Inferring COVID-19 spreading rates and potential change points for case number forecasts

Jonas Dehning,[1†] Johannes Zierenberg,[1†] F. Paul Spitzner,[1†]
Michael Wibral,[2] Joao Pinheiro Neto,[1] Michael Wilczek,[1†]Viola Priesemann[1†*]

[1] Max Planck Institute for Dynamics and Self-Organization, Göttingen, Germany,
[2] Campus Institute for Dynamics of Biological Networks, University Göttingen, Germany

*To whom correspondence should be addressed; E-mail: viola.priesemann@ds.mpg.de.
† Authors contributed equally

April 2, 2020

As COVID-19 is rapidly spreading across the globe, short-term modeling forecasts provide time-critical information for decisions on containment and mitigation strategies. A main challenge for short-term forecasts is the assessment of key epidemiological parameters and how they change as first governmental intervention measures are showing an effect. By combining an established epidemiological model with Bayesian inference, we analyze the time dependence of the effective growth rate of new infections. For the case of COVID-19 spreading in Germany, we detect change points in the effective growth rate which correlate well with the times of publicly announced interventions. Thereby, we can (a) quantify the effects of recent governmental measures to mitigating the disease spread, and (b) incorporate analogue change points to forecast future scenarios and case numbers. Our code is freely available and can be readily adapted to any country or region.

1 Introduction

During the initial outbreak of an epidemic, reliable short-term forecasts are key to estimate required medical capacities, and to inform and advice the public and the decision makers (*1*). During this initial phase, three tasks are of particular importance to provide time-critical in-

formation for crisis mitigation: (1) establishing central epidemiological parameters such as the basic reproduction number that can be used for short-term forecasting; (2) simulating the effects of different possible intervention measures aimed at mitigation of the outbreak; (3) estimating the actual effects of the measures taken to rapidly adjust them and to adapt short-term forecasts. Tackling these tasks is challenging due to the large statistical and systematic errors present during the initial stages of an epidemic with its low case numbers. Further complications arise from mitigation measures being taken rapidly, while the outbreak unfolds, but taking effect only after an unknown delay. To obtain sensible parameter estimates for short-term forecasting and policy evaluation despite these complications, any prior knowledge available needs to be integrated into modeling efforts to reduce uncertainties. This includes knowledge about basic mechanisms of disease transmission, recovery, as well as preliminary estimates of epidemiological parameters from other countries, or from closely related pathogens. The integration of prior knowledge, the quantitative assessment of the remaining uncertainties about epidemiological parameters, and the principled propagation of these uncertainties into forecasts is the domain of Bayesian modeling and inference.

Here, we draw on an established class of models for epidemic outbreaks: The Susceptible-Infected-Recovered (SIR) model (*2–5*) specifies the rates with which population compartments change over time, i.e., with which susceptible people become infectious, or infectious people recover. This simple model can be formulated in terms of coupled ordinary differential equations (in mean field), which enable analytical treatment (*6, 7*) or fast evaluation ideally suited for Bayesian inference. Accordingly, SIR-like models have been used to model epidemic spreads, from detailed scenario discussions (*8–10*) to Bayesian Markov-Chain Monte Carlo parameter estimation (*11–13*). Recently, this family of models also played a dominant role in the analyses of the global COVID-19 outbreak, from scenario forecast (*14–19*) to inference (*20–22*).

We combine the SIR model with Bayesian parameter inference and augment the model by a time-varying spreading rate λ. The time-varying λ is implemented via potential change points reflecting changes in λ driven by governmental intervention measures. Based on three distinct measures taken in Germany, we also expect to find three corresponding change points. We already detect two such change points from reported COVID-19 case numbers. We find a clear decrease of the spreading rate λ from 0.41 (CI [0.34,0.49]) to 0.23 (CI [0.20,0.29]), with this decrease initiated around March 7th (CI [4th,10th]), and a second decrease down to 0.13 (CI [0.10,0.17]) initiated around March 16th (CI [15th,18th]). This matches the timing of the first two governmental mitigation measures, i.e. first the cancellation of large public events such as trade fairs and soccer matches, and second the closing of schools, child-care facilities, and non-essential shops. We expect a further change point because a third, more stringent lock-down measure was introduced in the following week (March 22nd). However, at present case data do not provide the necessary information on the strength of that third change in λ due to the delay D between infection and report of the confirmed case of about 10 days. In a situation like this where we know that a change in the spreading rate λ has happened but its effect is unobservable yet, forecasts are necessarily difficult.

Our framework can help to infer the effectiveness of past measures as well as to explore po-

tential future scenarios with propagating the respective uncertainties. It can be readily adapted to other countries or regions. The code (already including data sources from many other countries), as well as the figures are all available on Github [1].

2 Background

2.1 Initial phase of the Covid-19 outbreak in Germany is well captured by a standard SIR model

An epidemic outbreak in the absence of mitigation measures can be described by an SIR model with a constant spreading rate λ. Effects of mitigation measures can be approximated as change points in λ, which only manifest after starting mitigation measures and an observation delay. Since first serious interventions in Germany occurred around March 8th, we thus restrict our first estimation to the time period 2020/03/01–2020/03/15.

We use Bayesian MCMC sampling to estimate the central epidemiological parameters for our stationary SIR model (see Methods), specified by a spreading rate λ, a recovery rate μ, a reporting delay D, and the number of initially infected people I_0 (Fig. 1). We obtain as median estimates for the spreading rate $\lambda = 0.41$, $\mu = 0.12$, $D = 8.7$, and $I_0 = 18$. Converted to the basic reproduction number R_0, a central epidemiological parameter, we find a median $R_0 = \lambda/\mu = 3.3$ (CI [2.4, 4.7]), which is consistent with previous reports that find values between 2.3 and 3.3 (*21,23,24*). Overall, the model shows good agreement for both new infections C_t (Fig. 1 A) and the cumulative infections $\sum_{t'=0}^{t} C_{t'}$ (Fig. 1 B) with the expected exponential growth (linear in lin-log plot). The absolute deviation between data and model (Fig. 1 C) is well captured by the case-number-dependent width of our likelihood (Methods) motivated by demographic noise in mean-field models of spreading processes (*25, 26*). However, we observe that for some model parameters, the distribution of estimated parameters (the posterior distribution, Fig. 1 D-H histogram) is largely determined by our initial choice of assumed parameters (the prior distribution, Fig. 1 D-H blue line). In particular, while λ and I_0 are sufficiently constrained by the data, μ an D are not. This is to be expected for the initial phase of an epidemic outbreak, which is dominated by exponential growth.

As long as the COVID-19 spread is still in the initial exponential growth phase, the SIR model can be approximated by an exponential function with effective growth rate $\lambda^* = \lambda - \mu$ (see Methods). As a consequence, λ and μ cannot be estimated independently by the MCMC sampling. This is further supported by a systematic scan of the model's log-likelihood in the λ–μ space showing an equipotential line for the maximum likelihood (Fig. 1 J). This verifies that the effective growth rate λ^* is the relevant free parameter with median $\lambda^* = 28\%$ from the complete MCMC sampling (Fig. 1 I). The control parameter of the dynamics in the exponential onset phase is thus the effective growth rate $\lambda^* = \lambda - \mu$ (Fig. 1 I): If the spreading rate is larger than the recovery rate, $\lambda > \mu$, case numbers grow exponentially. With $\lambda < \mu$, the

