The dynamics of entropy in the COVID-19 outbreaks

Ziqi Wang · **Marco Broccardo** ·
Arnaud Mignan · **Didier Sornette**

Abstract With the unfolding of the COVID-19 pandemic, mathematical modelling of epidemics has been perceived and used as a central element in understanding, predicting, and governing the pandemic event. However, soon it became clear that long-term predictions were extremely challenging to address. In addi-

Z. Wang (✉)
Earthquake Engineering Research and Test Center,
Guangzhou University, Guangzhou, China
e-mail: ziqidwang@yahoo.com

M. Broccardo (✉)
Department of Civil, Environmental and Mechanical
Engineering, University of Trento, Trento, Italy
e-mail: marco.broccardo@unitn.it

M. Broccardo
Institute for Risk and Uncertainties, University of
Liverpool, Liverpool, UK

A. Mignan · D. Sornette
Institute of Risk Analysis, Prediction and Management,
Southern University of Science and Technology, Shenzhen,
China
e-mail: mignana@sustech.edu.cn

A. Mignan
Department of Earth and Space Sciences, Southern
University of Science and Technology, Shenzhen, China

D. Sornette
Chair of Entrepreneurial Risks, Department of
Management, Technology, and Economics, ETH Zürich,
Zurich, Switzerland
e-mail: dsornette@ethz.ch

tion, it is still unclear which metric shall be used for a global description of the evolution of the outbreaks. Yet a robust modelling of pandemic dynamics and a consistent choice of the transmission metric is crucial for an in-depth understanding of the macroscopic phenomenology and better-informed mitigation strategies. In this study, we propose a Markovian stochastic framework designed for describing the evolution of entropy during the COVID-19 pandemic together with the instantaneous reproductive ratio. Then, we introduce and use entropy-based metrics of global transmission to measure the impact and the temporal evolution of a pandemic event. In the formulation of the model, the temporal evolution of the outbreak is modelled by an equation governing the probability distribution that describes a nonlinear Markov process of a statistically averaged individual, leading to a clear physical interpretation. The time-dependent parameters are formulated by adaptive basis functions, leading to a parsimonious representation. In addition, we provide a full Bayesian inversion scheme for calibration together with a coherent strategy to address data unreliability. The time evolution of the entropy rate, the absolute change in the system entropy, and the instantaneous reproductive ratio are natural and transparent outputs of this framework. The framework has the appealing property of being applicable to any compartmental epidemic model. As an illustration, we apply the proposed approach to a simple modification of the susceptible–exposed–infected–removed model. Applying the model to the Hubei region, South

Korean, Italian, Spanish, German, and French COVID-19 datasets, we discover significant difference in the absolute change of entropy but highly regular trends for both the entropy evolution and the instantaneous reproductive ratio.

1 Introduction

Coronaviruses are one of the most significant threats to human society [1–6]. Limited to short outbreaks in the recent past [7–9], their pandemic-level potential was well known [10,11], yet most countries proved unprepared to cope with the so-called coronavirus infectious disease of 2019 (COVID-19). Revealed in the Hubei province, China, the novel coronavirus has spread all over the world. China responded with massive containment measures starting at the end of January 2020, which limited further contamination on the mainland [8,12]. In Europe, most individual states have responded with similar containment measures. However, there has been a lack of common European action. Strict or soft containment measures have been applied with different timeframes and specialized to individual health and socio-cultural systems, showing very different pandemic evolutions. At the time of writing, the main episode of COVID-19 is (in general) under control in China, South Korea, and continental Europe [11], despite the possibility of multiple waves. On the contrary, North America and South America are still in the middle of the pandemic, and a clear picture of the evolution of events is not possible yet.

The amount of data available allows various modelling techniques to be tested more robustly than in previous epidemics. However, no model (from the physics-based to the purely data-driven) has been or is able to predict the long-term evolution of the pandemic accurately. (Conversely, short-term predictions are possible with some degree of accuracy [13,14].) There are several reasons behind this long-term unpredictability; an incomplete list includes the partial understanding of the phenomenon, the (many, or even infinite) missing variables, the high sensitivity of the model to parameters, the incomplete/inaccurate data acquisition scheme, and the lack of uniform measurement methods.

However, a profound reason that makes any long-term prediction difficult is the presence of endogenous variables (a well-known problem in social sciences [15]). The endogenous variables may involve local policies, socio-cultural aspects, human behaviours, and communication strategies, and they are typically difficult to model and measure. Since an epidemic evolves as a result of the interplay between the "natural evolution" of the disease and society/human interventions, a robust and generalizable microscopic model with complete characterizations of endogenous variables is challenging to build. Given this, in this study, we use macroscopic phenomenology-based modelling to gain insight into the epidemic dynamics.[1] Therefore, here, the goal is not to provide long-term numerical predictions, although the proposed modelling technique can be used for extrapolation.

Among various modelling options, susceptible–infected–removed (SIR) types of compartmental models have gained wide popularity due to their simplicity and straightforwardness in interpreting the macroscopic phenomenology. A significant amount of SIR-type model based studies have already been carried out to investigate the transmission properties of the COVID-19, and an incomplete list includes [8,12,16–20]. The spectrum of complexity of these models is broad. They can range from a minimum number of compartments (which offers a better generalization) to a large number of compartments (which offers a better local description). They can be deterministic (i.e. counting the deterministic number of individuals for each compartment) or stochastic (i.e. defining a joint probability measure of the number of individuals for each compartment). They can have different data acquisition schemes (from a simple frequentist analysis of the single parameters to a complete Bayesian inversion scheme). Finally, they can simply macroscopically describe the pandemic evolution of a given location (top-down approach) or include a spatial topological description (including mobility) and/or a different degree of spreading among individuals by including adjacent matrices (bottom-up approach).

Given a dataset, these models can be calibrated and offer new insights into the evolution of the pandemic. For example, they can shed light on how the pandemic

---

[1] This approach follows the famous sentence of Richard Hamming "The purpose of (scientific) computing is insight, not numbers."

developed by measuring the reproductive ratio (constant or time varying) and finally estimating the effectiveness of containment measures. This metric also allows for a comparison between different regions but does not provide a quantitative measure of the impact of the spread. On the other hand, the evolution of the number of infected and deaths provides a means of direct impact; however, they lack objectivity as they are strongly influenced by the different populations of the regions, the measurement strategies, and the unreliability of the data. Furthermore, they are not global metrics as they do not provide an objective and robust way to unify them into a single (scalar) measure. Therefore, there is a research gap on how to provide a macroscopic model-metric pair to compare different regions' performance and get new insights into various outbreaks.

This study aims to fill this gap by proposing a macroscopic stochastic model equipped with a global transmission metric based on entropy. In this context, the entropy evolution of the process is a metric that describes the degree of disorder (i.e. of impact) of an epidemic. This metric allows for an objective comparison between regions and provides a global measure of both the evolution and the impact of COVID-19 outbreaks. In particular, we propose a compartmental stochastic model that has the following characteristics. (i) *Stochastic*: the model describes a statistically averaged individual by a nonlinear Markov process with compartmental epidemic states. (ii) *Time-dependent*: the model parameters are decomposed onto generic basis functions (of time). (iii) *Parsimonious*: instead of conventional orthogonal basis functions (e.g. orthogonal polynomials, Fourier/wavelet series) the adaptive basis functions are adopted to achieve a representation with minimum number of basis functions. (iv) *Bayesian*: the time-dependent parameters are assumed to be random and are calibrated by full Bayesian inversion. Furthermore, we equip the model with a metric based on entropy which has the following characteristics. (i) *Meaningful*: the metric provides a physical and transparent measure of the COVID-19 impact in a given region; moreover, it is by definition the time integral of the entropy rate, which represents the temporal evolution of the epidemic. (ii) *Global*: the metric provides a global and average description of the pandemic event. (iii) *Consistent*: the metric is not influenced by the number of individuals and can be used objectively to compare different regions. (iv) *Robust* the metric is associated with an error that is a direct

output of the Bayesian inversion scheme used to calibrate the stochastic model. Finally, to have a reliable description of the events, we provide robust strategies to fill in missing information and to correct the numerous inconsistencies on the current datasets.

The paper is organized as follows. First, we develop general concepts of the proposed epidemic model, including governing equation, time-dependent parameterization, and Bayesian model calibration (Sect. 2). Second, we introduce the entropy-based metric in Sect. 3. Third, we apply the proposed approach to formulate a SEIR compartmental model for modelling the temporal evolution of COVID-19 (Sect. 4). Next, we apply the proposed approach to real datasets to the following regions: Hubei (China), South Korea, Italy, Spain, Germany, and France (Sect. 5). Finally, we conclude the study by identifying the limitations, conclusions, and future research directions.

2 The stochastic epidemic model

In the literature, the term "stochastic compartmental model" can refer to different formulations (see e.g. [21] for a review) with distinct underlying assumptions on the source of uncertainty. For instance, the noise-driven stochastic model is formulated by: (i) introducing additive noise process into the deterministic compartmental model; (ii) translating the noise into diffusion of probability distribution; and (iii) obtaining an equation of probability distribution (e.g. Fokker–Plank equation). Clearly, in the noise-driven model the source of uncertainty is the additive noise. An alternative and more popular stochastic formulation is the event-driven model, which can be summarized as a direct *stochastic simulation* of the deterministic model. Specifically, in the event-driven model the deterministic rate matrix is used to define the transition probability of event $X_m(t) \xrightarrow{t+\Delta t} X_m(t) \pm \Delta X$, where $X_m(t)$ denotes the population in a compartment and ΔX the intra-state increment. With the transition probability, a direct stochastic simulation (via e.g. Gillespie's Direct Method [22,23]) would yield a random scenario of the epidemic. The proposed model can also be classified as an event-driven approach in the sense that the source of uncertainty is also the aleatory variability of transitions between epidemic states. However, instead of a stochastic simulation without a governing equation of probability distribution, the proposed model strictly

follows an equation of probability distribution which describes a nonlinear Markov process. Consequently, the proposed model possesses clear physical interpretations within the mathematical framework of nonlinear Markov process theory.

Compartmental models with time-varying parameters have been widely studied in the literature [24–28]. A fundamental question to be addressed in time-dependent models is the trade-off between over-fitting and under-fitting, or equivalently, model bias versus model variance. In the extreme scenario, a point-wise kernel-based parameterization may lead to an almost exact calibration on epidemic observations, yet the explanatory/extrapolation capability would be minimized, and the model variance would be maximized. In an over-parameterized model, the non-local trend/structure, which is crucial to characterize/understand the epidemic dynamics, can hardly be identified. In this study, we attempt to discover non-local structures from the epidemic dataset using a parsimonious formulation with adaptive basis functions. Moreover, since the model calibration is formulated in a Bayesian framework, likelihood-based model selection, e.g. using the Bayesian information criterion (BIC) [29], can be conveniently applied to specify the number of adaptive basis functions.

2.1 The original deterministic model

Consider a generic compartmental epidemic model with a fixed[2] total population N and a classification of the population into M compartments $\boldsymbol{X} = [X_1, \ldots, X_M]^\top$. The compartmental epidemic model describes the temporal evolution of the state vector $\boldsymbol{X}$, where every component of $\boldsymbol{X}$ is by definition non-negative and $\boldsymbol{X}$ is subjected to the conservation law $\|\boldsymbol{X}\|_1 = N$.

For an infinitesimal incremental Δt, we study the following master equation of state vector.

$$\boldsymbol{X}(t + \Delta t) = (\boldsymbol{I} + \boldsymbol{H}(\boldsymbol{X}(t), t)\Delta t)\boldsymbol{X}(t), \tag{1}$$

where $\boldsymbol{I}$ is the identity matrix and $\boldsymbol{H}(\boldsymbol{X}(t), t)$ is a problem-specific rate matrix (infinitesimal *propagator*). Equation (1) is equipped with the assumption that

the evolution of $\boldsymbol{X}(t)$ is smooth,[3] i.e. without jumps. Setting $\Delta t \to 0$, Eq. (1) leads to

$$\frac{\mathrm{d}\boldsymbol{X}(t)}{\mathrm{d}t} = \boldsymbol{H}(\boldsymbol{X}(t), t)\boldsymbol{X}(t). \tag{2}$$

Similar to mechanics, the rate matrix $\boldsymbol{H}(\boldsymbol{X}(t), t)$ governs the dynamics of $\boldsymbol{X}(t)$. To preserve the conservation law $\|\boldsymbol{X}(t)\|_1 = N$, we must have $\boldsymbol{H}(\boldsymbol{X}(t), t)^\top \mathbf{1} = \mathbf{0}$, where $\mathbf{1}$ is a vector of ones and $\mathbf{0}$ the null vector.

Particularly, if $\boldsymbol{X}(t)$ eventually attains a stationary state $\boldsymbol{X}^*$ defined as

$$\boldsymbol{X}^* := \lim_{t \to +\infty} \boldsymbol{X}(t), \tag{3}$$

and define $\boldsymbol{H}^*$ as

$$\boldsymbol{H}^* := \lim_{t \to +\infty} \boldsymbol{H}(\boldsymbol{X}(t), t). \tag{4}$$

We obtain the stationarity condition

$$\boldsymbol{H}^* \boldsymbol{X}^* = \mathbf{0}, \tag{5}$$

where $\mathbf{0}$ is a column vector of zeros.

