Evaluation of Intermodel Equivalence for Misinformation Diffusion: Equation-Based vs. Agent-Based Models

Cristian Berceanu[a], Francesco Bertolotti[b], Nadia Arshad[c], Monica Patrascu[a,d]

[a]*Complex Systems Laboratory, Department of Automatic Control and Systems Engineering, University Politehnica of Bucharest, Splaiul Independenței 313, Bucharest, 060042, Romania*
[b]*School of Industrial Engineering, LIUC Università Cattaneo, Corso G. Matteotti 22, Castellanza, 21053, Italy*
[c]*Section for Epidemiology and Medical Statistics, Department of Global Public Health and Primary Care, University of Bergen, Årstadveien 17, Bergen, 5009, Norway*
[d]*Centre for Elderly and Nursing Home Medicine and Neuro-SysMed Center, Department of Global Public Health and Primary Care, University of Bergen, Årstadveien 17, Bergen, 5009, Norway*

Abstract

With the rise of ubiquitous internet-based communication, understanding the spread of false information has become crucial, particularly during global crises, such as the COVID-19 pandemic. This study delves into the diffusion mechanisms of misinformation (unintentionally false information) and disinformation (intentionally false information) by assessing the intermodel equivalence between two distinct techniques: equation-based models (EBMs), which use equations to represent system behavior, and agent-based models (ABMs), which focus on individual agents within the system and also

Email addresses: cristian.berceanu@stud.acs.upb.ro (Cristian Berceanu), fbertolotti@liuc.it (Francesco Bertolotti), nadia.arshad@uib.no (Nadia Arshad), monica.patrascu@uib.no (Monica Patrascu)

consider individualized traits. We developed two ABMs tailored to reflect the classical EBMs of information diffusion, and compared their outputs for various configurations. Results show a weak relationship between the two types of models across most parameter ranges, even if, under certain conditions, the outcomes of the two models are aligned. These findings remark on the effect of structural differences between EBMs and ABMs, emphasizing the importance of using ABMs when considering the psycho-social aspects of infodemics.

Keywords:

misinformation, infodemics, agent-based models, equation-based model, complex systems, model evaluation

---

1. Introduction

In 1995, a year that highlights the rising wave of internet-related development, Kevin Kelly was writing [1]: "The central act of the coming era is to connect everything to everything." Nearly 30 years later, we are witnessing an unprecedented upsurge in internet-based social media platforms and instruments for person-to-person and person-to-group communication. Humans, more so than before, are now connected in ways that foster high and ubiquitous access to information and to each other [2]. Our social networks are larger than in previous decades, in both number of actors (nodes) and ties (edges), which makes way for new questions and calls for new analysis methodologies [3]. The field of complex systems provides such opportunities in the form of network science, simulation of computational models, and specific concepts (e.g., small worlds, co-evolution, emergence, etc.) [4, 5].

Social communities can be considered multi-layered networks in which people interact through a variety of relationships. This interpretation allows them to be investigated as complex networks [6, 7]. When considering the available communication channels (phone, messaging apps, social media platforms, etc.), the complexity of these systems increases, both in size and relationship heterogeneity [8]. Topology-based modeling is already intricate [9], but when we must also account for the content of the information exchange between nodes, modeling social networks overlaid onto social media or communication networks becomes even more challenging. In a complex systems perspective, humans can be seen as systems with agency (capable of autonomous decision-making and reasoning), while the interactions between them are interpreted as signals: either information, energy, or matter over

time. For the specific case of information diffusion, this approach allows us to consider the exchanged content.

In complex system science, the chosen level of analysis typically depends on the modeling purpose and can significantly impact the overall results [10, 11]. To enhance scientific communication and ensure replicability, researchers often employ a simplifying interpretative method, which results into a model on one of three scales [12]: a) macroscopic, when the whole network is represented as a black-box (i.e., without understanding its internal behaviour), and only large-scale variations of outcomes are observed; b) mesoscopic, in which some parts of the network are differentiated, but the level of granularity does not permit modeling individual nodes; and c) microscopic, when each entity in the network has agency and manages its own dynamics and psycho-social interactions.

Given this framework, the study of social-network information diffusion on a microscopic level is of increasing interest. Network science has been investigating the emergence of social ties and how ideas could travel through internet-enabled media, such as political blog analysis [13] and election manipulation [14]. During the COVID-19 pandemic, the diffusion of untrue or malicious information resulted in an infodemic [15], and a co-evolution effect has been observed with the spread of the virus itself [16]. Consequently, analysing and predicting the effects of infodemics is of high interest [17], and efforts have been focusing on integrating epidemiological models of viral spread with those of information diffusion [16, 18].

In this paper we investigate two model types for the spread of mis- and disinformation: macroscopic equation-based models (EBMs) and microscopic

agent-based models (ABMs). Our main research question is whether these two model types are equivalent. We hypothesise that microscopic models are needed to capture the psycho-social dynamics of human interaction.

Thus, through this study we aim to examine the inter-model equivalence of EBMs and ABMs. First, we present the main epidemiology-inspired EBMs. Second, we describe a set of three ABMs designed to match the EBMs. We perform a detailed cross-correlation analysis over the entire parameter range for all models, and finally we present the advantages and disadvantages of EBMs and ABMs for this information diffusion phenomenon.

Based on six classical viral spread EBMs, we implemented two ABMs versions of the six models that can be parameterized to entail mis- and disinformation behaviors at microscopic level, defined by rules correlative to the EBMs. This paper presents the ABMs and a detailed cross-correlation analysis over the entire parameter range for all six models.

The paper is organized as follows. Section 2 covers the background. Section 3 describes the study methods and the ABM design. Section 4 presents the results of the cross-correlation analysis between EBMs and ABMs, with a discussion in Section 5. Finally, Section 6 concludes the paper.

2. Background

2.1. Information Diffusion Models

A diffusion model is framework, usually implemented mathematically or computationally, that represents the spreading of entities or features in a space [19], which can be either topological or relational (i.e., a network) [20]. Even if they were first developed in physical sciences, social sciences

have widely adopted them, assuming that different kind of features could spread within a population [21]. Diffusion models provide a systematic way to understand and analyze the dynamics of spread processes as a means of predicting their outcomes or investigating the factors that influence diffusion [22]. Diffusion models are well established in different disciplinary areas, such as epidemics [23, 24], information spread [25, 26], opinion and strategy dynamics [27], and the economics of innovation [28, 29]. In particular, the epidemic models derived from Kermack and McKendrick's seminal works [30, 31, 32] have received increased attention. The underlying assumptions of these models are: a) an entity within a population could exist in one specific state (i.e., susceptible, infected, recovered, or exposed); b) an infection could be transmitted by proximity to another individual (i.e., in the classic susceptible-infected-recovered SIR model the number of newly infected individuals at time t depends on the product of susceptible and infected individuals, which stands for the number of interactions); and c) any state could be reached by internal dynamics. The infodemic models equate "infection" with "knowledge", thus proposing an analogous interpretation [33].

Diffusion models are generally characterized by four elements: the structure of the interactions, dynamics of the interactions, possible states that model the diffused feature, and their dynamics. A classical approach to the diffusion model, for instance, the one adopted by Kermack and McKendrick [30], derived from Newtonian mechanics, consists of modeling a population as state variables, which typically varies over time according to a differential equation, such that the structure of the model is an oriented graph where each node is a population subset in a given state, and the connections depict

functional dependencies. This approach is often called equation-based modeling [34, 35]. In contrast, in the last 30 years, an opposite paradigm has emerged, where the atomic unit of the model is not the state but the individual decision-making entity itself [36], usually represented as a computational object [37]. When these computational entities are agents (i.e., systems with agency [38]), the methodology is known as agent-based modeling [39, 40].

In spread modeling, equation- and agent-based approaches are often seen as concurrent, with complementary strengths and weaknesses [41], and there has been a long discussion regarding the preference of each of these methods [42]. On the one hand, EBMs are much less computationally expensive, at the cost of assuming homogeneity of sub-populations and using a mean-field approximation to the structure of interactions [43]. On the other hand, the smaller the population, the more individual features matter, and employing ABMs adheres closer to reality [44, 45].

In epidemiology, there is an observable interest in assessing which methodology is better [41, 46, 47, 48, 49] and in identifying strategies to develop hybrid models that encompass the best of both [44, 50, 51]. However, very few studies establish a correlation between epidemics and infodemics [52].

2.2. Infodemics

According to the World Health Organization, "an infodemic is too much information including false or misleading information in digital and physical environments during a disease outbreak." [53] Based on the degree of misleading intentionality, diffused untruths can be classified into [54, 55]: a) misinformation, which spreads without the intention to mislead; b) disinformation, which is produced purposely to cause harm (e.g., reporting ma-

nipulated statistics); c) malinformation, which refers to broadcasting true information to cause harm, such as circulating a report without its original context; d) rumor, which concerns the distribution of unverified information; and e) fake news, which are fabricated information mimicking news content.

On the one hand, gossip and rumour sharing builds and maintains social ties. The drive to express views and perspectives in unreliable circumstances fuels the spread of mis- and disinformation [56], which has had detrimental effects on public health and political and economic issues [57]. Notably, the waves of misinformation associated with the COVID-19 pandemic have raised multiple psychological and psycho-social issues that led to inappropriate measures, political instability, and mistrust in governing bodies [58, 59, 60, 61]. Even when scientific inaccuracies are confined within seemingly closed communities [62], they have widespread and destructive effects on social groups, especially marginalized ones, leading to widescale societal changes and unrest [63].

On the other hand, misinformation and its siblings affect the reliability and trust in social networks as avenues for the dissemination of news or other important verified facts. In general, the spread of untruths intended to manipulate the perceptions of users has been recognized as a fundamental issue in democratic societies [64]. Subsequently, detecting falsity in diffused information has become an important concern, and studies are looking to minimize affected users and reduce propagation [65].

We surmise that proper analysis tools and models are necessary, to study not only the paths of misinformation but also its long-term impact on human psycho-social behavior.

