

Data-driven inference of the reproduction number for COVID-19 before and after interventions for 51 European countries

Karnakov Petr^a, Arampatzis Georgios^a, Kičić Ivica^a, Wermelinger Fabian^a, Wälchli Daniel^a, Papadimitriou Costas^b, Koumoutsakos Petros^a

^a Computational Science and Engineering Laboratory, ETH Zurich, Switzerland

^b Department of Mechanical Engineering, University of Thessaly, Greece

Summary

The reproduction number is broadly considered as a key indicator for the spreading of the COVID-19 pandemic. Its estimated value is a measure of the necessity and, eventually, effectiveness of interventions imposed in various countries. Here we present an online tool for the data-driven inference and quantification of uncertainties for the reproduction number, as well as the time points of interventions for 51 European countries. The study relied on the Bayesian calibration of the SIR model with data from reported daily infections from these countries. The model fitted the data, for most countries, without individual tuning of parameters. We also compared the results of SIR and SEIR models, which give different estimates of the reproduction number, and provided an analytical relationship between the respective numbers. We deployed a Bayesian inference framework with efficient sampling algorithms, to present a publicly available graphical user interface (<https://cse-lab.ethz.ch/coronavirus>) that allows the user to assess and compare predictions for pairs of European countries. The results quantified the rate of the disease's spread before and after interventions, and provided a metric for the effectiveness of non-pharmaceutical interventions in different countries. They also indicated how geographic proximity and the times of interventions affected the progression of the epidemic.

Keywords: COVID-19, Bayesian inference, SIR model, interventions

Introduction

The forecasting of the evolution of the ongoing coronavirus disease 2019 (COVID-19) and the effects of non-pharmaceutical interventions are critical components for decision makers across the world. A broad range of data analysis tools and forecasting models have been deployed since the beginning of 2020 to assess the spread of the disease, as well as the expected number of infections and numbers of deaths [1–3]. A metric that is often deployed to quantify the progress of the disease is the reproduction

number, which is the expected number of secondary cases caused by a single infected individual over the infectious period. The beginning of the epidemics is characterised by the basic reproduction number R_0 , whereas the effective reproduction number R_t describes its progression in time. R_t exhibits significant complexity [4], but it is broadly considered that values of R_t above 1.0 indicate a rapid expansion of the infections. To reduce this number below 1.0, governments have applied various kinds of measures, such as social distancing, travel restrictions, closing of public places, schools and nonessential production. The estimation of R_t hinges on the forecasting model and the data used to infer their parameters [5]. Moreover, an often overlooked fact is that the method by which inference is made plays an important role in these estimates. A well-established technique for such inference is Bayesian inference [6]. However, even though Bayesian inference is a very potent method to estimate uncertainties in model structure and parameters, model parameters may be unidentifiable for the chosen type and amount of data, as well as the choice of priors [7, 8]. In particular for epidemic models, we believe that inferences for which the sampling algorithm has not converged may indicate a mismatch between the model and the data and need to be reported.

In this work we deployed Bayesian inference to quantify the evolution of R_t , as well as the time-points of interventions for 51 European countries. The study relied on the Bayesian calibration of the simple, well-established SIR model [9] extended to account for interventions, with data from reported daily infections. We present an online interface that allows for entry of customised data and comparisons between countries. The parameters of the model included the reproduction number, the day of the first intervention and a reduction factor for the reproduction number. By inferring these parameters from data, we identified when interventions became effective and determined the reproduction number before and after the measures had become effective. This indicates how well the imposed restrictions were able to slow down the spread of the disease.

Correspondence:

Petros Koumoutsakos,
Computational Science and
Engineering Laboratory,
ETH Zurich, Sälimstrasse
101, CH-8092 Zürich, petros[at]ethz.ch

Very recently a related work [10] inferred the intervention points and the evolution of R_t in Germany. In the present work we applied our model to 51 European countries. Our model was able to fit the data for the majority of the countries without individual tuning of the parameters or the inference method for each country. Our results are in excellent agreement with the work previously mentioned for Germany [10]. Moreover, our results indicated that capturing accurately the results for one country does not generalise well with the same accuracy to all countries. We visualised the inferred quantities on a map and ranked the countries by the disease's spread rate and the effectiveness of the interventions. We found that the results of our model were consistent with those from related studies in Switzerland [11]. The results suggest that at the current stage of the pandemic, countries can be categorised according to the inferred R_t . Most countries imposed their restrictions within the first weeks of the epidemic starting in March 2020, and since then about 80% of them have managed to bring their country specific R_t below 1.0, which indicates a decaying epidemic. Other countries, such as Armenia, Sweden and Moldova, remain at a stage with larger R_t .

Methods

Data: daily confirmed cases

We calibrated the well established SIR model using data from daily confirmed cases reported in the open source repository Humanistic GIS Lab, University of Washington [12], whose main data source for European countries is the World Health Organization [13]. To avoid the stochastic regime of the system, for each country we considered data only after the total number of confirmed cases exceeded five and the number of cases per million of population exceeded two. Population figures for countries are retrieved from the open repository [14]. The analysis in this paper considers data up to 18 May 2020. Official times of interventions that we use for comparison in the section "Inference of interventions for 51 European countries" are taken from reference [15], which aggregates the official beginning of non-pharmaceutical interventions by various governments.

