Proceedings of Machine Learning Research vol 242:1796–1812, 2024 6th Annual Conference on Learning for Dynamics and Control

Uncertainty Informed Optimal Resource Allocation with Gaussian Process based Bayesian Inference

Samarth Gupta [*] SAMARGPT@AMAZON.COM
Amazon
Saurabh Amin AMINS@MIT.EDU
Massachusetts Institute of Technology

Editors: A. Abate, M. Cannon, K. Margellos, A. Papachristodoulou

Abstract

We focus on the problem of uncertainty informed allocation of medical resources (vaccines) to heterogeneous populations for managing epidemic spread. We tackle two related questions: (1) For a compartmental ordinary differential equation (ODE) model of epidemic spread, how can we estimate and integrate parameter uncertainty into resource allocation decisions? (2) How can we computationally handle both nonlinear ODE constraints and parameter uncertainties for a generic stochastic optimization problem for resource allocation? To the best of our knowledge current literature does not fully resolve these questions. Here, we develop a data-driven approach to represent parameter uncertainty accurately and tractably in a novel stochastic optimization problem formulation. We first generate a tractable scenario set by estimating the distribution on ODE model parameters using Bayesian inference with Gaussian processes. Next, we develop a parallelized solution algorithm that accounts for scenario-dependent nonlinear ODE constraints. Our computational experiments on two different non-linear ODE models (SEIR and SEPIHR) indicate that accounting for uncertainty in key epidemiological parameters can improve the efficacy of time-critical allocation decisions by 4-8%. This improvement can be attributed to data-driven and optimal (strategic) nature of vaccine allocations.

Keywords: non-linear epidemiological models, ODEs, Parameter estimation, Bayesian Inference, Gaussian Processes, Gradient matching, resource allocation.

1. Introduction

In this paper we study the problem of *uncertainty informed* optimal resource allocation to control the spread of an infectious disease such as Covid-19. We develop a *data-driven, scalable* and ODE *model agnostic* approach while *accounting for uncertainty* for the vaccine allocation problem. Our approach is flexible in that it can be easily adapted to other control strategies such as imposing lockdowns Birge et al. (2022); Cianfanelli et al. (2021) and allocation of other resources such as medical personnel, supplies, testing facilities Somers and Manchester (2023); Köhler et al. (2020) & etc.

The vaccine allocation problem has been well studied in the literature. This includes earlier works like Brøgger (1967); Becker (1975) to more recent optimization based methods like Bertsimas et al. (2020); Fu et al. (2021). Researchers have also studied ways to incorporate uncertainty through stochastic epidemiological modelling Clancy and Green (2007); Fu et al. (2021) , stochastic optimization with uncertain parameters Tanner et al. (2008); Yarmand et al. (2014) and robust optimization Han et al. (2015). However, prior works have two major limitations:

i) Most papers such as Tanner et al. (2008); Yarmand et al. (2014); Yin and Buyuktahtakin (2021); Clancy and Green (2007) which claim to account for uncertainty, do not provide a principled *data-driven* method to model (and estimate) uncertainty. They simply model the allocation problem as a stochastic program under the assumption that a scenario-set exists without outlining a principled procedure on how to generate or estimate this scenario-set from *data*. Clearly, this does not effectively solve the problem of uncertainty informed vaccine allocation.

[*] Work done prior to joining Amazon.

ii) The presence of product term between the susceptible (S) and infected (I) population is a key characteristic of most compartmentalized epidemiological ODE models Bertsimas et al. (2020); Hayhoe et al. (2021). Due to this non-linearity, the resource allocation problem with the discretized ODEs results in a non-convex quadratic program. This is difficult to solve even in the nominal case i.e. without accounting for uncertainty, let alone uncertainty informed. To avoid the product term, previous papers Tanner et al. (2008); Yarmand et al. (2014); Yin and Buyuktahtakin (2021); Yin et al. (2023) resort to using simple (linear) epidemiological models so that the discretized ODEs result in a linear program which is easy to solve. Such linear models are limited in their ability to capture the true underlying non-linear dynamics of disease transmission; hence the resulting allocation strategies are not globally optimal.

In this work, we address both of the above limitations by making following novel contributions:

i) We make progress in resolving the issue of incorporating parameter uncertainty in the resource (vaccine) allocation problem in a *data-driven manner*. We do this by making connections with the ODE parameter estimation literature with Bayesian inference using GPs with gradient matching methods. We show that the posterior-distributions can be used to represent uncertainty through a tractable scenario-set.

ii) We provide a novel formulation for the uncertainty informed vaccine allocation problem as a stochastic optimization problem. We develop technical results for the *feasibility* and *decomposability* of this stochastic program.

iii) We develop a *parallelized*, scalable iterative solution algorithm to solve the stochastic program while retaining the original *non-linear*, continuous-time ODE model constraints. Due to this ODE *model agnostic* nature of our approach, we are also able to account for different levels of mobility within different sub-populations and the temporal variations in the onset of the pandemic in each of these sub-populations.

iv) We provide extensive empirical results on two different ODE models i.e. the SEIR and the SEPIHR models. Our results demonstrate that with optimal vaccine allocation, peak infections can be reduced by around 35%. More importantly, a further gain of around **4** to **8**% can be achieved when incorporating uncertainty.

2. Epidemiological Modelling and Pitfalls of Classical Parameter estimation

Mathematical modelling of pandemics (including epidemics) has an extensive literature going back to 1960s Brøgger (1967); Becker (1975). A fairly recent and concise overview can be found at Brauer (2017). More recently, spread of epidemics over networks has been well studied Torres et al. (2017); Nowzari et al. (2016, 2017); Preciado et al. (2013, 2014); Somers and Manchester (2022). Throughout literature, modelling the spread of different diseases using a compartmentalized model through a set of time-dependent ordinary differential equations (ODEs) is common and widely used Brauer (2008); Paré et al. (2020); Newton and Papachristodoulou (2020). Following the recent literature on covid-19 Wang et al. (2021); Li et al. (2022); Acemoglu et al. (2021); Cianfanelli et al. (2021); Cramer et al. (2022) we also adopt the compartmentalized modelling approach.

A popular epidemiological model which we use is the SEIR model, shown in fig. 1. In this model the entire population (of size N) is divided into four states: Susceptible (S), Exposed (E), Infected (I) and Recovered (R). The evolution of each state or the system dynamics is governed by equations in (1).

In (1), $\alpha, \beta,$ and γ are the model parameters and control the rate at which fraction of the population moves from one compartment to another. These model parameters are to be estimated from the

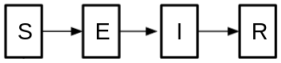


Fig. 1: SEIR model.

$$\left.\begin{aligned} \frac{dS(t)}{dt} &:= \dot{S}(t) = -\frac{\alpha}{N} S(t)I(t), \frac{dI(t)}{dt} := \dot{I}(t) = \beta E(t) - \gamma I(t) \\ \frac{dE(t)}{dt} &:= \dot{E}(t) = \frac{\alpha}{N} S(t)I(t) - \beta E(t), \frac{dR(t)}{dt} := \dot{R}(t) = \gamma I(t) \end{aligned}\right\} \quad (1)$$

available time-series data which we discuss subsequently. Mobility levels can be easily incorporated by adjusting the infection rate α accordingly.

Note that we use SEIR model only as a prototypical model, however, all our subsequent discussion including technical results and solution algorithm holds true for other ODE based models as well. In fact, in addition to the SEIR model, we also provide results on a second model, i.e. the SEPIHR model Rohith (2021) with additional states P (for protective quarantine) and H (for hospitalised quarantined) shown in fig. 2. The functional form of ODEs for this model is provided in supplementary information (SI).

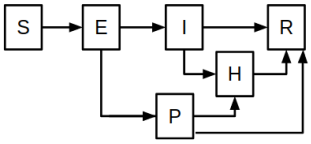


Fig. 2: SEPIHR model

Given the time-series data such as number of daily infections and deaths, the main question arises how to estimate SEIR model parameters i.e. $\alpha, \beta,$ and γ from this data. Therefore, we next discuss the commonly used non-linear least squares approach for ODE parameter estimation and its associated pitfalls, thus providing motivation for adopting Bayesian viewpoint.

2.1. Classical Parameter Estimation: Non-linear Least Squares (NLLS)

Before describing the NLLS approach, we briefly describe the *initial-value problem* (IVP) in the context of ODEs. For a given (or fixed) set of parameter values and initial conditions (denoted $\mathbf{x}_0$), a systems of ODEs can be numerically solved using an off-the-shelf ODE solver such as ODE45 in matlab or ODEINT in python. The solved system (also referred to as simulation) provides the value (or estimates) of different states at the specified time-stamps.

