based on data from April 25, 2020

The reproduction number of COVID-19 and its correlation with public health interventions

Kevin Linka · **Mathias Peirlinck** · **Ellen Kuhl**

Received: May 1, 2020

Abstract Throughout the past four months, no number has dominated the public media more persistently than the reproduction number of COVID-19. This powerful but simple concept is widely used by the public media, scientists, and political decision makers to explain and justify political strategies to control the COVID-19 pandemic. Here we explore the effectiveness of political interventions using the reproduction number of COVID-19 across Europe. We propose a dynamic SEIR epidemiology model with a time-varying reproduction number, which we identify using machine learning and uncertainty quantification. During the early outbreak, the reproduction number was 4.5±21.4, with maximum values of 6.5 and 5.9 in Spain and France. As of today, it has dropped to 0.7±20.2, with minimum values of 0.4 and 0.3 in Austria and France. We found a strong correlation between passenger air travel and the reproduction number with a time delay of 12.6±22.7 days. Our new dynamic SEIR model provides the flexibility to simulate various outbreak control and exit strategies to inform political decision making and identify safe solutions in the benefit of global health.

Keywords COVID-19 · epidemiology · SEIR model · reproduction number · machine learning

1 Motivation

Since the beginning of the new coronavirus pandemic in December 2020, no other number has been discussed more controversially than the reproduction number of COVID-19 [29]. Epidemiologists use the basic reproduction number R_0 to quantify how many new infectious a single infectious individual creates in an otherwise completely susceptible population [10]. The public media, scientists, and political decision makers across the globe have started to adopted the basic reproduction number as an illustrative metric to explain and justify the need for community mitigation strategies and political intervention [16]: An outbreak will continue for $R_0 > 1$ and come to an end for $R_0 < 1$ [19]. While the concept of R_0 seems fairly simple, the reported basic reproduction number for COVID-19 varies hugely depending on country, culture, calculation, stage of the outbreak [29]. Knowing the precise number of R_0 is important, but challenging, because of limited data and incomplete reporting [9]. It is difficult–if not impossible–to measure R_0 directly [41]. The earliest COVID-19 study that followed the first 425 cases of the Wuhan outbreak via direct contact tracing reported a basic reproduction number of 2.2 [27]. However, especially during the early stages of the outbreak, information was limited because of insufficient testing, changes in case definitions, and overwhelmed healthcare systems [39]. Most basic reproduction numbers of COVID-19 we see in the public media today are estimates of mathematical models that depend critically on the choice of the model, the initial conditions, and numerous other modeling assumptions [9]. To no surprise, the basic reproduction number predicted by mathematical models covers a wide range, from 2.2–3.6 for exponential growth models to 4.1–6.5 for more sophisticated compartment models [29].

Compartment models are a popular approach to simulate the epidemiology of an infectious disease [23]. A prominent compartment model is the SEIR model that represents the timeline of a disease through the interplay of four compartments that contain the susceptible, exposed, infectious, and recovered populations [4]. The SEIR model has three characteristic parameters, the transition rates β from the suscep-

Kevin Linka · Mathias Peirlinck · Ellen Kuhl
Department of Mechanical Engineering, Stanford University, Stanford, California, United States

Ellen Kuhl
E-mail: ekuhl@stanford.edu

tible to the exposed state, α from the exposed to the infectious state, and γ from the infectious to the recovered state [19]. The latter two are disease specific parameters associated with the inverses of the latent period $A = 1/\alpha$ during which an individual is exposed but not yet infectious, and the infectious period $C = 1/\gamma$ during which an individual can infect others [26]. For COVID-19, depending on the way of reporting, these two times can vary anywhere between $A = 2$ to 6 days and $C = 3$ to 18 days [33,35,36]. The most critical feature of any epidemiology model is the transition from the susceptible to the exposed state. This transition typically scales with the size of the susceptible and infectious populations S and I, and with the contact rate β, the inverse of the contact period $B = 1/\beta$ between two individuals of these populations [19]. The product of the infectious period and the contact rate defines the reproduction number $R = C\beta$ [9]. Community mitigation strategies and political interventions seek to reduce the contact rate β–and with it the reproduction number R–to control the outbreak of a pandemic [36].

The first official case of COVID-19 in Europe was reported on January 24, 2020. Within only 45 days, the pandemic spread across all 27 countries of the European Union [12]. On March 17, for the first time in its history, the European Union closed all its external borders to prevent a further spreading of COVID-19 [13]. Within the following two weeks, many local governments supplemented the European regulations with lockdowns and national travel restrictions. In response, passenger air travel within the European Union dropped by up to 95% [14]. These drastic measures have stimulated a wave of criticism, especially because initially, it was entirely unclear to which extent they would succeed in reducing the number of new infections [31].

In this study, six weeks into these constraints, we correlate the effect of Europe-wide travel restrictions to the outbreak dynamics of COVID-19. We introduce a dynamic SEIR model with a time-varying contact rate $\beta(t)$ that transitions smoothly from the initial contact rate β_0 at the beginning of the outbreak to the current contact rate β_t under global travel restrictions and local lockdown. We express the time-varying contact rate $\beta(t) = R(t)/C$ as a function of the effective reproduction number $R(t)$ and use machine learning [1] to learn the evolution of the reproduction number for each country of the European Union from its individual outbreak history [12]. Our model allows us to precisely quantify the initial basic reproduction number, the reduced current effective reproduction number, and the time to achieve this reduction, which are important quantitative metrics of the effectiveness of national public health intervention. Our model also specifies the exact time delay between the implementation of political actions and their effects on the outbreak dynamics of COVID-19. This time delay is particularly important to plan exit strategies and estimate risks associated with gradually or radically relaxing current local lockdowns and global travel restrictions.

2 Methods

2.1 Epidemiology modeling

We model the epidemiology of the COVID-19 outbreak using an SEIR model with four compartments, the susceptible, exposed, infectious, and recovered populations, governed by a set of ordinary differential equations [19],

$$\begin{aligned}
\dot{S} &= -\beta SI \\
\dot{E} &= +\beta SI - \alpha E \\
\dot{I} &= \qquad\quad + \alpha E - \gamma I \\
\dot{R} &= \qquad\qquad\qquad + \gamma I\,.
\end{aligned}$$

The transition rates between the four compartments, β, α, and γ, are inverses of the contact period $B = 1/\beta$, the latent period $A = 1/\alpha$, and the infectious period $C = 1/\gamma$. We interpret the latency rate α and the infectious rate γ as disease-specific for COVID-19, and assume that they are constant across all 27 countries of the European Union. We interpret the contact rate $\beta = \beta(t)$ as behavior specific, and assume that it is different for each country and can vary in time to reflect the effect of societal and political actions. For easier interpretation, we express the contact rate $\beta(t) = R(t)/C$ in terms of the time-varying effective reproduction number $R(t)$. For the effective reproduction number, we make an ansatz of hyperbolic tangent form,

$$R(t) = R_0 - \tfrac{1}{2}[\,1 + \tanh([t - t^*]/b)\,][R_0 - R_\mathrm{t}]\,.$$

This ansatz ensures a smooth transition from the basic reproduction number R_0 at the beginning of the outbreak to the current reproduction number R_t under travel restrictions and lockdown, where t^* is the transition time and b is the transition speed.

