

When Do We Need Massive Computations to Perform Detailed COVID-19 Simulations?

Christopher B. Lutz and Philippe J. Giabbanelli*

The COVID-19 pandemic has infected over 250 million people worldwide and killed more than 5 million as of November 2021. Many intervention strategies are utilized (e.g., masks, social distancing, vaccinations), but officials making decisions have a limited time to act. Computer simulations can aid them by predicting future disease outcomes, but they also require significant processing power or time. It is examined whether a machine learning model can be trained on a small subset of simulation runs to inexpensively predict future disease trajectories resembling the original simulation results. Using four previously published agent-based models (ABMs) for COVID-19, a decision tree regression for each ABM is built and its predictions are compared to the corresponding ABM. Accurate machine learning meta-models are generated from ABMs without strong interventions (e.g., vaccines, lockdowns) using small amounts of simulation data: the root-mean-square error (RMSE) with 25% of the data is close to the RMSE for the full dataset (0.15 vs 0.14 in one model; 0.07 vs 0.06 in another). However, meta-models for ABMs employing strong interventions require much more training data (at least 60%) to achieve a similar accuracy. In conclusion, machine learning meta-models can be used in some scenarios to assist in faster decision-making.

observed worldwide and continue to require action by government officials to limit the spread of disease.^[7] For example, France has implemented a COVID-19 “health pass” while Italy has a similar “Green Pass”; mandates for masks or vaccines are on the agenda across several US states; and lockdown as well as travel curbs are making a comeback in China. A broad set of intervention strategies is available to policymakers,^[8–11] including vaccines, preventative care (e.g., masks and social distancing), or lockdowns (e.g., remote work and travel restrictions). Implementing one of these strategies involves several logistical parameters: for instance, testing requires capacity for contact tracing and policies for quarantine (e.g., is a negative test required to leave quarantine? is a number of days required?); similarly, vaccination involves a complex logistical chain from shipping to administering doses.^[12]

In addition to the many possible combinations of interventions and parameter values, the “right set” of interventions can vary across places (e.g., based on disease incidence and hospital capacity), across

1. Introduction

As of November 2021, COVID-19 was directly responsible for an estimate of over 5 million deaths and over 250 million cases.^[1] Taking the United States as an example, these numbers translate to almost 750 000 deaths from about 46 million cases,^[2] while noting that fatality may be underestimated^[3] and differs across sub-groups based on factors such as socio-economic status or race and ethnicity.^[4] Although reports previously considered that “transition toward normalcy in the United States remains most likely in the second quarter of 2021,”^[5] the delta variant has effectively triggered a “new phase in the pandemic”^[6] as can be seen with a rebound of over 100 000 new cases per day in the United States in August and September.^[2] Similar phenomena can be

time (e.g., in response to a new wave of infections), and across individuals (e.g., priority for vaccination to those most at risk). This results in a very large search space of possible interventions.^[13] Although the necessity of immediate actions in the early days of the pandemic may have resulted in choosing an intervention based on minimal insight, there is now evidence for serious consequences in rolling a sub-optimal intervention: lives may be lost, the cost of future interventions may be heightened, or the adherence (hence the impact) of future interventions may be lowered.^[14–16]

Computer simulations using agent-based models (ABMs) can aid officials in making these decisions, by modeling the effects of specific interventions in specific places (e.g., small towns,^[17] educational institutions,^[18,19] supermarkets^[20]), populations (e.g., targeted vaccinations^[21]), or time windows (e.g., during a yearly mass pilgrimage^[22]). Several modeling frameworks are available to quickly run simulations for specific interventions, while accounting for individual heterogeneity in risk factors and contact patterns (e.g., by embedding agents across multiple networks such as community and work). However, significant computing resources are required to perform detailed simulations that track individual transmissions and evaluate various interventions.^[23,24] Our analysis across six projects showed that

C. B. Lutz, P. J. Giabbanelli
Department of Computer Science & Software Engineering
Miami University
205 Benton Hall, Oxford, OH 45056, USA
E-mail: lutzcb@miamioh.edu; giabbapj@miamioh.edu

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/adts.202100343>

DOI: 10.1002/adts.202100343

cloud computing or high-performance computing clusters were frequently needed.^[25] This is exemplified by the model from Chang et al., which ran on 4264 compute cores.^[23]

So far, the primary solution to perform resource-intensive agent-based COVID-19 simulations has been to find more computing power. For example, the COVID-19 High-Performance Computing Consortium (covid19-hpc-consortium.org) was created to make private computing resources available to COVID-19 researchers.^[26] Similarly, the Partnership for Advanced Computing in Europe (PRACE) offered a fast-track mechanism for access to supercomputers (prace-ri.eu/hpc-access/hpcsvirus), and national laboratories issues calls for rapid-response research.^[27] These computing requirements can be limiting in the context of pandemic responses, because officials are forced to wait for results before acting to prevent the spread of the disease. It also stresses inequities in simulation research, as some groups may struggle to perform their simulations in a timely manner due to the lack of resources at smaller research organizations.^[28]

In this paper, we examine whether massive computations are an absolute requirement to support decision-makers in comprehensively examining the expected consequences of various intervention scenarios in the context of COVID-19. Specifically, we assess whether we can replace a (computationally expensive) COVID-19 ABM with a (cheaper) machine learning model. The core idea is to assess whether it is possible to perform just enough simulations to train an accurate machine learning model, which can then be used to predict the remaining simulation results inexpensively. Our approach consists of generating data from four previously published and validated ABMs for COVID-19^[12,24,29,30] and using the data to train machine learning regression meta-models. By varying the amount of data used to train these meta-models, we characterize the relationship between how much data is used to train the meta-model and how accurate that meta-model is. Our specific contributions are as follows:

- We develop machine learning regression meta-models for four COVID-19 ABMs to predict the total proportion of the population that will become infected, in response to the intervention scenarios captured by the model's parameters.
- We examine the affects that different amounts of training data have on the overall accuracy of those meta-models, thus establishing the situations under which a COVID-19 ABM may require computing power to achieve accurate results.

The remainder of this paper is organized as follows. In Section 2, we summarize the key features of COVID-19 and interventions that are available in the four ABMs used in this study. Our background also briefly explains how ABMs are created for COVID-19 and how machine learning can be used to generate simulation meta-models. Then in Section 3, we cover our methods starting with the implementation and verification of the selected ABMs, and then detailing how we performed our machine learning regressions. In Section 4, we analyze the results produced by the verification data for our ABMs and the machine learning regressions. Next in Section 5, we discuss the interpretation and limitations of our results. Finally in Section 6, we provide concluding remarks

and suggest future work that could be undertaken based on our results.

2. Background

In this section, we will examine key details of COVID-19 and available interventions, as well as how ABMs are created for it. Then, we summarize how machine learning has been used previously to create meta-models for simulations.

2.1. COVID-19

COVID-19, caused by the SARS-CoV-2 virus, was first reported in Wuhan, China in December 2019^[31] with newer studies suggesting a first case as far back as mid-November 2019.^[32] The virus spreads through droplets that are released when an infected individual coughs or sneezes. As a respiratory disease, the primary mode of transmission is thus via exposure to droplets, either indirectly (e.g., via contact through contaminated objects or hands) or directly (air borne).^[33,34] The virus infects cells in an individual's lungs, interfering with the lungs' ability to function properly.^[35] Symptoms for the disease include fever, loss of smell, or cough.^[36,37] A systematic review of 45 studies reported that 73% of individuals experience at least one persistent symptom.^[38] Most commonly, symptoms such as fatigue, sleep disorders, or loss of smell can be experienced for months.^[39] Less common consequences include multi-organ damage,^[40] for example, in the the cardiovascular system,^[41] kidneys,^[42] nervous system,^[43] or immune system.^[44]

Prior to the development of vaccines, all interventions for COVID-19 were necessarily non-pharmaceutical. These interventions included the use of masks, social distancing, regional lockdowns, and contact tracing. Masks reduce the spread of COVID-19 by lowering the potential for infected particles from entering the environment, and higher mask compliance leads to more effective disease mitigation.^[45] Social distancing and the closing of restaurants, gyms, and other public locations led to a statistically significant reduction in the spread of COVID-19 as well.^[46] Contact tracing, which helps officials control the spread of the virus by tracking who may have been in contact with an infected individual, has also been found to be effective.^[47]

In December of 2020, the first pharmaceutical intervention (i.e., vaccine) for COVID-19 became available.^[48] The messenger RNA (mRNA) vaccines for COVID-19 contain a small piece of the SARS-CoV-2 virus's mRNA that triggers the immune system to start producing an immune response to the virus. The FDA found that the Pfizer vaccine was 95% effective against COVID-19,^[48] but effectiveness may be reduced by mutations.^[49] Mutations known as "variants of concern" impact aspects such as transmissibility and vaccine effectiveness.^[50] In particular, some variants can evade antibodies caused by infection and current vaccines,^[51] which calls for the development of next-generation vaccines and antibody therapies.^[52] The existence of repeat infections^[53] as well as breakthrough infections^[54,55] (i.e., infection of a fully vaccinated person) have led to question the possibility of herd immunity,^[56] that is, the assumption that most transmission would be blocked if a given threshold of the population gains immunity.

2.2. Agent-Based Models for COVID-19

In early 2020, given the urgency of the COVID-19 pandemic and the scarcity of data, the first generation of COVID-19 models was based on a compartmental approach in which individuals are aggregated into groups (e.g., susceptible, infected) and a simulation proceeds by applying flow equations between groups. These classic compartmental epidemiological models focused on estimating key epidemiological quantities such as R_0 , the number of new cases generated by each infected individual.^[57,58] Although these models were imperfect in different ways,^[59-61] expert predictions have still outperformed the general public^[62] and their models “have influenced public health policies and increased the familiarity of the general public as well as policymakers with the modeling process, its value, and its limitations.”^[25] The most commonly used epidemic models categorized the population into susceptible, infected, and recovered (SIR) or added an intermediate stage for exposure (SEIR).^[63] A review of models produced between January and November 2020 found that the SIR and SEIR approaches represented 46.1% of all models, whereas ABM only accounted for 1.3% of studies at the time.^[64]

As the evidence base progressed, modelers realized that the assumptions of compartmental models (e.g., treating individuals as part of homogeneous groups) were limiting. In the words of Tolk et al., “As the pandemic unfolded, it quickly became evident these were not valid assumptions: the virus does not impact all populations evenly, and the interaction among different groups is far from even.”^[65] The rationale for the use of

individual-level simulation models such as ABM thus centers on the notion of heterogeneity: heterogeneity of risk factors for individuals (e.g., age and underlying or “pre-existing” conditions that increase the risk of severe illness conditions^[66,67]), heterogeneity of behaviors (e.g., interest in vaccination,^[68] compliance with mask policies^[69,70]), heterogeneity in socio-ecological vulnerability across places (e.g., lower access to resources in rural counties,^[71] urban sub-populations at risk^[72]), and heterogeneity in contact patterns.^[73,74] Although the need was clear, the development of ABMs was initially challenged by a lack of data, limited understanding of the disease, and occasionally a limited skillset.^[25] The situation has changed with the growing evidence base on COVID-19, the availability of mobility data, and the development of reusable frameworks to instantiate ABMs specifically for COVID-19 (e.g., COVASIM,^[75] OpenABM-Covid19,^[76] various modeling pipelines^[77]).

