DEpiABS: Differentiable Epidemic Agent-Based Simulator

Zhijian Gao
Nanyang Technological University
Singapore, Singapore
zhijian001@e.ntu.edu.sg

Shuxin Li*
Nanyang Technological University
Singapore, Singapore
shuxin.li@ntu.edu.sg

Bo An
Nanyang Technological University
Singapore, Singapore
boan@ntu.edu.sg

ABSTRACT

The COVID-19 pandemic highlighted the limitations of existing epidemic simulation tools. These tools provide information that guides non-pharmaceutical interventions (NPIs), yet many struggle to capture complex dynamics while remaining computationally practical and interpretable. We introduce DEpiABS, a scalable, differentiable agent-based model (DABM) that balances mechanistic detail, computational efficiency and interpretability. DEpiABS captures individual-level heterogeneity in health status, behaviour, and resource constraints, while also modelling epidemic processes like viral mutation and reinfection dynamics. The model is fully differentiable, enabling fast simulation and gradient-based parameter calibration. Building on this foundation, we introduce a z-score-based scaling method that maps small-scale simulations to any real-world population sizes with negligible loss in output granularity, reducing the computational burden when modelling large populations. We validate DEpiABS through sensitivity analysis and calibration to COVID-19 and flu data from ten regions of varying scales. Compared to the baseline, DEpiABS is more detailed, fully interpretable, and has reduced the average normal deviation in forecasting from 0.97 to 0.92 on COVID-19 mortality data and from 0.41 to 0.32 on influenza-like-illness data. Critically, these improvements are achieved without relying on auxiliary data, making DEpiABS a reliable, generalisable, and data-efficient framework for future epidemic response modelling.

KEYWORDS

Epidemiology; Differentiable Agent-Based Model

1 INTRODUCTION

The COVID-19 pandemic exposed critical vulnerabilities in global public health systems, demonstrated by delayed and inappropriate implementation of non-pharmaceutical interventions (NPIs) [61]. The effectiveness of interventions depends fundamentally on the

*Corresponding author.

quality of information available to policymakers, such as the expected infection and mortality trends, to which epidemic modelling methods contribute immensely. To provide *insightful, timely,* and *trustworthy* information, epidemic models must represent real-world epidemic dynamics *faithfully* [40], *expeditiously* [10], and *transparently* [71]. Given the inadequacies in the control and mitigation of COVID-19 recorded in multiple countries [30, 38, 49], the development of epidemic models with *comprehensive advantages* over existing models in all 3 criteria, namely *simulation fidelity, computational efficiency,* and *mechanistic interpretability,* is crucial for preparing against future epidemics.

These criteria are inherently in tension, bringing trade-offs that have been addressed differently by multiple model paradigms and design techniques (**Table 1**), but never truly reconciled. *Traditional models* such as SIR [47] prioritise *efficiency over fidelity,* and researchers have attempted to enrich them with features like age structure or network-based transmission to improve realism [35, 74]. *Agent-based models (ABMs) reverse this priority* by explicitly representing individual heterogeneity; for example, Covasim [48] simulates disease transmission through diverse social networks while optimising runtime and memory usage to retain feasible efficiency. More recent advances attempt *introducing differentiability into ABMs,* enabling GPU acceleration and gradient-based optimisation without significant losses in realism. Two main approaches have emerged: *ABM with neural surrogates,* and *differentiable ABMs (DABMs). ABMs with neural surrogates* approximate model dynamics using artificial neural networks (ANN) [6, 7], boosting computational efficiency without losing simulation fidelity, yet *sacrificing mechanistic interpretability. DABM, as a paradigm,* reformulates model dynamics as tensor operations that preserve all details, therefore *theoretically reconciles* the existing trade-offs. However, most existing DABMs [19, 20, 87] integrate *neural components* to infer epidemiological parameters from auxiliary data, which reintroduces *opacity.* In conclusion, all existing models have non-negligible drawbacks in at least one criterion.

Therefore, we propose *DEpiABS,* a *structure-centric* DABM that achieves *comprehensive advantages* and fully leverages the theoretical potential of the paradigm. DEpiABS includes granular features such as *resource-constrained behaviours, viral mutation, and reinfection* with key mechanisms validated via *sensitivity analysis,* ensuring *high simulation fidelity.* The associated computational cost is reduced through *differentiable approximation and tensorisation.* Relying solely on its ANN-free, *fully transparent structure,* the model achieves *global mechanistic interpretability.* We further optimise the design so that runtime scales *linearly* with both population size and simulation length, and introduce a novel *z-score-based scaling method* that maps outputs from fixed-size populations to any scales while preserving statistical patterns, effectively *decoupling simulation cost from population size.* Finally, we validate DEpiABS's

Paradigm	Model	Simulation Fidelity				Mechanistic Interpretability	Computational Efficiency
		Contact Heterogeneity	Behavioural Autonomy Simulation	Biological Process Modelling	Real Data Validation	Full Transparency	Runtime (1M population, 30 days)
Non-ABM	FIGI-Net [78]	✗	✗	✗	✓	✗	<1s
	SIR [47]	✗	✗	✗	✓	✓	<1s
ABM	COVID-ABS [77]	✓	✓	✗	✗	✓	>10 days
	Covasim [48]	✓	✗	✗	✓	✓	<1min
	CityCOVID [67]	✓	✓	✓	✓	✓	<20s (HPC)
	ABM+NN [25]	✓	✓	✗	✓	✗	<1s
DABM	DeepABM [19]	✓	✗	✗	✗	✗	<5min
	JUNE-NZ [87]	✓	✗	✗	✓	✗	unspecified
	GradABM [20]	✓	✗	✗	✓	✗	<10s
	DEpiABS (ours)	✓	✓	✓	✓	✓	<1.5s

Table 1: Comparison of modelling paradigms and representative models across key evaluation criteria. Biological processes modelling refers to the modelling of epidemic dynamics beyond transmission, including mutation and symptom progression.

forecasting ability across *multi-regional* and *multi-epidemiological* datasets. By outperforming *GradABM* without using auxiliary data, we demonstrate superior *forecasting ability* and *data efficiency*. In summary, our contributions are:

- We introduce *DEpiABS, a fine-grained and fully interpretable DABM designed around a novel structure-centric principle*, that is scalable to any population sizes and outperforms the state-of-the-art in forecasting on multi-regional, multi-disease data with higher data efficiency, demonstrating comprehensive advantages over existing models in all three criteria.
- We resolve the scalability challenge of ABMs with negligible compromise in realism by *combining end-to-end differentiability with a z-score-based scaling method*, which enables simulating epidemics in large populations using a significantly smaller model population size.
- We demonstrate that *models under structure-centric design that prioritises feature granularity over data volume* can achieve performance comparable to, or even exceeding, data-centric benchmarks built around the opposite trade-off.

2 RELATED WORK

2.1 Traditional models

Traditional models generally fall into three categories—compartmental, metapopulation, and statistical [69]. Compartmental models in the paradigm of SIR [47] are popular for their tractability [3, 16, 23] but are deemed oversimplified [5]. Metapopulation models introduce geographic structures [22, 24, 68] that enable regional analyses, but still overlook individual heterogeneity [8]. Statistical models emphasise predictive accuracy [12, 28, 78], yet lack causal interpretability and do not explicitly model the epidemic processes. While all have provided valuable insights, they remain limited in realistically capturing complex epidemic dynamics.

2.2 Agent-based models

Agent-based models (ABMs) simulate disease spread through agents in structured environments [2] and are highlighted for their realism. They have supported epidemic planning since H1N1 [54]. Recent ABMs have incorporated increasingly realistic components, such as economic structures [77], digital tracing [41], and psychological

modules [60]. However, such realism incurs substantial computational cost, as individuals and their interactions must be modelled explicitly. Large-scale models like CityCOVID [67] thus require supercomputing resources. One remedy is static scaling, which aggregates multiple individuals into representative agents [85] but sacrifices output granularity by inflating rare events as scaling factors increase. Covasim [48] alleviates this through dynamic scaling based on output magnitude using predefined thresholds, but this demands careful tuning to prevent discontinuities between adjacent data points. Computational burden is further amplified during calibration, with the result that formal calibration is frequently omitted in many models [77]. Some models accelerate calibration with meta-heuristic algorithms [21, 56], but yield limited improvement.