Epidemic modelling using the SIR

The SIR model describes the evolution of susceptible (S), infected (I), and removed (R) population

$$\begin{aligned} \frac{dS}{dt} &= -\frac{R_t \gamma}{N} SI \\ \frac{dI}{dt} &= \frac{R_t \gamma}{N} SI - \gamma I \\ \frac{dR}{dt} &= \gamma I \end{aligned} \quad S(0) = N - I_0, \quad I(0) = I_0, \quad R(0) = 0, \quad (1)$$

where R_t is the reproduction number, γ is the removal rate (including recovery and mortality), I_0 is the initial number of infected individuals and N is the size of the population. To model an intervention or a series of measures taken by the government, we consider R_t to be a piece-wise linear function of time split into three phases

$$R_t(t) = \begin{cases} R_0, & t \leq t_{int} - \frac{1}{2} \delta_{int}, \\ \text{linear}, & t_{int} - \frac{1}{2} \delta_{int} \leq t \leq t_{int} + \frac{1}{2} \delta_{int}, \\ K_{int} R_0, & t \geq t_{int} + \frac{1}{2} \delta_{int}. \end{cases} \quad (2)$$

The three phases are: (i) uncontrolled disease outbreak before the first intervention takes place ($t \leq t_{int} - \frac{1}{2} \delta_{int}$), (ii) an adoption phase during which one or multiple measures take place ($t_{int} - \frac{1}{2} \delta_{int} \leq t \leq t_{int} + \frac{1}{2} \delta_{int}$), and (iii) the period after all interventions have become effective ($t \geq t_{int} + \frac{1}{2} \delta_{int}$).

In the first and last regime we assumed that the reproduction number R_t remained constant, and during the adoption phase we assumed a linear transition from R_0 to $k_{int}R_0$. The intervention time t_{int} , the reduction factor $k_{int} \in (0,1)$ and the duration of transition δ_{int} were inferred from the data.

Bayesian inference and sampling algorithms

We quantified the uncertainty in the extended SIR model, as described above, using data of daily confirmed cases for each country. We denoted this data with \hat{I}_i corresponding to day t_i . From the SIR model, the daily incidence was computed as

$$f(\vartheta, t_i) = \int_{t_i-1}^{t_i} \frac{R_t(\tau)\gamma}{N} S(\tau)I(\tau) d\tau = S(t_{i-1}) - S(t_i), \quad (3)$$

where $\vartheta = (R_0, \gamma, k_{int}, t_{int}, \delta_{int})$ are the model parameters of the SIR model including an intervention as described above. The initial number of infected individuals was set from the available data so that $I_0 = \hat{I}_0$. We consider the following generative model for the daily incidence data

$$\hat{I}_i \sim \text{NB}(f(\vartheta, t_i), g(\vartheta, t_i)), \quad (4)$$

where NB is the negative binomial parametrised by the mean $f(\vartheta, t_i)$ and the dispersion $g(\vartheta, t_i)$. Notice that using this parametrisation, the variance of an observation \hat{I}_i is given by $f(\vartheta, t_i) + f(\vartheta, t_i)^2/g(\vartheta, t_i)$. Here, the dispersion function is constant, i.e., $g(\vartheta, t) = r$, where the parameter vector has been extended to include the dispersion parameter r , $\vartheta = (R_0, \gamma, k_{int}, t_{int}, \delta_{int}, r)$.

We denoted the collection of observations by the vector $\hat{I} = (\hat{I}_1, \dots, \hat{I}_N)$ and the corresponding observation times by $t = (t_1, \dots, t_N)$. For the estimation of the probability of ϑ conditioned on the data, we apply Bayes' theorem,

$$p(\vartheta | \hat{I}; t) = \frac{p(\hat{I} | \vartheta; t)p(\vartheta)}{p(\hat{I}; t)} \quad (5)$$

where $p(\hat{I} | \vartheta; t)$ is the likelihood function, $p(\vartheta)$ is the prior probability distribution and $p(\hat{I}; t)$ is the model evidence.

Assuming the observations are independent, the likelihood function is given by

$$p(\hat{I} | \vartheta; t) = \prod_{i=1}^N \text{NB}(\hat{I}_i | f(\vartheta, t_i), g(\vartheta, t_i)) \quad (6)$$

We assumed uninformative priors for the model parameter $p(\vartheta) = \prod_{i=1}^N p(\vartheta_i)$. The lower and upper bounds of the priors are summarised in table 1. We assumed that the removal rate γ was the same in all countries and set it to $\gamma^{-1} = 5.2$ days following [16].

The inference of θ was performed using Korali [17, 18], a high-performance framework for uncertainty quantification and optimisation of computational models. The posterior distribution was estimated with Bayesian Annealed Sequential Importance Sampling (BASIS), a reduced variant of the Transitional Markov Chain Monte Carlo (TMCMC) algorithm [19], run with 5000 samples and default parameters taken from [20].

Graphical user interface for model predictions

We provide an online interface for real time evaluation of our predictive model. The interface allows for side-by-side comparison of two countries with default input data obtained from [12]. Modification of individual data points or substitution of different data sets can be achieved by changing the data in the input form. The accuracy of the Bayesian inference can be improved by increasing the number of samples used by the solver. User requests are forwarded from the web server to a remote workstation where the runs are executed. The back-end code of our model evaluation was implemented in Python and performance-critical parts were implemented in C++. The mean and confidence interval of the computed prediction for total and new daily infected cases were displayed together with input data. The inferred intervention date, t_{int} , was highlighted by the vertical line with corresponding reproduction number before and after the intervention. Figure 1 shows a screenshot of our online interface comparing the predictions for Switzerland and Sweden. Interventions introduced by the Swiss government starting on 17 March 2020 have been more successful in stopping the epidemic than the less strict approach of Sweden. The samples drawn from the posterior distribution are further shown in the columns of the selected countries.

Results and discussion

SIR model verification

We verified the presented intervention model using the results of a recent related study [10] that reported the results of a similar SIR model with interventions. The investigators performed Bayesian inference of the model parameters applied to Germany. That model had more parameters to described multiple interventions and considered a periodic modulation of the daily reported cases to account for trends such as under-reporting at weekends. We compared our results to this study in figure 2, where the reported quantities are the effective growth rate $\lambda^*(t) = \gamma(R_e(t)-1)$, the number of new daily cases and the total number of cases. The effective growth rate relates to one limiting case of the SIR model where the total number of infected individuals is small compared to the population such that $S/N \rightarrow 1$ and the model reduces to an exponential growth $I(t)$

Table 1: Parameters and the prior distributions for the model discussed in the section "Epidemic modelling using the SIR".