An ABM represents individuals (as agents) and their interactions with the environment as well as other agents. Each agent can have a state along with additional characteristics such as age or position within social networks. States in several ABMs are inspired by classic compartmental epidemiological models, hence it is common to use a SEIR model to represent the progression of each agent.^[29,78,79] Although earlier models may have relied on only four stages (SEIR), newer ones have introduced new stages to account for vaccination in two doses, disease severity, and the difference between asymptomatic and symptomatic individuals (Figure 1). In contrast with their mathematical predecessors, ABMs include the movements of agents over a space (e.g., a grid)

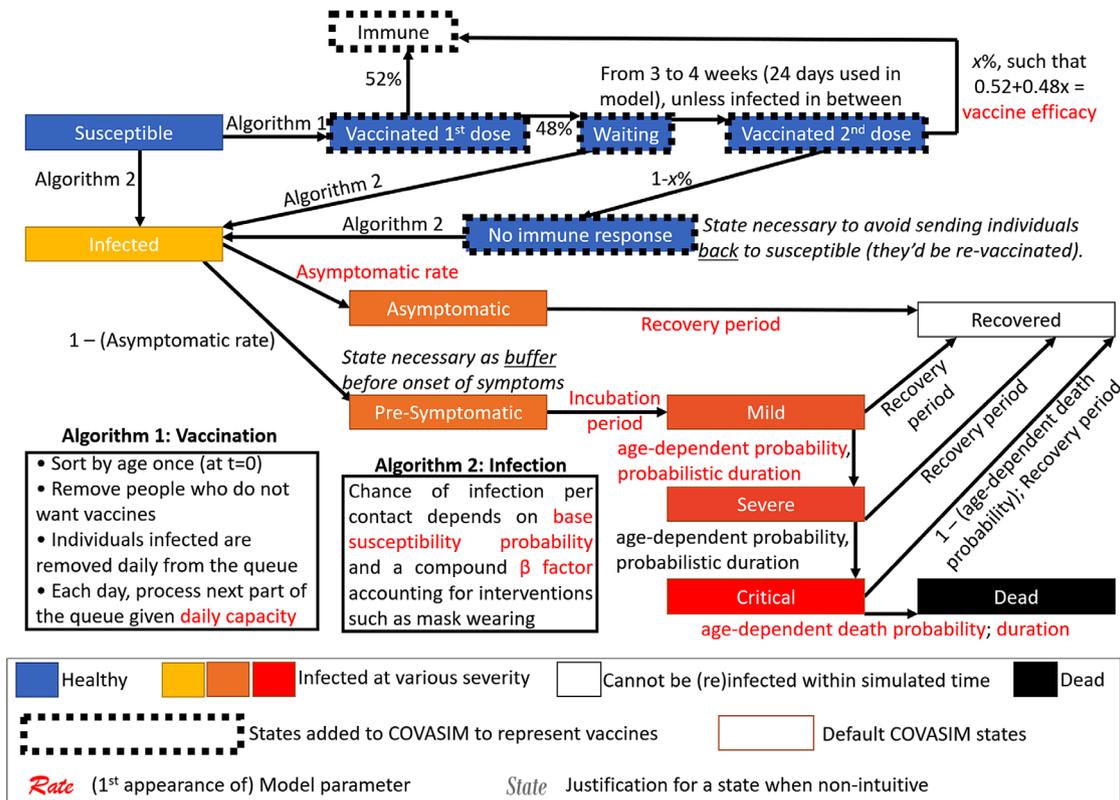


Figure 1. Flow diagram describing the possible journey of an agent, in the study of Li and Giabbanelli.^[12] Additional model logic is governed by algorithms, to identify agents who will be vaccinated or infected.

Table 1. Characteristics of the agent-based models.

Characteristic	Li ^[12]	Shamil ^[24]	Silva ^[29]	Badham ^[30]
Model states				
Healthy (susceptible)	Yes	Yes	Yes	Yes
Infected (not contagious)	Yes	Yes		Yes
Infected (contagious, asymptomatic)	Yes	Yes	Yes	
Infected (contagious, symptomatic)	Yes	Yes	Yes	Yes
Hospitalized			Yes	Yes
Mild infection	Yes			
Severe infection	Yes		Yes	
Critical infection	Yes			Yes
Recovered (immune)	Yes	Yes	Yes	Yes
Dead	Yes	Yes	Yes	Yes
Vaccinated (first dose)	Yes			
Vaccinated (waiting)	Yes			
Vaccinated (second dose)	Yes			
Fully vaccinated	Yes			
Model characteristics				
Time resolution	Days	Hours	Hours	Days
Work network	Yes	Yes	Yes	
School network	Yes	Yes		
Community network	Yes	Yes	Yes	Yes
Home network	Yes	Yes	Yes	
Number of agents	656 000	10 000	Varies	20 172

Table 2. Examples of wallclock time across hardware configurations for the four COVID-19 ABMs considered here.

Model	Wallclock time per simulation run [s]	Hardware setup
Li ^[12]	Between 900 and 1200	Cloud computing (Azure)
Shamil ^[24]	32 492.86 on average	High-performance computing cluster
Silva ^[29]	727.12 on average	High-performance computing cluster
Badham ^[30]	34.63 on average	Personal computer

and interactions with other agents or their environment (e.g., contaminated surfaces), which leads to the spread of the simulated disease. One of the factors that makes ABM agents unique is their ability to model individual characteristics such as age or other risk factors in the context of infectious disease.

States and characteristics found in the four chosen models for our study are expanded in **Table 1**. Although each model has been previously published with a detailed technical description, we provide a succinct overview of each model in the subsequent subsections to keep this manuscript self-contained. Since a key motivation for our approach (and the use of meta-modeling in general) is to create a cheaper proxy to a computationally expensive model, we exemplify the wallclock time typically required by the four chosen models, across platforms (**Table 2**). These platforms show the diversity of hardware that end users of models can access, from cloud-computing services (Microsoft Azure with AMD

EPYC platform for Li and Giabbanelli^[12]) and high-performance computing clusters (Intel Xeon processors with 96 Gb of memory for Shamil et al.^[24] and Silva et al.^[29]) to personal computers (AMD Ryzen with 16 Gb of memory for Badham et al.^[30]).

2.2.1. Li and Giabbanelli Model

The ABM developed by Li and Giabbanelli^[12] is built on the Covasim framework, by including several modifications that introduce vaccination. This ABM uses 656 000 agents and simulates the spread of COVID-19 for 180 days. This model contains states for susceptible agents, as well as several different states of infection, including asymptomatic infection, and three other levels of infection from mild to critical (Figure 1). It also includes the ability to simulate two vaccine doses for agents, which can remove them from the pool of susceptible agents. In addition to vaccine support, the model also inherits the ability to handle quarantines and contact tracing from Covasim. Finally, the Covasim platform embeds agents across different networks (e.g., community, work, school) to model how interventions have different impacts across settings (e.g., social distancing in the community, face masks at work).

2.2.2. Shamil et al. Model

The ABM created by Shamil et al.^[24] includes two possible agent configurations for different cities in the United States. One such configuration is a simplified representation for New York City, which contains 10 000 agents simulated for 90 days. It contains states representing different states of contagion and symptoms, with agents initially healthy and transitioning through a non-contagious state into a contagious asymptomatic state, then into a contagious symptomatic state, and finally into a recovered or dead state. This model includes representations of large-scale gatherings and various other day-to-day activities that could spread the virus, such as attending work or school. This ABM provides support for contact tracing and quarantines.

2.2.3. Silva et al. Model

The ABM released by Silva et al.^[29] models agents in infected states, as well as different levels of virus severity and hospitalization. Age and risk factors are also included for each agent, and these can be used when determining a lockdown policy. This simulation models the interactions of different agents from different home environments as well, and allows for lockdowns that prevent agents from traveling and interacting. This ABM provides support for quarantines, lockdowns, and masks.

2.2.4. Badham et al. Model

The ABM built by Badham et al.^[30] models agents in infected states, and contains states for hospitalization and critical infection. Community spread is represented as the agents move around their environment, although movement of agents can be restricted by several policies. This ABM provides support for social distancing, hospital simulation, isolation, and movement restrictions.



Figure 2. Flowchart for the process of training a meta-model.

2.3. Simulation Meta-Modeling

Meta-models, also known as surrogate models, are approximations of a more complex model.^[80] Although an approximation may be less accurate, this is usually tolerated in exchange for a significant improvement (i.e., reduction) in computational cost such as wallclock time.^[80] For example, a meta-model for hydrodynamic and thermal simulation reduced compute time by 93% while only reducing accuracy by 4%.^[81] In another simulation of social networks, simplifications resulted in an 85% decrease in runtime and a 32% decrease in memory requirements.^[82]

As the notion of “models” can be broadly conceptualized across fields, it is important to distinguish two settings. In pure mathematics, meta-models are mathematical functions that approximate the output of another, more complex mathematical function. In the simulation of interest here, meta-models are models that predict the results of a simulation with less computational requirements. Machine learning is one approach to create simulation meta-models. One common type of machine learning meta-model is a regression model,^[83–85] which is appropriate when the output of the simulation model (which we seek to approximate) is a discrete value based on its input parameters. Another common type of machine learning meta-model is a classification model, which produces a class (i.e., a group with similar characteristics) based on input parameters.^[86] In this paper, we focus on regressions, whereby the objective is to provide an accurate but faster proxy to the final result of an expensive simulation.

Figure 2 provides a high-level illustration for the process for constructing a simulation meta-model via machine learning. Gathering data for a meta-model involves running the simulation multiple times with varying input parameters. The parameters that should be varied during simulation runs are the desired meta-model input parameters.^[87] For example, to train a clinically relevant meta-model of the human immunodeficiency virus (HIV), a HIV simulation model may be run for several values of clinically relevant parameters such as the start and expected efficacy of treatment.^[88] The data produced by the simulation model then becomes the input or training data for the meta-model. The number of data points gathered for training depends on the computational requirements of the simulation itself.^[89] Once the meta-model has been trained, its results are compared against the simulation itself using an error metric such as the mean-square error (MSE), root-mean-square error (RMSE),^[87] or normalized RMSE.^[90]

Several sampling methods or “designed experiments” allow to produce the training data for the meta-model, that is, select values of the simulation parameters. Such techniques include random sampling, factorial sampling, and Latin hypercube sampling,^[91] which have a long track record when applied to simulations.^[92] Random sampling picks a certain number of entries or parameter values at random, which may lead to a cluster of points (i.e., oversampling in some areas) or an absence of points (i.e., under-

sampling). Factorial sampling explores every value of a parameter combined with every value of every other parameter.^[93] For example, if two different parameters could be either 0 or 1, the factorial set would be (0, 0), (0, 1), (1, 0), and (1, 1). Latin hypercube sampling uses each value for a parameter only once, but ensures an even spread over the domain of parameter values.^[94]

3. Experimental Section

In this section, we explain how we used each ABMs to perform the machine learning regressions. A table is provided to list the parameter values used in generating data from each model. When the parameter values require further explanations, a secondary table is also included; additional details can be found in the peer-reviewed manuscripts corresponding to each model. Apart from the Li and Giabbanelli model for which we already had data (as it was produced by our group), we accessed the public implementations for each of their other models, verified the code vis-a-vis the corresponding publication, and then produced the simulation data. A flowchart of the overall process is provided in **Figure 3**.

3.1. Data Generation and Collection

The datasets generated based on the following procedures can all be openly accessed by readers from our third-party repository, <https://osf.io/d7vqa/>. Our approach to data collection relied on a factorial design of experiments, thus gathering data for combinations of parameter values. We also assessed whether the models experienced bifurcations or noticeably different simulation paths, as each pathway in a model’s execution could then require dedicated data collection. As the models did not exhibit bifurcations (**Figure 4**), we focused on the creation of a comprehensive dataset in terms of combinations of parameter values.

3.1.1. Li and Giabbanelli Model

This model was produced by our research group and utilized as part of a factorial analysis (in the absence of vaccines)^[13] or via a grid search (with vaccines).^[12] It was thus the only case in which we had an existing and extensive simulated dataset to rely upon. For our meta-modeling purposes, we split the original dataset into three separate sets based on their vaccine plans: no vaccines in **Table 3** and the two vaccination plans in **Table 4**. The first subset of data captured simulations that did not include a vaccine, the second subset contained the simulations that used one vaccine plan (under the Trump administration), and the third subset contained was formed of vaccine simulations under the other vaccine plan (from the current Biden administration). These three sets of data were used to train three different meta-models, since the two vaccine datasets contained vaccination while the first, non-vaccine dataset was limited to non-pharmaceutical interventions. The number of simulation runs per combination of parameter values varies, as it was set to automatically spawn new runs until a 95% confidence interval would be reached (in contrast with a pre-determined fixed number of runs). The dataset had a total of

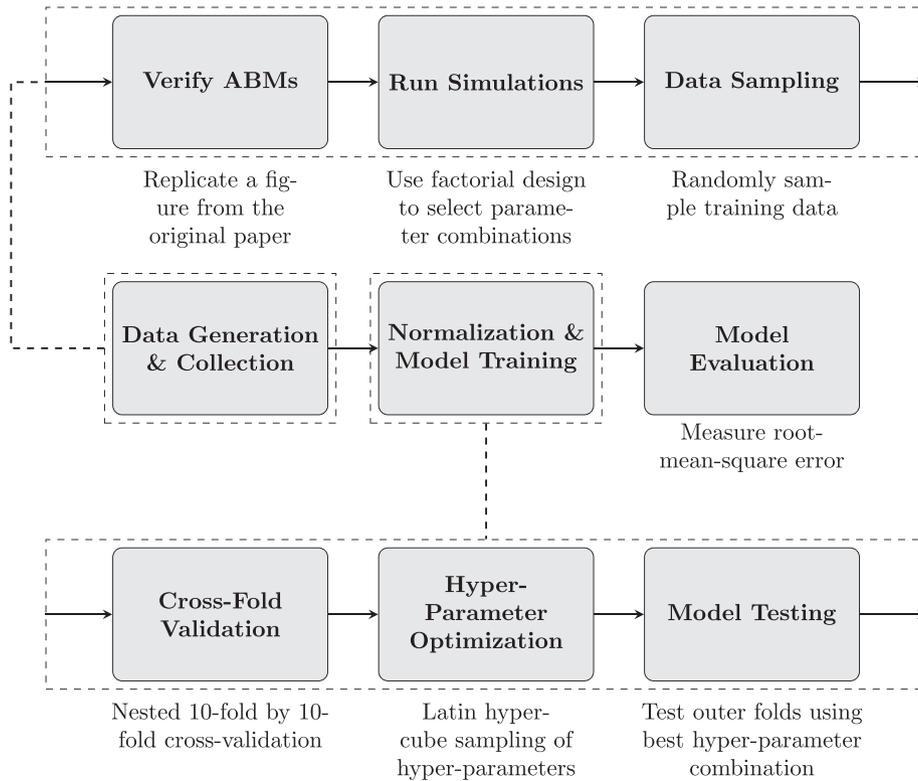


Figure 3. Our process for training a machine learning regression meta-model, detailing data generation from simulations (top) and machine learning (bottom).

