



Planning as Inference in Epidemiological Dynamics Models

Frank Wood^{1,2}, Andrew Warrington³, Saeid Naderiparizi¹, Christian Weilbach¹, Vaden Masrani¹, William Harvey¹, Adam Šcibior¹, Boyan Beronov¹, John Grefenstette⁴, Duncan Campbell⁴ and S. Ali Nasser^{1*}

¹Department of Computer Science, University of British Columbia, Vancouver, BC, Canada, ²Associate Academic Member and Canada CIFAR AI Chair, Mila Institute, Montreal, QC, Canada, ³Department of Engineering Science, University of Oxford, Oxford, United Kingdom, ⁴Epistemix Inc., Pittsburgh, PA, United States

In this work we demonstrate how to automate parts of the infectious disease-control policy-making process via performing inference in existing epidemiological models. The kind of inference tasks undertaken include computing the posterior distribution over controllable, via direct policy-making choices, simulation model parameters that give rise to acceptable disease progression outcomes. Among other things, we illustrate the use of a probabilistic programming language that automates inference in existing simulators. Neither the full capabilities of this tool for automating inference nor its utility for planning is widely disseminated at the current time. Timely gains in understanding about how such simulation-based models and inference automation tools applied in support of policy-making could lead to less economically damaging policy prescriptions, particularly during the current COVID-19 pandemic.

Keywords: public health preparedness, epidemiological dynamics, Bayesian inference, probabilistic programming, COVID-19

OPEN ACCESS

Edited by:

Weida Tong,
National Center for Toxicological
Research (FDA), United States

Reviewed by:

Theophane Weber,
DeepMind Technologies Limited,
United Kingdom
Dong Wang,
National Center for Toxicological
Research (FDA), United States

*Correspondence:

S. Ali Nasser
ali.nasser@ubc.ca

Specialty section:

This article was submitted to
Medicine and Public Health,
a section of the journal
Frontiers in Artificial Intelligence

Received: 09 April 2020

Accepted: 25 October 2021

Published: 31 March 2022

Citation:

Wood F, Warrington A, Naderiparizi S,
Weilbach C, Masrani V, Harvey W,
Scibior A, Beronov B, Grefenstette J,
Campbell D and Nasser SA (2022)
Planning as Inference in
Epidemiological Dynamics Models.
Front. Artif. Intell. 4:550603.
doi: 10.3389/frai.2021.550603

1 INTRODUCTION

Our goal in this paper is to demonstrate how the “*planning as inference*” methodology at the intersection of Bayesian statistics and optimal control can directly aid policy-makers in assessing policy options and achieving policy goals, when implemented using epidemiological simulators and suitable automated software tools for probabilistic inference. Such software tools can be used to quickly identify the range of values towards which controllable variables should be driven by means of policy interventions, social pressure, or public messaging, so as to limit the spread and impact of an infectious disease such as COVID-19.

In this work, we introduce and apply a simple form of planning as inference in epidemiological models to automatically identify policy decisions that achieve a desired, high-level outcome. As but one example, if our policy aim is to contain infectious population totals below some threshold at all times in the foreseeable future, we can condition on this putative future holding and examine the allowable values of controllable behavioural variables at the agent or population level, which in the framework of planning as inference is formalized in terms of a *posterior distribution*. As we already know, to control the spread of COVID-19 and its impact on society, policies must be enacted that reduce disease transmission probability or lower the frequency and size of social interactions. This is because we might like to, for instance, not have the number of infected persons requiring hospitalization exceed the number of available hospital beds.

Throughout this work, we take a Bayesian approach, or at the very least, a *probabilistic* consideration of the task. Especially in the early stages of a new outbreak, the infectious

dynamics are not known precisely. Furthermore, the spread of the disease cannot be treated deterministically, as a result of either fundamental variability in the social dynamics that drive infections, or of uncertainty over the current infection levels, or of uncertainty over appropriate models for analyzing the dynamics. Therefore, developing methods capable of handling such uncertainty correctly will allow for courses of action to be evaluated and compared more effectively, and could lead to “better” policies: a policy that surpasses the desired objective with 55% probability, but fails with 45% probability, may be considered “worse” than a policy that simply meets the objective with 90% probability. Bayesian analysis also offers a form of probabilistic inference in which the contribution of individual variables to the overall uncertainty can be identified and quantified. Beyond simply obtaining “the most effective policy measure,” this may be of interest to analysts trying to further understand *why* certain measures are more effective than others.

We first show how the problem of policy planning can be formulated as a Bayesian inference task in epidemiological models. This framing is general and extensible. We then demonstrate how particular existing software tools can be employed to perform this inference task in preexisting stochastic epidemiological models, without modifying the model itself or placing restrictions on the models that can be analyzed. This approach is particularly appealing, as it decouples the specification of epidemiological models by domain experts from the computational task of performing inference. This shift allows for more expressive and interpretable models to be expediently analyzed, and the sophistication of inference algorithms to be adjusted flexibly.

As a result, the techniques and tools we review in this paper are applicable to simulators ranging from simple compartmental models to highly expressive agent-based population dynamics models. In the former, the controls available to policy-makers are blunt—e.g., “reduce social interactions by some fractional amount”—but how best to achieve this is left as an exercise for policy-makers. In the latter, variables like “probability of individuals adhering to self-isolation” and “how long should schools be closed if at all” can be considered and evaluated in combination and comparison to others as potential fine-grained controls that could achieve the same policy objective more efficiently.

When governments impose any such controls, both citizens and financial markets want to know how draconian these measures must be and for how long they have to be in effect. Policy analysis based on models that reflect variability in resources such as healthcare facilities in different jurisdictions could hopefully make the answers to these questions more precise and the controls imposed more optimal in the utility maximizing sense. The same holds for the difference between models that can or cannot reflect variations in population mixing between rural and urban geographic areas. A person living in a farming county in central Illinois might reasonably wonder if it is as necessary to shelter in place there as it is in Chicago.

Current approaches to model-based policy-making are likely to be blunt. Simple models, e.g., compartmental models with few compartments, are rapid to fit to new diseases and easy to

compute, but are incapable of evaluating policy options that are more fine-grained than the binning used, such as regionalized measures. The net effect of being able to only consider blunt controls arguably has contributed to a collective dragging of feet, even in the face of the current COVID-19 pandemic. This delayed reaction, combined with brute application of control, has led to devastating socioeconomic impact, with many sectors such as education, investment markets and small-medium enterprises being directly impacted.

We can and should be able to do better. We believe, and hope to demonstrate, that models and software automation focused on planning as inference for policy analysis and recommendation is one path forward that can help us better react to this and future pandemics, and improve our public health preparedness.

We upfront note that the specific models that we use in this paper are far from perfect. First, the pre-existing models we use to demonstrate our points in this paper are only crudely calibrated to present-day population dynamics and specific COVID-19 characteristics. We have made some efforts on the latter point, in particular sourcing a COVID-19 adapted compartmental model (Ogilvy Kermack and McKendrick, 1927; Blackwood and Childs, 2018b; Hill et al., 2020) and parameters from (Bhalchandra and SnehalShekatkar, 2020; Ferguson et al., 2020; Magdon-Ismail, 2020; Massonnaud et al., 2020; Peng et al., 2020; Riou and Althaus, 2020; Rovetta et al., 2020; Russo et al., 2020; Traini et al., 2020; Wen et al., 2020), but we stress this limitation. In addition, in simple cases, the type of problems for which we discuss solutions in this paper may be solved with more straightforward implementations involving parameter sweeps and “manual” checking for desired policy effects, albeit at potentially higher human cost. In this sense, our goal is not to claim fundamental novelty, uniqueness or superiority of any particular inference technique, but rather to raise awareness for the practical feasibility of the Bayesian formulation of planning as inference, which offers a higher level of flexibility and automation than appears to be understood widely in the policy-making arena.

Note also that current automated inference techniques for stochastic simulation-based models (Toni et al., 2009; Kypraios et al., 2017; McKinley et al., 2018; Chatzilena et al., 2019; Minter and Retkute, 2019) are computationally demanding and are by their very nature approximate. The academic topic at the core of this paper and the subject of a significant fraction of our academic work (Paige et al., 2014; Wood et al., 2014; Tolpin et al., 2015; van de Meent et al., 2015; Paige and Wood, 2016; Rainforth et al., 2016; Rainforth et al., 2018; Naderiparizi et al., 2019; Zhou et al., 2019; Warrington et al., 2020) deals with this challenge. Furthermore, the basic structure of simulators currently available may lack important policy-influenceable interaction parameters that one might like to examine. If viewed solely in light of the provided examples, our contribution could reasonably be seen both as highlighting the utility of inference for planning in this application setting, and as automating the manual selection of promising policy parameters. The tools we showcase are capable of significantly more; however, for expediency and clarity, we have focused on control as inference, an application that has seen relatively little specific coverage in the literature, and

the simplest possible inference methods which do not require familiarity with the technical literature on approximate Bayesian inference. We leave other straightforward applications of automated inference tools in this application area, like parameter inference from observed outbreak data (Toni et al., 2009; Kypraios et al., 2017; McKinley et al., 2018; Chatzilena et al., 2019; Minter and Retkute, 2019), to others.

That being said, our hope is to inform field epidemiologists and policy-makers about an existing technology that could, right now, be used to support public policy planning towards more precise, potentially tailored interventions that ensure safety while also potentially leading to fewer economic ramifications. Fully probabilistic methods are apparently only relatively recently being embraced by the epidemiology community (Lessler et al., 2016; Funk and King, 2020), while the communities for approximate Bayesian inference and simulation-based inference have remained mostly focused on the tasks of parameter estimation and forecasting (Toni et al., 2009; Kypraios et al., 2017; McKinley et al., 2018), rather than control as inference. Beyond this demonstration, we hope to encourage timely and significant developments on the modeling side, and, if requested, to actually aid in the fight against COVID-19 by helping arm policy-makers with a new kind of tool and training them how to use it rapidly. Finally, we hope to engage the machine learning community in the fight against COVID-19 by helping translate between the specific epidemiological control problem and the more general control problem formulations on which we work regularly.

2 ASSUMPTIONS AND FINDINGS

We start with the assumption that the effectiveness of policy-making can be significantly improved by consulting the outputs of model-based tools which provide quantitative metrics for the ability of particular policy actions to achieve specific formalized goals. In particular, we imagine the following scenario. There exists some current population, and the health status of its constituents is only partially known. There exists a disease whose transmission properties may be only partially known, but whose properties cannot themselves be readily controlled. There exists a population dynamic that can be controlled in some limited ways at the aggregate level. There exists a “policy goal” or target which we will refer to as the allowable, allowed, or goal set of system trajectories. An example of this could be “the total number of infectious persons should not ever exceed some percentage of the population” or “the first date at which the total number of infectious persons exceeds some threshold is at least some number of days away.” We finally assume an implied set of allowable policy prescriptions, defined in the sense that population dynamics behaving according to such policies will be exactly the ones to attain the goal with high probability. In general, this set of allowable policies is intractable to compute exactly, motivating the use of automated tools implementing well established *approximate Bayesian inference* methods. We explicitly do not claim “completeness” of the stochastic dynamic models in any realistic sense, disregarding complications such as potential agent behaviour in strategic response to regionalized

policies, and do not attempt to quantify all costs and benefits of the considered policies, for example economic or cultural impacts. Rather, formulating the problem described above in terms of Bayesian inference results in a *posterior distribution* over policies which have been conditioned on satisfying the formalized policy desiderata within the formalized dynamical model. Effectively, this is to be understood as “scrutinizing,” “weighting,” “prioritizing,” or “focusing” potential policy actions for further consideration, rather than as an “optimal” prescription. When selecting a policy based on the posterior distribution, policy-makers are expected to account for additional, more complex socio-economic phenomena, costs and benefits using their own judgment.

The *only* things that may safely be taken away from this paper are the following:

- Existing compartmental models and agent-based simulators can be used as an aid for policy assessment via a Bayesian planning as inference formulation.
- Existing automated inference tools can be used to perform the required inferential computation.
- Opportunities exist for various fields to come together to improve both understanding of and availability of these techniques and tools.
- Further research and development into modeling and inference is recommended to be immediately undertaken to explore the possibility of more efficient, less economically devastating control of the COVID-19 pandemic.

What should *not* be taken away from this paper are any other conclusions, including in particular the following:

- Any conclusion or statements that there might exist less aggressive measures that could still be effective in controlling COVID-19.
- Any substantial novelty, uniqueness or performance claims about the particular numerical methods and software implementations for Bayesian inference which were used for the purpose of demonstrating the findings above. In particular, on the one hand, similar results could have been obtained using other software implementation strategies in principle, and on the other hand, more advanced inference methods could have been applied using the same software tools at the expense of rendering the conceptual exposition less accessible to audiences outside of the Bayesian inference community.

As scientists attempting to contribute “across the aisle,” we use more qualifying statements than usual throughout this work in an attempt to reduce the risk of misunderstandings and sensationalism.

3 APPROACH

In this section we formalize the policy-making task in terms of computing conditional probabilities in a probabilistic model of

the disease spread. While the technical description can get involved at times, we emphasize that in practice the probabilistic model is already defined by an existing epidemiological simulator and the probabilistic programming tools we describe in this paper provide the ability to compute the required conditional probabilities automatically, including automatically introducing required approximations, so the users only need to focus on identifying which variables to condition on, and feeding real-world data to the system. Readers familiar with framing planning as inference may wish to skip directly to **Section 4**.

Being able to perform probabilistic inference is crucial for taking full advantage of available simulators in order to design good policies. This is because in the real world, many of the variables crucially impacting the dynamics of simulations are not directly observable, such as the number of infectious but asymptomatic carriers or the actual number of contacts people make in different regions each day. These variables are called *latent*, as opposed to *observable* variables such as the number of deaths or the number of passengers boarding a plane each day, which can often be directly measured with high accuracy. It is often absolutely crucial to perform inference over some latent variables to achieve reliable forecasts. For example, the future course of an epidemic like COVID-19 is driven by the number of people currently infected, rather than the number of people currently hospitalized, while in many countries in the world currently only the latter is known.

While performing inference over latent variables is very broadly applicable, the scenario described above being but one example, in this paper we primarily address the problem of choosing good policies to reduce the impact of an epidemic which can also be formulated as an inference problem. This choice of problem was driven by the hypothesis that the search for effective controls may not in fact be particularly well-served by automation at the current time. In the epidemiological context, the questions we are trying to answer are ones like “when and for how long do we need to close schools to ensure that we have enough ventilators for everyone who needs them?” While obviously this is overly simplistic and many different policy decisions need to be enacted in tandem to achieve good outcomes, we use this example to illustrate tools and techniques that can be applied to problems of realistic complexity.

Our approach is not novel, it has been studied extensively under the name “control via planning as inference” and is now well understood (Todorov, 2008; Toussaint, 2009; Kappen et al., 2012; Levine, 2018). What is more, the actual computations that result from following the recipes for planning as inference can be, in some cases readily, manually replicated. Again, our aim here is to inform or remind a critically important set of policy-makers and modellers that these methodologies are extremely relevant to the current crisis. Moreover, at least partial automation of model-informed policy-guidance is achievable using existing tools, and, may even lead to sufficient computational savings to make their use in current policy-making practical. Again, our broader hope here is to encourage rapid collaborations leading to more targeted and less-economically-devastating policy recommendations.

3.1 An Abstract Epidemiological Dynamics Model

In this work we will look at both compartmental and agent-based models. An overview of these specific types of models appears later. For the purposes of understanding our approach to planning as inference, it is helpful to describe the planning as inference problem in a formalism that can express both types of models. The approach of conducting control via planning as inference follows a general recipe:

- 1) Define the latent and control parameters of the model and place a prior distribution over them.
- 2) Either or both define a likelihood for the observed disease dynamics data and design constraints that define acceptable disease progression outcomes.
- 3) Do inference to generate a posterior distribution on control values that conditions both on the observed data and the defined constraints.
- 4) Make a policy recommendation by picking from the posterior distribution consisting of effective control values according to some utility maximizing objective.

We focus on steps 1–3 of this recipe, and in particular do not explore simultaneous conditioning. We ignore the observed disease dynamics data and focus entirely on inference with future constraints. We explain the rationale behind these choices near the end of the paper.

Very generally, an epidemiological model consists of a set of global parameters and time dependent variables. Global parameters are (θ, η) , where θ denotes parameters that can be controlled by policy directives (e.g. close schools for some period of time or decrease the general level of social interactions by some amount), and η denotes parameters which cannot be affected by such measures (e.g. the incubation period or fatality rate of the disease).

The time dependent variables are (X_t, Y_t, Z_t) and jointly they constitute the full state of the simulator. X_t are the latent variables we are doing inference over (e.g. the total number of infected people or the spatio-temporal locations of outbreaks), Y_t are the observed variables whose values we obtain by measurements in the real world (e.g. the total number of deaths or diagnosed cases), and Z_t are all the other latent variables whose values we are not interested in knowing (e.g. the number of contacts between people or hygiene actions of individuals). For simplicity, we assume that all variables are either observed at all times or never, but this can be relaxed.

The time t can be either discrete or continuous. In the discrete case, we assume the following factorization

$$p(\theta, \eta, X_{0:T}, Y_{0:T}, Z_{0:T}) = p(\theta)p(\eta)p(X_0, Y_0, Z_0 | \theta, \eta) \cdot \prod_{t=1}^T p(X_t, Y_t, Z_t | X_{0:t-1}, Y_{0:t-1}, Z_{0:t-1}, \theta, \eta). \quad (1)$$

Note that we do not assume access to any particular factorization between observed and latent variables. We assume that a priori the controllable parameters θ are independent of non-controllable parameters η to avoid

situations where milder control measures are associated with better outcomes because they tend to be deployed when the circumstances are less severe, which would lead to erroneous conclusions when conditioning on good outcomes in **Section 3.3**.

3.2 Inference

The classical inference task (Toni et al., 2009; Kypraios et al., 2017; McKinley et al., 2018; Chatzilena et al., 2019; Minter and Retkute, 2019) is to compute the following conditional probability

$$p(\eta, X_{0:T} | Y_{0:T}, \theta) = \int p(\eta, X_{0:T}, Z_{0:T} | Y_{0:T}, \theta) dZ_{0:T}. \quad (2)$$

In the example given earlier X_t would be the number of infected people at time t and Y_t would be the number of hospitalized people at time t . If the non-controllable parameters η are known they can be plugged into the simulator, otherwise we can also perform inference over them, like in the equation above. This procedure automatically takes into account prior information, in the form of a model, and available data, in the form of observations. It produces estimates with appropriate amount of uncertainty depending on how much confidence can be obtained from the information available.

The difficulty lies in computing this conditional probability, since the simulator does not provide a mechanism to sample from it directly and for all but the simplest models the integral cannot be computed analytically. The main purpose of probabilistic programming tools is to provide a mechanism to perform the necessary computation automatically, freeing the user from having to come up with and implement a suitable algorithm. In this case, approximate Bayesian computation (ABC) would be a suitable tool. We describe it below, emphasizing again that its implementations are already provided by existing tools (Toni et al., 2009; Kypraios et al., 2017; McKinley et al., 2018; Chatzilena et al., 2019; Minter and Retkute, 2019).