2.2 Machine learning

To analyze the evolution of the effective reproduction number for each country, and predict possible exit strategy scenarios, we identify the initial exposed and infectious populations E_0 and I_0 and the effective reproduction number $R(t)$ using the reported COVID-19 cases in all 27 countries of the European Union [12]. For each country, our simulation window begins on the day at which the number of reported cases surpasses 100 individuals and ends on April 25, 2020. For all simulations, we select fixed latency and infectious periods of $A = 2.5$ days and $C = 6.5$ days [25, 27, 39]. To account for uncertainties in the initial exposed and infectious populations E_0 and I_0 and in the effective reproduction number

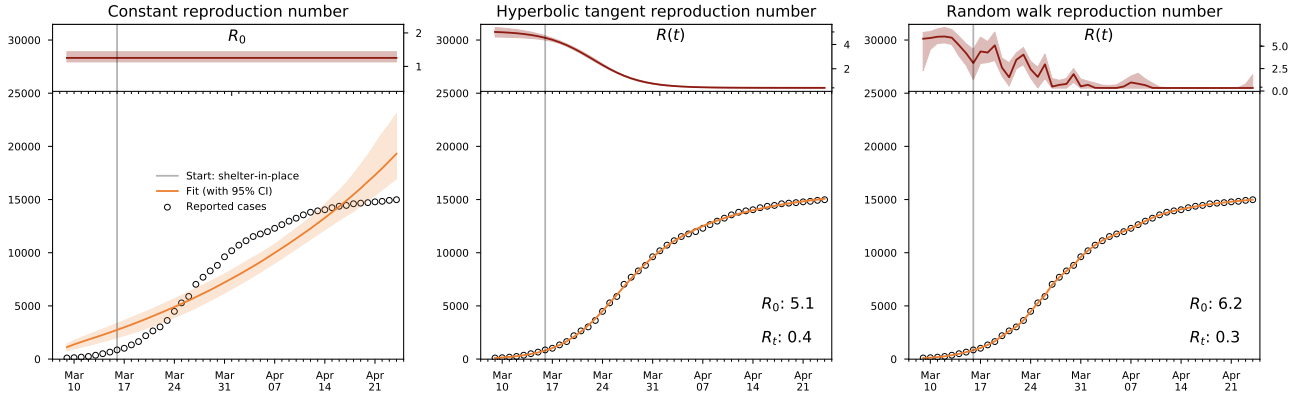

Fig. 1 Effects of time-varying effective reproduction number R(t). The constant effective reproduction number predicts an exponential increase in the number of cases that fits the initial but not for the later stages of the COVID-19 outbreak, left. The hyperbolic tangent reproduction number predicts a smooth early increase and later saturation of the number of cases, middle. A random walk reproduction number predicts a daily varying, non-smooth early increase and later saturation of the number of cases, right. Dots represent reported cases; orange curves illustrate fit with 95% confidence interval; red curves shows effective reproduction number with 95% confidence interval; here illustrated for the case of Austria.

$R(t)$, we use Bayesian inference with Markov-Chain Monte-Carlo to estimate the following set of model parameters $\vartheta = \{E_0, I_0, \sigma, R_0, R_t, t^*, b\}$. Here, σ represents the width of the likelihood $p(\hat{D}(t)\,|\,\vartheta)$ between the time-varying reported population $\hat{D}(t)$ and the simulated affected population $D(t,\vartheta)$. We adopt a Student's t-distribution for the likelihood between the data and the model predictions [8, 24] with a confirmed case number-dependent width,

$$p(\hat{D}(t)\,|\vartheta) \sim \text{StudentT}_{\nu=4}(\ \text{mean} = D(t,\vartheta),\ \text{width} = \sigma\sqrt{D(t,\vartheta)}).$$

We apply Bayes' rule to obtain the posterior distribution of the parameters [34,37] using the prior distributions in Tables 1 and 2 and the reported case numbers [12],

$$p(\vartheta\,|\,\hat{D}(t)) = \frac{p\left(\hat{D}(t)\,|\,D(t,\vartheta)\right)\,p(\vartheta)}{p\left(\hat{D}(t)\right)}.$$

We solve this distributon numerically using the NO-U-Turn sampler [20] implementation of the python package PyMC3 [38]. We use two chains. The first 500 samples are used to tune the sampler, and are later discarded. From the subsequent 1000 samples, we estimate the set of parameters ϑ. From the converged posterior distributions, we sample multiple combinations of parameters that describe the time evolution of reported cases. These posterior samples allow us to quantify the uncertainty on each parameter.

To probe the effect of different exit strategies, we explore three possible scenarios for the evolution of the effective reproduction number $R(t)$ for each posterior parameter sample set and predict the outbreak dynamics for the next month. The first scenario assumes a constant effective reproduction number $R(t) = R_t$, the second and third scenarios simulate the effect of a linear return from R_t to the country-specific basic reproduction number R_0, either rapidly within one month, or more gradually within three months.

Table 1 Prior distributions for the initial exposed and infectious populations E_0 and I_0.

Parameter	Distribution
E_0	LogNormal(log($D(t=A)$),1.5)
I_0	LogNormal(log($D(t=0)$),1.5)
σ	HalfCauchy($\beta = 1$)

Table 2 Prior distributions for the effective reproduction number $R(t)$ for three different scenarios.

$R(t)$ ... constant	
Parameter	**Distribution**
R_0	Normal(2.5,2)
$R(t)$ ... hyperbolic tangent	
Parameter	**Distribution**
R_0	Normal(2.5,2)
R_t	Normal(2.5,2)
t^*	Normal(10,10)
b	LogNormal(log(3),1.5)
$R(t)$ ... Gaussian random walk	
Parameter	**Distribution**
$R(t)$	GRW($\mu, \tau_1/(1.0-s)$)
μ	Normal(0,2)
τ_1	Exponential(1/2)
s	Uniform(0,1)

Figure 1 illustrates the effects of our our time-varying effective reproduction number, for three different cases, a static effective reproduction number that is equivalent to the basic reproduction number R_0, our smoothly decaying static effective reproduction number, and a daily varying effective reproduction number that follows a Gaussian random work with a drift μ and a daily stepwidth $\tau = \tau_1/(1.0-s)$, where τ_1 denotes the stepwidth precision and s the associated smoothing parameter. For the case of Austria, the three graphs illustrate the number of reported cases as dots, the model fit as orange curves with 95% confidence interval,