For a given set of parameters, using the estimated state values obtained by solving the IVP and time-series data, discrepancy or the least-squares error can be computed. This can be turned into a optimization problem where we want to find those values of the model parameters for which the least-squares error is minimized. L-BFGS is commonly used to solve such problems Li et al. (2022). Mathematically for SEIR model, the NLLS problem can be written as follows:

$$\min_{\alpha,\beta,\gamma} \sum_{t=1}^{N} \left((y_R^t - R(t))^2 + (y_I^t - I(t))^2 \right)$$
$$\text{s.t. } \{(1)\} \quad \forall\, t \in \{1, \ldots, N\} \text{ and } [S(0), E(0), I(0), R(0)] = \mathbf{x}_0$$

where y_R^t and y_I^t denote count data for infected and removed individuals at time t. The optimal parameters obtained after solving NLLS can then be used to re-solve the ODE system to make predictions for future time as well.

Why account for Uncertainty? NLLS discussed above can provide sufficiently reliable *point estimates* of the parameter values and predictions of new cases into the future provided the *time-series data is accurate*. Using these point estimates resource (vaccine) allocation problem is to be solved subsequently. The efficacy of the overall allocation solution in real-world is highly dependent on the accuracy of the predicted point estimates which are only as good as the data from which these estimates are generated. For Covid-19, the data reported by various private organizations and government agencies can be severely biased, under-reported Kobilov et al. (2021) and erroneous due to numerous reasons Angelopoulos et al. (2020). Reliance on these point estimates can result in severe region-wide inefficiencies. To address these issues and also account for potential modelling

errors, we incorporate *uncertainty* through Bayesian inference to estimate the *joint-distribution* of ODE model parameters from data, which we discuss next.

3. Bayesian Parameter Estimation

Bayesian inference for estimating ODE parameters has been well studied in literature Ramsay et al. (2007), however in the absence of closed form posterior and the requirement of solving the ODE system in each sampling iteration makes inference difficult. To overcome this limitation, Calderhead et al. (2009) proposed the use of Gaussian Processes (GPs) to model the evolution of a state over time while exploiting the fact that derivative of a Gaussian process is also a Gaussian process. This significantly helps in achieving tractability and allows Bayesian inference to be computationally feasible. Following Calderhead et al. (2009), numerous other related works like Dondelinger et al. (2013); Barber and Wang (2014); Macdonald et al. (2015); Niu et al. (2016); Gorbach et al. (2017); Wenk et al. (2019, 2020) have been proposed which also employ the use of GPs to efficiently estimate the parameters of a non-linear ODE system (for eg: SEIR model). We discuss some of these works, in particular the approach of Wenk et al. (2019) which is useful to our problem setting.

Consider a set of K time-dependent states denoted as $\mathbf{x}(t) = [x_1(t), \ldots, x_K(t)]^T$. The evolution of each of these K state over time is defined by a set of K time-dependent arbitrary differential equations denoted as follows:

$$\dot{x}_i(t) = \frac{dx_i(t)}{dt} = \mathrm{f}_i(x(t), \theta, t) \quad \forall \quad i \in \{1, \ldots, K\} \tag{3}$$

where the functional form of f_i is known (for eg. SEIR model). Noisy observations (i.e. the time-series data) of each of the K states (denoted $\mathbf{y}(t) = [y_1(t), \ldots, y_K(t)]^T$) at N different time points where $t_1 < \cdots < t_N$ are available, i.e.

$$\left.\begin{aligned} y_1(t) &= x_1(t) + \epsilon_1(t) \\ &\vdots \\ y_K(t) &= x_K(t) + \epsilon_K(t) \end{aligned}\right\}, \quad \begin{aligned} &\text{where } \epsilon_i(t) \sim \mathcal{N}(0, \sigma_i^2). \\ &\forall \ t \in \{t_1, \ldots, t_N\} \end{aligned}$$

Let $\boldsymbol{\epsilon}(t) = [\epsilon_1(t), \ldots, \epsilon_K(t)]^T$, then in vector notation we have $\mathbf{y}(t) = \mathbf{x}(t) + \boldsymbol{\epsilon}(t)$. As there are N observations for each of the K states, for a clear exposition we introduce matrices of size $K \times N$ as follows: $\mathbf{X} = [\mathbf{x}(t_1), \ldots, \mathbf{x}(t_N)]$ and $\mathbf{Y} = [\mathbf{y}(t_1), \ldots, \mathbf{y}(t_N)]$. We can then write :

$$P(\mathbf{Y}|\mathbf{X}, \sigma) = \prod_k \prod_t P(y_k(t)|x_k(t), \sigma) = \prod_k \prod_t \mathcal{N}(y_k(t)|x_k(t), \sigma^2) \tag{4}$$

Calderhead et al. (2009) proposed placing a Gaussian process prior on $\mathbf{x}_k$. Let $\boldsymbol{\mu}_k$ and $\boldsymbol{\phi}_k$ be the hyper-parameters of this Gaussian process, we can then write:

$$p(\mathbf{x}_k|\boldsymbol{\mu}_k, \boldsymbol{\phi}_k) = \mathcal{N}(\mathbf{x}_k|\boldsymbol{\mu}_k, \mathbf{C}_{\phi_k}) \tag{5}$$

In (5), $\mathbf{C}_{\phi_k}$, denotes the Kernel (or the covariance) matrix for a predefined kernel function with hyper-parameters $\boldsymbol{\phi}_k$. As differentiation is a linear operator therefore the derivative of a Gaussian process is also a Gaussian process (see ch-9 in Rasmussen (2006) and Solak et al. (2002)). Therefore a Gaussian process is closed under differentiation and the joint distribution of the state variables $\mathbf{x}_k$ and their derivatives $\dot{\mathbf{x}}_k$ is a multi-variate Gaussian distribution as follows:

$$\begin{bmatrix} \mathbf{x}_k \\ \dot{\mathbf{x}}_k \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \boldsymbol{\mu}_k \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{C}_{\phi_k}, {}' \mathbf{C}_{\phi_k} \\ \mathbf{C}'_{\phi_k}, \mathbf{C}''_{\phi_k} \end{bmatrix} \right) \tag{6}$$

where $\mathbf{C}_{\phi_k}$ and $\mathbf{C}''_{\phi_k}$ are the kernel matrices for the state $\mathbf{x}_k$ and its derivative $\dot{\mathbf{x}}_k$ respectively, while ${}'\mathbf{C}_{\phi_k}$ and $\mathbf{C}'_{\phi_k}$ are the cross-covariance kernel matrices between the states and their derivatives. Functional form of the entries of $\mathbf{C}_{\phi_k}, \mathbf{C}''_{\phi_k}, {}'\mathbf{C}_{\phi_k}$ and $\mathbf{C}'_{\phi_k}$ are provided in SI. Importantly,

this implies that using the Gaussian process defined on the state variables $\mathbf{x}_k$, we can also make predictions about their derivatives $\dot{\mathbf{x}}_k$. From (6), we can compute the conditional distribution of the state derivatives as:

$$p(\dot{\mathbf{x}}_k|\mathbf{x}_k, \boldsymbol{\mu}_k, \boldsymbol{\phi}_k) = \mathcal{N}(\dot{\mathbf{x}}_k|\mathbf{m}_k, \mathbf{A}_k) \tag{7}$$

where $\mathbf{m}_k =' \mathbf{C}_{\phi_k}\mathbf{C}_{\phi_k}{}^{-1}(\mathbf{x}_k - \boldsymbol{\mu}_k)$; $\mathbf{A}_k = \mathbf{C}''_{\phi_k} -' \mathbf{C}_{\phi_k}\mathbf{C}_{\phi_k}{}^{-1}\mathbf{C}'_{\phi_k}$. Note that $p(\dot{\mathbf{x}}_k|\mathbf{x}_k, \boldsymbol{\mu}_k, \boldsymbol{\phi}_k)$ corresponds to the second, i.e. GP part of the graphical model in fig. 3.

Using the functional form of the ODE system in (3) and with state specific Gaussian additive noise λ_k, we can write

$$p(\dot{\mathbf{x}}_k|\mathbf{X}, \boldsymbol{\theta}, \lambda_k) = \mathcal{N}(\dot{\mathbf{x}}_k|\mathbf{f}_k(\mathbf{X}, \boldsymbol{\theta}), \lambda_k\mathbf{I}) \tag{8}$$

where $\mathbf{f}_k(\mathbf{X}, \boldsymbol{\theta}) = [\mathrm{f}_k(\mathrm{x}(t_1), \boldsymbol{\theta}), \ldots, \mathrm{f}_k(\mathrm{x}(t_N), \boldsymbol{\theta})]^T$. Note that (8) corresponds to the ODE part of the graphical model in the fig. 3.

