Interpretability of Epidemiological Models: The Curse of Non-Identifiability

Ayush Deva, Siddhant Shingi, Avtansh Tiwari, Nayana Bannur, Sansiddh Jain,
Jerome White, Alpan Raval, Srujana Merugu
Wadhwani Institute for Artificial Intelligence

Abstract

Interpretability of epidemiological models is a key consideration, especially when these models are used in a public health setting. Interpretability is strongly linked to the *identifiability* of the underlying model parameters, i.e., the ability to estimate parameter values with high confidence given observations. In this paper, we define three separate notions of identifiability that explore the different roles played by the model definition, the loss function, the fitting methodology, and the quality and quantity of data. We define an epidemiological compartmental model framework in which we highlight these non-identifiability issues and their mitigation.

1 Introduction

The global COVID-19 pandemic has spurred intense interest in epidemiological compartmental models (Thompson, 2020; Brauer, 2008). The use of epidemiological models is driven by four main criteria: *expressivity* to faithfully capture the disease dynamics; *learnability* of parameters conditioned on the available data; *interpretability* to understand the evolution of the pandemic; and *generalizability* to future scenarios by incorporating additional information.

Compartmental models are popular because they are relatively simple and known to be highly expressive and generalizable. However, their interpretability and learnability depend strongly on the alignment between data observations and model complexity. Different choices of model parameters can often lead to (approximately) the same forecast case counts, leading to what is commonly referred to as non-identifiability (Raue et al., 2009; Jacquez & Perry, 1990). The problem of non-identifiability is only exacerbated with increased model complexity.

This lack of identifiability is detrimental because (a) parameter distributions estimated from the observed data tend to be biased with large variances, thus precluding easy interpretation, and (b) non-identifiable models typically have reduced accuracy on long-time forecasts due to the high parameter variance. This phenomenon is illustrated in Figure 1, where the forecasting errors between a non-identifiable model and its reparametrized version (later shown to be identifiable) are compared.

Non-identifiability in epidemiological models is rooted in the model dynamics, in the fitting loss function and methodology, and in the quality and quantity of data available. Identifiability is typically broadly classified into structural (i.e., purely model-dependent) (Reid, 1977; Massonis et al., 2020) and practical (i.e., data-, loss- and fitting methodology-dependent), with the latter often defined vaguely, and in the context of specific loss functions (Raue et al., 2009; Wieland et al., 2021).

Contributions. This paper delineates various general notions of model identifiability that are contextualized in compartmental epidemiological models, including a novel notion of *statistical identifiability* that depends on the loss function optimized in estimation, and an empirical framework to assess practical identifiability in terms of the highest posterior density intervals. We study these ideas in the specific context of a SEIR-like compartmental model (SEIARD) that was deployed in a major densely populated city in India for case forecasting to inform capacity and policy decisions.
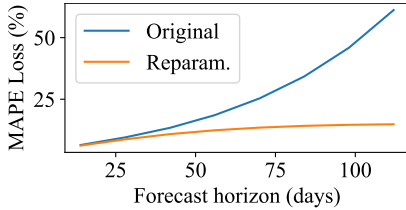
Figure 1: Forecast error. Section 3 defines "Original" and "Reparam."

	Structural	Statistical	Practical
Model form	✓	✓	✓
Loss function		✓	✓
Observation interval		✓	✓
Noisy data			✓
Fitting method			✓

Table 1: Notions of identifiability.

2 NOTIONS OF IDENTIFIABILITY

Consider a dynamical system M characterized by $\dot{\boldsymbol{x}} = f(\boldsymbol{x}, \boldsymbol{\theta})$, where $\boldsymbol{x}$ is the system state, $\dot{\boldsymbol{x}}$ the time derivative, and $\boldsymbol{\theta}$ the parameters. Let $\boldsymbol{y}(t) = g(\boldsymbol{x}(t), \boldsymbol{\theta})$ be the observation function that maps state $\boldsymbol{x}$ to observations at time t. Combining these, we may express $\boldsymbol{y}(t) = h(\boldsymbol{x}_0, \boldsymbol{\theta}, t)$, where $\boldsymbol{x}_0$ is the initial state. Epidemiological compartmental models are dynamical systems in which states correspond to compartmental population counts, with a subset or aggregates of these counts being observed. We consider three notions of identifiability for such a system (Table 1).