Parameter	Type	Value	Unit
R_0	Inferred	Uniform(1,4)	–
γ	Fixed	1/5.2 [16]	1/days
k_{int}	Inferred	Uniform (0, 1, 1)	–
t_{int}	Inferred	Uniform (0,80)	Days
$\bar{\delta}_{int}$	Inferred	Uniform (1,30)	Days
r	Inferred	Uniform (0,50)	–

Figure 1: Screenshot of the online interface comparing predictions for Switzerland and Sweden on 2 July 2020 with 2000 samples <https://cse-lab.ethz.ch/coronavirus>.

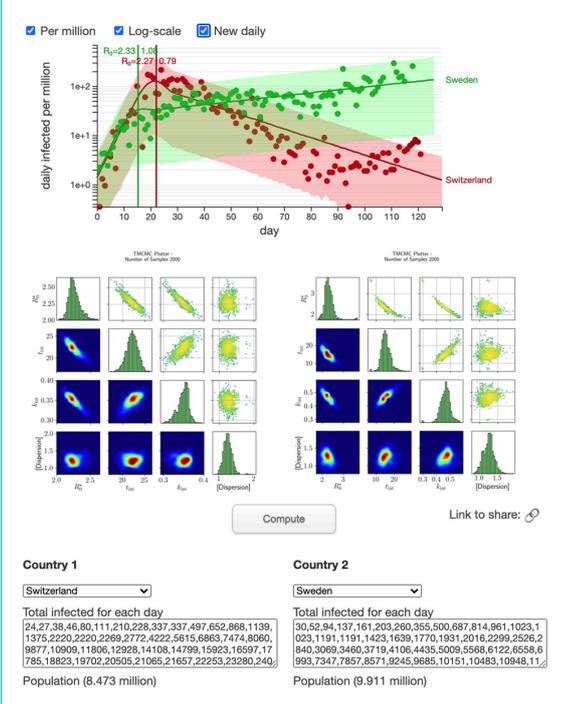
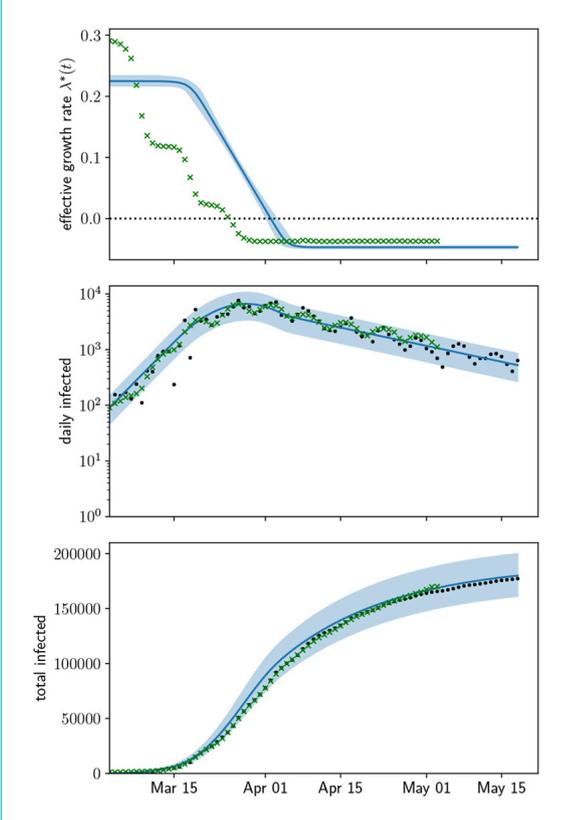


Figure 2: The effective growth rate (top), the number of new infections per day (middle) and the total number of infections in Germany (bottom) fitted by present model (blue lines and shades) compared to median predictions from [10] (green crosses). The solid lines show the mean prediction and the shades are the 90% confidence intervals.



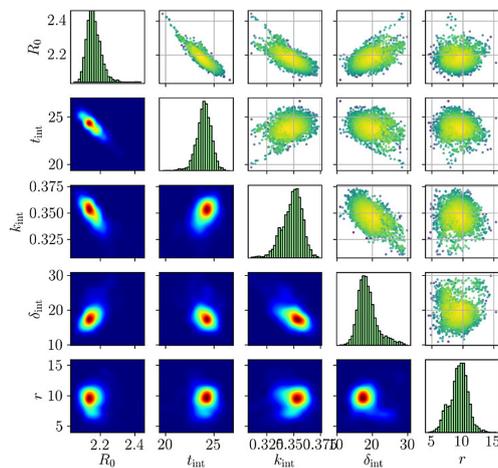
$\propto \exp(\lambda^* t)$, which is valid at early stages of the epidemic. Figure 3 shows the corresponding parameters drawn from the posterior distribution.

We want to highlight that our model was simpler and had fewer parameters than the one suggested in [10]. Also we observed that the fitted model agreed well with the reported data and provides an estimate for the intervention time in Germany.

Inference of interventions for 51 European countries

We applied our framework to 51 countries of the European continent and infer the model parameters for each country separately. The inferred reproduction number before and

Figure 3: Inferred parameters (the basic reproduction number R_0 , intervention time t_{int} , reduction factor after intervention k_{int} , duration of the transition δ_{int} and the dispersion parameter r of the negative binomial distribution [4] for Germany: samples drawn from the posterior distribution (upper triangle), marginal distributions of the individual parameter (diagonal) and likelihood heat map (lower triangle).