2.2 Probabilistic reformulation

Given a deterministic $\boldsymbol{H}(\boldsymbol{X}(t), t)$, Eq. (2) describes a deterministic trajectory of $\boldsymbol{X}(t)$. Since variabilities inevitably exist in the specification of $\boldsymbol{H}(\boldsymbol{X}(t), t)$ or/and the initial condition, the solution $\boldsymbol{X}(t)$ becomes a multivariate stochastic process. However, the aforementioned "randomization" is regarded as epistemic with respect to the model Eq. (2).[4] This section focuses on a more fundamental (aleatory) probabilistic reformulation of Eq. (2).

Adopting a *frequentist* point of view on probability, consider a normalization of $\boldsymbol{X}(t)$ by

$$\boldsymbol{P}(t) := \lim_{N \to +\infty} \frac{\boldsymbol{X}(t, N)}{\|\boldsymbol{X}(t, N)\|_1} = \lim_{N \to +\infty} \frac{\boldsymbol{X}(t, N)}{N}, \tag{6}$$

where $\boldsymbol{X}(t, N)$ is used to highlight that the compartmental population depends on N, and $\boldsymbol{P}(t)$ can be interpreted as the marginal probability distribution

---

[2] A fixed total population indicates a closed system, i.e. (i) the vital dynamics (natural birth/death) is neglected, assuming that the course of the epidemic is relatively short; (ii) the immigration and emigration are neglected, assuming that N is sufficiently large and the course of immigration/emigration is relatively slow.

[3] Here, we have to assume the population is a real number rather than an integer; later, in the probabilistic reformulation this assumption can be relaxed.

[4] The classification of epistemic or aleatory uncertainties should always be accompanied by the specification of a model universe. Here if we assume Eq. (2) is the correct underlying (deterministic) law, the uncertainties of $\boldsymbol{H}(\boldsymbol{X}(t), t)$ is epistemic (with respect to Eq. (2)) because if the form of $\boldsymbol{H}(\boldsymbol{X}(t), t)$ is fixed the law of Eq. (2) is deterministic.

of a discrete-state continuous-time stochastic process. Observe that despite $P(t)$ being equivalent to proportions in a deterministic model, the probabilistic individualistic interpretation leads to a fully stochastic dynamic interpretation of the problem. The underlying state associated with $P(t)$ is an epidemic state of a statistically averaged *individual*. Analogous to Eqs. (1) and (2), we obtain

$$P(t + \Delta t) = (I + Q(P(t), t)\Delta t)P(t), \tag{7}$$

and

$$\frac{\mathrm{d}P(t)}{\mathrm{d}t} = Q(P(t), t)P(t), \tag{8}$$

where $Q(P(t), t)$ is a rate matrix analogous to $H(X(t), t)$ in the deterministic model. Since $Q(P(t), t)$ explicitly depends on $P(t)$, Eq. (8) describes a *nonlinear Markov process* [30]. The conservation of probability is guaranteed by $Q(t)^\top \mathbf{1} = \mathbf{0}$.

Equation (7) provides a straightforward strategy for sampling random realizations of the process. In particular, for a fixed initial condition $P(t_0)$, the solution $P(t)$ is deterministic, and $Q(P(t), t)$ can be regarded as $Q(t)$ with $P(t)$ being a time-dependent *parameter* of $Q(t)$. The resulting *tangent non-homogeneous Markov process* has the following transient *stochastic matrix*

$$S(t, t + \Delta t) := I + Q(t)\Delta t. \tag{9}$$

In line with the macroscopic description [Eq. (6)], the initial condition $P(t_0)$ and the rate matrix $Q(P(t), t)$ are by definition exactly the same for all N individuals. This assumption corresponds *ad verbum* to fix a constant average number of contacts and other interaction parameters between persons per unit time. Given this, there is an implicit assumption of statistical independence among the N individuals. In an adiabatic system, this is equivalent to letting N particles following N-independent Brownian motions. Therefore, this macro-description is emerging from the micro-behaviour of individuals interacting according N-independent Brownian motions, and the virus is spreading according to a simple diffusive process.[5] Consequently, a macro-random scenario of an epidemic can be obtained via simulating N-independent and identically distributed processes from Eq. (8).

---

[5] At the micro-level, the spreading of the virus can be better described through a stochastic branching process. The simple diffusion process can be regarded as the result of a coarse graining.

In contrast to this macro-description, one could adopt a topological structure of the interactions between different individuals. This is generally done by including an adjacency operator which accounts for the different structure of the interactions among individuals (e.g. including mobility information, or considering the presence of superspreaders). As a consequence, each individual (or group of individuals) has a different average number of contacts and different interaction parameters. This leads to a heterogeneous compartmental model, which is inevitable dependent on a specific geographical area or social system. In this study, we focus on the general transmission trend of large regions, so that the trends can be more easily extrapolated and interpreted. Therefore, the simple macro-description is adopted. It is a specific choice which leads to a novel entropy-based measure to macroscopically compare the epidemic scenarios in different regions.

2.3 Time-dependent parameter model

We assume the "correct" model of $Q(P(t), t)$ cannot be discovered, and $Q(P(t), t)$ is replaced by a parametric model with a set of parameters $\boldsymbol{\alpha}(t)$, i.e.

$$Q(P(t), t) \approx Q(P(t), t; \boldsymbol{\alpha}(t)). \tag{10}$$

Let $\alpha(t)$ represent an arbitrary component of $\boldsymbol{\alpha}(t)$. A generic approach to parameterizing $\alpha(t)$ is to consider an expansion of the following form

$$\alpha(t) = \sum_{i=0}^{I} w_i \psi_i(t), \tag{11}$$

where w_i are coordinates of basis functions $\psi_i(t)$. A popular choice for the basis function is the orthogonal polynomials, e.g. Legendre/Hermite/Laguerre/Chebyshev polynomials. An issue with orthogonal polynomial basis is that it may require high-order terms to represent a complex function, and consequently, this leads to over-fitting and implausible extrapolations. A powerful alternative is to use adaptive basis functions with the form

$$\alpha(t) = \sum_{i=0}^{I} w_i \psi_i(t, \boldsymbol{w}_i'), \tag{12}$$

where $\boldsymbol{w}_i'$ are parameters of the adaptive basis. The benefit of using Eq. (12) instead of Eq. (11) is that a parsimonious representation can be formulated, at the cost of introducing additional parameters in bases. An

attractive choice for the adaptive basis $\psi_i(t, \boldsymbol{w}_i)$ is the sigmoid function, i.e.

$$\psi_i(t, \boldsymbol{w}_i') = \frac{1}{1 + \exp(w_{i1}' - w_{i2}'t)}. \tag{13}$$

The theoretical justification of using Eqs. (13) in (12) is the *universal approximation theorem* [31], and the resulting parametric function is in fact a feed-forward neural network with a single hidden layer.

In addition, the initial condition of Eq. (8) is unknown, and we parameterize $\boldsymbol{P}(t_0)$ by $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_{M-1}]$ (recall that M is the number of compartments). A natural parameterization of $\boldsymbol{P}(t_0)$ is

$$\boldsymbol{P}(t_0) = \left[\beta_1, \ldots, \beta_{M-1}, 1 - \sum_{m=1}^{M-1} \beta_m\right]^{\top}, \tag{14}$$

where β_m are nonnegative and subjected to the linear constraint $\sum_{m=1}^{M-1} \beta_m \in [0, 1]$. Note that $\boldsymbol{\beta}$ is time independent in the sense that the starting time point can be fixed. Therefore, the full parameter set of the epidemic model is written as $\boldsymbol{\theta} := \{\boldsymbol{w}, \boldsymbol{w}', \boldsymbol{\beta}\}$.

2.4 Model calibration

The goal of model calibration is to find the optimal $\boldsymbol{\theta}$ using real observation. We let $\mathcal{D}$ denote the dataset of observations collected for an epidemic up to some reference time point. The dataset $\mathcal{D}$ is composed by discrete measures on the number of persons in each observable compartment (e.g. infected, recovered, and dead), and $\mathcal{D}$ is a matrix of dimension $M_o \times T$, where M_o denotes the number of observable compartments and T denotes the number of observed unit time (e.g. days).

The likelihood function $\mathcal{L}(\mathcal{D}|\boldsymbol{\theta})$ measures the probability of observing $\mathcal{D}$ given the model specified by $\boldsymbol{\theta}$. Using Bayes rule on $\boldsymbol{\theta}$, we have

$$\pi(\boldsymbol{\theta}|\mathcal{D}) \propto \mathcal{L}(\mathcal{D}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}), \tag{15}$$

where $\pi(\boldsymbol{\theta}|\mathcal{D})$ is the *posterior* distribution of $\boldsymbol{\theta}$ conditional on the observed dataset $\mathcal{D}$ and $\pi(\boldsymbol{\theta})$ is the *prior* distribution of $\boldsymbol{\theta}$. The major challenge of using Eq. (15) in practice is to sample from the posterior, and typically, this can be handled by advanced Markov chain Monte Carlo methods.

The likelihood $\mathcal{L}(\mathcal{D}|\boldsymbol{\theta})$ may depend on both the observation error and the inherent variability of the epidemic model. Even by setting the observation error

to zero, for any specified $\boldsymbol{\theta}$ the prediction from the model is still random. If the accumulated numbers are of interest, e.g. the total number of recovered, for a large population size the variability in the prediction is expected to be small. Specifically, in a multinomial model the marginal coefficient of variation is proportional to $1/\sqrt{N P_m(t)}$. However, at the same time, the model prediction can be extremely sensitive to $\boldsymbol{\theta}$, and an almost negligible perturbation due to the (albeit small) randomness of $\boldsymbol{\theta}$ may lead to noticeably different predictions. Therefore, the Bayesian analysis is meaningful with or without the observation error.

To formulate the likelihood function, we first denote an individual (directed) random walk among various states as a Boolean operator $\boldsymbol{Y}^{(n)} = [\boldsymbol{y}_1^{(n)}, \ldots, \boldsymbol{y}_j^{(n)}, \ldots, \boldsymbol{y}_T^{(n)}]$, where $n \in [1, \ldots, N]$ and $j \in [1, \ldots, T]$ such that $t_{j+1} - t_j = 1$ [unit time].[6] The vectors $\boldsymbol{y}_j^{(n)}$ (of dimension $M_o \times 1$) represent the state of the n person at time t_j. Therefore, the components are all zero with exception of the current state, which takes the value of one. The joint probability density function of $\boldsymbol{Y}^{(n)}$, denoted by $f(\boldsymbol{Y}^{(n)}|\boldsymbol{\theta})$, is readily available from the governing equation Eq. (8). Next, we note that the observation $\mathcal{D}$ represents a collective scenario of the N-independent (under the assumptions of the macromodel) Markov processes $\boldsymbol{y}_j^{(n)}$.

Therefore, by brute force, the (observation-error-free) likelihood function has the following form

$$\mathcal{L}(\mathcal{D}|\boldsymbol{\theta}) = \sum_{y_{m,t}^{(n)}} \left[\mathbb{1}\left(\sum_n \boldsymbol{Y}^{(n)} = \mathcal{D}\right) \prod_n f\left(\boldsymbol{Y}^{(n)}|\boldsymbol{\theta}\right)\right], \tag{16}$$

where $\mathbb{1}(\cdot)$ is an indicator function and $\sum_{y_{m,t}^{(n)}}$ is a $M_o^{T \times N}$-fold summation. This brute force summation contains impossible paths that, however, are naturally excluded by the indicator function. Observe that this likelihood is fundamentally different from the deterministic compartmental models based on proportions rather than individual probabilities. Moreover, it is also different from the classical binomial (and related) likelihood approaches (used in direct Gillespie's methods). Equation (16) is clearly computationally intractable.

To formulate a computationally tractable likelihood function, we use the Markovian property and rewrite $\mathcal{L}(\mathcal{D}|\boldsymbol{\theta})$ as

---

[6] Although Eq. (8) is continuous in time, the observations are recorded in discrete time points; therefore, here t is discretized.

$$\mathcal{L}(\mathcal{D}|\boldsymbol{\theta}) = \mathcal{L}(\mathcal{D}(t_0)|\boldsymbol{\theta}) \prod_{j=1}^{T} \mathcal{L}(\mathcal{D}(t_j)|\mathcal{D}(t_{j-1}); \boldsymbol{\theta}),$$
(17)

where $\mathcal{D}(t_j)$ denotes the observation at time point t_j. The first term $\mathcal{L}(\mathcal{D}(t_0)|\boldsymbol{\theta})$ can be easily computed from a multinomial distribution with probability vector $\boldsymbol{P}(t_0)$. The specific expression of $\mathcal{L}(\mathcal{D}(t_j)|\mathcal{D}(t_{j-1}); \boldsymbol{\theta})$ varies with the adopted epidemic model, yet it is typically in the multinomial form. All the ingredients to compute $\mathcal{L}(\mathcal{D}(t_j)|\mathcal{D}(t_{j-1}); \boldsymbol{\theta})$ are included in the marginal distribution $\boldsymbol{P}(t)$, and the stochastic matrix $\boldsymbol{S}(t_j, t_{j+1}|\boldsymbol{\theta})$ is expressed as

$$\boldsymbol{S}(t_j, t_{j+1}|\boldsymbol{\theta}) = \exp\left(\int_{t_j}^{t_{j+1}} \boldsymbol{Q}(\tau|\boldsymbol{\theta}) \, d\tau\right).$$
(18)

Observe that due to the discretization of a continuous-time Markov process into a discrete-time Markov process, the matrix $\boldsymbol{S}(t_j, t_{j+1}|\boldsymbol{\theta})$ is "less sparse" than $\boldsymbol{Q}(t|\boldsymbol{\theta})$. For example, in a finite time interval, the impossible event $2 \rightarrow 4$ in matrix $\boldsymbol{Q}$ may have a finite probability of occurring in matrix $\boldsymbol{S}$ (through e.g. $2 \rightarrow 3 \rightarrow 4$). In fact, Eq. (18) can be interpreted as the result of applying Eq. (9) infinite times within the integration interval.

