Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Journal of Theoretical Biology 545 (2022) 111117

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

A multi-strain epidemic model for COVID-19 with infected and asymptomatic cases: Application to French data

Mathilde Massard [a], Raluca Eftimie [a], Antoine Perasso [b], Bruno Saussereau [a,*]

[a] Laboratoire de Mathématiques de Besançon, University of Bourgogne Franche-Comté, UFR ST 16 route de GRAY, 25030 Besançon, France
[b] Laboratoire Chrono-environnement, University of Bourgogne Franche-Comté, UFR ST 16 route de GRAY, 25030 Besançon, France

ARTICLE INFO

Article history:
Received 10 December 2021
Revised 22 March 2022
Accepted 29 March 2022
Available online 2 May 2022

Keywords:
COVID-19 Modeling
Variants of SARS-CoV-2
Basic reproduction number
COVID-19 data for France

ABSTRACT

Many SARS-CoV-2 variants have appeared over the last months, and many more will continue to appear. Understanding the competition between these different variants could help make future predictions on the evolution of epidemics. In this study we use a mathematical model to investigate the impact of three different SARS-CoV-2 variants on the spread of COVID-19 across France, between January-May 2021 (before vaccination was extended to the full population). To this end, we use the data from Geodes (produced by Public Health France) and a particle swarm optimisation algorithm, to estimate the model parameters and further calculate a value for the basic reproduction number R_0. Sensitivity and uncertainty analysis is then used to better understand the impact of estimated parameter values on the number of infections leading to both symptomatic and asymptomatic individuals. The results confirmed that, as expected, the alpha, beta and gamma variants are more transmissible than the original viral strain. In addition, the sensitivity results showed that the beta/gamma variants could have lead to a larger number of infections in France (of both symptomatic and asymptomatic people).

1. Introduction

Since the beginning of 2020, a pandemic of a novel coronavirus disease (COVID-19) has spread throughout the world and caused millions of deaths. COVID-19 is an infectious respiratory disease caused by the SARS-CoV-2 virus. The World Health Organization (WHO) learned of the existence of the SARS-CoV-2 on December 31, 2019 when an outbreak of "viral pneumonia" cases was notified in Wuhan, People's Republic of China (Zhou et al., 2020; Tang et al., 2020). This disease has then spread to the whole world and very quickly led to saturated hospitals.

The fast spread of SARS-CoV-2 also led to the emergence of different variants (i.e., viral genomes that may contain one or more mutations) for this virus, who have been circulating around the world since the beginning of this pandemic. While most of the genetic mutations observed in the circulating SARS-CoV-2 variants do not significantly change virus biology and its properties, some fitness-enhancing mutations (e.g., mutations leading to an increase in transmissibility, or in ability to evade the immune response) have been observed after the first months of virus spread (Harvey et al., 2021). The circulating variants are classified as: Variants Being Monitored (VBM) – those for which data indicate a potential impact on approved treatments but might not pose yet a significant and imminent risk to public health, Variants of Interest (VOI) – for which there is predicted increase in transmissibility of disease severity, and reduced efficacy of treatments, Variants of Concern (VOC) – for which there is evidence of an increase in transmissibility, increase in the severity of disease, and reduced effectiveness of treatments (Centers for Disease Control and Prevention, 2021; European Centre for Disease Prevention and Control, 2021). Variants can be re-classified based in their attributes and prevalence. As of October 2021, the European Centre for Disease Prevention and Control (ECDC) has listed three VOC (Beta, Gamma and Delta) and two VOI (Mu and Lambda) (European Centre for Disease Prevention and Control, 2021).

To investigate the spread of COVID-19, including the role of the different SARS-CoV-2 variants on this spread, and to propose measures to slow-down this spread, researchers have focused their attention on various mathematical compartmental models, either stochastic (He et al., 2020; Faranda and Alberti, 2020) or deterministic (Calafiore and Fracastoro, 2021; He et al., 2020; Tang et al., 2020; Bentout et al., 2021; Yavuz et al., 2021; Arruda et al., 2021; Gonzalez-Parra et al., 2021; Khyar and Allali, 2020; Sridhar et al., 2021; Yagan et al., 2021); see also references therein. The great majority of these models focus on one single variant and

* Corresponding author.
E-mail addresses: mathilde.massard@univ-fcomte.fr (M. Massard), raluca.eftimie@univ-fcomte.fr (R. Eftimie), antoine.perasso@univ-fcomte.fr (A. Perasso), bruno.saussereau@univ-fcomte.fr (B. Saussereau).

https://doi.org/10.1016/j.jtbi.2022.111117
0022-5193/

investigate, for example, the impact of hospitalisation and quarantine (He et al., 2020; Tang et al., 2020), the impact of vaccination (Yavuz et al., 2021), or the impact of different age classes (Bentout et al., 2021; Calafiore and Fracastoro, 2021). Over the last few months, when it became clear the importance of different VOC on the fast increase in the number of COVID-19 cases in different countries, a series of mathematical models have been derived to investigate the role of multiple variants on the spread of SARS-CoV-2 (Arruda et al., 2021; Gonzalez-Parra et al., 2021; Khyar and Allali, 2020; Sridhar et al., 2021; Yagan et al., 2021). For example, Sridhar et al. (2021) and Yagan et al. (2021) used network models to investigate the effectiveness of mask-wearing in limiting the spread of COVID-19 in the context of viral mutations. Arruda et al. (2021) considered a SEIR model that incorporated several viral strains and also reinfection due to vaning immunity. They studied time-varying control strategies in the context of lockdown measures, and so they focused on the cost of infection control (i.e., lockdown) vs. the cost of elevated infection levels to the healthcare system, over a two-year time interval. Khyar and Allali (2020) focused on the global stability analysis of a two-strain SEIR epidemic model with two general incidence rates. They also calculated the basic reproduction number for their epidemic model. Finally, Gonzalez-Parra et al. (2021) extended a two-strain SEIR model to include also asymptomatic, hospitalized and dead individuals, to investigate the transmission of COVID-19 in Columbia. They consider variations in the contagiousness of the two strains, to see their impacts on the number of infections, hospitalisations and deaths.

In this study we also focus on the role of different SARS-CoV-2 variants on the transmission of this virus across France in the absence of vaccination. However, unlike the previous studies, here we consider a simpler compartmental model (i.e., a generalisation of a SIR model) that includes also asymptomatic cases and dead individuals infected with the original variant, as well as cases infected with other variants. To calculate the basic reproductive number R_0, we first focus on a general model with $N > 0$ variants, and derive a formula for R_0. Then, we apply this formula for the case with two variants ($N = 2$), corresponding to the alpha and beta/gamma variants (see discussion on French data in Section 2). This model is parametrised (with the help of a particle swarm optimisation algorithm) using French data selected over the longest possible time interval that can take into account the emergence of different SARS-CoV-2 variants without including also the vaccination. Sensitivity analysis is performed to understand the impact of uncertainty in model parameters on the overall outcome (i.e., symptomatic and asymptomatic infections). This sensitivity analysis allows us to investigate the severity of infections with different variants in France, in the absence of any vaccination.

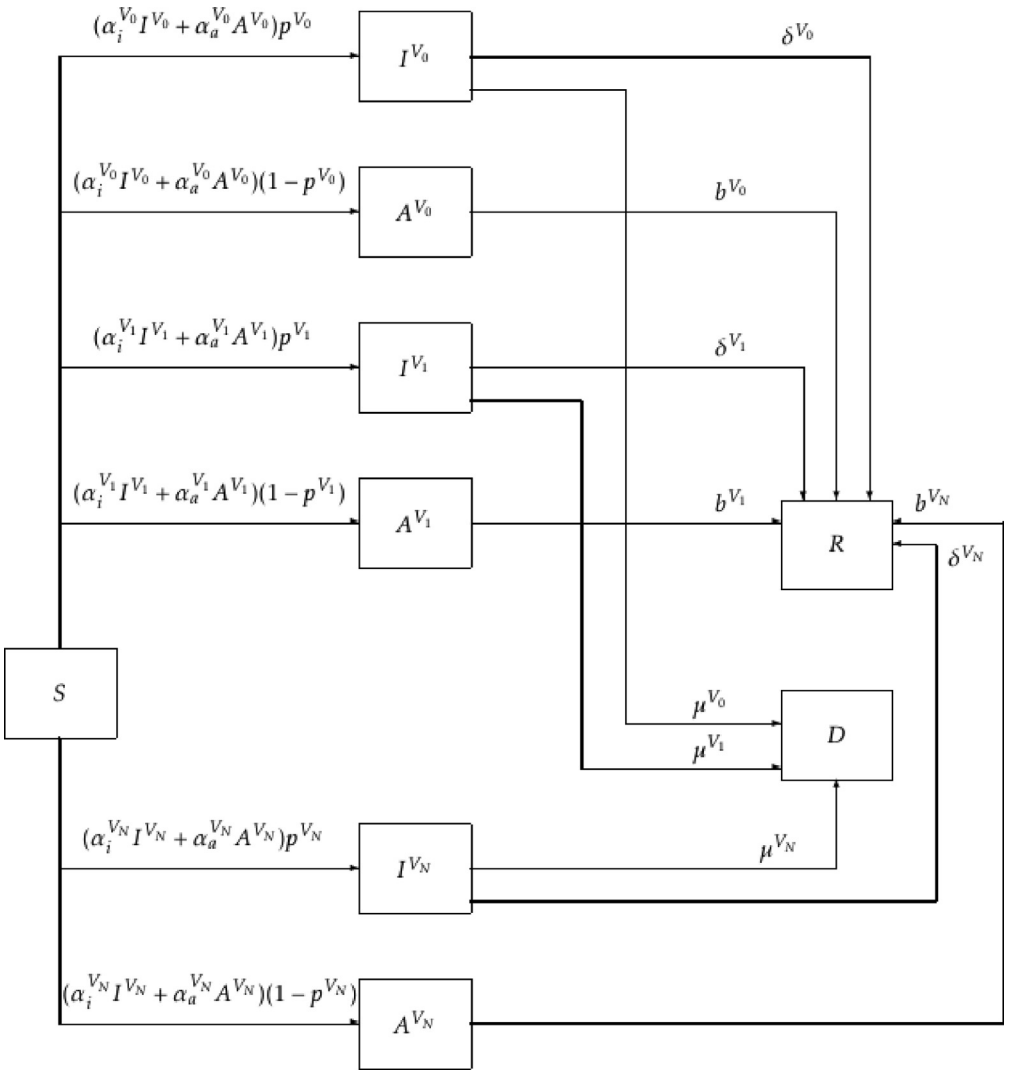

Fig. 1. Diagram describing the interactions incorporated into model (1).

The paper is structured as follows. In Section 2 we introduce a general compartment model that considers a generic number of N variants, that can lead to symptomatic as well as asymptomatic cases. We use this model to calculate a general formula for R_0 in terms of the parameters associated with different variants. In Section 3 we discuss SARS-CoV-2 data for France, and based on this data we focus on a simpler model with $N = 2$ variants: alpha and beta. Using available data, we parametrise this simpler model with the help of a particle swarm optimisation algorithm. In Section 4 we estimate the basic reproductive number for the three SARS-CoV-2 variants considered here, and we perform a sensitivity analysis to investigate the impact of uncertainties in the previously-identified model parameters on the model outcome (i.e., the number of infectious individuals, as well as R_0).

2. A general model with N variants

We start by presenting a general compartmental model that takes into account a generic number N of variants, with $N > 0$. We choose N variants because it allows us to include all current and possible future VOC variants. Symptomatic and asymptomatic cases are taken into consideration with the assumption that asymptomatic people do not die from the disease. Natural births and deaths can be overlooked, due to the large size of the French population. Loss of immunity can also be ignored because the number of individuals infected twice is still low (Brouqui et al., 2021), and this loss of immunity is not present for all variants. For example, variants with the E484K mutation in the spike protein may be responsible for immune escape (Nonaka et al., 2021; Wise, 2021).

2.1. Model description

For the model considered in this study, we denote the initial strain by V_0, the first variant by V_1, ..., the N-th variant by V_N. The population is divided into $2N + 5$ different classes: susceptible individuals (S) who are healthy and can contract the disease; individuals infected with the original virus and having symptoms (I^{V_0}); individuals infected with the original virus and having no symptoms (A^{V_0}); individuals infected with the variants V_1, ..., V_N and having symptoms (I^{V_1}, ..., I^{V_N}); individuals infected with the variants V_1, ..., V_N and having no symptoms (A^{V_1}, ..., A^{V_N}); dead individuals (D); recovered individuals (R) who are immune to the disease.