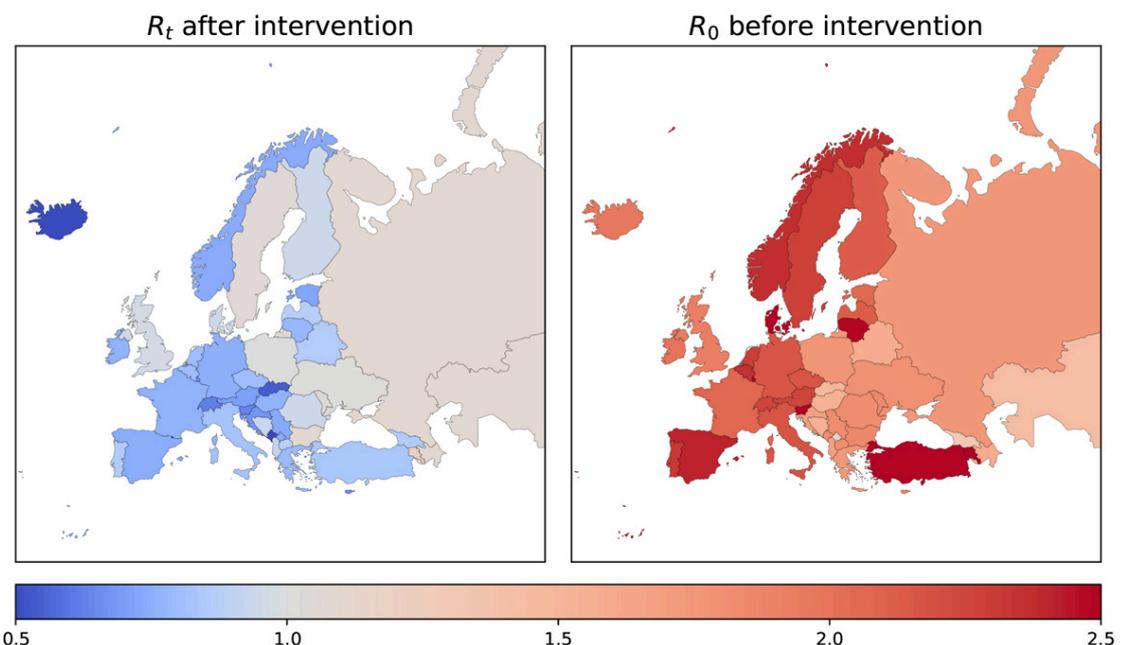
after intervention is shown in figures 4 and 5. The map and the scatter plot reveal the differences in the values of R_t and effects of interventions among various countries. Certain countries showed strong similarity in the inferred parameters. For example, the neighbouring Switzerland and Austria both have R_t dropping from 2.3 to 0.7 after interventions, even though the latter imposed more constraining measures. The same was evident for other pairs such as Italy and France, and also Poland and Ukraine. However, no single explanation for these similarities and differences can be suggested. Physical proximity of the countries did not always lead to similar progression of the epidemic. One example is Poland, where the epidemic started slower but the imposed restrictions proved less efficient than in adjacent Lithuania. The disease's development depended on the types of measures taken by governments, as well as the demographics and the particular features of the social interactions in the population. For instance, the UK and Sweden initially decided against imposing strict measures and therefore underwent a longer epidemic with more infections than their neighbors, Ireland and Norway, which corresponded to larger inferred values of R_t .

For Switzerland, one study with a more complex model [22] fitted on heterogeneous data for reported infections, hospitalisations and deaths, reported $R_0 = 2.8$ (95% confidence interval [CI] 2.1–3.8); our model gives a comparable estimate of $R_0 = 2.25$ (90% CI 2.16–2.37).

Another inferred parameter was the intervention time. Figure 6 shows the relation between the reproduction number and the intervention time measured from the beginning of the epidemic in each country. One trend is apparent from the values of R_0 before intervention: countries with larger R_0 tend to introduce the restrictions earlier.

The inferred times of the start of interventions $t_{int} - \frac{1}{2} \delta_{int}$ are shown in figure 7 and listed in table 2 together with the times when the restrictions were officially announced.

