



UNIVERSITÀ DEGLI STUDI DI MILANO - BICOCCA

Scuola di Scienze

Dipartimento di Informatica, Sistemistica e Comunicazione

Corso di Laurea Magistrale in Informatica

# AGENT-BASED MODELING AND LEARNING FOR EPIDEMIOLOGICAL STUDIES

**Relatore:** Prof. Antoniotti Marco

**Correlatore:** Prof.

**Relazione della prova finale di:**

Matteo Stievano

Matricola 829535

April 30, 2023

**Anno Accademico 2022-2023**

### **Abstract**

L'abstract, che probabilmente dovrà essere presentato in doppia lingua (ITA, ENG) rappresenta un breve sunto di tutto l'elaborato di cui fa da cappello. Ovviamente non dovrà entrare nel dettaglio, e anzi dovrà essere scorrevole e chiaro, senza soffermarsi troppo su alcun dettaglio, in quanto ci saranno le sezioni successive per espandere tali conoscenze. Ad occhio e croce potrebbe prendere al massimo mezza pagina (così che se necessario ci stiano le due versioni su una pagina).

# Contents

<b>1</b>	<b>Introduzione</b>	<b>1</b>
<b>2</b>	<b>Stato dell'Arte</b>	<b>4</b>
2.1	Modelli Compartmentali . . . . .	4
2.1.1	Modelli Deterministici . . . . .	10
2.1.2	Modelli Stocastici . . . . .	11
2.2	Modelli ad Agente . . . . .	12
2.2.1	Discretizzazione . . . . .	14
2.3	Epidemiologia . . . . .	16
2.3.1	Causalità in Epidemiologia . . . . .	16
2.4	Julia . . . . .	18
2.4.1	Agents.jl . . . . .	19
2.4.2	SciML.ai . . . . .	20
2.4.3	SafeBlues . . . . .	20
<b>3</b>	<b>Metodi e Modelli</b>	<b>23</b>
3.1	COVID-19 . . . . .	23
3.1.1	Modelli SEIR . . . . .	23
3.1.2	Modelli ad Agente . . . . .	23
3.2	Altre epidemie . . . . .	24
3.2.1	Ebola . . . . .	24
3.2.2	Aviaria . . . . .	24
3.2.3	Influenza . . . . .	24
3.3	Monitoraggio e Intervento . . . . .	25
<b>4</b>	<b>Risultati Ottenuti</b>	<b>26</b>
<b>5</b>	<b>Conclusioni</b>	<b>27</b>

# 1 Introduzione

L'utilizzo di metodi e tecniche sempre più sofisticati da parte della comunità scientifica, in particolare modo da parte di epidemiologi e medici, è sempre stato argomento di grande dibattito e interesse. Negli ultimi anni il mondo si è espanso esponenzialmente divenendo sempre più connesso, incrementando drasticamente la probabilità che un virus affligga a livello mondiale la popolazione creando un disastro senza precedenti.

La storia dell'uomo è costellata di epidemie, e solamente alcune tra loro si sono guadagnate il privilegio di essere ricordate e tra queste solamente una piccola parte ha ottenuto il primato di essere ricordata come una catastrofe. Forse è proprio questo che le ha rese così salde nell'immaginario comune aumentando la loro già imponente aurea di terrore.

Per fare degli esempi possiamo citare: la peste nera [49] che a partire dalla metà del 14esimo secolo ritornò in Europa uccidendo venti milioni di persone in soli sei anni, l'epidemia di tifo [54] che non solo fu fatale durante il periodo delle crociate, ma anche durante la seconda guerra mondiale all'interno dei campi di concentramento nazisti, oppure le varie influenze, prima tra tutte quella spagnola [45] la quale nel periodo del primo dopo guerra, ovvero tra il 1918 e il 1920, uccise 50 milioni di persone in tutto il mondo, oppure l'epidemia di AIDS [53], che ha all'attivo dal 1981 oltre 75 milioni di casi e 35 milioni di morti.