The dynamics of the population is described by the system below (see also Fig. 1 for a schematic description of the interactions between different model variables).

$$\begin{cases} \frac{dI^{V_0}}{dt} = Sp^{V_0}(\alpha_i^{V_0} I^{V_0} + \alpha_a^{V_0} A^{V_0}) - (\delta^{V_0} + \mu^{V_0}) I^{V_0} \\ \frac{dA^{V_0}}{dt} = S(1-p^{V_0})(\alpha_i^{V_0} I^{V_0} + \alpha_a^{V_0} A^{V_0}) - b^{V_0} A^{V_0} \\ \frac{dI^{V_1}}{dt} = Sp^{V_1}(\alpha_i^{V_1} I^{V_1} + \alpha_a^{V_1} A^{V_1}) - (\delta^{V_1} + \mu^{V_1}) I^{V_1} \\ \frac{dA^{V_1}}{dt} = S(1-p^{V_1})(\alpha_i^{V_1} I^{V_1} + \alpha_a^{V_1} A^{V_1}) - b^{V_1} A^{V_1} \\ \qquad \vdots \\ \frac{dI^{V_N}}{dt} = Sp^{V_N}(\alpha_i^{V_N} I^{V_N} + \alpha_a^{V_N} A^{V_N}) - (\delta^{V_N} + \mu^{V_N}) I^{V_N} \\ \frac{dA^{V_N}}{dt} = S(1-p^{V_N})(\alpha_i^{V_N} I^{V_N} + \alpha_a^{V_N} A^{V_N}) - b^{V_N} A^{V_N} \\ \frac{dR}{dt} = \sum_{j=0}^{N} (\delta^{V_j} I^{V_j} + b^{V_j} A^{V_j}) \\ \frac{dD}{dt} = \sum_{j=0}^{N} (\mu^{V_j} I^{V_j}) \\ \frac{dS}{dt} = -S \sum_{j=0}^{N} \left(\alpha_i^{V_j} I^{V_j} + \alpha_a^{V_j} A^{V_j} \right) \end{cases} \tag{1}$$

In the above system (1), the equation for $\frac{dI^{V_0}}{dt}$ describes the evolution of individuals infected with the initial strain and having symptoms (I^{V_0}). Susceptible persons S come into contact with persons infected by the initial strain with symptoms (I^{V_0}) or without

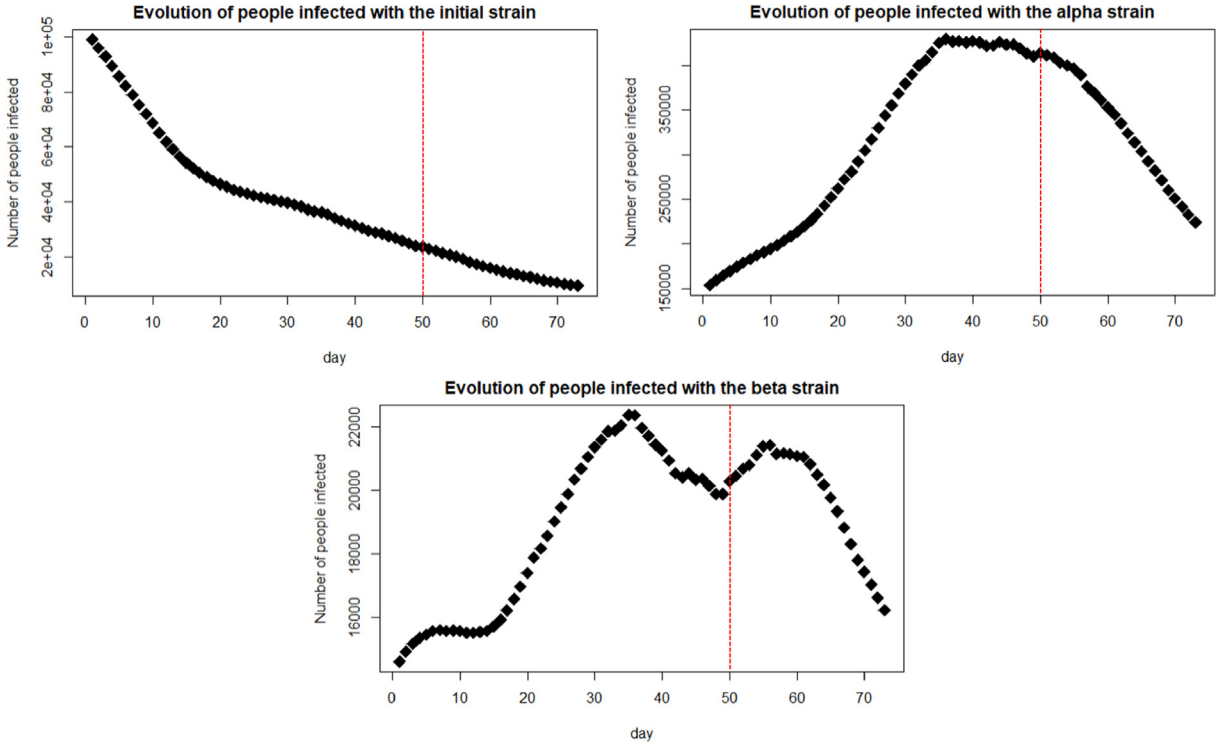

Fig. 2. The evolution of I^{V_0}, I^{V_1} and I^{V_2}, as given by data on Geodes (see also Tables 4–6). This is the rescaled data that includes the assumption of 13 days in the infected compartments.

Table 1
Summary of parameters involved in model (1). Parameters $\alpha_i^{V_k}$ and $\alpha_a^{V_k}$ are respectively equal to $\frac{1}{N*j_i^k}$ and $\frac{1}{N*j_a^k}$, with N the size of the population considered (hence the units for these parameters are $1/(day \times size)$). Here, j_i^k and j_a^k are the average duration times between two contaminating contacts (with the strain k) by a symptomatic person and by an asymptomatic person.

Symbol	Biological interpretation	Unit
$\mu^{V_0}, \mu^{V_1}, .., \mu^{V_N}$	Mortality rate due to the initial virus and to other N variant	day^{-1}
$\frac{1}{\delta^{V_0}}, \frac{1}{\delta^{V_1}}, \dots, \frac{1}{\delta^{V_N}}$	Average time of infection for symptomatic individuals	day
$\frac{1}{b^{V_0}}, \frac{1}{b^{V_1}}, \dots, \frac{1}{b^{V_N}}$	Average time of infection for asymptomatic individuals	day
$\alpha_i^{V_0}, \alpha_i^{V_1}, \dots, \alpha_i^{V_N}$	Transmission rate of variants and initial virus from symptomatic people	$size^{-1} \times day^{-1}$
$\alpha_a^{V_0}, \alpha_a^{V_1}, \dots, \alpha_a^{V_N}$	Transmission rates of variants and initial virus from asymptomatic people	$size^{-1} \times day^{-1}$
$p^{V_0}, p^{V_1}, \dots, p^{V_N}$	Probability that an individual infected with the initial virus, variant 1, .. variant N, is symptomatic	–

symptoms (A^{V_0}). There is a certain probability of being infected by the initial virus which results in the terms $S\alpha_i^{V_0} I^{V_0}$ and $S\alpha_a^{V_0} A^{V_0}$. A proportion p^{V_0} of these individuals will have symptoms. Finally, infected individuals can die at a rate μ^{V_0} or can recover at a rate δ^{V_0}. Similar terms can be found in the equations describing the evolution of the other symptomatic infected compartments $I^{V_1}, \dots, I^{V_N}$; see also Table 1.

The equation for $\frac{dA^{V_0}}{dt}$ in (1) describes the evolution of the individuals infected by the initial strain and having no symptoms (A^{V_0}). A proportion $(1 - p^{V_0})$ of susceptibles can become infected and have no symptoms, and as before there is a probability of infection following contact with other symptomatic individuals $S\alpha_i^{V_0} I^{V_0}$ or asymptomatic individuals $S\alpha_a^{V_0} A^{V_0}$. In addition, individuals can leave this asymptomatic class following recovery at a rate $+b^{V_0}$. We remind the reader that the asymptomatic persons cannot die. Similar terms can be found in the equations describing the evolution of the other compartments $A^{V_1}, \dots, A^{V_N}$; see also Table 1.

The equation for $\frac{dR}{dt}$ in (1) describes the evolution of recovered individuals. The terms that appear in this equation describe the recovery of infected and asymptomatic individuals, as discussed above.

The equation for $\frac{dD}{dt}$ in (1) describes the time-evolution of COVID-19-dead individuals. The terms in this equation are the death terms that we discussed above for the symptomatic infected compartments.

The equation for $\frac{dS}{dt}$ in (1) describes the time-evolution of the susceptible people. These individuals leave the compartment S when becoming infected (after contact with symptomatic or asymptomatic people).

All parameters involved in this general model (1) with N viral variants are summarized in the Table 1. In the following we discuss briefly the calculation of the basic reproduction number R_0 for this general model.

2.2. R_0 calculation

A fundamental tool in epidemiology is the basic reproduction number R_0, which represents the number of secondary infections resulting from a single infectious individual introduced into a fully susceptible population (Diekmann et al., 1990; Van den Driessche and Watmough, 2002; Perasso, 2018). Note that in general a population is not fully susceptible, with some individuals already having immunity/cross-immunity to the virus. Therefore, even if throughout this study we refer to the basic reproductive number R_0, we actually understand the effective reproduction number (i.e., the number of secondary infections resulting from a single infectious individual introduced into a population formed of susceptible and non-susceptible individuals); this will be more clear in Section 3 in the context of alpha and beta variants for COVID-19.

One method that can be used to calculate the basic reproduction number for the finite-dimensional system (1) makes use of the next generation matrix (Van den Driessche and Watmough, 2002). This method is based on the definition of R_0 as the dominant eigenvalue of the "next generation matrix" (Diekmann et al., 1990), i.e., a matrix that relates the numbers of newly infected individuals in various categories in consecutive generations (Diekmann et al., 2010). Applying this method to model (1) we obtain the next generation matrix M below:

$$M = \begin{pmatrix} \frac{S_0 p^{V_0} \alpha_i^{V_0}}{\delta^{V_0}+\mu^{V_0}} & \frac{S_0 p^{V_0} \alpha_a^{V_0}}{b^{V_0}} & 0 & 0 & \cdots & 0 & 0 \\ \frac{S_0 (1-p^{V_0}) \alpha_i^{V_0}}{\delta^{V_0}+\mu^{V_0}} & \frac{S_0 (1-p^{V_0}) \alpha_a^{V_0}}{b^{V_0}} & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \frac{S_0 p^{V_1} \alpha_i^{V_1}}{\delta^{V_1}+\mu^{V_1}} & \frac{S_0 p^{V_1} \alpha_a^{V_1}}{b^{V_1}} & \cdots & 0 & 0 \\ 0 & 0 & \frac{S_0 (1-p^{V_1}) \alpha_i^{V_1}}{\delta^{V_1}+\mu^{V_1}} & \frac{S_0 (1-p^{V_1}) \alpha_a^{V_1}}{b^{V_1}} & \cdots & 0 & 0 \\ \vdots & \vdots & & & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \frac{S_0 p^{V_N} \alpha_i^{V_N}}{\delta^{V_N}+\mu^{V_N}} & \frac{S_0 p^{V_N} \alpha_a^{V_N}}{b^{V_N}} \\ 0 & 0 & 0 & 0 & \cdots & \frac{S_0 (1-p^{V_N}) \alpha_i^{V_N}}{\delta^{V_N}+\mu^{V_N}} & \frac{S_0 (1-p^{V_N}) \alpha_a^{V_N}}{b^{V_N}} \end{pmatrix}.$$

Thus, we obtain the following formula for R_0:

$$R_0 = max\left\{\left|\frac{S_0 p^{V_j}\alpha_i^{V_j}}{\delta^{V_j}+\mu^{V_j}} + \frac{S_0(1-p^{V_j})\alpha_a^{V_j}}{b^{V_j}}\right| \quad ;j=0,1,..,N\right\}. \tag{2}$$

Each eigenvalue consists only of the parameters corresponding to a single strain. It is like there are several epidemics at the same time, and they don't interfere with each other. By obtaining a value of R_0 from the available data (see Section 3), we will be able to conclude which variant takes over the others.

3. Parameters estimation for a 2 variant-model ($N = 2$): alpha and beta/gamma variants

Throughout the rest of this paper we focus on a simpler model with only two variants, (i.e., $N = 2$), and we use the available data to obtain an estimate for R_0 (or rather the effective reproduction number, as discussed above). We estimate the 18 parameters involved in model (1) and in Eq. (2) using a Particle Swarm Optimization algorithm.

3.1. Data

The data we use in this study is from Geodes (i.e., the cartographic observatory of epidemiological indicators produced by Public Health France[1]), and covers the period between 12th February and 7th May 2021, because before the 12 February the different strains of the virus were not recorded (only virus presence was recorded). Moreover, we consider data only until 7th May to avoid the effects of vaccines (since in France the vaccination really started off in May). We focus on the percentages of positive RT-PCR tests identifying the initial strain, the alpha variant (which appeared in France since mid-December 2020) and the beta & gamma variants (present in France since end of December 2020 & beginning of February 2021, respectively), calculated over a 7-day period. This period allows to remove the effect of reduced testing on Sundays. The data can be seen in Tables 4–6, in Appendix A. Note that on Geodes, the data for the beta and gamma variants are presented together, and for this reason in this study we combine them into a single variant (called variant 2 in the next section).