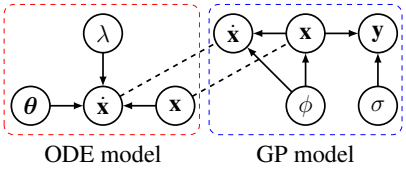


Fig. 3

The two models $p(\dot{\mathbf{x}}_k|\mathbf{x}_k, \boldsymbol{\mu}_k, \boldsymbol{\phi}_k)$ in (7) and $p(\dot{\mathbf{x}}_k|\mathbf{X}, \boldsymbol{\theta}, \lambda_k)$ in (8) are combined through two new random variables $\mathbf{F}_1$ and $\mathbf{F}_2$, resulting in the graphical model shown in fig. 4 Wenk et al. (2019). Considering a single state (for notational simplicity), for given values of $\mathbf{x}$ and $\boldsymbol{\theta}$, $\mathbf{F}_1$ in fig. 4, represents the deterministic output of the ODEs, i.e. $\mathbf{F}_1 = \mathbf{f}(\boldsymbol{\theta}, \mathbf{x})$. The value of $p(\mathbf{F}_1|\boldsymbol{\theta}, \mathbf{x})$ can be written using the Dirac-delta function (denoted $\delta(\cdot)$) as following:

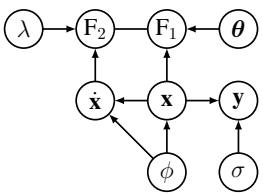


Fig. 4: Combined model.

$$p(\mathbf{F}_1|\boldsymbol{\theta}, \mathbf{x}) = \delta(\mathbf{F}_1 - \mathbf{f}(\boldsymbol{\theta}, \mathbf{x})) \tag{9}$$

Under the assumption that the GP model would be able to capture both, the true states and their derivatives perfectly, then it would imply that $\mathbf{F}_1$ is same as $\dot{\mathbf{x}}$, i.e. $\mathbf{F}_1 = \dot{\mathbf{x}}$. But clearly this assumption is unlikely to hold, therefore to account for any possible mismatch and small error in the GP states and GP derivatives, this condition is relaxed so that:

$$\mathbf{F}_1 = \dot{\mathbf{x}} + \epsilon =: \mathbf{F}_2, \quad \text{where} \quad \epsilon \sim \mathcal{N}(\mathbf{0}, \lambda\mathbf{I}) \tag{10}$$

The above argument regarding the the error in the states and derivatives of the GP model is captured in the graphical model (fig. 4) through the use of the random variable $\mathbf{F}_2$. From a given state-derivative $\dot{\mathbf{x}}$ obtained from the GP model, $\mathbf{F}_2$ is obtained after addition of Gaussian noise with standard deviation λ. The probability density of $\mathbf{F}_2$ can then be written as

$$p(\mathbf{F}_2|\dot{\mathbf{x}}, \lambda) = \mathcal{N}(\mathbf{F}_2|\dot{\mathbf{x}}, \lambda\mathbf{I})) \tag{11}$$

Note that the equality constraint in (10) is encoded in the graphical model using an un-directed edge between $\mathbf{F}_1$ and $\mathbf{F}_2$. For the purpose of inference, this equality constraint is incorporated in the joint density via the Dirac-delta function, i.e. $\delta(\mathbf{F}_1 - \mathbf{F}_2)$. The joint-density of the whole graphical model (fig. 4) is given as:

$$p(\mathbf{x}, \dot{\mathbf{x}}, \mathbf{y}, \mathbf{F}_1, \mathbf{F}_2, \boldsymbol{\theta}|\phi, \sigma, \lambda) = p(\boldsymbol{\theta})p(\mathbf{x}|\phi)p(\dot{\mathbf{x}}|\mathbf{x}, \phi)p(\mathbf{y}|\mathbf{x}, \sigma)p(\mathbf{F}_1|\boldsymbol{\theta}, \mathbf{x})p(\mathbf{F}_2|\dot{\mathbf{x}}, \lambda\mathbf{I})\delta(\mathbf{F}_1 - \mathbf{F}_2) \tag{12}$$

Finally, the marginal distribution of $\mathbf{x}, \boldsymbol{\theta}$ takes the following form:

$$p(\mathbf{x}, \boldsymbol{\theta}|\mathbf{y}, \boldsymbol{\phi}, \sigma, \lambda) = p(\boldsymbol{\theta}) \times \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \mathbf{C}_\phi) \times \mathcal{N}(\mathbf{y}|\mathbf{x}, \sigma^2\mathbf{I}) \times \mathcal{N}(\mathbf{f}(\mathbf{x}, \boldsymbol{\theta})|\mathbf{m}, \mathbf{A} + \lambda\mathbf{I}) \tag{13}$$

3.1. Empirical Sampling Results

We now provide sampling results on the two disease-transmission ODE models: 1) SEIR model (fig. 1) and 2) SEPIHR model (fig. 2).

SEIR: Using $\alpha = 0.9, \beta = 0.08$ and $\gamma = 0.1$ (as true parameter values) we simulate the SEIR model (eq:(1)) to get state values. We add zero-mean Gaussian noise with $\sigma = 0.1$ to each of the simulated state values to generate our dataset. Using the data only for first 15 days ($T =$

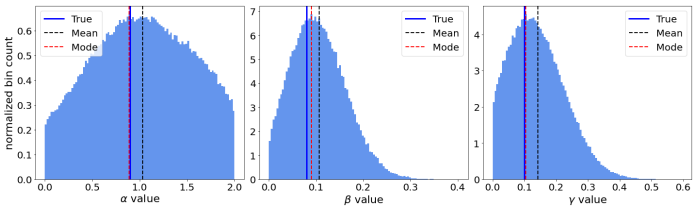


Fig. 5: SEIR: Empirical distribution after 3×10^5 samples from the MCMC sampling procedure.

$\{1, \ldots, 15\}$), we estimate the GP hyper-parameters for states using maximum-likelihood (see SI for details). We then run the Metropolis-Hastings MCMC sampling procedure using the density from eq: (13) to get our empirical posterior joint-distribution on α, β and γ. After removing the burn-in samples, for the remaining 3×10^5 samples, we plot the marginal distributions along with their mean and mode in fig. 5.

SEPIHR: This model has 5 parameters, i.e. $\alpha, \beta, \delta_1, \gamma_1$ and γ_2, for which the joint-distribution is to be estimated from data (see SI for model details). We use $\alpha = 1.1, \beta = 0.08, \delta_1 = 0.01, \delta_2 = 0.002, \delta_3 = 0.002, \gamma_1 = 0.1, \gamma_2 = 0.1$ and $\gamma_3 = 0.06$ as the true parameter values. Using the data for only first 15 days, we follow the same sampling procedure as described previously for the SEIR model. The marginal distributions along with their mean and mode are shown in fig. 6.

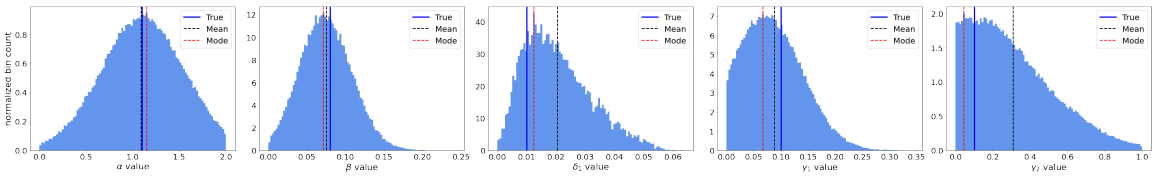


Fig. 6: SEPIHR: Empirical distribution after 3×10^5 samples from the MCMC sampling procedure.

We note that mode is very close to the true values in both the models, thus validating the capability of the sampling procedure in correctly estimating the parameter values. These empirical samples can be used to construct the scenario-set (denoted Ω) to represent parameter uncertainty in the vaccine allocation stochastic optimization formulation.