3. Methods

In this section we describe the equation- and agent-based models for mis- and disinformation diffusion, the comparative analysis method, and tools for model implementation and development. All models and results are available in the repository (will be de-anonymized after review): `https://anonymous.4open.science/r/results-big-data-soc`.

3.1. Preliminary Analysis: a Psychological Perspective on the Infodemic Infection Mechanisms

Infodemic models have been relying on the concept of "misinformation epidemic" [66], with subsequent deterministic models based on the epidemic SIR-class equations. While this assumptions is not without merit, information travels in an inherently different manner than biological viruses, being affected by how communication is carried out and by individualized factors such as the illusion of knowledge [67].

Thus, in this section we analyze whether the four states of susceptible, exposed, infected and recovered are suitable for describing the spread of mis- and disinformation. Research on the psychology of misinformation has provided much insight into the psychological processes underlying susceptibility to misinformation in multiple domains [68, 69]. Cognitive factors that contribute to supporting false beliefs include intuitive thinking (a lack of analytical thinking), cognitive failures (forget sources), and illusory truth (familiarity).

A recent model [70] proposes four psychological processes underlying susceptibility to health misinformation. Not only individuals with a capability to

reason accurately are less susceptible, but also their resilience increases with the motivation to reason accurately. In contrast, directionally- and identity-motivated reasoning increases susceptibility, derived from a desire to reach a preferred conclusion that is often consistent with one's pre-existing views. These intertwining mechanisms are not possible to be explicitly included in EBMs, but they are suitable for ABMs.

The exposed state is that in which a piece of information (e.g., news, blog posts, communication from social contacts) takes a while to be processed, especially when its complexity makes it difficult to grasp, or requires repeated exposures to generate belief. However, the mechanism behind the establishment of false and accurate belief is the same [71]. People are often biased to believe in the validity of information and "go with their gut" and intuition when deciding what is truth instead of deliberating [72]. For instance, 31% respondents in a U.S. survey ($n = 2023$) in March 2020 agreed that COVID-19 was purposefully created and spread, despite the absence of any plausible evidence for its intentional development [73]. People might have encountered conspiracy theories about the source of the virus many times, which might have contributed to this widespread belief because simply repeating a claim makes it more believable than presenting it only once [74]. Repetition increases belief in both misinformation and facts, and people get "infected". Regardless of cognitive ability and despite contradictory advice from an accurate source or accurate previous knowledge, there is a possibility that illusory truth persists months after the first exposure [75]. EBMs do not differentiate between repeated exposures, even though they do account for incubation time, whereas in an ABMs it would only be a matter of

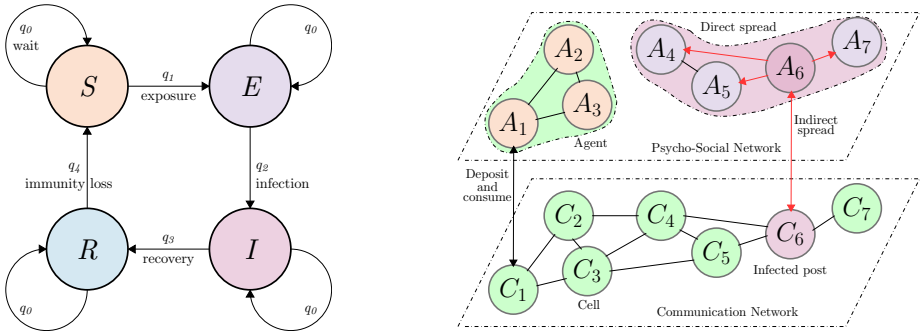
adding a condition to a state transition trigger. We thus put forward that the incubation rate should be redefined for infodemics.

But does belief really equate infection? [76] The exposure to false information is a strong contributor to the formation of false belief. Access to high-quality information is not necessary to spread misinformation. Instead, a range of precursors, cognitive and socio-affective drivers, influence the formation and storage of false beliefs [71], which draws a parallel with the properties of cells that make them or not hosts to viral multiplication. But misinformation and its siblings are not singular pieces that are passed on from person to person, nor are they easily identifiable by the "symptoms" of the "disease" they cause. Information and the knowledge it is stored as always come with related facts, untruths, beliefs, consequences, premises, and relationships that either support or restrain the "infection", thus affecting incubation, recovery, and loss of immunity. Where do we draw the line between infection and other beliefs that are not necessarily true but not harmful either? For instance, the belief that leaving scissors on the table leads to discord is not true, but it serves the purpose of avoiding accidents caused by exposed blades. Perhaps the definition of the infected state in infodemic models needs more interdisciplinary investigation.

Even with these limitations, there are enough arguments to analyse infodemics based on the four epidemiological states.

3.2. Equation-based Models (SEIRS)

The prevalent equation-based models (EBMs) for infodemics are based on epidemic interpretations of how information is transmitted throughout a population. They are macroscopic models, in which the persons engaged in



(a) *SEIRS* model representation with states (S susceptible, E exposed, R recovered, I infected) and transitions (q_0 wait for a trigger in current state, q_1 exposure, q_2 infection, q_3 recovery, q_4 immunity loss). All model combinations are listed in Appendix 1.

(b) ABM structure (conceptual representation) in which the cells ($C_1 - C_7$) form the communication network, while the agents ($A_1 - A_7$) form the psycho-social network; the spread of mis- and disinformation can be direct (between agents) or indirect (through stigmergy).

Figure 1: Conceptual representation of the two model types. The equation-based model (*SEIRS*) is structured as a finite-state machine in which population subsets mass-transition between states based on probabilities. The agent-based model (ABM) is structured as layered complex networks in which individuals transition between states based on local interactions.

receiving or relaying information lose their individuality. The main variables of the models become numbers reflecting portions of the population in one of several distinct states: susceptible, infected, recovered or exposed (figure 1a). In this study, we analyze six such diffusion models: *SI*, *SIS*, *SIR*, *SIRS*, *SEIR* and *SEIRS* [77]. They are differentiated by how many states are defined within the population and how many state transitions are allowed. For instance, in the *SI* model, the states are susceptible and infected, with the only transition $S \rightarrow I$; the *SIS* model also allows the reverse transition $I \rightarrow S$. For brevity, we only describe model *SEIRS* in this section; all six models are listed in Appendix I.

SEIRS [78] consists of four equations as dynamic representations of the timewise state transitions of the four variables (figure 1a), defined as:

$$\begin{cases} S'(t) = -\beta N^{-1} S(t) I(t) + \xi R(t) \\ E'(t) = \beta N^{-1} S(t) I(t) - \sigma E(t) \\ I'(t) = \sigma E(t) - \gamma I(t) \\ R'(t) = \gamma I(t) - \xi R(t) \end{cases} \tag{1}$$

where $S(t)$, $E(t)$, $I(t)$, $R(t)$ are the numbers of individuals with state susceptible, exposed, infected, or recovered, respectively, at time t. Notations $S'(t), E'(t), I'(t), R'(t)$ represent the first order derivatives of these variables and they model how the states of the population change between two moments in time. N is the total number of individuals in the population. The model parameters are: infection rate β, incubation rate σ, recovery rate γ, and immunity loss rate ξ.

Although *SEIRS*-class models can be expanded to include vital dynamics (asymmetrical birth and natural-cause death rates) [79], the total population is constant in the classic version: $N = S(t) + E(t) + I(t) + R(t)$.

The susceptible state is the nominal state of individuals and it models the population that can be either infected (models *SI(S)* and *SIR(S)*) or exposed (models *SEIR(S)*) to mis/disinformation. The population with exposed state cannot infect other individuals, but can become infected after an incubation period. Those with infected state are able to spread the "infection" to the susceptible individuals and can become recovered after a period of time. The individuals with recovered state are immune to the "infection" but can lose this immunity over time, at which point they transition to the susceptible state.

Parameter β represents the infection rate and is defined as the population percentage switching states from susceptible to exposed (or infected) in the unit of time. Parameter σ represents the incubation rate and is defined as the population percentage switching states from exposed to infected in the unit of time. Parameter γ represents the recovery rate and is defined as the population percentage switching states from infected to recovered in the unit of time. Parameter ξ represents the immunity loss rate and is defined as the population percentage switching states from recovered to susceptible in the unit of time.

For the *SEIRS* models, the unit of time is chosen to match the modeled epidemic, e.g. day for rapidly spreading infections, or month for slower dynamics. In this study, we consider the unit of time to be equivalent to the unit of time of the agent-based models, measured in *ticks*. Thus, we maintain the generality and scalability of the models.

3.3. Agent-based Models

Agent-based models are built with two types of entities: (a) an *agent* is a simulated system capable of perception, action, communication, and reasoning; (b) a *cell* is a simulated system part of the environment. The ABM unit of time is called a *tick*. Its real-world equivalence is selected based on problem context and dynamics. During one tick, all behaviors defined in the ABM (agents, cells, etc.) are executed in parallel asynchronously, by one step.

For this study, we propose an ABM structure to include the separation and dependence between the psycho-social component and the communication medium. We base this choice on the argument that one network alone

is not sufficient to embody these types of dynamics, put forth by studies on epidemic-infodemic interactions [80] and opinion dynamics [81]. Thus, agents model individuals (who can interact with the environment or each other) and the psycho-social network, while cells model the communication channels and/or network (figure 1b).

The psycho-social network is formed of all agents in the ABM. From an information diffusion perspective, this network appears as a graph with temporary edges, which are formed when two agents meet to directly transmit information to each other. This reflects the real world in which two persons open ad-hoc point-to-point communication sessions (e.g., face-to-face, phone call) with each other, forming a connection, but are not actively transmitting information to each other around-the-clock.

The social media/communication network (further referred to as the environment) is formed of all the cells in the ABM. In this study, we interpret each cell as a platform (e.g., blogging, micro-blogging, forums, social apps), broadcasting entity (e.g., newspaper website, television) or communication channel (e.g., VoIP apps). Equivalent to the real world, an agent "navigates" among these platforms; in this paper, we visualize this process by allowing agents to travel across the cell grid. While the cells are connected via adjacency to each other, we do not consider the network edges relevant in this situation, as platforms themselves do not exchange information; instead, the agents carry it across media.