On Geodes, the data included also percentages of a so-called "indeterminate variant". To remove this indeterminate variant (so we can have only the initial strain, the alpha variant and the beta/gamma variant), we considered the following approach. Denote the alpha strain by α, the beta strain by β and the initial strain by V_0. Then we calculate:

$$\text{new percentage of } \alpha = \frac{percentage\ of\ \alpha}{Total},$$

with

$$\text{Total} = \text{percentage of } \alpha + \text{percentage of } \beta + \text{percentage of } V_0.$$

Similar calculations were performed to obtain the new percentages for the beta/gamma variants, as well as the initial strain.

The number of new confirmed cases per day in France is available on the official french public data site [2]. These data are summarised in Fig. 7, in Appendix. To remove the effects of Sunday testing, when fewer RT-PCR tests are carried out, we take an average over 7 days at each time point, to correspond to the values of the percentages of the variants. These data are shown in Fig. 8 in Appendix.

From the percentages of each variant as well as the number of people infected at each time point t (i.e., day), we obtain three datasets corresponding to the number of new cases per day for each of the three strains. These three datasets are of size 85 corresponding to the period from February 12 to May 7. However, the different symptomatic infected compartments (I) in system (1) (with $N = 2$) count the current numbers of infected individuals and not the newly infected per day. To transform the data so that we can use it to parametrise our model, we assume that individuals stay an average of 13 days in a symptomatic compartment before recovering (George et al., 2021). By including this recovery delay, we reduced the size of the three datasets from 85 to 73 days:

$$I_{\text{new}}[i] = \Sigma_{k=0}^{12} I_{\text{originaldata}}[i+k], \qquad \forall i \in [1, 73] \tag{3}$$

with I_{new} denoting the rescaled cumulative data that consider the assumption of 13 days infections (i.e., we sum up all daily infections $I_{\text{originaldata}}$ that took place over the past 13 days), which is then used to parametrise the mathematical model.

The evolution of the three infected compartments can be seen in Fig. 2. Note that the curve for the infected population with the alpha variant (panel (b)) starts decreasing around day $t = 50$ (marked by the vertical dashed line). This corresponds to 2 April 2021, when the French government imposed a 4-weeks national lockdown (which lasted until 3 May). At the same time, the beta/gamma variant starts to increase slightly until $t \approx 60$, after which it decreases as well.

Data for initial population sizes. For the numerical simulations performed throughout the rest of the paper, we consider the following initial population sizes:

- Using the latest census data from mainland France (Décembre, 2020), in this study we assume that the total population of France is ≈ 64000000. We consider this initial condition because on 12 February 2021 (when we start our investigation into the dynamics of these different variants) there were 3,267,386 people in France who recovered from the disease (i.e., $R(0) = 3267386$) and 81,448 people who died (i.e., $D(0) = 81448$) (Santé publique france, 2021).
- For the infected sub-populations, we use the datasets in Tables 4–6: $I_{\text{initial}}[1] = 99128.5$ (the first value of the first dataset), $I_{\text{alpha}}[1] = 154239.5$ (the first value of the second dataset), $I_{\text{gamma/beta}}[1] = 14597.74$ (the first value of the third dataset). Regarding the infected people, we assume that 85% are symptomatic and 15% are asymptomatic (Xiao et al., 2020), and that the asymptomatic are not tested the same way as the symptomatic infected individuals and thus we take $I^{V_0}(0) = I_{\text{initial}}[1]$, $A^{V_0}(0) = I_{\text{initial}}[1] * \frac{0.15}{0.85}, I^{V_1}(0) = I_{\text{alpha}}[1], A^{V_1}(0) = I_{\text{alpha}}[1] * \frac{0.15}{0.85}$, $I^{V_2}(0) = I_{\text{gamma/beta}}[1], A^{V_2}(0) = I_{\text{gamma/beta}}[1] * \frac{0.15}{0.85}$.
- For the Susceptible population, we need to subtract from the total population the number of individuals infected with the different variants:

$$\begin{aligned} S(0) &= 64000000 - I_{\text{initial}}[1] - I_{\text{alpha}}[1] - I_{\text{gamma/beta}}[1] \\ &\quad - 3267386 - 81448 \\ &= 60383200.3 \end{aligned}$$

3.2. Particle Swarm Optimization

The Particle Swarm Optimization (PSO) algorithm was created in 1995 by Russel Eberhart, electrical engineer, and James Kennedy, social psychologist (Kennedy and Eberhart, 1995), and it is inspired by the collective behaviour of flocks of birds or schools of fish. For the PSO, the particle (i.e., potential solution of the model) moves by making allowances between getting closer to

[1] https://geodes.santepubliquefrance.fr.

[2] The data are available at the following address: https://www.data.gouv.fr/fr/datasets/synthese-des-indicateurs-de-suivi-de-lepidemie-covid-19/.

the optimal solution visited and getting closer to the solutions found in its neighborhood. This optimization algorithm performs well on parameter optimization for ordinary differential equation models (Akman et al., 2018), but convergence to the overall optimal solution is not always guaranteed. In the following we describe how we apply this PSO algorithm to identify the parameters of our model (1) with two SARS-CoV-2 variants, using the data described above.

3.2.1. Objective function and search bounds for PSO

For each viral strain we calculate the residual sum of squares

$$RSS^k(\theta) = \sum_{i=1}^{75} (y^k(t_i) - \hat{y}^k(t_i))^2, \quad \text{with} \quad k \in \{V_0, V_1, V_2\} \tag{4}$$

where

- $\theta = (\alpha_i^{v_0}, \alpha_a^{v_0}, p^{v_0}, b^{v_0}, \delta^{v_0}, \mu^{v_0}, \alpha_i^{v_1}, \alpha_a^{v_1}, p^{v_1}, b^{v_1}, \delta^{v_1}, \mu^{v_1}, \alpha_i^{v_2}, \alpha_a^{v_2}, p^{v_2}, b^{v_2}, \delta^{v_2}, \mu^{v_2})$.
- $y^k(t_i)$ describe the empirical observations at time t_i corresponding to the number of individuals infected by strain $k \in \{V_0, V_1, V_2\}$.
- $\hat{y}^k(t_i)$ describe the numerical predictions at time t_i for the individuals infected by strain $k \in \{V_0, V_1, V_2\}$ (as given by system (1) with $N = 2$).

The objective function to be minimized is therefore

$$RSS(\theta) = RSS^{V_0}(\theta) + RSS^{V_1}(\theta) + RSS^{V_2}(\theta). \tag{5}$$

Based on reasonable parameter bounds (see discussion below), we constrain our search space with lower and upper bounds for θ:

$$\begin{aligned}\text{lowerbound} = (&1/(64000000 * 25), 1/(64000000 * 25), 0.55, 1/13, 1/13, 1/13,\\ &1/(64000000 * 25), 1/(64000000) * 25), 0.55, 1/13, 1/13, 1/13,\\ &1/(64000000 * 25), 1/(64000000 * 25), 0.55, 1/13, 1/13, 1/13).\end{aligned}$$

$$\begin{aligned}\text{upperbound} = (&1/(64000000 * 1), 1/(64000000 * 1), 0.9, 1/3, 1/3, 1/5,\\ &1/(64000000 * 1), 1/(64000000 * 1), 0.9, 1/3, 1/3, 1/5,\\ &1/(64000000 * 1), 1/(64000000 * 1), 0.9, 1/3, 1/3, 1/5).\end{aligned}$$

This means that infected people take at most 25 days and at least 1 day to infect a person (Amiri Mehra et al., 2020). Indeed, the transmission rate has been estimated at $5.0981280 \times 10^{-12}$ for the whole world (Marinov et al., 2020) which corresponds to $1/(7,870,000,000 \times 5.0981280 \times 10^{-12}) = 24.9$ days to transmit the virus (because N = 7,870,000,000 is the world population). The symptomatic proportion is between 0.55 and 0.9 (Xiao et al., 2020), a symptomatically-infected person takes between 3 days (Pottier, 2020) and 13 days (George et al., 2021) to recover, and between 5 days (Fernández-Villaverde and Jones, 2020) and 13 days to die (also to match our infection hypothesis (George et al., 2021)).

3.3. Parameter estimation

In Fig. 3 we show the best fit between the data collected from Geodes (Geodes, 2021) on infected individuals (black dots), and the numerically-predicted number (red curve) of individuals infected with (a) initial strain, (b) alpha strain, (c) beta/gamma strain. The parameters for which the numerical solutions were obtained are summarised in Table 2, while Table 3 gives the biological interpretation of these parameters. The parameters $\alpha_i^{V_k}$ and $\alpha_a^{V_k}$ are respectively equal to $\frac{1}{N*j_i^k}$ and $\frac{1}{N*j_a^k}$, with N the size of the population considered, j_i^k and j_a^k the average duration between two contaminating contacts by a symptomatic person and by an asymptomatic person by the strain k. For the biological interpretation of the other parameters see Table 1.

The fit of the individuals infected with the initial and alpha strain is relatively good. In contrast, the fit of individuals infected

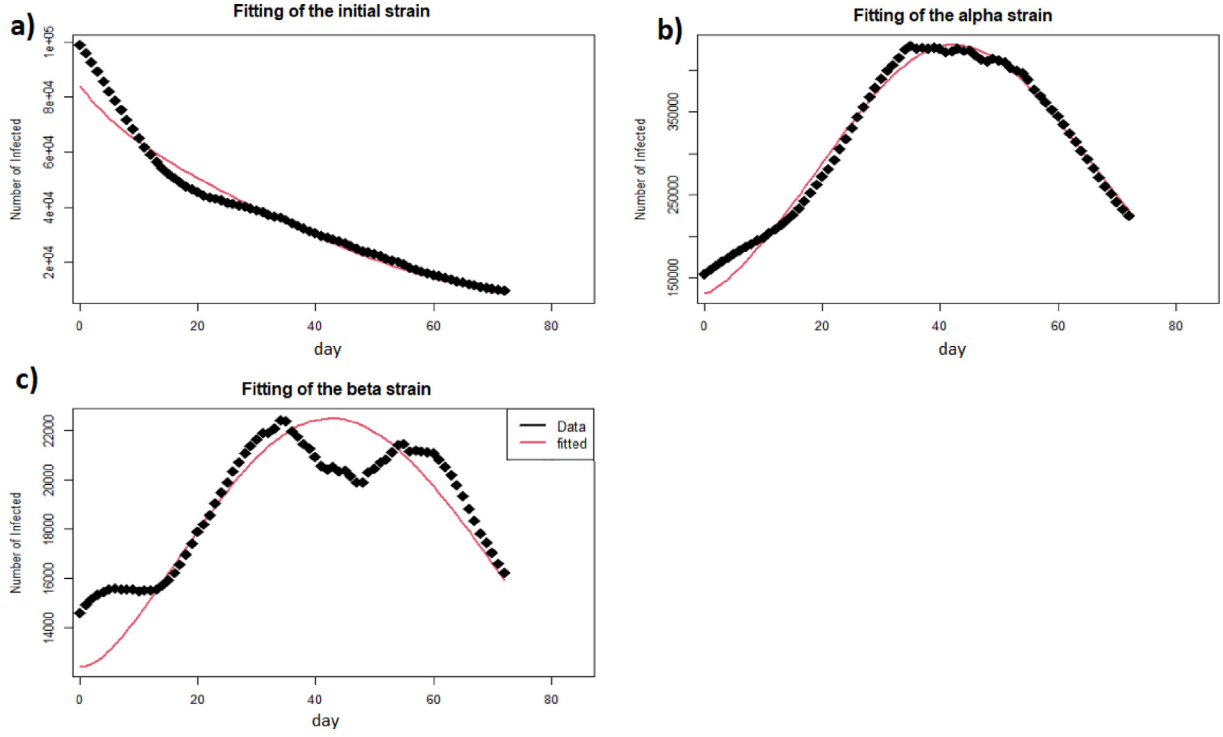

Fig. 3. Fitting the numerically-predicted infection curves (red curves) for the three strains, to the data collected from Geodes (black dots). The fit was performed using the Particle Swarm Optimisation (PSO) algorithm.

Table 2
Parameter values estimated from Geodes data using the PSO algorithm.

Symbol	Value	Symbol	Value	Symbol	Value
$\alpha_i^{V_0}$	2.198814e−09	$\alpha_i^{V_1}$	6.748454e−09	$\alpha_i^{V_2}$	1.970476e−09
$\alpha_a^{V_0}$	2.580832e−09	$\alpha_a^{V_1}$	1.099686e−08	$\alpha_a^{V_2}$	1.353279e−08
p^{V_0}	7.594355e−01	p^{V_1}	8.531318e−01	p^{V_2}	8.229163e−01
b^{V_0}	1.490951e−01	b^{V_1}	2.949318e−01	b^{V_2}	2.110977e−01
δ^{V_0}	7.926397e−02	δ^{V_1}	3.122869e−01	δ^{V_2}	1.194538e−01
μ^{V_0}	7.727916e−02	μ^{V_1}	1.184327e−01	μ^{V_2}	9.728384e−02

Table 3
Biological interpretation of parameters estimated here.