In a naive approach, each of the 3×10^5 samples can be used to represent a real-world scenario. However, working with such large sample size is computationally prohibitive and there is high redundancy in the samples. In literature, this issue has been well-studied as an optimal-transport problem Dupačová et al. (2003); Heitsch and Roemisch (2003); Rujeerapaiboon et al. (2022) and is commonly resolved using k-means clustering. For theoretical and other details behind k-means for optimal-transport formulation see SI. We also adopt the k-means approach and after performing k-means clustering on the empirical 3×10^5 samples, we generate our scenario-set Ω as follows:

$$\Omega = \left\{ (\alpha^j, \beta^j, \gamma^j, p_j) \; \forall \; j \in \{1, \ldots, \mathrm{k}\} \right\} \tag{14}$$

where $\alpha^j, \beta^j, \gamma^j$ denotes the location of the j-th centroid and p_j denotes its associated probability.

Related literature: Before concluding this section, we briefly mention related literature. Variational inference (VI) based approach of Gorbach et al. (2017) provides improvements over Dondelinger et al. (2013), however due to modelling assumptions is not suited for our work. The optimization based gradient matching approaches of Ramsay et al. (2007); Liang and Wu (2008); González et al. (2013); Niu et al. (2016); Wenk et al. (2020) and others like Gugushvili and Klaassen (2012) only provide point-estimates. The generative modelling approach of Barber and Wang (2014) suffers from identifiablity issues as explained by Macdonald et al. (2015). Approaches with different sampling methods would include Kramer et al. (2014); Paun and Husmeier (2022); Huang et al.

(2020); Calderhead and Girolami (2009, 2011) and approximation based methods would include Toni et al. (2009); Alahmadi et al. (2020); Dass et al. (2017). Other VI based methods would include Roeder et al. (2019); Ghosh et al. (2021). Probabilistic numerics Hennig et al. (2015) based methods include Tronarp et al. (2022) and Kersting et al. (2020). Teymur et al. (2018) showed the use of probabilistic integrators for ODEs in parameter estimation. Chkrebtii et al. (2016); Vanlier et al. (2013); Hug et al. (2013); Hauser et al. (2020) are other useful references.

4. Optimal Vaccine Allocation Formulation and Solution Algorithm

We now work towards formulating our optimization problem for vaccine allocation. Our goal is to allocate vaccines (on a daily basis) to a set of $\mathbb{K}$ sub-populations, such that the maximum number of total infections is minimized. This objective ensures that the peak of the pandemic is minimised as much as possible in order to reduce the burden on the healthcare services particularly medical personnel at the height of the pandemic. The $\mathbb{K}$ sub-populations correspond to different geographical regions such as nearby cities in a state. Let $\mathcal{K} = \{1, \ldots, \mathbb{K}\}$.

The spread of disease in each sub-population is modeled using a separate SEIR model. To account for the vaccinated individuals, the SEIR model in fig. 1 is updated with a new compartment (denoted by M) to represent the immune population and the updated model (fig. 7) is denoted by SEIRM. Let $V_k(t)$ represent the number of people vaccinated at time t in the k-th sub-population and η be efficacy of the vaccine, then the ODEs corresponding to the SEIRM model of the k-th sub-population are given by eq: (15).

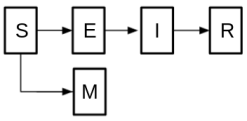


Fig. 7: SEIRM model

$$\left.\begin{aligned} \dot{S}_k(t) &= -\eta V_k(t) - \frac{u_k(t)\alpha}{N_k}\big(S_k(t) - \eta V_k(t)\big)\Big(\sum_{r=1}^{\mathbb{K}} \lambda_r^k I_r(t)\Big) \\ \dot{E}_k(t) &= \frac{u_k(t)\alpha}{N_k}\big(S_k(t) - \eta V_k(t)\big)\Big(\sum_{r=1}^{\mathbb{K}} \lambda_r^k I_r(t)\Big) - \beta E_k(t) \\ \dot{I}_k(t) &= \beta E_k(t) - \gamma I_k(t), \quad \dot{R}_k(t) = \gamma I_k(t), \quad \dot{M}_k(t) = \eta V_k(t) \end{aligned}\right\} \tag{15}$$

We also consider two important features of disease transmission. First, due to mobility there is contact between infected individuals of one sub-population with the susceptible individuals of another sub-population. Second, due to different levels of mobility between different sub-populations, onset of the pandemic in each of the sub-populations generally vary. Both of these are accounted in the updated states S_k and E_k in eq:(15), where λ_r^k denotes the mobility levels from sub-population r to sub-population k and $u_k(t)$ corresponds to a sigmoid function, $u_k(t) \coloneqq 1/(1 + e^{-c_1^k(t-c_2^k)})$ with parameters c_1^k and c_2^k. In particular, c_2^k controls the onset of the pandemic in the k-th sub-population, therefore we also account for uncertainty in $c_2^k \ \forall \ k \in \mathcal{K}$, by appropriately extending the scenario set Ω.

Let $\mathcal{T} = \{1, \ldots, T\}$, denote the simulation time period, $\mathcal{T}_v = \{t_s, \ldots, t_l\}$ denote the vaccination time-period where t_s and t_l are the first and last vaccination days, such that $\mathcal{T}_v \subseteq \mathcal{T}$. We can now write the *nominal (or non-stochastic)* optimization problem (denoted $\mathcal{NF}$) for vaccine allocation as (16), where B_t in (16e) denotes the total daily vaccine budget for all $\mathbb{K}$ sub-populations and U_t^k in (16f) denotes the vaccine budget for k-th sub-population. Equations (16b) represent the ODE constraints, (16c) & (16d) together computes the maximum (or peak) infection of the total population (denoted $\mathcal{I}$) and (16a) minimizes the peak infection.

We now provide the **uncertainty-informed**, i.e. stochastic counterpart (denoted $\mathcal{SF}$) of the nominal problem $\mathcal{NF}$ in (17), where Ω denotes the scenario-set, recall (14). Each state S,E,I,R,M in ODE constraints in (17b) now has an associated superscript ω corresponding to that scenario, $\mathcal{I}_\omega$ denotes the peak infection for scenario ω, (17a) computes the expected peak infection over all scenarios. The vaccine budget constrains in (17c) remain same as in $\mathcal{NF}$.

$$\mathcal{NF}: \min_{V} \quad \mathcal{I} \qquad \text{(Nominal)} \quad (16a)$$

s.t.

$$\{(15)\} \qquad \forall\ k \in \mathcal{K}, t \in \mathcal{T} \quad (16b)$$

$$\sum_{k=1}^{\mathbb{K}} I_k(t) = \tilde{I}(t) \qquad \forall\ t \in \mathcal{T} \quad (16c)$$

$$\tilde{I}(t) \leq \mathcal{I} \qquad \forall\ t \in \mathcal{T} \quad (16d)$$

$$\sum_{k=1}^{\mathbb{K}} V_k(t) \leq B_t \qquad \forall\ t \in \mathcal{T}_v \quad (16e)$$

$$0 \leq V_k(t) \leq U_t^k \quad \forall\ k \in \mathcal{K}, t \in \mathcal{T}_v \quad (16f)$$

$$V_k(t) = 0 \quad \forall\ k \in \mathcal{K}, t \in \mathcal{T} \setminus \mathcal{T}_v \quad (16g)$$

$$\mathcal{SF}: \min_{V} \quad \sum_{\omega \in \Omega} p_\omega \mathcal{I}_\omega \qquad \text{(Stochastic)} \quad (17a)$$

$$\left.\begin{aligned} \dot{S}_k^\omega(t) &= -\eta V_k(t) - \tfrac{u^\omega(t)\alpha^\omega}{N}\left(S_k^\omega(t) - \eta V_k(t)\right)\left(\textstyle\sum_{r=1}^{\mathbb{K}} \lambda_r^k I_r^\omega(t)\right) \\ \dot{E}_k^\omega(t) &= \tfrac{u^\omega(t)\alpha^\omega}{N}\left(S_k^\omega(t) - \eta V_k(t)\right)\left(\textstyle\sum_{r=1}^{\mathbb{K}} \lambda_r^k I_r^\omega(t)\right) - \beta^\omega E_k^\omega(t) \end{aligned}\right\}$$

$$\left.\begin{aligned} &\dot{I}_k^\omega(t) = \beta^\omega E_k^\omega(t) - \gamma^\omega I_k^\omega(t) \\ &\dot{R}_k^\omega(t) = \gamma^\omega I_k^\omega(t),\ \ \dot{M}_k^\omega(t) = \eta V_k(t) \\ &\sum_{k=1}^{\mathbb{K}} I_k^\omega(t) = \tilde{I}^\omega(t) \quad , \tilde{I}^\omega(t) \leq \mathcal{I}_\omega \end{aligned}\right\} \quad \begin{matrix} k \in \mathcal{K}, \\ \forall\ t \in \mathcal{T}, \\ \omega \in \Omega \end{matrix} \quad (17b)$$

$$\{(16e), (16f), (16g)\} \quad (17c)$$

Definition 1. A vaccine policy $\mathcal{V}$ is defined as: $\mathcal{V} = \{V_k(t)\ \forall\ k \in \mathcal{K}, t \in \mathcal{T}\}$.

