Hybrid metapopulation agent-based epidemiological models for efficient insight on the individual scale: A contribution to green computing

Julia Bicker [a, *], René Schmieding [a], Michael Meyer-Hermann [b], Martin J. Kühn [a, c]

[a] Institute of Software Technology, Department of High-Performance Computing, German Aerospace Center, Cologne, Germany
[b] Department of Systems Immunology and Braunschweig Integrated Centre of Systems Biology (BRICS), Helmholtz Centre for Infection Research, Brunswick, Germany
[c] Life and Medical Sciences Institute and Bonn Center for Mathematical Life Sciences, University of Bonn, Bonn, Germany

ARTICLE INFO

ABSTRACT

Emerging infectious diseases and climate change are two of the major challenges in 21st century. Although over the past decades, highly-resolved mathematical models have contributed in understanding dynamics of infectious diseases and are of great aid when it comes to finding suitable intervention measures, they may need substantial computational effort and produce significant CO_2 emissions. Two popular modeling approaches for mitigating infectious disease dynamics are agent-based and population-based models. Agent-based models (ABMs) offer a microscopic view and are thus able to capture heterogeneous human contact behavior and mobility patterns. However, insights on individual-level dynamics come with high computational effort that scales with the number of agents. On the other hand, population-based models (PBMs) using e.g. ordinary differential equations (ODEs) are computationally efficient even for large populations due to their complexity being independent of the population size. Yet, population-based models are restricted in their granularity as they assume a (to some extent) homogeneous and well-mixed population. To manage the trade-off between computational complexity and level of detail, we propose spatial- and temporal-hybrid models that use ABMs only in an area or time frame of interest. To account for relevant influences to disease dynamics, e.g., from outside, due to commuting activities, we use population-based models, only adding moderate computational costs. Our hybridization approach demonstrates significant reduction in computational effort by up to 98% — without losing the required depth in information in the focus frame. The hybrid models used in our numerical simulations are based on two recently proposed models, however, any suitable combination of ABM and PBM could be used, too. Concluding, hybrid epidemiological models can provide insights on the individual scale where necessary, using aggregated models where possible, thereby making a contribution to green computing.

* Corresponding author.
  E-mail addresses: julia.bicker@dlr.de (J. Bicker), martin.kuehn@dlr.de (M.J. Kühn).
  Peer review under the responsibility of KeAi Communications Co., Ltd.

1. Introduction

Two of the major 21st century's challenges are accelerating climate change (Callaghan et al., 2010) and emerging infectious diseases with an expected increase in frequency of epidemics and pandemics (Daszak et al., 2020). As these challenges are driven by the same underlying causes, sets of proposed counteractions might go hand-in-hand (Daszak et al., 2020; Epstein, 2001). In order to assess counteractions and support decision-makers, mathematical models are viable tools that have been used extensively in both domains. Models can help to understand the situation at hand and to analyze potential future developments, which often cannot be conducted in classical experiments. The recent COVID-19 pandemic has demonstrated the benefit of mathematical models to analyze ongoing infectious disease dynamics and to evaluate potential scenarios, see, e.g., (Kerr et al., 2021; Koslow et al., 2022; Reiner et al., 2020). Extensive testing strategies maintaining mobility as a social good, have, for instance, been studied in (Kühn et al., 2022). The set of mathematical models to address epidemiological emergencies to provide guidance to decision makers is diverse. It ranges from classical statistical models such as generalized regression models, through population-based (Medlock, 2004; Reiner et al., 2020), metapopulation (Kühn et al., 2021; Liu, Ong, & Pang, 2022), and agent-based (Bershteyn et al., 2018; Bracher et al., 2021; Kerkmann et al., 2024; Kerr et al., 2021; Müller et al., 2021) models to machine and deep learning models (Garnett, Cousens, Hallett, Steketee, & Walker, 2011; Schmidt, Zunker, Heinlein, & Kühn, 2024; Wynants et al., 2020). Population-based models using sets of ordinary (Kühn et al., 2021; Reiner et al., 2020) or integro-differential-equation-based models (Medlock, 2004; Wendler, Plötzke, Tritzschak, & Kühn, 2024) are well established methods that have been used by hundreds of work groups worldwide to analyze infectious disease dynamics. Their advantages are the ease of analysis, their well-understood character and the low computational requirements which do not depend on the number of individuals in the population but on the number of different subpopulations. However, their insight on the individual-level dynamics or on (singular) stochastic effects is limited and they also fail in the beginning of an outbreak when the number of infected is small. Agent- or individual-based models (ABMs or IBMs), on the other hand, allow the stochastic simulation of individual behavior and reaction to diseases, in a social and immunological dimension, see, e.g., (Kerr et al., 2021; Müller, Paltra, Rehmann, Nagel, & Conrad, 2023). Though, this comes at the cost of superlinear complexity. While the energy consumption corresponding to computational cost can be reduced through, e.g., optimizations on the cache level (Willem et al., 2015) or the optimal use of heterogeneous hardware structures such as SIMD registers (Kühn, Holke, et al., 2023), we here propose a complementary approach to directly avoid less impactful computations through the development of well-designed hybridization concepts.

The authors of (Müller et al., 2021) demonstrated that the inclusion of disease import, i.e., infections caused by travelers and commuters, is an important factor to explain disease dynamics. Via the metapopulation approach in (Kühn et al., 2022), it was shown how large the effects of disease import can become if testing is not done properly. Using simple statistics for disease import can be a suitable first step in improving model precision in a considered focus region. However, if disease import is subject to change through interventions, disease dynamics in coupled regions are advised to also be modeled dynamically.

In this work, we will present a framework for spatial- and temporal-hybrid epidemiological modeling, combining any pair of suitable fine-granular agent-based and coarse-granular population-based models. We will then suggest a particular combination of two suitable models introduced recently in (Winkelmann, Zonker, Schütte, & Conrad, 2021) and present results for three simplified examples thereby demonstrating potential areas of application and the benefit of the hybrid models. In this context, we will simulate a COVID-19-like disease with parameters motivated by (Kühn et al., 2021). However, as the scope of the current paper is the disease-agnostic introduction of novel hybrid models, no further parameter refinement or real-world application for COVID-19 has been performed. Let us note that any other disease model can be put in our framework. We will show that our hybrid modeling approach can account for individual developing disease dynamics, otherwise modeled with pure ABMs, while drastically reducing computational effort. We will provide discussions, limitations, and directions for future research before we eventually draw a conclusion.

2. Materials and methods

In this section, we first present related work on hybrid models. Then we present an agent-based and a metapopulation model, and eventually we introduce our novel, hybrid modeling framework which is applied to the two models introduced before.

2.1. Related work

The development of hybrid models is often driven by the intention to combine the advantages of two different modeling approaches. In the context of infectious disease dynamics, several suggestions for hybrid modeling have been made. While the authors of (Angione, Silverman, & Yaneske, 2022; Avegliano & Sichman, 2023; Robertson, Safta, Collier, Ozik, & Ray, 2024) used ABMs together with ordinary differential equation (ODE)-based and machine learning models, respectively, to improve parameter estimation, we focus on hybrid models where both model types are used in the simulation and prediction process. However, also in this context, the term *hybrid model* is used for many different realizations.

In (Banos, Corson, Gaudou, Laperrière, & Coyrehourcq, 2015; Kasereka, Kasoro, & Chokki, 2014; Bradhurst et al., 2015; Marilleau, Lang, & Giraudoux, 2018; Yoneyama, Das, & Krishnamoorthy, 2012), the authors used metapopulation or network-based models which are sometimes also denoted as hybrid models. These metapopulation models combined ODE-based models for different considered regions while the mobility between regions was modeled with an agent-based modeling principle. Similarly, although not coined hybrid, the PDMM of (Winkelmann et al., 2021) used ODE-based formulations for single regions and a stochastic jump process to model mobility between the regions. In (Kühn et al., 2021; Zunker et al., 2024), other examples of network-based metapopulation models were presented. There, ODE-based models were used for particular regions representing the nodes of a graph and deterministic mobility was realized through the edges of the graph. While also using a metapopulation model combined with an ABM in our hybridization framework, our understanding of hybrid goes beyond the aforementioned.

In (Bobashev, Goedecke, Yu, & Epstein, 2007; Hunter, Mac Namee, & Kelleher, 2020; Quang Nghi, Nguyen-Huu, Grignard, Xuan Huynh, & Drogoul, 2016), hybrid models were implemented that used different model types, namely agent-based and population-based models, in different temporal regimes. The authors used ABMs for low-incidence regimes and upon passing a threshold on the number of infected individuals, a switch to an ODE- or difference equation-based model was executed. We call this temporal coupling of model types *temporal hybridization*.

The authors of (Sewall, Wilkie, & Lin, 2011) implemented a hybrid model for simulating large-scale traffic. They coupled an ABM and an population-based model and provided rules for spatial exchange when one region was modeled with the ABM and a neighboring region with the population-based model. We call this method *spatial hybridization*. Similar to the temporal hybridization, they also implemented the conversion of a whole region from one model to another, here not dependent on time but on user input.

In this paper, we will introduce novel *spatial-hybrid models* for infectious disease dynamics which use an ABM in a focus region and macroscopic ODE-based models in all connected regions. Furthermore, we will extend the application of *temporal-hybrid models* where switching takes place upon exceeding a threshold.

2.2. High-performance computing, carbon footprint, and green computing

High-performance computing (HPC) can enable breakthrough research results for ABMs (Collier & North, 2013) but require great computing capability which naturally comes with high energy demands. The regularly updated Top500 list[1] of the 500 most powerful commercially available computer systems known to the providers, lists the (peak) power consumption for 199 out of the 500 supercomputers in November 2024. The total power consumptions for these 199 system sums up to 431,222.57 kW. Using an optimistic emission of 380g CO_2 for the production of 1 kWh power, provided by the German Federal Environment Agency for the 2023 German energy mix with 51.9% of renewable energy (Icha & Lauf, 2024), these systems alone have a daily emission of 3932 tons of CO_2. The author of (Portegies Zwart, 2020) situated the carbon footprint for the use of one million compute cores between an 8-h air travel and a launch of a Falcon 9 rocket into space. On a broader scale, green computing considers the entire product life cycle of technological devices from production through operation and recycling, resulting in several challenges as highlighted by (Wang, 2008). While substantial progress has been made on the hardware side to increase energy efficiency on the operational level, in particular through graphic processing units (Portegies Zwart, 2020; Rofouei, Stathopoulos, Ryffel, Kaiser, & Sarrafzadeh, 2008), carbon emission of HPC systems cannot be neglected. The manufacturing or recycling processes targeted by green computing have to be considered on another level, nevertheless, contributions to green computing can be made by scientists when designing large-scale (scientific) software with energy-awareness. Aside from classical performance engineering and code optimization to speed-up algorithms or achieve more efficient hardware usage, hybrid, multiscale, or adaptive frameworks can achieve substantial performance gains or energy consumption reductions. Through the principle of computing on a finest level where necessary and aggregating information where possible, green computing advancements can be made by scientists and software developers.

2.3. Explicit agent- and metapopulation-based models

In this section, we will introduce an ODE-based metapopulation model and an ABM, originally introduced in (Winkelmann et al., 2021). These models will be used in the applications of our hybrid modeling framework. We will extend the set of *infection states* to allow for a- and presymptomatic transmission. The description of the transmission and course of the disease model is provided in Section 2.3.3.

We start with the ABM, whose change over time is described by a so called master equation motivated by chemical reaction systems, cf., e.g., (Gillespie, 1977). While this does impose some restriction on what the agents' (inter)actions may be, it allows us to derive a piecewise deterministic metapopulation model (PDMM) that shares parameters and behavior with the ABM. The PDMM can be obtained by a model reduction of the presented ABM, see (Winkelmann et al., 2021). All models are implemented in C++, which is among the fastest and most efficient programming languages (Pereira et al., 2017; Portegies Zwart, 2020), and within the MEmilio software (Kühn, Abele, et al., 2023) for modular epidemics simulations.

---

[1] https://top500.org/lists/top500/list/2024/11/.

2.3.1. A diffusion- and drift-based ABM

The nature of agent-based modeling is to capture more intuitively and to model in a direct manner individual (human) properties and interactions. Corresponding models have already been explored before the availability of modern computers; see, e.g., (Schelling, 1971). The authors of (Hunter, Mac Namee, & Kelleher, 2017) have set up a so far nonexisting taxonomy for ABMs in infectious diseases and considered if particular ABMs included modules or submodels for society, transportation, and environment.

According to our understanding, an ABM consists of a finite number of unique *agents* $\alpha_1, \ldots, \alpha_{n_a}$ and an *environment* in which and to which an agent reacts. Furthermore,

- an agent is characterized by a finite number of *features* that determine its *state*,
- an agent interacts with other agents and their joint environment according to *interaction rules*,
- the state of an agent or the environment changes through interactions or with time.

