Using posterior predictive distributions to analyse epidemic models: COVID-19 in Mexico City

Ramsés H Mena, Jorge X Velasco-Hernández, Natalia B Mantilla-Beniers, Gabriel A Carranco-Sapiens, Luis Benet, Denis Boyer and Isaac Pérez Castillo

1 Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, México CDMX, Apartado Postal 20-726, 01000, México
2 Instituto de Matemáticas, Unidad Juriquilla, Universidad Nacional Autónoma de México, Unidad Juriquilla 76230, Querétaro, México
3 Facultad de Ciencias, Universidad Nacional Autónoma de México, 04510 CDMX, México
4 Instituto de Ciencias Físicas, Universidad Nacional Autónoma de México, Av. Universidad s/n, Col. Chamilpa, C.P. 62210 Cuernavaca, Morelos, México
5 Departamento de Sistemas Complejos, Instituto de Física, Universidad Nacional Autónoma de México, Apartado Postal 20-364, 01000 CDMX, México
6 Departamento de Física Cuántica y Fotónica, Instituto de Física, Universidad Nacional Autónoma de México, Apartado Postal 20-364, 01000 CDMX, México
7 London Mathematical Laboratory, 8 Margravine Gardens, London, W68RH, United Kingdom

Author to whom any correspondence should be addressed.
E-mail: isaacpc@fisica.unam.mx

Keywords: epidemiological models, COVID-19, Bayesian statistics, Monte Carlo methods

Abstract

Epidemiological models usually contain a set of parameters that must be adjusted based on available observations. Once a model has been calibrated, it can be used as a forecasting tool to make predictions and to evaluate contingency plans. It is customary to employ only point estimators of model parameters for such predictions. However, some models may fit the same data reasonably well for a broad range of parameter values, and this flexibility means that predictions stemming from them will vary widely, depending on the particular values employed within the range that gives a good fit. When data are poor or incomplete, model uncertainty widens further. A way to circumvent this problem is to use Bayesian statistics to incorporate observations and use the full range of parameter estimates contained in the posterior distribution to adjust for uncertainties in model predictions. Specifically, given an epidemiological model and a probability distribution for observations, we use the posterior distribution of model parameters to generate all possible epidemic curves, whose information is encapsulated in posterior predictive distributions. From these, one can extract the worst-case scenario and study the impact of implementing contingency plans according to this assessment. We apply this approach to the evolution of COVID-19 in Mexico City and assess whether contingency plans are being successful and whether the epidemiological curve has flattened.

1. Introduction

December 2019 saw the start of an outbreak of pneumonia of unknown etiology in Wuhan, China. This would be recognised as a disease caused by a new coronavirus able to infect humans and transmit within human populations. By January 23, Chinese authorities had taken severe measures to contain its spread: travel bans, mobility restrictions within Wuhan, isolation of suspected and confirmed cases, mass gathering bans and school and entertainment venue shutdowns. This did not prevent the virus from reaching several other countries and all regions of China quickly. On January 30, with 7711 confirmed cases in China and 83 in other countries, the World Health Organization declared SARS-CoV-2 a Public Health Emergency of International Concern [1, 2].

Mexico confirmed its first cases of COVID-19 on February 27 in travellers returning from Italy to Sinaloa and Mexico City, respectively. On March 15, the Mexican National Committee for Safety in Health (Comité Nacional para la Seguridad en Salud)
announced the start, on March 23, of distancing measures to mitigate the transmission of COVID-19, thus triggering the second phase of the epidemic. Phase three would be declared nearly a month later, on April 21. Distancing measures, implemented on March 23, included suspension of all non essential activities of public, private and social sectors, and was initially planned to last until April 30, but was later extended until May 17 or May 30, depending on the local situation of each municipality in the country. These measures were designed to lower disease incidence rates of COVID-19 and keep the number of hospitalized and critical cases manageable [3, 4].

Public health authorities admit that the number of actual infected cases in Mexico City, and in the country at large, is larger than those reported. The testing rate in Mexico is the lowest among the OECD countries [5] and the positivity rate for tests in Mexico City on the week ending on May 6, for example, ranged between 24.5% to 41.7% depending on the municipality [6]. Both of these factors likely contributed to a large sub-reporting of cases. The strain on the healthcare system is already important in Mexico City and other large population centers in the country. To the day of submission of this work, the model used by the Federal Government’s General Directorate of Epidemiology (Dirección General de Epidemiología) was not publicly released. It was finally made public in June [7], well past the period for which it was used to forecast the epidemic. There was no technical information available, before that date, on the model’s fundamental underlying assumptions on contact rates, initial conditions, percentage of asymptomatic carriers or basic reproduction number, among others. As far as we are aware, there is only one peer-reviewed model published previous to May 30 on the Mexican case, but it is centered on the analysis of the efficacy of mitigation strategies still in effect [8].

There is, therefore, the urgent need to have alternative models that project feasible scenarios of the epidemic in Mexico allowing, in turn, for comparisons of model predictions and the subsequent improvement of models, particularly in view of the foreseeable trajectory uniqueness and it may be expected that this type of model may be able to reproduce empirical data and predict future outcomes equally well.

In this scenario, mathematical models are a natural tool for identifying what needs to be done in order to avoid saturating the healthcare system. Models are often used to estimate, for example, the extent of the reduction in the effective transmission rate needed to control an epidemic. However, most of these models fall in a category commonly known as sloppy models [9]. These are models that depend on a large number of parameters and for which, once fitted to limited or noisy data, a broad range of certain parameter values produce similarly acceptable fits. This is clearly disconcerting, since using different parameter estimates one obtains widely different predictions from the same model, rendering its application to forecasting impractical, a problem which unfortunately is frequently overlooked at times when theoretical expectations and scientific rigor are in dire need. Indeed, as pointed out in [10], deterministic epidemiological models may give the false impression of trajectory uniqueness and it may be expected that this type of model may be able to reproduce empirical data and predict future outcomes equally well.

This pandemic has shown that estimates for various parameters vary wildly from country to country. Thus, comparing fitted parameters between different countries to either discard or validate a particular model may be misleading. It would be rather more sensible to run different scenarios for a given population and compare the results for that particular setting. The reasons as to why fitted parameters vary so much from country to country lie on the particular characteristics of each population or community, such as age distributions, risk factors, gross domestic product, access to healthcare, social norms, climate, to mention but a few.

