# Plateaus, rebounds and the effects of individual behaviours in epidemics

**Henri Berestycki**[1,2,✉], **Benoît Desjardins**[3,4], **Bruno Heintz**[4], **and Jean-Marc Oury**[4]

[1]École des hautes études en sciences sociales and CNRS, CAMS, Paris, France
[2]Senior visiting fellow, Institute for Advanced Study, Hong Kong University of Science and Technology
[3]Ecole normale supérieure Paris Saclay, France
[4]Geobiomics, Paris, France

**Covid-19 | plateau | rebound | shoulder | SIR system | reaction-diffusion system | wastewater-based epidemiology | digital PCR | behavior heterogeneity and variability | mesoscopic model | nonlinear dynamics | public health**

**Correspondence:** ***hb@ehess.fr***

**Plateaus and rebounds of various epidemiological indicators are widely reported in Covid-19 pandemics studies but have not been explained so far. Here, we address this problem and explain the appearance of these patterns. We start with an empirical study of an original dataset obtained from highly precise measurements of SARS-Cov-2 concentration in wastewater over nine months in several treatment plants around the Thau lagoon in France. Among various features, we observe that the concentration displays plateaus at different dates in various locations but at the same level. In order to understand these facts, we introduce a new mathematical model that takes into account the heterogeneity and the natural variability of individual behaviours. Our model shows that the distribution of risky behaviours appears as the key ingredient for understanding the observed temporal patterns of epidemics.**

## 1 Introduction

The onset of plateaus for various indicators of the current outbreak of Covid-19 such as incidence rate or hospitalisations appears to be a rather general feature of its dynamics, along with periods of exponential growth or decay, rebounds etc. Nonetheless, there are few theoretical explanations offered to understand this phenomenon and such plateaus hardly agree with the classical SIR paradigm of epidemics.

We show here that plateaus emerge *intrinsically* in the unfolding of an epidemic. That is, plateaus arise naturally if we take into account two elements: an underlying heterogeneity and a random variability of behaviours in the population. These features are of course more realistic than assuming that the population is perfectly homogeneous with an unwavering behaviour.

To shed light on this mechanism, we propose in this paper a new mathematical model. It takes the form of a system of reaction-diffusion equations, where one variable represents the behaviour of individuals (see Methods section). It is natural to consider that individuals may change their behaviour from one day to the next one. We assume here that individuals' behaviours move randomly according to Brownian motion among these classes. We show here that such a system that includes heterogeneity and variability of behaviours exhibits a richness of dynamics and in particular gives rise to intrinsic formation of plateaus, shoulders and rebounds.

Up to now, there are two main alternative explanations for the onset of plateaus. The first one is political (1). By managing the epidemic and keeping the exponential growth at bay, without destroying the economy, a plateau appears as some kind of optimal compromise under constraint. Another approach appears in a very recent work of Weitz et al. (2). It argues that plateaus are caused by change of behaviours due to awareness of fatalities and fatigue of the public facing regulatory mobility restrictions. In recent works, Arthur et al. (3) and Radicchi et al. (4) proposed models in the same spirit. More detailed discussion of related literature is provided below.

We then introduce and discuss the model, and illustrate the types of dynamics that we obtain in this model by numerical simulations. These simulations shed light on the key role of behavioural variability to obtain plateaus, shoulders and rebounds. There, we also describe the effect of introducing a second variant of the virus with a higher secondary epidemic peak .

To discuss the validity of our approach, we rely on observations stemming from a series of measurements, carried weekly over an extended period in the Thau lagoon area in South of France. These strikingly reveal the formation of plateaus, in some cases after a "shoulder" pattern.

Next, we report on the calibration of our model on the data of the Thau lagoon. It yields a remarkable fit. We further discuss our model in more detail in the light of the measurements in the Discussion section below. In the SI, we also show that this model also generates rebounds.

## 2 The Thau lagoon data

The measurement campaign concerned four wastewater treatment plants (WWTP) in the Thau lagoon area in France, serving the cities of Sète, Pradel-Marseillan, Frontignan and Mèze. The measurements were obtained by using digital PCR (5) (dPCR) to estimate the concentration of SARS-Cov-2 virus in samples taken weekly from 2020-05-12 to 2021-01-12. We provide further details about the measurement method in the Methods section.

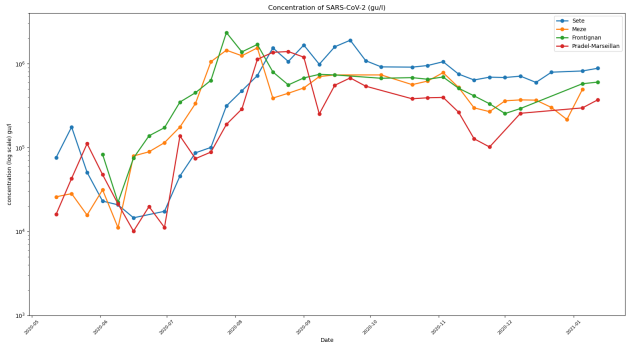


**Fig. 1.** Concentrations of SARS-Cov-2 (genome units per litre in logarithmic scale) from four WWTPs in Thau lagoon, measured weekly with dPCR technology from May $12^{th}$ 2020 to January $12^{th}$, 2021. The four WWTP serve the cities of Sète (top left graph), Mèze (top right), Frontignan (bottom left) and Pradel-Marseillan (bottom right). Note that there are some missing points.

Figure 1 shows the outcomes in a logarithmic scale over a nine months period. We summarise now their main features.

1. An exponential phase starts simultaneously in Mèze and Frontignan WWTPs in early June.

2. The genome units concentration curves in these two places reach, again simultaneously, a plateau. It has stayed essentially stable or slightly decreasing since then.

3. The evolution at Sète and Pradel-Marseillan remarkably followed the previous two zones in a parallel way, with a two weeks lag. The measurements at Sète and Pradel-Marseillan continued to grow linearly (recall that this is in log scale, thus exponentially in linear scale), while the Mèze and Frontignan figures have stabilised ; then, after two weeks, they too stabilised at a plateau with roughly the same value as for the other two towns.

4. The measurements seem to show a tendency to increase over the very last period.

## 3 Brief review of literature

Numerous papers (6–9) describe the number of daily social contacts as a key variable in the spread of infectious diseases like Covid-19 insofar as it is closely related to the transmission rate. Daily social contacts are usually described in terms of age, gender, income, type of job, household size (10)... Parameters that are particularly relevant in the context of viral outbreaks are also studied (11) such as cumulative duration of such contacts, social distance, indoor/outdoor environment... The very fine grain microscopic models (12) aim at identifying such parameters in the most precise way possible. A very recent study by Di Domenico et al. (13) takes up the data of hospitalisations in France for the past six months. Again, these exhibit striking epidemic plateaus since the beginning of 2021. The authors of this paper provide a microscopic insight of the propagation, emphasising the role of two different strains of the virus and the role of public health measures such as the curfew, school closing etc. These approaches are different from ours, as the point of view we adopt here can be seen as "mesoscopic".

Several earlier works have considered SIR type systems with heterogeneity. In particular, Arino et al. (14), and, more recently, Dolbeault and Turinici (15, 16), Magal et al. (17) have studied models with a finite number of different coefficients $\beta$. These systems are characterised by a discrete set of classes and do not involve variability. Almeida et al. (18) considered the case of continuous classes associated with a multidimensional trait $x$ to study the influence of variability of infectious individuals on the final size of an epidemic.

Weitz et al. (2) recently developed an *SEIR*-type compartmental model with variable transmission rate coefficient, in which two main competing psychological reactions to the epidemic generate plateaus. Namely, it involves awareness of fatalities and fatigue of the public facing mobility restrictions. In this interesting model, fatigue modulates the transmission rate $\beta$ in the *SEIR* system by the number of cumulative deaths (and in another version the daily number of deaths). Note that this value itself is an outcome of the model and thus this model is a kind of fixed point formulation. The authors show that their model yields plateaus, "shoulder" like patterns and oscillations for the dynamics of infectious individuals.

This paper of Weitz et al. is quite different from our model in that it assumes a homogeneous population: at a given time, all individuals have the same transmission rate. The common behaviour simply changes in time by reacting to the outcome of the epidemic and this change reflects in the evolution of the transmission rate. Moreover, we note that the model is calibrated with reported death of various U.S. states and does not involve wastewater concentration measurements. Thus, even though in a different manner from ours, this work also stresses the role of variability for the observed dynamics.

DiMarco et al. (19) have recently proposed a model close in spirit to ours but more complex. Based on kinetic theory, it describes the heterogeneity of individuals in terms of a variable $x \geq 0$ corresponding to the number of daily social contacts. Their model consists of a system of three *SIR* equations coupled with Boltzmann or Fokker-Planck type equations. The authors emphasise a collision type term in the resulting Boltzmann equations. This term represents changes of behaviours (that is of the values of $x$) when two individuals meet. Consequently, it ends up as quite different from the present work. As a closure of their model, DiMarco et al. (19) formally derive a so called *S-SIR* model. There, the variability of behaviours involves an explicit dependence of the transmission rates upon the current number of infectious individuals (20). Because of this feature, as a matter of fact, when applied to real data, it is eventually very similar to the approach in Weitz et al. (2)

## 4 The epidemiology model with heterogeneity and natural variability of population behaviour

The appearance of such plateaus and shoulders need not be explained neither by psychological reactions nor by public health policy effects. Indeed, the regulations were roughly constant during the measurement campaign and awareness or fatigue effects do not seem to have altered the dynamics over this long period of time. Witness to this is the fact that two groups of towns saw the same evolution, but two weeks apart one from the other. To understand this phenomena we propose a new model.

Given the complexity and multiplicity of behavioural factors favouring the spread of the epidemic, we assume that the transmission rate involves a normalised variable $a \in (0,1)$ that defines an aggregated indicator of risky behaviour within the susceptible population. Thus, we represent the *heterogeneity* of individual behaviours with this variable. We take $a$ as an implicit parameter that we do not seek to calculate. The classical $SIR$ model is macroscopic and the type of model we discuss here can be viewed as intermediate between macroscopic and microscopic.

The initial distribution of susceptible individuals $S_0(a)$ in the framework of a *SIR*-type compartmental description of the epidemic can be reasonably taken as a decreasing function of $a$. We take the infection transmission rate $a \mapsto \beta(a)$ to be an increasing function of $a$. In the SI Appendix, the reader will find a more general version of this model involving a probability kernel of transition from one state to another. The model here can be derived as a limiting case of that more general version.

Likewise, the behaviour of individuals usually changes from one day to another (21). Many factors are at work in this *variability*: social imitation, public health campaigns, opportunities, outings, the normal variations of activity (e.g. work from home certain days and use of public transportation and work in office on others) etc. Therefore, the second key feature of our model is *variability* of such behaviours: variations of the population density for a given $a$ do not only come from individuals becoming infected and leaving that compartment but also results from individuals moving from one stratus $a$ to another (21). In the simplest version of the model, variability is introduced as a diffusion term in the dynamics of susceptible individuals.

**The model.** We denote by $S(t,a)$ the density of individuals at time $t$ associated with risk parameter $a$, by $I(t)$ the total number of infected, and by $R(t)$ the number of removed individuals. We are then led to the following system:

$$\begin{cases} \dfrac{\partial S(t,a)}{\partial t} = d\,\dfrac{\partial^2 S(t,a)}{\partial a^2} - \beta(a)S(t,a)\dfrac{I(t)}{N}, \\ \dfrac{\mathrm{d}I(t)}{\mathrm{d}t} = \dfrac{I(t)}{N}\displaystyle\int_0^1 \beta(a)S(t,a)\,da - \gamma I(t), \\ \dfrac{\mathrm{d}R(t)}{\mathrm{d}t} = \gamma I(t), \end{cases} \tag{1}$$

where $\gamma^{-1}$ denotes the inverse of typical duration (in days) of the disease and $d$ a positive diffusion coefficient. System (1) is supplemented with initial conditions

$$S(0,a) = S_0(a), \quad I(0) = I_0, \quad \text{and} \quad R(0) = 0, \tag{2}$$

and with zero flux condition in $a$ at $a = 0,1$. In the SI Appendix, we describe more general versions. We also discuss the relation of this models with other current works in the Method section.

**A more general model.** In a more general version of our model, we can consider the population of infected as also structured by the parameter $a$. The equations are as before but now we keep track of the class $a$ in the infected population. The mechanism here is that a susceptible individual from class $a$ can be infected by infectious from any class $I(t,b)$ but then gives rise to an individual $I(t,a)$ of the same parent class. We also assume that there is a diffusion of the infected behaviours. We denote by $\mathfrak{B}(a,b)$ the transmission rate of $S(t,a)$ by $I(t,b)$. For simplicity and because it is natural, we will assume that it is of the form

$$\mathfrak{B}(a,b) = \beta(a)\beta(b)$$

where $\beta$ is as before. For full generality, we can also envision multi-dimensional parameters $a \in \mathbb{R}^d$, with $a_i \in (0,1)$. We are then led to the system:

$$\begin{cases} \dfrac{\partial S(t,a)}{\partial t} = d\,\Delta_a S(t,a) - S(t,a)\dfrac{\beta(a)}{N}\displaystyle\int_0^1 \beta(b)I(t,b)\,db, \\ \dfrac{\partial I(t,a)}{\partial t} = d'\Delta_a I(t,a) + S(t,a)\dfrac{\beta(a)}{N}\displaystyle\int_0^1 \beta(b)I(t,b)db \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad - \gamma I(t,a), \\ \dfrac{\mathrm{d}R(t)}{\mathrm{d}t} = \gamma\displaystyle\int_0^1 I(t,b)\,db, \end{cases} \tag{3}$$

In the SI we write further, more general, forms of this model, with $\mathfrak{B}(a,b)$ and more general diffusion of behaviours, that can include jumps or non-local variations. The type of models we discuss here may also shed light on the initial phase of the epidemic. We plan to investigate these questions in future work.

## 5 Patterns generated by the model

In the next section, we will discuss how the model fits the data observed in the Thau lagoon measurements. But before that, we start by showing that the above model Eq. (1) can generate the different patterns we mentioned. For this we rely on numerical simulations. And indeed we obtain plateaus, shoulders, and oscillations. The latter can be interpreted as epidemic rebounds.

The key parameter here is the diffusion coefficient $d$, which controls the amplitude of behavioural variability (see Figure 2). Large values of $d$ rapidly yield homogenised behaviours, leading to classical SIR-like dynamics of infectious individuals. For very small values of $d$, the system also has

a simple dynamics, in the sense that $I(t)$ has a unique maximum, and therefore has no rebounds. We derive this in the limit $d = 0$ (see SI) for which we show that there are neither plateaus nor rebounds.