Se ci si sofferma un attimo pensando proprio a questo tipo di pandemia, quella influenzale, ci viene da tirare un sospiro di sollievo, in quanto oramai come cittadini del primo mondo l'idea di influenza non ci fa più così paura. Eppure dovrebbe, e sfortunatamente lo abbiamo ricordato nel peggior modo possibile.

La pandemia di SARS-CoV-2 [48] scoppiata negli ultimi mesi del 2019 condizionò l'intera umanità per circa 3 anni, e nel momento che sto scrivendo queste righe continua a condizionarla. Questa pandemia si è macchiata di aver mietuto, allo stato attuale delle cose, quasi 7 milioni di vite accertate. Questa tragedia rimarrà impressa nella memoria umana in quanto capace di aver messo in crisi l'intero sistema governativo mondiale, creando uno stato di allarme, panico e alle volte perfino isteria, che pochi altri avvenimenti sono stati in grado di fare.

Osservando le statistiche proprie di questa epidemia, ciò che balza subito all'occhio è sicuramente il numero associato alle morti e agli infetti: quasi 7 milioni di morti e più di 700 milioni di infetti [30]. Numeri che da soli basterebbero a mettere a disagio qualsiasi lettore, eppure altri dati, nascosti a prima vista, possono fornire ulteriori macabre informazioni. Un esempio tra tutti è l'effetto che un'epidemia del genere ha avuto sull'economia [20], portando disagi generalizzati ovunque.

Legato al disagio economico vi è un altro dato preoccupante definito come poverty trap [8] [43]. Questo fenomeno nasce in quegli ambienti in cui le condizioni di povertà economica e la prevalenza di malattie possono imprigionare una società in uno stato persistente di bassa sanità e sempre maggiore povertà; questo fenomeno ciclico si auto sostiene e prende il nome di positive feedback [50]. Per ultimo, ma non per questo meno importante, l'effetto di una pandemia può portare ad instabilità governative le quali possono portare a un arretramento del sistema sanitario e di welfare [13], ricadendo come sopra descritto all'interno del fenomeno di positive feedback.

Questi sono solamente alcuni esempi dei problemi che possono sorgere, e che sono sorti con lo scoppio di una pandemia globale come è stata quella del COVID-19. Per questo la comunità scientifica, in

particolare gli epidemiologi cercano soluzioni sempre più efficaci e accurate per prevenire, arginare e contrastare avvenimenti del genere.

L'epidemiologia è una disciplina nata di recente evolutasi insieme alle esigenze della società ogni qualvolta una nuova emergenza sanitaria faceva irruzione nella quotidianità. La prima definizione di epidemiologia è stata data da Lilienfeld [15] nel 1978, e cita:

*l'epidemiologia è un modo di ragionare riguardo le malattie, e si occupa di effettuare inferenza biologica derivata dall'osservazione di fenomeni patologici all'interno di una popolazione.*

Con il tempo questa definizione ha subito molti cambiamenti, derivati anche e soprattutto dall'espansione degli ambiti relazionati all'epidemiologia; con l'aggiunta ad esempio della farmacoepidemiologia, dell'epidemiologia molecolare e dell'epidemiologia genetica. Non solo, ambiti come etica, filosofia ed epistemologia sono estremamente importanti ed influenti nella crescita e sviluppo di questa materia [15].

Attualmente con il termine epidemiologia si intende la disciplina biomedica che studia la distribuzione e la frequenza delle malattie ed eventi di rilevanza sanitaria nella popolazione. L'epidemiologia si occupa di analizzare le cause, il decorso e le conseguenze delle malattie [44]. Secondo Last et al (1998) l'epidemiologia è definita come:

*lo studio della distribuzione e dei determinanti delle situazioni o degli eventi collegati alla salute in una specifica popolazione, e l'applicazione di questo studio al controllo dei problemi di salute.*

Uno degli strumenti più utilizzati in epidemiologia sono le simulazioni software [51]. Dato lo scopo dell'epidemiologia, questa necessita di avvalersi di modelli matematici [36] che aiutano i ricercatori a trarre conclusioni sul sistema che analizza. Sistemi simili vengono definiti come complessi, [16] [24] ovvero sistemi dinamici a multicomponenti che tipicamente interagiscono tra loro, e che sono descrivibili tramite modelli matematici.