In section 4 we perform a global sensitivity analysis (GSA) of the model we use, in order to ascertain which parameters are most important in the variability of model output. We summarise our results and discuss future work in section 5.
2. On epidemiological models

The basic idea of epidemiological compartmental models is to split the host population (often assumed to be of constant size $N$) into $r$ compartments corresponding to infection stages, so that $\mathcal{N}_a(t)$ indicates the population in state $a$, $a = 1, \ldots, r$. We thus introduce vector $\mathcal{N}(t) = (\mathcal{N}_1(t), \ldots, \mathcal{N}_r(t))$ and assume the epidemic to follow a set of nonlinear ODEs

$$\frac{d\mathcal{N}(t)}{dt} = \mathcal{F}[\mathcal{N}(t), \theta],$$

where $\theta = (\theta_1, \ldots, \theta_p)$ is a set of $p$ model parameters. Let $\mathcal{N}(t, \theta)$ denote the solution for the set of equation (1) given $\theta$. Examples of simple epidemiological models are the susceptible–infected–recovered or the susceptible–exposed–infected–recovered models, for which the states are $\mathcal{N} = (S, I, R)$ or $\mathcal{N} = (S, E, I, R)$, respectively. More realistic models, as the one we will use here with the aim to estimate disease toll and burden, introduce additional states to follow hospitalized and critically-ill patients.

Suppose now that we have an observational dataset $D \equiv \{\mathcal{N}^{\text{obs}}(t)\}_{i=0}^{\infty}$, possibly with an observational time- and compartmental-correlation matrix. From here we can derive the likelihood $P(D|\theta)$ of observing this dataset given a set of parameters. Using Bayes’ rule, the posterior distribution of the parameters given the dataset is simply

$$P(\theta|D) \propto P(D|\theta)P_0(\theta),$$

where $P_0(\theta)$ is the prior distribution of the parameters. The standard way to calibrate the model is to find the set of parameters, denoted here as $\theta^*$, which maximizes the posterior distribution $P(\theta|D)$, that is, $\theta^* = \arg\max_\theta P(\theta|D)$. These are sometimes referred to as maximum a posteriori (MAP) estimators. When the prior distribution is flat, and the posterior distribution exists, $\theta^*$ coincides with the maximum likelihood estimator. Once the model has been calibrated using this point estimator, the evolution of the epidemic is given by $\mathcal{N}(t, \theta^*)$, which can then be used to make predictions.

Unfortunately, this method tends to fail for the so-called sloppy models [9], as the variances in parameter calibration can be rather large in certain directions of parameter space. In the present case, we find large deviations when using data only from the beginning of the epidemic curve. As a result, there is a large uncertainty in the conditions leading to a desired state, which renders the deterministic approach inadequate as a forecasting tool, e.g., to implement contingency plans. A full Bayesian approach considers the uncertainty captured by the whole posterior distribution $P(\theta|D)$, and not only the deterministic point estimator $\theta^*$ [10]. From this principle we can introduce various posterior predictive distributions. We start by considering the posterior predictive compartmental distribution given by:

$$P(n, t|D) = \int d\theta \ P(\theta|D)P[\mathcal{N}(t, \theta) = n|\theta],$$

where $P[\mathcal{N}(t, \theta) = n|\theta] = \delta[n - \mathcal{N}(t, \theta)]$, since the evolution equations modelling the epidemic (1) are deterministic. Here $P(n, t|D) = \text{Prob}[\mathcal{N}(t, \theta) = n|D]$ corresponds to the probability of observing a given state value, $n = (n_1, \ldots, n_r)$ at time $t$ given the data set $D$. Clearly, if $P(\theta|D)$ has a marked peak around $\theta^*$, with the extreme case being $P(\theta|D) = \delta(\theta - \theta^*)$, then $P(n, t|D)$ evolves deterministically according to $\mathcal{N}(t, \theta^*)$, that is $P(n, t|D) = \delta[n - \mathcal{N}(t, \theta^*)]$, which then recovers the previously mentioned standard approach. However, if the posterior distribution $P(\theta|D)$ is widely spread, so will be $P(n, t|D)$. Thus, we need to consider the whole distribution $P(n, t|D)$ as a forecasting tool, and use it to analyse the impact of contingency plans.

Generally, we expect the posterior predictive distribution $P(n, t|D)$ to have a compact support, since the host population is taken to be constant. With this in mind, we will denote as $\Omega^{\text{low}}(t)$ and $\Omega^{\text{up}}(t)$ its lower and upper boundaries, respectively. Consequently, $P(n, t|D)$ is zero for $n \notin [\Omega^{\text{low}}(t), \Omega^{\text{up}}(t)]$. The two boundaries, $\Omega^{\text{low}}_{\text{a}}(t)$ and $\Omega^{\text{up}}_{\text{a}}(t)$, which correspond fairly intuitively to the lower and upper envelopes of all possible epidemiological curves $\mathcal{N}_a(t)$ with $\theta$ drawn from $P(\theta|D)$, can be understood in epidemiological terms as the best- and worst-case scenarios of the epidemic for state $a$ at time $t$, respectively. Thus, they are fairly useful to determine the impact of the epidemic on a healthcare system. For instance, if we were to have a compartment $C$ modelling critically-ill patients, the corresponding upper boundary $\Omega_{\text{C}}^{\text{up}}(t)$ gives a bound for the worst-case scenario. Thus, if a particular healthcare system has a given maximum capacity, denoted here as $B$ [e.g. total intensive care units (ICU) available] to treat critically-ill patients, then having $\Omega_{\text{C}}^{\text{up}}(t) > B$ at some point indicates that the healthcare system has demands exceeding its capacity. A careful, and successful, contingency plan must consider the worst possible outcome of the epidemic, so that implemented measures guarantee that $\Omega_{\text{C}}^{\text{up}}(t) < B$.