Navigating. The environment thus represents the online (social) media network overlaid onto the direct point-to-point communication of the agents. The difference in behavior between direct communication and online media is

that the latter functions as a broadcast instead of a conversation. Blog posts, news articles, etc. are stored online and available for anyone in the network. In the real world, this type of broadcast is not received by all individuals at the same time, as navigating through personal media dashboards limits exposure (e.g., local news relayed within a region and not internationally, or social media clusters formed of family members, etc.): we achieve this effect by implementing the *agent movement* functions as they seek interactions and limiting their perception to a radius around themselves.

Engagement: consuming and generating information. The relationship humans have with information has been studied at length [82, 83]. The interpretation we adhere to in this paper associates this interaction process with foraging behaviors, through which people seek, select, consume and avoid information [84, 85]. For agents, cells are both information sources and sinks. Agents generate information by "posting" on cells (interpreted as either new content, or interactions with existing content such as commenting). When agents are infected, they might deposit mis- or disinformation, which is then picked up by other agents. In complex systems, this process of communicating through the environment is known as stigmergy [86]. Agents gain energy by consuming information and lose energy by posting, thus simulating engagement (e.g., scrolling through dashboards, commenting) or exposure fatigue (e.g., avoiding news applets).

Cell information growth. Aside from content deposited by agents on cells, we implement a generation mechanism through which information "grows". The underlying hypothesis is that information available in the environment can be generated by sources outside that environment. Thus, there is an

external component to the model which agents cannot affect, but that can influence the agent or its group. For instance, the social cluster of an individual being influenced by the media, over which single individuals do not have, in turn, an effect. Moreover, if cells are instances of online media platforms where agents interact with content, the information they carry can be influenced by recommendation engine algorithms, which promote any type of engagement, whether positive or negative [87], and thus become vectors for mis- and disinformation.

We developed two models: (a) the *simple model* is a direct translation of the finite-state machine of the EBMs into agent behaviors in which misinformation spreads point-to-point between agents; (b) the *enhanced model* is built upon the simple model by adding psycho-social components to the agent behaviors (e.g., preferential attachment), as well as the misinformation spread through the online (social) media network. All algorithms associated with the two models are detailed in Appendices II and III.

3.3.1. Simple ABM.

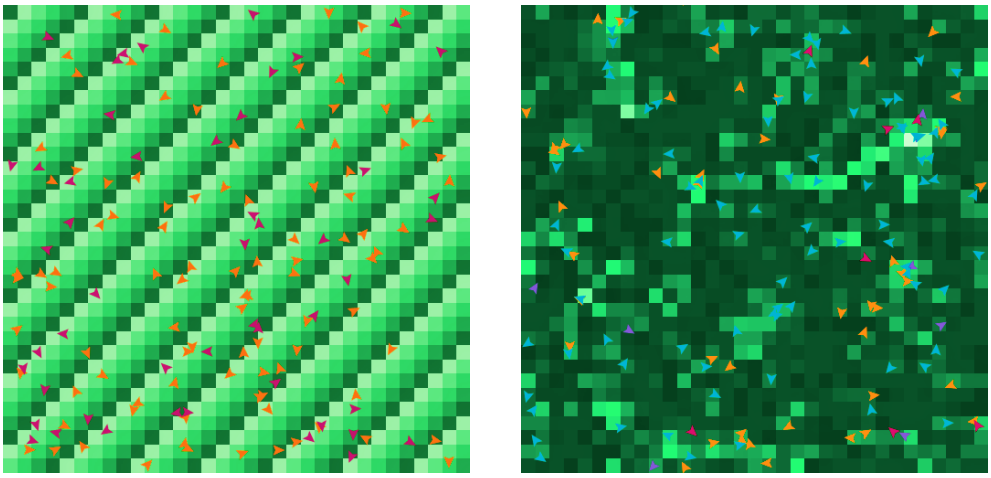
The state of an agent A_i ($i = 1..N$) is described by four variables: coordinates (x_i, y_i), orientation h_i, information status s_i, and energy level e_i. Coordinates are discrete (integer) and spatially define the position of the agent in the environment. Orientation is defined by the heading angle. Information status is a categorical variable $s \in$ {susceptible, exposed, infected, recovered} = {*S,E,I,R*} with four categories equivalent to the SEIRS model. Energy level is a continuous variable modeling the interest of agents to communicate (receive and relay information) and thus to move through the environment.

The state of a cell C_j ($j = 1..M^2$) is described by two variables: coordinates (x_j, y_j) and information quantity q_j. Cells are static, arranged on a torus, and visible in the model interface as an $M \times M$ square grid (agents leaving one border reappear on the opposite side). Together, cells form the communication network. Coordinates are discrete (integer) and define the position of the cell inside the grid. Information quantity is a continuous variable representing how much information is available to agents in one particular communication channel (displayed in shades of green: lighter for less information, darker for more). For this model all available information is considered true.

Figure 2 shows the initialization of the simple ABM at $t = 0$ *ticks* and the visualization of the model at $t = 50$ *ticks*.

Agent movement. Agents move through the environment toward a new position with a specified heading. The agent's target is computed based on vicinity: choosing the cell (patch) with the highest amount of information in agent radius r and viewing angle θ. The meeting of two agents on the same cell represents two individuals simultaneously using the same instance of a communication channel (e.g., a phone call, a messaging app, face-to-face conversation, etc.).

Agent information status. The categorical agent state variable $s \in \{S,E,I,R\}$ defines the set of actions it can perform, i.e. its behaviors (algorithms in Appendix III). The transitions between the four categories follow the rules of the EBMs finite-state machine and are implemented via threshold tests with the probabilities of infection β (relative to world and population size), successful incubation σ, recovery γ, and immunity loss ξ.



(a) t = 0 ticks

(b) t = 50 ticks

Figure 2: Visual representation of the agent-based implementation of the diffusion model at two different time steps. The color of each cell (shades of green) represents the level of information present in that specific area, while agents are depicted by arrows in a visual representation that shows both the position and the current orientation: orange susceptible, magenta infected, blue recovered, and violet exposed.

The simple model implements diffusion using point-to-point communication between agents, meaning that mis- and disinformation can only be transmitted when agents talk directly to each other. The point-to-point communication is established by infected agents with closest susceptible agents. This is the epidemiological interpretation in which biological viruses are transmitted through direct contact. Agents lose energy by moving through the world. This ABM is designed to allow for disinterest or sudden interest, and thus implement vital dynamics (e.g., "death" by leaving the network or "birth" by joining the network). However, to match the ABM to the classical SEIRS EBMs, we deactivated these functions. Because in this case the environment does not provide or store false information, it does not affect the dynamics of the infection spread.

3.3.2. Enhanced ABM.

The state of an agent A_i $(i = 1..N)$ is described by five variables: coordinates, orientation, information status, energy level, and group g_i, which is a categorical variable that tags the agent as belonging to one of two social groups: $g \in \{G1, G2\}$.

The state of a cell C_j $(j = 1..M)$ is described by three variables: coordinates, information quantity, and information type c_j, which is a categorical variable $c \in \{c_{true}, c_{false}\}$. Truthful information c_{true} is displayed as *green*, while falsity c_{false} is displayed as *red*, both in shades based on quantity. Cells are arranged in the same grid as the simple model.

Agent movement and *information status* are similar to the simple model. The difference is that the $S \rightarrow I$ and $S \rightarrow E$ transitions can be triggered externally by false information from the environment. The infection rate β

is also adjusted to account for two infection sources (other agents and cells).

Preferential attachment. In this study we implement a simple mechanic based on group adherence: agents in groups $G1$ and $G2$ will only believe false information from agents in the same group. Thus, the $S \rightarrow I$ and $S \rightarrow E$ transitions become conditional.

3.3.3. Implementation.

Both the simple and enhanced ABMs presented in the previous subsections function on the same base structure (scheduling), which is them presented in this separated section for clarity reasons, and to avoid repetitions. Appendix II presents in the details the models scheduling and their interfaces. The scheduling of the model can be summarized into phases. The initialization phase loads the agent list and positions, and the cell distribution with information quantities into the model (here, we choose the same initialization for all model runs). The next phases execute agent behaviors: first is the the movement phase, in which an agent A_i chooses a target to move toward. The decision-making mechanism considers a given radius d^e around agent A_i. Agents deplete a fixed amount of energy e_c for each movement. Second, the spreading phase, in which an agent A_i with status $s_i = I$ spreads mis- or disinformation to the nearest neighbor A_k with status $s_k = S$. To be infected or exposed, A_k should be at a distance $d_{ik} < d^c$ (preset radius). Third, agent A_k changes status to $s_k = E$ for models *SEIR(S)* or to $s_k = I$ for the others. Fourth, an agent A_i with $s_i = I$ can recover to $s_i = R$ for *SIR(S), SEIR(S)*. Model selection is possible through the ABM interface. Fifth, a recovered agent A_i with $s_i = R$ can become susceptible in the case of models *SIS, SIRS, SEIRS*. (Deactivated function: an agent

A_i with negative energy e_i is removed from the agent list, i.e., "dies", which occurs whenever it expends more energy than it collects and stores.) The cell behaviors execution phase consists of the information q_j on each cell C_j replenishing with a preset growth rate up to a maximum limit.

3.4. Comparative Analysis

The comparative analysis we perform in this study has two outcome measures which describe how well the ABM generated variables $y_a \in \{S(t), E(t), I(t), R(t)\}$ match the corresponding EBM generated variables y_e.

The Pearson correlation coefficient ρ [88] is given by:

$$\rho = \frac{\mathrm{cov}(y_a, y_e)}{\sigma_{y_a}\sigma_{y_e}}, \tag{2}$$

where σ_{y_a} and σ_{y_e} are the standard deviations. This measure shows how well the two variables match longitudinally (shape over time) and is defined over the interval $[-1; 1]$, where 1 is best match, -1 mirrored evolution, and 0 is complete mismatch.