For some intermediate range of the parameter $d$, plateaus may appear after an exponential growth, like in the initial phase of the SIR model. A small amplitude oscillation, called "shoulder", precedes a temporary stabilisation on a plateau, followed by a large time convergence to zero of infectious population. We also show that for small enough $d$, time oscillations of the infectious population curve, i.e. epidemic rebounds, may be generated by Model Eq. (1). Such oscillations also appear after a plateau, in a similar way to what one can see in observations.

Simulations in Figure 2 illustrate the various patterns obtained on the dynamics of infected population as a function of the diffusion parameter. For small enough $d$, in the top left graph of Figure 2, one can see oscillations of the fraction of infectious individuals. These oscillations cannot be achieved in the classical SIR model. In fact, the two lower graphs of that figure, for somewhat larger values of $d$, exhibit the SIR model outcomes. Indeed, for sufficiently large $d$, the system becomes rapidly homogeneous (i.e. constant with respect to $a$), as we can see in the corresponding graph in the SI. Yet, such oscillations are standard in the dynamics of actual epidemics, like the current Covid-19 pandemic. The intermediate value of $d$, represented in the upper right corner of Figure 2 shows the typical onset of a plateau at a rather high value of $I$. Note that this plateau is preceded by a first small dip and then a characteristic "shoulder-like" oscillation.

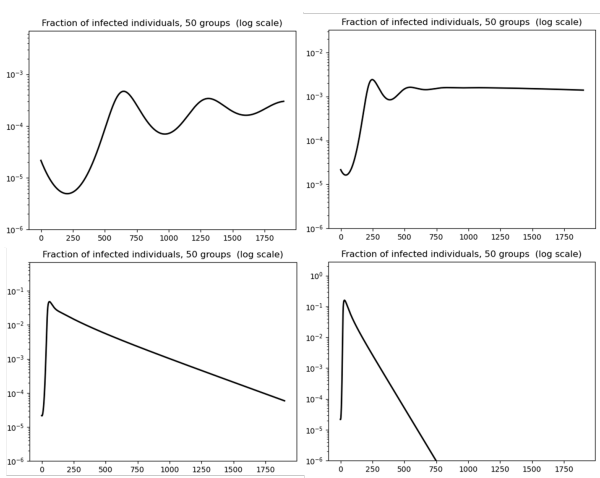


**Fig. 2.** Model behaviour depending on diffusion parameter values: infected rate dynamics in logarithmic scale. From left to right and then top to bottom, graphs are associated with $d = 10^{-5}$, $d = 5 \cdot 10^{-5}$, $d = 10^{-3}$ and $d = 5 \cdot 10^{-3}$ (in $day^{-1}$ unit).

Secondary epidemic peaks are of lower amplitude than the first one, as shown in the top graphs of Figure 2. This empirical observation leads us to conjecture that, at least in many cases, it is a general property of this model (with $\beta$ independent of time). This property would then reflect a kind of dissipative nature of Model (1). It is natural to surmise that a change of behaviours in time may generate oscillations with higher secondary peaks. Such changes result for instance from lifting social distancing measures or from fatigue effects in the population.

We illustrate this with numerical simulations in Figure 3. We assume a collective time modulation of the $\beta(a)$ transmission profile. That is, we replace $\beta(a)$ by $\beta(a)\varphi(t)$ for some time dependent function $\varphi$, the other parameters are the same as in the simulations shown in Figure 2. We look at the effect of a "lockdown exit" type effect. Then, $\varphi(t)$ takes two constant values, 1 from $t = 0$ to $t = 1000$ and 1.2 after $t = 1100$. In between, that is, for $t \in (1000, 1100)$, $\varphi(t)$ changes linearly from the value 1 to 1.2.

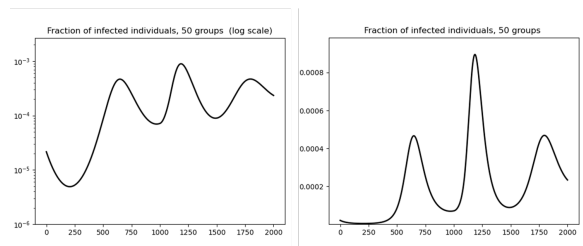


**Fig. 3.** Multiple epidemic rebounds: susceptible individuals is divided into 50 discrete groups in the case where relaxation of social distancing measures starts on Day $t = 1000$ and end up on Day $t = 1100$. The fraction of infected individuals in the population is represented in the left graph in logarithmic scale and in linear scale in the right graph.

One can clearly see a secondary peak with higher amplitude than the first one. Note that after this peak, a third one occurs, with a lower amplitude than the second one. This third peak happens in the regime when $\beta$ is again constant in time.

**The effect of variants.** Another important factor that yields secondary peaks with higher amplitudes is the appearance of variants. Consider the situation with two variants. We denote by $I_1(t)$ and $I_2(t)$ the corresponding infected individuals. The first variant, which we call the historical strain, is associated with $\beta_1$ and $I_1(0)$ and starts at $t = 0$. The variant strain corresponds to $\beta_2$ and $I_2$ and starts at Day $t = 1000$. In this situation, Eq. (1) is extended by the following system:

$$\begin{cases} \dfrac{\partial S(t,a)}{\partial t} = d\,\dfrac{\partial^2 S(t,a)}{\partial a^2} - (\beta_1(a)I_1(t) + \beta_2(a)I_2(t))\,\dfrac{S(t,a)}{N}, \\ \dfrac{\mathrm{d}I_1(t)}{\mathrm{d}t} = \dfrac{I_1(t)}{N}\displaystyle\int_0^1 \beta_1(a)S(t,a)\,da - \gamma_1 I_1(t), \\ \dfrac{\mathrm{d}I_2(t)}{\mathrm{d}t} = \dfrac{I_2(t)}{N}\displaystyle\int_0^1 \beta_2(a)S(t,a)\,da - \gamma_2 I_2(t), \\ \dfrac{\mathrm{d}R(t)}{\mathrm{d}t} = \gamma_1 I_2(t) + \gamma_1 I_2(t), \end{cases} \tag{4}$$

The total infected population is $I(t) = I_1(t) + I_2(t)$. Figure 4 shows a simulation of this system. Before the onset of the second variant, i.e. for $t < 1000$, we observe a peak, followed by a small shoulder and a downward tilted plateau. The second variant corresponds to a higher transmission coefficient: namely, we take here $\beta_2(a) \equiv \frac{3}{2}\beta_1(a)$. When it appears at time $t = 1000$, initially there is no effect, because the initial number of infectious with variant 2 is very small. Then, there is an exponential growth caused by this second variant gaining strength. The secondary peak is then higher

than the first one. A very small shoulder precedes another stabilisation on a downward plateau.

Figure 4 also shows the dynamics of fractions of infected with each one of the variants. Note that the infectious with variant 1 very rapidly all but disappear at the onset of the second exponential growth phase. One might have expected that the historical strain would be gradually replaced by the new strain, merely tilting further downward the plateau. But that does not happen. Thus, it is remarkable that the historical strain gets nearly wiped out at the very beginning of the second exponential growth.

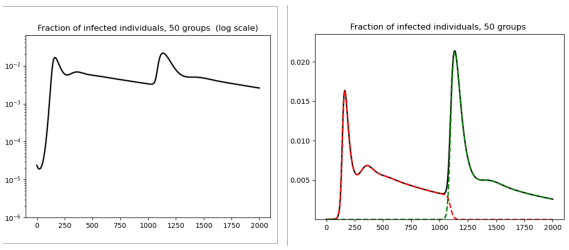


**Fig. 4.** Multiple epidemic rebounds due to a variant virus: susceptible individuals is divided into $50$ discrete groups in the case where a new variant appears at Day $t = 1000$. The transmission rate $\beta_2$ is taken as $\beta_2(a) = 1.5\,\beta_1(a)$, $d = 0.0002$, $\gamma_1 = 0.1$ and $\gamma_2 = 0.05$. The fraction of infected individuals in the population is represented in the left graph in logarithmic scale. The total infected population is represented in linear scale in the right graph (black curve), variant 1 in red and variant 2 in green.

## 6 Application to the Thau lagoon measurements

Model (1) describes the dynamics of infected fraction of population $t \mapsto I(t)/N$. In order to represent the dynamics of concentration of SARS-Cov-2 as measured in WWTPs, the effective proportionality coefficient $\lambda$ minimising the least square distance between measured concentrations and $\lambda$ multiplied by the rate of infected individuals is computed in the calibration process over the time interval from 9 June 2020 to 12 January 2021. Calibration of model (1) also requires fitting the values of $\gamma$, the profiles $a \mapsto \beta(a)$ and the initial distribution of susceptible individuals in terms of $a$.

We carried this procedure and the resulting fitted curve is displayed in Figure 5. Note that the outcome correctly captures the shoulder and plateau patterns.

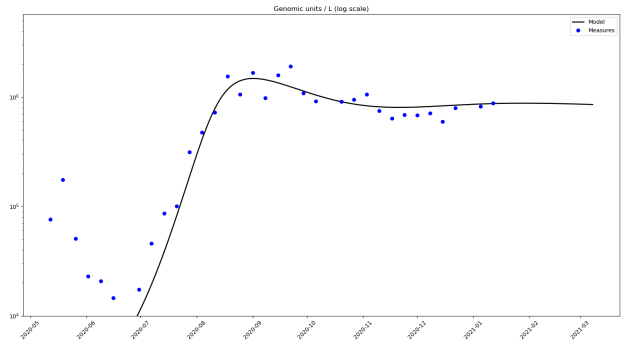


**Fig. 5.** Calibrated model on Sète area: blue dots are measures of SARS-CoV-2 genome units and black curve represents the total infected individuals as an output of the model discretized into $n_g = 20$ groups in $a$. Initial distribution of susceptible individuals and $\beta$ function are taken as described in supplementary information. Parameters $d$ and $\gamma$ are taken as follows: $d = 2.5 \cdot 10^{-4}\ day^{-1}$, and $\gamma = 0.1\ day^{-1}$.

The underlying dynamics of the rate of susceptible individuals is given in Figure 6 below for $n_g = 20$ groups. The lower curve illustrates the competition phenomenon between diffusion and sink term due to new infections, depending on the level of risk $a$ of each stratus.

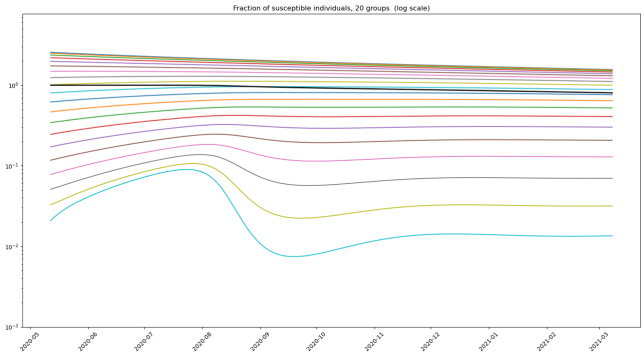


**Fig. 6.** Calibrated model on Sète WWTP: density of susceptible individuals of each group for $n_g = 20$. The densities of susceptible of each group is represented in colour curves as functions of time. The curves are ordered from top to bottom according to increasing $a$. The resulting average total susceptible population is represented in black. Susceptible individuals of highest $a$ trait, which are represented in the bottom light blue curve exhibit a non monotonic behaviour.

## 7 Discussion

We claim that Model (1) explains the formation of plateaus and rebounds in the dynamics of the outbreak through the heterogeneity and variability of population behaviour with respect to epidemiological risk. Figure 6 shows that susceptible individuals with the riskiest behaviour, characterised by the highest $\beta$ transmission coefficient, are rapidly transferred to the infected compartment. Variability of behaviours modelled by diffusion with respect to $a$ parameter then refeeds the fringe of riskiest susceptible individuals. Parameter regimes where those two phenomena have the same order of magnitude generate patterns such as plateaus, shoulders and rebounds. We explored above the types of patterns arising when we vary the diffusion coefficient $d$. We note that plateaus occur for an *intermediate range* of values of $d > 0$, which represents the amplitude of variability. Namely, large

values of $d$ lead to standard SIR-type dynamics, since diffusion quickly homogenises the behaviours. When $d$ decreases, plateaus starting with shoulder-like patterns appear. However, for even smaller values of $d$, oscillations arise, which can be interpreted as epidemic rebounds. From numerical experiments, the amplitude of rebounds always seem to be of smaller amplitude than the first epidemic peak. However, higher secondary peaks arise when we significantly modulate in time the transmission rate. This may represent a progressive exit from lockdown or the effect of new and more contagious variants

The dynamics of the Covid-19 outbreak in the cities of Thau lagoon appears almost insensitive to public health regulations. In particular the second lockdown in France from 28 October to 14 December 2020 had hardly any effect. Likewise, the Christmas Holiday season also seemed to have had little influence on the observed plateaus.

The number of hospitalised individuals in France over the last quarter of 2020 is represented in Figure 7. There, the dynamics shows a growing phase followed by a shoulder and a plateau, very similar to the pattern observed in Thau lagoon.

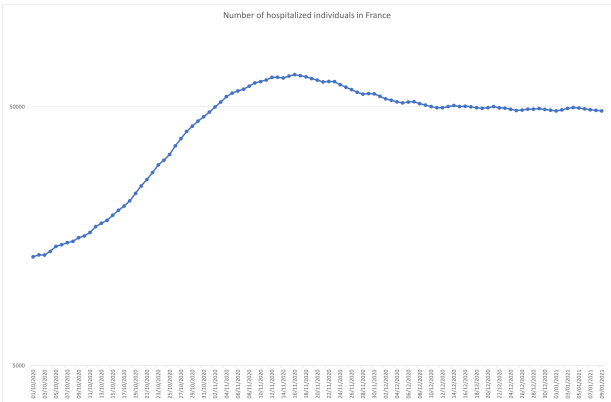


**Fig. 7.** Number of hospitalised individuals in France (log scale, As from October 2020)

Several noteworthy observations came out from the Thau lagoon data. First, we can see two distinct exponential growths in two separate subsets of two towns. The two graphs are parallel with a constant delay of two weeks. The first group reaches a plateau at a certain level of infected and stays thereafter essentially flat, while the second group continues to grow until it reaches about the same value of infected and then becomes essentially flat too. These remarkable observations call for interpretations. Indeed, first, they cannot simply result from public policy measures as these would have affected all these neighbouring towns in a similar fashion. Second, there likely is a spatial diffusion effect that triggered the growth in the second group coming from the first one. However, this diffusion would not explain the fact that the two groups reached the same level of plateaus.