The main problem in this model is that we do not have access to the likelihood $p(Y_t | X_t, \theta, \eta)$ so we cannot apply the standard importance sampling methods. To use ABC, we extend the model with auxiliary variables $Y_{0:T}^{obs}$, which represent the actual observations recorded, and use a suitably chosen synthetic likelihood $p(Y_t^{obs} | Y_t)$, often Gaussian. Effectively, that means we're solving the following inference problem,

$$p(X_{0:T} | Y_{0:T}^{obs}, \theta) = \iiint p(\eta, X_{0:T}, Y_{0:T}, Z_{0:T} | Y_{0:T}^{obs}, \theta) dY_{0:T} dZ_{0:T} d\eta, \quad (3)$$

which we can solve by importance sampling from the prior. Algorithmically, this means independently sampling a large number N of trajectories from the simulator

$$\left(\eta^{(i)}, X_{0:T}^{(i)}, Y_{0:T}^{(i)}, Z_{0:T}^{(i)} \right) \stackrel{i.i.d.}{\sim} p(\eta, X_{0:T}, Y_{0:T}, Z_{0:T} | \theta) \quad \text{for } i \in \{1, \dots, N\}, \quad (4)$$

computing their importance weights

$$w_i = \frac{p(Y_{0:T}^{obs} | Y_{0:T}^{(i)})}{\sum_{j=1}^N p(Y_{0:T}^{obs} | Y_{0:T}^{(j)})}, \quad (5)$$

and approximating the posterior distribution

$$p(X_{0:T} | Y_{0:T}^{obs}, \theta) \approx \hat{p}(X_{0:T} | Y_{0:T}^{obs}, \theta) = \sum_{i=1}^N w_i \delta_{X_{0:T}^{(i)}}(X_{0:T}), \quad (6)$$

where δ is the Dirac delta putting all the probability mass on the point in its subscript. In more intuitive terms, we are approximating the posterior distribution with a collection of weighted samples where weights indicate their relative probabilities.

3.3 Control as Inference: Finding Actions That Achieve Desired Outcomes

In traditional inference tasks we condition on data observed in the real world. In order to do control as inference, we instead condition on what we *want* to observe in the real world, which tells us which actions are likely to lead to such observations. This is accomplished by introducing auxiliary variables that indicate how desirable a future state is or is not. In order to keep things simple, here we restrict ourselves to the binary case where $Y_t \in \{0, 1\}$, where 1 means that the situation at time t is acceptable and 0 means it is not. This indicates which outcomes are acceptable, allowing us to compute a distribution over those policies, while leaving the choice of which specific policy is likely to produce an acceptable outcome to policy-maker(s). For example, Y_t can be 1 when the number of patients needing hospitalization at a given time t is smaller than the number of hospital beds available and 0 otherwise.

To find a policy θ that is likely to lead to acceptable outcomes, we need to compute the posterior distribution

$$p(\theta | \forall_{t>0}: Y_t = 1). \quad (7)$$

Once again, probabilistic programming tools provide the functionality to compute this posterior automatically. In this case, rejection sampling would be a simple and appropriate inference algorithm. The rejection sampling algorithm repeatedly samples values of θ from the prior $p(\theta)$, runs the full simulator using θ , and keeps the sampled θ only if all Y_t are 1. The collection of accepted samples approximates the desired posterior. We use rejection sampling in our agent-based modeling experiments, but emphasize that other, more complex and potentially more computationally efficient, approaches to computing this posterior exist.

This tells us which policies are most likely to lead to a desired outcome but not how likely a given policy is to lead to that outcome. To do that, we can evaluate the conditional probability $p(\forall_{t>0}: Y_t = 1 | \theta)$, which is known as the model evidence, for a particular θ . A more sophisticated approach would be to condition on the policy leading to a desired outcome with a given probability p_0 , that is

$$p(\theta | p(\forall_{t>0}: Y_t = 1 | \theta) > p_0). \quad (8)$$

For example, we could set $p_0 = 0.95$ to find a policy that ensures availability of hospital beds for everyone who needs one with at least 95% probability. The conditional probability in **Eq. 8** is more difficult to compute than the one in **Eq. 7**. It can be approximated by methods such as nested Monte Carlo (NMC) (Rainforth et al., 2018), which are natively available in advanced probabilistic programming systems such as Anglican (Tolpin et al., 2016) and WebPPL (Goodman and Stuhlmüller, 2014a) but in specific

cases can also be implemented on top of other systems, such as PyProb (Le et al., 2017), with relatively little effort, although using NMC usually has enormous computational cost.

To perform rejection sampling with nested Monte Carlo, we first draw a sample $\theta_i \sim p(\theta)$, then draw N samples of $Y_{0:T}^{(j)} \sim p(Y_{0:T} | \theta_i)$ and reject θ_i if fewer than $p_0 N$ of sampled sequences of Y s are all 1s, otherwise we accept it. This procedure is continued until we have a required number K of accepted θ s. For sufficiently high values of N and K , this algorithm approximates the posterior distribution Eq. 8 arbitrarily well.

However we compute the posterior distribution, it contains multiple values of θ that represent different policies that, if implemented, can achieve the desired result. In this setup it is up to the policy-maker(s) to choose a policy θ^* that has support under the posterior, i.e., yields the desired outcomes, taking into account some notion of utility.

Crucially, despite their relative simplicity, the rejection sampling algorithms we have discussed evaluate randomly sampled values of θ . This is a fundamental difference from the commonly used grid search over a deterministic array of parameter values. In practical terms, this is important because “well distributed” random samples are sufficient for experts to gauge the quantities of interest, and avoid grid searches that would be prohibitively expensive for θ with more than a few dimensions.

3.4 Stochastic Model Predictive Control: Reacting to What’s Happened

During an outbreak governments continuously monitor and assess the situation, adjusting their policies based on newly available data. A convenient theoretical and general framework to formalize this is that of model predictive control (Camacho and Alba, 2013). In this case, Y_t consists of variables Y_t^{data} that can be measured as the epidemic unfolds (such as the number of deaths) and the auxiliary variables Y_t^{aux} that indicate whether desired goals were achieved, just like in Section 3.3. Say that at time $t = 0$ the policy-maker(s) choose a policy to enact θ_0^* based on the posterior distribution

$$\begin{aligned} p(\theta | Y_0^{data}, \forall_{t>0}: Y_t^{aux} = 1, \eta) \\ = \iint p(\theta | X_0, Z_0, Y_0^{data}, \forall_{t>0}: Y_t^{aux} = 1, \eta) \hat{p}(X_0, Z_0 | Y_0^{data}, \eta) dX_0 dZ_0. \end{aligned} \quad (9)$$

Then at time $t = 1$ they will have gained additional information Y_1^{data} , leading to a new belief over the past and current states that we denote as $\hat{p}(X_{0:1}, Z_{0:1} | Y_{0:1}^{data}, \theta_0^*, \eta)$, for which we give a formula in the general case in Eq. 11. The policy-maker(s) then choose the policy θ_1^* from the posterior distribution $p(\theta | Y_{0:1}^{data}, \forall_{t>1}: Y_t^{aux} = 1, \theta_0^*, \eta)$.

Generally, at time t we compute the posterior distribution conditioned on the current state and on achieving desirable outcomes in the future

$$\begin{aligned} p(\theta | Y_{0:t}^{data}, \forall_{t'>t}: Y_{t'}^{aux} = 1, \theta_{0:t-1}^*, \eta) \\ = \iint p(\theta | X_{0:t}, Z_{0:t}, Y_{0:t}^{data}, \forall_{t'>t}: Y_{t'}^{aux} = 1, \theta_{0:t-1}^*, \eta) \\ \hat{p}(X_{0:t}, Z_{0:t} | Y_{0:t}^{data}, \theta_{0:t-1}^*, \eta) dX_{0:t} dZ_{0:t}. \end{aligned} \quad (10)$$

Policy-maker(s) then can use this distribution to choose the policy θ_t^* that is actually enacted. The current belief state is computed by inference

$$\begin{aligned} \hat{p}(X_{0:t}, Z_{0:t} | Y_{0:t}^{data}, \theta_{0:t-1}^*, \eta) \propto p(X_t, Z_t, Y_t^{data} | X_{0:t-1}, Z_{0:t-1}, Y_{0:t-1}^{data}, \theta_{0:t-1}^*, \eta) \\ \hat{p}(X_{0:t-1}, Z_{0:t-1} | Y_{0:t-1}^{data}, \theta_{0:t-2}^*, \eta). \end{aligned} \quad (11)$$

Eq. 11 can be computed using methods described in Section 3.2, while Eq. 10 can be computed using methods described in Section 3.3. As such, the policy enacted evolves over time, as a result of re-solving for the optimal control based on new information.

We note that what we introduce here as MPC may appear to be slightly different from what is commonly referred to as *model predictive control* (García et al., 1989). Firstly, instead of solving a finite-dimensional optimization problem over controls at each step, we perform a Bayesian update step and sample from the resulting posterior distribution over controls. Secondly, in more traditional applications of MPC a *receding horizon* (Jacob et al., 2011) is considered, where a finite and fixed-length window is considered. The controls required to satisfy the constraint over that horizon are then solved for and applied—without consideration of timesteps beyond this fixed window. At the next time step, the controls are then re-solved for. In this work, we rather consider a constant policy for the remaining $T - t$ time steps, as opposed to allowing a variable policy (discussed below) over a fixed window. We note that time-varying controls are fully permissible under the framework we present, as we demonstrate in Section 3.5. Furthermore, we could easily consider a fixed horizon, just by changing the definition of $Y_{t'}^{aux}$ to be equal to 1 for $t' > t + p$, where p is the length of the window being considered. Both of these extensions are provisioned for under the framework we provide, and could be used to implement what might be considered as more conventional stochastic MPC, but with all the auxiliary benefits of Bayesian inference.

Before we proceed, it is critical to note that the posterior $p(\theta | \forall_{t>0}: Y_t^{aux} = 1)$ describes the probability, given the dynamic model and the currently available information, for the parameters θ to achieve the policy goals. As such, our approach provides a method for “screening” policies such that candidate policies that achieve the required outcomes are found. What it does not take into account is the relative “cost” of each policy, or, how best to achieve the required parameter values. For instance, in the SEIR examples we present later, we learn the reduction in transmission required to avoid exceeding an infection threshold. The reduction in transmission can be achieved (for instance) by encouraging hand-washing and social distancing. While the method will identify whether or not a particular level of transmission will achieve the desired outcome, it does not indicate the optimal trade-off between the cost of a particular policy, the amount that the policy exceeds the required outcome, and the probability of the outcome. A human must therefore select, from the set of compliant policies identified, the policy that maximizes utility while minimizing cost. Therefore, this method is an important component of a wider policy making toolkit, as opposed to an oracle that can dictate the optimal policy decisions.

3.5 Time-Varying Control: Long Term Planning

It is also possible to explicitly model changing policy decisions over time, which enables more detailed planning, such as when to enact certain preventive measures such as closing schools. Notationally, this means instead of a single θ there is a separate θ_t for each time t . We can then find a good sequence of policy decisions by performing inference just like in **Section 3.3** by conditioning on achieving the desired outcome

$$p(\theta_{0:T} \mid \forall_{t>0}: Y_t = 1). \quad (12)$$

The inference problem is now more challenging, since the number of possible control sequences grows exponentially with the time horizon. Still, the posterior can be efficiently approximated with methods such as Sequential Monte Carlo.

It is straightforward to combine this extension with model predictive control from **Section 3.4**. The only required modification is that in **Eq. 10** we need to compute the posterior over all future policies,

$$\begin{aligned} & p(\theta_{t:T} \mid Y_{0:t}^{data}, \forall_{t'>t}: Y_{t'}^{aux} = 1, \theta_{0:t-1}^*) \\ &= \iint p(\theta_{t:T} \mid X_{0:t}, Z_{0:t}, Y_{0:t}^{data}, \forall_{t'>t}: Y_{t'}^{aux} = 1, \theta_{0:t-1}^*) \\ & \quad \hat{p}(X_{0:t}, Z_{0:t} \mid Y_{0:t}^{data}, \theta_{0:t-1}^*, \eta) dX_{0:t} dZ_{0:t}. \end{aligned} \quad (13)$$

At each time t the policy-maker(s) only choose the current policy θ_t^* , without committing to any future choices. This combination allows for continuously re-evaluating the situation based on available data, while explicitly planning for enacting certain policies in the future. While here we explicitly consider only open-loop control policies, this re-planning allows new information to be taken into account, and facilitates reactionary policy decisions to evolving situations.

In models with per-timestep control variables θ_t , it is very important that in the model (but not in the real world) the enacted policies must not depend on anything else. If the model includes feedback loops for changing policies based on the evolution of the outbreak, it introduces positive correlations between lax policies and low infection rates (or other measures of severity of the epidemic), which in turn means that conditioning on low infection rates is more likely to produce lax policies. This is a known phenomenon of reversing causality when correlated observational data is used to learn how to perform interventions (Pearl, 2000).

We note another potential pitfall that may be exacerbated by time-varying control: if the model used does not accurately reflect reality, any form of model-based control is likely to lead to poor results. To some extent, this can be accounted for with appropriate distributions reflecting uncertainty in parameter values. However, given the difficulty of modeling human behavior, and especially of modeling people's reactions to novel policies, there is likely to be some mismatch between modeled and real behavior. To give an example of a possible flaw in an agent-based model: if a proposed fine-grained policy closed one park while keeping open a second, nearby, park, the model may not account for the likely increase in visitors to the

second park. The larger and more fine-grained (in terms of time or location) the space of considered policies is, the more such deficiencies may exist. We therefore recommend that practitioners restrict the space of considered policies to those which are likely to be reasonably well modeled.

3.6 Automation

We have intentionally not really explained how one might actually computationally characterize any of the conditional distributions defined in the preceding section. For the compartmental models that follow, we provide code that directly implements the necessary computations. Alternatively we could have used the automated inference facilities provided by any number of packages or probabilistic programming systems. Performing inference as described in existing, complex simulators is much more complex and not nearly as easy to implement from scratch. However, it can now be automated using the tools of probabilistic programming.

3.6.1 Probabilistic Programming

Probabilistic programming (van de Meent et al., 2018) is a growing subfield of machine learning that aims to build an analogous set of tools for automating inference as automatic differentiation did for continuous optimization. Like the gradient operator of languages that support automatic differentiation, probabilistic programming languages introduce observe operators that denote conditioning in the probabilistic or Bayesian sense. In those few languages that natively support nested Monte Carlo (Goodman and Stuhlmüller, 2014b; Tolpin et al., 2016), language constructs for defining conditional probability objects are introduced as well. Probabilistic programming languages (PPLs) have semantics (Staton et al., 2016) that can be understood in terms of Bayesian inference (Bishop, 2006; Gelman et al., 2013; Ghahramani, 2015). The major challenge in designing useful PPL systems is the development of general-purpose inference algorithms that work for a variety of user-specified programs. The work in the paper uses only the very simplest, and often least efficient, general purpose inference algorithms, importance sampling and rejection sampling. Others are covered in detail in (van de Meent et al., 2018).

Of all the various probabilistic programming systems, only one is readily compatible with inference in existing stochastic simulators: PyProb (Le et al., 2019). Quoting from its website¹ “PyProb is a PyTorch-based library for probabilistic programming and inference compilation. The main focus of PyProb is on coupling existing simulation codebases with probabilistic inference with minimal intervention.” A textbook, technical description of how PyProb works appears in (van de Meent et al., 2018, Chapt. 6).

Recent examples of its use include inference in the standard model of particle physics conditioning on observed detector outputs (Baydin et al., 2018; Baydin et al., 2019), inference about internal composite material cure processes through a composite material cure simulator conditioned on observed

¹<https://github.com/pyprob/pyprob>

surface temperatures (Munk et al., 2020), and inference about malaria spread in a malaria simulator (Gram-Hansen et al., 2019).

3.6.2 Alternative Approaches

It is important to note upfront that our ultimate objective is to understand the dependence of the policy outcome on the controllable parameters (for instance, as defined by Eqs 7, 8). There are a litany of methods to quantify this dependency. Each method imposes different constraints on the model family that can be analyzed, and the nature of the solution obtained. For instance, the fully Bayesian approach we take aims to quantify the entire distribution, whereas traditional *optimal control* methods may only seek a pointwise maximizer of this probability. Therefore, before we introduce the models we analyze and examples we use to explore the proposed formulation, we give a brief survey of alternative approaches one could use to solve this problem, and give the benefits and drawbacks of our proposed formulation compared to these approaches.

The traditional toolkits used to analyze problems of this nature are *optimal control* (Vinter and Vinter, 2010; Lewis et al., 2012; Sharomi and Malik, 2017; Bertsekas, 2019) and *robust control* (Zhou and Doyle, 1998; Hansen and Sargent, 2001; Green and Limebeer, 2012). Most generally, optimal control solves for the control inputs (here denoted θ) such that some outcome is achieved (here denoted $\forall_{t>0}: Y_t^{aux} = 1$) while minimizing the associated cost of the controls. An example of this may be successfully maintaining stable flight while minimizing fuel expenditure. However, for the controls to be optimal, this assumes the model is correct. Alternatively, robust control does not assume that the model is correct, and instead solves for controls that are maximally robust to the uncertainty in the model, while still achieving the desired outcome.

Many traditional control approaches can exploit a mathematical model of the system to solve for controls that consider multiple timesteps, referred to as *model predictive control* (MPC) (Morari and Lee, 1999; Rawlings, 2000; Camacho and Alba, 2013). Alternatively, control methods can be *model-free*, where there is no notion of the temporal dynamics of the system being controlled, such as the canonical PID control. Model-based methods often solve for more effective control measures with lower overall computational costs, by exploiting the information encoded in the model. However, model predictive control methods are only applicable when an (accurate) model is available.

One could re-frame the objective as trying to maximize the *reward* of some applied controls. Here, reward may be a 0–1 indicator corresponding to whether or not the hospital capacity was exceeded at any point in time. The reward may also be more sophisticated, such as by reflecting the total number of people requiring hospital treatment, weighted by how critical each patient was. Reward may also internalize some notion of the *cost* of a particular control. Shutting workplaces may prevent the spread of infection, but has economic and social implications that (at least partially) counteract the benefit of reducing the infection rate. This can be implemented as earning progressively more negative reward for stronger policies.

This type of analysis is formalized through *reinforcement learning* (RL) (Sutton, 1992; Kaelbling et al., 1996; Sutton and Barto, 2018; Bertsekas, 2019; Warrington et al., 2021). In RL, a *policy*, here θ , is solved for that maximizes the expected reward. RL is an incredibly powerful toolkit that can learn highly expressive policies that vary as a function of time, state, different objectives etc. RL methods can also be described as model-based (Moerland et al., 2021) and model-free (Bertsekas, 2019) analogously to traditional control methods. While we do not delve into RL in detail here, it is important to highlight as an alternative approach, and we refer the reader to Levine (2018) for a more detailed introduction of RL, to Bertsekas (2019) for a discussion of the relationship between RL and control methods, and to Levine (2018) for a discussion of the relationship between RL and planning-as-inference.

