

RESEARCH

Open Access



An epidemic model integrating direct and fomite transmission as well as household structure applied to COVID-19

Karunia Putra Wijaya^{1*}, Naleen Ganegoda², Yashika Jayathunga¹, Thomas Götz¹, Moritz Schäfer¹ and Peter Heidrich¹

*Correspondence:

karuniaputra@uni-koblenz.de

¹Mathematical Institute, University of Koblenz, DE-56070 Koblenz, Germany

Full list of author information is available at the end of the article

Abstract

This paper stresses its base contribution on a new SIR-type model including direct and fomite transmission as well as the effect of distinct household structures. The model derivation is modulated by several mechanistic processes inherent from typical airborne diseases. The notion of minimum contact radius is included in the direct transmission, facilitating the arguments on physical distancing. As fomite transmission heavily relates to former-trace of sneezes, the vector field of the system naturally contains an integral kernel with time delay indicating the contribution of undetected and non-quarantined asymptomatic cases in accumulating the historical contamination of surfaces. We then increase the complexity by including the different transmission routines within and between households. For airborne diseases, within-household interactions play a significant role in the propagation of the disease rendering countrywide effect. Two steps were taken to include the effect of household structure. The first step subdivides the entire compartments (susceptible, exposed, asymptomatic, symptomatic, recovered, death) into the household level and different infection rates for the direct transmission within and between households were distinguished. Under predefined conditions and assumptions, the governing system on household level can be raised to the community level. The second step then raises the governing system to the country level, where the final state variables estimate the total individuals from all compartments in the country. Two key attributes related to the household structure (number of local households and number of household members) effectively classify countries to be of low or high risk in terms of effective disease propagation. The basic reproductive number is calculated and its biological meaning is invoked properly. The numerical methods for solving the DIDE-system and the parameter estimation problem were mentioned. Our optimal model solutions are in quite good agreement with datasets of COVID-19 active cases and related deaths from Germany and Sri Lanka in early infection, allowing us to hypothesize several unobservable situations in the two countries. Focusing on extending minimum contact radius and reducing the intensity of individual activities, we were able to synthesize the key parameters telling what to practice.

Keywords: Mathematical model; Direct transmission; Fomite transmission; COVID-19; Parameter estimation

1 Introduction

In December 2019, a number of residents were diagnosed with pneumonia in Wuhan, China alarming a viral outbreak. Later in early January 2020, it was identified that the cause of this pneumonia be a novel coronavirus (SARS-CoV-2) and the disease is officially named as COVID-19 [1]. World Health Organization (WHO) declared COVID-19 as a pandemic on March 11, 2020 and has been conveying technical guidance to mitigate the disease burden with over 20 million confirmed cases and over 125,000 fatalities worldwide as of mid April, 2020 [2]. Quantification measures such as the basic reproductive number and case fatality rate direct scientists, authorities, and general public toward testing and simulating, bearing interventions and evaluating healthcare capacity, and preventive habits, respectively.

Signs and symptoms of the infection start from dry cough and fever, and then some other mild clinical manifestations would be possible until severe levels such as acute respiratory distress syndrome persist [3]. According to early surveillance, Huanan seafood market was the originating place and thereafter human-to-human transmission has escalated in an exponential pace. As per RNA-sequence analysis, the novel coronavirus may be arisen from cross-species transmission between bats and snakes [4]. Three variants of the virus were identified in a phylogenetic network analysis that used bat virus as an outgroup [5]. Two of them are common in Europeans and Americans, while the other one is common in East Asia, showing the vulnerability in genetic changes. Li et al. [6] as per their analysis of early dynamics, reported that the basic reproductive number \mathcal{R}_0 of COVID-19 in China was 2.2 and incubation period 5.2 days. In another early estimation, Zhao et al. [7] reported that \mathcal{R}_0 ranges from 2.24 to 3.58. Amidst those varying estimations, transmission risk of COVID-19 is noteworthy as compared to previous two outbreaks SARS in 2003 and MERS in 2012 since individuals can be infectious before being symptomatic [1, 8].

After the early signs of the outbreak, health authorities worldwide have taken preventive actions with the patronage of WHO [9]. These actions range over good practices in household level such as washing hands, wearing masks and undertaking social distancing in public level, quarantine at home or designated centers, travel restrictions in international level, and finally lockdowns. Countrywise variations can be observed in motivation, compliance, and success toward effective control strategies [10]. There is notable difference in hospitalization, for instance China practiced it as an isolation procedure as well, while USA and UK recommended it only for severe cases [11]. Whether to extend or relax distancing measures is still a dilemma as there is an emerging threat of second wave resurgence [12].

Four levels of transmission designated by WHO (1—no cases, 2—sporadic cases, 3—clusters of cases, 4—community spread) set the need of basic interventions and intensity [13]. Countries are given a more detailed classification, indicating lower to higher risk as interrupted transmission, under investigation, imported cases only, local transmission and community transmission. However, it should be scaled down to workplace and household level. Then compartmental aspect of how many are susceptible, symptomatic, asymptomatic, recovered, etc. will decide the immediate risk to the community [14]. Both susceptibility and fatality risks have age dependency, where older group shows higher vulnerability [15, 16]. In addition, sneezing, touching contaminated surfaces, individual immunity, number of members sharing the space necessarily come into effect. To be more successful, all preventive measures should be communicated in a practically sounding way,

irrespective of the socio-economic background of community or country. Under dreadful circumstances, we should focus on what are expected from each individual to mitigate further risk of COVID-19.

In this work, we are concerned with a certain rule of thumb that can be applied on the individual level. Two measures are put under investigation: minimum contact radius during physical distancing and the average number of daily cross-household encounters. The first measure has been appearing in the media from various countries, steering the possibility of getting infection via direct transmission (i.e. by sneezing). The second measure represents how many encounters with infected persons per day a susceptible person makes on average. This number is unobservable, yet is known to drive the force of infection in standard epidemic models, therefore also subject to treatment. After model derivation and data fitting, our final plan will be to compare the performance of extending the minimum contact radius and reducing the number of cross-household encounters from the model solution in terms of minimizing the number of predicted active cases. Our model bases its design on a standard SIR-type model with additional features highlighting minimum contact radius, direct and fomite transmission, also different household structures incorporating district-specific attributes. Via the latter we can be more specific on selecting transmission routines within and between households. Many factors are infused into a basic model through parameters in the form of rates, probabilities, and ratios, as seen later.

2 Model

We consider a region Ω that is endemic to COVID-19, occupied by humans of a constant total population N . The population is subdivided into several compartments based on the infection status and severity. The first compartment is *susceptible* S , determining a group of humans who are free from the virus but vulnerable to infections. An S -individual can have a first contact to the virus and it takes a few days from such an onset until symptoms appear, called *incubation period* θ^{-1} [d], during which the person belongs to the *exposed* compartment E . As COVID-19 also contracts asymptomatic cases [17], at the termination of the incubation period, the person can still be asymptomatic A or symptomatic I depending on the immune response (fitness), age, and possible co-morbidities [1]. We assume that the asymptomatic cases share the proportion of a to the exposed cases. An asymptomatic case goes further to either of the two cases: *detected* pA or *undetected* $(1-p)A$. Determination of the average probability measure p is heavily contingent on whether the person initiates a self-report to the medical department or not. Having been designated infection-positive, the asymptomatic person initiates self-quarantine at a certain rate. When the person undergoes a symptomatic infection, we assume that (s)he is automatically directed to a hospital. At this stage, all humans occupying pA and I are reported in the media as *active cases*. The moment when the infection ends, all the corresponding active cases become *closed cases* in which infection-free humans either recover or die, who are then compiled into R and D compartment, respectively.

Let us assume that no short movements from A - to I -individuals or vice versa during the illness periods are expected and that mild health impacts on asymptomatic cases produce no deaths. Our model then reads as

$$S' = -\beta \left[1 + \frac{N_h - 1}{H_C} \right] S \cdot p_A A - \beta \cdot S \cdot p_I I - \alpha \cdot S \cdot \int_{t-\sigma}^t p_{AA}(\xi) d\xi,$$

$$\begin{aligned}
 E' &= \beta \left[1 + \frac{N_h - 1}{H_C} \right] S \cdot p_{AA} + \beta \cdot S \cdot p_{AI} + \alpha \cdot S \cdot \int_{t-\sigma}^t p_{AA}(\xi) d\xi - \theta E, \\
 A' &= a \cdot \theta E - \gamma_A A, \\
 I' &= (1 - a) \cdot \theta E - \gamma_I I, \\
 R' &= \gamma_A A + (1 - \mu(t)) \cdot \gamma_I I, \\
 D' &= \mu(t) \cdot \gamma_I I.
 \end{aligned} \tag{1}$$

The parameter β [$\text{ind}^{-1} \cdot \text{d}^{-1}$] denotes the infection rate related to direct transmission, containing some parameters driving the virus transmission, two of which are the minimum contact radius r and the average number of daily cross-household encounters between a susceptible and an infected person k_{ch} (see model derivation in Appendix A). The other infection rate α [$\text{ind}^{-1} \cdot \text{d}^{-1}$] relates to the fomite transmission. The integral term gives the accumulated contaminated surfaces attributed to non-quarantined and undetected asymptomatic cases. The parameters N_h and H_C denote the average number of household members in Ω and regional level by the number of households for which the average encounters k_{ch} is possible, respectively. The probabilistic term $p_A := \nu p + (1 - p)$, where ν^{-1} is the quarantine duration in the day scale, denotes the proportion of non-quarantined and undetected asymptomatic cases. Finally, γ_A^{-1} [d], γ_I^{-1} [d], p_I , σ [d] denote the average illness periods for A -individuals and I -individuals, the correction factor for within-hospital encounters, and the maximum duration of virus survival on a surface, respectively.

One can easily verify that every unit normal of the boundary of the nonnegative orthant in \mathbb{R}^6 possesses a right-to-obtuse angle against the vector field of (1) in the corresponding boundary. This elucidates nonnegativity of the system trajectories for all possible parameter values, invoking a certain biological meaningfulness. The model also annexes practical relevance as the recovered R and death compartment D follow the cumulative mode owing to the fractional online case fatality rate $\mu(t)$. The initial conditions for all the state variables are either zero or positive depending on the initial time $t = 0$ taken.

2.1 Case fatality rate

Case fatality rate (CFR) usually is used not only to measure the deadliness of a certain infectious disease but also to describe the general patients' fitness as well as the tenacity of the health system in the observed region. A blunt connection to our modeling study is the fact that CFR dissevers the proportions of recovered and dead humans after average illness periods γ_A^{-1} for A -individuals and γ_I^{-1} for I -individuals. The actual CFR is to be calculated when the entire course of epidemics is completed, leaving out *only* closed cases (deaths and recoveries). Then, the value takes the ratio between total deaths and closed cases [18]. If the disease is still ongoing, computation of CFR becomes tricky. The widely-used estimate for its simplicity is called *naïve estimator*, calculated based on the ratio between accumulated deaths and confirmed cases over the entire history up to the current time. This ratio is biased and misleading, given that the outcome of part of the denominator is unknown, making the denominator grows much faster than the numerator [19]. This fallacy has invited (statistical) researchers to come up with alternatives in the modification of the numerator and denominator; one claims advantages against the others. Speculations are based on whether concentrating on active cases and/or just closed

cases [18, 19], age-dependence [11], time delay in between the real incidences and report [20], and different countrywise treatments in recording incidences [21].

Our present study is not in the business of coming up with a correct, unbiased CFR. The fitting is heavily bound to data with no requirement for secondary analysis. Therefore, our task is to identify how the corresponding publisher reports the data in terms of estimating deaths and recoveries. In the model (1), the portion $\gamma_I I$ who move out from hospitals per day are distributed to deaths and recoveries based on the online case fatality rate $\mu(t)$. Unlike usual CFR measures that relate to accumulated masses, this $\mu(t)$ corresponds to the division into deaths and recoveries only at the given time. These daily data can be found from the difference of the closed cases at the present and those at the previous time. If both accumulated deaths and recoveries do not advance after a day before, then we assign $\mu = 0$ on that day. Finally, our prediction interprets how the data in the publisher would appear, not how the actual cases would behave.