The features of the used ABM are infection states given by the overall infection dynamics and agents' positions determined by agents' movements. Hence, an agent is represented as a tuple $(x, z) \in \Omega \times \mathscr{Z}$, where $\Omega \subset \mathbb{R}^2$, the environment, is a compact domain and $\mathscr{Z} = (z_1, z_2, \ldots, z_{n_i})$ a set of infection states. Considering n_a agents, the system state is defined as a vector $Y := (X, Z) \in \mathbb{Y} := \Omega^{n_a} \times \mathscr{Z}^{n_a}$, where an agent α (with its two features, position and infection state) is represented by the α-th component of the system state $Y_\alpha = (x_\alpha, z_\alpha) = (X, Z)_\alpha$. The evolution of the system state over time is modeled as a continuous-time Markov process $(Y(t),\ t \geq 0)$. The evolution of the process is determined by the master equation, that describes the relevant probability measure by the change of its probability density

$$\partial_t p(X, Z; t) = G p(X, Z; t) + L p(X, Z; t). \tag{1}$$

The operator G defines the infection state adoptions and only acts on Z, while L defines location changes, only acting on X.

For the formal definition of G and L the reader is referred to (Winkelmann et al., 2021). Here, we will only describe the realization of the operators used for the later applications. Starting with the infection state dynamics, let $i \to j$ be an infection state adoption, where $i \in \mathscr{Z}$ is the current infection state, and $j \in \mathscr{Z} \setminus \{i\}$ the infection state to adopt. For any particular agent, we assume that the likelihood of this adoption to happen in a given amount of time is determined by some rate $f_{i,j}^{(\alpha)}(Y(t))$ depending on the system state at the current time t. We call this rate *(infection state) adoption rate*. A suitable method to realize this rate-dependent infection dynamic is using independent inhomogeneous Poisson processes $\mathscr{P}_{i,j}^{(\alpha)}(t) := \mathscr{P}_{f_{i,j}^{(\alpha)}(Y(t))}$ for infection state adoption $i \to j$ and agent α. As a natural extension of $f_{i,j}^{(\alpha)}$, for $j = i$, we define $f_{i,i}^{(\alpha)} \equiv 0$. These rates should be interpreted as rates to change the current state.

For a given infection state adoption, various compartments can influence the adoption rate. The influence of an adoption $i \to j$ is defined as an index set $\Psi_{i,j} \subset \mathscr{Z} \setminus \{i\}$. We use these influences to differentiate between two types of adoptions.

1. First-order adoption: An adoption event that does not require interactions with other agents, i.e., spontaneous infection state change, like recovery or death from the disease. Its adoption rate only depends on the compartment the agent is currently in, thus $\Psi_{i,j} = \varnothing$.
2. Second-order adoption: An adoption based on pairwise interactions, i.e., infection via a contact with an infectious agent. The rate depends on the compartment of origin of an agent as well as all influences in $\Psi_{i,j} \neq \varnothing$.

Therefore, if there are no influences, the adoption event is of first-order, and of second-order if there are any.

The adoption rates are given by adoption rate functions $f_{i,j} : \mathbb{Y} \to [0, \infty)^{n_a}$, which in general depend on the whole system state. For each agent α, the α-th component of $f_{i,j}$ is given by the function $f_{i,j}^{(\alpha)} : \mathbb{Y} \to [0, \infty)$, which depending on the order of the adoption $i \to j$ is defined as

$$f_{i,j}^{(\alpha)}(X, Z) = \delta_i(z_\alpha) \gamma_{i,j}(x_\alpha) \tag{2}$$

for first-order adoption events, or

$$f_{i,j}^{(\alpha)}(X, Z) = \delta_i(z_\alpha) \sum_{\tau \in \Psi_{i,j}} \sum_{\beta=1}^{n_a} \delta_\tau(z_\beta) \gamma_{i,j,\tau}(x_\alpha, x_\beta) \tag{3}$$

for second-order adoptions. Here $\delta_i(z_\alpha)$ is the Kronecker-delta, being one when agent α has infection state i and zero otherwise. The functions $\gamma_{i,j}: \Omega \to [0, \infty)$ and $\gamma_{i,j,\tau}: \Omega^2 \to [0, \infty)$ give the magnitude of the adoption rate, only depending on an agent's position. We choose

$$\gamma_{i,j}(x) = c_{i,j} \text{ and } \gamma_{i,j,\tau}(x,y) = c_{i,j,\tau} \mathbb{1}_{\|x-y\|_\Omega \leq r}, \tag{4}$$

with a cutoff for second-order adoptions beyond the interaction radius $r > 0$. The radius determines how close two agents have to be for being considered contacts, and we use rate constants $c_{i,j} \geq 0$, $c_{i,j,\tau} \geq 0$. We always set $c_{i,i} = c_{i,i,\tau} = 0$ as adoptions from one state to itself are not possible. Note that rate constants $c_{i,j}$ and $c_{i,j,\tau}$ could also be time-varying $c_{i,j}(t)$ and $c_{i,j,\tau}(t)$.

Agents' movement trajectories y (omitting a dependency on infection state i) are determined by independent diffusion processes. They are given by stochastic differential equations of the form

$$\frac{dx(t)}{dt} = b(t, x(t)) + \sigma(t, x(t))\xi(t). \tag{5}$$

Here, $b : [0, T] \times \mathbb{R}^2 \to \mathbb{R}^2$ is called drift coefficient or potential on the domain, and $\sigma : [0, T] \times \mathbb{R}^2 \to \mathbb{R}^{2 \times m}$ is called diffusion coefficient or noise. The potential can be given by functions $F(t) : \Omega \to \mathbb{R}$ in which case $b(t) := -\nabla F(t)$. As b is assumed to be deterministic, all random behavior is solely caused by the noise. The magnitude of the noise determines the influence of the white noise process $\xi : [0, T] \to \mathbb{R}^m$ on the diffusion. Formally, it is defined as $\xi = \frac{dW}{dt}$, the derivative of the Brownian motion in $\mathbb{R}^m$. The parameter m can be used to stratify agents' movement, for example with respect to infection state or region.

In practice, we use an agent's current position as initial value for Eq. (5), and only integrate a small time step, using the Euler-Maruyama method, to get its new position.

2.3.2. The corresponding piecewise deterministic metapopulation model

The corresponding piecewise deterministic metapopulation model (PDMM) simplifies the ABM by using two approximations. In the following, we will shortly describe these approximations. The full derivation of the PDMM from the ABM introduced above can be found in (Winkelmann et al., 2021).

First, computational effort is reduced by discretizing the domain. The domain is split into subregions $\Omega_1, ..., \Omega_{n_R}$ with $\Omega = \bigcup_{k=1}^{n_R} \Omega_k$ and agents are aggregated to subpopulations in the subregions. Hence, an agent's position is now defined only by its subregion index $k \in \{1, ..., n_R\}$ instead of an exact position $x \in \Omega$. Consequently, all agents in one subregion have the same position. The system state is defined as the matrix $N = (N_i^{(k)})_{i \in \mathcal{Z}, k=1,...,n_R} \in \mathbb{R}_{\geq 0}^{n_I \times n_R}$, with $N_i^{(k)}$ the number of agents in region Ω_k and infection state i. Let $e_i^{(k)}$ be the matrix where the entry for (i, k) is one and all others are zero. The movement between subpopulations is modeled by Poisson processes $\mathcal{L}_i^{(k,l)}$ with rates $\lambda_i^{(k,l)}$ for spatial transitions from subregion k to l, $k, l \in \{1, ..., n_R\}$. These rates depend on the transitioning agent's infection state i and are of the form $N \to N + e_i^{(l)} - e_i^{(k)}$. Formally, the spatial transition rates are given by the location change operator L through

$$\lambda_i^{(k,l)} := \frac{\int_\Omega \delta_{\Omega_l}(x) L_i[\delta_{\Omega_k}(x)] dx}{\int_\Omega \delta_{\Omega_k}(x) dx}, \tag{6}$$

but in practice they may be easier to sample than to calculate analytically. The space of all possible system states is given by

$$\mathbb{M}_{n_a} = \left\{ N \,\middle|\, \sum_{i=1}^{n_I} \sum_{k=1}^{n_R} N_i^{(k)} = n_a \right\}, \tag{7}$$

where n_a is again the total number of agents in the model.

Secondly, model and computational cost is reduced by approximating the infection state adoption processes by a deterministic system of equations. The underlying assumption is that the populations are large such that infection state adoptions are relatively rapid and spatial transitions are comparatively rare. Using region-dependent adoption rate functions $\hat{f}_{i,j}^{(k)} : \mathbb{R}_{\geq 0}^{n_I} \to \mathbb{R}_{\geq 0}$, the adoption dynamics are given by ODEs of the form

$$\frac{d}{dt} N_i^{(k)}(t) = \sum_{j \in \mathcal{Z}, i \neq j} \left(\hat{f}_{j,i}^{(k)}\left(N^{(k)}(t)\right) - \hat{f}_{i,j}^{(k)}\left(N^{(k)}(t)\right) \right), \tag{8}$$

without accounting for spatial transitions. The evolution of the system state $N(t)$ over time is described by a continuous-time Markov jump process. Writing $(N(t))_{t \geq 0}$ as a path, we obtain

$$N(t) = N(0) + \sum_{\substack{k,l=1 \\ k \neq l}}^{n_R} \sum_{i=1}^{n_I} \mathscr{L}_i^{(k,l)} \left(\int_0^t \lambda_i^{(k,l)} N_i^{(k)}(s) \mathrm{d}s \right) \left(e_i^{(l)} - e_i^{(k)} \right) + \sum_{k=1}^{n_R} \sum_{i,j=1}^{n_I} \int_0^t \widehat{f}_{i,j}^{(k)}(N(s)) \mathrm{d}s \left(e_j^{(k)} - e_i^{(k)} \right). \tag{9}$$

For the full derivation from the ABM through a stochastic metapopulation model to the PDMM using convergence results for Markov processes, see again (Winkelmann et al., 2021).

2.3.3. Transmission and course of the disease

As course of the disease model, we use a Susceptible-Exposed-Carrier-Infected-Recovered-Dead model, thus $\mathscr{Z} = (S, E, C, I, R, D)$. Individuals in the *Carrier* compartment may be either pre- or asymptomatic, hence they do not show symptoms while the *Infected* compartment encompasses all individuals showing symptoms ranging from mild to severe. Both compartments are infectious to *Susceptibles*. There is only one second-order adoption (see Fig. 1) which is $S \rightarrow E$ with influences $\Psi_{SE} = \{C, I\}$. As there is no outflow from the R compartment, this model does not consider the possibility of reinfection.

Given the courses of the disease defined here, the local equations giving the infection state dynamics in region Ω_k of the PDMM from Section 2.3.2 write

$$\frac{d}{dt} N_S^{(k)} = -\frac{\rho^{(k)} \left(\xi_C^{(k)} N_C^{(k)} + \xi_I^{(k)} N_I^{(k)} \right)}{n_a^k} N_S^{(k)},$$

$$\frac{d}{dt} N_E^{(k)} = \frac{\rho^{(k)} \left(\xi_C^{(k)} N_C^{(k)} + \xi_I^{(k)} N_I^{(k)} \right)}{n_a^k} N_S^{(k)} - \gamma_{E,C}^{(k)} N_E^{(k)},$$

$$\frac{d}{dt} N_C^{(k)} = \gamma_{E,C}^{(k)} N_E^{(k)} - \left(\gamma_{C,R}^{(k)} + \gamma_{C,I}^{(k)} \right) N_C^{(k)},$$

$$\frac{d}{dt} N_I^{(k)} = \gamma_{C,I}^{(k)} N_C^{(k)} - \left(\gamma_{I,R}^{(k)} + \gamma_{I,D}^{(k)} \right) N_I^{(k)}, \tag{10}$$

$$\frac{d}{dt} N_R^{(k)} = \gamma_{C,R}^{(k)} N_C^{(k)} + \gamma_{I,R}^{(k)} N_I^{(k)},$$

$$\frac{d}{dt} N_D^{(k)} = \gamma_{I,D}^{(k)} N_I^{(k)}.$$

2.4. A framework of hybrid modeling and its application

In this section, we introduce the general concept of two different types of hybrid epidemiological models: spatial-hybrid and temporal-hybrid models. These approaches do not exclude each other so that also spatio-temporal-hybrid models can be realized. Both hybridization approaches are shown in Fig. 2. For both approaches, we will provide explicit applications with the models presented in the previous section.