However, a critical drawback of both control and RL is the strong dependence on the loss functions or specific models considered, both in terms of convergence and the optimal solution recovered (Adam and DeJong, 2003). This dependence may therefore mean that the solution recovered is not representative of the “true” optimal solution. Furthermore, RL approaches can be incredibly expensive to train, especially in high-dimensional time series models, with a large array of controls and sparse rewards. More traditional control-based methods can be limited in the models that they can analyze, both in terms of the transition distribution and the range of controls that can be considered. Finally, both RL and control-based approaches offer limited insight into *why* certain controls were proposed. This can limit the ability of researchers and modellers to expediently investigate the emergent properties of complex, simulation-based models. Furthermore, this lack of transparency may be disadvantageous for policy-makers, who may be required to justify why certain policy decisions were made, and the confidence with which that policy decision was believed to be optimal.

All of these methods, at least in terms of their core intentions, are very similar. Different methods impose different restrictions on the models that can be analyzed, and the nature of the solution obtained. In this work, we take the approach of framing control and planning as fully Bayesian inference with a binary objective. We choose this formulation as it places very few restrictions on the model class that can be analyzed, allows uncertainty to be succinctly handled throughout the inference, a broad range of objectives to be defined under the same framework, and correctly calibrated posterior distributions over all random variables in the model to be recovered. Furthermore, this formulation allows us to access the ever-growing array of powerful inference algorithms to perform the inference. While naive approaches such as grid-search or random sampling may be performant in low-dimensional applications, they do not scale to high dimensions through the *curse of dimensionality* (Köppen, 2000; Owen, 2013). Therefore, we focus on methods that can scale to high-dimensional applications (Owen, 2013). Finally, Bayesian inference methods are, arguably, the most complete formulation. Once the full posterior or joint distribution have been recovered, many different methodologies, constraints, objectives etc. can be formulated and evaluated *post facto*. This

flexibility allows for fine-grained analysis to be conducted on the inferred distributions to provide analysts with a powerful and general tool to further analyze and understand the model. Less myopically, this analysis and the understanding garnered may be more beneficial to our broader understanding of outbreaks and the feasibility and efficacy of particular responses.

Finally, we note that while throughout the experiments presented in **Section 5.2** we use the PyProb framework, PyProb is not the *only* framework available. Any inference methodology could be used to compute the probability in **Eq. 7**. However, PyProb is a natural choice as it allows for greater flexibility in interfacing between black-box simulator code, and powerful and efficiently implemented inference algorithms, without modifying either simulator or inference code. This reduces the implementation overhead to the analyst, and reduces the scope for implementation bugs and oversights to be introduced.

4 MODELS

Epidemiological dynamics models can be used to describe the spread of a disease such as COVID-19 in society. Different types span vastly different levels of fidelity. There are classical compartmental models (SIR, SEIR, etc.) (Blackwood and Childs, 2018a) that describe the bulk progression of diseases of different fundamental types. These models break the population down to a series of compartments (e.g. susceptible (S), infectious (I), exposed (E), and recovered (R)), and treat them as continuous quantities that vary smoothly and deterministically over time following dynamics defined by a particular system of ordinary differential equations. These models are amenable to theoretical analyses and are computationally efficient to forward simulate owing to their low dimensionality. As policy-making tools, they are rather blunt unless the number of compartments is made large enough to reflect different demographic information (age, socio-economic info, etc.), spatial strata, or combinations thereof.

At the other end of the spectrum are agent-based models (Hunter et al., 2017; Badham et al., 2018; Hunter et al., 2018; Tracy et al., 2018) (like the University of Pittsburgh Public Health Dynamics Laboratory's FRED (Grefenstette et al., 2013) or the Institute for Disease Modeling's EMOD (Anna et al., 2018)) that model populations and epidemiological dynamics via discrete agent interactions in "realistic" space and time. Imagine a simulation environment like the game *Sim-CityTM*, where the towns, populations, infrastructure (roads, airports, trains, etc.), and interactions (go to work, school, church, etc.) are modeled at a relatively high level of fidelity. These models exist only in the form of stochastic simulators, i.e., software programs that can be initialized with disease and population characteristics, then run forward to show in a more fine-grained way the spread of the disease over time and space.

Both types of models are useful for policy-making. Compartmental models are usually more blunt unless the number of compartments is very high and it is indexed by spatial location, demographics and age categories. Increasing the number of compartments adds more unknown parameters

which must be estimated or marginalized. Agent-based models are complex by nature, but they may be more statistically efficient to estimate, as they are parameterized more efficiently, often directly in terms of actual individual and group behaviour choices. In many cases, predictions made by such models are more high fidelity, certainly more than compartmental models with few compartments, and this has implications for their use as predictive tools for policy analysis. For instance, policies based on simulating a single county in North Dakota with excellent hospital coverage and a highly dispersed, self-sufficient population could lead to different intervention recommendations compared to a compartmental model of the whole of the United States with only a few compartments.

4.1 A Compartmental Model of COVID-19

We begin by introducing a low-dimensional compartmental model to explore our methods in a well-known model family, before transitioning to a more complex agent-based simulator. The model we use is an example of a classical SEIR model (Ogilvy Kermack and McKendrick, 1927; C Blackwood and Childs, 2018; Hill et al., 2020). In such models, the population is subdivided into a set of compartments, representing the susceptible (uninfected), exposed (infected but not yet infectious), infectious (able to infect/expose others) and recovered (unable to be infected). Within each compartment, all individuals are treated identically, and the full state of the simulator is simply the size of the population of each compartment. Our survey of the literature found a lack of consensus about the compartmental model and parameters which most faithfully simulate the COVID-19 scenario. Models used range from standard SEIR (Liu et al., 2020a; Massonnaud et al., 2020; Rovetta et al., 2020; Traini et al., 2020), SIR (Jia et al., 2020; Pujari and Shekatkar, 2020; Teles, 2020; Traini et al., 2020; Weber et al., 2020), SIRD (Anastassopoulou et al., 2020; Liu et al., 2020b; Caccavo, 2020), QSEIR (Liu et al., 2020c), and SEAIHRD (Arenas et al., 2020). The choice depends on many factors, such as how early or late in the stages of an epidemic one is, what type of measures are being simulated, and the availability of real word data. We opted for the model described in this section, which seems to acceptably represent the manifestation of the disease in populations. Existing work has investigated parameter estimation in stochastic SEIR models (Lekone and Finkenstädt, 2006; Roberts et al., 2015). Although we will discuss how we set the model parameters, we emphasize that our contribution is instead in demonstrating how a calibrated model could be used for planning.

4.1.1 Model Description

We use an SEI³R model (Hill et al., 2020), a variation on the standard SEIR model which allows additional modeling freedom. It uses six compartments: susceptible (S), exposed (E), infectious with mild (I_1), severe (I_2) or critical infection (I_3), and recovered (R). We do not include baseline birth and death rates in the model, although there is a death rate for people in the critically infected compartment. The state of the simulator at time $t \in [0, T]$ is $X_t = \{S_t, E_t, I_{1,t}, I_{2,t}, I_{3,t}, R_t\}$ with $S_t, E_t, I_{1,t}, I_{2,t}, I_{3,t}$ and R_t indicating the population sizes (or

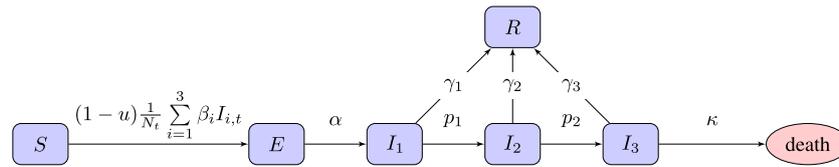


FIGURE 1 | Flow chart of the SEI³R model we employ. A member of the susceptible population *S* moves to exposed *E* after being exposed to an infectious person, where “exposure” is defined as the previous susceptible person contracting the illness. After some incubation period, a random duration parameterized by α , they develop a mild infection (I_1). They may then either recover, moving to *R*, or progress to a severe infection (I_2). From I_2 , they again may recover, or else progress further to a critical infection (I_3). From I_3 , the critically infected person will either recover or die.

proportions) at time t . The unknown model parameters are $\eta = \{\alpha, \beta_1, \beta_2, \beta_3, p_1, p_2, \gamma_1, \gamma_2, \gamma_3, \kappa\}$, each with their own associated prior. To the model we add a free control parameter, denoted $u \in [0, 1]$, that acts to reduce the transmission of the disease. Since u is the only free parameter, $\theta = u$. An explanation of u is given later in the text. There are no internal latent random variables (Z_t) in this model. In this paper we do not demonstrate inference about θ given Y^{obs} within this model, and so do not consider Y^{obs} here. We do, however, consider Y^{aux} to perform policy selection, and discuss the form of Y^{aux} later. Defining the total live population (i.e., the summed population of all compartments) at time t to be N_t , the dynamics are given by the following equations, and shown in **Figure 1**.

$$\frac{d}{dt}S_t = -(1 - u) \frac{1}{N_t} \sum_{i=1}^3 \beta_i I_{i,t} S_t \quad (14)$$

$$\frac{d}{dt}E = (1 - u) \frac{1}{N_t} \sum_{i=1}^3 \beta_i I_{i,t} S_t - \alpha E_t \quad (15)$$

$$\frac{d}{dt}I_{1,t} = \alpha E_t - p_1 I_{1,t} - \gamma_1 I_{1,t} \quad (16)$$

$$\frac{d}{dt}I_{2,t} = p_1 I_{1,t} - p_2 I_{2,t} - \gamma_2 I_{2,t} \quad (17)$$

$$\frac{d}{dt}I_{3,t} = p_2 I_{2,t} - \kappa I_{3,t} - \gamma_3 I_{3,t} \quad (18)$$

$$\frac{d}{dt}R_t = \gamma_1 I_{1,t} + \gamma_2 I_{2,t} + \gamma_3 I_{3,t}. \quad (19)$$

For the purposes of simulations with this model, we initialize the state with 0.01% of the population having been exposed to the infection, and the remaining 99.99% of the population being susceptible. The population classified as infectious and recovered are zero, i.e., $X_0 = \{0.9999, 0.0001, 0, 0, 0, 0\}$ and $N_t = 1$.

4.1.2 Example Trajectories

Before explaining how we set the SEI³R model parameters, or pose inference problems in the model, we first verify that we are able to simulate feasible state evolutions. As we will describe later, we use parameters that are as reflective of current COVID-19 epidemiological data as possible at the time of writing. **Figures 2A,B** show deterministic simulations from the model with differing control values u . Shown in green is the susceptible population, in blue is the exposed population, in red is the infectious population, and in purple is the recovered population. The total live population is shown as a black dotted line. All populations are normalized by the initial total population, N_0 . The dashed black line represents a threshold

under which we wish to keep the number of infected people at all times. The following paragraph provides the rationale for this goal.

4.1.3 Policy Goal

As described in **Section 3.4**, parameters should be selected to ensure that a desired goal is achieved. In all scenarios using the SEI³R model, we aim to maintain the maximal infectious population proportion requiring healthcare below the available number of hospital beds per capita, denoted C . This objective can be formulated as an auxiliary observation, $Y_{0:T}^{aux}$, introduced in **Section 3**, as:

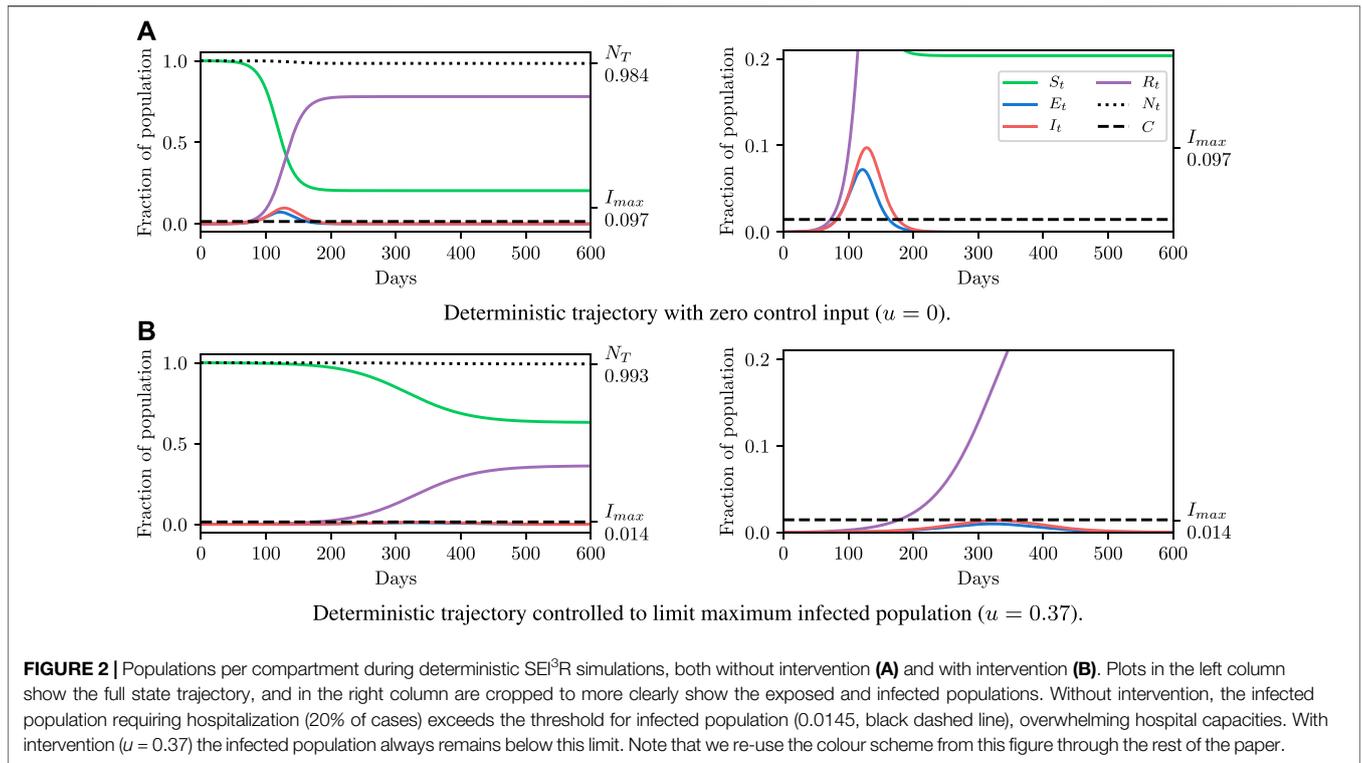
$$Y_{0:T}^{aux} = \mathbb{I} \left[\left(\max_{t \in \{0, \dots, T\}} (I_{1,t} + I_{2,t} + I_{3,t}) \right) < C \right], \quad (20)$$

where $I_{1,0:T}$, $I_{2,0:T}$ and $I_{3,0:T}$ are sampled from the model, conditioned on a θ value. This threshold value we use was selected to be 0.0145, as there are 0.0029 hospital beds per capita in the United States (World Bank, 2020), and roughly 20% of COVID-19 cases require hospitalization. This constraint was chosen to represent the notion that the healthcare system must have sufficient capacity to care for all those infected who require care, as opposed to just critical patients. However, this constraint is only intended as a demonstrative example of the nature of constraints and inference questions one can query using models such as these, and under the formalism used here, implementing and comparing inferences under different constraints is very straightforward. More complex constraints may account for the number of critical patients differently to those with mild and severe infections, model existing occupancy or seasonal variations in capacity, or, target other metrics such as the number of deceased or the duration of the epidemic.

The constraint is not met in **Figure 2A**, but is in **Figure 2B**, where a greater control input u has been used to slow the spread of the infection. This is an example of the widely discussed “flattening of the curve.” As part of this, the infection lasts longer but the death toll is considerably lower.

4.1.4 Control Input

As noted before, we assume that only a single “controllable” parameter affects our model, u . This is the reduction in the “baseline reproductive ratio,” R_0 , due to policy interventions. Increasing u has the same effect as reducing the infectiousness parameters β_1, β_2 and β_3 by the same proportion. u can be interpreted as the effectiveness of policy choices to prevent new infections. Various policies could serve to increase u ,



since it is a function of both, for example, reductions in the “number of contacts while infectious” (which could be achieved by social distancing and isolation policy prescriptions), and the “probability of transmission per contact” (which could be achieved by, e.g., eye, hand, or mouth protective gear policy prescriptions). It is likely that both of these kinds of reductions are necessary to maximally reduce u at the lowest cost.

For completeness, the baseline reproductive ratio, R_0 , is an estimate of the number of people a single infectious person will in turn infect and can be calculated from other model parameters (Hill et al., 2020). R_0 is often reported by studies as a measure of the infectiousness of a disease, however, since R_0 can be calculated from other parameters we do not explicitly parameterize the model using R_0 , but we will use R_0 as a convenient notational shorthand. We compactly denote the action of u as controlling the baseline reproductive rate to be a “controlled reproductive rate,” denoted \hat{R}_0 , and calculated as $\hat{R}_0 = (1 - u)R_0$. This is purely for notational compactness and conceptual ease, and is entirely consistent with the model definition above.

