OPPORTUNITIES AND CHALLENGES IN DEVELOPING COVID-19 SIMULATION MODELS: LESSONS FROM SIX FUNDED PROJECTS

Philippe J. Giabbanelli

Dept. Computer Science & Software Engineering
Miami University
205 Benton Hall
Oxford, OH, USA
giabbapj@miamioh.edu

Jennifer Badham
Brian Castellani

Department of Sociology
Durham University
29-32 Old Elvet, Durham, UK
brian.c.castellani@durham.ac.uk
research@criticalconnections.com.au

Hamdi Kavak

Dept. Computational and Data Sciences
George Mason University
227 Research Hall, Fairfax Campus
Fairfax, VA, USA
hkavak@gmu.edu

Vijay Mago

Department of Computer Science
Lakehead University
AT 5027, 955 Oliver Rd
Thunder Bay, ON, Canada
vmago@lakeheadu.ca

Ashkan Negahban

School of Graduate Professional Studies
The Pennsylvania State University
30 E. Swedesford Rd.
Malvern, PA, USA
anegahban@psu.edu

Samarth Swarup

Biocomplexity Institute & Initiative
University of Virginia
P.O. Box 400298
Charlottesville, VA, USA
swarup@virginia.edu

ABSTRACT

The COVID-19 pandemic showed us the importance of modeling and forecasting efforts to guide decision makers. However, a year into the COVID-19 pandemic, the computational science literature lacks a proper internal exploration of the modeling journey of researchers around the world, including how they responded to the shared challenges our community faced such as data limitations, model fitting and working with public stakeholders. The current paper is a detailed examination of the internal processes of six research teams, which were funded in several countries to model COVID-19. Each team was asked to reflect on the research question and how they solved their respective modeling challenges, as well as how, looking back, they would do things differently.

Keywords: Agent-Based Model, Compartmental Model, COVID-19, Data Limitations, Training.

1 INTRODUCTION

From its emergence in 2019, the coronavirus disease (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has claimed over 3.4 million lives at the time of writing (May 2021). Countries such as the USA and the UK are particularly affected, with well over 1,700 deaths per million residents[1]. To exemplify this impact in the USA (Andrasfay and Goldman 2021), we note that 2020 saw a decrease in life expectancy of more than one year and witnessed the largest single-year increase in mortality since 1918, which had both a flu pandemic and the first world war. Although the vaccination campaigns underway in several countries are significantly improving our ability to manage this infectious disease, COVID-19 is a global public health issue grounded in the complexities of human social and health behaviors. Examples include the varying effectiveness of public health campaigns to address the pandemic; the struggle of citizens to comply with lockdown rules, social distancing measures and adjusting their health behaviors; and how the disease played itself out across major social inequalities, particularly amongst the poor, immigrants and ethnicity minorities. Consequently, public health and its arsenal of non-pharmaceutical interventions (e.g., social distancing, mask wearing) will continue to serve an essential public health mission, particularly in light of the possibilities of similar pandemics in the near future. However, these interventions must be carefully designed, given their potential impact on other facets of health (e.g., adverse mental health symptoms) or economic dislocation (Morgan et al. 2021, Bueno-Notivol et al. 2021). The design and evaluation of these interventions is also a complex endeavor given the multiplicity of possible actions (e.g., social distancing in the community, at work, or in schools) and associated parameters (e.g., compliance level), uncertainty in human behaviors, as well as temporal (when do we start and end?), geographical (do we differentiate the intervention across places?), and demographic (e.g., age category) considerations.

Simulation models are one tool used extensively during the pandemic to assist with policy decisions. Early highly-cited modeling studies include the LSHTM or Imperial models in the UK (Flaxman et al. 2020, Davies et al. 2020), or the IHME model (IHME 2020) used by federal and local governments in the USA. Such models were admittedly “rushed” (Adam 2020) given the necessity of making immediate decisions and the paucity of evidence. Later examinations resulted in mixed findings (James et al. 2021), as some initial models “fared very poorly” when their predictions were compared to actual case numbers (Chin et al. 2020) while others had an accurate forecast (Rice et al. 2020). Differences are due to a host of factors, including data limitations (Jewell et al. 2020), the presence of coding bugs (Rahman and Farhana 2020), or limitation of model expressiveness. Although models are imperfect, decision-makers need them to make informed and transparent decisions. Further, policy support may not require precise predictions, but rather qualitative assessment as to the level of risk associated with different decisions. Consequently, *early models* may be imperfect but the practice of modeling continues to flourish and raise the bar in the quality of models (Currie et al. 2020). This increase in standards and expectations is achieved in part through detailed guidance documents on how to improve models over time by promoting practices such as unit testing (Lucas et al. 2020), ensemble modeling (Buckee and Johansson 2020), quantifying and disclosing uncertainties (Ioannidis et al. 2020), or transparency (Barton et al. 2020).