The motivation to use population-based models combined with ABMs for disease dynamics is to reduce the computational effort of individual-level models which scale with the number of considered agents or individuals. In particular, we here provide the computational cost for the two explicit models introduced before. For both models, we need to consider computational costs for spatial transitions or movements and infection state adoption dynamics. The movement of agents in the ABM is given by a diffusion process that has to be evaluated for every agent. Assuming a constant time for the evaluation, the cost to calculate movement dynamics in the ABM lies in $\mathscr{O}(n_a)$. Furthermore, the infection state adoption rates have to be evaluated in every iteration. As first order adoption rates are computed per agent, the complexity as well is in $\mathscr{O}(n_a)$. For second-order adoptions, pairwise comparison of agents is necessary, hence, we obtain a superlinear complexity and in the worst case one iteration has computational cost in $\mathscr{O}(n_a^2)$. For the PDMM the cost for infection state adoption dynamics in one
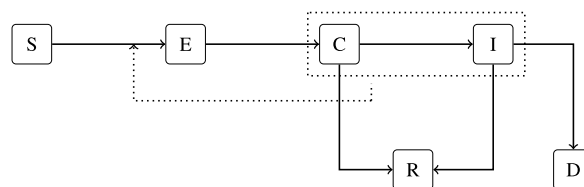

Fig. 1. Flow chart of the infection state adoption model. The infection states are Susceptible (S), Exposed (E), Carrier (C), Infected (I), Recovered (R) and Dead (D). The dotted lines shows the influences for the second-order adoption.
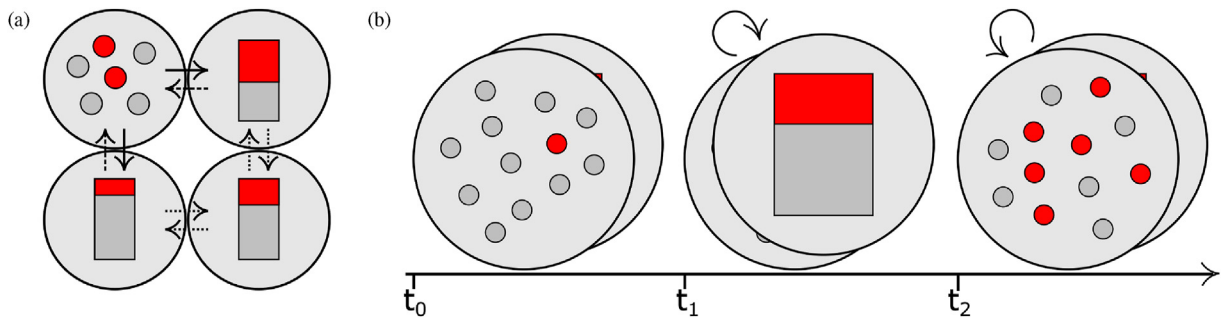
Fig. 2. Design sketch of a spatial-hybrid model (a) and a temporal-hybrid model (b). Large gray disks represents (sub)regions, with possible population exchanges indicated by arrows. The population is represented by small disks for agents and bar charts for ODE compartments. Red indicates infected agents or population shares, gray indicates noninfected agents or population shares. The exchange in (a) only moves small parts of a local population to another region, the exchange in (b) moves the entire population to the other model.

iteration is given by integrating the system of differential equations, thus it is independent of n_a. Technically, the cost for integration scales with the number of equations meaning the number of regions and compartments, but their size is negligibly small compared to n_a. Therefore, the relevant complexity and runtime factor in the PDMM is given by the spatial transition events. The frequency of the spatial transitions depends on the magnitude of the rates. As the transition rates are multiplied with the corresponding subpopulation, see Eq. (9), the complexity of the spatial dynamics lies in $\mathcal{O}(n_a)$. A summary of the model characteristics is given in Table 1.

A requirement to the two models used is that we can define a mapping between them. In its simplest form, the coarse-granular model may be obtained by projection from the fine model. In a more elaborated setting, we need to map fine-granular, continuous features like viral load to a discrete, coarse-granular definition. Naturally, going from fine-to-coarse is easier as information gets reduced. Going from coarse-to-fine either needs sampling or the replication of retained information to realize a realistic model over the whole simulation horizon (spatial or temporal). While in the current paper we only needed to sample a corresponding position in a 2D plane, more sophisticated sampling (based on retained information on travelers) will be necessary for more realistic ABMs. The corresponding sampling then could include information on households or workplaces and agents' immune histories. In general, the coarse (and hybrid) results should approximate the fine result. One condition for this would be that the models' mean results converge. In addition, both models should match in the sense that there is a projection from the fine to the coarse model population, with a suitable mapping from coarse to fine such that the information lost when retrieving the fine model population from the coarse model, does not exceed the approximation error. Nevertheless, to determine whether the results of a particular hybrid model will be accepted depends on individual evaluation of the trade-off between approximation error and speed-up.

2.4.1. Spatial hybridization

The idea of a *spatial-hybrid model* originates from research questions on fine-granular infection spread, like household or workplace transmission in a particular region. To consider only this (focus) region would neglect the influence of connected regions on infection dynamics through, e.g., commuting activities. However, the use of fine-granular models for all connected regions might be prohibitively expensive or individual-level data is either exclusively available in the focus region, or substantially costly to collect for outside regions.

Consequently, we model only the focus region by a fine-granular model and use coarse-granular models in any connected region. With this approach we get results with the desired resolution in the focus region while considering the dynamic, time-dependent influence of the connected regions in a runtime efficient manner. In the following, we will denote this approach as *spatial hybridization* and will refer to hybrid models with higher granularity in a focus region as *spatial-hybrid models*.

For the spatial hybridization, we require a disjoint partition of the modeled domain into at least two subregions. The exchange of populations between the different regions can be realized with different mobility models, however, sending agents to a coarser model does cause loss of information (e.g. sending an agent to a simple ODE-based model will only retain its infection state and lose all spatial information). To the best of our knowledge, there is only little theory on this exchange problem and the exchange is highly dependent on the concrete models used. Some discussions and suggested models can be found in (Hunter et al., 2020; Sewall et al., 2011; Bradhurst et al., 2015). While a particular goal is that the hybridization approaches the ABM outcome, for instance, ODE-based models are not able to let a virus go completely extinct while ABMs can simulate this extinction. This does not mean that a spatial-hybrid of ODE-ABM cannot be used if infection numbers are small in a region where the potential of extinction is real. This only means that the results may differ from the results of a fully fine-resolved (stochastic) model and hence that this limitation and its impact needs to be considered when evaluating simulation results. For further discussion on this effect, see e.g. (Schaller & Meyer-Hermann, 2006; Shnerb, Louzoun, Bettelheim, & Solomon, 2000). More suitable combinations could use stochastic differential equation-based models or a mixed spatial-temporal-hybrid (in combination with the next section).

Table 1
Model comparison of ABM and PDMM. The term $\mathscr{O}(n_a)^*$ describes the runtime scaling due to spatial transition rates and can be replaced by $\mathscr{O}(n_r n_l)$, with $n_r n_l \ll n_a$, under the assumption that spatial transitions are several orders of magnitude lower than infection state adoptions.

	ABM	PDMM
Complexity	superlinear, up to $\mathscr{O}(n_a^2)$	$\mathscr{O}(n_a)^*$
Spatial domain	continuous	discrete
Adoption dynamics	stochastic	deterministic

Two ways to obtain compatible models are to fit the two models against each other, or to derive the coarse model (parameters) from the fine one. However, in both cases it is important to understand the model behavior very well to be aware of inter-dependencies. For instance, location and infection state changes may depend on each other and cannot always be fitted separately. An interesting discussion can be found in (Niemann, Uram, Wolf, Djurdjevac Conrad, & Weiser, 2024).

In the following, we will provide some further details on the precise realization of the spatial hybridization for the ABM and PDMM introduced in Section 2.3. Let Ω_1 be the focus region. For the precise implementation, ABM and PDMM are set up for the whole domain while both models have no movement or infection state adoptions in the non-relevant regions meaning the focus region for the PDMM and all regions outside the focus regions for the ABM, i.e., the corresponding rates are zero. Then, both models are run independently from each other as long as no agents are leaving or entering the focus region. On time points for population exchange, models will be synchronized. As mentioned before, the exchange of agents and populations between the models is a crucial aspect of the hybridization. As the mobility process of the PDMM is also based on agents and not on shares of the whole population and as we set the adoption rates in the PDMM in the focus region to zero, it holds $N_i^{(1)} \in \mathbb{N}$ for $i \in \mathscr{Z}$. The exchange between the models is done as follows.

- ABM $\rightarrow$ PDMM: Let t be the time point of the last exchange and $Y^-(t, t') \subseteq Y(t')$ be the set of agents that need to be exchanged from ABM to PDMM, i.e., agents with positions $x \notin \Omega_1$, at time point $t' > t$. As we switch from fine-granular to the corresponding coarse-granular model, the exchange can be done by a trivial projection. Agents in $Y^-(t, t')$ are just added to the subpopulation $N_i^{(\phi(x))}(t')$ in the PDMM according to their infection state $i \in \mathscr{Z}$ and a projection $\phi: \Omega \rightarrow \{2, ..., n_R\}$ with $\phi(x) = k$ for $x \in \Omega_k$. Subsequently, this agent is removed from the ABM.
- PDMM $\rightarrow$ ABM: The exchange from coarse-granular to fine-granular model is not as trivial, since the PDMM only has spatial information aggregated to subregions and, hence, misses concrete positions inside the subregions. For the exchange at time point t', $\sum_{i \in \mathscr{Z}} N_i^{(1)}(t')$ new agent tuples are created and added to the ABM system state $Y(t')$. The infection state $z \in \mathscr{Z}$ of a newly created agent is directly given by the PDMM subpopulation index. Its position $x \in \Omega_1$ is sampled from a distribution $\mathscr{P}$ which depends in general on the ABM movement operator L, see Eq. (1). As we use a diffusion process in our applications, $\mathscr{P}$ depends on the potential b, see Eq. (5), and will be precisely defined in Section 3.1 for the two examples provided there.

To get the initial populations for the spatial-hybrid model, we first only add agents to the ABM and then make one exchange step to set the initial system state $N(0)$ in the PDMM. With this, all agents outside Ω_1 are removed from the ABM, and in the PDMM the population in Ω_1 is set to 0.

The implementation of the spatial hybridization is shown in Algorithm 1.

Algorithm 1. Spatial hybridization

---

1 Create ABM and PDMM for $\Omega = \bigcup_{k=1}^{n_R} \Omega_k$ and $t = t_0$;
2 Set and restrict rates and populations, see Section 2.4.1;
3   ABM for Ω_1;
4   PDMM for $\Omega_2, ..., \Omega_{n_R}$;
5 **While** $t \in [t_0, t_{\max}]$ **do**;
6   Define next synchronization point t';
7   Advance ABM and PDMM from t to t';
8   /* Exchange agents between models */
9   **Foreach** $\alpha \in Y^-(t, t')$;
10   Set $N_{z_\alpha}^{(\phi(x_\alpha))}(t') += 1$;
11   Remove α from $Y(t')$;
12   **Foreach** $i \in \mathscr{Z}$;
13   **While** $N_i^{(1)}(t') > 0$ **do**;
14   Sample position $x \sim \mathscr{P}$;
15   Add agent (x, i) to $Y(t')$;
16   Set $N_i^{(1)}(t') -= 1$;
17   Set $t = t'$;

2.4.2. Temporal hybridization

The idea of *temporal-hybrid models* is based on the assumption that the lower the case numbers, the higher is the impact of stochastic events to the (simulation) outcome. On the other hand, if case numbers are high, individual behavior and stochastic events become less influential. This means that single simulation results of stochastic models come closer to averaged outcomes and hence mean-field models deliver a suitable approximation also for individual simulations.

The temporal hybridization is based on the same requirements as the spatial-hybrid model, except that we only use one model for the whole domain and switch between coarse- and fine-granular model during the course of the simulation. Then, instead of exchanging agents or populations on spatial transitions, we exchange the entire model at certain time points. These time points depend on the current system state and are thus selected dynamically. As a consequence, we need fine granular data for the whole domain.

In Section 3.2, we use an ABM to model disease spread when the number of infected individuals is below a predefined threshold and switch to a coarser, hence ODE-based or metapopulation model, as soon as the number of infected individuals exceeds the threshold. We denote hybrid models that switch between fine and coarse model using a predefined, time-dependent criterion as *temporal-hybrid models*. In literature this kind of hybrid model is also referred to as stage-based hybrid model, see e.g. (Bradhurst et al., 2015) and (Bobashev et al., 2007). In case of the virus-extinction example from above, we could change from an ODE-based or metapopulation model to an ABM whenever the number of virus carriers drops below a threshold that would realistically allow an extinction to occur and switch back whenever the virus will, with some certainty or probability, spread to a larger population. Note that this approach needs careful evaluation of the switching condition.

Like for the spatial-hybrid model the exchange of information between models is crucial and dependent on the concrete models used. A benefit of a temporal-hybrid model is that it can effectively save computational cost in periods of diffusive spread while maintaining the required level of detail in highly stochastic periods.

The setup for the temporal-hybrid model is similar to the implementation of the spatial hybridization described in Section 2.4.1, in so far as both models are setup with matching parameters. Instead of restricting the models to spatial regions, both are used on the whole domain. However, only one model runs at a time. To decide which model that is, we define a function Γ that returns, based on time, the system state at t and predefined criteria, a model {ABM, PDMM}, i.e., Γ is a function that determines with which model to continue from a given time point. In the application presented in Section 3.2, the model choice depends on the number of infectious agents, which makes most sense as infectious disease dynamics are driven by the stochasticity in the behavior of infectious agents. The procedure of the temporal hybridization is shown in Algorithm 2.