Viral strain	Parameter interpretation	Value
V_0	Number of days for a contaminating contact with a symptomatic	7.11
	Number of days for a contaminating contact with a asymptomatic	6.05
	Percentage of symptomatic individuals	0.76
	Asymptomatics: number of days to recover	6.71
	Symptomatics: number of days to recover	12.62
	Number of days before dying	12.94
V_1 (alpha)	Number of days for a contaminating contact with a symptomatic	2.31
	Number of days for a contaminating contact with a asymptomatic	1.42
	Percentage of symptomatic individuals	0.85
	Asymptomatics: number of days to recover	3.39
	Symptomatics: number of days to recover	3.20
	Number of days before dying	8.44
V_2 (beta/ gamma)	Number of days for a contaminating contact with a symptomatic	7.93
	Number of days for a contaminating contact with a asymptomatic	1.15
	Percentage of symptomatic individuals	0.82
	Asymptomatics: number of days to recover	4.73
	Symptomatics: number of days to recover	8.37
	Number of days before dying	10.28

with the beta strain is quite poor for time $t \in [0, 10]$ days and $t \in [40, 60]$ days.

The evolution of asymptomatic infected people is presented in the Fig. 4. We see that the number of asymptomatic individuals infected with the initial strain increases at the beginning and then decreases (following the trend of symptomatic infected individuals). The asymptomatic individuals infected with the other two variants (alpha and beta) follow the same general pattern as the symptomatic individuals but with fewer cases.

Looking at the numbers in Tables 3 and 2 we see that the people who die the fastest are those affected by the alpha variant. Indeed, $\mu^{V_1} > \mu^{V_2} > \mu^{V_0}$. (Note that this result is also supported by the values of the case fatality rate (CFR) for each of the three variants: $CFR^0 = 5.94\%; CFR^1 = 9.48\%; CFR^2 = 7.98\%$.) In addition, people who take the least time to transmit the virus are those affected by the alpha variant which is consistent with the fact that it is said to be more transmissible. There are a higher number of days for the transmission of the disease to symptomatic people than for the asymptomatic people, and this is true for all three cases.

4. Results

4.1. Estimation of the basic reproduction number

To calculate the value of R_0, we have computed three non-zero eigenvalues for the next generation matrix: $\lambda_0 = 0.898367$ (corresponding to the initial strain), $\lambda_1 = 1.141363$ (corresponding to the alpha strain), $\lambda_2 = 1.140814$ (corresponding to the beta/gamma strain). It is clear that the two variants (with $R_0 = 1.14$) are more contaminating than the initial strain (with $R_0 = 0.898$). Both alpha and beta variants have very close values, so we can conclude at first that they have a similar intensity, at least for France, in the absence of vaccination. Note that this R_0 value is consistent with the value calculated by other French studies that investigated COVID-19 transmission during the same period of time (Gaymard et al., 2020).

4.2. Local sensitivity analysis

We now investigate the impact on R_0 when we vary by one day the biological parameter estimates given in Table 3. For example, if we estimated at 7.11 the number of days required for the transmission of the disease by a symptomatic person, we now consider the interval $[6.11, 8.11]$ days. Each time we vary one parameter, while fixing all other parameters. The results are presented in Fig. 5.

First, we see in Fig. 5(a) that the parameters of the initial strain do not impact R_0 because the corresponding value is never the maximum among the three eigenvalues of the next generation matrix. For the alpha variant (Fig. 5(b)), we see that R_0 is especially sensitive to $\alpha_i^{V_1}$ and $\alpha_a^{V_1}$, parameters describing the transmission of infection following contacts with symptomatic and asymptomatic infected individuals. Thus, reducing the contacts (e.g., via lockdown) could lead to a reduction in R_0. Also notice that none of the parameters related to the alpha variant increase R_0 above 2. For the beta/gamma variants (Fig. 5(c)), we see that R_0 is very sensitive to $\alpha_a^{V_2}$. Changes in this parameter can increase R_0 up to $R_0 \approx 6$ (unlike the case for the alpha variant, where R_0 was always below 2). This shows us the importance of controlling the $\alpha_a^{V_2}$ parameter to be able to control the outbreak of the beta/gamma variant. Another parameter that impacts R_0 is b^{V_2}, the recovery rate of asymptomatic individuals. This shows the key role of the asymptomatic people in the rapid spread of the beta/gamma strain.

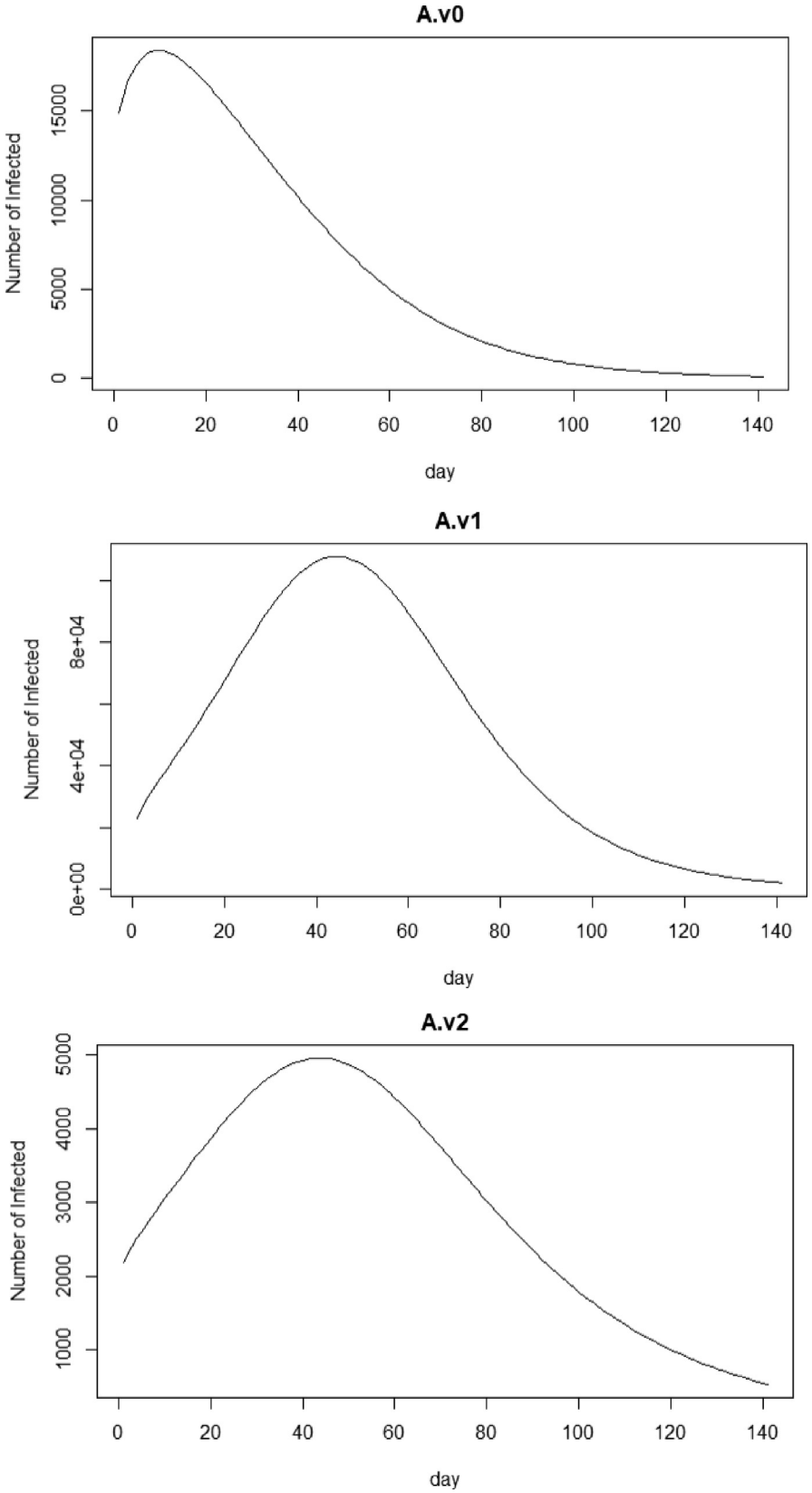

Fig. 4. Time-evolution of the three asymptomatic compartments of model (1).

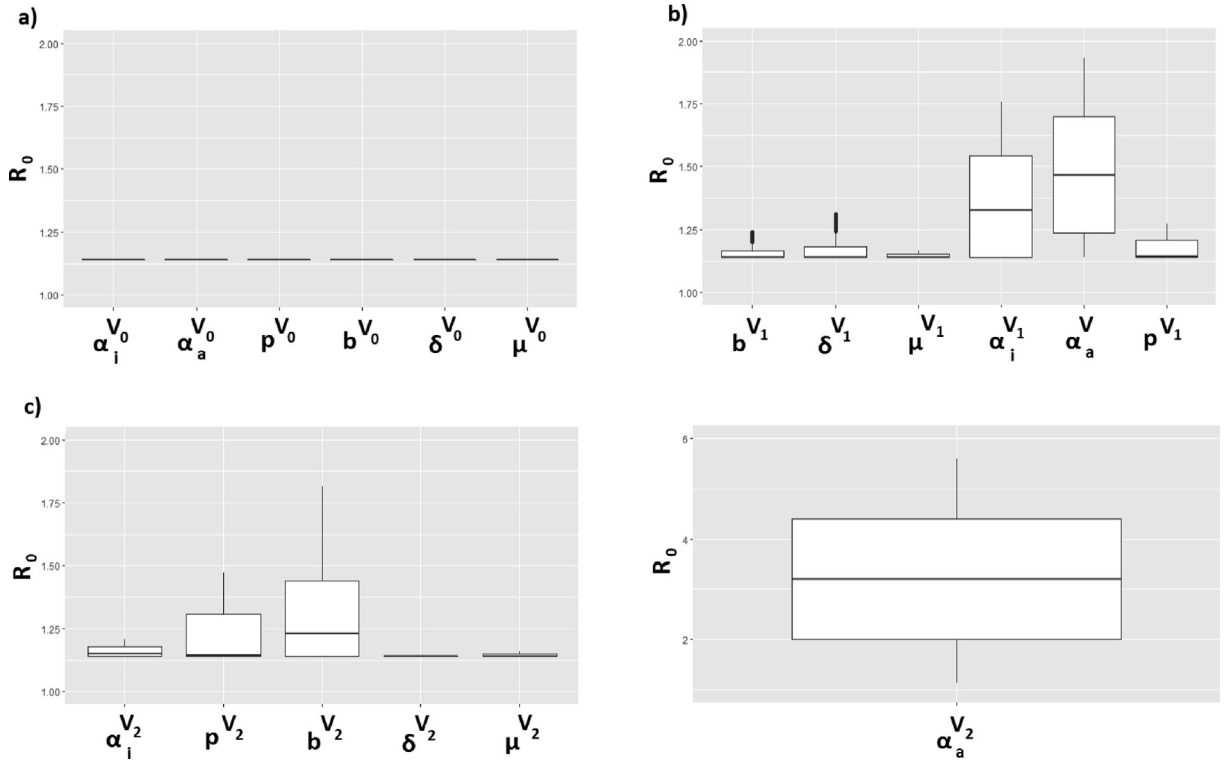

Fig. 5. Local sensitivity of R_0 when we vary: (a) parameters associated with the original strain V_0; (b) parameters associated with the alpha strain V_1; (c) parameters associated with the beta/gamma strain V_2. The right-hand sub-panel in (c) shows, on a much larger vertical axis, R_0 as we vary $\alpha_a^{V_2}$.

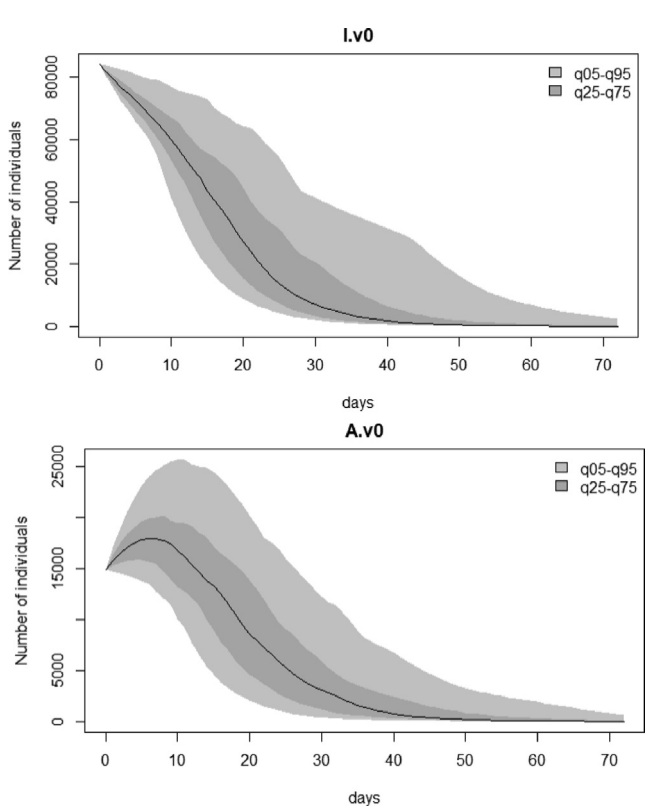

Fig. 6. Uncertainty in the $I^{V_0}(t)$ and $A^{V_0}(t)$, as we vary by a maximum of ± 1 day all model parameters (according to the LHS scheme).