At least four major advances have been made in simulation modeling of Covid-19. First, although modelers often work in isolation and cross-modal comparison is a sporadic practice even in well-defined fields, we have witnessed efforts to promote comparison, sharing, and reuse. For example, the Reich Lab COVID-19 hub[2] provides an accessible repository of model outputs, from early ones to more recent cases. Such repositories make it possible to compare models and identify what works better (thus facilitating the design of future models) as well as create ensembles, which reduces the risk of substantial errors in the decision inputs. Second, despite their imperfections, early models (e.g., SEIR) have influenced public health policies and increased the familiarity of the general public as well as policy-makers with the modeling process, its

[1] See https://www.statista.com/statistics/1104709/coronavirus-deaths-worldwide-per-million-inhabitants/.

[2] See https://viz.covid19forecasthub.org/.

value, and its limitations: our otherwise obscure daily work now shaped national politics and was the focus of the daily press. As a result, societies have become more educated consumers of models, particularly for decision support and 'what-if' analysis. Third, the simulation community has organized, from listservs and other forums to coalitions (Tolk et al. 2020) or consortia[3]. This encouraged an exchange of resources, skills, and ideas, in a spirit of support and comradery while being simultaneously most needed and used. Fourth, there have been major improvements in data, both in gathering and making available case data, and in using fine-grained cellphone-based mobility data in forecasting epidemic spread.

As improvements are realized, our expectations can continue to rise. However, this requires a shift from broad goals (e.g., sharing models, forming working groups, educating a target audience) to more specific aspects of model development, implementation, and application (Giabbanelli et al. 2019, Giabbanelli et al. 2021). Consequently, we need to complement existing high-level guidance (e.g., modelers should perform unit testing, decisions should involve several models) with a detailed assessment of the challenges currently faced by COVID-19 modelers, together with proposed solutions. Although there is a quickly growing number of reviews on COVID-19 models (Mohamadou et al. 2020, Latif et al. 2020, Swapnarekha et al. 2020, Guan et al. 2020), they are limited to an *external* examination that only accounts for the culmination of a modeling project (as conveyed in a publication) rather than the journey. The proliferation of books including COVID-19 models[4] provides numerous case studies and occasionally exposes modeling challenges, but the scattered nature of the studies prevents an identification of *shared* challenges and solutions.

In this paper, we report shared challenges and potential solutions through a detailed examination of internal processes in six research teams, which were funded in several countries to model COVID-19. To complement the abundance of COVID-19 models and the multiplicity of books or reviews, we focus on the experience or 'journey' of the modelers. Each team was asked to reflect on challenges and solutions, both in the past (what *did* they do?) and in retrospect (what *would* they do differently if they were to start the project all over again?). In section 2, we briefly explain how the journey started for each team based on the *goals* of their projects, as it eventually influences the challenges that they encounter and the determination of an appropriate solution such that the model remains fit for purpose. Shared challenges and current solutions regarding data limitations are discussed in section 3. Section 4 is devoted to the relation between modelers, policy makers, and the general public. Concluding perspectives are provided in section 5.

2 BACKGROUND: DEVELOPING SIMULATION MODELS FOR COVID-19

The challenges faced by a modeling team are shaped by a few early determining factors, such as the goals of the model, the techniques used, available resources, or the ability to reuse code. In this section, we briefly cover these aspects across our six projects to delineate the situations through which our experiences were formed, as well as to provide a succinct overview of recurring themes in COVID-19 modeling research.

The emphasis of a COVID-19 simulation model is often on the *policy* goal, which generally consists of predicting population level health outcomes (e.g., number of cases, hospitalizations) as a function of health behaviors (e.g., mobility behaviors, compliance with interventions) and policy interventions (e.g., shelter-in-place ordinances). However, there are numerous other reasons for which we create models (Edmonds et al. 2019). A *scientific* goal may be to understand how a disease is shaped by certain features, even if they are not directly changeable through policy interventions. An *engineering* goal may focus on the creation of a framework or platform, which may then be used by others; for example, this can consist of creating

[3]E.g., the COVID-19 HPC Consortium at https://www.xsede.org/covid19-hpc-consortium.

[4]A few titles on the topic from over 10 books include "Predictive and Preventive Measures for Covid-19 Pandemic", "Artificial Intelligence for COVID-19", "Computational Intelligence Techniques for Combating COVID-19", "Predictive Models for Decision Support in the COVID-19 Crisis", or "COVID-19: Prediction, Decision-Making, and its Impacts".

a scalable solution to perform data-driven simulations with complex behaving agents. Collectively, the six projects in which the authors are involved cover all these goals (Table 1).

A detailed comparison of the projects is beyond the scope of this study, as our interest lies in identifying challenges and solutions. For such comparisons, we refer the readers to (Panovska-Griffiths et al. 2021, Adiga et al. 2020). In each project, we asked modelers *why* they chose a certain technique. Three reasons were frequently mentioned, in line with other studies on modeling practices (Voinov et al. 2018):
(1) *familiarity* with the technique, through experience and training. Challenges may be different for teams that opted for a new modeling paradigm and then sought to develop new competencies.
(2) the amount of *data* available. Our teams cover different situations, as some chose an ABM partly to leverage an abundance of data (e.g., granular visit data, Cuebiq's mobility data) while others opted for a compartmental model due to data limitations. These initial differences in data are important as they translate to contrasting modeling choices and eventually to different challenges.
(3) the ability to *represent* policy-relevant constructs. For example, an agent-based model allows to analyze the effect of individual-level interactions and control strategies such as quarantines.