To improve efficiency, surrogates have been introduced. Compartmental and Bayesian emulators [29, 83] accelerate simulation, but often oversimplify the dynamics of ABMs. ANN-based surrogates [6, 7, 25, 57] better approximate dynamics and enable efficient gradient-based calibration, but at the cost of mechanistic interpretability.

2.3 Differentiable agent-based models

DABMs achieve end-to-end differentiability through smooth approximations and tensorisations, enhancing computational efficiency while preserving mechanistic interpretability, as it avoids opaque ANN modules. The minor fidelity loss from smooth approximations is typically negligible. However, most DABMs are data-centric models that employ inferential ANNs to utilise auxiliary data, reintroducing black-box structures. One of the first epidemiological applications, DeepABM [19], models agent interactions with a graph neural network (GNN) and achieves a 200-fold acceleration over equivalent Mesa implementations, but lacks calibration. A subsequent model, JUNE-NZ [87], integrates GNNs for contact structures with LSTM networks for temporal dynamics, fitting the 2019 New Zealand measles outbreak in Auckland well but showing limited generalisability. Besides, they do not model agent autonomy or biological processes. The most outstanding model so far is GradABM [20] and its variants, which captures at least as many details as expert-designed ABMs [1, 73] and outperforms them on COVID-19 mortality and influenza-like-illness data across ten Massachusetts counties.

3 PROBLEM DESCRIPTION

This paper aims to contribute to *pandemic mitigation* by addressing the inadequacies in information-providing epidemic models that support the design and implementation of intervention policies. Successes in mitigating the COVID-19 pandemic through intervention policies before biological and medical research released sufficient results had averted hundreds of thousands of cases in regions like South Korea [55] and China [82]. Summarising from the literature, an adequate epidemic model should have:

- **Simulation fidelity:** The model represents the real-world epidemic dynamics *faithfully* to provide information useful for understanding and control [40].
- **Computational efficiency:** The model executes *expeditiously* to cope with the time constraints of a public health emergency [10].
- **Mechanistic interpretability:** The model operates *transparently* to ensure the provided information is trustworthy to decision-makers [71].

The challenge that remains open lies in designing a model *without significant shortcomings* with respect to the three key criteria, because the properties that enhance one tend to degrade another: fidelity demands *detail*, efficiency demands *abstraction*, and interpretability *restricts the use of black-box structures* that can dissolve the tension between the other two.

Traditional models [22, 47] and *simple ABMs* [41, 48] are *computationally efficient*, but their *poor simulation fidelity* limits the reliability of the information they provide [79, 81]; *complex ABMs* [67, 77] give *faithful simulations*, but their *poor computational efficiency* can make timely simulation and calibration costly, and sometimes impossible [9, 43]; *ABM with NN surrogates* [7, 57] are good in both criteria, but they *compromise mechanistic interpretability*. The *DABM* paradigm holds significant promise for addressing this challenge, as it does not possess non-negligible shortcomings on a paradigmatic level. However, existing *data-centric DABMs* [19, 20, 87] remain limited in *simulation fidelity* and *mechanistic interpretability* due to *specific design choices*.

Therefore, we aim to fully leverage the paradigm's potential by following a novel *structure-centric DABM-design principle*. We first construct a *fine-grained agent-based model* that incorporates extensive details without regard for computational cost, thus ensuring *simulation fidelity*. We then convert it into a *DABM* through differentiable approximations to accelerate computation and parameter optimisation. Finally, we propose a *z-score-based scaling method* that enables simulation under large data–model scale differences (up to 2000 times) while preserving output granularity, thus ensuring *computational efficiency*. Unlike data-centric DABMs, we *reject black-box components*, thus ensuring *mechanistic interpretability*.

4 MODEL STRUCTURE

In this section, we demonstrate how we operationalise *simulation fidelity* and *mechanistic interpretability* as the core of our novel structure-centric design. The model is constructed from fine-grained, fully transparent mechanisms that explicitly mirror real-world epidemiological and social processes. It contains three components: **Society**, which includes all non-epidemiological context configurations; **Epidemic**, which includes all epidemiological configurations; and **Population**, the configurations of the agents.

4.1 Society

The Society component defines the model's foundational operational context through three modules: the **temporal horizon**, **spatial context**, and **a simple economic system**.

The temporal horizon is discretised into daily time steps, matching the granularity of most epidemiological surveillance data. Each day is labelled as a weekday or weekend and grouped into months using modulo operations. For simplicity, each month is approximated as 30 days, which avoids unnecessary structural complexity while preserving temporal periodicity.

The spatial context, following existing ABMs, is a map composed of three facility types: offices [41, 48], markets [64, 76], and hospitals [41, 70]. These facilities capture the main venues of employment, consumption, and healthcare, supporting three interaction types: within-office, worker–customer, and doctor–patient. The map contains multiple clusters, each comprising one of each facility type, and the number of clusters is configurable to simulate polycentric urban planning [26]. Agents must select a location to visit, and encounters occur probabilistically among co-located agents. Staying at home is modelled as remaining outside the map and achieving full isolation [84].

The economic system builds on the above mechanisms. Agents balance income generation and expenditure: each receives a monthly salary and must pay recurring bills at the end of each month. Absenteeism accumulates when agents skip work on weekdays, and exceeding an employer-defined threshold results in a temporary salary reduction. Each day, agents consume one unit of essential supplies and can only purchase replacements in markets. These mechanisms jointly create financial incentives that may conflict with health-preserving behaviours, compelling agents to trade off infection avoidance against income maintenance, reproducing empirically observed presenteeism due to financial pressures [31, 32].

4.2 Epidemic

The Epidemic component captures disease dynamics across three levels: **between-host transmission, within-host progression**, and **pathogen mutation.** To represent these processes, we transform the classical *SEIR* framework into an *SAID* formulation that accounts for partial immunity and reinfection. Specifically, agents who recover from infection re-enter the *S* (Susceptible) class. An additional Asymptomatic Infected (*A*) class is introduced to represent infectious carriers without symptoms, supported by empirical evidence of widespread asymptomatic transmission [52, 66]. The Exposed (*E*) class is omitted as exposure can be inferred directly from expected encounters among agents. The Infected (*I*) class is redefined to represent symptomatic individuals. A Deceased (*D*) class captures agents whose symptom severity surpasses a lethal threshold. The traditional Recovered (*R*) class is removed as permanent immunity is replaced by cumulative partial immunity. The transition is shown in **Figure 1**.

The *between-host transmission* mechanism governs how infections propagate through encounters between infectious and susceptible agents. Transmission occurs probabilistically upon contact,
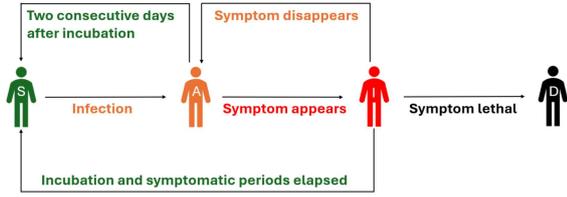
Figure 1: Health class transition

consistent with the biological process described by the dose–response model [14, 37]. Building on subsequent analysis of infection initiation dynamics [58, 80], we model each encounter as a single *infection-initiation* event, where infection occurs when cumulative infection initiations surpass the host's immune tolerance. Accordingly, partial immunity is modelled by increasing the number of transmissions required for infection establishment.

The *within-host progression* mechanism quantifies the temporal dynamics of symptom development for each infected agent. Following infection, an agent undergoes an incubation period (0 severity), followed by a symptomatic period. Agents naturally recover when the total incubation and symptomatic durations elapse. The incubation and symptomatic durations, denoted t_{inc}^j and t_{sym}^j for agent j at time step t, are independently resampled for each infection episode based on empirical results reported in [59, 62]:

$$t_{inc}^j \sim \log \mathcal{N}(1.63, 0.5), t_{sym}^j \sim \frac{age^j}{43.11 \cdot immune_{t-1}^j} \cdot \mathcal{N}(7.86, 6.46).$$

Agents remain in class A during incubation and move to I once symptoms appear. Symptom severity follows a Normal distribution, a standard assumption for ordinal health outcomes [15, 42], with age-dependent mean and variance reflecting greater severity and variability in older individuals [27]. If severity falls below zero during the symptomatic phase, the agent reverts to A, representing asymptomatic remission [53]. Recovery occurs early after two consecutive non-positive days [50]. Agents whose symptom severity exceeds 6 are classified as D.