The normalized root mean of square error (NRMSE) is defined as:

$$\mathrm{NRMSE} = \frac{\sqrt{\frac{1}{q}\sum_{k=1}^{q}(y_{a_k} - y_{e_k})^2}}{y_{e_{\max}} - y_{e_{\min}}} \tag{3}$$

where q is the number of samples, with $y_{e_{\max}}$ and $y_{e_{\min}}$ the maximum and minimum values of y_e. NRMSE is a dimensionless score for scale-independent comparisons of model outcomes, defined over the interval $[0; 1]$, where 0 means best match and 1 worst [89].

In this study, we use the combined outcome measures ρ and NRMSE to assess the similarity of the model outcomes as dicrete signals: ρ quantifies

how well y_a and y_e match in shape, whereas NRMSE offers an estimate of the relative differences between model outcomes.

Note: ρ cannot be calculated when one of the contributing variables is zero; to avoid skewing the summary calculations, these were adjusted to match (1) or not-a-match (0) based on the NRMSE of the pair.

3.5. Fitting to real-world data: case study on vaccine acceptance

We evaluate the two types of models in a fitting experiment. The real world data we choose for a case study describes vaccine acceptance and disapproval during the pandemic. The model type is SIS, in which we consider the susceptible state as "accepting the vaccine" and the infected state as "not accepting the vaccine", with the possibility to recover to susceptible state.

Study design and participants. The data was collected via a global survey on COVID-19 beliefs, described in [90] from 23 countries between June 2020 and March 2021. One of the questions asked: "If a vaccine for COVID-19 becomes available, would you choose to get vaccinated?". Martinelli and Veltri [91] then conducted a study on COVID-19 vaccine acceptance, which produced a dataset expressing the percentage of the populating accepting of the vaccine. The dataset is longitudinal over 36 weeks, with a sample of 2 weeks between measurements. From this dataset, we choose Romania as a case study.

Fitting method. We search for the enhanced ABM and the EBM parameters so that the model outputs $S_i(k)$ and $I_i(k)$, with $i \in \{\text{ABM}, \text{EBM}\}$, match the real world data $S_{RW}(k)$ and $I_{RW}(k)$. Considering that the two signals are mirrored, for a parameter set π_i and discrete time step $k = 0..19$,

the multi-objective optimization problem is:

$$\min_k \left(\rho(k)|_{S_i(k),S_{RW}(k)} \, , \, \mathrm{NRMSE}^{-1}(k)\big|_{S_i(k),S_{RW}(k)} \right) \tag{4}$$

The ABM parameter set is $\pi_{\mathrm{ABM}} = \{\beta, \gamma, P_M, T_F, I_R, L_M^{S,I}, L_P^{S,I}, D_P^{S,I}\}$, where $\beta \in [0,1]$ is the infection rate and $\gamma \in [0,1]$ is the recovery rate. $P_M \in [1,15]$ is the population multiplier and scales the number of agents from a initialization baseline of 100. $T_F \in [1,336] \cap \mathbb{N}$ is the time factor and determines the size of an ABM tick relative to the real world data sampling time; thus, T_F represents how many ABM ticks pass for each week in the real world. I_R is the information regrowth rate. $L_M^{S,I} \in [1,100] \cap \mathbb{N}$ and $L_P^{S,I} \in [1,100] \cap \mathbb{N}$ represent the energy loss of agents from movement and from posting information, for susceptible and infected agents, respectively. Finally, $D_P^{S,I} \in [1,25] \cap \mathbb{N}$ represent the delay between agents consuming and posting information, for susceptible and infected agents, respectively.

The EBM parameter set is $\pi_{\mathrm{EBM}} = \{\beta, \gamma, T_F\}$, where $\beta \in [0,1]$ is the infection rate and $\gamma \in [0,1]$ is the recovery rate. T_F represents how many data points we are considering for each week in the real world, $T_F \in [1,336] \cap \mathbb{N}$.

For both, the population size is normalized to percentages and the two signals are initialized to approximate the first datapoint of the real world data. To calculate the criteria ρ and NRMSE, we bin the model outputs to 19 points by averaging.

The optimziation procedure is performed using the Python hyper-optimization library Optuna, which provides samplers and pruners and determines the importance of parameters and their interdependence. In this case study, we choose the Tree-structured Parzen Estimator (TPE) sampler and the hyperband pruner.

3.6. Tools, Frameworks, and Software

The ABMs were implemented and simulated using NetLogo 6.3.0 [92]. The EBMs were implemented in Python 3.11.0 in discrete form, with a time step equivalent to one ABM tick. All visualizations, plots, heatmaps and calculations were also performed in Python 3.11.0. The following Python packages were used: matplotlib 3.7.1, numpy 1.24.3, scikit-learn 1.2.2, scipy 1.10.1, optuna 3.5.0, optuna-dashboard 0.14.0.

4. Results

4.1. Evaluation of intermodel equivalence

In this section we present a selection of relevant results (limited by space considerations), with more examples in Appendices IV and V, and a full set of simulation results in the repository. Results are generated from three instances of the ABMs: a) a small variant of the simple ABM with a world size of $M = 33 \times 33$ cells (the default world size in NetLogo), b) a large variant of the simple ABM with $M = 99 \times 99$ cells, and c) the enhanced ABM with $M = 99 \times 99$ cells. These environment sizes fall within research findings on the numbers of news outlets [93] in North America. For each of the three ABM instances, we set up two experiments: Case I and Case II, with a difference in initialization values for E (zero for I, nonzero for II). These cases along with the full lists of initialization and configuration parameters are included in the repository at `https://anonymous.4open.science/r/results-big-data-soc`.

Figures 3 and 4 show two examples of matching and non-matching model outcomes, respectively, as variations over 200 *ticks*, for the *SEIRS* model us-

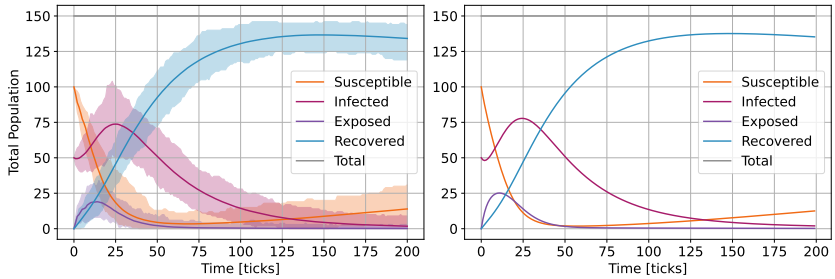


Figure 3: Example of matching outcomes for the *SEIRS* model, ABM (left) vs. EBM (right), $\beta = 0.2$, $\gamma = 0.03$, $\sigma = 0.155$, $\xi = 0.001$ (means and deviations for 1000 runs of the ABM)

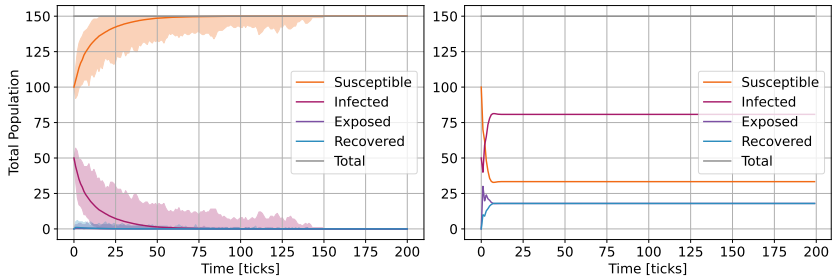


Figure 4: Example of non-matching outcomes for the *SEIRS* model, ABM (left) vs. EBM (right), $\beta = 0.9$, $\gamma = 0.2$, $\sigma = 0.9$, $\xi = 0.9$ (means and deviations for 1000 runs of the ABM)

ing 1000 runs of the simple ABM (small variant). In the matching case, the means of the ABM runs closely follow the outcomes of the EBM (equivalence measures for all six models are in Appendix IV). Even though the ABM was initialized in the same manner at every run, the deviations of models outcomes show that infodemic dynamics are sensitive to small individual variations happening at local agent level (caused by the probabilistic conditions), whereas the EBM applies the same probabilities of state transitions at group level, i.e., mean.

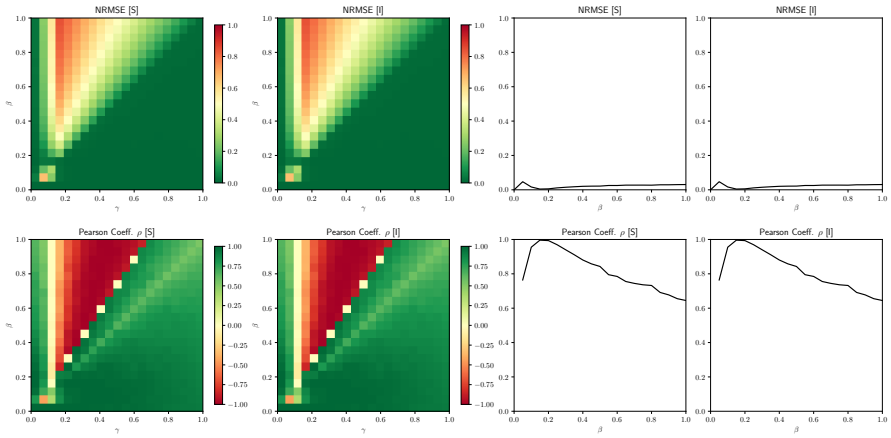
Table 1 shows cumulative results (means and standard deviations) for parameter variations: $\beta, \gamma, \sigma, \xi \in [0.1; 1]$ with a step of 0.1 over 2000 *ticks*, resulting in 10 experiments for *SI*, 100 *SIS* and *SIR* each, 1000 *SIRS* and *SEIR* each, 10000 *SEIRS*. Results are consistent for model size in the simple ABM case. Some models fair better (e.g., *SIR*) than others (e.g., *SIS*) on average, but even so, none of the EBMs reproduce exactly the ABM response across the entire parameter range. The enhanced ABM shows similar results, except for *SIR*, which is considerably less well matched.