2.2 Mass control

Many governments worldwide have ordered their citizens to stay at home and cancel nonessential businesses to reduce the risk of infection. Curfew, lockdown, workplace clearance, and school closure are colloquial terms that converge to the same aim: avoiding mass gathering. Only then every term is uniquely determined by certain types of action. For example, lockdown is always associated with restricting people's movement (avoiding going out from home), transportations, and business operations. In this study, we model these actions as to reduce the average number of daily cross-household encounters k_{ch} . Added with law enforcement, we assume that people react in an impulsive way following the government's order as per the commencement of either of those actions. The actions remain in effect unless they are closed by the government. We can then model β as a piecewise constant function over time, for a given minimum contact radius r . The constancy remains when there is no new action introduced in the population. Additionally, attention must be paid to the fact that the effect of a certain action can only be seen from incidence data given some time delay for people to get admitted to hospitals (such that they can be recorded) and for the recorded data to appear in the media. Such time delays can be fixed according to the incubation period.

3 The basic reproductive number

The key idea in calculating the basic reproductive number lies in finding a certain threshold at the interface of different stability statuses of the disease-free equilibrium [22]. The basic reproductive number related to our model (1) is given by

$$\mathcal{R}_0 := \frac{(\beta[1 + \frac{N_h - 1}{H_C}] + \alpha\sigma)p_A N a}{\gamma_A} + \frac{\beta N p_I (1 - a)}{\gamma_I}. \quad (2)$$

For details on the derivation, see Appendix B. The first term in the formula (2) represents a certain proportion $(\beta[1 + \frac{N_h - 1}{H_C}] + \alpha\sigma)p_A N a$ of new infection cases in the susceptible compartment attributed to an A -individual happening per an illness period γ_A^{-1} . Observe that both direct and fomite transmission play roles in increasing the proportion, which is enhanced as the virus would have survived longer (large σ) and there are more humans in the standard household (large N_h). Surely is the transmission compromised as the proportion

of asymptomatic cases is negligible (small a). The second term represents a similar contribution from an I -individual in a hospital, which of course is regulated by the probabilistic term p_I and their proportion $(1 - a)$.

Two important remarks are enumerated. First, the formulation of the basic reproductive number is restricted to the assumptions inherent from the earliest take-off period, i.e., $\mu = 0$ and $R(t) = D(t) = 0$ for $-\sigma \leq t \leq 0$. Therefore, the formula cannot afford its use on which the disease already advances through returning more deaths and intervention measures are overwhelming. Nonetheless, this early estimation is important for the understanding of the disease progression without any way of obstruction. How this value changes thus relies on the magnitude and tenacity of intervention measures taken in the next step. Second, the role of household structure is now more apparent. Two countries of the same choice of the number of local households H_C were predicted to exhibit different numbers of cross-household encounters k_{ch} . When both high-risk and low-risk country are similar in other parameters except k_{ch} and N_h —which is almost the usual case—the values of βN from both countries should be significantly different: the high-risk country takes a larger value than the low-risk country. Amplified by N_h , the basic reproductive number in the high-risk country becomes much larger than that in the low-risk country.

4 Study areas

4.1 Germany

The first COVID-19 confirmed case was reported on January 27, 2020 in Bavaria. On March 4, Germany alarmed on “extraordinary urgency” and appropriate attitudes against chaos, prohibiting exports of protective equipment and taking actions to order more. On March 16, Germany followed measures imposed to Bavaria on public movement that includes restricted movements in sports, religious meetings, playgrounds [23]. The chancellor of Germany announced a contact ban on March 22 late afternoon. On April 4, 200,000 stranded German vacationers return to Germany [24].

4.2 Sri Lanka

The first COVID-19 positive case reported in January 27, 2020 in Sri Lanka was a tourist from Hubei province, China. There were no reported cases until a first Sri Lankan was tested positive on March 11 [25]. Then, the government took immediate actions to prevent further transmission by first closing the schools on March 12 followed by declaring public and mercantile holidays from March 16 and announcing work from home since March 20 [26]. All passenger arrivals were banned on March 18 as all preliminary cases were reported among those returnees from Europe, Middle-East and India. All returnees have been undergoing compulsory 2-week quarantine at designated centers operated by military under the instruction of health authorities. Those who completed quarantine have been further requested to undergo another 2-week self-quarantine too. Most critical restriction in public level was the island-wide curfew imposed from March 20 [27]. Although the curfew has been lifting time-to-time, several districts including highly populated Western province have been under continuous curfew [26].

5 Numerical implementations

5.1 Problem definition

The data used in our investigation are collected by Johns Hopkins University [28], which can be accessed via a GitHub repository [29]. Three types of data are available: the accu-

mulated number of deaths reported on a daily basis (Dead), accumulated recovered cases, and accumulated confirmed cases. The latter fold both accumulated deaths and recovered cases as well as active cases (Active) represented by $pA + I$ in our model; see also further explanation in [30]. Observe that the model (1) already governs the timely accumulated deaths due to the absence of outflows from the compartment. Let us denote t_s, t_f as the starting and final time for the assimilation, which are flexibly chosen according to one's future need. Then, our aim is to find a set of parameter values ρ that solves the following problem

$$\min_{\rho} \|p\hat{A}(\rho) + \hat{I}(\rho) - \text{Active}\|_{\ell^2(t_s, t_f)}^2 + \omega \|\hat{D}(\rho) - \text{Dead}\|_{\ell^2(t_s, t_f)}^2,$$

where \hat{A} 's, \hat{I} 's, and \hat{D} 's behavior are regulated by our model and ρ . The hat symbol, e.g. \hat{I} , represents the numerically approximated value of I -individuals from the model, picked on a daily basis. Note that in one time scale [d], the numerical approximate also comprises evaluations at different time points in between. The symbol $\ell^2(t_s, t_f)$ denotes the usual least-square for discrete entities. The regularization parameter ω accounts for the trade-off between the fitting to active cases and that to deaths. Typically, ω is driven by a belief to what extent either of the two datasets is more reliable than the other. The balance is gained from training the two expressions in the objective function such that

$$\omega \approx \frac{\|p\hat{A}(\rho) + \hat{I}(\rho) - \text{Active}\|_{\ell^2(t_s, t_f)}^2}{\|\hat{D}(\rho) - \text{Dead}\|_{\ell^2(t_s, t_f)}^2}. \quad (3)$$

More details of the computation are provided in Sect. 5.2.

On the one hand, to what extent the infection rate β can make the model solution fits given data depends on some parameters, see (4) in Appendix A. Covering unknowns, the infection rate has always been a good candidate for the fitting parameter [31]. On the other hand, it still remains unclear if fomite transmission would also lead to new infection despite apparent indication of virus survival on surfaces [32]. Good empirical evidence was reported by Centers for Disease Control and Prevention (CDC). They confirmed that COVID-19 RNA was found on various surfaces in the Diamond Princess's cabins of both symptomatic and asymptomatic cases even up to 17d posterior to clearance [33]. Given this fact, human–virus contact tracing technically is an infeasible task, in which case the explanation continues with

“... data cannot be used to determine whether transmission occurred from contaminated surfaces.”

We just see that the other infection rate α also deems unobservability. A relation between β and α is as yet beyond far from trivial.

The proportion of detected cases p for the case of China has been estimated around 14% [34]. However, the value might vary from country to country due to different social awareness and mentality. Moreover, the proportion of asymptomatic cases a is naturally unobservable notwithstanding the previous assumption on age structure, i.e., that a contains information related to the ratio between young and older humans. The correction factor in the virus transmission between medical staffs and patients p_I is also unknown. Recall

that γ_A^{-1} and γ_I^{-1} define the average illness periods spent by asymptomatic and symptomatic cases respectively. All asymptomatic and hospitalized persons most likely leave their infection history uniquely, depending on their fitness, age, and available medical access. This information is both personal- and country-sensitive, in which case the average illness periods may also turn unobservable. In addition, the initial conditions for the exposed compartment $\eta_E(t)$ and asymptomatic compartment $\eta_A(t)$ where $t_s - \sigma \leq t \leq t_s$ are also unknown. We then chose the available data values in the past for the hospitalized (based on active cases, $\text{Active} - p\eta_A$), recovered η_R , and death compartment η_D . The susceptible compartment clearly follows $\eta_S = N - \eta_E - \eta_A - \text{Active} - p\eta_A - \eta_R - \eta_D$. The fitting parameters in this study are then summarized as

$$\rho = (\beta, \alpha, p, p_I, a, \gamma_A, \gamma_I, \eta_E(t), \eta_A(t)), \quad t_s - \sigma \leq t \leq t_s.$$

Finding optimal time-dependent parameters can be exhaustive since the number of actual parameters in the final discrete representation of the problem is dependent on the discretization taken. We avoid a possible curse of dimensionality by specifying parameter values on a daily basis, i.e., where historical data are defined, and determining those at the time points in between using the spline interpolation.

5.2 Solving the parameter estimation problem

This study engineers a multi-parent genetic algorithm without the selection procedure [35]. All the fitting parameters are represented via binary blocks (whose lengths depend on the boundary values), then concatenated to form a long binary chain representing a “player”. In every iteration, 50 players for ρ were used and the regularization parameter ω associated to the minimal objective function value is evaluated via (3). The mean from the whole iterations is then used to update the ω -value in the objective function, until then iterations are repeated. We then stop the iteration when the best objective function value gets stagnant after 25% of the total iterations that lead to the last constant value.

6 Results and interpretations

6.1 Fitting results

To accompany our modeling, we declare known parameter values in Table 1. The number of time grids in the domain for initial condition $[-\sigma, 0]$ was chosen 300, meaning that every day we have 100 solution points with the starting point indicating the day. The time step used is then $h = 0.01$. For the comparison with data, we solely assign the time points that correspond to the days. The number of local households was specified 100 for both Germany and Sri Lanka. Based on the average number of household members N_h , Sri

Table 1 Known and pre-specified parameter values

Par	Description	Unit	Range	Germany	Sri Lanka	Reference
N	total population as of April 12, 2020	[ind]	–	83,726,366	21,395,610	[36]
N_h	average number of household members	[ind]	–	2.1	3.9	[37, 38]
ν^{-1}	quarantine period	[d]	5.2–20	14	14	[39]
θ^{-1}	incubation period	[d]	0–20	5.2	5.2	[6]
σ	duration of virus survival on a surface	[d]	0–9	3	3	[32]
H_C	number of local households	–	–	100	100	
M	number of time grids in $[-\sigma, 0]$	–	≥ 3	300	300	

Lanka and Germany resemble the two dummy countries portrayed in Fig. 8. It is then expected that in a neighborhood of 100 households, Sri Lankans can easily encounter strangers/neighbours as compared to Germans or in particular, $k_{ch}^{Sri Lanka} > k_{ch}^{Germany}$. Unless irregularities are prevailing from the other parameters covered by β , we hypothesize that $\beta^{Sri Lanka} > \beta^{Germany}$.

The unknown parameter values must be equipped with feasible ranges that are not too narrow (shorter binary chains or fast computation, but suboptimal) or too wide (slow computation, more optimal). We have designated for the earliest take-off period that $0 \leq \beta, \alpha \leq 3/N$, according to numerical experiments. Soon after a new action is introduced in the population, we assume that both β and α decrease reflecting people’s awareness and means of obeying government’s order. All the probabilistic terms must satisfy $0 \leq p, p_I, a \leq 1$ for all time. According to [40], the median of hospitalization duration for the case of patients in China was 12d. For our numerical investigation, we impose $1 \leq \gamma_A^{-1} \leq 9$ and $9 \leq \gamma_I^{-1} \leq 20$ so the latter may vary around the cited 12d.