Structural Identifiability. We say that a parameter θ_i in a model M is structurally identifiable if for any $\boldsymbol{\theta}$ in the domain, $h(\hat{x}_0, \hat{\boldsymbol{\theta}}, t) = h(x_0, \boldsymbol{\theta}, t) \ \forall t \implies \hat{\theta}_i = \theta_i$, i.e., distinct parameter choices result in distinct observation series. Note that in partially observable systems, this notion of identifiability also applies to components of the initial state x_0 that are not observed. In linear time-invariant systems (LTI), where $\dot{\boldsymbol{x}} = B(\boldsymbol{\theta})\boldsymbol{x}$ and $\boldsymbol{y} = C(\boldsymbol{\theta})\boldsymbol{x}$, the resulting solution takes the form $\boldsymbol{x} = e^{B(\theta)t}\boldsymbol{x}_0$. Structural identifiability for this case is characterized in terms of the properties of $B(\boldsymbol{\theta})$ and $C(\boldsymbol{\theta})$ (Kalman, 1959; Sontag, 2013). Recent works (Martinelli, 2020; Villaverde, 2019; Massonis et al., 2020) extend these results to nonlinear systems and SEIR model variants.

Statistical Identifiability. We propose a new notion of identifiability that depends both on the parametric form of the model and the statistical estimation process. Let $y_M(t) = h(\boldsymbol{x}_0, \boldsymbol{\theta}, t)$ denote the model output and $y_D(t)$ the observations over a finite time horizon $t \in [t_b, t_e]$. Parameter estimation given the observations is carried out by optimizing a loss function $\mathscr{L}(\cdot)$, i.e., $\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \mathscr{L}_D(\boldsymbol{\theta}) = \arg\min_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \sum_{t=t_b}^{t_e} \mathscr{L}(y_D(t), y_M(t))$. We consider a parameterwise loss function, called the *profile-likelihood* (PL) of θ_i (Raue et al., 2009), defined as $\mathscr{L}_D^i(\theta_i) := \min_{\boldsymbol{\theta}_{-i}} \mathscr{L}_D(\theta_i, \boldsymbol{\theta}_{-i})$, where $\boldsymbol{\theta}_{-i} = \boldsymbol{\theta} \setminus \theta_i$ denotes the complementary set of parameters. We say a parameter θ_i is *statistically identifiable* within a domain Θ_i if $\mathscr{L}_D^i(\theta_i)$ is strictly convex with respect to $\theta_i \in \Theta_i$. Though this ensures a unique global optimum, it does not necessarily imply joint convexity of $\mathscr{L}_D$ with respect to $\boldsymbol{\theta}$.

Clearly, statistical identifiability implies structural identifiability, but the converse is not always true. For example, a loss function with a flat extended minimum will lead to lack of statistical identifiability even though the model itself is structurally identifiable. However, structural identifiability does imply statistical identifiability for a LTI system with a strictly convex loss function, when the number of observations at distinct times in D is at least equal to the rank of the observability matrix (Villaverde, 2019; Kalman, 1959). The proof follows from the strict convexity of the loss function, the convexity of the solution $e^{B(\theta)t}\boldsymbol{x}_0$ itself, and the fact that the composition of a convex function with a convex non-decreasing function remains convex (Boyd et al., 2004).

Practical Identifiability Intervals. Models could be both structurally and statistically identifiable, but parameter fitting could nevertheless present practical challenges, related to application-specific tolerance to error on the fitted parameters (Roosa & Chowell, 2019). This situation is captured through the notion of *practical identifiability*. Given a model M, data D, loss function $\mathscr{L}_D$, and fitting method H, let $p_D(\theta_i)$ be the resulting posterior distribution of θ_i. Following Raue et al. (2009), we characterize practical identifiability in terms of the α-identifiability intervals of the parameters, in two different ways.