4.1.5 Using Point Estimates of Model Parameters

We now explain how we set the model parameters to deterministic estimates of values which roughly match COVID-19. The following section will consider how to include uncertainty in the parameter values. Specifically, the parameters are the incubation period α^{-1} ; rates of disease progression p_1 and p_2 ; rates of recovery from each level of infection, γ_1 , γ_2 , and γ_3 ; infectiousness for each level of infection, β_1 , β_2 , and β_3 ; and a death rate for critical infections, κ . $u \in [0, 1]$ is a control parameter, representing the strength of action taken to prevent

new infections (Boldog et al., 2020). To estimate distributions over the uncontrollable model parameters, we consider their relationships with various measurable quantities

$$\text{incubation period} = \alpha^{-1} \tag{21}$$

$$\text{mild duration} = \frac{1}{\gamma_1 + p_1} \tag{22}$$

$$\text{severe duration} = \frac{1}{\gamma_2 + p_2} \tag{23}$$

$$\text{critical duration} = \frac{1}{\gamma_3 + \kappa} \tag{24}$$

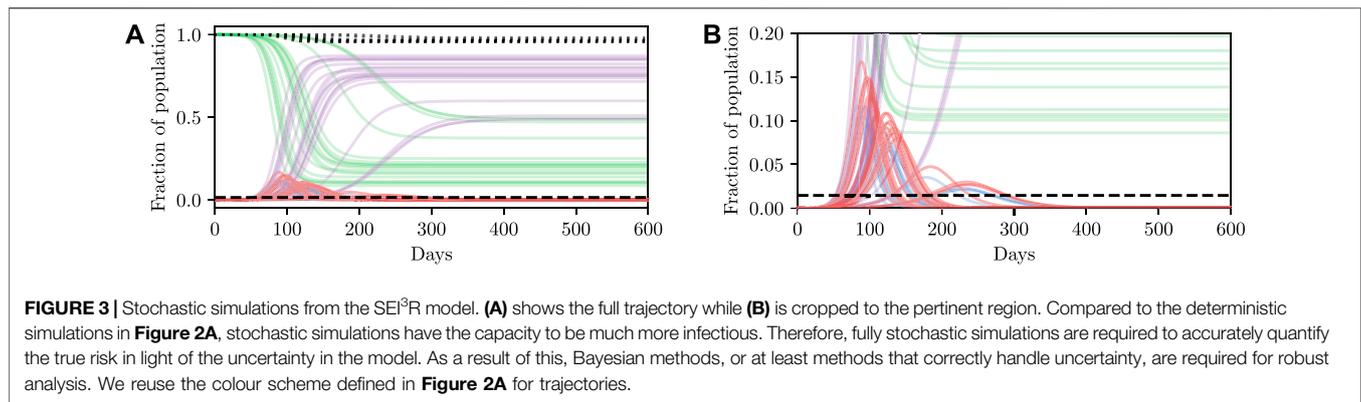
$$\text{mild fraction} = \frac{\gamma_1}{\gamma_1 + p_1} \tag{25}$$

$$\text{severe fraction} = \frac{\gamma_2}{\gamma_2 + p_2} \cdot (1 - \text{mild fraction}) \tag{26}$$

$$\text{critical fraction} = 1 - \text{severe fraction} - \text{mild fraction} \tag{27}$$

$$\text{fatality ratio} = \frac{\kappa}{\gamma_3 + \kappa} \cdot \text{critical fraction}. \tag{28}$$

Given the values of the left-hand sides of each of Eqs 21–28, (as estimated by various studies) we can calculate model parameters α , p_1 , p_2 , γ_1 , γ_2 , γ_3 and κ by inverting this system of Equations. These parameters, along with estimates for β_1 , β_2 , and β_3 , and a control input u , fully specify the model. Hill et al. (2020) use such a procedure to deterministically fit parameter values. Given the parameter values, the simulation is entirely deterministic. Therefore, setting parameters in this way enables us to make deterministic simulations of “typical” trajectories, as shown in Figure 2. Specifying parameters in this way and



running simulations in this system provides a low overhead and easily interpretable environment, and hence is an invaluable tool to the modeller.

4.1.6 Dealing With Uncertainty About Model Parameter Values

Deterministic simulations are easy to interpret on a high level, but they require strong assumptions as they fix the values of unknown parameters to point estimates. We therefore describe how we can perform inference and conditioning in a stochastic model requiring less strict assumptions, and show that we are able to provide meaningful confidence bounds on our inferences that can be used to inform policy decisions more intelligently than without this stochasticity. As described in **Section 3**, stochasticity can be introduced to a model through a distribution over the latent global parameters η . Examples of stochastic simulations are shown in **Figure 3A**. Clearly there is more capacity in this model for representing the underlying volatility and unpredictability of the precise nature of real-world phenomena, especially compared to the deterministic model.

However, this capacity comes with the reality that increased effort must be invested to ensure that the unknown latent states are correctly accounted for. For more specific details on dealing with this stochasticity please refer back to **Section 3**, but, in short, one must simulate for multiple stochastic values of the unknown parameters, for each value of the controllable parameters, and agglomerate the many individual simulations appropriately for the inference objective. When asking questions such as “will this parameter value violate the constraint?” there are feasibly some trajectories that are slightly above and some slightly below the trajectory generated by the deterministic simulation due to the inherent stochasticity (aleatoric uncertainty) in the real world. This uncertainty is integrated over in the stochastic model, and hence we can ask questions such as “what is the probability that this parameter will violate the constraint?” Using confidence values in this way provides some measure of how certain one can be about the conclusion drawn from the inference—if the confidence value is very high then there is a measure of “tolerance” in the result, compared to a result with a much lower confidence.

We define a joint distribution over model parameters as follows. We consider the 95% confidence intervals of β_1 , β_2 , and β_3 and the values in the left-hand sides of **Eqs 21–24**, and assume that their true values are uniformly distributed across these confidence intervals. Then at each time t in a simulation, we sample these values and then invert the system of **Eqs 21–28** to obtain a sample of the model parameters. More sophisticated distributions could easily be introduced once this information becomes available. We now detail the nominal values used for typical trajectories (and the confidence intervals used for sampling). The nominal values are mostly the same as those used by (Hill et al., 2020). We use: an incubation period of 5.1 days (4.5–5.8) (Lauer et al., 2020); a mild infection duration of 6 days (5.5–6.5) (Wölfel et al., 2020); a severe infection duration of 4.5 days (3.5–5.5) (Sanche et al., 2020); a critical infection duration of 6.7 days (4.2–10.4); fractions of mild, severe, and critical cases of 81, 14, and 5% (Wu and McGoogan, 2020); and a fatality ratio of 2% (Wu and McGoogan, 2020). We also use $\beta_1 = 0.33/\text{day}$ (0.23–0.43), and $\beta_2 = 0/\text{day}$ (0–0.05), and $\beta_3 = 0/\text{day}$ (0–0.025). Where possible, the confidence intervals are obtained from the studies which estimated the quantities. Where these are not given, we use a small range centred on the nominal value to account for possible imprecision.

4.2 Agent-Based Simulation

While compartmental models, such as the SEIR model described in **Section 4.1**, provide a mathematically well understood global approximation to disease dynamics, due to their coarse-grained statistical nature they cannot capture many important aspects and local details of the physical and social dynamics underlying the spread of a disease. These aspects include geographic information, spatio-temporal human interaction patterns in social hubs such as schools or workplaces, and the impact of individual beliefs on transmission events. To address these limitations, agent-based simulators (ABS) have been introduced. Such simulators have practically no restrictions in terms of expressiveness, i.e., they can make use of all features of modern Turing-complete programming languages, at the significant computational cost of simulating all details involved.

4.2.1 FRED: Fine-Grained Simulation of Disease Spreading

FRED² (Grefenstette et al., 2013) is an instance of the class of epidemiological agent-based simulators that are currently available for use in policy-making. FRED is an agent-based modeling language and execution platform for simulating changes in a population over time. FRED represents individual persons, along with social contacts and interactions with the environment. This enables the model to include individual responses and behaviors that vary according to the individual's characteristics, including demographics (age, sex, race, etc.), as well as the individual's interactions with members of various social interaction groups, such as their neighborhood, school or workplace. The FRED user can define and track any dynamic condition for the individuals within the population, including diseases (such as COVID-19), attitudes (such as vaccine acceptance), and behaviors (such as social distancing).

FRED captures demographic and geographic heterogeneities of the population by modeling every individual in a region, including realistic households, workplaces and social networks. Using census-based models available for every state and county in the US and selected international locations, FRED simulates interactions within the population in discrete time steps of 1 h. Transmission kernels model the spatial interaction between infectious places and susceptible agents. These characteristics enable FRED to provide much more fine-grained policy advice at either the regional or national level, based on socio-economic and political information which cannot be incorporated into compartmental models.

We chose to use FRED in this work because it has been used to evaluate potential responses to previous infectious disease epidemics, including vaccination policies (Lee et al., 2011), school closure (Potter et al., 2012), and the effects of population structure (Kumar et al., 2015) and personal health behaviors (Kumar et al., 2013; Liu et al., 2015).

After 10 years of development as an academic project, FRED has been licensed by the University of Pittsburgh to Epistemix³, to develop commercial applications of the FRED modeling technology. In turn, Epistemix has developed a detailed COVID-19 model in FRED, which is used in the experiments described here.

The FRED COVID-19 model includes three interconnected components: 1) The natural history of COVID-19; 2) The social dynamics/behavior of individuals; and 3) The Vaccination Program. The COVID-19 model was designed using the latest scientific data, survey information from local health authorities, and in consultation with expert epidemiologists. This model has been used to project COVID-19 cases in universities, K-12 school districts, large cities, and offices.

The FRED COVID-19 natural history model represents the period of time and trajectory of an individual from infection or onset to recovery or death. In the current version of the model, when an individual, or an agent, is exposed to SARS-CoV-2, the virus that causes COVID-19, the individual enters a 2-day latent period before they become infectious. In the infectious state, individuals can either be asymptomatic, symptomatic, or hospitalized. The probability of entering any of these infectious states is based on the individual's age, infection history, and vaccination history. Individuals have a duration of illness (i.e., number of days they can transmit the virus) which is dependent on infectious state, a severity of disease (i.e., magnitude of transmissibility), and a disease outcome (recovery or death).

When an agent is exposed to SARS-CoV-2 and becomes symptomatic, the individual chooses whether or not to isolate themselves from normal activities. Approximately 20% of individuals continue regular daily activities while symptomatic. Agents who are exposed and develop an asymptomatic infection do not isolate themselves and go about their regular activities. This introduces both symptomatic and asymptomatic forms of transmission into the model.

Prevalence of mask wearing and adherence to social distancing are unique to each location and change over time. The level of compliance to these behaviors is set based on the number of active infections that were generated from reported cases in the previous 2 weeks. Social distancing is assumed to reduce the number of contacts between agents in each place the agent attends.

Following the general recipe for framing planning as inference in **Section 3.1**, the following section defines what a prior on controls θ is in terms of FRED internals, how FRED parameters relate to η , and how to condition FRED on desirable future outcomes $Y_{0:T}^{aux}$. **Section 4.2.3** describes results of performing automated inference in this stochastic simulation-based model using the probabilistic programming system PyProb. The main point of this section is to illustrate how a stochastic simulator can be seen as a probabilistic model and, when integrated with an appropriate probabilistic programming system, can be repurposed to perform automatic inference, in this instance for planning as inference. The FRED-specific recipe we provide below should be read with the understanding that it is easy to apply the same recipe to most if not all existing epidemiological simulators, as their fundamental computational structure is regular.

4.2.2 Turning FRED Into a Probabilistic Program

The FRED simulator has a parameter file which stipulates the values of θ and η . In other words, both the controllable and non-controllable parameters live in a parameter file. FRED, when run given a particular parameter file, produces a sample from the distribution $p(X_{0:T}, Z_{0:T} | \theta, \eta)$. Changing the random seed and re-running FRED will result in a new sample from this distribution.

The difference between $X_{0:T}$ and $Z_{0:T}$ in FRED is largely in the eye of the beholder. One way of thinking about it is that $X_{0:T}$ are all the values that are computed in a run and saved in an output file and $Z_{0:T}$ is everything else.

In order to turn FRED into a probabilistic programming model useful for planning via inference several small but

²<https://fred.publichealth.pitt.edu/>

³<http://www.epistemix.com/>

consequential changes must be made to it. These changes can be directly examined by browsing one of the public source code repositories accompanying this paper.⁴ First, the random number generator and all random variable samples must be identified so that they can be intercepted and controlled by PyProb. Second, any variables that are determined to be controllable (i.e., part of θ) need to be identified and named. Third, in the main stochastic simulation loop, the state variables required to compute Y_t^{aux} and Y_t^{obs} must be extracted. Fourth, these variables must be given either synthetic ABC likelihoods or given constraints in the form of likelihoods. Finally, a mechanism for identifying, recording, and or returning $X_{0:T}$ to the probabilistic programming system must be put in place. FRED, like many stochastic simulators, includes the ability to write-out results of a run of the simulator to the filesystem. This, provided that the correspondence between a sample $\theta^{(i)}$ and the output file or files that correspond to it is established and tracked, is how $X_{0:T}^{(i)}$ is implicitly defined.

In the interest of time and because we were familiar with the internals of PyProb and knew that we would not be using inference algorithms that were incompatible with this choice, the demonstration code does not show a full integration in which all random variables are controlled by the probabilistic programming system. Instead, it only controls the sampling of θ and the observation of Y_t^{aux} . Notably, this means that inference algorithms like lightweight Metropolis Hastings (Wingate et al., 2011), which are also included in PyProb, cannot be used with the released integration code.

4.2.3 Details of FRED+PyProb Integration

Our integration of PyProb into FRED required only minor modifications to FRED's code base, performed in collaboration with the FRED developers at Epistemix. More details about the integration of FRED and PyProb include:

- 1) The simulator is connected to PyProb through a cross-platform execution protocol (PPX⁵). This allows PyProb to control the stochasticity in the simulator, and requires FRED to wait, at the beginning of its execution, for a handshake with PyProb through a messaging layer.
- 2) PyProb overwrites the policy parameter values θ with random draws from the user-defined prior. While PyProb internally keeps track of all random samples it generates, we also decided to write out the updated FRED parameters to a parameter file in order to make associating $\theta^{(i)}$ and $X_{0:T}^{(i)}$ easy and reproducible.
- 3) For each daily iteration step in FRED's simulation, we call PyProb's observe function with a likelihood corresponding to the constraint we would like to hold in that day.

With these connections established, we are able to select an inference engine implemented by PyProb to compute the posterior. We use a particularly simple algorithm, namely rejection sampling, in order to focus our exposition on the

conceptual framework of planning as inference. PyProb implements multiple other, more complex, algorithms, which may be able to better approximate the posterior with a given computational budget. However for the inference task we consider, in which we attempt to infer only a small fraction of the random variables in the simulator, we find that rejection sampling is sufficiently performant.

We also remind the reader that, like in **Section 3.5**, more complex controls can be considered, in principle allowing for complex time-dependent policies to be inferred. We do not examine this here, but note that this extension is straightforward to implement in the probabilistic programming framework, and that PyProb is particularly well adapted to coping with the additional complexity. Compared to sampling parameter values for FRED at the beginning of the simulation, such time-varying policies could be implemented through changing the FRED model source code directly. This approach will be explored in future research.

5 EXPERIMENTS

We now demonstrate how inference in epidemiological dynamics models can be used to inform policy-making decisions. We organize this section according to a reasonable succession of steps of increasing complexity that one might take when modeling a disease outbreak. We again stress that we are not making COVID-19 specific analyses here, but instead highlight how framing the task as in **Section 3** allows existing machine learning machinery to be leveraged to enhance analysis and evaluation of outcomes; and avoid some potential pitfalls.

We begin by showing how a simple, deterministic compartmental SEIR-based model can be used to inform policy-making decisions, and show how analysis derived from such a deterministic model can fail to achieve stated policy goals in practice. Next, we demonstrate how using a stochastic model can achieve more reliable outcomes by accounting for the uncertainty present in real world systems. While these stochastic models address the limitations of the deterministic model, low-fidelity SEIR models are, in general, not of high enough fidelity to provide localized, region-specific policy recommendations. To address this we conclude by performing inference in an existing agent-based simulator of infectious disease spread and demonstrate automatic determination of necessary controls.

5.1 SEI³R Model

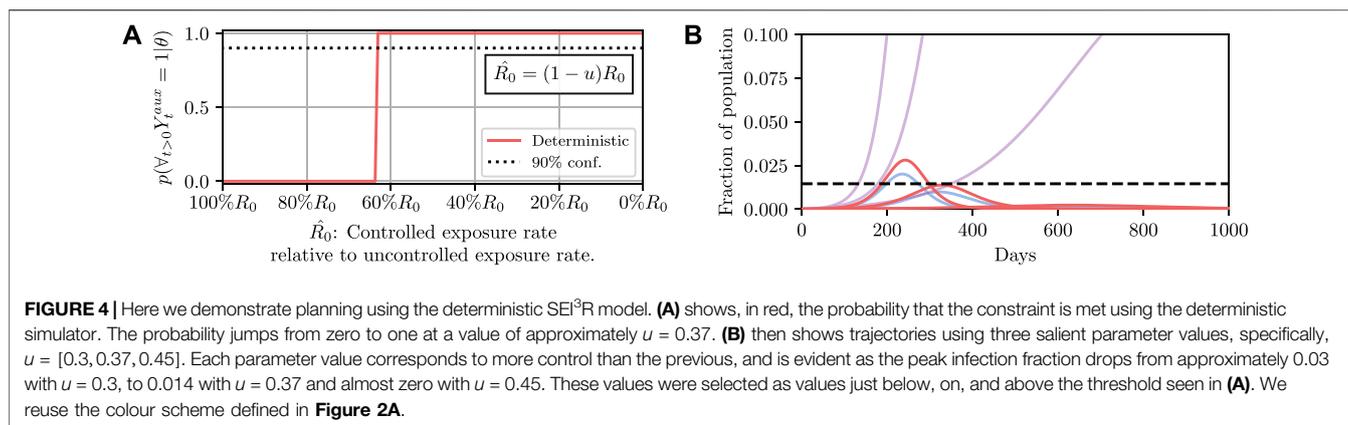
The most straightforward approach to modeling infectious diseases is to use low-dimensional, compartmental models such as the widely used susceptible-infectious-recovered (SIR) models, or the SEI³R variant introduced in **Section 4.1**. These models are fast to simulate and easy to interpret, and hence form a powerful, low-overhead analysis tool.

5.1.1 Deterministic Model

The system of equations defining the SEI³R model form a deterministic system when global parameter values, such as the mortality rates or incubation periods, are provided.

⁴<https://github.com/plai-group/FRED>

⁵<https://github.com/pyprob/ppx>



However, the precise values of these parameter values are unknown, and instead only confidence intervals for these parameters are known, i.e., the incubation period is between 4.5 and 5.8 (Lauer et al., 2020). This variation may be due to underlying aleatoric uncertainty prevalent in biological systems, or epistemic uncertainty due to the low-fidelity nature of SIR-like models. We do not discuss them here, but work exists automatically fitting point-wise estimates of model parameter values directly from observed data (Wearing et al., 2005; Mamo and Koya, 2015).

Regardless of whether one obtains a point estimate of the parameter values by averaging confidence intervals, or by performing parameter optimization, the first step is to use these values to perform fully deterministic simulations, yielding simulations such as those shown in **Figure 2A**. Simulations such as these are invaluable for understanding the bulk dynamics of systems, investigating the influence of variations in global parameter values or investigating how controls affect the system. However, the ultimate utility in these models is to use them to inform policy decisions to reduce the impact of outbreaks. As eluded to above, this is the primary thrust of this work, combining epidemiological simulators with automated machine learning methodologies to model policy outcomes, by considering this problem as *conditioning* simulations on outcomes.

To demonstrate such an objective, we consider maintaining the infected population below a critical threshold C at all times. In a deterministic system there are no stochastic quantities and hence whether the threshold is exceeded is a deterministic function of the controlled parameters, i.e., the value of $p(\forall_{t>0}: Y_t^{aux} = 1 | \theta)$ (related to **Eq. 7** via Bayes rule) is binary in a deterministic system and hence takes a value of either 0 or 1. Therefore, we can simply simulate the deterministic system for a finite number of θ values, and select those parameter values that do not violate the constraint. We vary the free parameter $u \in [0, 1]$, where u is a scalar value that reduces the baseline reproduction rate as $\hat{R}_0 = (1 - u)R_0$. We define u in this way such that u represents an *intervention*, or change from normal conditions. The parameter u is the only parameter we have control over, and hence $\theta = u$.

Results for this are shown in **Figure 4**. It can then be read off that under the deterministic model \hat{R}_0 must be reduced by at least 37.5% of R_0 to satisfy the constraint. **Figure 4** shows trajectories simulated using insufficient intervention with $u = 0.3$ ($\hat{R}_0 = 70\%R_0$), acceptable intervention of $u = 0.375$ ($\hat{R}_0 = 62.5\%R_0$), and excessive intervention of $u = 0.45$ ($\hat{R}_0 = 55\%R_0$), and show that these parameters behave as expected, violating the constraint, remaining just under the threshold and remaining well beneath the threshold respectively.