The *pathogen mutation* mechanism introduces temporal variation in transmission and progression. At each time step, the pathogen mutates with a fixed probability. When a mutation occurs, agents with partial immunity lose one level of protection, allowing reinfection consistent with observed immune escape [18, 65], and symptom parameters (μ, σ) are rescaled to 0.5–1.5 times their prior values to capture variant-induced heterogeneity in symptom trajectories.

4.3 Population

The Population characterises agents along three aspects: *basic configurations, observation interfaces,* and *behavioural rules.* Shared structural forms define the population collectively, while parameter randomisation within these forms yields individual heterogeneity.

The *basic configurations* define each agent's static attributes, including demographic and personal characteristics. Demographics cover age, workplace address, preferred market and hospital, income (with salary-reduction thresholds), and expenditure levels

(modelled as fixed monthly bills). Personal characteristics include irrationality factors represented by randomised decision rules [39] (**A.4**, the appendix is available in the full version[*]) and judgment thresholds [36] for interpreting observed situations, following common practices.

The *observation interfaces* include an internal and an external interface. The internal observation interface retrieves the current savings and number of days of absenteeism, supplies (in daily quotas), and symptom severity, which inform judgment on financial, supply, and health states, respectively. The external observation interface observes the number of agents with each health class and aggregates them into the proportion of asymptomatic infections in the infected population, the proportion of cumulative infections in the whole population, and the proportion of cumulative infections resulting in death, which constitutes judgments on epidemic severity.

The *behavioural rule* maps observations to location choices. Following the fast-and-frugal heuristics framework [33, 34], it consists of if–then rules that compare observations against judgment thresholds (basic configurations) and output a distribution over four decisions: *stay at home, go to work, go shopping,* or *go to the hospital.* Observations are first translated into drives (intensity levels 1, 2, or 3) to each decision, with the following decision-drive correspondences:

- **Stay at home:** Epidemic severity and health state (mild).
- **Go to work:** Financial state.
- **Go shopping:** Supply state.
- **Go to the hospital:** Health state (severe).

The emergency level is determined by the extent to which the drive to one decision exceeds the others in intensity, while the decisions sorted with respect to their drives define a plausibility ranking that forms a priority heuristic, a widely used behavioural model[11, 13]. Ties are resolved by prioritising shopping, hospital visits, staying home, and working, in that order, reflecting survival priorities across supplies, health, infection avoidance, and income, aligning with observed behaviours such as panic buying and delayed care-seeking during epidemics [17, 63].

Subsequently, a distribution over the four decisions is determined using the initialised randomisation rules representing irrationality, which define three sets of probability mass functions with different skewness. Higher emergency levels correspond to more skewed distributions, simulating bounded rationality and decision salience [86], and decisions with higher plausibility ranks will be assigned with larger probability masses. The decision will be sampled from the determined categorical distribution, and a location will be selected correspondingly.

4.4 Model Workflow

Integrating the components described above, the complete ABM workflow is illustrated in **Figure 2**. At each time step, every living agent observes epidemic severity via the external observation interface and retrieves its financial, supply, and health states via the internal interface. These inputs are processed by the behavioural rules to select a location on the map(or stay at home). Agents who go out may encounter others making the same choice. When a susceptible

---
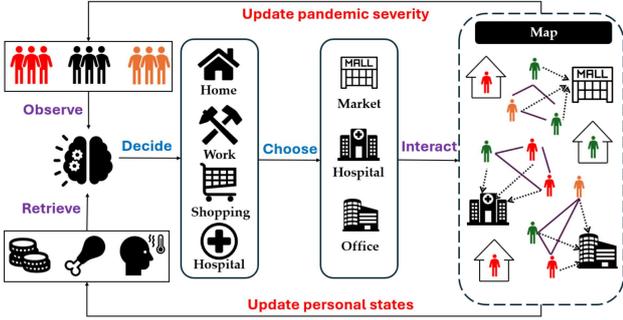[*]Full version available at https://arxiv.org/pdf/2602.12102

Figure 2: Model Workflow

agent meets an infectious one, transmission occurs probabilistically, and infection is established once cumulative transmissions exceed the agent's immune tolerance. Transmission outcomes and decisions will then jointly update each agent's personal states. Then the ABM iterates this process over successive time steps until the simulation reaches the predefined temporal horizon.

5 PARAMETER CALIBRATION METHOD

In this section, we present our parameter calibration method that aims to align the fine-grained ABM with real-world epidemic data while maintaining computational efficiency. Existing calibration methods for agent-based epidemic models often simplify the model structure, aggregate agent behaviours, or employ sampling approximations to reduce computational cost. Although effective for scaling, these approaches weaken the model's mechanistic fidelity and make parameter calibration against empirical data unreliable. To address these limitations, we propose to *combine end-to-end differentiability with a z-score-based scaling method*, forming a unified optimisation framework that preserves micro-level interpretability while enabling efficient data-driven calibration. Specifically, the differentiable approximation reformulates the ABM into a continuous computational graph that supports gradient-based parameter learning directly from empirical epidemic data. At the same time, the z-score-based scaling normalises and rescales outputs across populations to stabilise optimisation and generalise learned dynamics to large-scale epidemics. Together, these techniques establish a hybrid framework that seamlessly integrates fine-grained mechanistic simulation with data-driven calibration, ensuring that the ABM remains both computationally tractable and empirically faithful.

5.1 Differentiable Approximation

Most mechanisms in the ABM are inherently non-differentiable, such as the discrete decision rules and categorical state transitions that define agent behaviours. While some components, such as if–then decision rules, can only be relaxed, others, such as categorical variables (e.g., health classes), can be tensorised into differentiable representations. To construct *an end-to-end differentiable structure that supports gradient-based calibration*, we approximate these discrete processes with differentiable surrogate functions. This reformulation transforms the original rule-based ABM into

a differentiable computational graph, enabling GPU-accelerated parallelisation and allowing parameters to be calibrated directly against empirical epidemic data.

5.1.1 Relaxation. Three types of non-differentiable mechanisms require relaxation: **if–then rules, modulo functions, and random variables.** To this end, we introduce a novel *selective relaxation method* to relax if-then rules, design a *differentiable periodic indicator following the idea of sinusoidal mappings of time* [45] to relax the modulo function, and use the *reparametrisation trick* [44, 51] to relax random variables.

For if-then rules (of the generic form "if $a > A$ then x," equivalent to the Heaviside step function $x\mathrm{I}(a > A)$), relaxation inevitably trades off condition precision against optimisation stability, as closer approximations to the step function yield higher precision but increase the risk of vanishing gradients during optimisation. To balance this trade-off, we exploit the fact that different mechanisms demand different levels of precision. Our selective relaxation method applies distinct continuous relaxations under three categories: *precision-critical, moderately precise, and fuzzy conditions.* For precision-critical cases (e.g., health state transitions), we use:

$$h(a) = x \cdot \frac{\mathrm{ReLU}(a - A)}{\mathrm{ReLU}(a - A) + \xi} \begin{cases} = 0 \text{ if } a \leq A \\ \approx x \text{ otherwise} \end{cases}$$

where ξ is a small slack variable. This closely approximates the Heaviside step function, but yields an extremely small gradient $x \frac{\xi}{(\mathrm{I}(a>A)+\xi)^2} \mathrm{I}(a > A)$. Therefore, for moderate-precision scenarios (e.g., symptom progression), we use a smoother variant:

$$h(a) = x \cdot \mathrm{ReLU}(\tanh(a - A)) = \begin{cases} 0 & \text{if } a \leq A \\ x \cdot \tanh(a - A) & \text{otherwise} \end{cases}$$

which still outputs zero when the condition is unmet, but increases smoothly when satisfied. Its gradient, $x\mathrm{I}(a > A)\mathrm{sech}^2(a - A) \in [-|x|, |x|]]$, help maintain better stability during gradient descent. For fuzzy scenarios such as behavioural decisions, we adopt the logistic approximation, following the literature [4]:

$$h(a) = x \cdot \sigma(k \cdot (a - A))$$

where k is the steepness factor. It provides a continuous transition between 0 and 1, reflecting boundedly rational decision processes.