$n = 727\,706$ data points, which are approximately evenly divided between the vaccine subsets ($n = 209\,042$ for the Trump administration and $n = 207\,990$ for the Biden administration) and the non-pharmaceutical subset ($n = 310\,674$).

3.1.2. Shamil et al. Model

For the Shamil et al. model,^[24] we ran the New York city configuration with four different values for two parameters

using a factorial design, where every possible combination is used. We also performed ten replications on each combination. The parameters we used are explained in **Table 5**. The 85% smartphone ownership level is based on a survey by the Pew Research Center from February 8, 2021.^[95] The 14-day quarantine is based on the recommendations of the CDC at a time close to when the model was developed, in December 2020.^[96] In total, we generated $n = 240$ data points from this model.

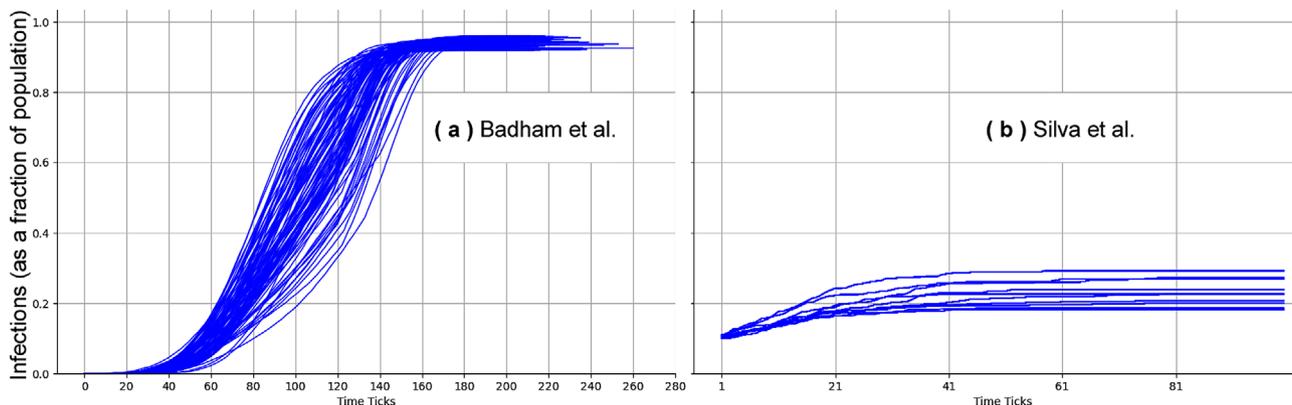


Figure 4. Examples of paths of 100 individual simulation runs over time in two of the models. Simulation parameters were: a) distancing policy: by contact, short/long movement reductions: 0.25, contact reduction 0.4; b) 500 agents, grid size 500, half lockdown, masks required, contagion distance 0.5. The 95% confidence intervals were: a) 0.89204 ± 0.04045 (0.85159 to 0.93248), b) 0.23349 ± 0.00691 (0.22657 to 0.24040).

Table 3. Parameters and values explored by Li and Giabbanelli in the absence of vaccines. When vaccines are present, only 6 combinations to provide broadly different scenarios in Table 4.

Parameter	Values used
Mask wearing	In all networks, work and school only, community only
Reduction in social contacts	At work and in schools: 5%, 30%. In the community: 10%, 30%
When to test quarantined individuals	Start of quarantine, end of quarantine, both
Test sensitivity	55%, 100%
Delay in contact tracing	None, 1 week
% of contacts of a positive case that can be traced	20%, 100%
When to start contact tracing for an exposed individual	Immediately, upon confirmation from a positive case

Table 4. Parameters and values used by Li and Giabbanelli. The two intervention settings (Trump administration, Biden administration) are run separately. Each intervention setting is run with all combination of the other parameter values.

Parameter	Values used
Vaccination	Trump administration: 20 million vaccines in Jan'21, 30 in Feb'21, 50 million monthly thereafter Biden Administration: 1 million vaccines every day
Non-pharmaceutical interventions	6 scenarios, including a no-intervention case. The dataset in Table 3 explores the scenarios. (For a complete description of the six scenarios, see Table 2 in Li & Giabbanelli ^[12])
% of population seeking vaccines	From 20% to 60%. Intermediate values are automatically determined to maximize information.
Vaccine efficacy	From 88% to 99%. Intermediate values are automatically determined to maximize information.

3.1.3. Silva et al. Model

For the Silva et al. model,^[29] we ran the model with a factorial design on five different parameters, running each for 60 days in the model. We also performed 30 replications on each combination. The parameters we used and their values are listed in **Table 6**. To verify that we used this model correctly, we replicated **Figure 5a** from the original paper^[29] that used tracked agent states over 60 days with no interventions and a population and grid size of 300. In total, we generated $n = 13\,500$ data points from this model.

Table 5. Parameters and values used in the Shamil et al. model dataset.

Parameter	Values used
Days required for contact tracing	1, 3, 5, 7
Percentage of contacts that can be traced	0.5, 0.6, 0.7, 0.8
Percentage of the population who own a smartphone	0.85
Days required in quarantine	14

Table 6. Parameters and values used in the Silva et al. model dataset.

Parameter	Values used
Population size	500, 750, 1000
Grid size	500, 750, 1000
Type of lockdown	none, total, half, conditional, vertical (explanations in Table 7)
Masks required?	True, False
Contagion distance	0.05, 0.25, 0.5, 0.75, 1

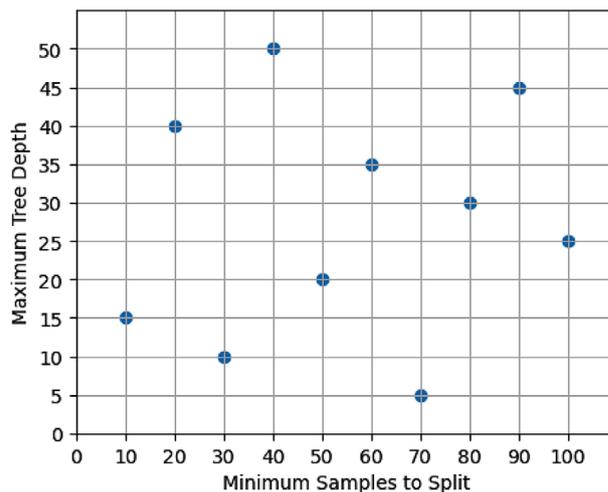


Figure 5. Latin hypercube for TreeLearner hyperparameters.

3.1.4. Badham et al. Model

For the Badham et al. model,^[30] we ran the model with a factorial design on 4 different parameters, performing 30 replications. All other settings are set to the default values for the model's 1.1 version. The parameters we altered are explained in **Table 8**. To verify that we used this model correctly, we replicated **Figure 4** from the original paper^[30] that tracked daily hospital admissions per day. In total, we generated $n = 2430$ data points from this model.

Table 7. Explanations of values for the lockdown type parameter (c.f. Table 6).

Lockdown type	Description
none	No lockdown
total	Lockdown required for everyone
half	Lockdown required for random 50% of the population
conditional	Lockdown required for anyone in a contagious environment
vertical	Lockdown required for anyone in an at-risk category or in a contagious environment

Table 8. Parameters and values used in the Badham et al. model dataset.

Parameter	Values used
Social distancing policy	None, AllPeople, ByContact (explanations in Table 9)
Contact reduction	0.2, 0.4, 0.6
Short movement reduction	0, 0.25, 0.5
Longer movement reduction	0, 0.25, 0.5

Table 9. Explanations of values for the social distancing policy parameter (c.f. Table 8).

Distancing policy	Description
None	No distancing
AllPeople	Everyone reduces their contacts by the “contact reduction” amount
ByContact	Probability of contact is reduced by the “contact reduction” amount

3.2. Normalization and Model Training

To perform our regressions, we used TreeLearner regressors from the Orange data mining library (orange3 version 3.26.0) running on Python (version 3.8.3). When we loaded our datasets, we converted their cumulative infection numbers to proportions of the total population, so that the results from the four different models are comparable. Then, we used a tenfold cross-validation method to actually train our meta-models. Cross-validation involves separating data into ten different “folds.” One by one, these folds are used as testing data for our model, while the other nine folds are used to train the model. We also needed to optimize the parameters of our regressor, so we used hyper-parameter optimization to maximize our meta-model accuracy.

Hyperparameter optimization ensures that the regressor is using the most accurate set of model parameters, and it does this by performing a second set of cross-validations, which is known as a nested cross-fold validation design. The training data from the first set of folds is divided into a further ten folds, where one fold at a time is used as a validation set. Then, the remaining nine folds are used to train a model using the different combinations of parameters to select the best combination. To select our parameters, we used a Latin hypercube design, which ensures that the sample space for parameters is covered effectively. We optimized the two parameters that have the most impact on the decision tree regressor:

- the maximum depth of the tree. A smaller depth forces the meta-model to be simplified. We considered values from 5 to 50 by steps of 5.
- the minimum number of samples to split a node. The machine learning algorithm stops sooner and delivers a simpler model if we raise the minimum number of samples. We examined values from 10 to 100 by steps of 10.

For details on the impact of these parameters on a decision tree, we refer the reader to two recent examples of optimization.^[97,98]

The optimization process was conducted using a Latin hypercube with ten different samples (Figure 5).

3.3. Model Evaluation

Once the hyperparameter optimization selected the most accurate set of parameters, these parameters were used to build the model for the outer folds. Then, we were able to analyze how the accuracies of the models changed in comparison with the amount of data that was used to train them by calculating the RMSE for their predicted infected proportions (Equation 1). RMSE is a common error metric for regression models^[99–101] and it is useful because the units for RMSE are the same as the units for the simulation output (i.e., our errors for this paper are measured in proportions of the population).

$$\text{RMSE} = \sqrt{\frac{\sum_{k=0}^n (\hat{Y}_k - Y_k)^2}{n}} \quad (1)$$

While the RMSE is informative for the quality of the regression, it does not suffice to compare two models, especially when the inputs are auto-correlated in the simulation model and are assumed to be independent in the meta-model. Consequently, we complement the RMSE by comparing the point estimates of the meta-model and simulation model on different outputs (number of deaths, number of infections) and across levels of policy-relevant control parameters (e.g., reduction in contacts).

Finally, we measure the time necessary to train the meta-model (wallclock time measured in seconds). This measure is important to assess whether the *complete* pipeline is truly time-saving for the end user, as the training time may otherwise be a hidden and potentially prohibitive cost for some machine learning methods. In sum, if we need few expensive simulations to create a meta-model that achieves a low RMSE and the meta-model can be created quickly, then there is a demonstrated incentive to use our solution instead of current practices.

4. Results

The ABMs used for this study were chosen as we could access their code and verify our use of the model vis-a-vis published figures in peer-reviewed manuscripts. The first subsection focuses on this verification effort, by contrasting the results that we obtained from the model with the authors’ published results; figures were all reproduced with the authors’ permission. The second subsection is centered on the machine learning results. To provide full transparency and support replicability efforts regarding machine learning, the training data used in this subsection is openly accessible on the third-party repository <https://osf.io/d7vqa/>, provided by the Open Science Framework.