Figure 5 shows the intermodel equivalence outcomes NRMSE and ρ for models *SIS* and *SI*, obtained using parameters variations $\beta, \gamma \in [0.1; 1]$ with a step of 0.05 over 10 ABM runs each and 2000 *ticks*. The results to S and I are similar due to the mirror-effect in these two outcomes, which is expected. While there are parameter combinations and intervals for which results match, this is not consistent over their entire range. The heatmaps of the *SIS* model shows combinations for which the ABM and the EBM are profoundly distinct in dynamics: an infodemic with high infection rate β and mid-range recovery rate γ will cause the ABM and EBM models to behave differently. The addition of the online social media network in the enhanced ABM visibly shifts the matching/non-matching coverages.

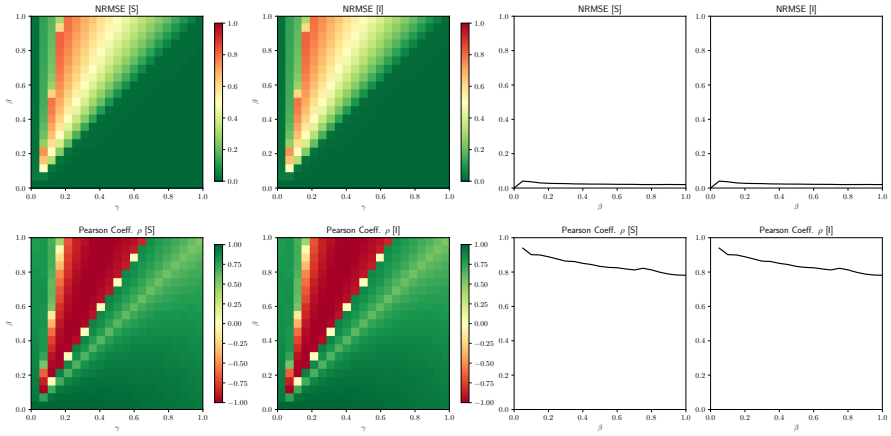
Figure 6 shows the effect of preferential attachment (conditional infection) on the enhanced ABM vs. simple ABM outcomes for groups of different sizes $N_{G1} = 80\%$, $N_{G2} = 20\%$ (total $N = 1350$ agents), over 200 *ticks* and 1000 runs (means and deviations). While the exposed outcome seems similar enough, the others (susceptible, infected, recovered) show considerable change in infodemic dynamics. Group sizes do not show an effect (see

Model	NRMSE				ρ			
	S	E	I	R	S	E	I	R
Simple ABM (small variant) vs. EBM ($N = 150$ agents); mean (SD)								
SI	0.03 (0.01)	-	0.03 (0.01)	-	0.74 (0.16)	-	0.74 (0.16)	-
SIS	0.18 (0.24)	-	0.18 (0.24)	-	0.21 (0.75)	-	0.21 (0.75)	-
SIR	0.25 (0.18)	-	0.01 (0.01)	0.25 (0.18)	0.68 (0.36)	-	0.84 (0.10)	0.90 (0.07)
SIRS	0.19 (0.24)	-	0.12 (0.18)	0.07 (0.09)	0.30 (0.67)	-	0.45 (0.63)	0.32 (0.50)
SEIR	0.25 (0.15)	0.01 (0.00)	0.01 (0.00)	0.25 (0.15)	0.75 (0.35)	0.83 (0.12)	0.91 (0.06)	0.94 (0.04)
SEIRS	0.21 (0.26)	0.06 (0.08)	0.09 (0.14)	0.06 (0.08)	0.47 (0.61)	0.65 (0.42)	0.59 (0.56)	0.42 (0.47)
Simple ABM (large variant) vs. EBM ($N = 1350$ agents); mean (SD)								
SI	0.04 (0.01)	-	0.04 (0.01)	-	0.70 (0.05)	-	0.70 (0.05)	-
SIS	0.18 (0.24)	-	0.18 (0.24)	-	0.14 (0.79)	-	0.14 (0.79)	-
SIR	0.30 (0.20)	-	0.01 (0.01)	0.30 (0.20)	0.53 (0.39)	-	0.82 (0.09)	0.89 (0.07)
SIRS	0.19 (0.24)	-	0.12 (0.18)	0.07 (0.09)	0.24 (0.73)	-	0.39 (0.69)	0.28 (0.58)
SEIR	0.31 (0.17)	0.01 (0.00)	0.01 (0.01)	0.31 (0.17)	0.58 (0.40)	0.81 (0.11)	0.90 (0.06)	0.93 (0.04)
SEIRS	0.22 (0.27)	0.06 (0.09)	0.10 (0.15)	0.06 (0.09)	0.41 (0.68)	0.65 (0.45)	0.55 (0.62)	0.39 (0.54)
Enhanced ABM vs. EBM ($N = 1350$ agents); mean (SD)								
SI	0.03 (0.01)	-	0.03 (0.01)	-	0.75 (0.13)	-	0.75 (0.13)	-
SIS	0.11 (0.14)	-	0.11 (0.14)	-	0.35 (0.55)	-	0.35 (0.55)	-
SIR	0.20 (0.15)	-	0.14 (0.23)	0.26 (0.18)	0.32 (0.29)	-	0.36 (0.61)	0.62 (0.29)
SIRS	0.12 (0.16)	-	0.09 (0.12)	0.06 (0.07)	0.37 (0.51)	-	0.46 (0.48)	0.36 (0.43)
SEIR	0.19 (0.14)	0.01 (0.00)	0.17 (0.27)	0.28 (0.21)	0.55 (0.24)	0.81 (0.11)	0.38 (0.69)	0.69 (0.31)
SEIRS	0.14 (0.18)	0.06 (0.08)	0.11 (0.15)	0.05 (0.06)	0.48 (0.49)	0.64 (0.38)	0.54 (0.50)	0.43 (0.44)

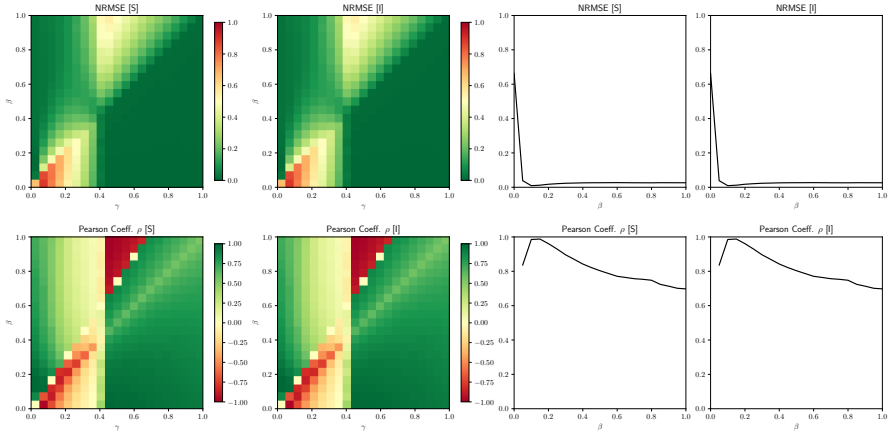
Table 1: Cumulative results for intermodel equivalence across parameter variations: $\beta, \gamma, \sigma, \xi \in [0.1; 1]$ with a step of 0.1, resulting in 10 experiments for *SI*, 100 *SIS* and *SIR*, 1000 *SIRS* and *SEIR*, 10000 *SEIRS*. Case II configuration. Notations: NRMSE normalized root mean of square error, ρ Pearson's correlation coefficient, SD standard deviation, ABM agent-based model, EBM equation-based model, S susceptible, E exposed, I infected, R recovered.



(a) Effect of β and γ on NRMSE and ρ for the simple ABM (small variant, $N = 150$ agents): heatmaps for *SIS* (left) and graphs for *SI* (right)



(b) Effect of β and γ on NRMSE and ρ for the simple ABM (large variant, $N = 1350$ agents): heatmaps for *SIS* (left) and graphs for *SI* (right)



(c) Effect of β and γ on NRMSE and ρ for the enhanced ABM ($N = 1350$ agents): heatmaps for *SIS* (left) and graphs for *SI* (right)

Figure 5: Intermodel equivalence outcomes for the *SIS* and *SI* models (result exemplification), where NRMSE is the normalized root mean of square error, ρ is Pearson's correlation coefficient, β is the infection rate, and γ is the recovery rate. Case I configuration.

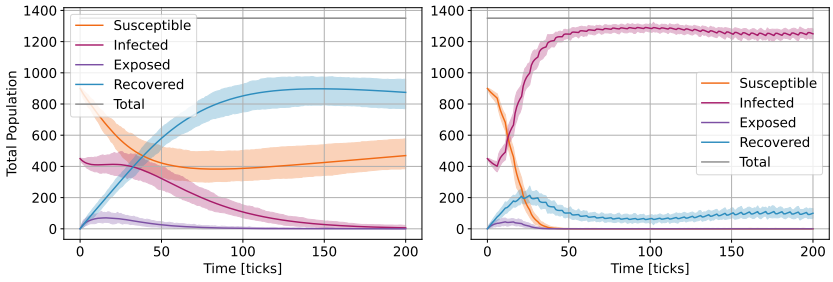


Figure 6: Effect of preferential attachment on the ABM model outcomes with $N_{G1} = 80\%$, $N_{G2} = 20\%$ (total $N = 1350$ agents): simple model large variant (left) and enhanced model (right) for 1000 runs (means and deviation).

Appendix V), but this observation is enough to raise questions regarding the many social group dynamics of the real world.

4.2. Fitting to real world data: results

First, we applied the fitting procedure to the enhanced ABM. After 700 iterations, the best result has $\text{NRMSE}|_{S_{\text{ABM}},S_{RW}} = 0.055$ and $\rho|_{S_{\text{ABM}},S_{RW}} = 0.872$. The model parameter are: $\beta = 0.633$, $\gamma = 0.073$, $I_R = 0.41$, $P_M = 10.35$, $T_F = 19$, $L_M^S = 28$, $L_M^I = 4$, $D_P^S = 23$, $D_P^I = 25$, $L_P^S = 96$, $L_P^I = 22$. Second, we applied the fitting procedure to the EBM. After 700 iterations, the best result has $\text{NRMSE}|_{S_{\text{EBM}},S_{RW}} = 0.086$ and $\rho|_{S_{\text{EBM}},S_{RW}} = 0.714$. The model parameter are: $\beta = 338 \cdot 10^{-6}$, $\gamma = 672 \cdot 10^{-6}$, $T_F = 30$. Figure 7 shows the results of these two fitting procedures against the real world data. While the EBM outputs seem to follow an overall trend of decrease or increase, it does not capture the dynamics of people changing their opinions over time. In this case the enhanced ABM manages to show how the proportions of the population invert their ratio at weeks 18-22. This period corresponds to October 2020 when Romania faced a considerable increase in cases.