Modulo functions are used to make calendar-like time step classification into months and weekdays or weekends. Considering only integer t, we design the following differentiable periodic indicator as a relaxation:

$$\rho(t; m, n) = \mathrm{I}(\cos^2(\frac{(t - m)\pi}{n}) > \cos(\frac{\pi}{n})) = \begin{cases} 1 \text{ if } t \equiv m \pmod{n} \\ 0 \text{ otherwise} \end{cases}$$

where the Heaviside step function is approximated using the precision-oriented relaxation from **(2)**. Using this function, weekends can be labelled by $\rho(t; 6, 7) + \rho(t; 0, 7)$ and month ends by $\rho(t; 0, 30)$.

Relaxation of random variables using the reparameterisation trick has been well studied and widely used in the field of Variational Auto-Encoders (VAEs) [44, 51, 72]. Since well-established reparameterisation techniques already exist for the random variables used in our model, i.e. the Normal and Categorical random variables (with the Bernoulli random variable treated as a special case of the latter), we directly adopt these standard formulations.

5.1.2 Tensorisation. The Tensorisation mainly follows established DABM practices (details in **A.2–A.4**). We elaborate on specific adjustments for three mechanisms: categorical variables, agents' encounters, and the transmission mechanism. Categorical variables are represented by one-hot or multi-hot tensors. Agents' encounters are tensorised as an adjacency representation, following the same idea as [20] but operationalised differently due to the difference in the model's spatial context. Transmission is then computed using the tensorised health classes and the encounter matrix.

Four types of categorical variables are tensorised: health classes, decisions, location choices, and location preferences. Health classes and decisions are represented by one-hot vectors $h_t^j \in \{0,1\}^4$ and $d_t^j \in \{0,1\}^3$ (staying at home excludes further interactions), respectively, and location choices are represented by one-hot matrices $l_t^j \in \{0,1\}^{n \times 3}$ with n being the number of clusters of locations. Similarly, location preferences are represented by multi-hot tensors $L^j \in \{0,1\}^{n \times 3 \times 3}$ with three entries of 1, because there must be a preference (and for work, workplace) for each of the three decisions, excluding staying at home.

Encounters are derived from location choice matrices l_t^j. Let P be the population size and m the encounter probability. The encounter matrix E_t is defined as:

$$l_{t_1} = \sum_{j=1}^{P} l_t^j, \quad l_{t_2} = \frac{1}{\sqrt{2}}(\|l_t^i - l_t^j\|_{\mathcal{F}})_{i=1,j=1}^{P,P}$$

$$E_t = (l_{t_1} l_{t_1}^T \odot (1_P - l_{t_2})) \odot (1_P - I_P) \cdot m$$

where l_{t_1} indicates agents who leave home, and l_{t_2} specifies whether agent pairs choose the same or different locations, or if one remains at home. Entry-wise multiplication of $l_{t_1} l_{t_1}^T$ and $(1 - l_{t_2})$ identifies co-located pairs, while $(1_P - I_P)$ removes self-encounters. Multiplying by m yields pairwise encounter probabilities.

Transmissions are then computed. Let β denote the transmission probability and $h_t = (h_t^j)_{j=1}^{P} \in \{0,1\}^{4 \times P}$ the stacked health labels. The expected number of exposure events per healthy agent is:

$$E_t^I = (E_t \odot h_{t-1 1}) \cdot (h_{t-1 2} + h_{t-1 3}) \cdot \beta$$

Here, $E_t \odot h_{t-1 1}$ selects encounters involving susceptible agents, while $h_{t-1 2} + h_{t-1 3}$ represents all asymptomatic and symptomatic infected agents. Their product gives expected close contacts, and multiplying by β yields expected exposure events.

5.2 Output Scaling

While differentiable approximation significantly accelerates computations by enabling gradient-based parallelisation, it contributes little to reducing memory consumption. To further enhance computational efficiency, particularly with respect to memory, an additional mechanism is required. Output scaling addresses this by allowing a small simulated population to represent a much larger real-world one. However, existing methods are limited: static scaling, which scales between model and real populations, is too rigid and reduces output granularity when population differences are large, while dynamic scaling, which scales outputs solely by their magnitudes, demands careful hyperparameter tuning.

To overcome these limitations, we propose a novel z-score-based scaling method, a forecasting-oriented scaling method that directly scales between the model outputs and the empirical data, rather than between populations or with respect to output magnitudes. Forecasting-orientation is justified because forecasting, which relies on calibrating model outputs to real data, is the task most sensitive to scalability. Compared with static scaling, our method allows a larger difference between model and real population scales, since key epidemic indices (e.g., daily infections and mortalities) constitute only small fractions of the total population, which means that the ratio between a small-scale model output and empirical data is much smaller than the ratio between their populations. Compared with dynamic scaling, our approach requires no hyperparameter tuning, as demonstrated below.

Consider the model as a function $f(w; x)$, where w represents the non-learnable parameters and x represents the learnable parameters. Let the output time series be $\tilde{y} = f(w; x)$. Our scaling method will first compute the z-score $\tilde{y}_z = \frac{\tilde{y} - \tilde{y}_\mu}{\tilde{y}_\sigma}$ where $\tilde{y}_\mu$ and $\tilde{y} - \sigma$ are the mean and standard deviation of $\tilde{y}$. After that, $\tilde{y}_z$ will be rescaled through an inverse computation of z-score via y_μ and y_σ, the mean and standard deviation of the real data for calibration, and translated so that the smallest entries of the transformed output and the original output can align, i.e.

$$\tilde{y}' = \tilde{y}_z y_\sigma + y_\mu, \quad \hat{y} = \tilde{y}' - \tilde{y}'_{min} 1 + y_{min} 1$$

The translation ensures non-negative values and aligns the minimum of the transformed output with that of the real data.

This method directly matches the first two statistical moments (scale and location) of the simulated and real series, guaranteeing a consistent and reliable projection of the model output without distorting its trend, yet the transformation parameters are automatically derived from statistical properties, thus eliminating the need for manual hyperparameter tuning.

6 EXPERIMENT

In this section, we demonstrate approaches we take to further justify our model's simulation fidelity and computational efficiency. For simulation fidelity, we conducted sensitivity analysis to validate the realism of the model's key mechanisms, and performance evaluation using GradABM [20] as a baseline to demonstrate the model's forecasting capability, generalisability, and data efficiency. For computational efficiency, we demonstrate how the model's runtime varies with population size and forecast horizon.

6.1 Sensitivity Analysis

6.1.1 One-at-a-time Analysis. This analysis shows how key simulations vary when parameters involved in the related mechanisms are changed. The evaluated simulations include the cumulative infections, cumulative deaths, and the number of infected agents in critical conditions (defined as agents with symptom severities over 4). For cumulative infections, we examine the influences of transmission probability, mean monthly bill and maximum amount of daily purchasable supplies, which are key factors of the pathogen's infectiousness, financial pressure, and supply stability, respectively. For the cumulative deceased and the number of agents in critical condition, we examined the influences of the mean and standard deviation of daily severity changes. As illustrated in **Figures 3, 4, 5**, key mechanisms of the model operate realistically: the infection
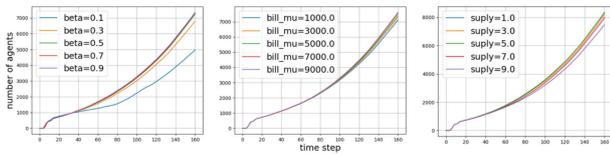
Figure 3: Cumulative infections under different transmission probabilities (left), mean monthly bill (middle), and maximum purchasable daily quotas of supplies (right)
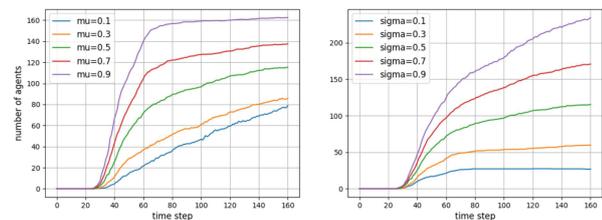
Figure 4: Cumulative mortality under different means (left) and standard deviations (right) of symptom progression
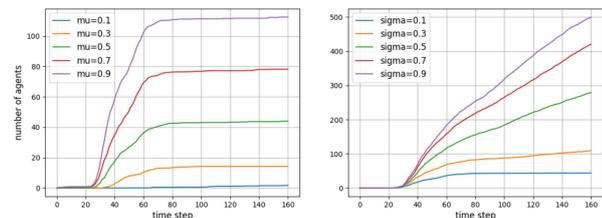
Figure 5: Cumulative agents in critical conditions under different means (left) and standard deviations (right) of symptom progression

growth rates increase as the pathogen's infectiousness and financial pressure increase, and as the supply stability decreases.