Equally important is to derive the posterior predictive distribution of times at which the epidemic curve will peak. Indeed, let $t_{\text{peak}} = \arg\max_\theta \mathcal{N}_a(t, \theta)$ be the time at which the epidemic reaches its peak for compartment $a$, and let us further denote $t_{\text{peak}} = (t_{\text{peak}}^{(1)}, \ldots, t_{\text{peak}}^{(r)})$. The corresponding posterior predictive distribution of times at which the peaks occur reads:

$$P(\{t_{\text{peak}}|D\} = \int d\theta \ P(\theta|D)\delta[t_{\text{peak}} - \arg\max_\theta \mathcal{N}(t, \theta)].$$

Notice that one would be tempted to predict the peak of the epidemic based on (2) by first
obtaining the mean value for a given compartment, \( \langle n_a(t) \rangle_{n_a[D]} \), and then looking for the time at which the mean curve peaks, \( \arg \max_t = \langle n_a(t) \rangle_{n_a[D]} \).

Clearly this is not necessarily equal to \( \langle t_{\text{peak}} \rangle_{n_a[D]} \), so it is more appropriate to use the posterior predictive distribution to correctly assess the probability of the peak occurring at a given time.

3. Model selection, and resulting analysis for COVID-19 in Mexico City

3.1. Model selection

For the compartment model used to analyse COVID-19 data for Mexico City, we have chosen to follow the one used in [15, 16] (and references therein). Here, susceptible individuals \( S \) become exposed \( (E) \) to the virus through contact with infected individuals \( I \). Exposed individuals progress towards the symptomatic state \( I \) within an average time \( \tau_I \). As usual, mixing is assumed to be homogeneous. Infected individuals \( I \) cause an average of \( R_0 \) secondary infections over their infectious period. After an average time \( \tau_I \) (days), infected individuals either recover or progress towards hospitalization. In turn, hospitalized individuals \( H \) either recover or worsen towards a critical state after a time \( \tau_H \). Critical individuals \( C \) allow us to model ICU demand. They either return to state \( H \) or die, moving to \( D \), after a time \( \tau_C \). Recovered individuals \( R \) are assumed to be immune. The dynamics of this model is given by the following set of differential equations:

\[
\frac{dS(t)}{dt} = -\beta(t) \frac{S(t)I(t)}{N} \tag{4}
\]

\[
\frac{dE(t)}{dt} = \beta(t) \frac{S(t)I(t)}{N} - \frac{E(t)}{\tau_I} \tag{5}
\]

\[
\frac{dI(t)}{dt} = \frac{E(t)}{\tau_I} - \frac{I(t)}{\tau_I} \tag{6}
\]

\[
\frac{dH(t)}{dt} = (1 - m) \frac{I(t)}{\eta} + (1 - f) \frac{C(t)}{\tau_C} - \frac{H(t)}{\tau_H} \tag{7}
\]

\[
\frac{dC(t)}{dt} = c \frac{H(t)}{\tau_H} - \frac{C(t)}{\tau_C} \tag{8}
\]

\[
\frac{dR(t)}{dt} = m \frac{I(t)}{\tau_I} + (1 - c) \frac{H(t)}{\tau_H} \tag{9}
\]

\[
\frac{dD(t)}{dt} = f \frac{C(t)}{\tau_C} \tag{10}
\]

The fraction of infections that are mild is \( m \), the fraction of cases that turn critical is \( c \), and the fraction of critical cases with fatal outcome is \( f \). Other variants of the model consider, for instance, a recovery time for mild infections which is different from \( \tau_I \), or a fraction of those infected that are asymptomatic. Equations (4)–(10) provide a relatively simple description of epidemic dynamics, including entry to and exit from the hospital, that allows us to focus on the number of hospitalized and critical cases, and foresee whether health services will be saturated. The transmission parameter in the model is taken to be

\[
\beta(t) = \frac{R_0 M(t)}{\tau_I}, \tag{11}
\]

where \( R_0 \) is the basic reproduction number, and \( M(t) \) models mitigation measures, where \( M(t) = 1 \) means that no such measures are taken. While, generally speaking, pathogens affect populations in an uneven way, due to heterogeneity in the risk experienced by age, comorbidities or other factors (e.g. behaviour, nutrition and so on), for simplicity we assume a population homogeneous in all respects. A generalization to include how a particular age distribution affects model evolution is straightforward [15], and is ongoing work.

3.2. Estimation of the parameters’ posterior distribution

To construct the parameters’ posterior distribution \( P(\theta|\mathcal{D}) \), we first need to discuss the choice of the model likelihood \( P(\mathcal{D}|\theta) \) and the prior distribution \( P_0(\theta) \). Taking into account a possibly rather large value \( N \) of the host population, and since it is not possible to have the empirical correlation matrix for the observational data, we take the model’s likelihood to be:

\[
P(\mathcal{D}|\theta) = \frac{1}{\sqrt{(2\pi)^{N_{\text{max}}}}^{\sum_{a=1}^{N_{\text{max}}}}} \times \exp \left[ \frac{1}{2} \sum_{a=0}^{N_{\text{max}}} \sum_{a'=1}^{N_{\text{max}}} \left( \mathcal{N}_a(t, \theta) - \mathcal{N}_a(\text{obs})(t) \right)^2 \right], \tag{12}
\]

where we have assumed that the observational dataset is independent, that Gaussian variables are identically distributed, with unit variance, and \( N_{\text{max}} \) is the time of the last recorded data. Granted, this is a gross oversimplification as, in principle, these data are time and compartment correlated and one would need the corresponding empirical correlation matrix, for which we have no information. Our approach is therefore reminiscent of an approximate Bayesian calculation, a method of statistical inference which does not require exact likelihood calculations and is often used, for instance, in population genetics [17].

Regarding the choice of \( P_0(\theta) \), it would seem appropriate in principle to consider an empirical \( a \) priori distribution by gathering the resulting parameter values \( \theta \) obtained in studies from other countries. While this approach is indeed tempting from a statistical point of view, it assumes that health conditions, comorbidities and other important aspects are the same in different countries, which is clearly untrue. Thus a cautious approach, aiming to avoid confirmation bias, is to use a flat prior within a range of parameters \( \theta_a \in [\theta_a^{\text{min}}, \theta_a^{\text{max}}] \), where the boundaries of the parameters are reasonably wide, enough to
be compatible with data available from other studies. That is

$$P_0(\theta) = \frac{1}{Z_0} \prod_{a=1}^{r} \int \{ \theta_a \in [\theta_a^{(\text{min})}, \theta_a^{(\text{max})}] \} ,$$  

(13)

where $Z_0$ is a normalization constant of the prior distribution and $\mathbb{I}$ is an indicator function. Given that the explicit solution of the evolution equations of the model is in general not available, the resulting parameters’ posterior distribution $P(\theta|\mathcal{D})$ is a rather complicated function of $\theta$, so that the natural way to carry out the integrals appearing in the expressions of the predictive posteriors is by numerically estimating $P(\theta|\mathcal{D})$ by a Monte Carlo method.