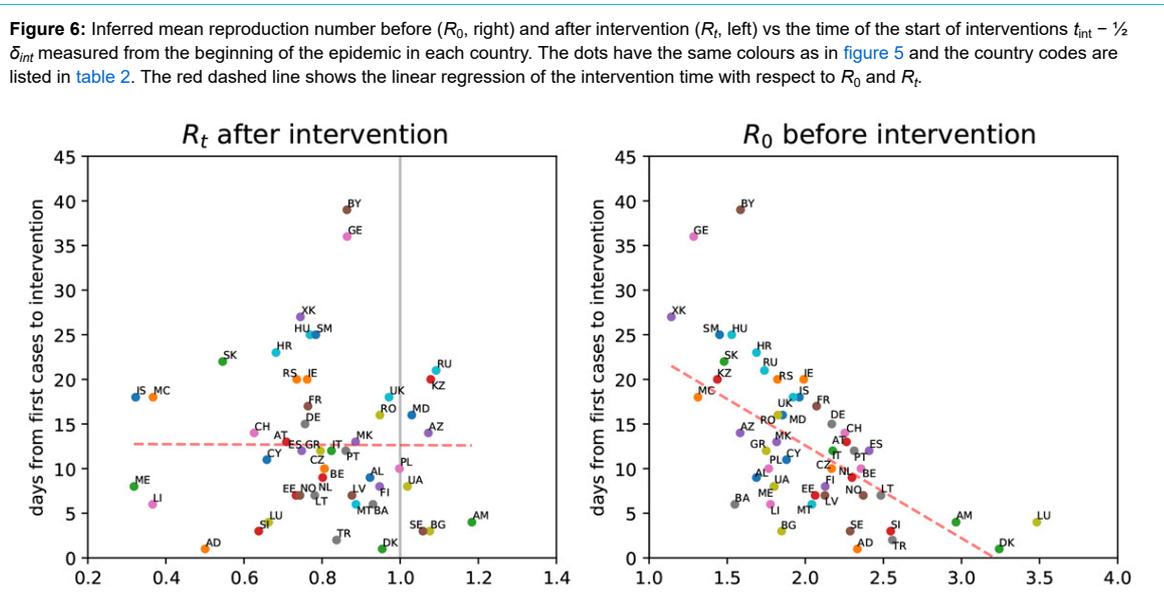
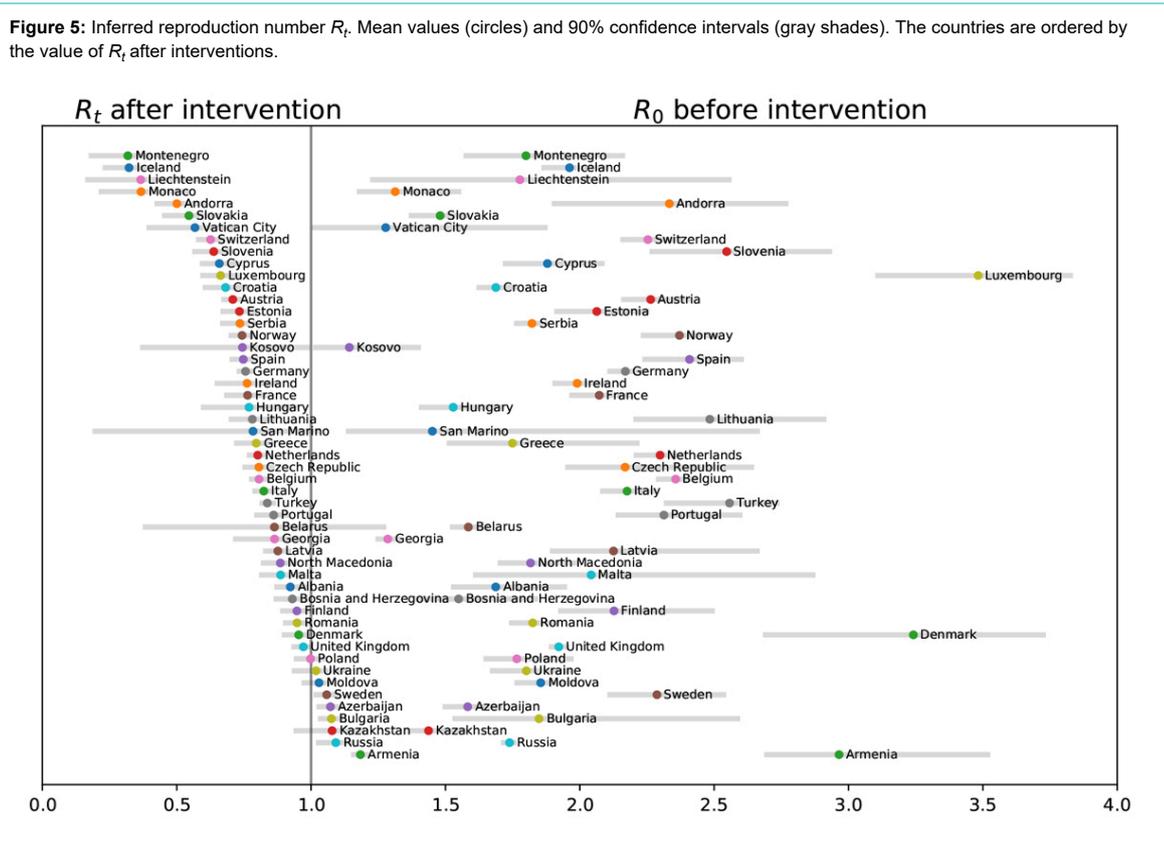
Figure 4: Inferred reproduction number R_t before (right) and after interventions (left) shown in color on a map of Europe [21].



The countries are ordered by the inferred intervention time, which helps to identify countries with similar inferred parameters. For certain countries, such as Switzerland, Germany and Greece, the inferred intervention time matched the first official announcements with a delay of at most one week. In other cases, the delays reached up to one month. This may indicate that the measures were imposed gradually and did not show an immediate effect.

A comparison of the SIR and SEIR models

A limitation of the SIR model is that its calibration is performed only on confirmed cases, which may be only a fraction of those infected in a certain population. More complex models also account for hospitalised cases and reported deaths. However, such data were not readily available for all 51 countries examined in this study. Nevertheless, we provide a comparison of the dynamics of the SIR model with the more complex SEIR model that also accounts for an incubation period, during which an infected individual is not yet infectious [23].



We used the data of infected individuals in Switzerland and the results indicated that the SIR and the SEIR model have very similar fits for the daily infections and exhibit similar dynamics. At the same time, the SEIR model resulted in a value of R_t larger than the one estimated by SIR. This difference can also be quantified by a simple relationship (equation 7) that hinges on the fact that while models have the same definition of R_t , they estimate different

growth rates of $I(t)$ and therefore of $N - S(t)$. Compared with SIR (equation 1), the SEIR model adds a parameter α

$$\frac{dS}{dt} = -\frac{R_t \gamma I S}{N},$$

$$\frac{dE}{dt} = \frac{R_t \gamma I S}{N} - \alpha E,$$

$$\frac{dI}{dt} = \alpha E - \gamma I,$$

Table 2: Inferred time of the start of interventions $t_{int} - \frac{1}{2} \delta_{int}$ (mean and standard deviation) compared to the official beginning of non-pharmaceutical interventions [15] in each country.