4.1. Agent-Based Model Verification

As noted in Section 3.1.1, the model by Li and Giabbanelli^[12] was not subject to verification, since we directly used data produced by the model. That is, we already had access to its full simulation

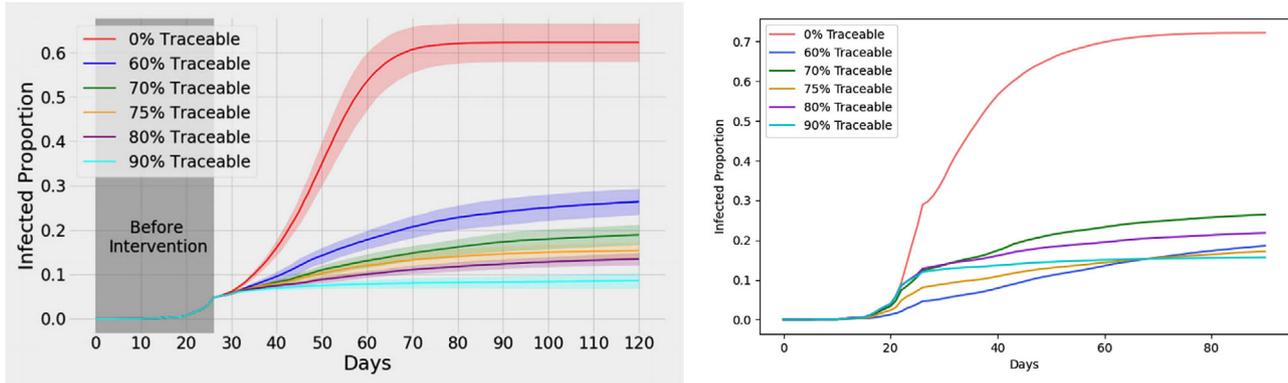


Figure 6. Figure 5a from Shamil, et al.^[24] (left) with our replication (right) using five repetitions. The left figure is reproduced with permission.^[24] Copyright 2021, The Authors, published under exclusive licence to Springer Science+Business Media, LLC part of Springer Nature.

results. In contrast, the other two models had to be verified since we did not have a spreadsheet of results to use and hence we ran the code provided by the authors. The verification sought to replicate the published work of the authors, to ensure that we were using their model adequately.

For the Shamil model, we replicated Figure 5a from Shamil et al.^[24] at six different contact tracing levels (0, 0.6, 0.7, 0.75, 0.8, and 0.9). Qualitatively, we note that shapes and trajectories of the infections are similar in our replication (Figure 6). However, values occasionally differ and some of the curves for the higher trace levels also overlapped in our replication, which is not something found in the original. Hence, results are not reproduced in terms of outcome, but they are “analysis reproducible”^[102] since the same conclusions about the effect of a COVID-19 interventions are reached based on the authors’ original results and the results obtained here. Based on previous attempts at replicating simulation models,^[103] a possibility is that the high variability in the model’s output results in noticeable differences across individual runs and only a large number of runs (e.g., to achieve a 95% confidence interval) would allow to create comparable curves. In other words, the discrepancies are a likely consequence of output variability in the published work.

For the Silva model, we replicated Figure 5a from Silva et al.^[29] using a population size and grid size of 300 (Figure 7). The state counts in our replication closely follow the average state counts

in the original figure. For the Badham model, we replicated Figure 4 from Badham et al.^[30] using the model’s atJul13 scenario (Figure 8). The shape and values in our replication closely follow the original figure. Both models were thus reproduced at the level of the outcome.

4.2. Machine Learning Results: RMSE and Point Predictions

The RMSE for the Li and Giabbanelli model^[12] without vaccines is shown in Figure 9. The RMSE is around 0.0795 regardless of the sample size, which shows that running as little as 5% of the experiments from the ABM suffices to then infer the rest via machine learning. However, we note that the error margins do benefit from an early increase in data, of up to 30% of the sample size. The situation changes when vaccines are introduced. For both the Trump administration’s vaccine plan and the Biden administration’s plan, we see that the RMSE strictly decreases as the sample size increases (Figure 10). The error can be as low as 0.0034, while noting that the decrease is essentially from a low initial error (RMSE of about 0.0041 to 0.0043 at 5% of the sample size) to a slightly lower error.

Different situations are observed among the other three models. For the Shamil et al. model,^[24] results show an error around 0.065 when using most of the data (Figure 11). Most interestingly,

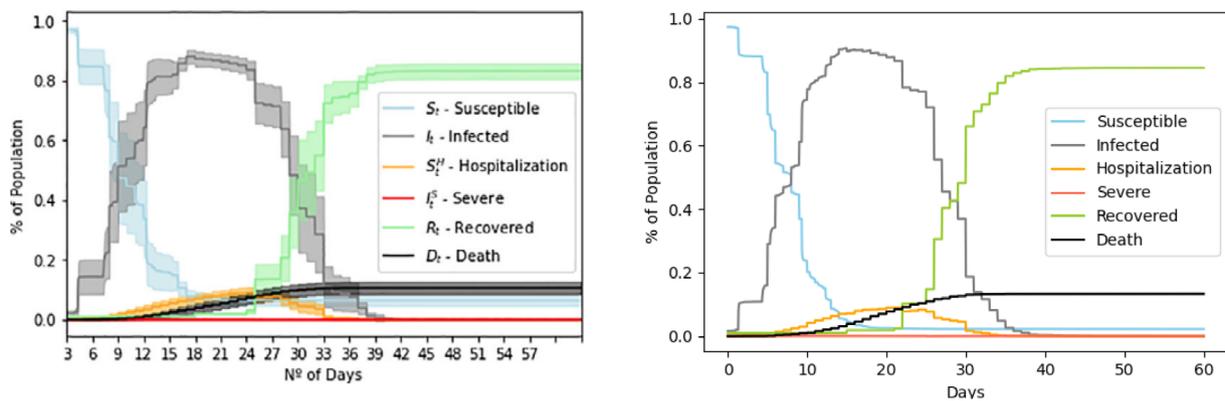


Figure 7. Figure 5a from Silva et al.^[29] (left) with our replication (right) using ten repetitions. The left figure is reproduced with permission.^[29] Copyright 2021, Elsevier.

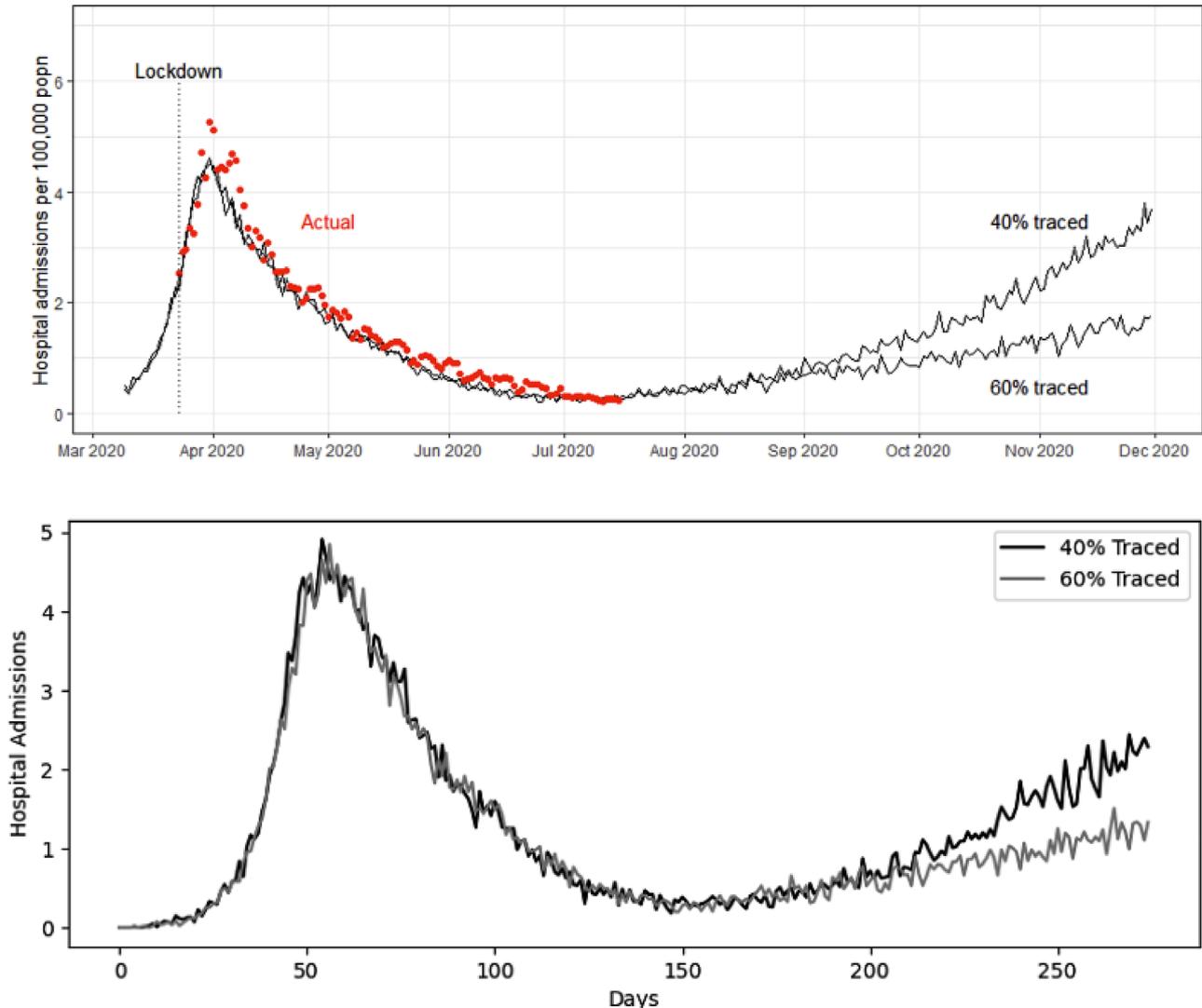


Figure 8. Figure 4 from Badham et al.^[30] (top) with our replication (bottom) using 1000 repetitions. The top figure is reproduced under the terms of the Creative Commons CC-BY License.^[30] Copyright 2021, The Authors, published by JASSS.

the error has a non-monotonic relation with the sample size: it is lowest at the smallest sample size (0.1%) and at about half of the sample (0.45%), but higher in between. For the model by Silva et al.,^[29] we see a decrease in RMSE as the sample size increases (Figure 12). Again, we note that the gains should be contextualized given the small range of the RMSE: 10% of the dataset suffices to achieve an error of 0.028 and even taking the full dataset only brings it down to 0.0225 at the best. Finally, the model by Badham et al.^[30] shows a decrease of RMSE as the sample size increases up to 50% of the dataset, and then the error plateaus (Figure 13). Relative to its scale, we emphasize that the total decrease in RMSE between the 5% and 100% sample sizes is only around 0.01.

To further assess the errors, we compared the outcomes of the meta-model and the four simulation models on two variables of interest (death cases, infections) across several values of policy-relevant parameters such as the tracing percentage or contact reduction. Results show that the meta-model follows the simu-

lation model very closely across several intervention scenarios (Figure 14-a,c,d). Although Figure 14d stands out, the tight y-axis means that the meta-model has settled on one value (0.19) whereas the simulation model still makes very small changes (by at most 5×10^{-5}). The main discrepancy is seen in Figure 14e, where the meta-model follows the overall trend of the simulation but point estimates can deviate by as much as 0.025.

4.3. Training Times

For each ABM, we measured how long it took to create the machine learning meta-model based on 100% of the data. This is an upper bound, as the previous section has shown that comparable estimations are obtained when the meta-model is trained with less data. Results (Table 10) were obtained on a personal computer, showing that the meta-model can be built within minutes. The training time depends on the complexity of the search space,

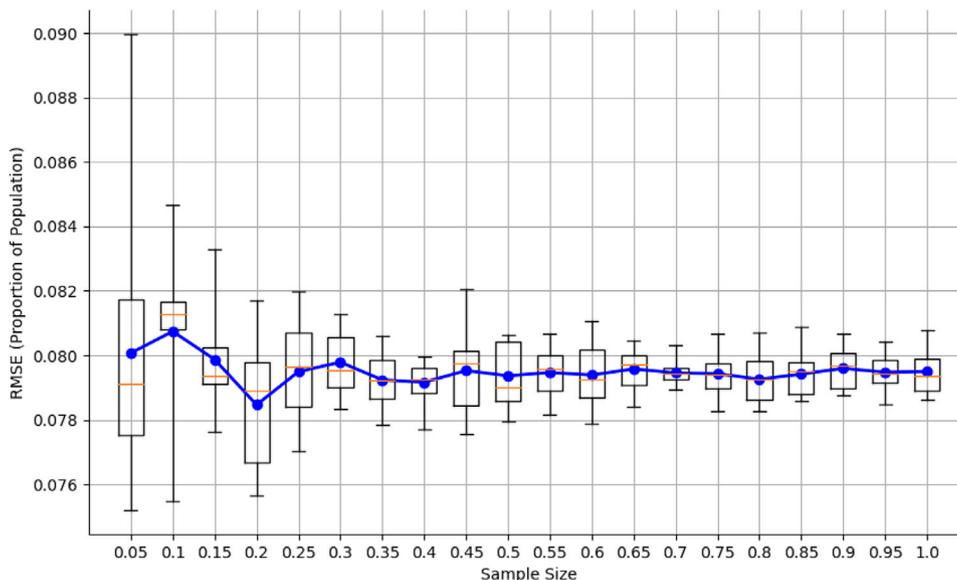


Figure 9. Graph of the root-mean-square error of the predicted proportion of the population that was infected at different sample sizes for the Li and Giabbanelli dataset without vaccines.

