Epidemics 51 (2025) 100836

Contents lists available at ScienceDirect

Epidemics

journal homepage: www.elsevier.com/locate/epidemics

Integrative modeling of the spread of serious infectious diseases and corresponding wastewater dynamics

Nina Schmid [a,b], Julia Bicker [c], Andreas F. Hofmann [d], Karina Wallrafen-Sam [a,b], David Kerkmann [e], Andreas Wieser [f,g,h,i], Martin J. Kühn [a,b,c], Jan Hasenauer [a,b,*]

[a] Life & Medical Sciences Institute (LIMES), Bonn, Germany
[b] Bonn Center for Mathematical Life Sciences, University of Bonn, Bonn, Germany
[c] Institute of Software Technology, Department of High-Performance Computing, German Aerospace Center, Cologne, Germany
[d] tandler.com GmbH, Buch am Erlbach, Germany
[e] Helmholtz Centre for Infection Research, Brunswick, Germany
[f] Institute of Infectious Diseases and Tropical Medicine, LMU University Hospital Munich, Munich, Germany
[g] German Centre for Infection Research (DZIF), Partner Site Munich, Munich, Germany
[h] Max von Pettenkofer Institute, Faculty of Medicine, LMU Munich, Munich, Germany
[i] Immunology, Infection and Pandemic Research (IIP), Fraunhofer Institute for Translational Medicine and Pharmacology (ITMP), Munich, Germany

ARTICLE INFO

Dataset link: https://github.com/SciCompMod/memilio/tree/inside-demonstrator-final, https://doi.org/10.5281/zenodo.14046493, https://github.com/inside-consortium/inside_demonstrator

Keywords:
Infectious diseases
Wastewater
Agent-based model
Hydrodynamic model
Sewer network

ABSTRACT

The COVID-19 pandemic has emphasized the critical need for accurate disease modeling to inform public health interventions. Traditional reliance on confirmed infection data is often hindered by reporting delays and under-reporting, while antigen or antibody testing of a full cohort can be costly and impractical. Wastewater-based surveillance offers a promising alternative by detecting viral concentrations from fecal shedding, potentially providing a more accurate estimate of true infection prevalence. However, challenges remain in optimizing sampling protocols, locations, and normalization strategies, particularly in accounting for environmental factors like precipitation.

We present an integrative model that simulates the spread of serious infectious diseases by linking detailed infection dynamics with wastewater processes through viral shedding curves. Through comprehensive simulations, we examine how virus characteristics, precipitation events, measurement protocols, and normalization strategies affect the relationship between infection dynamics and wastewater measurements. Our findings reveal a complex relationship between disease prevalence and corresponding wastewater concentrations, with key variability sources including upstream sampling locations, continuous rainfall, and rapid viral decay. Notably, we find that flow rate normalization can be unreliable when rainwater infiltrates sewer systems. Despite these challenges, our study demonstrates that wastewater-based surveillance data can serve as a leading indicator of disease prevalence, predicting outbreak peaks before they occur. The proposed integrative model can thus be used to optimize wastewater-based surveillance, enhancing its utility for public health monitoring.

1. Introduction

The COVID-19 pandemic has highlighted the need for effective real-time monitoring and prediction of infectious disease dynamics to support timely and informed intervention policies. In this context, various mathematical modeling studies have been conducted to assess vaccine distribution strategies (Castonguay et al., 2023; Gudina et al., 2024) or non-pharmaceutical interventions like telework suggestions, prohibitions of private gatherings of certain sizes, contact tracing, or partial lock-downs (Chae et al., 2023; Abernethy and Glass, 2022; Oraby et al., 2021; Endo et al., 2021).

The spread of infectious diseases can be modeled using a broad range of approaches, including statistical and machine learning models (Siedner et al., 2020; Latsuzbaia et al., 2020; Núñez et al., 2023), compartmental and meta-population models (Contento et al., 2023; Castonguay et al., 2023; Zunker et al., 2024), agent-based models (Chae et al., 2023; Lorch et al., 2022; Kerr et al., 2021; Hinch et al., 2021), and hybrid approaches (Hunter et al., 2020; Bicker et al., 2025). Among these approaches, agent-based models (ABMs) allow for the most detailed description of disease dynamics and can provide insight into the causal mechanisms underlying observed outcomes (Tracy et al.,

* Corresponding author.
E-mail address: jan.hasenauer@uni-bonn.de (J. Hasenauer).

https://doi.org/10.1016/j.epidem.2025.100836
Received 18 December 2024; Received in revised form 28 March 2025; Accepted 7 May 2025
Available online 31 May 2025

2018). ABMs simulate the spread of infectious diseases on an individual level, thereby facilitating the incorporation of comprehensive information about localization, interaction, and behavior. The models are intrinsically stochastic and are often based on discrete- or continuous-time Markov processes. While ABMs have clear advantages over other infectious disease modeling methods, their advancement remains an active field of research, with a key challenge being the choice of appropriate model parameters.

The predictive power of models for the spread of infectious diseases depends on the quality of the data available to inform the model parameters. The number of confirmed infections is the most common data source. However, confirmed infections are affected by reporting delays (Larremore et al., 2021) and subject to under-reporting (Lau et al., 2021; Merkt et al., 2024), limiting the reliability of the resulting models (Raimúndez et al., 2021). Meanwhile, the use of antigen or antibody tests to study large, representative population cohorts is resource intensive even when single tests are cheap, and such cohorts may still be subject to sampling bias (Millard et al., 2023; Howe et al., 2013). Wastewater-based surveillance presents a promising solution to these issues by capturing viral concentrations from all infected individuals within a catchment area, including those who are asymptomatic or undetected by traditional individual testing methods. Via detection of viral RNA in sewage, outbreaks can be identified before clinical cases are reported. This capability was demonstrated during the COVID-19 pandemic, when several national wastewater-based surveillance programs were established, e.g. the AMELAG project in Germany (Fachgebiet 32, Robert Koch-Institut, 2024). Wastewater-based surveillance is also applicable to a variety of other diseases detectable in sewage, including poliovirus, hepatitis, norovirus, and influenza (Kilaru et al., 2023), as well as to the monitoring of antimicrobial resistance (Chau et al., 2022), with some studies dating back several decades (Ranta et al., 2001). Integrating wastewater data with traditional infectious disease models holds promise for improving prediction accuracy and has been attempted in several studies (Champredon et al., 2024; Nourbakhsh et al., 2022). Yet, infectious disease monitoring based on wastewater-based surveillance data still presents several significant challenges and various sources of uncertainty (Wade et al., 2022): The choice of the sampling location can yield different viral concentrations due to the interplay of population density, human mobility patterns, and sewer infrastructure. The impact of in-network viral degradation and variations in flow-time on wastewater measurements is not well understood. Environmental conditions such as rainfall can also impact measurements due to sample dilution, while the effectiveness of established normalization strategies remains unclear.

These challenges associated with wastewater-based surveillance data can in principle be addressed using comprehensive computational models that provide in-depth descriptions of the spread of infectious diseases as well as wastewater dynamics. Yet, to the best of our knowledge, most published studies use relatively simple approaches. For example, Wu et al. (2020) calculated a rough estimate of SARS-CoV-2 prevalence upstream of a wastewater treatment facility using the normalized viral load measured in twelve wastewater samples and assumptions about the sewer system flow volume, stool sizes, and average viral concentration in stool among infected persons. The authors concluded that prevalence in their population of interest was much higher than the confirmed case count, even under conservative assumptions. However, they also noted that their estimate was subject to considerable uncertainty, as they did not account for the timeline of viral shedding or the loss of viral copies along sewer lines, among other factors. Hart and Halden (2020) used a simplified hydrodynamic model of a city sewer network to estimate SARS-CoV-2 detectability in wastewater under different temperature-driven decay scenarios. This study highlighted the importance of appropriately accounting for viral decay when analyzing wastewater data, but the authors assumed exclusively dry weather conditions and their consideration of the relationship between SARS-CoV-2 prevalence and viral load entering the sewer system was limited; accounting for variations in viral shedding across individuals and over time was out of the study's scope. Wang et al. (2020) conducted a simulation study of typhoid pathogen shedding and transport through a synthetic sewer system inspired by that of Kolkata, India; they concluded that optimal sampling site placement depends on pathogen decay rates, assay sensitivity, disease incidence, and spatial clustering of shedding persons. However, this study did not account for the effects of rainfall or human mobility, instead assuming that each case shed into a single location for their entire infection period. Peccia et al. (2020) used a distributed lag time series model to conclude that wastewater-based surveillance data led positive COVID-19 tests and hospital admissions by several days in one urban U.S. community. This approach highlighted the potential of wastewater data to provide early warnings of outbreaks, but relied on the assumption that the observed wastewater measurements provided an unbiased estimate of the true viral concentration trajectory. Finally, Nourbakhsh et al. (2022) coupled an SEIR-type compartment model of SARS-CoV-2 transmission with a simple advection–dispersion-decay model of virus concentration dynamics in a sewer system to estimate cumulative incidence in several cities based on empirical wastewater measurements and reported case data. However, because of a limited sewer system model and the use of an ordinary differential equation-based compartment model, the model was not well adapted for scenarios such as small communities or low-prevalence settings. To the best of our knowledge, no existing study on wastewater-based infectious disease monitoring uses advanced methods from the well-established field of sewage network modeling (Schütze et al., 2002). This is problematic as wastewater dynamics in sewage networks are highly complex and require sophisticated modeling tools for the prediction of pollutant load (Bach et al., 2014). Accordingly, advances beyond existing approaches to jointly model infection and wastewater dynamics could significantly enhance our understanding of disease spread and lead to more effective public health interventions.

Our work contributes to the field of infectious disease modeling by integrating an ABM for infection dynamics, a model of viral shedding, and a detailed hydrodynamic model of sewage flow and viral load. Through simulation studies for a respiratory virus (Section 3.1), we investigate the effects of measurement protocols (Sections 3.3 and 3.4), precipitation events (Section 3.5), viral decay (Section 3.6), and normalization strategies (Section 3.7) on the relationship between infection dynamics and wastewater measurements. The controlled setting of our simulation study enables an in-depth model-based analysis and practical recommendations for real-world wastewater-based surveillance, and our model furthermore provides a basis for coherent data integration. Our findings advance the development of detailed integrative models informed by data, enhancing the accuracy and reliability of infectious disease monitoring and prediction.

2. Mathematical model

To study wastewater-based surveillance data, we combine state-of-the-art models for the spread of infectious diseases and wastewater dynamics (Fig. 1). The link is established using a viral shedding model. All models are dynamic and executed on the same time axis, such that a coupling with comparisons of outputs is possible. The combined model's modular structure makes it feasible to update or replace individual component models to suit different use cases without disrupting the overall framework. In this section, we outline the mathematical formulations of the individual models and their simulation algorithms, as well as their integration.

2.1. Overview

Our proposed overall integrative model consists of the following three linked models:

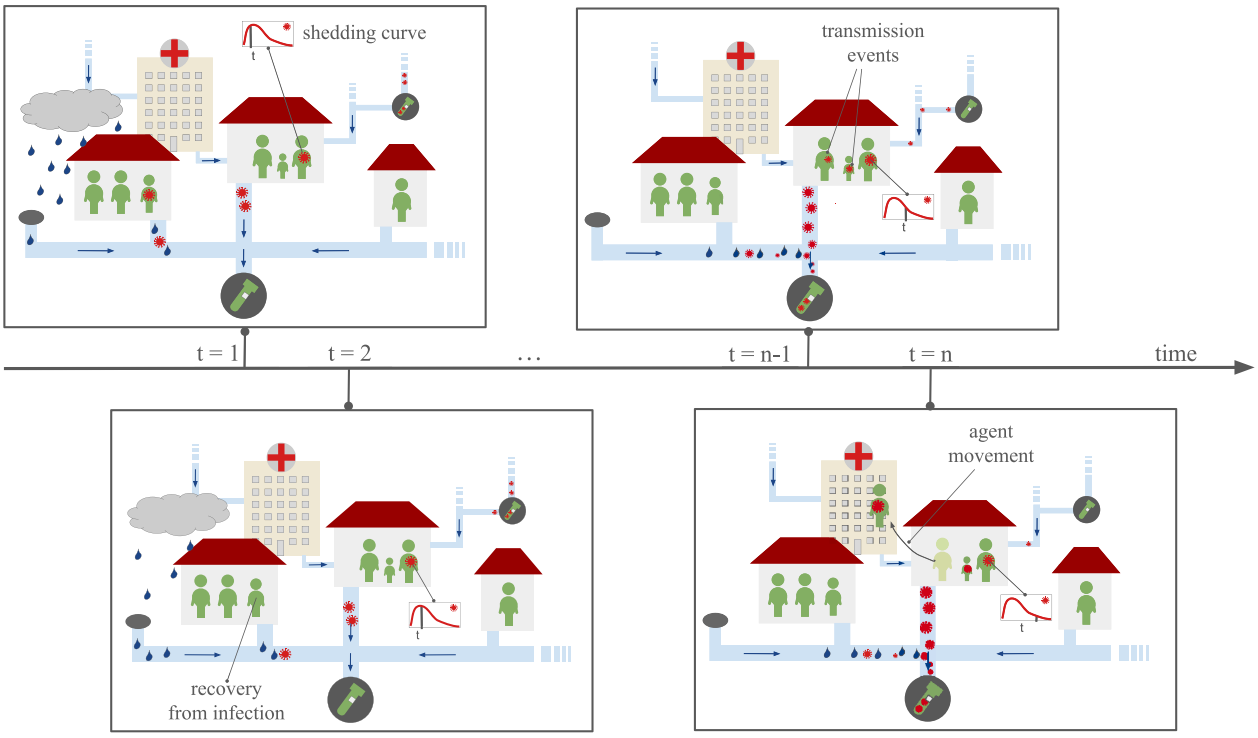

Fig. 1. Visualization of the Integrative Model. The *infection dynamics model* simulates the movement of agents as well as transmission and disease progression events. Based on the outputs of the infection dynamics model, RNA is shed at specific locations and time points into the sewer system according to the *shedding model*. This information is propagated to the *wastewater dynamics model*, which simulates the concentration measurements in wastewater over time and by location. Among other elements, the wastewater dynamics model can account for precipitation events, sampling protocols, viral load dynamics, and RNA degradation. This figure only depicts a simplified neighborhood with two types of locations (homes and a hospital), while the integrative model used for this study allows for various types of agent movements (Section 2.2.1).

The **infection dynamics model** describes the time-dependent location and infection state of individual persons, hereafter also referred to as agents. The infection state of agents can change due to events such as virus transmission, worsening of symptoms, or recovery. The likelihood of an agent infecting others is determined by its viral load, which varies throughout its infection course, and the length of contact.