Country	Inferred	Official	Inferred - Official
Albania (AL)	2020-03-19 ± 5.5	2020-03-13	+6
Andorra (AD)	2020-03-17 ± 3.9		
Armenia (AM)	2020-03-16 ± 2.5	2020-03-24	-8
Austria (AT)	2020-03-16 ± 3.5	2020-03-16	+0
Azerbaijan (AZ)	2020-03-29 ± 4.5	2020-03-31	-2
Belarus (BY)	2020-04-21 ± 7.2		
Belgium (BE)	2020-03-14 ± 1.5	2020-03-18	-4
Bosnia and Herzegovina (BA)	2020-03-26 ± 5.1		
Bulgaria (BG)	2020-03-15 ± 5.3		
Croatia (HR)	2020-03-26 ± 3.4	2020-03-18	+8
Cyprus (CY)	2020-03-22 ± 4.2		
Czech Republic (CZ)	2020-03-18 ± 5.5	2020-03-16	+2
Denmark (DK)	2020-03-05 ± 4.6	2020-03-11	-6
Estonia (EE)	2020-03-13 ± 2.9		
Finland (FI)	2020-03-13 ± 4.3	2020-03-27	-14
France (FR)	2020-03-18 ± 3.2	2020-03-17	+1
Georgia (GE)	2020-04-11 ± 5.4	2020-03-31	+11
Germany (DE)	2020-03-18 ± 1.8	2020-03-23	-5
Greece (GR)	2020-03-17 ± 7.6	2020-03-23	-6
Hungary (HU)	2020-04-08 ± 7.3	2020-03-28	+11
Iceland (IS)	2020-03-20 ± 2.8		
Ireland (IE)	2020-03-25 ± 3.1	2020-03-12	+13
Italy (IT)	2020-03-06 ± 2.0	2020-03-09	-3
Kazakhstan (KZ)	2020-04-08 ± 8.4		
Kosovo (XK)	2020-04-21 ± 15.8	2020-03-14	+38
Latvia (LV)	2020-03-16 ± 4.7		
Liechtenstein (LI)	2020-03-22 ± 6.4		
Lithuania (LT)	2020-03-20 ± 3.3	2020-03-16	+4
Luxembourg (LU)	2020-03-15 ± 3.6	2020-03-18	-3
Malta (MT)	2020-03-17 ± 6.9		
Moldova (MD)	2020-03-30 ± 4.9		
Monaco (MC)	2020-04-02 ± 5.9		
Montenegro (ME)	2020-03-27 ± 4.3	2020-03-24	+3
Netherlands (NL)	2020-03-13 ± 1.5	2020-03-16	-3
North Macedonia (MK)	2020-03-23 ± 3.7		
Norway (NO)	2020-03-07 ± 2.1	2020-03-12	-5
Poland (PL)	2020-03-24 ± 5.0	2020-03-13	+11
Portugal (PT)	2020-03-20 ± 4.7	2020-03-19	+1
Romania (RO)	2020-03-27 ± 5.1	2020-03-25	+2
Russia (RU)	2020-04-11 ± 1.8	2020-03-28	+14
San Marino (SM)	2020-03-29 ± 20.1	2020-03-14	+15
Serbia (RS)	2020-03-31 ± 2.3	2020-03-15	+16
Slovakia (SK)	2020-04-03 ± 5.6	2020-03-16	+18
Slovenia (SI)	2020-03-09 ± 2.9		
Spain (ES)	2020-03-14 ± 4.2	2020-03-14	+0
Sweden (SE)	2020-03-06 ± 2.2		
Switzerland (CH)	2020-03-14 ± 1.8	2020-03-17	-3
Turkey (TR)	2020-03-21 ± 3.0	2020-04-23	-33
Ukraine (UA)	2020-04-01 ± 4.4	2020-03-17	+15
United Kingdom (UK)	2020-03-24 ± 1.6	2020-03-23	+1
Vatican City (VA)	2020-03-27 ± 5.9		

$$\frac{dI}{dt} = \gamma I.$$

Given that the total number of reported infections in most countries does not exceed 1%, the models are well approximated respectively by the linearised SIR

$$\frac{dS}{dt} = -R_t \gamma I,$$

$$\frac{dI}{dt} = (R_t - 1) \gamma I,$$

$$\frac{dR}{dt} = \gamma I,$$

and the linearised SEIR

$$\frac{dS}{dt} = -R_t \gamma I,$$

$$\frac{dE}{dt} = R_t \gamma I - \alpha E,$$

$$\frac{dI}{dt} = \alpha E - \gamma I,$$

$$\frac{dR}{dt} = \gamma I.$$

By solving these equations, we obtained two asymptotic growth rates for $I(t) \sim e^{\lambda t}$

$$\lambda^{SIR} = \gamma(R_t - 1)$$

$$\lambda^{SEIR} = -\frac{\alpha}{2} - \frac{\gamma}{2} + \frac{\sqrt{4R_t\alpha\gamma + \alpha^2 - 2\alpha\gamma + \gamma^2}}{2}$$

Fitting these linearised models to the same data implies equal growth rates, i.e., $\lambda^{SIR} = \lambda^{SEIR}$. In turn this leads to a simple analytical relationship between the respective R_t

$$R_t^{SEIR} = R_t^{SIR} \left(1 + \frac{(R_t^{SIR} - 1)\gamma}{\alpha} \right) \quad (7)$$

An example in figure 8 shows that both models can achieve the same growth rates before and after interventions, but for the SEIR model that requires a larger R_0 (2.25

for SIR and 3.82 for SEIR) and a smaller reduction factor k_{int} (0.27 and 0.124, respectively). Likewise, the inference with both models on data for Switzerland in figure 9 gives similar predictions and intervention times, but different estimates of R_0 : 2.25 for SIR and 4.0 for SEIR. For the incubation period in the SEIR, we used $\alpha^{-1} = 2.9$ days [24]. Results in figure 9 were obtained using the nested sampling algorithm [25] implemented in Korali.

Outlook

The present Bayesian inference framework integrated data from reported daily infection cases with a simple, well-established epidemiology model, to provide forecasts for the evolution of the disease with quantified uncertainties. Moreover, it provided data-driven estimates for the time-

Figure 8: Solutions of SIR and SEIR models. Parameters of SIR: $R_0 = 2.25$, $\gamma^{-1} = 5.2$ days, $t_{int} = 25$, $\delta_{int} = 25$, and $k_{int} = 0.27$. Parameters of SEIR: $R_0 = 3.82$, $\gamma^{-1} = 5.2$ days, $\alpha^{-1} = 2.9$ days [24], $t_{int} = 25$, $\delta_{int} = 25$, and $k_{int} = 0.124$. Values of R_0 and k_{int} were chosen to match the growth rates of both models before and after interventions following (7). Black dashed lines show the transition range $t_{int} \pm \frac{1}{2} \delta_{int}$.