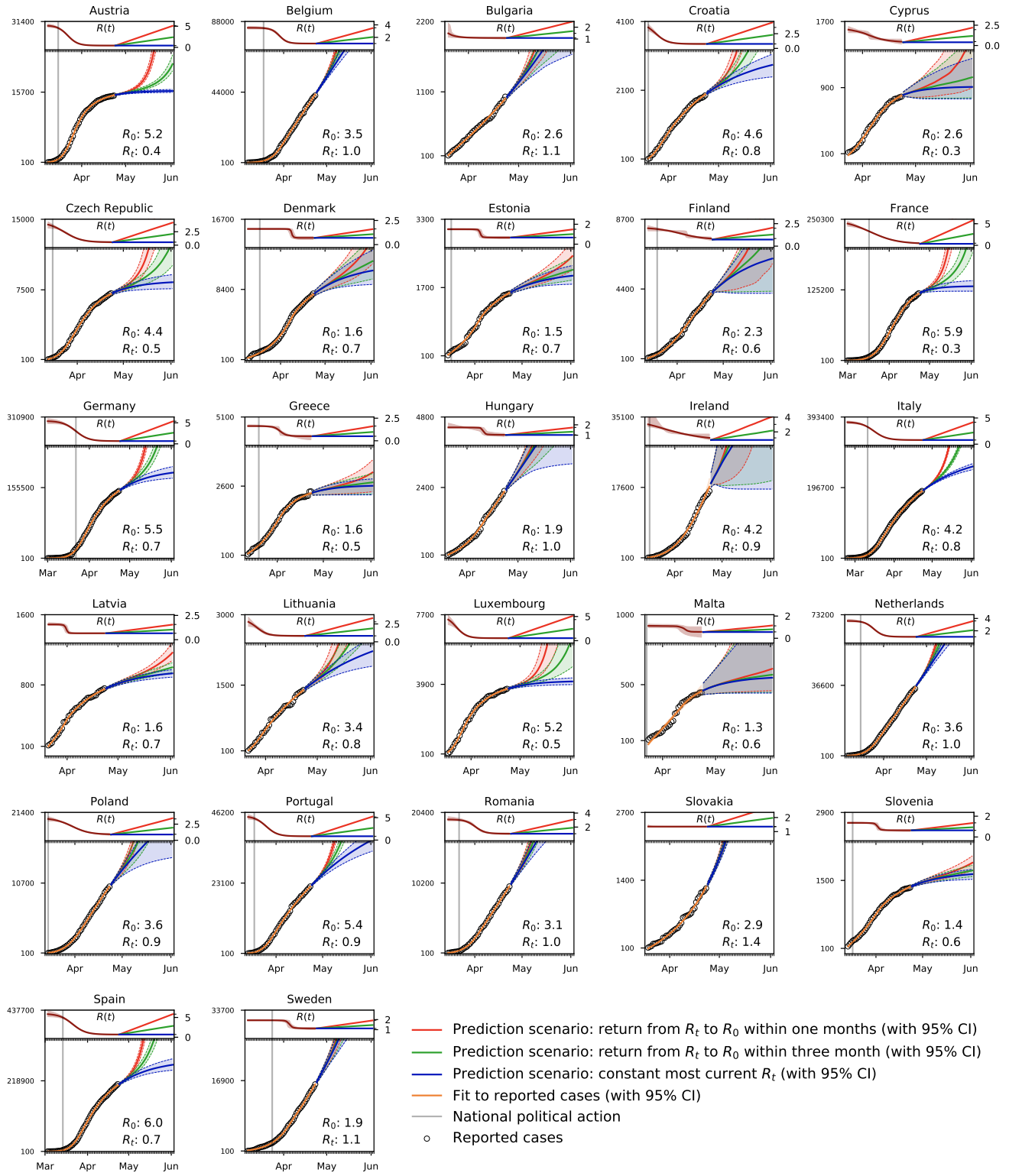

Fig. 2 Outbreak dynamics of COVID-19 across Europe and prediction of different exit strategies. Dots represent reported cases; orange curves illustrate fit of the SEIR model; red curves show effective reproduction number. Reported numbers are the basic reproduction number R_0 and the current effective reproduction number R_t. Red, green, and blue lines illustrate the projections for three possible exit strategies: a rapid and gradual return to the basic reproduction number R_0 within one month and three months, and a continuation of the current state at R_t; current date: April 24, 2020.

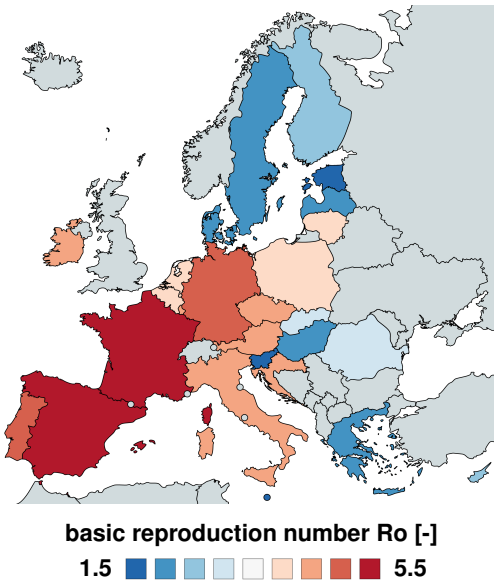

Fig. 3 Basic reproduction number R_0 of the COVID-19 outbreak across Europe. The basic reproduction number characterizes the number of new infectious created by one infectious individual at the beginning of the outbreak. It has maximum values in Spain, France, and Germany with 6.0, 5.9, and 5.5 and minimum values in Estonia, Slovenia, and Malta with 1.5, 1.4, and 1.3.

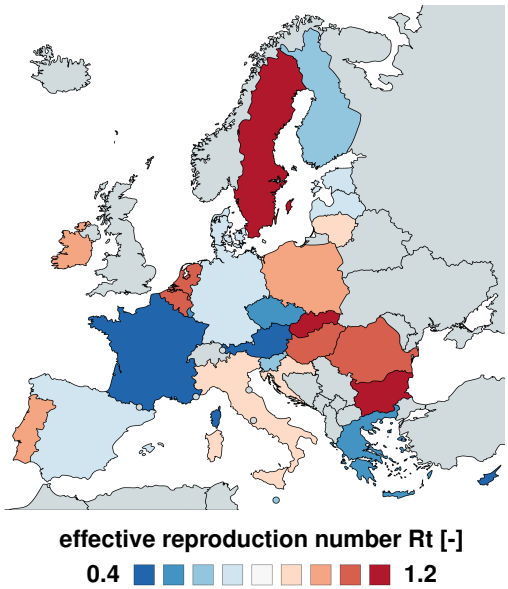

Fig. 4 Effective reproduction number R_t of the COVID-19 outbreak across Europe. The effective reproduction number characterizes the number of new infectious created by one infectious individual at the current stage of the outbreak. It has maximum values in Slovakia, Bulgaria, and Sweden with 1.4, 1.1, and 1.1 and minimum values in Austria, Cyprus, and France with 0.4, 0.3, and 0.3; current date: April 24, 2020.

and the effective reproduction numbers as red curves with 95% confidence interval. Of these three methods, the first method can fit the early exponential increase of the COVID-19 outbreak, but not the later saturation; the third method can fit both the early exponential increase and the later saturation, but not with a closed form expression. Only the second method based on the smooth hyperbolic tangent ansatz provides both a good fit and a closed functional form to compare the time lines of the outbreak in different countries and make informed predictions.

3 Results

Figure 2 illustrates the outbreak dynamics of COVID-19 for all 27 countries of the European Union. The dots represent reported cases, the orange curve illustrates the fit of the SEIR model with time-varying contact rate $\beta(t)$, and the red curves show effective reproduction number R_t. The vertical gray line highlights the beginning of political countermeasures. The graphs also report the basic reproduction number R_0 and the current effective reproduction number R_t of the current date, April 24, 2020. For the time period after that date, the red, green, and blue lines show the projected case numbers for a quick and slow return to the basic reproduction number R_0 and a continuation of the current state at R_t.