Algorithm 2.   Temporal hybridization

---
1 Create ABM and PDMM for Ω at $t = t_0$;
2 Select one model as $m_{current} \in$ {ABM, PDMM};
3 Initialize population only for $m_{current}$;
4 Define time step Δt;
5 **While** $t \in [t_0, t_{max}]$ **do;**
6 $m_{next} = \Gamma(m_{current}, t)$;
7   **If** $m_{current} \neq m_{next}$;
8   Move the population from $m_{current}$ to m_{next};
9 $m_{current} = m_{next}$;
10   Set $t = t + \Delta t$;
11   Advance $m_{current}$ to t;

Moving the population between both models follows the same principle as the agent exchange described in Section 2.4.1, applied to all agents in all regions at once, i.e., for changing from ABM to PDMM agents are just aggregated to subpopulations according to their position and infection state and for the change from PDMM to ABM, agents' concrete positions on the domain are sampled from a distribution $\mathscr{P}$ while their infection state is given by the subpopulation index. The distribution $\mathscr{P}$ we used for our examples will be defined in Section 3.2.

3. Results

In this section, we provide simulation results for the agent-based and the metapopulation model presented in Section 2.3. We will compare these results against our spatial- and temporal-hybrid models. The parameters related to infection dynamics are motivated by wild-type COVID-19 and based on (Kühn et al., 2021) for all examples (see Table D.4) − apart from the transmission rate $\rho^{(k)}$ from Susceptible to Exposed, which varies between the examples and also between regions of one example. The chosen values for the transmission rate $\rho^{(k)}$ are within the range presented in (Kühn et al., 2021) and lead to initial reproduction numbers R_0 between 0.8 and 5, roughly corresponding to estimates for the effective reproduction number for wild-type COVID-19 in different German Federal states between March and August 2020, see (Khailaie et al., 2021). However, the examples presented use simplified models as the scope of the current paper is the disease-agnostic introduction of novel hybrid models. Therefore no further parameter refinement for COVID-19 and comparison to real-world data has been done. All models are simulated using a temporal Gillespie algorithm, allowing us to obtain stochastically exact results without

sampling the distribution (Djurdjevac Conrad, Helfmann, Zonker, Winkelmann, & Schütte, 2018; Vestergaard & Génois, 2015). We will highlight differences and similarities in the particular model outcomes and compare model runtimes as a proxy indicator for computational effort. All simulations were conducted on a Intel Xeon "Skylake" Gold 6132 (2.60 GHz) with four nodes with 14 CPU cores each and 384 GB DDR4 memory, using separate cores to conduct Monte Carlo runs in parallel.

3.1. Spatial hybridization

In the following, we present the results for two applications of the spatial hybridization. The first application naturally extends the theoretical setting of (Winkelmann et al., 2021) by using four instead of two potential wells (i.e., local minima of the potential) and significantly more agents. In the second application, we consider a potential depicting the city of Munich and its surrounding counties. For both applications, we performed a sensitivity analysis to determine the most influential parameters, see Appendix A. Additionally, the implicit influence of parameters on the runtime of ABM, PDMM, and spatial-hybrid model was analyzed in Appendix B.

3.1.1. Application 1 - A mathematical quadwell example

In our first application, we use a function with four minima - a quadwell potential - and a constant noise term $\sigma \in \mathbb{R}$ for the diffusion process modeling agent movement in the ABM, see Eq. (5). The quadwell potential is constant in time and given by

$$F(x,y) = \left(x^2 - 1\right)^2 + \left(y^2 - 1\right)^2; \tag{11}$$

see Fig. 3. This function naturally defines four subregions for the PDMM that are given by the potential wells as $\Omega_1 = (-\infty, 0) \times (0, \infty)$, $\Omega_2 = (0, \infty) \times (0, \infty)$, $\Omega_3 = (-\infty, 0) \times (-\infty, 0)$ and $\Omega_4 = (0, \infty) \times (-\infty, 0)$. This corresponds to numbering the wells in Fig. 3b from 1 to 4 starting with the upper left and ending with the lower right well, numbered line by line. The magnitude of the noise term determines the number of transitions between metaregions, see Fig. E.1. As the infection dynamics in the focus region should be influenced by the dynamics in the other regions, we do want spatial transitions between the metaregions, which do not occur for $\sigma = 0.3$ (Fig.E.1a) and only very few for $\sigma = 0.5$ (Fig.E.1b). However, the spatial transitions should not happen as frequent as for $\sigma = 0.6$ (Fig.E.1c) as the assumption for the runtime advantage of the PDMM is that spatial transitions are rare compared to infection state adoptions (see Table 1). Therefore, we chose $\sigma = 0.55$ for our application, see Fig. 4.

Given the ABM noise term σ, the spatial transition rates for the PDMM can be sampled by calculating the average number of transitions in a unit-time step relative to the population in the subregion. We chose $\rho^{(k)} = 0.1$ for $k = 1, 3, 4$ corresponding to $R_0 = 0.8$ in region Ω_1, Ω_3 and Ω_4 and $\rho^{(2)} = 0.3$ corresponding to $R_0 = 2.4$ in region Ω_2. The parameters used for the quadwell potential are summarized in Table D.4 and Table D.5. The system state at four different time points for one ABM realization with 1000 agents in Fig. E.15 shows that, due to the higher transmission rate, the infection spreads much faster in Ω_2 compared to the other regions. We define Ω_1 as the focus region for the spatial-hybrid model and simulate the ABM, the PDMM and the spatial-hybrid model for $t_{\max} = 150$ days. When agents transition to the focus region in the spatial-hybrid model, their exact position has to be sampled from a distribution $\mathscr{P}$, see Algorithm 1. For this application we use a normal distribution whose parameters, i.e., mean and variance depend on the region an agent transitions from. The mean and variance were calculated from the transitions of 2850 agents shown in Fig. E.16. Denoting the region the transitioning agent came from as $\Omega_\alpha^{(old)}$, this yields for the new position $x_\alpha^{(new)} \in \Omega_1$ of agent α, $x_\alpha^{(new)} \sim \mathscr{P}_\alpha$ with

$$\mathscr{P}_\alpha = \begin{cases} \left(\mathscr{N}\left(-0.1, 0.03^2\right), \mathscr{N}\left(0.94, 0.19^2\right)\right), & \text{if } \Omega_\alpha^{(old)} = \Omega_2 \\ \left(\mathscr{N}\left(-0.94, 0.19^2\right), \mathscr{N}\left(0.1, 0.03^2\right)\right), & \text{if } \Omega_\alpha^{(old)} = \Omega_3 \\ \left(\mathscr{N}\left(-0.1, 0.03^2\right), \mathscr{N}\left(0.1, 0.03^2\right)\right), & \text{if } \Omega_\alpha^{(old)} = \Omega_4 \end{cases} \tag{12}$$

We chose $n_a = 8000$ agents with 1% of the population being initially infected in every well (0.2% Exposed, 0.3% Carrier and 0.5% Infected). The results of 500 simulations are shown in Fig. 5. We consider the ABM - the detailed model - as ground truth and calculate the *Mean Absolute Percentage Error (MAPE)*, the *Mean Absolute Error (MAE)* and the *Mean Squared Error (MSE)* between the ABM mean and the mean of PDMM and spatial-hybrid model. Comparing Fig. 5b and Fig. 5c, it can be seen that region 2 is the main driver of the infection dynamics i.e., most infectious agents are in Ω_2. Fig. 5a shows that in the focus region the spatial-hybrid model approximates the ABM outcomes much better than the PDMM as the percentiles of the hybrid model - in contrast to the PDMM percentiles - almost completely overlap the ABM percentiles. Furthermore, all error metrics calculated on the mean of simulation outcomes are smaller for the spatial-hybrid model. As Ω_2 is not the focus region, this region is modeled with the PDMM in the spatial-hybrid model. Therefore the PDMM and hybrid model curves for Ω_2,

especially the percentiles, have a similar shape, however the hybrid model has smaller errors than the PDMM, see Fig. 5c. For the sake of completeness, the results for region 3 and region 4 are shown in Fig. E.17.

The runtime of the spatial-hybrid model for the chosen setup is dominated by the ABM, but is still one order of magnitude smaller than the ABM runtime for 400 agents and even two orders of magnitude for 16, 000 agents, see Fig. 6a. For 40, 000 agents, the runtime for the described scenario can be reduced by 98% using the spatial-hybrid model.

In addition to the scaling behavior dependent on the number of agents, we consider the runtime scaling with respect to the number of initially infected and the value of the transmission rate $\rho^{(2)}$ in region 2 with, otherwise, the same setup as above. The runtime of ABM and spatial-hybrid model increases with an increasing proportion of initially infected while the runtime of the PDMM is independent of the proportion of initially infected, see Fig. 7a. The mean time of the ABM increases by 30% when increasing the initially infected from 0.1% to 1% and by nearly 50% when increasing from 0.1% to 10% initially infected. The mean time of the spatial-hybrid model increases by 7% and 42% when increasing the initial number of infected from 0.1% to 1 and 10%, respectively. For an increasing value of $\rho^{(2)}$, a similar behavior can be observed; see Fig. 7b. The mean runtime of the ABM rises by 175% when increasing $\rho^{(2)}$ from 0.1 to 0.2 and by 38% when increasing $\rho^{(2)}$ from 0.2 to 0.4. The decreasing gradient can be explained by the increasing number of recovered and dead agents when choosing $\rho^{(2)} = 0.4$. For recovered and dead agents, no infection state adoptions - needing computations and thus runtime - happen any more. The same behavior can be observed for the spatial-hybrid model with an even earlier saturation. The PDMM runtime is again not influenced by a varying value of $\rho^{(2)}$.

3.1.2. Application 2 - A potential for Munich and its surrounding counties

Secondly, we consider a less theoretical example, in particular a potential for the city of Munich, defined as our focus region Ω_1, and its surrounding counties.

To realize a difference between intercounty and intracounty mobility patterns, we define a potential on the political map of the considered region. We apply a Gaussian curve on the discretized county borders (in pixels) to define a potential $F(x)$. More precisely, $F(x) = h$ if $x \in \mathbb{R}$ is on a border, $F(x) \in (0, h)$ modeled through a Gaussian curve if x is in a given distance from the border and $F(x) = 0$ otherwise, i.e., x inside a county. The resulting potential directly determines eight regions for the PDMM, aligning with the eight counties considered. Fig. 8 shows the potential with a zoomed in snipped to the gradient on the border.The noise term σ and the border height h determine the number of agents crossing a border. A drawback of modeling agent movements with potential and diffusion process is the limited control on (intercounty) mobility and, in particular, that mostly agents next to the borders transition to neighboring regions in an appropriate time. However, in reality individuals that are further away from the borders should also transition to neighboring counties in a moderate time, e.g., on a daily basis caused by commuting activities. To that end, we modify the agents' movements. For short-distance travel, we use the diffusion process given by Eq. (5), where we set the border height h for the potential F to 10 and additionally prevent crossing borders through random walks. For long-distance travel, we introduce two normally distributed random variables T_d^α and T_r^α as additional attributes for agent α. Samples of these variables are only used when α is a commuter on the current day $\lfloor t \rfloor$. We use the random variable $T_d^\alpha \sim \mathcal{N}\left(\mu_d, \sigma_d^2\right)$ as the time point an agent α leaves its home county on day $\lfloor t \rfloor$ and $T_r^\alpha \sim \mathcal{N}\left(\mu_r, \sigma_r^2\right)$ as the time point an agent α returns on day $\lfloor t \rfloor$ – ignoring and resampling values beyond three standard deviations to ensure that samples are in $[0, 1]$. We introduce a commuting term $K(t)$, which is added onto L from Eq. (1), allowing us to better control the desired number of border crossings. Let $K_\alpha(t)$ be the α-th component of $K(t)$. This entry is zero if agent α is not a commuter on day $\lfloor t \rfloor$. If α is a commuter, $K_\alpha(t)$ is given by

$$K_\alpha(t) = \begin{cases} \pi(x_\alpha) - x_\alpha, & \text{if } t_d^\alpha + \lfloor t \rfloor \in (t, t + \delta t) \vee t_r^\alpha + \lfloor t \rfloor \in (t, t + \delta t) \\ 0, & \text{else} \end{cases}, \tag{13}$$

where $\lfloor t \rfloor$ marks the beginning of the day and t_d^α and t_r^α are realizations of T_d^α and T_r^α. Furthermore, $\pi: \Omega \rightarrow \Omega$ is a random variable on the domain using the probability function
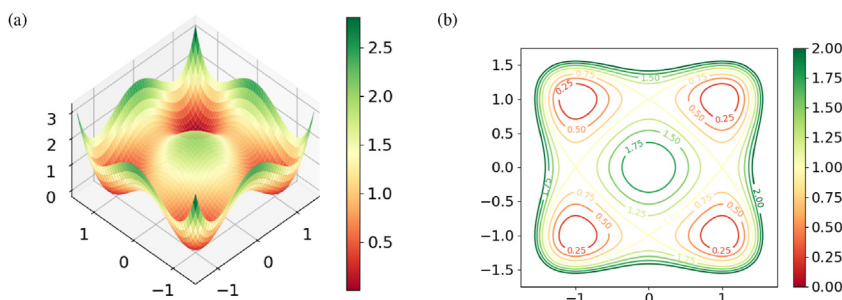