5.1.2 Stochastic Simulation

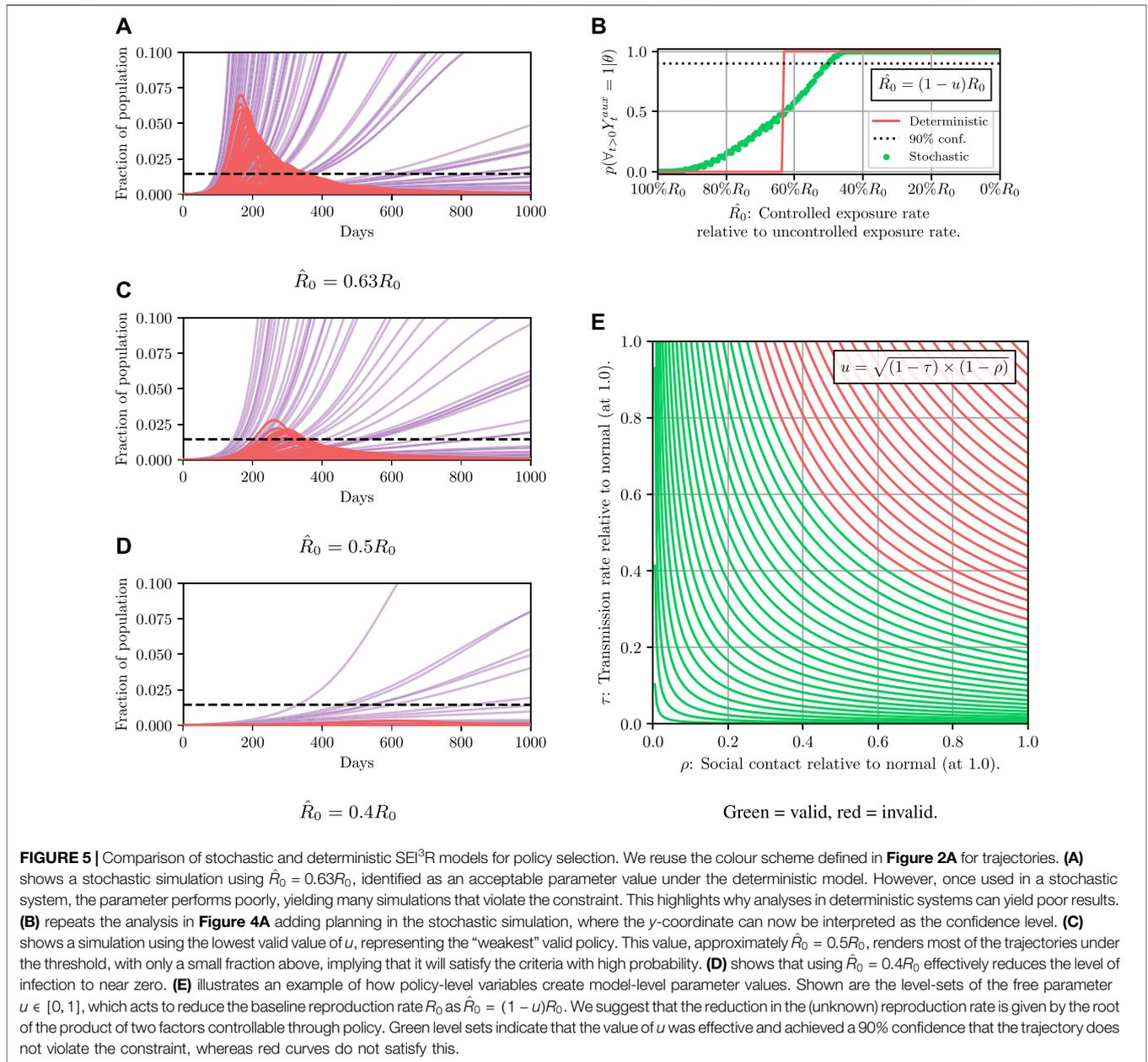
While the above example demonstrates how parameters can be selected by conditioning on desired outcomes, we implicitly made a critical modeling assumption. While varying the free parameter u , we fixed the *other* model parameter values (α^{-1} , γ_1 , etc.) to single values. We therefore found a policy intervention in an unrealistic scenario, namely one in which we (implicitly) claim to have certainty in all model parameters except u .

To demonstrate the pitfalls of analyzing deterministic systems and applying the results to an inherently stochastic system such as an epidemic, we use the permissible value of u solved for in the deterministic system, $u = 0.375$, and randomly sample values of the remaining simulation parameters. This “stochastic” simulator is a more realistic scenario than the deterministic variant, as each randomly sampled η represents a unique, plausible epidemic being rolled out from the current world state.

The results are shown in **Figure 5A**. Each line represents a possible epidemic. We can see that using the previously found value of u results in a large number of epidemics where the infectious population exceeds the constraint, represented by the red trajectories overshooting the dotted line. Simply put, the control parameter we found previously fails in an unacceptable number of simulations.

This detail highlights the shortcomings of the deterministic model: in the deterministic model a parameter value was either accepted or rejected with certainty. There was no notion of the variability in outcomes, and hence we have no mechanism to concretely evaluate the risk of a particular configuration.

Instead, we can use a stochastic model which at least does account for some aleatoric uncertainty about the world. We

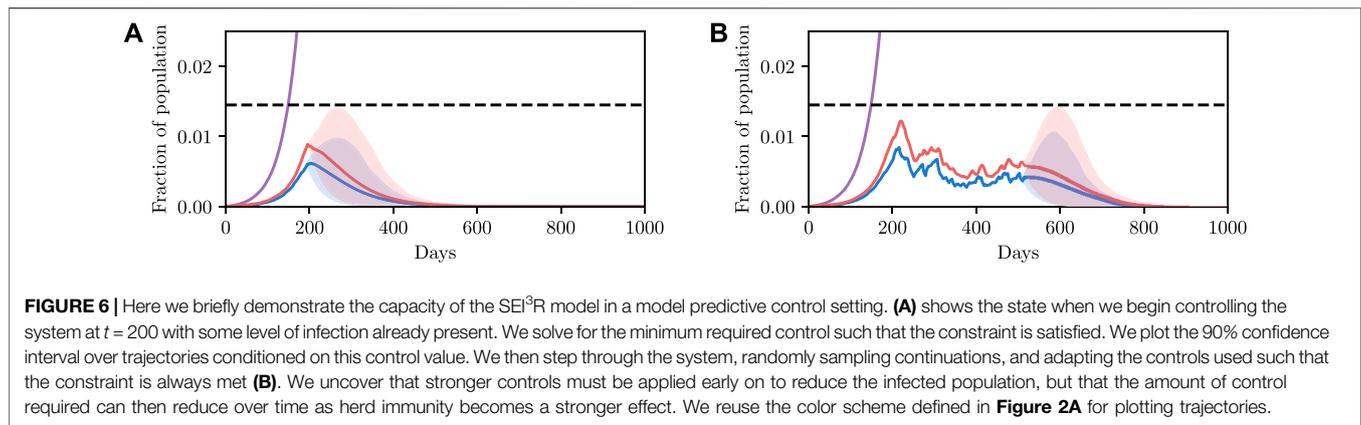


repeat the analysis picking the required value of u , but this time using the stochastic model detailed in **Section 4.1**. In practice, this means the (previously deterministic) model parameters detailed in **Eqs 21–28** are randomly sampled for each simulation according to the procedure outlined following the equations.

To estimate the value of $p(\forall_{t>0}: Y_t^{aux} = 1 | \theta)$, for a given u value, we sample M stochastic trajectories from the system. We then simply count the number of trajectories for which the condition $\forall_{t>0}: Y_t^{aux} = 1$ holds, and divide this count by M . Intuitively, this operation is simple: for a given θ , simulate a number of possible trajectories, and, as the number of simulations M tends to infinity, the fraction that satisfy the constraint is the desired probability value. We note that this operation corresponds to an “inner”

Monte Carlo expectation, sampling under the distribution of simulator trajectories conditioned on θ , evaluating the expected number of trajectories that do not violate the threshold. This value is then passed through a non-linear indicator function extracting those parameters that yield a confidence above a certain threshold. We are then free to use any method we please for exploring θ space, or, evaluating additional Monte Carlo expectations under the resulting θ distribution. As such, this system is a nested Monte Carlo sampler (Rainforth et al., 2018).

The results are shown in **Figure 5B**. The likelihood of the result under the stochastic model is not a binary value like in the deterministic case, and instead occupies a continuum of values



representing the confidence of the results. We see that the intersection between the red and green curves occurs at approximately 0.5, explaining the observation that approximately half of the simulations in **Figure 5A** exceed the threshold. We can now ask questions such as: “what is the parameter value that results in the constraint not being violated, with 90% confidence?” We can read off rapidly that we must instead reduce the value of \hat{R}_0 to 50% of its original value to satisfy this confidence-based constraint. Repeating the stochastic simulations using these computed values confirms that very few simulations violate the constraint (**Figure 5C**). The ability to tune the outcome based on a required level of confidence is paramount for safety-critical applications, as it informs how sensitive the system is to the particular parameter choice and is more resilient to model misspecification.

5.1.3 Model Predictive Control

We have shown how one can select the required parameter values to achieve a desired objective. To conclude this example, we apply the methodology to iterative planning. The principal idea underlying this is that policies are not static and can be varied over time conditioned on the current observed state. Under the formalism used here, re-evaluating the optimal control to be applied, conditioned on the new information, is as simple as re-applying the planning algorithm at each time step. Note that here we consider constant control. However, more complex, time-varying control policies can easily be considered under this framework. For instance, instead of recovering a fixed control parameter, the parameter values of a polynomial function defining time-varying control input could be recovered, or, a scalar value determining the instantaneous control input at each time-step. This is a benefit of the fully Bayesian and probabilistic programming-based approach we have taken: the model class that can be analyzed is not fixed and can be determined (and easily changed and iterated on) by the modeller, and the inference back-end cleanly and efficiently handles the inference.

We show a demonstration of this in **Figure 6**. In this example, we begin at time $t = 200$ with non-zero infection rates. We solve for a policy that satisfies the constraint with 90% certainty, and

show this confidence interval over trajectories as a shaded region. We then simulate the true evolution of the system for a single step sampling from the conditional distribution over state under the selected control parameter. We then repeat this process at regular intervals, iteratively adapting the control to the new world state. We see that the confidence criterion is always satisfied and that the infection is able to be maintained at a reasonable level. We do not discuss this example in more detail, and only include it as an example of the utility of framing the problem as we have, inasmuch as iterative re-planning based on new information is a trivial extension under the formulation used.

5.1.4 Policy-Based Controls

We have illustrated how simulations can be used to answer questions about the suitability of parameter values we can influence, while marginalizing over those parameters we do not have control over. However, u is not something that is *directly* within our control. Instead, the value of u is set through changing policy level factors. As an exploratory example we suggest that the value of u is the square root of the product of two policy-influenceable factors: the fractional reduction in social contact, ρ , below its normal level (indicated as a value of 1.0), and the transmission rate relative to the normal level, τ , where we again denote normal levels as 1.0. This relationship is shown in **Figure 5E**.

We indicate u level sets that violate the constraint in red, and valid sets in green. We suggest taking the least invasive, valid policy, being represented by the highest green curve. Once the analysis above has been performed to obtain a value of u , that satisfies the required infection threshold, it defines the set of achievable policies. Any combination of τ and η along this curve renders the policy valid. Here, additional factors may come into consideration that make particular settings of τ and η more or less advantageous. For instance, wearing more PPE may be cheaper to implement and less economically and socially disruptive than social distancing, and so higher values of τ may be selected relative to η . This reduces to a simple one-dimensional optimization of the cost surface along the level-set.

While we have simply hypothesized this as a potential relationship, it demonstrates how policy level factors influence

TABLE 1 | Prior over FRED control parameter $\theta = (\theta_1, \dots, \theta_5)$ for influenza simulations.

Prior	Control
$\theta_1 \sim \text{Uniform}(0, 14)$	shelter-in-place duration
$\theta_2 \sim \text{Uniform}(0, 1)$	shelter-in-place compliance rate
$\theta_3 \sim \text{Uniform}(0, 1)$	isolation rate
$\theta_4 \sim \text{Uniform}(0.01, 0.21)$	school closure attack rate threshold
$\theta_5 \sim \text{Uniform}(0, 1)$	hand washing compliance rate

simulations and outcomes. While the SEIR model family is an invaluable tool for analyzing and understanding the bulk dynamics of outbreaks, it is too coarse-grained for actual, meaningful, localized policy decisions to be made, especially when those policy decisions are directly influencing populations. Further, these notions of “policy” are somewhat abstract here because of the high-level nature of the SEIR model used. We now go on to resolve these issues by using the more sophisticated agent-based simulator FRED, where simulations are able to represent localized variations, and where real policy measures are more easily defined.

5.2 FRED Simulator

In this section, we turn to agent-based simulators. We showcase how control as inference might possibly be used to inform regional policy decisions through two scenarios presented in Sections 5.2.1, 5.2.2. In the first scenario, we consider an influenza outbreak and in the second a COVID-19 outbreak is simulated. In both scenarios the policy makers wish to “flatten the curve” by limiting some statistic of the infected population under a certain threshold.

5.2.1 Influenza Simulation

In this section, we consider a scenario where an *influenza* (not COVID-19) outbreak has occurred in Allegheny County (similarly to (Grefenstette et al., 2013)), and policy makers wish to limit the number of infections to less than 10% of the county’s population. To achieve this hypothetical goal, policy-makers might consider the following controls among others (corresponding to the θ parameter defined in Table 1):

- A “social distancing” policy which mandates all citizens stay home for a fixed period of time. Here, policy makers must influence:
 - 1) θ_1 , a *shelter-in-place duration*, or the length of time a social distancing policy must be in place.
 - 2) θ_2 , a *shelter-in-place compliance rate*, or the percentage of the population to which this policy applies.
- θ_3 , a *symptomatic isolation rate*, the fraction of symptomatic individuals that self-isolate during an epidemic.
- θ_4 , a *school closure attack rate threshold*, a threshold on the total percentage of people infected that automatically triggers a 3-week school closure when exceeded.
- θ_5 , a *hand washing compliance*, the percentage of the population that washes their hands regularly.

For simplicity of results interpretation we have put uniform priors, appropriately continuous or discrete, on intervals of interest for these controllable parameters.

Also relative to Eq. 7 we choose Y_t to be a binary variable indicating if the proportion of the county’s infected population is below 10% on day t . By inferring $p(\theta | \forall t > 0 : Y_t = 1)$, we characterize which control values will lead to this desired outcome.

Using the model described in Section 4.2 with parameters as defined in this section, we used PyProb to perform automated inference over the policy parameters described in Table 1. To reiterate, we infer the posterior distribution over control parameters θ which satisfy the goal of limiting the instantaneous number of infections to less than 10% of the county’s population at all times up to a maximum of 150 days. In conducting our experiments we generated one million samples, 40% of which satisfied the policy goal.

5.2.1.1 Simulator Configuration

We simulate Allegheny County, Pennsylvania, using 2010 U.S. Synthetic Population data. We use an older, publicly available version of FRED⁶ and its default parameters, which simulates a model of influenza “in which infectivity starts 1 day before symptoms and lasts for 7 days. Both symptoms and infectivity ramp up and ramp down.”⁷ All simulations begin with 10 random agents seeded with the virus at day zero. Person-to-person contact rates for various environments (household, office, etc.) were calibrated to specifically model Allegheny County. Parameter files for the four policies under consideration are available online⁸.

5.2.1.2 Results

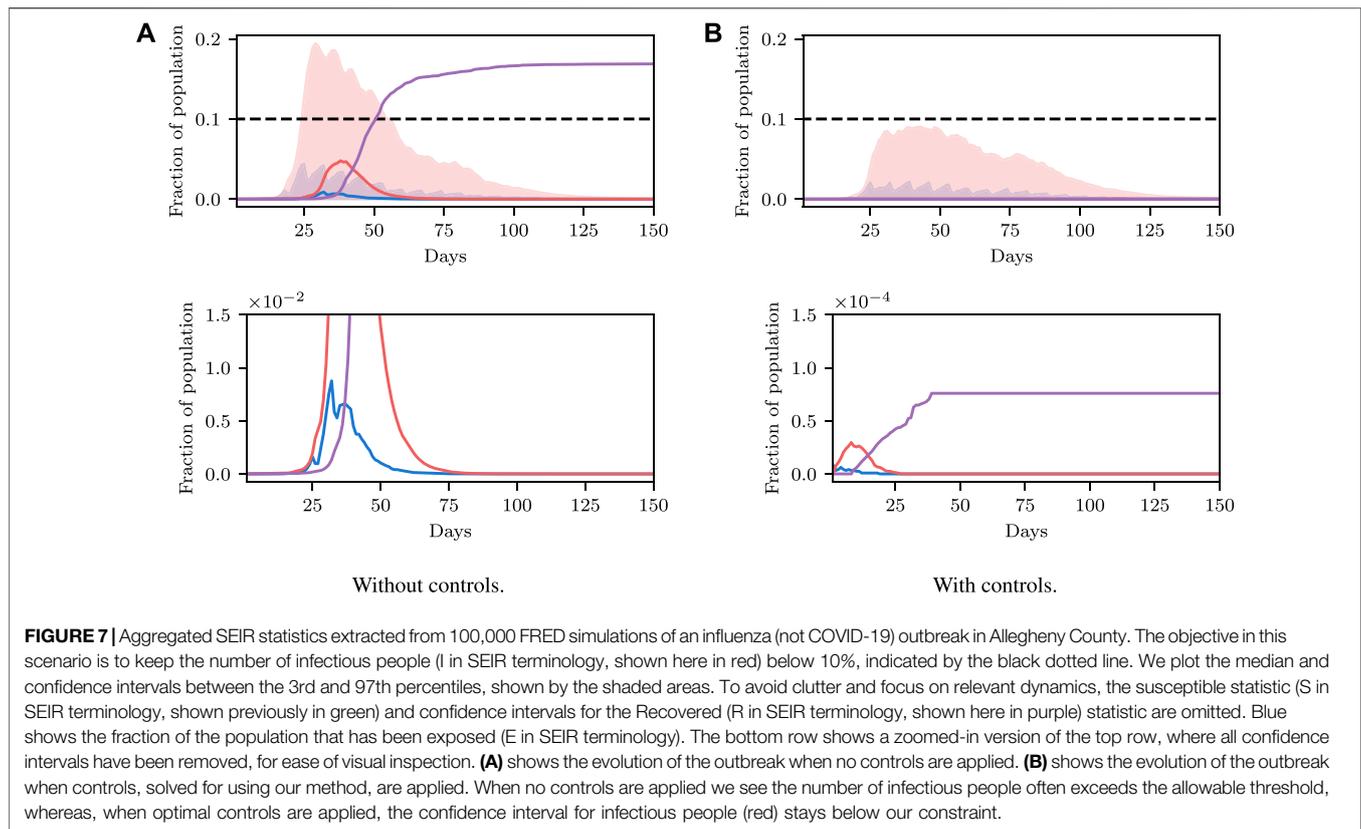
Figure 7 shows the SEIR statistics for both the uncontrolled (left) and controlled (right) simulations. We see that the confidence interval in red stays beneath our infection rate constraint indicated by the dashed black line, while in the uncontrolled scenario the confidence band well exceeds the constraint. We also observed the overall number of exposed and infectious populations is much lower using samples of control variables from the inferred posterior, which indicates that control variable values drawn from this posterior are indeed effective in limiting the spread of the virus.