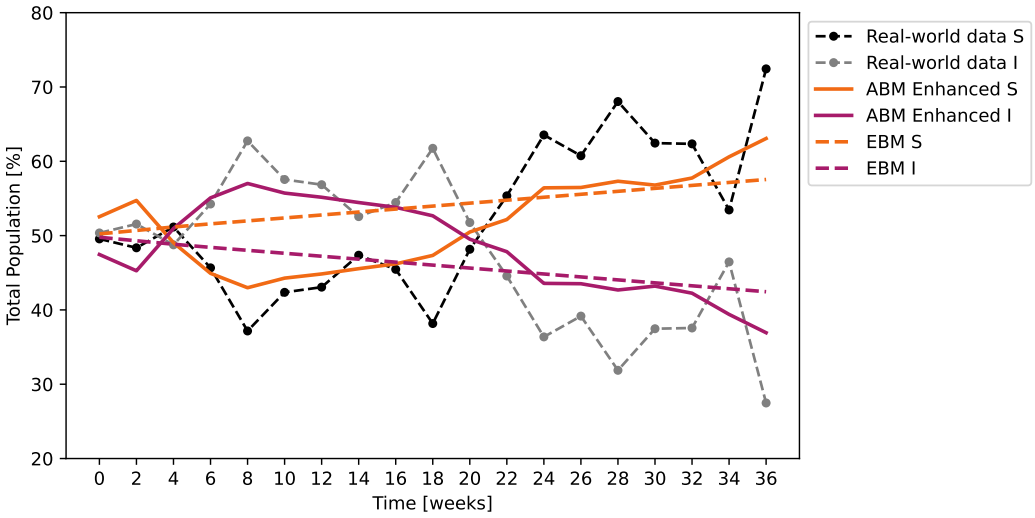


Figure 7: Results of the SIS model fitting against the real world data for the enhanced ABM ($\beta = 0.633$, $\gamma = 0.073$) averaged over 100 runs (variance 5498.17) and for the EBM ($\beta = 338 \cdot 10^{-6}$, $\gamma = 672 \cdot 10^{-6}$).

Further, we perform a comparative test with the parameters resulting from the enhanced ABM fitting procedure (where applicable). Figure 8 shows the responses of the three models (enhanced and simple ABM, EBM) against real world data. Table 2 presents the two outcomes (ρ, NRMSE), with the best fit for the enhanced ABM, which shows that even when the infection and recovery rates are determined to match the dynamics of real world data, the EBM response does not manage to illustrate these dynamics. The simple ABM response comes closer, but the fit is still poor, highlighting the importance of the online communication network as a misinformation facilitator.

	ABM Enhanced	ABM Simple	EBM
ρ	**0.872**	0.391	-0.067
NRMSE	**0.055**	0.103	0.418

Table 2: Outcome measures (Pearson correlation coefficient ρ and normalized root mean of squared error NRMSE) comparing the real-world data with the outputs of the enhanced ABM, simple ABM and EBM, simulated with the parameters $\beta = 0.633$, $\gamma = 0.073$, resulted from the enhanced ABM fitting procedure. Notations: ABM agent-based model, EBM equation-based model.

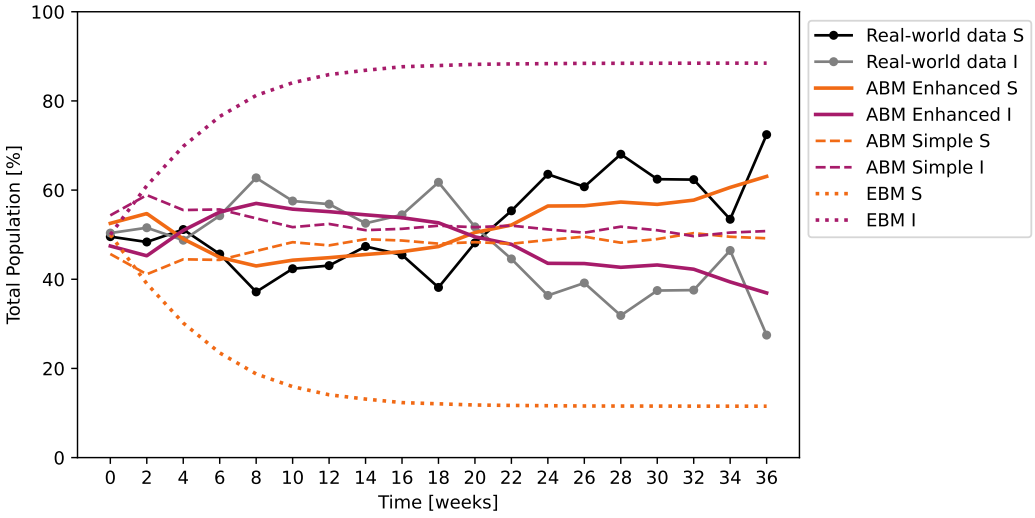


Figure 8: Enhanced ABM outputs fitted against real world data, compared with the simple ABM and the EBM outputs obtained using the same parameters (where applicable). $\beta = 0.633$, $\gamma = 0.073$. Both ABM outputs are averaged over 100 runs and have a variance of 5498.17 for the enhanced and 1577.25 for the simple versions.

5. Discussion

In this study we investigated EBMs and ABMs for the spread of mis- and disinformation. For this, we designed two models: a) a simple ABM as a direct translation of *SEIRS*-type EBMs' underlining logic, and b) an enhanced ABM to reflect communication and online media networks, as well as preferential attachment. Then, we performed an intermodel equivalence analysis. Results show that EBMs and ABMs display both matching and non-matching outcomes, depending on parameter ranges. Our hypothesis that microscopic models (e.g., ABMs) are necessary to capture elements of the human psycho-social context is confirmed.

The critical difference between these types of models is that EBMs assume individual homogeneity, whereas the ABM structure allows for more complex reasoning, such as preferential attachment, multiple pieces/types of information, believing views opposite their own group, etc. While there are attempts to rework EBMs for various epidemics [94, 48], in some cases combining epidemics and infodemics in one state machine [95], these models have the same issue: losing individuality.

The main reason for the mismatch of these types of models is that the effects observed at population level (e.g., infection rate) are not observed at individual level (e.g., probability of infection) because humans, as psycho-social networks, behave like a complex system in which the infodemic is an emergent process, and to paraphrase Kevin Kelly [1], even made of the same building blocks, we would not find the beehive in the bee. The top-down approach of equation-based modelling transforms a population into one black-box mass, whereas the bottom-up design of complex systems allows

for these diffusion processes to emerge naturally, from interactions between individuals. Strong emergence is the "magic" that can explain multi-layered human interactions at large scale [96]. This is why, when the conditions allow it, the ABM approach would be preferable to analyse, for instance, the effects of individual vs. mass-level interventions of information manipulation (e.g., propaganda).

The *SEIRS*-type EBMs we analyzed here are not necessarily entirely wrong, but they are based on assumptions that do not consider the whole nature of information diffusion. Epidemics driven by biological viruses assume some level of physical contact, whereas information spreads through entirely different media and mechanics. Epidemic research has modified these basic models to better represent various diseases, and the same should happen for infodemics. As they are now, the *SEIRS* EBM models are missing the communication and online media networks, as well as psycho-social attachments and beliefs, meaning that each person would have different infection rates based on information coming from different sources.

So, then, what? Models cannot be perfect clones of the real world [97], but their level of wrongness matters when humans are involved. Infodemic and epidemic models are useful for testing in a safe environment (simulation) various interventions before deployment or adoption, and it is critical that the model predictions are as accurate as possible.

Concerning the corrective interventions on mis- and disinformation, EBMs do not currently offer ways to simulate or evaluate these mechanisms, as they only show the outcomes of a population (or subset). Efficiency estimations of local interventions require ABMs. Often, interventions to resist

misinformation must overcome various cognitive and socio-affective barriers. The most common types of correction are individual. Fact-based corrections directly address inaccuracies and provide accurate information [98]. Broader protection against different types of misleading tactics is offered by addressing the logical fallacies common in disinformation [99] or by challenging the plausibility of the misinformation or the credibility of its source. Multiple approaches can be combined into a single correction [100]; however they all have to be applied locally. In fact, effective regulatory actions must be implemented at individual level. Even though some corrections can propagate through specific social groups, there is always an element of internally-generated intent if coercive action at the population level is to be avoided. This is what in systems science is known as decentralized control, as opposed to centralized, in which an authority body applies interventions at societal level.

As it currently stands, EBMs can still be useful. Because they provide mean trends within a population, EBMs can supplement predictive ABM simulations and act as baselines for identifying outlier, unusual, or unexpected behaviors. However, we suggest that EBMs should be revised when it comes to misinformation diffusion, so that the parameters of the wide-scale infodemic would reflect the emergence from local behaviors. As the case study on vaccine acceptance shows, the enhanced ABM fits the real world data better than the simple version. This suggests that adding elements to a model can be helpful and sometimes even necessary to correctly interpret the behavior of the modeled system, especially when it is generated from non-trivial interactions, such as information diffusion in an infodemic. Fu-

ture studies should consider the limitations and advantages of both ABMs and EBMs, and keep in mind that their outcomes are not always equivalent.

6. Conclusions and future developments

In this study, we evaluated the intermodel equivalence for infodemics. We designed agent-based models (ABMs) and compared their outcomes with classical equation-based models (EBMs) inspired by viral epidemics. We found low equivalence over the entire model parameter range, although the outcomes were similar for specific values. We also found that ABMs can capture the dynamics of real world data better than EBMs. We surmise that ABMs and EBMs serve different purposes with widely different structures (one is microscopic, the other macroscopic) and the choice of choosing one over the other should be informed, with awareness to their limitations and the fact that they are not interchangeable. Moreover, EBMs for infodemics should be revised from their counterparts modeling biological viral spreads.