Fig. 3. Quadwell potential Eq. (11) used for the diffusion process in the ABM. The four metaregions are separated by the axes $x = 0$ and $y = 0$. The figure shows an (a) isometric view and (b) a contour plot of the potential.
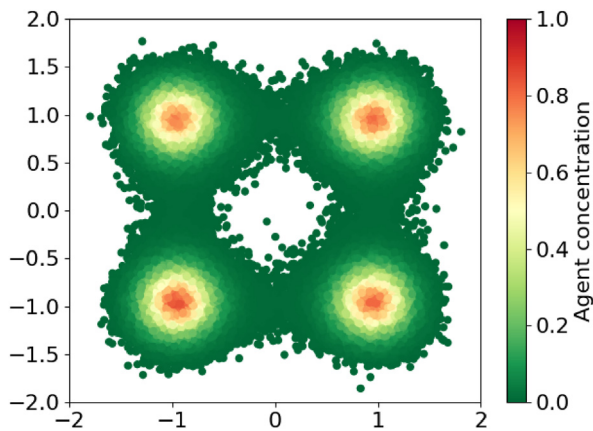
Fig. 4. Position distribution for a simulation of 800 agents for 50 days in the quadwell scenario for noise term $\sigma = 0.55$. Agents are initialized with positions having 0.3 distance from the axes $x = 0, \pm 2$ and $y = 0, \pm 2$.
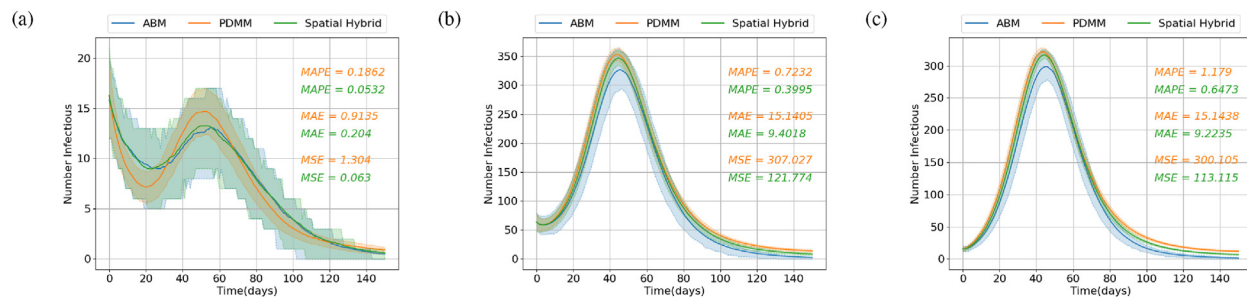

Fig. 5. Spatial hybridization for the quadwell potential: Focus region, sum of all regions and Region 2. Number of infectious agents (compartments C and I) for (a) the focus region Ω_1 (b) the sum of all regions and (c) Region 2. The figures show the mean outcomes in solid lines with a partially transparent face between the p25 and p75 percentiles from 500 runs. Additionally, MAPE, MAE, and MSE between the ABM mean and the mean of PDMM and spatial-hybrid model are displayed.
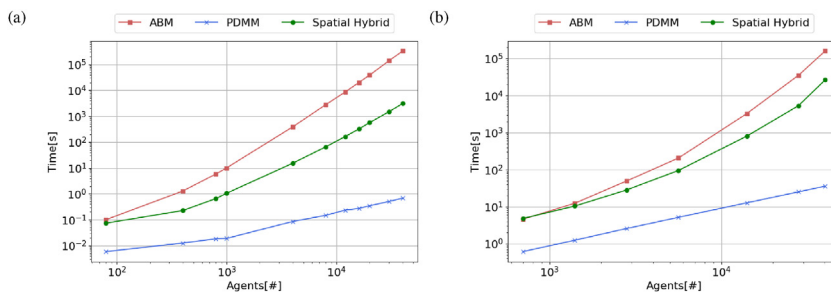
Fig. 6. Log-scaled runtime (in seconds) for ABM, PDMM and spatial-hybrid model. Shown is the mean runtime of 56 runs (one compute node) for (a) the quadwell example with the setup according to Table D.8, Table D.9 and (b) the Munich example with the setup according to Table D.8, Table D.10.

$$\mathbb{P}(\pi(x_\alpha) = y) = \mathbb{P}(y|x_\alpha \in \Omega_k; y \in \Omega_l) \cdot \lambda^{(k,l)} = \begin{cases} \dfrac{\lambda^{(k,l)}}{\nu(\Omega_l)}, & l \neq k \\ \lambda^{(k,k)}, & l = k \wedge x_\alpha = y \\ 0, & l = k \wedge x_\alpha \neq y \end{cases} \tag{14}$$

where $\lambda^{(k,l)}$ is the relative number of commuters from region Ω_k to region Ω_l and ν a counting measure on the discretized map (see Fig. 8) ignoring border pixels. Hence, $\nu(\Omega_l)$ is the number of pixels inside region Ω_l. When using this adapted movement
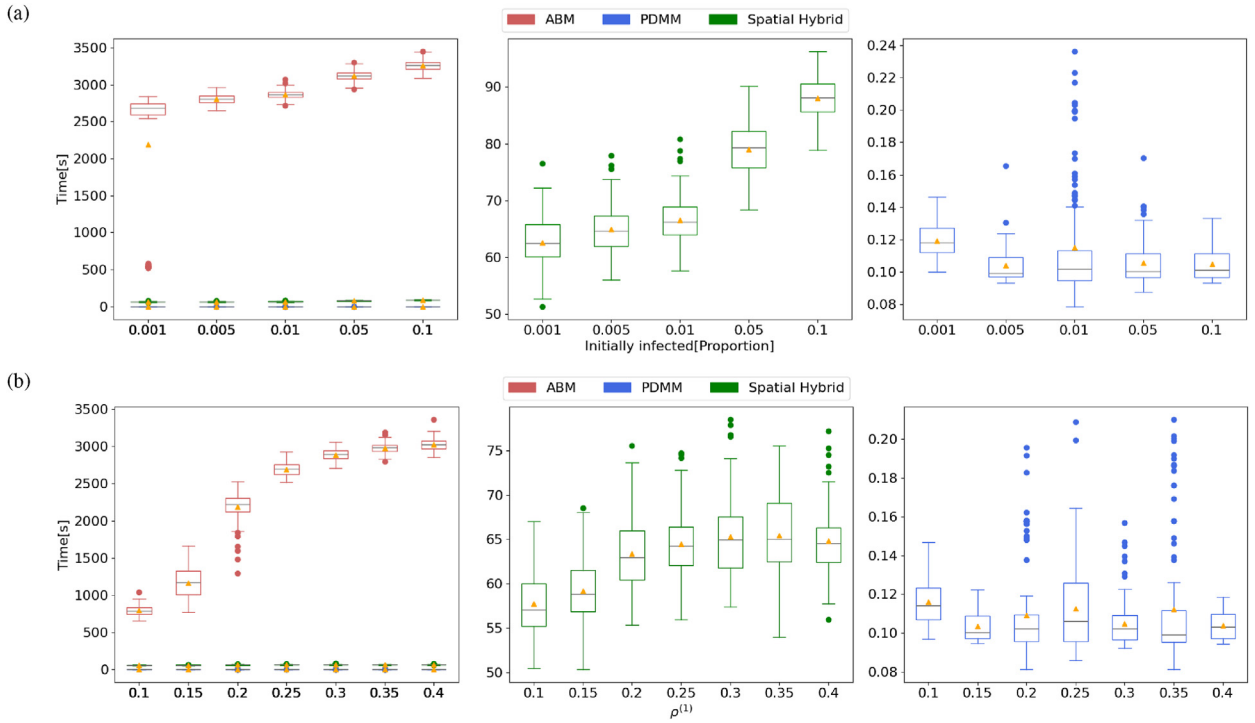
(a)

(b)

Fig. 7. Runtime distribution for ABM, PDMM and spatial-hybrid model for a varying proportion of initially infected and varying values of $\rho^{(2)}$. Shown are the runtimes of 112 runs for (a) five different values for the proportion of initially infected with the corresponding proportion distributed equally to compartments E, C and I and (b) seven different values for the transmission rate in Ω_2. In (a) $\rho^{(2)} = 0.3$ for all runs and in (b) 1% of the population (0.2% Exposed, 0.3% Carrier and 0.5% Infected) is initially infected in all runs.

for the ABM, we need to adapt the exchange of agents from PDMM to ABM for the spatial-hybrid model as well. When we exchange an agent α from PDMM to ABM, we (uniformly) sample a new position $x_\alpha^{(new)}$ for it in the focus region Ω_1, hence, $\mathscr{P}$ has the probability function

$$\mathbb{P}\left(x_\alpha^{(new)} = y \in \Omega_1\right) = \frac{1}{\nu(\Omega_1)}. \tag{15}$$

In addition to sampling the new position in the focus region, we decide whether the agent has started its commute or is returning from it, depending on the current time t. In case of the former, i.e., if $t \leq \lfloor t \rfloor + \mu_d + z_{99.5} \cdot \sigma_d$ with $z_{99.5} = 2.5758$ the 99.5% quantile of the standard normal distribution, a return time $t_r^\alpha \sim \mathcal{N}(\mu_r, \sigma_r^2)$ is drawn.

We expect commuters to leave their home county on average at 9 a.m. with the 99.5% quantile between 5 a.m. and 1 p.m., resulting in $\mu_d = \frac{9}{24}$ and $\sigma_d = \left(\frac{13}{24} - \mu_d\right)/z_{99.5} = 0.0647$. Further, we expect commuters to return on average at 6 p.m. with the 99.5% quantile between 1 p.m. and 11 p.m., hence $\mu_r = \frac{18}{24}$ and $\sigma_d = \left(\frac{23}{24} - \mu_r\right)/z_{99.5} = 0.0809$. The spatial transition rates of the PDMM are equal to the commute weights $\lambda^{(k,l)}$, $k, l = 1, \ldots, 8$ used in Eq. (14). Their values are based on (Kühn et al., 2021) and are given in Eq. (C.1). Fig. 9 shows that through the incorporation of the commuting term K, the number of daily transitions in Munich City matches the static daily commuting data from (BMAS, 2020) used in (Kühn et al., 2021), and scaled according to the number of agents n_a.

For this application, we chose the same transmission rate for all regions with $\rho^{(k)} = 0.2$ corresponding to $R_0 = 1.6$. All parameters used for the Munich potential are summarized in Table D.4 and Table D.6. We chose $n_a = 14{,}064$ agents with initially infected (0.05% Exposed, 0.05% Carrier, and 0.1% Infected) only in Ω_5, see Fig. E.18. In all other regions, all agents are susceptible at the beginning. We run 500 simulations of ABM, PDMM and spatial-hybrid model for $t_{max} = 100$ days, see Fig. 10. We again consider MAPE, MAE, and MSE of the mean outcomes as error metrics. The MAPE is not defined for the focus region because the ABM mean is zero on the first few days, see Fig. 10. In general, the spatial-hybrid model and PDMM have both low errors. In the focus region the MAE of the PDMM is even slightly lower than the one for the spatial-hybrid model. However, the MSE of the spatial-hybrid model is lower and the percentiles also fit the ABM percentiles better than the PDMM percentiles. Yet, the p25 percentile of the ABM displays that in at least 25% of all runs, the virus dies out which is not captured by neither spatial-hybrid model nor PDMM. The reason for this is that ODE-based models are not able to let a virus go completely
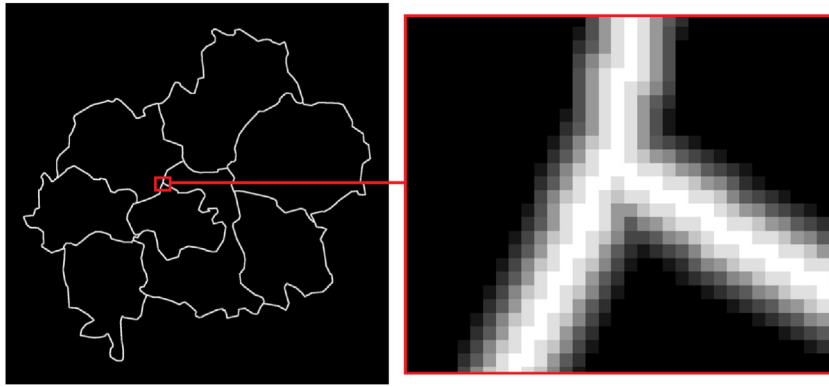
Fig. 8. Potential defined for the ABM in and around Munich. Black areas have value zero and white parts value one. We obtain the potential *F* by discretizing the map, resulting in a matrix $P \in [0,1]^{p \times p}$ and then interpolating between matrix entries.

extinct and as the region with initially infected is not the focus region, it is modeled with the PDMM in the spatial-hybrid. To capture this behavior, a temporal-hybrid model could be used, see Section 3.2.