It is worth noting that these observations are inconsistent with the classical *SIR* model: in this model, once an epidemic reaches a peak, it then decreases steadily by another exponential factor. In other words, a plateau requires an effective $\mathcal{R}_t$ number (22, 23) of approximately one: $\mathcal{R}_t \sim 1$. However this would be hard to sustain over such a long period of time as observed because the susceptible population is depleted. It also appears difficult to explain that this occurs exactly at the same plateau level for two distinct populations.

**Conclusions.** In this paper we reported on precise concentration measurements of SARS-CoV-2 in wastewater upstream of four WWTPs around the Thau lagoon in South of France during the nine months period from May 2020 to January 2021. These observations exhibit plateaus and shoulder patterns. Such characteristics, along with rebounds, are widely observed in epidemics. We provided here an explanation by considering that the population is heterogeneous in the level of risk in the individual behaviours, which random changes from day to day. Indeed, we show here that the combination of heterogeneity and variability leads to a constant replenishment of the population of susceptible individuals with higher risks, thus feeding as it were the epidemic. This mechanism explains the formation of plateaus in the dynamics of the epidemic and also accounts for oscillations and rebounds.

To substantiate this claim, we proposed a mathematical model for epidemics that explicitly involves heterogeneity and variability. The model takes the form of an *SIR* system with diffusion of behaviours. To get a model as parsimonious as possible (24), we only assumed the diffusion of risks among susceptible and that a single continuous and one-dimensional risk variable $a$ characterises the behaviour type. Numerical simulations of this system indeed show exactly this type of shape that we observed, with shoulders preceding long plateaus, and rebounds. Without any change in the behaviours, the secondary peaks, although they can be important, are of a lesser amplitude than the first peaks. Here we show further that a one time change in the transmission coefficients may generate rebounds with higher amplitudes. Such changes typically occur when social distancing measures are lifted.

We also explored the effect of introducing a variant strain of the virus, with a slightly higher transmission rate. We also achieved higher amplitude rebounds in this framework. The analysis of the presence of the two strains shows that the historical strain is replaced quite abruptly by the more transmissible one, earlier than one would guess from its intrinsic evolution.

To further validate this model, we fitted it on the data collected in the Thau lagoon area. Here we did not invoke any external effects such as fatigue or change of public health measures. Indeed, over the concerned period, there were arguably few and minor changes. Furthermore, the two weeks delay between the two curves we observed precludes such effects. Indeed, two towns reached the plateau two weeks earlier than the other two, the latter continued to see an increase, until roughly the same level was reached two weeks later. The calibration of our model on the Thau lagoon observations yielded a very good agreement with the data.

**Perspectives and extensions of the model.** There are several directions to extend our work. First, one can con-

sider more general forms of the model as the one introduced in Eq. (3). In the SI, we also indicate further general forms.

Then, one might envision a more precise approach to the parameter $a$ beyond the notion of unit of contact. One can take into account e.g., duration and circumstances of contacts. It appears natural to use "risky sociability" in computing the transmission rate. One may also consider multi-dimensional versions of the parameter. Understanding quantitatively the change of behaviour as a function of this variable $a$ may require an interdisciplinary avenue of research.

Likewise, it is quite natural to assume that the variability of behaviour also depends on $a$ rather than being uniformly distributed. We discuss such an extension in the SI where we show that it leads to an equation with drift terms. We leave developments in this direction to further work.

Another important aspect that transpires in the Thau lagoon data, is the spatial spreading that takes place in the epidemic. In a recent paper, the first author of this study and collaborators (25) have proposed a model at the country level with a quantitative approach to spatial diffusion in France. The study of diffusion at a smaller scale that we might call mesoscopic and its inclusion in the framework we propose here are promising perspectives.

# Methods

**The epidemiological model with heterogeneity and variability.** The model we develop here extends the classical *SIR* compartmental approach recalled in SI appendix by taking into account heterogeneity and variability of behaviours on the susceptible population. The total population is assumed to be constant equal to $N$. That is, we do not take into account incoming or outgoing populations, nor demographic changes. This is a standard assumption in the Covid-19 studies. Actually, one might want to dispense with it if one is to consider a significant amount of travellers, especially during vacation periods or because there is a lockdown that brings many people to leave large cities.

At time $t$, the population of susceptible is structured by $a$ and $t$ and is described by its density $S(t,a)$. We assume that the total number of infected is given by $I(t)$ and that of the removed by $R(t)$. We do not distinguish from which population strata the infected individuals come from (see SI for extensions).

One can think of $a$ as a trait parameter roughly describing the level of risk a given susceptible individual is taking with respect to Covid-19. It is normalised so that $0 \leq a \leq 1$, $a = 0$ being associated with very cautious people and $a = 1$ to people with very risky behaviour. This implicit variable represents for instance the number of social contacts per day of an individual, taking into account their length, whether social distancing is observed, if it takes place indoor or outdoor, in a more or less crowded environment, or if individuals wear a mask etc. Thus, it can be seen as a lumped variable that represents a global risk score. In this context, the impact of political restrictions can easily be represented by a modification of the distribution law of the behaviours.

We observe that in this model, we adopt the point of view that the behaviour distribution affects the *risk takers* in the susceptible population rather than the behaviour of infected individuals. Indeed, it appears more natural to consider that the susceptible face an ambient distribution $I(t)$. For instance, the choice to go to a crowded bar where there *might be* a superspreader, is reflected in variable $a$. Nonetheless, we can also write a similar model where the risky behaviour structures the infected population (see SI). Individuals often vary their behaviour, because for instance of fatigue effects for people who have heeded too strongly social distancing calls or, on the contrary, for people who have been reckless and see other people fall sick and consequently become somewhat more cautious. Thus, the potential reservoir of individuals for a given stratus level $a$ is not static and it is more important than would appear at first glance.

Hence, we consider that the variations of $S(t,a)$ do not only come from individuals becoming infected and leaving that compartment but also results from individuals moving from one stratus to another. Here we assume that this shuffling of behaviours follows Brownian motion. We are then led to System (1) presented in the Result section above. At the endpoints of the interval $(0,1)$ for $a$, we impose homogeneous Neumann (zero flux assumption) boundary conditions:

$$\frac{\partial S(t,0)}{\partial a} = \frac{\partial S(t,1)}{\partial a} = 0. \tag{5}$$

Note that in this system the total population $N$ is conserved by the dynamics:

$$N(t) := \int_0^1 S(t,a)\,da + I(t) + R(t) = N(0).$$

In the SI, we derive this model from more basic considerations. We also describe some mathematical properties of this system. There, we further discuss more general systems. In particular we consider the framework where the variability itself depends on the trait $a$.

It is enlightening to keep track of the fraction of infected coming from specific strata. To this end, we can introduce a variable $I(t,a)$ representing the number of infected that came from stratus $a$. It is given by the solution of the equation:

$$\begin{cases} \dfrac{\partial I(t,a)}{\partial t} = \beta(a) S(t,a) \dfrac{I(t)}{N} - \gamma I(t,a), \\ I(0,a) = I_0(a), \end{cases} \tag{6}$$

It is straightforward that this is consistent with the definition of the total population of infected, that is:

$$I(t) = \int_0^1 I(t,a)\,da.$$

We may include then here the effect that infectious individuals with high $a$'s are more likely to infect other susceptible. This just amounts to replace $I(t)$ in system Eq. (1)–(6) by

$$J(t) := \int_0^1 \beta(a) I(t,a)\,da$$

in (1)-(6). We are then led to Model Eq. (3).

**SARS-Cov-2 concentration measurement from wastewater with digital PCR.** The teams of the local authorities of the Thau lagoon, *Syndicat Mixte du Bassin de Thau* (or SMBT) collected samples every Tuesday from each of the four WWTPs. Each sample consisted of a compound of 24 hourly samples.

I.A.G.E. (INGENIERIE ET ANALYSE EN GENOME EDITING, 2700 route de Mende 34980 Montferrier-sur-Lez, France) developed a diagnostic method to detect very low concentrations of SARS-CoV-2 in such wastewater samples. This method combines an optimised extraction process with a DNA quantification based on a digital PCR (dPCR) targeting region of the RdRp (IP2/IP4) (5) (This method has been submitted by IAGE to the European Patent Office the 31st of December 2020 under the application number EP20306715.2). The measures produced identical results within three significant digits between the two targets.

In contrast to classical quantitative real-time PCR (qRT-PCR), dPCR allows the *absolute* quantification of low concentration levels of target sequences of nucleic acid molecules from DNA or RNA samples. dPCR outperforms qRT-PCR with respect to accuracy (26) and repeatability of measurements; it is also has a much lower detection threshold (about 20 times lower) (5).

Among recently achieved wastewater measurement campaigns in sewers (27–34), it seems that none of them exploit these measurements in a dynamical model. This is probably due to the uncertainties associated with qRT-PCR measurements and to the difficulty to translate genome unit concentrations into numbers of infected individuals. As outlined in Ahmed et al. (27), the rate of infected individuals within the population served by the instrumented WWTP may be related to the measured genome unit concentration through

$$\frac{\textit{Infected persons}}{\textit{Total population}} = \lambda \cdot \frac{\textit{genome units}}{\textit{litre wastewater}} \tag{7}$$

where

$$\lambda = \frac{\dfrac{\textit{litre wastewater}}{\textit{day} \cdot \textit{total population}}}{\left(\dfrac{\textit{g faeces}}{\textit{day} \cdot \textit{person}}\right) \cdot \left(\dfrac{\textit{genome units}}{\textit{g faeces}}\right)} \tag{8}$$

Still, the individual variability of each parameter of Equation (8) is very high, as pointed out in several works (27, 35, 36) (see also references therein). Moreover, transport of wastewater from the emission point to the WWTP involves additional significant phenomena identified in Hart et al. (37): virus degradation over time, usually modelled by exponential decay law where the half life depends on temperature.

The approach developed here is different from these earlier works and attempts to take advantage of an original modelling approach of population heterogeneity in order to estimate the effective parameter $\lambda$ relating the rate of infected people to the measured concentration of SARS-Cov-2. A satisfactory capture of the dynamics of concentrations by the model, including the exponential growth followed by the formation of shoulder-like and plateau pattern leads to such effective parameter estimation, as described in the Result section. In particular, we were led to the choice of $\lambda = 111,230,001$ (in litre per genome unit).

Such virus concentration time series, essentially proportional to the fraction of people infected by Covid-19, provides an accurate quantitative method to monitor the epidemic. Furthermore, as well known (29, 32, 34), the appearance of the virus in wastewater precedes the observations of the disease and therefore yields a remarkable early warning system, ahead of hospital counts. Moreover, it reflects all infectious people regardless of whether they are symptomatic or asymptomatic.

# Bibliography


1. N. Haug et al. Ranking the effectiveness of worldwide COVID-19 government interventions. *Nature human behaviour*, 4:1303–1312, 2020. doi: https://doi.org/10.1038/s41562-020-01009-0.
2. J. Weitz, S.W. Park, C. Eksin, and J. Dushoff. Awareness-driven behavior changes can shift the shape of epidemics away from peaks and toward plateaus, shoulders, and oscillations. *PNAS*, 117(51), 2020. doi: https://doi.org/10.1073/pnas.2009911117.
3. R.F. Arthur et al. Adaptive social contact rates induce complex dynamics during epidemics. *BioRxiv*, 1, 2020. doi: https://doi.org/10.1101/2020.04.14.028407.
4. F. Radicchi and G. Bianconi. Epidemic plateau in critical susceptible-infected-removed dynamics with nontrivial initial conditions. *Physical Review E*, 102, 2020. doi: https://arxiv.org/abs/2102.03836.
5. Institut Pasteur, Paris. Protocol: Real-time rt-pcr assays for the detection of sars-cov-2. Technical Report 1, Institut Pasteur, Paris, 2020.
6. P. Stroud, S. del Valle, S. Sydoriak, J. Riese, and S. Minszewski. Spatial dynamics of pandemic influenza in a massive artificial society. *Journal of Artificial Societies and Social Simulation*, 10(4-9), 2007.
7. Y. Ibuka et al. Social contacts, vaccination decisions and influenza in japan. *J. Epidemiol. Community Health*, 70:152–167, 2016. doi: http://dx.doi.org/10.1136/jech-2015-205777.
8. K. Leung et al. Social contact patterns relevant to the spread of respiratory infectious diseases in Hong Kong. *Nature scientific reports*, 7:7974, 2017. doi: DOI:10.1038/s41598-017-08241-1.
9. O. Diekmann, H. Heesteerbeek, and T. Britton. *Mathematical tools for understanding infectious diseases dynamics*. Princeton University Press, 2013. ISBN 9780691155395.
10. J. Zhang et al. Changes in contact pattern shape the dynamics of the Covid-19 outbreak in china. *Science*, 368:1481–1486, 2020.
11. G. Béraud et al. The French Connection: The First Large Population-Based Contact Survey in France Relevant for the Spread of Infectious Diseases. *PLOS ONE*, 1, 2015. doi: http://dx.doi.org/10.6084/m9.figshare.1466917.
12. E.S. Knock. The 2020 sars-cov-2 epidemic in england: key epidemiological drivers and impact of interventions. *MedRxiv*, 1, 2020. doi: https://doi.org/10.1101/2021.01.11.21249564.
13. L. Di Domenico, C.E. Sabbatini, G. Pullano, D. Lévy-Bruhl, and V. Colizza. Impact of january 2021 curfew measures on sars-cov-2 b.1.1.7 circulation in france. *MedRXiv*, 1, 2021. doi: https://doi.org/10.1101/2021.02.14.21251708.
14. J. Arino, J.R. Davis, D. Hartley, and R. Jordan. A multi-species epidemic model with spatial dynamics. *Mathematical Medicine and Biology*, 22(2):129–142, 2005. doi: https://doi.org/10.1093/imammb/dqi003.
15. J. Dolbeault and G. Turinici. Social heterogeneity and the Covid-19 lockdown in a multi-group SEIR model. *MedRxiv*, 1, 2021. doi: https://doi.org/10.1101/2020.05.15.20103010.
16. J. Dolbeault and G. Turinici. Heterogeneous social interactions and the COVID-19 lockdown outcome in a multi-group SEIR model. *HAL*, 1, 2020. doi: https://hal.archives-ouvertes.fr/hal-02559938v2.
17. P. Magal, O. Seydi, and G. Webb. Final size of a multigroup sir epidemic model: Irreducible and non-irreducible modes of transmission. *Mathematical Biosciences*, 301:59–67, 2018. doi: https://doi.org/10.1016/j.mbs.2018.03.020.
18. L. Almeida, P.A. Bliman, G. Nadin, B. Perthame, and N. Vauchelet. Final size and convergence rate for an epidemic in heterogeneous population. *arXiv*, 1, 2020.
19. G. Dimarco, B. Perthame, G. Toscani, and M. Zanella. Social contacts and the spread of infectious diseases. *arXiv*, 1, 2020.
20. A. Korobeinikov and P.K. Maini. Non-linear incidence and stability of infectious disease models. *Mathematical Medicine and Biology*, 22:113–128, 2005. doi: doi:10.1093/imammb/dqi001.
21. H.F Chan, A. Skali, D.A. Savage, D. Stadelmann, and B. Torgler. Risk attitudes and human mobility during the covid-19 pandemic. *Nature scientific reports*, 10, 2020. doi: https://doi.org/10.1038/s41598-020-76763-2.
22. L Roques, E.K. Klein, J Papaïx, A. Sar, and S. Soubeyrand. Using early data to estimate the actual infection fatality ratio from covid-19 in france. *Biology*, 9(97), 2020. doi: https://doi.org/10.3390/biology9050097.
23. H.R. Thieme. Spectral Bound and Reproduction Number for Infinite-Dimensional Population Structure and Time Heterogeneity. *SIAM Journal on Applied Mathematics*, 70(1):188–211, 2009. doi: https://doi.org/10.1137/080732870.
24. A Bertozzi et al. The challenges of modeling and forecasting the spread of covid-19. *PNAS*, 117(29):16732–16738, 2020. doi: https://doi.org/10.1073/pnas.2006520117.
25. L. Roques, O. Bonnefon, V. Baudrot, S. Soubeyrand, and H. Berestycki. A parsimonious

approach for spatial transmission and heterogeneity in the covid-19 propagation. *Royal Society Open Science*, 12, 2020. doi: https://doi.org/10.1098/rsos.201382.