Theorem 1. *Feasibility of $\mathcal{V}$: The feasibility of a vaccine policy $\mathcal{V}$ in $\mathcal{SF}$ is only decided by the budget constraints in (17c) and not by the ODE constraints in (17b).*

Theorem 1 holds true because of the budget constraints (17c), $V_k(t)$ is non-negative and finite. Thus the existence and uniqueness of a solution to ODEs in (17b) is guaranteed and can be shown analytically using the Picard-Lindelöf theorem with appropriate initial conditions Coddington and Levinson (1955); Sastry (2013); Sowole et al. (2019).

Lemma 1. *Decomposability w.r.t Ω: For a given (fixed) vaccine policy $\mathcal{V}$, the ODE constraints in (17b) become decomposable, i.e. the set of ODE constraints in scenario ω_i can be solved independently of the set of ODE constraints in scenario $\omega_j\ \forall j \in \Omega \setminus i$.*

Lemma (1) follows from the fact that for a given scenario (say ω_i), constraints in (17b) require parameters only corresponding to scenario ω_i. This has major computational implications as it allows for parallel evaluation of scenarios in Ω. Due to the additive nature of the objective function (17a) w.r.t. to Ω, we can compute the objective function value after parallel computation of scenarios. Therefore, we can efficiently solve $\mathcal{SF}$ using an iterative heuristic based optimization procedure described in algorithm 1. Details on heuristics are provided in SI. Note that due to the total vaccine budget constraint (16e) in $\mathcal{NF}$ and $\mathcal{SF}$, using approaches like Bayesian optimization (BO) is not feasible.

Algorithm 1 Optimization procedure to solve $\mathcal{NF}$ or $\mathcal{SF}$

1: Randomly sample a batch of Vaccine policies of size B, i.e. $\bar{\mathcal{V}}_0 = \{\mathcal{V}_1, \ldots, \mathcal{V}_B\}$ and set $i = 0$.
2: **while** $i \leq N_{\text{opt}}$ **do**
3: **for** $k \leftarrow 1$ to B **do**
4: Evaluate constraint violation (denoted C_k) of $\bar{\mathcal{V}}_i[k]$ using (17c) .
5: In *parallel*, simulate all $|\Omega|$ scenarios for $\bar{\mathcal{V}}_i[k]$ using an ODE solver to compute $\mathcal{I}_\omega$.
6: Compute f_{obj}^k(17a): $f_{obj}^k \leftarrow \sum_{\omega \in \Omega} p_\omega \mathcal{I}_\omega$
7: **end for**
8: Update the batch of vaccine policies $\bar{\mathcal{V}}_i$ with heuristic rules using $\{f_{obj}^1, \ldots, f_{obj}^B\}$ and $\{C_1, \ldots, C_B\}$ to generate next batch of vaccine policies $\bar{\mathcal{V}}_{i+1}$.
9: $i \leftarrow i + 1$
10: **end while**
11: **return** feasible vaccine policy $\mathcal{V}$ with lowest f_{obj}.

5. Experimental (Simulation) Results

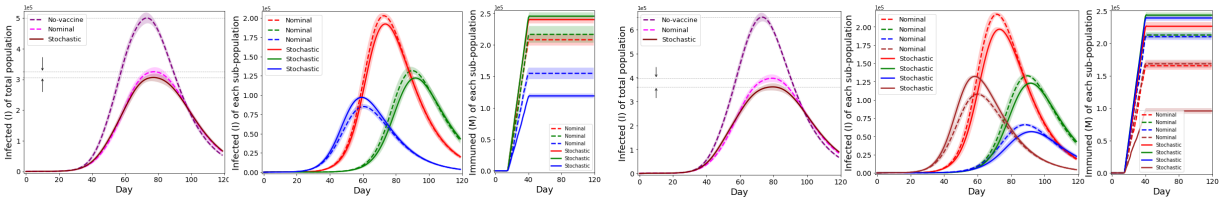
In this section, we show the efficacy of our proposed approach on two different disease transmission models, i.e. the SEIR and the SEPIHR models. For all experiments we report average of 5 runs. In addition to the experiments in this section, various other numerical experiments under different setups are provided in SI.

Recall that in section 3.1, we have already discussed the details and sampling results for both the SEIR and SEPIHR model including figures 5 & 6 respectively. We also outlined how to obtain the scenario-set Ω from the samples to account for uncertainty in vaccine allocation. Therefore our main goal in this section is to show the benefit of incorporating uncertainty by comparing the vaccine allocation policy (denoted $\mathcal{V}_{\mathcal{N}}$) obtained from solving the nominal formulation $\mathcal{NF}$ against the vaccine allocation policy (denoted $\mathcal{V}_{\mathcal{S}}$) obtained from solving the stochastic solution $\mathcal{SF}$. We also benchmark against a zero or no-vaccination policy denoted (denoted $\mathcal{V}_{\phi}$), where $\mathcal{V}_{\phi} = \{V_k(t) = 0 \; \forall \; t \in \mathcal{T}, k \in \mathcal{K}\}$.

SEIR model: We use a total simulation time horizon of $T = 120$ days, vaccination period of 25 days starting on $t_s = 16$ and ending on $t_l = 40$ with daily available vaccine budgets $B_t = 24 \times 10^3$ and $U_t^k = 10^4$. Importantly, note that in section 3.1 for parameter estimation, we used data only for first 15 days, i.e. $T = \{1, \ldots, 15\}$, thus maintaining consistency for real-world applicability. We perform experiments in two different settings, in the **first setting** we work with $\mathbb{K} = 3$ i.e. three sub-populations of sizes $7.5 \times 10^5, 5 \times 10^5$ and 10^6 respectively and in the **second setting** we increase $\mathbb{K}$ to $\mathbb{K} = 4$, with an additional sub-population of size 6×10^5. Numerical values of other parameters like $\lambda_r^k, c_1^k, c_2^k, \eta$ and additional experiments to evaluate their effect are provided in SI.

For each setting i.e. $\mathbb{K} = 3$ and $\mathbb{K} = 4$, using algorithm 1 and the nominal estimates of α, β and γ, we solve the $\mathcal{NF}$ to get the nominal vaccine policy $\mathcal{V}_{\mathcal{N}}$. Using the scenario-set Ω (generated from the discrete-parameter distribution) we solve $\mathcal{SF}$ to get the uncertainty-informed vaccine policy $\mathcal{V}_{\mathcal{S}}$.

We next evaluate the efficacy of all the three vaccine policies i.e. $\mathcal{V}_{\phi}, \mathcal{V}_{\mathcal{N}}$ and $\mathcal{V}_{\mathcal{S}}$. For each of these policies, we simulate all the scenarios in the scenario-set Ω and compute the expected values of all the states i.e. S,E,I,R and M over the time horizon T.



(a) $\mathbb{K} = 3$ sub-populations (b) $\mathbb{K} = 4$ sub-populations

Fig. 8: **SEIR:** Evaluation of diff. vaccine policies i.e. no-vaccine $\mathcal{V}_{\phi}$, nominal $\mathcal{V}_{\mathcal{N}}$ & stochastic $\mathcal{V}_{\mathcal{S}}$.

The evolution of the infected state (I) of the *total population* and the infected (I) and immuned (M) states of each *sub-population* are shown in fig. 8(*a*) and 8(*b*) for $\mathbb{K} = 3$ and 4 respectively. We note that for $\mathbb{K} = 3$ (fig 8(*a*)) , the expected peak infection is reduced from around 501k (with no-vaccination i.e. $\mathcal{V}_{\phi}$) to 324k with nominal vaccination policy $\mathcal{V}_{\mathcal{N}}$. This reduction of peak infection by 35.3% is expected due to vaccination. More importantly, we observe that with the stochastic vaccination policy $\mathcal{V}_{\mathcal{S}}$ the peak infection is further reduced to around 308k, which is an improvement of around **4.9%** over $\mathcal{V}_{\mathcal{N}}$ and 38.56% over $\mathcal{V}_{\phi}$. This improvement of $\mathcal{V}_{\mathcal{S}}$ over $\mathcal{V}_{\mathcal{N}}$ by **4.9%** is also referred to as the *value of stochastic solution* (VSS) or equivalently the benefit of accounting for uncertainty.

