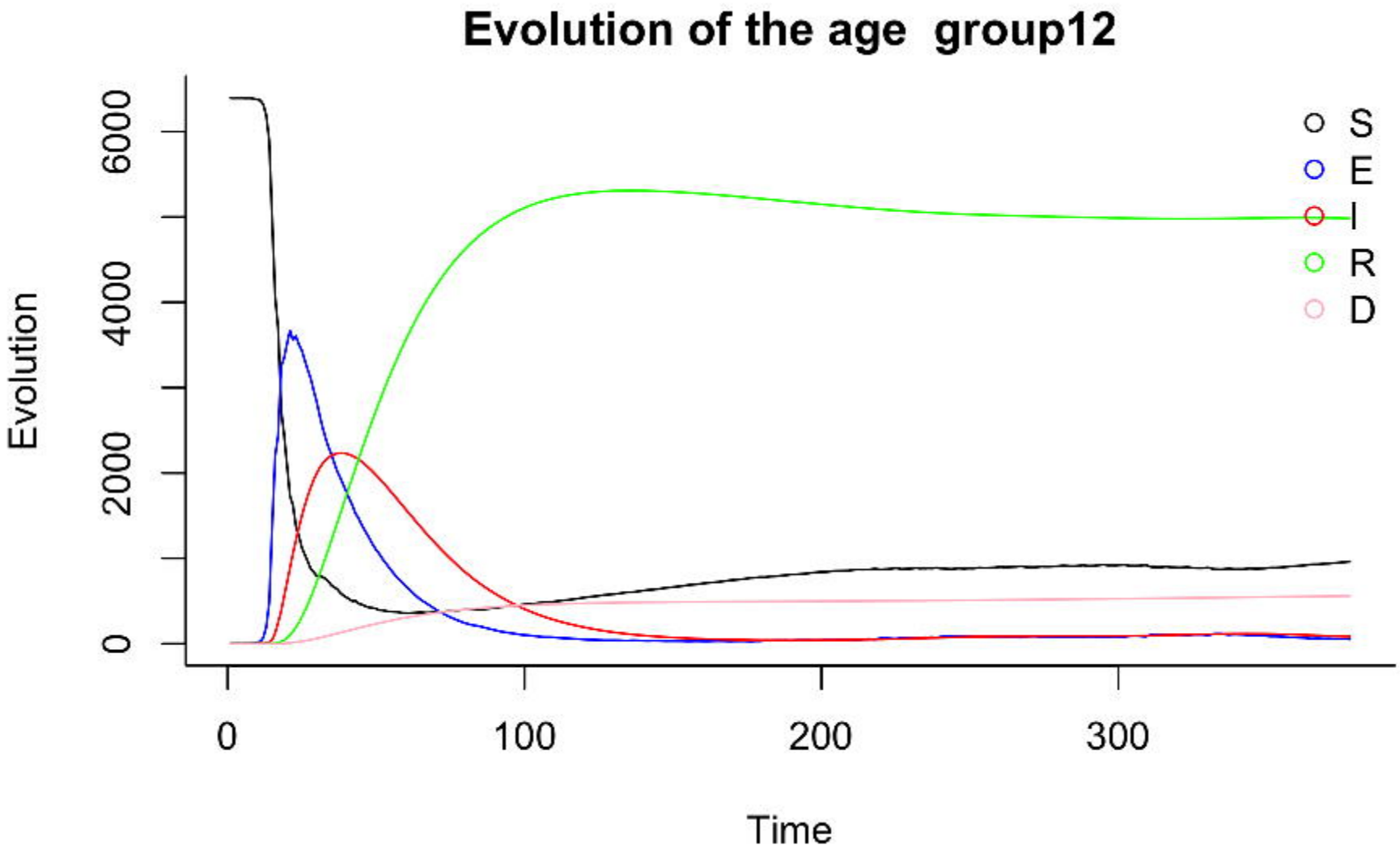

expo_conv
exp
weibull
gamma
lognormal
0.00
0.05
0.10
0.15
0.20
0
5
10
15
x

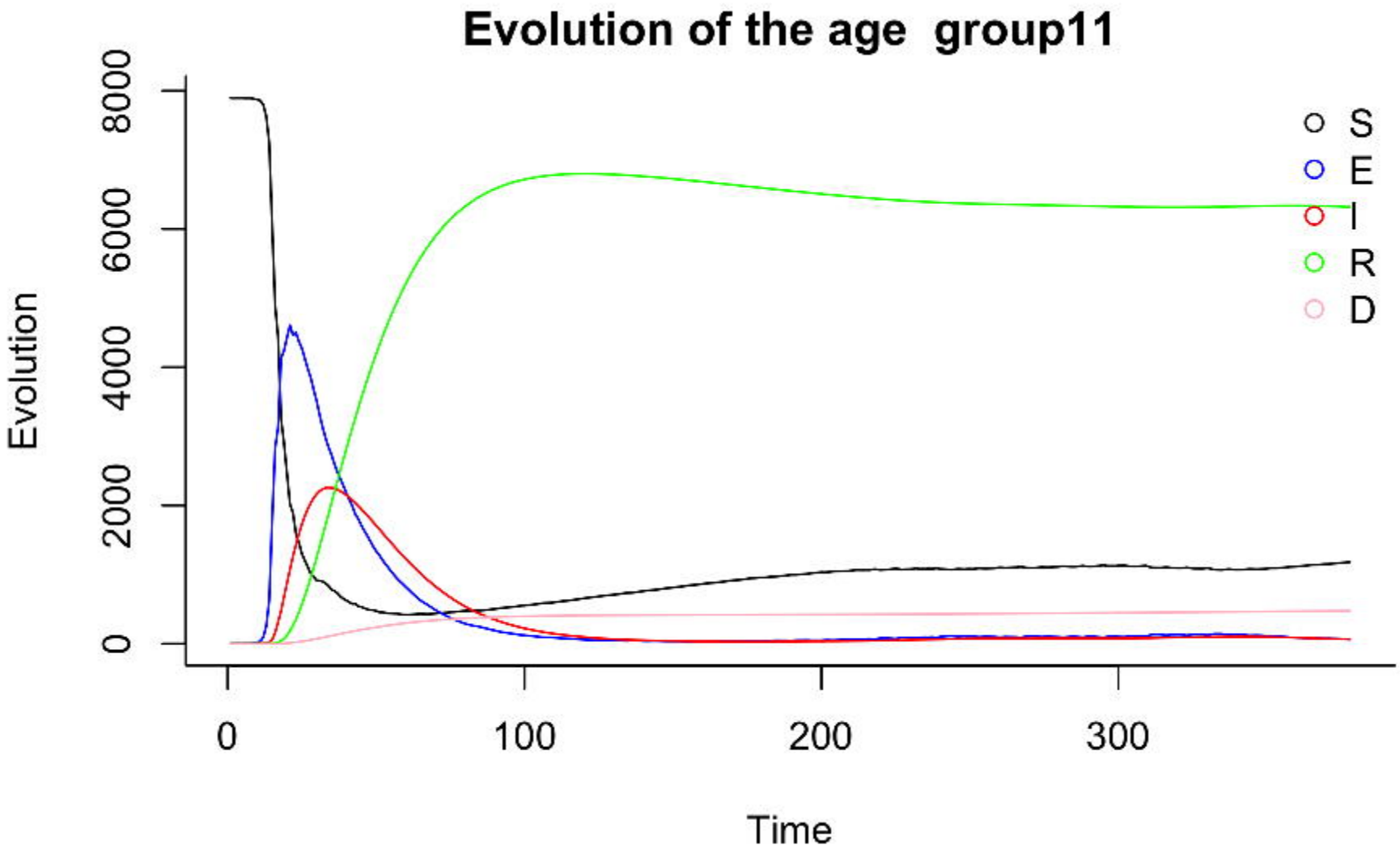

Evolution of the age group16
S
E
I
R
D
Evolution
0
500
1000
2000
0
100
200
300
Time