A phase is taken until there is a new countrywide action introduced in the population and its effect can be seen a few days after the commencement. We took 5 days for such a delay, which is approximately the incubation period. Citing Sect. 4, we divide the time domain into four phases for Germany and three phases for Sri Lanka, each of which represents a subdomain where a certain action remains in effect, before added by another one. The summary is presented in Fig. 1.

The fitting results for Germany and Sri Lanka can be seen in Figs. 2 and 3, respectively. Notwithstanding different numerical values yielded on every phase, our results support the hypothesis that $\beta^{Sri Lanka} > \beta^{Germany}$ for all the overlapping time. The role of fomite transmission, which is steered by α , is much pronounced in Germany especially on earlier phases as compared to Sri Lanka. Our results allow us to argue that Germans could get infected through the comparable roles of direct and fomite transmission, whereas Sri Lankans are more overwhelmed by direct transmission through sneezing. The proportion of detected cases p for Germany (0.7976) is generally higher than that for Sri Lanka (0.5549). The correction factor related to the virus transmission from patients into medical staffs p_I generally starts from a small value during the earlier take-off period (0.0276 for Germany, 0.0048 for Sri Lanka). A hypothetical reason is due to a limited number of COVID-19 patients. Then, it waves to a significantly larger value (almost 1) and decreases thereafter. This means that, the within-hospital transmission entails almost no difference

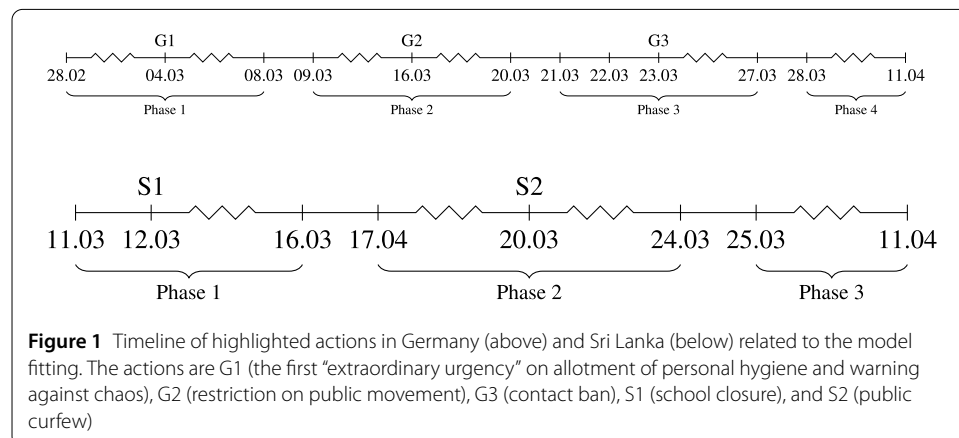
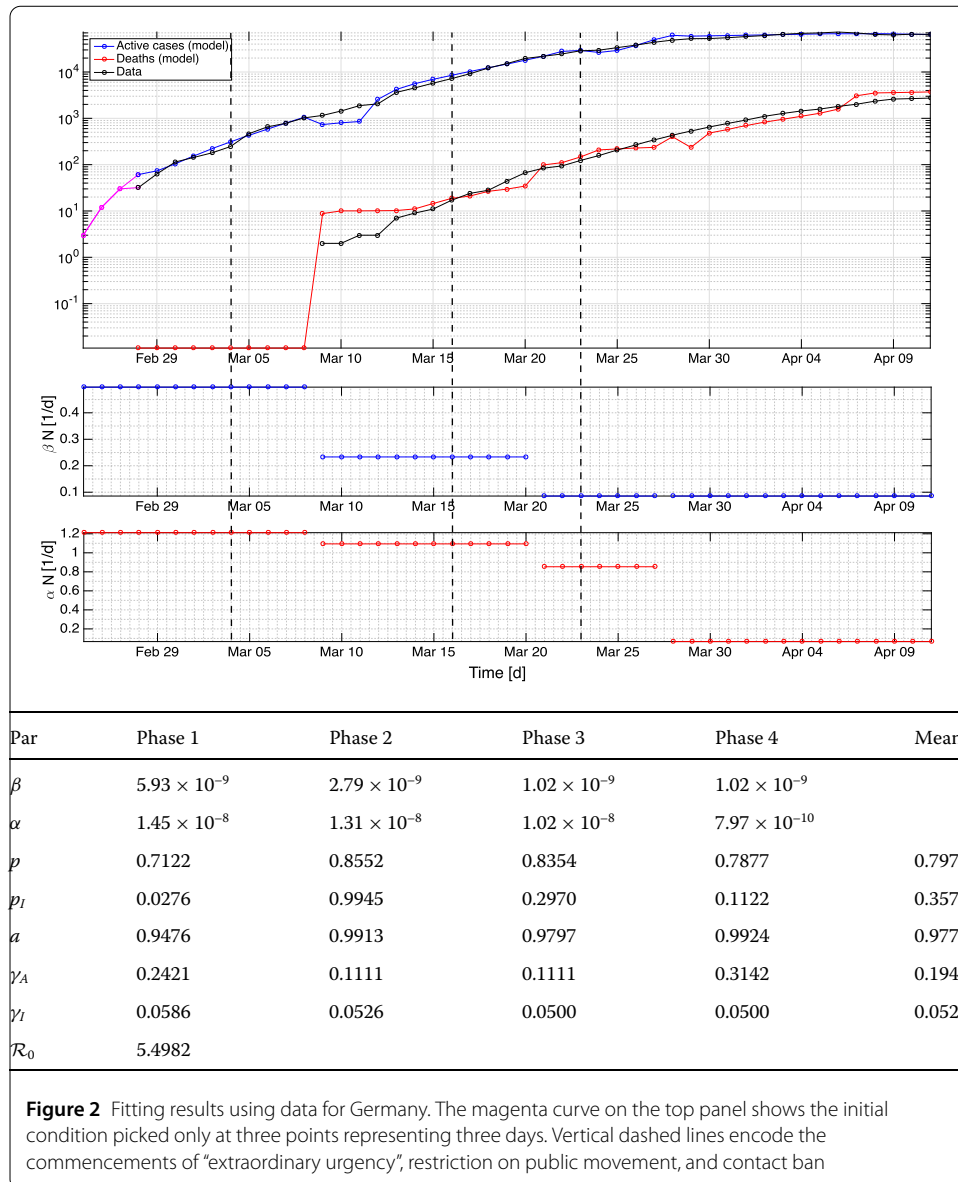
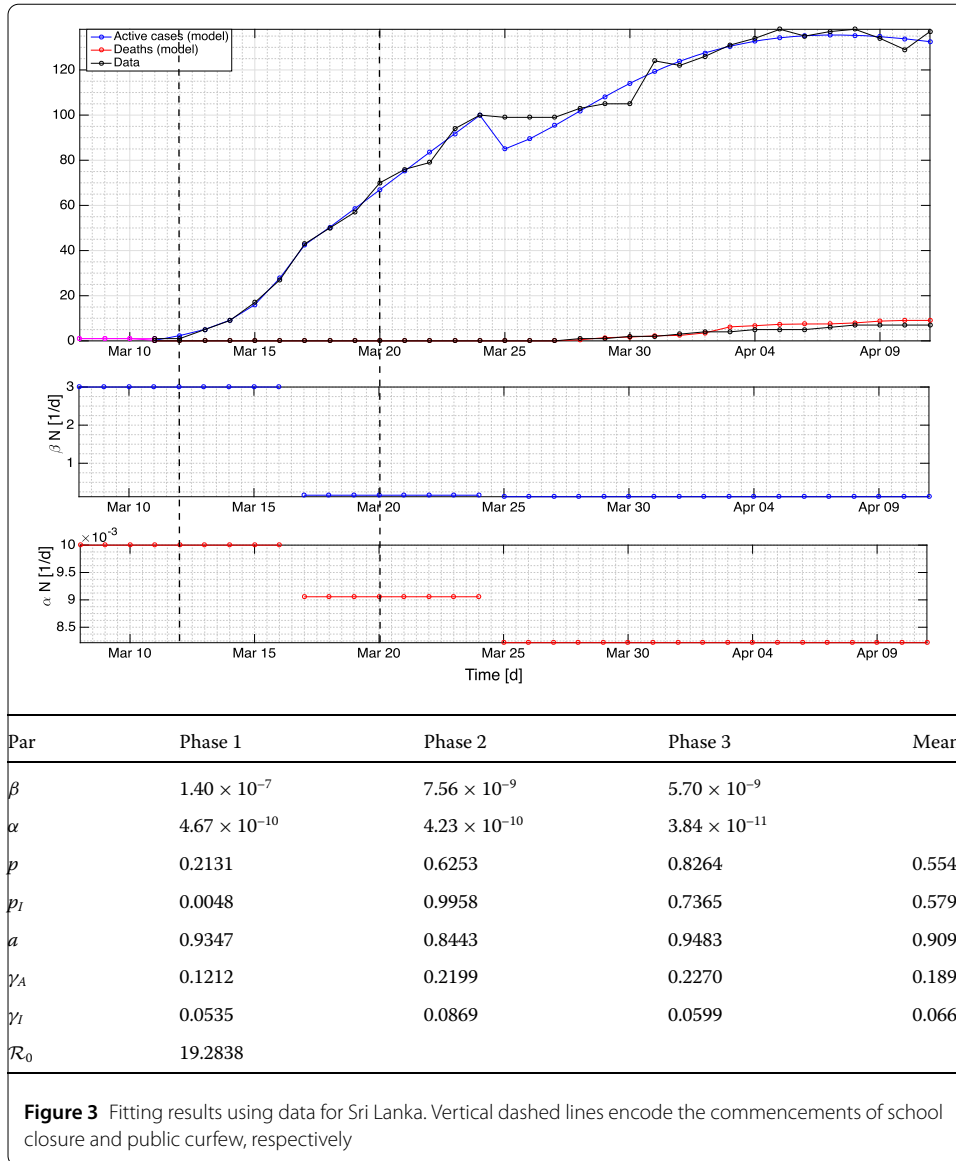


Figure 1 Timeline of highlighted actions in Germany (above) and Sri Lanka (below) related to the model fitting. The actions are G1 (the first “extraordinary urgency” on allotment of personal hygiene and warning against chaos), G2 (restriction on public movement), G3 (contact ban), S1 (school closure), and S2 (public curfew)



than the usual susceptible–asymptomatic transmission outside of hospitals in the later phases. The mean values (0.3578 for Germany, 0.5790 for Sri Lanka) lead to the hypothesis that Sri Lankan medical staffs pose higher risk of getting infected in hospitals. A possible reason may be related to the reported irresponsible behavior of several COVID-19 patients who have been admitted to the hospitals without providing real information about their symptoms, historical whereabouts, and close connections to COVID-19 exposed individuals and even hiding from the health officials in backtracking [41–43]. Moreover, the proportion of asymptomatic cases a is generally high for both countries (0.9778 for Germany, 0.9091 for Sri Lanka). This number weighs the exposed humans to be assigned to either the asymptomatic A or symptomatic I compartment in the optimal way such that the active cases $pA + I$ match with data. This process is done without scrutinizing what happen to the A and I compartment individually. This modeling study owes a possible correction to the value of a due to the unavailability of the hospitalized data for I



with exact timing of admission and discharge. In case of presence, the objective function should be adjusted by including the hospitalized terms, and a should account for the tradeoff between active and hospitalized cases. Moreover, the findings on recovery rates lead to the hypothesis that the asymptomatic Germans suffer from illness for 5.14d on average, while the symptomatic ones for 18.94d. The asymptomatic and symptomatic Sri Lankans suffer from illness for 5.28d and 14.97d, respectively. Finally, the values of \mathcal{R}_0 were computed only on the earlier take-off period. As mentioned in Sect. 3, the different household structures in Germany and Sri Lanka were shown to influence \mathcal{R}_0 significantly. Our numerical results confirm this. Moreover, it should be noted that our \mathcal{R}_0 was based on aggregating two proportions of new infection cases in the susceptible compartment from an undetected/non-quarantined A -individual and an I -individual, each per the corresponding illness period. Therefore, such a definition is not to be compared with the conventional geometric multiplier on the occurrence of infection cases per day attributed to a single infected person in a “virgin” population. For some more references on studies