6.1.2 Sobol Analysis. To explore the influences of the mutual effects of the parameters, we employed the Sobol variance decomposition method [75], using the total-effect Sobol indices with 95% CI, and Saltelli sampling size $N(2d + 2)$, $N = 1024$.

The sensitivity analysis demonstrated a clear and expected hierarchy in parameter influences on epidemic dynamics, even with mutual effects considered. For simulated infections, transmission and encounter probabilities dominate with Sobol indices at least 45% higher than other groups, consistent with their fundamental role of impact in disease spread. Health progression parameters and behavioural parameters form a secondary tier, reflecting the model's dual biological-behavioural architecture. For simulated deceased cases, parameters associated with individual health emerge as the dominant factors, demonstrating that the within-host progression

mechanisms function as intended, given that agent mortality depends solely on these health dynamics.

We also compared the influence of the parameters on the number of agents making each decision. The decision-making parameters prove most influential across all outputs except for the number of working agents. In this latter case, transmission parameters retain dominance, albeit narrowly surpassing decision-making factors. The persistent impact of transmission and health-related parameters aligns with expectations for epidemic-driven decision-making. An unexpected finding concerns the differential impact of health-related parameters on decision-making: their effect on shopping and staying home appears stronger than on hospital visits. A preliminary hypothesis attributes this to overly restrictive parameter ranges to determine mean symptom severity change, resulting in too few agents developing symptoms severe enough to warrant hospitalisation. To investigate this further, the range was expanded in subsequent experiments. After allowing parameters that represent the spread of more fatal diseases to be sampled, the influences of health-related parameters become reasonable: the health-related factors are most influential to the number of agents going to the hospital, and second most influential to the number of agents staying at home.

6.2 Performance Evaluation

6.2.1 Baseline. We use the three versions of GradABM [20] as baselines. GradABM is a differentiable agent-based model designed to efficiently simulate large-scale epidemic dynamics and is among the most widely cited open-source DABMs with well-documented calibration and forecasting results. It was validated using COVID-19 mortality data from 10 counties in Massachusetts, with population scales from 70000 to 1600000, and influenza-like illness (ILI) data for the state. The three versions are: C-GradABM, which doesn't use any auxiliary data; DC-GradABM, which uses a deep neural network to infer simulation parameters from regional heterogeneous features such as mobility and census data; and JDC-GradABM, which generalises this approach by leveraging heterogeneous data from all regions jointly. They all use 7 parameters for the COVID-19 simulation and 8 parameters for the influenza simulation. In order to make a reliable comparison, we adopted the same data, data processing method, evaluation lengths, and metrics. More details are explained in **B.2**.

6.2.2 Results. **Table 2** demonstrates the results obtained by our model and the comparisons to the results of the baselines. For COVID-19 mortality, the mean and standard deviation of forecasting errors over 5 runs given by DEpiABS are both smaller than JDC-GradABM, the most powerful version of GradABM, in all metrics. For influenza, DEpiABS gives a smaller mean normal deviation than JDC-GradABM, and mean RMSE and MAE similar to DC-GradABM. The standard deviations of all norms are between JDC-GradABM and DC-GradABM. These results demonstrate our model's good forecasting ability and generalisability. We also illustrate the comparisons between the model's forecast and the real data of Norfolk County in **Figure 6**. Illustrations of results on other regions can be found in **B.2**.

This yields three key insights. First, our model is data-efficient, as it achieves performance comparable to or exceeding JDC-GradABM,

Model	COVID-19			Flu		
	ND	RMSE	MAE	ND	RMSE	MAE
C-GradABM	2.39 ± 0.35	205.24 ± 42.56	73.66 ± 10.88	0.88 ± 0.14	2.97 ± 0.44	2.64 ± 0.43
DC-GradABM	1.15 ± 0.24	67.09 ± 23.89	35.50 ± 7.37	0.50 ± 0.19	1.78 ± 0.62	1.50 ± 0.57
JDC-GradABM	0.97 ± 0.18	50.99 ± 12.12	30.02 ± 5.60	0.41 ± 0.02	1.47 ± 0.06	1.22 ± 0.06
DEpiABS	$\mathbf{0.92 \pm 0.05}$	$\mathbf{18.95 \pm 4.39}$	$\mathbf{12.97 \pm 2.25}$	$\mathbf{0.32 \pm 0.05}$	1.80 ± 0.28	1.47 ± 0.23

Table 2: Comparison between the performances of GradABM and DEpiABS. Given real and simulated time series y and $\hat{y}$ of length T, the metrics used are defined as follows: ND= $\frac{\Sigma_{t=1}^{T}|y_t - \hat{y}_t|}{\Sigma_{t=1}^{T}|y_t|}$, RMSE= $\frac{1}{T}\sqrt{\Sigma_{t=1}^{T}|y_t^2 - \hat{y}_t^2|}$, MAE= $\frac{1}{T}\Sigma_{t=1}^{T}|y_t - \hat{y}_t|$
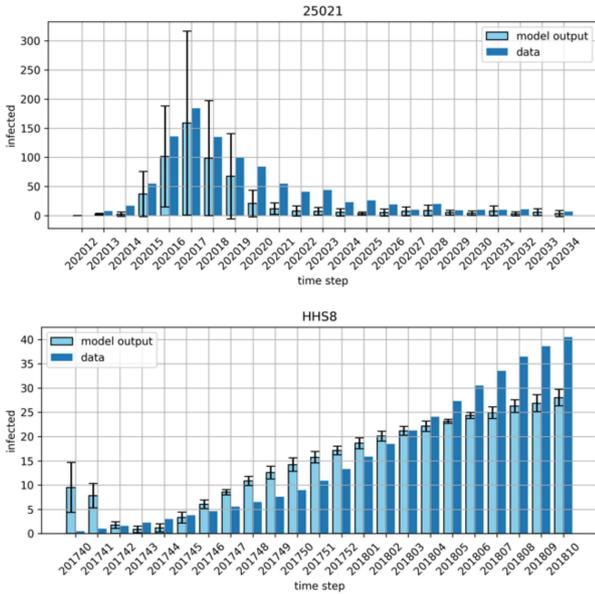

Figure 6: Forecasted and real COVID mortality (top) and forecasted and real ILI (bottom) in Norfolk County, Massachusetts
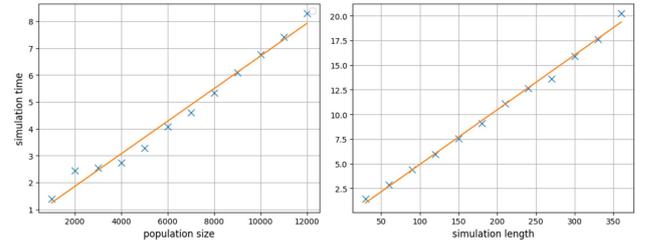

Figure 7: Runtime vs. number of simulated agents. A straight line of $y = 0.3078x - 0.4338$ can be fitted with a coefficient of determination $R = 0.89$

the run time with a Mesa equivalence, which in our case is the unoptimised model. Results show that runtime grows linearly with respect to the number of simulated agents (**Figure 7**), aligning with the models mentioned above, and also with respect to the simulation length. The model is 200-250 times faster than the Mesa equivalence, reaching a simulation time of 1.40 seconds for 1000 simulated agents and 30 time steps, comparable to [19] and [20].

7 CONCLUSION

We present DEpiABS, a scalable and interpretable structure-centric differentiable agent-based model (DABM). By enhancing feature granularity and rejecting the ANN component, DEpiABS fully leverages the potential of the DABM paradigm. It surpasses the state-of-the-art data-centric DABM in simulation fidelity and outperforms it with higher data efficiency while maintaining full mechanistic interpretability and comparable computational efficiency. We also propose and validate a z-score-based scaling method that decouples simulation cost from population size, enabling ABM forecasting for large-scale populations efficiently. As one of the first structure-centric DABMs, it demonstrates the competitiveness of this design philosophy over data-centric alternatives.

the strongest variant using aggregated auxiliary data from all counties, while using only the same amount of data as C-GradABM, the weakest variant without auxiliary data. Second, our z-score-based scaling method proves reliable, as the reported results are obtained from a 500-agent simulation employing this method. Third, our structure-centric DABM-design principle is competitive, since our model, one of the first of its kind, outperforms the state-of-the-art data-centric DABM.