While Figure 7 indicates that the inferred controls can achieve the desired aims, it doesn’t indicate *which* policy to choose. To answer this, we plot the two-dimensional marginal distributions over controlled policy parameters in Figure 8. Policy-makers could use a figure like this, coupled with a utility function, to generate policy recommendations. Interpreting Figure 8, first we see the importance of influencing hand washing compliance and quick school closures. This plot indicates that hand washing compliance must be driven above 50% and the school closure attack rate threshold beneath 7% in order to achieve our stated

⁶<https://github.com/PublicHealthDynamicsLab/FRED>

⁷https://github.com/PublicHealthDynamicsLab/FRED/blob/FRED-v2.12.0/input_files/defaults

⁸<https://github.com/plai-group/FRED/tree/FRED-v2.12.0/params>



goal. We also note the correlation between hand washing compliance and isolation rate. If we can achieve only 70% hand washing compliance then the isolation rate must be driven high, however, if hand washing compliance is very good then a lower isolation rate is tolerable. The importance of hygiene and long-term school closures has also been noted in the epidemiology literature (Lee et al., 2010; Saunders-Hastings et al., 2017; Germann et al., 2019).

A final interpretation of the results of **Figure 8** can be provided in terms of the outcome of a hypothetical influenza policy-making scenario. A recommendation one might extract from this visualization of the posterior inference results is a conjunction of the following controls:

- 1) Ensure that all schools close as soon as 2% of the population contracts the virus.
- 2) Attempt to drive the hand washing compliance to 80%.
- 3) Attempt to drive the symptomatic isolation rate to 50%.
- 4) No amount of sheltering in place is required.

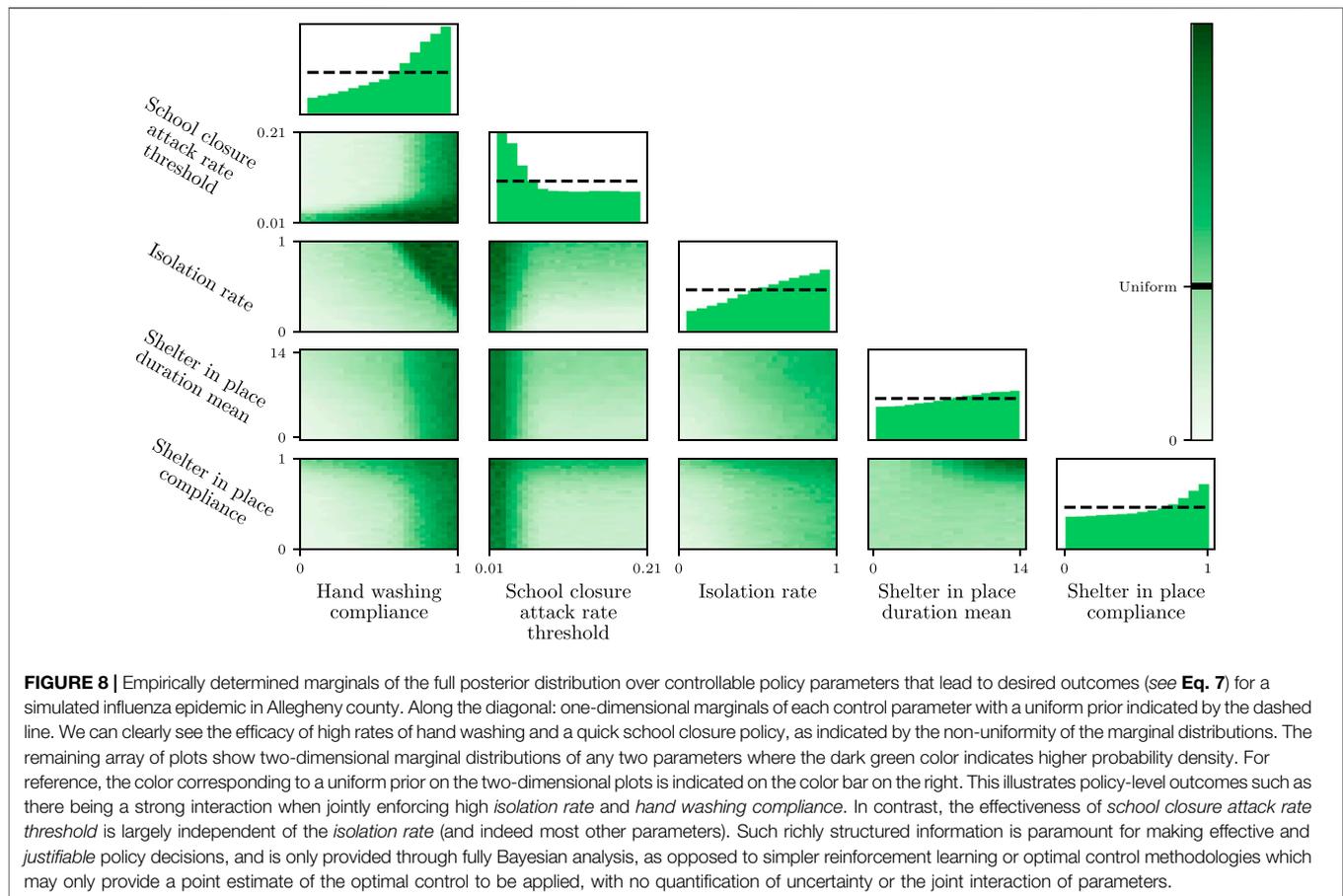
Such a recommendation corresponds to what one could imagine would be, in comparison to more draconian options, a relatively mild, economically speaking, policy response that still attains the objective. Of course, in the presence of a real influenza outbreak, vaccination would be among the most critical controls to consider. We have intentionally excluded it here in order to be in some small way more reflective of the COVID-19 pandemic which is emerging at the time of conducting this experiment.

The results of applying this policy are shown in **Figure 9**, where we show a single simulation of the spread of influenza in Allegheny county over time in both the uncontrolled (left) and controlled (right) cases. Each red dot shows the household of an infectious person. This policy recommendation can be seen to work as it reduces the maximum number of infected households from 215,799 on day 29 in the uncontrolled case to 65,997 on day 83 in the uncontrolled case. The maximum in the controlled case actually occurs during a second “outbreak” after schools reopen (noticeable around day 90 in **Figure 9C**). While the controlled policy results in two “spikes,” our inference procedure accounts for this and correctly controls the maximum number of infectious persons to remain below 10% of the total at all times.

5.2.2 COVID-19 Simulation

In this section, we consider COVID-19 outbreaks of/in the Seattle metropolitan area. The controls we consider in these experiments are the following (corresponding to the θ parameter defined in **Table 2**):

- θ_1 , a *social distancing*, or the fraction of population practising social distancing.
- θ_2 , a *workplace closure*, or the fraction of businesses deemed essential and exempt from workplace closure mandates.
- θ_3 , a *school closure attack rate threshold*, a threshold on fraction of population hospitalized per day. Schools get closed if the daily hospitalization rate exceeds this threshold.



Similar to Section 5.2.1, we have put uniform priors on these controllable parameters, as described in Table 2.

We put thresholds on the number of daily hospitalizations and daily deaths in different experiments. Hence, Y_t in Eq. 7 is a binary variable indicating if any of the thresholds are exceeded.

5.2.2.1 Simulator Configuration

Our simulations presented here are on a synthetic population of the Seattle metropolitan area incorporated in the FRED simulator (a population size of 3,416,570). The synthetic population closely matches the 2010 census data for the United States with high spatial resolution, differing from the American Community Survey by less than 1% for large counties. A detailed description of the methodology and further comparisons are provided in (Wheaton et al., 2009). The parameters of the disease are calibrated to COVID-19, as explained in Section 4.2. We have conducted simulations in three different time periods and the initial conditions of each simulation (number of active infections, total hospitalizations, total deaths and vaccinated people) is set according to the numbers reported by U.S. health officials for the chosen dates (see Table 3).

5.2.2.2 Results

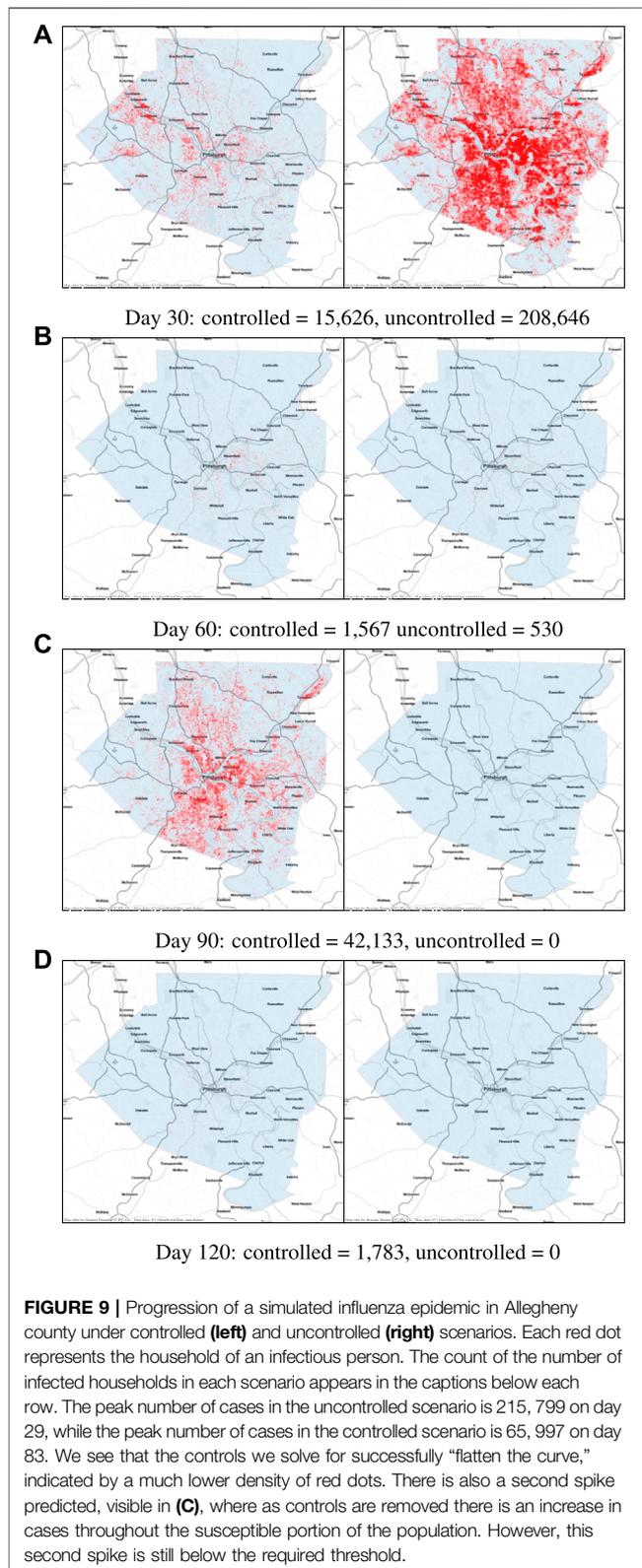
Figure 10 shows the number of daily cases, hospitalizations, and deaths (note that this is not the same as SEIR) for all time

periods considered in this experiment. Similar to SEIR plots, it shows results for both the uncontrolled (left) and controlled (right) simulations.

Figure 11 shows samples from two-dimensional marginals of the distribution over controlled policy parameters. According to these results, social distancing is the most important control among the three and policy makers should make sure more than 70% of the population practice it. Although enforcing a more strict school or workplace closure helps as well and allows a slightly wider range for social distancing compliance, they are far less effective than social distancing. Adding more relative control parameters (for example, public mask wearing, recommended vaccine booster shots, restricting large events, imposing travel restrictions, etc.) are planned extensions of future research.

We have conducted COVID-19 simulations for three different scenarios:

- The 6 months following March 1, 2020 when the outbreak is in its early stages with a few cases.
- The 6 months following December 1, 2020 when there is a large number of cases, hospitalizations and deaths.
- The 12 months following June 1, 2021 when the number of hospitalizations is the largest among our simulations and there are more than 3000 active infections and near



3000 deaths. However, in this case 47% of the population are vaccinated at the beginning of the simulation.

TABLE 2 | Prior over FRED control parameter $\theta = \{\theta_1, \theta_2, \theta_3\}$ for COVID-19 simulations.

Prior	Control
$\theta_1 \sim \text{Uniform}(0, 1)$	social distancing
$\theta_2 \sim \text{Uniform}(0, 1)$	workplace closure
$\theta_3 \sim \text{Uniform}(1, 1.5 \times 10^{-4})$	school closure attack rate threshold

The goal in the first two scenarios is to avoid the number of daily hospitalizations and daily deaths exceeding 409 and 100 people, respectively, while these numbers are 102 and 20 for the third scenario. For the third scenario our results in **Figure 11** show the goals are achieved even with a slightly looser set of controls over the population despite the large number of active infections and hospitalizations and tighter goals. This is indeed the effect of higher vaccination rates in the initial conditions for the third scenario.

Finally, as a part of this project we have created a website⁹ where users can perform planning as inference in COVID-19 simulations for different locations in the U.S. and Canada. It allows users to choose initial conditions of the simulations, policy goals and the set of control parameters (or interventions) to explore. The website automatically runs simulations on the population of the specified location and produces an interactive version of **Figures 10, 11**. It also provides the location of infections on the selected location’s map, similar to **Figure 9**. We are working on expanding the features and set of controllable parameters on the website as well.

6 SOFTWARE

In this paper, we have introduced and reviewed control as inference in epidemiological dynamics models and illustrated how this technique can be used to substantially increase the level of automation and precision of some aspects of policy-making using models at two extremes of expressivity and fidelity—compartmental and agent-based. We are releasing accompanying source code for our SEIR-type COVID-19 model along with bespoke inference code that researchers in the machine learning, control, policy-making, and approximate Bayesian inference communities can use immediately in their own research. We are also releasing code that demonstrates how a complex, existing, agent-based epidemiological dynamics model can be quickly interfaced to an existing probabilistic programming system so as to automate, even in such a complex model, the inference tasks resulting from the problem formulation of control as inference. Again, we stress that we do not report methodological innovation per se, as there are other ways of recombining existing inference packages and existing models to approach the control as inference

⁹<https://covid19ideas.cs.ubc.ca/>

TABLE 3 | Initial conditions and policy goals used in the COVID-19 simulations in the Seattle metropolitan area.

Simulation dates	Active infections	Total hospitalizations	Total deaths	% Vaccinated	Policy goals	
					Hospitalizations	Deaths
1 Mar. 2020–1 Sep. 2020	400	0	1	0	409	100
1 Dec. 2020–1 Jun. 2021	8017	2096	1459	0	409	100
1 Jun. 2021–1 Jun. 2022	3024	8313	2867	47	102	20

problem. In this sense, our principal goal is to demonstrate the feasibility of, and raise interdisciplinary awareness for, such approaches, rather than to champion any specific implementation.

The first part of the software release accompanying this paper is a pure Python implementation of the SEI³R model adapted to COVID-19 from **Section 4.1**. It is contained inside an online software version control repository and consists of the model, “custom” inference code that implements control as inference via approximate nested Monte Carlo (NMC), and code that shows how to perform stochastic model predictive control using the model, i.e., the code used to produce **Figure 6**.

The second part of the software release accompanying this paper encompasses an influenza model implemented using an older, open source version of the agent-based simulation platform FRED (Grefenstette et al., 2013), as well as the modifications necessary for the latter to be interfaced with the probabilistic programming system PyProb for simulator-based approximate Bayesian inference. This code is contained in three separate repositories: a fork of the FRED repository with PyProb integration¹⁰, a fork of the PyProb code with slight additions required for FRED integration¹¹, and a collection of scripts for orchestrating and plotting artifacts from simulation and inference¹². This last repository also contains a Singularity container (Kurtzer et al., 2017) image bundling all of the necessary software dependencies, including the repositories above. This should significantly reduce the setup time for interested parties, as Singularity images are supported at a large number of existing high performance computing facilities. While the current FRED source code is proprietary, we are happy to discuss the process of PyProb+FRED integration with any interested parties.

We believe that model-inference combinations such as FRED+PyProb could provide formidable policy analysis tools with potentially significant societal benefits, particularly because they would allow high-fidelity assessment of region-specific targeted policies. We expect that the set of effective policy interventions might differ across the regional characteristics of counties, states or countries, as well as their transient epidemiological situations. Testing this hypothesis depends upon expanding the number of counties, provinces, and countries with simulation profiles in FRED, or in any other agent-based epidemiological model that can be interfaced with

a universal and language-agnostic probabilistic programming system such as PyProb.

7 DISCUSSION

Our experience in conducting this research has led us to identify a number of opportunities for improvement in the fields of simulation-based inference and control.

7.1 Software Tools

Building a simple SEIR-type model with very few compartments is a project of the scope of graduate homework. Building and maintaining a simulator like FRED or the US National Large Scale Agent Model (Parker and Epstein, 2011) is a massive undertaking, which would be prohibitive to replicate or significantly extend in a crisis situation. As far as we could find with limited search effort when conducting this work in March 2020, there was neither a central repository of up-to-date open-source agent-based epidemiological models, nor an organizing body we could interface to immediately.