In four projects, the code was entirely written from scratch. One project used a library (Java 2APL) and another project re-used a modeling platform developed by another group (Covasim). We note that this tendency to create implementations from scratch is not limited to COVID-19: our recent study on practices in developing artificial societies (Vendome et al. 2020) found that most projects do not use libraries, while a minority uses libraries primarily to deal with data types (e.g., spatial, network) rather than to support model development (e.g., agent cognition). It is thus possible that some of the challenges experienced by modelers in the next sections stem from the high cost of re-creating solutions and re-discovering issues, instead of seeking and integrating building blocks developed for a similar purpose. As *reuse* is a known challenge for social-behavioral simulations, we refer the reader to (Carley 2019) for a separate discussion.

Large-scale simulations can require a significant amount of resources, as illustrated by several of our projects. For example, the ABM for the New York City case study includes more than 8 million agents and ran on Google cloud computing resources with 96 vCPUs and 624GB of RAM. Similarly, a US-wide ABM utilized Microsoft Azure and 1TB of RAM for each simulation run. Since all projects were funded, the necessary computational resources were made available hence challenges during the modeling journey do not include computational limitations. However, we acknowledge that other groups may face limitations either in securing resources or in using them for the first time. For these topics, we refer the reader to the literature on high-performance computing and Covid-19 (Hack and Papka 2020, Machi et al. 2021).

3 ADDRESSING DATA LIMITATION ACROSS STAGES OF THE MODELING PROCESS

Shelton's commentary summarizes the paradox of COVID-19 research: although there is a "dizzying amount of data being generated [leading] to calls to proclaim this the first 'data-driven pandemic', [we face the] general inadequacy of our data infrastructures and assemblages to solving pressing social issues" (Shelton 2020). Data issues challenged all of our teams, affecting our ability to estimate the probability distribution or even obtain a reliable point-estimate for the disease-related parameters such as transmission rate, hospitalization period, and fatality rate per age group or preexisting condition type. As a collective, we observed an impact of data limitation on each step of the modeling and simulation process. To **initialize a simulation** of a disease spread, we need to 'seed it' with infected cases. In the case of Thunder Bay, there were *almost no active* cases for many weeks. Data from comparable cities could not be obtained, since the city of Thunder Bay is geographically quite isolated and spread out unlike other metropolitan cities of Canada. For example, the numbers from the city of Toronto would not be directly applicable (even after scaling) to the situation of Thunder Bay. Initialization also encompasses the creation of a virtual population that is heterogeneous in terms of disease-relevant factors (e.g., preexisting condition to infer health outcomes; household size to infer

Author	Model objectives	Target population	Technique	Refs
PJG	Assess how the number and timing of cases depends on vaccines (efficacy, capacity, interest from the population) and non-pharmaceutical interventions (face masks, working remotely, stay-at-home orders, testing, contact tracing, and quarantining)	US nationwide	Agent-Based Modeling (ABM). Coded in Python with Covasim	5
JB, BC	Estimate hospital admissions and occupied beds over time based on scenarios of implemented policy measures, particularly activity restrictions and self-isolation on symptoms to assist planners to understand likely staffing level needs and whether proposed measures would achieve the desired outcome.	North East region of England, but not specifically customized to that population	ABM. Coded in NetLogo	6
HK	How can the inclusion of the spatial and temporal variation of human behavior (e.g., mobility behavior) improve simulation models of disease spread?	US. Proof-of-concept at county level; can be scaled to larger resolutions e.g. states.	ABM. Coded in Python	7
VM	Help public health agencies of Thunder Bay, ON to plan required resources at the Thunder Bay Regional Health Science Centre	City of Thunder Bay and Indigenous people of NAN communities	Compartmental. Coded in Java with Laravel for web portal	8
AN	Developing integrated mobility and epidemic vulnerability measures and assessing their efficacy for allocation of intervention resources such as random testing sites in communities (e.g., census tracts) and nodes within a transportation network (e.g., subway stations)	Any mobility network and metropolitan area. Used New York City for case studies	ABM. Coded in Repast Simphony	9
SS	Behaviors are generally modeled using a compliance rate, which implicitly assumes independent decisions by each agent without regard to any normative reasoning. We seek to simulate the drivers (e.g., peer perception, demographics) and effects of peer compliance onto epidemic spread.	US State of Virginia. Methods can apply to any region.	ABM. Coded in Python	10

Table 1: Overarching goals, context, and techniques for the six simulation studies performed by the authors.
[5]See (Li, Giabbanelli, et al. 2021) and (Giabbanelli and Li 2020) [6]See (Badham et al. 2021) and (Castellani and Caiado 2020) [7]See (Pesavento et al. 2020) [8]See (Savage et al. 2020) [9]See (Speir and Negahban 2020) [10]See (de Mooij et al. 2021)

spreading patterns) such that an accurate distribution of results can eventually be given to policy-makers. While related data is publicly available for cities, counties, and states in the U.S., there is no single database that provides the data all in one place. There is thus a major challenge of *identifying and integrating datasets*.