The first approach defines this interval in terms of level sets of the loss function, i.e., $J_{PL}^i(\alpha, \mathscr{L}_D, H) = \{\theta_i | \mathscr{L}_D^i(\theta_i) \leq \mathscr{L}_D^\alpha\}$, where $\mathscr{L}_D^\alpha$ is the α-quantile of $\mathscr{L}_D(\cdot)$, defined as the level set with probability mass α as given by the posterior distribution. The special case of the
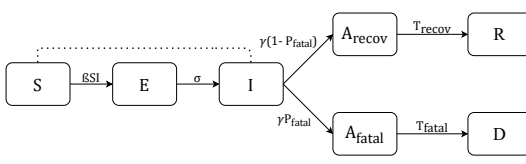
Figure 2: SEIARD Model.

$$\begin{bmatrix} 0 & 0 & -\beta & 0 & 0 & 0 & 0 \\ 0 & -1/T_{\text{inc}} & \beta & 0 & 0 & 0 & 0 \\ 0 & 1/T_{\text{inc}} & 1/T_{\text{inf}} & 0 & 0 & 0 & 0 \\ 0 & 0 & (1 - P_{\text{fatal}})/T_{\text{inf}} & -1/T_{\text{recov}} & 0 & 0 & 0 \\ 0 & 0 & P_{\text{fatal}}/T_{\text{inf}} & 0 & -1/T_{\text{fatal}} & 0 & 0 \\ 0 & 0 & 0 & 1/T_{\text{recov}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/T_{\text{fatal}} & 0 & 0 \end{bmatrix}$$

Figure 3: SEIARD Dynamics Matrix (LTI when $S \simeq N$)

squared loss where $\mathscr{L}_D^\alpha$ is the sum of $\mathscr{L}_D(\theta^*)$ and the α-quantile of χ^2 distribution with a single degree of freedom is discussed in (Wieland et al., 2021; Raue et al., 2013).

The second, Bayesian approach defines the α-identifiability interval $J_{PP}^i(\alpha, \mathscr{L}_D, H)$ as a *highest posterior density interval* (HPDI) (Wright, 1986) with probability mass α. When $\mathscr{L}_D^i(\cdot)$ is strictly convex in θ_i and the marginal posterior distribution $p_D^i(\theta_i)$ is monotonically decreasing with respect to $\mathscr{L}_D^i(\cdot)$, the two definitions can be shown to be equivalent based on the convexity properties and one-one association of the corresponding level sets (Boyd et al., 2004). Empirical analysis in Raue et al. (2013) suggests a similar relationship. See Section A.4 for a detailed analysis.

3   NON-IDENTIFIABILITY IN EPIDEMIOLOGICAL MODELS

SEIARD Model. Figure 2 shows the SEIARD model, which is an extension of the well-known SEIR model (Hethcote, 2000). The **S** (Susceptible), **E** (Exposed), and **I** (Infectious) compartments and their associated parameters (transmission rate β, incubation period $T_{\text{inc}} = \sigma^{-1}$, and infectious period $T_{\text{inf}} = \gamma^{-1}$) are defined as in the SEIR model. The post-infectious stage consists of parallel paths through $\mathbf{A}_{\text{fatal}}$ or $\mathbf{A}_{\text{recov}}$, depending on the eventual outcome: fatality (**D**) with probability P_{fatal} or recovery (**R**), with T_{fatal} and T_{recov} being the respective time durations. Observed case counts are mapped to compartment populations as follows: recovered cases $(R) = |\mathbf{R}|$, deceased cases $(D) = |\mathbf{D}|$, active cases $(A) = |\mathbf{A}_{\text{recov}}| + |\mathbf{A}_{\text{fatal}}|$. Defining state $\boldsymbol{x} = (S, E, I, A_{\text{recov}}, A_{\text{fatal}}, R, D)$ and observations $\boldsymbol{y} = (A_{\text{recov}} + A_{\text{fatal}}, R, D)$, the model dynamics (Figure 3) reduces to that of an LTI (Kalman, 1959) in the early stages of the epidemic, because one can then set $S \simeq N$ to first approximation, thus yielding $\dot{S} = -\beta SI/N \simeq -\beta I$.