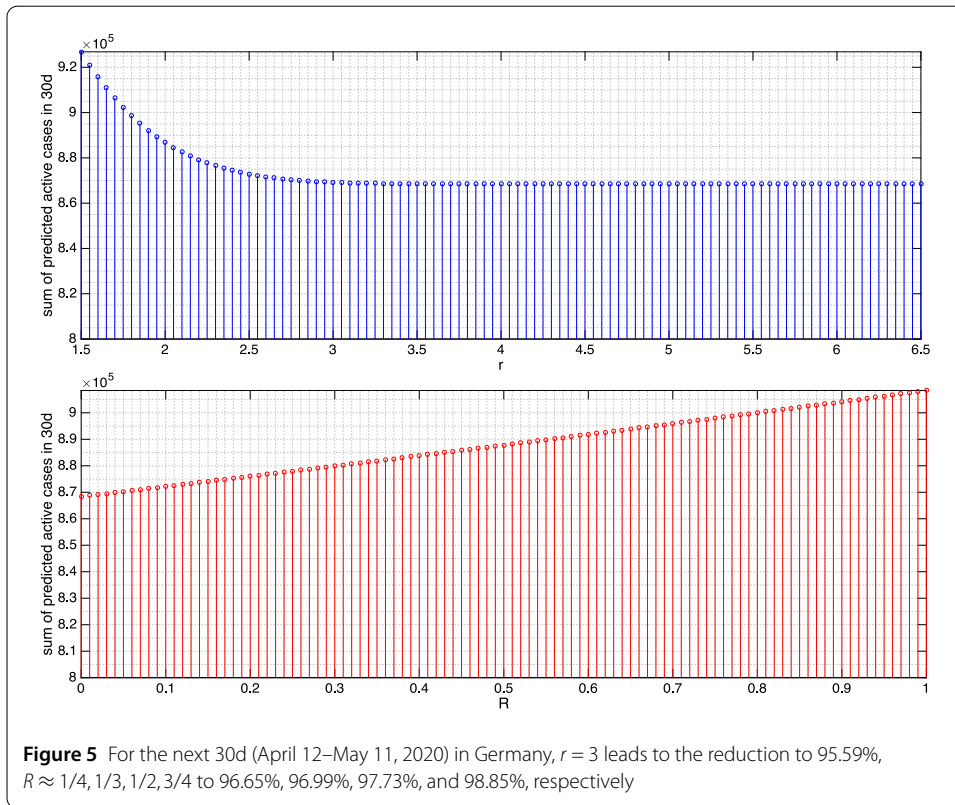
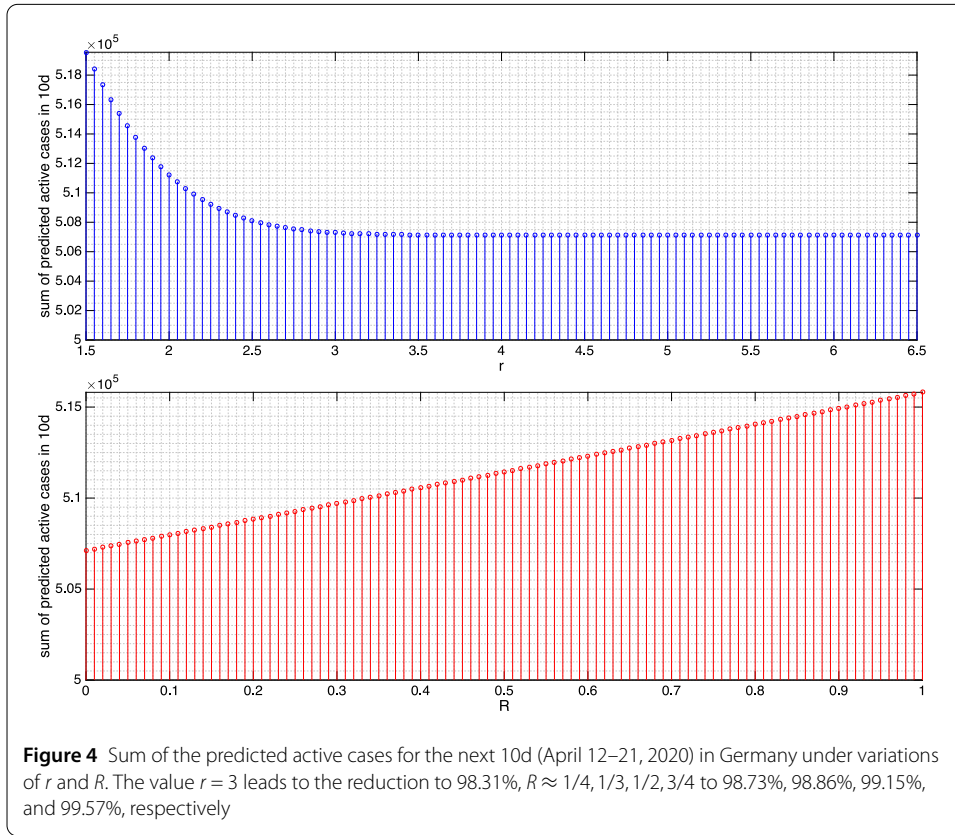
related to COVID-19's \mathcal{R}_0 , see [7, 44]. We shall also note that \mathcal{R}_0 value is both model- and data-sensitive. At least, there has been no general proof that any distinct SIR-type models fitting equally well to the same data would produce identical values.

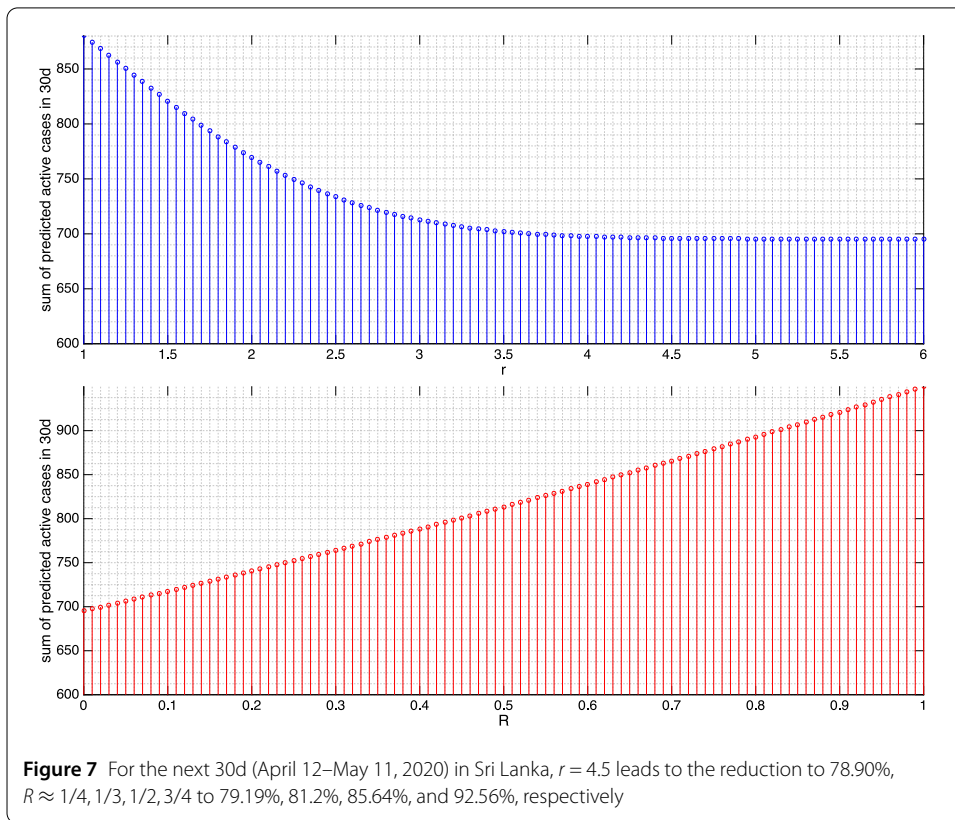
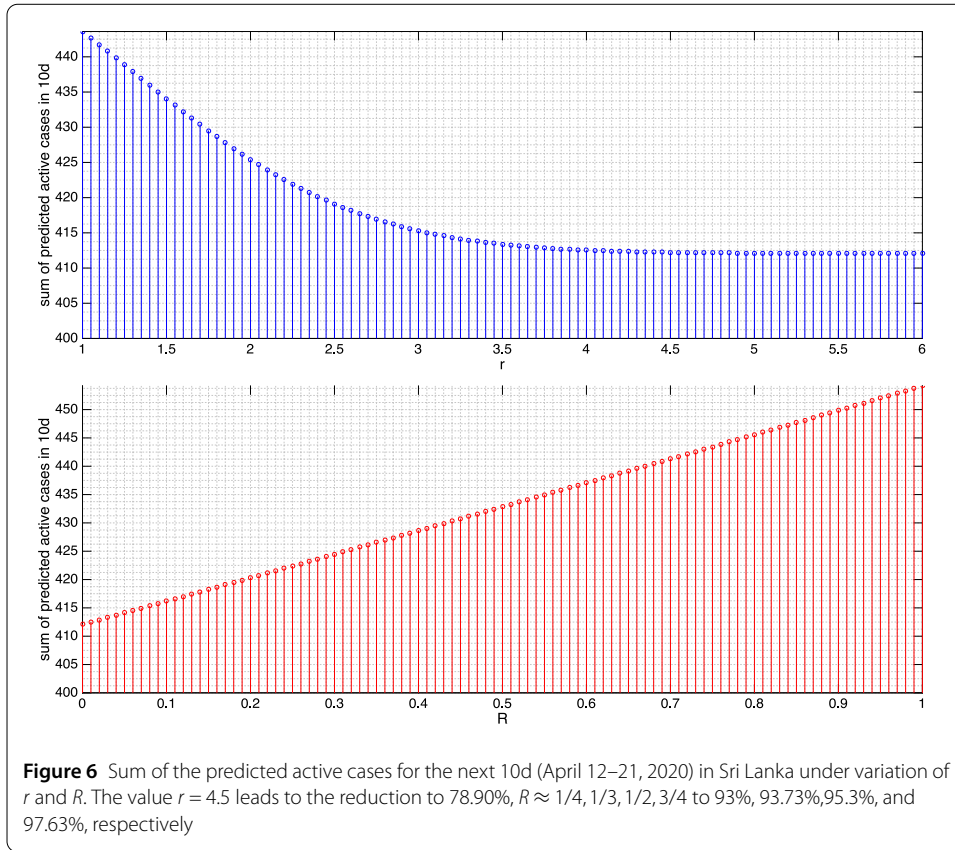
6.2 Readjustment of physical distancing and mass control

In Germany, public curfew and the application of 1.5 m minimum contact radius were officially announced on March 22, 2020 late afternoon, commencing effectively the day after and remaining in effect five days later according to our hypothesis, which is the beginning of Phase 4. On the other hand, Phase 1 represents the earliest take-off period, where citizens were not bound to any official regulation regarding the physical distancing, inducing the minimum contact radius $r = 0$. We can now compute the deviation β_1 in (5) in Appendix A the identity $\beta_{\text{Phase 4}} = \beta_{\text{Phase 1}} \cdot \exp(-1.5^2/2\beta_1^2)$, returning $\beta_1 \approx 0.7997$. For Sri Lanka, official announcement regarding the public curfew was issued on March 20, 2020 [27], which lies in Phase 2. The minimum contact radius was assigned 1 m in that announcement. We can then use a similar identity $\beta_{\text{Phase 3}} = \beta_{\text{Phase 2}} \cdot \exp(-1^2/2\beta_1^2)$ to reveal $\beta_1 \approx 1.3299$.