6.3 Scalability Test

With our z-score-based method, the population size required to maintain reasonable output granularity is low, but a larger population size is still beneficial for making more accurate forecasts. Thus, the scalability test is important. Following the evaluation method incorporated by many existing works [19, 20, 48], we conducted the scalability test in two ways: finding how the model's runtime varies with population size and forecast horizon, and comparing

REFERENCES

[1] Matthew Abueg, Robert Hinch, Neo Wu, Luyang Liu, William Probert, Austin Wu, Paul Eastham, Yusef Shafi, Matt Rosencrantz, Michael Dikovsky, et al. 2021. Modeling the effect of exposure notification and non-pharmaceutical interventions on COVID-19 transmission in Washington state. *NPJ digital medicine* 4, 1 (2021), 49.

[2] Marco Ajelli, Bruno Gonçalves, Duygu Balcan, Vittoria Colizza, Hao Hu, José J Ramasco, Stefano Merler, and Alessandro Vespignani. 2010. Comparing large-scale computational approaches to epidemic modeling: agent-based versus structured metapopulation models. *BMC infectious diseases* 10 (2010), 1–13.

[3] G Gómez Alcaraz and C Vargas-De-León. 2012. Modeling control strategies for influenza A H1N1 epidemics: SIR models. *Revista Mexicana de Física* 58, 1 (2012), 37–43.

[4] Philipp Andelfinger. 2021. Differentiable agent-based simulation for gradient-guided simulation-based optimization. In *Proceedings of the 2021 ACM SIGSIM Conference on Principles of Advanced Discrete Simulation*. 27–38.

[5] Roy M Anderson and Robert M May. 1991. *Infectious diseases of humans: dynamics and control*. Oxford university press.

[6] Claudio Angione, Eric Silverman, and Elisabeth Yaneske. 2022. Using machine learning as a surrogate model for agent-based simulations. *Plos one* 17, 2 (2022), e0263150.

[7] Rushil Anirudh, Jayaraman J Thiagarajan, Peer-Timo Bremer, Timothy Germann, Sara Del Valle, and Frederick Streitz. 2022. Accurate calibration of agent-based epidemiological models with neural network surrogates. In *Workshop on Healthcare AI and COVID-19*. PMLR, 54–62.

[8] Diepreye Ayabina, Hasan Sevil, Adam Kleczkowski, and M Gabriela M Gomes. 2025. A note on metapopulation models. *arXiv preprint arXiv:2506.04284* (2025).

[9] Luděk Berec, Tomáš Diviák, Aleš Kuběna, René Levínský, Roman Neruda, Gabriela Suchopárová, Josef Šlerka, Martin Šmíd, Jan Trnka, Vít Tuček, et al. 2023. On the contact tracing for COVID-19: A simulation study. *Epidemics* 43 (2023), 100677.

[10] Anna Bershteyn, Hae-Young Kim, and R Scott Braithwaite. 2022. Real-time infectious disease modeling to inform emergency public health decision making. *Annual Review of Public Health* 43, 1 (2022), 397–418.

[11] Michael H Birnbaum. 2008. Evaluation of the priority heuristic as a descriptive model of risky decision making: Comment on Brandstätter, Gigerenzer, and Hertwig (2006). (2008).

[12] Yacong Bo, Cui Guo, Changqing Lin, Yiqian Zeng, Hao Bi Li, Yumiao Zhang, Md Shakhaoat Hossain, Jimmy WM Chan, David W Yeung, Kin On Kwok, et al. 2021. Effectiveness of non-pharmaceutical interventions on COVID-19 transmission in 190 countries from 23 January to 13 April 2020. *International Journal of Infectious Diseases* 102 (2021), 247–253.

[13] Eduard Brandstätter, Gerd Gigerenzer, and Ralph Hertwig. 2006. The priority heuristic: making choices without trade-offs. *Psychological review* 113, 2 (2006), 409.

[14] Andrew F Brouwer, Mark H Weir, Marisa C Eisenberg, Rafael Meza, and Joseph NS Eisenberg. 2017. Dose-response relationships for environmentally mediated infectious disease transmission models. *PLoS computational biology* 13, 4 (2017), e1005481.

[15] Mark Cauchi, Andrew R Mills, Allan Lawrie, David G Kiely, and Visakan Kadirkamanathan. 2024. Individualized survival predictions using state space model with longitudinal and survival data. *Journal of the Royal Society Interface* 21, 216 (2024), 20230682.

[16] Wenzhi Chen. 2015. A mathematical model of Ebola virus based on SIR model. In *2015 International Conference on Industrial Informatics-Computing Technology, Intelligent Technology, Industrial Information Integration*. IEEE, 213–216.

[17] Yiqun Chen, Abbas Rajabifard, Soheil Sabri, Katie Elisabeth Potts, Farhad Laylavi, Yuke Xie, and Yibo Zhang. 2020. A discussion of irrational stockpiling behaviour during crisis. *Journal of Safety Science and Resilience* 1, 1 (2020), 57–58. https://doi.org/10.1016/j.jnlssr.2020.06.003

[18] Ying Chen, Wenhui Zhu, Xinyue Han, Miaoshuang Chen, Xin Li, Haiping Huang, Mengyuan Zhang, Rongjie Wei, Huadong Zhang, Changhong Yang, et al. 2024. How does the SARS-CoV-2 reinfection rate change over time? The global evidence from systematic review and meta-analysis. *BMC Infectious Diseases* 24, 1 (2024), 339.

[19] Ayush Chopra, Ramesh Raskar, Jayakumar Subramanian, Balaji Krishnamurthy, Esma S Gel, Santiago Romero-Brufau, Kalyan S Pasupathy, and Thomas C Kingsley. 2021. DeepABM: scalable and efficient agent-based simulations via geometric learning frameworks-a case study for COVID-19 spread and interventions. In *2021 Winter simulation conference (WSC)*. IEEE, 1–12.

[20] Ayush Chopra, Alexander Rodríguez, Jayakumar Subramanian, Arnau Quera-Bofarull, Balaji Krishnamurthy, B Aditya Prakash, and Ramesh Raskar. [n.d.]. Differentiable Agent-based Epidemiology. ([n. d.]).

[21] Robert Chase Cockrell and Gary An. 2018. Examining the controllability of sepsis using genetic algorithms on an agent-based model of systemic inflammation. *PLoS computational biology* 14, 2 (2018), e1005876.

[22] Pietro Coletti, Pieter Libin, Oana Petrof, Lander Willem, Steven Abrams, Sereina A Herzog, Christel Faes, Elise Kuylen, James Wambua, Philippe Beutels, et al. 2021. A data-driven metapopulation model for the Belgian COVID-19 epidemic: assessing the impact of lockdown and exit strategies. *BMC infectious diseases* 21 (2021), 1–12.

[23] Ian Cooper, Argha Mondal, and Chris G Antonopoulos. 2020. A SIR model assumption for the spread of COVID-19 in different communities. *Chaos, Solitons & Fractals* 139 (2020), 110057.

[24] Guilherme S Costa, Wesley Cota, and Silvio C Ferreira. 2020. Metapopulation modeling of COVID-19 advancing into the countryside: an analysis of mitigation strategies for Brazil. *MedRxiv* (2020), 2020–05.

[25] Francesco Cozzi, Marco Pangallo, Alan Perotti, André Panisson, and Corrado Monti. 2025. Learning Individual Behavior in Agent-Based Models with Graph Diffusion Networks. *arXiv preprint arXiv:2505.21426* (2025).

[26] Ben Derudder, Evert Meijers, John Harrison, Michael Hoyler, and Xingjian Liu. 2022. Polycentric urban regions: conceptualization, identification and implications. , 6 pages.

[27] Melina Farshbafnadi, Sara Kamali Zonouzi, Mohammadmahdi Sabahi, Mahsa Dolatshahi, and Mohammad Hadi Aarabi. 2021. Aging & COVID-19 susceptibility, disease severity, and clinical outcomes: The role of entangled risk factors. *Experimental gerontology* 154 (2021), 111507.

[28] Seth Flaxman, Swapnil Mishra, Axel Gandy, H Juliette T Unwin, Thomas A Mellan, Helen Coupland, Charles Whittaker, Harrison Zhu, Tresnia Berah, Jeffrey W Eaton, et al. 2020. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* 584 (2020), 257.

[29] Luis L Fonseca, Lucas Böttcher, Borna Mehrad, and Reinhard C Laubenbacher. 2025. Optimal control of agent-based models via surrogate modeling. *PLoS Computational Biology* 21, 1 (2025), e1012138.