Table 3 and Figures 3 and 4 summarize the basic reproduction number R_0 at the beginning of the COVID-19 outbreak and the effective reproduction number R_t at the current date, April 24, 2020. The basic reproduction number R_0 has maximum values in Spain, France, and Germany with 6.0, 5.9, and 5.5 and minimum values in Estonia, Slovenia, and Malta with 1.5, 1.4, and 1.3. The population weighted mean of the basic reproduction number across the European Union is $R_0 = 4.50 \pm 1.44$. The current effective reproduction number R_t is significantly lower than the initial basic reproduction number R_0. In most countries, it is well below the critical value of $R_t = 1.0$. It has has maximum values in Slovakia, Bulgaria, and Sweden with 1.4, 1.1, and 1.1 and minimum values in Austria, Cyprus, and France with 0.4, 0.3, and 0.3. The population weighted mean of the basic reproduction number across the European Union is $R_t = 0.72 \pm 0.24$.

Figure 5 provides a direct correlation between the reduction in passenger air travel and the effective reproduction number of the COVID-19 outbreak across Europe. The black dots illustrate the reduction in passenger air travel, the red curves show effective reproduction number with 95% confidence interval. Spearman's rank correlation ρ, a measure of the statistical dependency between both variables, reveals the strongest correlation in the Netherlands and Sweden with 0.97 and 0.95, where the p-value of zero confirms a strong correlation. Only in Bulgaria and Slovakia, where the number of cases has not yet plateaued and the effective reproduction number does not show a clear smoothly decaying trend, there is no significant correlation between air travel and the effective reproduction number.

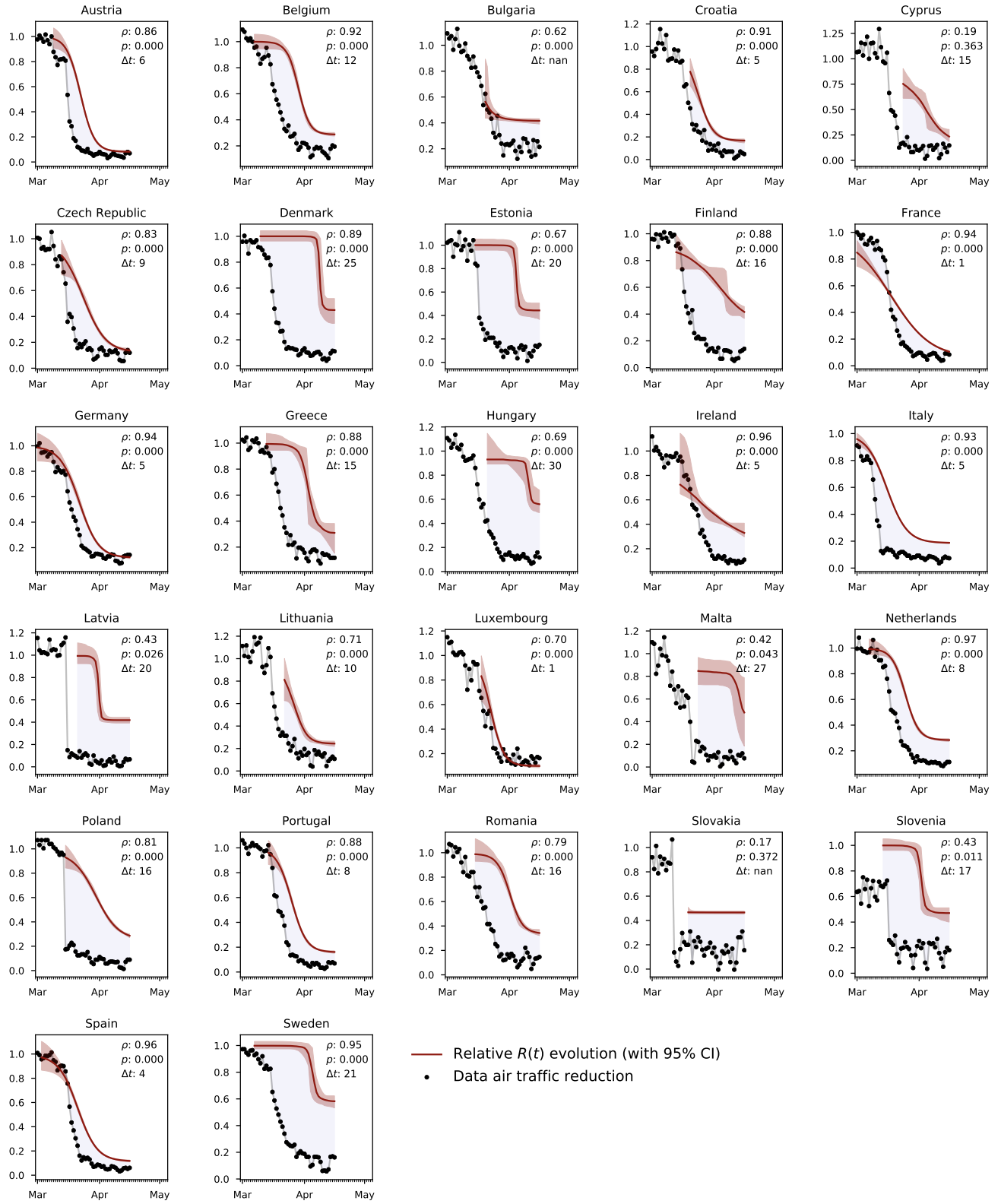

Fig. 5 Correlation between the reduction in passenger air travel and the effective reproduction number of the COVID-19 outbreak across Europe. Dots represent reduction in passenger air travel; red curves show effective reproduction number with 95% confidence interval. Spearman's rank correlation ρ and the p-value, measures of the statistical dependency between both variables, reveal the strongest correlation in the Netherlands and Sweden with 0.97 and 0.95. The time delay Δt highlights the temporal delay between the passenger air travel and the reproduction number curves; current date: April 24, 2020.

Table 3 Parameters of the COVID-19 outbreak across Europe. Basic reproduction number R_0, effective reproduction number R_t, adaptation time t^*, adaptation speed b, and time delay Δt for fixed latency period $A = 2.5$ days and infectious period $C = 6.5$ days.