For a simple illustration of concept in constructing the likelihood function, we consider a two-state system where state 1 can either move to state 2 or stay still, while state 2 can only stay still. We assume $\mathcal{D}(t_{j-1})$ records [100, 50] in occupations of states 1 and 2 (for a total of 150 Markov chains), and $\mathcal{D}(t_j)$ records [90, 60]. Given the aforementioned transition structure, we know 10 out of 100 chains at t_j moves from state 1 to state 2. Therefore, the likelihood $\mathcal{L}(\mathcal{D}(t_j)|\mathcal{D}(t_{j-1}); \boldsymbol{\theta})$ is simply the binomial $\binom{100}{10} P_{1\rightarrow 2}^{10} P_{1\rightarrow 1}^{90}$, where the transition probability $P_{i \rightarrow j}$ can be directly read from Eq. (18). One may not be able to observe the populations in all compartments; in this case, the total probability theorem can be used to integrate the unobservable states out (see Sect. 4 for an example).

2.5 Addressing data unreliability

The likelihood function introduced above only considers the inherent stochastic variability of the model. In reality, on top of the inherent stochastic variability, the underlying errors/uncertainties of a reported dataset involve multiple alternative sources. A rigorous way to treat such unreliability of reported data is to introduce a distribution assumption on the error $\boldsymbol{\epsilon}$, and the likelihood function can be written as

$$\mathcal{L}(\mathcal{D}|\boldsymbol{\theta}) = \int_{\boldsymbol{\epsilon} \in \Omega_{\boldsymbol{\epsilon}}} \mathcal{L}(\mathcal{D}|\boldsymbol{\theta}, \boldsymbol{\epsilon}) \pi(\boldsymbol{\epsilon}) \, d\boldsymbol{\epsilon},$$
(19)

where $\mathcal{L}(\mathcal{D}|\boldsymbol{\theta}, \boldsymbol{\epsilon})$ is the likelihood with a specified error, $\pi(\boldsymbol{\epsilon})$ is the probability distribution of the error, and $\Omega_{\boldsymbol{\epsilon}}$ represents the feasible domain of the error. Note that in general $\boldsymbol{\epsilon}$ represents a set of discretized stochastic processes. Apart from the technical challenge of integrating the high-dimensional Eq. (19), the major challenge of incorporating the error is the specification of $\pi(\boldsymbol{\epsilon})$. Clearly, an assumption on $\pi(\boldsymbol{\epsilon})$ would reshape the likelihood function towards the shape of $\pi(\boldsymbol{\epsilon})$, and an inappropriate assumption would generate artificial and even misleading transmission properties. Therefore, we adopt an indirect path to incorporate the unreliability of reported data. Specifically, we apply a kernel function $\kappa(\cdot)$ to the original error-free likelihood function, i.e.

$$\hat{\mathcal{L}}(\mathcal{D}|\boldsymbol{\theta}) = \kappa\left(\mathcal{L}(\mathcal{D}|\boldsymbol{\theta})\right).$$
(20)

The kernel function is selected to "flatten" the likelihood function so that the unreliability in the reported data can be, to some extent, captured. In this study, we consider an exponential kernel, and Eq. (20) is rewritten as

$$\hat{\mathcal{L}}(\mathcal{D}|\boldsymbol{\theta}) = \exp \frac{\log \mathcal{L}(\mathcal{D}|\boldsymbol{\theta})}{n_{\epsilon}},$$
(21)

where $n_{\epsilon} > 1$ is a scaling factor. Clearly, if $n_{\epsilon} = 1$, $\hat{\mathcal{L}}(\mathcal{D}|\boldsymbol{\theta})$ is identical to the original likelihood $\mathcal{L}(\mathcal{D}|\boldsymbol{\theta})$, and if $n_{\epsilon} \rightarrow \infty$, $\hat{\mathcal{L}}(\mathcal{D}|\boldsymbol{\theta})$ approaches uniform.

Instead of a specific error distribution, in practice it is more likely to have a crude idea on the possible magnitude of the errors in the reported dataset. For a further simplification, we focus on the errors in the infected cases, since the causal structure of infected and recovered/dead would let the errors in infected eventually flow into recovered/dead. Therefore, the question left is to relate "the magnitude of errors in infected cases" to the n_{ϵ} in Eq. (21). It turns out, as a consequence of a sequence of qualitative reasoning, a reasonable choice of n_{ϵ} is to let

$$n_{\epsilon} \propto \frac{\Delta_{\epsilon}^2}{\Delta_{\text{infected}}},$$
(22)

where Δ_{infected} represents the maximum increment of infected and Δ_{ϵ} represents the possible error in the maximum increment of infected. Note that Eq. (22)

is proposed as a crude guidance for setting the magnitude of n_ϵ. The reasoning of Eq. (22) is described as follows. (i) In Eq. (21), if $\mathcal{L}(\mathcal{D}|\boldsymbol{\theta})$ is Gaussian, the effect of applying $1/n_\epsilon$ is to introduce a scaling factor of n_ϵ to the covariance of $\mathcal{L}(\mathcal{D}|\boldsymbol{\theta})$. (ii) The likelihood $\mathcal{L}(\mathcal{D}|\boldsymbol{\theta})$ is a product of multinomial kernels (see Eq. (17)), which can be approximated by Gaussian with the maximum variance (of the infected compartment) in the size of Δ_{infected}. (iii) Equation (22) is obtained as one assumes the scaled variance (scaled by factor n_ϵ) has a similar magnitude as Δ_ϵ^2.[7] For example, if one has a crude idea that the error of infected can be 30% of the reported infected, using Eq. (22) one could set $n_\epsilon \propto 0.09\Delta_{\text{infected}}$.

3 Entropy as a global transmission metric

In the literature, only a few studies have investigated the application of entropy in epidemics modelling [32–34]. Moreover, in the previous works the motivation and formulation of entropy, as well as the adopted epidemics modelling framework, are entirely different from the current study. The key feature of the proposed stochastic model is that entropy-based transmission measures can be naturally developed. Specifically, for a discretized time grid $\{t_j, j = 1, \ldots, T\}$ and stochastic matrix $\boldsymbol{S}(t_j, t_{j+1})$ (Eq. (18)), we consider the Shannon entropy rate expressed as

$$
\begin{aligned}
&\mathcal{H}(t_j|t_{j-1}) \\
&= -\sum_{m=1}^{M}\sum_{n=1}^{M} P_n(t_{j-1})S_{m,n}(t_{j-1}, t_j) \log(S_{m,n}(t_{j-1}, t_j)).
\end{aligned} \tag{23}
$$

For $j = 0$, $\mathcal{H}(t_0|t_{-1}) \equiv \mathcal{H}(t_0) = -\sum_{m=1}^{M} P_m(t_0)$ $\log P_m(t_0)$. Recall that the marginal distribution $\boldsymbol{P}(t)$ and the stochastic matrix $\boldsymbol{S}(t_j, t_{j+1})$ vary with the initial condition $\boldsymbol{P}(t_0)$. Therefore, Eq. (23) and $\mathcal{H}(t_0)$ should be averaged over the posterior distribution of the initial condition (obtained from the Bayesian analysis). In evaluation of Eq. (23), the convention $0 \log 0 \equiv 0$ is adopted.

In a homogeneous Markov process, the entropy rate is constant, and one has the important theoretical result

$\lim_{T\to\infty} \frac{1}{T}\mathcal{H}(t_0, t_1, \ldots, t_T) = \mathcal{H}(t_1|t_0)$. In the proposed epidemic model, the Markov process is nonlinear and non-homogeneous. Therefore, the evolution of the entropy rate $\mathcal{H}(t_j|t_{j-1})$ within a specified duration should be considered, and they characterize the evolution of the degree of disorder.

Using the Markovian property of the epidemic model in conjunction with the additive property of entropy,[8] the entropy $\mathcal{H}(t_0, t_1, \ldots, t_T)$ has the concise form

$$
\mathcal{H}(t_0, t_1, \ldots, t_T) = \sum_{j=0}^{T} \mathcal{H}(t_j|t_{j-1}). \tag{24}
$$

The entropy $\mathcal{H}(t_0, t_1, \ldots, t_T)$ is a scalar, and it provides a global measure on the total degree of disorder for an epidemic scenario. An important feature (shared by the reproductive ratio) of the entropy rate and the total entropy is that they are quantitatively comparable across different regions. This is because the entropy-based measures are associated with the statistically averaged individual, which is similar to measuring the mean-field approximation of the complex epidemic dynamics system.

Qualitative speaking, a large $\mathcal{H}(t_0, t_1, \ldots, t_T)$ may be contributed by: (i) a large pulse-like $\mathcal{H}(t_j|t_{j-1})$, i.e. the entropy rate reaches high values but stays (in high values) for a short period; (ii) a moderate flat $\mathcal{H}(t_j|t_{j-1})$, i.e. the entropy rate evolves with moderate values for a long period. In an epidemic scenario, a large pulse-like evolution of the entropy rate implies that the virus reaches a significant proportion of population but damped out (through the accumulation of recovered/dead) fast, and a flat evolution implies that the epidemic spreads in a moderate severe state for a long time. To quantitatively analyse whether the entropy rate evolution is pulse-like or flat, we introduce a concentration measure to $\mathcal{H}(t_j|t_{j-1})$. Specifically, we again adopt the concept of Shannon entropy such that the concentration measure of $\mathcal{H}(t_j|t_{j-1})$ is defined as the inverse of the Shannon entropy of the normalized $\mathcal{H}(t_j|t_{j-1})$, i.e.

$$
\begin{aligned}
\mathcal{C}(\mathcal{H}) &= \frac{1}{\mathcal{H}(\bar{\mathcal{H}}(t_j|t_{j-1}))} \\
&= -\frac{\mathcal{H}(t_0, t_1, \ldots, t_T)}{\sum_{j=0}^{T}\mathcal{H}(t_j|t_{j-1})(\log\mathcal{H}(t_j|t_{j-1}) - \log\mathcal{H}(t_0, t_1, \ldots, t_T))},
\end{aligned} \tag{25}
$$

where $\bar{\mathcal{H}}(t_j|t_{j-1}) = \mathcal{H}(t_j|t_{j-1}))/\mathcal{H}(t_0, t_1, \ldots, t_T)$ is the normalized $\mathcal{H}(t_j|t_{j-1})$. Note that the total entropy

---

[7] This assumption implies if $\Delta_\epsilon = \sqrt{\Delta_{\text{infected}}}$, n_ϵ is 1. In other words, if the error is in the size of the standard deviation of multinomial distribution, one cannot tell it is error or inherent stochastic variability.

[8] Note that the entropy is additive when the underlying distributions are statistically independent, and in Markov processes, the consecutive conditional distributions are independent.

Fig. 1 Diagram of the modified SEIR model. The transition from susceptible to exposed involves the term $\alpha_1(t)P_2(t)$, indicating the exposed is contagious
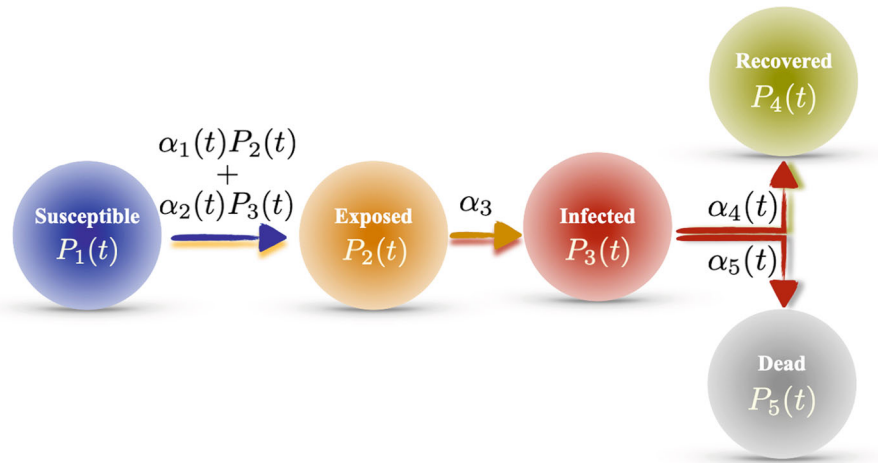
$\mathcal{H}(t_0, t_1, \ldots, t_T)$ appears in Eq. (25) as the normalizing constant of $\mathcal{H}(t_j|t_{j-1})$ (when $\mathcal{H}(t_j|t_{j-1})$ is normalized into a probability mass function). Also note that the Shannon entropy instead of the variance-based measures is adopted since a large variance does not necessarily reflect a large dispersion (e.g. a mixture model with highly concentrated component densities could produce a large variance).