The **shedding model** describes the release of virus and viral fragments from infected individuals into their surroundings. It is used in the infection dynamics model to determine the transmission probability as well as to describe the release through urine and stool. The shedding curve is assumed to be dependent on the individual's viral load, which is time-dependent and initially increases before declining as the host's immune response takes effect. The shedding model uses the infection state of agents and their time since transmission, which are provided by the infection dynamics model.

The **wastewater dynamics model** describes the transport and degradation of viral RNA within the sewage network. Using the viral shedding input from the shedding model and the agents' locations from the infection dynamics model, this model calculates the RNA concentrations at various points in the network, accounting for factors such as viral decay, flow rates, and the architecture of the sewer system. The output of this model is the RNA concentration at different sampling points, which provides a comprehensive picture of the data to be expected from wastewater-based surveillance.

In the current model, the wastewater does not influence the infection dynamics; accordingly, the integrative model possesses a hierarchical structure. In the following, we discuss the individual models in more detail.

2.2. Modeling infection dynamics

In this study, we use an ABM implemented in the C++-based software framework MEmilio (Kühn et al., 2023) to simulate disease states and mobility patterns at the agent level, providing a fine-grained view of disease dynamics. The mobility processes (Section 2.2.1) use an hourly timescale while the disease progression and transmission processes (Section 2.2.2 and Section 2.2.3) use a daily timescale. Throughout the following sections, we use $\tilde{t}$ to denote time in hours and t to denote time in days.

The properties of an agent α are defined via an m-tuple $(a_1, \ldots, a_m) \in \Omega$ with m different attributes a_i, $i = 1, \ldots, m$. The attributes can be static – meaning that they do not change over the course of the simulation – or dynamic. The static attributes are:

- An agent's age group $\mathcal{A}^{(\alpha)} \in \{1, \ldots, n_A\}$, with n_A denoting the total number of age groups.
- An agent's set of locations $\mathcal{L}^{(\alpha)} = \{l_j\}_{j \in K^{(\alpha)}}$, with l_j denoting a particular location and $K^{(\alpha)} \subset \{1, \ldots, n_L\}$ defining the subset of the indices of all locations that the agent can potentially move to during a simulation. n_L denotes the total number of locations in the global location graph.

The dynamic attributes are:

- An agent's current location $l^{(\alpha)} \in \mathcal{L}^{(\alpha)}$.
- An agent's current infection state $s^{(\alpha)} \in \mathcal{S}$, with $\mathcal{S}$ denoting a set of infection states.
- An agent's time since virus exposure $\tilde{\tau}^{(\alpha)}$ in hours, which is set to NaN if the agent has not been infected.

The simulation of the ABM provides information about the agent's trajectory in space and its infection state. In the following, we present additional details on the ABM, providing the basis for the simulation of mobility (Algorithm 1) and the full population dynamics (Algorithm 2).

Algorithm 1: ABM Mobility.

Input: Agent α, time point $\tilde{t}_k$, agent's location $l^{(\alpha)}(\tilde{t}_k)$
Output: Agent's location $l^{(\alpha)}(\tilde{t}_{k+1})$
Set $l_{\text{new}} = l^{(\alpha)}(\tilde{t}_k)$
Forall mobility rules $m_i \in \mathcal{M}$
  If $m_i(\alpha, \tilde{t}_k) \neq T(l_{\text{new}})$
    Get location $l \in \mathcal{L}^{(\alpha)}$ such that $T(l) = T^{(i)}_{\text{to}}$
    If capacity(l) not reached
      $l_{\text{new}} = l$
      return l_{new}

2.2.1. Mobility model

The ABM models mobility using a location graph with locations $l_1, \dots, l_{n_L}$, which each have a type $T \in \mathcal{T}$. For this study, $\mathcal{T}$ contains the location types *Home*, *School*, *Work*, *Recreation*, *Shop*, *Hospital*, and *Intensive Care Unit (ICU)*. There can be multiple locations of the same type. In addition to the type, a location can also have a user-defined capacity specifying the maximum number of agents that can enter the location and/or a user-defined maximum number of contacts an agent can have at the location. Every agent α has a set of locations $\mathcal{L}^{(\alpha)} = \{l_j\}_{j \in K^{(\alpha)}}$ assigned to it, meaning that it is restricted to a subgraph of the global location graph (see Supplementary Figure S1a). We denote the current location of an agent α at a given time point $\tilde{t} \in [\tilde{t}_0, \tilde{t}_{\max}]$, with $\tilde{t}$ given in hours, as $l^{(\alpha)}(\tilde{t}) \in \mathcal{L}^{(\alpha)}$. Movements, i.e. location transitions, are modeled by an ordered set of mobility rules $\mathcal{M} = \{m_1, \dots, m_{n_M}\}$ that have probabilistic components, which cause stochasticity between simulations. These mobility rules include daily regular behavior like going to work or school on weekdays, irregular behavior like occasionally attending a social event, and behavior related to the infection state of an agent, e.g. going to hospital when having severe symptoms. For Ω containing all potential states of any agent and $[\tilde{t}_0, \tilde{t}_{max}]$ denoting the simulation period, a mobility rule $m_i : \Omega \times [\tilde{t}_0, \tilde{t}_{\max}] \to \mathcal{T}$ is given by

$$m_i(\alpha, \tilde{t}) = \begin{cases} T^{(i)}_{\text{to}}, & \text{if } \delta^{(i)}(\tilde{t}, \alpha)\ \delta_{T^{(i)}_{\text{from}}}\left[T(l^{(\alpha)}(\tilde{t}))\right] X_i = 1 \\ T(l^{(\alpha)}(\tilde{t})), & \text{else} \end{cases} \tag{1}$$

with $T(l) \in \mathcal{T}$ denoting the type of location l, $X_i \in \{0, 1\}$ denoting a Bernoulli distributed random variable with probability p_i, which has a different value for every mobility rule, and $\delta_*[\cdot]$ denoting binary-valued functions defined as

$$\delta_{T^{(i)}_{\text{from}}}\left[T(l^{(\alpha)})\right] = \begin{cases} 1, & \text{if } T(l^{(\alpha)}) = T^{(i)}_{\text{from}}, \\ 0, & \text{else.} \end{cases} \tag{2}$$

Furthermore, $T^{(i)}_{\text{from}}, T^{(i)}_{\text{to}} \in \mathcal{T}$ are location types and $\delta^{(i)}(\tilde{t}, \alpha)$ are binary-valued functions controlling when transitions from $T^{(i)}_{\text{from}}$ to $T^{(i)}_{\text{to}}$ happen for agent α. The definitions for $\delta^{(i)}(\tilde{t}, \alpha)$ and the values of p_i are provided in Appendix A for each mobility rule. The location $l^{(\alpha)}$ of agent α at time point $\tilde{t}$ is given by an initial location at $\tilde{t}_0$ and then by Algorithm 1, which is evaluated at discrete time points $\tilde{t}_0, \dots, \tilde{t}_{\max}$ given a time step $\Delta\tilde{t}$ and $\tilde{t}_{k+1} = \tilde{t}_k + \Delta\tilde{t}$.

2.2.2. Disease progression

An agent α has a time-dependent infection state $s^{(\alpha)}(t)$, with t in days, from a set of infection states S. The infection states used for this study are *Susceptible* (S), *Exposed* (E), *Non-symptomatically Infected* (I_{ns}), *Symptomatically Infected* (I_{sy}), *Severely Infected* (I_{sev}), *Critically Infected* (I_{cri}), *Recovered* (R), and *Dead* (D). Infection state I_{ns} consists of infectious pre- and asymptomatic agents, while infection state I_{sev} consists of agents requiring hospital treatment and I_{cri} consists of agents requiring ICU treatment. We call an agent infected if it has infection state $s^{(\alpha)}(t) \in \{E, I_{\text{ns}}, I_{\text{sy}}, I_{\text{sev}}, I_{\text{cri}}\}$ and formerly infected if $s^{(\alpha)}(t) \in \{R, D\}$. Transitions between infection states are stochastic and possible either through virus transmission ($S \to E$) (see Section 2.2.3) or disease progression ($E \to I_*, I_* \to I_{**}, I_{**} \to \{R, D\}$) (see Supplementary Figure S1b). For a (formerly) infected agent α, the course of infection is defined as $\mathcal{I}^{(\alpha)} = \{(t^{(\alpha)}_1, s^{(\alpha)}_1), \dots, (t^{(\alpha)}_{h_\alpha}, s^{(\alpha)}_{h_\alpha})\}$ containing the time points $t^{(\alpha)}_1, \dots, t^{(\alpha)}_{h_\alpha}$ at which the agent changes or changed its infection state and the corresponding infection states $s^{(\alpha)}_1, \dots, s^{(\alpha)}_{h_\alpha}$. Hence, $t^{(\alpha)}_1$ is the time point at which the agent is exposed and $t^{(\alpha)}_{h_\alpha}$ the time point at which the agent recovers or dies, i.e. $s_{h_\alpha} \in \{R, D\}$. The intermediate time values $t^{(\alpha)}_2, \dots, t^{(\alpha)}_{h_\alpha - 1}$ are the time points at which an agent changes to one of the infectious states of the agent's individual course. The length h_α of the course of infection differs between non-symptomatic and (severe or critical) symptomatic courses and is therefore agent-dependent. The stay times in infection states $E, \dots, I_{\text{cri}}$ are log-normally distributed and the transitions $E \to I_*, I_* \to I_{**}, I_{**} \to \{R, D\}$ between all infection states, apart from virus transmission ($S \to E$), are Bernoulli distributed. See Supplementary Table S1 for the values of all disease progression-related parameters used in the Results section. These parameter values are motivated by wild-type COVID-19 and based on Kühn et al. (2021), Kerr et al. (2021), and Kerkmann et al. (2024).

2.2.3. Disease transmission

Infected agents can transmit the virus to susceptible agents if they are at the same location. For a susceptible agent in location l, the waiting time until transmission is exponentially distributed with rate $\Lambda_l(t)$. Assuming that agents change their locations only at discrete time points, $t_0, \dots, t_{\max}$, the number of agents at a location is constant in $[t_k; t_{k+1})$ and $\Lambda_l(t)$ is given by

$$\Lambda_l(t) = \sum_{\alpha \in l} \lambda_\alpha(t). \tag{3}$$

for $t \in [t_k; t_{k+1})$ given in days. We use the casual notation $\alpha \in l$ to iterate over the infected agents at location l in the interval $[t_k; t_{k+1})$. The agent-dependent rate $\lambda_\alpha(t)$ is given by the infectiousness curve described in Section 2.3, Eq. (8). The location-specific infection rate (3) builds on the assumption of homogeneous mixing within the location. If the waiting time until transmission is longer than the time until the susceptible agent leaves location l, no transmission occurs. Hence, for a fixed time step $\Delta t = t_{k+1} - t_k$ given in days, the probability that a susceptible agent at location l gets exposed is

$$1 - e^{-\Delta t \int_{t_k}^{t_{k+1}} \Lambda_l(t) dt}. \tag{4}$$

2.3. Modeling viral load, infectiousness, and shedding

The viral load of individual patients determines their infectiousness and shedding. Here, we model viral shedding using established models (Jones et al., 2021; Kerkmann et al., 2024) that account for the temporal variability of shedding, which has been shown to be highly relevant to wastewater-based epidemiology (Hoffmann and Alsing, 2023). The viral load for an agent α in RNA copies per nasopharyngeal swab sample on the $\log_{10}$ scale at a given time t in days (see Supplementary Figure S2a) is defined as

$$\upsilon_\alpha(t) = \begin{cases} \dfrac{\upsilon^{\text{sy}}_{\max}}{\tau_E + \tau_{I_{\text{ns}}}} \cdot (t - t_E), & \text{if } t \in [t_E, t_{\upsilon_{\max}}] \\ \upsilon_{\max} + -\dfrac{\upsilon_{\max}}{t_{\text{R/D}} - t_{\upsilon_{\max}}} \cdot (t - t_{\upsilon_{\max}}), & \text{if } t \in (t_{\upsilon_{\max}}, t_{\text{R/D}}] \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

with t_E, $t_{\upsilon_{\max}}$, and $t_{\text{R/D}}$ denoting the time points in days of virus exposure, maximal viral load, and recovery or death, respectively. Thus, we assume that viral shedding begins increasing from 0 copies per swab at the time of exposure and that it reaches 0 copies per swab again at the time of recovery or death. Furthermore, τ_E and $\tau_{I_{\text{ns}}}$ denote

Algorithm 2: ABM Simulation.

Input: Start time t_0, end time $t_{\max}$, time step Δt, parameters as described above
Initialize locations
  Create desired number of locations for every type
  Set capacity and maximum number of contacts
Initialize agents
  Set age group
  Assign locations and set initial location
  Set initial infection state and sample course of infection for initially infected
Set $t = t_0$
While $t < t_{\max}$
  Forall agents α
    If $s^{(\alpha)} = S$
      Calculate $\Lambda_{I^{(\alpha)}}(t)$
      Draw waiting time $\upsilon \sim \text{Exp}(\Lambda_{I^{(\alpha)}}(t + \frac{\Delta t}{2}))$
      If $\upsilon \leq \Delta t$
        Sample course of infection $\mathcal{I}^{(\alpha)}$
      Else continue
    Forall mobility rules m_i
      If $T(l^{(\alpha)}) \neq m_i(\alpha, 24\frac{\text{hours}}{\text{day}} \cdot t)$
        Get location $l_{new} \in \mathcal{L}^{(\alpha)}$ such that $T(l_{new}) = m_i(\alpha, 24\frac{\text{hours}}{\text{day}} \cdot t)$
        If capacity(l_{new}) not reached
          $l^{(\alpha)} = l_{new}$
          break
  $t = t + \Delta t$

the times in days spent in the *Exposed* and *Non-Symptomatically Infected* states, respectively, and $\upsilon_{\max}^{\text{sy}}$ denotes the peak viral load in $\log_{10}$ RNA copies per swab for symptomatic infections. The time points are agent specific, with $t_E = t_1^{(\alpha)} \leq t_{\upsilon_{\max}} \leq t_{\text{R/D}} = t_{h_\alpha}^{(\alpha)}$. As the infection state trajectories are sampled at the time point of exposure, $t_{\text{R/D}}$ is readily available.