Before embarking the difficult task of calibrating the model using open data of SARS-CoV-2 cases in Mexico City, we sought to develop some intuition of the model’s behaviour presented in section 2. Consequently, we addressed how calibration behaves depending on the time window considered for the observational data. For an arbitrary, yet realistic, choice of parameters, we generate a synthetic data set $\{X^{(\text{syn})}(t)\}_{t=0}^{t_{\text{max}}}$ and explore the MAP problem in terms of $t_{\text{max}}$. We observed that for $t_{\text{max}} < t_{\text{peak}}$ the posterior distribution is indeed rather flat while for $t_{\text{max}} > t_{\text{peak}}$ a more defined maximum appears and the model is more easily calibrated. We must note that these observations are neither new nor remarkable for this type of models. However, they emphasise that whichever analysis is performed on these models, particularly if they are intended as a forecasting tool, using data from the beginning of the epidemic curve alone renders great variability of possible outcomes and must be carried out reckoning these properties [10].

Thus, as we have observational data only from the beginning of the epidemic, and since we are dealing here with sloppy models, the parameters’ posterior distribution is rather flat in most directions of parameter space. This indicates that the most efficient approach is first to solve the minimization problem

$$\theta^* = \arg \max_{\theta \in [\theta^{(\text{min})}, \theta^{(\text{max})}]} P(\mathcal{D}|\theta) ,$$  

(14)

and then to explore the space of parameters around the point $\theta^*$ uniformly, by perturbing it with a random variable $\epsilon$, that is $\theta^* \rightarrow \theta = \theta^* + \eta \epsilon$, with $\eta$ controlling the spread of the exploration in the parameter space. In this way, the two posterior predictive distributions presented above are approximated by

$$P(n, t|\mathcal{D}) = \frac{1}{V(\Omega, \theta^*)} \int_{\Omega} d\theta \ P[N(t, \theta) = n|\theta] ,$$  

(15)

$$P(t_{\text{peak}}|\mathcal{D}) = \frac{1}{V(\Omega, \theta^*)} \int_{\Omega} d\theta \ \delta[t_{\text{peak}} - \arg \max_{\theta \in \Omega} N(t, \theta)] ,$$  

(16)

which can be easily estimated by direct Monte Carlo sampling. Here $V_{\Omega, \theta^*} = \Omega(\theta^*) d\theta$, and $\Omega(\theta^*)$ stands for the integration region in the parameter space, centered at $\theta^*$ and of width $\eta$.

3.3. Analysis and results for COVID-19 in Mexico City

We have applied this approach to study the evolution of SARS-CoV-2 spread in Mexico City using the public database provided by the Federal Health Secretary (Secretaría de Salud) on June 26 [18, 19]. We considered data starting on February 27 (which we denote as $t = 0$) up to June 18, to allow for delays in reporting of cases due to either delays in seeking medical attention, reporting or test processing. The database allows us to obtain incidence time series (new cases in the last 24 h), as well as those newly hospitalized and critically-ill (complicated hospitalizations including the use of mechanical ventilators). It also includes the total number of deceased patients, both those who were lab confirmed for SARS-CoV-2, and those suspect, awaiting results of RT-PCR tests. Note that new cases do not correspond to the number of cases in each compartment, a piece of information which does not appear in the data. When calibrating the model we have considered a cautious approach to add half of the suspected cases to those confirmed for each of the infected compartments, based on estimations of the positivity test rate for suspect cases. Clearly, not all suspected cases will be confirmed as SARS-CoV-2, since this epidemic is happening concurrently with other seasonal diseases and therefore we are describing an aggregate of all respiratory diseases. However, we believe it is important to include some suspected cases since they anyway may increase demand on the healthcare system.

Further, another clarification is in order: while the data provide the initial conditions for some compartments, we do not have information for others, in particular, for the initial condition $S(0)$. Thus we consider $S(0)$ to be also a parameter of the model. In conclusion, when calibrating the model daily new cases is fitted to $E(t)/\tau_f$, the daily new hospitalized cases is fitted to $(1 - m)I(0)\frac{\beta}{\gamma} + (1 - f)\frac{\beta}{\gamma}$, the daily new critically-ill patients is fitted to $\frac{\beta}{\gamma}$, and finally, total death numbers is fitted to $D(t)$. As for the initial conditions, according to data $(S(0), E(0) = I(0) \times \tau_f, I(0) = 3, H(0) = 1, C(0) = 1, R(0) = 0, D(0) = 0)$, with the initial time $t = 0$ set to February 27. For the contingency plan, we take $M(t)$ a step function representing the decrease in effective contact rate induced by the mitigation measures (which has been smoothed for computational convenience) equal to 1 for $t$ smaller than 25 (corresponding to March 23, the date when the mitigation plan began in Mexico) and equal to $\gamma$ for $t > 25$. The value of $N$ is obtained as

$$N = S(0) + E(0) + I(0) + H(0) + C(0) + R(0) + D(0).$$
All in all, given the data, we follow the steps described in the previous section 3.2, namely: we firstly estimate the parameters’ posterior distribution \( P(\theta | D) \), and, secondly, we use it, according to (15) and (16), to estimate predictive posteriors.

To solve numerically equation (14) we used a Monte Carlo method, which consists in integrating the equations of motion with different parameters a few million times, and then using a black box optimization method written in Julia [20]. The resulting values for the optimal parameters \( \theta^* \) are summarised in table 1 (third column), where we also show the interval range \( \theta_\sigma \in [\theta_\sigma^{(\min)}, \theta_\sigma^{(\max)}] \) (second column) in which we looked for an optimal solution \( \theta^* \). The interval range for each parameter \( \theta_\sigma \) was chosen wide enough to be compatible with parameter values estimated in other studies (which we summarise in appendix).