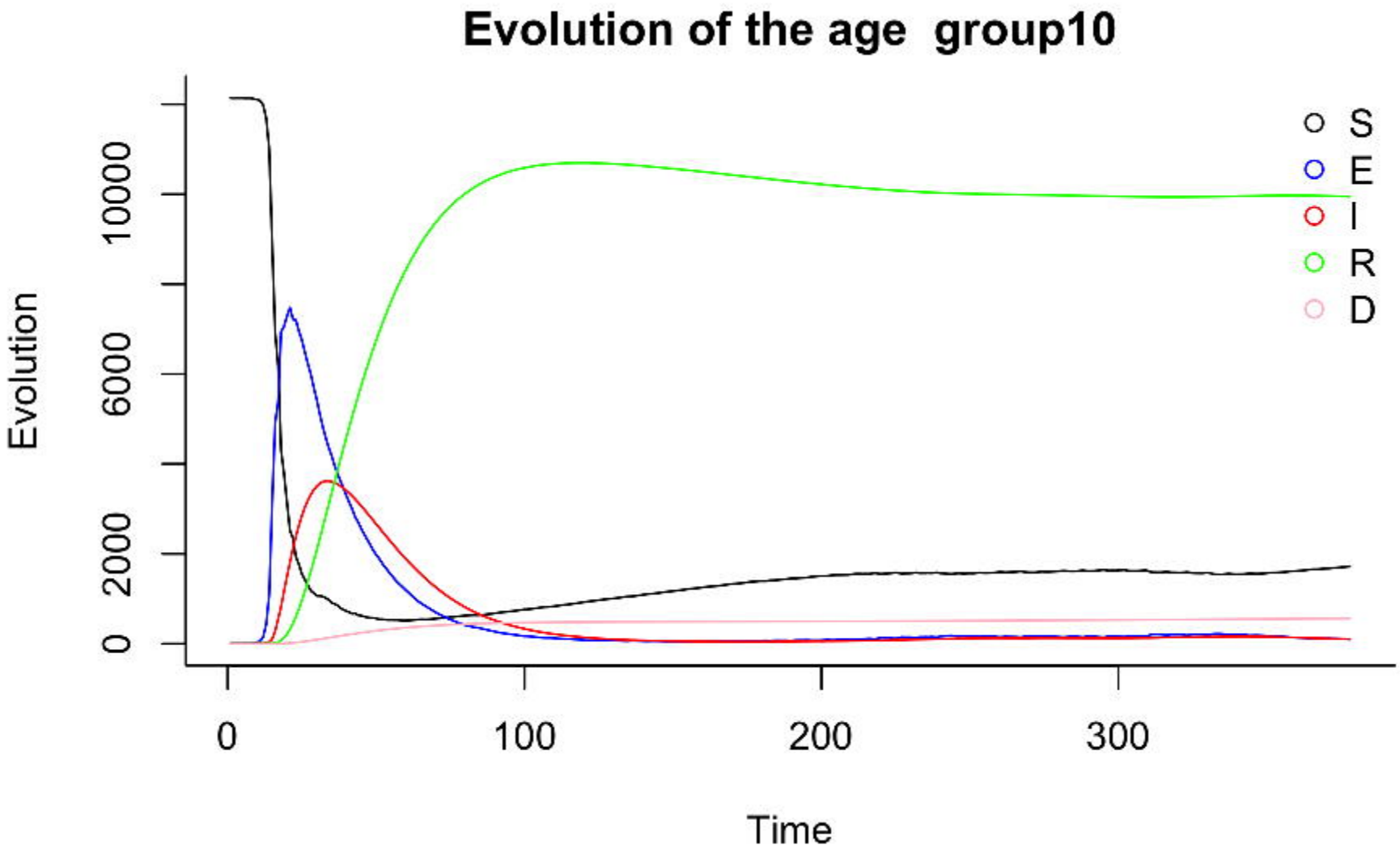

Evolution of the age group15
S
E
I
R
D
Evolution
0
500
1000
1500
0
100
200
300
Time

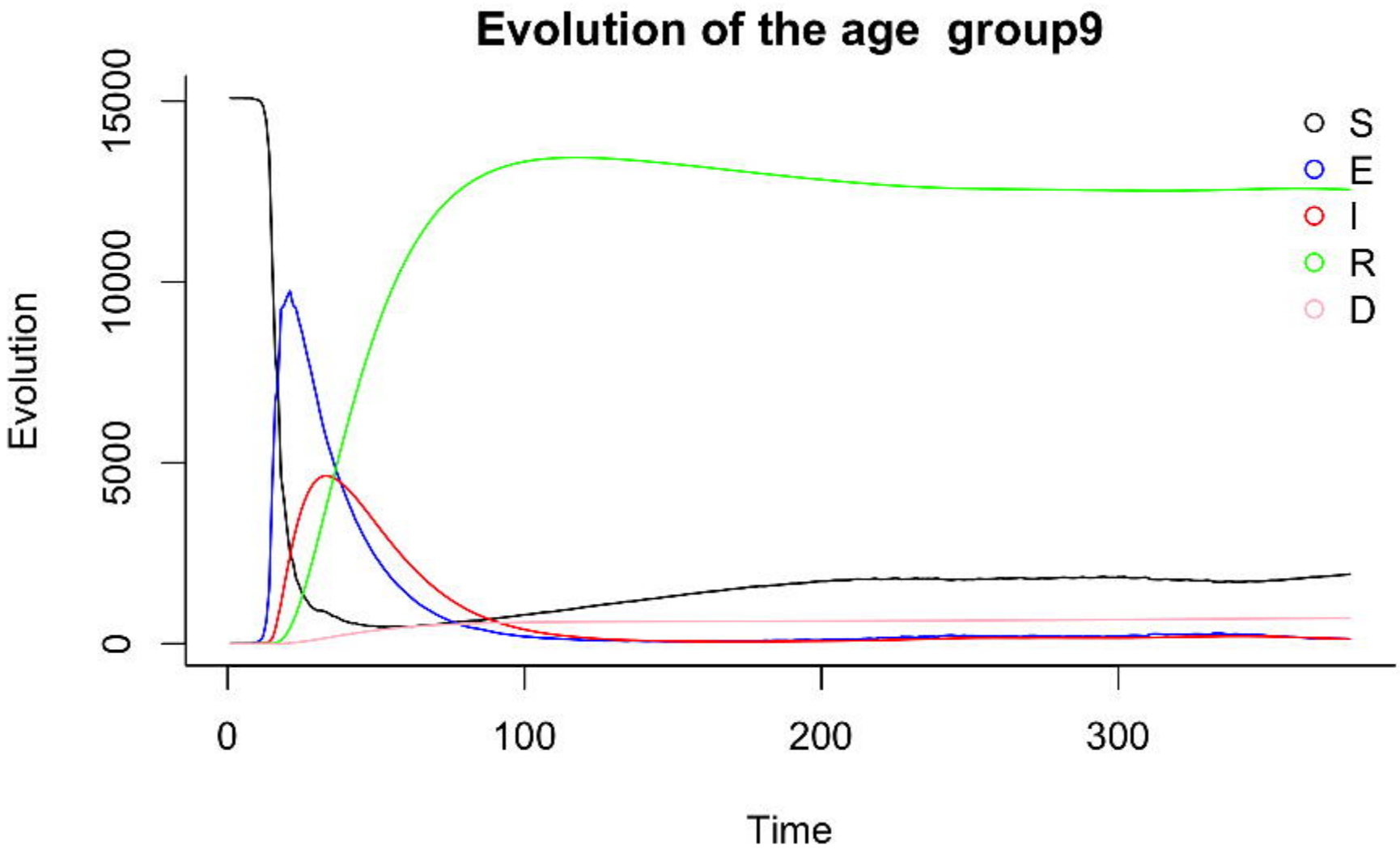

Evolution of the age group14
S
E
I
R
D
Evolution
0
500
1500
2500
0
100
200
300
Time

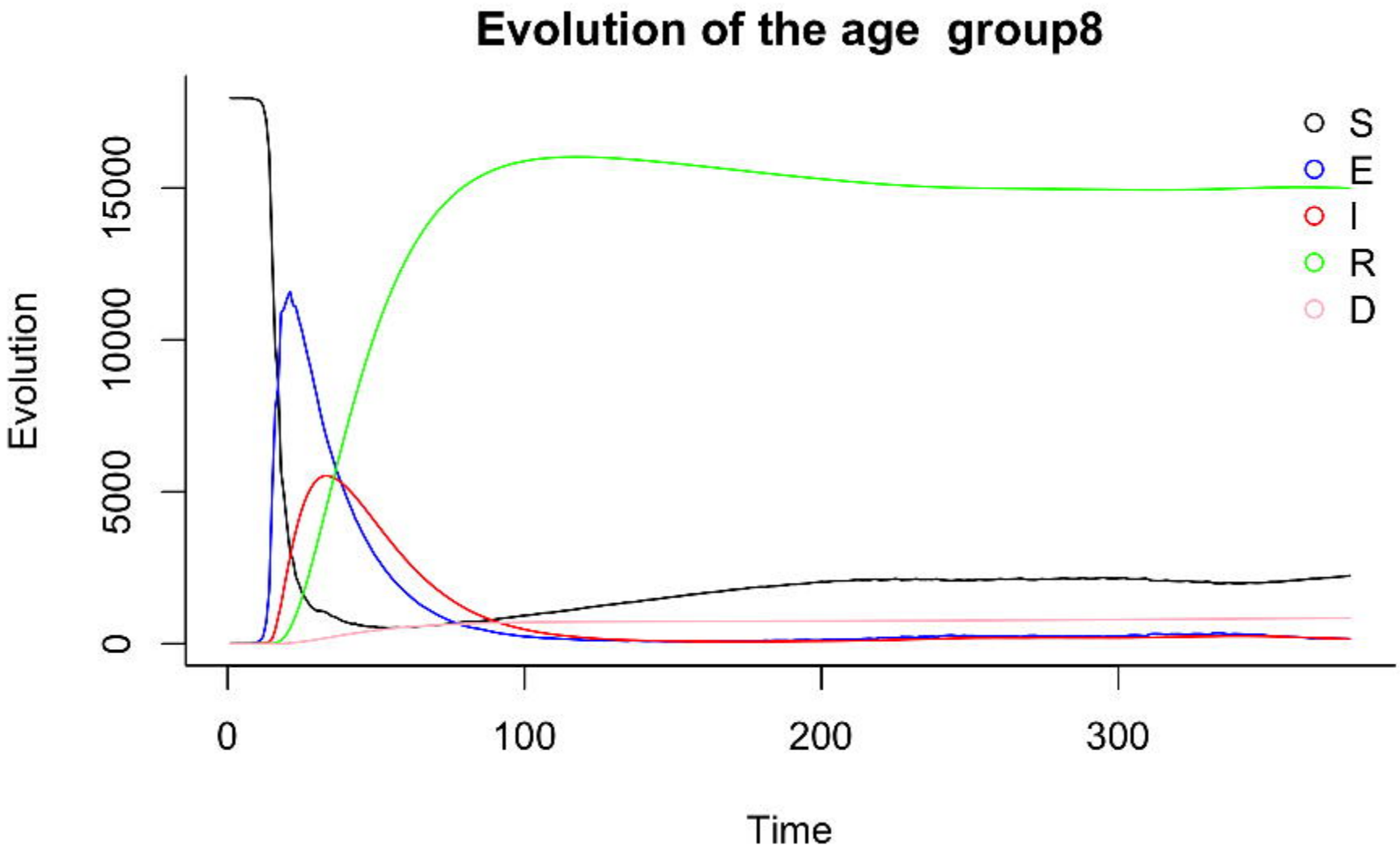

Evolution of the age group13
Evolution
0
1000
2000
3000
4000
0
100
200
300
Time
S
E
I
R
D