In this paper, we propose the entropy rate $\mathcal{H}(t_j|t_{j-1})$, the entropy $\mathcal{H}(t_0, t_1, \ldots, t_T)$, and the concentration factor $\mathcal{C}(\mathcal{H})$ as complements to the conventional reproductive ratio. Appendix A illustrates various attractive features of the entropy-based measures. In practice, instead of computing entropy-based measures for the original distribution vector $\boldsymbol{P}$ and the stochastic matrix $\boldsymbol{S}$, one may need to reshape $\boldsymbol{P}$ and $\boldsymbol{S}$ to obtain measures of different scales. For example, a typical epidemic model may involve the states of recovered and dead. Naturally, one would prefer the scenario of "a large recovery probability and a small death probability" over "a large death probability and a small recovery probability." However, Eq. (23) or Eq. (24) does not differentiate between recovery and death, and the aforementioned two scenarios can have exactly the same entropy (rate). The conventional reproductive ratio measure has the same issue. To let the entropy-based measures incorporate the concept of "high recovery probability is preferable over high death probability", one could reshape the distribution vector $\boldsymbol{P}$ and the stochastic matrix $\boldsymbol{S}$ by merging the recovery state with the infected state. Consequently, the entropy (rate) of the reshaped system would diminish the contribution

from the recovered state and highlight the contribution from the dead state (see Appendix A for an example).

4 Application to COVID-19

In the light of the general framework introduces in Sects. 2 and 3, this section introduces a simple modification of the SEIR model with the exposed being also contagious.

4.1 Modified SEIR

The modified SEIR has a five-dimensional probability state vector $\boldsymbol{P}(t)$ described as follows:

- $P_1(t)$: the (instantaneous probability of being) susceptible.
- $P_2(t)$: the exposed.
- $P_3(t)$: the infected.
- $P_4(t)$: the recovered.
- $P_5(t)$: the dead.

The rate matrix $\boldsymbol{Q}(\boldsymbol{P}(t), t)$ is written as

$$\boldsymbol{Q}(\boldsymbol{P}(t), t)$$

$$= \begin{bmatrix} -(\alpha_1(t)P_2(t) + \alpha_2(t)P_3(t)) & 0 & 0 & 0 & 0 \\ \alpha_1(t)P_2(t) + \alpha_2(t)P_3(t) & -\alpha_3(t) & 0 & 0 & 0 \\ 0 & \alpha_3(t) & -(\alpha_4(t) + \alpha_5(t)) & 0 & 0 \\ 0 & 0 & \alpha_4(t) & 0 & 0 \\ 0 & 0 & \alpha_5(t) & 0 & 0 \end{bmatrix},$$

(26)

where $\boldsymbol{\alpha}(t) = [\alpha_1(t), \ldots, \alpha_5(t)]$ are non-negative parameters to be calibrated. Note that here for generality we write every parameter as time dependent; however, in practice it is typically sufficient to set only a few

of them as time dependent. Figure 1 illustrates the flow between compartments of the modified SEIR model.

4.2 Likelihood function

The likelihood function can be derived as a simple application of the concepts introduced in Sect. 2.4. First, we introduce a compound state, denoted by $1 \vee 2$, to represent the state of being in either susceptible or exposed. The most important property of the compound state $1 \vee 2$ is that it is an observable, i.e. if an individual is not at the state of infected nor at the state of recovered/dead, it is in the compound state. We let $P_{1 \vee 2 \rightarrow m}(t_j, t_{j+1})$ represent the transition probability from the compound state at t_j to other states m, $m = 3$ (infected), 4 (recovered), 5 (dead) at t_{j+1}. Using the total probability theorem, $p_{1 \vee 2 \rightarrow m}(t_j, t_{j+1})$ can be expressed as

$$
\begin{aligned}
& P_{1 \vee 2 \rightarrow m}(t_j, t_{j+1}) \\
& = \frac{P_1(t_j)}{P_1(t_j) + P_2(t_j)} P_{1 \rightarrow m}(t_j, t_{j+1}) \\
& \quad + \frac{P_2(t_j)}{P_1(t_j) + P_2(t_j)} P_{2 \rightarrow m}(t_j, t_{j+1}),
\end{aligned}
\tag{27}
$$

where $P_1(t_j)$ and $P_2(t_j)$ are solution of Eq. (8), and $P_{1 \rightarrow m}$ and $P_{2 \rightarrow m}$ can be obtained from Eq. (18). Next, we arrange the dataset vector $\mathcal{D}(t_j)$ in the form $\mathcal{D}(t_j) = \left[\mathcal{D}_{1 \vee 2}(t_j), \mathcal{D}_3(t_j), \mathcal{D}_4(t_j), \mathcal{D}_5(t_j) \right]$ to, respectively, represent the instantaneous number of compound state, instantaneous number of infected, accumulative number of recovered, and accumulative number of dead. Let $\Delta \mathcal{D}(t_j, t_{j+1}) := |\mathcal{D}(t_{j+1}) - \mathcal{D}(t_j)|$ represent the absolute difference between two consecutive dataset vectors.

Before introducing the likelihood function, we introduce an additional assumption that the $\Delta \mathcal{D}_{1 \vee 2}(t_j, t_{j+1})$ number of Markov chains all transit to the state 3 (the infected). This assumption can be always (made) correct since: (i) if t_j is sufficiently close to t_{j+1}, naturally one cannot jump to the recovered/dead state from susceptible/exposed; (ii) if $t_{j+1} - t_j$ is large, one could re-mesh the timescale and perform interpolation on the dataset, so that t_j can always be close to t_{j+1} *by construction*. Finally, the conditional likelihood function $\mathcal{L}(\mathcal{D}(t_{j+1}) | \mathcal{D}(t_j); \boldsymbol{\theta})$ can be written as

$$
\mathcal{L}(\mathcal{D}(t_{j+1}) | \mathcal{D}(t_j); \boldsymbol{\theta}) = \mathcal{L}_1 \mathcal{L}_2,
\tag{28}
$$

where

$$
\mathcal{L}_1 = \frac{\mathcal{D}_{1 \vee 2}!}{(\mathcal{D}_{1 \vee 2} - \Delta \mathcal{D}_{1 \vee 2})! \Delta \mathcal{D}_{1 \vee 2}! 0! 0!}
$$

$$
(P_{1 \vee 2 \rightarrow 1 \vee 2})^{\mathcal{D}_{1 \vee 2} - \Delta \mathcal{D}_{1 \vee 2}}
$$
$$
(P_{1 \vee 2 \rightarrow 3})^{\Delta \mathcal{D}_{1 \vee 2}} (P_{1 \vee 2 \rightarrow 4})^0 (P_{1 \vee 2 \rightarrow 5})^0
\tag{29}
$$

and

$$
\begin{aligned}
\mathcal{L}_2 = {} & \frac{\mathcal{D}_3!}{(\mathcal{D}_3 - \Delta \mathcal{D}_4 - \Delta \mathcal{D}_5)! \Delta \mathcal{D}_4! \Delta \mathcal{D}_5!} \\
& (P_{3 \rightarrow 3})^{\mathcal{D}_3 - \Delta \mathcal{D}_4 - \Delta \mathcal{D}_5} \\
& (P_{3 \rightarrow 4})^{\Delta \mathcal{D}_4} (P_{3 \rightarrow 5})^{\Delta \mathcal{D}_5}.
\end{aligned}
\tag{30}
$$

In Eqs. (29) and (30), the notations are simplified to drop t_j, t_{j+1} and $\boldsymbol{\theta}$. To avoid possible ambiguity, the simplification rules are: $\mathcal{D}_m \equiv \mathcal{D}_m(t_j)$, $\Delta \mathcal{D}_m \equiv \Delta \mathcal{D}_m(t_j, t_{j+1})$, $P_m \equiv P_m(t_j | \boldsymbol{\theta})$, and $P_{m \rightarrow m'} \equiv P_{m \rightarrow m'}(t_j, t_{j+1} | \boldsymbol{\theta})$. Substituting Eqs. (28), (29) and (30) into (17), one obtains the complete likelihood function.

4.3 Transmission measures

To obtain the entropy-based measures, we reshape the five-dimensional vector $\boldsymbol{P}$ into $[P_{1 \vee 2}, P_{3 \vee 4}, P_5]^\top$, and the corresponding stochastic matrix $\boldsymbol{S}$ is also reshaped (via the total probability theorem) accordingly. The reason to consider the compound state $1 \vee 2$ is because as a whole the state $1 \vee 2$ is an observable, therefore the possible errors in identifying the exposed can be marginalized out. The reason to consider the compound state $3 \vee 4$ is discussed in Sect. 3, and as a result, the concept "high recovery probability is preferable over high death probability" is correctly incorporated.

In addition to the entropy-based measures, using the next-generation matrix approach [35], the instantaneous reproductive ratio of the modified SEIR model can be defined as

$$
R_0(t) := \frac{\alpha_1(t)}{\alpha_3(t)} + \frac{\alpha_2(t)}{\alpha_4(t) + \alpha_5(t)}.
\tag{31}
$$

Note that the instantaneous reproductive ratio $R_0(t)$ can be understood as the basic reproductive ratio of a tangent model defined as a constant model with parameters $\boldsymbol{\alpha}$ equal to the instantaneous parameters $\boldsymbol{\alpha}(t)$ at the reference time point t.

4.4 Modelling and computational details

As aforementioned, the whole parameter set $\boldsymbol{\theta}$ not only involves parameters of the rate matrix $\boldsymbol{Q}(\boldsymbol{P}(t), t)$, i.e. $\boldsymbol{\alpha}(t)$ (represented by $\boldsymbol{w}, \boldsymbol{w}'$ of basis functions), but also parameters to represent the initial state, i.e. $\boldsymbol{\beta}$.

In the modelling practice, except α_3, which is related to the mean incubation period, we calibrate all the other parameters (including the initial conditions) with Bayesian analysis. The mean incubation period, which is $1/\alpha_3$ in the model, is reported in various previous studies [11,36], and typically, it is around 5 and in the range of [3, 7] days. Therefore, we set $1/\alpha_3$ as an epistemic random variable within [3, 7]. The time-dependent parameters are modelled with sigmoid basis functions. The number of function basis for each parameter is determined in an additive manner. Specifically, we start with constant α and iteratively increase the number of basis functions until the variation in likelihood function value (Eq. (17)) or BIC index becomes small.

The Gibbs sampling with a uniform proposal distribution for each component of $\boldsymbol{\theta}$ is adopted to sample from the posterior distribution. The step size of the Gibbs sampling is adaptively tuned using the acceptance rate of the Markov chain [37,38]. The seed samples for the Gibbs sampler are selected in the neighbourhood of the posterior mode. This is obtained by sequential Monte Carlo method [39,40] combined with deterministic trust region optimization [41,42].

5 Modelling results on real datasets of COVID-19

5.1 Datasets

For the studied regions, the time series of the populations of infected, recovered, and dead during January to May 2020 are used in model calibration. The data are collected from WHO, European CDC and Chinese CDC [11,43,44]. The regions considered in this study include: Hubei province, South Korea, Italy, Germany, Spain, and France.[9] For each region, the population size N is fixed to the most recent value reported by Worldometer [45]. We choose these countries/regions because they have the same order of population size (this is irrelevant to the entropy-based measure which is N independent), they applied different containment strategies, and they represent different cultures. Moreover, at the time of writing of this article the peak of the epidemic waves is passed. A complete and thorough analysis of a large number of regions is out of the scope

of the current study. In fact, here we focus primarily on the model and metric definition and their use.

5.2 Data correction

Due to abrupt counting policy changes and various corrections, the COVID-19 datasets for Hubei, Spain, and France not only violate the smoothness assumption[10] of the proposed modelling framework, but also contradict the fundamental fact that the accumulative number can only be non-decreasing. Therefore, the datasets must be corrected. It is obvious that the cluster/jump of data has a missing information, which is the (correct) time of occurrence. To obtain a consistent dataset we fill this missing information by using the expected time of occurrence with respect to the distribution of the previous events. Since the dataset is recorded daily, marginally they form a multinomial distribution along the discrete time axis. It follows that the missing time information can be filled by using the daily expected number of events. Specifically, let t_J represent the time point when the jump/drop happens (for a specified compartment), and let $\Delta\mathcal{D}_J$ represent the magnitude of the data jump/drop. We perform a postprocess of the dataset expressed as follows:

$$\mathcal{D}(t_i) \leftarrow \left(1 + \frac{\Delta\mathcal{D}_J}{\mathcal{D}(t_J)}\right) \mathcal{D}(t_i), \tag{32}$$

where $t_i = t_0, t_1, \ldots, t_J$, and $\mathcal{D}(t_i)$ represents the cumulative[11] number at t_i. Note that $\Delta\mathcal{D}_J$ could be negative. For an illustration of the correction, Fig. 2 shows the raw and the corrected datasets for Hubei province.

5.3 The overall epidemic dynamics of various regions

After performing model calibrations on datasets of various regions, we present a comparison analysis based on various transmission measures. Figure 3 shows the evolution of the entropy rate of COVID-19 outbreaks for each of the regions considered. This graph represents the time evolution of the degree of disorder (in

---

[9] We planned to include also UK; however, the data on the recovered patients are not available yet (unfortunately).

[10] Due to the inherent stochastic variability, the random samples drawn from the model are naturally non-smooth. However, here the "violence of smoothness assumption" indicates that an **artificial manipulation** of data will corrupt the smoothness of propagator $\boldsymbol{Q}$.

[11] In Eq. (32), $\mathcal{D}(t_i)$ has to be the cumulative number instead of instantaneous one.