[1] https://github.com/Priesemann-Group/covid19_inference_forecast

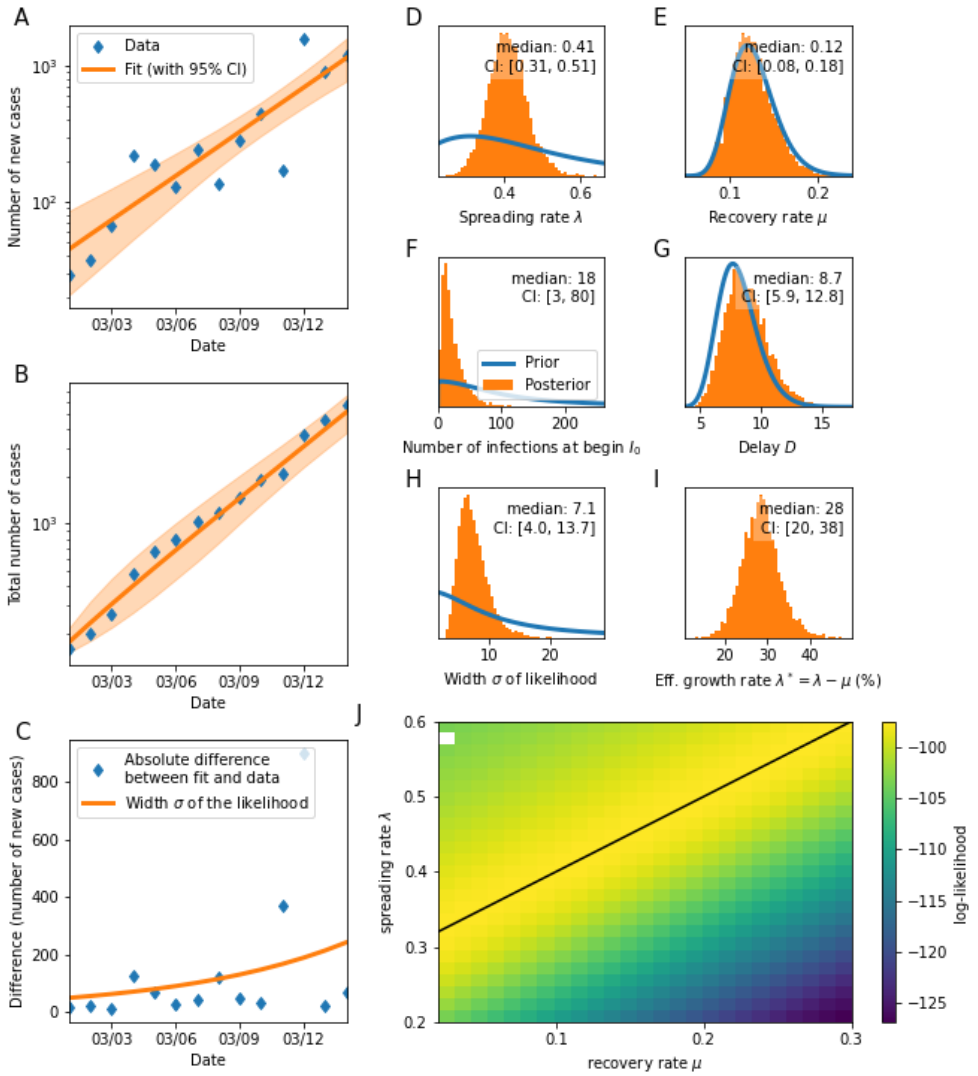

Figure 1: Central epidemiological parameters of the SIR model are constrained by the initial onset period (March 1st–15th). **A:** Number of new cases over time. **B:** Same as A but for total number of cases over time (cumulative). **C:** Absolute difference between model and data (blue line) is captured by the demographic-noise width scale-factor σ we chose for the likelihood (Student's t-distribution). The error of the likelihood function scales with the number of new cases as $\sigma\sqrt{C_t}$. **D-I:** Prior (blue) and posterior (orange) distributions for all model parameters: estimated spreading rate λ, recovery rate μ, reporting delay D between infection date and reporting date, number of cases I_0 at the start of the simulation, width scale-factor σ of the likelihood distribution and the effective growth rate $\lambda^* = \lambda - \mu$ (calculated from λ and μ). **J:** Log-likelihood distribution for different combinations of λ and μ. A linear combination of λ and μ yield the same maximal likelihood (black line).

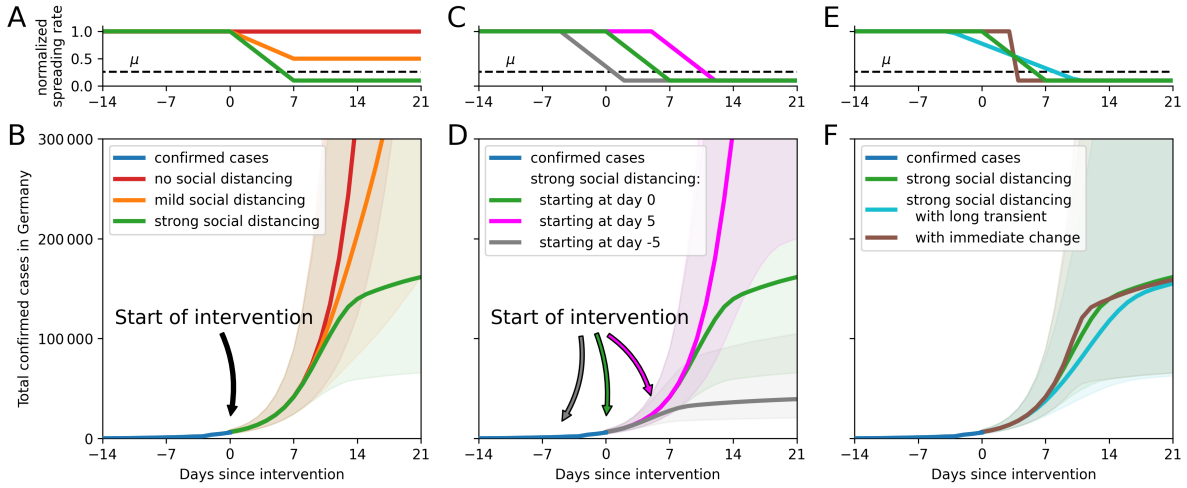

Figure 2: The timing and effectiveness of interventions strongly impact future COVID-19 cases in Germany. **A, B:** We assume three differently scenarios for interventions starting on 2020/03/15: (I) no social distancing — red, (II) mild social distancing — orange, or (III) strict social distancing — green. **C, D:** We also analyzed how a delayed restriction impacts case numbers: Strict restrictions starting on 2020/03/15 (green), or five days later or earlier. A delay (or advance) of five days in implementing restrictions has a major impact on the expected case numbers. **E, F:** Comparison of the time span over which interventions ramp up to full effect. For all ramps that are *centered around the same day*, the resulting case numbers are fairly similar.

recovery dominates and the spread is diminished. The two different dynamics (supercritical and subcritical, respectively) are separated by a critical point at $\lambda = \mu$ (*26*).

2.2 Magnitude and timing of interventions matters for the mitigation of the outbreak

One ulterior motivation for the parameter estimation from past disease spread is to forecast future case numbers, and how they are impacted by political interventions. By modeling different interventions, we show that both, the amount of change in behavior (leading to a change in spreading rate λ, Fig. 2 A,B) and the exact timing of the change (Fig. 2 C,D) determine the future development.

After parameter distributions were inferred on the real-world data up until 2020/03/15, hypothetical interventions were implemented by starting a transition from the past (inferred) spreading rate λ_0 to a new value λ_1 on 2020/03/15. With such a change point, we model three

potential scenarios of public behaviour: **(I) No social distancing;** Public behaviour is unaltered and the spread continues with the inferred rate ($\lambda_1 = \lambda_0$). **(II) Mild social distancing;** The spreading rate decreases to 50%, ($\lambda_1 = \lambda_0 \,/\, 2$). Although people effectively cut the number of contacts in half, the exponential increase in the total number of reported cases continues for another 8 days, before any effect is visible. This duration reflects the reporting delay D between exposure (transmission of the virus to a new susceptible person) and the reporting of the case. **(III) Strong social distancing;** The spreading rate decreases to 10%, ($\lambda_1 = \lambda_0 \,/\, 10$). Contacts are severely limited, but even when people stay at home as much as possible, *some* contacts are unavoidable. Even under such drastic policy changes, no effect is visible until the reporting delay D is over. Thereafter, a quick decrease in daily new infections manifests within two weeks, and the total number of cases reaches a stable plateau. In this scenario, a plateau is reached because the new spreading rate $\lambda_1 \approx 0.04$ is smaller than the recovery rate $\mu = 0.125$.

Timing matters: Apart from the strength of an intervention, its *onset* time has great impact on the total case number (Fig. 2 C,D). For example, focusing on the strong intervention (III) — where a stable plateau is reached — the effect of advancing or delaying the change $\lambda_0 \to \lambda_1$ by just five days leads to more than a 3-fold difference in cumulative cases. While we find that the timing of an intervention has great effect on case numbers, the duration over which the change takes place has only minor effect — if the intervals of change are centered around the same date. In Fig. 2 E,F we illustrate the adjustment of $\lambda_0 \to \lambda_1$ for durations of 14, 7 and 1 day(s). Note that the duration of the adjustment is a simple choice to incorporate variability in individual behaviour, and is not linked to the reporting delay D.

3 Results

3.1 Change point detection reveals the effect of the two first governmental interventions on the outbreak of COVID-19 in Germany

As long as a disease spreads basically unnoticed by the community, the model parameters can be considered stationary (fixed). However, the COVID-19 spread in Germany has by mid-March led to a considerable change in policy and in the behavior of individuals, starting from washing hands more thoroughly and stricter self-isolation upon suspicious symptoms to formal measures like closing public events and places, schools and even introducing a contact ban. The aim of all these measures was to reduce the effective spreading rate $\lambda^* = \lambda - \mu$. As soon as the recovery rate μ is larger than (absolute) spreading rate λ, the number of new confirmed infections should diminish (after the respective delay). As the recovery rate is hard to influence (clinical intervention, immunization), it is expected that these measures dominantly lead to changes in the spreading rate. Hence, detecting change points in the spreading rate — and quantifying the amount of change as quickly as possible — becomes a central modelling challenge, especially with respect to short-term forecasts. Ideally, detected changes can be related to specific mitigation measures, so that one gains an understanding about the effectiveness of different measures.