Next, we use expressions (15) and (16) to estimate predictive posteriors, where we have chosen a value of \( \eta = 0.1 \), meaning that we explore the parameter space uniformly within each parameter’s interval around \( \theta^* \) but with the standard deviation of the uniform distribution scaled by \( \eta \). We have chosen this particular value of \( \eta \) to demonstrate that, even with a small value, a rather wide spread is obtained from the original deterministic curve. Later on, in section 4 we quantify precisely the impact of each parameter on model behaviour using Sobol’s total indices in a GSA. All the results for these distributions are summarised in the plots appearing in figure 1. The first row of plots in this figure shows the predictive posterior for daily new cases of infected, hospitalized, and critically-ill patients. In all cases, the solid red line corresponds to the model calibrated with \( \theta^* \), white markers correspond to observational data used to calibrate the model, density plots show values of the predictive posterior distribution and, finally, dashed black lines limit the enveloping region of all possible predictive scenarios for each variable. In order to assess the predictive power of our calibration we used the public database of July 24 and plotted data from June 18 to July 17 using blue markers. In this way, the blue, dashed vertical line separates the calibrated model (to its left) and the predictive model (to its right).

The first row of plots is fairly informative and it is worth discussing them in detail. We first notice that by using the posterior distribution of all parameters, model predictions spread fairly widely, with all possible epidemic curves encapsulated within the dashed black lines. Thus, the deterministic solution is very sensitive to parameter changes, which makes it unsuitable as a forecasting tool by itself. Secondly, density plots show that certain epidemic curves tend to accumulate in specific regions (that appear in dark). Interestingly enough, there is an increase in density symmetrically distributed above and below the deterministic curve at the beginning of the epidemic. It turns out that the increased density above the red solid line corresponds to the epidemic that would have resulted if no contingency plan had been implemented. The latter is indicated by a solid green line only on the first plot. We thus conclude that the contingency plan was rather successful in flattening the curve and shifting its peak to the right. In fact, from the parameters’ posterior distribution, one can show that before activating the contingency plan, the basic reproduction rate \( R_0 \) was 3.26, which was lowered to 1.60 after the plan was activated on March 23. Similarly, the increased density of curves below the deterministic line indicates what would have happened if mitigation had been more successful. We finally observe that the deterministic curve for daily new critically-ill patients obtained is above the data, suggesting that we are overestimating the total toll for deceased patients in our analysis. More importantly, it is necessary to point out that this discrepancy may be due to the difficulty in obtaining detailed compartmental data from the open database. In general, the open database contains records of the status of the patients taken during the epidemiological interview made at their admission as suspect cases, and it does not contain the specific dates of the change in their state during their hospital stay [18, 21, 22].

The colors of the middle row of figure 1 indicate the cumulative distribution function for newly infected, hospitalized, and critical cases per day. Thus, in this case, the color scale in the density plot goes from zero (white) to one (black). The solid black line on these three plots corresponds to the median curve, while the lower and upper curves (marked in solid red lines) are the 5% and 95% percentiles. In other words, the probability that all epidemiological curves generated by the calibrated model are comprised between the two solid red lines is 90%. Notice that one shortcoming of using only point estimators in compartmental models is that they yield epidemic curves which are fairly symmetric around their maximum, a feature that is not observed in actual data from many countries; instead, epidemic decline is slower than symmetric and fatter tails are predicted. However, using Bayesian statistics one can produce more realistic epidemic curves, with fattened tails, as can be appreciated in the median curves reported in the second row of figure 1.

In the first two rows of figure 1, as mentioned earlier, we also checked model predictions by comparing them to data from June 18 until July 17. The prediction is reasonable for daily new infected and daily new hospitalized cases, granted that in the calibration of the model we did not consider the progressive lifting of mitigation strategies currently taking place in Mexico City. We remark that our approach makes it possible to capture dynamic phenomena not captured by classical deterministic Kermack–McKendrick models, such as the plateau-like shape observed in the epidemic curve for Mexico City, lasting from May 20 to early August, approximately.
Table 1. Values of optimal parameters found by maximizing the parameters’ posterior distribution. Here $N_{\text{pop}}$ represents the population of Mexico City, which corresponds roughly to nine million individuals.

| Model’s parameters $\theta$ | $\theta \in [\theta_{a}^{\text{min}}, \theta_{a}^{\text{max}}]$ | Optimal parameters $\theta^*$ |
|-----------------------------|-----------------|----------------------------|
| $\tau_i$                   | [4, 9]          | 4.0 d                      |
| $\tau_f$                   | [4, 6]          | 5.99 d                     |
| $\tau_h$                   | [3, 15]         | 3.0 d                      |
| $\tau_c$                   | [3, 15]         | 3.0 d                      |
| $m$                        | [0.2, 0.8]      | 0.79                       |
| $c$                        | [0.2, 0.7]      | 0.69                       |
| $f$                        | [0.2, 0.8]      | 0.79                       |
| $R_0$                      | [0.8, 5]        | 3.26                       |
| $\gamma$                   | [0, 1]          | 0.49                       |
| $S(0)$                     | [0, 0.9 \times N_{\text{pop}}] | 111 569 susceptibles |

Figure 1. Top row: results for newly infected, hospitalized, and critical cases per day. In the first plot we indicate the meaning of each curve: the solid red line corresponds to the deterministic prediction with the mitigation plan; the solid green line is the deterministic prediction without mitigation. Dashed, solid lines limit the predictive posterior, while density plots give the actual value of the predictive posterior for the corresponding compartment in each plot; darker regions correspond to the accumulation of epidemic curves. Finally, white markers portray data for Mexico City used to calibrate, while blue markers correspond to data used for contrast against model predictions. Middle row: density plots correspond to the CDF for newly infected, hospitalized, and critical cases per day. Solid black lines indicate the median values, while solid red lines mark the 5% and 95% percentiles, respectively. In the middle figure of this row, an inset plot shows a cut of the CDF for a particular day. Bottom row: posterior predictive distribution of times at which the epidemic curves peak for daily new infected (left panel), hospitalized (middle panel), and critical (right panel) cases. Vertical lines indicate the date on which the peak would have occurred without mitigation. Note that the offset in dates among the three compartments in these plots can be explained largely as the mean time a patient takes to be hospitalized after getting infected and to deteriorate to a critical state once hospitalized.