Fig. 2 Raw and corrected datasets of Hubei province. There are two policy changes regarding the dataset: (i) in February 12, 2020, the diagnosis criterion was temporarily relaxed, and as a result, there is an artificial jump in the number of infected; (ii) in April 17, 2020, the cumulative number of infected and dead is altered by a constant jump, and the cumulative number of recovered is altered by a constant drop. The jumps/drops are marked by a rectangular in the figure. The populations of infected, recovered, and dead are corrected using Eq. (32). Note that for infected the correction is made on cumulative numbers, and then, the instantaneous infected is obtained by subtracting the accumulative recovered and dead
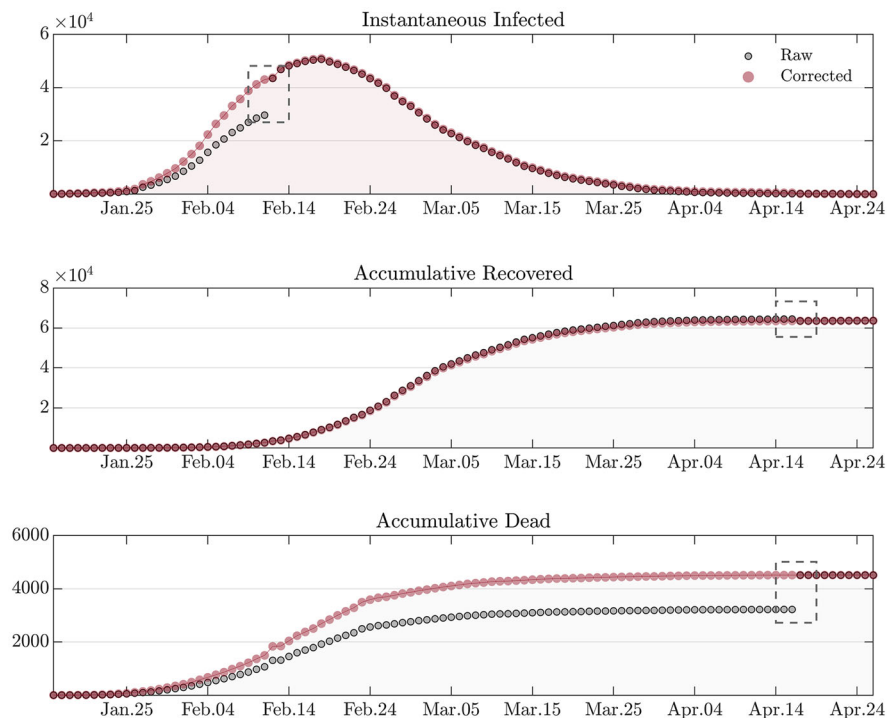
Fig. 3 The entropy rates for various regions. The figure shows the temporal evolution of entropy rate for various regions. The solid lines correspond to the posterior mean estimations, and the shaded areas correspond to $\{10\%, 20\%, \ldots, 99\%\}$ credible intervals (around the posterior mean)
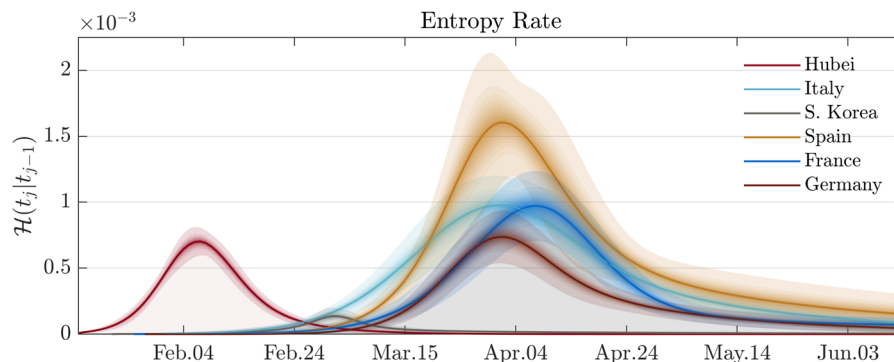
terms of infections and deaths) introduced by the virus in an average statistical individual of the region. This graph reflects features of the daily evolution of infection and recovered/deaths, but it is fundamentally different from the evolution of each compartment. In fact, it has the key property of being objective and comparable between regions. Interestingly, the evolution of the entropy rate has a similar form for each region, but a significant difference in the magnitude of the disorder. In particular, the cumulative integral of the entropy rate represents the change of entropy in the system and, therefore, the total impact in a region. In Fig. 4—top

panel, we report this impact measure for each of the regions considered. Based on this metric, Spain was the most affected region despite the epidemic wave hit the country later than Italy. On the opposite side, South Korea is the country with the least change in entropy, highlighting an effective combination of policies and cultural habits that limited the impact of the epidemic. This is probably due to the experience gained during the recent 2015 Middle East Respiratory Syndrome coronavirus (MERS-CoV) outbreak [46]. Also, Hubei's reaction, with extreme containment measures, has overall limited the impact of the epidemic. Germany has the

Fig. 4 The entropies and concentration factors for various regions. The figure shows a comparison of total entropies and concentration factors for various regions, with the violins illustrating the posterior distribution
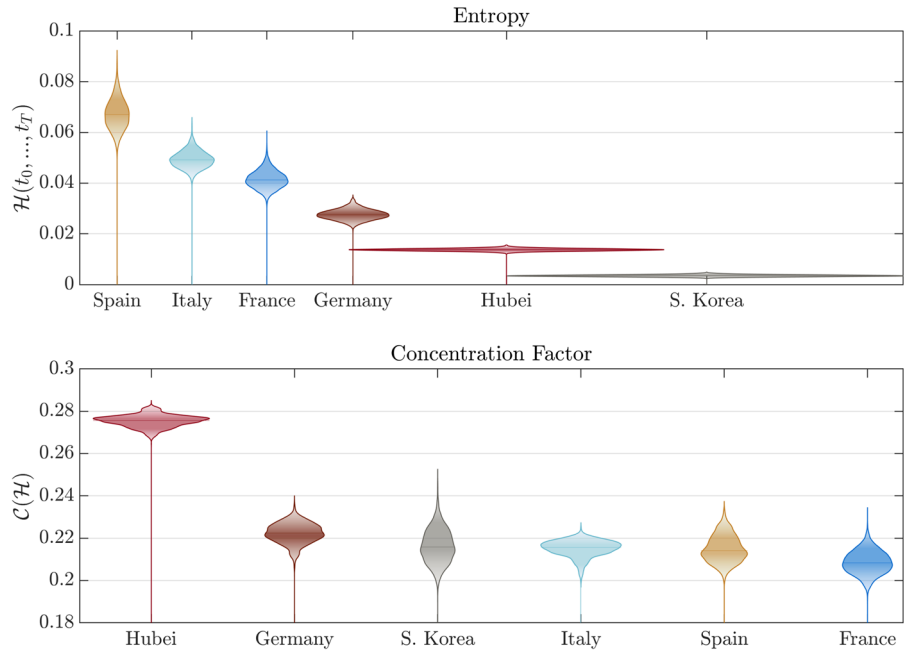
smallest total entropy among studied European countries.

Interestingly, the peak of entropy rate for Spain, Italy, and Germany occurred in about the same period but with a different left tail behaviour (i.e. in the growing phase). On the other hand, the behaviour of the right tail (i.e. the descent phase) is similar, showing a fatter and longer tail. A similar asymmetry can also be observed in Hubei and South Korea. A deviation from this "classic" behaviour is represented by Hubei, which does not show this long tail behaviour but has a rather compact and almost symmetric shape. A surprising result is shown in Fig. 4—bottom panel. Although the impact in each country is significantly different, the concentration factor is similar to support the fact that the evolution of COVID-19 is similar for all outbreaks. The Hubei region is slightly deviating from this trend, showing a higher concentration factor corroborating the lack of a right fat tail and, therefore, showing a higher prevalence as an impulse.

Figure 5 shows a comparison of the instantaneous reproductive ratio and death rate, together with the date of lockdown in each region. One can infer that the lockdown reduced $R_0(t)$ effectively. However, surprisingly, the most effective decrease has been observed in South Korea where *no* national lockdown has been imple-

mented, but only local containment measures, and massive early-stage testing.

It is important to note that the modelling results are associated with the optimized parsimonious model for each region. Specifically, in an optimized parsimonious model the number of time-dependent variables as well as the number of adaptive basis functions for each time-dependent variable is optimized, in the sense that increasing the number would not noticeably improve the calibration accuracy and decreasing the number would significantly degrade the accuracy. Finally, for an illustration on the degree of accuracy the model has achieved, the model calibration results of Hubei are shown in Fig. 6. The calibration for the other countries and their limitations are reported in B.

5.4 Robustness on the transmission trend

A natural concern regarding the discovered transmission trend is that if the trend is a genuine underlying structure of the epidemic, or it is merely some artificial/superficial structures from the specific time-dependent model. It is challenging to (perfectly) resolve this concern because a compartmental model (or any mathematical model) is inevitably an approxi-

Fig. 5 The instantaneous reproductive ratio, recovery, and death rates for various regions. The lockdown date for each region is shown as vertical dashed line. Note that South Korea does not have a lockdown policy
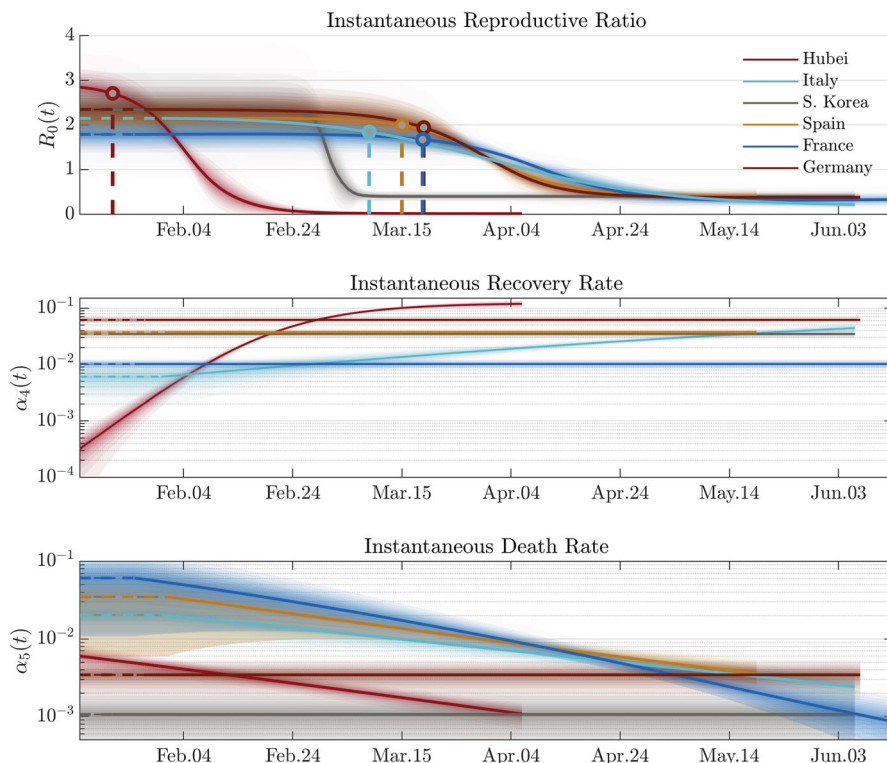

Fig. 6 Modelling the overall epidemic dynamics of Hubei province with the modified SEIR model. The red line corresponds to the posterior mean estimation. The shaded area corresponds to $\{10\%, 20\%, \ldots, 99\%\}$ credible intervals around the posterior mean. The parameters $\alpha_1(t)$, $\alpha_4(t)$ and $\alpha_5(t)$ are modelled with a single sigmoid basis, and α_2 is modelled with a constant variable. The n_ϵ in Eq. (21) is fixed to 100, assuming the error in the increment of infected is of the order of a few hundreds. The figure suggests a highly accurate calibration on data using at most one adaptive basis function for each parameter
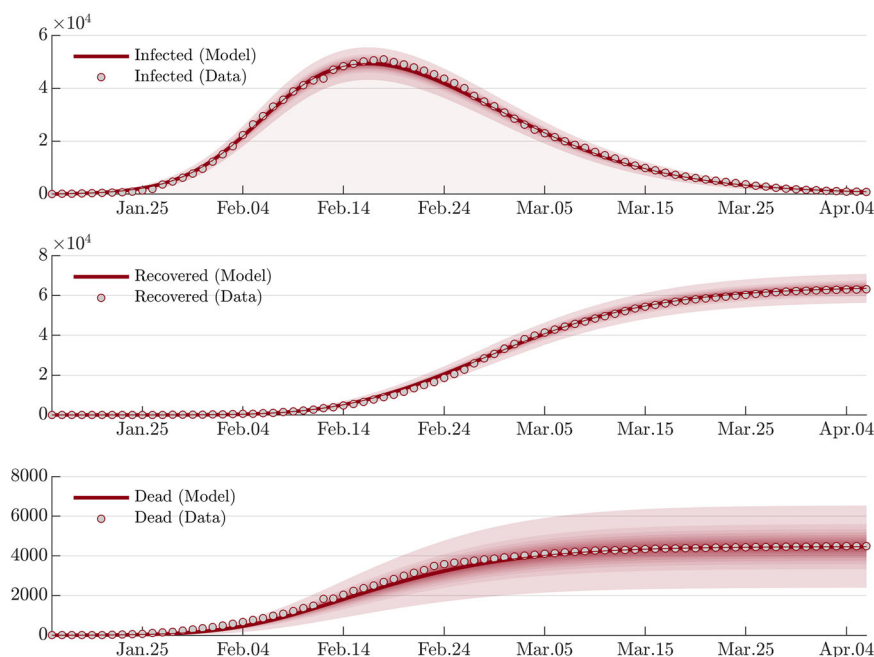
mation on the real epidemic. Moreover, even an exact model exists, it is still challenging (if not impossible) to accurately identify the model due to the presence of endogenous variables. However, at least we could show that the proposed framework is self-consistent. In C, we simulate artificial epidemics from analytical SEIR laws and investigate whether the proposed modelling approach could identify correct transmission trends.