As the minimum contact radius is related to physical distancing, we seek to find a certain rule of thumb related to mass control (public curfew, lockdown, school closure, workplace clearance). As discussed in Sect. 2.2, the control can be implemented through reducing the cross-household encounters k_{ch} . We introduce a ratio $R := k_{\text{ch}}^{\text{new}}/k_{\text{ch}}^{\text{old}} = \beta^{\text{new}}/\beta^{\text{old}} \in [0, 1]$ where $\beta^{\text{old}} = \beta_{\text{Phase 4}}$ for Germany and $\beta^{\text{old}} = \beta_{\text{Phase 3}}$ for Sri Lanka, also $R = 1$ directs not to impose any reduction on β^{old} . As k_{ch} denotes the 'average' number of cross-household encounters leading to possible infection, R can determine reduction on the intensity of all individual activities. For example, $R = 1/2$ means that every household member who goes for shopping twice a week must reduce to once a week, goes to office work six days per week must reduce to three days, does sport/outdoor exercise four times a week should change to twice a week, gets out from home twice a day changes to once a day, and so on.

The results from varying r and R in changing the sum of predicted active cases can be found in Figs. 4 and 5 for Germany and Figs. 6 and 7 for Sri Lanka. Several findings are highlighted. We did not attempt to increase r to be arbitrarily large as, apparently for both 10d- and 30d-prediction, the sums remain essentially constant after $r = 3$ m for Germany and $r = 4.5$ m for Sri Lanka. In the context of our model, it is hypothesized that the minimum contact radius of 1.5 m in Germany and 1 m in Sri Lanka can still be improved to 3 m and 4.5 m, respectively, in order to gain significant reduction in the number of active cases. We understood that even though β can be reduced to zero by means of a lockdown, the residues in the portions of exposed and asymptomatic as well as symptomatic compartment eliminate exponentially but hardly jump to zero in an impulsive manner. This is why we can only afford reduction to the least of 95.59% for Germany and 78.90% for Sri Lanka. In Germany, increasing the minimum contact radius to 3 m is equivalent in the outcome of predicted active cases to giving the intensity reduction ratio 1/4. This means that, every person who cannot take up 3 m minimum contact radius can reduce his/her intensity of daily/weekly activities to 1/4 or vice versa. In Sri Lanka, predictions of the active cases in the next 10d and 30d give significantly different outcomes. For a short term (10d), extending the minimum contact radius to 4.5 m gives more significant reduction (by 78.90%) as compared to giving the intensity reduction ratio 1/4 (93%). For a longer term (30d), both actions are equivalent in terms of the outcomes.





7 Conclusion

A micro-level approach on individual behavior and household structure in viewing COVID-19 transmission is proposed in this work. It caters linking infected status in more realistic way before compromising into macro-level. For instance, sneezing and surface touching are treated in a detailed manner. Meanwhile, we design the infection rate β in a unique way by amalgamating many factors relevant to direct transmission through sneezing. Manipulation of a decreasing function β of the minimum contact radius r bridges the scientific argument in propounding the physical distancing. Additionally, several notable factors are taken in modeling fomite transmission through surface touching. We enumerate the final effect of these as an accumulation over a period σd in the past that asymptomatic individuals walked around and freely sneezed. Moreover, with the aim of tolerating differences to and fro households, we introduce the effect of different household structures into the model involving the number of local households and the average number of household members. This leads to easier identification of low-risk and high-risk country in terms of the basic reproductive number on the earliest take-off period. Our model solutions using the final attribution (1) are in quite good agreement with available data from Germany and Sri Lanka on early infection. Unobservable effects that may represent reality, for example increasing minimum contact radius and reducing the intensity of individual activities, have been hypothesized. Both measures can also be applied in combination flexibly, in the sense that those who cannot stick to one can apply the other. For some alternating choices between actions and outcomes, reduced magnitudes of the intensity of individual activities to $1/3$, $1/2$, $3/4$ are also presented.

We shall also highlight several limitations and outlooks from our current modeling. First, as previously cited, the model of fomite transmission has not been accompanied by field evidence indicating if the transmission indeed happens. The corresponding functional structure however mediates other possible transmission mechanisms which can be due to asymptomatic persons, but in a more extended version where distance no longer matters, for example aerosol transmission. Due to lightweight sneeze micro-particles (those $<5 \mu\text{m}$ in size), scientists argue that aerosol transmission intensifies indoor virus transmission [45] and related infection cases from poorly ventilated rooms have been evidenced [46]. Whether this concept conflicts with the distancing measure emanated from direct transmission will be a factor that needs to be analyzed and treated properly. The question is, whether aerosol transmission has actually replaced the direct transmission in its role as the main transmission of the disease. In this sense, increasing minimum contact radius during close contact to reduce infection probability is no longer relevant. Second, we realize that the data used in this study as well as corresponding results will not infer a longer period ahead. In the sense that data jump to larger values after April 11 than predicted, our findings should remain actual and the actions might even need emphasis. When the data devolve to lower values, we have been let to observe based on the historical progression of the active cases that the jump cannot be so radical as well. Third, the model has not been compared with the use of personal protective measures. Translating possible measures such as wearing mask in more secure way into equal reduction on β without having to stay away from friends/colleagues and minimize outdoor activities could be a possible outlook. Intensive cleaning of surfaces is also an additional intervention to mitigate fomite transmission. Fourth, due to the model complexity and a relatively large extent of unobservable parameters and data, we cannot afford Bayesian inference

that should have mediated robustness of the optimal parameter values against data or model uncertainty. As a result, no confidence intervals of the parameter values are presented. We have infused prior knowledge for the ranges of the parameter values and used a genetic algorithm where, although no guarantee for finding the global minimum, there is also no analytical evidence that it always converges to a local minimum (whereas this is an inherent feature of gradient-based methods). Our future consideration includes model specification toward complexity and data extent reduction to provide a trade-off between reasonable prediction and proper Bayesian inference.

Appendix A: Model derivation

Since studies employing SIR-type models are ubiquitous, it brings no cost to revisit how such a typical model physically interprets. Extended details will help readers walk through the model derivation, such that they are able to critically assess the assumptions and possible caveats.

A.1 Direct transmission

We assume that all the human individuals are homogenous, meaning that those with certain ages and social ranks are not treated differently than the others. The time scale (unit) is assumed to be 1 day [d]. The window of observation is considered short enough that the resultant of births and natural deaths is ignorable. Focusing on certain countries, we seek to simulate our model starting after global travel restriction, where migrations and imports in the population are negligible. Therefore, the total human population is considered constant. We assume that all the undetected individuals $(1 - p)A$ are free to move around like the susceptible humans, while those who undergo quarantine among pA are not. All the I -individuals clearly stay in hospitals. Infection happens when a susceptible person unrelated to medical department encounters an A -individual. A susceptible medical staff can also get infected through contact with an I -individual. Suppose that all the A - and I -individuals express certain sneezing behavior during infection with a uniform *sneezing rate* s [d^{-1}]. This number determines how often they sneeze per day on average. Other important parameters for the successful virus transmission are an *effectivity measure* e_s , determining how probable a sneeze (respiratory droplets) leading to a successful infection, and a *probability measure* p_s , determining how probable an infected person sneezes during contact with a susceptible person. Another parameter leading to successful infection is related to how proximate a susceptible and an infected person during contact are. We define a decreasing function $\phi(r)$ of the minimum contact radius r [m] describing how essential certain contact radii leading to the transfer of water droplets from an infected to a susceptible person. Moreover, we note that A/N determines the probability of a susceptible person to meet with an infected person from A , equalling to the number of A -individuals divided by the total population. The frequency multiplied by the probability, $S \cdot A/N$, determines the 'expected' number of encounters that would probably lead to infection. This number, however, assumes that every S -individual encounters an A -individual only at maximum one time per day. If k denotes the average number of encounters with A -individuals experienced by any S -individual, then $k \cdot S \cdot A/N$ denotes the average number of encounters that possibly cause infection. This seemingly large k (can also be greater than N/A) is neutralized by p_s , and essentially the remaining parameters mentioned earlier, such that not

all S -individuals experience infection per day. Therefore, attributed to direct transmission is the loss of individuals from the susceptible compartment S of daily magnitude

$$\beta(r) \cdot S \cdot (A + p_I I), \tag{4}$$

where $\beta(r) := s \cdot e_s \cdot p_s \cdot \phi(r) \cdot k \cdot 1/N$. One can think of the importance of $s, e_s, p_s, \phi(r)$ by asking if infection still happens when every of them is assumed vanishing. For example, if the sneezing rate $s = 0$, then no infection is prevailing since nobody sneezes during contact. Even when an infected person sneezes frequently, no infection can be expected if the sneezing is not done in front of a susceptible person, i.e., $p_s = 0$. The appearance of p_I in (4) stems from the preserved manner and medical restrictions imposed on I -individuals in hospitals.

A.2 Model for the infection rate

The new *infection rate* β clearly follows the behavior of ϕ with respect to the minimum contact radius r . Here, the function β is modeled in the sense of substantiating certain statistical sampling. A sufficiently large number of contact cases of preferably different contact radii are pooled and the classified patients are distributed with respect to the minimum contact radius. We assume that the resulting distribution is superimposed by a normal distribution

$$\beta(r) := \beta_0 \cdot \exp\left(-\frac{r^2}{2\beta_1^2}\right), \quad r \geq 0. \tag{5}$$

The parameter β_0 defines a reference maximal infection rate obtained when the minimum contact radius r is 0. This parameter can be computed e.g. during an earlier take-off period where only a small number of humans are well-acknowledged with the aid of physical distancing. The other parameter β_1 accounts for the switch from a significant decrement ($r < \beta_1$ regime) to a slow decrement ($r > \beta_1$ regime), indicating an interface between essential and non-essential physical distancing.

A.3 Fomite transmission

Experts hypothesize that the virus transmission is attributed to not only direct transmission but also *fomite transmission*. The latter represents transmission resulted from the act of touching the former-trace of sneezes from infected persons on surfaces. A recent study suggests that SARS-CoV-2 can survive on surfaces like plastic or stainless steel for up to $\sigma \approx 3d$ [32] or, assuming similarity with SARS-CoV-1 and MERS-CoV, even as long as 9d [47]. Suppose that A -individuals in the past σd walked around and sneezed where every individual created the average total sneezing coverage $c_\sigma = s \cdot D_s \cdot c_f$ [$m^2 \cdot d^{-1}$] per day, where s was the sneezing rate. The parameter D_s [m^2] denotes the average unit area covered by a single sneeze. Samples for D_s can possibly be found on desks, chairs, beds, clothing, or door handles, cf. [33]. The correction factor c_f clears out the possibility that the same individual or different individuals sneezed on the same spot. It thus is not expected that the accumulated area of the sneezes exceed the total area $|\Omega|$. Due to restricted movement and isolation, we assume $c_\sigma \simeq 0$ corresponding to I -individuals in hospitals. With such composition, the following formula

$$\frac{c_\sigma}{|\Omega|} \int_{t-\sigma}^t A(\xi) d\xi < 1 \tag{6}$$

estimates the probability of a susceptible person in the current time t to encounter (not yet touch) the accumulated area for the former-trace of the sneezes from A -individuals from the past σ . A sample upper bound for c_σ is gained from the fact that $A \leq N$ for all time, such that $c_\sigma N \sigma / |\Omega| < 1$ implies (6). The expression in (6) multiplied by the frequency S , the rate representing the daily intensity of touching surfaces f_t , the probability of touching nose or mouth shortly after touching surfaces p_{nm} , and the effectivity measure e_s , returns the expected number of infection cases attributed to touching the former-trace of sneezes additional to the direct transmission. For brevity, we introduce a folder $\alpha := f_t \cdot p_{nm} \cdot e_s \cdot c_\sigma / |\Omega|$. Accordingly, the number of new infections in the susceptible compartment S undergoes the following correction

$$\beta \cdot S \cdot (A + p_I I) + \alpha \cdot S \cdot \int_{t-\sigma}^t A(\xi) d\xi. \quad (7)$$

Covering unknowns, it is traditional to appoint the infection rates β and α for the fitting parameters.