Parameter Estimation. In our experiments, synthetic data was simulated from the SEIARD model using realistic parameters (Section A.3). We treat the initial state values of **E** (E_0) and **I** (I_0) compartments as unobserved parameters. We estimate model parameters (including E_0 and I_0) by optimizing the mean absolute percentage error (MAPE) between the true and predicted values for each of the three observed time series in $\boldsymbol{y}$ as well as that of the total count (sum of these three). Fitting is performed using two different approaches: Sequential Model Based Optimization (SMBO) based on Tree-structured Parzen Estimators as implemented in the HyperOpt library (Bergstra et al., 2013), and an MCMC-based sampling approach (see Section A.2).

Identifiability in SEIARD Model. Structural identifiability analysis similar to that in Massonis et al. (2020) indicates that the SEIARD model is structurally non-identifiable. This may be attributed to the unobserved compartments (E, I) as well as the split between A_{recov} and A_{fatal}. A larger fatality rate with a larger delay T_{fatal} would be indistinguishable from a smaller fatality with a smaller delay just by observing the active, deceased and recovery counts. This structural non-identifiability also manifests as *statistical non-identifiability*, as can be seen in the shape of the profile likelihood curves corresponding to the non-reparametrized (Original) SEIARD model in Figure 4(a) and 4(b). Similarly, the confidence intervals of the parameters of this model, as shown in Figure 4(a), 4(b), and 4(d), are broad, indicating considerable *practical non-identifiability*.

Improving Identifiability in SEIARD Model. Non-identifiability in dynamical systems can often be attributed to correlated parameters. Figure 4(e) shows the correlation between parameters constructed using samples from the joint posterior distribution. A common way (Joubert et al., 2020) to reduce non-identifiability is thus through reparameterization of models by constraining a parameter subset. The choice of parameters to constrain may be informed by secondary information sources. For epidemiological models, parameters such as the incubation period T_{inc}, infectious period T_{inf},
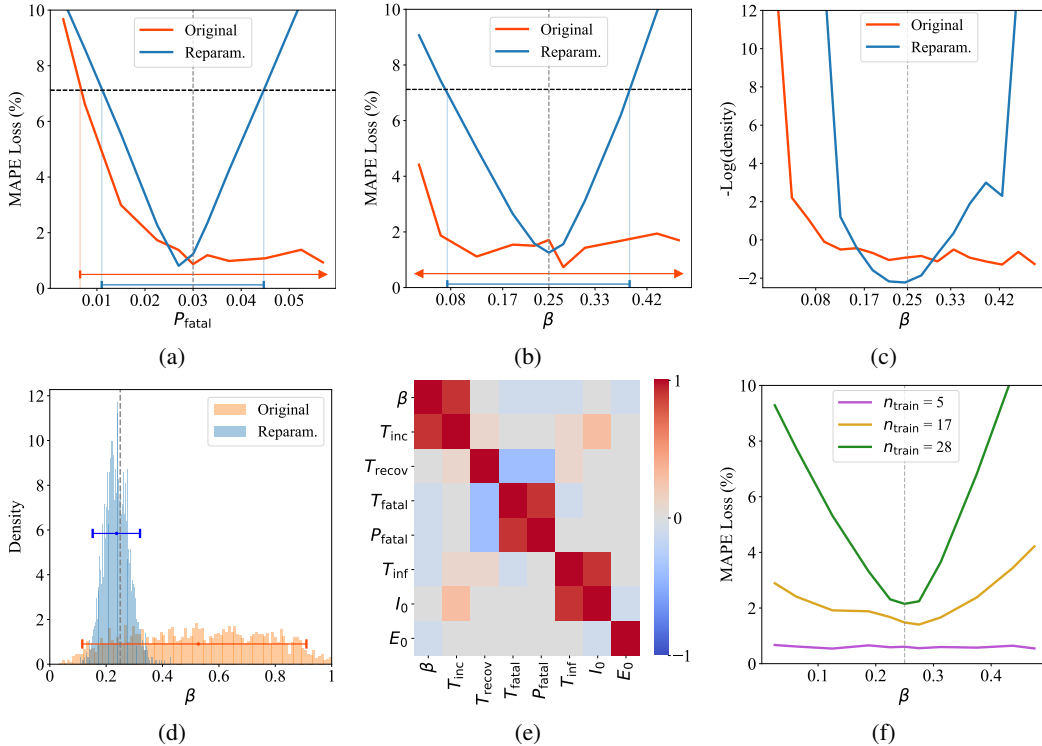
Figure 4: Panels (a) and (b) show PL curves for P_{fatal} and β respectively. The horizontal black line depicts the 95% quantile loss as defined in Section 2, while the vertical gray lines are the true values given in Section A.3. Error bars indicate confidence intervals based on where the curves intersect the black line; an arrow head indicates that the interval extends beyond the plotted range. Panel (c) shows the negative logarithm of the posterior density of β; the corresponding posterior density is shown in panel (d). Error bars in (d) refer to the 95%-HPDI (Section 2) of β. Panel (e) is the correlation matrix of parameter samples obtained via MCMC sampling for the Original model. Panel (f) shows PL curves of β for the Reparam. model for different training durations.