6 Limitations and future research directions

6.1 Incorporating the undetected cases

In Sect. 5, the reported/observed population in each compartment is used to calibrate the model, and the kernel function in Eq. (21) only flattens the likelihood function instead of altering its intrinsic shape. Consequently, the model describes an epidemic scenario consistent with but also confined by the reported cases. An important missing issue to address is to incorporate the undetected cases to fully uncover the magnitude of the epidemic. A practical modelling strategy is to introduce a probability distribution assumption on the (possibly time-dependent) ratio between reported and undetected cases and rewrite the likelihood function similar to Eq. (19). Clearly, the critical ingredient is the model assumption on the undetected. The ongoing studies on blood test for antibodies of SARS-CoV-2 [47] can be useful for this future research direction.

6.2 Application to more complex compartmental models

Depending on the modelling purposes, one could introduce additional compartments, e.g. the tested/suspected, the ICU case, the female and male, the old and young, etc., to study the interactions between different groups. It is also straightforward to include spatial distributed information by including adjacency and incidence matrices. However, one should be aware that the model variance and the possibility of converging to local insignificant likelihood modes in general would increase with model complexity. Therefore, it would be crucial to collect robust prior knowledge regarding the modelling parameters.

7 Conclusions

In this study, we have proposed a stochastic compartmental modelling framework of epidemics equipped with entropy-based metrics to measure both the impact and the evolution of a pandemic event. The model belongs to the nonlinear Markov processes class, which allows a robust formulation and a natural setting for developing entropy-based metrics. In addition, we have provided a complete Bayesian inversion scheme to calibrate the model parameters with related uncertainties. Subsequently, we specialized the proposed structure to a modified SEIR model and the COVID-19 pandemic. In particular, we used the framework to investigate six regions: Hubei, South Korea, Italy, Spain, Germany, and France. We showed that the change in entropy in the selected areas (which is associated with the impact of an epidemic) is significantly different. However, it is surprising to note that the dynamic evolution of pandemic waves shows very regular trends and very similar concentration measures.

A Illustration on the entropy-based measures

To avoid unnecessary complications, we consider a simple stochastic SIR model with the constant parameters $\alpha = [\alpha_1, \alpha_2, \alpha_3]$ associated with the infection, recovery and death rates, respectively. First, we consider two systems with: (i) $[\alpha_1, \alpha_2, \alpha_3] = [0.06, 0.015, 0.005]$ for System 1, and (ii) $[\alpha_1, \alpha_2, \alpha_3] = [0.3, 0.075, 0.025]$ for System 2. The two systems have an identical basic reproductive ratio of 3.0. The initial

Fig. 7 Illustration of entropy-based measures for systems with the same basic reproductive ratio. The final stage of $\boldsymbol{P}(t)$ for the two systems is identical since they have an identical reproductive ratio. However, System 2 evolves more rapidly than System 1, and this is reflected by the entropy rate and concentration factor. System 1 has a longer active period than System 2, and consequently, the total entropy of System 1 is larger
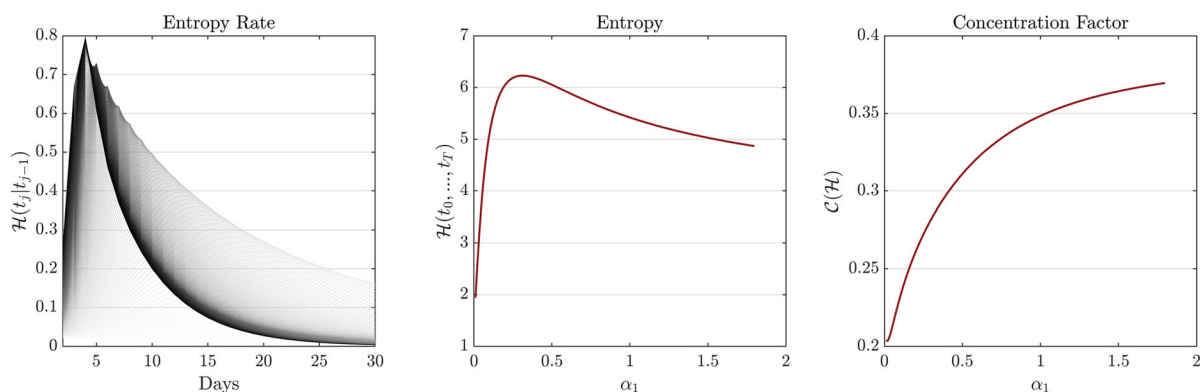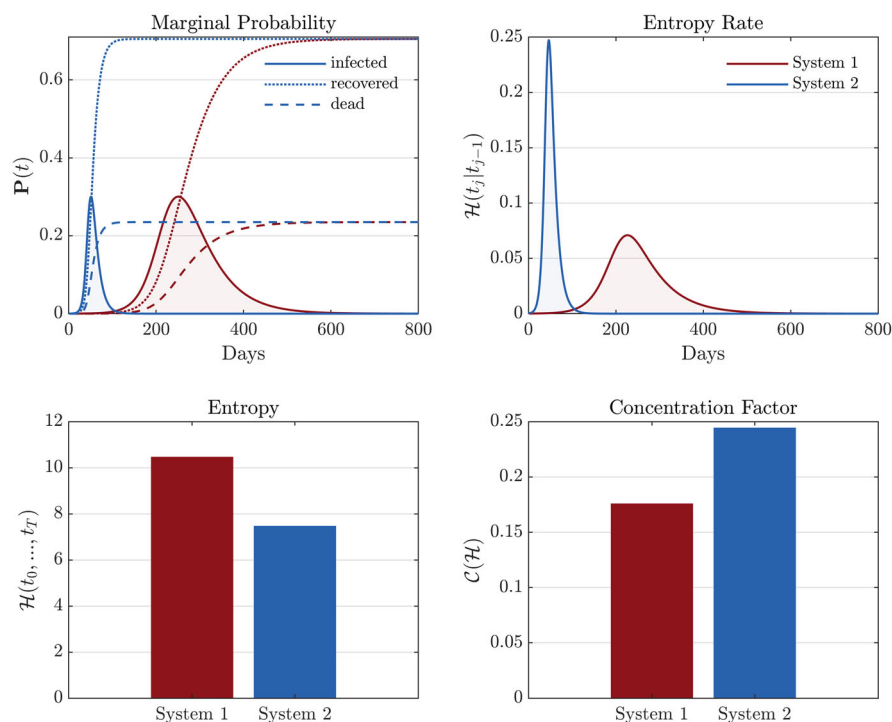
Fig. 8 Illustration of entropy-based measures with varying α_1 values in a stochastic SIR model. The left figure illustrates the time evolution of entropy rate with different α_1 values, and the lighter to darker curves are associated with lower to higher α_1 values. The middle and right figures illustrate the total entropy and concentration factor with respect to α_1. It is seen from the figure that as the force of infection grows (while the recovery and death rate are fixed), the evolution of the entropy rate becomes more and more pulse like. A more important observation is that the entropy may have a peak value due to the trade-off between strong-phase magnitude and duration of the entropy rate evolution

condition is set to $\boldsymbol{P}(t_0) = [0.99, 0.01, 0, 0]$. Figure 7 illustrates entropy-based measures for the two systems.

Next, we fix $\alpha_2 = 0.15$ and $\alpha_3 = 0.05$, and let α_1 vary within $[0.2, 2]$, so that the basic reproductive ratio varies within $[1, 10]$. The initial condition is set to $\boldsymbol{P}(t_0) = [0.99, 0.01, 0, 0]$. Figure 8 illustrates the entropy rate, total entropy, and concentration factor.

Both Figs. 7 and 8 imply that the basic reproductive ratio and the entropy-based measures describe different aspects of the epidemic dynamics, and they do not have a one-on-one mapping. Crudely speaking, the entropy rate characterizes the instantaneous intensity and flatness of the epidemics, while the total entropy measures the overall impact. Note that in Figs. 7 and
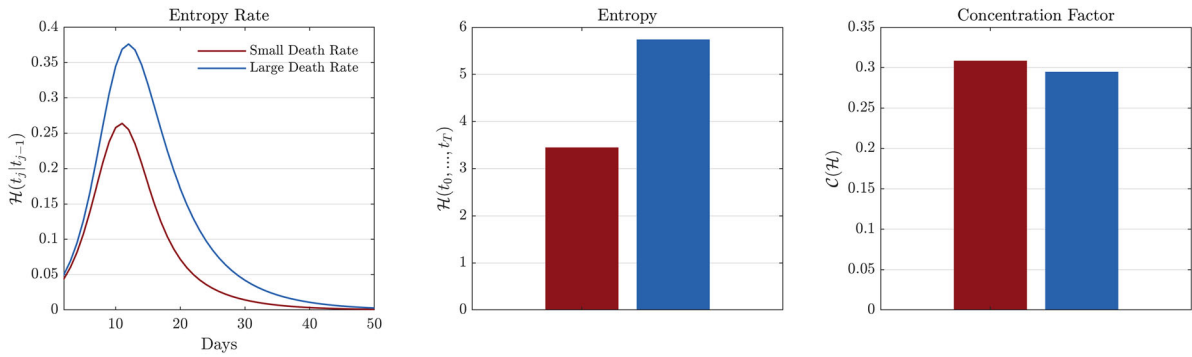
Fig. 9 Illustration of the entropy-based measures using the reshaped P and S. The figure suggests that when the infected and recovered states are merged into a compound state, the contribution (to the entropy rate and entropy) from the death cases is increased. The concentration factor is not sensitive to this technique. Note that without using a reshaped P and S, the entropy-based measures for the two systems will be exactly the same

8, the entropy-based measures are computed using the original distribution vector P and stochastic matrix S. In Sects. 3 and 4.3, it is mentioned that P and S can be reshaped to highlight the contribution from the death cases. Figure 9 illustrates the effect of this technique. Specifically, two SIR models with the same initial conditions $P(t_0) = [0.99, 0.01, 0, 0]$ are considered. The first model has $\alpha = [0.6, 0.19, 0.01]$, and the second model is obtained by swapping the recovery and death rates of the first model. The basic reproductive ratios for the two models are both 3.0.

B Results of model calibration

The model calibration results for Italy, South Korea, Spain, France, and Germany are shown in Figs. 10, 11, 12, 13, and 14. In the figures, the solid lines correspond to posterior mean estimations, and the shaded areas correspond to {10%, 20%, ..., 99%} credible intervals around the posterior mean.