For $\mathbb{K} = 4$ (fig 8(*b*)), we observe that the peak infection with no-vaccine policy $\mathcal{V}_\phi$ is around 653k, and is reduced to 393k with $\mathcal{V}_\mathcal{N}$ and is further reduced to 361k with $\mathcal{V}_\mathcal{S}$, i.e. $\mathcal{V}_\mathcal{S}$ provides a reduction of around **8%** over $\mathcal{V}_\mathcal{N}$. This higher VSS of **8%** for $\mathbb{K} = 4$ compared to **4.9%** for $\mathbb{K} = 3$ is due to the fact that the size of scenario set $|\Omega|$ is directly proportional to the number of sub-populations $\mathbb{K}$. Recall that we also account for the uncertainty in the onset of the pandemic in each sub-population through the parameter c_2^k.

Note that since the immuned sub-population size is directly proportional to vaccines allocated to that sub-population, therefore the third figure in 8(*a*) and 8(*b*) also shows how many vaccines are allocated to each sub-population relative to each other. We observe that there is a clear difference between the nominal and the stochastic allocations. This significant difference in nature of the vaccine policies explain the reduction of **4.9%** and **8%** respectively, providing validity to our results in the sense that the reductions obtained are not simply due to minor numerical changes in solution values. Elaborate discussions on the differences of the two policies ($\mathcal{V}_\mathcal{N}$ vs $\mathcal{V}_\mathcal{S}$) are provided in SI.

SEPIHR model: We next evaluate our approach on the SEPIHR model with additional states P (for protective quarantine) and H (for hospitalised quarantined). Corresponding optimization formulations (i.e. nominal and stochastic) are provided in SI. Importantly here as the number of hospitalisations (H) is modeled explicitly, therefore we minimize the peak (maximum) hospitalisations.

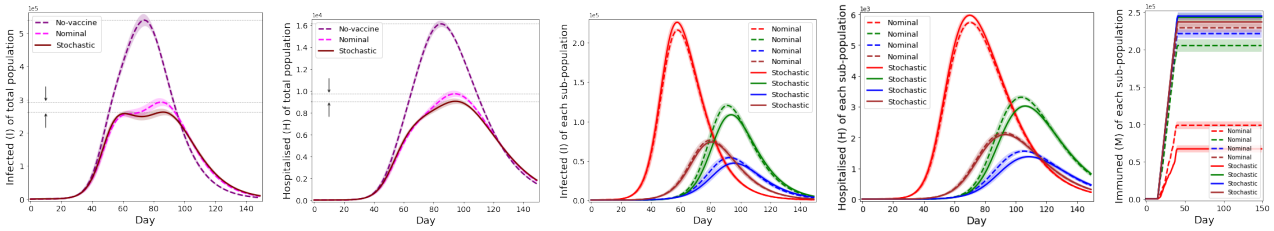


Fig. 9: **SEPIHR:** Evaluation of different policies: $\mathcal{V}_\phi$, $\mathcal{V}_\mathcal{N}$ and $\mathcal{V}_\mathcal{S}$ with $\mathbb{K} = 4$ sub-populations.

In fig. 9, we show the evolution of the infected (I) and hospitalised (H) states of the total population, along with the I,H and immuned (M) for each of the 4 sub-populations. We note that the peak infections (I) for the three policies (i.e. $\mathcal{V}_\phi$, $\mathcal{V}_\mathcal{N}$ and $\mathcal{V}_\mathcal{S}$) are around 539k, 280k and 262k and the peak hospitalisations are around 16k, 9.4k and 9k respectively. Therefore, $\mathcal{V}_\mathcal{S}$ provides a reduction of **6.3%** in peak infections (I) over $\mathcal{V}_\mathcal{N}$ and a **4.4%** reduction in peak hospitalisations (H). Interestingly, from the fifth plot in fig. 8, we note that despite its largest size and earliest the onset of the pandemic, red population is allocated the least vaccines. This can be explained by the fact that we aim to minimize the peak of the total population (see SI for detailed discussion).

The above results on SEIR and SEPIHR models clearly demonstrate the benefit of uncertainty-informed vaccine allocation using Bayesian inference over using nominal estimates. Our improvements of **4-8%** are either consistent with prior works in literature such as Yarmand et al. (2014) or much better Thul and Powell (2023).

6. Concluding Remarks and Future Work

In this paper, we proposed first, an uncertainty informed vaccine allocation problem as a stochastic optimization problem, for which the tractable scenario-set is constructed in a novel data-driven manner using Bayesian inference for ODEs with GPs and second a scalable solution algorithm to solve the stochastic program and showed that a significant gain can be achieved by accounting for uncertainty. For future work, a natural extension would be to systematically investigate equity and fairness of allocation through additional constraints and different objective functions.

References

D. Acemoglu, A. Fallah, A. Giometto, D. Huttenlocher, A. Ozdaglar, F. Parise, and S. Pattathil. Optimal adaptive testing for epidemic control: combining molecular and serology tests, 2021. URL https://arxiv.org/abs/2101.00773.

Amani A. Alahmadi, Jennifer A. Flegg, Davis G. Cochrane, Christopher C. Drovandi, and Jonathan M. Keith. A comparison of approximate versus exact techniques for bayesian parameter inference in nonlinear ordinary differential equation models. *Royal Society Open Science*, 7(3): 191315, 2020. doi: 10.1098/rsos.191315.

Anastasios Nikolas Angelopoulos, Reese Pathak, Rohit Varma, and Michael I. Jordan. On identifying and mitigating bias in the estimation of the covid-19 case fatality rate. *Harvard Data Science Review*, 7 2020. doi: 10.1162/99608f92.f01ee285. URL https://hdsr.mitpress.mit.edu/pub/y9vc2u36. https://hdsr.mitpress.mit.edu/pub/y9vc2u36.

David Barber and Yali Wang. Gaussian processes for bayesian estimation in ordinary differential equations. In Eric P. Xing and Tony Jebara, editors, *Proceedings of the 31st International Conference on Machine Learning*, volume 32 of *Proceedings of Machine Learning Research*, pages 1485–1493, Bejing, China, 22–24 Jun 2014. PMLR. URL https://proceedings.mlr.press/v32/barber14.html.

NG Becker. The use of mathematical models in determining vaccination policies. *Bulletin of the International Statistics Institute*, 46:478–490, 1975.

Dimitris Bertsimas, Joshua Ivanhoe, Alexandre Jacquillat, Michael Li, Alessandro Previero, Omar Skali Lami, and Hamza Tazi Bouardi. Optimizing vaccine allocation to combat the covid-19 pandemic. *medRxiv*, 2020. doi: 10.1101/2020.11.17.20233213. URL https://www.medrxiv.org/content/early/2020/11/18/2020.11.17.20233213.

John R. Birge, Ozan Candogan, and Yiding Feng. Controlling epidemic spread: Reducing economic losses with targeted closures. *Management Science*, 68(5):3175–3195, 2022. doi: 10.1287/mnsc.2022.4318.

Fred Brauer. *Compartmental Models in Epidemiology*, pages 19–79. Springer Berlin Heidelberg, 2008.

Fred Brauer. Mathematical epidemiology: Past, present, and future. *Infectious Disease Modelling*, 2(2):113–127, 2017. ISSN 2468-0427. doi: https://doi.org/10.1016/j.idm.2017.02.001. URL https://www.sciencedirect.com/science/article/pii/S2468042716300367.

Suzanne Brøgger. Systems analysis in tuberculosis control: a model. *The American review of respiratory disease*, 95 3:419–34, 1967.

Ben Calderhead and Mark Girolami. Estimating bayes factors via thermodynamic integration and population mcmc. *Computational Statistics and Data Analysis*, 53(12):4028–4045, 2009. ISSN 0167-9473. doi: https://doi.org/10.1016/j.csda.2009.07.025.

Ben Calderhead and Mark Girolami. Statistical analysis of nonlinear dynamical systems using differential geometric sampling methods. *Interface Focus*, 1(6):821–835, 2011. doi: 10.1098/rsfs.2011.0051.

Ben Calderhead, Mark Girolami, and Neil Lawrence. Accelerating bayesian inference over nonlinear differential equations with gaussian processes. In D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, editors, *Advances in Neural Information Processing Systems*, volume 21. Curran Associates, Inc., 2009. URL https://proceedings.neurips.cc/paper/2008/file/07563a3fe3bbe7e3ba84431ad9d055af-Paper.pdf.