I primi modelli chiamati compartimentali [7] basano il proprio funzionamento sullo studio di gruppi di individui disgiunti che interagiscono tra loro, analizzabili tramite un sistema di equazioni ordinarie differenziali (ODE) [9]. Questo approccio è stato teorizzato da Kermack e McKendrick nel 1927 applicando una modellazione matematica al comportamento delle malattie infettive su un gruppo di individui, tenendo in considerazione la variabile del tempo. Da qui è nato il famoso modello Susceptible-Infectious-Recovered (SIR) [42], il quale tuttora viene utilizzato ampiamente insieme alle sue varianti. Successivamente con l'unione di svariate discipline come: la teoria dei giochi, i sistemi complessi, il comportamento emergente, la sociologia computazionale, i sistemi multiagente e la programmazione evolutiva sono nati i modelli ad agente [39] [6], modelli computazionali autonomi per la simulazione di sistemi complessi.

Grazie alla loro teorizzazione negli anni 1940 ma soprattutto grazie al loro uso concreto dagli anni 1990 il mondo della simulazione ha avuto un grande balzo in avanti. Ciò che rende questi sistemi molto flessibili e potenti è la capacità di far emergere spontaneamente dei comportamenti complessi da un insieme di regole semplici. Ovviamente però la richiesta di risorse e capacità computazionale è estremamente elevata rispetto alla controparte di modelli puramente matematici.

Ogni approccio ha dei pro e dei contro e la vera potenza di questi modelli sta principalmente nella acutezza di chi deve poi farne uso. In base ai compromessi e le assunzioni fatte durante la fase

di modellazione ogni approccio può rivelarsi vincente. Uno dei compromessi maggiori che viene generalmente applicato a questi modelli è quello della discretizzazione dell'ambiente [23]. La realtà è a tutti gli effetti un insieme continuo di eventi, ma essendo tutti i dispositivi di calcolo discreti è impossibile simulare avvenimenti continui (siano essi nel tempo o nello spazio) in maniera diretta e perciò bisogna fare dei compromessi, trasformando il proprio spazio di lavoro in uno più adatto alle macchine che lo devono simulare.

Un'altra assunzione generalista e approssimativa per definizione, ma necessaria per la costruzione di un modello che analizzi il decorso di un'epidemia all'interno di una società, è quella legata alla tipologia di comportamento che verrà mostrato dagli esseri umani in condizioni di pericolo [38] [14]. Queste sono solo alcune delle assunzioni e compromessi che bisogna fare quando ci si avvicina al mondo della simulazione. Tuttavia applicare delle assunzioni, alle volte anche forti e controintuitive, non sempre è sinonimo di errore. Alle volte tramite lo studio della causalità degli eventi [16] [31] è possibile astrarre un set minimale di assunzioni che se applicate danno la capacità al modello di rappresentare molto bene il comportamento desiderato.

Certamente ci saranno alcune discrepanze soprattutto in casi estremi, ma è un'eventualità che viene tenuta in considerazione ogni qualvolta si parla di simulazione, e che non è possibile eliminare del tutto.

*La causalità o causa effetto, è quell'influenza per cui un evento, un processo, uno stato o un oggetto contribuiscono nella produzione di un nuovo evento, processo, stato o effetto, dove la causa è parzialmente responsabile dell'effetto e l'effetto è parzialmente dipendente dalla causa.*

A prima vista non sembra una tematica molto complessa o di difficile approccio, complice il fatto che in quanto esseri umani siamo una specie che si è evoluta per trovare una correlazione tra gli eventi, ma correlazione non significa, e soprattutto non implica, causalità [3]. Questo tema si riferisce all'incapacità legittima di dedurre la relazione di causa - effetto tra due eventi o variabili solamente sulla base di una osservazione della loro associatività o correlazione.

Diventa chiaro come una delle parti più complesse dell'epidemiologia sia proprio quella di stabilire le cause di un dato fenomeno e comprendere come un determinato intervento su di esse influenzi quest'ultimo [16] [31].