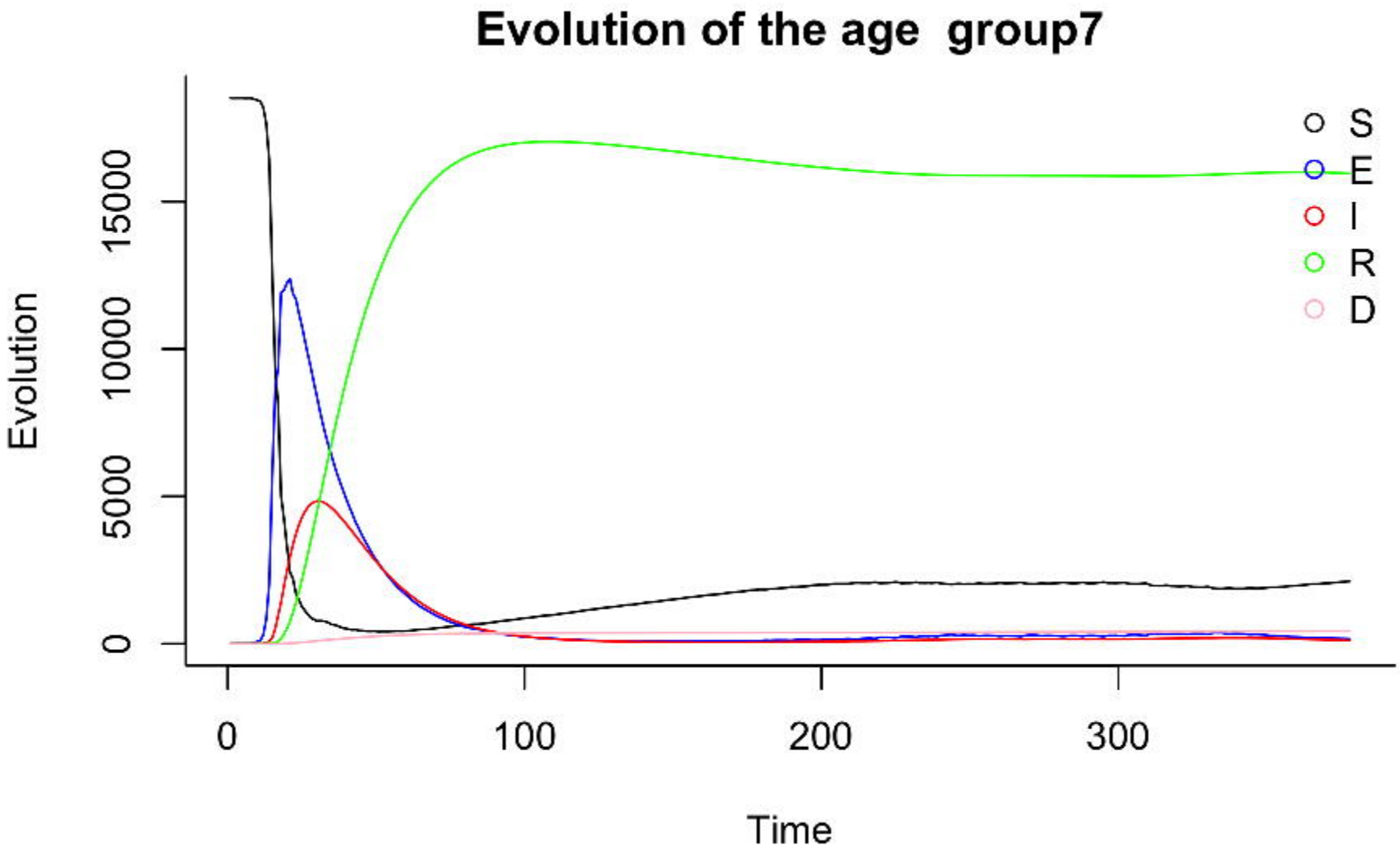

Evolution of the age group12
Evolution
Time
0
2000
4000
6000
0
100
200
300
S
E
I
R
D

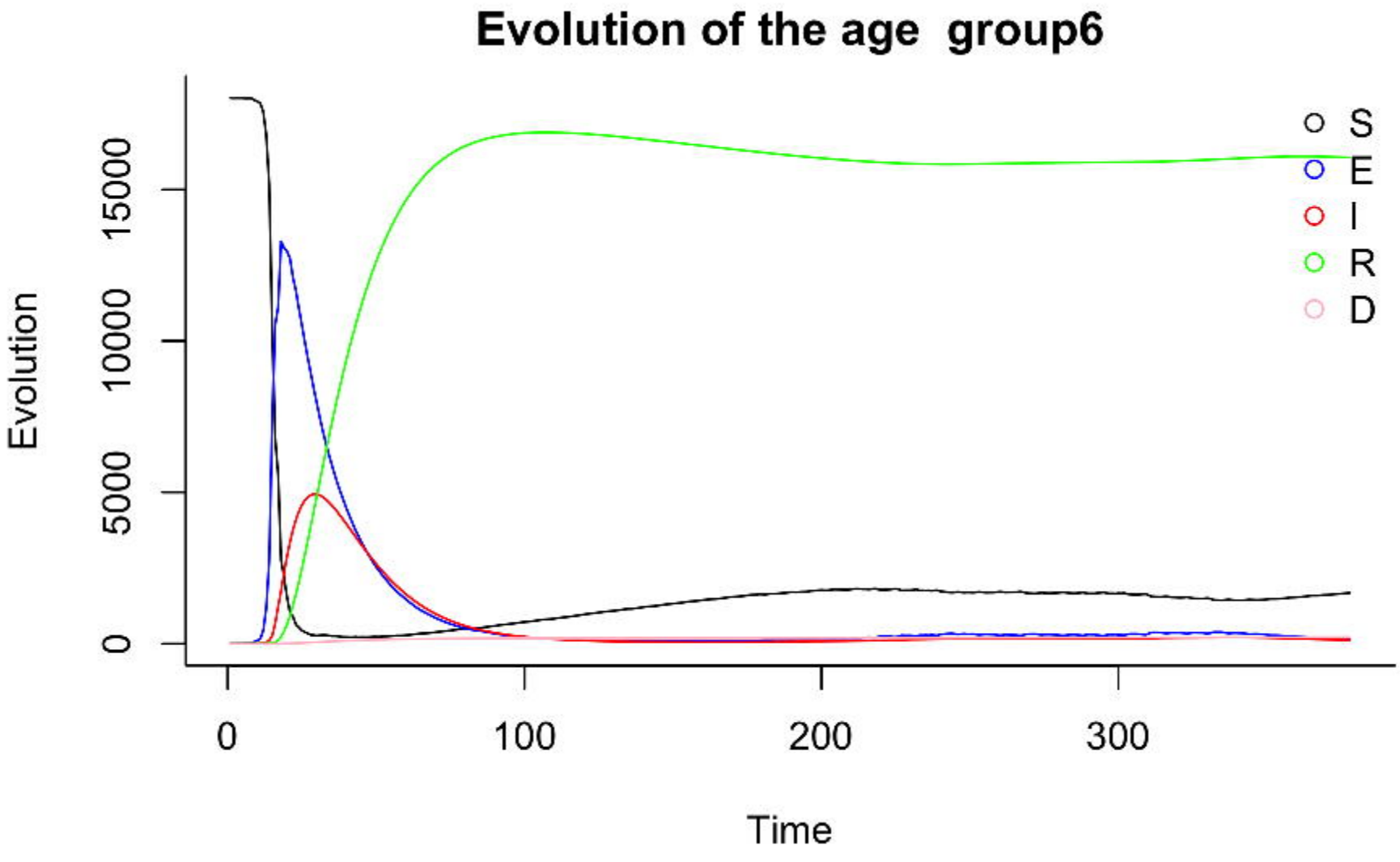

Evolution of the age group11
Evolution
Time
0
2000
4000
6000
8000
0
100
200
300
S
E
I
R
D