Country	Population	R_0	R_t	t^*	b	Δt
Austria	8.840.521	5.21±0.26	0.42±0.02	13.47± 0.51	6.83± 0.54	6.45±14.47
Belgium	11.433.256	3.53±0.11	1.01±0.03	21.93± 0.48	5.60± 0.87	11.97±16.93
Bulgaria	7.025.037	2.63±0.76	1.07±0.07	5.35± 5.53	13.02±30.02	±
Croatia	4.087.843	4.60±0.84	0.76±0.05	4.01± 1.62	7.31± 1.41	5.18±23.01
Cyprus	1.189.265	2.59±1.00	0.32±0.37	9.31± 5.32	12.78± 9.87	15.41±40.31
Czech Republic	10.629.928	4.39±0.75	0.53±0.07	10.68± 2.31	11.13± 2.10	8.60±19.68
Denmark	5.793.636	1.62±0.03	0.69±0.08	29.59± 0.86	0.86± 0.79	24.51±30.59
Estonia	1.321.977	1.50±0.03	0.66±0.06	20.58± 0.86	1.11± 1.43	19.56±32.58
Finland	5.515.525	2.30±0.70	0.62±0.42	19.83± 6.20	19.06±16.01	15.87±25.83
France	66.977.107	5.94±0.77	0.34±0.10	16.91± 2.59	19.93± 2.57	0.95± 0.91
Germany	82.905.782	5.46±0.35	0.65±0.04	20.85± 0.95	9.54± 1.01	4.89± 0.15
Greece	10.731.726	1.64±0.06	0.47±0.14	21.08± 1.07	4.45± 2.75	15.25±26.08
Hungary	9.775.564	1.88±0.41	0.99±0.16	19.09± 4.63	4.52±10.04	29.70±43.09
Ireland	4.867.309	4.17±0.62	0.85±0.54	10.36± 5.21	25.11±11.63	4.64±23.36
Italy	60.421.760	4.24±0.15	0.79±0.01	20.89± 0.55	10.09± 0.58	4.56± 1.11
Latvia	1.927.174	1.58±0.08	0.66±0.02	10.42± 0.69	1.15± 0.98	20.09±34.42
Lithuania	2.801.543	3.43±0.81	0.83±0.05	4.53± 2.39	6.34± 2.15	9.57±31.53
Luxembourg	607.950	5.25±0.92	0.52±0.04	4.88± 1.22	6.40± 0.98	1.23±23.88
Malta	484.630	1.32±0.39	0.57±0.28	12.48±12.43	9.98±33.78	27.03±36.48
Netherlands	17.231.624	3.62±0.11	1.02±0.02	18.10± 0.46	6.66± 0.79	8.14±15.10
Poland	37.974.750	3.56±0.40	0.88±0.11	14.79± 1.75	11.63± 2.86	15.65±28.79
Portugal	10.283.822	5.37±0.39	0.85±0.03	11.51± 0.80	7.41± 0.81	7.62±20.51
Romania	19.466.145	3.12±0.16	1.05±0.06	17.12± 0.77	6.38± 1.49	16.10±24.12
Slovakia	5.446.771	2.90±0.92	1.35±0.02	14.55± 8.08	2.83± 5.39	±
Slovenia	2.073.894	1.36±0.03	0.63±0.04	19.16± 0.78	2.06± 1.70	17.14±29.16
Spain	46.796.540	6.00±0.45	0.68±0.03	18.03± 0.97	9.65± 0.79	4.07± 9.03
Sweden	10.175.214	1.94±0.03	1.12±0.05	29.11± 0.93	2.48± 1.89	21.03±18.11
European Union	446.786.293	4.50±1.44	0.72±0.24	17.54± 6.43	10.57± 5.35	12.61±22.67

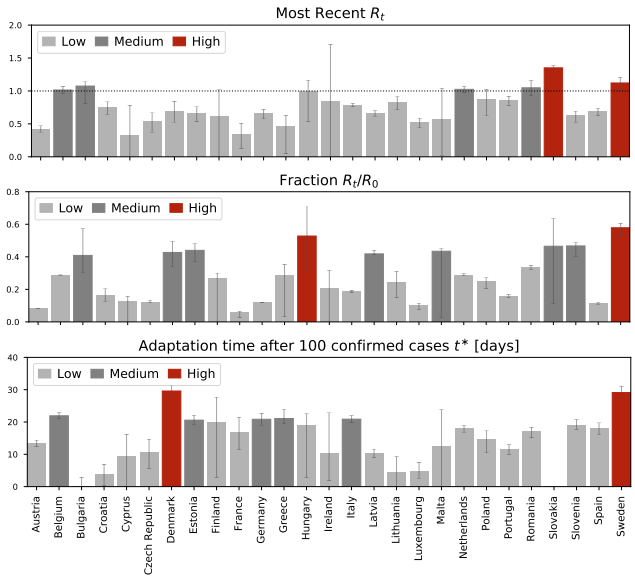

Fig. 6 Parameters of the COVID-19 outbreak across Europe. Effective reproduction number R_t, relative reduction in reproduction number R_t/R_0, and adaptation time t^*. The relative reduction in the reproduction number R_t/R_0 is largest in Sweden, Hungary, and Denmark with 0.58, 0.53, and 0.44, and smallest in Luxembourg, Austria, and France with 0.10, 0.08, and 0.05; current date: April 24, 2020.

Figure 6 summarizes the learnt effective reproduction number R_t, the relative reduction in the reproduction number R_t/R_0, and the adaptation time t^* for all 27 countries of the European Union. The relative reduction in the reproduction number R_t/R_0 is largest in Sweden, Hungary, and Denmark with 0.58, 0.53, and 0.44, and smallest in Luxembourg, Austria, and France with 0.10, 0.08, and 0.05. The mean of the relative reduction in the reproduction number across all 27 countries is $R_t/R_0 = 0.28 \pm 0.16$.

Figure 7 summarizes the time delay Δt between the reduction of air travel and reduction of the effective reproduction number in Figure 5. As such, it is a direct measure between cause and effect. The time delay has maximum values in Hungary and Malta with 30 and 27 days and minimum values in France and Luxembourg both with 1 day. The population-weighted of the mean time delay across the European Union is $\Delta t = 12.6 \pm 22.7$ days.

4 Discussion

Mathematical models can inform political interventions. As many countries begin to explore safe exit strategies from total lockdown, shelter in place, and national travel restrictions to manage the COVID-19 pandemic, political decision makers are turning to mathematical

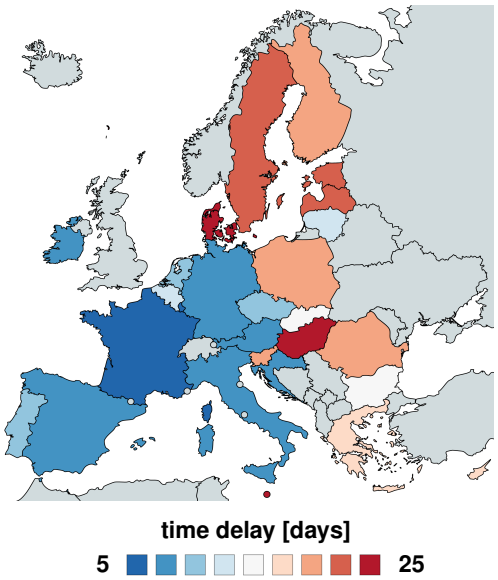

Fig. 7 Time delay Δt between reduction of air travel and reduction of the effective reproduction number across Europe. The time delay characterizes the temporal difference between the mobility and the reproduction number curves in Figure 5. It has maximum values in Hungary and Malta with 30 and 27 days and minimum values in France and Luxembourg both with 1 day.

models for advise [7]. A powerful quantitative concept to characterize the contagiousness and transmissibility of the new coronavirus is the basic reproduction number R_0 [41]. This number explains–in simple terms–how many new infections are caused by a single one infectious individual in an otherwise completely susceptible population [10]. However, against many false claims, the *basic* reproduction number does not measure the effects of public health interventions [9]. Here, we quantify these effects, for every point in time, for every country, using the *effective* reproduction number $R(t)$, a time-dependent metric that changes dynamically in response to community mitigation strategies and political actions. We learn the effective reproduction number from case data of the COVID-19 outbreak across Europe using machine learning and uncertainty quantification and systematically correlate it to political interventions.