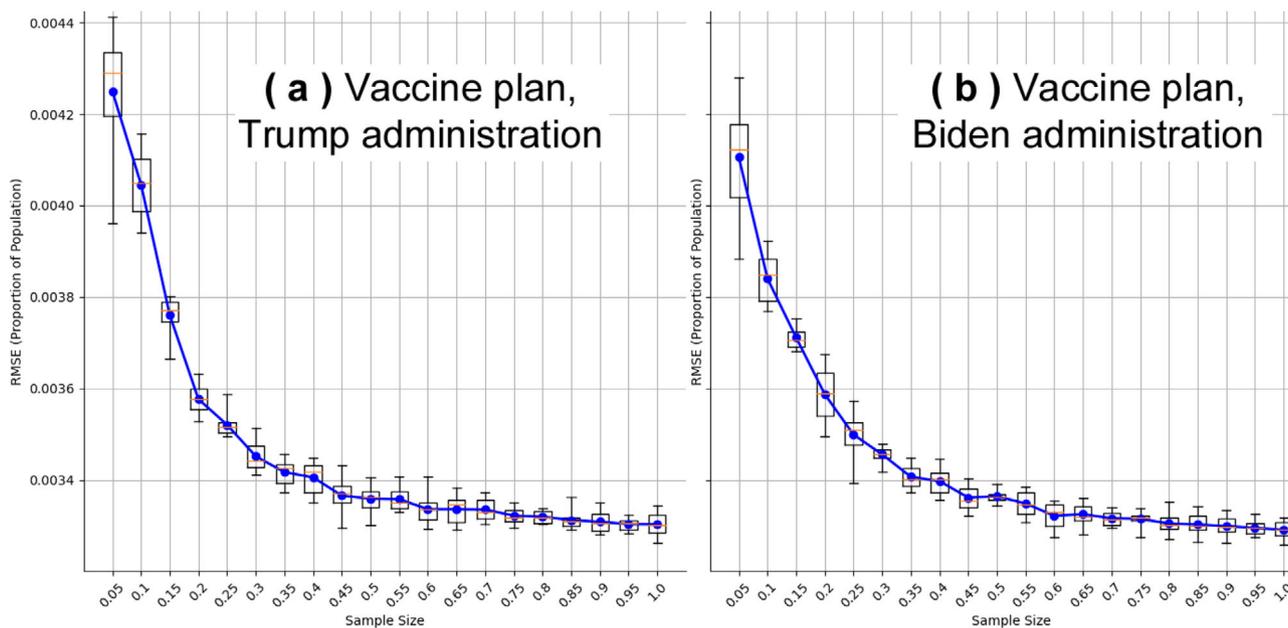


Figure 10. Graph of the root-mean-square error of the predicted proportion of the population that was infected at different sample sizes for the Li and Giabbanelli dataset using a) the Trump administration vaccine plan and b) the Biden administration vaccine plan.

as evidenced by variations for a single model based on either the policy scenario (Li and Giabbanelli) or the desired output (Silva et al.). These time estimates can be contextualized based on Table 2, which showed the time to run a single simulation from each ABM. The contrast reveals that the costs of building the meta-model are less than running a single simulation from the corresponding ABM, which often needed a computing cluster. In conclusion, there is no “hidden cost” in building the meta-models as this operation is computationally inexpensive compared to the simulations.

5. Discussion

As real-world data does not directly support what-if analyses and policy evaluation of alternatives, high-resolution models such as ABMs are needed to support policymakers in running detailed scenarios and building trust in their outcomes. However, these ABMs can be computationally intensive, thus requiring resources that policymakers may not have readily at their disposal, or taking too long given the urgent need for actions. In this case, meta-models are needed to support analysts and decision-makers

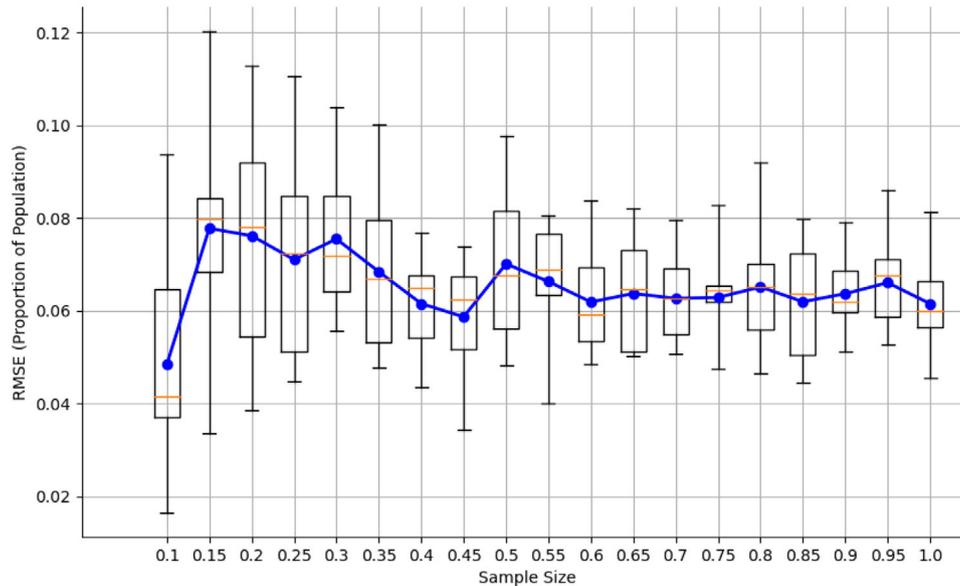


Figure 11. Graph of the root-mean-square error of the predicted proportion of the population that was infected at different sample sizes for the Shamil et al. dataset.

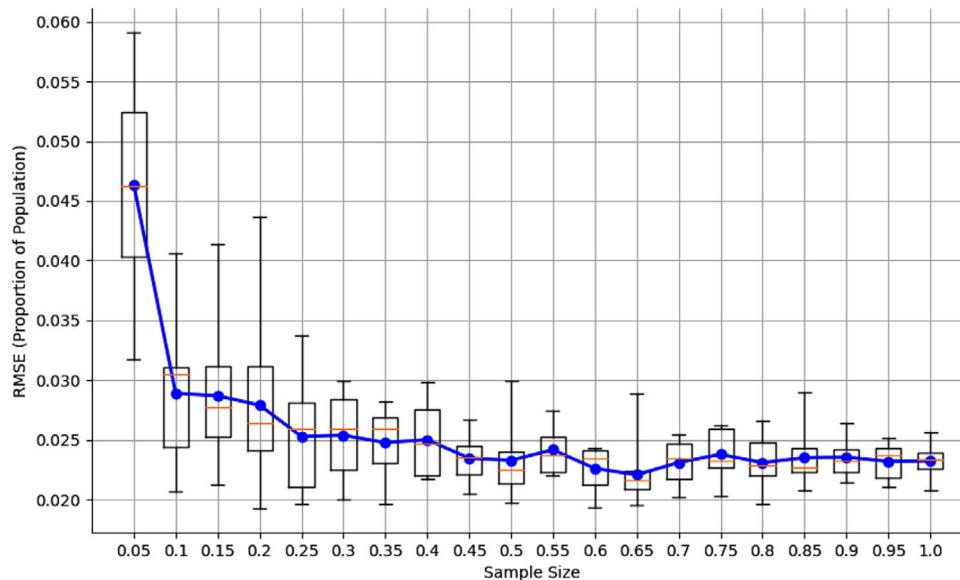


Figure 12. Graph of the root-mean-square error of the predicted proportion of the population that was infected at different sample sizes for the Silva et al. dataset.

by generating results faster, on a ubiquitous hardware. While many other works have been devoted to either creating ABMs for COVID-19 or applying machine learning to real-world data,^[65] our study is the first to combine these techniques to examine whether access to COVID-19 ABMs could be democratized by making them more computationally affordable for end users. Specifically, we analyzed how the amount of data used to train a regression meta-model affects its accuracy and differentiated situations where more data is required from situations where small amounts of data are sufficient. This analysis contributes to the broader literature on assessing and simplifying COVID-19 mod-

els, which has already established that the number of parameters could be decreased significantly.^[59]

Our results shows that models which do not have strong interventions like lockdowns or vaccines do not require as much training data, hence it is possible to run few computationally expensive simulations and then switch to an inexpensive meta-model. The Covasim model with no vaccines and the Shamil model showed no discernible downward trend in their RMSE values, so adding more training data would not guarantee an increase in model accuracy. While the Badham model showed a slight decrease in RMSE, it only decreased by 0.01 from 5% to

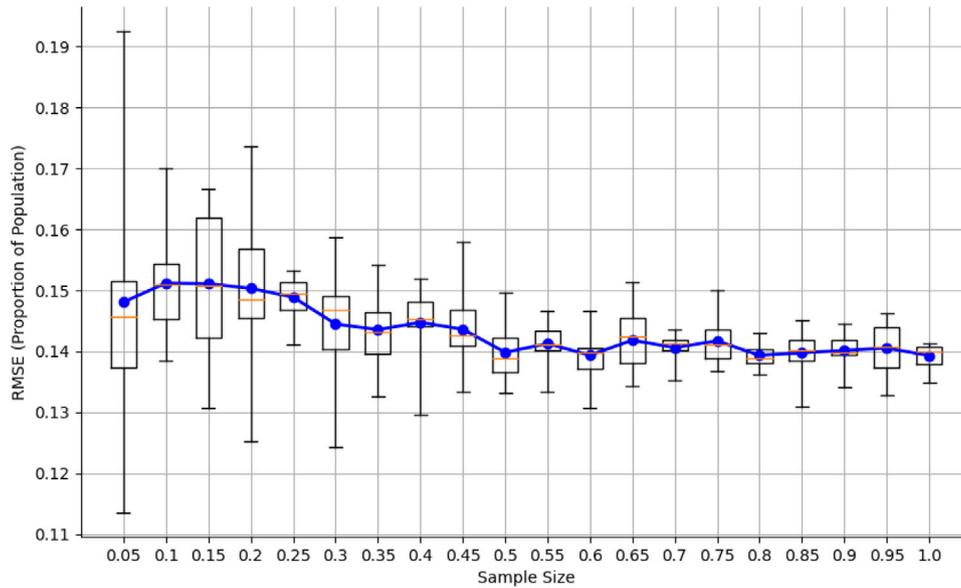


Figure 13. Graph of the root-mean-square error of the predicted proportion of the population that was infected at different sample sizes for the Badham et al. dataset.