Future developments include building an interaction topology between agents on a network and subsequently assessing how the structure of interactions (i.e., the features of the network) and its dynamics (creating or destroying links between agents) affect the results of the equation- and agent-based diffusion models. Moreover, we will implement and investigate time- or location-variant parameters (e.g., variable infection rate), introduce vital dynamics to both types of models and mitigation mechanisms, as well as different pieces of mis- and disinformation spreading concurrently through the agent population.

7. Acknowledgements

No direct funding is declared for this article. MP would like to thank the GC Rieber Foundation, the Norwegian Government, and the Research Council of Norway (288164) for supporting their work at the Centre for Elderly and Nursing Home Medicine, University of Bergen, Norway.

References

[1] K. Kelly, Out of control: the new biology of machines, social systems, and the economic world (1995).

[2] K. T. Gradoń, J. A. Hołyst, W. R. Moy, J. Sienkiewicz, K. Suchecki, Countering misinformation: A multidisciplinary approach, Big Data & Society 8 (1) (2021) 20539517211013848.

[3] J. E. Smith, N. Pinter-Wollman, Observing the unwatchable: Integrating automated sensing, naturalistic observations and animal social network analysis in the age of big data, Journal of Animal Ecology 90 (1) (2021) 62–75.

[4] C. Hooker, Introduction to philosophy of complex systems: A: Part a: Towards a framework for complex systems, in: Philosophy of complex systems, Elsevier, 2011, pp. 3–90.

[5] S. Roman, F. Bertolotti, A master equation for power laws, Royal Society Open Science 9 (2022) 220531.

[6] M. Mitchell, Complex systems: Network thinking, Artificial intelligence 170 (18) (2006) 1194–1212.

[7] A. D. Broido, A. Clauset, Scale-free networks are rare, Nature communications 10 (1) (2019) 1017.

[8] A. F. Siegenfeld, Y. Bar-Yam, An introduction to complex systems science and its applications, Complexity 2020 (2020) 6105872.

[9] P. Sen, Complexities of social networks: A physicist's perspective, arXiv e-prints (2006) physics–0605072.

[10] B. Edmonds, C. L. Page, M. Bithell, E. Chattoe-Brown, V. Grimm, R. Meyer, C. Montañola-Sales, P. Ormerod, H. Root, F. Squazzoni, Different modelling purposes, JASSS (2019). doi:10.18564/jasss.3993.

[11] L. Serena, M. Marzolla, G. D'Angelo, S. Ferretti, A review of multilevel modeling and simulation for human mobility and behavior, Simulation Modelling Practice and Theory (2023) 102780.

[12] M. Lachowicz, Microscopic, mesoscopic and macroscopic descriptions of complex systems, Probabilistic Engineering Mechanics 26 (1) (2011) 54–60.

[13] S. Meraz, The democratic contribution of weakly tied political networks: Moderate political blogs as bridges to heterogeneous information pools, Social Science Computer Review 31 (2) (2013) 191–207.

[14] J. P. Baptista, A. Gradim, Online disinformation on facebook: the spread of fake news during the portuguese 2019 election, Journal of contemporary European studies 30 (2) (2022) 297–312.

[15] Z. Peng, M. Li, Y. Wang, G. T. Ho, Combating the covid-19 infodemic using prompt-based curriculum learning, Expert Systems with Applications 229 (2023) 120501.

[16] J. Agley, Y. Xiao, Misinformation about covid-19: evidence for differential latent profiles and a strong association with trust in science, BMC Public Health 21 (2021) 1–12.

[17] E. Raja, B. Soni, C. Lalrempuii, S. K. Borgohain, An adaptive cyclical learning rate based hybrid model for dravidian fake news detection, Expert Systems with Applications (2023) 122768.

[18] K.-C. Yang, F. Pierri, P.-M. Hui, D. Axelrod, C. Torres-Lugo, J. Bryden, F. Menczer, The covid-19 infodemic: twitter versus facebook, Big Data & Society 8 (1) (2021) 20539517211013861.

[19] A. Fick, Ueber diffusion, Annalen der Physik 170 (1) (1855) 59–86.

[20] J. R. Anderson, A spreading activation theory of memory, Journal of verbal learning and verbal behavior 22 (3) (1983) 261–295.

[21] M. E. Newman, Spread of epidemic disease on networks, Physical review E 66 (1) (2002) 016128.

[22] M. Morris, Epidemiology and social networks: Modeling structured diffusion, Sociological methods & research 22 (1) (1993) 99–126.

[23] A. Vespignani, H. Tian, C. Dye, J. O. Lloyd-Smith, R. M. Eggo, M. Shrestha, S. V. Scarpino, B. Gutierrez, M. U. Kraemer, J. Wu,

et al., Modelling covid-19, Nature Reviews Physics 2 (6) (2020) 279–281.

[24] M. Sorapure, Anthropographics in covid-19 simulations, Big Data & Society 9 (1) (2022) 20539517221098414.

[25] D. Gruhl, R. Guha, D. Liben-Nowell, A. Tomkins, Information diffusion through blogspace, in: Proc. of the 13th international conference on World Wide Web, 2004, pp. 491–501.

[26] M. Cinelli, A. Pelicon, I. Mozetič, W. Quattrociocchi, P. K. Novak, F. Zollo, Dynamics of online hate and misinformation, Scientific reports 11 (1) (2021) 22083.

[27] F. J. León-Medina, Social learning and the complex contagion of political symbols in twitter: The case of the yellow ribbon in catalonia, Big Data & Society 10 (2) (2023) 20539517231180569.

[28] F. M. Bass, A new product growth for model consumer durables, Management science 15 (5) (1969) 215–227.

[29] W. Wei, Z. Xin-Gang, W. Jieying, Can support policies promote the innovative diffusion of waste-to-energy technology?, Environmental Science and Pollution Research 29 (37) (2022) 55580–55595.

[30] W. O. Kermack, A. G. McKendrick, A contribution to the mathematical theory of epidemics, Proceedings of the Royal Society of London Series A 115 (772) (1927) 700–721.

[31] W. O. Kermack, A. G. McKendrick, Contributions to the mathematical theory of epidemics. ii.—the problem of endemicity, Proceedings of the Royal Society of London Series A 138 (834) (1932) 55–83.

[32] W. O. Kermack, A. G. McKendrick, Contributions to the mathematical theory of epidemics. iii.—further studies of the problem of endemicity, Proceedings of the Royal Society of London Series A 141 (843) (1933) 94–122.

[33] J. Woo, H. Chen, Epidemic model for information diffusion in web forums: experiments in marketing exchange and political dialog, SpringerPlus 5 (1) (2016) 1–19.

[34] S. Daun, J. Rubin, Y. Vodovotz, G. Clermont, Equation-based models of dynamic biological systems, Journal of critical care 23 (4) (2008) 585–594.

[35] H. Sayama, A partial integro-differential equation-based model of adaptive social network dynamics, in: ALIFE 2022: The 2022 Conference on Artificial Life, MIT Press, 2022.

[36] E. Bonabeau, Agent-based modeling: Methods and techniques for simulating human systems, Proceedings of the national academy of sciences 99 (suppl_3) (2002) 7280–7287.

[37] N. R. Jennings, On agent-based software engineering, Artificial intelligence 117 (2) (2000) 277–296.

[38] C. Berceanu, I. Banu, B. S. Husebo, M. Patrascu, Predictive agent-based crowd model design using decentralized control systems, IEEE Systems Journal 17 (1) 1383–1394.

[39] S. C. Bankes, Agent-based modeling: A revolution?, Proceedings of the National Academy of Sciences 99 (2002) 7199–7200.

[40] F. Bertolotti, A. Locoro, L. Mari, Sensitivity to initial conditions in agent-based models, in: Multi-Agent Systems and Agreement Technologies EUMAS 2020, Springer, 2020, pp. 501–508.

[41] P. Milling, N. Schieritz, Modeling the forest or modeling the trees-a comparison of system dynamics and agent-based simulation, in: Proceedings of the 21st international conference of the System Dynamics Society, System Dynamics Soc., 2003, pp. 1–15.

[42] H. Van Dyke Parunak, R. Savit, R. L. Riolo, Agent-based modeling vs. equation-based modeling: A case study and users' guide, in: Multi-Agent Systems and Agent-Based Simulation, 1998. Proceedings, Springer, 1998, pp. 10–25.

[43] Y. Sun, L. Cheng, A survey on agent-based modelling and equation-based modelling, Citeseerx, Department of Computer Science, Georgia State University, Atlanta (2005).

[44] G. V. Bobashev, D. M. Goedecke, F. Yu, J. M. Epstein, A hybrid epidemic model: combining the advantages of agent-based and equation-based approaches, in: 2007 winter simulation conference, IEEE, 2007, pp. 1532–1537.

[45] M. Marini, C. Brunner, N. Chokani, R. S. Abhari, Enhancing response preparedness to influenza epidemics: Agent-based study of 2050 influenza season in switzerland, Simulation Modelling Practice and Theory 103 (2020) 102091.

[46] H. Rahmandad, J. Sterman, Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models, Management science 54 (5) (2008) 998–1014.

[47] S. R. Sukumar, J. J. Nutaro, Agent-based vs. equation-based epidemiological models: A model selection case study, in: 2012 ASE/IEEE International Conference on BioMedical Computing (BioMedCom), IEEE, 2012, pp. 74–79.

[48] S. K. Kasereka, G. N. Zohinga, V. M. Kiketa, R.-B. M. Ngoie, E. K. Mputu, N. M. Kasoro, K. Kyandoghere, Equation-based modeling vs. agent-based modeling with applications to the spread of covid-19 outbreak, Mathematics 11 (1) (2023) 253.

[49] G. Schneckenreither, N. Popper, G. Zauner, F. Breitenecker, Modelling sir-type epidemics by odes, pdes, difference equations and cellular automata–a comparative study, Simulation Modelling Practice and Theory 16 (8) (2008) 1014–1023.