Oksana A. Chkrebtii, David A. Campbell, Ben Calderhead, and Mark A. Girolami. Bayesian Solution Uncertainty Quantification for Differential Equations. *Bayesian Analysis*, 11(4):1239 – 1267, 2016. doi: 10.1214/16-BA1017. URL https://doi.org/10.1214/16-BA1017.

Leonardo Cianfanelli, Francesca Parise, Daron Acemoglu, Giacomo Como, and Asuman Ozdaglar. Lockdown interventions in sir models: Is the reproduction number the right control variable? In *2021 60th IEEE Conference on Decision and Control (CDC)*, pages 4254–4259, 2021. doi: 10.1109/CDC45484.2021.9682977.

Damian Clancy and Nathan Green. Optimal intervention for an epidemic model under parameter uncertainty. *Mathematical biosciences*, 205:297–314, 02 2007. doi: 10.1016/j.mbs.2006.08.023.

E. A. Coddington and N. Levinson. *Theory of ordinary differential equations*. McGraw-Hill New York, 1955.

Estee Y. Cramer, Evan L. Ray, Velma K. Lopez, Johannes Bracher, Andrea Brennen, Alvaro J. Castro Rivadeneira, ..., and Nicholas G. Reich. Evaluation of individual and ensemble probabilistic forecasts of covid-19 mortality in the united states. *Proceedings of the National Academy of Sciences*, 119(15), 2022. doi: 10.1073/pnas.2113561119. URL https://www.pnas.org/doi/abs/10.1073/pnas.2113561119.

Sarat C. Dass, Jaeyong Lee, Kyoungjae Lee, and Jonghun Park. Laplace based approximate posterior inference for differential equation models. *Statistics and Computing*, 27(3):679–698, may 2017. ISSN 0960-3174. doi: 10.1007/s11222-016-9647-0. URL https://doi.org/10.1007/s11222-016-9647-0.

Frank Dondelinger, Dirk Husmeier, Simon Rogers, and Maurizio Filippone. Ode parameter inference using adaptive gradient matching with gaussian processes. In Carlos M. Carvalho and Pradeep Ravikumar, editors, *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics*, volume 31 of *Proceedings of Machine Learning Research*, pages 216–228, Scottsdale, Arizona, USA, 29 Apr–01 May 2013. PMLR. URL https://proceedings.mlr.press/v31/dondelinger13a.html.

Jitka Dupačová, N. Gröewe-Kuska, and Werner Röemisch. Scenario reduction in stochastic programming: An approach using probability metrics. *Mathematical Programming*, 95:493–, 01 2003.

Chenyi Fu, Melvyn Sim, and Minglong Zhou. Robust epidemiological prediction and optimization. 2021. doi: http://dx.doi.org/10.2139/ssrn.3869521.

Sanmitra Ghosh, Paul Birrell, and Daniela De Angelis. Variational inference for nonlinear ordinary differential equations. In Arindam Banerjee and Kenji Fukumizu, editors, *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, volume 130 of *Proceedings of Machine Learning Research*, pages 2719–2727. PMLR, 13–15 Apr 2021.

Javier González, Ivan Vujačić, and Ernst Wit. Reproducing kernel hilbert space based estimation of systems of ordinary differential equations. *Pattern Recognition Letters*, 45, 11 2013. doi: 10.1016/j.patrec.2014.02.019.

Nico S Gorbach, Stefan Bauer, and Joachim M Buhmann. Scalable variational inference for dynamical systems. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL https://proceedings.neurips.cc/paper/2017/file/e71e5cd119bbc5797164fb0cd7fd94a4-Paper.pdf.

Shota Gugushvili and Chris A. J. Klaassen. $\sqrt{n}$-consistent parameter estimation for systems of ordinary differential equations: bypassing numerical integration via smoothing. *Bernoulli*, 18(3): 1061–1098, 2012. ISSN 1350-7265. doi: 10.3150/11-BEJ362.

Shuo Han, Victor M. Preciado, Cameron Nowzari, and George J. Pappas. Data-driven network resource allocation for controlling spreading processes. *IEEE Transactions on Network Science and Engineering*, 2(4):127–138, 2015. doi: 10.1109/TNSE.2015.2500158.

Anthony Hauser, Michel J. Counotte, Charles C. Margossian, Garyfallos Konstantinoudis, Nicola Low, Christian L. Althaus, and Julien Riou. Estimation of sars-cov-2 mortality during the early stages of an epidemic: a modeling study in hubei, china, and six regions in europe. *medRxiv*, 2020. doi: 10.1101/2020.03.04.20031104. URL https://www.medrxiv.org/content/early/2020/07/12/2020.03.04.20031104.

Mikhail Hayhoe, Francisco Barreras, and Victor M. Preciado. Multitask learning and nonlinear optimal control of the covid-19 outbreak: A geometric programming approach. *Annual Reviews in Control*, 52:495–507, 2021. ISSN 1367-5788. doi: https://doi.org/10.1016/j.arcontrol.2021.04.014. URL https://www.sciencedirect.com/science/article/pii/S1367578821000389.

Holger Heitsch and Werner Roemisch. Scenario reduction algorithms in stochastic programming. *Computational Optimization and Applications*, 24:187–206, 01 2003. doi: 10.1023/A:1021805924152.

Philipp Hennig, Michael Osborne, and Mark Girolami. Probabilistic numerics and uncertainty in computations. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 471, 06 2015. doi: 10.1098/rspa.2015.0142.

Hanwen Huang, Andreas Handel, and Xiao Song. A bayesian approach to estimate parameters of ordinary differential equation. *Computational Statistics*, 35, 09 2020. doi: 10.1007/s00180-020-00962-8.

S. Hug, A. Raue, J. Hasenauer, J. Bachmann, U. Klingmüller, J. Timmer, and F.J. Theis. High-dimensional bayesian parameter estimation: Case study for a model of jak2/stat5 signaling. *Mathematical Biosciences*, 246(2):293–304, 2013. ISSN 0025-5564. doi: https://doi.org/10.1016/j.mbs.2013.04.002.

Hans Kersting, Nicholas Krämer, Martin Schiegg, Christian Daniel, Michael Tiemann, and Philipp Hennig. Differentiable likelihoods for fast inversion of 'likelihood-free' dynamical systems. In *Proceedings of the 37th International Conference on Machine Learning*, ICML'20. JMLR.org, 2020.

Botir Kobilov, Ethan Rouen, and George Serafeim. Predictable country-level bias in the reporting of covid-19 deaths. *Journal of Government and Economics*, 2:100012, 2021. ISSN 2667-3193. doi: https://doi.org/10.1016/j.jge.2021.100012. URL https://www.sciencedirect.com/science/article/pii/S2667319321000124.

Andrei Kramer, Ben Calderhead, and Nicole Radde. Hamiltonian monte carlo methods for efficient parameter estimation in steady state dynamical systems. *BMC bioinformatics*, 15:253, 07 2014. doi: 10.1186/1471-2105-15-253.

Johannes Köhler, Lukas Schwenkel, Anne Koch, Julian Berberich, Patricia Pauli, and Frank Allgöwer. Robust and optimal predictive control of the covid-19 outbreak. *Annual Reviews in Control*, 51, 12 2020. doi: 10.1016/j.arcontrol.2020.11.002.

Michael Lingzhi Li, Hamza Tazi Bouardi, Omar Skali Lami, Thomas A. Trikalinos, Nikolaos Trichakis, and Dimitris Bertsimas. Forecasting covid-19 and analyzing the effect of government interventions. *Operations Research*, 2022. doi: 10.1287/opre.2022.2306.

Hua Liang and Hulin Wu. Parameter estimation for differential equation models using a framework of measurement error in regression models. *Journal of the American Statistical Association*, 103: 1570–1583, 12 2008. doi: 10.1198/016214508000000797.

Benn Macdonald, Catherine Higham, and Dirk Husmeier. Controversy in mechanistic modelling with gaussian processes. In Francis Bach and David Blei, editors, *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pages 1539–1547, Lille, France, 07–09 Jul 2015. PMLR.

Matthew Newton and Antonis Papachristodoulou. Network lyapunov functions for epidemic models. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pages 1798–1803, 2020. doi: 10.1109/CDC42340.2020.9304021.

Mu Niu, Simon Rogers, Maurizio Filippone, and Dirk Husmeier. Fast parameter inference in nonlinear dynamical systems using iterative gradient matching. In Maria Florina Balcan and Kilian Q. Weinberger, editors, *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 1699–1707, New York, New York, USA, 20–22 Jun 2016. PMLR.