Questa rapida introduzione all'epidemiologia e alla simulazione di sistemi complessi tramite l'utilizzo in particolare di modelli ad agente sarà l'argomento cardine dell'intero elaborato. Nelle sezioni successive verrà proposta un'analisi dello stato dell'arte dell'attuale panorama epidemiologico e simulativo, con alcune rapide digressioni sul problema della discretizzazione, problema assai sentito nell'ambito della simulazione.

Successivamente verranno portati alla luce alcuni esempi e modelli di sistemi ad agente che si occupano di modellare differenti tipi di epidemie, tutte caratterizzate però dal fatto di essere epidemie infettive a diffusione principalmente aerea, come ad esempio l'ebola, l'aviaria e l'influenza.

Successivamente vi sarà un'analisi più dettagliata riguardo la pandemia da COVID-19 che recentemente ci ha colpito. Infine verranno analizzati alcuni metodi di monitoraggio di queste simulazioni con un focus su alcuni metodi di intervento.

## 2 Stato dell'Arte

### 2.1 Modelli Compartmentali

In epidemiologia i modelli compartimentali sono una tecnica di modellazione generica che si pre-dispone molto bene allo studio complessivo del comportamento di una malattia infettiva [42]. Questa tecnica di modellazione si applica anche ad altre branche della scienza, come ad esempio la finanza.

Questa tecnica di modellazione matematica basa il proprio funzionamento sull'assunzione che, data una popolazione di individui, questi vengano etichettati in maniera differente, in base allo stato di progressione della malattia che hanno, o non hanno, contratto. Così facendo si vanno a definire dei compartimenti ben separati che possono interagire tra loro, ma che rimangono comunque chiaramente distinti l'uni dagli altri.

Il modello che tutt'ora viene usato come riferimento e come base per lo studio e modellazione è il così detto modello SIR: Susceptible, Infectious, Recovered.

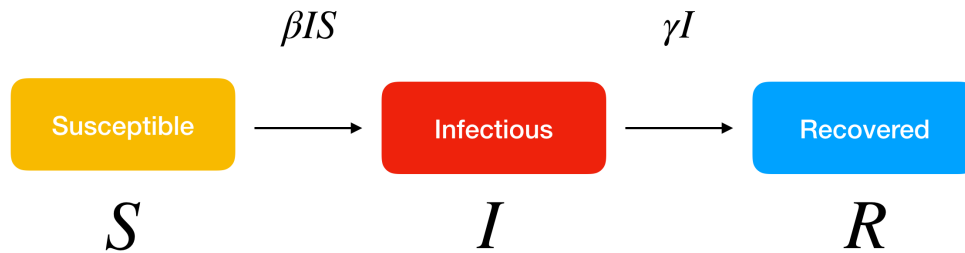


Figure 1: Struttura modello SIR

Questo modello è stato ideato all'inizio del 20esimo secolo, più precisamente nel 1917, da Kermack e McKendrick. Come introdotto questo modello si basa sull'assunzione che all'interno di una popolazione durante il decorso di una malattia vi possano esistere solamente tre stadi in cui un individuo può essere inserito:

- **Susceptible:** Questo stadio rappresenta lo stato iniziale per la maggior parte degli individui all'interno di una popolazione. Rappresenta il numero di persone che possono contrarre la malattia.
- **Infectious:** Questo stadio rappresenta tutti quegli individui che dallo stato di Susceptible, dopo essere venuti in contatto con un individuo infetto, diventano a loro volta individui infetti.
- **Recovered:** Questo stadio rappresenta una duplice categoria, quella degli individui che alla fine del decorso della malattia sopravvivono ad essa, e quelli che invece muoiono a causa di questa. Generalmente questo stato viene anche definito come Removed.