The runtime gain for the Munich potential is smaller than for the quadwell potential but the spatial-hybrid model is still 6-times faster than the ABM for 40,000 agents, which is a runtime gain of more than 80%; see Fig. 6b. The reason for this reduced effect is the very disadvantageous setup for the hybrid model with Munich City as focus region, whose population makes up approximately 75% of all agents in the modeled counties. Consequently, in the spatial-hybrid model, the number of agents modeled with the ABM is roughly 75% of n_a while in the quadwell potential it is only 25% of all agents. Secondly, the chosen initialization also has a relevant influence on the runtime gain. As we have an overall low number of infectious agents in the simulation (on average at most 1.5% of the total population compared to approximately 4% in the quadwell), there are fewer infection state adoptions in the ABM per time step, which leads to less computational costs. One can observe that the runtime for the spatial-hybrid model for Munich is even slightly higher than the ABM runtime for $n_a < 700$. Until that value, the exchange of agents from PDMM to ABM - which is slightly more complex for the Munich potential compared to the quadwell potential - produces overhead which only amortized for $n_a > 700$.

We furthermore considered the scaling behavior with respect to the number of initially infected in region 5, Munich Land, and a changing value for $\rho^{(k)}$ with otherwise the same setup as above. The corresponding results are provided in Fig. E.20 and align with the corresponding results from the quadwell potential.

3.2. Temporal hybridization

For the temporal hybridization we consider a single well potential given by

$$F(x,y) = \frac{x^4 + y^4}{2};$$ (16)

see Fig. 11. By design we only have one region and the PDMM reduces to a classical ODE-based model without spatial resolution. We use a constant noise $\sigma = 0.5$ for the diffusion process and a relatively small contact radius $r = 0.1$, thus σ has a crucial influence on contacts between agents in the ABM and consequently on second-order adoptions. All parameters and their values for the single well application are listed in Table D.4 and Table D.7.

We chose a scenario with exactly one initially infected agent in compartment *I* with an otherwise susceptible population. The transmission rate $\rho^{(1)} = 0.6$ leading to $R_0 = 4.8$ is set such that in around 30% of all simulations, the virus dies out. Otherwise, we get an epidemic outbreak and introduce a nonpharmaceutical intervention (NPI) which is implemented through a reduction of $\rho^{(1)}$ by $\kappa = 80\%$ after $t_{NPI} = 0.5(t_0 + t_{\max})$. This NPI then leads to a slow decline of infections over the remaining simulation time.

Since we are considering only one region, the derived PDMM is a deterministic ODE-model in which the virus never dies out. Obviously, we cannot use this PDMM to meaningfully simulate this specific scenario in its entirety. However, we can use the PDMM to speed up calculation once an outbreak reaches a certain size or after the virus has become extinct. Therefore, the temporal-hybrid model always starts with the ABM, and then switches to the PDMM once an extinction or an outbreak reaches a certain size. Note that after extinction, we could even stop the simulation but we continue with the ODE-based model until $t_{\max}$. This is done for practical reasons of comparison and since the runtime of this model is negligible. The switch is defined by a threshold $s > 1$ to be compared against the sum of agents in infection states *E*, *C* and *I*. If this number is larger than *s*, we switch from ABM to PDMM, projecting the whole population of the ABM. A suitable value for the threshold *s* depends on the specific setup, especially on the transmission rate ρ. It is not trivial to be determined, particularly for real-world scenarios.
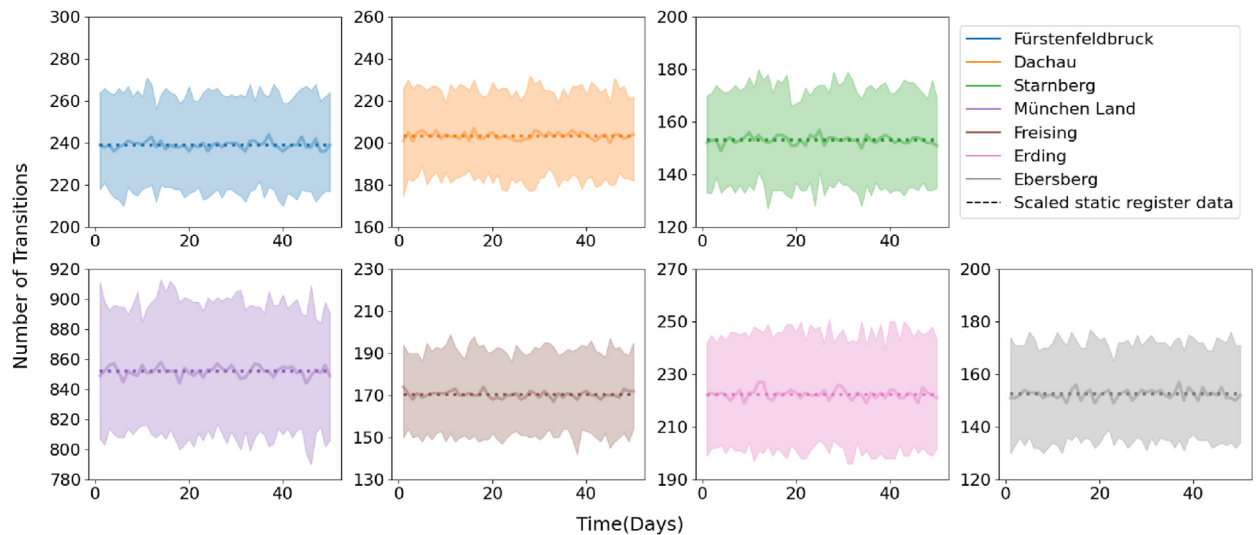
Fig. 9. Daily commuting from and to Munich City for different surrounding counties. 5% and 95% percentiles of daily transitions in simulations shown as colored area together with median as solid line. The static register data from Federal Agency of Work, scaled to number of simulated agents, is shown as dotted lines.

A second threshold $0 \leq s' < s$ could be introduced to switch from PDMM to ABM as soon as the number of infected agents falls below s'. This second threshold indicates a number of infected, where stochasticity is (again) important to determine the development of the disease dynamics. This can be of particular importance if disease mitigation, e.g., through NPIs, had been conducted successfully. When switching from PDMM to ABM, agents' positions and infection states have to be sampled according to the current PDMM system state. The infection state of an agent can be sampled according to the distribution given by the PDMM subpopulations $N_i^{(k)}$. The position is sampled from a distribution $\mathscr{P}$ which could for the considered single well potential be a uniform distribution on $[-1, 1] \times [-1, 1]$, see Fig. E.21. In the presented scenario, we set s' to zero.

Here, the model selection function Γ from Algorithm 2 returns the PDMM if the number of infected agents exceeds s, and the ABM otherwise. We chose $n_a = 10{,}000$, $t_0 = 0$, $t_{\max} = 40$, and a fixed step size of $\Delta t = 0.2$ for this Algorithm. The single initially infected agent is chosen randomly from the entire population. Fig. 12 shows the results of 10,000 simulations of ABM, PDMM, and the temporal-hybrid model with $s = 2$ and $s = 5$.

The reduction of $\rho^{(1)}$ at $t_{NPI} = 20$ leads to a kink in the number of infected agents (compartments E, C and, I), see Fig. 12a. Since in 30% of all ABM simulations the virus dies out, the 25th-percentile is almost entirely zero, except for right after the initialization. We can also see that the percentiles as well as the means of the temporal-hybrid models with $s = 2$ and $s = 5$ are close together and the errors of both hybrid models are lower than the errors of the PDMM. As expected, the temporal-hybrid model with threshold $s = 5$ has even lower errors than the one with threshold $s = 2$. As the PDMM is deterministic, the percentiles and means of it are all the same and have a MAPE of 0.0247 for all runs. While Fig. 12a shows the results for all simulations, Fig. 12b shows the results for all simulations where the virus becomes extinct and Fig. 12c the corresponding curves for all simulations where the virus survives, hence, an outbreak occurs. The PDMM is not shown in Fig. 12b as it is not able to capture the extinction scenario. In Fig. 12c, both temporal-hybrid models and ABM curves have shifted upwards compared to the results of all simulations in Fig. 12a. The PDMM result remains trivially the same and therefore underestimates the mean number of infected in the survival scenarios. For completeness, we also provided the results for the later compartment I, see Fig. E.22, where we see a rather smooth transition compared to the kink in Fig. 12.

Both switching thresholds seem to be a good choice as the resulting temporal-hybrid models are very close to the ABM results, i.e., they have low errors - the mean as well as the percentiles. Note, however, that the low thresholds are due to a large transmission rate $\rho^{(1)}$ and dense contact network realized through the diffusion-drift process. The temporal-hybrid model with $s = 2$ does not capture the increase of the ABM mean in the extinction runs until t_{NPI}, see Fig. 12b while the hybrid model with $s = 5$ does.

The ABM runs significantly faster in extinction runs than in survival runs (see Table 3), as there are no pairwise interactions to consider once there are no infected agents anymore. Furthermore, the PDMM is - as expected - multiple orders of magnitude faster than the ABM, making it desirable to use it whenever possible. On average, the temporal-hybrid model for $s = 5$ is 20-times and the one for $s = 2$ even 40-times faster than the ABM in the results for all simulations (see Table 2), 8-to 9-times for the extinction simulations (see Table 3), and 22- ($s = 5$) and 49-times ($s = 2$) for the survival simulations (see Table 3). When comparing the total simulation time for the different switching thresholds, all simulations show a time gain when using the lower threshold of $s = 2$ while for the survival simulations the time gain is the greatest. For all stochastic models (ABM and temporal-hybrid), the difference in minimum and maximum time to the mean time correlates with the large
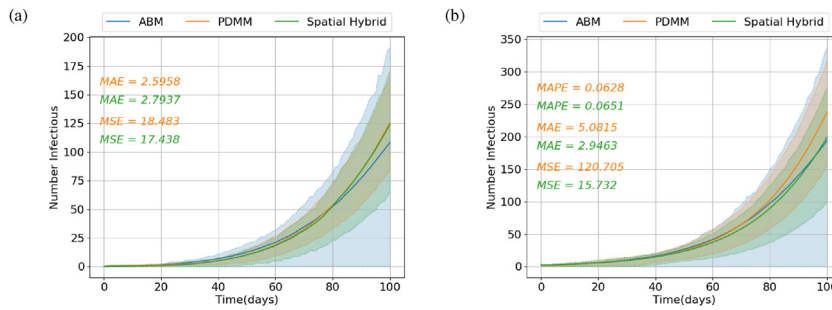
Fig. 10. Spatial hybridization for the Munich potential: Focus region and sum of all regions. Number of infectious agents (compartments *C* and *I*) for (a) Munich City (focus region) and (b) the sum of all regions. The figures show the mean outcomes in solid lines with a partially transparent face between the p25 and p75 percentiles from 500 runs. If possible, MAPE, MAE, and MSE between the ABM mean and the mean of PDMM and spatial-hybrid model are displayed. Since the number of infected agents in the focus region is zero at the beginning, the MAPE cannot be calculated for the left figure.
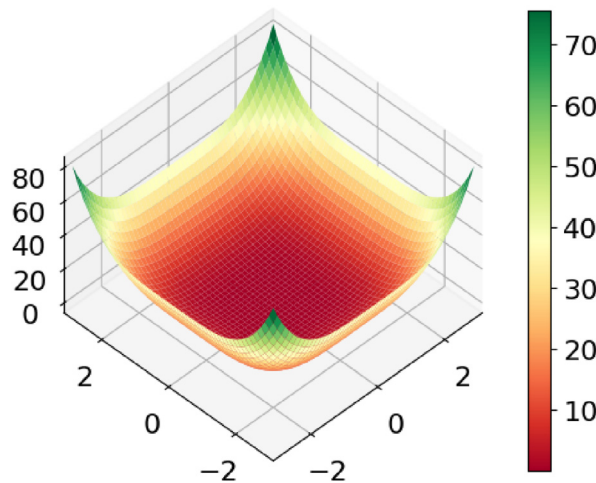

Fig. 11. Single well potential Eq. (16) used for the diffusion process in the ABM of the temporal-hybrid model.

variation in simulation outcomes; see the percentiles in Fig. 12. For this application, we do not further examine the runtime scaling behavior, as it does not behave significantly different to the results for the spatial-hybrid model, see Section 3.1 and Fig. 6. The runtime of the temporal-hybrid model is dominated by the proportion of time (instead of population) in which the ABM is used for the simulation.