The classical SEIR model can predict a natural equilibrium and herd immunity. The SEIR model has advanced to the model of choice for the outbreak dynamics of COVID-19 [29]. It belongs to a class of infectious disease models that epidemiologists characterize as compartment models [11]. Compartment models represent the population via a sequence of compartments through which the population passes as the disease progresses [46]. Out of the many different compartment models, the SEIR model seems best suited to mimic the epidemiology of COVID-19 via four compartments: the susceptible, exposed, infectious, and recovered populations. For more than three decades [4], epidemiologists have successfully applied the SEIR model to understand the outbreak dynamics of the measles, chickenpox, mumps, polio, rubella, pertussis, and smallpox [19]. For this class of diseases, the outbreak ends as the number of daily new cases, βSI, decreases. As such, the classical SEIR model is self-regulating: It naturally converges to an endemic equilibrium, at which either the susceptible group S, or the infectious group I, or both have become small enough to prevent new infections [26]. In epidemiology, this equilibrium is known as herd immunity [17]. In a homogeneous, well-mixed population, herd immunity occurs once a fraction of $(1 - 1/R_0)$ of the population has become immune, either through the disease itself or through vaccination. For the basic reproduction number of $R_0 = 4.50 \pm 1.44$ we found in this study, the herd immunity level would be 78%. This value is lower than 94% for the measles, 89% for chickenpox with, 86% for mumps and rubella, and 80% for polio [2], but significantly higher than the values of 16% to 27% for the seasonal flu [5]. Knowing the precise basic reproduction number of COVID-19 is therefore critical to estimate the conditions for herd immunity and predict the success of vaccination strategies.

The dynamic SEIR model can predict the effects of public health interventions. The classical SEIR model is a valuable tool to understand the interplay of the susceptible, exposed, infectious, and recovered populations under unconstrained conditions. However, for the current COVID-19 pandemic, similar to SARS, MERS, or Ebola, the dynamics of these four populations are tightly regulated by public health interventions [7]. This implies that model parameters like the contact rate β, the time it takes for an infectious individual to come into contact and infect others, are not constant, but modulated by social behavior and political action [3]. Here we explicitly account for a dynamic contact rate $\beta(t)$ and express it as a function of the time-varying effective reproduction number $R(t)$ [44]. This allows us to "bend the curve" and predict temporary equilibrium states, far away from the equilibrium state of herd immunity, but stable under current conditions [26]. Yet, these states can quickly become unstable again once the current regulations change. Our dynamic SEIR model allows us to study precisely these scenarios.

The time-varying effective reproduction number reflects the strength of public health interventions. To model temporal changes in the reproduction number, we propose a hyperbolic tangent type ansatz for the effective reproduction number $R(t)$. This functional form can naturally capture the basic reproduction number R_0, the converged reproduction number under the current constraints R_t, the time point between both conditions t^*, and the transition speed b. Figure 1 illustrates how

our hyperbolic tangent based model compares against a constant reproduction number and against a Gaussian random walk. The constant reproduction number in Figure 1, left, nicely captures the exponential increase at the early stages of the outbreak, but fails to "bend the curve" before herd immunity occurs. Nonetheless, several recent studies have successfully used an SEIR model with a constant reproduction number to model the outbreak dynamics of COVID-19 in China [35] and in Europe [28] by explicitly reducing the total population N to an affected population $N^* = \eta N$. The scaling coefficient $\eta = N^*/N$ is essentially a fitting parameter that indirectly quantifies the level of confinement [3]. For example, when averaged over 30 Chinese provinces, the mean affected population was $\eta = 5.19 \cdot 10^{-5} \pm 2.23 \pm 10^{-4}$, suggesting that the effect of COVID-19 was confined to only a very small fraction of the total population [35]. The Gaussian random walk in in Figure 1, left, naturally captures the effects of public health interventions, however, in a daily varying, rather unpredictable way. It is a valuable method to analyze case data retrospectively, but since it does not allow for a closed functional form, it is not very useful to make informed predictions. From Figure 1, we conclude that the hyperbolic tangent based ansatz with four physically meaningful parameters is the most useful approach to represent the time-varying effective reproduction number $R(t)$ for our current purposes.

Machine learning identifies basic and effective reproduction numbers from reported cases. Unfortunately, we can neither measure the basic nor the effective reproduction number directly. However, throughout the past four months, the COVID-19 pandemic has probably generated more quantitative data than any infectious disease in history. Machine learning offers incredible opportunities to evaluate these data and learn correlations and trends [32, 47]. Here we learn the effective reproduction number $R(t)$ directly from the reported COVID-19 cases in all 27 countries of the European Union, starting from the day of the first reported case on January 24, until today, April 25, 2020. This not only allows us to identify the model parameters and confidence intervals, but also to quantify correlations between travel restrictions and reduced effective reproduction numbers. Table 3 and Figures 2 to 4 summarize our basic reproduction numbers R_0 and the reduced effective reproduction numbers R_t for all 27 countries. Our mean basic reproduction number of $R_0 = 4.50 \pm 1.44$ exceeds the first estimates of 1.4 to 2.5 from the World Health Organization based on a tracing study that reported a value of 2.2 during the early outbreak in Wuhan [27]. However, our results agree well with the more recent values of 5.7 for the Wuhan outbreak [39] and with a recent review that suggested values from 4.1 to 6.5 calculated with SEIR models [29]. Our basic reproduction number of 4.5 is lower than the numbers of 18 for measles, 9 for chickenpox, 7 for mumps, 7 for rubella, and 5 for poliomyelitis [2]. Compared to the SARS coronavirus with a range from 2 to 5 [29], our values of SARS-CoV-2 in Table 3 are rather on the high end, suggesting that the new coronavirus would spread more rapidly than SARS. Knowing the precise basic reproduction number will be critical to estimate the number of contacts to trace, if we want to successfully control the dynamics of COVID-19 through contact trancing [18].

Political mitigation strategies reduce the effective reproduction number with a time delay of two weeks. Freedom of movement is the fundamental principle of the European Union. On March 13, 2020, the World Health Organization declared Europe the epicenter of the COVID-19 pandemic with more reported cases and deaths than the rest of the world combined [42]. To prevent a further spreading of the pandemic, four days later, for the first time in history, the European Union closed all its external borders [13]. In the following two weeks, the local governments augmented the European regulations with local lockdowns and national travel restrictions. Figure 5 shows that these measures had an enormous effect on the mobility within the European Union: By March 22, 2020, the average passenger air travel in Europe was cut in half, and as of today, April 25, it is reduced by 86% in Germany, 92% in France, 93% in Italy, and 95% in Spain [14]. These drastic actions have triggered an ongoing debate about the effectiveness of different outbreak strategies and the appropriate level of constraints [31]. Table 3 and Figures 5 to 7 summarize our time-varying effective reproduction number $R(t)$ and highlight the time delay of its reduction with respect to the European travel restrictions. An important socio-economical metric is mean time delay of $\Delta t = 12.6 \pm 22.7$ days between the effect of the current travel restrictions and the inflection point of the reproduction number curve. Figures 6 and 7 show that this time delay varies hugely across Europe with the fastest response of 1 day in France and Luxembourg, followed by Spain with 4 days, Germany, Ireland, and Italy with 5 days, and Austria with 6 days. These fast response times naturally also reflect decisions on the national level. France had the first reported COVID-19 case in Europe on January 24, 2020 and acted rigorously and promptly by introducing the first national measures on March 16. Similarly, Italy, Germany, and Spain had introduced their national measures on March 9, March 13, and March 9 [43]. Figures 6 and 7 clearly highlight the special role of Sweden, where the government focusses efforts on encouraging the right behavior and creating social norms rather than mandatory restrictions: The time delay of 21 days is one of the longest in the European Union, and Sweden is one of the few countries where the effective reproduction number

has not yet decreased below one. Taken together, these results confirm that, especially during the early stages of an outbreak, controlling mobility can play a critical role in spreading a disease [6]. This has stimulated an active ongoing debate when and how it would be safe to lift these restrictions.