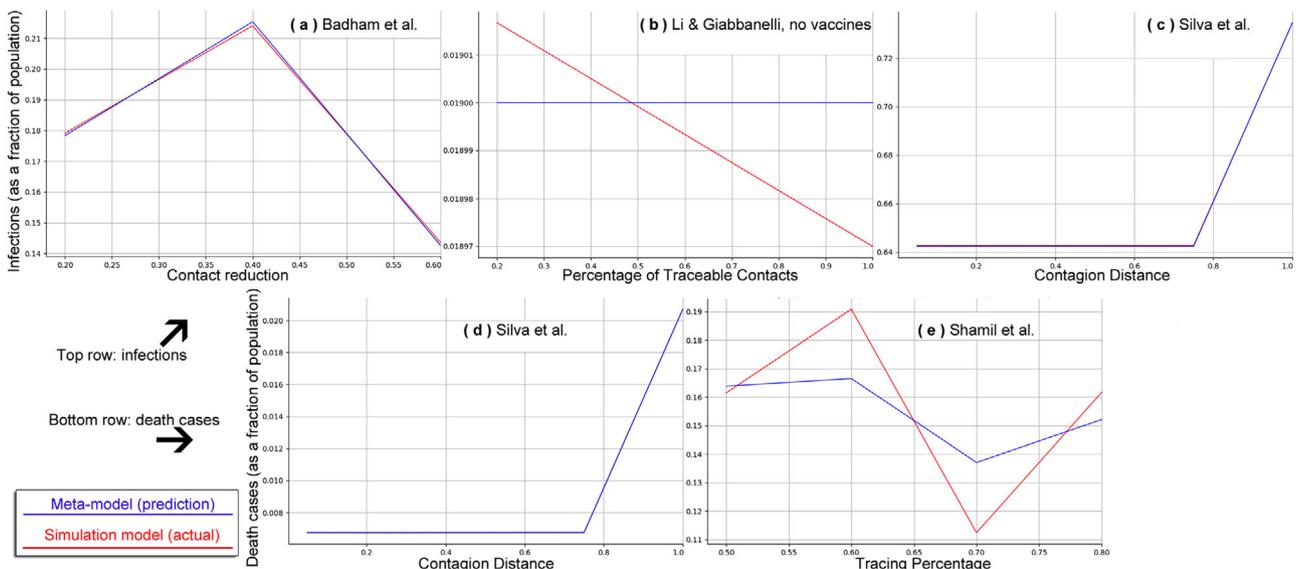


Figure 14. Comparison of outcomes from the meta-model (blue) with the simulation model (red) across variables of interest (top row: infections; bottom row: death cases) and policy scenarios for the four models. Other parameters were fixed as follows: a) distancing policy: by contact, short/long movement reductions: 0.25; b) layers impacted: all, contact reduction in work/school/community: 0.7, daily tests: 1.1M, quarantine: both, test sensitivity 0.55, time to contact trace 7 days, no presumptive quarantine; c–d) 500 agents, grid size 500, conditional lockdown, masks; e) 3 days required for contact tracing.

100% sample size. Conversely, the models which did have strong interventions showed a strong downward trend in their RMSE values. Both Covasim models that used vaccines only began to stabilize after 60% of the data was used in training, and while the Silva model stabilized slightly faster, it took around 45% of the data to do so. Note that the RMSE values were low to start with (e.g., less than 0.0045 for Covasim with vaccines), so the initial error may already be tolerable for some end users as part of a tradeoff between accuracy and computational needs.

There are three main limitations to our work. First, we focused on peer-reviewed COVID-19 ABMs that we could run to obtain data and for which we could verify our use of the model vis-a-vis published results. As shown in previous calls for transparency in COVID-19 modeling, modelers do not systematically provide their code,^[104] despite the existence of several platforms in which scientists can openly share code and data.^[105] In a transparency assessment, Jalali et al. found that most models do not share their code,^[106] which echoes similar observations about practices

Table 10. The maximum time to train a model on a personal computer shows that the cost is less or equal than performing a single simulation run (Table 2), which confirms that our overall approach can be time-saving for decision-makers.

Model	Scenario	Time [s] to build meta-model on 100% data	Data points
Li and Giabbanelli	No vaccine	217.5882	310 674
	Vaccine, Trump administration	375.5181	209 042
	Vaccine, Biden administration	396.3571	207 990
Silva et al.	Predicting infections	41.0124	13 500
	Predicting deaths	34.1787	
Badham et al.	N/A	10.1618	2 430
Shamil et al.	N/A	4.7391	240

in agent-based modeling across application domains.^[107,108] Our criteria thus meant that we could only assess a subset of existing models and it is possible that different trends or initial error levels are observed in other models. We note that projects that shared their code were also transparent regarding how their computational results were produced,^[109] hence we were able to perform verification and we applied the same level of transparency when conveying the model's parameters.

Second, several key parameters and assumptions regarding COVID-19 continue to change. For example, a study on almost 100 000 volunteers in July 2021 found a vaccine effectiveness of 49%, which is less than the lowest value assumed in some of the previous modeling studies.^[12,110] Other studies have shown that vaccinated individuals have a comparable viral load to unvaccinated ones within the first few days,^[111,112] whereas the flow diagrams used in many models have considered that vaccinated individuals were fully removed from the population. Our conclusions are limited to COVID-19 ABMs developed so far, since future models may exhibit markedly different dynamics. In particular, bifurcation may be present in future models, which would require an analysis of models by clusters of trajectories rather than average dynamics.^[113] Although in reality there are rare events whose large impact (e.g., super-spreader events, highly publicized case of a celebrity) can set the rest of the simulation on a very different trajectories, most current ABMs do not (yet) account for these events, hence they are easier to approximate with a meta-model. Many of the COVID-19 models exhibiting bifurcations are built from nonlinear differential equations rather than ABMs; these models provide abundant illustrations of the existence of bifurcations of various types,^[114–118] such as Hopf bifurcation (and the related Neimark–Sacker bifurcation) or period-doubling bifurcation.

Finally, the size of the datasets that we generated were limited by the high computational needs of the COVID-19 ABMs. For example, the Shamil model required around 8 h per run on a node of a high-performance computing cluster. Given this limitation, our conclusions focus on the trend (e.g., is there a reduction in error as more simulations are used?) rather than on a precise point-estimate of the error (e.g., exactly how many simulations are required to achieve a given error level).

There are several potential avenues for future work. We focused on predicting the final result of the simulation, but there would also be merit to predicting the time series of outputs. Although the use of meta-modeling of simulations for time series is more common for financial simulations,^[119] there could also be an opportunity for future studies to apply these techniques in the case of COVID-19 ABMs. In particular, several such ABMs (e.g., COVID-Town^[120]) have been developed to perform a joint analysis of economic and epidemic dynamics, hence they are particularly interested in the shape^[121] of the economic recovery over time (i.e., the time series).

Another possibility would be to explore the prediction of multiple outputs. Indeed, various stakeholders may be interested in the effects of a COVID-19 intervention on different outcomes ranging from epidemiological (e.g., number of new cases) to logistical (e.g., spare capacity in intensive care units) and economical (e.g., loss in productivity or revenues). Creating a multi-value regression model would thus address the multi-value optimization problem whereby decision-makers seek a balance between the effects of an intervention on health, the economy, and other (possibly non-correlated) issues. While we may expect that more simulations are necessary to train an accurate meta-model able to predict several partially correlated outputs, the specific relationship is still a question. That is, we still need to establish how the number of predicted outputs raises the need for simulation data. In the interim, modelers can follow our approach to predict the effect of an intervention onto different evaluation metrics, but they would create one regression model per metric rather than a single model predicting all metrics at the same time. **Figure 15** exemplifies the performance of a model trained to predict the number of deaths rather than infections on the Silva et al. ABM, showing that the model is already accurate with a very small fraction of the data and can be further improved if more data is provided.

Finally, the amount of data necessary to train an accurate model may be further reduced depending on the design of experiments.^[122] Indeed, when a high-resolution simulation system is used to train a meta-model, the design of experiments to produce the data is essential for the success of the meta-model.^[123,124] In ABMs that have a very large number of parameters, an important first step would be to assess which ones play a role in determining the output, either by themselves or in contribution with other parameters (second order effect, third order effects, etc). This can be approached by a factorial design of experiments and an ensuing analysis to decompose the variance in the output onto the action of parameters. As aforementioned, such analyses have shown that the number of parameters could be decreased significantly.^[59] This approach may be less applicable to the ABMs studied here, as they have only a handful of parameters, which play a statistically significant role either directly or in contribution with other parameters.^[13] In this case, an adaptive design of experiments may be more helpful^[125–127] than a fixed design.

6. Conclusion

In this paper, we analyzed how the amount of data used to train a simulation meta-model affected the accuracy of the model. We found that for models with no strong interventions such as vaccines or lockdowns, a small amount of data could generate a

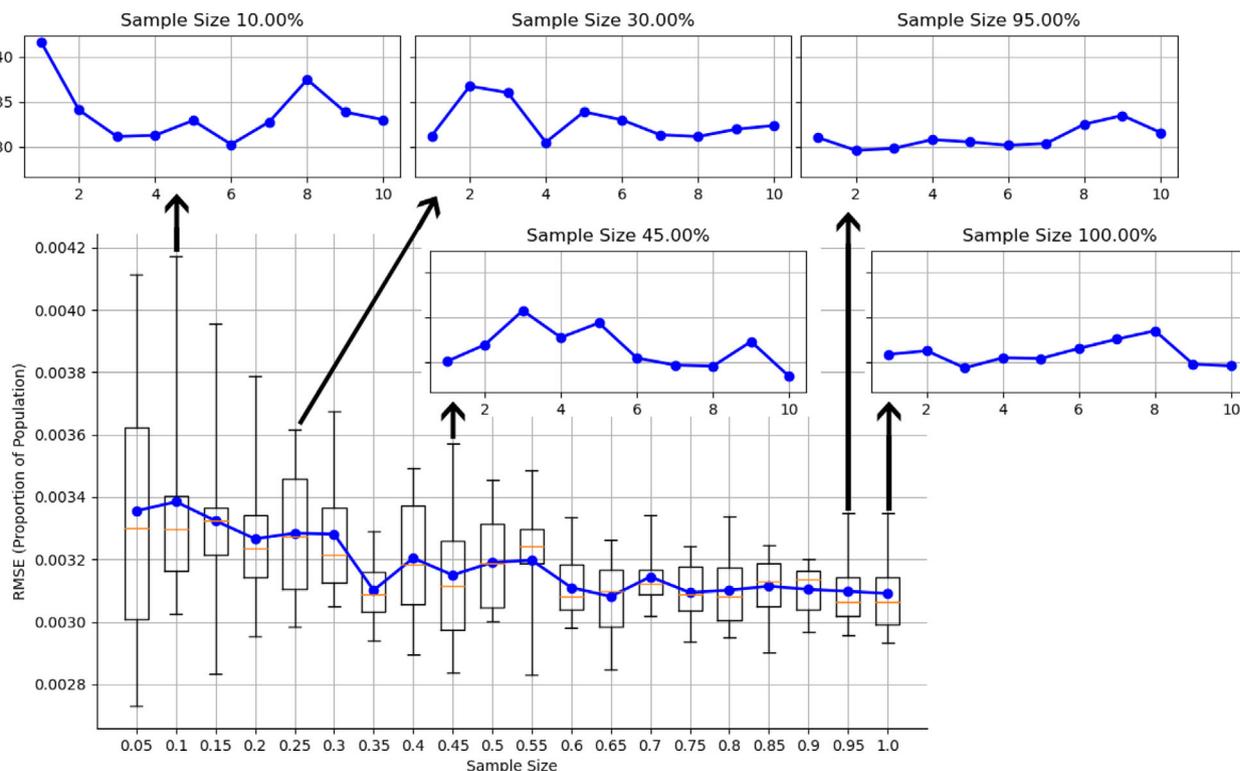


Figure 15. Root-mean-square error of the predicted proportion of the population that died (rather than be infected) as a function of the fraction of data used to train the meta-model. Insets provide a per-fold graph showing the same RMSE (y-axis) across folds in a 10 cross-fold validation (x-axis); note that folds are unordered hence insets cannot be used as trend graphs.

model with similar accuracy to one trained on a much larger amount of data. However, models which had strong interventions took large amounts of data to train a model that achieved a stable accuracy. These results indicate that modeling the spread of COVID-19 without strong interventions can be done with very little data, but when stronger interventions are considered, much more data is required to train an accurate model.

Acknowledgements

For each of the three models used in this paper, the authors thank Dr. Mohammad Sohel Rahman, Prof. Petrônio Cândido de Lima e Silva, Dr. Jennifer Badham for giving them permission to reuse their images in their verification process. The authors also appreciate the assistance of Dr. Badham in operating the model.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are openly available in the Open Science Framework repository <https://osf.io/d7vqa/>, under "COVID-19 Metamodeling - Training Datasets".

Keywords

agent-based models, COVID-19, machine learning, meta-modeling, surrogate model

Received: August 26, 2021
Revised: November 1, 2021
Published online: November 23, 2021

- [1] World Health Organization, Coronavirus disease (covid-19) pandemic, **2021**. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>
- [2] Centers for Disease Control and Prevention, United states covid-19 cases, deaths, and laboratory testing (naats) by state, territory, and jurisdiction, **2021**. https://covid.cdc.gov/covid-data-tracker/#cases_casesper100klast7days
- [3] A. D. Iuliano, H. H. Chang, N. N. Patel, R. Threlkel, K. Kniss, J. Reich, M. Steele, A. J. Hall, A. M. Fry, C. Reed, *The Lancet Regional Health-Americas* **2021**, *1*, 100019.
- [4] S. Miller, L. R. Wherry, B. Mazumder, *Health Affairs* **2021**, 10–1377.
- [5] S. Charumilind, M. Craven, J. Lamb, A. Sabow, M. Wilson, When will the covid-19 pandemic end? An update, **2021**.
- [6] K. Kupferschmidt, M. Wadman, *Science* **2021**, *372*, 1375.
- [7] A. Fretheim, K. G. Brurberg, F. Forland, *Eurosurveillance* **2020**, *25*, 2000687.
- [8] D. Pradhan, P. Biswasroy, P. K. Naik, G. Ghosh, G. Rath, *Arch. Med. Res.* **2020**, *51*, 363.
- [9] T. Hale, N. Angrist, R. Goldszmidt, B. Kira, A. Petherick, T. Phillips, S. Webster, E. Cameron-Blake, L. Hallas, S. Majumdar, H. Tatlow, *Nat. Hum. Behav.* **2021**, *5*, 529.
- [10] B. Nussbaumer-Streit, V. Mayr, A. I. Dobrescu, A. Chapman, E. Persad, I. Klerings, G. Wagner, U. Siebert, D. Ledinger, C. Zachariah, G. Gartlehner, *Cochrane Database Syst. Rev.* **2020**, *9*, CD013574.