7.2 Methodology

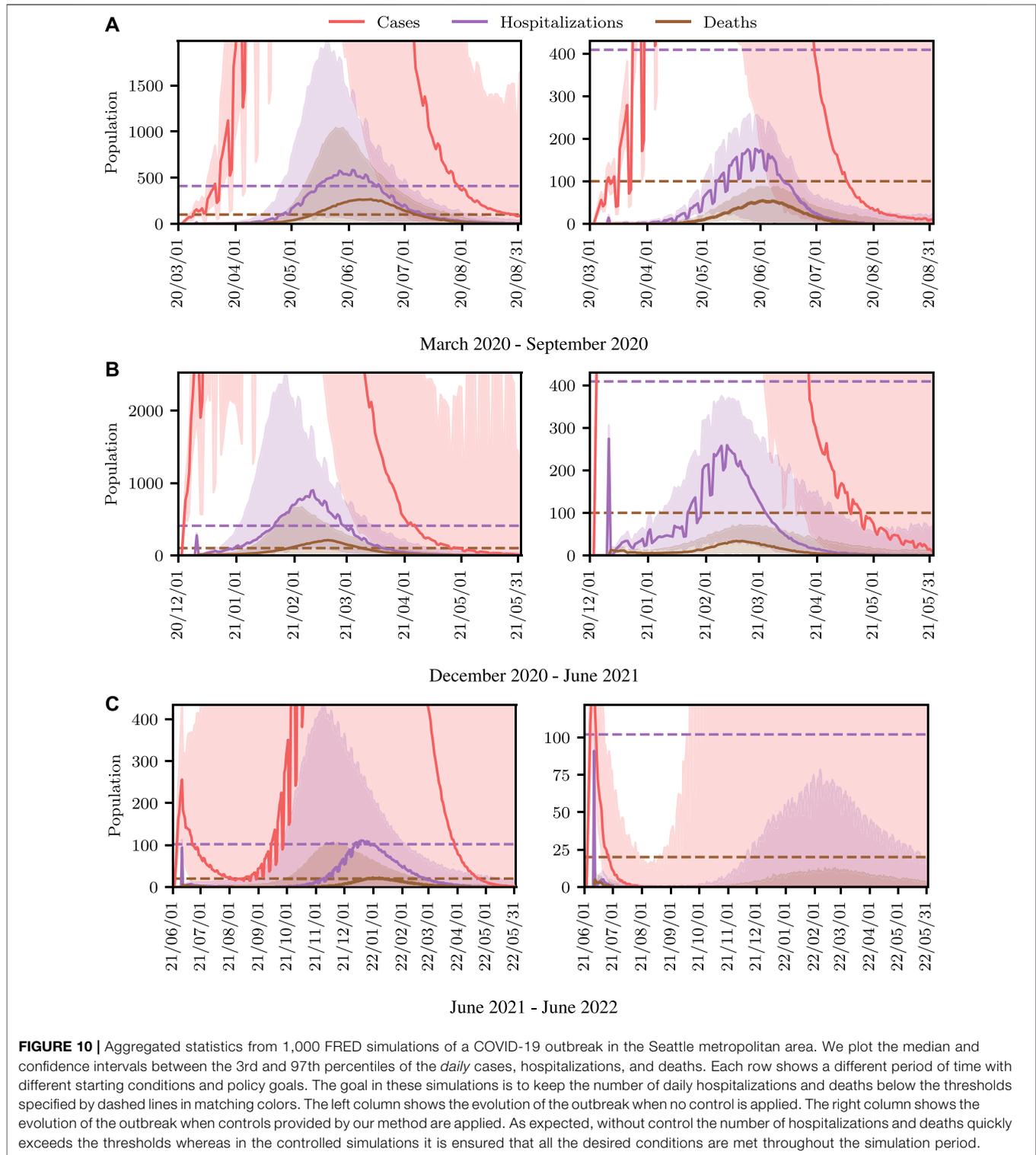
There appear to be practically very consequential conceptual gaps between the fields of control, epidemiology, statistics, policy-making, and probabilistic programming. To quote Lessler et al. (Lessler et al., 2016), “the historic separation between the infectious disease dynamics and ‘traditional’ epidemiologic methods is beginning to erode.” We, like them, see our work as a new opportunity for “cross pollination between fields.” As discussed earlier, the most closely related work that we found in the literature is all focused on automatic model parameter estimation from observational data (Toni et al., 2009; Kypraios et al., 2017; McKinley et al., 2018; Chatzilena et al., 2019; Minter and Retkute, 2019). These methods and the models to which they have been applied could be repurposed for planning, as we have demonstrated, simply by changing the random variables they condition on to include safety or utility metrics. Our emphasis therefore lies on the feasibility of planning as inference using existing software tools for approximate Bayesian inference.

Looking closer at the implementation choices, we found at least two existing papers that explore using probabilistic programming coupled to epidemiological simulators; (Funk and King, 2020) which used Libbi (Murray, 2015) and (Gram-Hansen et al., 2019) which used PyProb. The latter is an example of work that “hijacks” a malaria simulator in the same way we “hijacked” the FRED

¹⁰<https://github.com/plai-group/FRED>

¹¹<https://github.com/plai-group/pyprob>

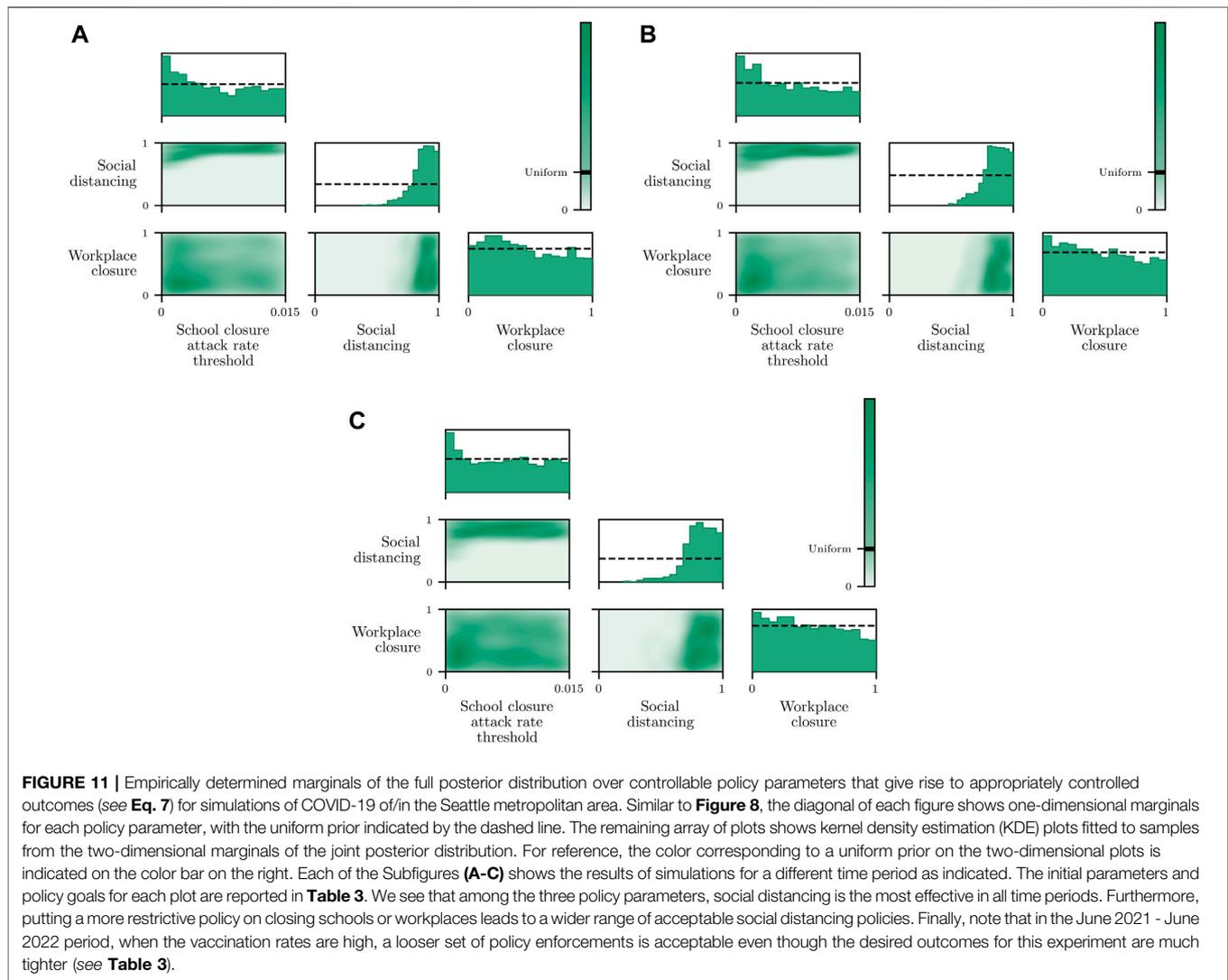
¹²<https://github.com/plai-group/covid>



simulator in this paper. Neither explicitly addresses control. Chatzilena et al., (2019) use the probabilistic programming system STAN Carpenter et al. (2017) to address parameter estimation in SEIR-type models, but it too does not explore control, nor could it be repurposed to control an agent-based

model with non-differentiable joint densities or, more generally, any external simulator not directly defined in the STAN language.

Like (Toni et al., 2009; Kypraios et al., 2017; McKinley et al., 2018; Chatzilena et al., 2019; Minter and Retkute, 2019) we too could have demonstrated automated parameter estimation in both of the



models that we consider, for instance to automatically adjust uncertainty in disease-specific parameters by conditioning on observable quantities such as death counts that are measured online during a breakout. However, as the epidemiological community already relies upon long-standing methods established for estimating confidence intervals for model parameters during evolving pandemics, we exclusively restricted ourselves to demonstrating how to achieve control via inference, assuming that the disease parameter priors have been obtained from established parameter estimation techniques. Combining the two kinds of observations in a single inference task is technically straightforward but does require care in interpretation.

8 FINAL THOUGHTS

At the time of writing all authors, apart from JG and DC at Epistemix Inc., Pittsburgh, were principally affiliated with the “Programming Languages for Artificial

Intelligence” (PLAI) research group at UBC, Vancouver, which is primarily involved in developing next-generation probabilistic AI tools and techniques. We felt, however, that the acute circumstances demanded we lend whatever we could to the global fight against COVID-19. Beyond the specific contributions outlined above, our secondary aim in writing this paper was to encourage other researchers to contribute their expertise to the fight against COVID-19 as well. We believe the world will be in a better place more quickly if they do.

DATA AVAILABILITY STATEMENT

The FRED simulator and model configurations used for the COVID-19 experiments are proprietary products of Epistemix Inc. All the other source codes, parameters and datasets presented in this study can be found at the open source repository <https://github.com/plai-group/covid>.

AUTHOR CONTRIBUTIONS

FW - conception, formulation, SEIR, FRED, writing; AW - conception, formulation, SEIR, writing; SN - conception, formulation, FRED, writing; CW - conception, formulation, FRED, writing; VM - conception, formulation, FRED, writing; WH - conception, formulation, SEIR, writing; AS - conception, formulation, writing; BB - conception, formulation, writing; DC - FRED; JG - FRED, writing; SAN - conception, writing.

ACKNOWLEDGMENTS

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canada CIFAR AI Chairs Program, and the Intel Parallel

Computing Centers program. This material is based upon work supported by the United States Air Force Research Laboratory (AFRL) under the Defense Advanced Research Projects Agency (DARPA) Data Driven Discovery Models (D3M) program (Contract No. FA8750-19-2-0222). Additional support was provided by UBC's Composites Research Network (CRN), Data Science Institute (DSI), Support for Teams to Advance Interdisciplinary Research (STAIR) Grants, the Innovation for Defence Excellence and Security (IDEaS) program, under its COVID-19 Challenges (CPCA-0397), and a CIFAR AI and COVID-19 Catalyst Grant. This research was enabled in part by technical support and computational resources provided by WestGrid (<https://www.westgrid.ca/>) and Compute Canada (www.computecanada.ca). This article was submitted as a preprint to arXiv: Wood et al. 2020, available at <https://arxiv.org/abs/2003.13221>.

REFERENCES

- Adam, L., and DeJong, G. (2003). "The Influence of Reward on the Speed of Reinforcement Learning: An Analysis of Shaping," in Proceedings of the 20th International Conference on Machine Learning (ICML-03), Washington, DC, August 21–24, 2003, 440–447.
- Anastassopoulou, C., Russo, L., Tsakris, A., and Siettos, C. (2020). Data-based Analysis, Modeling and Forecasting of the Novel Coronavirus (2019-nCoV) Outbreak. medRxiv [Preprint]. doi:10.1101/2020.02.11.20022186
- Anna, B., Gerardin, J., Bridenbecker, D., Lorton, C. W., Bloedow, J., Baker, R. S., et al. (2018). Implementation and Applications of EMOD, An Individual-Based Multi-Disease Modeling Platform. *Pathog. Dis.* 76 (5), fty059. doi:10.1093/femspd/fty059
- Arenas, A., Cota, W., Gómez-Gardeñes, J., Gómez, S., Granell, C., Matamalas, J. T., et al. (2020). A Mathematical Model for the Spatiotemporal Epidemic Spreading of COVID19. medRxiv [Preprint]. doi:10.1101/2020.03.21.20040022
- Badham, J., Chattoe-Brown, E., Gilbert, N., Chalabi, Z., Kee, F., and Hunter, R. F. (2018). Developing Agent-Based Models of Complex Health Behaviour. *Health & Place* 54, 170–177. doi:10.1016/j.healthplace.2018.08.022
- Baydin, A. G., Heinrich, L., Bhimji, W., Gram-Hansen, B., Louppe, G., Shao, L., et al. (2018). "Efficient Probabilistic Inference in the Quest for Physics Beyond the Standard Model," in Thirty-second Conference on Neural Information Processing Systems (NeurIPS), Vancouver, British Columbia, December 8–14, 2018.
- Baydin, A. G., Shao, L., Bhimji, W., Heinrich, L., Meadows, L., Liu, J., et al. (2019). "Etalumis: Bringing Probabilistic Programming to Scientific Simulators at Scale," in Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, Denver, USA, November 17–22, 2018, 1–24.
- Bertsekas, D. P. (2019). *Reinforcement Learning and Optimal Control*. Belmont, MA: Athena Scientific.
- Bhalchandra, S. P., and SnehalShekatkar, M. (2020). Multi-city Modeling of Epidemics Using Spatial Networks: Application to 2019-nCov (COVID-19) Coronavirus in India. medRxiv [Preprint]. doi:10.1101/2020.03.13.20035386
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- Blackwood, J. C., and Childs, L. M. (2018a). An Introduction to Compartmental Modeling for the Budding Infectious Disease Modeler. *Lett. Biomathematics* 5 (1), 195–221. doi:10.1080/23737867.2018.1509026
- Blackwood, J. C., and Childs, L. M. (2018b). An Introduction to Compartmental Modeling for the Budding Infectious Disease Modeler. *Lett. Biomathematics* 5 (1), 195–221. doi:10.30707/lib5.1blackwood
- Boldog, P., Tekeli, T., Vizi, Z., Dénes, A., Bartha, F. A., and Röst, G. (2020). Risk Assessment of Novel Coronavirus Covid-19 Outbreaks Outside China. *Jcm* 9 (2), 571. doi:10.3390/jcm9020571
- Caccavo, D. (2020). Chinese and Italian COVID-19 Outbreaks Can Be Correctly Described by a Modified SIRD Model. medRxiv [Preprint]. doi:10.1101/2020.03.19.20039388
- Camacho, E. F., and Alba, C. B. (2013). *Model Predictive Control*. Springer science & business media.
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., et al. (2017). Stan: A Probabilistic Programming Language. *J. Stat. Softw.* 76 (1), 1–32. doi:10.18637/jss.v076.i01
- Chatzilena, A., van Leeuwen, E., Ratmann, O., Baguelin, M., and Demiris, N. (2019). Contemporary Statistical Inference for Infectious Disease Models Using Stan. *Epidemics* 29, 100367. doi:10.1016/j.epidem.2019.100367
- Ferguson, N. M., Laydon, D., Nedjati-Gilani, G., Imai, N., Ainslie, K., Baguelin, M., et al. (2020). *Impact of Non-pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand*. Available at: <https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf>.
- Funk, S., and King, A. A. (2020). Choices and Trade-Offs in Inference with Infectious Disease Models. *Epidemics* 30, 100383. doi:10.1016/j.epidem.2019.100383
- García, C. E., Prett, D. M., and Morari, M. (1989). Model Predictive Control: Theory and Practice-A Survey. *Automatica* 25 (3), 335–348. doi:10.1016/0005-1098(89)90002-2
- Gelman, A., Stern, H. S., Carlin, J. B., Dunson, D. B., Aki, V., and Rubin, D. B. (2013). *Bayesian Data Analysis*. Chapman and Hall/CRC.
- Germann, T. C., Gao, H., Gambhir, M., Plummer, A., Biggerstaff, M., Reed, C., et al. (2019). School Dismissal as a Pandemic Influenza Response: When, Where and for How Long. *Epidemics* 28, 100348. doi:10.1016/j.epidem.2019.100348
- Ghahramani, Z. (2015). Probabilistic Machine Learning and Artificial Intelligence. *Nature* 521 (7553), 452–459. doi:10.1038/nature14541
- Goodman, N. D., and Stuhlmüller, A. (2014). *The Design and Implementation of Probabilistic Programming Languages*. Available at: <http://dippl.org> (Accessed March 25, 2020).
- Goodman, N. D., and Stuhlmüller, A. (2014). The Design and Implementation of Probabilistic Programming Languages. Available at: <http://dippl.org> (Accessed March 26, 2020).
- Gram-Hansen, B., Schröder de Witt, C., Rainforth, T., Torr, P. H. S., Teh, Y. W., and Baydin, A. G. (2019). "Hijacking Malaria Simulators with Probabilistic Programming," in AI for Social Good Workshop, International Conference on Machine Learning, Long Beach, June 9–15, 2019.
- Green, M., and Limebeer, D. J. N. (2012). *Linear Robust Control*. North Chelmsford, Massachusetts: Courier Corporation.
- Grefenstette, J. J., Brown, S. T., Rosenfeld, R., DePasse, J., Stone, N. T., Cooley, P. C., et al. (2013). FRED (A Framework for Reconstructing Epidemic Dynamics): An Open-Source Software System for Modeling Infectious Diseases and Control Strategies Using Census-Based Populations. *BMC public health* 13 (1), 940. doi:10.1186/1471-2458-13-940
- Hansen, L. P., and Sargent, T. J. (2001). Robust Control and Model Uncertainty. *Am. Econ. Rev.* 91 (2), 60–66. doi:10.1257/aer.91.2.60
- Hill, A., Levy, M., Xie, S., Sheen, J., Shinnick, J., Gheorghe, A., et al. (2020). Modeling COVID-19 Spread vs Healthcare Capacity. Available at: <https://alhill.shinyapps.io/COVID19seir/> (Accessed March 25, 2020).