26. T. Suo et al. ddpcr: a more sensitive and accurate tool for sars-cov-2 detection in low viral load specimens. *MedRxiv*, 1, 2020. doi: https://doi.org/10.1101/2020.02.29.20029439.
27. W. Ahmed et al. First confirmed detection of sars-cov-2 in untreated wastewater in australia: a proof of concept for the wastewater surveillance of covid-19 in the community. *Sci. Total Environment*, 728, 2020. doi: https://doi.org/10.1016/j.scitotenv.2020.138764.
28. E. Haramoto et al. First environmental surveillance for the presence of SARS-CoV-2 RNA in wastewater and river water in Japan. *Sci. Total Environment*, 737, 2020. doi: https://doi.org/10.1016/j.scitotenv.2020.140405.
29. G. Medema et al. Presence of SARS-Coronavirus-2 RNA in sewage and correlation with reported covid-19 prevalence in the early stage of the epidemic in the Netherlands. *Environ. Sci. Technol. Lett.*, 7:511–516, 2020. doi: https://dx.doi.org/10.1021/acs.estlett.0c00357.
30. B. Miyani et al. SARS-Cov-2 in Detroit wastewater. *J. Environ. Eng.*, 146(11), 2020. doi: https://doi.org/10.1061/(ASCE)EE.1943-7870.0001830.
31. S. Wurtzer et al. Evaluation of lockdown impact on SARS-Cov-2 dynamics through viral genome quantification in paris wastewater. *www.eurosurveillance.org*, 1, 2020. doi: https://doi.org/10.1101/2020.04.12.20062679.
32. W. Randazzo et al. SARS-CoV-2 RNA titers in wastewater anticipated covid-19 occurrence in a low prevalence area. *Water Research*, 181, 2020. doi: https://dx.doi.org/10.1016%2Fj.watres.2020.115942.
33. P.M. d'Aoûst et al. Quantitative analysis of SARS-Cov-2 RNA from wastewater solids in communities with low covid-19 incidence and prevalence. *Water Research*, 188, 2021. doi: https://doi.org/10.1101/2020.08.11.20173062.
34. J. Peccia et al. Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. *Nature biotechnology*, 38:1164–1167, 2020. doi: https://doi.org/10.1038/s41587-020-0684-z.
35. P. Foladori et al. Sars-cov-2 from faeces to wastewater treatment: What do we know? a review. *Science of The Total Environment*, 743, 2020. doi: https://doi.org/10.1016/j.scitotenv.2020.140444.
36. R. Wölfel et al. Virological assessment of hospitalized patients with COVID-2019. *Nature*, 581:465–469, 2020. doi: https://doi.org/10.1038/s41586-020-2196-x.
37. O.E. Hart and R.U. Halden. Computational analysis of SARS-CoV-2/COVID-19 surveillance by wastewater-based epidemiology locally and globally: Feasibility, economy, opportunities and challenges. *Science of the Total Environment*, 730, 2020. doi: https://doi.org/10.1016/j.scitotenv.2020.138875.
38. F. Bauer, P. van den Driessche, and J. Wu. Mathematical epidemiology. *Lecture Notes in Math.*, 1945, 2008.
39. Francesco Biscani and Dario Izzo. A parallel global multiobjective framework for optimization: pagmo. *Journal of Open Source Software*, 5(53):2338, 2020. doi: https://doi.org/10.21105/joss.02338.
40. H. Berestycki, J.-M. Roquejoffre, and L. Rossi. The periodic patch model for population dynamics with fractional diffusion. *Discrete and Continuous Dynamical Systems - series S*, 4:1–13, 2011. doi: http://dx.doi.org/10.3934/dcdss.2011.4.1.
41. M. Lorenzo and Y. Picó. Wastewater-based epidemiology: current status and future prospects. *Current Opinion in Environmental Science & Health*, 9:77–84, 2019. doi: https://doi.org/10.1016/j.coesh.2019.05.007.
42. G. Campi et al. Metastable states in plateaus and multi-wave epidemic dynamics of Covid-19 spreading in Italy. *arxiv*, 1, 2021. doi: https://arxiv.org/abs/2102.03836.
43. Reuters. *"Fauci says U.S. coronavirus infections may be plateauing, vaccines should protect against new variants"*. Press release 2021-01-21, January 2021. https://www.reuters.com/article/us-usa-biden-fauci-idUSKBN29Q2XH.
44. SIG France, Press conference by Prime Minister J. Casteix. *"Mais, je vous l'ai dit, cette amélioration marque le pas depuis une semaine. Nous sommes sur une sorte de plateau."* . Press release 2020-12-10, December 2020. https://www.gouvernement.fr/partage/11951-conference-de-presse-du-premier-ministre-point-de-situation-sur-la-lutte-contre-la-covid-19.
45. Westdeutsche Zeitung. *"Mit Hilfe der geltenden einschneidenden Maßnahmen sei es zwar geschafft worden, die Zahlen auf einem gewissen Plateau zu halten."*. Press release 2021-01-15, January 2021. https://www.wz.de/nrw/bund-laender-treffen-zu-schaerferen-corona-massnahmen-fuer-dienstag-angesetzt_aid-55699681.
46. Le Progrès, Interview of E. Macron. *"Aujourd'hui, nous sommes sur un plateau."*. Press release 2021-02-02, February 2021. https://www.leprogres.fr/sante/2021/02/02/coronavirus-la-france-doit-valider-ce-mardi-le-vaccin-astrazeneca.

# Supplementary Information

**The classical *SIR* model with time dependent transmission rate.** The classical *SIR* compartmental model (38) uses a dynamics governed by mass-action laws. It assumes that individuals are homogeneously mixed and that every individual is equally likely to interact with every other individual. It reads:

$$\begin{cases} \dfrac{dS}{dt} = -\dfrac{\beta I S}{N} \\ \dfrac{dI}{dt} = \dfrac{\beta I S}{N} - \gamma I \\ \dfrac{dR}{dt} = \gamma I \end{cases} \tag{9}$$

***Single location growth and behavioural changes..*** Changes of behaviour driven by phenomena such as awareness of fatalities or fatigue with respect to mobility restrictions can be modelled by a time modulation of the infection rate $\beta$.

To start with, we analyse this problem on synthetic data. Namely, we consider an exact dynamics given by the $SIR$ model with parameter changes on two given dates. We now take part of the resulting points as given observations. The goal is to carry an optimisation procedure where we try to identify the dates and magnitudes of these changes i.e. the different values of $\beta$ that come into play. Later on we will consider the Thau lagoon data.

Figure 8 represents a piece-wise constant modulation in time with reduction of social interactions down to $70\%$ starting Day $t = 25$ followed by an increase to $50\%$ on Day $t = 50$.

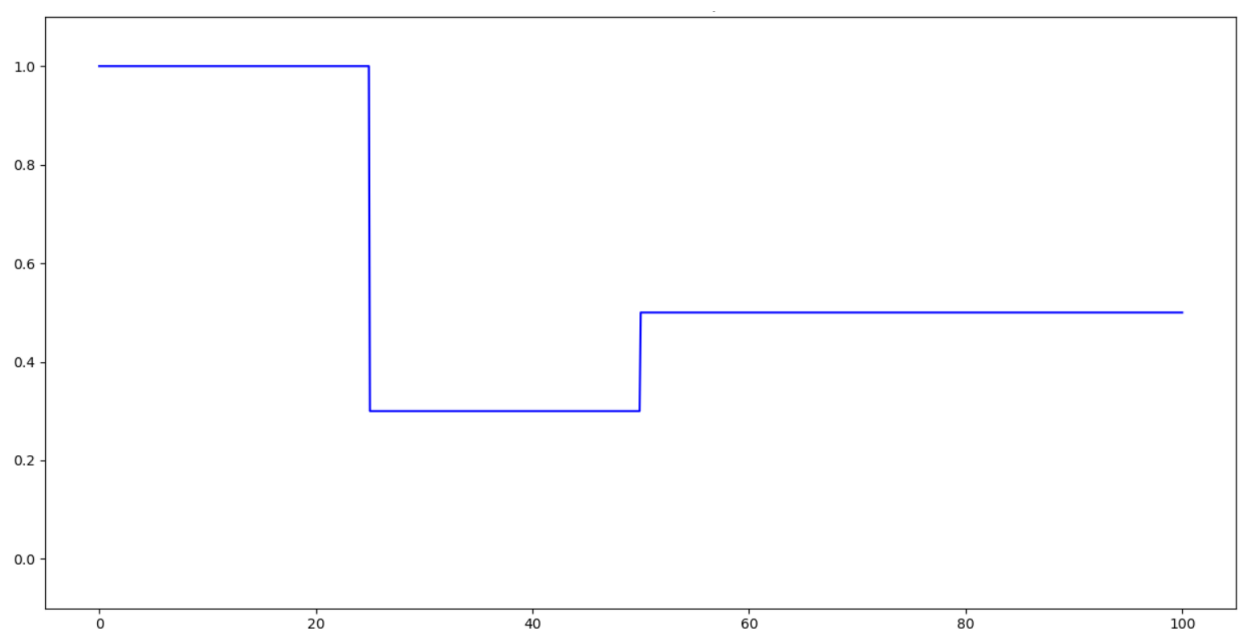


**Fig. 8.** Time modulation function associated with collective behaviour modulation as a function of time (day numbers)

The effect on such modulation leads to the interruption of the growth phase followed by a lower decrease for later times.

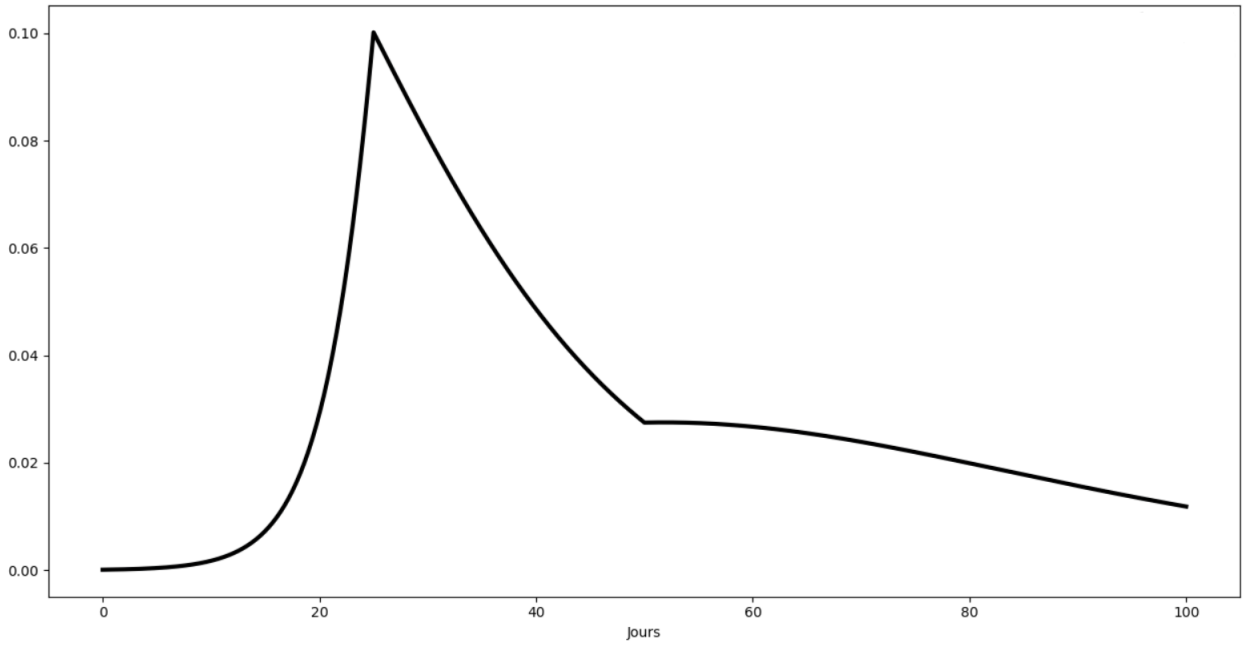


**Fig. 9.** Rate of infected individuals as a function of days

We developed a toolbox based upon PYGMO (39) particle swarm optimisation algorithms in order to solve the type of inverse problem we encounter here. We consider a time sampled version of infection rates in the total population. We allow a given number of transition times associated with behaviour changes, in between which the coefficient $\beta$ is constant. The problem then is to determine in an optimal way the *SIR* model parameters, initial conditions, the transition times and the values of the $\beta$'s in between. Using a least square objective function, we obtained a satisfactory convergence towards the dynamics of the initial model and associated modulation function.