A.4 Assumption on age structure

Age structure of COVID-19 dynamics abides by two mechanisms as variations of exposure to the virus and variations in clinical risk. Wu et al. [15] investigate age-specific susceptibility to infections and fatality risk of symptomatic cases. Both measures showed a notable increase with age indicating highest portions for age-classes 60–69, 70–79 and >79 years. Zhou et al. [16] verify older age as a risk factor that leads to identifying vulnerable patients. They analyzed 191 patients in two hospitals in Wuhan to reveal that the median age was 56 years (IQR 46–67), while that of non-survivors and survivors were 69 (IQR 63–76) and 52 (IQR 45–58) respectively. These estimates also indicate higher exposure and fatality risk in higher ages. Verity et al. [11] also report something similar in the case fatality rate. A considerable difference was evident in two age groups <60 and ≥ 60 both from their parametric (1.4% and 4.5%) and non-parametric analysis (1.5% and 12.8%). This investigation was based on cases from both inside and outside of mainland China.

Knowing that the influence of age on the endemicity level cannot be disregarded—on the one hand—and the model can grow immensely in dimension due to age stratification—on the other hand—we infuse the information about age structure into the proportion of asymptomatic cases a . The preceding observations indicate that most symptomatic cases originate from older humans. The larger a means the more younger humans outside of hospitals spreading the virus.

A.5 Immunity

Researchers are still struggling to determine whether COVID-19 leaves out (partial) immunity to the recovered humans. The primary obstacle is as yet the acquired short course of the pandemic, as of time this paper is written. Research is still ongoing to see if COVID-19 is connected to SARS (2003) and MERS (2012) in terms of reaped immunity; if connections between both diseases extrapolate to certain behavior of COVID-19 is however not yet confirmed [48]. Considering the time obstacle, it thus is justifiable to assume that all recovered humans gain immunity during the relatively short observation in this study. As a (physical) consequence, there is no outflow from the recovered compartment directing to the susceptible compartment due to loss of immunity.

A.6 Household structure

We account for the difference in the virus transmission within and between households. The key idea is to break down all the state variables in the model into the household level. Let i denote the household index and N_h the average number of household members whose data can be accessed from the official website of the United Nations and that carrying census data in different countries [37, 38]. We assume there are H households occupying the observed region Ω . It is straightforward to see that $N = N_h H$. Like in the population level, every individual can experience being susceptible, infected, recovered, or dead. We thus have $S_i, E_i, A_i, I_i, R_i, D_i$ for all $i = 1, \dots, H$. For brevity, we denote $S := \sum_i S_i$, which also applies for the other state variables. We assume that close interactions within households define a larger number of encounters k_{wh} (wh stands for within-household) such that $k_{wh} > k_{ch}$ (ch for cross-household). Consequently, the within-household β_h and cross-household infection rate β hold the relation $\beta_h > \beta$ for all minimum contact radii r . For the sake of lowering the degree of unknowns in the model, we postulate

$$\frac{\beta_h}{\beta} = \frac{k_{wh}}{k_{ch}} \approx N_h. \tag{8}$$

The approximation (8) describes certain behavioral overtones. We suppose that k_{ch} is relatively small and less sensitive to the change in the infection status due to an apparent border imposed on strangers rather than on household members. So, we fix k_{ch} . When a standard household only contains $N_h = 1$ individual, then there is nobody in the household to encounter that would lead to new infection other than strangers. We then assume by (8) that the number of encounters per day doubles as now $N_h = 2$ individuals are present, triples as $N_h = 3$, and so on. An important remark is that this approximation may underestimate the real pictures as asymptomatic persons might still behave as if no infection occurs and if indeed so, the other members might take care of them incautiously.

As all the I -individuals stay in hospitals, it brings no relevance to cluster them based on household. However, they may still infect medical staffs. Decomposing the population on the household level gives us

$$S'_i = - \underbrace{\beta_h \cdot S_i \cdot A_i}_{\text{within-household}} - \underbrace{\beta \cdot S_i \cdot \sum_{j \neq i} (A_j + p_I I_j)}_{\text{cross-household}} - \alpha \cdot S_i \cdot \sum_j \int_{t-\sigma}^t A_j(\xi) d\xi. \tag{9}$$

The third expression on the right-hand side of the equation is manifested from touching surfaces inside and outside home. Under a sufficiently large H , the sum in part of the second and in the third expression can nicely be folded into $p_I I$ and $\int_{t-\sigma}^t A(\xi) d\xi$, respectively. Violation to such an assumption might bear a certain outlier indexed by i where $\sum_{j \neq i} p_I I_j < p_I I$. We further consider over the course of first three pandemic-months, that $A_i \ll S_i$ for almost all i . This assumption infers the situation where susceptible humans still dominate in most of the households. As a result, $\sum_i S_i A_i \approx [(1/H) \sum_i A_i] S$. Making use of the estimate (8), we can sum up the expressions in (9) for all i to have

$$S' = -\beta \left[1 + \frac{N_h - 1}{H} \right] S \cdot A - \beta \cdot S \cdot p_I I - \alpha \cdot S \cdot \int_{t-\sigma}^t A(\xi) d\xi. \tag{10}$$

The effect of household structure can now be seen from the expression $(N_h - 1)/H$, which grows with N_h . However, this expression turns to be negligible in the countrywide scale. Taking the case for Germany, we have $(N_h - 1)/H \approx 1/(4 \times 10^7)$. This negligible outcome was produced by violating our assumption that strangers with different health statuses would always have the possibility to encounter k_{ch} times a day on average, which is a serious defect. Practically, those from remote districts would have almost no chance to meet, yet they contribute in accelerations of the disease spread in their neighborhoods, which render a countrywide effect. Therefore, a correction is necessary to account for this issue. Let us denote $\Omega := \bigcup_{n \leq C} \Omega_n$ as a collection of different C localizations of the same number of households H_C and

$$\mathcal{I}_n := \{i : i\text{th household belongs to } \Omega_n\},$$

where $C = C(k_{ch})$ and essentially (Ω_n) have to be chosen such that the number of cross-household encounters k_{ch} is realistic. We acquire $H_C = H/C$ and may also denote

$$\begin{aligned} S_n &:= \sum_{i \in \mathcal{I}_n} S_i, & E_n &:= \sum_{i \in \mathcal{I}_n} E_i, \\ N_n &:= S_n + E_n + A_n + I_n + R_n + D_n, & \beta_{h,n} &:= s \cdot e_s \cdot p_s \cdot \phi \cdot k_{wh} \cdot \frac{1}{N_n}, \\ \beta_n &:= s \cdot e_s \cdot p_s \cdot \phi \cdot k_{ch} \cdot \frac{1}{N_n}, & \alpha_n &:= f_t \cdot p_{nm} \cdot e_s \cdot c_\sigma / |\Omega_n|. \end{aligned}$$

If we repeat the process as before, then we obtain a correction in (9) for $i \in \mathcal{I}_n$ as follows

$$S'_i = -\beta_{h,n} \cdot S_i \cdot A_i - \beta_n \cdot S_i \cdot \sum_{j \in \mathcal{I}_n; j \neq i} (A_j + p_I I_j) - \alpha_n \cdot S_i \cdot \sum_{j \in \mathcal{I}_n} \int_{t-\sigma}^t A_j(\xi) d\xi. \tag{11}$$

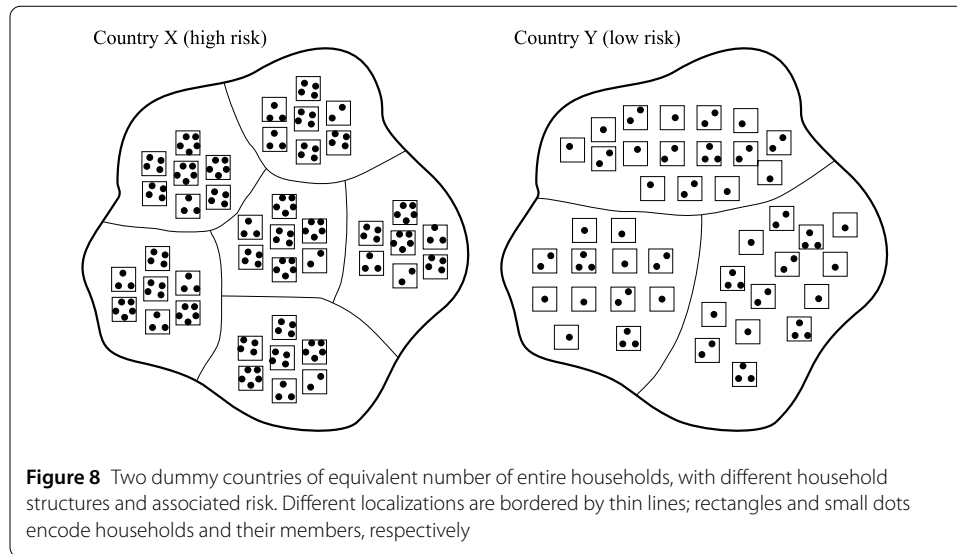
Under the dominance of susceptible humans, summing up the preceding equation over all $i \in \mathcal{I}_n$ yields

$$\begin{aligned} S'_n &= -\beta_n \cdot S_n \cdot A_n - \beta_n(N_h - 1) \cdot \sum_{i \in \mathcal{I}_n} S_i A_i - \beta_n \cdot S_n \cdot p_I I_n - \alpha_n \cdot S_n \cdot \int_{t-\sigma}^t A_n(\xi) d\xi \\ &\approx -\beta_n \left[1 + \frac{N_h - 1}{H_C} \right] \cdot S_n \cdot A_n - \beta_n \cdot S_n \cdot p_I I_n - \alpha_n \cdot S_n \cdot \int_{t-\sigma}^t A_n(\xi) d\xi. \end{aligned} \tag{12}$$

We can appoint β and α such that the summation of (12) over all n approximates

$$S' = -\beta \left[1 + \frac{N_h - 1}{H_C} \right] S \cdot A - \beta \cdot S \cdot p_I I - \alpha \cdot S \cdot \int_{t-\sigma}^t A(\xi) d\xi. \tag{13}$$

If N_n were constant for all n , then we have $(\beta, \alpha) = (\beta_n, \alpha_n)/C$, which trace back our original infection rates. When k_{ch} is predetermined, surely can designating C as to represent blocks, subdistricts, districts, cities, provinces, or a whole country be handed over to decision-makers who know better to what extent most of the local inhabitants only interact with their own such that k_{ch} is realistic. Yet, this information might still be hard to get. From the point of view of data fitting, we can simply appoint the number of local



households H_C as any constant. Then, k_{ch} adjusts accordingly together with the other parameters as we seek to optimize the folder β . There is another aspect that we can extract from (13). Suppose that we compare two (dummy) countries X, Y of the same number of entire households H , see Fig. 8. Country X has citizens who have bigger families (large N_h), like to stay around home, work nearby home, and are less dynamic in terms of meeting people other than staying home and working. This is the case where daily citizens' spatial coverage is relatively low, inducing small $|\Omega_n| \sim 1/C \sim H_C$. Country Y has more dynamic citizens of the same cross-household encounters k_{ch} (facilitated by adequate transportation system) as that of country X, also small families (small N_h), work in more distant locations from home. The larger spatial coverage induces a large H_C . In comparison, country X has a larger population as compared to country Y such that the same k_{ch} has to be acquired by country X and country Y with “less” and “much” effort, respectively. Our model (13) hypothesizes that country X poses higher risk of accelerated infection propagation due to its packed-type household structure and less dynamics as compared to country Y. The risk can also be understood in a different way. We can impose the same number of local households H_C but different H 's. In this setting, country X produces a larger k_{ch} than what country Y does. Prominent examples could be Sri Lanka–Germany or China–Finland. Looking at the same neighborhood of e.g. $H_C = 100$ households in those countries in comparison would justify the intuition on having different k_{ch} 's.