We assume that symptomatically infected agents reach the peak viral load the moment they show symptoms, while agents not showing symptoms reach the peak in the middle of their infection period. Furthermore, we assume agents that share the same duration $\tau_E + \tau_{I_{\text{ns}}}$ also share the same linear increase until $t_{\upsilon_{\max}}$. This yields for the variables $t_{\upsilon_{\max}}$ and $\upsilon_{\max}$:

$$t_{\upsilon_{\max}} = \begin{cases} t_E + \tau_E + \tau_{I_{\text{ns}}}, & \text{if } I_{\text{sy}} \in \{s_1^{(\alpha)}, \dots, s_{h_\alpha}^{(\alpha)}\} \\ t_E + 0.5 \cdot (t_{\text{R/D}} - t_E), & \text{else.} \end{cases}$$

$$\upsilon_{\max} = \begin{cases} \upsilon_{\max}^{\text{sy}}, & \text{if } I_{\text{sy}} \in \{s_1^{(\alpha)}, \dots, s_{h_\alpha}^{(\alpha)}\} \\ \dfrac{\upsilon_{\max}^{\text{sy}}}{\tau_E + \tau_{I_{\text{ns}}}} \cdot (t_{\upsilon_{\max}} - t_E), & \text{else.} \end{cases}$$

Following the observation of Jones et al. (2021), we model the shape of an agent α's shedding by a sigmoid function of their viral load, i.e.,

$$\zeta_\alpha(t) = \frac{1}{1 + \exp\left(-(a + b \cdot \upsilon_\alpha(t))\right)} \tag{6}$$

with shape parameters $a, b > 0$. As in Jones et al. (2021), since there is no information about when shedding starts after exposure, we make the assumption that shedding is zero or close to it as long as the agent is still in the *Exposed* state. Therefore we introduce a time shift of $\tau_{\text{shift}} = 0.6 \cdot \tau_E$. Thus, the scaled and shifted shedding curve for an agent α is given by:

$$\gamma_\alpha(t) = \begin{cases} \kappa_\gamma \cdot \zeta_\alpha(t - \tau_{\text{shift}}), & \text{if } t \in [t_E + \tau_{\text{shift}}, t_{\text{R/D}}] \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

where κ_γ is a scaling factor that translates $\zeta_\alpha(t - \tau_{\text{shift}})$ into an RNA shedding rate. The RNA shedding into the sewage network is the number of RNA copies shed in total per day and has to be normalized by the water flushed into the wastewater system to receive a unit of copies per liter.

Similarly, the corresponding unitless infectiousness curve for agent α at time point t (see Supplementary Figure S2b) is given by

$$\lambda_\alpha(t) = \begin{cases} \kappa_\lambda \cdot \zeta_\alpha(t - \tau_{\text{shift}}), & \text{if } t \in [t_E + \tau_{\text{shift}}, t_{\text{R/D}}] \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

with κ_λ translating $\zeta_\alpha(t - \tau_{\text{shift}})$ into a transmission rate (see Section 2.2.3). See Supplementary Table S2 for the value of all shedding-related parameters used in the Results section. For all parameters except κ_λ, wild-type COVID-19-like values were chosen based on Jones et al. (2021), Kerkmann et al. (2024); the value of κ_λ was chosen to ensure that all simulated outbreaks had a defined peak and that proportional prevalence stayed below 5%.

2.4. Modeling the sewage system and its hydrodynamics

We use a comprehensive model of a wastewater network to simulate the sewage flow and reactive transport of dissolved chemical substances in wastewater while avoiding oversimplifications that could lead to incorrect interpretations of wastewater measurements. In this model, a sewer system is represented as a directed acyclic graph defined by n edges $\mathcal{E} = \{e_1, \dots, e_n\}$ and m nodes $\mathcal{N} = \{n_1, \dots, n_m\}$. The edges represent sewage pipes and the nodes represent junctions as well as entry and exit points. Edges and nodes are characterized by several parameters, including total volume and height. The state of the system is the amount of water contained in the edges and nodes, its flow rate, and the concentrations of the relevant substances. The simulation of water inflow to the system is based on two phenomena: hydraulic surface runoff during precipitation events and water usage of industry and citizens.

At their core, the hydrodynamic calculations are based on the Saint-Venant Equation (Hodges, 2019). The equation assumes one-dimensional flows in open channels and mass and momentum conservation, which yields

$$\frac{1}{g}\frac{\partial \upsilon}{\partial t} + \frac{\upsilon}{g}\frac{\partial \upsilon}{\partial x} + \frac{\upsilon}{g \cdot A}q + \frac{\partial h}{\partial x} + \frac{\lambda}{4 \cdot R}\frac{\upsilon|\upsilon|}{2g} = 0, \tag{9}$$

with flow velocity υ in $\frac{\text{m}}{\text{s}}$, time t in s, gravitational acceleration g in $\frac{\text{m}}{\text{s}^2}$, vertical position along the pipe x in m, cross-sectional flow area A in m^2, lateral inflow corresponding to the precipitation of a specified time interval q in $\frac{\text{m}^2}{\text{s}}$, water levels h in m, pipe friction coefficient λ (unitless) and hydraulic radius R in m (i.e. $\frac{A}{\text{Circumference}}$).

Replacing the differential operators with differential quotients for specified time points $t_1 < t_2$ and locations $x_1 < x_2$ yields

$$\frac{\partial \upsilon}{\partial t} \to \frac{\upsilon(t_2, x) - \upsilon(t_1, x)}{t_2 - t_1}, \tag{10}$$

$$\frac{\partial \upsilon}{\partial x} \to \frac{\upsilon(t, x_2) - \upsilon(t, x_1)}{x_2 - x_1}, \tag{11}$$

$$\frac{\partial h}{\partial x} \to \frac{h(t, x_2) - h(t, x_1)}{x_2 - x_1}, \tag{12}$$

for $t \in [t_1, t_2]$ and $x \in [x_1, x_2]$. By interpreting $\Delta x = x_2 - x_1$ as the length of an edge of the wastewater network, Eq. (9) after rearrangement becomes a quadratic equation of the form

$$a\upsilon(t, x)^2 + b\upsilon(t, x) + c = 0, \tag{13}$$

where

$$a = \text{sign}(\upsilon(t_1, x)) \cdot \frac{\lambda}{8R},$$

$$b = \frac{1}{t_2 - t_1} + \frac{\upsilon(t, x_2) - \upsilon(t, x_1)}{\Delta x} + \frac{\text{q}}{\text{A}} \quad \text{and}$$

$$c = -\frac{v(t_1, x)}{t_2 - t_1} + \frac{h(t, x_2) - h(t, x_1)}{\Delta x}\, g\ .$$

The quadratic equation has two complex solutions, where the flow velocity v is equal to the real part of those solutions:

$$v(t, x) = \mathrm{RE}\left(-b + \frac{\sqrt{b^2 - 4ac}}{2a}\right). \tag{14}$$

Using this solution and the boundary conditions $v(t, x_1)$, $v(t, x_2)$, $h(t, x_1)$, and $h(t, x_2)$, as well as the initial condition $v(t_1, x)$, a solution of the flow velocity can be calculated for arbitrary time points and edges of the network system. The state of the nodes of the system define the boundary and initial conditions. Based on the solutions from Eq. (9) for edges connected to a node, the in- and outflow to the node is given for each edge. At the center of the node, the sum of in- and outflow equals zero (mass conservation). Based on this assumption, the flow rates and heights at the edge borders can be calculated.

While Δx can be defined based on the length of an edge, $\Delta t = t_2 - t_1$ has to be chosen carefully. A too large value of Δt will yield inaccurate numerical solutions, while a too small value will yield unnecessarily high computation times. A complex solution of Eq. (13) with an imaginary component larger than zero indicates the transition to an oscillatory state. To ensure a meaningful numerical solution, Δt is chosen to be smaller than the corresponding oscillation period. Further considerations like the maximum total change of volume provide an equation to set Δt for each iteration of the numerical solution scheme, such that stable solutions of v are ensured even for extreme hydraulic scenarios. By combining Eq. (13) with the known physical principles of energy loss along a pipe (using the Prandtl-Colebrook equation) and energy loss due to inelastic collisions in manhole structures (using the Borda-Carnot equation), the simulation precision is further refined.

Based on the calculated flow velocities and other time-varying edge state variables, the substance concentration per location and time point can be calculated. Viral loads generated by agents of the ABM enter the sewer system as concentrations in the respective amount of domestic wastewater generated for every time step of the model simulation, at the location the agent currently occupies. Viral fragments are then transported through the system according to the pre-calculated flow rates, potentially taking chemical reactions in the form of viral decay into account. The outputs of the hydraulic simulation are time- and location-dependent concentration curves. The time step for the viral load calculation is chosen prior to the calculations and a suitable value depends on the specifics of the viral decay dynamics, where faster changing dynamics suggest choosing a smaller time step.

After defining the wastewater network, its connected surfaces and corresponding runoffs, as well as substance characteristics, the simulation proceeds in two steps (see Supplementary Figure S3). First, the flow velocities and volumes are calculated with a numerical solver. Second, the viral load over time is simulated.

For the numerical simulation, we use the urban water management modeling and simulation environment ++SYSTEMS, developed by the company tandler.com GmbH. Utilizing the mathematical principles described above, ++SYSTEMS with its backend and calculation kernel DYNA forms a fully dynamic, geospatial modeling and management software for waste- and rainwater (individually or combined) sewer systems. Details on the implementation are available in the ++SYSTEMS documentation (tandler.com GmbH, 2024).

3. Results

3.1. Demonstrator setup

To address open questions and challenges related to the interpretation of wastewater-based surveillance data, we performed a simulation study, which allows us to evaluate counterfactual scenarios without missing data or data uncertainty. To this end, we developed a synthetic, yet realistic model of a city neighborhood and corresponding sewer system (Fig. 2); we then traced infectious disease outbreaks in this controlled setting using our sequence of three linked models, from the MEmilio-based infection dynamics model to the shedding model to the ++SYSTEMS-based wastewater dynamics model. The synthetic neighborhood used in our study contains residential, recreational, university, mixed shopping and business, and mixed residential and industry surface areas. The residential and residential/industry surface areas together contain between 400 and 450 ABM *Home* locations; the recreational surface areas contain 12 ABM *Recreation* locations; the university, shopping/business, and residential/industry areas together contain between 20 and 30 ABM *Work* locations; and the shopping/business areas also contain between 10 and 20 ABM *Shop* locations as well as one *School*, one *Hospital*, and one *ICU* ABM location. This neighborhood is populated by between 840 and 910 agents, with the exact population size determined stochastically at the start of each simulation; the agents' simulated movements, both on weekdays and weekends, remain completely inside the model. The model's synthetic sewer system was designed such that realistic sewer conditions are maintained during all simulation scenarios: No sanitary sewer overflow occurs, all pipes are at a maximum of 90% of their hydraulic capacity, and gravity flow is realized throughout the whole system. At any given moment, wastewater flow into the sewer system from a particular surface area is proportional to the number of agents currently located there.

Following the official reporting standards for COVID-19 cases in Germany, we considered $n_A = 6$ age groups: 1 - small children, 2 - school-age children, 3 - younger working-age adults, 4 - older working-age adults, 5 - younger retirement-age adults, and 6 - older retirement-age adults. Households – i.e. groups of agents that share their assigned *Home* location – were created for each residential area based on its total number of agents. We considered 1- to 5-person households which each had at least one member of the adult age groups $\mathcal{A} \in \{3, 4, 5, 6\}$. The household distribution was motivated by household composition data averaged across Germany as measured by the German microcensus 2019 (Statistisches Bundesamt, 2024). Each agent was assigned non-*Home* locations it could potentially move to during a simulation: Every agent was assigned a location of type *Shop*, *Recreation*, *Hospital*, and *ICU*, while a *School* location was only assigned to agents in age group 2 and a *Work* location only to agents in age groups 3 and 4. A location of a given type was chosen for an agent from a uniform distribution of all locations of that type. For all locations, the capacity and maximum number of contacts were set high enough to ensure that agents were never turned away from one of their assigned locations because it was too full and that agents could always contact anyone with whom they were co-located. Finally, the initial infection states were allotted to agents by independent and identical sampling from the initial infection state distribution (0.2% E, 0.5% I_{ns}, 0.29% I_{sy}, and 0.01% I_{sev}), i.e. on average, 1% of the modeled agents were initially infected. All agents began each simulation at their assigned *Home* location.

The integrative model contains various parameters describing characteristics of the virus, which allows for the modeling of a broad spectrum of communicable respiratory or, in particular, COVID-19-like diseases. The ++SYSTEMS files defining the exact shape of the sewage system, area characteristics, etc. are provided as supplementary material for each experiment and in the subsequent sections we only mention settings that differ between the experiments. An overview of all ABM parameters as well as the experiment-specific parameters is provided in Appendix C.

Since most cluster systems are based on Linux, we facilitate our model using Ubuntu. ++SYSTEMS is a Windows program, hence, we created a headless virtual machine, which can be started and navigated through via a command line interface. Running 250 ABM simulations parallelized across 8 cores took about 1.5 h, while running 250 corresponding ++SYSTEMS simulations using the same setup took between about 12.5 and 16.5 h, depending on the rain and viral decay scenario (CPU specifics: AMD Ryzen 9 5900HX). The flow of information between model components via input and output files is illustrated in Supplementary Figure S4.

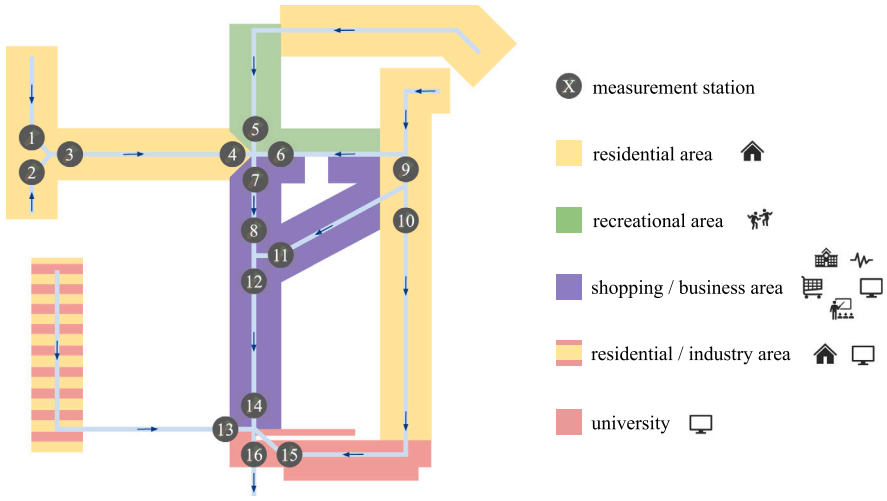

Fig. 2. Demonstrator Neighborhood. The synthetic neighborhood on which we base our simulation study constitutes areas of different types. All wastewater inflow into the sewer system comes from the shaded surface areas. Surface areas depicted as flush against a particular pipe contribute wastewater laterally into that pipe. Flow rates and substance concentrations are simulated for 16 different measurement stations.