3.3. Energy consumption

To quantify the effects of the proposed hybrid models on the estimated CO_2 emissions, we measured the energy consumption of the ABMs and the hybrid models for the chosen setups in the quadwell application Section 3.1.1,Fig. 5, the Munich application Section 3.1.2,Fig. 10 and the single well application Section 3.2,Fig. 12. We used *LIKWID* (Treibig, Hager, & Wellein, 2010) to measure the energy consumption in Joules for each model using 56 cores in parallel, see Table 4. According to extrapolations from the German Federal Environment Agency (Icha & Lauf, 2024), 1 kWh power produced 380g CO_2 emissions in 2023. Using this extrapolation, we calculated the CO_2 savings achieved for all applications by replacing a full ABM through a hybrid ABM-ODE model. For the quadwell application, the spatial-hybrid model emitted 1180g CO_2 less than the full ABM which is a saving of 97.75%. For the Munich potential, the savings are - like the runtime saving - less than for the quadwell potential, but we still reduced the CO_2 emissions by a factor of roughly 3.4. The reductions of the temporal-hybrid models for the single well application are nearly as high as of the spatial-hybrid model for the quadwell application with 94.67% ($s = 5$) and 97.28% ($s = 2$) corresponding to a saving of 2524g CO_2 and 2594g CO_2, respectively.

4. Discussion and conclusion

While the number of research articles on hybrid models for infectious disease dynamics in the sense of our understanding is very small, several authors have published pioneering works in (Bobashev et al., 2007; Hunter et al., 2020; Kasereka et al.,
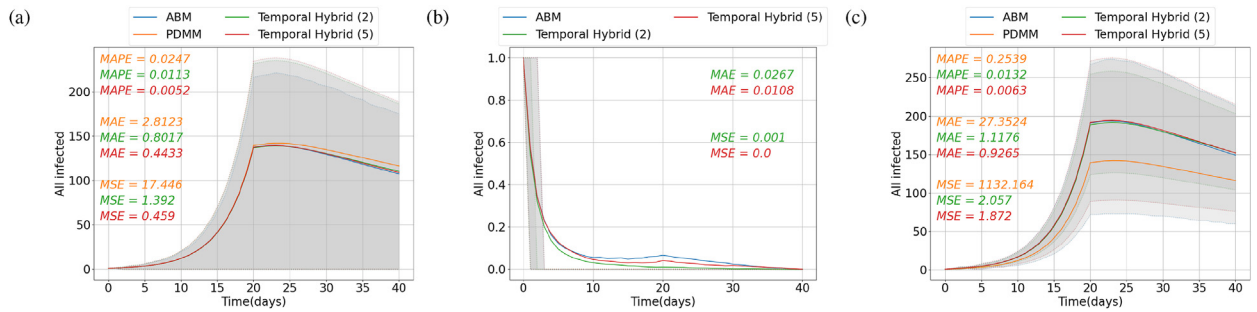
Fig. 12. Temporal hybridization for the single well example. Shown is the sum of compartments *E*, *C* and *I* for (a) all runs, (b) extinction runs with virus extinction, and (c) survival runs without virus extinction for ABM, PDMM and temporal-hybrid models with $s = 2$ and $s = 5$. The figures show the mean outcomes in solid lines with a partially transparent face between the p25 and p75 percentiles from 10,000 runs with 10,000 agents.

2014; Bradhurst et al., 2015; Yoneyama et al., 2012) over the last two decades. The approaches coming closest to our works are given in (Bobashev et al., 2007; Hunter et al., 2020). However, runtime and computational cost have, e.g., not been considered in (Bobashev et al., 2007). In (Hunter et al., 2020) various interesting results have been published for a model running on a laptop computer.

In this work, we formalized two hybridization approaches and introduced a general framework of spatial- and temporal-hybrid models, combining agent-based and ODE- or metapopulation-based models. In particular, the temporal-hybrid model with its dynamic model choice is also an *adaptive* model and future works may also consider adaptivity in time and space through spatio-temporal-hybrid models.

We presented two suitable models recently introduced, which were adapted for our examples and combined in the suggested manner. The examples provided show the application of the proposed hybridization framework in different theoretical settings, thus demonstrating its runtime benefit while maintaining a high resolution in a specific area or time frame. Although the parameters used in the examples are motivated by COVID-19, the application is not limited to a particular pathogen. Moreover, since this paper is intended to serve as a proof of concept, we did not consider a true real-world application. The critical aspect of combining the two model types is the definition of exchange rules, an issue that has not been discussed a lot yet, but which has an essential impact on the accuracy and the performance of the hybridization. Models that are derived from one another in a theoretical way, like in (Schmidtchen, Tse, & Wackerle, 2018; Winkelmann et al., 2021) offer a more direct way to implement exchanges. Nevertheless, nuances in projections or mappings can become important as it has been observed in our simulations. With the presented random walk movement of agents, the same agent may cross region boundaries multiple times in a short time, which might drastically change the hybrid model's outcome compared to the outcome of the fine model if this effect is not accounted for in the coarse model.

The movement pattern of the ABM, given by a potential and a Brownian motion is a rather theoretical movement pattern that delivered us two suitable models to be combined, but, especially when wanting to develop an actual real-world application, a more realistic human mobility pattern or an ABM using contact networks should be considered. Additionally, the transmission model of the ABM could be enhanced by incorporating more features such as an individual viral load. In future work, we will investigate more complex ABMs realizing discrete locations such as households, schools, and workplaces. For these, more sophisticated methods for projection and, also, reconstruction of information will be needed. For instance, human-to-human contact networks or contact tracing used to mitigate virus dynamics suffer from loss of information when exchanging between models of different granularity.

As shown, hybrid models can combine the best of both worlds and deliver fine-granular insights through the use of ABMs, using substantially reduced computational resources. Both hybridization approaches were able to capture disease dynamics of the ABM better than the PDMM.

For the spatial hybridization, the runtime gain depends on the total number of agents in the focus region, relative to all other regions modeled by metapopulations. Similarly, the temporal hybridization runtime gain depends mostly on the proportion of simulation time that the ABM is used. From our simulations, we have seen that up to 98 % of the simulation time

Table 2
Summarized simulation timings (in sec) for ABM, PDMM and temporal-hybrid models for all 10,000 simulations.

Model	All simulations		
	min	mean	max
ABM	41.80	331.76	1835.93
PDMM	0.00040	0.00047	0.01140
Temporal-hybrid $s = 2$	0.5587	8.4082	92.6604
Temporal-hybrid $s = 5$	0.5591	16.6405	104.497

Table 3
Summarized simulation timings (in sec) for ABM, PDMM and temporal-hybrid models for simulations with and without virus extinction.

Model	Extinction simulations			Survival simulations		
	min	mean	max	min	mean	max
ABM	41.80	48.77	109.147	91.07	442.20	1835.93
PDMM	–	–	–	0.00040	0.00047	0.01140
Temporal-hybrid s = 2	0.5587	5.4414	85.2773	0.8115	9.5341	92.6604
Temporal-hybrid s = 5	0.5591	6.2384	99.2354	2.9843	20.7604	104.497

Table 4
Energy consumption of all models. The measurements for the quadwell potential represent 500 runs as in Section 3.1.1, the measurements for the Munich potential represent 500 runs as in Section 3.1.2 and the measurements for the single well potential represent 10, 000 runs as in Section 3.2. Power in kilo-Watt-hours (kWh) computed from measured energy and CO_2 emissions[g] computed from (Icha & Lauf, 2024). Reduction (in percent) computed for hybrid model compared to the full, reference ABM.

Model	Power[kWh]	CO_2 emissions[g]	CO_2 reduction
ABM quadwell	3.17889	1207.97915	–
Spatial-hybrid quadwell	0.07162	27.21423	97.75%
ABM Munich	4.15400	1578.51937	–
Spatial-hybrid Munich	1.20890	459.38063	70.90%
ABM single well	7.01801	2666.84380	–
Temporal-hybrid (5) single well	0.37380	142.04358	94.67%
Temporal-hybrid (2) single well	0.19095	72.56005	97.28%

can be saved by replacing an ABM by a spatial-hybrid model. Although this number depends highly on the chosen models and the individual setting, a reduction of 90 % of the simulation time or a speedup of 10 can be achieved easily.

CRediT authorship contribution statement

Julia Bicker: Writing − review & editing, Writing − original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **René Schmieding:** Writing − review & editing, Writing − original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Michael Meyer-Hermann:** Writing − review & editing, Validation, Supervision, Resources, Funding acquisition. **Martin J. Kühn:** Writing − review & editing, Writing − original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Data availability

The MEmilio code repository is publicly available on Github under https://github.com/SciCompMod/memilio, the developed models are to be found on the fork release https://github.com/reneSchm/memilio/releases/tag/hybrid-paper-v2 (based on the branch hybrid-model-omp); The models can be found in the directory/cpp/hybrid_paper, with simulation algorithms under/cpp/models/mpm. All input data can be generated with or is in the code repository. Visualized simulation data can be shared upon request.

Declaration competing interest

All authors declare to not have any conflict of interest.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.idm.2024.12.015.

References

Angione, C., Silverman, E., & Yaneske, E. (2022). Using machine learning as a surrogate model for agent-based simulations. *PLoS One, 17*(2), Article e0263150. https://doi.org/10.1371/journal.pone.0263150. publisher: Public Library of Science https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0263150.

Avegliano, P., & Sichman, J. S. (2023). Equation-based versus agent-based models: Why not embrace both for an efficient parameter calibration? *The Journal of Artificial Societies and Social Simulation, 26*(4), 3. https://doi.org/10.18564/jasss.5183. URL https://www.jasss.org/26/4/3.html.

Banos, A., Corson, N., Gaudou, B., Laperrière, V., & Coyrehourcq, S. (2015). The importance of being hybrid for spatial epidemic models:A multi-scale approach. *Systems, 3*(4), 309–329. https://doi.org/10.3390/systems3040309. URL https://www.mdpi.com/2079-8954/3/4/309.

Bershteyn, A., Gerardin, J., Bridenbecker, D., Lorton, C. W., Bloedow, J., Baker, R. S., et al. (2018). For the Institute for Disease Modeling, Implementation and applications of EMOD, an individual-based multi-disease modeling platform. *Pathogens and Disease, 76*(5). https://doi.org/10.1093/femspd/fty059. URL https://academic.oup.com/femspd/article/doi/10.1093/femspd/fty059/5050059.

BMAS. (2020). Pendlerverflechtungen der sozialversicherungspflichtig Beschäftigten nach Kreisen - Deutschland (Jahreszahlen), publication Title. www.bmas.de. https://statistik.arbeitsagentur.de/SiteGlobals/Forms/Suche/Einzelheftsuche_Formular.html?topic_f=beschaeftigung-sozbe-krpendd.

Bobashev, G. V., Goedecke, D. M., Yu, F., & Epstein, J. M. (2007). A hybrid epidemic model: Combining the advantages of agent-based and equation-based approaches. In *2007 winter simulation conference* (pp. 1532–1537). Washington, DC, USA: IEEE. https://doi.org/10.1109/WSC.2007.4419767. URL http://ieeexplore.ieee.org/document/4419767/.

Bracher, J., Wolffram, D., Deuschel, J., Görgen, K., Ketterer, J. L., Ullrich, A., et al. (2021). A pre-registered short-term forecasting study of COVID-19 in Germany and Poland during the second wave. *Nature Communications, 12*(1), 5173. https://doi.org/10.1038/s41467-021-25207-0. URL https://www.nature.com/articles/s41467-021-25207-0.

Bradhurst, R. A., Roche, S. E., East, I. J., Kwan, P., & Garner, M. G. (2015). A hybrid modeling approach to simulating foot-and-mouth disease outbreaks in Australian livestock. *Frontiers in Environmental Science, 3*. https://doi.org/10.3389/fenvs.2015.00017. URL http://www.frontiersin.org/Environmental_Informatics/10.3389/fenvs.2015.00017/abstract.

Callaghan, T. V., Bergholm, F., Christensen, T. R., Jonasson, C., Kokfelt, U., & Johansson, M. (2010). A new climate era in the sub-Arctic: Accelerating climate changes and multiple impacts. *Geophysical Research Letters, 37*(14), Article 2009GL042064. https://doi.org/10.1029/2009GL042064. URL https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2009GL042064.

Collier, N., & North, M. (2013). Parallel agent-based simulation with repast for high performance computing. *SIMULATION, 89*(10), 1215–1235. https://doi.org/10.1177/0037549712462620. publisher: SAGE Publications Ltd STM.

Daszak, P., Amuasi, J., das Neves, C., Hayman, D., Kuiken, T., Roche, B., et al. (2020). Workshop report on biodiversity and pandemics of the intergovernmental platform on biodiversity and ecosystem services (IPBES). *Tech. rep.*. https://doi.org/10.5281/zenodo.7432079

Djurdjevac Conrad, N., Helfmann, L., Zonker, J., Winkelmann, S., & Schütte, C. (2018). Human mobility and innovation spreading in ancient times: A stochastic agent-based simulation approach. *EPJ Data Science, 7*(1), 24. https://doi.org/10.1140/epjds/s13688-018-0153-9. URL https://epjdatascience.springeropen.com/articles/10.1140/epjds/s13688-018-0153-9.

Epstein, P. R. (2001). Climate change and emerging infectious diseases. *Microbes and Infection*, 747–754, 3, 2001 © 2001 Éditions scientifiques et médicales Elsevier SAS. All rights reserved S1286457901014290/REV.