[50] C. Gräbner, C. S. Bale, B. A. Furtado, B. Alvarez-Pereira, J. E. Gentile, H. Henderson, F. Lipari, Getting the best of both worlds? developing complementary equation-based and agent-based models, Computational Economics 53 (2019) 763–782.

[51] J. Kelleher, A hybrid agent-based and equation based model for the spread of infectious diseases (2020).

[52] A. G. Dunn, T. D. Purnat, A. Ishizumi, T. Nguyen, S. Briand, Measuring the burden of infodemics with a research toolkit for connecting information exposure, trust, and health behaviours, Archives of Public Health 81 (1) (2023) 1–8.

[53] World Health Organization, WHO Infodemic, www.who.int/health-topics/infodemic, accessed: 2024-02-02.

[54] E. Carmi, S. J. Yates, E. Lockley, A. Pawluczuk, Data citizenship: Rethinking data literacy in the age of disinformation, misinformation, and malinformation, Internet Policy Review 9 (2) (2020) 1–22.

[55] D. M. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, et al., The science of fake news, Science 359 (6380) (2018) 1094–1096.

[56] L. Wu, F. Morstatter, K. M. Carley, H. Liu, Misinformation in social media: definition, manipulation, and detection, ACM SIGKDD explorations newsletter 21 (2) (2019) 80–90.

[57] L. Yang, Z. Li, Deadline-aware misinformation prevention in social networks with time-decaying influence, Expert Systems with Applications 238 (2024) 121847.

[58] W. Dong, J. Tao, X. Xia, L. Ye, H. Xu, P. Jiang, Y. Liu, et al., Public emotions and rumors spread during the covid-19 epidemic in

china: Web-based correlation study, Journal of Medical Internet Research 22 (11) (2020) e21933.

[59] W. Pian, J. Chi, F. Ma, The causes, impacts and countermeasures of covid-19 "infodemic": A systematic review using narrative synthesis, Information processing & management 58 (6) (2021) 102713.

[60] H. Garrett, Why does fake news work? on the psychosocial dynamics of learning, belief, and citizenship, Unpacking fake news: An educator's guide to navigating the media with students (2019) 15–29.

[61] M. Charquero-Ballester, J. G. Walter, I. A. Nissen, A. Bechmann, Different types of covid-19 misinformation have different emotional valence on twitter, Big Data & Society 8 (2) (2021) 20539517211041279.

[62] C. Conrad, Inside a cult: gender critical, www.caelanconrad.ca/insideacult, accessed: 2023-06-28 (2022).

[63] S. G. Horne, M. McGinley, N. Yel, M. R. Maroney, The stench of bathroom bills and anti-transgender legislation: Anxiety and depression among transgender, nonbinary, and cisgender lgbq people during a state referendum., Journal of Counseling Psychology 69 (1) (2022) 1.

[64] J. Franceschi, L. Pareschi, M. Zanella, From agent-based models to the macroscopic description of fake-news spread: the role of competence in data-driven applications, Partial Differential Equations and Applications 3 (6) (2022) 68.

[65] A. Zareie, R. Sakellariou, Minimizing the spread of misinformation in

online social networks: A survey, Journal of Network and Computer Applications 186 (2021) 103094.

[66] V. d'Andrea, O. Artime, N. Castaldo, P. Sacco, R. Gallotti, M. De Domenico, Epidemic proximity and imitation dynamics drive infodemic waves during the covid-19 pandemic, Physical Review Research 4 (1) (2022) 013158.

[67] C. Berceanu, N. Arshad, M. Patrascu, Contagion propagation with rule-based reasoning and decentralized control in an agent-based susceptible-infected-recovered-susceptible infodemic model, in: 2023 International Conference on Big Data, Knowledge and Control Systems Engineering (BdKCSE), IEEE, 2023, pp. 1–6.

[68] S. Van Der Linden, Misinformation: susceptibility, spread, and interventions to immunize the public, Nature Medicine 28 (3) (2022) 460–467.

[69] U. K. Ecker, S. Lewandowsky, J. Cook, P. Schmid, L. K. Fazio, N. Brashier, P. Kendeou, E. K. Vraga, M. A. Amazeen, The psychological drivers of misinformation belief and its resistance to correction, Nature Reviews Psychology 1 (1) (2022) 13–29.

[70] X. Nan, Y. Wang, K. Thier, Why people believe health misinformation and who are at risk? a systematic review of individual differences in susceptibility to health misinformation, Social Science & Medicine (2022) 115398.

[71] E. J. Marsh, A. D. Cantor, N. M. Brashier, Believing that humans swallow spiders in their sleep: False beliefs as side effects of the processes that support accurate knowledge, in: Psychology of learning and motivation, Vol. 64, Elsevier, 2016, pp. 93–132.

[72] T. Prike, M. M. Arnold, P. Williamson, The relationship between anomalistic belief, misperception of chance and the base rate fallacy, Thinking & Reasoning 26 (3) (2020) 447–477.

[73] J. E. Uscinski, A. M. Enders, C. Klofstad, M. Seelig, J. Funchion, C. Everett, S. Wuchty, K. Premaratne, M. Murthi, Why do people believe covid-19 conspiracy theories?, Harvard Kennedy School Misinformation Review 1 (3) (2020).

[74] C. Unkelbach, A. Koch, R. R. Silva, T. Garcia-Marques, Truth by repetition: Explanations and implications, Current Directions in Psychological Science 28 (3) (2019) 247–253.

[75] L. K. Fazio, Repetition increases perceived truth even for known falsehoods, Collabra: Psychology 6 (1) (2020).

[76] E. Bonnevie, J. Sittig, J. Smyser, The case for tracking misinformation the way we track disease, Big Data & Society 8 (1) (2021) 20539517211013867.

[77] W. Duan, Z. Fan, P. Zhang, G. Guo, X. Qiu, Mathematical and computational approaches to epidemic modeling: a comprehensive review, Frontiers of Computer Science 9 (2015) 806–826.

[78] O. N. Bjørnstad, K. Shea, M. Krzywinski, N. Altman, The seirs model for infectious disease dynamics., Nature methods 17 (6) (2020) 557–559.

[79] O. M. Otunuga, Estimation of epidemiological parameters for covid-19 cases using a stochastic seirs epidemic model with vital dynamics, Results in Physics 28 (2021) 104664.

[80] X. Hong, Y. Han, G. Tanaka, B. Wang, Co-evolution dynamics of epidemic and information under dynamical multi-source information and behavioral responses, Knowledge-Based Systems 252 (2022) 109413.

[81] H. D. Aghbolagh, M. Ye, L. Zino, Z. Chen, M. Cao, Coevolutionary dynamics of actions and opinions in social networks, IEEE Transactions on Automatic Control (2023).

[82] C. E. Shannon, A mathematical theory of communication, The Bell system technical journal 27 (3) (1948) 379–423.

[83] N. Gontier, S. Hartmann, M. Pleyer, E. D. Rodrigues, Combinatoriality and compositionality in communication, skills, tool use, and language, International Journal of Primatology (2024) 1–9.

[84] S. Whittaker, Personal information management: from information consumption to curation, Annual review of information science and technology 45 (1) (2011) 1.

[85] N. A. Cooke, Posttruth, truthiness, and alternative facts: Information behavior and critical information consumption for a new age, The library quarterly 87 (3) (2017) 211–221.

[86] O. Holland, C. Melhuish, Stigmergy, self-organization, and sorting in collective robotics, Artificial life 5 (2) (1999) 173–202.

[87] L. V. Bryant, The youtube algorithm and the alt-right filter bubble, Open Information Science 4 (1) (2020) 85–90.

[88] I. Cohen, Y. Huang, J. Chen, J. Benesty, Pearson correlation coefficient, Noise reduction in speech processing (2009) 1–4.

[89] M. V. Shcherbakov, A. Brebels, N. L. Shcherbakova, A. P. Tyukov, T. A. Janovsky, V. A. Kamaev, et al., A survey of forecast error measures, World applied sciences journal 24 (24) (2013) 171–176.

[90] A. Collis, K. Garimella, A. Moehring, M. A. Rahimian, S. Babalola, N. H. Gobat, D. Shattuck, J. Stolow, S. Aral, D. Eckles, Global survey on covid-19 beliefs, behaviours and norms, Nature Human Behaviour 6 (9) (2022) 1310–1317.

[91] M. Martinelli, G. A. Veltri, Covid-19 vaccine acceptance: A comparative longitudinal analysis of the association between risk perception, confidence, and the acceptance of a covid-19 vaccine, Risk Analysis (2023).

[92] U. Wilensky, NetLogo. http://ccl.northwestern.edu/netlogo/, Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL (1999).

[93] Harvard Institute for Quantitative Social Science, The Future of Media Project, www.projects.iq.harvard.edu/futureofmedia/about, accessed: 2024-02-12.

[94] E. Hunter, B. Mac Namee, J. Kelleher, A hybrid agent-based and equation based model for the spread of infectious diseases, Journal of Artificial Societies and Social Simulation 23 (4) (2020) 14.

[95] S. M. Owoeye, Epidemiological game-theory dynamics of the coevolution of infodemic and pandemic with vaccination, Georgia State University, Thesis. (2022).

[96] M. A. Bedau, Weak emergence, Philosophical perspectives 11 (1997) 375–399.

[97] G. Box, All models are wrong, but some are useful, Robustness in Statistics 202 (1979) (1979) 549.

[98] N. M. Brashier, G. Pennycook, A. J. Berinsky, D. G. Rand, Timing matters when correcting fake news, Proceedings of the National Academy of Sciences 118 (5) (2021) e2020043118.

[99] J. Cook, S. Lewandowsky, U. K. Ecker, Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence, PloS one 12 (5) (2017) e0175799.

[100] D. MacFarlane, L. Q. Tay, M. J. Hurlstone, U. K. Ecker, Refuting spurious covid-19 treatment claims reduces demand and misinformation sharing, Journal of Applied Research in Memory and Cognition 10 (2) (2021) 248–258.