Once a simulation is initialized, we apply **rules** stipulating how individuals behave and how the disease affects them. Although some COVID-19 models may ignore mobility or abstract it as entirely random (e.g., agents are spheres drifting in a vaccuum), accurately capturing *mobility and commute patterns* helps to track the spread of a pathogen and also to estimate the response of people to lockdowns or other activity restrictions. We used datasets such as the free data from SafeGraph COVID-19 Data Consortium on foot traffic. Directly using it in a simulation introduces a bias, as the data only captures smartphone users. This bias matters for COVID-19 simulations because the elderly are at risk for serious health consequences, but their smartphone use is much lower than in the young population. Without a reliable way to do age-based re-sampling from aggregated data, mobility patterns have to be corrected based on subject matter expert suggestions or other data sources yet to come. While mobility is a classic aspect of epidemic simulations, challenges specific to COVID-19 were also encountered when developing the behavioral rules. In particular, about 40% of COVID-19 cases are *asymptomatic*, which can impact the overall progression of the epidemic and population-level outcomes. Asymptomatic carriers do not realize they are infectious, so they might put more people at risk if they do not comply with behavioral norms such as mask-wearing and physical distancing. Susceptible individuals may feel safer if they think that others are not infected, thus also potentially impacting their compliance with interventions. Although CDC data estimates the proportion of asymptomatic cases, and case data may help to infer the relative infectivity of asymptomatic infectious people, there are still glaring data gaps on the behavioral side.

If a simulation can be initialized and updated through evidence-based rules, then we can produce results and examine their quality through a **validation** process. Comparison with observed real-world outcomes/data is a common way of validating simulation models. However, this is challenging for simulation models of COVID-19 as an ongoing pandemic. Due to the many asymptomatic cases, data on number of cases is not reliable and at best represents the portion of cases with mild to severe symptoms that actually sought medical help and *were tested*. Access to tests was limited, especially in the early stages, and logistics of large-scale testing also limited roll-out, especially in more rural areas. Estimates of the *true burden* of disease have varied enormously, with an antibody experiment suggesting a case number ten times higher than reported confirmed cases[11] while CDC estimates suggest that only one in seven cases was initially reported (Reese et al. 2020). Hospitalization data is more reliable than case data (and adequate to estimate hospital burden), but still has limitations since hospitals were selective in admitting COVID-19 patients[12].

To facilitate the identification of solutions, we group these data-centric challenges into two categories. First, values can be *unknown*, such as the exact number of people who are infected in a specific area. This is a salient problem for models which support local decisions, such as forecasting staffing demands at a given hospital. Second, some aspects may be measured but the data is *unavailable*, for example due to privacy restrictions for health data (e.g., individual level hospital records could be used to calculate the distribution of length of stay in hospital conditional on survival or transfer to critical care) or a lack of coordination among producers and users of information[13]. To deal with unknowns, we identified and adopted three main solutions, detailed from least to most complex.

[11] See news coverage at https://www.nytimes.com/2020/04/23/nyregion/coronavirus-new-york-update.html.

[12] Reports documented that U.S. hospitals turned away patients who later developed severe symptoms and eventually died from COVID-19 complications after they were finally admitted into the hospital. Sample coverage includes https://www.nytimes.com/2020/04/19/nyregion/new-york-new-jersey-coronavirus-hospitals.html.

[13] The Canadian military's intelligence had gathered information on COVID-19, but it remained classified hence unavailable to the Public Health Agency of Canada that was in charge of guiding the national response to the pandemic in its early days. See e.g. https://www.cbc.ca/news/politics/covid-military-medical-intelligence-1.5866627.

① Use the closest available high-quality data proxy or *generator*. For example, there are pre-calibrated synthesized population databases for most counties and states in the U.S. that are specifically suitable for agent-based models (Wheaton et al. 2009) and can be used as a starting point for simulating infectious disease outbreaks[14]. A recent machine learning study showed that real-world and synthetic datasets could serve to predict COVID-19 death cases with a similar accuracy (El Emam 2021). To create detailed simulation models, specialized generators for COVID-19 including age-adjusted contact patterns are available via COVASIM, OpenABM-Covid19, or COMOKIT (Hinch et al. 2020, Gaudou et al. 2020, Kerr et al. 2020). These micro-based proxies can be further calibrated with macro-level proxies. For example, the hospitalisation rate is observable and serves as a proxy for the unobservable prevalence in a simulated epidemic.