A.7 Self-quarantine

All the humans who are detected to have been exposed to the virus proceed to quarantine for the period of ν^{-1} on the day scale. A study of 425 first patients in Wuhan, China has found that the mean incubation period from the distribution was 5.2d [6], but quarantine is suggested to take up 14d worldwide [39], which is at the 99th percentile of the overall patients in the study. This means that they are obliged to stay at home and practice good hand during those days. Details such as keeping the body and hygiene-related personal belongings away from other members in the household as well as not having visitors should be apparent. Due to quarantine, as many as $(1 - \nu)pA$ detected infected individuals per day literally never go out for walk or even meet other household members and strangers.

A wishful consequence is that now only $\nu p_A + (1 - p)A$ non-quarantined and undetected individuals are able to transmit the virus. Denoting $p_A := \nu p + (1 - p)$, the infection term (13) changes to

$$S' = -\beta \left[1 + \frac{N_h - 1}{H_C} \right] S \cdot p_A A - \beta \cdot S \cdot p_I I - \alpha \cdot S \cdot \int_{t-\sigma}^t p_A A(\xi) d\xi.$$

For a far-from-realistic case where the standard quarantine takes place ∞ d, nobody from detected cases contributes to new infection via either direct or fomite transmission.

Appendix B: The basic reproductive number

The following descriptions comprise details on the derivation of the basic reproductive number during an earlier take-off period. Due to the constancy of the total population N , we can exclude S from the analysis by denoting $S = N - E - A - I - R - D$. We then seek to compute the basic reproductive number on an earlier take-off period, where one can assume that the online case fatality rate μ is relatively constant or simply $\mu = 0$, and $R(t) = D(t) = 0$ for all $-\sigma \leq t \leq 0$. As a result, the subsystem $(R', D') = (0, 0)$ degenerates into the equilibrium state $(R, D) = 0$. For an easy treatment, let us fold $(E', A', I')|_{R=D=0}$ into

$$u' = f \left(t, u, \int_{t-\sigma}^t g(u(\xi)) d\xi \right), \quad u(t) = \eta(t), \quad -\sigma \leq t \leq 0. \tag{14}$$

Substituting $E = A = I = 0$ into the right-hand side of (14), we obtain the disease-free equilibrium $u_0 = (0, 0, 0)$. The error $\bar{u} := u - u_0$ satisfies the equation

$$\bar{u}' = f \left(u_0, \underbrace{\int_{t-\sigma}^t g(u_0) d\xi}_{=0} \right) + \partial_u f \left(u_0, \int_{t-\sigma}^t g(u_0) d\xi \right)^T \bar{u} + \mathcal{O}(\|\bar{u}\|^2).$$

When a solution trajectory u is sufficiently close to u_0 , then the second-order term of the error is negligible and the Jacobian matrix

$$\partial_u f \left(u_0, \int_{t-\sigma}^t g(u_0) d\xi \right) = \begin{pmatrix} -\theta & (\beta [1 + \frac{N_h - 1}{H_C}] + \alpha \sigma) p_A N & \beta p_I N \\ a\theta & -\gamma_A & 0 \\ (1 - a)\theta & 0 & -\gamma_I \end{pmatrix}$$

takes the lead in determining whether the trajectory will stay around u_0 for all the remaining time. The trajectories can confine to u_0 providing that all the eigenvalues of the Jacobian matrix have strictly negative real part. Apparently, the eigenvalues satisfy the following cubic equation

$$\begin{aligned} &\lambda^3 + (\gamma_A + \gamma_I + \theta)\lambda^2 \\ &+ (\theta\gamma_A + \theta\gamma_I + \gamma_A\gamma_I) \left[1 - \frac{(\beta [1 + \frac{N_h - 1}{H_C}] + \alpha \sigma) p_A N a + \beta N p_I (1 - a)}{\gamma_A + \gamma_I + \gamma_A\gamma_I/\theta} \right] \lambda \\ &+ (\theta\gamma_A\gamma_I) \left[1 - \frac{(\beta [1 + \frac{N_h - 1}{H_C}] + \alpha \sigma) p_A N a \gamma_I + \beta N p_I (1 - a) \gamma_A}{\gamma_A\gamma_I} \right] = 0. \end{aligned}$$

It is not difficult to show that all the coefficients of the cubic polynomial remain positive if the constant term is positive or the *basic reproductive number* $\mathcal{R}_0 < 1$, where

$$\mathcal{R}_0 := \frac{(\beta[1 + \frac{N_h-1}{H_C}] + \alpha\sigma)p_A Na}{\gamma_A} + \frac{\beta N p_I(1-a)}{\gamma_I}. \tag{15}$$

A direct consequence from having positive coefficients is that the polynomial intersects y -axis at a positive value and has positive first and second derivative for all positive λ , meaning there cannot be any root with positive real part. As the equilibrium state u_0 is asymptotically stable in (14), the disease-free equilibrium $(E, A, I, R, D) = (0, 0, 0, 0, 0)$ is also asymptotically stable in the coupled system (1), see [49, Theorem 3.1].

Appendix C: Numerical treatment for the DIDE system

Our model system can be studied in the framework of a system of delay integro-differential equations (DIDEs) abstracted like in (14). We are in a good position to know that this type of system has been under previous investigations [50–53], even from the field of population dynamics [54]. A practical method for solving (14) is the direct adaptation of the (one-step) Runge–Kutta method, cf. [50]. Without loss of generality, let us suppose that $t_s = 0$. We depart from the usual time discretization where the time domain for the initial condition $[-\sigma, 0]$ is discretized into a set of, say, M sub-intervals of equal length. The time step and resulting time points of observation are then determined by $h := \sigma/M$ and $t_n := hn, n \in \mathbb{Z}$. We assume equal length only for the sake of simplicity. The state variable u automatically undertakes the discretization $u \approx (u_0, u_1, \dots)$, containing six rows. A d -stage Runge–Kutta method features two vectors $(b_i), (c_i) \in \mathbb{R}^d$ and one matrix $(a_{ij}) \in \mathbb{R}^{d \times d}$ such that

$$U_{n,i} = u_n + h \sum_{j=1}^d a_{ij} f(t_n + c_j h, U_{n-1,j}, G_{n,j}),$$

$$G_{n+1,i} = h \left[\sum_{j=1}^d a_{ij} g(U_{n,j}) + \sum_{k=1}^M \sum_{j=1}^d b_j g(U_{n-k,j}) - \sum_{j=1}^d a_{ij} g(U_{n-M,j}) \right],$$

$$u_{n+1} = u_n + h \sum_{i=1}^d b_i f(t_n + c_i h, U_{n,i}, G_{n+1,i}), \quad n = 0, 1, \dots$$

The initial conditions for the intermediate values $U_{n,i}$ and approximate of the integral term $G_{0,i}$ are given by

$$U_{n,i} = \eta(t_n + c_i h), \quad -M \leq n \leq 0,$$

$$G_{0,i} = h \left[\sum_{j=1}^d a_{ij} g(U_{0,j}) + \sum_{k=1}^M \sum_{j=1}^d b_j g(U_{-k,j}) - \sum_{j=1}^d a_{ij} g(U_{-M,j}) \right]$$

for all $1 \leq i \leq d$. The scheme given above is of Pouzet-type where $(b_i), (c_i), (a_{ij})$ are taken from the Runge–Kutta for ODE counterpart, and has been proven to be convergent [55, 56].

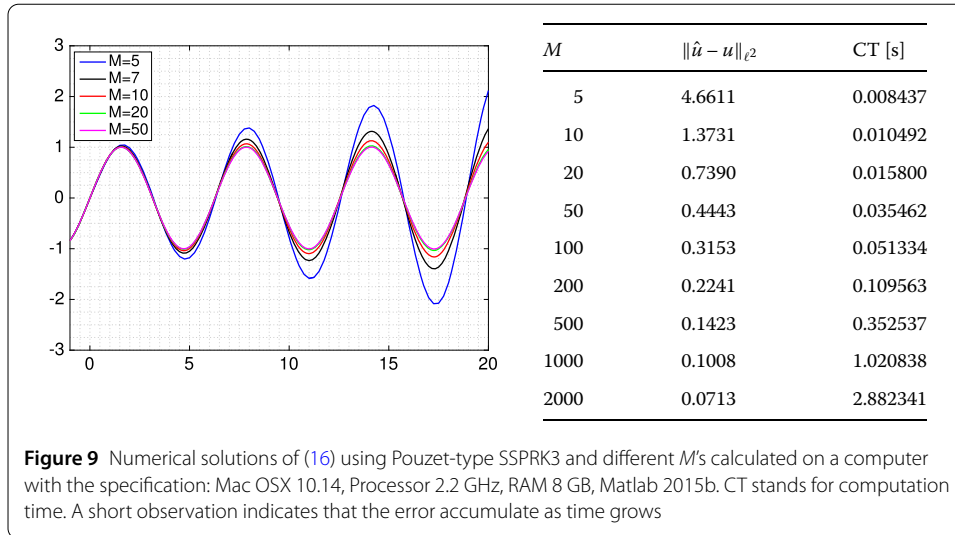


Figure 9 Numerical solutions of (16) using Pouzet-type SSPRK3 and different M 's calculated on a computer with the specification: Mac OSX 10.14, Processor 2.2 GHz, RAM 8 GB, Matlab 2015b. CT stands for computation time. A short observation indicates that the error accumulate as time grows

Before applying the scheme to our model, we shortly test its performance. We consider a test function $u(t) = \sin(t)$ from which computations of the derivative and kernel term return the relation

$$u' = \frac{\sin(\sigma)}{1 - \cos(\sigma)}u - \frac{1}{1 - \cos(\sigma)} \int_{t-\sigma}^t u(\xi) d\xi, \quad \sigma \notin 2\pi\mathbb{Z}, \tag{16}$$

$$u(t) = \sin(t), \quad -\sigma \leq t \leq 0.$$

The time delay is specified as $\sigma = 1$ and starting and final time are $t_s = 0, t_f = 20$. The adopted scheme, also for the overall simulations in this paper, is Third-order Strong Stability Preserving Runge–Kutta (SSPRK3) represented by

$$(b_i) = \begin{pmatrix} 1/6 \\ 1/6 \\ 2/3 \end{pmatrix}, \quad (c_i) = \begin{pmatrix} 0 \\ 1 \\ 1/2 \end{pmatrix}, \quad (a_{ij}) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1/4 & 1/4 & 0 \end{pmatrix}.$$

The computation results for different M 's are presented in Fig. 9. The numerical solution is named \hat{u} .

Acknowledgements

Not applicable.

Funding

Not applicable.

Abbreviations

COVID-19, Corona Virus Disease 2019; WHO, World Health Organization; SARS, Severe acute respiratory syndrome; MERS, Middle East respiratory syndrome; CoV, Corona Virus; CDC, Centers for Disease Control and Prevention; SIR, Susceptible–Infected–Recovered; RNA, Ribonucleic acid; IQR, Interquartile range; CFR, Case fatality rate; RKI, Robert Koch Institute; ODE, Ordinary differential equation; DIDE, Delay integro-differential equation; SSPRK, Strong Stability Preserving Runge–Kutta.

Availability of data and materials

The data were recorded by Johns Hopkins University [28], which can be accessed via a GitHub repository [29].

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors contributed equally, have read and approved the final manuscript.

Authors' information

Not applicable.