3.2. Non-trivial relation of prevalence in catchment area and measurement concentrations

For a first assessment of the process dynamics, we considered the total number of upstream agents, the number of upstream infected agents, and the upstream RNA influx for a single measurement station for an outbreak scenario (Fig. 3). The chosen measurement station, Station 3, is downstream of residential areas, so the number of agents in its catchment area – and, by extension, the wastewater inflow – changes over the course of a day and from weekdays to weekends (Fig. 3a). For example, there are large drops in the number of upstream agents during daytime hours on each weekday due to work- and school-age agents leaving their homes, with smaller drops on weekends due to visits to shops or recreational events. Since the subset of agents who are infected may also leave their homes for work, school, or recreation (Fig. 3b), agent mobility results in substantial changes to the upstream RNA influx over time (Fig. 3c). The measurable virus levels at a particular station are a function of both the RNA influx and the sewage volume – which, as noted previously, depends on the total number of agents in the catchment area. Thus, the upstream RNA influx divided by the number of upstream agents (Fig. 3d) closely mirrors the measurable virus levels at Station 3 (Fig. 3e) – although the two trajectories are not exactly the same since both the RNA and wastewater must travel through the sewer system to reach the measurement station. Since agents e.g. return from their places of work or study at slightly different time points, the virus levels can show large deviations from an average value for only a few simulation minutes. Overall, the model highlights the impact of mobility on wastewater-based surveillance results.

3.3. Characteristics of catchment area influence wastewater measurement dynamics

Establishing a monitoring system for wastewater is time- and effort-consuming. Legal permits for measurements and access to the locations have to be organized, the sampling stations have to be set up, and the samples have to be collected and transported to laboratories for further analysis on a regular basis. Hence, it is not surprising that even established monitoring networks have limited sampling locations and time schedules, often reporting 1–3 values per week and neighborhood or city. This renders the optimal placement of sampling locations and the selection of appropriate sampling strategies critical. Here, we investigated the impact of the choice of sampling location using our fine-grained integrative model, as a corresponding real-world study would be infeasible.

We considered 16 possible sampling locations in the synthetic neighborhood. These sampling locations correspond to catchment areas with a broad spectrum of different properties: First, the corresponding catchment areas of the sampling stations differ with respect to their area types, i.e. primarily include residential areas (stations 1, 2, 3, 4, 9, 10), recreational areas (station 6), shopping/business areas (station 11), or mixed areas. Second, the catchment areas differ with respect to their sizes. Further upstream locations (particularly stations 1, 2, 3, 9, 10, and 13) summarize the dynamics of a smaller catchment area than downstream locations (particularly stations 7, 8, 12, 14, and 16). In this neighborhood, sewer flow times to the furthest downstream stations are at most around 90 min. Supplementary Table S5 provides an overview of the upstream area types, catchment area population sizes, and sewer flow times for each station.

To assess the information content of the wastewater-based surveillance data, we conducted a comprehensive simulation study. A total of 250 simulations of the proposed integrative model were used to account for the inherent stochasticity of infection processes; see Supplementary Figure S5 for a visualization of the prevalence over the 14-day simulation time frame. We began by assuming no viral decay and no effects of precipitation on the sewer system, and unless stated otherwise, all following results assume a 3 min temporal model resolution. Assessment of the simulation results shows that sampling locations downstream of residential areas (e.g. station 1, Fig. 4a) provide reproducible daily and weekly trends in the measured virus levels. For example, each week there is a dip in viral levels during the daytime on weekdays but not on weekends, since agents – including infected agents – leave their homes upstream of station 1 to go to work or school elsewhere in the neighborhood during standard working hours. In contrast, sampling locations downstream of regions containing recreational areas (e.g. station 6 & 11, Fig. 4b & d) show more variability between simulations. Sampling stations near the endpoint of the network, which have large catchment areas (e.g. station 16, Fig. 4c) and would in practice fall closer to a wastewater treatment plant, yield smoother curves with no or less extreme daily and weekly trends.

To evaluate how representative the viral load in the wastewater is at the different sampling locations, we computed the temporal cross-correlations between the RNA copies per liter in wastewater samples and (i) the true viral shedding into the wastewater (Fig. 5a1) and (ii) the true overall prevalence (Fig. 5a2). We found the highest cross-correlation values are reached for the stations with larger catchment areas, in particular stations 7, 8, 12, 14, and 16; see Supplementary Table S5 for a full list of cross-correlation values by station. The correlation coefficient is as high as 0.90 for the amount of virus shed

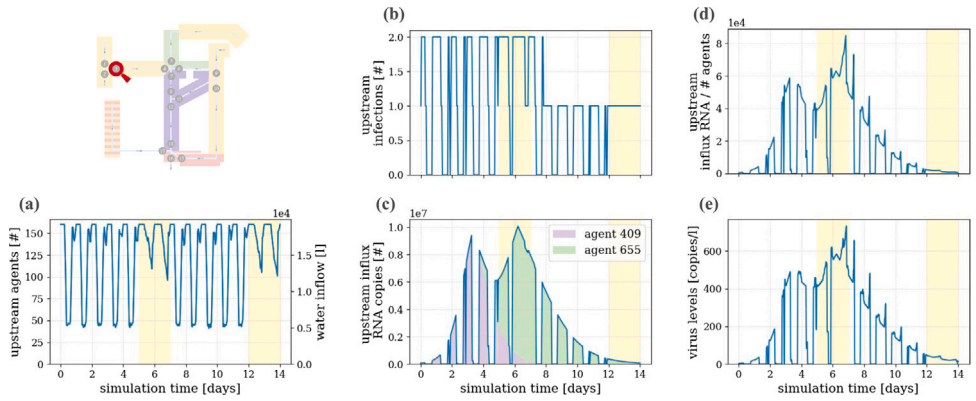

Fig. 3. Exemplary (processed) Simulation Output. Output for viral measurements in a scenario without precipitation or viral decay at measurement station 3. The yellow shading indicates weekends.

(Fig. 5b1) and 0.56 for the true prevalence (Fig. 5b2). These correlation coefficients are surprisingly high given the variability between the (stochastic) simulation runs. Indeed, if the initial infections were not distributed randomly in space across the neighborhood, the pattern would be more pronounced and large catchment areas would be even more beneficial (results not shown). If wastewater measurements from a particular sampling station are compared to the infection prevalence only in surface areas upstream of that sampling station instead of the total prevalence, then for upstream stations, correlations are highest when the prevalence is shifted by a multiple of 24 h, due to the previously discussed daily patterns in agent mobility (Supplementary Figure S6). This effect becomes less pronounced the further downstream a sampling station is located. Correlations between wastewater measurements and the number of incident infections were generally negligible (Supplementary Figure S7).

The integrative model also shows that the temporal cross-correlation is generally higher when a negative time lag is applied to the (true) prevalence – i.e., when the prevalence is shifted *earlier* in time relative to the wastewater virus concentrations. A sensitivity analysis shows that these results are generally robust to reasonable changes to the shedding model parameters (Supplementary Figure S9 and Figure S10). Since a high virus concentration in the wastewater indicates that the level of infectiousness across the population is also high and that a wave of new infections will therefore likely soon follow, the virus level at sampling stations with large catchment areas was most predictive for the prevalence 10 to 40 h later. This time lag likely depends on the incubation time of the virus and its replication rate in the human body. As the reported prevalence is delayed compared to the true one, the time shift observed in practice will be even larger (Fig. 5c). Overall, our results suggest that in the absence of complicating factors such as viral decay – the impact of which would be limited in this particular sewer due to the relatively short travel times – choosing a wastewater sampling location far enough downstream to be unaffected by daily and weekly trends may help predict increases in prevalence before they occur.

3.4. Temporal sampling design has minimal impact on wastewater monitoring results if samples are taken downstream

The sampling design differs between wastewater monitoring studies. The most common setups are one grab sample collected per day of interest (usually during the morning flush) or a 24-h compound sample based on a collection of one sample per hour (Kilaru et al., 2023). To assess the impact of these sampling strategies and their benefits and disadvantages, we simulated both strategies using the same setup as in the previous section (in which we assumed the use of discrete grab samples every three minutes). The analysis of the simulation results indicates that in the case of a clear daily trend, the choice of sampling protocol can influence the results and e.g. lead to systematically biased estimates of the general dynamics (Fig. 6a &b). Stations further downstream are less influenced by daily and weekly agent movement patterns (see Section 3.3) and different sampling protocols yield comparable results (Fig. 6c &d), i.e. a maximal cross-correlation of 0.56 between measured viral load and time-lagged prevalence.

3.5. Rain influx impacts reliability of wastewater monitoring non-linearly

Rainwater infiltration influences the amount of fluid in the wastewater system, the fluid velocity, and the concentration of particles in the overall wastewater. Yet, many currently available models simply disregard rain events and the associated dataset. This approach is an oversimplifying assumption that might lead to unreliable results, as it is unclear how long-lasting the effects of rain might be. To provide a fine-grained analysis of the impact of rain on wastewater measurements, we simulated three scenarios: No precipitation (which was also used for the previous results), moderate gentle rain, and moderate rain showers. Here, we followed the rain intensity definitions from Germany's National Meteorological Service (DWD) (Deutscher Wetterdienst, 2024): "moderate gentle rain" means between 0.1 mm and 0.5 mm in 60 min and "moderate rain" means between 2.5 mm and 10.0 mm in 60 min.

In order to run simulations with a two-week duration with realistic time-dependent rain intensities, a suitable two-week period from a synthetic rain series (Bayerisches Landesamt für Umwelt (LfU), 2024) generated by the Bavarian Environment Agency (LfU) was used. The intensities were adapted such that the above-mentioned rain definition criteria were met. As a result, the two rain scenarios only differ in their intensities; the temporal profile of rain peaks is the same for both scenarios, ensuring comparability (Fig. 7a).

Due to effects like evaporation, the filling of water basins (e.g. uptake by the ground), and permeable and non-permeable surface fractions, only a small proportion of rainfall ends up in the hydraulic system and a minimum amount of precipitation is necessary to have an effect at all. This net hydraulic surface runoff increases the water volume, and hence, can increase the flow rates (Fig. 7b) and reduce the concentration of RNA in wastewater significantly (Fig. 7c). RNA copies per liter close to zero correspond to a proportion close to 1 of rainwater in the sewage. If the flow rates increase by several orders of magnitude for a short amount of time (as is e.g. the case in the moderate gentle rain scenario presented here), the sewage containing large fractions of rainwater is flushed out of the system very quickly, yielding a state comparable to the non-precipitation state afterwards. The moderate gentle rain scenario showcases that there is a minimum amount of precipitation necessary to have an effect on the flow rates and virus concentrations. The moderate rain scenario highlights how precipitation influences measurements if the rainwater inflow to the sewage is larger than zero for several neighboring time steps.

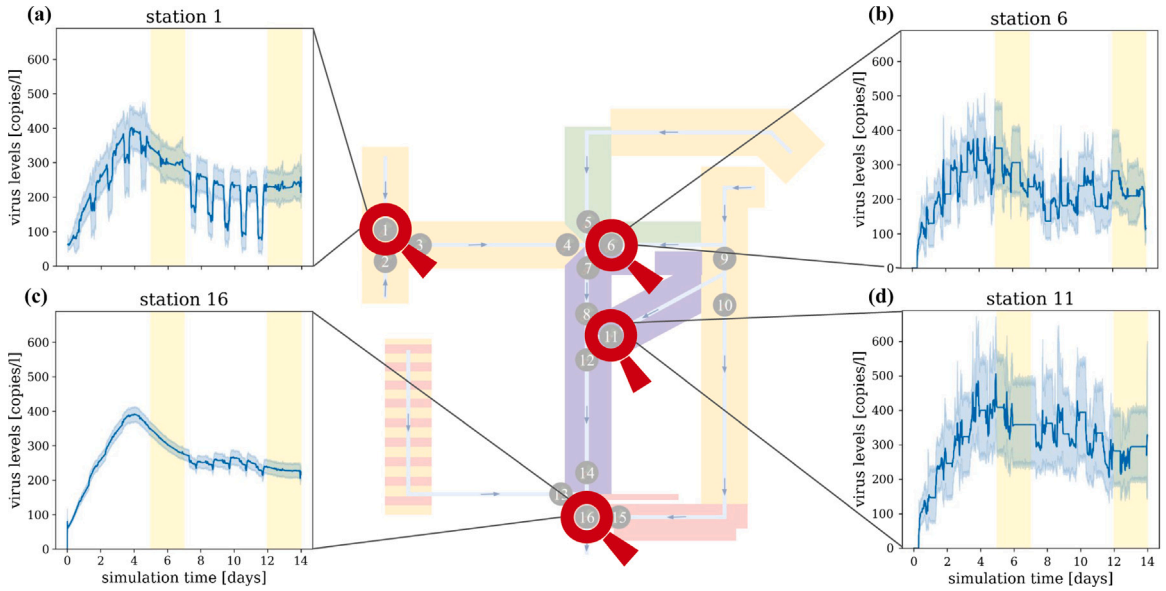

Fig. 4. Variability of Measurements for Sampling Locations. Comparison of virus levels in wastewater at different sampling locations. Shown are the mean (solid line) and 95% confidence intervals (shaded area) of the 250 simulation results per location. The yellow shading indicates weekends. See Supplementary Figure S8 for all stations.