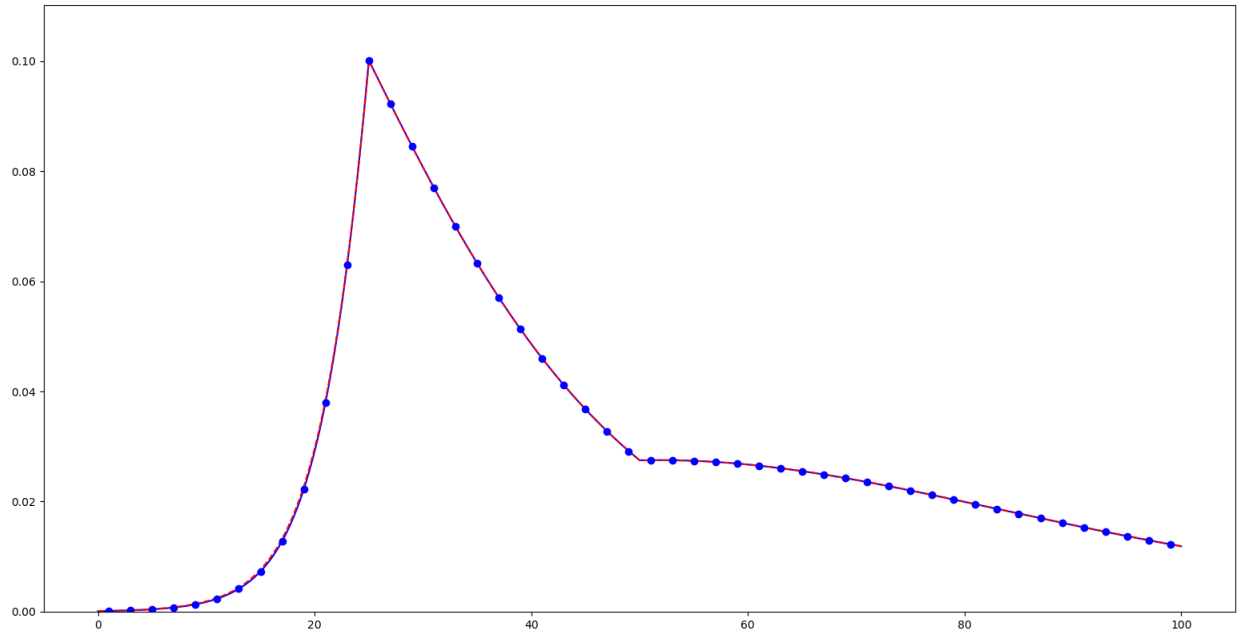


**Fig. 10.** Rate of infected individuals as a function of days as a result of the inverse problem: blue dots correspond to daily sampling of synthetic data, red line to the resulting *SIR* model.

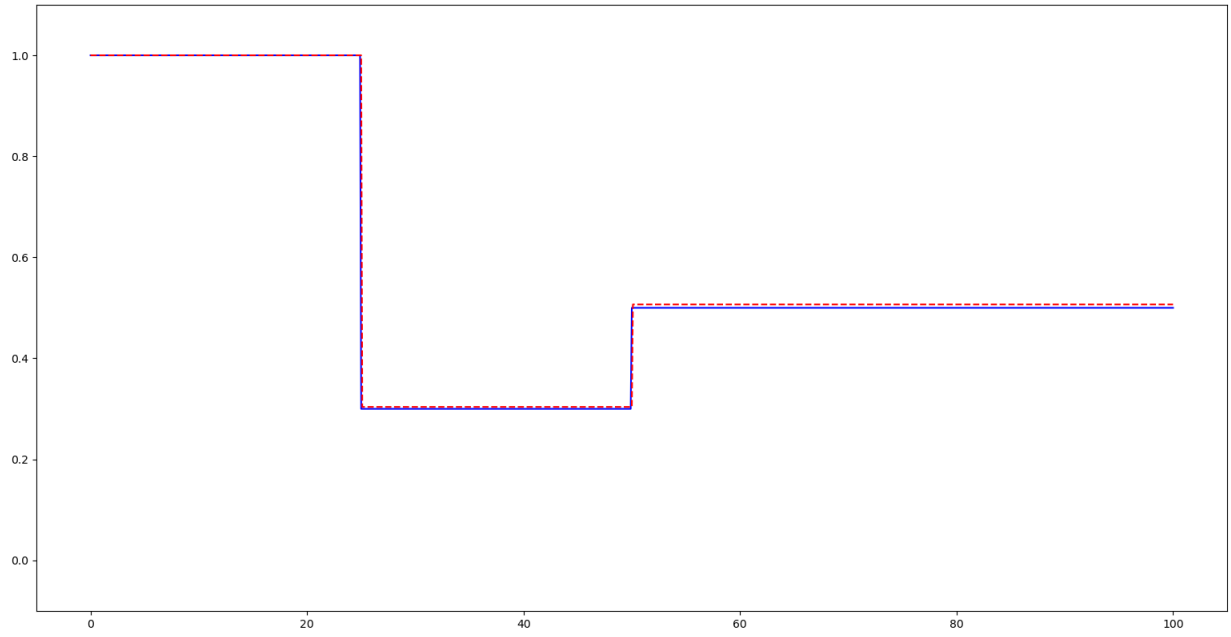


**Fig. 11.** Modulation of collective behaviour as a result of the inverse problem

We developed a similar approach with noise added to sampled synthetic data, with satisfactory convergence both in terms of dynamics of infection rate (Figure 12) and modulation functions (levels and time changes in Figure 13).

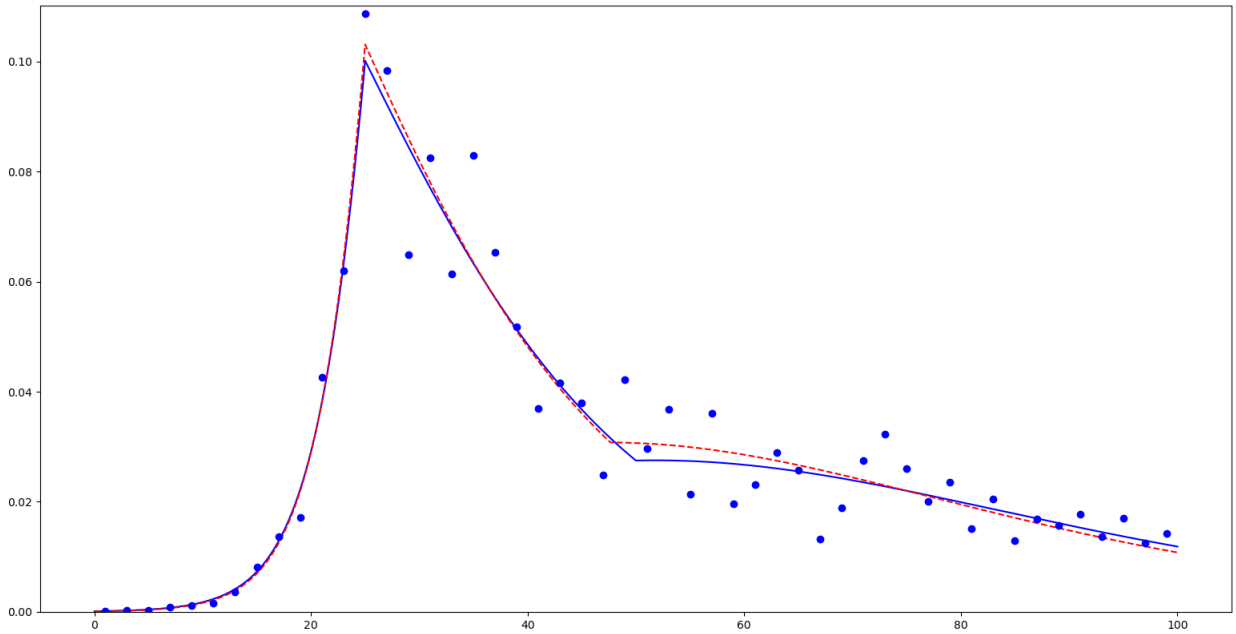


**Fig. 12.** Inverse problem with noisy synthetic data: blue dots correspond to daily sampled synthetic data with noise, blue and red curves respectively to the initial synthetic model and to the result of the inverse problem.

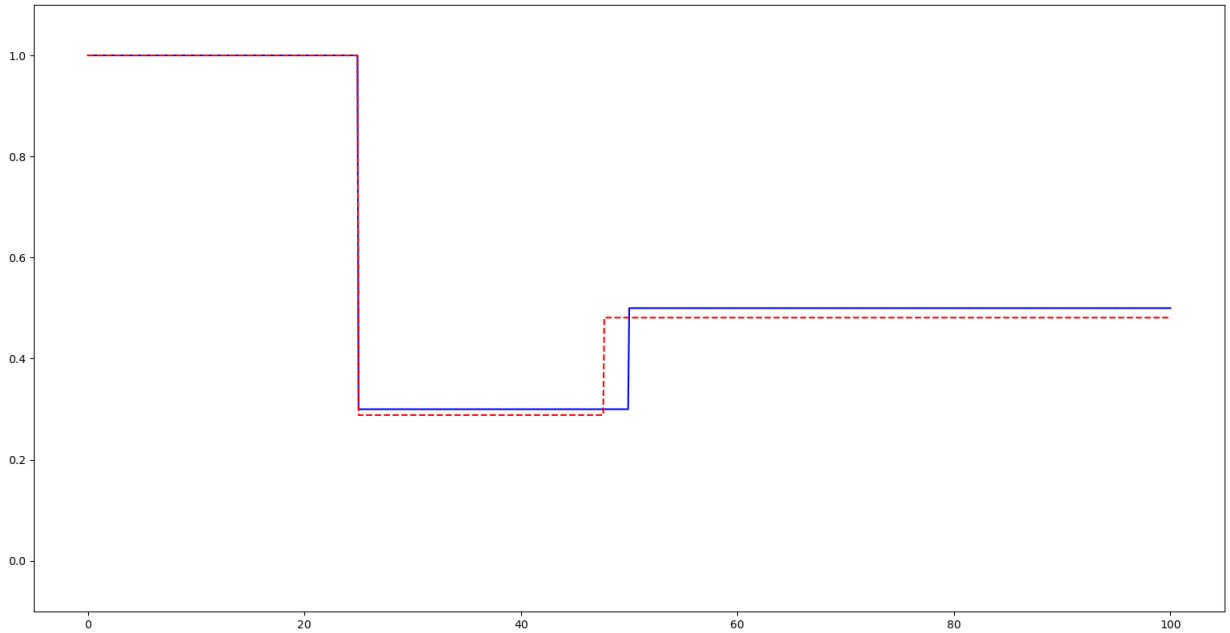


**Fig. 13.** Original modulation function in blue and result of the inverse problem in red.

***A model with change of behaviour at discrete times in Thau lagoon..*** We applied the preceding algorithmic approach on real data of the four WWTP measurement campaigns of the Thau lagoon. In order to keep the model parsimonious, we used the same time modulation function of $\beta$ for all the four cities together. The transition dates and levels have to be determined.

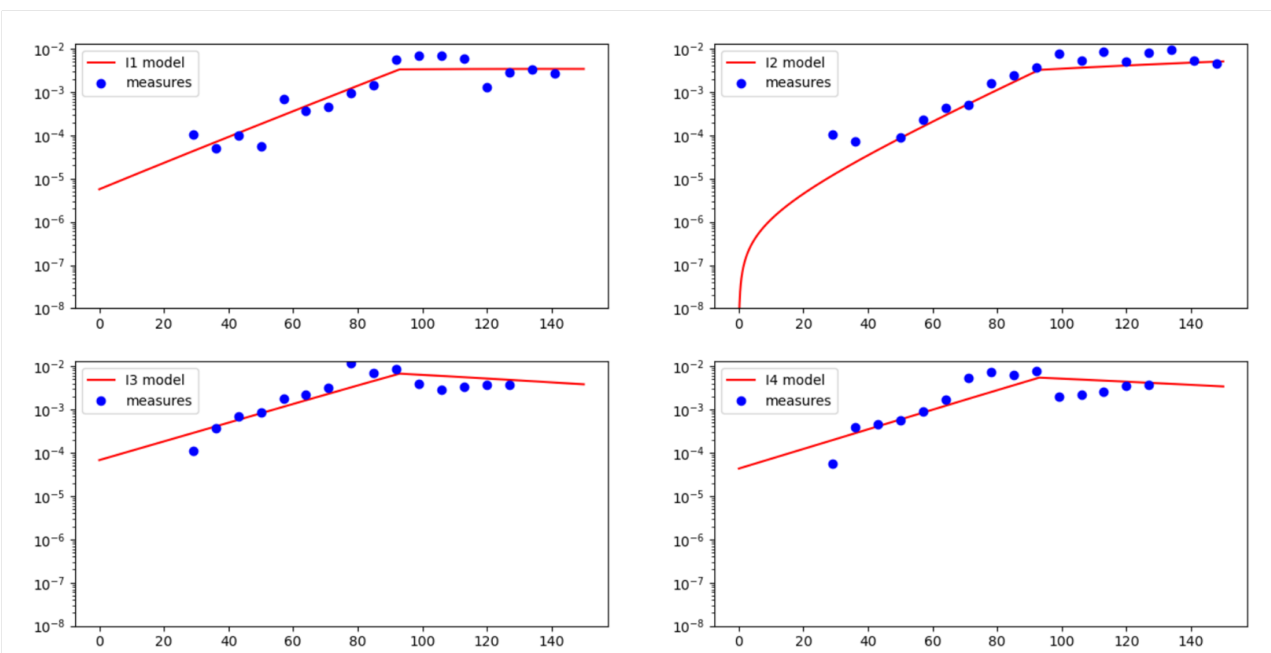


**Fig. 14.** Change of behaviour in logarithmic scale in the Thau lagoon (initial day $t = 0$ being associated with May 12$^{th}$, 2020). The red curve represents the model output (rate of infected population) and the blue dots correspond to measurements. Zones $I_1$, $I_2$, $I_3$ and $I_4$ respectively correspond to Sète, Mèze, Frontignan and Pradel-Marseillan.

The above results show that the optimal capture of the transition from exponential growth to plateau type dynamics occurs on August 10$^{th}$, 2020. The resulting piecewise constant modulation function in time is represented in Figure 15 below.