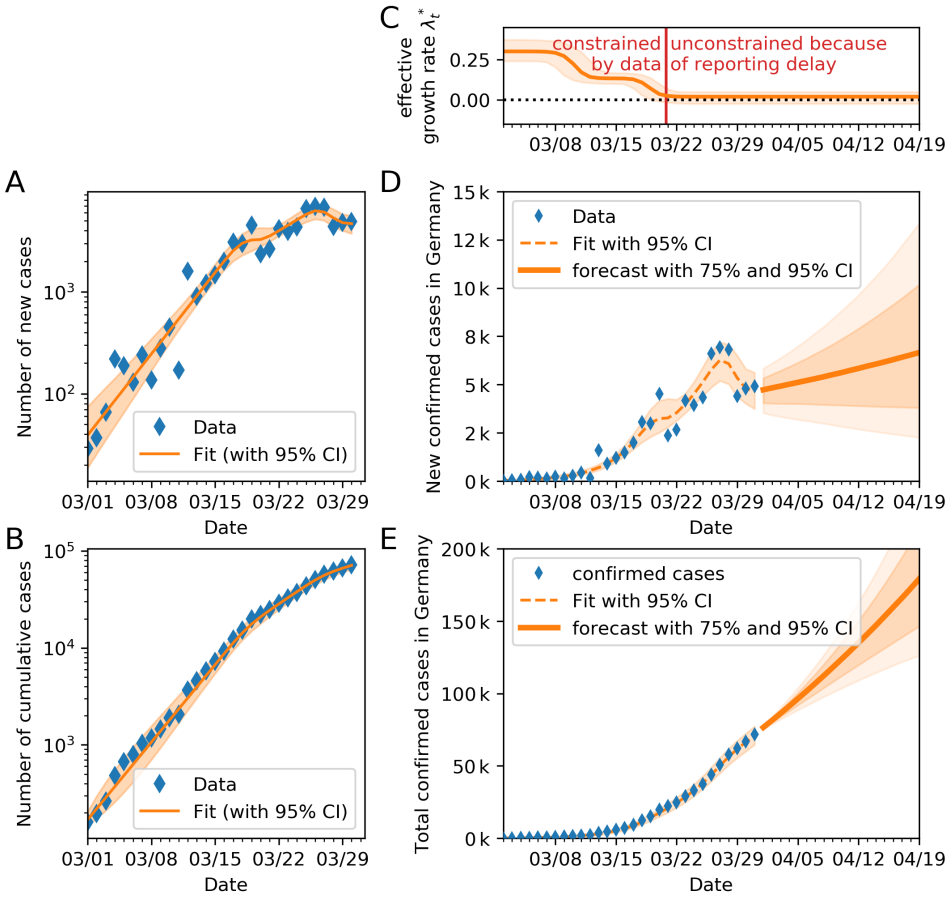

Figure 3: The data until March 31 indicates two change points, consistent with the first two governmental interventions. Assuming a model with two change points, we infer the spreading rates λ_i from the observed COVID-19 case numbers in Germany (blue dots). We then forecast the future development of case numbers (orange line), assuming no further change points here. For the same approach with one or three change points, please see the corresponding figures in the SI (Fig. S1 and S3). Prior and posterior distributions of all model parameters are reported in Fig. 4. **A:** Comparison of daily recorded new cases and the model, log scale. **B:** Comparison of total recorded cases and the model, log scale. **C:** Time-dependent model-estimate of the effective spreading rate λ_t^*. **D:** Model forecast of new cases, based on the inferred λ_t, linear scale. **E:** Model forecast of total cases, based on the inferred λ_t, linear scale.

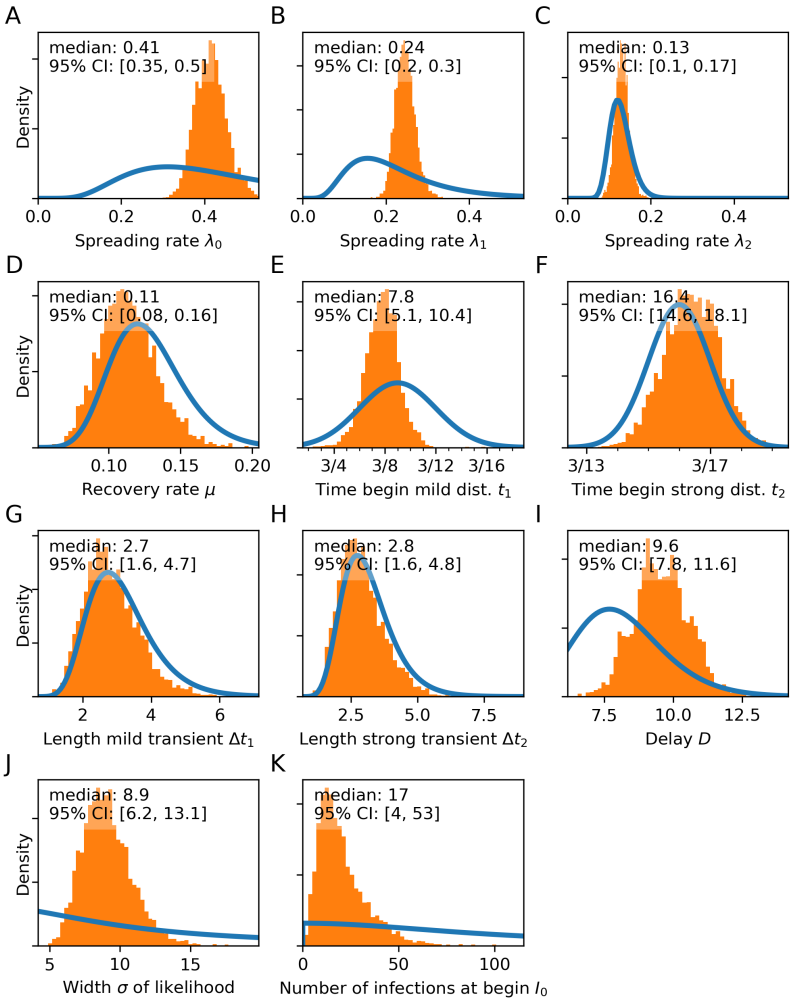

Figure 4: For the SIR model with change points, the data from March 1st–31st constrains the first date of intervention (t_1) in addition to central epidemiological parameters. In **A-K**, all parameters are plotted: the estimated spreading rates λ_0, λ_1, and λ_2, recovery rate μ, the timing of the two change points t_1 and t_2, durations Δt_1, Δt_2 over which the changes unfold, the reporting delay D between infection date and reporting date, width scale-factor σ of the likelihood distribution, and number of cases I_0 at the start of the simulation. For the same model with one or three change points, please see the corresponding figures in the SI (Fig. S2 and S4).

Table 1: Model comparison using leave-one-out (LOO) cross-validation.

Scenario	LOO-score	Effective number of parameters (pLOO)
zero change points	491(16)	4.80
one change point	469(17)	10.67
two change points	449(16)	7.85
three change points	449(16)	8.00

To detect potential changes in λ, we assumed up to three potential changes, starting from the initial spreading rate λ_0 during the exponential growth phase: $\lambda_0 \rightarrow \lambda_1$ at time t_1, $\lambda_1 \rightarrow \lambda_2$ at time t_2, $\lambda_2 \rightarrow \lambda_3$ at time t_3 (Fig. 4 A,B, also see Methods). We expect each change to unfold over a few days (Fig. 4 G,H). The first change point is expected to result from the official changes in policies around March 8th where large gatherings like soccer matches and fairs were banned. A second change point is expected around March 15th, where schools and many shops were closed. A third change point is expected around March 22nd, where all non essential shops were closed, and a contact ban was enacted. As described above, the behavioral changes introduced at these change points take a certain time to show an effect in the observed data. We chose priors for all parameters based on the information we had on March 28th (see Methods) that constrained t_1 at this time. Henceforth, we kept our priors unaltered, which enabled us — with more data — to constrain t_2 in the next step.