We can similarly explore the posterior predictive distribution of dates on which the epidemic peak occurs. These are shown in the bottom row of figure 1 for newly infected, hospitalized and critically-ill patients per day, which were obtained according to equation (3). These distributions are again very informative: in all of them the vertical, dashed line indicates the date on which the epidemic would have peaked with no contingency measures. Interestingly enough, the support of the distribution of times is compact, meaning that one could provide a rather hard and robust interval within which the peak actually happens, though admittedly it is a wide range. We can also provide the mean date when the peak occurs. For instance, the mean date for new infected cases is June 2, with a standard deviation of 22 d
Figure 2. Density plots for the PDF (left panel) and the CDF (right panel) for the total number of deaths. In both cases, white markers correspond to the data used to calibrate the model while blue markers are data that provide a sense of the model’s forecasting ability (both regimes are separated by a vertical blue dashed line). The rest of the lines have the same meaning as those in figure 1.

Figure 3. Sobol’s total order indices estimated for our model. Top row: impact of ten parameters on the maximum value of daily new infected, hospitalized, critically ill, and deceased patients (from left to right). Bottom row: likewise for the time at which the maximum occurs. In both cases Sobol’s total indices are color coded and plotted as a function of time, which appears along the vertical axis.

which roughly coincides with the estimates of the 30 d epidemic peak interval centered around May 30 reported in [8]. One may argue that such a large standard renders predictions for the peak of the epidemic worthless. However, the total span of the first wave of the epidemic, until it finishes, is around nine months, while variability in our estimates is only half a month. A similar analysis follows for the other two posterior distributions for daily new hospitalized and critical cases.

In figure 2 we show model results for total death numbers. Unlike the data shown in figure 1, total death numbers contain inherently fewer fluctuations, since it is cumulative, not daily numbers, as those reported in figure 1, and this smoothens the graph somewhat. Other than that, we can see that, with respect to this time series, the model is not only well calibrated (white markers), but also the prediction obtained from June 18 until July 17 is remarkably good when compared to data (blue markers).

4. Global sensitivity analysis

To have a full picture of the model used to describe the evolution of COVID-19 in Mexico City, we performed GSA [23–25]. Briefly, this framework quantifies the sensitivity on an output (or several) of the

8 Here, a clarification is in order. Both databases used for calibration and prediction are not consistent with each other because of diagnostic and reporting delays. As new information is obtained, the database is updated, thus changing the previous values. The number of suspected deaths in the first one is therefore smaller than those in the second one. Given that model calibration used total death numbers augmented by half of those suspected, when checking model predictions only the number of confirmed deaths was used. This removes the artificial jump that would have been obtained otherwise.
model considering its parameters as input factors. As epidemiological compartmental models are fairly intuitive models we can foresee qualitatively what to expect from this type of analysis. For instance, we expect parameters $S(0), R_0$, and $\gamma$ to have the largest impact, regardless of the time window to which one applies GSA, while other parameters will be relevant only in certain time windows, e.g. $\tau_1$ and $\tau_2$ are relevant at the beginning of the epidemic curve, but hardly important when the peak has passed.

Thus, following [23, 24] we use a variance-based method due to Sobol. As model outputs we take the maximum values of daily new values of $(I(t), H(t), C(t), D(t))$ and the time at which they occur. We then estimate Sobol’s total order indices using as input parameters those found in the first column of table 1, which are allowed to vary within the interval given by the second column of the same table. As we expected, parameters have different impacts on model behaviour depending on phase of the epidemic; we estimated Sobol’s indices as a function of time $t \in [0, t_{\text{final}}]$ and show these results in figure 3. Finally, the Julia package DifferentialEquations.jl [26] was used to numerically estimate these indices, performing a Monte Carlo simulation over $10^5$ trajectories.

Results of the GSA for Sobol’s total indices are shown in figure 3 as density plots, in which the $x$-axis corresponds to the ten parameters of the model and the $y$-axis is the time $t_{\text{final}}$ on which the run ended. More precisely, the first row of figures shows total indices for the model output using the maximum values of the indices up to time $t_{\text{final}}$, while the second row corresponds to the time at which that maximum occurs. Looking at these results, and considering the type of compartmental model we use in this work, in which there are no long feedback mechanisms, we daresay that the results are fairly intuitive, in tune with what one would expect. Indeed, Sobol’s total indices for $\max_{t \leq t_{\text{final}}} I_{\text{daily}}(t)$ (first plot in the first row) suggest that parameters $\gamma, \tau_1, S_0, R_0$, and $\gamma$ play a significant role in model behaviour, while the others, which are relevant on later stages of the epidemic, play no role. This is precisely what makes this type of models difficult, if not impossible, to calibrate solely considering data from the beginning of an epidemic. The other plots in figure 3 can be interpreted in a similar manner.

5. Conclusions and future work

Contingency plans based on epidemiological models must be formulated, analysed and carried out very carefully. Even with fairly accurate observational data, the weight of stochasticity inherent to the start of an epidemic means that parameter estimates based on data from the beginning of an outbreak will be quite uncertain. In turn, models parametrized with such data will carry great uncertainty in long term forecasts. On the other hand, this uncertainty can be quantified using techniques from Bayesian statistics, which may then be used to consider worst-case scenarios.

Even though the model analysed here is simple, the main conclusion of this work is that extrapolating results without accounting for sensitivity to changes in parameters can result in predictions way off the mark. We believe that the same conclusion would hold for more detailed models, e.g., those which include specific details of the population, since most of them are also sloppy.

In terms of the mitigation measures implemented in Mexico City, our results show that they have so far managed to flatten the curve and shift the peak for newly infected cases per day to the right. However, this and other compartmental models are rather sensitive to parameter calibration. Access to richer data containing more epidemiological and clinical information would help to better control model predictions.