Author details

¹Mathematical Institute, University of Koblenz, DE-56070 Koblenz, Germany. ²Department of Mathematics, University of Sri Jayawardenepura, SL-10250 Nugegoda, Sri Lanka.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 9 September 2020 Accepted: 22 December 2020 Published online: 05 January 2021

References

- Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z, Yu T, Xia J, Wei Y, Wu W, Xie X, Yin W, Li H, Liu M, Xiao Y, Gao H, Guo L, Xie J, Wang G, Jiang R, Gao Z, Jin Q, Wang J, Cao B. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020;395(10223):497–506.
- World Health Organization. Country & technical guidance—coronavirus disease (COVID-19). 2020. <https://covid19.who.int/>. Accessed: 11.04.2020.
- Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, Qiu Y, Wang J, Liu Y, Wei Y, Xia J, Yu T, Zhang X, Zhang L. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet*. 2020;395(10223):507–13.
- Ji W, Wang W, Zhao X, Zai J, Li X. Cross-species transmission of the newly identified coronavirus 2019-nCoV. *J Med Virol*. 2020;92:433–40.
- Forstera P, Forster L, Renfrew C, Forster M. Phylogenetic network analysis of SARS-CoV-2 genomes. *Proc Natl Acad Sci*. 2020;117(17):9241–3.
- Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KS, Lau EH, Wong JY, Xing X, Xiang N, Wu Y, Li C, Chen Q, Li D, Liu T, Zhao J, Liu M, Tu W, Chen C, Jin L, Yang R, Wang Q, Zhou S, Wang R, Liu H, Luo Y, Liu Y, Shao G, Li H, Tao Z, Yang Y, Deng Z, Liu B, Ma Z, Zhang Y, Shi G, Lam TT, Wu JT, Gao GF, Cowling BJ, Yang B, Leung GM, Feng Z. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med*. 2020;382(13):1199–207.
- Zhao S, Lin Q, Ran J, Musa SS, Yang G, Wang W, Lou Y, Gao D, Yang L, He D, Wang MH. 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. *Int J Infect Dis*. 2020;92:214–7.
- Hellewell J, Abbott S, Gimma A, Bosse NI, Jarvis CI, Russell TW, Munday JD, Kucharski AJ, Edmunds WJ. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *Lancet Glob Health*. 2020;8(4):e488–96.
- World Health Organization. Coronavirus disease (COVID-2019) situation reports. 2020. <https://www.who.int/emergencies/>. Accessed: 11.04.2020.
- United Nations. Coronavirus global health emergency: coverage from UN news. 2020. <https://news.un.org/en/events/>. Accessed: 15.04.2020.
- Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, Cuomo-Dannenburg G, Thompson H, Walker PGT, Han Fu AD, Griffin JT, Baguelin M, Bhatia S, Boonyasiri A, Cori A, Cucunub Z, FitzJohn R, Gaythorpe K, Green W, Hamlet A, Hinsley W, Laydon D, Nedjati-Gilani G, Riley S, van Elsland S, Volz E, Wang H, Xi YWX, Donnelly CA, Ghani AC, Ferguson NM. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect Dis*. 2020;20(6):669–77.
- Leung K, Wu JT, Liu D, Leung GM. First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment. *Lancet*. 2020;395(10233):1382–93.
- World Health Organization. Coronavirus disease (COVID-2019) situation reports. 2020. <https://www.who.int/docs/>. Accessed: 15.04.2020.
- Chan JF-W, Yuan S, Kok K-H, To KK-W, Chu H, Yang J, Xing F, Liu J, Yip CC-Y, Poon RW-S, Tsoi H-W, Lo SK-F, Chan K-H, Poon VK-M, Chan W-M, Ip JD, Cai J-P, Cheng VC-C, Chen H, Hui CK-M, Yuen K-Y. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet*. 2020;395(10223):514–23.
- Wu JT, Leung K, Bushman M, Kishore N, Niehus R, de Salazar PM, Cowling BJ, Lipsitch M, Leung GM. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nat Med*. 2020;26:506–10.
- Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, Xiang J, Wang Y, Song B, Gu X, Guan L, Wei Y, Li H, Wu X, Xu J, Tu S, Zhang Y, Chen H, Cao B. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*. 2020;395(10229):1054–62.
- Rothe C, Schunk M, Sothmann P, Bretzel G, Froeschl G, Wallrauch C, Zimmer T, Thiel V, Janke C, Guggemos W, Seilmaier M, Drosten C, Vollmar P, Zwirgmaier K, Zange S, Wölfel R, Hoelscher M. Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. *N Engl J Med*. 2020;382(10):970–1.
- Ghani AC, Donnelly CA, Cox DR, Griffin JT, Fraser C, Lam TH, Ho LM, Chan WS, Anderson RM, Hedley AJ, Leung GM. Methods for estimating the case fatality ratio for a novel, emerging infectious disease. *Am J Epidemiol*. 2005;162(5):479–86.
- Angelopoulos AN, Pathak R, Varma R, Jordan MI. Identifying and correcting bias from time- and severity-dependent reporting rates in the estimation of the COVID-19 case fatality rate. *arXiv preprint*. 2003.08592v3 (2020).
- Baud D, Qi X, Nielsen-Saines K, Musso D, Pomar L, Favre G. Real estimates of mortality following COVID-19 infection. *Lancet Infect Dis*. 2020;20(7):773.
- Johns Hopkins University. Mortality analyses. 2020. <https://coronavirus.jhu.edu/data/mortality>. Accessed: 22.04.2020.

22. van den Driessche P, Watmough J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math Biosci.* 2002;180:29–48.
23. Wikipedia. COVID-19 pandemic in Germany. 2020. https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Germany#1. Accessed: 20.05.2020.
24. Weiss T. Ein Virus breitet sich aus. 2020. <https://www.faz.net/aktuell/>. Accessed: 20.04.2020.
25. Epidemiology Unit—Sri Lanka. COVID-19 national epidemiological report—Sri Lanka. 2020. <http://www.epid.gov.lk>. Accessed: 09.04.2020.
26. Department of Government Information—Sri Lanka. Latest news. 2020. <https://www.dgi.gov.lk/>. Accessed: 09.04.2020.
27. Presidential Secretariat—Sri Lanka. More stern measures to curb COVID-19. 2020. <https://www.presidentsoffice.gov.lk>. Accessed: 23.04.2020.
28. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis.* 2020;20(5):533–4.
29. GitHub. COVID-19 dashboard. 2020. <https://github.com/github/covid19-dashboard>. Accessed: 23.04.2020.
30. Johns Hopkins University. Covid-19 dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). 2020. <https://coronavirus.jhu.edu/map.html>. Accessed: 23.04.2020.
31. Lin Q, Zhao S, Gao D, Lou Y, Yang S, Musa SS, Wang MH, Cai Y, Wang W, Yang L, He D. A conceptual model for the coronavirus disease 2019 (COVID-19) outbreak in Wuhan, China with individual reaction and governmental action. *Int J Infect Dis.* 2020;93:211–6.
32. van Doremalen N, Bushmaker T, Morris DH, Holbrook MG, Gamble A, Williamson BN, Tamin A, Harcourt JL, Thornburg NJ, Gerber SI, Lloyd-Smith JO, de Wit E, Munster VJ. Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *N Engl J Med.* 2020;382:1564–7.
33. CDC. Public health responses to COVID-19 outbreaks on cruise ships—worldwide, February–March 2020. 2020. <https://www.cdc.gov/mmwr/volumes/69/wr/mm6912e3.htm>. Accessed: 08.04.2020.
34. Li R, Pei S, Chen B, Song Y, Zhang T, Yang Y, Shaman J. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science.* 2020;368:489–93.
35. Ting C. On the mean convergence time of multi-parent genetic algorithms without selection. In: Caparrere M, et al, editors. *Advances in artificial life. ECAL 2005. Lecture notes in computer science.* vol. 3630. Berlin: Springer; 2005. p. 403–12.
36. Worldometer. Countries in the world by population (2020). 2020. <https://www.worldometers.info/world-population/population-by-country/>. Accessed: 12.04.2020.
37. United Nations. Household size & composition, 2019. 2019. <https://population.un.org/Household/index.html>. Accessed: 08.04.2020.
38. Department of Census and Statistics—Sri Lanka. Population and housing. 2020. <http://www.statistics.gov.lk/>. Accessed: 15.04.2020.
39. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, Azman AS, Reich NG, Lessler J. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Ann Intern Med.* 2020. <https://doi.org/10.7326/M20-0504>.
40. Guan W, Ni Z, Hu Y, Liang W, Ou C, He J, Liu L, Shan H, Lei C, Hui D, Du B, Li L, Zeng G, Yuen K-Y, Chen R, Tang C, Wang T, Chen P, Xiang J, Li S, Wang JL, Liang Z, Peng Y, Wei L, Liu Y, Hu YH, Peng P, Wang JM, Liu J, Chen Z, Li G, Zheng Z, Qiu S, Luo J, Ye C, Zhu S, Zhong N. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med.* 2020;382:1708–20.
41. Arachchi RJ. Irresponsible social behaviour links to lack of risk perception—Prof Kalinga Tudor Silva. 2020. <http://www.sundayobserver.lk>. Accessed: 23.04.2020.
42. Kariyawasam R. Lessons learnt from case of Mt Lavinia gem businessman. 2020. <https://ceylontoday.lk/print-more/55447>. Accessed: 23.04.2020.
43. Arachchi RJ. Sri Lanka battles on with Covid-19. 2020. <http://www.sundayobserver.lk/2020/04/05>. Accessed: 23.04.2020.
44. Liu Y, Gayle AA, Wilder-Smith A, Rocklöv J. The reproductive number of COVID-19 is higher compared to SARS coronavirus. *J Travel Med.* 2020;27(2):taaa021.
45. Tang S, Mao Y, Jones RM, Tan Q, Ji JS, Li N, Shen J, Lv Y, Pan L, Ding P, Wang X, Wang Y, MacIntyre CR, Shi X. Aerosol transmission of SARS-CoV-2? Evidence, prevention and control. *Environ Int.* 2020;144:106039.
46. Lancet-Editorial. COVID-19 transmission—up in the air. *Lancet Respir Med.* 2020;8(12):1159.
47. Kampf G, Todt D, Pfaender S, Steinmann E. Persistence of coronaviruses on inanimate surfaces and their inactivation with biocidal agents. *J Hosp Infect.* 2020;104:246–51.
48. Prompetchara E, Ketloy C, Palaga T. Immune responses in COVID-19 and potential vaccines: lessons learned from SARS and MERS epidemic. *Asian Pac J Allergy Immunol.* 2020;38:1–9.
49. Vidyasagar M. Decomposition techniques for large-scale systems with nonadditive interactions: stability and stabilizability. *IEEE Trans Autom Control.* 1980;25(4):773–9.
50. Koto T. Stability of Runge–Kutta methods for delay integro-differential equations. *J Comput Appl Math.* 2002;145:483–92.
51. Zhang C, Vandewalle S. Stability analysis of Runge–Kutta methods for nonlinear Volterra delay-integro-differential equations. *IMA J Numer Anal.* 2004;24:193–214.
52. Khasawneh FA, Mann BP. Stability of delay integro-differential equations using a spectral element method. *Math Comput Model.* 2011;54:2493–503.
53. Shakourifar M, Enright WH. Reliable approximate solution of systems of Volterra integro-differential equations with time-dependent delays. *SIAM J Sci Comput.* 2011;33(3):1134–58.
54. Kuang Y. *Delay differential equations—with applications in population dynamics.* Mathematics in science and engineering. vol. 191. San Diego: Academic Press; 1993.
55. Brunner H, Hairer E, Norsett S. Runge–Kutta theory for Volterra integral equations of the second kind. *Math Comput.* 1982;39:147–63.
56. Lubich C. Runge–Kutta theory for Volterra integrodifferential equations. *Numer Math.* 1982;40:119–35.