3.2 The models with two or three change points fit the observed data better than those with less change points.

Formal, leave-one-out (LOO) cross-validation based Bayesian model comparison (*27*) indicated that the models with two and three change points describe the data better than the models with zero or one change point (Table 1). The model with two change points nominally was the best model, yet the difference between the two best models was considerably smaller than the standard errors of the LOO-score estimates, and is thus not reliably interpretable. Based on the known timing of the third intervention, and our estimates for the delay D (posterior median, 9.5 days), however, we still favor the model with two change points for a description of the current situation. The model with none or a single change point, however, have a LOO-score that are at least about one standard deviation lower than than those of the best models and can be discounted.

With the inclusion of these potential change points into our family of models, we indeed found evidence for two change points in the posterior distributions of the model parameters: First, λ decreased from $\lambda_0 = 0.41$ (CI [0.34, 0.49]) to $\lambda_1 = 0.25$ (CI [0.2, 0.3]). The effective spreading rate $\lambda^* = \lambda - \mu$ decreased by more than a factor 2, from 0.29 to 0.13. The date of the change point was inferred to be the March 7th/8th (95% CI [4th, 11th])]; this inferred date matches the timing of the first political actions like cancelling soccer games, as well as increased awareness.

As of the beginning of April, the data starts to be informative about the second change point from λ_1 to λ_2. Yet, differences between prior and posterior distributions for λ_2 and t_2 are still small. Our prior assumption was that the reduction from λ_1 to λ_2 follows a similar factor of 2 as the first change from λ_0 to λ_1, and that the change point was located around March 15th (see above).

Because we have clear evidence for a first change point, matching the first political intervention, and we start to also have evidence for a second one, matching the second intervention, we will further below also consider the potential future action of a third change point, corresponding to a "contact ban" intervention on March 22nd.

Overall, λ_0 inferred from the model with change points is very similar to λ inferred above from the simpler model, which assumed a stationary λ and was fitted only with data until March 15th (Fig. 1 A). As before, the data provide little information on μ (Fig. 4 D). When interpreting the λ values, remember that the relevant parameter describing the newly confirmed cases is the effective spreading rate $\lambda^* = \lambda - \mu$, as $\lambda^* > 0$ (or $\lambda^* < 0$) determines the exponential growth (or decay) rate.

The durations over which the changes are taking place (transients at t_1 and t_2) are not expected to have a major impact on the results (see Fig. 2 E,F, scenarios). Also the scale factor σ and the number of initial infections I_0 are completely consistent with the initial model inference during the exponential onset phase above (cf Fig. 1). When comparing our inference based on the time-varying λ and its change-points to the number of confirmed cases, we find them to largely match (Fig. 3 A,B). Note that the tell-tale kink in the data is more evident in the raw number of newly confirmed cases (Fig. 3 A) than in the cumulative report (Fig. 3 B). The inferred temporal decrease of new cases, before increasing again, comes from changing an exponential growth rate over small time-interval in the model. It is consistent with the observed temporary drop in newly confirmed cases and suggests a rapid implementation of mitigation measures by the public.

We also observed a spread in the data points that was somewhat larger than expected by the model. We assign this to the fact that we did not incorporate an additional prior describing uncertainty and noise that is introduced by fluctuations in reporting (less reports on weekends, availability of test kits, etc.). Given this caveat, we consider the match of model and data convincing.

3.3 We expect the effects of the "contact ban" of March 22nd to show effect on novel case numbers about two weeks later.

Whereas the data provide sufficient information about the first change point (Fig. 4 A,B,E), and the second one starts to be well constrained, the third one (i.e. the start of the "contact ban" of March 22nd) must be considered unconstrained (Figs. 4 C,F, SI Figs. S1–S4). Displaying all three scenarios together (Fig. 1 on 1 page summary) enables one to compare forecasts resulting from the models with one, two or three change points. It becomes clear that in the current phase, where we expect a third change point, but know very little about its effect on λ, the

future development of case numbers exhibits a large uncertainty. Nonetheless, the observed case numbers already suggest that the second change point brought the disease spread close to the critical transition that separates exponential growth from a decline of new cases. Together with the third potential change point that is bound to further reduce λ, this leaves us with some optimism that the disease spread can be brought under control, if the prescribed social distancing measures are obeyed.

4 Discussion

We here presented a SIR model with a potential change points for spreading rate combined with Bayesian parameter estimation. This model allowed us to estimate the unperturbed central basic epidemiological parameters in the earliest days of the COVID19 outbreak in Germany, but also the detection of two change points in the spreading rate. We could show that change points in the spreading rate will be reflected in confirmed case numbers with a median delay of $D = 9.5$ days and could, this way, link the observed changes to two measures taken by the German government: (1) travel warnings, and the ban on all large events with more than 1000 participants (effected around 2020/03/08), and (2) the closing of schools, childcare centers and many shops (announced 2020/03/15).

The introduction of explicit modeling of change points has a considerable advantage in terms of data efficiency in that all available data could be used for estimating change-point independent model parameters. As long as it is plausible that additional measures will not change model parameters other than λ our approach can be extend by adding further change points and then be reused for further short-term forecasts. Under novel developments, such as the introduction of novel, effective therapies influencing the recovery rate μ our approach provides the algorithmic blueprint for modeling and detecting such change points as well.

Our model comparison also ruled out models without a change point or with only one change point. While this may seem trivial, it has important consequences for making the short term forecasts that decision makers rely on. Demonstrating and quantifying the effect of change points in the past enables us to project the effects of change points in the recent past on the future development of case numbers, even if they are not apparent yet in the observed case numbers. Hence, it is important to look out for and identify potential change points as early as possible, and incorporate them appropriately into the forecast model.

Our results rely on a straight-forward implementation of time-varying spreading rates, that assumed spreading rates to be constant in time, except for rate changes occurring only at the time of interventions – the change points – where we formulated broad prior distributions for the new spreading rate and the time-point of the change. Our results seem to indicate that this modeling approach is sufficient at present for Germany: While we introduced fairly broad priors on the spreading rates we obtained a fairly narrow posterior distribution for the changed spreading rates (Fig. 4 B), indicating that the assumption of discrete steps in the spreading rates is viable. With respect to the posterior distributions for the dates of the change points, we also

found the data to provide valuable information in narrowing the posterior distribution for the first change point, compared to the prior, and in slightly moving the posterior mode for the second change point.

Change point detection and interpretation hinges on proper estimation on the delay D between a new infection and recording a newly confirmed case. Thus, it seems important to assess whether the posterior median value of $D = 9.5$ days is at least compatible with what is known. In our model D sums up at least three separate factors, i.e. the biological incubation period, and an additional delay from first very mild symptoms to symptoms warranting testing under the constraint of limiting testing capacities, and a possible delay before a testing slot is available. At present, the incubation is reported by the WHO as being in a broad range from 1-14 days with a mean of 5 days (*28*), and the WHO states that 'symptoms are usually mild and begin gradually'. A gradual onset of symptoms may delay testing because people tend to wait before asking to be tested, or may be asked to wait as testing in more urgent cases is prioritized. Adding a mean of 5 days of incubation period to the patient waiting for three days after gradual symptom onset before asking for a test, and waiting 1 days for a test would be a plausible scenario in Germany. The resulting 9 days are compatible with the median and spread of the posterior median of $D = 9.5$ we found, given the wide range of durations for the incubation period. It should be noted in addition, that once mass-testing for asymptomatic members of the general population is available, or even mandatory, a change point for D will certainly have to be added to the model.

Since most of the forecast uncertainty comes from scenario assumptions and uncertainties in model parameters, we purposely kept the dynamic model as simple as possible for it to remain tractable. Focusing on a simple model, we have excluded many details that are relevant for precise modelling of epidemics, especially at very late stages or in small populations. Such details include explicit (modeling of) incubation times (*20, 29*), spatial heterogeneity (*15, 20*), isolation (*14, 29*), or subsampling effects hiding undetected cases even beyond the reporting delay (*30*). However, we argue that most of these effects are small compared to the scale of uncertainty between different forecast scenarios that may help decision makers take action, and for short term forecasts.

In conclusion, with our model we could detect and quantify the effect of the two recent governmental interventions from the German COVID-19 data. Furthermore, we forecasted the effects of the most recent intervention. Together, our work showed how important the precise timing and magnitude interventions are for the forecasted case numbers, and how to incorporate the substantial delay D between the date of an infection and the date of the confirmed case. Hence, the development of case numbers in the coming week mainly depends on our behavior in the past week. Likewise, our behavior now determines whether or not we will continue to see exponential growth of case numbers, or whether we achieve a transition to exponential decay by reducing our contact and consequently the spreading rate sufficiently.