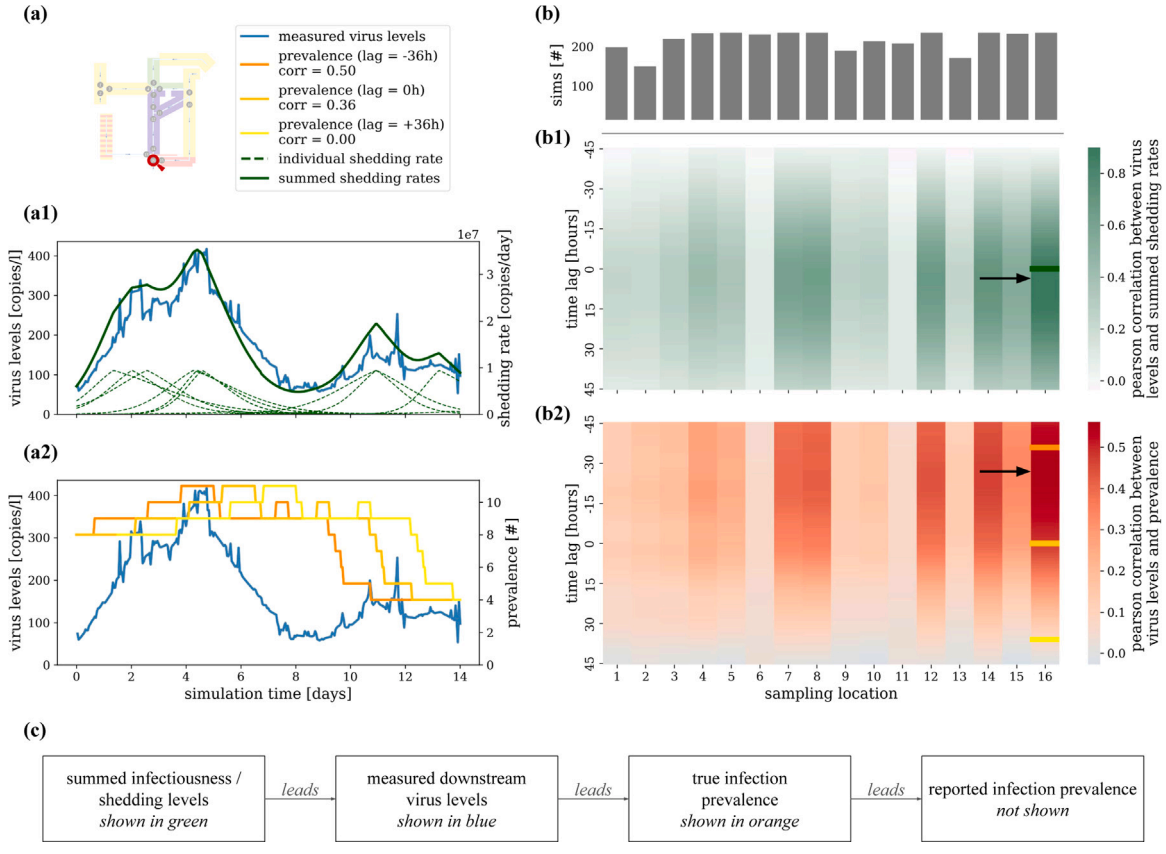

Fig. 5. Cross-Correlations Between Wastewater Measurements and Prevalence. (a) Trajectory of the wastewater viral load for sampling location 16 compared to the summed shedding rates across all prevalent infections (a1) or the total true prevalence shifted with lags -36, 0, and 36 h (a2), for one simulation. (b) Pearson cross-correlations between the wastewater viral load at the 16 different locations and the summed shedding rates (b1) or the total true prevalence (b2), averaged over 250 simulations. (Supplementary Figure S11 shows the correlations of the mean trajectories; the means of the per-simulation correlations are shown here.) The time lag describes the shift in prevalence or shedding rates. The maximum cross-correlations are marked with black arrows. The upper bar plot indicates how many simulations were used to calculate the correlations for each location; simulations in which no virus was ever measured at a particular location were removed. (c) Schematic illustrating the temporal relationship between the different outcomes.

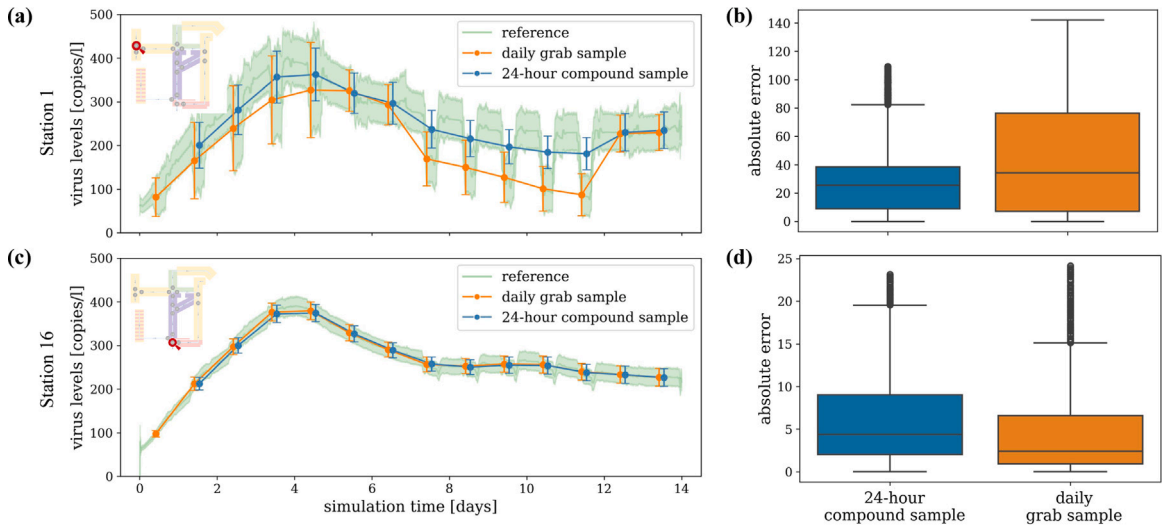

Fig. 6. Sampling Protocols. RNA concentration in wastewater with 24-h compound sampling or daily grab sampling (at 10:00 am each day) compared to the reference scenario (grab sampling every three minutes). Shown in subfigures a and c are the mean (solid lines) and 95% confidence intervals (shaded areas or error bars) of the 250 simulation results for stations 1 and 16. Shown in subfigures b and d, again for stations 1 and 16, are the distributions of the absolute error between the linearly interpolated average results for 24-h compound or daily grab sampling and the average results for the reference scenario.

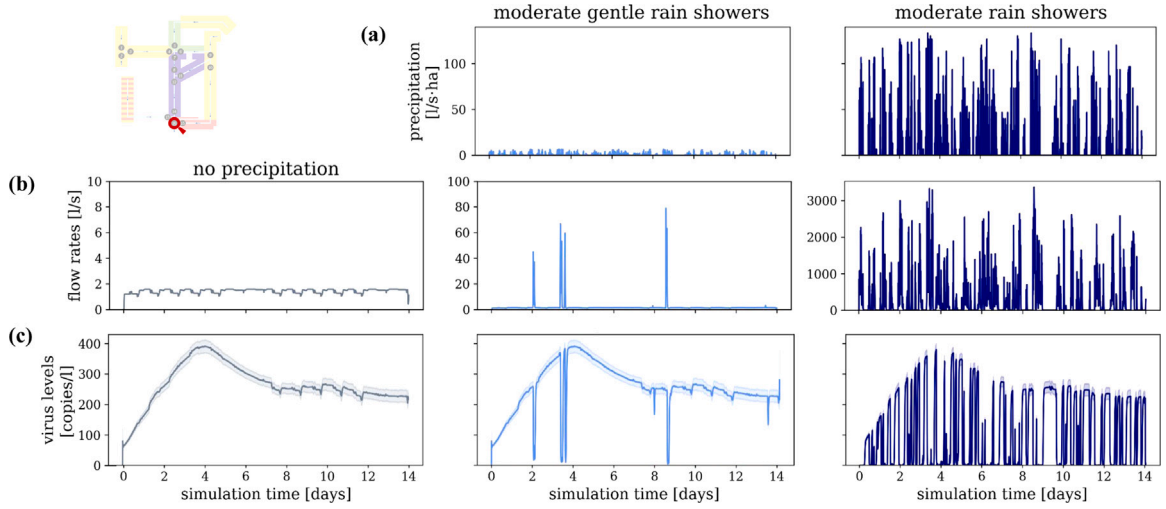

Fig. 7. Influence of Precipitation. Comparison of different rain scenarios (no precipitation, moderate gentle rain showers, and moderate rain showers) at sampling location 16 visualizing (a) the precipitation, (b) the flow rates in liters per second, and (c) the RNA concentration in wastewater in copies per liter.

For the model results presented here, rolling-ball normalization – effectively connecting the dry-weather measurements – would likely recover the shape of the viral load trajectory in the no-precipitation reference scenario, but this is an artifact of the model's three-minute temporal resolution. In reality, taking samples every three minutes would be infeasible; samples might instead be taken once per hour at most, so when a subset is affected by precipitation, the trends in the viral load trajectory that are due to changes in disease prevalence can easily be obscured. Thus, the analysis of the nonlinear influence of precipitation events on measurements showcases the importance of normalizing observations to compare measurements of precipitation and dry weather time points. For more details on normalization strategies, see the results presented in Section 3.7.

3.6. Viral decay characteristics impact wastewater monitoring non-linearly

The interplay of virus and host immune response determine virus shedding and transmission and, hence, the prevalence and influx of virus particles into the wastewater system. Yet, the virus particles are not necessarily stable but can decay. For SARS-CoV-2, estimates of the 90% reduction times in wastewater at ambient temperatures range between 5.5 and 28.8 days (Burnet et al., 2023; Guo et al., 2023; Ahmed et al., 2020). As the sewer of the synthetic neighborhood has a relatively short maximum flow time, we have been able to reasonably neglect viral decay so far. We now assume that our virus of interest has much faster virus reduction times than SARS-CoV-2, in order to study the impact of rapid decay processes on wastewater monitoring results and their relation to prevalence. To this end, we compare three different temporal models for the decay rate υ of [RNA] (in copies per liter) in wastewater:

- no decay: $\upsilon = 0 \ \frac{\text{copies}}{l \cdot s}$
- linear decay: $\upsilon = k_1$, where $k_1 = \begin{cases} -0.1 \frac{\text{copies}}{l \cdot s}, & \text{for [RNA]} > 0 \\ 0 \frac{\text{copies}}{l \cdot s}, & \text{for [RNA]} \leq 0 \end{cases}$

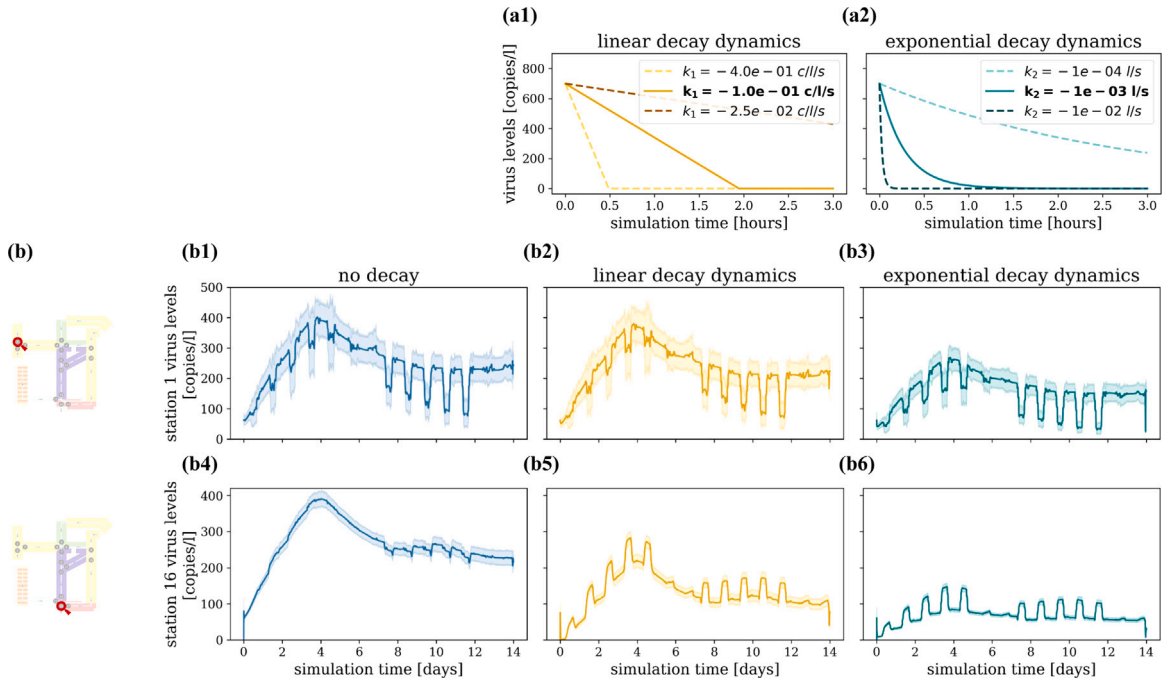

Fig. 8. Viral Decay Dynamics. (a) Illustration of the effects of linear decay (a1) vs. exponential decay (a2) with 3 example parameter settings each on a starting virus concentration of 700 copies per liter. (b) The virus concentration measured at station 1 (top) and station 16 (bottom) over time without rain and with no decay (b1/b4), linear decay with $k_1 = -0.1 \frac{\text{copies}}{l \cdot s}$ (b2/b5), and exponential decay with $k_2 = -1 \cdot 10^{-3} \frac{1}{s}$ (b3/b6).

- exponential decay: $v = k_2 \cdot [\text{RNA}]$, where $k_2 = -1 \cdot 10^{-3} \frac{1}{s}$ (corresponding to a 90% reduction time of about 2300 s or about 0.6 h)

The no decay scenario assumes that viral particles remain intact and serves as a reference indicating the upper bound for viral concentration measurements. The linear decay scenario (Fig. 8a1) is motivated by a potential interaction between a virus and a certain enzyme or environmental condition in the wastewater that degrades the virus at constant rate. The exponential decay scenario (Fig. 8a2) captures the most commonly observed viral decay dynamics, translating to a constant decay probability per unit time.

We analyzed the simulation results for the three decay scenarios, without precipitation, at two sampling locations: 1 and 16 (Fig. 8b). At the upstream station 1, compared to the no decay setting (Fig. 8b1), the measured viral loads are proportionally slightly lower for the linear decay setting (Fig. 8b2) and considerably lower for the exponential decay setting (Fig. 8b3). This is because exponential decay with a 90% reduction time of 2300 s is considerably faster than linear decay of −0.1 copies per liter per second for the relevant concentrations. However, the general shape of the viral load trajectory at station 1 over time is unaffected by the decay setting: For all three decay scenarios, there are periodic dips in the virus level on weekdays and a defined peak around the fourth day. At the downstream station 16, compared to the no decay setting (Fig. 8b4), not only are the measured viral loads lower for the linear decay setting (Fig. 8b5) and especially the exponential decay setting (Fig. 8b6), but the shape of the viral load trajectory over time is also affected, with defined peaks during weekday daytime periods. When the viral decay is non-negligible, virus copies shed from the upstream residential areas tend to decay before they reach the furthest downstream station, so station 16 primarily measures the copies shed from the nearest areas: a university and a shopping/business region, which are only inhabited during working hours. This phenomenon indicates that the interaction between viral load and viral decay can be complex and that a detailed integrative model is necessary to adequately describe its impact on the reliability of wastewater viral loads as an indicator of disease prevalence.