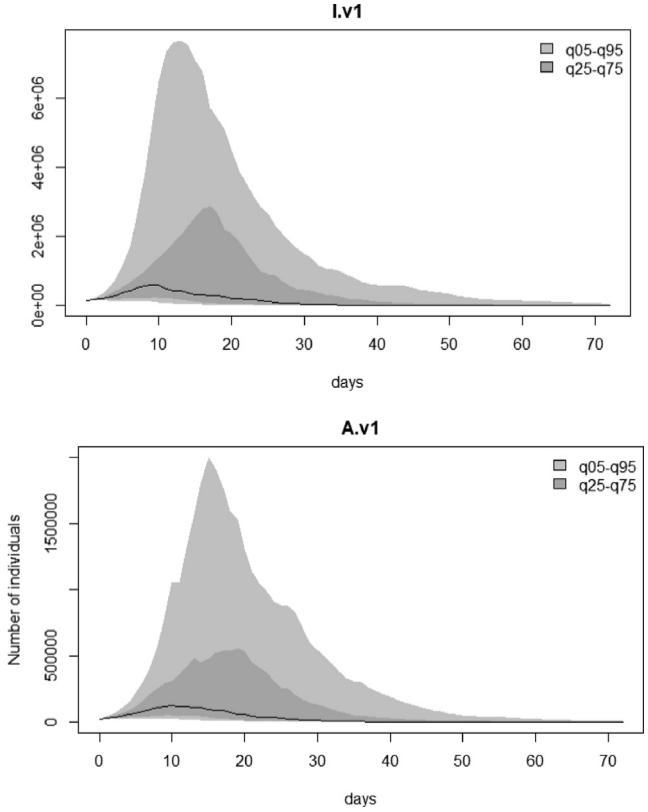

Fig. 7. Uncertainty in the $I^{V_1}(t)$ and $A^{V_1}(t)$ as we vary by a maximum of ± 1 day all model parameters (according to the LHS scheme).

4.3. Global sensitivity analysis for the infected compartments

Since knowledge about model parameters is incomplete (especially knowledge about the role of asymptomatic people on disease transmission), we are now performing a global sensitivity and uncertainty analysis of the different infected compartments, where we vary all parameters at the same time. To this end, we consider a classical approach that combines the Latin Hypercube Sampling (LHS) with the Partial Rank Correlation Coefficient (PRCC) (Marino et al., 2008). The parameters are varied in the same way as before (i.e., by one day for each of the rates), and the sample size for the LHS is 100.

In Fig. 6 we show the variation in the number of individuals infected with the original strain, as we vary all model parameters. The black curves show the median time evolution of $I^{V_0}(t)$ and $A^{V_0}(t)$, while the gray regions show the range between quantiles $q25 - q75$ and $q5 - q95$. We can see that both I^{V_0} and A^{V_0} are very sensitive for $t \in [10, 30]$days. For example, a 1-day change in the various transition rates incorporated into the model, leads on day $t = 10$ (i.e. February 22) to a variation between 50,000 to 80,000 in the symptomatic infected people (I^{V_0}), and a variation between 10,000 to 25,000 asymptomatic infected people approximately. In Fig. 7 we show the variation in the number of individuals infected with the alpha strain, as we vary all model parameters. Note here the change in the maximum peak of both I^{V_1} and A^{V_1}: not only the amplitude but also the shift in time: from day $t \approx 9$ for the median black curve, to day $t \approx 13$ for the $q5 - q95$ quantiles. If we compare these results with the results in Fig. 8 for the beta/gamma strain, we observe that while the maximum for I^{V_1} is reached around

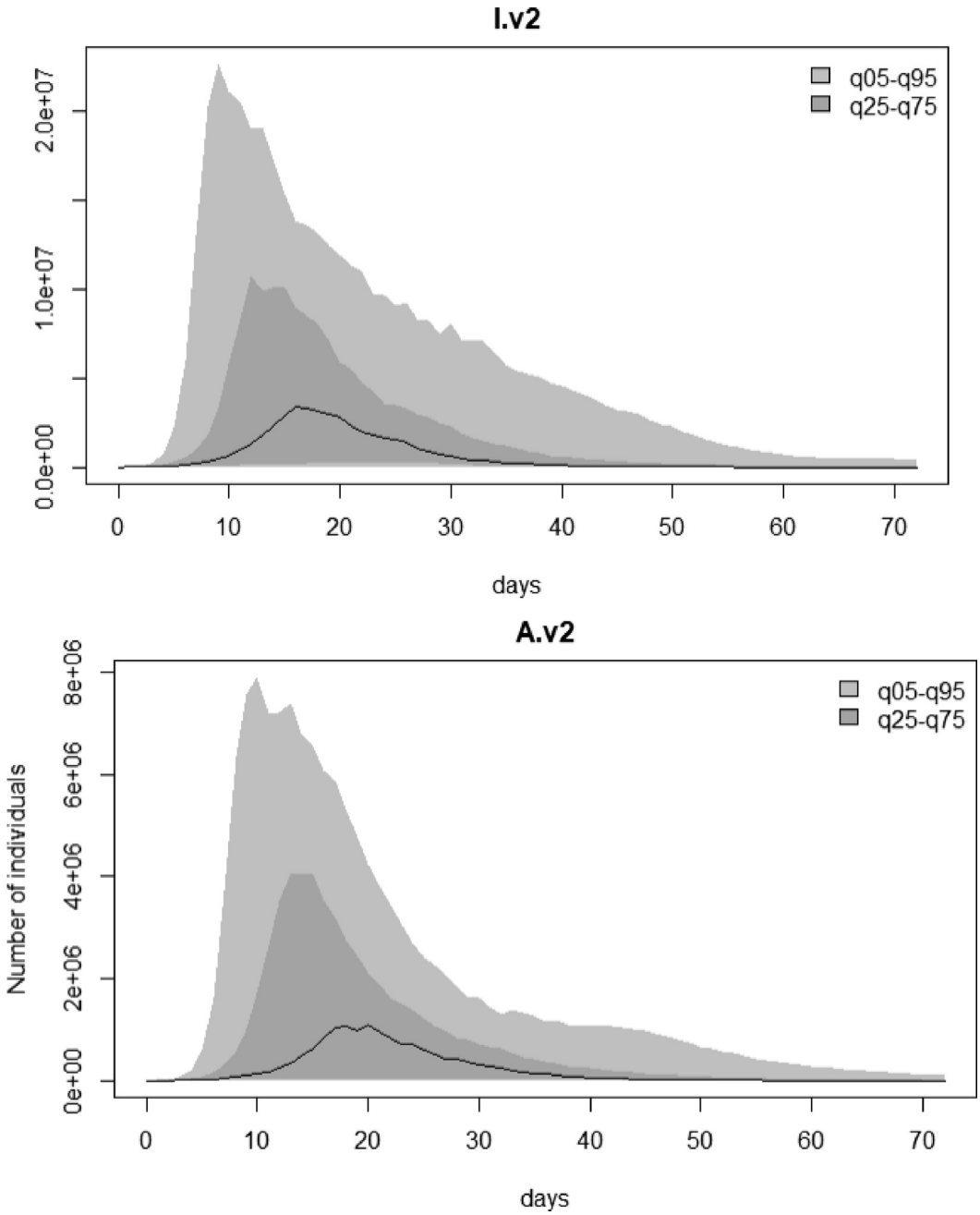

Fig. 8. Uncertainty in the $I^{V_2}(t)$ and $A^{V_2}(t)$ as we vary by a maximum of ± 1 day all model parameters (according to the LHS scheme).

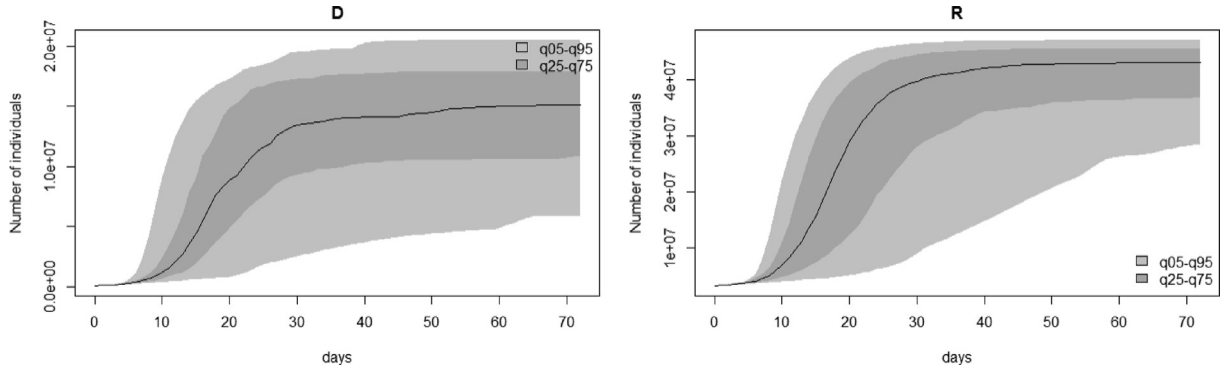

Fig. 9. Uncertainty in the $R(t)$ and $D(t)$ as all model parameters are varied by one day.

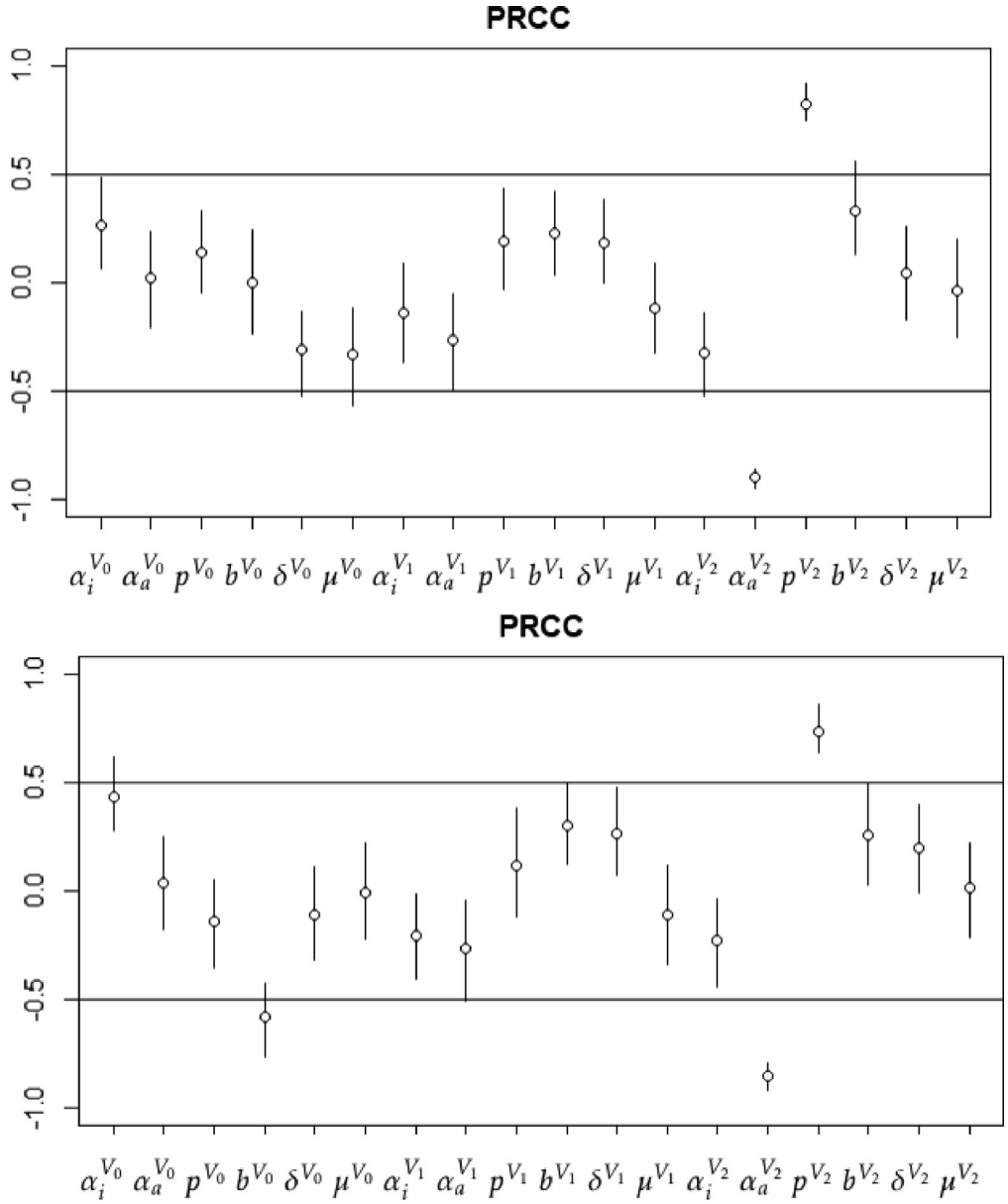

Fig. 10. PRCC of I^{V_0} and A^{V_0}.