② Account for the vagaries of testing and the ever-changing nature of vaccination or non-pharmaceutical interventions via well-crafted *scenarios*. As detailed guidance on scenario development already exists (Amer et al. 2013), we stress its relevance to COVID-19 simulation. An intervention involves a multiplicity of variables: testing alone encompasses a daily capacity for testing, a delay for test results, whether another test is needed to end a quarantine, etc. Future values for some of these variables may be unknown and relying on a single scenario is problematic, as exemplified by the federal U.S. 'Operation Warp Speed' vaccine scenario that pledged to administer 20 million doses in December but ultimately managed 3 million doses. It is thus essential to craft a set of scenarios accounting for a *range* of possible outcomes, ideally consisting of an upper- and lower-bound as well as equidistant outcomes within these bounds. Practically, a modeler can simulate the outcomes (e.g., total number of death) across a large number of combinations of possible values and then select as scenarios a small number of configurations spanning markedly different outcomes[15]. This approach can be challenging if a model accounts for many variables, since the number of combinations to generate grows exponentially – the next approach handles this situation.

③ Navigating an enormous search space can be done efficiently with *designed experiments* (Sanchez 2005). A binary factorial design (also known as a 2^k) first identifies two values for each parameter that will have the most different impacts (e.g., least/most amount of daily tests) and then simulates all combinations. Results are analyzed to identify the contribution of each parameter *and interactions* to the variance, which allows to then carefully fix the value of parameters with a low impact while dedicating more computing power to a fine-grained assessment of parameters with a larger impact. There are two challenges specific to COVID-19 models in this process. First, we may not know which values have the most different impacts; for example, which two values should be selected for the network in which mask wearing is enforced (home, school, work, community?) or which values yield the most different results for tests in quarantine (testing at the beginning? at the end? both?). Second, not all combinations may be valid as interventions impact each other (e.g., whether a person wears mask in public is superseded by quarantining). Statisticians would thus have a role to play in devising the right scenarios for COVID-19 models as part of interdisciplinary teams.

To address unavailable data, we are *constantly on the lookout for new research* findings and changes in intervention scenarios while creating a model. We thus supplement our evidence base by reports and other information published by people who do have access, including digitizing figures. This requires a frequent screening across multiple aggregation platforms[16]. Although this strategy can be effective to identify helpful pre-prints on epidemiological data such as (Ssentongo et al. 2020, Linton et al. 2020), there are caveats when modeling interventions as they are extremely fluid and news can conflict within days. For example, on a Tuesday, the US government vowed to release a stockpile of vaccines, which would have been captured in a model by increased vaccine capacity. By Friday, the administration acknowledged that such a stockpile did not exist, as the vaccine reserve was already exhausted[17]. Numbers can also be revised long after their

[14]For a generator extended to handle COVID-19, see https://github.com/synthetichealth/synthea by MITRE.

[15]See (Li, Giabbanelli, et al. 2021) as an example to identify six non-pharmaceutical scenarios and two vaccine scenarios.

[16]Research papers are tracked via https://twitter.com/COVID_Papers or searchable by category via https://outbreak.info/.

[17]See e.g. https://www.washingtonpost.com/health/2021/01/15/trump-vaccine-reserve-used-up/.

publication, as illustrated by the under-reporting of nursing home COVID deaths in New York[18], hence models that passed validation based on earlier data may no longer be valid. Both cases provide cautionary tales: it may be necessary to use reports or news to deal with data unavailability while modeling an ongoing pandemic, but the uncertainty about the numbers should be stated and the sensitivity of the results should be assessed (Edeling et al. 2021) via scenarios as explained in item ② above.

4 CREDIBILITY AND COMMUNICATION

As detailed in the 'Aqua Book', the development of a model for high-level decision-making activities is an interdisciplinary endeavor that involves different *roles*, each with a set of expectations. There is at least a customer or decision-maker (also known as a 'commissioner' of a model) and a modeler. During an engagement phase, the commissioner(s) and the modeler(s) are expected to jointly develop a scope of work and ensure appropriate resources such as technical skills, subject-matter domain expertise, or data (HM Treasury 2015, p.21). In several COVID-19 projects (not limited to those outlined in this paper), we note the absence of a commissioner hence modelers are directly aiming at influencing public policy. This is challenging for both the public and modelers. Modelers may not be accustomed to their ideas being directly (mis)used in the public domain and have a limited experience in working with public stakeholders or managing a public presence. As they have not commissioned the model, the general public may struggle to understand its purpose and limitations; for example, they may not be aware of the differences between a 'toy model' and an 'empirical model.' As a result, there is a misalignment between expectations (e.g., 'a model is an oracle') and achievements (e.g., a model as a virtual laboratory). This is manifested in misguided commentaries that equate imperfect with incorrect, without considering whether the imperfections actually changed the policy decisions that had to be made at the time. It should be stressed that waiting for a better model is also a decision *not* to act to restrict viral transmission until that better model is available.

To address this situation, the modeling community needs upskilling in public policy and science communication. This can be achieved through workshops or seminars, conducted with public policy evaluators and policy makers as well as modelers who are experienced in working at this interface. As the general audience is unlikely to examine a model by parsing its code or reading an ODD specification, modelers also need efficient practices in visual communication, which starts by assessing which data visualizations have been helpful or misleading in getting their points across (Bowe et al. 2020). The need to train modelers does not stop there, as COVID-19 research also requires competencies in computational epidemiology[19]. Many universities have initiatives to connect researchers to policy-makers based on shared interests, thus providing researchers with an opportunity to engage legislators at state and federal levels directly[20]. To increase our overall preparedness in computational epidemiology and thus have the capacity to address future emergencies, we would have to either ensure that students are already trained in modeling and/or rely on the reuse of building blocks instead of learning to develop models from scratch.