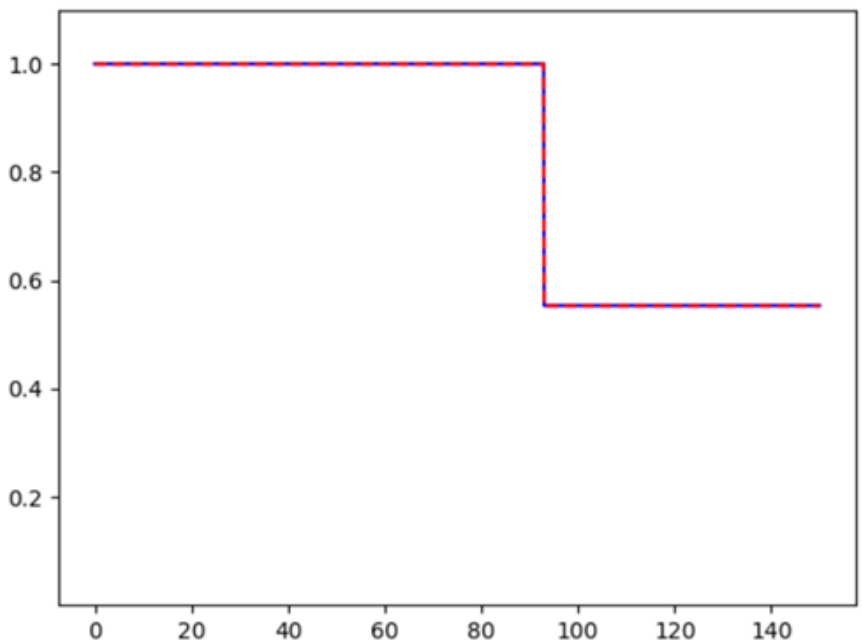


**Fig. 15.** Time modulation of collective behaviour in Thau lagoon as a result of the inverse problem.

Even though the plateau dynamics seems to be correctly represented in Figure 14, the main shortcoming of this is that it does not really explain the observations of the Thau lagoon data. Indeed, the model does not explain why the four zones reach the same infection rate level associated with the plateau regime and why there is a two weeks delay. Furthermore, it does not yield a shoulder effect, as seen in the data.
One may use spatial diffusion to describe the delay between the zones (Frontignan, Mèze) on the one hand and (Sète, Pradel-Marseillan) on the other hand. But it also fails to explain why there are two parallel curves. We plan to come back to spatial diffusion in future work.

**Derivation of the Model.** This section is concerned with the formal derivation of the additional diffusion term in the dynamics of susceptible individuals. As previously emphasised, this diffusion term stems from the combination of the heterogeneity of behaviours and their variability in time. A natural case to consider here is to assume that the shuffling of behaviours happens according to Brownian motion. Namely, individual risk traits move according to the process

$$da_t \;=\; \sqrt{2}\sigma\, dW_t,$$

where $\sigma > 0$ is the possibly $a$-dependent diffusion, and $W_t$ is a Brownian motion in $(0,1)$ with reflection conditions at the end-points of the interval. By the Fokker-Planck equation, in the absence of epidemic, an initial distribution of population $S(0,a)$ gives rise to $S(t,a) = \mathbb{E}[S(0,a_t)]$ governed by

$$\frac{\partial S}{\partial t} \;=\; \frac{\partial^2[\sigma^2(a)\,S]}{\partial a^2}. \tag{10}$$

To consider the dynamics of the epidemic, we are thus led to the following more general version:

$$\begin{cases} \dfrac{\partial S(t,a)}{\partial t} = d\,\dfrac{\partial^2[\sigma^2(a)\,S(t,a)]}{\partial a^2} - \beta(a)S(t,a)\dfrac{I(t)}{N}, \\ \dfrac{\mathrm{d}I(t)}{\mathrm{d}t} \;=\; \dfrac{I(t)}{N}\displaystyle\int_0^1 \beta(a)S(t,a)\,da - \gamma I(t), \\ \dfrac{\mathrm{d}R(t)}{\mathrm{d}t} = \gamma I(t). \end{cases} \tag{11}$$

The first equation involves drift and other additional terms. The numerical and theoretical analysis of the case where $\sigma$ depends on a, and the corresponding extensions of the Fokker-Planck equation with drift terms, seem relevant from a modelling viewpoint.

In the simple case where $\sigma^2(a) = d$ where $d$ is a constant, one ends up with the first equation of our system (see System (12) below).
One could also envision a multi-dimensional variable $a = (a_1, \ldots, a_p)$ to encompass various behavioural characteristics that could possibly be related to sociological observations. In this case, one would be led to a higher dimensional partial differential equation. where the term $\partial^2_{aa} S$ is replaced by the Laplace operator $\Delta_a S$ as we did in the general model presented in Methods section. Such developments would be useful to improve the model and its applicability.

**Simple properties of Model (12).** For the convenience of the reader, we recall the model introduced in the main paper

$$\begin{cases} \dfrac{\partial S(t,a)}{\partial t} = d\dfrac{\partial^2 S(t,a)}{\partial a^2} - \beta(a)S(t,a)\dfrac{I(t)}{N}, \\ \dfrac{\mathrm{d}I(t)}{\mathrm{d}t} = \dfrac{I(t)}{N}\displaystyle\int_0^1 \beta(a)S(t,a)\,da - \gamma I(t), \\ \dfrac{\mathrm{d}R(t)}{\mathrm{d}t} = \gamma I(t), \end{cases} \tag{12}$$

The case $d = 0$ (heterogeneous but without variability) writes:

$$\begin{cases} \dfrac{\partial S(t,a)}{\partial t} = -\dfrac{\beta(a)S(t,a)I(t)}{N}, \\ \dfrac{\mathrm{d}I(t)}{\mathrm{d}t} = \dfrac{I(t)}{N}\displaystyle\int_0^1 \beta(a)S(t,a)\,da - \gamma I(t), \\ \dfrac{\mathrm{d}R(t)}{\mathrm{d}t} = \gamma I(t). \end{cases} \tag{13}$$

We list below some elementary mathematical properties of the model. First, we note that this model contains the traditional *SIR* model when the initial profile of susceptible $a \mapsto S_0(a)$ is uniform in $a$. Indeed, it is straightforward to see that then, $S(t,a)$ also does not depend on $a$.
The dynamics of total infectious individuals is governed by

$$\frac{dI(t)}{dt} = I(t)\left(\frac{1}{N}\int_0^1 \beta(a)S(t,a)\,da - \gamma\right), \tag{14}$$

so that the equivalent of the so called "effective R" coefficient in traditional SIR model, namely $\beta S(t)/(\gamma N)$, is replaced by

$$\frac{1}{N\gamma}\int_0^1 \beta(a)S(t,a)\,da. \tag{15}$$

Unlike traditional *SIR* model, in which the evolution of infectious $t \mapsto I(t)$ begins by an exponentially growing phase, has a unique maximum and tends to zero for large time, Model (12) may exhibit more sophisticated behaviours such as rebounds in $I$ with multiple local maxima as well as plateaus. This property can be intuitively understood by analysing the evolution of the growth factor in (14):

$$\frac{1}{N}\frac{d}{dt}\int_0^1 \beta(a)S(t,a)\,da = -\frac{I(t)}{N^2}\int_0^1 \beta^2(a)S(t,a)\,da - \frac{d}{N}\int_0^1 \beta'(a)\frac{\partial S(t,a)}{\partial a}\,da. \tag{16}$$

In the absence of diffusion, $d = 0$, this growth factor is non increasing in time. For heterogeneous profiles and non zero diffusion, the first term of the right hand side is always negative whereas the second term may be positive if for instance $a \mapsto \beta(a)$ is increasing and $a \mapsto S(t,a)$ is a decreasing function. Below we show that it is always the case under some conditions. Thus, depending on the amplitude of the above two terms, the growth factor may have increasing and decreasing phases. It naturally leads one to consider initial conditions characterised by the property $S_0'(a)\beta'(a) < 0$ such as

$$\beta(a) = \beta_0 \exp\left(\frac{a^2}{\eta^2}\right) \quad \text{for some } \eta > 0, \tag{17}$$

$$S_0(a) = S_1 + (S_0 - S_1)\exp\left(-\frac{a^2}{\kappa^2}\right) \quad \text{for some } \kappa > 0 \text{ and } 0 < S_1 < S_0. \tag{18}$$

It means that individuals are sorted by increasing infection rate, according to the variable $a$, and that the vast majority of individuals have low infectiousness whereas few individual close to $a = 1$ have extremely risky behaviour. To some extent,

social contact surveys (6, 7, 21) justify an initial distribution of susceptible individuals like (18). However, we are not aware of a rigorous justification of the expression (17) of the sharply increasing profile in $\beta$.
Let us now show that the evolution preserves the property of being decreasing in $a$ for the profiles of the susceptible population.

**Proposition 1.** *If $S_0$ is a decreasing function of $a$, then $S(t,\cdot)$ is also decreasing in $a$ for all time $t$.*
To prove this property, we simply note that the derivative $v(t,a) := \partial S(t,a)/\partial a$ of $S$ with respect to $a$ satisfies the following equation

$$\begin{cases} \dfrac{\partial v}{\partial t} + \beta \dfrac{I(t)}{N} v - d \dfrac{\partial^2 v}{\partial a^2} = -\beta'(a) S(t,a) \dfrac{I(t)}{N} \\ v(t,0) = v(t,1) = 0, \\ v(0,a) = \dfrac{\partial S(0,a)}{\partial a}. \end{cases} \tag{19}$$

Since the right hand side of the first equation of (19) is negative and the initial condition $v(0,a) \leq 0$, the parabolic maximum principle then shows that $v(t,a) \leq 0$ for all further time.

**Further mathematical properties of Model (12).** Let us first notice that we can describe the dynamics of susceptible individuals in (12) in terms of its probability density $f(t,a) = S(t,a)/S(t)$ where $S(t) = \int_0^1 S(t,a)\,da$:

$$\frac{\partial f(t,a)}{\partial t} = -\frac{I(t)}{N} f(t,a) \left(\beta(a) - \bar{\beta}(t)\right) + d \frac{\partial^2 f(t,a)}{\partial a^2}, \tag{20}$$

where we denote $\bar{\psi}(t) = \int_0^1 \psi(t,a) f(t,a)\,da$. This equation shows the effect of infections on the distribution of susceptible population as a function of $a$. Indeed, it shows that, for instance in the absence of diffusion ($d = 0$), the proportion of the population of high $a$ goes down as a result of infection affecting it more in relative terms than the remaining of the population. The presence of diffusion ($d > 0$) mitigates this effect by fuelling as it were the epidemic with individuals who had initial lower risk trait.
The equivalent of (16) for the average transmission rate $\bar{\beta}$ is

$$\frac{d\bar{\beta}(t)}{dt} = -\frac{I(t)}{N} Var(\beta)(t) - d \int_0^1 \beta'(a) \frac{\partial f(t,a)}{\partial a}\,da, \quad \text{where} \quad Var(\beta) = \overline{\beta^2} - \bar{\beta}^2. \tag{21}$$

In other words, in the absence of behavioural variability ($d = 0$), the average of the transmission rate decreases under the effect of its variance, the latter being directly linked to behavioural heterogeneity. Thus a higher heterogeneity promotes a faster decay of the average transmission rate. This is a remarkable property of the system in absence of diffusion or with small diffusion.
In the presence of diffusion $d > 0$, this effect is in competition with the second term of the right hand side of (21) which may be positive if for instance the profile $a \mapsto \beta(a)$ is increasing while the distribution $f$ decreases along variable $a$.
Our next result concerns the effect of heterogeneity on herd immunity. We analyze it in the absence of diffusion.

**Theorem 1.** *Let $(t,a) \mapsto (S(t,a), I(t), R(t))$ be the solution of (13) (that is when $d = 0$) with initial conditions $S(0,a) = \tilde{S}_0 f_0(a)$, $I(0) = \tilde{I}_0 > 0$ and $R(0) = 0$ (for some positive constants $\tilde{S}_0$ and $\tilde{I}_0$). Then, $t \mapsto S(t,\cdot)$ tends as $t \to +\infty$ to a limit profile $a \mapsto S_\infty(a)$ in such a way that*

$$\int_0^1 S_\infty(a)\,da \geq \tilde{S}_\infty,$$

*where $\tilde{S}_\infty > 0$ is the limit value as $t$ goes to $+\infty$ of susceptible individuals $\tilde{S}(t)$ associated to the homogeneous SIR model (9) with initial conditions $(\tilde{S}_0, \tilde{I}_0, 0)$ and transmission rate $\bar{\beta}_0 = \int_0^1 \beta(a) f_0(a)\,da$.*
Thus, this theorem asserts that heterogeneity lowers the herd immunity rate needed to stop an epidemic. Note that Several works study the asymptotic states of similar models without diffusion, in a discrete class framework. We mention in particular Magal et al. (17), Dolbeault and Turinici (15) and Almeida et al. (18).
**Sketch of proof:** The existence of the limit profile $a \mapsto S_\infty(a)$ follows from the monotonicity of $t \mapsto S(t,a)$ for fixed $a \in (0,1)$. Since $t \mapsto I(t)$ can be easily proved to vanish for large time, the recovered compartment $R(t)$ also tends to a limit $R_\infty$. Using $R$ as a time variable leads to:

$$I(R) = \tilde{I}_0 - R - \int_0^1 S_0(a) \left( \exp\left( -\frac{\beta(a) R}{N\gamma} \right) - 1 \right) da,$$

so that

$$S_\infty(a) = S_0(a) \exp\left( -\frac{\beta(a) R_\infty}{N\gamma} \right)$$

and

$$\tilde{I}_0 = R_\infty + \int_0^1 S_0(a)\left(\exp\left(-\frac{\beta(a)R_\infty}{N\gamma}\right)-1\right)da. \tag{22}$$

Equation (22) expresses $\tilde{I}_0$ as an increasing function $F_\beta$ of $R_\infty$, which is therefore uniquely defined. Application of Jensen's convexity inequality then leads to

$$G\left(\int_0^1 \beta(a)f_0(a)\,da\right) \le \int_0^1 G(\beta(a))f_0(a)\,da, \quad \text{where} \quad G(Z) = \exp(-ZR_\infty/(\gamma N)). \tag{23}$$

We deduce that $F_\beta(x) \ge F_{\bar{\beta}_0}(x)$ for all $x > 0$, where $F_{\bar{\beta}_0}$ is associated with the homogeneous SIR model (9) with initial conditions $(\tilde{S}_0, \tilde{I}_0, 0)$ and homogeneous parameter $\bar{\beta}_0$, which allows to conclude.
An interesting property of System (13) is the conservation of the following entropy-like quantity

$$V(t) = \int_0^1 S(t,a)\,da + I(t) - \frac{\gamma N}{\beta_m}\int_0^1 \log S(t,a)\,da, \quad \text{where} \quad \beta_m = \int_0^1 \beta(a)\,da. \tag{24}$$

The property $V(t) = V(0)$ can be deduced from (13).
Next, we consider more specifically the role of diffusion $d > 0$. First, we note that in the presence of diffusion, the above entropy $V^d$ based on solutions $(S^d(t,a), I^d(t), R^d(t))$ of (12) is non-decreasing in a similar manner as in the framework of viscous perturbations of scalar conservation laws:

$$\frac{dV^d(t)}{dt} = -\frac{\gamma N d}{\beta_m}\int_0^1 \frac{(\partial_a S^d(t,a))^2}{S^d(t,a)}\,da \le 0. \tag{25}$$

The following result emphasises the differences between the case with diffusion ($d > 0$) and the case without ($d = 0$).