and the time to death T_{fatal} can be estimated from medical literature or line lists of representative case data. A model where these parameters are fixed to their true values (see Section A.3) is referred to as the reparameterized (Reparam.) SEIARD model, which can be shown to be structurally identifiable (Massonis et al., 2020). The PL curves for the reparameterized model in Figure 4(a) and 4(b), and the posterior density plot in Figure 4(d) have relatively tighter confidence intervals, indicating an improvement in practical identifiability. Additionally, Figure 4(c) empirically indicates that the PL curves and the negative logarithm of the posterior density give similar information.

Data Dependence of Identifiability Figure 4(f) shows the PL curves for the reparameterized model with various training durations, indicating increased practical identifiability with more observations. A larger set of observations may therefore serve to further fine-tune parameters in such a setting.

4 CONCLUSIONS AND FUTURE WORK

We presented a novel framework to analyze identifiability of model parameters in epidemiological models, using the various lenses of model structure, loss functions and fitting methodology, and data. While the notions of structural and statistical identifiability are useful for detecting and resolving non-identifiability in an analytic fashion via reparameterization, practical identifiability intervals function as an effective tool for fine-tuning the parameter estimation process. We present these ideas and empirical results in the specific context of the SEIARD compartmental model. In the future, we plan to explore connections between different types of identifiability and empirically analyze the identifiability of multiple SEIR variants on real and synthetic data.

REFERENCES

J. Bergstra et al. Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In *Intl. conf. on machine learning*, pp. 115–123, 2013.

Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

Fred Brauer. Compartmental models in epidemiology. In *Mathematical epidemiology*, pp. 19–79. Springer, 2008.

Herbert W Hethcote. The mathematics of infectious diseases. *SIAM review*, 42(4):599–653, 2000. doi: 10.1137/S0036144500371907.

John A Jacquez and Timothy Perry. Parameter estimation: local identifiability of parameters. *American Journal of Physiology-Endocrinology and Metabolism*, 258(4):E727–E736, 1990.

D Joubert, JD Stigter, and J Molenaar. An efficient procedure to assist in the re-parametrization of structurally unidentifiable models. *Mathematical biosciences*, 323:108328, 2020.

Rudolf Kalman. On the general theory of control systems. *IRE Transactions on Automatic Control*, 4(3): 110–110, 1959.