3.7. Virus normalization with PMMoV can outperform normalization with flow rates

Wastewater systems can have – as outlined above – variable flow rates depending on factors like precipitation and water usage. Hence, wastewater samples have to be normalized to ensure accurate and reliable data interpretation. The two most commonly used normalization strategies are based on either flow rates or additional indicators like the concentration of Pepper Mild Mottle Virus (PMMoV). PMMoV is a plant virus commonly found in human feces at relatively stable concentrations and hence, serves as a good indicator for the amount of human waste in the sample. Here, we use the model to assess which normalization strategy yields corrected viral load values closest to those one would measure if there was no precipitation event.

The normalization is calculated using one of the following two strategies:

- normalization with flow rates: $[\text{RNA}]_{\text{normalized}} = \frac{Q}{\text{mean}(Q_{\text{dry}})} \cdot [\text{RNA}]$
- normalization with PMMoV: $[\text{RNA}]_{\text{normalized}} = \frac{\text{mean}([\text{PMMoV}]_{\text{dry}})}{[\text{PMMoV}]} \cdot [\text{RNA}]$

where Q is the current flow rate, $\text{mean}(Q_{\text{dry}})$ the mean flow rate on dry days, [PMMoV] the current PMMoV concentration, and $\text{mean}([\text{PMMoV}]_{\text{dry}})$ the mean concentration of PMMoV on dry days (Mitranescu et al., 2022). Rainwater infiltration into a sewer system dilutes PMMoV while increasing flow rates, so correcting wastewater measurements using the ratio between the expected and measured PMMoV concentration or between the measured and expected flow rate can help reduce unwanted variability in wastewater-based data.

To simulate measurements of PMMoV we assumed a constant PMMoV shedding per agent throughout the simulations. Since the 90% reduction time of PMMoV is generally estimated to be on the scale of weeks or months (Mercier et al., 2024), and therefore far longer than our maximum sewer travel time, we also assumed in our simulations that PMMoV is never subject to viral decay. A few observations (between 0.15% and 0.28% per scenario), for which the measured PMMoV concentration was zero copies per liter, were removed. As seen

in Fig. 9a, flow rate normalization (right column) is not effective for correcting wastewater measurements in our model for the effects of either moderate gentle or moderate rain. In both the moderate gentle and moderate rain scenarios, infiltration of rainwater into the model sewer system causes greater proportional increases to the flow rates than decreases to the virus concentrations. Thus, normalization with flow rates leads to notable over-corrections (see Fig. 9b). In contrast, normalization with PMMoV (see Fig. 9a, center column) appears highly effective. Rainwater infiltration affects PMMoV and the virus of interest similarly, so the PMMoV-normalized measurements, while noisier, generally match the no-rain reference scenario in terms of both shape and scale. Since our PMMoV normalization approach corrects only for rainfall and not for viral decay, PMMoV normalization only marginally improves the median absolute error compared to the reference scenario for the moderate rain, exponential decay and the moderate rain, linear decay scenarios (see Fig. 9c and Supplementary Table S6). However, PMMoV appears even less effective when the viral decay is linear rather than exponential. This is likely because in the linear decay scenario, unlike the exponential decay scenario, viral concentrations can degrade to zero, which renders normalization useless for certain observations.

Rain reduces the cross-correlations between overall prevalence in the catchment area and the measured RNA copies per liter in wastewater samples over time (see Fig. 9d). For the scenarios with rain but without viral decay, applying PMMoV normalization to the wastewater measurements restores the cross-correlation coefficients to very close to their levels in the reference scenario. For the scenarios with both rain and viral decay, PMMoV normalization partially restores the general trends in viral load measurements over time, and therefore has a more limited but non-negligible impact on cross-correlations.

4. Discussion

In this study, we presented a new integrative model for the fine-grained description of infectious disease dynamics and wastewater surveillance. The model couples a stochastic model of individual mobility, infection transmission, and disease progression with a highly detailed hydrodynamic model of viral RNA transport through wastewater networks. Using this model for the description of a synthetic neighborhood and corresponding sewer system, we were able to investigate the influence of sampling protocols, precipitation events, viral decay characteristics, and normalization strategies on the relationship between infection dynamics and resulting wastewater measurements. In the context of our relatively small and socioeconomically homogeneous simulated study area, we found that locations for sampling stations should be chosen carefully, so that they lie downstream of a sufficient number of agents and diverse location types; that precipitation and viral decay can have unexpected, nonlinear impacts on wastewater viral load that require detailed integrative modeling approaches to be understood; and that flow rate normalization should only be implemented with caution as it can lead to large over-corrections if there has recently been precipitation. Overall, our study suggests that if appropriate sampling, normalization, and analysis techniques are used, then wastewater may serve as a leading indicator of disease prevalence.

Sampling station characteristics - including the size of the upstream population, the times of day during which the upstream areas are generally populated, and the distribution of sewer travel times to the station - can both qualitatively and quantitatively affect wastewater measurement dynamics. To produce the highest correlations with prevalence, our study indicates that from among the set of feasible sampling stations, public health officials should select locations that are far enough downstream to receive wastewater from a representative sample of the population of interest as well as from a mix of homes, places of work and study, and recreational areas. This aligns well with current sampling strategies, which often utilize the inflow into wastewater treatment plants, but additional locations may sometimes be useful, particularly if large heterogeneous regions are being monitored. In such cases, the selection of ancillary sampling locations must account for limitations such as accessibility, availability of electricity (if automatic sampling devices are used), and budget. For downstream locations, we find that the choice of temporal sampling design is less crucial, as 24-h compound sampling and daily grab sampling approaches provide similar results for larger and more diverse catchment areas in our study. If infections are distributed throughout the area of interest, then 24-h compound sampling can help alleviate some of the disadvantages of upstream sampling location placement, but if infections are localized, downstream placement becomes more crucial, as a sampling station too far upstream may miss the outbreak entirely. However, these considerations must be balanced against others: Our study showed how the effects of viral decay become more pronounced the further downstream a sampling station is. Furthermore, in real-world scenarios, inflow of industrial wastewater – which can dilute viral loads and contain chemicals that affect viral decay – may also be a concern. Thus, while we generally argue for mid- or downstream sampling locations, we recommend that public health officials collaborate with mathematical modelers who utilize a detailed modeling approach like ours in order to choose the optimal sampling location for a particular region of interest.

Our study also illustrates how precipitation and viral decay, separately or together, can have complex and nonlinear impacts on the measured viral load in wastewater, which need to be accounted for when using wastewater data as a public health indicator. Due to the interactions between evaporation, water retention, and other factors, the influence of rainfall on sewer flow rates and viral concentrations within different pipes can be difficult to predict; for example, our model shows how a continuous drizzle can potentially lead to discrete drops in measured viral loads in wastewater. Although using sensors to measure sewer flow rates at a sampling station and then normalizing wastewater measurement accordingly is possible, this approach performed suboptimally in our study and often led to dramatic over-corrections. We find that normalization with a human fecal indicator such as PMMoV – or, alternatively, human mitochondrial DNA, which is not diet-dependent and therefore varies less across individuals (Liu et al., 2024) – may be preferable in sewer systems that have recently been infiltrated by rainwater. This finding is in contrast to the conclusions of Rainey et al. (2023), who recommended normalization with flow rates to account for variations in the size of a sewershed's service population, but who did not consider the effects of precipitation. Future work – for example, on an expanded version of our model that includes several cities of varying size – should explore how best to account for both precipitation and population size when normalizing wastewater measurements for comparisons over time and across sampling locations.

Like precipitation, viral decay in our study led to both qualitative and quantitative changes in the viral load trajectory measured at a particular station over time. Our decay scenarios were intentionally exaggerated – we considered exponential decay with a 90% reduction time of about 0.6 h, whereas the actual 90% reduction time of SARS-CoV-2 in wastewater is temperature-dependent but usually estimated to be upwards of five days – but this extreme case illustrates how the signal from surface areas far from the sampling location can be lost if viral decay is not appropriately accounted for. In our study, we found that even in the presence of rapid viral decay, PMMoV normalization could restore non-negligible correlations between prevalence and wastewater measurements, but it could only slightly reduce the absolute error relative to a no-rain, no-decay reference scenario. Since PMMoV is very stable in wastewater, normalization with PMMoV alone cannot account for the effects of viral decay on wastewater measurements. Instead, an appropriate viral decay model should be chosen based on the characteristics of the virus of interest and the sewer system, including temperature (Hart and Halden, 2020) and biofilms (Zhang et al., 2023) – both of which are likely to be affected in turn by the amount of precipitation in the pipes. Due to its level of detail, our study provides new insights into the interactions between rain, viral

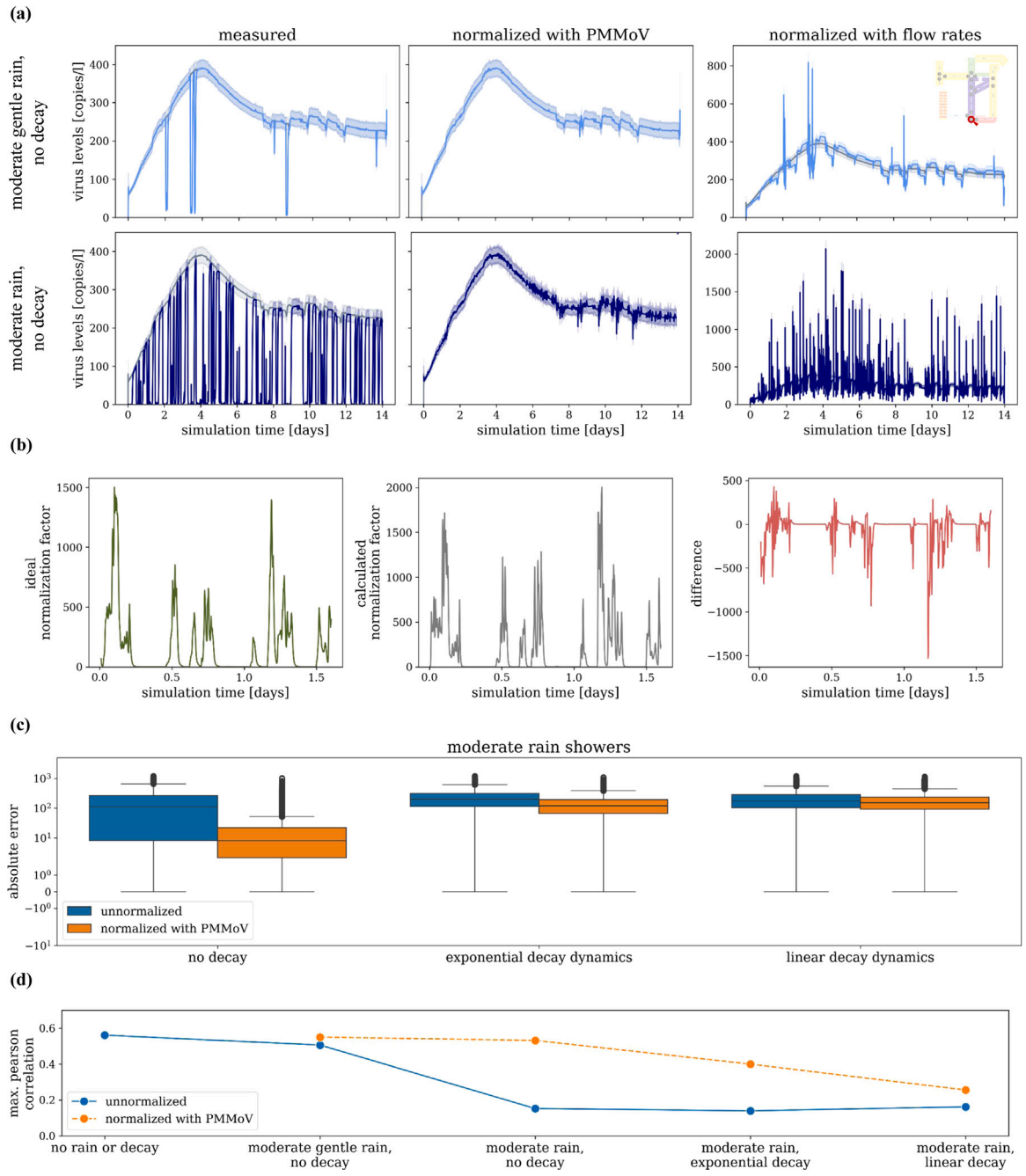

Fig. 9. Performance of Normalization Strategies. (a) The virus concentration at station 16 in the moderate gentle (top) or moderate (bottom) rain scenarios, as measured (left), after PMMoV normalization (center), and after flow rate normalization (right). The virus concentration in the no-rain, no-decay reference scenario is indicated in the background in gray. (b) Comparison of the "ideal" normalization factor that would transform the moderate rain results to the reference results, versus the actual flow-rate-based normalization factor, for one simulation. (c) The distribution over all simulations and time points of the absolute error between the moderate rain results with no, exponential, or linear decay and the reference scenario. (d) The maximal Pearson cross-correlation between wastewater viral load and prevalence for various scenarios.

decay, and wastewater measurements and underscores the importance of appropriate normalization and analysis of wastewater data.

Overall, our study indicates that despite potential confounding factors, if appropriate sampling, normalization, and analysis techniques are utilized, then wastewater-based surveillance data can provide insights into trends in disease prevalence and possibly predict outbreak peaks shortly before they occur. Viral load measurements in wastewater depend not only on the total number of people shedding, but also on the viral load – and, by extension, the infectiousness – of each prevalent infection. High viral loads in wastewater indicate high infection potential across the catchment area, meaning that new infections are likely to occur soon; thus, in our model, the peak in viral load measurements at the furthest downstream station tend to occur about 30 h before the corresponding peak in overall prevalence. Previous studies have found that epidemiological measures such as positive test counts and hospital admissions tend to lag several days behind wastewater measurements (Peccia et al., 2020; Kumar et al., 2021; Wu et al., 2022; D'Aoust et al., 2021), and our study suggests that this may not be entirely explained by reporting delays. Thus, our results support these studies' conclusion that wastewater-based surveillance data can help guide public health officials in deciding when to implement or ease infection control measures.

4.1. Limitations and future work

We assume each agent's water usage to be distributed uniformly across the day, but in reality, an individual would only produce wastewater – and, if infected, shed into the sewer system – at certain time points (e.g. while using the bathroom at home after waking up in the morning). Furthermore, water usage may vary based on the socioeconomic differences between locations (Balata et al., 2022), which we did not consider in this study. We expect that realistic patterns of water usage over time would likely create additional daily trends in wastewater measurements and highlight the advantages of 24-h compound sampling over grab samples. Future work on this model should incorporate the circadian rhythm of water consumption and shedding behavior to confirm this effect.