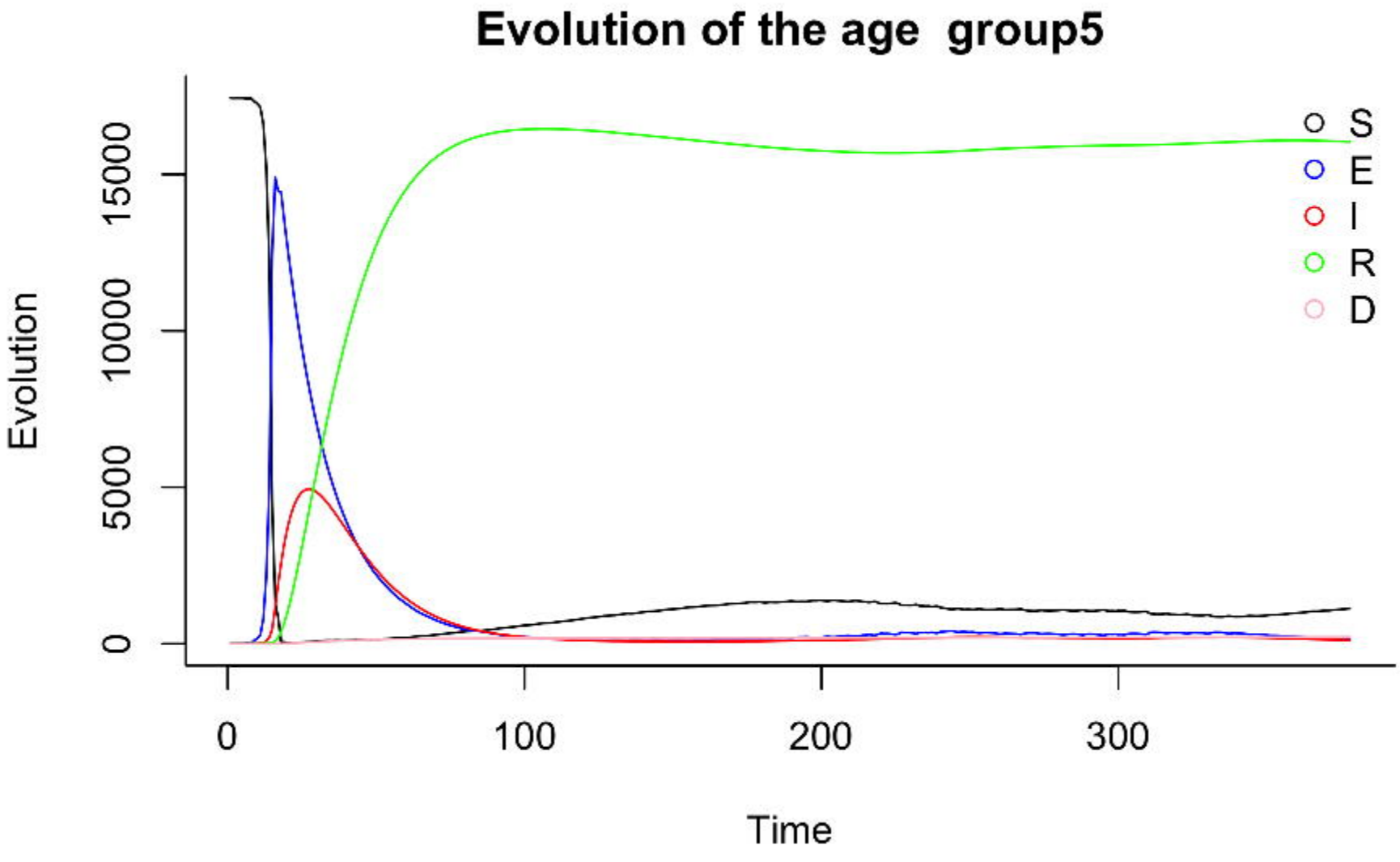

Evolution of the age group10
S
E
I
R
D
Evolution
0
2000
6000
10000
0
100
200
300
Time

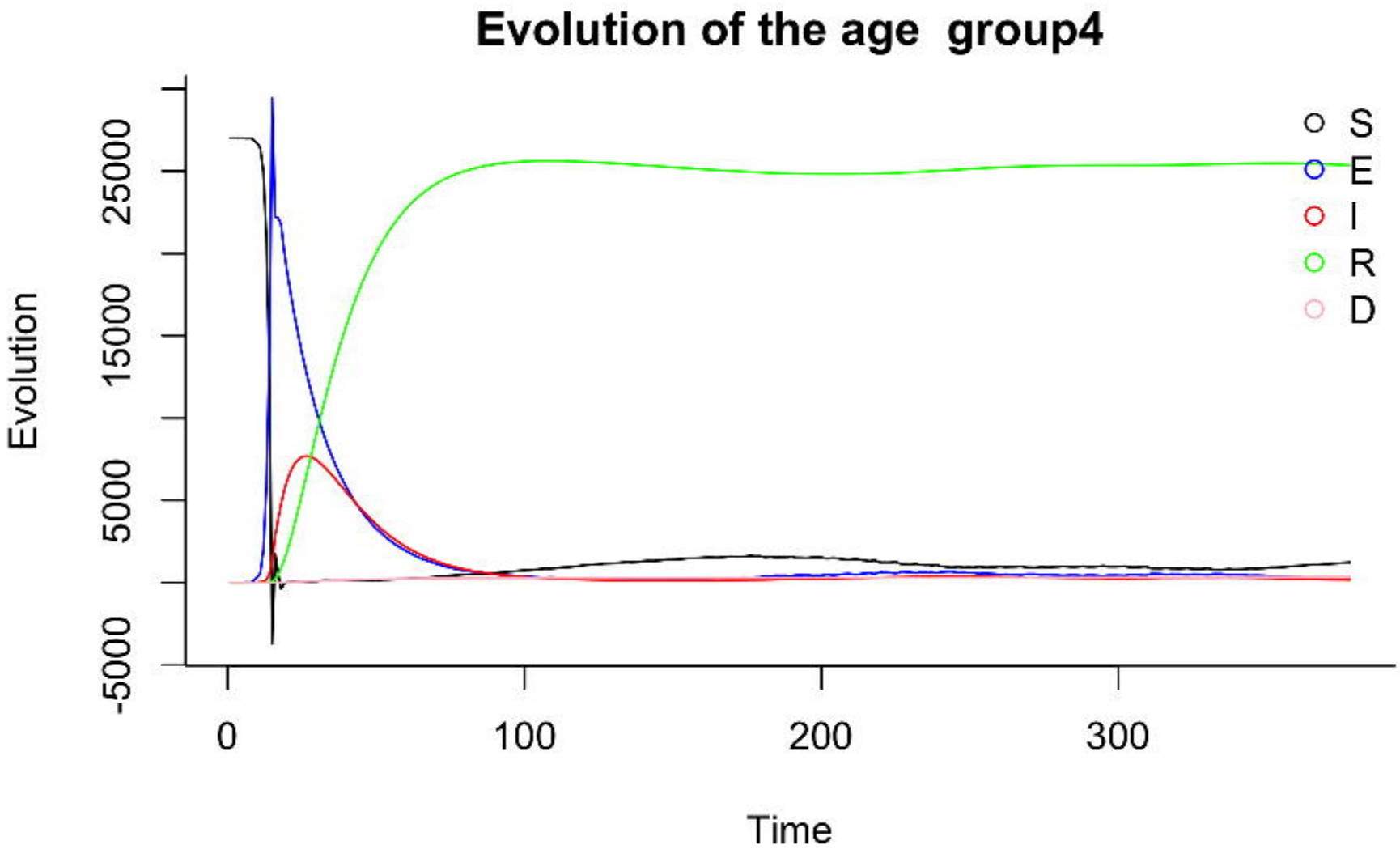

Evolution of the age group9
Evolution
15000
10000
5000
0
0
100
200
300
Time
S
E
I
R
D

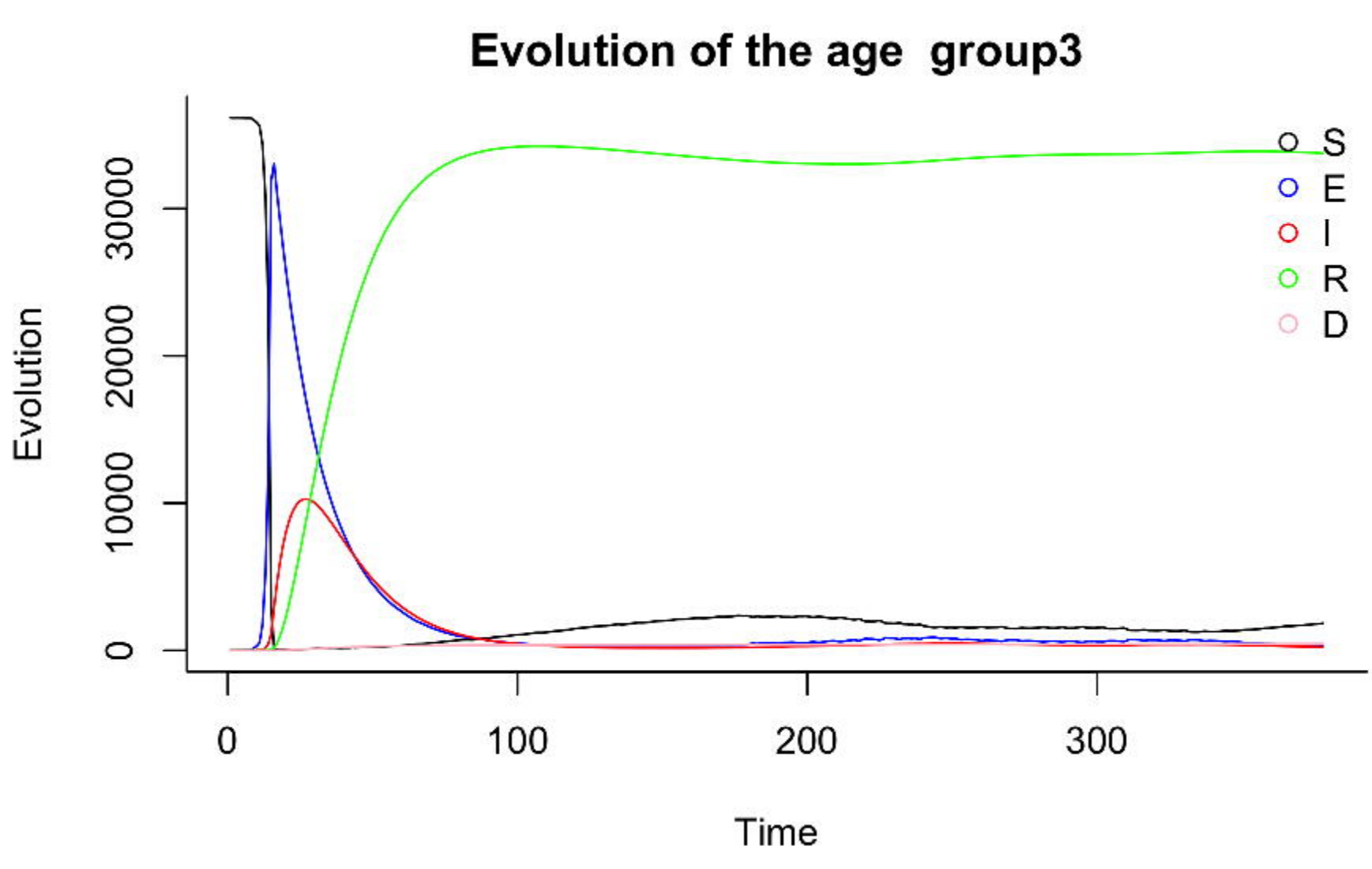

Evolution of the age group8
Evolution
0
5000
10000
15000
0
100
200
300
Time
S
E
I
R
D