Garnett, G. P., Cousens, S., Hallett, T. B., Steketee, R., & Walker, N. (2011). Mathematical models in the evaluation of health programmes. *The Lancet, 378*(9790), 515–525. https://doi.org/10.1016/S0140-6736(10)61505-X

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry, 81*(25), 2340–2361. https://doi.org/10.1021/j100540a008

Hunter, E., Mac Namee, B., & Kelleher, J. D. (2017). A taxonomy for agent-based models in human infectious disease epidemiology. *The Journal of Artificial Societies and Social Simulation, 20*(3), 2. https://doi.org/10.18564/jasss.3414. URL http://jasss.soc.surrey.ac.uk/20/3/2.html.

Hunter, E., Mac Namee, B., & Kelleher, J. (2020). A hybrid agent-based and equation based model for the spread of infectious diseases. *The Journal of Artificial Societies and Social Simulation, 23*(4), 14. https://doi.org/10.18564/jasss.4421. URL http://jasss.soc.surrey.ac.uk/23/4/14.html.

Icha, P., & Lauf, T. (2024). Entwicklung der spezifischen Treibhausgas-Emissionen des deutschen Strommix in den Jahren 1990–2023. https://www.umweltbundesamt.de/publikationen/entwicklung-der-spezifischen-treibhausgas-10.

Kasereka, S., Kasoro, N., & Chokki, A. P. (2014). A hybrid model for modeling the spread of epidemics: Theory and simulation. In *2014 4th international symposium ISKO-maghreb: Concepts and tools for knowledge management (ISKO-maghreb)* (pp. 1–7). Algiers, Algeria: IEEE. https://doi.org/10.1109/ISKO-Maghreb.2014.7033457. URL http://ieeexplore.ieee.org/document/7033457/.

Kerkmann, D., Korf, S., Nguyen, K., Abele, D., Schengen, A., Gerstein, C., et al. (2024). Agent-based modeling for realistic reproduction of human mobility and contact behavior to evaluate test and isolation strategies in epidemic infectious disease spread. https://arxiv.org/abs/2410.08050.

Kerr, C. C., Stuart, R. M., Mistry, D., Abeysuriya, R. G., Rosenfeld, K., Hart, G. R., et al. (2021). Covasim: An agent-based model of COVID-19 dynamics and interventions. *PLoS Computational Biology, 17*(7), 1–32. https://doi.org/10.1371/journal.pcbi.1009149

Khailaie, S., Mitra, T., Bandyopadhyay, A., Schips, M., Mascheroni, P., Vanella, P., et al. (2021). Development of the reproduction number from coronavirus SARS-CoV-2 case data in Germany and implications for political measures. *BMC Medicine, 19*(1), 1–16, publisher: BioMed Central.

Koslow, W., Kühn, M. J., Binder, S., Klitz, M., Abele, D., Basermann, A., et al. (2022). Appropriate relaxation of non-pharmaceutical interventions minimizes the risk of a resurgence in SARS-CoV-2 infections in spite of the Delta variant. *PLoS Computational Biology, 18*(5), Article e1010054. https://doi.org/10.1371/journal.pcbi.1010054. URL https://dx.plos.org/10.1371/journal.pcbi.1010054.

Kühn, M. J., Abele, D., Binder, S., Rack, K., Klitz, M., Kleinert, J., et al. (2022). Regional opening strategies with commuter testing and containment of new SARS-CoV-2 variants in Germany. *BMC Infectious Diseases, 22*(1), 333. https://doi.org/10.1186/s12879-022-07302-9. URL https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-022-07302-9.

Kühn, M. J., Abele, D., Kerkmann, D., Korf, S., Zunker, H., Wendler, A., et al. (2023). MEmilio - a high performance Modular EpideMIcs simuLatIOn software. https://doi.org/10.5281/zenodo.10412635.

Kühn, M. J., Abele, D., Mitra, T., Koslow, W., Abedi, M., Rack, K., et al. (2021). Assessment of effective mitigation and prediction of the spread of SARS-CoV-2 in Germany using demographic information and spatial resolution. *Mathematical Biosciences*. , Article 108648. https://doi.org/10.1016/j.mbs.2021.108648. https://www.sciencedirect.com/science/article/pii/S0025556421000845

Kühn, M. J., Holke, J., Lutz, A., Thies, J., Röhrig-Zöllner, M., Bleh, A., et al. (2023). SIMD vectorization for simultaneous solution of locally varying linear systems with multiple right-hand sides. *The Journal of Supercomputing, 79*(13), 14684–14706.

Liu, J., Ong, G. P., & Pang, V. J. (2022). Modelling effectiveness of COVID-19 pandemic control policies using an Area-based SEIR model with consideration of infection during interzonal travel. *Transportation Research Part A: Policy and Practice, 161*, 25–47. https://doi.org/10.1016/j.tra.2022.05.003. URL https://linkinghub.elsevier.com/retrieve/pii/S0965856422001197.

Marilleau, N., Lang, C., & Giraudoux, P. (2018). Coupling agent-based with equation-based models to study spatially explicit megapopulation dynamics. *Ecological Modelling, 384*, 34–42. https://doi.org/10.1016/j.ecolmodel.2018.06.011. URL https://linkinghub.elsevier.com/retrieve/pii/S0304380018302163.

Medlock, J. P. (2004). *Integro-differential-equation models in ecology and epidemiology*. University of Washington.

Müller, S. A., Balmer, M., Charlton, W., Ewert, R., Neumann, A., Rakow, C., et al. (2021). Predicting the effects of COVID-19 related interventions in urban settings by combining activity-based modelling, agent-based simulation, and mobile phone data. *medRxiv*. https://doi.org/10.1101/2021.02.27.21252583

Müller, S. A., Paltra, S., Rehmann, J., Nagel, K., & Conrad, T. O. (2023). Explicit modeling of antibody levels for infectious disease simulations in the context of SARS-CoV-2. *iScience, 26*(9), Article 107554. https://doi.org/10.1016/j.isci.2023.107554. https://linkinghub.elsevier.com/retrieve/pii/S2589004223016310

Niemann, J.-H., Uram, S., Wolf, S., Djurdjevac Conrad, N., & Weiser, M. (2024). Multilevel optimization for policy design with agent-based epidemic models. *Journal of Computational Science, 77*, Article 102242. https://doi.org/10.1016/j.jocs.2024.102242. https://www.sciencedirect.com/science/article/pii/S1877750324000358

Pereira, R., Couto, M., Ribeiro, F., Rua, R., Cunha, J., Fernandes, J. P., et al. (2017). Energy efficiency across programming languages: How do energy, time, and memory relate?. In *Proceedings of the 10th ACM SIGPLAN international conference on software language engineering* (pp. 256–267). Vancouver BC Canada: ACM. https://doi.org/10.1145/3136014.3136031. URL https://dl.acm.org/doi/10.1145/3136014.3136031.

Portegies Zwart, S. (2020). The ecological impact of high-performance computing in astrophysics. *Nature Astronomy, 4*(9), 819–822. https://doi.org/10.1038/s41550-020-1208-y. publisher: Nature Publishing Group https://www.nature.com/articles/s41550-020-1208-y.

Quang Nghi, H., Nguyen-Huu, T., Grignard, A., Xuan Huynh, H., & Drogoul, A. (2016). Toward an agent-based and equation-based coupling framework. In P. C. Vinh, & L. Barolli (Eds.), *Nature of computation and communication* (Vol. 168, pp. 311–324). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-46909-6_28. series Title: Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering http://link.springer.com/10.1007/978-3-319-46909-6_28.

Reiner, R. C., Barber, R. M., Collins, J. K., Zheng, P., Adolph, C., Albright, J., Antony, C. M., Aravkin, A. Y., Bachmeier, S. D., Bang-Jensen, B., & IHME COVID-19 Forecasting Team. (2020). Modeling COVID-19 scenarios for the United States. *Nature Medicine*, 1–12. https://doi.org/10.1038/s41591-020-1132-9

Robertson, C., Safta, C., Collier, N., Ozik, J., & Ray, J. (2024). Bayesian calibration of stochastic agent based model via random forest. http://arxiv.org/abs/2406.19524.

Rofouei, M., Stathopoulos, T., Ryffel, S., Kaiser, W., & Sarrafzadeh, M. (2008). Energy-aware high performance computing with graphic processing units. In *Proceedings of the 2008 conference on Power aware computing and systems, HotPower'08* (p. 11). USA: USENIX Association.

Schaller, G., & Meyer-Hermann, M. (2006). Continuum versus discrete model: A comparison for multicellular tumour spheroids. *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences, 364*(1843), 1443–1464. https://doi.org/10.1098/rsta.2006.1780. URL https://royalsocietypublishing.org/doi/10.1098/rsta.2006.1780.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology, 1*(2), 143–186. https://doi.org/10.1080/0022250X.1971.9989794

Schmidt, A., Zunker, H., Heinlein, A., & Kühn, M. J. (2024). Towards graph neural network surrogates leveraging mechanistic expert knowledge for pandemic response. https://arxiv.org/abs/2411.06500.

Schmidtchen, M., Tse, O. T. C., & Wackerle, S. (2018). A multiscale approach for spatially inhomogeneous disease dynamics. *Communication in Biomathematical Sciences, 1*(2), 62–87. https://doi.org/10.5614/cbms.2018.1.2.1. number: 2 https://journals.itb.ac.id/index.php/cbms/article/view/5768.

Sewall, J., Wilkie, D., & Lin, M. C. (2011). Interactive hybrid simulation of large-scale traffic. In *Proceedings of the 2011 SIGGRAPH asia conference* (pp. 1–12). Hong Kong China: ACM. https://doi.org/10.1145/2024156.2024169. URL https://dl.acm.org/doi/10.1145/2024156.2024169.

Shnerb, N. M., Louzoun, Y., Bettelheim, E., & Solomon, S. (2000). The importance of being discrete: Life always wins on the surface. *Proceedings of the National Academy of Sciences, 97*(19), 10322–10324. https://doi.org/10.1073/pnas.180263697. URL https://pnas.org/doi/full/10.1073/pnas.180263697.

Treibig, J., Hager, G., & Wellein, G. (2010). Likwid: A lightweight performance-oriented tool suite for x86 multicore environments. In *Proceedings of PSTI2010, the first international workshop on parallel software tools and tool infrastructures*. San Diego CA.

Vestergaard, C. L., & Génois, M. (2015). Temporal Gillespie algorithm: Fast simulation of contagion processes on time-varying networks. *PLoS Computational Biology, 11*(10), Article e1004579. https://doi.org/10.1371/journal.pcbi.1004579. URL https://dx.plos.org/10.1371/journal.pcbi.1004579.

Wang, D. (2008). Meeting green computing challenges. In *2008 10th electronics packaging technology conference* (pp. 121–126). https://doi.org/10.1109/EPTC.2008.4763421. URL https://ieeexplore.ieee.org/abstract/document/4763421.

Wendler, A. C., Plötzke, L., Tritzschak, H., & Kühn, M. J. (2024). *A nonstandard numerical scheme for a novel secir integro differential equation-based model with nonexponentially distributed stay times*. Submitted for publication. URL https://arxiv.org/abs/2408.12228.

Willem, L., Stijven, S., Tijskens, E., Beutels, P., Hens, N., & Broeckhove, J. (2015). Optimizing agent-based transmission models for infectious diseases. *BMC Bioinformatics, 16*(1), 183. https://doi.org/10.1186/s12859-015-0612-2. URL https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-015-0612-2.

Winkelmann, S., Zonker, J., Schütte, C., & Conrad, N. D. (2021). Mathematical modeling of spatio-temporal population dynamics and application to epidemic spreading. *Mathematical Biosciences, 336*, Article 108619. https://doi.org/10.1016/j.mbs.2021.108619. https://www.sciencedirect.com/science/article/pii/S0025556421000614

Wynants, L., Van Calster, B., Collins, G. S., Riley, R. D., Heinze, G., Schuit, E., et al. (2020). Prediction models for diagnosis and prognosis of covid-19: Systematic review and critical appraisal. *BMJ.* , Article m1328. https://doi.org/10.1136/bmj.m1328. URL https://www.bmj.com/lookup/doi/10.1136/bmj.m1328.

Yoneyama, T., Das, S., & Krishnamoorthy, M. (2012). A hybrid model for disease spread and an application to the SARS pandemic. *The Journal of Artificial Societies and Social Simulation, 15*(1), 5. https://doi.org/10.18564/jasss.1782. URL http://jasss.soc.surrey.ac.uk/15/1/5.html.

Zunker, H., Schmieding, R., Kerkmann, D., Schengen, A., Diexer, S., Mikolajczyk, R., et al. (2024). Novel travel time aware metapopulation models and multi-layer waning immunity for late-phase epidemic and endemic scenarios. *PLoS Computational Biology, 20*(12), 1–34. https://doi.org/10.1371/journal.pcbi.1012630, 10.1371/journal.pcbi.1012630.

Further readings

Twitter. (2020). *Twitter docs.* https://developer.twitter.com/.

Morris, M.D. (1991). Factorial sampling plans for preliminary computational experiments. Technometrics, 33(2), 161–174. doi:10.1080/00401706.1991.10484804. URL http://www.tandfonline.com/doi/abs/10.1080/00401706.1991.10484804.