- [11] S.-C. Chung, S. Marlow, N. Tobias, A. Alogna, I. Alogna, S.-L. You, K. Khunti, M. McKee, S. Michie, D. Pillay, *BMJ Open* **2021**, *11*, e047832.
- [12] J. Li, P. Giabbanelli, *JMIR Med. Inform.* **2021**, *9*, e27419.
- [13] J. Li, P. J. Giabbanelli, In *International Conference on Computational Science*, Springer, **2021**, 655–662.
- [14] A. Muscillo, P. Pin, T. Razzolini, *PLoS One* **2020**, *15*, e0237057.
- [15] N. Harring, S. C. Jagers, Å. Löfgren, *World Development* **2021**, *138*, 105236.
- [16] E. N. Iftekhar, V. Priesemann, R. Balling, S. Bauer, P. Beutels, A. C. Valdez, S. Cuschieri, T. Czypionka, U. Dumpis, E. Glaab, E. Grill, C. Hanson, P. Hotulainen, P. Klimek, M. Kretzschmar, T. Krüger, J. Krutzinna, N. Low, H. Machado, C. Martins, M. McKee, S. B. Mohr, A. Nassehi, M. Perc, E. Petelos, M. Pickersgill, B. Prainsack, J. Rocklöv, E. Schernhammer, A. Staines, et al., *The Lancet Regional Health-Europe* **2021**, *8*, 100185.
- [17] A. Truszkowska, B. Behring, J. Hasanyan, L. Zino, S. Butail, E. Caroppo, Z.-P. Jiang, A. Rizzo, M. Porfiri, *Adv. Theory Simul.* **2021**, *4*, 2000277.
- [18] U. K. Mukherjee, S. Bose, A. Ivanov, S. Souyris, S. Seshadri, P. Sridhar, R. Watkins, Y. Xu, *Sci. Rep.* **2021**, *11*, 6264.
- [19] R. Zafarnejad, P. M. Griffin, *Comput. Biol. Med.* **2021**, 104518.
- [20] F. Ying, N. O'Clery, *PLoS One* **2021**, *16*, e0249821.
- [21] B. Jahn, G. Sroczynski, M. Bicher, C. Rippinger, N. Mühlberger, J. Santamaria, C. Urach, M. Schomaker, I. Stojkov, D. Schmid, G. Weiss, U. Wiedermann, M. Redlberger-Fritz, C. Druml, M. Kretzschmar, M. Paulke-Korinek, H. Ostermann, C. Czasch, G. Endel, W. Bock, N. Popper, U. Siebert, *Vaccines* **2021**, *9*, 434.
- [22] A. M. Al-Shaery, B. Hejase, A. Tridane, N. S. Farooqi, H. A. Jassmi, *Sustainability* **2021**, *13*, 6923.
- [23] S. L. Chang, N. Harding, C. Zachreson, O. M. Cliff, M. Prokopenko, *Nat. Commun.* **2020**, *11*, 5710.
- [24] M. S. Shamil, F. Farheen, N. Ibtehad, I. M. Khan, M. S. Rahman, *Cognit. Comput.* **2021**, <https://doi.org/10.1007/s12559-020-09801-w>.
- [25] P. Giabbanelli, J. Badham, B. Castellani, H. Kavak, V. Mago, A. Negahba, S. Swarup, in *Proc. of the 2021 Annual Modeling and Simulation Conference (ANNSIM)*, Fairfax, VA, USA, July **2021**.
- [26] J. J. Hack, M. E. Papka, *Comput. Sci. Eng.* **2020**, *22*, 75.
- [27] T. L. Bauer, W. E. Beyeler, P. D. Finley, R. F. Jeffers, C. D. Laird, M. Makvandi, A. V. Outkin, C. Safta, K. M. Simonson, Sandia's research in support of covid-19 pandemic response: Computing and information sciences., Technical report, Sandia National Lab. (SNL-NM), Albuquerque, NM (United States), **2020**.
- [28] N. Ahmed, M. Wahed, arXiv:2010.15581, **2020**.
- [29] P. C. Silva, P. V. Batista, H. S. Lima, M. A. Alves, F. G. Guimarães, R. C. Silva, *Chaos, Solitons Fractals* **2020**, *139*, 110088.
- [30] J. Badham, P. Barbrook-Johnson, C. Caiado, B. Castellani, *J. Artif. Soc. Soc. Simul.* **2021**, <https://doi.org/10.18564/jasss.4532>.
- [31] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. Leung, E. H. Lau, J. Y. Wong, X. Xing, N. Xiang, Y. Wu, C. Li, Q. Chen, D. Li, T. Liu, J. Zhao, M. Liu, W. Tu, C. Chen, L. Jin, R. Yang, Q. Wang, S. Zhou, R. Wang, H. Liu, Y. Luo, Y. Liu, G. Shao, et al., *N. Engl. J. Med.* **2020**, *382*, 1199.
- [32] D. L. Roberts, J. S. Rossman, I. Jarić, *PLoS Pathog.* **2021**, *17*, e1009620.
- [33] T. R. Sosnowski, *Curr. Opin. Colloid Interface Sci.* **2021**, *54*, 101451.
- [34] K. Al Huraimel, M. Alhosani, S. Kunhabdulla, M. H. Stietiya, *Sci. Total Environ.*, **2020**, *744*, 140946.
- [35] A. McKeever, *National Geographic*, **2020**.
- [36] E. Fraser, *BMJ* **2020**, *370*, m3001.
- [37] P. W. Blair, D. M. Brown, M. Jang, A. A. Antar, J. C. Keruly, V. S. Bachu, J. L. Townsend, J. A. Tornheim, S. C. Keller, L. Sauer, D. L. Thomas, Y. C. Manabe, Ambulatory COVID Study Team. in *Open Forum Infectious Diseases*, Vol. 8, Oxford University Press, Oxford **2021**, p. ofab007.
- [38] T. Nasserie, M. Hittle, S. N. Goodman, *JAMA Netw. Open* **2021**, *4*, e2111417.
- [39] A. Raveendran, R. Jayadevan, S. Sashidharan, *Diabetes Metab. Syndr.: Clin. Res. Rev.* **2021**, *15*, 869.
- [40] Z. Yan, M. Yang, C.-L. Lai, *Biomedicines* **2021**, *9*, 966.
- [41] J. Ishigami, M. Kou, N. Ding, K. Matsushita, *Curr. Epidemiol. Rep.* **2021**, *8*, 1.
- [42] M. Legrand, S. Bell, L. Forni, M. Joannidis, J. L. Koyner, K. Liu, V. Cantaluppi, *Nat. Rev. Nephrol.* **2021**, *17*, 751.
- [43] M. K. M. Karuppan, D. Devadoss, M. Nair, H. S. Chand, M. K. Lakshmana, *Mol. Neurobiol.* **2021**, *58*, 2465.
- [44] S. Ahmed, O. Zimba, A. Y. Gasparian, *Clin. Rheumatol.* **2021**, *40*, 2611.
- [45] J. Howard, A. Huang, Z. Li, Z. Tufekci, V. Zdimal, H.-M. van der Westhuizen, A. von Delft, A. Price, L. Fridman, L.-H. Tang, V. Tang, G. L. Watson, C. E. Bax, R. Shaikh, F. Questier, D. Hernandez, L. F. Chu, C. M. Ramirez, A. W. Rimoin, *Proc. Natl. Acad. Sci.* **2021**, *118*, 4.
- [46] C. Courtemanche, J. Garuccio, A. Le, J. Pinkston, A. Yelowitz, *Health Affairs.* **2020**, *39*, 1237.
- [47] Y. J. Park, Y. J. Choe, O. Park, S. Y. Park, Y.-M. Kim, J. Kim, S. Kweon, Y. Woo, J. Gwack, S. S. Kim, J. Lee, J. Hyun, B. Ryu, Y. S. Jang, H. Kim, S. H. Shin, S. Yi, S. Lee, H. K. Kim, H. Lee, Y. Jin, E. Park, S. W. Choi, M. Kim, J. Song, S. W. Choi, D. Kim, B.-H. Jeon, *Emerging Infect. Dis.* **2020**, *26*, 2465.
- [48] U. Food, D. Administration, Fda takes key action in fight against covid-19 by issuing emergency use authorization for first covid-19 vaccine, **2020**. www.fda.gov/news-events/press-announcements/fda-takes-key-action-fight-against-covid-19-issuing-emergency-use-authorization-first-covid-19
- [49] Z. Wang, F. Schmidt, Y. Weisblum, F. Muecksch, C. O. Barnes, S. Finkin, D. Schaefer-Babajew, M. Cipolla, C. Gaebler, J. A. Lieberman, T. Y. Oliveira, Z. Yang, M. E. Abernathy, K. E. Huey-Tubman, A. Hurlley, M. Turroja, K. A. West, K. Gordon, K. G. Millard, V. Ramos, J. Da Silva, J. Xu, R. A. Colbert, R. Patel, J. Dizon, C. Unson-O'Brien, I. Shmelyovich, A. Gazumyan, M. Caskey, P. J. Bjorkman, et al., *Nature* **2021**, *592*, 616.
- [50] W. T. Harvey, A. M. Carabelli, B. Jackson, R. K. Gupta, E. C. Thomson, E. M. Harrison, C. Ludden, R. Reeve, A. Rambaut, S. J. Peacock, D. L. Robertson, *Nat. Rev. Microbiol.* **2021**, *19*, 409.
- [51] M. Hoffmann, P. Arora, R. Groß, A. Seidel, B. F. Hörnich, A. S. Hahn, N. Krüger, L. Graichen, H. Hofmann-Winkler, A. Kempf, M. S. Winkler, S. Schulz, H.-M. Jäck, B. Jahrsdörfer, H. Schrezenmeier, M. Müller, A. Kleger, J. Münch, S. Pöhlmann, *Cell* **2021**, *184*, 2384.
- [52] M. Yuan, D. Huang, C.-C. D. Lee, N. C. Wu, A. M. Jackson, X. Zhu, H. Liu, L. Peng, M. J. van Gils, R. W. Sanders, D. R. Burton, S. M. Reincke, H. Prüss, J. Kreye, D. Nemazee, A. B. Ward, I. A. Wilson, *Science* **2021**, *373*, 818.
- [53] M. Bongiovanni, A. M. Marra, F. Bini, B. D. Bodini, D. di Carlo, G. Giuliani, *J. Infect.* **2021**, *82*, e4.
- [54] M. Bongiovanni, A. M. Marra, F. Bini, B. D. Bodini, D. di Carlo, G. Giuliani, *Morb. Mortal. Wkly. Rep.* **2021**, *70*, 792.
- [55] M. Bergwerk, T. Gonen, Y. Lustig, S. Amit, M. Lipsitch, C. Cohen, M. Mandelboim, E. Gal Levin, C. Rubin, V. Indenbaum, I. Tal, M. Zavitán, N. Zuckerman, A. Bar-Chaim, Y. Kreiss, G. Regev-Yochay, *N. Engl. J. Med.* **2021**, *385*, 1474.
- [56] O. Dyer, *BMJ* **2021**, *374*, n1933.
- [57] S. He, Y. Peng, K. Sun, *Nonlinear Dyn.* **2020**, *101*, 1667.
- [58] A. B. Gumel, E. A. Iboi, C. N. Ngonghala, E. H. Elbasha, *Infect. Dis. Model.* **2021**, *6*, 148.