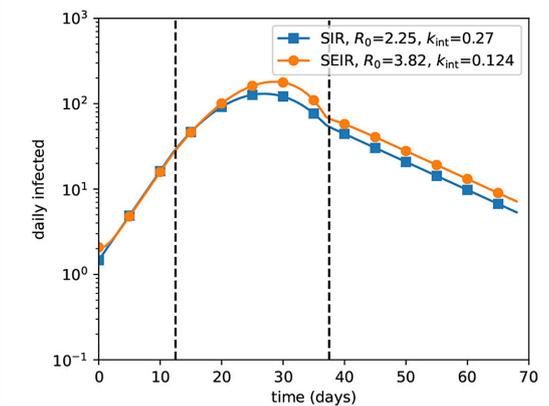


Figure 7: Inferred mean time of the start of interventions $t_{int} - \frac{1}{2} \delta_{int}$ (circles) with 90% confidence intervals (gray shades) compared with the official beginning (vertical bars) of non-pharmaceutical interventions [15] in each country. Empty circles indicate missing data for the official time. The countries are ordered by the inferred intervention time such that countries with similar parameters are located closer.

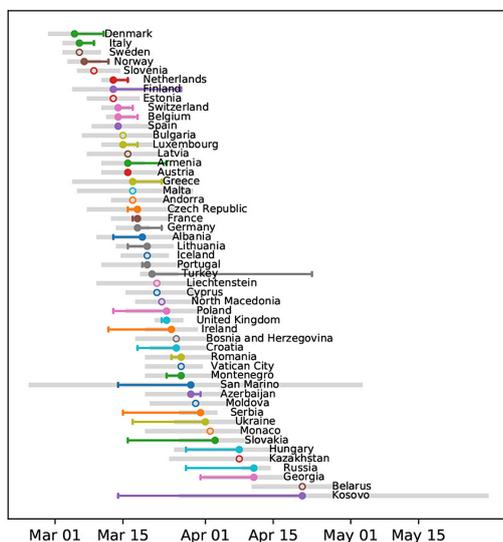
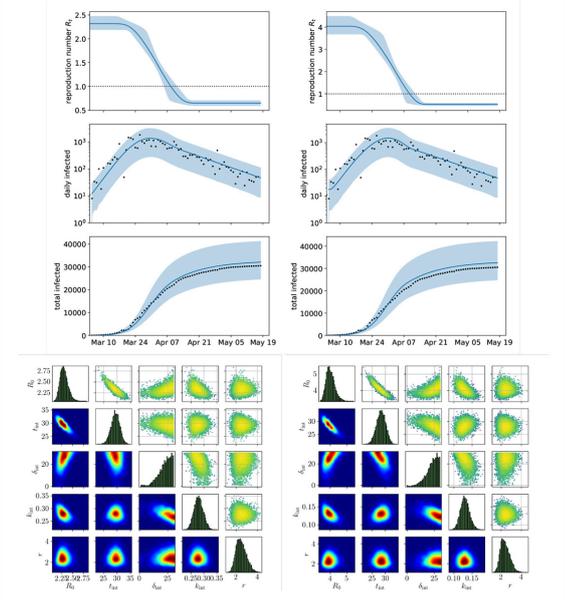


Figure 9: Inference with SIR (left) and SEIR (right) on data for Switzerland: predictions (top) and sampling (bottom). Inferred parameters are R_0 , t_{int} , δ_{int} and k_{int} . The incubation period in the SEIR model is 2.9 days.



point of interventions across different countries that are in agreement with the first announcement of interventions for a number of European countries. Although the model was able to infer accurately timepoints of interventions for some countries (for example, Austria, Azerbaijan, Denmark, Greece, France, Luxembourg, Norway, Switzerland), it did not capture the correct date for others (for example, Georgia, Ireland, Poland, Russia, Serbia). Understanding the success and failures of the model is a subject of our ongoing investigations. Such differences may lead to a better integration of models and data that may need to account for factors such as demographics, social interactions and geographical proximity between countries.

The present study relied on the SIR model using as the only type of data the reported infections. Models with more parameters fitted on heterogeneous datasets, such as hospitalisations and reported deaths, may provide a more detailed description of the epidemic at the cost of added complexity. At the same time, such data are not readily available for all countries examined in this study. Moreover, increasing the complexity of models does not necessarily lead to more useful predictions [26].

The present framework is modular and can be easily modified to incorporate advanced epidemiology models [11, 27], different types of data [28] and advanced sampling methodologies [29]. Our ongoing efforts include the handling of heterogeneous data to inform agent-based models for the evolution of COVID-19 as well as studies for optimal testing in order to increase the quantity and veracity of the data regarding the reported cases.

We believe that the present study indicated the need for Bayesian inference, with properly integrated data, models and sampling algorithms, to forecast of the evolution of COVID-19. Forecasts based on models with quantified uncertainties help to better assess the risk involved by the spreading of the disease and guide future interventions.

Disclosure statement

No financial support and no other potential conflict of interest relevant to this article was reported.

Acknowledgments

We acknowledge insightful discussions with Lucas Amoudruz, Martin Boden, Michalis Chatzimanolakis and Pascal Weber (ETHZ). Sergio Martin (ETHZ) provided valuable technical assistance with the Korali framework. We thank the anonymous Reviewers for their insightful and constructive comments. We gratefully acknowledge funding by ETH Zürich and computing resources by the Swiss Supercomputing Centre (CSCS).