$t = 13$ days, for I^{V_2} the maximum is reached for $t \approx 9$ days. Therefore, the beta/gamma strains are more aggressive. Moreover, by comparing Figs. 7,8, we see that I^{V_2} and A^{V_2} could reach much higher amplitudes (i.e., 10-fold higher) compared to I^{V_1} and A^{V_1}. We will return to this aspect in the Discussion Section 5.

Finally, for completeness, in Fig. 9 we also show the time-variation in the recovered $R(t)$ and dead $D(t)$ people. Again, we see a large variability in the outcome between days $t = 10$ and $t = 30$ which, for $D(t)$ lasts also for $t > 30$ days.

To quantify the sensitivity of model outcomes (i.e., number of symptomatic and asymptomatic infected individuals) to changes in model parameters, and identify the critical parameters, in Figs. 10–12 we show the Partial Rank Correlation Coefficient (PRCC) – a sampling-based measure for nonlinear but monotonic relationships between inputs and outputs (Marino et al., 2008). We see here that the parameters with the greatest impact on the outcome (i.e., parameters with PRCC indexes above 0.5 and below −0.5) are:

- For I^{V_0}: parameters $\mu^{V_0}, \alpha_a^{V_2}, p^{V_2}$ and b^{V_2}; while for A^{V_0}: parameters $\alpha_i^{V_0}, b^{V_0}, \alpha_a^{V_2}$, and p^{V_2};
- For I^{V_1}: parameters $b^{V_1}, \alpha_a^{V_2}, p^{V_2}$ and b^{V_2}; while for A^{V_1}: parameters $b^{V_1}, \alpha_a^{V_2}, p^{V_2}$ and b^{V_2};
- For I^{V_2}: parameter p^{V_2} and $\alpha_a^{V_2}$; while for A^{V_2}: parameters b^{V_2} and p^{V_2}.

First we observe that due to the interactions between the model variables, parameters associated with the second variant (V_2) impact the infections with the original (V_0) strain and the infections with the first variant (V_1). For example, p^{V_2} impacts not only I^{V_2} and A^{V_2}, but also $I^{V_0}, A^{V_0}, I^{V_1}$ and A^{V_1}, which suggest that the people infected by beta/gamma strain could have a strong impact on the overall evolution of epidemics. This is despite the fact that for R_0, the impact of beta/gamma variant was decoupled from the impact of the original variant and alpha variant (see Eqn. (2)).

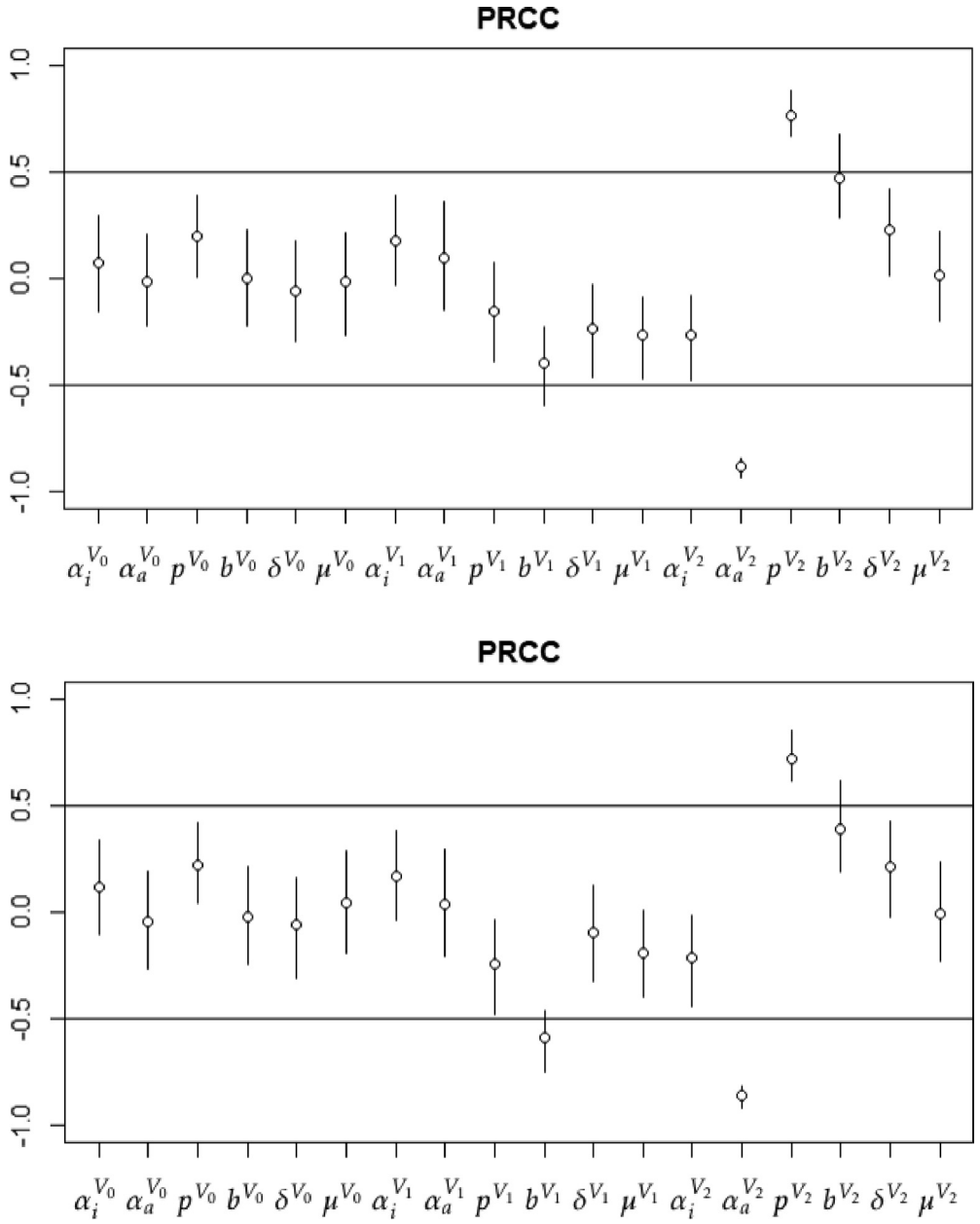

Fig. 11. PRCC of I^{V_1} and A^{V_1}.

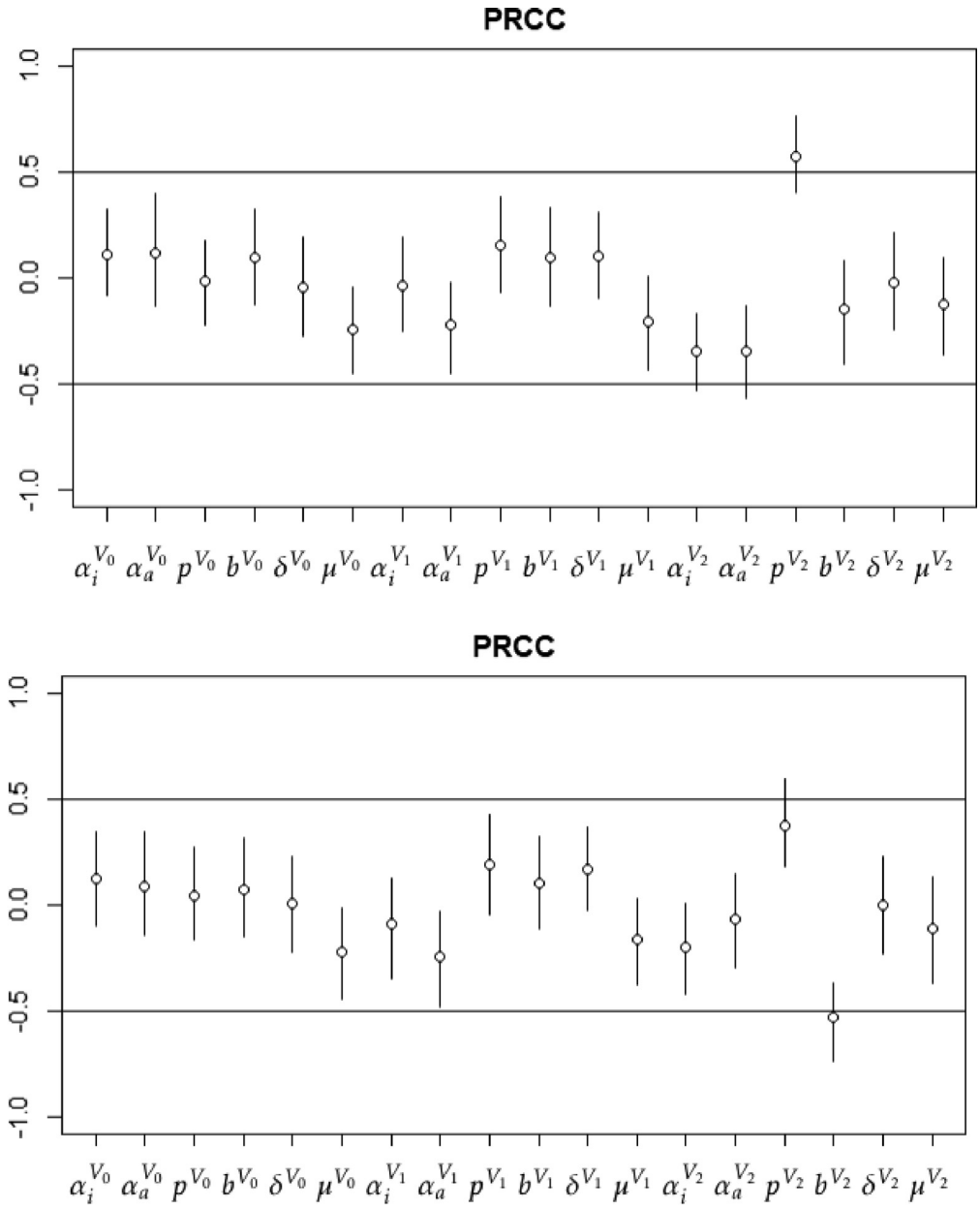

Fig. 12. PRCC of I^{V_2} and A^{V_2}.

5. Discussion & conclusion

In this paper we investigated the evolution of COVID-19 epidemic in France between February-May 2021 in the presence of multiple viral strains, while estimating the parameters that characterise this epidemic. The time interval for the data was chosen to reduce the impact of vaccination on the results (although other control measures such as masks and social distancing were still in place in France during this period). We focused on three viral variants identified in the data: the original (i.e., March 2020) SARS-CoV-2 variant, the alpha variant (first detected in the UK) and the beta/gamma variant (first detected in South Africa and in Brazil). The last two were considered together here, because in France the tests did not discriminate between them.

First, we showed that the value for R_0 was given by the maximum of three basic reproduction values corresponding to the three different variants. This suggested a sort of decoupling between the evolution of different variants, the most transmissible one having the greatest influence on the evolution of the outbreak. Using French data on Geodes (2021) and a PSO algorithm, we estimated $R_0 = 1.14$, consistent with the value estimated by other French studies (Gaymard et al., 2020). Moreover, this R_0 value was similar for the alpha and beta/gamma variants, but larger than the value corresponding to the original strain (thus suggesting that between February-May the alpha and beta/gamma variants were more transmissible than the original strain). A local sensitivity analysis for R_0 showed that the parameters with the largest impact were the transmission rates of symptomatic and asymptomatic individuals; see Fig. 5.

In regard to parameter estimation (see Table 3) we also found a higher number of days necessary to create contaminating contacts for the symptomatic people than for the asymptomatic people (for all three variants). This unexpected result could be explained by the fact that asymptomatic people come into contact with more peers ignoring distancing rules. The results in Table 3 show also that the number of days to recover when symptomatic is slightly lower than for asymptomatic people for the alpha variant, while it is almost twice more for the other strains. It would be interesting

to investigate whether there is a specificity of the alpha variant in contrast to the other variants. At this point we could not find any epidemiological data to test the validity of this theoretical result.

Second, a global sensitivity and uncertainty analysis identified the variations in the amplitude of symptomatic and asymptomatic infections for all three strains, as well as the day when these infections peak (see Figs. 6–8). Moreover, the PRCC analysis identified some critical parameters for the evolution of the epidemics (see Figs. 10–12). The interesting result was that parameters associated with the V_2 variant (p^{V_2}) seemed to impact also the evolution of infections with the V_0 and V_1 variants, thus suggesting that the infections with the different viral strains are not really decoupled (as suggested by the R_0 formula (2)).

The uncertainty analysis also showed that the evolution of those infected with the beta/gamma strain i.e. V^2 has more uncertainty compared to the other two strains (when we vary the interaction rates by one day).

Also in regard to the uncertainty of the results in this study, we need to mention the fact that since we could not find mortality data for each of the different strains discussed here, the exact identification of parameters μ^{V_i} actually depends on all other model parameters listed in Table 2.

Throughout this study we chose the average infection period to be 13 days, since this was similar to other published studies (George et al., 2021). Even reducing this period by 1–2 days did not seem to lead to significant changes. However, in the future it would be interesting to investigate model dynamics when we change (increase/decrease) this infection period in a more significant way.