We also leave for future work a more complete sensitivity analysis of how relaxing the assumptions underlying our shedding and mobility models might affect our results. Reinfections, vaccinations, and cross-talk between co-circulating pathogens are out of the scope of this initial model, and we therefore do not consider, for example, how the effects of vaccinations on viral shedding (Garcia-Knight et al., 2022) or the emergence of new virus variants (Li et al., 2023) could complicate the relationship between disease dynamics and wastewater measurements. Due to data limitations, we assume that viral shedding in urine and feces is proportional to respiratory shedding, and to preclude transmissions from recovered agents, we assume that an infected individual's viral shedding stops at the time they recover – meaning that the viral load of infected agents in our model often declines more rapidly than expected based on the empirical measurements for SARS-CoV-2 in Jones et al. (2021). As more literature on SARS-CoV-2 becomes available, future work on our model should align it with the latest data on respiratory versus fecal shedding processes. Furthermore, we assume that the peak viral load value is the same for all symptomatic infections. Realistically, this value may vary significantly across individuals based on factors such as immune response (Challenger et al., 2022), and such heterogeneity might decrease the cross-correlations between prevalence and wastewater measurements when the number of infections is small.

Finally, the methods we have so far only applied in the context of a proof-of-concept study of a synthetic neighborhood will, in the future, be adapted to real-world scenarios. Preliminary tests have proven that our modeling framework can be scaled up to the level of a large city over a months-long time frame. This requires, for example, optimizing the data transfer process between model components so that information on hundreds of thousands or millions of agents is shared efficiently; performing Bayesian estimation on key unknown parameter values to ensure that modeled disease dynamics reflect empirically observed outbreaks; and introducing an appropriate noise model to account for the effects of wastewater measurement uncertainty and detection limits. Our model's potential ability to map real-world wastewater measurements back to underlying prevalence remains to be fully tested.

4.2. Conclusions

Our study illustrates the value of sophisticated models of infection and wastewater dynamics and highlights the potential of wastewater-based surveillance data to reflect trends in prevalence without being influenced by sampling bias, reporting delays, or under-ascertainment. While applications to real-world data remain for future work, our simulation study provided key insights into the advantages of downstream sampling location placement, 24-h compound sampling, models for the effects of viral decay, and PMMoV normalization.

CRediT authorship contribution statement

Nina Schmid: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Conceptualization. **Julia Bicker:** Writing – review & editing, Writing – original draft, Software, Methodology. **Andreas F. Hofmann:** Writing – review & editing, Writing – original draft, Software, Methodology, Funding acquisition. **Karina Wallrafen-Sam:** Writing – review & editing, Writing – original draft, Visualization, Formal analysis. **David Kerkmann:** Writing – review & editing, Software, Methodology. **Andreas Wieser:** Writing – review & editing, Validation, Supervision, Funding acquisition. **Martin J. Kühn:** Writing – review & editing, Validation, Supervision, Funding acquisition. **Jan Hasenauer:** Writing – review & editing, Validation, Supervision, Project administration, Funding acquisition.

Implementation and availability

The version of MEmilio used in this study, including all input files, is publicly available at https://github.com/SciCompMod/memilio/tree/inside-demonstrator-final. Access to the software ++SYSTEMS and its documentation (tandler.com GmbH, 2024) is available upon reasonable request to andreas.hofmann@tandler.com; all parameter files and simulation outputs used in this study are available at https://doi.org/10.5281/zenodo.14046493. Results were analyzed using the Python code available at https://github.com/inside-consortium/inside_demonstrator.

Funding

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (EXC 2047 – 390685813, EXC 2151 – 390873048 and under the Project MESID – 528702961), by the German Federal Ministry of Education and Research (BMBF) (INSIDe – grant numbers 031L0297A, 031L0297B, 031L0297D, and 031L0297E), and by the University of Bonn (via the Schlegel Professorship of J.H.). It was furthermore supported by the Initiative and Networking Fund of the Helmholtz Association (grant agreement number KA1-Co-08, Project LOKI-Pandemics).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.epidem.2025.100836.

Data availability

Implementation and Availability

The version of MEmilio used in this study, including all input files, is publicly available at https://github.com/SciCompMod/memilio/tree/inside-demonstrator-final. Access to the software ++SYSTEMS and its documentation tandler.com GmbH (2024) is available upon reasonable request to andreas.hofmann@tandler.com; all parameter files and simulation outputs used in this study are available at https://doi.org/10.5281/zenodo.14046493. Results were analyzed using the Python code available at https://github.com/inside-consortium/inside_demonstrator.

References

Abernethy, Gavin M., Glass, David H., 2022. Optimal COVID-19 lockdown strategies in an age-structured SEIR model of Northern Ireland. J. R. Soc. Interface 19 (188), 20210896.

Ahmed, Warish, Bertsch, Paul M., Bibby, Kyle, Haramoto, Eiji, Hewitt, Joanne, Huygens, Flavia, Gyawali, Pradip, Korajkic, Asja, Riddell, Shane, Sherchan, Samendra P., Simpson, Stuart L., Sirikanchana, Kwanrawee, Symonds, Erin M., Verhagen, Rory, Vasan, Seshadri S., Kitajima, Masaaki, Bivins, Aaron, 2020. Decay of SARS-CoV-2 and surrogate murine hepatitis virus RNA in untreated wastewater to inform application in wastewater-based epidemiology. Environ. Res. 191, 110092.

Bach, Peter M., Rauch, Wolfgang, Mikkelsen, Peter S., McCarthy, David T., Deletic, Ana, 2014. A critical review of integrated urban water modelling – Urban drainage and beyond. Environ. Model. Softw. 54, 88–107.

Balata, Edson Elidio, Pinto, Hugo, da Silva, Manuela Moreira, 2022. Latent dimensions between water use and socio-economic development: A global exploratory statistical analysis. Reg. Sustain. 3 (3), 269–280.

Bayerisches Landesamt für Umwelt (LfU), 2024. NiedSimBy: Synthetische Niederschlagsreihen - LfU Bayern. https://www.lfu.bayern.de/wasser/abwasser_synthetische_niederschlagsreihen/index.htm. (Accessed 20 February 2024).

Bicker, Julia, Schmieding, René, Meyer-Hermann, Michael, Kühn, Martin J., 2025. Hybrid metapopulation agent-based epidemiological models for efficient insight on the individual scale: A contribution to green computing. Infect. Dis. Model. 10 (2), 571–590.

Burnet, Jean-Baptiste, Cauchie, Henry-Michel, Walczak, Cécile, Goeders, Nathalie, Ogorzaly, Leslie, 2023. Persistence of endogenous RNA biomarkers of SARS-CoV-2 and PMMoV in raw wastewater: Impact of temperature and implications for wastewater-based epidemiology. Sci. Total Environ. 857 (Part 2), 159401.

Castonguay, François M., Blackwood, Julie C., Howerton, Emily, Shea, Katriona, Sims, Charles, Sanchirico, James N., 2023. Optimal spatial evaluation of a pro rata vaccine distribution rule for COVID-19. Sci. Rep. 13 (1), 2194.

Chae, Min-Kyung, Hwang, Dong-Uk, Nah, Kyeongah, Son, Woo-Sik, 2023. Evaluation of COVID-19 intervention policies in South Korea using the stochastic individual-based model. Sci. Rep. 13 (1), 18945.

Challenger, Joseph D., Foo, Cher Y., Wu, Yue, Yan, Ada W.C., Marjaneh, Mahdi Moradi, Liew, Felicity, Thwaites, Ryan S., Okell, Lucy C., Cunnington, Aubrey J., 2022. Modelling upper respiratory viral load dynamics of SARS-CoV-2. BMC Med. 20, 25.

Champredon, David, Papst, Irena, Yusuf, Warsame, 2024. Ern: An r package to estimate the effective reproduction number using clinical and wastewater surveillance data. PLoS ONE 19 (6), e0305550.

Chau, K.K., Barker, L., Budgell, E.P., Vihta, K.D., Sims, N., Kasprzyk-Hordern, B., Harriss, E., Crook, D.W., Read, D.S., Walker, A.S., Stoesser, N., 2022. Systematic review of wastewater surveillance of antimicrobial resistance in human populations. Environ. Int. 162, 107171.

Contento, Lorenzo, Castelletti, Noemi, Raimúndez, Elba, Le Gleut, Ronan, Schälte, Yannik, Stapor, Paul, Hinske, Ludwig Christian, Hoelscher, Michael, Wieser, Andreas, Radon, Katja, Fuchs, Christiane, Hasenauer, Jan, 2023. Integrative modelling of reported case numbers and seroprevalence reveals time-dependent test efficiency and infectious contacts. Epidemics 43, 100681.

D'Aoust, Patrick M., Graber, Tyson E., Mercier, Elisabeth, Montpetit, Danika, Alexandrov, Ilya, Neault, Nafisa, Baig, Aiman Tariq, Mayne, Janice, Zhang, Xu, Alain, Tommy, Servos, Mark R., Srikanthan, Nivetha, MacKenzie, Malcolm, Figeys, Daniel, Manuel, Douglas, Jüni, Peter, MacKenzie, Alex E., Delatolla, Robert, 2021. Catching a resurgence: Increase in SARS-CoV-2 viral RNA identified in wastewater 48 h before COVID-19 clinical tests and 96 h before hospitalizations. Sci. Total Environ. 770, 145319.

Deutscher Wetterdienst, 2024. Wetter und Klima - Deutscher Wetterdienst - Glossar - N - Niederschlagsintensität. https://www.dwd.de/DE/service/lexikon/Functions/glossar.html?lv3=101906&lv2=101812. (Accessed 20 February 2024).

Endo, Akira, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Leclerc, Quentin J., Knight, Gwenan M., Medley, Graham F., Atkins, Katherine E., Funk, Sebastian, Kucharski, Adam J., 2021. Implication of backward contact tracing in the presence of overdispersed transmission in COVID-19 outbreaks. Wellcome Open Res. 5, 239.

Fachgebiet 32, Robert Koch-Institut, 2024. Abwassersurveillance AMELAG. https://zenodo.org/records/14137715. (Accessed 20 November 2024).

Garcia-Knight, Miguel, Anglin, Khamal, Tassetto, Michel, Lu, Scott, Zhang, Amethyst, Goldberg, Sarah A., Catching, Adam, Davidson, Michelle C., Shak, Joshua R., Romero, Mariela, Pineda-Ramirez, Jesus, Diaz-Sanchez, Ruth, Rugart, Paulina, Donohue, Kevin, Massachi, Jonathan, Sans, Hannah M., Djomaleu, Manuella, Mathur, Sujata, Servellita, Venice, McIlwain, David, Gaudiliere, Brice, Chen, Jessica, Martinez, Enrique O., Tavs, Jacqueline M., Bronstone, Grace, Weiss, Jacob, Watson, John T., Briggs-Hagen, Melissa, Abedi, Glen R., Rutherford, George W., Deeks, Steven G., Chiu, Charles, Saydah, Sharon, Peluso, Michael J., Midgley, Claire M., Martin, Jeffrey N., Andino, Raul, Kelly, J. Daniel, 2022. Infectious viral shedding of SARS-CoV-2 Delta following vaccination: A longitudinal cohort study. PLoS Pathog. 18 (9), e1010802.

Gudina, Esayas Kebede, Elsbernd, Kira, Yilma, Daniel, Kisch, Rebecca, Wallrafen-Sam, Karina, Abebe, Gemeda, Mekonnen, Zeleke, Berhane, Melkamu, Gerbaba, Mulusew, Suleman, Sultan, Mamo, Yoseph, Rubio-Acero, Raquel, Ali, Solomon, Zeynudin, Ahmed, Merkt, Simon, Hasenauer, Jan, Chala, Temesgen Kabeta, Wieser, Andreas, Kroidl, Arne, 2024. Tailoring COVID-19 vaccination strategies in high-seroprevalence settings: Insights from ethiopia. Vaccines 12 (7), 745.

Guo, Ying, Liu, Yanchen, Gao, Shuhong, Zhou, Xu, Sivakumar, Muttucumaru, Jiang, Guangming, 2023. Effects of temperature and water types on the decay of coronavirus: A review. Water 15 (6), 1051.

Hart, Olga E., Halden, Rolf U., 2020. Computational analysis of SARS-CoV-2/COVID-19 surveillance by wastewater-based epidemiology locally and globally: Feasibility, economy, opportunities and challenges. Sci. Total Environ. 730, 138875.

Hinch, Robert, Probert, William J.M., Nurtay, Anel, Kendall, Michelle, Wymant, Chris, Hall, Matthew, Lythgoe, Katrina, Bulas Cruz, Ana, Zhao, Lele, Stewart, Andrea, Ferretti, Luca, Montero, Daniel, Warren, James, Mather, Nicole, Abueg, Matthew, Wu, Neo, Legat, Olivier, Bentley, Katie, Mead, Thomas, Van-Vuuren, Kelvin, Feldner-Busztin, Dylan, Ristori, Tommaso, Finkelstein, Anthony, Bonsall, David G., Abeler-Dörner, Lucie, Fraser, Christophe, 2021. OpenABM-Covid19 – An agent-based model for non-pharmaceutical interventions against COVID-19 including contact tracing. PLoS Comput. Biol. 17 (7), e1009146.

Hodges, Ben R., 2019. Conservative finite-volume forms of the Saint-Venant equations for hydrology and urban drainage. Hydrol. Earth Syst. Sci. 23 (3), 1281–1304.

Hoffmann, Till, Alsing, Justin, 2023. Faecal shedding models for SARS-CoV-2 RNA among hospitalised patients and implications for wastewater-based epidemiology. J. R. Stat. Soc. 72 (2), 330–345.

Howe, Laura D., Tilling, Kate, Galobardes, Bruna, Lawlor, Debbie A., 2013. Loss to follow-up in cohort studies: Bias in estimates of socioeconomic inequalities. Epidemiology 24 (1), 1–9.

Hunter, Elizabeth, Mac Namee, Brian, Kelleher, John, 2020. A hybrid agent-based and equation based model for the spread of infectious diseases. J. Artif. Soc. Soc. Simul. 23 (4), 14.

Jones, Terry C., Biele, Guido, Mühlemann, Barbara, Veith, Talitha, Schneider, Julia, Beheim-Schwarzbach, Jörn, Bleicker, Tobias, Tesch, Julia, Schmidt, Marie Luisa, Sander, Leif Erik, Kurth, Florian, Menzel, Peter, Schwarzer, Rolf, Zuchowski, Marta, Hofmann, Jörg, Krumbholz, Andi, Stein, Angela, Edelmann, Anke, Corman, Victor Max, Drosten, Christian, 2021. Estimating infectiousness throughout SARS-CoV-2 infection course. Science 373 (6551), eabi5273.