5 Materials and methods

As a basis for our forecast scenarios, we use the differential equations of the well-established SIR (Susceptible-Infected-Recovered) model. Case data comes from the COVID-19 data repository maintained by the Johns Hopkins University Center for Systems Science and Engineering [2]. While the model dynamics is well understood in general, here our main challenge is to estimate model parameters specifically for the COVID-19 outbreak. To that end, we combined a Bayesian approach — to incorporate prior knowledge — with Markov Chain Monte Carlo (MCMC) sampling — to explore the parameters. Put simply, we first estimate the parameter distribution that best describes the observed situation, and then we use many samples from this parameter distribution to evolve the model equations and thus forecast future developments.

5.1 Simple model: SIR model with stationary spreading rate λ

We consider a time-discrete version of the standard SIR model. In short, we assume that the disease spreads at rate λ from the infected population stock (I) to the susceptible population stock (S), and that the infected stock recovers (R) at rate μ. This well-established model for disease spreading can be described by the following set of (deterministic) ordinary differential equations (see, e.g., Refs (*3, 4, 14*)). Within a population of size N,

$$\begin{aligned} \frac{\mathrm{d}S}{\mathrm{d}t} &= -\lambda \frac{SI}{N} \\ \frac{\mathrm{d}I}{\mathrm{d}t} &= \lambda \frac{SI}{N} - \mu I \\ \frac{\mathrm{d}R}{\mathrm{d}t} &= \mu I \,. \end{aligned} \tag{1}$$

As a remark, during the onset phase of an epidemic only a very small fraction of the population is infected (I) or recovered (R), and thus $S \approx N \gg I$ such that $S/N \approx 1$. Therefore, the differential equation for the infected reduces to a simple linear equation, exhibiting an exponential growth

$$\frac{dI}{dt} = (\lambda - \mu) I \quad \text{solved by} \quad \mathrm{I(t)} = \mathrm{I(0)}\, \mathrm{e}^{(\lambda-\mu)\mathrm{t}} \,. \tag{2}$$

Because our data set is discrete in time (Δt =1 day), we solve the above differential equations with a discrete time step ($\mathrm{d}I/\mathrm{d}t \approx \Delta I/\Delta t$), such that

$$\begin{aligned} S_t - S_{t-1} &= -\lambda \Delta t \tfrac{S_{t-1}}{N} I_{t-1} &&=: -I_t^{\mathrm{new}} \\ R_t - R_{t-1} &= \mu \Delta t I_{t-1} &&=: \; R_t^{\mathrm{new}} \\ I_t - I_{t-1} &= \left(\lambda \tfrac{S_{t-1}}{N} - \mu\right) \Delta t I_{t-1} = I_t^{\mathrm{new}} - R_t^{\mathrm{new}} \,. \end{aligned} \tag{3}$$

Importantly, I_t models the number of all (currently) active infected people, while I_t^{new} is the number of new infections that will eventually be reported according to standard WHO convention. Importantly, we explicitly include a reporting delay D between new infections I_t^{new} and

[2] https://github.com/CSSEGISandData/COVID-19

newly reported cases (C_t) as

$$C_t = I_{t-D}^{\text{new}}. \tag{4}$$

We begin our simulations at time $t = 0$ with I_0 infected cases and start including real-word data of reported cases $\hat{C}_t$ from day $t > D$ (see below for a parameterization).

Full model: SIR model with change points in λ

Our change-point detection builds on a generalization of the simple SIR model with stationary spreading rate. Instead, we now assume that the spreading rate λ_i, $i = 1, ..., n$, may change at certain time points t_i from λ_{i-1} to λ_i, linearly over a time window of Δt_i days. Thereby, we account for policy changes to reduce λ, which were implemented in Germany step-by-step. Thus, the parameters t_i, Δt_i, and λ_i are added to the parameter set of the simple model above, and the differential equations are augmented by the time-varying λ_i.

Estimating model parameters with Bayesian MCMC

We estimate the set of model parameters $\theta = \{\lambda_i, t_i, \mu, D, \sigma, I_0\}$ using Bayesian inference with Markov-chain Monte-Carlo (MCMC). The parameter σ is the scale factor for the width of the likelihood $P\left(\hat{C}_t|\theta\right)$ between observed data and model (see below). Our implementation relies on the python package pymc3 (*31*) with NUTS (No-U-Turn Sampling) (*32*). The structure of our approach is the following:

Choose random initial parameters and evolve according to model equations. Initially, we choose parameters θ from prior distributions that we explicitly specify below. Then, time integration of the model equations generates a (fully deterministic) time series of new infected cases $C(\theta) = \{C_t(\theta)\}$ of the same length as the observed real-world data $\hat{C} = \left\{\hat{C}_t\right\}$.

Iteratively update the parameters using MCMC. The drawing of new candidate parameters and the time integration of the SIR model is repeated in every MCMC step. The idea is to probabilistically draw parameter updates and to accept them such that the deviation between the model outcome and the available real-world time-series is likely to reduce. We quantify the deviation between the model outcome for one time point t, $C_t(\theta)$ and the corresponding real-world data point $\hat{C}_t$ with the local likelihood

$$p\left(\hat{C}_t|\theta\right) \sim \text{StudentT}_{\nu=4}\left(\text{mean} = \text{C}_\text{t}(\theta),\ \text{width} = \sigma\sqrt{\text{C}_\text{t}(\theta)}\right).$$

We chose the Student's t-distribution because it resembles a Gaussian distribution around the mean but features heavy tails, which make the MCMC more robust with respect to outliers (*33*). The case-number-dependent width models the demographic noise of typical mean-field solutions for epidemic spreading, e.g., $\dot{\rho}(t) = a\rho(t) - b\rho^2(t) + \sqrt{\rho(t)}\eta(t)$, where ρ is the activity

Table 2: Priors on the model parameters for the SIR model with stationary spreading rate.

Parameter	Variable	Prior distribution
Spreading rate	λ	$\mathrm{LogNormal}(\log(0.4), 0.5)$
Recovery rate	μ	$\mathrm{LogNormal}(\log(1/8), 0.2)$
Reporting delay	D	$\mathrm{LogNormal}(\log(8), 0.2)$
Initially infected	I_0	$\mathrm{HalfCauchy}(100)$
Scale factor	σ	$\mathrm{HalfCauchy}(10)$

and $\eta(t)$ is Gaussian white noise (*25, 26*). This choice is consistent with our data (Fig. 1 A-C). The overall deviation is then simply the product of local likelihoods over all time points.

For each MCMC step, the new parameters are drawn so that a set of parameters that minimizes the previous deviation is more likely to be chosen. In our case, this is done with an advanced gradient-based method (NUTS (*32*)). To reiterate, every time integration that is performed has its own set of parameters and yields one complete model time series. If the time integration describes the data well the parameter set is accepted, and this yields one Monte Carlo sample of the model parameters for the posterior distribution; the MCMC step is then repeated to create more samples from the posterior. Eventually, the majority of accepted parameter samples will describe the real-world data well, so that consistent forecasts are possible in the forecast phase.

Forecast using Monte Carlo samples. For the forecast, we take all samples from the MCMC step and continue time integration according to different forecast scenarios explained below. Note that the overall procedure yields an ensemble of forecasts — as opposed to a single forecast that would be solely based on one set of (previously optimized) parameters.

Priors that constrain model parameters

As forecasts are needed rapidly at the onset of an epidemic, the available real-world data is typically not informative enough to identify all free parameters, or to empirically find their underlying distributions. We therefore chose informative priors on initial model parameters where possible and complemented with uninformative priors else. Our choices are summarized in Tab. 2 for the simple model, SIR model with stationary spreading rate for the exponential onset phase, and in Tab. 3 for the full model with change points, and justified in the following.

Priors for the simple model (Table 2): In order to constrain our simple model, an SIR model with stationary spreading rate for the exponential onset phase, we chose the following informative priors. Because of the ambiguity between the spreading and recovery rate in the exponential onset phase (see Simple model), we chose a narrow log-normal prior for the recovery rate $\mu \sim \mathrm{LogNormal}(\log(1/8), 0.2)$ with median recovery time of 8 days (*14*). Note that, our implementation of μ accounts for the recovery of infected people and isolation

Table 3: Priors on the model parameters for the SIR model with change points.