Control of the epidemic curve of SARS-CoV-2 in Mexico City requires evaluating the mitigation strategies that are implemented in the country. Mathematical models are central to this effort, but certain conditions need to be considered and evaluated for their efficient application. Mexico has the lowest testing rate among the OECD countries [5]. A high testing rate is recommended to adequately manage mitigation measures. Moreover, testing is necessary to estimate the true size of the epidemic. In Mexico, several hundreds of health units constitute the country’s sentinel surveillance system where cases are detected and traced to identify possible contacts and to gather other relevant information [21, 22]. A case detected by symptomatic surveillance has to be confirmed by testing, but, to obtain a concrete, workable estimation of the true dimension of the epidemic, tests must be widely applied to the general population, not only to suspect cases already detected by the surveillance system.

The positivity test rate for SARS-CoV-2 in the various municipalities of Mexico City was around 20%–40% on May 8, 2020 [6]. This high positivity rate and the limited number of tests performed to date, may prevent obtaining accurate estimates of both epidemic size and the actual growth rate of the epidemic, including the likely dates on which the epidemic curve starts to decline. Since testing is insufficient and, for the particular situation of the Mexican economy, increasing testing rates is unfeasible, mathematical modelling projections can help evaluate different scenarios that are consistent with the observed epidemic curve. Our model provides projections based on confirmed cases, corrected for under-reporting, that put the more likely dates of maximum incidence towards the end of May or early June, 2020. Earlier dates are possible, too, but with lower probability. These findings are important because the sanitary emergency measures
implemented in late March in Mexico City, were lifted on May 30, 2020. If our scenarios are correct, the risk of a new outbreak is high, given that the date for ending confinement would coincide with the dates predicted to be of maximum incidence. Since SARS-CoV-2 is a new virus, there is yet no significant herd immunity in the population. In Mexico, April 30 (Children’s Day) and May 10 (Mother’s Day) are significant dates for family gatherings and celebrations. The effect of these perturbations on the epidemic curve has been addressed [27] and their likely role on the initiation of the plateau-like stages mentioned before has been determined. However, our modelling approach allows for these considerations and for projecting the impact of mitigation or other interventions because of its probabilistic nature.

As for future work, there are a number of avenues we are currently exploring, both theoretical from the modelling side and practical, as a predictive tool. For instance, we will shortly explore the likely impact for Mexico City of increased population mobility resulting from lifting the sanitary emergency measures too soon. Clearly, we plan to extend this analysis to other regions of Mexico.

Acknowledgments

The authors are grateful for the support from UNAM-DGAPA-PAPIIT IV100220. RHM is grateful for the support of CONTEX project 2018-9B. JXVH acknowledges support from Grant UNAM-DGAPA-PAPIIT IN115720. GACS kindly acknowledges support from UNAM-DGAPA-PAPIIT IN114717. LB acknowledges support from UNAM-DGAPA-PAPIIT IG100819. We also thank SECTEI-CDMX for providing data on the evolution of COVID-19 in Mexico City. We thank William Lee and Héctor Benitez, UNAM, for their unwavering support during the elaboration of this work.

Appendix. Estimates from other countries

An important initial issue is how to define the prior distribution $P_0(\theta)$. For simplicity we consider a flat distribution in each of the parameters defining $\theta$, whose ranges we base in estimates made in other countries; see table A1.

In general, most countries have resorted to studies based on epidemic data from China, in order to understand and manage their outbreaks. Recent work [29] has reported that the mean duration from onset of symptoms to death is about 17.8 d (95% CI 16.9–19.2 d), and the mean duration from symptom onset to hospital discharge is 22.6 d (95% CI 21.1–24.4 d). The same paper has estimated the overall infection fatality rate for China to be 0.66% (0.39–1.33) with higher numbers for older ages. Also, they reported that the percentage of individuals likely to be hospitalized increased with age to a maximum of 18.4% (11.0–7.6). Mizumoto et al [33] report from their study of the Diamond Princess cruise ship a proportion of asymptomatic cases of 17.9% (95% CI 15.5–20.2), an estimate which is sensitive to the mean incubation period assumed. Nishiura et al [34], using data from Wuhan, China, estimate the asymptomatic ratio (percentage of carriers with no symptoms) to be higher, at 30.8% (95% CI 7.7–53.8). Wang et al [35] report a varying average daily attack rate per million people for Wuhan, depending on the stage of the epidemic: 2.2 (95% CI 2.0–2.4) before January 11, 44.9 (43.6–46.2) between January 11 and 22, 150.9 (148.3–153.5) between January 23 and February 1, and 54.1 (52.9–55.3) after February 2. However, differences in this rate were found to depend on age and risk group (healthcare professionals). As for the severity of the disease, these authors report 49.9% mild, 27.4% moderate and 19.7% severe, although they point out that this last percentage decreased gradually to reach 14.7% in the last phase of the epidemic.

Health authorities in Mexico are apparently estimating the overall attack rate to be well below the highest attack rate reached in China (Jan 23rd–Feb 1st) of 150.9 per million people. However, the percentage of asymptomatic cases is in accordance with Nishiura et al [34] estimate. The percentage of people requiring hospitalization is lower (9.8%) than the maximum of 18.4% reported (also for China) in [29].

### Table A1. Examples of estimated parameters from studies carried out in different countries.

| Parameter | Median | 95% credible interval or range |
|-----------|--------|-------------------------------|
| Infection → onset of symptoms $\tau_1$ | 5.1 d | 4.5–5.8 d |
| Onset of symptoms → death | 17.8 d | 16.9–19.2 d |
| Onset of symptoms → hospital discharge | 24.7 d | 22.9–28.1 d |
| Serial interval ($\approx \tau_1$) | 6.5 d | 3–8 (range) d |
| Prob. severe symptoms → ICU | 0.36 adults, 0.2 seniors | — |
| Hospitalized → $R$ rate | 0.072 adults, 0.022 seniors (1/d) | — |
| Hospitalized → $D$ rate | 0.0042 adults, 0.014 seniors (1/d) | — |
| ICU → $R$ rate | 0.05 adults, 0.036 seniors (1/d) | — |
| ICU → $D$ rate | 0.0074 adults, 0.029 seniors (1/d) | — |
| $R_0$ | 2.4 | 2–4.5 |
References