**Proposition 2.** *Assume $d > 0$ and consider solutions $(S^d(t,a), I^d(t), R^d(t))$ of (12). The following properties hold:*

*(i)* *$t \mapsto S^d(t,a)$ is not necessarily decreasing in time for $a \in (0,1)$ given.*

*(ii)* *If $\mathcal{R}_0 = \int_0^1 S_0^d(a)\beta(a)\,da/(N\gamma) > 1$, then there exists $t_0 > 0$ such that $t \mapsto I^d(t)$ is increasing for $t \in (0, t_0)$.*

*(iii)* *If $\mathcal{R}_0 < 1$, $t \mapsto I^d(t)$ is decreasing in time over a finite time interval. Still, there are situations where $I^d$ can be increasing over some later time intervals.*

We will give the proof of this result in a forthcoming work.

We now compare the combined impact of heterogeneity and variability compared with the case of homogeneous populations.

**Theorem 2.** *For given $d > 0$, let $(S^d(t,a), I^d(t), R^d(t))$ be solution of (12) with initial conditions $(S_0^d(a) = S_0(a), I^d(0) = I_0, R^d(0) = 0)$. The following results holds*

*(i)* *$S^d(t,\cdot)$ converges as $t$ goes to $+\infty$ to a positive constant $S_\infty^d$ independent of $a$.*

*(ii)* *Let $(\bar{S}(t), \bar{I}(t), \bar{R}(t))$ be solution of the homogeneous SIR model (9) with initial conditions $(\bar{S}(0) = \int_0^1 S_0(a)\,da, \bar{I}(0) = I_0, \bar{R}(0) = 0)$ and transmission rate $\beta_m = \int_0^1 \beta(a)\,da$. Then, denoting $\bar{S}_\infty$ the large time limit of $t \mapsto \bar{S}(t)$, one has $S_\infty^d > \bar{S}_\infty$.*

We already know from Theorem 1 above that heterogeneity is beneficial in terms of herd immunity. From this point of view, the effect of variability is a priori not so clear because the constant shuffling of population keeps fuelling as it were the epidemic. However, part (ii) in the previous result shows that heterogeneity and variability still have a positive effect on herd immunity (that is $S_\infty^d > \bar{S}_\infty$).
The convergence of solutions to states that do no depend on $a$ is natural because of the role of diffusion. Indeed, the diffusion term takes over once the $I(t)$ term becomes small and thus the first equation describes a diffusion. But this only happens in the long term and the relevance here is rather at intermediate times.
**Sketch of proof:** Assuming that the initial profile $S_0$ and $\beta$ are regular enough in the $a$ variable, integration by parts of the $S$ equation of (12) multiplied by $\partial_t S$ leads to:

$$\int_0^1 \frac{S^d(t,a)^2}{2}da + \int_0^t I^d(\tau)\int_0^1 \frac{\beta(a)S^d(\tau,a)^2}{N}d\tau da + \int_0^t\int_0^1 d\,|\partial_a S^d(\tau,a)|^2\,d\tau da$$

$$= \int_0^1 \frac{S_0^d(a)^2}{2} da \quad \text{for all} \quad t \geq 0. \tag{26}$$

Moreover,

$$\int_0^t \int_0^1 |\partial_t S^d(\tau,a)|^2 \, d\tau da + \int_0^1 d \, \frac{|\partial_a S^d(t,a)|^2}{2} da = -\int_0^t \int_0^1 \frac{\beta(a) S^d(\tau,a) I^d(\tau)}{N} \partial_t S^d(\tau,a) \, d\tau da$$

$$+ \int_0^1 d \, \frac{|\partial_a S_0^d(a)|^2}{2} da. \tag{27}$$

Hence,

$$\int_0^t \int_0^1 |\partial_t S^d(\tau,a)|^2 \, d\tau da + \int_0^1 d \, \frac{|\partial_a S^d(t,a)|^2}{2} da \leq \frac{1}{2} \int_0^t \int_0^1 |\partial_t S^d(\tau,a)|^2 \, d\tau da$$

$$+ \sup_{\tau \geq 0} \frac{I^d(\tau)}{N} \sup_{a \in (0,1)} \beta(a) \int_0^t \int_0^1 \frac{\beta(a) S^d(\tau,a)^2 I^d(\tau)}{2N} \, d\tau da + \int_0^1 d \, \frac{|\partial_a S_0^d(a)|^2}{2} da \quad \text{for all} \quad t \geq 0, \tag{28}$$

so that we have

$$\int_0^t \int_0^1 |\partial_t S^d(\tau,a)|^2 \, d\tau da + \int_0^1 d \, |\partial_a S^d(t,a)|^2 \, da \leq \int_0^1 d \, |\partial_a S_0^d(a)|^2 \, da$$

$$+ \sup_{a \in (0,1)} \beta(a) \int_0^t \int_0^1 \frac{\beta(a) S^d(\tau,a)^2 I^d(\tau)}{N} \, d\tau da. \tag{29}$$

It follows that $\partial_a S^d(t,\cdot)$ converges to 0 in $L^2(0,1)$: if it does not hold, there exists $(t_n)_{n \in \mathbb{N}} \to +\infty$ and $\alpha > 0$ such that $\|\partial_a S(t_n,\cdot)\|^2_{L^2(0,1)} \geq \alpha$. On the other hand, there exists $T$ such that

$$\sup_{a \in (0,1)} \beta(a) \int_T^{+\infty} \int_0^1 \frac{\beta(a) S^d(\tau,a)^2 I^d(\tau)}{N} \, d\tau da \leq \frac{\alpha}{2}.$$

From (29) generalised between two positive times, for all $t > T$, there exists $n$ such that $t_n > t$ and

$$\int_0^1 d \, |\partial_a S^d(t_n,a)|^2 \, da \leq \int_0^1 d \, |\partial_a S^d(t,a)|^2 \, da + \frac{\alpha}{2}, \quad \text{so that} \quad \int_0^1 d \, |\partial_a S^d(t,a)|^2 \, da \geq \frac{\alpha}{2},$$

which contradicts the integrability of $\partial_a S$ in $L^2(\mathbb{R}^+ \times (0,1))$. Next, from (29), we deduce that $S^d(t,\cdot)$ is bounded in $H^1(0,1)$ uniformly in time, so that it is compact in $C^{0,1/2}(0,1)$: a sequence $(t_n)_{n \in \mathbb{N}} \to +\infty$ exists such that $S^d(t_n,\cdot)$ converges in $C^{0,1/2}(0,1)$ to some $S^d_\infty \in H^1(0,1)$ as $n$ goes to $+\infty$. Since $\partial_a S^d_\infty = 0$, $S^d_\infty$ is a constant. In order to prove the uniqueness of $S^d_\infty$, we observe that for all $(t,t') \in \mathbb{R}^{+2}$

$$\frac{1}{2} \int_0^1 \left( S^d(t,a)^2 - S^d(t',a)^2 \right) da = -d \int_{t'}^t \int_0^1 |\partial_a S^d(\tau,a)|^2 \, d\tau da - \int_{t'}^t \int_0^1 \frac{\beta(a) I^d(\tau) S^d(\tau,a)}{N} \, d\tau da, \tag{30}$$

and use the fact that $\partial_a S \in L^2(\mathbb{R}^+ \times (0,1)$, $S/N \leq 1$, and $I \in L^1(\mathbb{R}^+)$.

In order to prove (ii), using the fact that $dR^d(t)/dt = \gamma I^d(t) \geq$ and $R^d(t) \leq N$, we deduce that $R^d(t)$ converges to some $R^d_\infty > 0$. The conservation of total population $\int_0^1 S^d(t,a) \, da + I^d(t) + R^d(t) = 1$ then yields the convergence of $I^d(t)$ to 0 as $t \to +\infty$.

Integrating (25) between 0 and $t$, and letting $t$ go to $+\infty$ leads to

$$S^d_\infty - \frac{\gamma N}{\beta_m} \log S^d_\infty = V^d(0) - d \int_0^{+\infty} \int_0^1 \frac{|\partial_a S^d(t,a)|^2}{S^d(t,a)} dt da, \tag{31}$$

which yields the estimate $S^d_\infty > \bar{S}_\infty$. The detailed proofs and further developments are postponed to a forthcoming work.

**Some simulations with rebounds and plateaus.** Some simulations show that the model adequately captures dynamical features such as epidemiological rebounds and plateaus. The initial conditions considered here are (18) where $\kappa = 0.44$, $S_0 = 90$ and $S_1 = 119,000$ (total initial susceptible individuals are around $46,398$), initial infected people $I(0) = 1$, and $\beta$ is given by (17) for $\beta_0 = 0.008$ and $\eta$ is taken such that $\beta(1) = 26$.

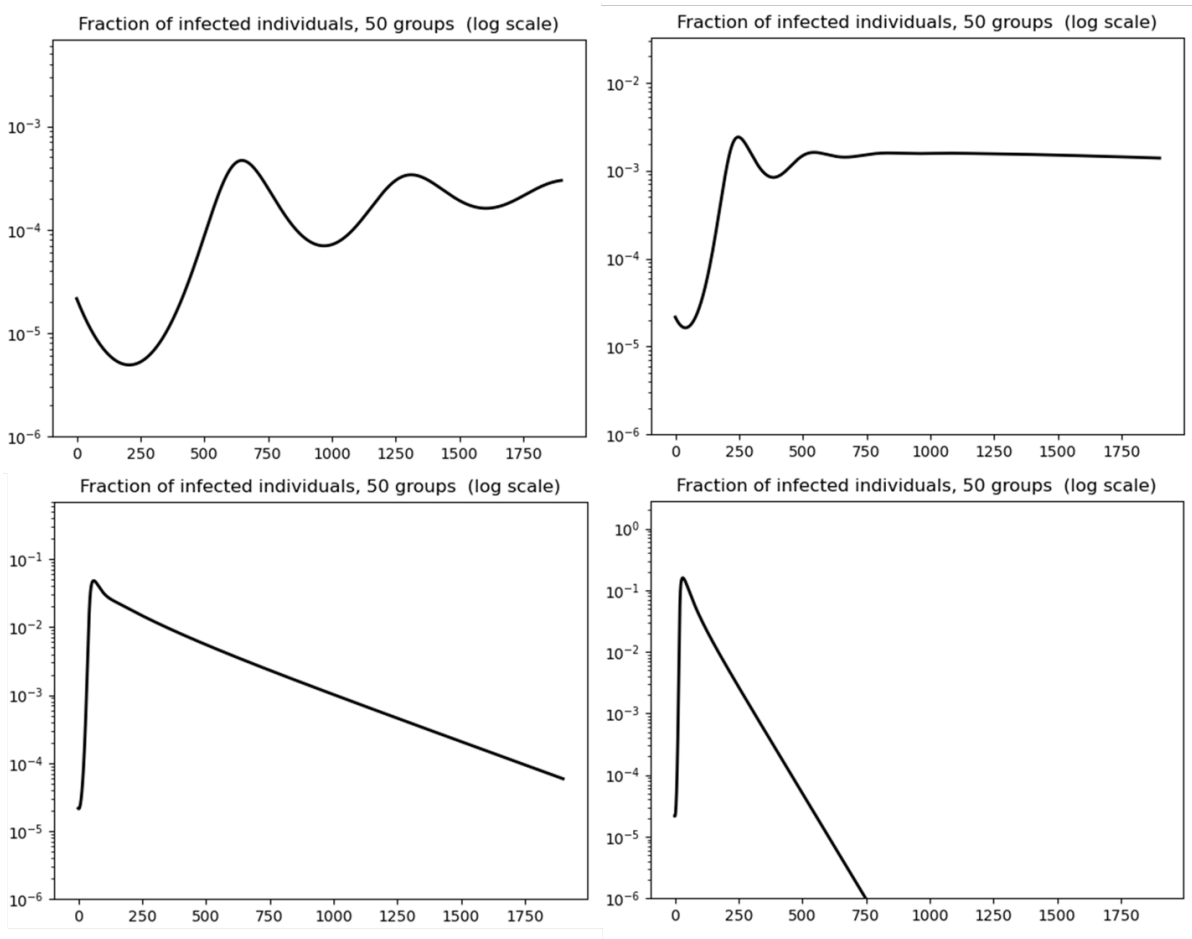


**Fig. 16.** Model behaviour depending on diffusion parameter values: infected rate dynamics in logarithmic scale. From left to right and then top to bottom, graphs are associated with $d = 10^{-5}$, $d = 5 \cdot 10^{-5}$, $d = 10^{-3}$ and $d = 5 \cdot 10^{-3}$ (in $day^{-1}$ unit)

Diffusion of susceptible individuals seem to be the most sensitive parameter in order to obtain either traditional *SIR* type behaviour or rebounds and plateau-type dynamics. Figure 16 illustrates the changes of dynamics depending on diffusion parameter $d$. The associated dynamics of susceptible individuals are represented in Figure 17 when variable $a \in (0,1)$ is discretized into $n_g = 50$ groups through a finite difference scheme. Coloured curves represent each discretized group and the black curve the average of the groups (total susceptible individuals).