Parameter	Variable	Prior distribution
Change points	t_1	$\mathrm{Normal}(2020/03/09, 3)$
	t_2	$\mathrm{Normal}(2020/03/15, 1)$
	t_3	$\mathrm{Normal}(2020/03/22, 1)$
Change duration	Δt_i	$\mathrm{LogNormal}(\log(3), 0.3)$
Spreading rates	λ_0	$\mathrm{LogNormal}(\log(0.4), 0.5)$
	λ_1	$\mathrm{LogNormal}(\log(0.2), 0.5)$
	λ_2	$\mathrm{LogNormal}(\log(1/8), 0.2)$
	λ_3	$\mathrm{LogNormal}(\log(1/16), 0.2)$
Recovery rate	μ	$\mathrm{LogNormal}(\log(1/8), 0.2)$
Reporting delay	D	$\mathrm{LogNormal}(\log(8), 0.2)$
Initially infected	I_0	$\mathrm{HalfCauchy}(100)$
Scale factor	σ	$\mathrm{HalfCauchy}(10)$

measures, because it describes the duration during which a person can infect others. For the spreading rate, we assume a broad log-normal prior distribution $\lambda \sim \mathrm{LogNormal}(\log(0.4), 0.5)$ with median 0.4. This way, the prior for $\lambda - \mu$ has median 0.275 and the prior for the base reproduction number($R_0 = \lambda/\mu$) has median 3.2 consistent with the broad range of previous estimates (*21, 23, 24*). In addition, we chose a log-normal prior for the reporting delay $D \sim \mathrm{LogNormal}(\log(8), 0.2)$ to incorporate both the incubation time between 1–14 days with median 5 (*28*) plus a delay from infected people waiting to contact the doctor and get tested.

The remaining model parameters are constrained by uninformative priors, in practice the Half-Cauchy distribution (*34*). The half-Cauchy distribution $\mathrm{HalfCauchy}(\mathrm{x}, \beta) = 2/\pi\beta[1 + (\mathrm{x}/\beta)^2]$ is essentially a flat prior from zero to $O(\beta)$ with heavy tails beyond. Thereby, β merely sets the order of magnitude that should not be exceeded for a given parameter. We chose for the number of initially infected people in the model (16 days before first data point) $I_0 \sim \mathrm{HalfCauchy}(100)$ assuming an order of magnitude $O(100)$ and below. In addition, we chose of the scale factor of the width of the likelihood function $\sigma \sim \mathrm{HalfCauchy}(10)$, which is further multiplied to the number of new cases.

Priors for the full model (Table 3): In order to constrain our full model, an SIR model with change points in the spreading rate, we chose the same priors as for the simple model but added the required priors associated with the change points.

For the timing of change points, we chose normal distributed priors. In particular, we chose for the first change point $t_1 \sim \mathrm{Normal}(2020/03/09, 3)$ because on the weekend of March 8th, large public events, like visits to soccer matches or fairs, were cancelled. We chose for the second change point $t_2 \sim \mathrm{Normal}(2020/03/15, 1)$, because on March 15th, the closing of schools and other educational institutions along with the closing of non-essential stores were announced and implemented on the following day. Restaurants were allowed to stay open until 6 pm. We

chose for the third change point $t_3 \sim \mathrm{Normal}(2020/03/22, 1)$, because on March 22nd, a far-reaching contact ban ("Kontaktsperre"), which includes the prohibition of even small public gatherings as well as complete closing of restaurants and non-essential shops was imposed by the government authorities. Further policy changes may occur in future; however, for now, we do not include more change points. We model the time dependence of λ as change points, and not as continuous changes, because the policy changes were incurred in these three discrete steps, and in our observations were adhered by the public. Continuous changes, originating e.g. from increased awareness of the population can be accounted for by the discrete steps as well, within the precision of reported cases we have.

The change points take effect over a certain time period Δt_i for which we choose $\Delta t_i \sim \mathrm{LogNormal}(\log(3), 0.3)$ with a median of 3 days over which the spreading rate changes continuously as interventions have to become effective. The precise duration of the transition has hardly any affect on the cumulative number of cases (Fig. 2 E-F). We assumed a duration of three days, because some policies were not announced at the same day for all states within Germany; moreover, the smooth transition also can absorb continuous changes in behavior.

For the spreading rates, we chose log-normal distributed priors as in the simple model. In particular, we chose for the initial spreading rate the same prior as in the simple model, $\lambda_0 \sim \mathrm{LogNormal}(\log(0.4), 0.5)$; after the first change point $\lambda_1 \sim \mathrm{LogNormal}(\log(0.2), 0.5)$, assuming the first intervention to reduce the spreading rate by 50% from our initial estimates ($\lambda_0 \approx 0.4$) with a broad prior distribution; after the second change point $\lambda_2 \sim \mathrm{LogNormal}(\log(1/8), 0.2)$, assuming the second intervention to reduce the spreading rate to the level of the recovery rate, which would lead to a stationary number of new infections. This corresponds approximately to a reduction of λ at the change point by 50%; and after the third change point $\lambda_3 \sim \mathrm{LogNormal}(\log(1/16), 0.2)$, assuming the third intervention to reduce the spreading rate again by 50%. With that intervention, λ_3 is smaller than the recovery rate μ, causing a decrease in new case numbers and a saturation of the cumulative number of infections. In general, we assume that each package of governmental interventions (together with the increasing awareness) leads to a reduction (and not an increase) of λ at a change point. As we cannot know yet the precise reduction factor, we adhere to assume a reduction by 50%, but always with a fairly wide uncertainty, so that in principle even an increase at the change point would be possible.

Using change points has the advantage that extrapolation to future behavior assumes no further change in λ apart from the change points. That facilitates the understanding of future scenarios. In a future iteration of the model, we will compare continuous versus change-point behavior.

Model comparison

Since change point detection entails evaluating models with different numbers of parameters, some form of fair model comparison needs to be performed. Here, we compared the models with different numbers of change points by their pointwise out-of-sample prediction accuracy using the log-likelihood evaluated at the posterior simulations of the parameter values obtained

Table 4: Overview of model parameters.

Variable	Parameter
$\theta = \{\lambda_i, t_i, \mu, \sigma, I_0\}$	Set of model parameters that are optimized
λ	Spreading rate
μ	Recovery rate
$\lambda^* = \lambda - \mu$	Effective spreading rate
λ_i	Spreading rate after i-th intervention
t_i	Time of i-th intervention
σ	Scale factor of the width of Student's t-distribution
N	Population size (83.700.000)
S_t	Susceptible at time t
I_t	Infected at time t
R_t	Recovered at time t
Δt	Time step
$R_t^{\text{new}} = \mu \Delta t I_{t-1}$	New recoveries at time t
$I_t^{\text{new}} = \lambda \Delta t \frac{S_{t-1}}{N} I_{t-1}$	New infections at time t
$C_t = I_{t-D}^{\text{new}}$	New reported cases at time t
D	Delay of case detection

from the fitted models. Out-of-sample accuracy was approximated using Leave-one-out cross-validation (LOO) (*27*).

References

1. M. Enserink, K. Kupferschmidt, *Science* **367**, 1414 (2020).
2. W. O. Kermack, A. G. McKendrick, G. T. Walker, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **115**, 700 (1927).
3. H. Hethcote, *SIAM Rev.* **42**, 599 (2000).
4. J. Anderson, I. Lampl, I. Reichova, M. Carandini, D. Ferster, *Nat. Neurosci.* **3**, 617 (2000).
5. N. C. Grassly, C. Fraser, *Nat Rev Microbiol* **6**, 477 (2008).
6. R. Parshani, S. Carmi, S. Havlin, *Phys. Rev. Lett.* **104**, 258701 (2010).
7. T. Harko, F. S. N. Lobo, M. K. Mak, *Applied Mathematics and Computation* **236**, 184 (2014).
8. B. Shulgin, L. Stone, Z. Agur, *Bull. Math. Biol.* **60**, 1123 (1998).

9. O. N. Bjørnstad, B. F. Finkenstädt, B. T. Grenfell, *Ecol. Monogr.* **72**, 169 (2002).
10. L. Hufnagel, D. Brockmann, T. Geisel, *Proc. Natl. Acad. Sci. USA* **101**, 15124 (2004).
11. T. Britton, P. D. O'Neill, *Scand J Stat* **29**, 375 (2002).
12. J. Lourenço, *et al.*, *eLife* **6**, e29820 (2017).
13. N. R. Faria, *et al.*, *Sci Rep* **7**, 1 (2017).
14. B. F. Maier, D. Brockmann, *arXiv:2002.07572* (2020).
15. P. Bittihn, R. Golestanian, *arXiv:2003.08784* (2020).
16. R. M. Anderson, H. Heesterbeek, D. Klinkenberg, T. D. Hollingsworth, *The Lancet* **395**, 931 (2020).
17. J. R. Fauver, *et al.*, *medRxiv* p. 2020.03.25.20043828 (2020).
18. A. Arenas, *et al.*, *medRxiv* p. 2020.03.21.20040022 (2020).
19. O. Mitjà, *et al.*, *The Lancet* **0** (2020).
20. R. Li, *et al.*, *Science* (2020).
21. A. J. Kucharski, *et al.*, *The Lancet Infectious Diseases* (2020).
22. J. Lourenco, *et al.*, *medRxiv* p. 2020.03.24.20042291 (2020).
23. J. Zhang, *et al.*, *medRxiv* (2020).
24. Y. Liu, A. A. Gayle, A. Wilder-Smith, J. Rocklöv, *J Travel Med* **27** (2020).
25. S. di Santo, P. Villegas, R. Burioni, M. A. Muñoz, *Phys. Rev. E* **95** (2017).
26. M. A. Muñoz, *Rev. Mod. Phys.* **90**, 031001 (2018).
27. A. Vehtari, A. Gelman, J. Gabry, *Stat Comput* **27**, 1413 (2017).
28. S. A. Lauer, *et al.*, *Ann Intern Med* (2020).
29. L. Peng, W. Yang, D. Zhang, C. Zhuge, L. Hong, *arXiv:2002.06563* (2020).
30. J. Wilting, V. Priesemann, *Nat. Commun.* **9**, 2325 (2018).
31. J. Salvatier, T. V. Wiecki, C. Fonnesbeck, *PeerJ Comput. Sci.* **2**, e55 (2016).
32. M. D. Hoffman, A. Gelman, *J. Mach. Learn. Res.* **15**, 1593 (2014).
33. K. L. Lange, R. J. A. Little, J. M. G. Taylor, *J. Am. Stat. Assoc.* **84**, 881 (1989).
34. A. Gelman, *Bayesian Anal.* **1**, 515 (2006).