[1] World Health Organization 2020 Novel coronavirus (2019-nCoV) Situation Report 1 https://who.int/docs/default-source/coronaviruse/situation-reports/20200121-sitrep-1-2019-ncov.pdf?sfvrsn=20200121

[2] World Health Organization 2020 Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV) https://who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-nCoV)

[3] Unidad de Inteligencia Epidemiológica y Sanitaria 2020 Comunicado Técnico Diario Nuevo Coronavirus en el Mundo (COVID-19) https://www.gob.mx/salud/documentos/datos-abiertos-152127 (in Spanish)

[4] Dirección General de Epidemiología 2020 Informe Epidemiológico de la Situación de COVID-19 Mayo de 2020 https://oecd.org/coronavirus/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-nCoV)

[5] OCDE 2020 Testing for COVID-19: a way to lift confinement restrictions https://oecd.org/coronavirus/policy-responses/testing-for-covid-19-a-way-to-lift-confinement-restrictions-89736248/ (in Spanish)

[6] Gobierno de la Ciudad de México 2020 Datos abiertos https://datos.cdmx.gob.mx/explore/dataset/base-covid-sinave/table/ (in Spanish)

[7] Capistran M A, Capella A and Andres Christen J Forecasting hospital demand during COVID-19 pandemic outbreaks (arXiv:2006.01873)

[8] Acuña-Zegarra M A, Santana-Clibran M and Velasco-Hernandez J X 2020 Modeling behavioral change and COVID-19 containment in Mexico: a trade-off between lockdown and compliance Math. Biosci. 323 108370

[9] Gutenkunst R N, Waterfall J J, Casey F P, Brown K S, Myers E N, Jacobs-Austedal B, Alon U, Simon I, King A A 2007 Bounding parameter estimates and sensitivities in systems biology models PLoS Comput. Biol. 3 e189

[10] Castro M, Ares S, Cuesta J A and Manrubia S 2020 Predictability: Can the turning point and end of an expanding epidemic be precisely forecast? (arXiv:2004.09842)

[11] Maier B F and Brockmann D 2020 Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China Science 368 742–6

[12] Chiarelli M et al 2020 The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak Science 368 395–400

[13] Kucharski A J et al 2020 Early dynamics of transmission and control of COVID-19: a mathematical modelling study Lancet Infect. Dis. 20 553–8

[14] Kerrack W O and McKendrick A G 1927 A contribution to the mathematical theory of epidemics Proc. R. Soc. A 115 700–21

[15] Maslov S and Goldenfeld N 2020 Window of opportunity for mitigation to prevent overflow of ICU capacity in Chicago by COVID-19 (arXiv:2003.09564)

[16] Noll N B, Aksamitov I, Druelle V, Badenhorst A, Ronzani B, Jefferies G, Albert J and Neher R 2020 COVID-19 scenarios: an interactive tool to explore the spread and associated morbidity and mortality of SARS-CoV-2 medRxiv (https://doi.org/10.1101/2020.05.05.20091363)

[17] Beaumont M A, Zhang W and Balding D J 2002 Approximate Bayesian computation in population genetics Genetics 162 205–35

[18] Dirección General de Epidemiología, Secretaría de Salud 2020 Datos abiertos https://gob.mx/salud/documentos/datos-abiertos-152127 (in Spanish)

[19] Carranco-Sapién G A et al 2020 Mexico-COVID-19 https://github.com/carranco-sapien/Mexico-COVID-19

[20] Feldt R and Stukalov A 2018 BlackBoxOptim.jl https://github.com/robertfeldt/BlackBoxOptim.jl

[21] Dirección General de Epidemiología 2020 Lineamiento Estandarizado para la Vigilancia Epidemiológica y por Laboratorio de COVID-19 FEBRERO DE 2020 Technical Report Secretaría de Salud Gobierno de México (in Spanish)

[22] Dirección General de Epidemiología 2020 Lineamiento Estandarizado para la Vigilancia Epidemiológica y por Laboratorio de Enfermedad Respiratoria Viral ABRIL DE 2020 Technical Report Secretaría de Salud Gobierno de México (in Spanish)

[23] Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M and Tarantola S 2008 Global Sensitivity Analysis: The Primer (New York: Wiley)

[24] Pianosi F, Beven K, Freer J, Hill J W, Rougier J, Stephenson D B and Wagener T 2016 Sensitivity analysis of environmental models: a systematic review with practical workflow Environ. Modelling Softw. 79 214–32

[25] Servadio J L and Convertino M 2018 Optimal information networks: application for data-driven integrated health in populations Sci. Adv. 4 e1701088

[26] Rackauckas C and Nie Q 2017 DifferentialEquations.jl—a performant and feature-rich ecosystem for solving differential equations in Julia J. Open Res. Softw. 5 15

[27] Santana-Clibran M, Acuna-Zegarra M A and Velasco-Hernandez J X 2020 Flattening the curve and the effect of atypical events on mitigation measures in Mexico: a modelling perspective medRxiv (https://doi.org/10.1101/2020.05.21.20109678)

[28] Lauer S A, Grantz K H, Bi Q, Jones F K, Zheng Q, Meredith H R, Azman A S, Reich N G and Lessler J 2020 The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application Ann. Intern. Med. 172 577–82

[29] Verity R et al 2020 Estimates of the severity of COVID-19 disease medRxiv (https://doi.org/10.1101/2020.03.09.20033557)

[30] Ferguson N M et al 2020 Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand Imperial College COVID-19 Response Team 16-03-2020 (London: Imperial College London) (https://doi.org/10.25561/77482)

[31] Di Domenico L, Pullano G, Sabbatini C E, Boëlle P Y and Ferguson N M 2020 Report 13: Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries Imperial College COVID-19 Response Team 30-03-2020
[33] Mizumoto K, Kagaya K, Zarebski A and Chowell G 2020 Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020 *Eurosurveillance* **25** 2000180

[34] Nishiura H, Jung S M, Linton N M, Kinoshita R, Yang Y, Hayashi K, Kobayashi T, Yuan B and Akhmetzhanov A R 2020 The extent of transmission of novel coronavirus in Wuhan, China, 2020 *J. Clin. Med.* **9** 330

[35] Pan A et al 2020 Association of public health interventions with the epidemiology of the COVID-19 outbreak in Wuhan, China *JAMA* **323** 1915–23