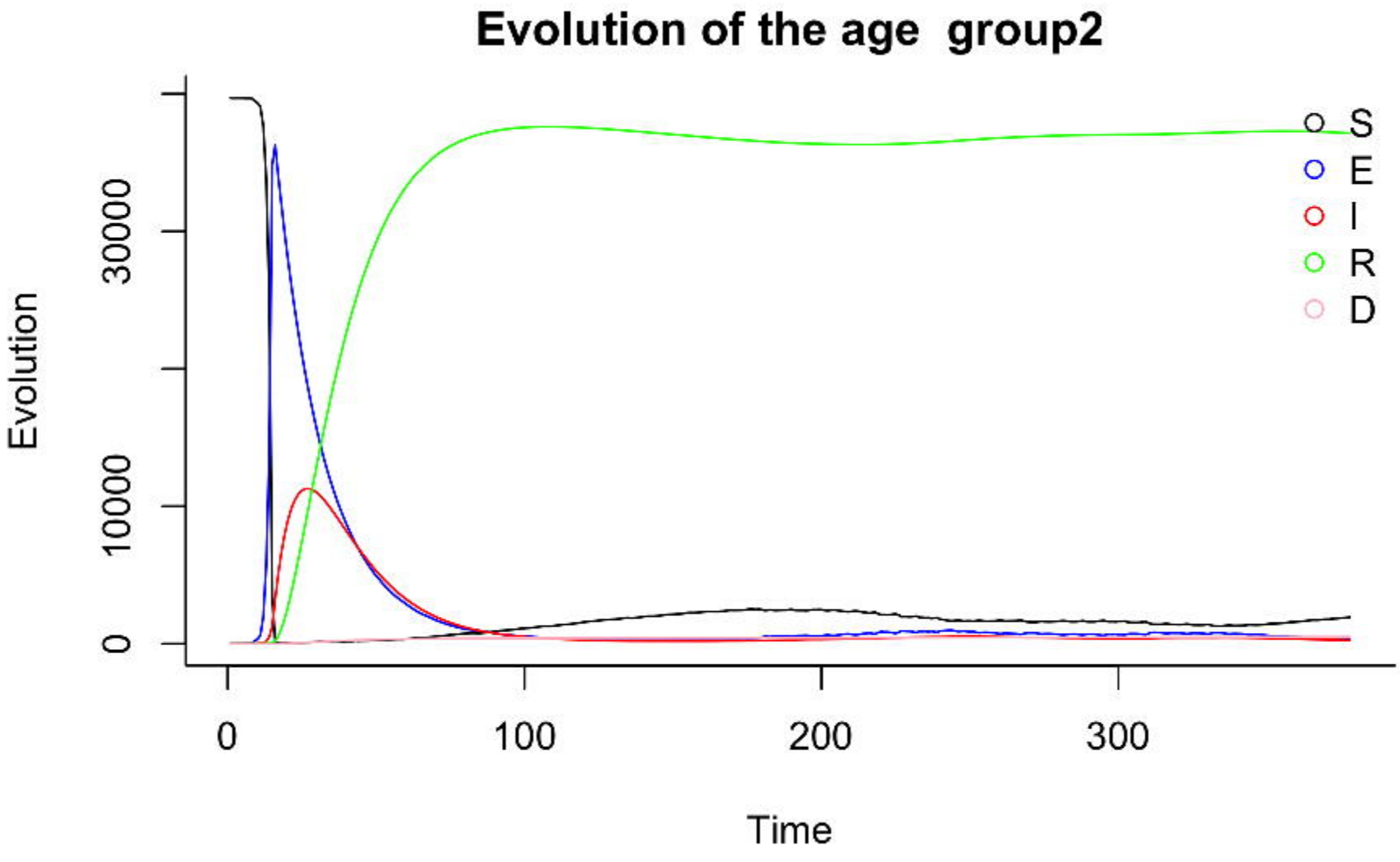

Evolution of the age group7
S
E
I
R
D
Evolution
0
5000
10000
15000
0
100
200
300
Time

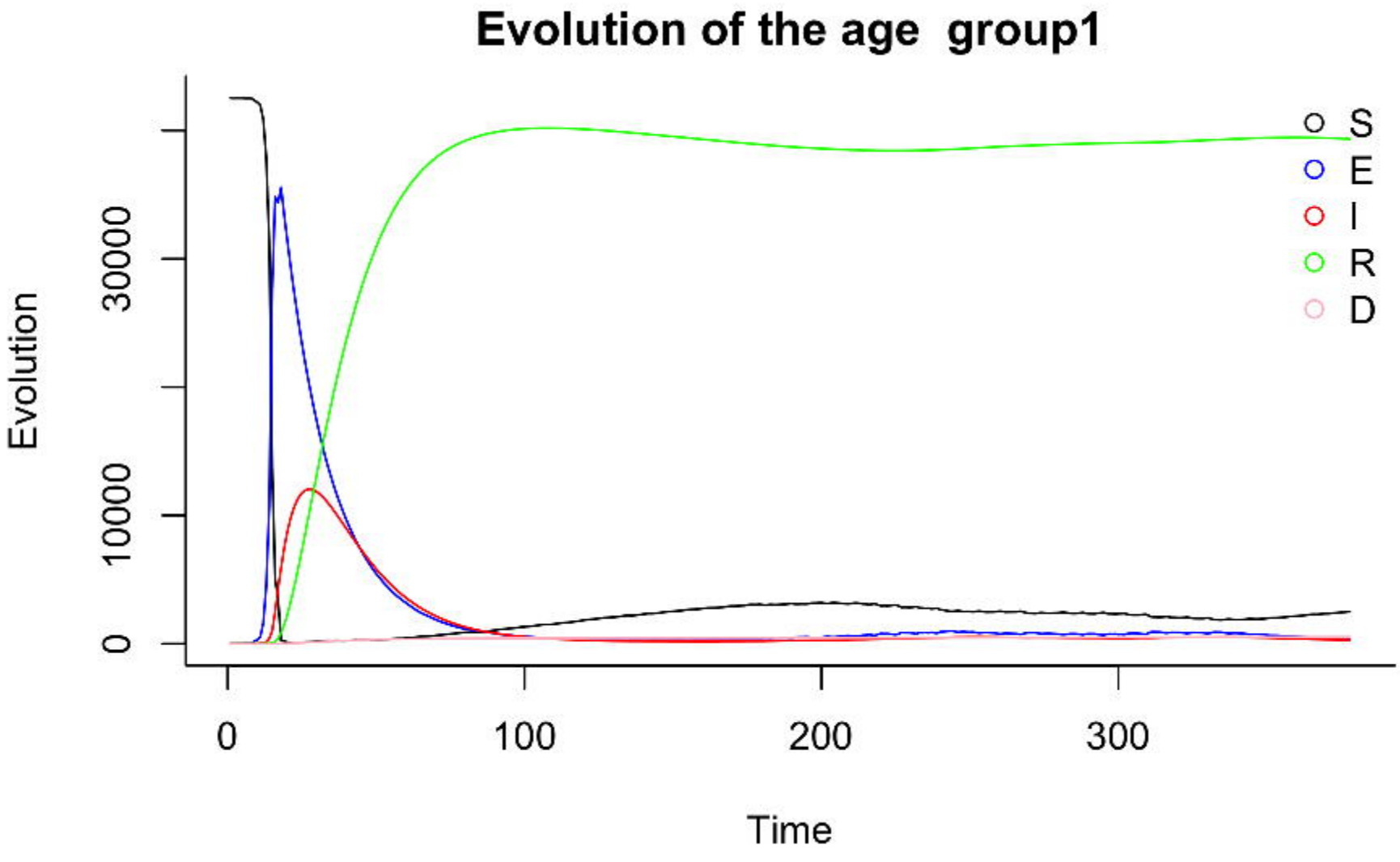

Evolution of the age group6
Evolution
0
5000
10000
15000
0
100
200
300
Time
S
E
I
R
D