Acknowledgments

We thank Tim Friede, Theo Geisel, Moritz Linkmann, Matthias Loidolt and Vladimir Zykov for carefully and promptly reviewing our work internally. We thank the Priesemann group - especially Daniel Gonzalez Marx, Fabian Mikulasch, Lucas Rudelt & Andreas Schneider - for exciting discussions and for their valuable comments. We thank the colleagues of the Göttingen Campus, with whom we were discussing the project and the COVID-19 case forecast in the past weeks very intensively: Heike Bickeböller, Eberhard Bodenschatz, Wolfgang Brück, Alexander Ecker, Andreas Leha, Ramin Golestanian, Helmut Grubmüller, Stephan Herminghaus, Reinhard Jahn, Norbert Lossau & Simone Scheithauer. JD and PS received financial support from the Bernstein network via the smartstart program. JZ received finanicial support from the Joachim Herz Stiftung.

Supplementary materials

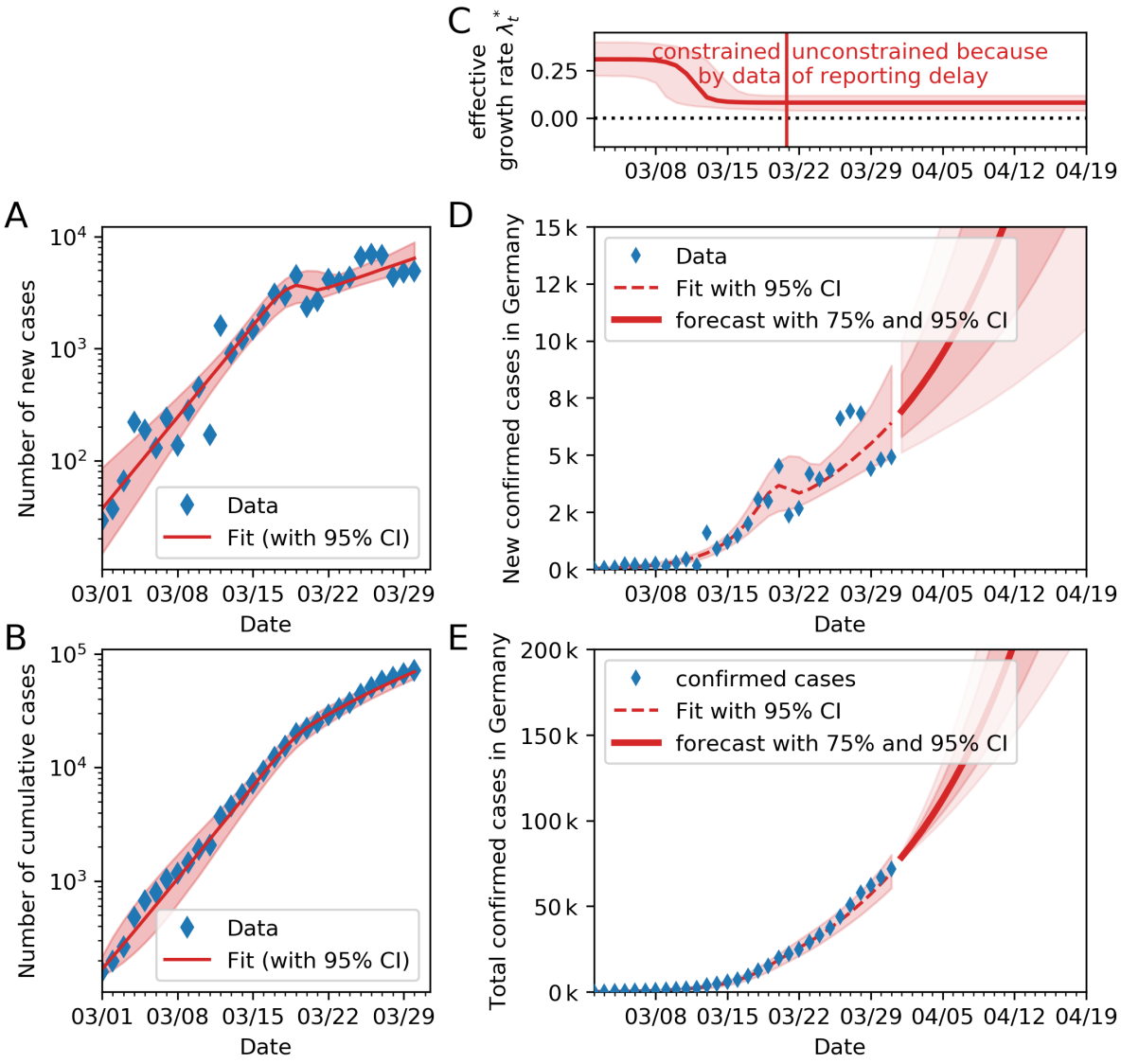

Figure S1: Inferring (the time of) the change point in spreading rate λ, **assuming only one change point**. Please refer to Fig. 3 for the case with two change points. **A:** Comparison of daily recorded new cases and the model, log scale. **B:** Comparison of total recorded cases and the model, log scale. **C:** Time-dependent model-estimate of the spreading rate λ_t. **D:** Model forecast of new cases, based on the inferred λ_t, linear scale. **E:** Model forecast of total cases, based on the inferred λ_t, linear scale.

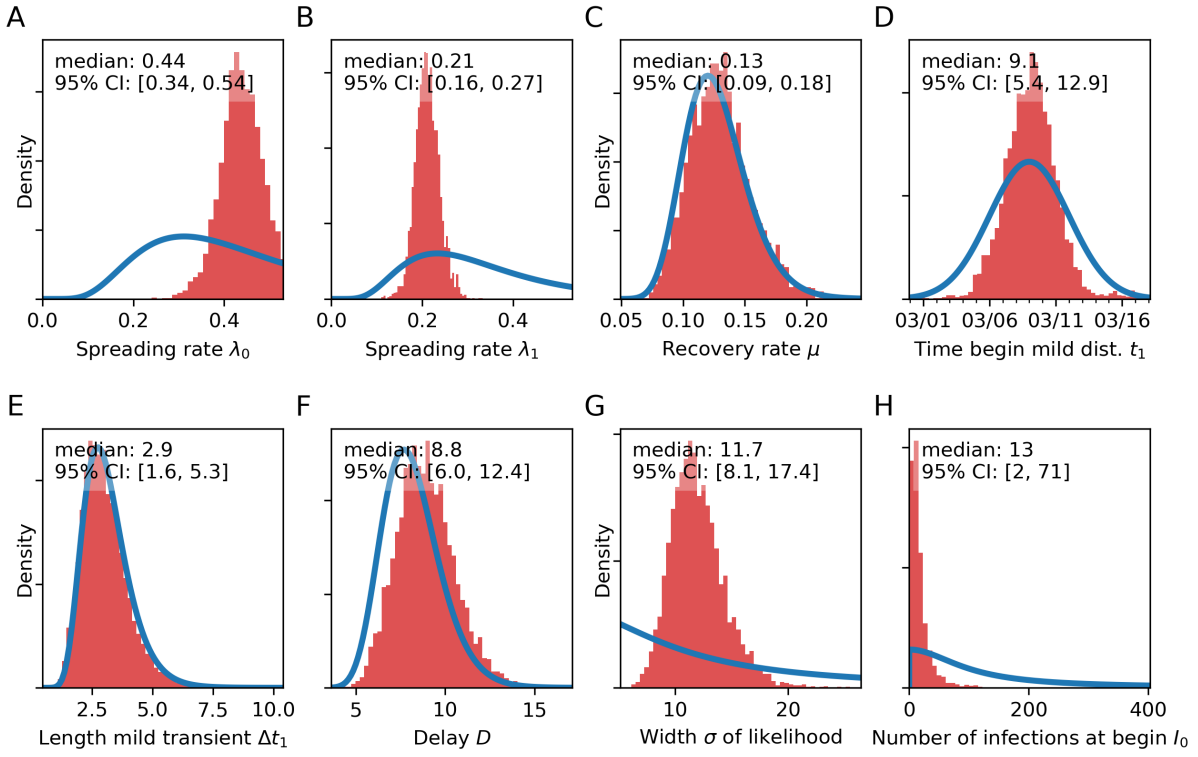

Figure S2: Posterior distributions from the change-point detection (red) with one change point compared to prior distributions (blue). Please refer to Fig. 4 for a more detailed description of the distributions.