This study focused on the dynamics of COVID-19 epidemics in the presence of alpha and beta/gamma strains, with distancing measures (and masks worn indoor and outdoor) but no vaccinations (as the vaccination program in France really took off in May 2021). Our results about the potential large amplitude in the infections with beta and gamma strains, can be understood in this context: these strains would have caused a disastrous outbreak situation if the situation had not changed with the arrival of the vaccine. The subsequent emergence of the delta variant also changed the dynamics of the epidemics, and if we include it into our model (1) the simulation results and predictions will inevitably change.

The study can be further extended to the new SARS-CoV-2 variants that are appearing, which might avoid the anti-viral immune responses generated by the vaccination. In regard to this, a very recent review (Malik et al., 2022) discussed vaccine effectiveness in the context of the multiple SARS-CoV-2 variants that emerged over the last months, and concluded that since a large number of variants have mutations mainly associated with the spike protein, which is also a key component of most of the vaccines on the market, vaccine efficacy needs to be assessed for each variant. While some vaccine efficacy studies have been published for earlier variants (Bian et al., 2021), many more such studies are ongoing.

CRediT authorship contribution statement

Mathilde Massard: Conceptualization, Writing – original draft, Writing – review & editing. **Raluca Eftimie:** Supervision. **Antoine Perasso:** Supervision. **Bruno Saussereau:** Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

MM, RE, AP and BS work was supported by a MODCOV19 project. In addition, AP work was supported by an EpiMaths project.

Appendix A. Geodes data set

In Tables 4–8 below we summarise the Geodes data (Geodes, 2021) used throughout this study.

Table 4
Percentage of initial strain.

Date	Value	Date	Value	Date	Value
2021-02-12-2021-02-18	42,9	2021-02-13-2021-02-19	40,7	2021-02-14-2021-02-20	39,7
2021-02-15-2021-02-21	39,5	2021-02-16-2021-02-22	37,3	2021-02-17-2021-02-23	35,3
2021-02-18-2021-02-24	33,6	2021-02-19-2021-02-25	32,0	2021-02-20-2021-02-26	30,3
2021-02-21-2021-02-27	29,5	2021-02-22-2021-02-28	29,4	2021-02-23-2021-03-01	27,2
2021-02-24-2021-03-02	25,7	2021-02-25-2021-03-03	24,3	2021-02-26-2021-03-04	22,9
2021-02-27-2021-03-05	21,6	2021-02-28-2021-03-06	20,9	2021-03-01-2021-03-07	20,7
2021-03-02-2021-03-08	19,1	2021-03-03-2021-03-09	18,1	2021-03-04-2021-03-10	17,0
2021-03-05-2021-03-11	15,7	2021-03-06-2021-03-12	14,7	2021-03-07-2021-03-13	14,2
2021-03-08-2021-03-14	14,1	2021-03-09-2021-03-15	13,0	2021-03-10-2021-03-16	12,2
2021-03-11-2021-03-17	11,5	2021-03-12-2021-03-18	11,0	2021-03-13-2021-03-19	10,4
2021-03-14-2021-03-20	10,2	2021-03-15-2021-03-21	10,1	2021-03-16-2021-03-22	9,4
2021-03-17-2021-03-23	8,9	2021-03-18-2021-03-24	8,4	2021-03-19-2021-03-25	8,0
2021-03-20-2021-03-26	7,6	2021-03-21-2021-03-27	7,4	2021-03-22-2021-03-28	7,3
2021-03-23-2021-03-29	6,9	2021-03-24-2021-03-30	6,7	2021-03-25-2021-03-31	6,5
2021-03-26-2021-04-01	6,3	2021-03-27-2021-04-02	6,1	2021-03-28-2021-04-03	6,0
2021-03-29-2021-04-04	6,0	2021-03-30-2021-04-05	5,9	2021-03-31-2021-04-06	5,6
2021-04-01-2021-04-07	5,4	2021-04-02-2021-04-08	5,4	2021-04-03-2021-04-09	5,3
2021-04-04-2021-04-10	5,2	2021-04-05-2021-04-11	5,2	2021-04-06-2021-04-12	5,1
2021-04-07-2021-04-13	4,9	2021-04-08-2021-04-14	4,8	2021-04-09-2021-04-15	4,6
2021-04-10-2021-04-16	4,3	2021-04-11-2021-04-17	4,3	2021-04-12-2021-04-18	4,2
2021-04-13-2021-04-19	4,0	2021-04-14-2021-04-20	3,9	2021-04-15-2021-04-21	3,8
2021-04-16-2021-04-22	3,8	2021-04-17-2021-04-23	3,7	2021-04-18-2021-04-24	3,7
2021-04-19-2021-04-25	3,7	2021-04-20-2021-04-26	3,6	2021-04-21-2021-04-27	3,5
2021-04-22-2021-04-28	3,4	2021-04-23-2021-04-29	3,4	2021-04-24-2021-04-30	3,4
2021-04-25-2021-05-01	3,4	2021-04-26-2021-05-02	3,4	2021-04-27-2021-05-03	3,3
2021-04-28-2021-05-04	3,4	2021-04-29-2021-05-05	3,3	2021-04-30-2021-05-06	3,4
2021-05-01-2021-05-07	3,4	2021-05-02-2021-05-08	3,4	2021-05-03-2021-05-09	3,4
2021-05-04-2021-05-10	3,3	2021-05-05-2021-05-11	3,2	2021-05-06-2021-05-12	3,3
2021-05-07-2021-05-13	3,2				

Table 5
Percentage of alpha variant.

Date	Value	Date	Value	Date	Value
2021-02-12-2021-02-18	44,3	2021-02-13-2021-02-19	46,0	2021-02-14-2021-02-20	46,9
2021-02-15-2021-02-21	47,1	2021-02-16-2021-02-22	49,2	2021-02-17-2021-02-23	51,0
2021-02-18-2021-02-24	52,8	2021-02-19-2021-02-25	54,3	2021-02-20-2021-02-26	55,9
2021-02-21-2021-02-27	56,6	2021-02-22-2021-02-28	56,8	2021-02-23-2021-03-01	59,0
2021-02-24-2021-03-02	60,6	2021-02-25-2021-03-03	62,1	2021-02-26-2021-03-04	63,5
2021-02-27-2021-03-05	64,8	2021-02-28-2021-03-06	65,6	2021-03-01-2021-03-07	65,8
2021-03-02-2021-03-08	67,2	2021-03-03-2021-03-09	68,1	2021-03-04-2021-03-10	69,1
2021-03-05-2021-03-11	70,1	2021-03-06-2021-03-12	71,2	2021-03-07-2021-03-13	71,6
2021-03-08-2021-03-14	71,7	2021-03-09-2021-03-15	73,0	2021-03-10-2021-03-16	73,9
2021-03-11-2021-03-17	74,6	2021-03-12-2021-03-18	75,2	2021-03-13-2021-03-19	75,6
2021-03-14-2021-03-20	75,9	2021-03-15-2021-03-21	75,9	2021-03-16-2021-03-22	76,5
2021-03-17-2021-03-23	76,9	2021-03-18-2021-03-24	77,7	2021-03-19-2021-03-25	78,1
2021-03-20-2021-03-26	78,8	2021-03-21-2021-03-27	79,2	2021-03-22-2021-03-28	79,3
2021-03-23-2021-03-29	80,1	2021-03-24-2021-03-30	80,6	2021-03-25-2021-03-31	80,8
2021-03-26-2021-04-01	81,2	2021-03-27-2021-04-02	81,4	2021-03-28-2021-04-03	81,5
2021-03-29-2021-04-04	81,5	2021-03-30-2021-04-05	81,4	2021-03-31-2021-04-06	81,8
2021-04-01-2021-04-07	82,1	2021-04-02-2021-04-08	82,2	2021-04-03-2021-04-09	82,3
2021-04-04-2021-04-10	82,6	2021-04-05-2021-04-11	82,6	2021-04-06-2021-04-12	82,7
2021-04-07-2021-04-13	82,5	2021-04-08-2021-04-14	82,3	2021-04-09-2021-04-15	82,4
2021-04-10-2021-04-16	82,4	2021-04-11-2021-04-17	82,3	2021-04-12-2021-04-18	82,3
2021-04-13-2021-04-19	82,2	2021-04-14-2021-04-20	82,5	2021-04-15-2021-04-21	82,6
2021-04-16-2021-04-22	82,5	2021-04-17-2021-04-23	82,4	2021-04-18-2021-04-24	82,3
2021-04-19-2021-04-25	82,3	2021-04-20-2021-04-26	82,2	2021-04-21-2021-04-27	81,9
2021-04-22-2021-04-28	81,6	2021-04-23-2021-04-29	81,3	2021-04-24-2021-04-30	80,8
2021-04-25-2021-05-01	80,5	2021-04-26-2021-05-02	80,5	2021-04-27-2021-05-03	80,2
2021-04-28-2021-05-04	79,9	2021-04-29-2021-05-05	79,7	2021-04-30-2021-05-06	79,4
2021-05-01-2021-05-07	79,1	2021-05-02-2021-05-08	79,1	2021-05-03-2021-05-09	79,1
2021-05-04-2021-05-10	78,2	2021-05-05-2021-05-11	77,9	2021-05-06-2021-05-12	77,6
2021-05-07-2021-05-13	77,4				

Table 6
Percentage of beta variant.

Date	Value	Date	Value	Date	Value
2021-02-12-2021-02-18	4,5	2021-02-13-2021-02-19	4,6	2021-02-14-2021-02-20	4,6
2021-02-15-2021-02-21	4,6	2021-02-16-2021-02-22	4,6	2021-02-17-2021-02-23	4,8
2021-02-18-2021-02-24	4,9	2021-02-19-2021-02-25	5,0	2021-02-20-2021-02-26	5,2
2021-02-21-2021-02-27	5,4	2021-02-22-2021-02-28	5,4	2021-02-23-2021-03-01	5,5
2021-02-24-2021-03-02	5,4	2021-02-25-2021-03-03	5,3	2021-02-26-2021-03-04	5,2
2021-02-27-2021-03-05	5,0	2021-02-28-2021-03-06	4,9	2021-03-01-2021-03-07	4,8
2021-03-02-2021-03-08	4,7	2021-03-03-2021-03-09	4,7	2021-03-04-2021-03-10	4,9
2021-03-05-2021-03-11	5,0	2021-03-06-2021-03-12	5,0	2021-03-07-2021-03-13	5,0
2021-03-08-2021-03-14	5,0	2021-03-09-2021-03-15	5,0	2021-03-10-2021-03-16	5,0
2021-03-11-2021-03-17	4,9	2021-03-12-2021-03-18	4,8	2021-03-13-2021-03-19	4,7
2021-03-14-2021-03-20	4,6	2021-03-15-2021-03-21	4,6	2021-03-16-2021-03-22	4,6
2021-03-17-2021-03-23	4,5	2021-03-18-2021-03-24	4,4	2021-03-19-2021-03-25	4,4
2021-03-20-2021-03-26	4,4	2021-03-21-2021-03-27	4,3	2021-03-22-2021-03-28	4,3
2021-03-23-2021-03-29	4,3	2021-03-24-2021-03-30	4,2	2021-03-25-2021-03-31	4,2
2021-03-26-2021-04-01	4,1	2021-03-27-2021-04-02	4,1	2021-03-28-2021-04-03	4,0
2021-03-29-2021-04-04	4,1	2021-03-30-2021-04-05	4,1	2021-03-31-2021-04-06	4,0
2021-04-01-2021-04-07	3,9	2021-04-02-2021-04-08	3,9	2021-04-03-2021-04-09	3,9
2021-04-04-2021-04-10	3,8	2021-04-05-2021-04-11	3,8	2021-04-06-2021-04-12	3,8
2021-04-07-2021-04-13	3,9	2021-04-08-2021-04-14	4,0	2021-04-09-2021-04-15	4,0
2021-04-10-2021-04-16	4,1	2021-04-11-2021-04-17	4,2	2021-04-12-2021-04-18	4,2
2021-04-13-2021-04-19	4,4	2021-04-14-2021-04-20	4,6	2021-04-15-2021-04-21	4,7
2021-04-16-2021-04-22	4,9	2021-04-17-2021-04-23	5,0	2021-04-18-2021-04-24	5,1
2021-04-19-2021-04-25	5,1	2021-04-20-2021-04-26	5,1	2021-04-21-2021-04-27	5,1
2021-04-22-2021-04-28	5,2	2021-04-23-2021-04-29	5,3	2021-04-24-2021-04-30	5,4
2021-04-25-2021-05-01	5,5	2021-04-26-2021-05-02	5,5	2021-04-27-2021-05-03	5,6
2021-04-28-2021-05-04	5,7	2021-04-29-2021-05-05	5,7	2021-04-30-2021-05-06	5,7
2021-05-01-2021-05-07	5,7	2021-05-02-2021-05-08	5,7	2021-05-03-2021-05-09	5,7
2021-05-04-2021-05-10	5,9	2021-05-05-2021-05-11	6,0	2021-05-06-2021-05-12	5,9
2021-05-07-2021-05-13	6,0				

Table 7
Number of new confirmed cases in France per day.