Cameron Nowzari, Victor M. Preciado, and George J. Pappas. Analysis and control of epidemics: A survey of spreading processes on complex networks. *IEEE Control Systems Magazine*, 36(1): 26–46, 2016. doi: 10.1109/MCS.2015.2495000.

Cameron Nowzari, Victor M. Preciado, and George J. Pappas. Optimal resource allocation for control of networked epidemic models. *IEEE Transactions on Control of Network Systems*, 4(2): 159–169, 2017. doi: 10.1109/TCNS.2015.2482221.

Philip E. Paré, Carolyn L. Beck, and Tamer Başar. Modeling, estimation, and analysis of epidemics over networks: An overview. *Annual Reviews in Control*, 50:345–360, 2020. ISSN 1367-5788. doi: https://doi.org/10.1016/j.arcontrol.2020.09.003. URL https://www.sciencedirect.com/science/article/pii/S1367578820300614.

L. Mihaela Paun and Dirk Husmeier. Emulation-accelerated hamiltonian monte carlo algorithms for parameter estimation and uncertainty quantification in differential equation models. *Statistics and Computing*, 32(1), feb 2022. ISSN 0960-3174. doi: 10.1007/s11222-021-10060-4.

Victor Preciado, Michael Zargham, Chinwendu Enyioha, Ali Jadbabaie, and George Pappas. Optimal resource allocation for network protection against spreading processes. *Control of Network Systems, IEEE Transactions on*, 1:99–108, 03 2014. doi: 10.1109/TCNS.2014.2310911.

Victor M. Preciado, Michael Zargham, Chinwendu Enyioha, Ali Jadbabaie, and George Pappas. Optimal vaccine allocation to control epidemic outbreaks in arbitrary networks. In *52nd IEEE Conference on Decision and Control*, pages 7486–7491, 2013. doi: 10.1109/CDC.2013.6761078.

J. O. Ramsay, G. Hooker, D. Campbell, and J. Cao. Parameter estimation for differential equations: a generalized smoothing approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(5):741–796, 2007. doi: https://doi.org/10.1111/j.1467-9868.2007.00610.x. URL https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9868.2007.00610.x.

Carl Edward Rasmussen. Gaussian processes for machine learning. MIT Press, 2006.

Geoffrey Roeder, Paul Grant, Andrew Phillips, Neil Dalchau, and Edward Meeds. Efficient amortised Bayesian inference for hierarchical and nonlinear dynamical systems. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 4445–4455. PMLR, 09–15 Jun 2019.

G. Rohith. An augmented seir model with protective and hospital quarantine dynamics for the control of covid-19 spread. *medRxiv*, 2021. doi: 10.1101/2021.01.08.21249467. URL https://www.medrxiv.org/content/early/2021/01/09/2021.01.08.21249467.

Napat Rujeerapaiboon, Kilian Schindler, Daniel Kuhn, and Wolfram Wiesemann. Scenario reduction revisited: fundamental limits and guarantees. *Mathematical Programming*, 191:207–242, 2022.

S. Sastry. *Nonlinear Systems: Analysis, Stability, and Control*. Interdisciplinary Applied Mathematics. Springer New York, 2013. ISBN 9781475731088. URL https://books.google.com/books?id=j_PiBwAAQBAJ.

E. Solak, R. Murray-smith, W. Leithead, D. Leith, and Carl Rasmussen. Derivative observations in gaussian process models of dynamic systems. In S. Becker, S. Thrun, and

K. Obermayer, editors, *Advances in Neural Information Processing Systems*, volume 15. MIT Press, 2002. URL https://proceedings.neurips.cc/paper/2002/file/5b8e4fd39d9786228649a8a8bec4e008-Paper.pdf.

Vera L. J. Somers and Ian R. Manchester. Multi-stage sparse resource allocation for control of spreading processes over networks. In *2022 American Control Conference (ACC)*, pages 3632–3639, 2022. doi: 10.23919/ACC53348.2022.9867834.

Vera L. J. Somers and Ian R. Manchester. Minimizing the risk of spreading processes via surveillance schedules and sparse control. *IEEE Transactions on Control of Network Systems*, 10(1): 394–406, 2023. doi: 10.1109/TCNS.2022.3203359.

Samuel Sowole, Daouda Sangare, Abdullahi Ibrahim, and Isaac Paul. On the existence, uniqueness, stability of solution and numerical simulations of a mathematical model for measles disease. 2019:84–111, 07 2019.

Matthew Tanner, Lisa Sattenspiel, and Lewis Ntaimo. Finding optimal vaccination strategies under parameter uncertainty using stochastic programming. *Mathematical biosciences*, 215:144–51, 08 2008. doi: 10.1016/j.mbs.2008.07.006.

Onur Teymur, Ben Calderhead, Han Cheng Lie, and T. J. Sullivan. Implicit probabilistic integrators for odes. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, page 7255–7264, Red Hook, NY, USA, 2018. Curran Associates Inc.

Lawrence Thul and Warren Powell. Stochastic optimization for vaccine and testing kit allocation for the covid-19 pandemic. *European Journal of Operational Research*, 304(1):325–338, 2023. ISSN 0377-2217. doi: https://doi.org/10.1016/j.ejor.2021.11.007. URL https://www.sciencedirect.com/science/article/pii/S037722172100953X. The role of Operational Research in future epidemics/ pandemics.

Tina Toni, David Welch, Natalja Strelkowa, Andreas Ipsen, and Michael Stumpf. Approximate bayesian computation scheme for parameter inference and model selection in dynamical systems. j r soc interface 6: 187-202. *Journal of the Royal Society, Interface / the Royal Society*, 6:187–202, 03 2009. doi: 10.1098/rsif.2008.0172.

Jackeline Abad Torres, Sandip Roy, and Yan Wan. Sparse resource allocation for linear network spread dynamics. *IEEE Transactions on Automatic Control*, 62(4):1714–1728, 2017. doi: 10.1109/TAC.2016.2593895.

Filip Tronarp, Nathanael Bosch, and Philipp Hennig. Fenrir: Physics-enhanced regression for initial value problems, 2022. URL https://arxiv.org/abs/2202.01287.

J. Vanlier, C.A. Tiemann, P.A.J. Hilbers, and N.A.W. van Riel. Parameter uncertainty in biochemical models described by ordinary differential equations. *Mathematical Biosciences*, 246(2):305–314, 2013. ISSN 0025-5564. doi: https://doi.org/10.1016/j.mbs.2013.03.006.

Rui Wang, Danielle Maddix, Christos Faloutsos, Yuyang Wang, and Rose Yu. Bridging physics-based and data-driven modeling for learning dynamical systems. 2021.

Philippe Wenk, Alkis Gotovos, Stefan Bauer, Nico S. Gorbach, Andreas Krause, and Joachim M. Buhmann. Fast gaussian process based gradient matching for parameter identification in systems of nonlinear odes. In Kamalika Chaudhuri and Masashi Sugiyama, editors, *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, volume 89 of *Proceedings of Machine Learning Research*, pages 1351–1360. PMLR, 16–18 Apr 2019. URL https://proceedings.mlr.press/v89/wenk19a.html.

Philippe Wenk, Gabriele Abbati, Michael Osborne, Bernhard Schölkopf, Andreas Krause, and Stefan Bauer. Odin: Ode-informed regression for parameter and state inference in time-continuous dynamical systems. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34:6364–6371, 04 2020. doi: 10.1609/aaai.v34i04.6106.

Hamed Yarmand, Julie S. Ivy, Brian Denton, and Alun L. Lloyd. Optimal two-phase vaccine allocation to geographically different regions under uncertainty. *European Journal of Operational Research*, 233(1):208–219, 2014. ISSN 0377-2217. doi: https://doi.org/10.1016/j.ejor.2013.08.027. URL https://www.sciencedirect.com/science/article/pii/S0377221713006929.

Xuecheng Yin and I. Esra Buyuktahtakin. A multi-stage stochastic programming approach to epidemic resource allocation with equity considerations. *Health Care Management Science*, 2021.

Xuecheng Yin, I. Esra Buyuktahtakin, and Bhumi P. Patel. Covid-19: Data-driven optimal allocation of ventilator supply under uncertainty and risk. *European Journal of Operational Research*, 304(1):255–275, 2023. ISSN 0377-2217. doi: https://doi.org/10.1016/j.ejor.2021.11.052. URL https://www.sciencedirect.com/science/article/pii/S0377221721010031.