Several actions would improve how the general public perceives the value of models. Simple, education oriented models are needed for the public to help understand difficult and counter-intuitive notions, such as why there is a delay between restrictions and improvements. While there have been some excellent initiatives, few resources have been allocated and these are not widely promoted. Modelers also need greater consistency in their public communication. The systematic inclusion of scenarios and a *range* of possible outcomes would thus convey the important message that several possible futures are presented, hence shifting away from the notion of an oracle with a unique answer that is either valid or a failure. It is also essential to stress to the public that we need several models.

[18] See e.g. https://www.nytimes.com/2021/03/04/nyregion/cuomo-nursing-home-deaths.html.

[19] Researchers such as https://www.youtube.com/user/NathanielOsgood have released recordings on teaching ABM in AnyLogic for several years. The Santa Fe Institute also provides training on modeling, under the broader umbrella of complexity.

[20] C.f. Penn State University's "Research to Policy Collaboration" at https://www.research2policy.org/covid-19

5 CONCLUSION

Paving the way for the development of the next COVID-19 models requires an assessment of challenges and an identification of solutions. In this paper, the assessment was performed by modelers who independently worked on six COVID-19 projects. Each project has resulted in publications (Table 1), which can provide technical details for readers interested in a specific application context. The six modeling teams varied widely on several aspects, such as what was represented in a model (e.g., vaccines), model resolution, and policy engagement. Nonetheless, they faced similar challenges in many respects, particularly dealing with either unknown or unavailable data, and managing a public presence. To shift from a problem-focused assessment to a solution-oriented approach, we discussed several opportunities, including the use of rapidly emerging centralized data repositories, processes to deal with unknowns at several levels, and how to build credibility. We note that our generalizable experiences are *not* about the similarity of the mathematics or the contents of the models: the themes of data access and credibility are about using models in policy, when insights are needed immediately to inform public health measures. Our experiences thus contribute to guiding and enhancing modeling efforts for future pandemics and other global crises that require urgent response and decision-making informed by models. As such applications require an interdisciplinary effort, our perspectives should not be construed as the only way to create better models and improve their impact. Rather, our views should be complemented with guidance from epidemiologists and public health officials.

ACKNOWLEDGMENTS

PJG is supported by a philanthropic grant from Microsoft AI for Health. In the UK, JB and BC thank the Medical Research Council (MC_PC_18045) and the Economic and Social Research Council; team members were also supported by the Economic and Social Research Council (ES/S000402/1) while publication fees were made possible via an IBM grant for supporting open access COVID-19 research. In the US, HK is supported by the National Science Foundation (DEB-2030685) and internal funds at George Mason University for summer research (ASSIP 2020-21, STIP 2021). In Canada, VM was funded by the Lakehead University (Romeo #1467916) and the Northern Ontario Academic Medical Association Fund (C-14-2-18). In the USA, AN is supported by the Institute for Computational and Data Sciences at The Pennsylvania State University and computational resources provided by a Google Cloud COVID-19 Research Grant. SS thanks the National Science Foundation (CCF-1918656) and a Defense Threat Reduction Agency subcontract/ARA (S-D00189-15-TO-01-UVA). The authors are collectively indebted to many students, colleagues, and collaborators who contributed to the Covid-19 projects and enabled the reflections in this paper.

REFERENCES

Adam, D. 2020. "Special report: The simulations driving the world's response to COVID-19.". *Nature* vol. 580 (7803), pp. 316.

Adiga, A. et al. 2020. "Mathematical models for covid-19 pandemic: a comparative analysis". *Journal of the Indian Institute of Science*, pp. 1–15.

Amer, M. et al. 2013. "A review of scenario planning". *Futures* vol. 46, pp. 23–40.

Andrasfay, T., and N. Goldman. 2021. "Reductions in 2020 US life expectancy due to COVID-19 and the disproportionate impact on the Black and Latino populations". *PNAS* vol. 118 (5).

Badham, J. et al. 2021. "Justified Stories with Agent-Based Modelling for Local COVID-19 Planning.". *Journal of Artificial Societies & Social Simulation* vol. 24 (1).

Barton, C. M. et al. 2020. "Call for transparency of COVID-19 models". *Science* vol. 368 (6490), pp. 482–483.

Bowe, E. et al. 2020. "Learning from lines: Critical COVID data visualizations and the quarantine quotidian". *Big data & society* vol. 7 (2), pp. 2053951720939236.

Buckee, C. O., and M. A. Johansson. 2020. "Individual model forecasts can be misleading, but together they are useful". *European Journal of Epidemiology* vol. 35 (8), pp. 731–732.

Bueno-Notivol, J. et al. 2021. "Prevalence of depression during the COVID-19 outbreak: a meta-analysis of community-based studies". *Int J Clinical and health psychology* vol. 21 (1), pp. 100196.