[30] CA Francisco. 2020. Understanding the US failure on coronavirus—an essay by Drew Altman. *bmj* 370 (2020), m3417.

[31] Susan Gallagher, Asha Roy, Sandra J Domeracki, Todd Mohrmann, Vicki Missar, Janet Jule, Shreela Sharma, and Ryan DeWitt. 2021. The low-wage essential worker: Occupational concerns and needs in the COVID-19 pandemic—A round table. *Workplace Health & Safety* 69, 4 (2021), 154–160.

[32] Juan Jesús García-Iglesias, Juan Gómez-Salgado, Joao Apostolo, Rogério Rodrigues, Emília Isabel Costa, Carlos Ruiz-Frutos, Santiago Martínez-Isasi, Daniel Fernández-García, and Ángel Vilches-Arenas. 2023. Presenteeism and mental health of workers during the COVID-19 pandemic: a systematic review. *Frontiers in public health* 11 (2023), 1224332.

[33] Gerd Gigerenzer. 2004. Fast and frugal heuristics: The tools of bounded rationality. *Blackwell handbook of judgment and decision making* 62 (2004), 88.

[34] Gerd Gigerenzer and Wolfgang Gaissmaier. 2011. Heuristic decision making. *Annual review of psychology* 62, 2011 (2011), 451–482.

[35] Lara Goscé, Andrew Phillips, Paula Spinola, Rishi K Gupta, and Ibrahim Abubakar. 2020. Modelling SARS-COV2 spread in London: approaches to lift the lockdown. *Journal of Infection* 81, 2 (2020), 260–265.

[36] Mark Granovetter. 1978. Threshold models of collective behavior. *American journal of sociology* 83, 6 (1978), 1420–1443.

[37] Charles N Haas, Joan B Rose, and Charles P Gerba. 2014. *Quantitative microbial risk assessment*. John Wiley & Sons.

[38] Pedro C Hallal and Cesar G Victora. 2021. Overcoming Brazil's monumental COVID-19 failure: an urgent call to action. *Nature Medicine* 27, 6 (2021), 933–933.

[39] A Brett Hauber, Juan Marcos González, Catharina GM Groothuis-Oudshoorn, Thomas Prior, Deborah A Marshall, Charles Cunningham, Maarten J IJzerman, and John FP Bridges. 2016. Statistical methods for the analysis of discrete choice experiments: a report of the ISPOR conjoint analysis good research practices task force. *Value in health* 19, 4 (2016), 300–315.

[40] Hans Heesterbeek, Roy M Anderson, Viggo Andreasen, Shweta Bansal, Daniela De Angelis, Chris Dye, Ken TD Eames, W John Edmunds, Simon DW Frost, Sebastian Funk, et al. 2015. Modeling infectious disease dynamics in the complex landscape of global health. *Science* 347, 6227 (2015), aaa4339.

[41] Robert Hinch, William JM Probert, Anel Nurtay, Michelle Kendall, Chris Wymant, Matthew Hall, Katrina Lythgoe, Ana Bulas Cruz, Lele Zhao, Andrea Stewart, et al. 2021. OpenABM-Covid19—An agent-based model for non-pharmaceutical interventions against COVID-19 including contact tracing. *PLoS computational biology* 17, 7 (2021), e1009146.

[42] John L Hopper and Graeme P Young. 1988. A random walk model for evaluating clinical trials involving serial observations. *Statistics in medicine* 7, 5 (1988), 581–590.

[43] Elizabeth Hunter and John D Kelleher. 2022. Validating and testing an agent-based model for the spread of COVID-19 in Ireland. *Algorithms* 15, 8 (2022), 270.

[44] Eric Jang, Shixiang Gu, and Ben Poole. 2016. Categorical reparameterization with gumbel-softmax. *arXiv preprint arXiv:1611.01144* (2016).

[45] Seyed Mehran Kazemi, Rishab Goel, Sepehr Eghbali, Janahan Ramanan, Jaspreet Sahota, Sanjay Thakur, Stella Wu, Cathal Smyth, Pascal Poupart, and Marcus Brubaker. 2019. Time2vec: Learning a vector representation of time. *arXiv*

preprint arXiv:1907.05321 (2019).

[46] Jackie Kazil, David Masad, and Andrew Crooks. 2020. Utilizing Python for Agent-Based Modeling: The Mesa Framework. In *Social, Cultural, and Behavioral Modeling*, Robert Thomson, Halil Bisgin, Christopher Dancy, Ayaz Hyder, and Muhammad Hussain (Eds.). Springer International Publishing, Cham, 308–317.

[47] William Ogilvy Kermack and Anderson G McKendrick. 1927. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character* 115, 772 (1927), 700–721.

[48] Cliff C Kerr, Robyn M Stuart, Dina Mistry, Romesh G Abeysuriya, Katherine Rosenfeld, Gregory R Hart, Rafael C Núñez, Jamie A Cohen, Prashanth Selvaraj, Brittany Hagedorn, et al. 2021. Covasim: an agent-based model of COVID-19 dynamics and interventions. *PLOS Computational Biology* 17, 7 (2021), e1009149.

[49] Veria Khosrawipour, Hien Lau, Tanja Khosrawipour, Piotr Kocbach, Hirohito Ichii, Jacek Bania, and Agata Mikolajczyk. 2020. Failure in initial stage containment of global COVID-19 epicenters. *Journal of medical virology* 92, 7 (2020), 863–867.

[50] Min-Chul Kim, Chunguang Cui, Kyeong-Ryeol Shin, Joon-Yong Bae, Oh-Joo Kweon, Mi-Kyung Lee, Seong-Ho Choi, Sun-Young Jung, Man-Seong Park, and Jin-Won Chung. 2021. Duration of culturable SARS-CoV-2 in hospitalized patients with Covid-19. *New England Journal of Medicine* 384, 7 (2021), 671–673.

[51] Diederik P Kingma, Max Welling, et al. 2013. Auto-encoding variational bayes.

[52] Lauren M Kucirka, Stephen A Lauer, Oliver Laeyendecker, Denali Boon, and Justin Lessler. 2020. Variation in false-negative rate of reverse transcriptase polymerase chain reaction–based SARS-CoV-2 tests by time since exposure. *Annals of internal medicine* 173, 4 (2020), 262–267.

[53] Jeffrey C Kwong. 2016. Editorial commentary: symptoms and viral shedding in naturally acquired influenza infections. , 438–439 pages.

[54] Bruce Y Lee, Shawn T Brown, Philip Cooley, Maggie A Potter, William D Wheaton, Ronald E Voorhees, Samuel Stebbins, John J Grefenstette, Shanta M Zimmer, Richard K Zimmerman, et al. 2010. Simulating school closure strategies to mitigate an influenza epidemic. *Journal of public health management and practice* 16, 3 (2010), 252–261.

[55] David Lee and Jaehong Lee. 2020. Testing on the move: South Korea's rapid response to the COVID-19 pandemic. *Transportation Research Interdisciplinary Perspectives* 5 (2020), 100111.

[56] Ria Lestari, Adi Pancoro, Adang Suwandi Ahmad, Afwarman Manaf, Munawar Ahmad, Chairul A Nidom, and Yusuf Budiana. 2011. Pandemic Influenza simulation based on Genetic Algorithm and Agent Based Modeling. In *Proceedings of the 2011 International Conference on Electrical Engineering and Informatics*. IEEE, 1–6.

[57] Varun Madhavan, Adway Mitra, and Partha Pratim Chakrabarti. 2022. Deep learning-based spatially explicit emulation of an agent-based simulator for pandemic in a city. *arXiv preprint arXiv:2205.14396* (2022).

[58] Bryan T Mayer, James S Koopman, Edward L Ionides, Josep M Pujol, and Joseph NS Eisenberg. 2011. A dynamic dose–response model to account for exposure patterns in risk assessment: a case study in inhalation anthrax. *Journal of The Royal Society Interface* 8, 57 (2011), 506–517.