Agostino Martinelli. Rank conditions for observability and controllability for time-varying nonlinear systems. *arXiv preprint arXiv:2003.09721*, 2020.

Gemma Massonis, Julio R Banga, and Alejandro F Villaverde. Structural identifiability and observability of compartmental models of the COVID-19 pandemic. *Annual reviews in control*, 2020.

Andreas Raue, Clemens Kreutz, Thomas Maiwald, Julie Bachmann, Marcel Schilling, Ursula Klingmüller, and Jens Timmer. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(15):1923–1929, 2009.

Andreas Raue, Clemens Kreutz, Fabian Joachim Theis, and Jens Timmer. Joining forces of Bayesian and frequentist methodology: a study for inference in the presence of non-identifiability. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(1984):20110544, 2013.

J Reid. Structural identifiability in linear time-invariant systems. *IEEE Transactions on Automatic Control*, 22 (2):242–246, 1977.

Kimberlyn Roosa and Gerardo Chowell. Assessing parameter identifiability in compartmental dynamic models using a computational approach: application to infectious disease transmission models. *Theoretical Biology and Medical Modelling*, 16(1):1–15, 2019.

Eduardo D Sontag. *Mathematical control theory: deterministic finite dimensional systems*, volume 6. Springer Science & Business Media, 2013.

Robin N Thompson. Epidemiological models are important tools for guiding COVID-19 interventions. *BMC medicine*, 18(1):1–4, 2020.

Alejandro F Villaverde. Observability and structural identifiability of nonlinear biological systems. *Complexity*, 2019, 2019.

Franz-Georg Wieland, Adrian L Hauber, Marcus Rosenblatt, Christian Tönsing, and Jens Timmer. On structural and practical identifiability. *arXiv preprint arXiv:2102.05100*, 2021.

DE Wright. A note on the construction of highest posterior density intervals. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 35(1):49–53, 1986.

A APPENDIX

A.1 SEIARD MODEL DYNAMICS

The dynamical equations governing the transitions in SEIARD model (Figure 2) are

$$\frac{dS}{dt} = -\beta \frac{IS}{N}; \ \frac{dE}{dt} = \beta \frac{IS}{N} - \sigma E; \ \frac{dI}{dt} = \sigma E - \gamma I, \tag{1}$$

$$\frac{dA_{\text{recov}}}{dt} = (1 - P_{\text{fatal}}) \cdot \gamma I - \frac{A_{\text{recov}}}{T_{\text{recov}}}, \tag{2}$$

$$\frac{dA_{\text{fatal}}}{dt} = P_{\text{fatal}} \cdot \gamma I - \frac{A_{\text{fatal}}}{T_{\text{fatal}}}, \tag{3}$$

$$\frac{dR}{dt} = \frac{A_{\text{recov}}}{T_{\text{recov}}}; \ \frac{dD}{dt} = \frac{A_{\text{fatal}}}{T_{\text{fatal}}}, \tag{4}$$

where N is the city population, β, σ, and γ are the standard epidemiological parameters for a SEIR model, P_{fatal} is the transition probability to the mortality branch, and T_{recov} and T_{fatal} are timescales that govern transitions out of the $\mathbf{A}_{\text{recov}}$ and $\mathbf{A}_{\text{fatal}}$ compartments. Variables S, E, I, A_{recov}, A_{fatal}, R, D denote the populations of the similarly named compartments.

In the early stages of the epidemic, when $S \simeq N$, the model dynamics reduces to that of an LTI, with $\dot{\boldsymbol{x}} = B(\boldsymbol{\theta})\boldsymbol{x}$ and $\boldsymbol{y} = C(\boldsymbol{\theta})\boldsymbol{x}$. For our model, we have $B(\boldsymbol{\theta})$ and $C(\boldsymbol{\theta})$ as shown below:

$$B(\boldsymbol{\theta}) = \begin{bmatrix} 0 & 0 & -\beta & 0 & 0 & 0 & 0 \\ 0 & -1/T_{\text{inc}} & \beta & 0 & 0 & 0 & 0 \\ 0 & 1/T_{\text{inc}} & 1/T_{\text{inf}} & 0 & 0 & 0 & 0 \\ 0 & 0 & (1 - P_{\text{fatal}})/T_{\text{inf}} & -1/T_{\text{recov}} & 0 & 0 & 0 \\ 0 & 0 & P_{\text{fatal}}/T_{\text{inf}} & 0 & -1/T_{\text{fatal}} & 0 & 0 \\ 0 & 0 & 0 & 1/T_{\text{recov}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/T_{\text{fatal}} & 0 & 0 \end{bmatrix}, C(\boldsymbol{\theta}) = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

A.2 MCMC IMPLEMENTATION DETAILS

Let $X[t] = [X_h[t]]_{h \in \mathcal{H}}$ be a multivariate time series with $X_h[t]$ denoting the hth compartment time-series, and $\mathcal{H}$ be the set of indices of components. Let the fitting period be given by $[t_i, t_j]$. The key components of our MCMC-within-Gibbs sampling are as follows.

Likelihood function. We assume a likelihood function of the form

$$P(X \mid \boldsymbol{\theta}, s) = \prod_{h \in \mathcal{H}} \prod_{t=t_i+1}^{t_j} \mathcal{N}(\hat{z}_{h,\boldsymbol{\theta}}[t] \mid z_h[t], s), \tag{5}$$

where $\mathcal{N}(z \mid \mu, \sigma^2)$ denotes the Normal distribution pdf with mean μ and variance σ^2 following an appropriate conjugate prior. Further, $z_h[t] = \log(X_h[t]) - \log(X_h[t-1])$, and $\hat{z}_{h,\boldsymbol{\theta}}[t]$ is the forecast equivalent of $z_h[t]$ and s is the variance of the normal likelihood function, which is discussed later in this section.

Proposal distribution. At iteration k, we generate the samples from the proposal distribution for accept-reject step as

$$\boldsymbol{\theta} \sim Q(\boldsymbol{\theta}_{k-1}, \Sigma_{\text{prop}}, \boldsymbol{\theta}_{\min}, \boldsymbol{\theta}_{\max}) \tag{6}$$

where $\boldsymbol{\theta}_{k-1}$ is the parameter vector chosen at $k-1$, and $Q(\cdot)$ is the pdf of a multivariate truncated Gaussian with the parameter range $[\boldsymbol{\theta}_{\min}, \boldsymbol{\theta}_{\max}]$ and covariance matrix Σ_{prop} as listed in Table 2 (MCMC Proposal Variance).

We further assume that s has the conjugate prior $s \sim \text{InvGamma}(u, v)$ where $u = 40$ and $v = 2/700$ are hyperparameters. Thus, it is straightforward to show, by multiplying the Normal likelihood with the prior, that if the MCMC chain has sampled parameters $\boldsymbol{\theta}_k$, the sample $s_k \sim P(s_k \mid \boldsymbol{\theta}_k, X)$ is also drawn from an Inverse Gamma distribution with parameters

$$u_k = u + 2(t_j - t_i - 1), \ \ v_k = v + \sum_{h \in \mathcal{H}} \sum_{t=t_i+1}^{t_j} \frac{(z_h[t] - \hat{z}_{h,\boldsymbol{\theta}_k}[t])^2}{2}.$$

A.3 Hyperparameters

Table 2 contains the parameter ranges and values used for all experiments in Section 3. The first row enumerates the true value of parameters used to simulate the synthetic dataset while the second lists the search space for each parameter while fitting. The proposal variance used in Equation 6 is listed in the bottom row of Table 2. Table 3 contains all the hyperparameters used to generate the synthetic dataset. R_0, A_0, D_0 denote the initial values of recovered, active and deceased cases and N denotes total population.