Carley, K. M. 2019. "Social-Behavioral Simulation: Key Challenges". *Social-Behavioral Modeling for Complex Systems*.

Castellani, B., and C. Caiado. 2020, may. "Coronavirus: why we need local models to successfully exit lockdown".

Chin, V. et al. 2020. "A case study in model failure? COVID-19 daily deaths and ICU bed utilisation predictions in New York State". *European Journal of Epidemiology* vol. 35 (8), pp. 733–742.

Currie, C. S. et al. 2020. "How simulation modelling can help reduce the impact of COVID-19". *Journal of Simulation* vol. 14 (2), pp. 83–97.

Davies, N. G. et al. 2020. "Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study". *The Lancet Public Health* vol. 5 (7), pp. e375–e385.

de Mooij, J. et al. 2021. "Quantifying the Effects of Norms on COVID-19 Cases Using an Agent-based Simulation". In *Proc. 22nd Int. Workshop on Multi-Agent-Based Simulation*.

Edeling, W. et al. 2021. "The impact of uncertainty on predictions of the CovidSim epidemiological code". *Nature Computational Science* vol. 1 (2), pp. 128–135.

Edmonds, B. et al. 2019. "Different Modelling Purposes". *Journal of Artificial Societies and Social Simulation* vol. 22 (3), pp. 6.

El Emam, K. a. 2021. "Evaluating the utility of synthetic COVID-19 case data". *JAMIA Open* vol. 4 (1), pp. ooab012.

Flaxman, S. et al. 2020. "Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in European countries: technical description update". *arXiv preprint arXiv:2004.11342*.

Gaudou, B. et al. 2020. "Comokit: A modeling kit to understand, analyze, and compare the impacts of mitigation policies against the covid-19 epidemic at the scale of a city". *Front. public health* vol. 8.

Giabbanelli, P. J. et al. 2019. "Ideal, best, and emerging practices in creating artificial societies". In *2019 Spring Simulation Conference (SpringSim)*, pp. 1–12. IEEE.

Giabbanelli, P. J. et al. 2021. "The application of modeling and simulation to public health: Assessing the quality of Agent-Based Models for obesity". *Simul Model Pract Th* vol. 108, pp. 102268.

Giabbanelli, P. J., and J. Li. 2020. "Identifying Synergistic Interventions to Address COVID-19 Using a Large Scale Agent-Based Model". *medRxiv*.

Guan, J. et al. 2020. "Modeling the transmission dynamics of COVID-19 epidemic: a systematic review". *Journal of Biomedical Research* vol. 34 (6), pp. 422.

Hack, J. J., and M. E. Papka. 2020. "The US High-Performance Computing Consortium in the Fight Against COVID-19". *Computing in Science & Engineering* vol. 22 (6), pp. 75–80.

Hinch, R. et al. 2020. "OpenABM-Covid19-an agent-based model for non-pharmaceutical interventions against COVID-19 including contact tracing". *medRxiv*.

HM Treasury 2015. "The Aqua Book: guidance on producing quality analysis for government".

IHME 2020. “Forecasting COVID-19 impact on hospital bed-days, ICU-days, ventilator-days and deaths by US state in the next 4 months”. *MedRxiv*.

Ioannidis, J. P. et al. 2020. “Forecasting for COVID-19 has failed”. *International journal of forecasting*.

James, L. P. et al. 2021. “The use and misuse of mathematical modeling for infectious disease Policymaking: lessons for the COVID-19 pandemic”. *Medical Decision Making* vol. 41 (4), pp. 379–385.

Jewell, N. P. et al. 2020. “Predictive mathematical models of the COVID-19 pandemic: underlying principles and value of projections”. *Jama* vol. 323 (19), pp. 1893–1894.

Kerr, C. C. et al. 2020. “Covasim: an agent-based model of COVID-19 dynamics and interventions”. *medRxiv*.

Latif, S. et al. 2020. “Leveraging data science to combat covid-19: A comprehensive review”. *IEEE Transactions on Artificial Intelligence*.

Li, J., P. Giabbanelli et al. 2021. “Returning to a Normal Life via COVID-19 Vaccines in the United States: A Large-scale Agent-Based Simulation Study”. *JMIR Medical Informatics* vol. 9 (4), pp. e27419.

Linton, N. M. et al. 2020. “Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: a statistical analysis of publicly available case data”. *Journal of clinical medicine* vol. 9 (2), pp. 538.

Lucas, T. C. et al. 2020. “Responsible modelling: Unit testing for infectious disease epidemiology”. *Epidemics*, pp. 100425.

Machi, D. et al. 2021. “Scalable epidemiological workflows to support covid-19 planning and response”. *medRxiv*.

Mohamadou, Y. et al. 2020. “A review of mathematical modeling, artificial intelligence and datasets used in the study, prediction and management of COVID-19”. *Applied Intelligence* vol. 50 (11), pp. 3913–3925.

Morgan, A. K. et al. 2021. “The effects of COVID-19 on global economic output and sustainability: evidence from around the world and lessons for redress”. *Sustainability: Science, Practice and Policy* vol. 17 (1), pp. 77–81.