- [59] W. Edeling, H. Arabnejad, R. Sinclair, D. Suleimenova, K. Gopalakrishnan, B. Bosak, D. Groen, I. Mahmood, D. Crommelin, P. V. Coveney, *Nat. Comput. Sci.* **2021**, *1*, 128.
- [60] V. Chin, N. I. Samia, R. Marchant, O. Rosen, J. P. Ioannidis, M. A. Tanner, S. Cripps, *Eur. J. Epidemiol.* **2020**, *35*, 733.
- [61] N. P. Jewell, J. A. Lewnard, B. L. Jewell, *JAMA* **2020**, *323*, 1893.
- [62] G. Recchia, A. L. Freeman, D. Spiegelhalter, *PLoS One* **2021**, *16*, e0250935.
- [63] P. Harjule, V. Tiwari, A. Kumar, *J. Interdiscip. Math.* **2021**, *24*, 259.
- [64] J. Gnanvi, K. V. Salako, B. Kotanmi, R. G. Kakaï, *Infect. Dis. Model.* **2021**, *6*, 258.
- [65] A. Tolk, C. Glazner, J. Ungerleider, *Comput. Sci. Eng.* **2020**, *23*, 17.
- [66] B. Gallo Marin, G. Aghagoli, K. Lavine, L. Yang, E. J. Siff, S. S. Chi-ang, T. P. Salazar-Mather, L. Dumenco, M. C. Savaria, S. N. Aung, T. Flanigan, I. C. Michelow, *Rev. Med. Virol.* **2021**, *31*, e2146.
- [67] A. Venkatakrishnan, C. Pawlowski, D. Zemmour, T. Hughes, A. Anand, G. Berner, N. Kayal, A. Puranik, I. Conrad, S. Bade, R. Barve, P. Sinha, J. C. O'Horo, A. D. Badley, J. Halamka, V. Soundararajan, *NPJ Digit. Med.* **2021**, *4*, 117.
- [68] A. Fridman, R. Gershon, A. Gneezy, *PLoS One* **2021**, *16*, e0250123.
- [69] E. Y. Chan, *Pers. Individ. Dif.* **2021**, *171*, 110463.
- [70] M. H. Haischer, R. Beilfuss, M. R. Hart, L. Opielinski, D. Wrucke, G. Zirgaitis, T. D. Uhrich, S. K. Hunter, *PLoS One* **2020**, *15*, e0240785.
- [71] J. T. Mueller, K. McConnell, P. B. Burow, K. Pofahl, A. A. Merdjanoff, J. Farrell, *PNAS* **2021**, *118*, 2019378118.
- [72] L. M. Andersen, S. R. Harden, M. M. Sugg, J. D. Runkle, T. E. Lundquist, *Sci. Total Environ.* **2021**, *754*, 142396.
- [73] J. A. Firth, J. Hellewell, P. Klepac, S. Kissler, A. J. Kucharski, L. G. Spurgin, *Nat. Med.* **2020**, *26*, 1616.
- [74] T. R. Pollmann, S. Schönert, J. Müller, J. Pollmann, E. Resconi, C. Wiesinger, C. Haack, L. Shtembari, A. Turcati, B. Neumair, S. Meighen-Berger, G. Zattera, M. Neumair, U. Apel, A. Okolie, *EPJ Data Sci.* **2021**, *10*, 37.
- [75] C. C. Kerr, R. M. Stuart, D. Mistry, R. G. Abeysuriya, K. Rosenfeld, G. R. Hart, R. C. Núñez, J. A. Cohen, P. Selvaraj, B. Hagedorn, L. George, A. Izzo, A. Palmer, D. Delport, C. Bennette, B. Wagner, S. Chang, J. Panovska-Griffiths, M. Jastrzebski, A. P. Oron, E. Wenger, M. Famulare, D. J. Klein, *PLoS Comput. Biol.* **2021**, *17*, e1009149.
- [76] R. Hinch, W. J. Probert, A. Nurtay, M. Kendall, C. Wymant, M. Hall, K. Lythgoe, A. Bulas Cruz, L. Zhao, A. Stewart, L. Ferretti, D. Montero, J. Warren, N. Mather, M. Abueg, N. Wu, A. Finkelstein, D. G. Bonsall, L. Abeler-Dörner, C. Fraser, *PLoS Comput. Biol.* **2021**, *17*, e1009146.
- [77] J. C. Lemaitre, K. H. Grantz, J. Kaminsky, H. R. Meredith, S. A. Trulove, S. A. Lauer, L. T. Keegan, S. Shah, J. Wills, K. Kaminsky, J. Perez-Saez, J. Lessler, E. C. Lee, *Sci. Rep.* **2021**, *11*, 7534.
- [78] T. Kano, K. Yasui, T. Mikami, M. Asally, A. Ishiguro, *Proc. R. Soc. A* **2021**, *477*, 20200604.
- [79] J. Gomez, J. Prieto, E. Leon, A. Rodríguez, *PLoS One* **2021**, *16*, e0245787.
- [80] R. R. Barton, in *Proc. Winter Simulation Conf.* IEEE, Lake Buena Vista, FL, USA, December **1994**, pp. 237–244.
- [81] M. Saadatpour, A. Afshar, J. E. Edinger, *Water Resour. Manage.* **2017**, *31*, 2729.
- [82] A. Tregubov, J. Blythe, in *Int. Workshop on Multi-Agent Systems and Agent-Based Simulation*, Springer, Berlin **2020**, pp. 81–93.
- [83] A. Prada, A. Gasparella, P. Baggio, *Appl. Energy* **2018**, *225*, 814.
- [84] J. Park, K.-Y. Kim, *Appl. Soft Comput.* **2017**, *51*, 354.
- [85] A. MacAllister, A. Kohl, E. Winer, *Expert Syst. Appl.* **2020**, *139*, 112830.
- [86] F. Osisanwo, J. Akinsola, O. Awodele, J. Hinmikaiye, O. Olakanmi, J. Akinjobi, *Int. J. Comput. Trends Technol.* **2017**, *48*, 128.
- [87] R. R. Barton, in *2015 Winter Simulation Conference (WSC)*. IEEE, **2015**, pp. 1765–1779.
- [88] C. B. Lutz, P. J. Giabbanelli, A. Fisher, V. K. Mago, in *Proc. 2021 Annual Modeling and Simulation Conf. (ANNSIM)*, Fairfax, VA, USA, July **2021**.
- [89] T. Østergård, R. L. Jensen, S. E. Maagaard, *Appl. Energy* **2018**, *211*, 89.
- [90] A. Fisher, B. Adhikari, C. Zhai, J. E. Morgan, V. K. Mago, P. J. Giabbanelli, in *2020 Spring Simulation Conf. (SpringSim)*. IEEE, Fairfax, VA, USA, May **2020**, pp. 1–12.
- [91] G. E. Box, N. R. Draper, *Empirical Model-Building and Response Surfaces*, John Wiley & Sons, Hoboken, NJ **1987**.
- [92] J. R. Wilson, in *Proc. 19th Conf. Winter Simulation*, December **1987**, pp. 378–381.
- [93] L. M. Collins, J. J. Dziak, K. C. Kugler, J. B. Trail, *Am. J. Prev. Med.* **2014**, *47*, 498.
- [94] N. A. Butler, *Biometrika* **2001**, *88*, 847.
- [95] P. R. Center, Mobile fact sheet, **2021**. www.pewresearch.org/internet/fact-sheet/mobile/
- [96] C. for Disease Control, Prevention, Options to reduce quarantine using symptom monitoring and diagnostic testing, **2020**. www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/scientific-brief-options-to-reduce-quarantine.html
- [97] M. K. Wozniak, P. J. Giabbanelli, in *Proc. 2021 ACM SIGSIM Conf. Principles of Advanced Discrete Simulation*, May **2021**, pp. 13–25.
- [98] N. Rosso, P. Giabbanelli, *JMIR Public Health Surveill.* **2018**, *4*, e9536.
- [99] J. A. Bartholomai, H. B. Frieboes, in *2018 IEEE Int. Symposium on Signal Processing and Infor Technology (ISSPIT)*. IEEE, Louisville, KY, USA, December **2018**, pp. 632–637.
- [100] S. H. Shah, Y. Angel, R. Houborg, S. Ali, M. F. McCabe, *Remote Sens.* **2019**, *11*, 920.
- [101] M. Zeraatpisheh, S. Ayoubi, A. Jafari, S. Tajik, P. Finke, *Geoderma* **2019**, *338*, 445.
- [102] O. Erik Gundersen, *Philos. Trans. R. Soc. A* **2021**, *379*, 20200210.
- [103] P. J. Giabbanelli, C. Freeman, J. A. Devita, N. Rosso, Z. L. Brumme, in *Proc. 2019 ACM SIGSIM Conf. Principles of Advanced Discrete Simulation*, May **2019**, pp. 103–114.
- [104] C. M. Barton, M. Alberti, D. Ames, J.-A. Atkinson, J. Bales, E. Burke, M. Chen, S. Y. Diallo, D. J. Earn, B. Fath, Z. Feng, *Science* **2020**, *368*, 482.
- [105] P. V. Coveney, R. R. Highfield, *Philos. Trans. R. Soc. A* **2021**, *379*, 20200067.
- [106] M. S. Jalali, C. DiGennaro, D. Sridhar, *Lancet Glob. Health* **2020**, *8*, e1459.
- [107] M. A. Janssen, C. Pritchard, A. Lee, *Environ. Model. Softw.* **2020**, *134*, 104873.
- [108] P. J. Giabbanelli, B. Tison, J. Keith, *Simul. Model. Pract. Theory* **2021**, *108*, 102268.
- [109] M. Krafczyk, A. Shi, A. Bhaskar, D. Marinov, V. Stodden, *Philos. Trans. R. Soc. A* **2021**, *379*, 20200069.
- [110] P. Elliott, D. Haw, H. Wang, O. Eales, C. Walters, K. Ainslie, C. Atchison, C. Fronterre, P. Diggle, A. Page, A. J. Trotter, S. J. Prosolek, D. Ashby, C. A. Donnelly, W. Barclay, G. Taylor, G. Cooke, H. Ward, A. Darzi, S. Riley, S. C. Robson, N. J. Loman, T. R. Connor, T. Golubchik, R. T. M. Nunez, C. Ludden, S. Corden, I. Johnston, D. Bonsall, C. P. Smith, et al. **2021**, <https://doi.org/10.1101/2021.09.02.21262979>.
- [111] K. K. Riemersma, B. E. Grogan, A. Kita-Yarbro, G. E. Jeppson, D. H. O'Connor, T. C. Friedrich, K. M. Grande, *medRxiv* **2021**, <https://doi.org/10.1101/2021.07.31.21261387>.
- [112] C. M. Brown, *MMWR. Morb. Mortal. Wkly. Rep.* **2021**, *70*, 1059.
- [113] A. Tolk, in *Proc. of the Modeling and Simulation of Complexity in Intelligent, Adaptive and Autonomous Systems 2016 (MSCIAAS 2016) and Space Simulation for Planetary Space Exploration (SPACE 2016)*, April **2016**, pp. 1–7.
- [114] M. S. Khan, M. Samreen, M. Ozair, T. Hussain, J. Gómez-Aguilar, *Eur. Phys. J. Plus* **2021**, *136*, 853.

- [115] H. Lu, Y. Ding, S. Gong, S. Wang, *Math. Biosci. Eng.* **2021**, *18*, 3197.
- [116] G. Rohith, K. Devika, *Nonlinear Dyn.* **2020**, *101*, 2013.
- [117] S. S. Nadim, J. Chattopadhyay, *Chaos, Solitons Fractals* **2020**, *140*, 110163.
- [118] F. Nazarimehr, V.-T. Pham, T. Kapitaniak, *Nonlinear Dyn.* **2020**, *101*, 1681.
- [119] S. L. Rosen, C. P. Saunders, S. K. Guharay, in *Proc. 2012 Winter Simulation Conf. (WSC)*. IEEE, Berlin, Germany, December **2012**, pp. 1–12.
- [120] P. Mellacher, arXiv:2011.06289, **2020**.
- [121] D. Sharma, J.-P. Bouchaud, S. Gualdi, M. Tarzia, F. Zamponi, *PLoS One* **2021**, *16*, e0247823.
- [122] S. M. Sanchez, in *Proc. Winter Simulation Conf. 2014*. IEEE, Savannah, GA, USA, December **2014**, pp. 805–816.
- [123] C. N. Madu, C.-H. Kuei, *Simul. Pract. Theory* **1994**, *2*, 27.
- [124] S. B. Crary, *Analog Integr. Circuits Signal Process.* **2002**, *32*, 7.
- [125] R. B. Gramacy, H. K. Lee, *Technometrics* **2009**, *51*, 130.
- [126] V. Aute, K. Saleh, O. Abdelaziz, S. Azarm, R. Radermacher, *Struct. Multidiscipl. Optim.* **2013**, *48*, 581.
- [127] H. Liu, Y.-S. Ong, J. Cai, *Struct. Multidiscipl. Optim.* **2018**, *57*, 393.