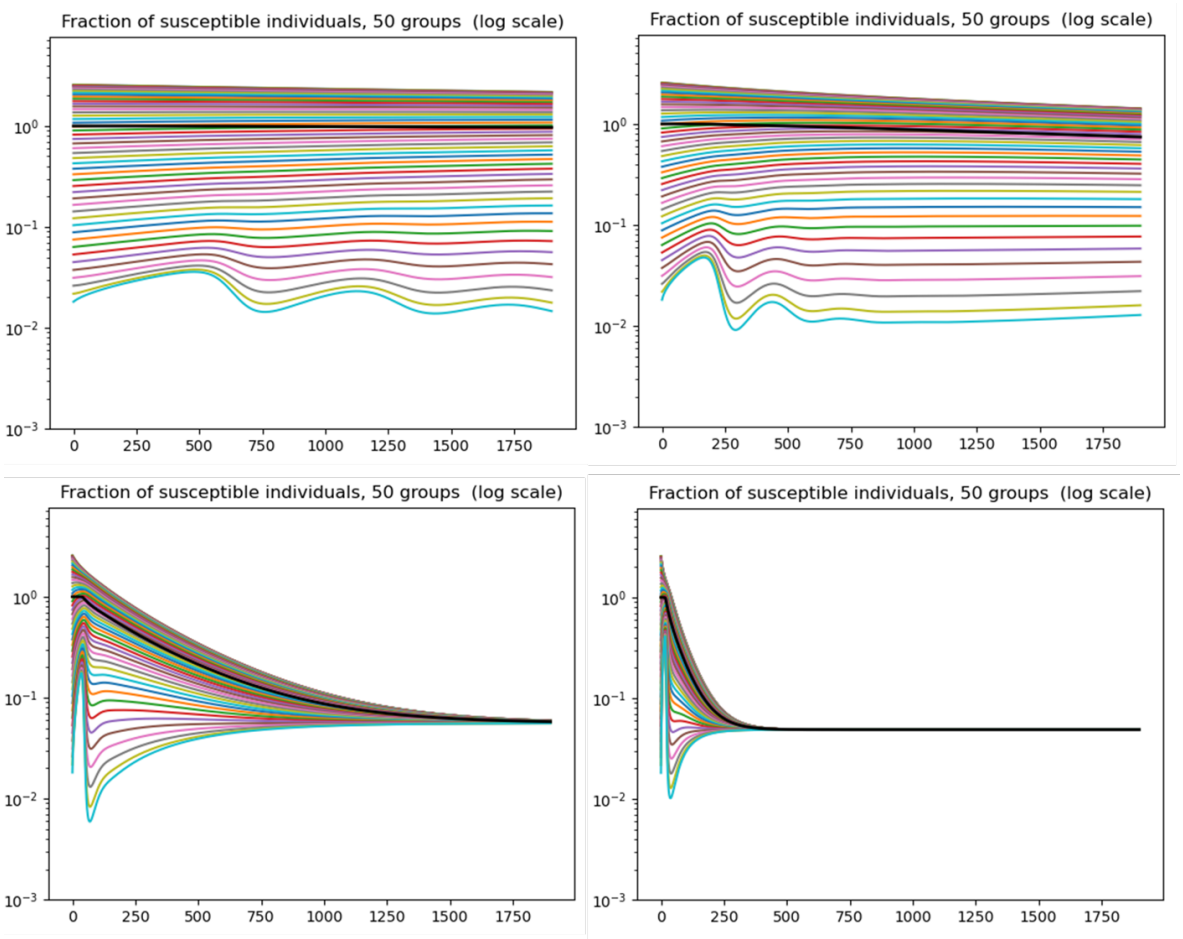


**Fig. 17.** Model behaviour depending on diffusion parameter values: 50 groups of susceptible individuals, in logarithmic scale. The black curve corresponds to total susceptible individuals and the bottom light blue curve is associated with the most risky behaviour ($a$ close to 1) and may not be monotonically decreasing in time. From left to right and then top to bottom, graphs are associated with $d = 10^{-5}$, $d = 5 \cdot 10^{-5}$, $d = 10^{-3}$ and $d = 5 \cdot 10^{-3}$ (in $day^{-1}$ unit)

One can observe the effect of diffusion: while individuals with the riskiest behaviour are rapidly consumed, their group is renewed through diffusion from the majority of less infectious individuals. Depending on the parameter range of the diffusion coefficient $d$ in (12), it leads to multiple epidemic waves, plateaus, or classical single wave *SIR*-like dynamics.

**Stability of plateaus: the key role of diffusion.** Let us illustrate in a heuristic manner the impact of diffusion on the existence and stability of plateaus in the dynamics of infectious individuals in Model (12). First, the second derivative of $x = \log I$ satisfies

$$\frac{d^2x}{dt^2}(t) = -I(t)\int_0^1 \frac{\beta^2(a)S(t,a)}{N^2}\,da - d\int_0^1 \beta'(a)\frac{\partial S(t,a)}{\partial a}\,da. \tag{32}$$

In the case $d = 0$, we deduce from (32) that $d^2x/dt^2(t) < 0$ for all $t \in \mathbb{R}^+$, so that $x = \log I$ is a strictly concave function of time. The consequence is that only slow decrease at the beginning of the decay phase can be obtained. Plateaus or shoulder-like patterns cannot arise in the case $d = 0$, nor rebounds since $x(t)$ has a single maximum after which it keeps decreasing.
In the case $d \neq 0$, Equation (32) can be reformulated as

$$\frac{d^2x}{dt^2}(t) = \eta(t)\left(1 - \frac{I(t)}{I_p(t)}\right), \tag{33}$$

where

$$\eta(t) = -\frac{d}{N}\int_0^1 \beta'(a)\frac{\partial S(t,a)}{\partial a}\,da \quad \text{and} \quad I_p(t) = \frac{N^2\eta(t)}{\displaystyle\int_0^1 \beta^2(a)S(t,a)\,da}. \tag{34}$$

Assuming that $\beta$ and $S_0$ are respectively increasing and decreasing functions of $a$, we deduce from Proposition 1 that $\eta(t)$ and

$I_p(t)$ are always positive. Introducing $x_p(t) = \log I_p(t)$, Equation (33) can be rewritten as

$$\frac{d^2x}{dt^2}(t) + \eta(t)\left(\exp[x(t) - x_p(t)] - 1\right) = 0. \tag{35}$$

It means that $d^2x/dt^2(t)$ may be positive (resp. negative) depending on whether $x(t)$ is lower (resp. greater) than $x_p(t)$. In particular, in the neighborhood of times $t$ such that $x(t)$ is close to $x_p(t)$,

$$\exp[x(t) - x_p(t)] - 1 \approx x(t) - x_p(t),$$

so that oscillations may arise, leading to patterns such as the ones seen on Figure 16. These oscillations also explain rebounds, or shoulders-like patterns before plateaus. Of course, this is not a proof but an indication to this effect and it warrants further mathematical investigation.

**More general systems.** First, it should be noted that the coefficient $\beta(a)$ can be time dependent: $\beta = \beta(t,a)$. We can thus incorporate public health policies changes such as social distancing, lockdowns etc. For instance, we can consider a transition between two profiles $\beta_1(a)$ and $\beta_2(a)$ at a given date.

Then, we observe that System (12) is a particular case of a more general class of models. Indeed, we may want to also consider heterogeneity in the compartment of infected. This for instance reflects the division between symptomatic and asymptomatic individuals, or other differences. In particular, in terms of behaviour one can then include effects such as "super-spreaders". We then get a model where the population of infected also depends on the same parameter $a$, that is $I = I(t,a)$. We assume that the probability of interaction between a susceptible of class $a$ and an infected of class $b$ is given by a coefficient of transmission $\mathfrak{B}(a,b)$. These considerations lead us to the following general system.

$$\begin{cases} \dfrac{\partial S(t,a)}{\partial t} - d\dfrac{\partial^2 S(t,a)}{\partial a^2} = -\dfrac{S(t,a)}{N}\displaystyle\int_0^1 \mathfrak{B}(a,b)I(t,b)\,db, \\ \dfrac{\partial I(t,a)}{\partial t} = \dfrac{S(t,a)}{N}\displaystyle\int_0^1 \mathfrak{B}(a,b)I(t,b)\,db - \gamma I(t,a), \\ \dfrac{\mathrm{d}R(t)}{\mathrm{d}t} = \gamma\displaystyle\int_0^1 I(t,a)\,da. \end{cases} \tag{36}$$

We observe that our model above (12) is derived from this more general class by assuming that $\mathfrak{B}(a,b)$ is independent of $b$. In fact, the same is true for a finite number of states. This type of modelling of heterogeneity can be found in Magal et al.(17) and Arino et al. (14) in the discrete setting ($a$ has a finite number of values) and without diffusion ($d = 0$). These works do not seem to yield shoulder, plateau and rebound-type patterns.

For simplicity we assume that the decay rate $\gamma$ is uniform. Note that the second equation states that new infected individuals of class $a$ result from an infection of a susceptible individual of the same class. Other choices are possible, but then, in order to be consistent, one would need to introduce new compartments such as symptomatic/asymptomatic, hospitalised etc. Likewise, we could consider other variants of the $SIR$ model, such as the $SEIR$ model. It would then be natural to assume that exposed individuals also move around. We are then led to a system of the following kind.

$$\begin{cases} \dfrac{\partial S(t,a)}{\partial t} - d\dfrac{\partial^2 S(t,a)}{\partial a^2} = -\dfrac{S(t,a)}{N}\displaystyle\int_0^1 \mathfrak{B}(a,b)I(t,b)\,db, \\ \dfrac{\partial E(t,a)}{\partial t} - d\dfrac{\partial^2 E(t,a)}{\partial a^2} = \dfrac{S(t,a)}{N}\displaystyle\int_0^1 \mathfrak{B}(a,b)I(t,b)\,db - \rho E, \\ \dfrac{\partial I(t,a)}{\partial t} = \rho E - \gamma I(t,a), \\ \dfrac{\mathrm{d}R(t)}{\mathrm{d}t} = \gamma\displaystyle\int_0^1 I(t,a)\,da. \end{cases} \tag{37}$$

We can also consider a more general diffusion process on the behavioural variables than Brownian diffusion. We already mentioned above Eq. (11) the version involving a diffusion $\partial^2_{aa}[\sigma^2(a)S(t,a)]$ rather than $d\,\partial^2_{aa}S(t,a)$.

With a somewhat different point of view, we can assume that there is a probability kernel $K(a,b)$ that represents the probability distribution that an individual with trait $b$ jumps to behaviour $a$. Then, the variation of $S$ involves an integral operator rather

than a heat operator. We are then led to the following system.

$$\begin{cases} \dfrac{\partial S(t,a)}{\partial t} = \displaystyle\int_0^1 K(a,b)S(t,b)\,db - S(t,a) - \beta(a)S(t,a)\dfrac{I(t)}{N}, \\ \dfrac{\mathrm{d}I(t)}{\mathrm{d}t} = \dfrac{I(t)}{N}\displaystyle\int_0^1 \beta(a)S(t,a)\,da - \gamma I(t), \\ \dfrac{\mathrm{d}R(t)}{\mathrm{d}t} = \gamma I(t), \end{cases} \tag{38}$$

System (38) is supplemented with the same type of initial conditions as above (12). Note that there are no boundary conditions in this case.

It is classical that one can recover System (12) as a certain limit of the more general class of systems (38) (compare e.g. the article by the first author et al. (40)). However, here it is interesting to look at more behavioural assumptions on the kernel $K$. For instance, it is natural to assume that $K(a,b)$ is singular when $b$ is close to $a$, that $K(a,b)$ is rather large when $b$ is close to 1, whereas it is very small when $b$ is small. We leave the study of such kernels for future study.

**Digital PCR in the framework of wastewater based epidemiology.** Wastewater-based surveillance has been used over the past two decades to assess in real time the exposure of people to chemicals (in particular pollutants, licit and illicit or misused therapeutic or other drugs) and pathogens (bacteria or viruses) at the community level (41). It provides a global picture of population health at the scale of an urban area connected to wastewater treatment plants (WWTP). Fully complying with the privacy of individuals, this approach rests on an explicit, objective observation. It also has the advantage to serve as an early warning system. Nonetheless, there are also uncertainties in the quantitative interpretation of these measurements. In particular, dilution due to storm-water, infiltration of parasite inflow water, varying population, bio-marker degradation, varying time spent in the flow..., all generate a variety of biases.

In spite of these uncertainties, there is a renewed interest in wastewater based epidemiology (WBE), in the context of the Covid-19 outbreak. There are numerous initiatives over the world, in Australia (27), Japan (28), Netherlands (29), Detroit (30), Paris (31), Spain (32) and Canada (33). The recent periodic sampling campaigns in WWTPs rely on classical quantitative real time PCR (qRT-PCR) to estimate the concentration of SARS-Cov-2. At this stage, these programs serve as early warning platforms based on the observation that they detect SARS-Cov-2 in sewers at least seven days ahead of individual testings (30, 32, 34).

Only a few of them also used digital PCR on sub-samples to assess the measurement uncertainties of qRT-PCR measurements. Even if the benefit of dPCR in particular for low viral loads is mentioned (33), it is not currently used routinely for systematic measurements, mainly for economic reasons.

**Plateau-type patterns at country level.** In most European countries, Covid-19 data exhibit similar plateau-type features, as illustrated in the so called effective $R$ coefficient computed from the daily number of deaths (25). This effective $R$ coefficient exhibits phases where it stabilises around 1, which corresponds to an epidemic plateau.

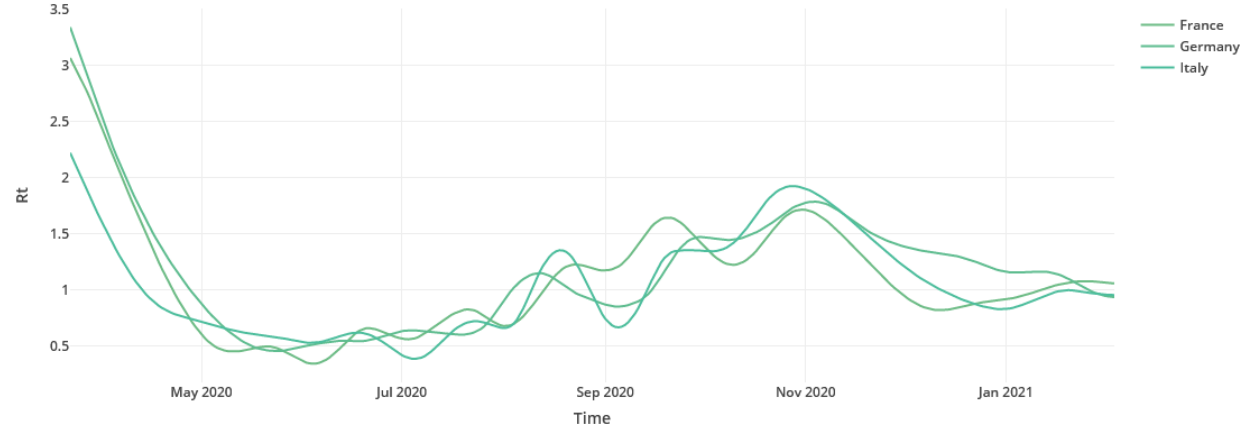


**Fig. 18.** Effective $R$ coefficient as a function of time (22) https://shiny.biosp.inrae.fr/app_direct/mapCovid19/.

Campo et al. (42) study more specifically the dynamics of Covid-19 in Italy and put emphasis on plateaus and multi-wave patterns, for which they discuss meta-stability properties.

Press articles illustrate the political leaders' observation that the Covid-19 outbreak has reached a plateau phase: Reuters, January $21^{st}$, 2021 (43); French government information service, December $10^{th}$, 2020 (44); A. Merkel, Westdeutsche Zeitung, January $21^{st}$, 2021 (45); E. Macron, February $2^{nd}$, 2021, interview on TF1 TV channel (46)).