Kerkmann, David, Korf, Sascha, Nguyen, Khoa, Abele, Daniel, Schengen, Alain, Gerstein, Carlotta, Göbbert, Jens Henrik, Basermann, Achim, Kühn, Martin J., Meyer-Hermann, Michael, 2024. Agent-based modeling for realistic reproduction of human mobility and contact behavior to evaluate test and isolation strategies in epidemic infectious disease spread. arXiv2410.08050 Prepublished.

Kerr, Cliff C., Stuart, Robyn M., Mistry, Dina, Abeysuriya, Romesh G., Rosenfeld, Katherine, Hart, Gregory R., Núñez, Rafael C., Cohen, Jamie A., Selvaraj, Prashanth, Hagedorn, Brittany, George, Lauren, Jastrzebski, Michal, Izzo, Amanda S., Fowler, Greer, Palmer, Anna, Delport, Dominic, Scott, Nick, Kelly, Sherrie L., Bennette, Caroline S., Wagner, Bradley G., Chang, Stewart T., Oron, Assaf P., Wenger, Edward A., Panovska-Griffiths, Jasmina, Famulare, Michael, Klein, Daniel J., 2021. Covasim: An agent-based model of COVID-19 dynamics and interventions. PLoS Comput. Biol. 17 (7), e1009149.

Kilaru, Pruthvi, Hill, Dustin, Anderson, Kathryn, Collins, Mary B., Green, Hyatt, Kmush, Brittany L., Larsen, David A., 2023. Wastewater surveillance for infectious disease: A systematic review. Am. J. Epidemiol. 192 (2), 305–322.

Kühn, Martin J., Abele, Daniel, Kerkmann, David, Korf, Sascha A., Zunker, Henrik, Wendler, Anna C., Bicker, Julia, Nguyen, Khoa, Schmieding, René, Plötzke, Lena, Lenz, Patrick, Betz, Maximilian F., Gerstein, Carlotta, Schmidt, Agatha, Johannssen, Paul, Klitz, Margrit, Koslow, Wadim, Binder, Sebastian, Siggel, Martin, Kleinert, Jan, Rack, Kathrin, Lutz, Annette, Meyer-Hermann, Michael, 2023. MEmilio v1.0.0 - A high performance modular EpideMIcs simulation software. https://zenodo.org/records/10412635. (Accessed: 20 November 2024).

Kühn, Martin J., Abele, Daniel, Mitra, Tanmay, Koslow, Wadim, Abedi, Majid, Rack, Kathrin, Siggel, Martin, Khailaie, Sahamoddin, Klitz, Margrit, Binder, Sebastian, Spataro, Luca, Gilg, Jonas, Kleinert, Jan, Häberle, Matthias, Plötzke, Lena, Spinner, Christoph D., Stecher, Melanie, Zhu, Xiao Xiang, Basermann, Achim, Meyer-Hermann, Michael, 2021. Assessment of effective mitigation and prediction of the spread of SARS-CoV-2 in Germany using demographic information and spatial resolution. Math. Biosci. 339, 108648.

Kumar, Manish, Joshi, Madhvi, Patel, Arbind Kumar, Joshi, Chaitanya G., 2021. Unravelling the early warning capability of wastewater surveillance for COVID-19: A temporal study on SARS-CoV-2 RNA detection and need for the escalation. Environ. Res. 196, 110946.

Larremore, Daniel B., Wilder, Bryan, Lester, Evan, Shehata, Soraya, Burke, James M., Hay, James A., Tambe, Milind, Mina, Michael J., Parker, Roy, 2021. Test sensitivity is secondary to frequency and turnaround time for COVID-19 screening. Sci. Adv. 7 (1), eabd5393.

Latsuzbaia, Ardashel, Herold, Malte, Bertemes, Jean-Paul, Mossong, Joël, 2020. Evolving social contact patterns during the COVID-19 crisis in Luxembourg. PLoS One 15 (8), e0237128.

Lau, H., Khosrawipour, T., Kocbach, P., Ichii, H., Bania, J., Khosrawipour, V., 2021. Evaluating the massive underreporting and undertesting of COVID-19 cases in multiple global epicenters. Pulmonology 27 (2), 110–115.

Li, Fanglin, Deng, Jiayi, Xie, Canbin, Wang, Guyi, Xu, Min, Wu, Chenfang, Li, Jinxiu, Zhong, Yanjun, 2023. The differences in virus shedding time between the Delta variant and original SARS-CoV-2 infected patients. Front. Public Heal. 11, 1132643.

Liu, Pengbo, Sablon, III, Orlando, Wang, Yuke, Hilton, Stephen Patrick, Khalil, Lana, Ingersoll, Jessica Mae, Truell, Jennifer, Edupuganti, Sri, Alaaeddine, Ghina, Naji, Amal, Monarrez, Eduardo, Wolfe, Marlene, Rouphael, Nadine, Kraft, Colleen, Moe, Christine L., 2024. Longitudinal fecal shedding of SARS-CoV-2, pepper mild mottle virus, and human mitochondrial DNA in COVID-19 patients. Front. Med. 11, 1417967.

Lorch, Lars, Kremer, Heiner, Trouleau, William, Tsirtsis, Stratis, Szanto, Aron, Schölkopf, Bernhard, Gomez-Rodriguez, Manuel, 2022. Quantifying the effects of contact tracing, testing, and containment measures in the presence of infection hotspots. ACM Trans. Spat. Algorithms Syst. 8 (4).

Mercier, Élisabeth, D'Aoust, Patrick M., Eid, Walaa, Hegazy, Nada, Kabir, Pervez, Wan, Shen, Pisharody, Lakshmi, Renouf, Elizabeth, Stephenson, Sean, Graber, Tyson E., MacKenzie, Alex E., Delatolla, Robert, 2024. Sewer transport conditions and their role in the decay of endogenous SARS-CoV-2 and pepper mild mottle virus from source to collection. Int. J. Hyg. Environ. Heal. 263, 114477.

Merkt, Simon, Ali, Solomon, Gudina, Esayas Kebede, Adissu, Wondimagegn, Gize, Addisu, Muenchhoff, Maximilian, Graf, Alexander, Krebs, Stefan, Elsbernd, Kira, Kisch, Rebecca, Betizazu, Sisay Sirgu, Fantahun, Bereket, Bekele, Delayehu, Rubio-Acero, Raquel, Gashaw, Mulatu, Girma, Eyob, Yilma, Daniel, Zeynudin, Ahmed, Paunovic, Ivana, Hoelscher, Michael, Blum, Helmut, Hasenauer, Jan, Kroidl, Arne, Wieser, Andreas, 2024. Long-term monitoring of SARS-CoV-2 seroprevalence and variants in Ethiopia provides prediction for immunity and cross-immunity. Nat. Commun. 15 (1), 3463.

Millard, Louise A.C., Fernández-Sanlés, Alba, Carter, Alice R., Hughes, Rachael A., Tilling, Kate, Morris, Tim P., Major-Smith, Daniel, Griffith, Gareth J., Clayton, Gemma L., Kawabata, Emily, Davey Smith, George, Lawlor, Deborah A., Borges, Maria Carolina, 2023. Exploring the impact of selection bias in observational studies of COVID-19: a simulation study. Int. J. Epidemiol. 52 (1), 44–57.

Mitranescu, Alexander, Uchaikina, Anna, Kau, Anna-Sonia, Stange, Claudia, Ho, Johannes, Tiehm, Andreas, Wurzbacher, Christian, Drewes, Jörg E., 2022. Wastewater-based epidemiology for SARS-CoV-2 biomarkers: Evaluation of normalization methods in small and large communities in southern Germany. ACS ES& T Water 2 (12), 2460–2470.

Nourbakhsh, Shokoofeh, Fazil, Aamir, Li, Michael, Mangat, Chand S., Peterson, Shelley W., Daigle, Jade, Langner, Stacie, Shurgold, Jayson, D'Aoust, Patrick, Delatolla, Robert, Mercier, Elizabeth, Pang, Xiaoli, Lee, Bonita E., Stuart, Rebecca, Wijayasri, Shinthuja, Champredon, David, 2022. A wastewater-based epidemic model for SARS-CoV-2 with application to three Canadian cities. Epidemics 39, 100560.

Núñez, Matías, Barreiro, Nadia L., Barrio, Rafael A., Rackauckas, Christopher, 2023. Forecasting virus outbreaks with social media data via neural ordinary differential equations. Sci. Rep. 13 (1), 10870.

Oraby, Tamer, Tyshenko, Michael G., Maldonado, Jose Campo, Vatcheva, Kristina, Elsaadany, Susie, Alali, Walid Q., Longenecker, Joseph C., Al-Zoughool, Mustafa, 2021. Modeling the effect of lockdown timing as a COVID-19 control measure in countries with differing social contacts. Sci. Rep. 11 (1), 3354.

Peccia, Jordan, Zulli, Alessandro, Brackney, Doug E., Grubaugh, Nathan D., Kaplan, Edward H., Casanovas-Massana, Arnau, Ko, Albert I., Malik, Amyn A., Wang, Dennis, Wang, Mike, Warren, Joshua L., Weinberger, Daniel M., Arnold, Wyatt, Omer, Saad B., 2020. Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. Nature Biotechnol. 38 (10), 1164–1167.

Raimúndez, Elba, Dudkin, Erika, Vanhoefer, Jakob, Alamoudi, Emad, Merkt, Simon, Fuhrmann, Lara, Bai, Fan, Hasenauer, Jan, 2021. COVID-19 outbreak in wuhan demonstrates the limitations of publicly available case numbers for epidemiological modeling. Epidemics 34, 100439.

Rainey, Andrew L., Liang, Song, Bisesi, Jr., Joseph H., Sabo-Attwood, Tara, Maurelli, Anthony T., 2023. A multistate assessment of population normalization factors for wastewater-based epidemiology of COVID-19. PLoS ONE 18 (4), e0284370.

Ranta, J., Hovi, T., Arjas, E., 2001. Poliovirus surveillance by examining sewage water specimens: studies on detection probability using simulation models. Risk Anal. 21 (6), 1087–1096.

Schütze, Manfred R., Butler, David, Beck, M. Bruce, 2002. Modelling, Simulation and Control of Urban Wastewater Systems. Springer, London, pp. 129–178.

Siedner, Mark J., Harling, Guy, Reynolds, Zahra, Gilbert, Rebecca F., Haneuse, Sebastien, Venkataramani, Atheendar S., Tsai, Alexander C., 2020. Social distancing to slow the US COVID-19 epidemic: Longitudinal pretest–posttest comparison group study. PLoS Med. 17 (8), e1003244.

Statistisches Bundesamt, 2024. Mikrozensus. https://www-genesis.destatis.de/genesis/online?operation=statistic&levelindex=&levelid=&code=12211. (Accessed 31 July 2024).

tandler.com GmbH, 2024. DYNA - Komplexes Parallelschrittverfahren - Verfahrensbeschreibung. Technical Report.

Tracy, Melissa, Cerdá, Magdalena, Keyes, Katherine M., 2018. Agent-based modeling in public health: Current applications and future directions. Annu. Rev. Public. Health 39, 77–94.

Wade, Matthew J., Jacomo, Anna Lo, Armenise, Elena, Brown, Mathew R., Bunce, Joshua T., Cameron, Graeme J., Fang, Zhou, Farkas, Kata, Gilpin, Deidre F., Graham, David W., Grimsley, Jasmine M.S., Hart, Alwyn, Hoffmann, Till, Jackson, Katherine J., Jones, David L., Lilley, Chris J., McGrath, John W., McKinley, Jennifer M., McSparron, Cormac, Nejad, Behnam F., amd Marcos Quintela-Baluja, Mario Morvan, Roberts, Adrian M.I., Singer, Andrew C., Souque, Celia, Speight, Vanessa L., Sweetapple, Chris, Walker, David, Watts, Glenn, Weightman, Andrew, Kasprzyk-Hordern, Barbara, 2022. Understanding and managing uncertainty and variability for wastewater monitoring beyond the pandemic: Lessons learned from the United Kingdom national COVID-19 surveillance programmes. J. Hazard. Mater. 424, 127456.

Wang, Yuke, Moe, Christine L., Dutta, Shanta, Wadhwa, Ashutosh, Kanungo, Suman, Mairinger, Wolfgang, Zhao, Yichuan, Jiang, Yi, Teunis, Peter F.M., 2020. Designing a typhoid environmental surveillance study: A simulation model for optimum sampling site allocation. Epidemics 31, 100391.

Wu, Fuqing, Xiao, Amy, Zhang, Jianbo, Moniz, Katya, Endo, Noriko, Armas, Federica, Bonneau, Richard, Brown, Megan A., Bushman, Mary, Chai, Peter R., Duvallet, Claire, Erickson, Timothy B., Foppe, Katelyn, Ghaeli, Newsha, Gu, Xiaoqiong, Hanage, William P., Huang, Katherine H., Lee, Wei Lin, Matus, Mariana, McElroy, Kyle A., Nagler, Jonathan, Rhode, Steven F., Santillana, Mauricio, Tucker, Joshua A., Wuertz, Stefan, Zhao, Shijie, Thompson, Janelle, Alm, Eric J., 2022. SARS-CoV-2 RNA concentrations in wastewater foreshadow dynamics and clinical presentation of new COVID-19 cases. Sci. Total Environ. 805, 150121.

Wu, Fuqing, Zhang, Jianbo, Xiao, Amy, Gu, Xiaoqiong, Lee, Wei Lin, Armas, Federica, Kauffman, Kathryn, Hanage, William, Matus, Mariana, Ghaeli, Newsha, Endo, Noriko, Duvallet, Claire, Poyet, Mathilde, Moniz, Katya, Washburne, Alex D., Erickson, Timothy B., Chai, Peter R., Thompson, Janelle, Alm, Eric J., 2020. SARS-CoV-2 titers in wastewater are higher than expected from clinically confirmed cases. MSystems 5 (4), e00614–20.

Zhang, Shuxin, Sharma, Elipsha, Tiwari, Ananda, Chen, Yan, Sherchan, Samendra P., Gao, Shuhong, Zhou, Xu, Shi, Jiahua, Jiang, Guangming, 2023. The reduction of SARS-CoV-2 RNA concentration in the presence of sewer biofilms. Water 15 (11), 2132.

Zunker, Henrik, Schmieding, René, Kerkmann, David, Schengen, Alain, Diexer, Sophie, Mikolajczyk, Rafael, Meyer-Hermann, Michael, Kühn, Martin J., 2024. Novel travel time aware metapopulation models: A combination with multi-layer waning immunity to assess late-phase epidemic and endemic scenarios. MedRxiv 2024.03.01.24303602 Prepublished.