Fig. 10 Modelling the overall epidemic dynamics of Italy with the modified SEIR model. The $\alpha_1(t)$, $\alpha_4(t)$ and $\alpha_5(t)$ are modelled with a single sigmoid basis, and $\alpha_2(t)$ is modelled as a constant variable. The n_ϵ in Eq. (21) is fixed to 1000, assuming the error in the increment of infected is in the order of a few thousands
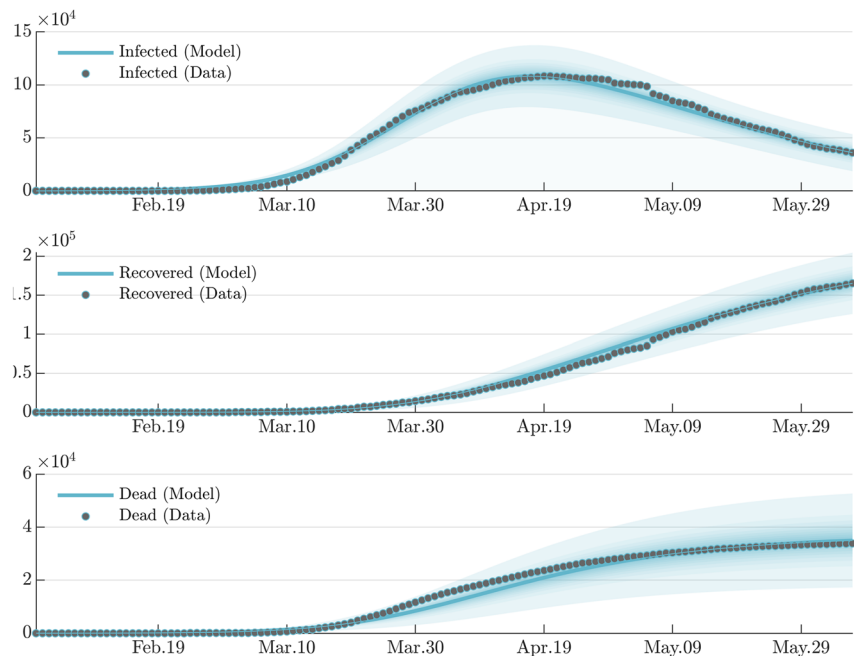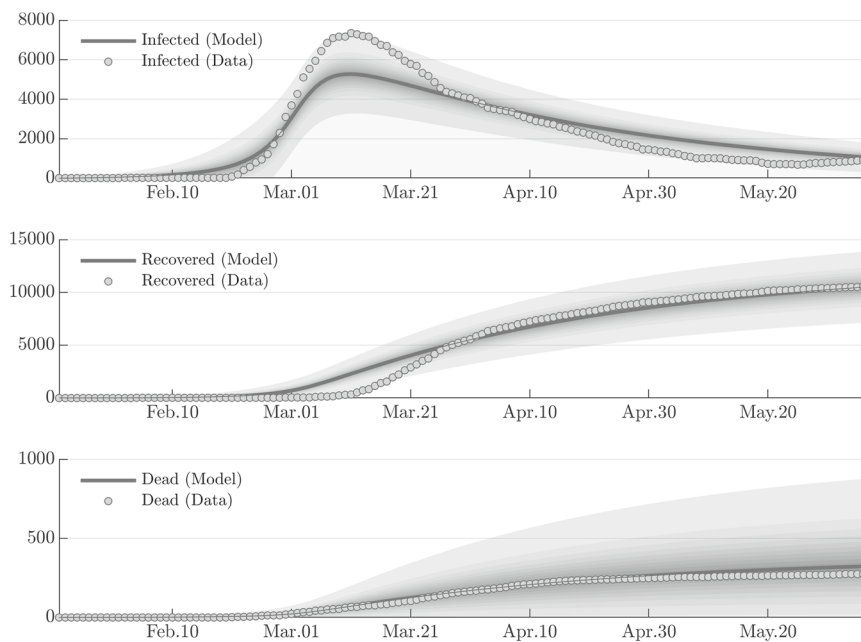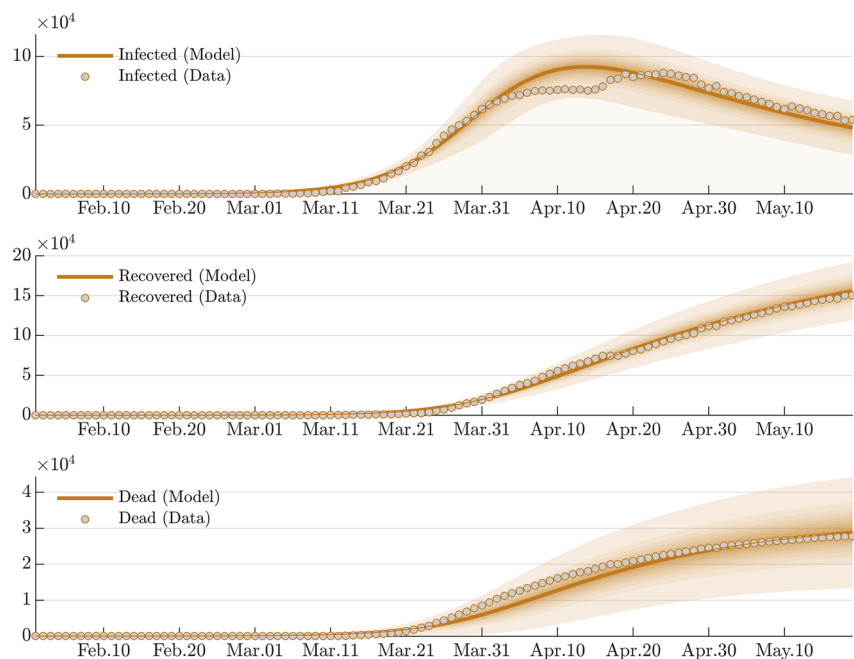
Fig. 11 Modelling the overall epidemic dynamics of South Korea with the modified SEIR model. The $\alpha_1(t)$ is modelled with a single sigmoid basis, and $\alpha_2(t)$, $\alpha_4(t)$, and $\alpha_5(t)$ are modelled as constant variables. The n_ϵ in Eq. (21) is fixed to 100, assuming the error in the increment of infected is in the order of a few hundreds
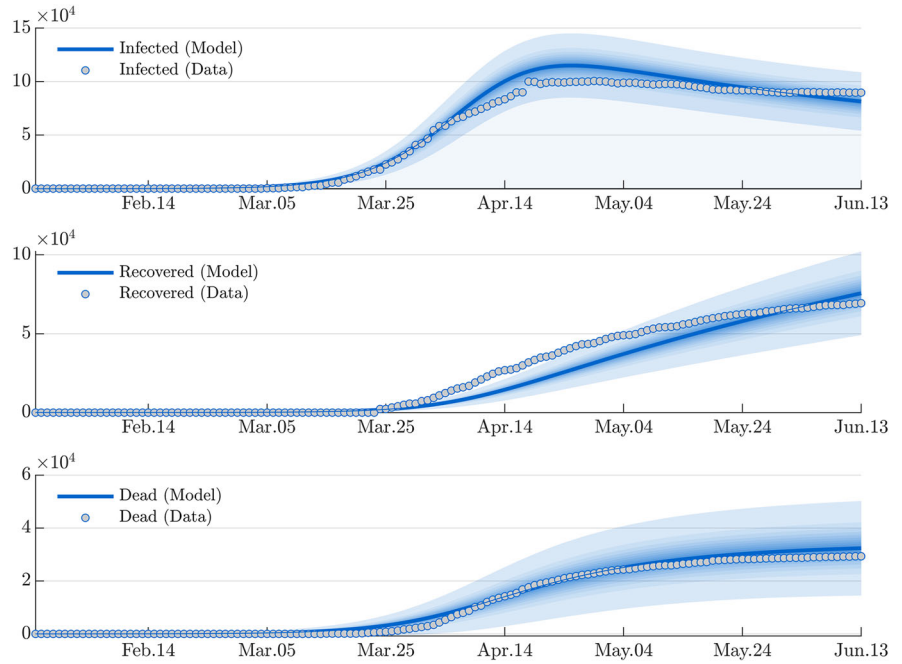

Fig. 12 Modelling the overall epidemic dynamics of Spain with the modified SEIR model. The $\alpha_1(t)$ and $\alpha_5(t)$ are modelled with a single sigmoid basis, and α_2 and α_4 are modelled as constant variables. The n_ϵ in Eq. (21) is fixed to 1000, assuming the error in the increment of infected is in the order of a few thousands

Fig. 13 Modelling the overall epidemic dynamics of France with the modified SEIR model. The $\alpha_1(t)$ and $\alpha_5(t)$ are modelled with a single sigmoid basis, and α_2 and α_4 are modelled as constant variables. The n_ϵ in Eq. (21) is fixed to 1000, assuming the error in the increment of infected is in the order of a few thousands

Fig. 14 Modelling the overall epidemic dynamics of Germany with the modified SEIR model. The $\alpha_1(t)$ is modelled with a single sigmoid basis, and $\alpha_2(t)$, $\alpha_4(t)$, and $\alpha_5(t)$ are modelled as constant variables. The n_ϵ in Eq. (21) is fixed to 1000, assuming the error in the increment of infected is in the order of a few thousands
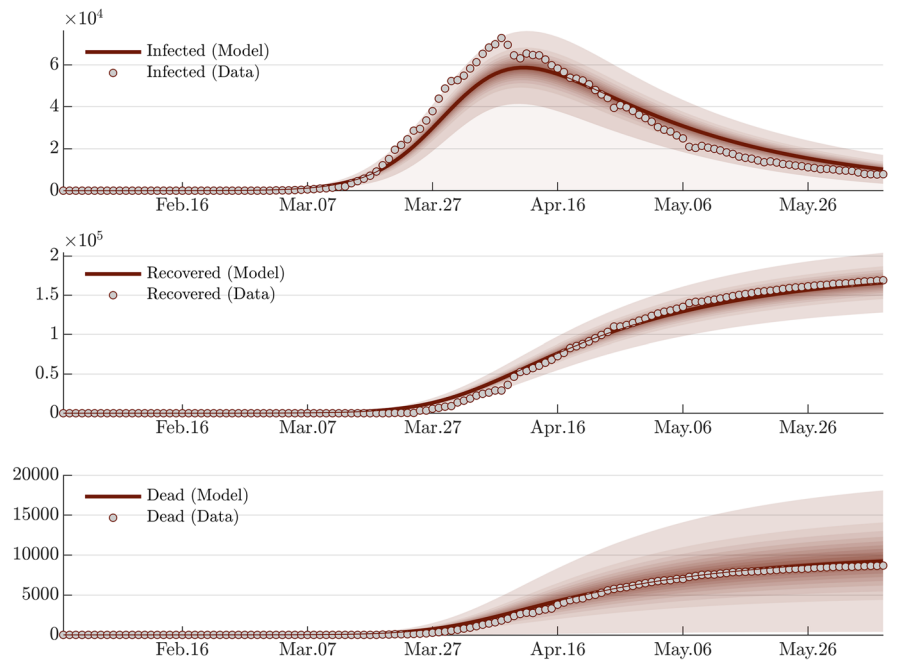
C Results on robustness/self-consistent test on the transmission trend

We simulate a random epidemic scenario from a modified stochastic SEIR model with parameters $\boldsymbol{\alpha} =$ $[\alpha_1(t), 0.10, 0.20, \alpha_4(t), \alpha_5(t)]$. The duration of the simulation is set to 40 days. The time-dependent parameters are specified as $\alpha_1(t) = 0.6 - 0.5t/40$, $\alpha_4(t) = 0.05 + 0.30t/40$, and $\alpha_5(t) = 0.15 - 0.10t/40$, and consequently, the instantaneous repro-

Fig. 15 The calibrated time-dependent modified SEIR model. The $\alpha_1(t)$, $\alpha_2(t)$, $\alpha_4(t)$, and $\alpha_5(t)$ are modelled with a single sigmoid basis, and α_3 is modelled as a constant variable. The n_ϵ in Eq. (21) is fixed to 1, assuming no error in the dataset. The figure suggests an accurate calibration. Even the unobserved exposed population and its initial condition are accurately identified. However, this is because the dataset is generated from an analytical model. For real dataset where a mathematical model is only an approximation, it is preferable to use real clinical data to specify the mean incubation period and its epistemic uncertainty (as it is performed in the main text)
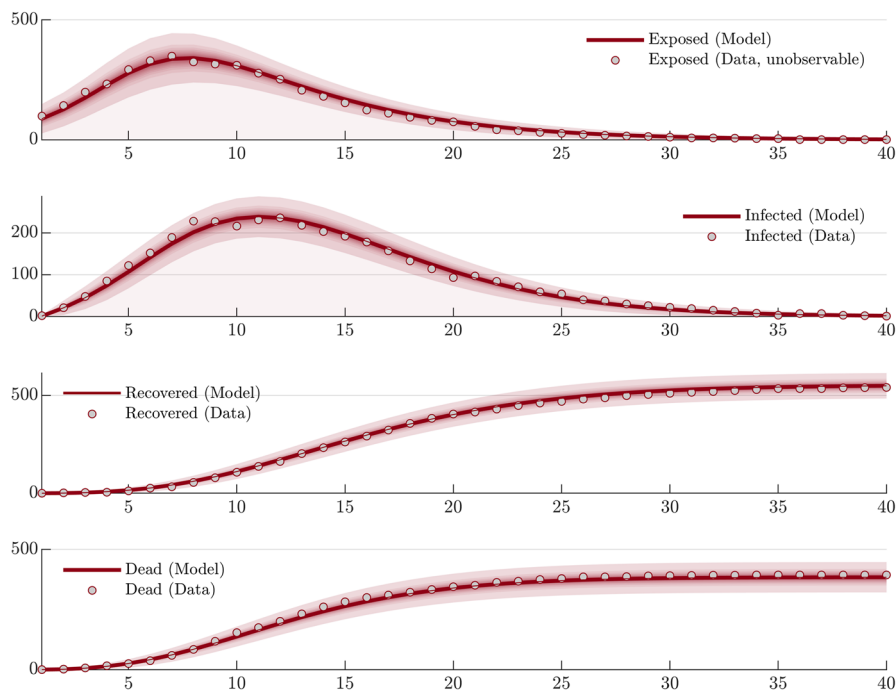
Fig. 16 The transmission properties of the time-dependent SEIR model compared with the exact values. The figure suggests that the transmission trends (in terms of the posterior mean estimation) are in general identified accurately. The moderate bias is mainly caused by the inherent variability of the stochastic simulation (see the following verification)
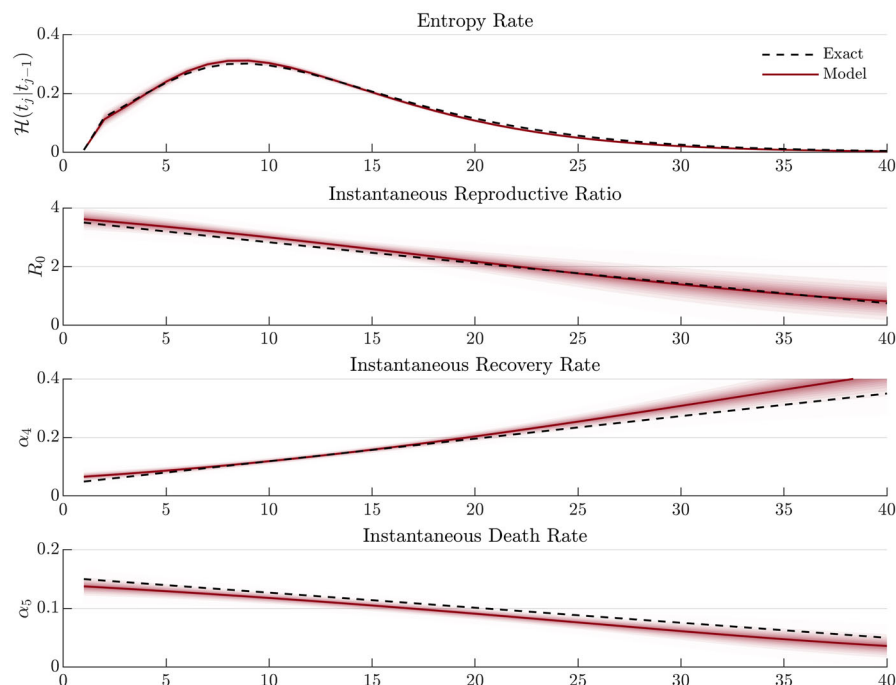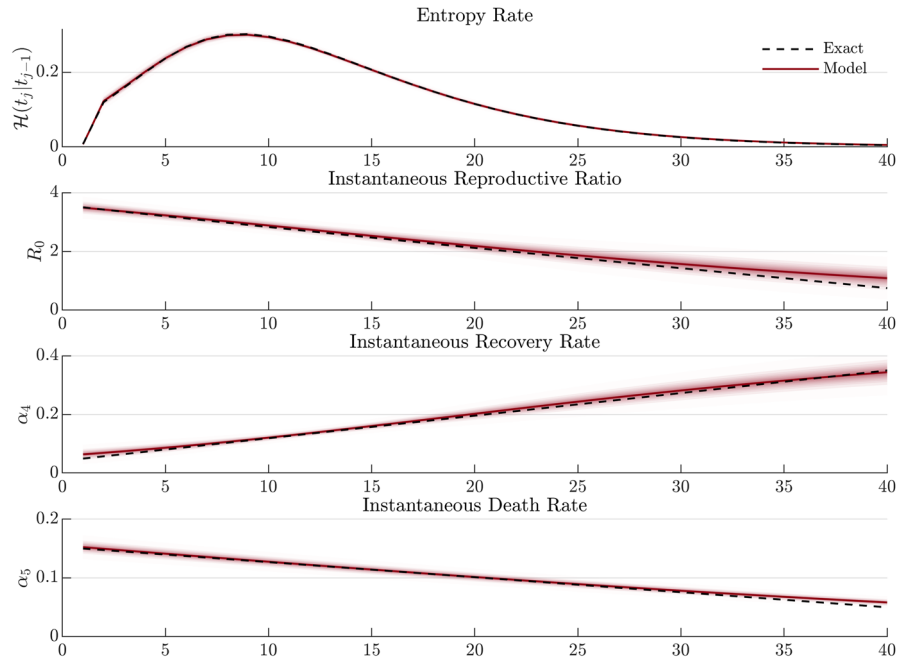
Fig. 17 The transmission properties of the time-dependent SEIR model calibrated on the expectation of the analytical model. Compared with Fig. 16, it is seen that the bias is noticeably decreased