	Parameter							
	β	T_{inc}	T_{inf}	T_{recov}	T_{fatal}	P_{fatal}	E_0	I_0
True Value	0.25	5.10	6.60	14.00	10.00	0.03	1.00	1.00
Search Space	[0, 1]	[1, 100]	[1, 100]	[1, 100]	[1, 100]	[0, 1]	[0, 5]	[0, 5]
MCMC Proposal Variance	0.10	4.00	4.00	4.00	4.00	0.01	0.50	0.50

Table 2: Parameters used for experiments in Section 3

	Hyper Parameter				
	R_0	A_0	D_0	N	Horizon
Value	0	5	0	10^7	400 days

Table 3: Hyperparameters used to generate the synthetic dataset.

A.4 Discussion on Practical Identifiability Intervals

We briefly outline the connection between the Bayesian and the loss function-based notions of practical identifiability intervals in Section 2 for clarity.

1. For a strictly convex Lipschitz continuous real-valued function $\mathscr{L}_D : \boldsymbol{\Theta} \mapsto \mathbb{R}$, with a global minimum c^*, its level sets $\{\boldsymbol{\theta} \in \boldsymbol{\Theta} | \mathscr{L}_D(\boldsymbol{\theta}) \leq c\}$ for $c \in [c^*, \infty)$ are (possibly multi-dimensional) convex regions nested within each other (Boyd et al., 2004).

2. Given a probability distribution $q(\boldsymbol{\theta})$ that is non-zero on $\boldsymbol{\Theta}$, every level set of $\mathscr{L}_D(\boldsymbol{\theta})$ has a well-defined probability mass resulting in a bijection $Q_{\mathscr{L}} : [0, 1] \mapsto [c^*, \infty)$. In other words, for any $\alpha \in [0, 1]$, there is a unique level and a unique level set of $\mathscr{L}_D$ with probability mass equal to α.

3. For any function $h : \boldsymbol{\Theta} \mapsto \mathbb{R}$ that can be expressed as $h(\boldsymbol{\theta}) = g(\mathscr{L}_D(\boldsymbol{\theta}))$ for a strictly monotonically increasing function $g(\cdot)$, the level set $\{\boldsymbol{\theta} \in \boldsymbol{\Theta} | \mathscr{L}_D(\boldsymbol{\theta}) \leq c\}$ is the same as $\{\boldsymbol{\theta} \in \boldsymbol{\Theta} | (\boldsymbol{\theta}) \leq g(c)\}$ for $c \in [c^*, \infty)$. Due to the mapping between the level sets, the uniqueness of regions with probability mass equaling α holds true even for $h(\cdot)$.

4. When $\mathscr{L}_D(\boldsymbol{\theta})$ is the sum of the negative log-likelihood and appropriate regularization terms corresponding to the prior on $\boldsymbol{\theta}$, then the posterior distribution $p_D(\boldsymbol{\theta})$ is strictly monotonically decreasing with respect to $\mathscr{L}_D(\boldsymbol{\theta})$.

5. With the above definition of $\mathscr{L}_D(\boldsymbol{\theta})$, since $-p_D(\boldsymbol{\theta})$ turns out to be a strictly monotonically increasing function of $\mathscr{L}_D(\boldsymbol{\theta})$, the level sets map to each other. Hence, for a given $\alpha \in [0, 1]$, it is the same unique convex region that corresponds to a level set of $\mathscr{L}_D$ and the α-HPDI (highest posterior density interval) of $p_D(\boldsymbol{\theta})$.

6. Since $\mathscr{L}_D(\cdot)$ is strictly convex, the level sets corresponding to parameter-wise loss or profile likelihood $\mathscr{L}_D^i(\cdot)$ are contiguous intervals that are the just the 1-D projections of the level sets of $\mathscr{L}_D(\cdot)$ along the ith dimension for the same level value .

7. On the other hand, the projection of the credible intervals from $p_D(\boldsymbol{\theta})$ on to the ith dimension do not exactly correspond to that of the marginal posterior $p_D^i(\theta_i)$. For α close to 1, these intervals are, however, approximately equal as reflected in the empirical analysis with $\alpha = 0.95$.