Date	Value	Value	Value	Value	Value
2021-02-12-2021-02-16	20701	21231	16546	4376	19590
2021-02-17-2021-02-21	25018	22501	24116	22371	22046
2021-02-22-2021-02-26	4646	20064	31519	25403	25207
2021-02-27-2021-03-03	23996	19952	4703	22857	26788
2021-03-04-2021-03-08	25279	23507	23306	21825	5327
2021-03-09-2021-03-13	23302	30303	27166	25229	29759
2021-03-14-2021-03-18	26343	6471	29975	38501	34998
2021-03-19-2021-03-23	35088	35327	30581	15792	14678
2021-03-24-2021-03-28	65373	45641	41869	42619	37014
2021-03-29-2021-04-02	9094	30702	59038	50659	46677
2021-04-03-2021-04-07	13917	66794	10793	8045	12951
2021-04-08-2021-04-12	84999	41243	43284	34895	8536
2021-04-13-2021-04-17	39113	43505	38045	36442	35861
2021-04-18-2021-04-22	29344	6696	43098	34968	34318
2021-04-23-2021-04-27	32340	32633	24465	5952	30317
2021-04-28-2021-05-02	31539	26538	24299	25670	9888
2021-05-03-2021-05-07	3760	24371	26000	21712	19124
2021-05-08-2021-05-12	20745	9128	3292	19791	21498
2021-05-13	19461				

Table 8
Number of new confirmed cases in France per day, averaged over 7 days.

Date	Value	Date	Value	Date	Value
12/02-18/02	18566,14286	13/02-19/02	19054,00000	14/02-20/02	19216,85714
15/02-21/02	20002,57143	16/02-22/02	20041,14286	17/02-23/02	20108,85714
18/02-24/02	21037,57143	19/02-25/02	21452,14286	20/02-26/02	21608,00000
21/02-27/02	21840,14286	22/02-28/02	21541,00000	23/02-01/03	21549,14286
24/02-02/03	21948,14286	25/02-03/03	21272,28571	26/02-04/03	21254,57143
27/02-05/03	21011,71429	28/02-06/03	20913,14286	01/03-07/03	21180,71429
02/03-08/03	21269,85714	03/03-09/03	21333,42857	04/03-10/03	21835,57143
05/03-11/03	22105,14286	06/03-12/03	22351,14286	07/03-13/03	23273,00000
08/03-14/03	23918,42857	09/03-15/03	24081,85714	10/03-16/03	25035,14286
11/03-17/03	26206,28571	12/03-18/03	27325,14286	13/03-19/03	28733,57143
14/03-20/03	29529,00000	15/03-21/03	30134,42857	16/03-22/03	31466,00000
17/03-23/03	29280,71429	18/03-24/03	33119,57143	19/03-25/03	34640,00000
20/03-26/03	35608,71429	21/03-27/03	36650,42857	22/03-28/03	37569,42857
23/03-29/03	36612,57143	24/03-30/03	38901,71429	25/03-31/03	37996,71429
26/03-01/04	38713,57143	27/03-02/04	39400,42857	28/03-03/04	35300,14286
29/03-04/04	39554,42857	30/03-05/04	39797,14286	31/03-06/04	36560,42857
01/04-07/04	29976,57143	02/04-08/04	34882,28571	03/04-09/04	34106,00000
04/04-10/04	38301,28571	05/04-11/04	33744,28571	06/04-12/04	33421,85714
07/04-13/04	37860,14286	08/04-14/04	42225,00000	09/04-15/04	35517,28571
10/04-16/04	34831,42857	11/04-17/04	33771,00000	12/04-18/04	32978,00000
13/04-19/04	32715,14286	14/04-20/04	33284,42857	15/04-21/04	32064,85714
16/04-22/04	31532,42857	17/04-23/04	30946,42857	18/04-24/04	30485,28571
19/04-25/04	29788,28571	20/04-26/04	29682,00000	21/04-27/04	27856,14286
22/04-28/04	27366,28571	23/04-29/04	26254,85714	24/04-30/04	25106,14286
25/04-01/05	24111,42857	26/04-02/05	22029,00000	27/04-03/05	21715,85714
28/04-04/05	20866,42857	29/04-05/05	20075,14286	30/04-06/05	19385,71429
01/05-07/05	18646,42857	02/05-08/05	17942,85714	03/05-09/05	17834,28571
04/05-10/05	17767,42857	05/05-11/05	17113,14286	06/05-12/05	16470,00000
07/05-13/05	16148,42857				

We note that the values in Table 7, corresponding to the number of cases on Mondays, are lower because fewer tests are carried out on Sundays and therefore fewer new cases are confirmed the next day. To remove this effect, we take an average over 7 days at each time, and obtain the data summarised in Table 8.

References

Akman, D., Akman, O., Schaefer, E., 2018. Parameter estimation in ordinary differential equations modeling via particle swarm optimization. J. Appl. Math.

Amiri Mehra, A.H., Shafieirad, M., Abbasi, Z., Zamani, I., 2020. Parameter estimation and prediction of covid-19 epidemic turning point and ending time of a case study on sir/sqair epidemic models. Comput. Math. Methods Med.

Arruda, E.F., Das, S.S., Dias, C.M., Pastore, D.H., 2021. Modelling and optimal control of multi strain epidemics, with application to covid-19. PLoS ONE 16, (9) e0257512.

Bentout, S., Tridanee, A., Djilaliand, S., Touaoula, T.M., 2021. Age-structured modeling of covid-19 epidemic in the usa, uae and algeria. Alexandria Eng. J. 60 (1).

Bian, Lianlian, Gao, Qiushuang, Gao, Fan, Wang, Qian, He, Qian, Xing, Wu., Mao, Qunying, Miao, Xu., Liang, Zhenglun, 2021. Impact of the delta variant on vaccine efficacy and response strategies. Expert Rev. Vaccines 20 (10), 1201–1209.

Brouqui, P., Colson, P., Melenotte, C., Houhamdi, L., Bedotto, M., Devaux, C., Gautret, P., Million, M., Parola, P., La Didier Stoupan, B., Scola, J.-C. Lagier, Raoult, D., 2021. Covid-19 re-infection. Eur. J. Clin. Invest. 51, e13537.

Calafiore, G.C., Fracastoro, G., 2021. Age class structure in sird models for the covid-19–an analysis of tennessee data. In: 2021 29th Mediterranean Conference on Control and Automation (MED).

Centers for Disease Control and Prevention (CDC). Sars-cov-2 variant classifications and definitions. https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html#print, October 2021.

Décret n2020-1706 du 24 décembre 2020 authentifiant les chiffres des populations de métropole, des départements d'outre-mer de la guadeloupe, de la guyane, de la martinique et de la réunion, et des collectivités de saint-barthélemy, de saint-

martin et de saint-pierre-et-miquelon. https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000042742102.
Diekmann, O., Heesterbeek, J.A.P., Metz, J.A.J., 1990. On the definition and the computation of the basic reproduction ratio r_0) in models for infectious diseases in heterogeneous populations. J. Math. Biol. 28 (4), 365–382.
Diekmann, O., Heesterbeek, J.A.P., Roberts, M.G., 2010. The construction of next-generation matrices for compartmental epidemic models. J. R. Soc. Interface 7, 873–885.
European Centre for Disease Prevention and Control (ECDC). Sars-cov-2 variants of concern as of 21 october 2021. https://www.ecdc.europa.eu/en/covid-19/variants-concern, October 2021.
Faranda, D., Alberti, T., 2020. Modeling the second wave of covid-19 infections in france and italy via a stochastic seir model. Chaos 30(11).
Fernández-Villaverde, J., Jones, C.I., 2020. Estimating and simulating a sird model of covid-19 for many countries, states, and cities (Technical report). National Bureau of Economic Research.
Gaymard, A., Bosetti, P., Feri, A., Destras, G., Enouf, V., Andronico, A., Burrel, S., Behillil, S., Sauvage, C., Bal, A., Morfin, F., Van Der Werf, S., Josset, L., French viro COVID group, Blanquart, F., Coignard, B., Cauchemez, S., Lina, B., 2021. Early assessment of diffusion and possible expansion of sars-cov-2 lineage 20i/501y. v1 (b.1.1.7, variant of concern 202012/01) in france, january to March 2021. Euro Surveill 26(9):pii=2100133.
Geodes. https://geodes.santepubliquefrance.fr/#view=map2&c=indicator.
George, N., Tyagi, N.K., Prasad, J.B., 2021. Covid-19 pandemic and its average recovery time in indian states. Clin. Epidemiol. Global Health.
Gonzalez-Parra, G., Martínez-Rodríguez, D., Villanueva-Micó, R.J., 2021. Impact of a new sars-cov-2 variant on the population: A mathematical modeling approach. Math. Comput. Appl. 26(2).
Harvey, W.T., Carabelli, A.M., Jackson, B., Gupta, R.K., Thomson, E.C., Harrison, E.M., Ludden, C., Reeve, R., Rambaut, A., 2021. COVID-19 Genomics UK (COG-UK) Consortium, and S.J. Peacock andD.L. Robertson. SARS-CoV-2 variants, spike mutations and immune escape. Nat. Rev. Microbiol. 19, 409–424.
He, S., Peng, Y., Sun, K., 2020. SEIR modeling of the COVID-19 and its dynamics. Nonlinear Dyn. 101 (3).
He, S., Tang, S., Rong, L., 2020. A discrete stochastic model of the covid-19 outbreak: Forecast and control. Math. Biosci. Eng. 17(4).
Kennedy, J., Eberhart, R., 1995. Particle swarm optimization. In: Proceedings of ICNN'95-international conference on neural networks, vol. 4.
Khyar, O., Allali, K., 2020. Global dynamics of a multi-strain seir epidemic model with general incidence rates: application to covid-19 pandemic. Nonlinear Dyn. 102 (1).
Malik, J.A., Ahmed, S., Mir, A., Shinde, M., Bender, O., Alshammari, F., Ansari, M., Anwar, S., 2022. The sars-cov-2 mutations versus vaccine effectiveness: new opportunities to new challenges. J. Infect. Public Health 15, 228–240.
Marino, S., Hogue, I.B., Ray, C.J., Kirschner, D.E., 2008. A methodology for performing global uncertainty and sensitivity analysis in systems biology. J. Theor. Biol. 254 (1), 178–196.
Marinov, T.T., Marinova, R.S., 2020. Dynamics of covid-19 using inverse problem for coefficient identification in sir epidemic models. Chaos Solitons Fractals: X 5, 100041.
Nonaka, C.K.V., Franco, M.M., de Lorenzo Barcia, C.A., Gräf, T., de Ávila Mendonça, R. N., de Sousa, K.A.F., Neiva, L.M.C., Fosenca, V., Mendes, A.V.A., de Aguiar, R.S., Giovanetti, M., de Freitas Souza, B.S., 2021. Genomic evidence of sars-cov-2 reinfection involving e484k spike mutation, brazil. Emerg. Infect. Diseases 27 (5).
Perasso, A., 2018. An introduction to the basic reproduction number in mathematical epidemiology. ESAIM: Proceedings and Surveys, 62.
Pottier, L., 2020. Simulation of the covid19 epidemic in France. In: Working paper or preprint.
Santé publique france. https://www.santepubliquefrance.fr/.
Sridhar, A., Yagan, O., Eletreby, R., Levin, S.A., Plotkin, J.B., Poor, H.V., 2021. Leveraging a multiple-strain model with mutations in analyzing the spread of covid-19.
Tang, B., Wang, X., Li, Q., Bragazzi, N.L., Tang, S., Xiao, Y., Wu, J., 2020. Estimation of the transmission risk of the 2019-ncov and its implication for public health interventions. J. Clin. Med. 9 (2).
Tang, X., Wu, C., Li, X., Song, Y., Yao, X., Wu, X., Duan, Y., Zhang, H., Wang, Y., Qian, Z., Cui, J., Lu, J., 2020. On the origin and continuing evolution of sars-cov-2. Natl. Sci. Rev. 7 (6).
Van den Driessche, P., Watmough, J., 2002. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math. Biosci. 180 (1–2).
Wise, J., 2021. Covid-19: The e484k mutation and the risks it poses. Br. Med. J. 372.
Xiao, C., Huang, Ziyue, Wang, Jingxuan, Zhao, Shi, Wong, Martin C.S., Chong, Marc K. C., He, Daihai, Li, Jinhui, 2020. The rate of asymptomatic covid-19 infection: A systematic review and meta-analysis including 12,713 infections from 136 studies.
Yagan, O., Sridhar, A., Eletreby, R., Levin, S., Plotkin, J.B., Poor, H.V., 2021. Modelling and analysis of the spread of covid-19under a multiple-strain. Harvard Data Sci. Rev. 1.
Yavuz, M., Coşar, F.Ö., Günay, F., Özdemir, F.N., 2021. A new mathematical modeling of the covid-19 pandemic including the vaccination campaign. Open J. Model. Simul. 9 (3).
Zhou, P., Yang, X.L., Wang, X.G., Hu, B., Zhang, L., Zhang, W., et al., 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579 (7798).