ductive ratio decreases from 3.50 to 0.75 as t grows from 0 to 40 (days). The population size is fixed to $N = 10^3$, and the initial condition is set to $\boldsymbol{P}(t_0) = [1 - 101/N, 100/N, 1/N, 0, 0]$. The specification of a relatively large initial condition for the unobservable exposed state is to "challenge" the proposed approach and investigate whether the exposed population can be accurately identified even without observing it. We calibrate a time-dependent modified SEIR model, and the results are illustrated in Figs. 15 and 16.

Finally, to verify that the bias in Fig. 16 is mainly caused by the inherent stochastic variation of the model, we calibrate the model on the expectation of the analytical model, and the results of transmission properties are illustrated in Fig. 17.

References

1. Atchison, C.J., Bowman, L., Vrinten, C., et al.: Perceptions and behavioural responses of the general public during the covid-19 pandemic: a cross-sectional survey of uk adults. medRxiv (2020). https://doi.org/10.1101/2020.04.01.20050039

2. Huang, C., Wang, Y., Li, X., et al.: Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet **395**(10223), 497–506 (2020). https://doi.org/10.1016/S0140-6736(20)30183-5

3. Qiu, H., Wu, J., Hong, L., et al.: Clinical and epidemiological features of 36 children with coronavirus disease 2019 (covid-19) in Zhejiang, China: an observational cohort study. Lancet Infect. Dis. (2020). https://doi.org/10.1016/S1473-3099(20)30198-5

4. Verity, R., Okell, L.C., Dorigatti, I., et al.: Estimates of the severity of coronavirus disease 2019: a model-based analysis. Lancet Infect. Dis. (2020). https://doi.org/10.1016/S1473-3099(20)30243-7

5. Wang, D., Hu, B., Hu, C., et al.: Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. JAMA **323**(11), 1061–1069 (2020). https://doi.org/10.1001/jama.2020.1585

6. Yu, N., Li, W., Kang, Q., et al.: Clinical features and obstetric and neonatal outcomes of pregnant patients with covid-19 in wuhan, china: a retrospective, single-centre, descriptive study. The Lancet Infectious Diseases (2020). https://doi.org/10.1016/S1473-3099(20)30176-6

7. Ghinai, I., McPherson, T.D., Hunter, J.C., et al.: First known person-to-person transmission of severe acute respiratory syndrome coronavirus 2 (sars-cov-2) in the USA. Lancet **395**(10230), 1137–1144 (2020). https://doi.org/10.1016/S0140-6736(20)30607-3

8. Leung, K., Wu, J.T., Liu, D., et al.: First-wave covid-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment. Lancet **395**(10233), 1382–1393 (2020). https://doi.org/10.1016/S0140-6736(20)30746-7

9. Pullano, G., Pinotti, F., Valdano, E., et al.: Novel coronavirus (2019-ncov) early-stage importation risk to Europe, January 2020. Eurosurveillance **25**(4), 2000057 (2020). https://doi.org/10.2807/1560-7917.ES.2020.25.4.2000057

10. Kandel, N., Chungong, S., Omaar, A., et al.: Health security capacities in the context of covid-19 outbreak: an analysis of international health regulations annual report data from 182 countries. Lancet **395**(10229), 1047–1053 (2020). https://doi.org/10.1016/S0140-6736(20)30553-5

11. W. H. Organization: Coronavirus disease (covid-2019) situation reports (2020). https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports

12. Prem, K., Liu, Y., Russell, T.W., et al.: The effect of control strategies to reduce social mixing on outcomes of the covid-19 epidemic in Wuhan, China: a modelling study. Lancet Public Health (2020). https://doi.org/10.1016/S2468-2667(20)30073-6

13. I. C. London: Short-term forecasts of covid-19 deaths in multiple countries. https://mrc-ide.github.io/covid19-short-term-forecasts/index.html#introduction (2020)

14. Wu, K., Darcet, D., Wang, Q., Sornette, D.: Generalized logistic growth modeling of the covid-19 outbreak in 29 provinces in China and in the rest of the world. https://medrxiv.org/cgi/content/short/2020.03.11.20034363v1 (2020)

15. Favero, C., Ichino, A., Rustichini, A.: Restarting the Economy while Saving lives Under Covid-19. CEPR Discussion Paper No. DP14664 **48** (2020). https://ssrn.com/abstract=3594296. Accessed 10 April 2020

16. Gatto, M., Bertuzzo, E., Mari, L., et al.: Spread and dynamics of the covid-19 epidemic in Italy: effects of emergency containment measures. In: Proceedings of the National Academy of Sciences (2020). https://doi.org/10.1073/pnas.2004978117. https://www.pnas.org/content/early/2020/04/22/2004978117

17. Read, J.M., Bridgen, J.R., Cummings, D.A., et al.: Novel coronavirus 2019-ncov: early estimation of epidemiological parameters and epidemic predictions. medRxiv (2020). https://doi.org/10.1101/2020.01.23.20018549

18. Wang, H., Wang, Z., Dong, Y., et al.: Phase-adjusted estimation of the number of coronavirus disease 2019 cases in Wuhan, China. Cell Discov. **6**(1), 10 (2020). https://doi.org/10.1038/s41421-020-0148-0

19. Wu, J.T., Leung, K., Leung, G.M.: Nowcasting and forecasting the potential domestic and international spread of the 2019-ncov outbreak originating in Wuhan, China: a modelling study. Lancet **395**(10225), 689–697 (2020). https://doi.org/10.1016/S0140-6736(20)30260-9

20. Yang, Z., Zeng, Z., Wang, K., et al.: Modified seir and ai prediction of the epidemics trend of covid-19 in China under public health interventions. J. Thorac. Dis. **12**(3), 165 (2020)

21. Keeling, M.J., Rohani, P.: Modeling Infectious Diseases in Humans and Animals. Princeton University Press, Princeton (2008)

22. Gillespie, D.T.: A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J. Comput. Phys. **22**(4), 403–434 (1976). https://doi.org/10.1016/0021-9991(76)90041-3

23. Gillespie, D.T., Petzold, L.R.: Improved leap-size selection for accelerated stochastic simulation. J. Chem. Phys. **119**(16), 8229–8234 (2003). https://doi.org/10.1063/1.1613254

24. Chen, Y.C., Lu, P.E., Chang, C.S.: A time-dependent sir model for covid-19. https://arxiv.org/abs/2003.00122. Accessed 28 May 2020 (2020).

25. Chowell, G., Hengartner, N., Castillo-Chavez, C., Fenimore, P., Hyman, J.: The basic reproductive number of ebola and the effects of public health measures: the cases of congo and uganda. J. Theor. Biol. **229**(1), 119–126 (2004). https://doi.org/10.1016/j.jtbi.2004.03.006

26. Lekone, P.E., Finkenstädt, B.F.: Statistical inference in a stochastic epidemic seir model with control intervention: Ebola as a case study. Biometrics **62**(4), 1170–1177 (2006). https://doi.org/10.1111/j.1541-0420.2006.00609.x

27. Pollicott, M., Wang, H., Weiss, H.: Extracting the time-dependent transmission rate from infection data via solution of an inverse ODE problem. J. Biol. Dyn. **6**(2), 509–523 (2012)

28. Teles, P.: A time-dependent seir model to analyse the evolution of the sars-covid-2 epidemic outbreak in Portugal and evolution (2020). arXiv: Populations. Accessed 9 Apr 2020

29. Liddle, A.R.: Information criteria for astrophysical model selection. Mon. Not. R. Astron. Soc.: Lett. **377**(1), L74–L78 (2007). https://doi.org/10.1111/j.1745-3933.2007.00306.x

30. Kolokoltsov, V.N.: Nonlinear Markov Processes and Kinetic Equations. Cambridge Tracts in Mathematics, Cambridge University Press, Cambridge (2010). https://doi.org/10.1017/CBO9780511760303

31. Cybenko, G.: Approximation by superpositions of a sigmoidal function. Math. Control Signal. Syst. **2**(4), 303–314 (1989). https://doi.org/10.1007/BF02551274

32. Artalejo, J., Lopez-Herrero, M.: The sis and sir stochastic epidemic models: a maximum entropy approach. Theor. Popul. Biol. **80**(4), 256–264 (2011). https://doi.org/10.1016/j.tpb.2011.09.005

33. Scarpino, S.V., Petri, G.: On the predictability of infectious disease outbreaks. Nat. Commun. **10**(1), 898 (2019). https://doi.org/10.1038/s41467-019-08616-0

34. Waezizadeh, T., Fatehi, F.: Entropy for DTMC SIS epidemic model. J. Mahani Math. Res. Cent. **5**, 59–67 (2017)

35. Roddam, A.W.: Mathematical epidemiology of infectious diseases: model building, analysis and interpretation: Diekmann, O., Heesterbeek, J.A.P., 2000. Wiley, Chichester, pp. 303, 39.95. ISBN 0-471-49241-8. Int. J. Epidemiol. **30**(1), 186–186 (2001). https://doi.org/10.1093/ije/30.1.186

36. Lauer, S.A., Grantz, K.H., Bi, Q., et al.: The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: estimation and application. Annals of internal medicine pp. M20–0504 (2020). https://doi.org/10.7326/M20-0504. https://pubmed.ncbi.nlm.nih.gov/32150748. 32150748[pmid]

37. Haario, H., Saksman, E., Tamminen, J.: An adaptive Metropolis algorithm. Bernoulli pp. 223–242 (2001)

38. Wang, Z., Broccardo, M., Song, J.: Hamiltonian Monte Carlo methods for subset simulation in reliability analysis. Struct. Saf. **76**, 51–67 (2019). https://doi.org/10.1016/j.strusafe.2018.05.005

39. Cérou, F., Moral, P.D., Furon, T., Guyader, A.: Sequential monte carlo for rare event estimation. Stat. Comput. **22**(3), 795–808 (2012)

40. Moral, P.D., Doucet, A., Jasra, A.: Sequential Monte Carlo samplers. J. R. Stat. Soc. Ser. B-Stat. Methodol. **68**(3), 411–436 (2006)

41. Moré, J.J.: The Levenberg–Marquardt algorithm: implementation and theory. In: Watson, G.A. (ed.) Numerical Analysis, pp. 105–116. Springer, Berlin, Heidelberg (1978)

42. Yuan, Y.: Recent advances in trust region algorithms. Math. Progr. **151**(1), 249–281 (2015). https://doi.org/10.1007/s10107-015-0893-2

43. CDC, C.: Covid-19 report. http://2019ncov.chinacdc.cn/2019-nCoV/ (2020)

44. CDC, E.: Covid-19 report. https://www.ecdc.europa.eu/en/covid-19-pandemic (2020)

45. Worldometers: World population. https://www.worldometers.info/world-population/ (2020)

46. Oh, M.D., Park, W.B., Park, S.W., Choe, P.G., Bang, J.H., Song, K.H., Kim, E.S., Kim, H.B., Kim, N.J.: Middle east respiratory syndrome: what we learned from the 2015 outbreak in the Republic of Korea. Korean J. Intern. Med. **33**(2), 233 (2018)

47. N. I. of Health: Nih begins study to quantify undetected cases of coronavirus infection. https://www.nih.gov/news-events/news-releases/nih-begins-study-quantify-undetected-cases-coronavirus-infection (2020)