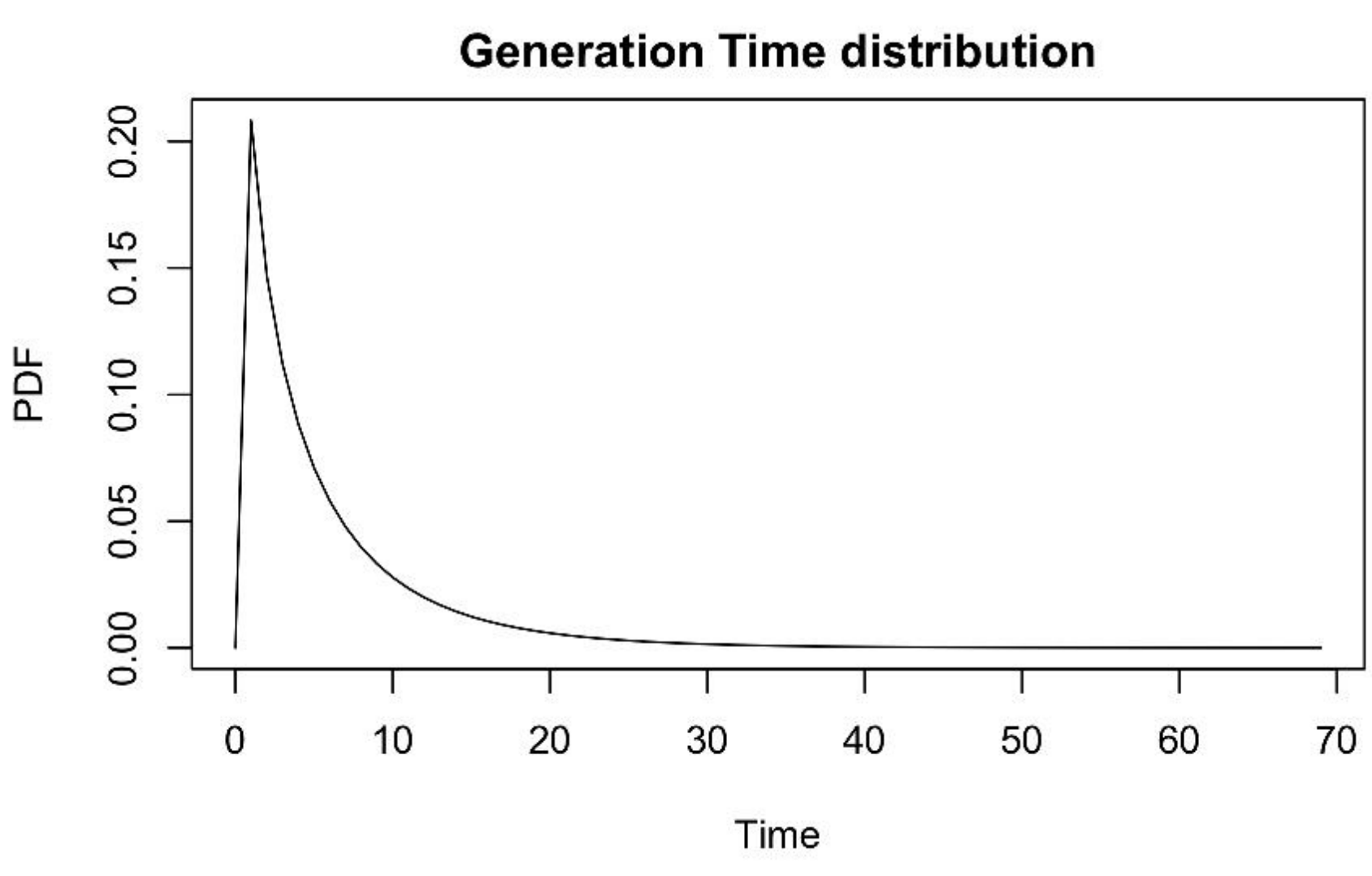

Evolution of the age group5
Evolution
0
5000
10000
15000
0
100
200
300
Time
S
E
I
R
D

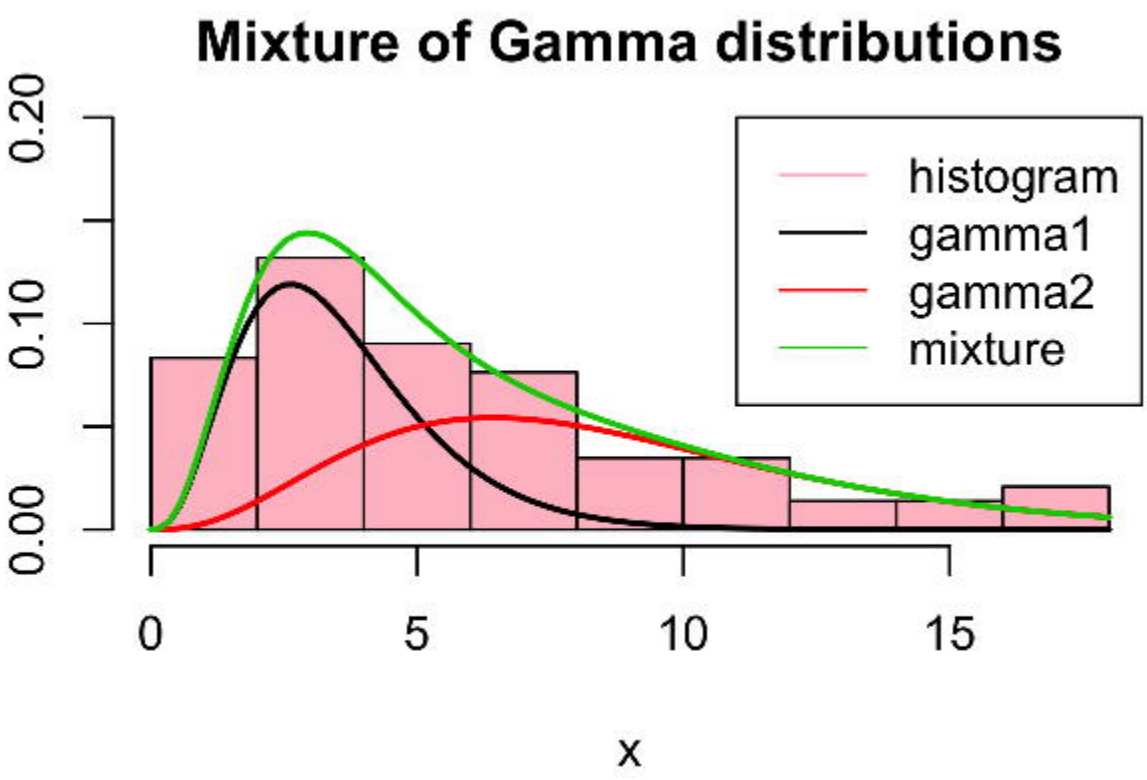

Evolution of the age group4
Evolution
Time
S
E
I
R
D
-5000
5000
15000
25000
0
100
200
300

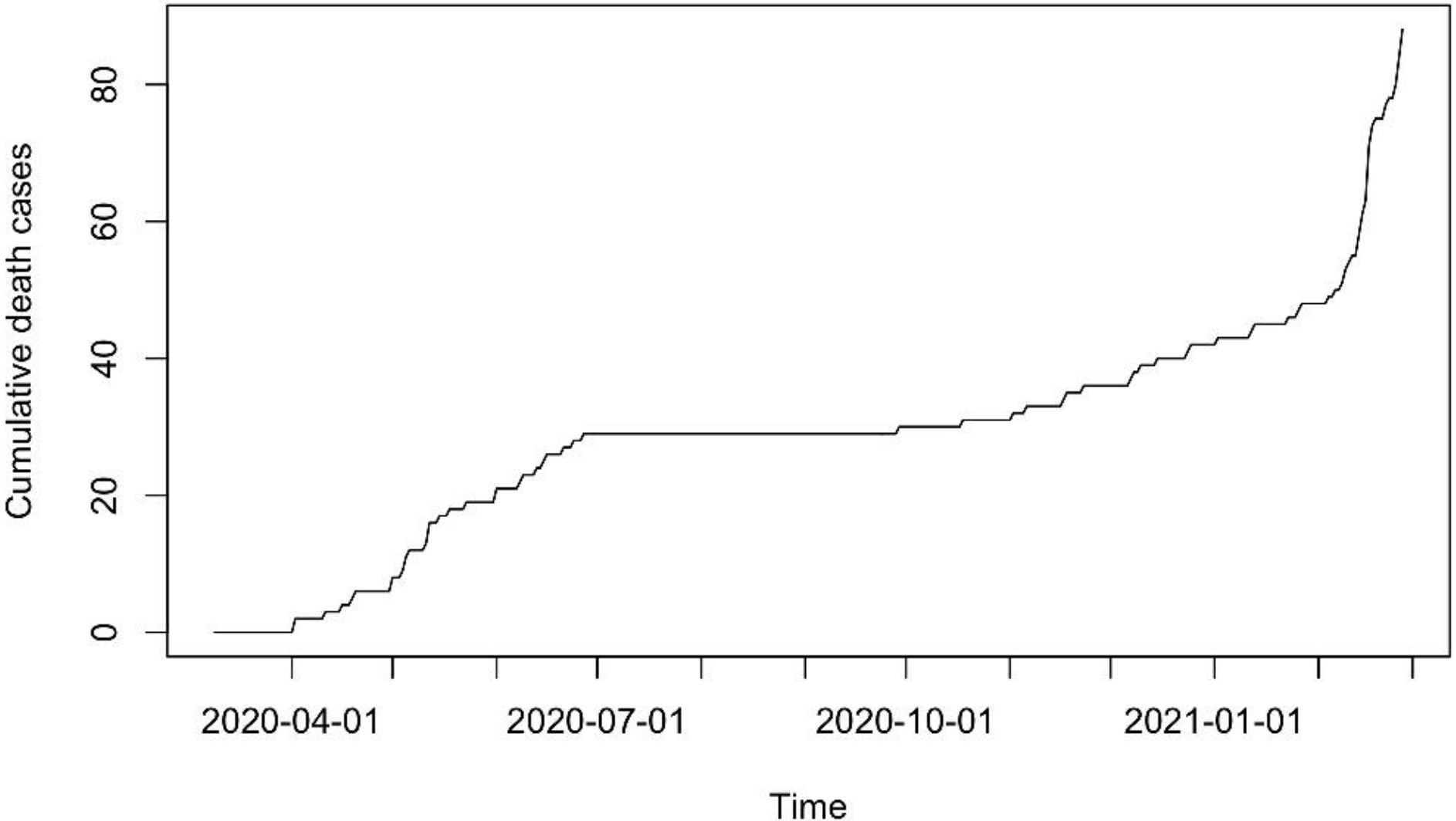

Evolution of the age group3
Evolution
0
10000
20000
30000
0
100
200
300
Time
S
E
I
R
D