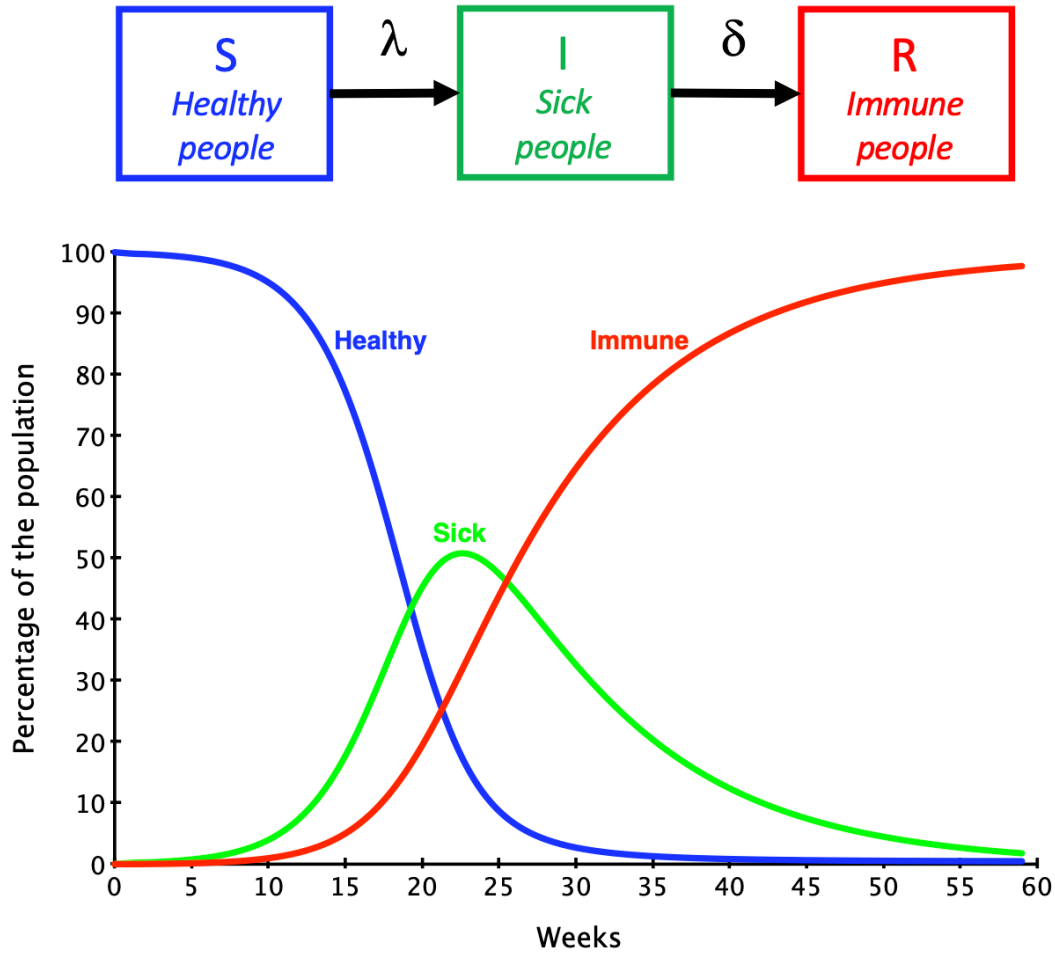


Figure 2: Visualizzazione grafico modello SIR

Questo semplice modello funziona grazie all'utilizzo di un sistema di Equazioni Ordinarie Differenziali (ODE) [9]. Una volta definiti i parametri iniziali utili per la simulazione del decorso di una malattia, è pressochè immediato trovare il risultato al tempo  $T$  del sistema. Questo modello può essere calcolato tenendo in considerazione un andamento più caotico del sistema, utilizzando invece che un sistema di ODE, un sistema di SDE [2] ovvero di Equazioni Differenziali Stocastiche.

Con il tempo questo sistema è stato espanso per tenere in considerazione comportamenti differenti sia della popolazione che delle malattie, andando a definire una moltitudine di modelli utili a differenti scopi. In epidemiologia il modello di riferimento maggiormente utilizzato è il modello SEIR (Susceptible, Exposed, Infectious, Recovered) con le sue varianti proprie di ogni approccio.