References

- Riou J, Althaus CL. Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020. *Euro Surveill.* 2020;25(4):2000058. doi: <http://dx.doi.org/10.2807/1560-7917.ES.2020.25.4.2000058>. PubMed.
- Wu JT, Leung K, Bushman M, Kishore N, Niehus R, de Salazar PM, et al. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nat Med.* 2020;26(4):506–10. doi: <http://dx.doi.org/10.1038/s41591-020-0822-7>. PubMed.
- Zhao C, Tepekule B, Criscuolo NG, Wendel Garcia PD, Hilty MP, Fumeaux T, et al.; Risc-Icu Consortium Investigators In Switzerland. icumonitoring.ch: a platform for short-term forecasting of intensive care unit occupancy during the COVID-19 epidemic in Switzerland. *Swiss Med Wkly.* 2020;150:w20277. PubMed.
- Delamater PL, Street EJ, Leslie TF, Yang YT, Jacobsen KH. Complexity of the Basic Reproduction Number (R_0). *Emerg Infect Dis.* 2019;25(1):1–4. doi: <http://dx.doi.org/10.3201/eid2501.171901>. PubMed.
- Ridenhour B, Kowalik JM, Shay DK. Unraveling R(0): Considerations for Public Health Applications. *Am J Public Health.* 2018;108(Suppl 6):S445–54. doi: <http://dx.doi.org/10.2105/AJPH.2013.301704r>.
- Jaynes ET. *Probability Theory: The Logic of Science.* Bretthorst GL, editor. Cambridge, UK: Cambridge University Press; 2003.
- Beck JL. Bayesian system identification based on probability logic. *Struct Contr Health Monit.* 2010;17(7):825–47. doi: <http://dx.doi.org/10.1002/stc.424>.
- Simoen E, Papadimitriou C, Lombaert G. On prediction error correlation in Bayesian model updating. *J Sound Vibrat.* 2013;332(18):4136–52. doi: <http://dx.doi.org/10.1016/j.jsv.2013.03.019>.
- Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. *Proc R Soc Lond, A Contain Pap Math Phys Character.* 1927;115(772):700–21. doi: <http://dx.doi.org/10.1098/rspa.1927.0118>.
- Dehning J, Zierenberg J, Spitzner FP, Wibral M, Neto JP, Wilczek M, et al. Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science.* 2020;eabb9789. doi: <http://dx.doi.org/10.1126/science.abb9789>. PubMed.
- Althaus C. Real-time modeling and projections of the COVID-19 epidemic in Switzerland. Available at: <https://ispmbern.github.io/covid-19/swiss-epidemic-model/> [Cited 2020 May 18]
- Humanistic GIS Lab, University of Washington. COVID19: Novel Coronavirus Infection Map. 2020. Available at: <https://github.com/jakobzhao/virus> [Cited 2020 May 18]
- World Health Organization. Coronavirus disease (COVID-2019) situation reports; 2020. Available at: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/>.
- A simple data of the world by country each in JSON format. Available at: <https://github.com/samayo/country-json>. [Cited 2020 May 18]
- COVID-19 pandemic lockdowns. https://en.wikipedia.org/wiki/COVID-19_pandemic_lockdowns. [Cited 2020 May 18]
- Ganyani T, Kremer C, Chen D, Torneri A, Faes C, Wallinga J, et al. Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, March 2020. *Euro Surveill.* 2020;25(17). doi: <http://dx.doi.org/10.2807/1560-7917.ES.2020.25.17.2000257>. PubMed.
- Martin SM, Wa'Ichli D, Arampatzis G, Koumoutsakos P, Korali: A High-Performance Framework for Uncertainty Quantification and Optimization. Available at: <https://cse-lab.ethz.ch/korali/>. [Cited 2020 May 18]
- Wa'Ichli D, Martin SM, Economides A, Amoudruz L, Arampatzis G, Bian X, et al. Load Balancing in Large Scale Bayesian Inference. In: *Proceedings of the Platform for Advanced Scientific Computing Conference.* ACM; 2020.
- Ching J, Chen Y. Transitional Markov Chain Monte Carlo Method for Bayesian Model Updating, Model Class Selection, and Model Averaging. *J Eng Mech.* 2007;133(7):816–32. doi: [http://dx.doi.org/10.1061/\(ASCE\)0733-9399\(2007\)133:7\(816\)](http://dx.doi.org/10.1061/(ASCE)0733-9399(2007)133:7(816)).
- Wu S, Angelikopoulos P, Papadimitriou C, Moser R, Koumoutsakos P. A Hierarchical Bayesian Framework for Force Field Selection in Molecular Dynamics Simulations. *Philosophical Transactions A.* 2015;295:189–208.
- Natural Earth. http://naciscdn.org/naturalearth/packages/natural_earth_vector.zip. [Cited 2020 May 18]
- Lemaitre JC, Perez-Saez J, Azman AS, Rinaldo A, Fellay J. Assessing the impact of non-pharmaceutical interventions on SARS-CoV-2 transmission in Switzerland. *Swiss Med Wkly.* 2020;150:w20295. doi: <http://dx.doi.org/10.4414/sm.w.2020.20295>. PubMed.
- Carcione JM, Santos JE, Bagaini C, Ba J. A Simulation of a COVID-19 Epidemic Based on a Deterministic SEIR Model. *Front Public Health.* 2020;8:230. doi: <http://dx.doi.org/10.3389/fpubh.2020.00230>. PubMed.
- He X, Lau EHY, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat Med.* 2020;26(5):672–5. doi: <http://dx.doi.org/10.1038/s41591-020-0869-5>. PubMed.
- Feroz F, Hobson M, Bridges M. MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics. *Mon Not R Astron Soc.* 2009;398(4):1601–14. doi: <http://dx.doi.org/10.1111/j.1365-2966.2009.14548.x>.
- Siegenfeld AF, Taleb NN, Bar-Yam Y. Opinion: What models can and cannot tell us about COVID-19. *Proc Natl Acad Sci USA.* 2020;202011542. doi: <http://dx.doi.org/10.1073/pnas.2011542117>. PubMed.
- Noll NB, Aksamentov I, Druelle V, Badenhorst A, Ronzani B, Jefferies G, et al. COVID-19 Scenarios: an interactive tool to explore the spread and associated morbidity and mortality of SARS-CoV-2. *medRxiv.* 2020. Available at: <https://www.medrxiv.org/content/early/2020/05/12/2020.05.05.20091363>.

- 28 Krammer F, Simon V. Serology assays to manage COVID-19. *Science*. 2020;368(6495):1060–1. doi: <http://dx.doi.org/10.1126/science.abc1227>. PubMed.
- 29 Arampatzis G, Wa'Ichli D, Angelikopoulos P, Wu S, Hadjidoukas P, Koumoutsakos P. Langevin Diffusion for Population Based Sampling with an Application in Bayesian Inference for Pharmacodynamics. *SIAM J Sci Comput*. 2018;40(3):B788–811. doi: <http://dx.doi.org/10.1137/16M1107401>.