Panovska-Griffiths, J. et al. 2021. “Mathematical modeling as a tool for policy decision making: Applications to the COVID-19 pandemic”. *Handbook of Statistics*.

Pesavento, J. et al. 2020. “Data-driven mobility models for COVID-19 simulation”. In *Proceedings of the 3rd ACM SIGSPATIAL Int. Workshop Advances in Resilient and Intelligent Cities*, pp. 29–38.

Rahman, A., and E. Farhana. 2020. “An Exploratory Characterization of Bugs in COVID-19 Software Projects”. *arXiv preprint arXiv:2006.00586*.

Reese, H. et al. 2020. “Estimated incidence of COVID-19 illness and hospitalization–United States, February–September, 2020”. *Clinical infectious diseases*, pp. ciaa1780.

Rice, K. et al. 2020. “Effect of school closures on mortality from coronavirus disease 2019: old and new predictions”. *bmj* vol. 371.

Sanchez, S. M. 2005. “Work smarter, not harder: guidelines for designing simulation experiments”. In *Proc. Winter Simulation Conf.*, pp. 14–pp. IEEE.

Savage, D. et al. 2020. “Investigating the implications of COVID-19 for the rural and remote population of Northern Ontario using a mathematical model”. *medRxiv*.

Shelton, T. 2020. “A post-truth pandemic?”. *Big Data & Society* vol. 7 (2), pp. 2053951720965612.

Speir, C., and A. Negahban. 2020. “Analyzing Covid-19 Control Strategies in Metropolitan Areas: A Customizable Agent-Based Simulation Tool”. In *Proceedings of the 20 Winter Simulation Conference*.

Ssentongo, P. et al. 2020. "Association of cardiovascular disease and 10 other pre-existing comorbidities with COVID-19 mortality: A systematic review and meta-analysis". *PloS one* vol. 15 (8), pp. e0238215.

Swapnarekha, H. et al. 2020. "Role of intelligent computing in COVID-19 prognosis: A state-of-the-art review". *Chaos, Solitons & Fractals* vol. 138, pp. 109947.

Tolk, A. et al. 2020. "Computational Decision Support for the COVID-19 Healthcare Coalition". *Computing in Science & Engineering*.

Vendome, C. et al. 2020. "How Do Modelers Code Artificial Societies? Investigating Practices and Quality of Netlogo Codes from Large Repositories". In *2020 Spring Simulation Conf.*, pp. 1–12. IEEE.

Voinov, A. et al. 2018. "Tools and methods in participatory modeling: Selecting the right tool for the job". *Environmental Modelling & Software* vol. 109, pp. 232–255.

Wheaton, W. D. et al. 2009. "Synthesized population databases: A US geospatial database for agent-based models". *Methods report (RTI Press)* vol. 2009 (10), pp. 905.

AUTHOR BIOGRAPHIES

PHILIPPE J. GIABBANELLI, Ph.D., is an Associate Professor in the Department of Computer Science & Software Engineering at Miami University (USA). His research interests include network science, machine learning, and simulation and applied to human health behaviors. His email address is giabbapj@miamioh.edu. His website is https://www.dachb.com.

JENNIFER BADHAM, Ph.D., is a research associate in the Department of Sociology at Durham University (UK). She received her PhD from University of New South Wales. Her research uses computational methods to investigate how individual decisions, social structure and social phenomena mutually influence and constrain each other. Her email address is research@criticalconnections.com.au.

BRIAN CASTELLANI, Ph.D., FAcSS, is a Professor in the Department of Sociology, Durham University (UK) and Adjunct Professor of Psychiatry at Northeastern Medical University (USA). His research involves the development of complexity theory and method in application to public health, the psychology of public opinion, and public policy. His email address is brian.c.castellani@durham.ac.uk.

HAMDI KAVAK is an Assistant Professor in the Computational and Data Sciences Department and co-director of the Center for Social Complexity at GMU. His research interests lie at the intersection of data science and modeling and simulation. His email and website addresses are hkavak@gmu.edu and http://www.hamdikavak.com.

VIJAY MAGO, Ph.D. is an Associate Professor and Chair in the Department of Computer Science at Lakehead University (Canada). He obtained his PhD. from Panjab University (India) and leads a research lab (https://www.datalab.science) focusing on text analytics, data mining, machine learning and natural language processing. His email address is vmago@lakeheadu.ca.

ASHKAN NEGAHBAN, Ph.D., is an Assistant Professor of Engineering Management at the School of Graduate Professional Studies at Penn State University (USA). He has a PhD and MS from Auburn University and a BS from University of Tehran. His research involves stochastic simulation methods. His email and web addresses are anegahban@psu.edu and http://ashkannegahban.com.

SAMARTH SWARUP, Ph.D., is a Research Associate Professor in the Biocomplexity Institute & Initiative at the University of Virginia (USA). He received his Ph.D. from the University of Illinois at Urbana-Champaign. His research interests are in large-scale data-driven agent-based simulations, simulation analytics, machine learning, and multi-agent systems. His email address is swarup@virginia.edu.