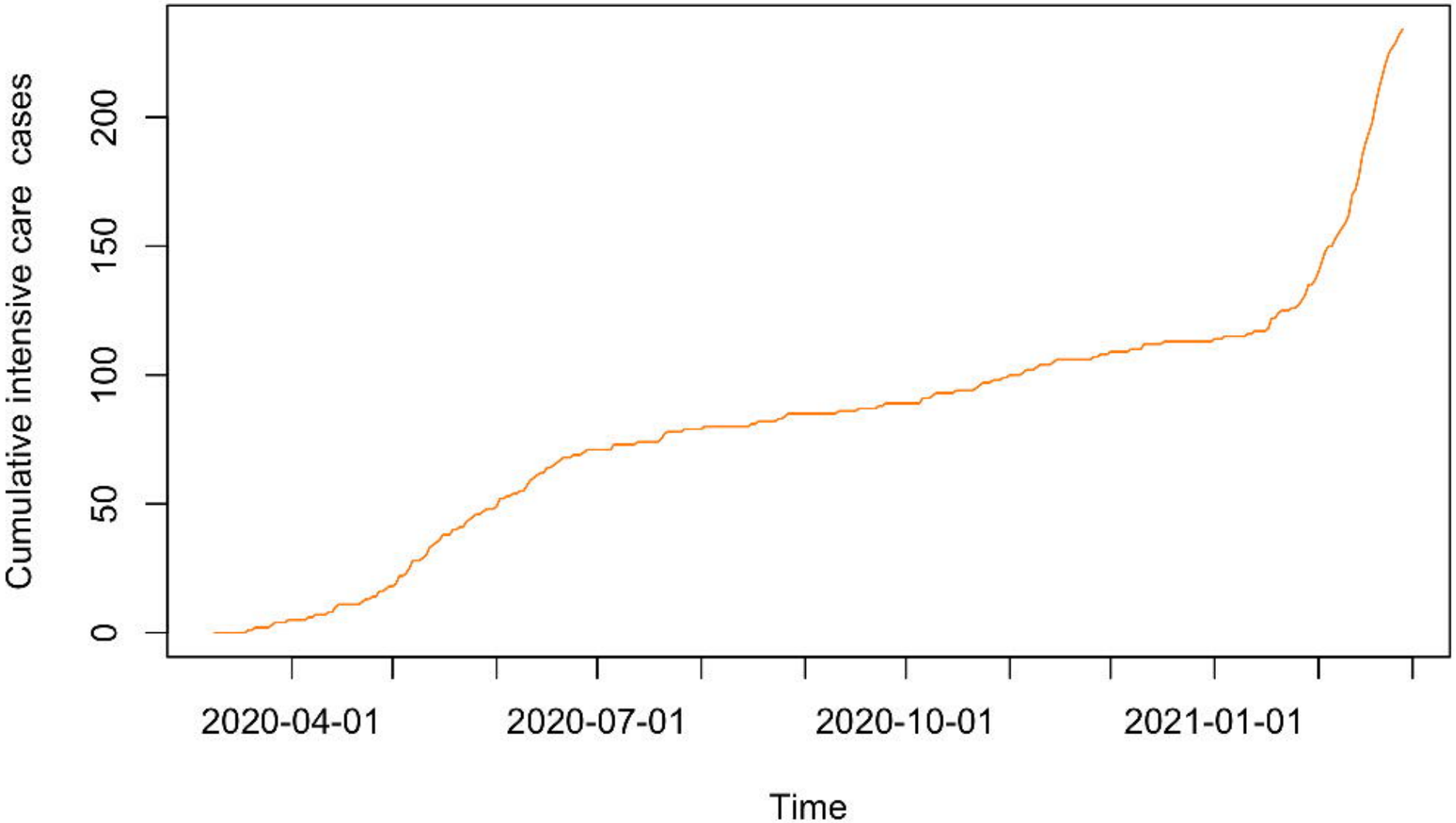

Evolution of the age group2
Evolution
0
10000
30000
0
100
200
300
Time
S
E
I
R
D

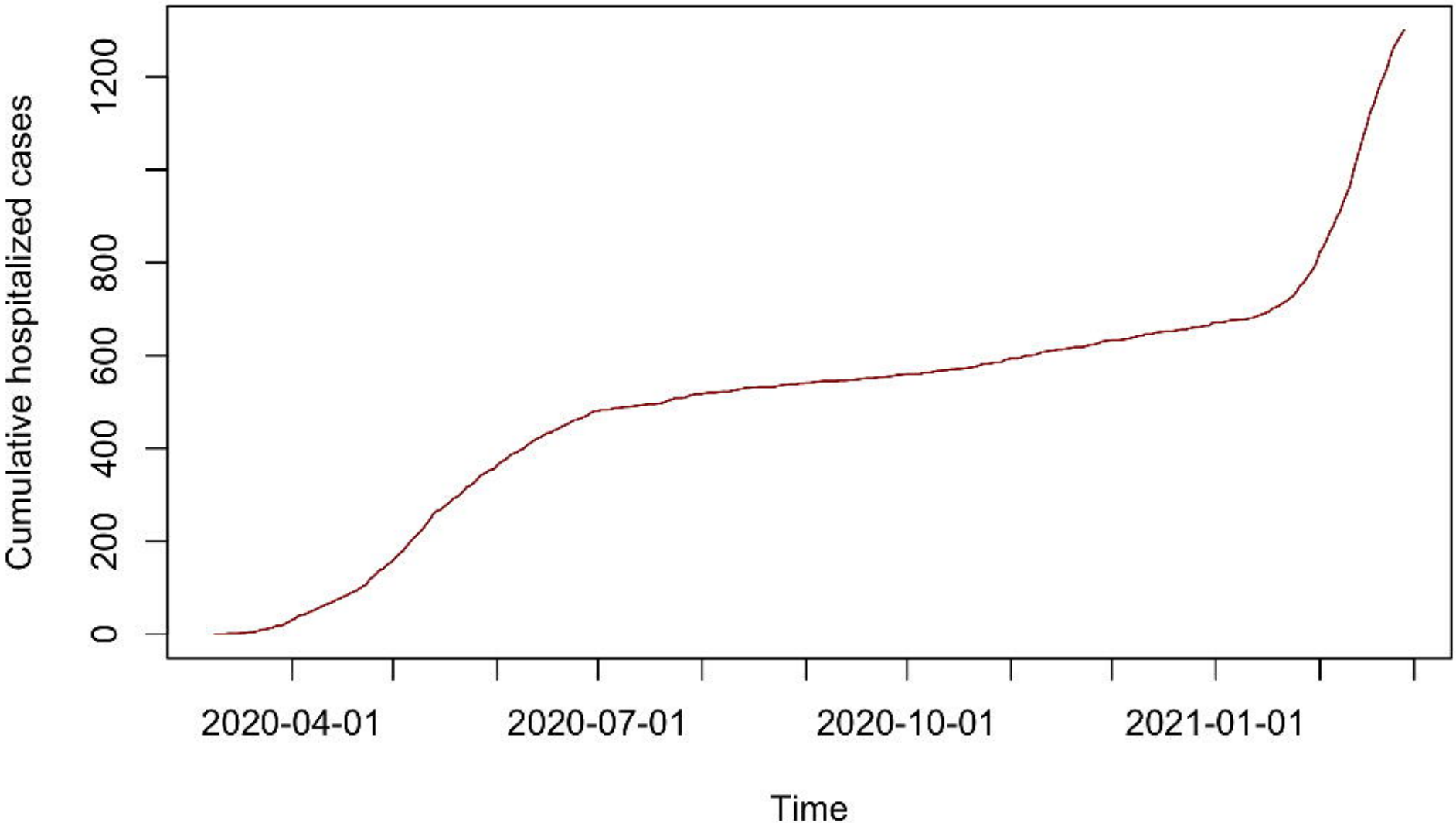

Evolution of the age group1
S
E
I
R
D
Evolution
0
10000
30000
0
100
200
300
Time

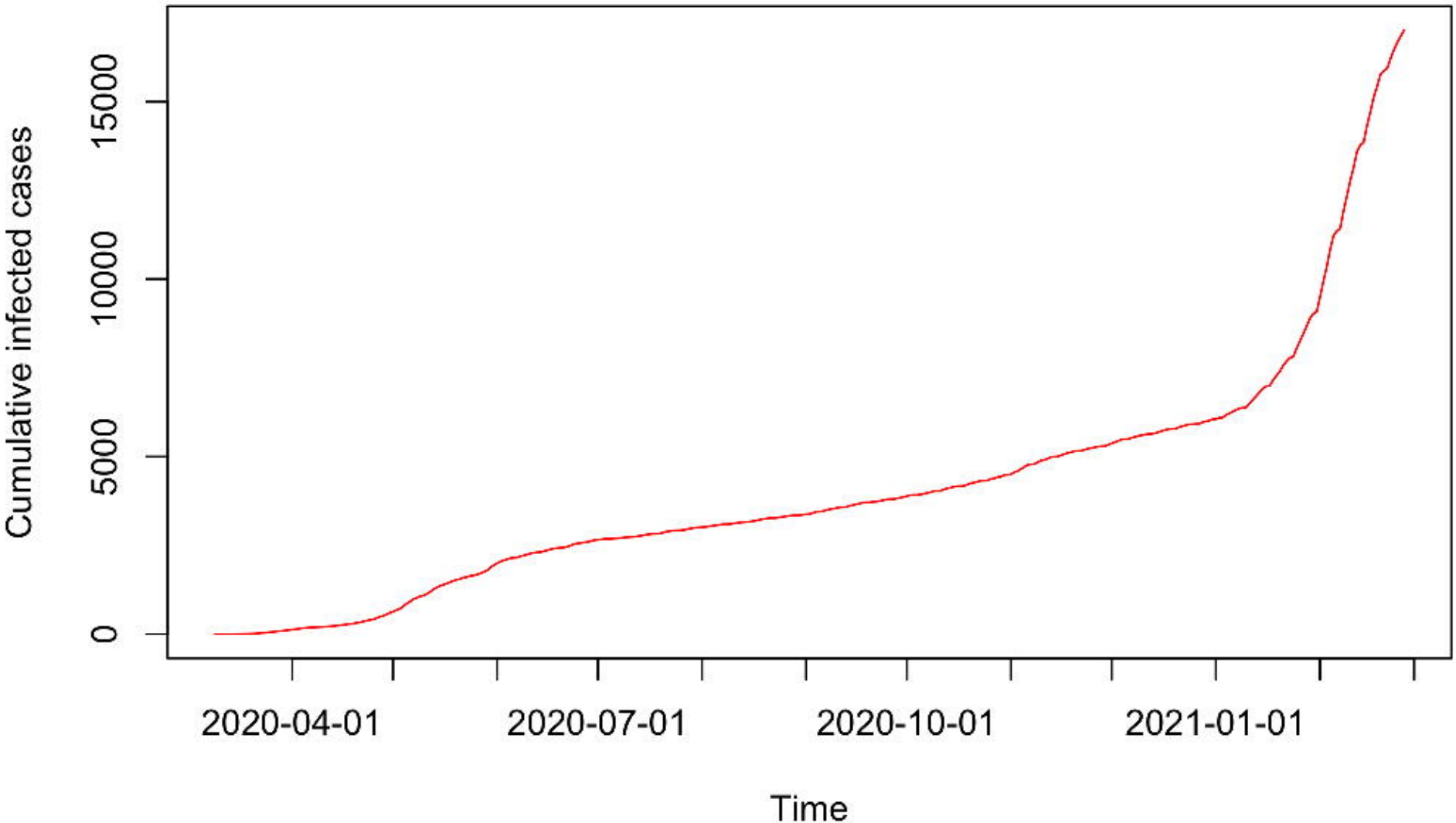

Generation Time distribution
PDF
0.00
0.05
0.10
0.15
0.20
0
10
20
30
40
50
60
70
Time

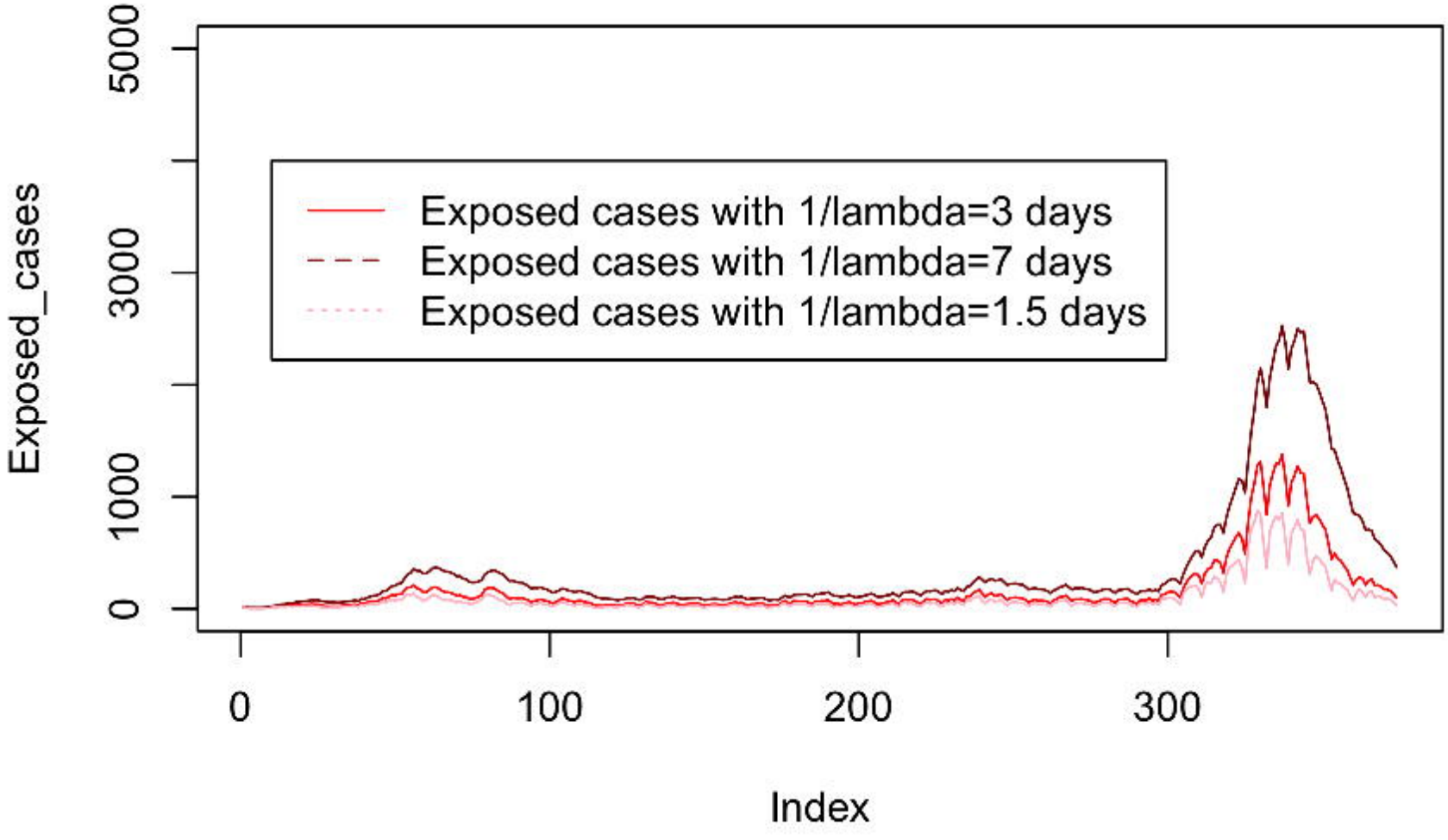

Mixture of Gamma distributions
0.20
0.10
0.00
histogram
gamma1
gamma2
mixture
0
5
10
15
x

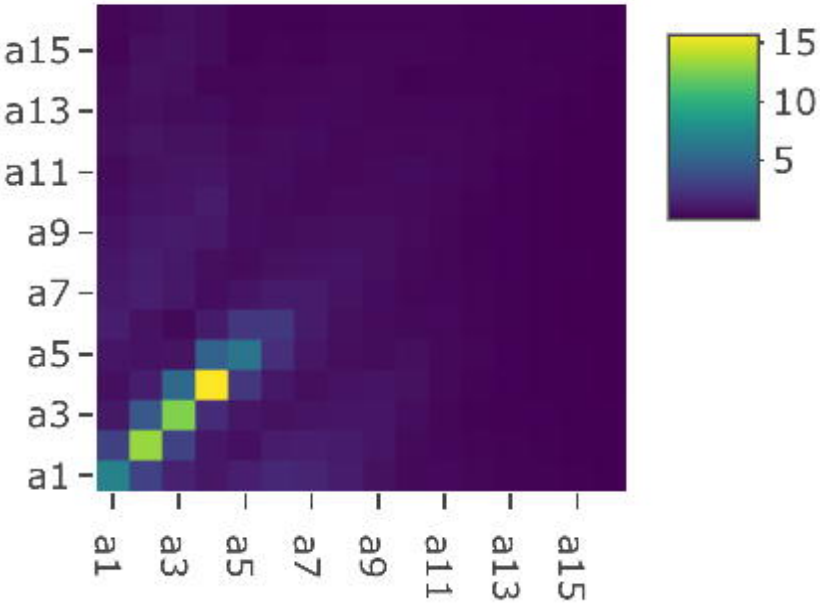

Cumulative death cases
0
20
40
60
80
2020-04-01
2020-07-01
2020-10-01
2021-01-01
Time


Cumulative intensive care cases
0
50
100
150
200
2020-04-01
2020-07-01
2020-10-01
2021-01-01
Time


Cumulative hospitalized cases
0
200
400
600
800
1200
2020-04-01
2020-07-01
2020-10-01
2021-01-01
Time


Cumulative infected cases
0
5000
10000
15000
2020-04-01
2020-07-01
2020-10-01
2021-01-01
Time


Exposed cases with 1/lambda=3 days
Exposed cases with 1/lambda=7 days
Exposed cases with 1/lambda=1.5 days
Exposed_cases
0
1000
3000
5000
0
100
200
300
Index


a15
a13
a11
a9
a7
a5
a3
a1
a1
a3
a5
a7
a9
a11
a13
a15
15
10
5