- Hunter, E., Mac Namee, B., and Kelleher, J. (2018). An Open-Data-Driven Agent-Based Model to Simulate Infectious Disease Outbreaks. *PLOS ONE* 13 (12), e0208775. doi:10.1371/journal.pone.0208775
- Hunter, E., Mac Namee, B., and Kelleher, J. D. (2017). A Taxonomy for Agent-Based Models in Human Infectious Disease Epidemiology. *Jasss* 20 (3), 2. doi:10.18564/jasss.3414
- Jacob, M., Wang, Y., and Boyd, S. (2011). Receding Horizon Control. *IEEE Control Syst. Mag.* 31 (3), 52–65. doi:10.1109/CACSD.2010.5612665
- Jia, W., Han, K., Song, Y., Cao, W., Wang, S., Yang, S., et al. (2020). Extended SIR Prediction of the Epidemics Trend of COVID-19 in Italy and Compared with Hunan, China. medRxiv [Preprint]. doi:10.1101/2020.03.18.20038570
- Kaelbling, L. P., Littman, M. L., and Moore, A. W. (1996). Reinforcement Learning: A Survey. *Jair* 4, 237–285. doi:10.1613/jair.301
- Kappen, H. J., Gómez, V., and Opper, M. (2012). Optimal Control as a Graphical Model Inference Problem. *Mach Learn.* 87 (2), 159–182. doi:10.1007/s10994-012-5278-7
- Köppen, M. (2000). “The Curse of Dimensionality,” in 5th Online World Conference on Soft Computing in Industrial Applications (WSC5), September 4–18, 2000, 1, 4–8.
- Kumar, S., Piper, K., Galloway, D. D., Hadler, J. L., and Grefenstette, J. J. (2015). Is Population Structure Sufficient to Generate Area-Level Inequalities in Influenza Rates? an Examination Using Agent-Based Models. *BMC public health* 15 (1), 947. doi:10.1186/s12889-015-2284-2
- Kumar, S., Grefenstette, J. J., Galloway, D., Albert, S. M., and Burke, D. S. (2013). Policies to Reduce Influenza in the Workplace: Impact Assessments Using an Agent-Based Model. *Am. J. Public Health* 103 (8), 1406–1411. doi:10.2105/ajph.2013.301269
- Kurtzer, G. M., Sochat, V., and Bauer, M. W. (2017). Singularity: Scientific containers for mobility of compute. *PLoS One* 12, e0177459.
- Kypraios, T., Neal, P., and Prangle, D. (2017). A Tutorial Introduction to Bayesian Inference for Stochastic Epidemic Models Using Approximate Bayesian Computation. *Math. biosciences* 287, 42–53. doi:10.1016/j.mbs.2016.07.001
- Lauer, S. A., Grantz, K. H., Bi, Q., Jones, F. K., Zheng, Q., Meredith, H. R., et al. (2020). The Incubation Period of Coronavirus Disease 2019 (COVID-19) from Publicly Reported Confirmed Cases: Estimation and Application. *Ann. Intern. Med.* 172, 577–582. doi:10.7326/m20-0504
- Le, T. A., Kosiorok, A. R., Siddharth, N., Teh, Y. W., and Wood, F. (2019). “Revisiting Reweighted Wake-Sleep for Models with Stochastic Control Flow,” in Proceedings of the conference on Uncertainty in Artificial Intelligence (UAI). Available at: <https://arxiv.org/abs/1805.10469>.
- Le, T. A., Baydin, A. G., and Wood, F. (2017). “Inference Compilation and Universal Probabilistic Programming,” in Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS), volume 54 of Proceedings of Machine Learning Research (Fort Lauderdale, FL: PMLR), 1338–1348.
- Lee, B. Y., Brown, S. T., Bailey, R. R., Zimmerman, R. K., Potter, M. A., McGlone, S. M., et al. (2011). The Benefits to All of Ensuring Equal and Timely Access to Influenza Vaccines in Poor Communities. *Health Aff.* 30 (6), 1141–1150. doi:10.1377/hlthaff.2010.0078
- Lee, B. Y., Brown, S. T., Cooley, P., Potter, M. A., Wheaton, W. D., Voorhees, R. E., et al. (2010). Simulating School Closure Strategies to Mitigate an Influenza Epidemic. *J. Public Health Manag. Pract. JPHMP* 16 (3), 252–261. doi:10.1097/phh.0b013e3181ce594e
- Lekone, P. E., and Finkenstädt, B. F. (2006). Statistical Inference in a Stochastic Epidemic SEIR Model with Control Intervention: Ebola as a Case Study. *Biometrics* 62 (4), 1170–1177. doi:10.1111/j.1541-0420.2006.00609.x
- Lessler, J., Azman, A. S., Grabowski, M. K., Salje, H., and Rodriguez-Barraquer, I. (2016). Trends in the Mechanistic and Dynamic Modeling of Infectious Diseases. *Curr. Epidemiol. Rep.* 3 (3), 212–222. doi:10.1007/s40471-016-0078-4
- Levine, S. (2018). *Reinforcement Learning and Control as Probabilistic Inference: Tutorial and Review*. arXiv [Preprint]. Available at: <https://arxiv.org/abs/1805.00909> (Accessed May 2, 2018).
- Lewis, F. L., Vrabie, D., and Szymos, V. L. (2012). *Optimal Control*. John Wiley & Sons.
- Liu, F., Enanoria, W. T., Zippich, J., Blumberg, S., Harriman, K., Ackley, S. F., et al. (2015). The Role of Vaccination Coverage, Individual Behaviors, and the Public Health Response in the Control of Measles Epidemics: An Agent-Based Simulation for California. *BMC public health* 15 (1), 447. doi:10.1186/s12889-015-1766-6
- Liu, Q., Liu, Z., Zhu, J., Zhu, Y., Li, D., Gao, Z., et al. (2020). Assessing the Global Tendency of COVID-19 Outbreak. medRxiv [Preprint]. doi:10.1101/2020.03.18.20038224
- Liu, X., Hewings, G., Wang, S., Qin, M., Xiang, X., Zheng, S., et al. (2020). Modeling the Situation of COVID-19 and Effects of Different Containment Strategies in China with Dynamic Differential Equations and Parameters Estimation. medRxiv [Preprint]. doi:10.1101/2020.03.09.20033498
- Liu, Z., Magal, P., Seydi, O., and Webb, G. (2020). Predicting the Cumulative Number of Cases for the COVID-19 Epidemic in China from Early Data. arXiv [Preprint]. Available at: <https://arxiv.org/abs/2002.12298> (Accessed February 27, 2020).
- Magdon-Ismail, M. (2020). Machine Learning the Phenomenology of COVID-19 from Early Infection Dynamics. medRxiv [Preprint].
- Mamo, D. K., and Koya, P. R. (2015). Mathematical Modeling and Simulation Study of SEIR Disease and Data Fitting of Ebola Epidemic Spreading in West Africa. *J. Multidisciplinary Eng. Sci. Technol.* 2 (1), 106–114.
- Massonnaud, C., Roux, J., and Crépey, P. (2020). COVID-19: Forecasting Short Term Hospital Needs in France. medRxiv [Preprint]. doi:10.1101/2020.03.16.20036939
- McKinley, T. J., Vernon, I., Andrianakis, I., McCreesh, N., Oakley, J. E., Nsubuga, R. N., et al. (2018). Approximate Bayesian Computation and Simulation-Based Inference for Complex Stochastic Epidemic Models. *Stat. Sci.* 33 (1), 4–18. doi:10.1214/17-STS618
- Minter, A., and Retkute, R. (2019). Approximate Bayesian Computation for Infectious Disease Modeling. *Epidemics* 29, 100368. doi:10.1016/j.epidem.2019.100368
- Moerland, T. M., Broekens, J., and Jonker, C. M. (2021). Model-based Reinforcement Learning: A survey. arXiv [Preprint]. Available at: <https://arxiv.org/abs/2006.16712> (Accessed Jun 30, 2020).
- Morari, M., and Lee, J. H. (1999). Model Predictive Control: Past, Present and Future. *Comput. Chem. Eng.* 23 (4–5), 667–682. doi:10.1016/s0098-1354(98)00301-9
- Munk, A., Adam, Š., Baydin, A. G., Stewart, A., Fernlund, A., Poursartip, A., et al. (2020). “Deep Probabilistic Surrogate Networks for Universal Simulator Approximation,” in The second International Conference on Probabilistic Programming (PROBPROG), October 20–22, 2021.
- Murray, L. M. (2015). Bayesian State-Space Modeling on High-Performance Hardware Using LibBi. *J. Stat. Soft.* 67, 1–36. doi:10.18637/jss.v067.i10
- Naderiparizi, S., Šcibior, A., Munk, A., Ghadiri, M., Baydin, A. G., Gram-Hansen, B., et al. (2019). *Amortized Rejection Sampling in Universal Probabilistic Programming*. Available at: <https://arxiv.org/abs/1910.09056>.
- Ogilvy Kermack, W., and McKendrick, A. G. (1927). A Contribution to the Mathematical Theory of Epidemics. *Proc. R. Soc. Lond. Ser. A, Containing Pap. a Math. Phys. character* 115 (772), 700–721.
- Owen, A. B. (2013). *Monte Carlo Theory, Methods and Examples*.
- Paige, B., Wood, F., Doucet, A., and Teh, Y. W. (2014). “Asynchronous Anytime Sequential Monte Carlo,” in *Advances in Neural Information Processing Systems*, 3410–3418.
- Paige, B., and Wood, F. (2016). “Inference Networks for Sequential Monte Carlo in Graphical Models,” in Proceedings of the 33rd International Conference on Machine Learning (Cambridge, MA: PMLR).
- Parker, J., and Epstein, J. M. (2011). A Distributed Platform for Global-Scale Agent-Based Models of Disease Transmission. *ACM Trans. Model. Comput. Simul.* 22 (1), 1–25. doi:10.1145/2043635.2043637
- Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. New York, NY: Cambridge University Press.
- Peng, L., Yang, W., Zhang, D., Zhuge, C., and Hong, L. (2020). *Epidemic Analysis of COVID-19 in China by Dynamical Modeling*. medRxiv [Preprint]. doi:10.1101/2020.02.16.20023465
- Potter, M. A., Brown, S. T., Cooley, P. C., Sweeney, P. M., Hershey, T. B., Gleason, S. M., et al. (2012). School Closure as an Influenza Mitigation Strategy: How Variations in Legal Authority and Plan Criteria Can Alter the Impact. *BMC public health* 12 (1), 977–1011. doi:10.1186/1471-2458-12-977
- Pujari, B. S., and Shekatkar, S. (2020). *Multi-city Modeling of Epidemics Using Spatial Networks: Application to 2019-nCoV (COVID-19) Coronavirus in India*. medRxiv [Preprint]. doi:10.1101/2020.03.13.20035386
- Rainforth, T., Cornish, R., Yang, H., Warrington, A., and Wood, F. (2018). “On Nesting Monte Carlo Estimators,” in *International Conference on Machine Learning* (Stockholmsmässan: PMLR), 4267–4276.
- Rainforth, T., Naesseth, C. A., Lindsten, F., Paige, B., van de Meent, J. W., Doucet, A., et al. (2016). “Interacting Particle Markov Chain Monte Carlo,” in Proceedings of the 33rd International Conference on Machine Learning (Cambridge, MA: PMLR).
- Rawlings, J. B. (2000). Tutorial Overview of Model Predictive Control. *IEEE Control Syst. Mag.* 20 (3), 38–52.

- Riou, J., and Althaus, C. L. (2020). Pattern of Early Human-To-Human Transmission of Wuhan 2019 Novel Coronavirus (2019-nCoV), December 2019 to January 2020. *Euro Surveill.* 25 (4), 2020. doi:10.2807/1560-7917.ES.2020.25.4.2000058
- Roberts, M., Andreasen, V., Lloyd, A., and Pellis, L. (2015). Nine Challenges for Deterministic Epidemic Models. *Epidemics* 10, 49–53. doi:10.1016/j.epidem.2014.09.006
- Rovetta, A., Bhagavathula, A. S., and Castaldo, L. (2020). *Modeling the Epidemiological Trend and Behavior of COVID-19 in Italy*. medRxiv [Preprint]. doi:10.1101/2020.03.19.20038968
- Russo, L., Anastassopoulou, C., Tsakris, A., Bifulco, G. N., Campana, E. F., Toraldo, G., et al. (2020). Tracing Day-Zero and Forecasting the COVID-19 Outbreak in Lombardy, Italy: A Compartmental Modeling and Numerical Optimization Approach. *PLoS One* 15, e0240649. doi:10.1371/journal.pone.0240649
- Sanche, S., Lin, Y. T., Xu, C., Romero-Severson, E., Hengartner, N. W., and Ruian, Ke. (2020). *The Novel Coronavirus, 2019-nCoV, Is Highly Contagious and More Infectious than Initially Estimated*. arXiv [Preprint]. Available at: <https://arxiv.org/abs/2002.03268> (Accessed Feb 9, 2020).
- Saunders-Hastings, P., CrispoCrispo, J. A. G., Sikora, L., and Krewski, D. (2017). Effectiveness of Personal Protective Measures in Reducing Pandemic Influenza Transmission: A Systematic Review and Meta-Analysis. *Epidemics* 20, 1–20. doi:10.1016/j.epidem.2017.04.003
- Sharomi, O., and Malik, T. (2017). Optimal Control in Epidemiology. *Ann. Operations Res.* 251 (1–2), 55–71. doi:10.1007/s10479-015-1834-4
- Staton, S., Wood, F., Yang, H., Heunen, C., and Kammar, O. (2016). “Semantics for Probabilistic Programming: Higher-Order Functions, Continuous Distributions, and Soft Constraints,” in 2016 31st Annual ACM/IEEE Symposium on Logic in Computer Science (LICS) (New York: ACM), 1–10.
- Sutton, R. S., and Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. MIT press.
- Sutton, R. (1992). “Reinforcement Learning,” in *The Springer International Series in Engineering and Computer Science* (Springer US).
- Teles, P. (2020). Predicting the Evolution of COVID-19 in Portugal Using an Adapted SIR Model Previously Used in South Korea for the MERS Outbreak. medRxiv [Preprint]. doi:10.1101/2020.03.18.20038612
- Todorov, E. (2008). “General Duality Between Optimal Control and Estimation,” in 2008 47th IEEE Conference on Decision and Control, Cancun, Mexico, December 9–11, 2008 (IEEE), 4286–4292. doi:10.1109/cdc.2008.4739438
- Tolpin, D., van de Meent, J.-W., Frank, P., and Wood, B. (2015). Output-Sensitive Adaptive Metropolis-Hastings for Probabilistic Programs. *ECML PKDD*, 311–326. doi:10.1007/978-3-319-23525-7_19
- Tolpin, D., van de Meent, J. W., Yang, H., and Wood, F. (2016). “Design and Implementation of Probabilistic Programming Language Anglican,” in Proceedings of the 28th Symposium on the Implementation and Application of Functional programming Languages, Leuven, August 31–September 2, 2016, 1–12. doi:10.1145/3064899.3064910
- Toni, T., Welch, D., Strelkowa, N., Ipsen, A., and Stumpf, M. P. H. (2009). Approximate Bayesian Computation Scheme for Parameter Inference and Model Selection in Dynamical Systems. *J. R. Soc. Interf.* 6 (31), 187–202. doi:10.1098/rsif.2008.0172
- Toussaint, M. (2009). “Robot Trajectory Optimization Using Approximate Inference,” in Proceedings of the 26th annual international conference on machine learning, 1049–1056. doi:10.1145/1553374.1553508
- Tracy, M., Cerdá, M., and Keyes, K. M. (2018). Agent-based Modeling in Public Health: Current Applications and Future Directions. *Annu. Rev. Public Health* 39 (1), 77–94. doi:10.1146/annurev-publhealth-040617-014317
- Traini, M. C., Caponi, C., and De Socio, G. V. (2020). Modeling the Epidemic 2019-nCoV Event in Italy: A Preliminary Note. medRxiv [Preprint]. doi:10.1101/2020.03.14.20034884
- van de Meent, J.-W., Yang, H., Mansinghka, V., and Wood, F. (2015). “Particle Gibbs with Ancestor Sampling for Probabilistic Programs,” in *Artificial Intelligence and Statistics* (Cambridge, MA: PMLR).
- van de Meent, J. W., Brooks, P., Yang, H., and Wood, F. (2018). *An Introduction to Probabilistic Programming*. Available at: <https://arxiv.org/abs/1809.10756>.
- Vinter, R. B., and Vinter, R. B. (2010). *Optimal Control*. Springer.
- Warrington, A., Lavington, J. W., Adam, S., Schmidt, M., and Wood, F. (2021). “Robust Asymmetric Learning in POMDPs,” in International Conference on Machine Learning, July 18–24, 2021 (Cambridge, MA: PMLR), 11013–11023.
- Warrington, A., Naderiparizi, S., and Wood, F. (2020). “Coping with Simulators that Don’t Always Return,” in The 23rd International Conference on Artificial Intelligence and Statistics (AISTATS 2020), August 26–28, 2020.
- Wearing, H. J., Rohani, P., and Keeling, M. J. (2005). Appropriate Models for the Management of Infectious Diseases. *Plos Med.* 2 (7), e174. doi:10.1371/journal.pmed.0020174
- Weber, A., Ianelli, F., and Goncalves, S. (2020). Trend Analysis of the COVID-19 Pandemic in China and the Rest of the World. arXiv [Preprint]. Available at: <https://arxiv.org/abs/2003.09032> (Accessed March 19, 2020).
- Wen, Y., Wei, L., Li, Y., Tang, X., Feng, S., Leung, K., et al. (2020). Epidemiological and Clinical Characteristics of COVID-19 in Shenzhen, the Largest Migrant City of China. medRxiv [Preprint]. doi:10.1101/2020.03.22.20035246
- Wheaton, W., Cajka, J., Chasteen, B., Wagener, D., Cooley, P., Ganapathi, L., et al. (2009). Synthesized Population Databases: A Us Geospatial Database for Agent-Based Models. *Methods Report (RTI Press)*. 2009 (10), 905. doi:10.3768/rtipress.2009.mr.0010.0905
- Wingate, D., Stuhlmüller, A., and Goodman, N. (2011). “Lightweight Implementations of Probabilistic Programming Languages via Transformational Compilation,” in Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, Fort Lauderdale, FL, April 1–13, 2011, 770–778.
- Wölfel, R., Corman, V. M., Guggemos, W., Seilmaier, M., Zange, S., Müller, M. A., et al. (2020). Clinical Presentation and Virological Assessment of Hospitalized Cases of Coronavirus Disease 2019 in a Travel-Associated Transmission Cluster. medRxiv [Preprint]. doi:10.1101/2020.03.05.20030502
- Wood, F., van de Meent, J. W., and Mansinghka, V. (2014). “A New Approach to Probabilistic Programming Inference,” in *Artificial Intelligence and Statistics*, 1024–1032.
- Wood, F., Warrington, A., Naderiparizi, S., Weilbach, C., Masrani, V., Harvey, W., et al. (2020). Planning As Inference In Epidemiological Models. arXiv preprint arXiv:2003.13221.
- World Bank (2020). Hospital Beds Per Capita in the united states. Available at: <https://data.worldbank.org/indicator/SH.MED.BEDS.ZS?locations=US> (Accessed March 25, 2020).
- Wu, Z., and McGoogan, J. M. (2020). Characteristics of and Important Lessons from the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases from the Chinese Center for Disease Control and Prevention. *JAMA* 323 (13), 1239–1242. doi:10.1001/jama.2020.2648
- Zhou, K., and Doyle, J. C. (1998). *Essentials of Robust Control*, 104. Upper Saddle River, NJ: Prentice-Hall.
- Zhou, Y., Gram-Hansen, B. J., Kohn, T., Rainforth, T., Yang, H., and Frank, W. L. F.-P. P. L. (2019). “A Low-Level First Order Probabilistic Programming Language for Non-differentiable Models,” in Proceedings of the Twentieth International Conference on Artificial Intelligence and Statistics (AISTATS), Naha, Okinawa, April 16–18, 2019.

Disclaimer: Frontiers Media SA remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Conflict of Interest: Authors JG and DC were employed by the company Epistemic Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Wood, Warrington, Naderiparizi, Weilbach, Masrani, Harvey, Scibior, Beronov, Grefenstette, Campbell and Nasser. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.