[59] Conor McAloon, Áine Collins, Kevin Hunt, Ann Barber, Andrew W Byrne, Francis Butler, Miriam Casey, John Griffin, Elizabeth Lane, David McEvoy, Patrick Wall, Martin Green, Luke O'Grady, and Simon J More. 2020. Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research. *BMJ Open* 10, 8 (2020). https://doi.org/10.1136/bmjopen-2020-039652 arXiv:https://bmjopen.bmj.com/content/10/8/e039652.full.pdf

[60] Konstantinos Mitsopoulos, Ritwik Bose, Brodie Mather, Archna Bhatia, Kevin Gluck, Bonnie Dorr, Christian Lebiere, and Peter Pirolli. 2023. Psychologically-valid generative agents: A novel approach to agent-based modeling in social sciences. In *Proceedings of the AAAI Symposium Series*, Vol. 2. 340–348.

[61] Zahra Mohammadi, Monica Gabriela Cojocaru, and Edward Wolfgang Thommes. 2022. Human behaviour, NPI and mobility reduction effects on COVID-19 transmission in different countries of the world. *BMC Public Health* 22, 1 (2022), 1594.

[62] Monira Mollazehi, Mohammad Mollazehi, and Abdel-Salam Abdel-Salam. 2020. Modeling Survival Time to Recovery from COVID-19: A Case Study on Singapore. *Research Square* (2020). https://doi.org/10.21203/rs.3.rs-18600/v2

[63] Mohadeseh Motamed-Jahromi, Zahra Meshkani, Seyed Masood Mosavi-Negad, Victoria Momenabadi, and Mahdieh Sadat Ahmadzadeh. 2021. Factors affecting panic buying during COVID-19 outbreak and strategies: A scoping review. *Iranian Journal of Public Health* 50, 12 (2021), 2473.

[64] Christian Nitzsche and Stefan Simm. 2024. Agent-based modeling to estimate the impact of lockdown scenarios and events on a pandemic exemplified on SARS-CoV-2. *Scientific Reports* 14, 1 (2024), 13391.

[65] Carolina KV Nonaka, Marília Miranda Franco, Tiago Gräf, Camila Araújo de Lorenzo Barcia, Renata Naves de Ávila Mendonça, Karoline Almeida Felix De Sousa, Leila MC Neiva, Vagner Fosenca, Ana VA Mendes, Renato Santana De Aguiar, et al. 2021. Genomic evidence of SARS-CoV-2 reinfection involving E484K spike mutation, Brazil. *Emerging infectious diseases* 27, 5 (2021), 1522.

[66] Daniel P Oran and Eric J Topol. 2021. The proportion of SARS-CoV-2 infections that are asymptomatic: a systematic review. *Annals of internal medicine* 174, 5 (2021), 655–662.

[67] Jonathan Ozik, Justin M Wozniak, Nicholson Collier, Charles M Macal, and Mickaël Binois. 2021. A population data-driven workflow for COVID-19 modeling and learning. *The International Journal of High Performance Computing Applications* 35, 5 (2021), 483–499.

[68] Sen Pei, Sasikiran Kandula, and Jeffrey Shaman. 2020. Differential effects of intervention timing on COVID-19 spread in the United States. *Science advances* 6, 49 (2020), eabd6370.

[69] Nicola Perra. 2021. Non-pharmaceutical interventions during the COVID-19 pandemic: A review. *Physics Reports* 913 (2021), 1–52.

[70] Thi Mui Pham, Hannan Tahir, Janneke HHM van de Wijgert, Bastiaan R Van der Roest, Pauline Ellerbroek, Marc JM Bonten, Martin CJ Bootsma, and Mirjam E Kretzschmar. 2021. Interventions to control nosocomial transmission of SARS-CoV-2: a modelling study. *BMC medicine* 19, 1 (2021), 211.

[71] Darya Pokutnaya, Bruce Childers, Alice E Arcury-Quandt, Harry Hochheiser, and Willem G Van Panhuis. 2023. An implementation framework to improve the transparency and reproducibility of computational models of infectious diseases. *PLOS Computational Biology* 19, 3 (2023), e1010856.

[72] Danilo Jimenez Rezende, Shakir Mohamed, and Daan Wierstra. 2014. Stochastic backpropagation and approximate inference in deep generative models. In *International conference on machine learning*. PMLR, 1278–1286.

[73] Santiago Romero-Brufau, Ayush Chopra, Alex J Ryu, Esma Gel, Ramesh Raskar, Walter Kremers, Karen S Anderson, Jayakumar Subramanian, Balaji Krishnamurthy, Abhishek Singh, et al. 2021. Public health impact of delaying second dose of BNT162b2 or mRNA-1273 covid-19 vaccine: simulation agent based modeling study. *bmj* 373 (2021).

[74] Henrik Salje, Cecile Tran Kiem, Noemie Lefrancq, Noemie Courtejoie, Paolo Bosetti, Juliette Paireau, Alessio Andronico, Nathanael Hoze, Jehanne Richet, Claire-Lise Dubost, et al. 2020. Estimating the burden of SARS-CoV-2 in France (vol 368, eabd4246, 2020). American Association for the Advancement of Science.

[75] Andrea Saltelli, Karen Chan, and E Marian Scott. 2000. *Sensitivity analysis: Gauging the worth of scientific models*. John Wiley & Sons.

[76] Raunak Sarbajna, Karima Elgarroussi, Hoang D Vo, Jianyuan Ni, and Christoph F Eick. 2024. COVID19-CBABM: A City-Based Agent Based Disease Spread Modeling Framework. *arXiv preprint arXiv:2409.05235* (2024).

[77] Petrônio CL Silva, Paulo VC Batista, Hélder S Lima, Marcos A Alves, Frederico G Guimarães, and Rodrigo CP Silva. 2020. COVID-ABS: An agent-based model of COVID-19 epidemic to simulate health and economic effects of social distancing interventions. *Chaos, Solitons & Fractals* 139 (2020), 110088.

[78] Tzu-Hsi Song, Leonardo Clemente, Xiang Pan, Junbong Jang, Mauricio Santillana, and Kwonmoo Lee. 2025. Fine-grained forecasting of COVID-19 trends at the county level in the United States. *NPJ Digital Medicine* 8, 1 (2025), 204.

[79] Yot Teerawattananon, KC Sarin, Y-Ling Chi, Saudamini Dabak, Joseph Kazibwe, Hannah Clapham, Claudia Lopez Hernandez, Gabriel M Leung, Hamid Sharifi, Mahlet Habtemariam, et al. 2022. Recalibrating the notion of modelling for policymaking during pandemics. *Epidemics* 38 (2022), 100552.

[80] PFMrivm Teunis and AHrivm Havelaar. 2000. The Beta Poisson dose-response model is not a single-hit model. *Risk Analysis* 20, 4 (2000), 513–520.

[81] Jason Thompson, Roderick McClure, Nick Scott, Margaret Hellard, Romesh Abeysuriya, Rajith Vidanaarachchi, John Thwaites, Jeffrey V Lazarus, John Lavis, Susan Michie, et al. 2022. A framework for considering the utility of models when facing tough decisions in public health: a guideline for policy-makers. *Health Research Policy and Systems* 20, 1 (2022), 107.

[82] Huaiyu Tian, Yonghong Liu, Yidan Li, Chieh-Hsi Wu, Bin Chen, Moritz UG Kraemer, Bingying Li, Jun Cai, Bo Xu, Qiqi Yang, et al. 2020. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* 368, 6491 (2020), 638–642.

[83] Ian Vernon, Jonathan Owen, Joseph Aylett-Bullock, Carolina Cuesta-Lazaro, Jonathan Frawley, Arnau Quera-Bofarull, Aidan Sedgewick, Difu Shi, Henry Truong, Mark Turner, et al. 2022. Bayesian emulation and history matching of JUNE. *Philosophical Transactions of the Royal Society A* 380, 2233 (2022), 20220039.

[84] Ross Williams, Niyousha Hosseinichimeh, Aritra Majumdar, and Navid Ghaffarzadegan. 2023. Epidemic modeling with generative agents. *arXiv preprint arXiv:2307.04986* (2023).

[85] Sarah Wise, Sveta Milusheva, and Sophie Ayling. 2022. The importance of scaling for an agent based model: An illustrative case study with covid-19 in Zimbabwe. In *International Conference on Computational Science*. Springer, 259–265.

[86] Nick Yeung and Christopher Summerfield. 2012. Metacognition in human decision-making: confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367, 1594 (2012), 1310–1321.

[87] Sijin Zhang, Alvaro Orsi, and Lei Chen. 2024. An AI-enabled Agent-Based Model and Its Application in Measles Outbreak Simulation for New Zealand. *arXiv preprint arXiv:2403.03434* (2024).