Exit strategies will have different effects in individual countries. Political decision makers around the globe are currently trying to identify safe exit strategies from global travel restrictions and local lockdown. Mathematical models can provide guidelines and answer what-if scenarios. Our predictions in Figure 2 show projections of the number of total cases, for three different scenarios: a rapid release of all measures over a period of one month, a gradual release over a period of three months, or a maintenance of the current status quo. Naturally, the case numbers increase in all three cases, with the steepest increase for the most rapid return. Interestingly, our method provides significantly different confidence intervals for different countries suggesting that a controlled return will be more predictable in some countries like Austria and less in others. Our projections suggest that in Sweden, were policy makers had encouraged each individual to take responsibility for their own health rather than enforcing political constraints, the projected case numbers will follow the current curve, without major deviations.

Limitations. Just like any infectious disease model, our model inherently faces limitations associated with data uncertainties from differences in testing, inconsistent diagnostics, incomplete counting, and delayed reporting. For our specific study of COVID-19, we encounter a few additional limitations: First, although a massive amount of data are freely available through numerous well-documented public databases, the selection of the model naturally limits what we can predict and it remains challenging to map the available information into the format of the SEIR model. Second, the initial conditions for our exposed and infectious populations will always remain unknown and many new first cases have been reported throughout the past couple of weeks. To reduce the influence of unknown initial conditions, our machine learning algorithm learns these populations alongside the effective reproduction number. Third, in its current state, our model does not distinguish between community mitigation strategies, local public health recommendations, and global political actions. We are currently integrating the current approach into a global network model that will provide more granularity of individual mitigation strategies. Fourth, and probably most importantly, our current knowledge limits our ability to make firm predictions about the recovered group, which will be critical to estimate the return to normal. Recent studies have shown that there is a huge unreported asymptomatic population, up to an order of magnitude larger than the reported symptomatic population traced in our study. More targeted tests will be needed to identify the size of this population and explore whether it behaves differently in terms of contact rate and infectious period, which would both radically change the overall reproduction number. As more data become available, we are confident that we will learn from uncertainty quantification, become more confident in our model predictions, and learn how to quickly extract important trends.

5 Conclusion

We quantified the effectiveness of public health interventions using the effective reproduction number R, the time-varying reproduction number of the COVID-19 pandemic, across all 27 countries of the European Union. We adopted an SEIR epidemiology model with a dynamic effective reproduction number, which we learned for each country from its individual reported cases using machine learning and uncertainty quantification. We found that, during the early stages of the COVID-19 outbreak, the effective reproduction number across Europe was 4.5 ± 1.4. Today, on April 25, 2020, massive public health interventions have successfully reduced the effective reproduction number to 0.7 ± 0.2. Strikingly, this reduction displays a strong correlation with mobility in the form of air travel, with a time delay of 12.6 ± 22.7 days. This time delay is an important metric as we seek to identify safe exit strategies from the current lockdown and travel restrictions. To highlight the potential of our model, we predicted of the outbreak dynamics of COVID-19 for different exit strategies that either maintain the current status quo, gradually return to normal, or rapidly return to the early outbreak exponential growth. We showed that our dynamic epidemiology model provides the flexibility to simulate the effects and timelines of various outbreak control and exit strategies to inform political decision making and identify solutions that minimize the impact of COVID-19 on global health.