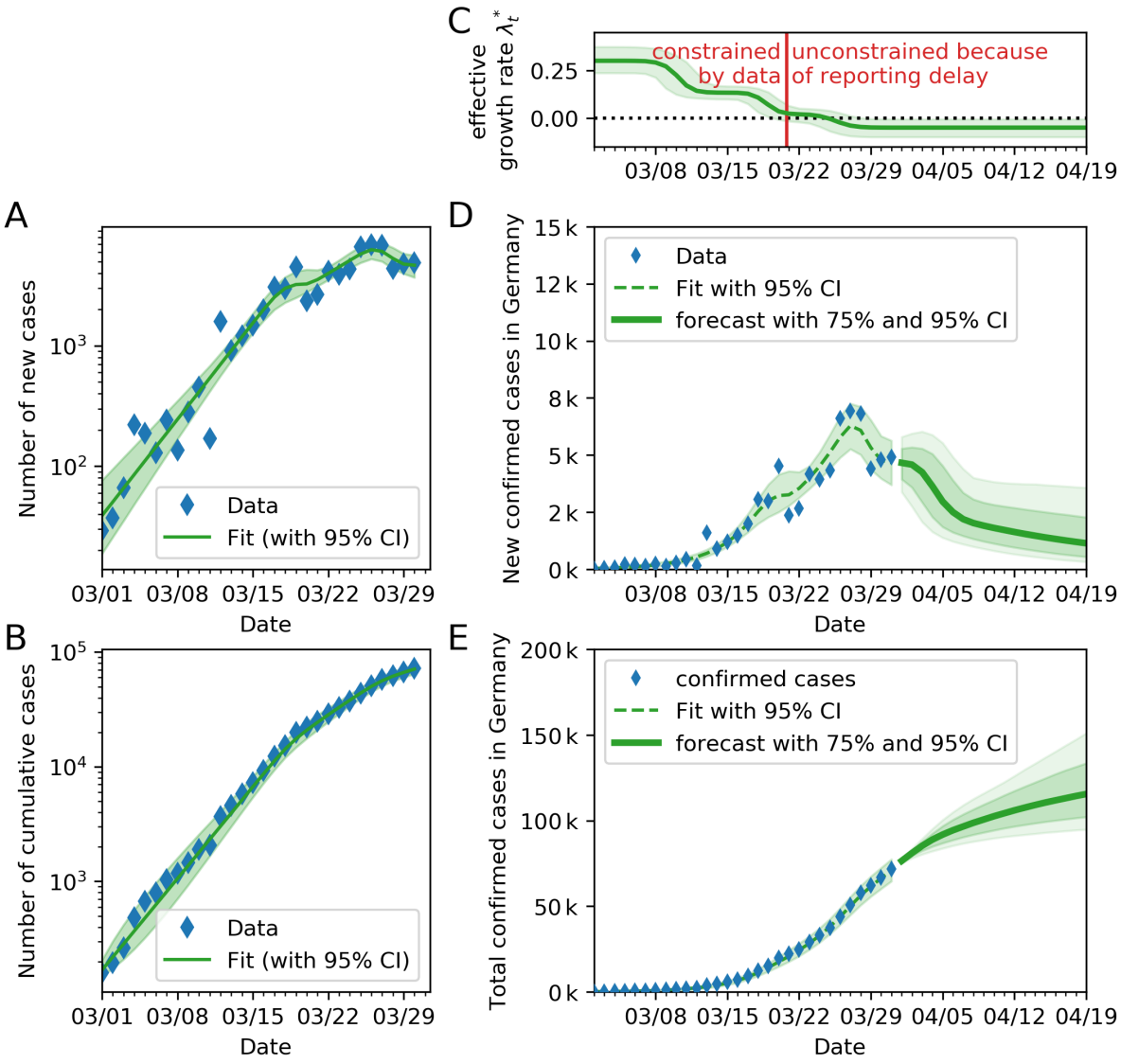

Figure S3: Inferring and forecasting (the time of) the change point in spreading rate λ, **assuming three change points**. Please refer to Fig. 3 for the case with two change points. **A:** Comparison of daily recorded new cases and the model, log scale. **B:** Comparison of total recorded cases and the model, log scale. **C:** Time-dependent model-estimate of the spreading rate λ_t. **D:** Model forecast of new cases, based on the inferred λ_t, linear scale. **E:** Model forecast of total cases, based on the inferred λ_t, linear scale.

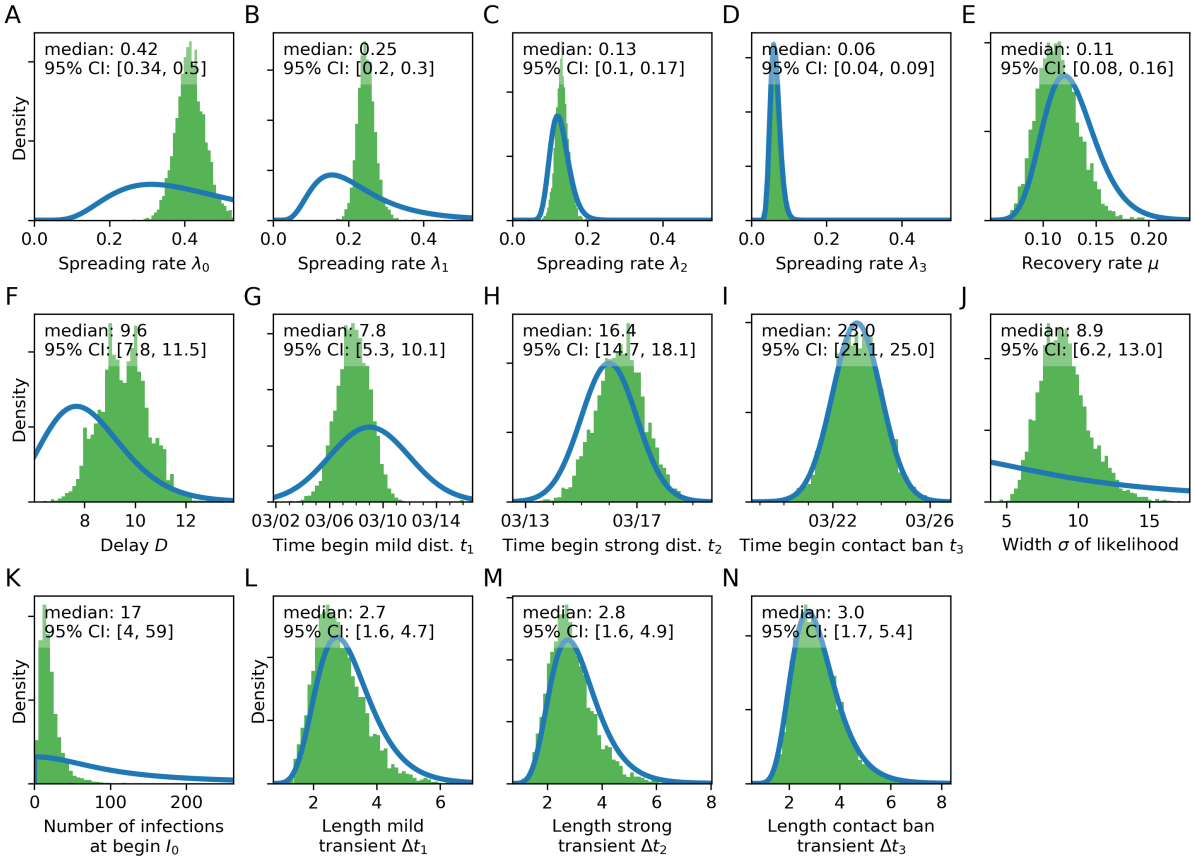

Figure S4: Posterior distributions from the change-point detection with three change points (green) compared to prior distributions (blue). Please refer to Fig. 4 for a more detailed description of the distributions.