Acknowledgements We acknowledge stimulating discussions with Dr. Francisco Sahli Costabal. This work was supported by a DAAD Fellowship to Kevin Linka and a Stanford Bio-X IIP seed grant to Mathias Peirlinck and Ellen Kuhl.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. M. Alber, A. Buganza Tepole, W. Cannon, S. De, S. Dura-Bernal, K. Garikipati, G. Karniadakis, W.W. Lytton, P. Perdikaris, L. Petzold, E. Kuhl. Integrating machine learning and multiscale modeling: Perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences. npj Digital Medicine; 2019; 2:115.
2. R. M. Anderson, R. M. May. Directly transmitted infectious diseases: control by vaccination. Science 215 (1982) 1053-1060.
3. A. Arenas, W. Cota, J. Gomez-Gardenes, S. Gomez, C. Granell, J.T. Matamalas, D. Soriano-Panos, B. Steinegger. Derivation of the effective reproduction number R for COVID-19 in relation to mobility restrictions and confinement. medRxiv (2020) doi:10.1101/2020.04.06.20054320.
4. J. L. Aron, I. B. Schwartz. Seasonality and period-doubling bifurcation in an epidemic model. J. Theor. Bio 110 (1984) 665-679.
5. M. Biggerstaff, S. Cauchemez, C. Reed, M. Gambhir, L. Finelli. Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: a systematic review of the literature. BMC Infectious Disease 14 (2014) 480.
6. M. Chinazzi, J. T Davis, M. Ajelli, C. Gioanni, ... A. Vespignani. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science (2020) doi:10.1126/science.aba9757.
7. S. Cobey. Modeling infectious disease dynamics. Science (2020) doi:20.1126/science.aba5659.
8. J. Dehning, J. Zierenberg, F.P. Spitzner, M. Wibral, J. Pinheiro Neto, M. Wilczek, V. Priesemann, Inferring COVID-19 spreading rates and potential change points for case number forecasts. arXiv (2020) 2004.01105.
9. P. L. Delamater, E. J. Street, T. F. Leslie, Y. T. Yang, K. H. Jacobsen. Complexity of the basic reproduction number (R_0). Emerg. Infect. Disease 25 (2019) 1-4.
10. K. Dietz. The estimation of the basic reproduction number for infectious diseases. Stat. Meth. Med. Res. 2 (1993) 23-41.
11. P van den Diessche, J. Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartment models of disease transmission. Math. Biosci. 180 (2002) 29-48.
12. European Centre for Disease Prevention and Control. Situation update worldwide. `https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases.` assessed: April 13, 2020.
13. European Commission. COVID-19: Temporary restriction on non-essential travel to the EU. Communication from the Commission to the European Parliament, the European Council and the Council. Brussels, March 16, 2020.
14. Eurostat. Your key to European statistics. Air transport of passengers. `https:// ec.europa.eu/eurostat`; assessed: April 13, 2020.
15. Y. Fang, Y. Nie, M. Penny. Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: a data-driven analysis. J. Med. Virol. (2020) 1-15.
16. A.S. Fauci, H.C. Lane, R.R. Redfield. Covid-19–Navigating the uncharted. New Engl. J. Med. (2020) 382:1268-1269.
17. P. E. M. Fine. Herd immunity: history, theory, practice. Epidemiologic Reviews 15 (1993) 265-302.
18. J. Hellewell, S. Abbott, A. Gimma, N.I. Bosse, C.I. Jarvis, T.W. Russell, J.D. Munday, A.J. Kucharski, W.J. Edmunds. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. Lancet Global Health (2020) 8:e488-496.
19. H. W. Hethcote. The mathematics of infectious diseases. SIAM Review 42 (2000) 599-653.
20. M.D. Hoffman, A. Gelman. The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. J. Machine Learning Res. 15 (2014) 1593–1623.
21. J. Hsu. Here's how computer models simulate the future spread of new coronavirus. Scientific American (2020) February 23, 2020.
22. D. S. Hui, E. I. Azhar, T. A. Madani, F. Ntoumi, R. Kock, O. Dar, G. Ippolito, T. D. Mchogh, Z. A Memish, C. Drosten, A. Zumla, E. Petersen. The continuing 2019-nCoV epidemic thread of novel coronaviruses to global health - The latest 2019 novel coronavirus outbreak in Wuhan, China. Int. J. Infect. Diseases 91 (2020) 264-266.
23. W. O. Kermack, G. McKendrick. Contributions to the mathematical theory of epidemics, Part I. Proc. Roy. Soc. London Ser. A 115 (1927) 700-721.
24. K.L. Lange, R.J.A. Little, J.M.G. Taylor, Robust statistical modeling using the T distribution. J. Am. Stat. Ass. 84 (1989) 881-896.
25. S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, J. Lessler. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Ann. Int. Med. (2020) doi:10.7326/M20-0504.
26. M. Y. Li, J. S. Muldowney. Global stability for the SEIR model in epidemiology. Math. Biosci. 125 (1995) 155-164.
27. Q. Li, X. Guan, P. Wu, X. Wang, ... Z. Feng. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. New Eng. J. Med. (2020) doi:10.1056/NEJMoa2001316.
28. K. Linka, M. Peirlinck, F. Sahli Costabal, E. Kuhl. Outbreak dynamics of COVID-19 in Europe and the effect of travel restrictions. Comp Meth Biomech Biomed Eng; in press. doi:10.1080/10255842.2020.1759560; medRxiv doi:10.1101/2020.04.18.20071035v1.
29. Y. Liu, A. A. Bayle, A. Wilder-Smith, J. Rocklov. The reproductive number of COVID-19 is higher compared to SARS coronavirus. J. Travel Medicine (2020) 1-4.
30. B. F. Maier, D. Brockmann. Effective containment explains sub-exponential growth in confirmed cases of recent COVID-19 outbreak in mainland China. medRxiv (2020) doi:10.1101/2020.02.18.20024414.
31. B. Mason Meier, R. Habibi, Y. Tony Yang. Travel restrictions violate international law. Science 367 (2020) 1436.
32. G.C.Y. Pang, M. Alber, A. Buganza Tepole, W. Cannon, S. De, S. Dura-Bernal, K. Garikipati, G. Karniadakis, W.W. Lytton, P. Perdikaris, L. Petzold, E. Kuhl. Multiscale modeling meets machine learning: What can we learn? Arch. Comp. Meth. Eng. (2020) doi:10.1007/s11831-020-09405-5.
33. S. W. Park, B. M. Bolker, D. Champredon, D. J. D. Earn, M. Li, J. S. Weitz, B. T. Grenfell, J. Dushoff. Reconciling early-outbreak estimates of the basic reproductive number and its uncertainty: framework and applications to the novel coronavirus outbreak. medRxiv (2020) doi:10.1101/2020.01.30.20019877.
34. M. Peirlinck, F. Sahli Costabal, K.L. Sack, J.S. Choy, G.S. Kassab, J.M. Guccione, M. De Beule, P. Segers, E. Kuhl. Using machine learning to characterize heart failure across the scales. Biomech. Model. Mechanobio. 18 (2019) 1987-2001.
35. M. Peirlinck, K. Linka, F. Sahli Costabal, E. Kuhl. Outbreak dynamics of COVID-19 in China and the United States. Biomech. Model. Mechanobio. (2020) doi:10.1007/s10237-020-01332-5.
36. K. Prem, Y. Liu, A.J. Kucharski, R.M. Eggo, N. Davies. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modeling study. Lancet Public Health. doi:10.1016/S2468-2667(20)30073-6.
37. F. Sahli Costabal, K. Matsuno, J. Yao, P. Perdikaris, E. Kuhl. Machine learning in drug development: Characterizing the effect of 30 drugs on the QT interval using Gaussian process regression, sensitivity analysis, and uncertainty quantification. Comp. Meth. Appl. Mech. Eng. 348 (2019) 313-333.
38. J. Salvatier, T.V. Wiecki, C. Fonnesbeck. Probabilistic programming in Python using PyMC3. PeerJ Computer Science 2 (2016) e55.

39. S. Sanche, Y.T. Lin, C. Xu, E. Romero-Severson, N. Hengartner, R. Ke. High contagiousness and rapid spread of severe acute respiratory syndrome coronavirus 2. Emerg. Infect. Disease. (2020) doi:10.3201/eid2607.200282.
40. B. Tang, F. Xia, N.L. Bragazzi, Z. McCarthy, X. Wang, S. He, X. Sun, S. Tang, Y. Xiao, J. Wu. Lessons drawn from China and South Korea for managing COVID-19 epidemic: insights from a comparative modeling study. Bull World Health Organ (2020) doi:10.2471/blt.20.257238.
41. G. Viceconte, N. Petrosillo. COVID-19 R0: Magic number or conundrum? Infect. Dis. Rep. 12 (2020) 8516.
42. World Health Organization. WHO Virtual press conference on COVID-19. https://www.who.int.docs/default-source/coronaviruse/transcripts/who-audio-emergencies-coronavirus-press-conference-full-andfinal-11mar2020.pdf?sfvrsn=cb432bb3_2 recorded: March 11, 2020; accessed: March 16, 2020.
43. Wikipedia. 2020 Coronavirus pandemic in Europe. https://en.wikipedia.org/wiki/2020_corona-virus_pandemic_in_Europe; accessed: April 28, 2020.
44. J. Yuan, M. Li, G. Lv, Z.K. Lu. Monitoring transmissibility and mortality of COVID-19 in Europe. Int. J. Infectious Disease. doi:10.1016/j.ijid.2020.03.050.
45. S. Zhao, Q. Lin, J. Ran, S. S. Musa, G. Yang, W. Wang, Y. Lou, D. Gao, L. Yang, D. He, M. H. Wang, Maggie. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. bioRxiv (2020) doi:10.1101/2020.01.23.916395.
46. T.I. Zohdi. An agent-based computational framework for simulation of competing hostile planet-wide populations. Comp. Meth. Appl. Mech. Eng. 314 (2017) 513-526.
47. T.I. Zohdi. A machine-learning framework for rapid adaptive digital-twin based fire-propagation simulation in complex environments. Comp. Meth. Appl. Mech. Eng. 363 (2020) 112907.