Alcuni modelli definiscono i propri stati in maniera da considerare come stato interno al sistema anche l'agente patogeno, così da poter modellare e simulare l'andamento dell'infettività della pandemia in relazione alle contromisure prese, siano esse farmaceutiche o non. Ne è un esempio il modello proposto da [26] nel quale il modello viene proposto con l'idea di incorporare le misure di distanziamento sociale come variabili per misurare la loro efficacia contro la recente pandemia da COVID-19.

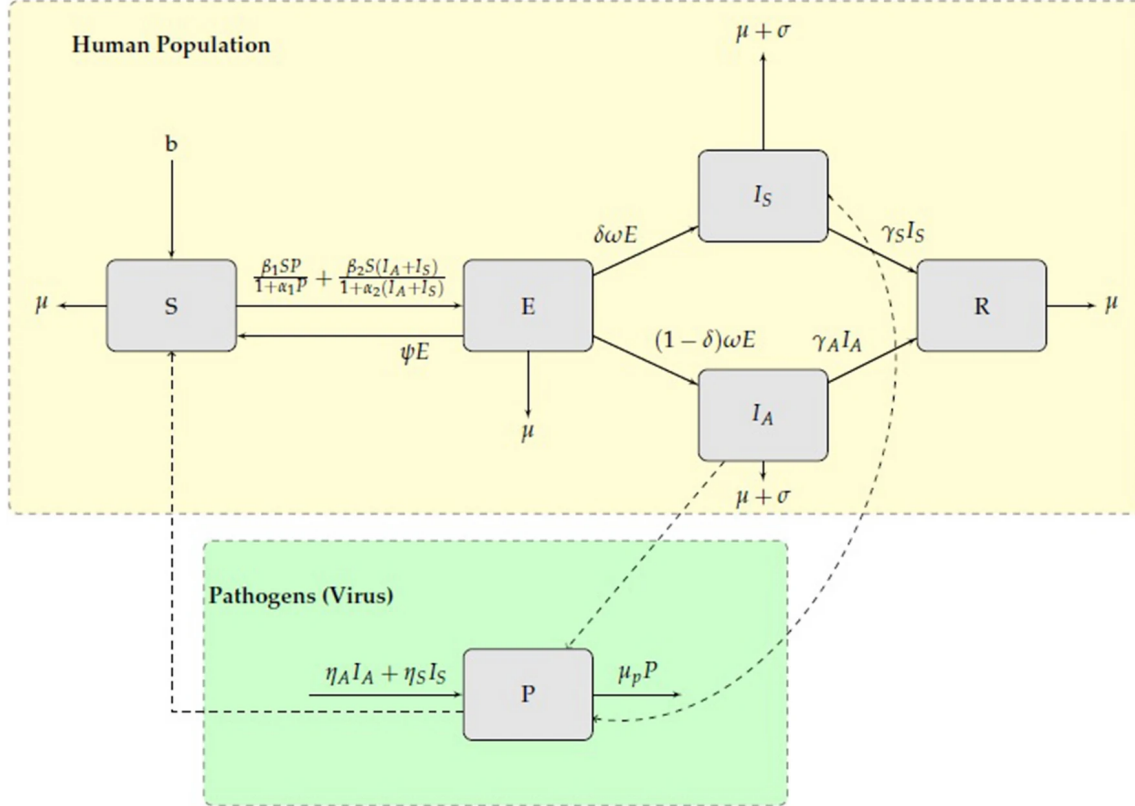


Figure 3: Esempio di modello SEIR preso dall'articolo [26]

Altri modelli, come quello proposto da [18] mantengono la stessa filosofia, ovvero quella di analizzare l'efficacia delle misure di prevenzione non farmaceutiche sull'andamento di un'epidemia, ma non modellano esplicitamente l'agente patogeno come stato del modello, bensì variando i parametri di infettività e contagio, arrivano allo stesso risultato. Un'altra differenza tra i due approcci è quella della tipologia di equazioni differenziali utilizzate, [26] hanno utilizzato delle ODE mentre [18] delle SDE.

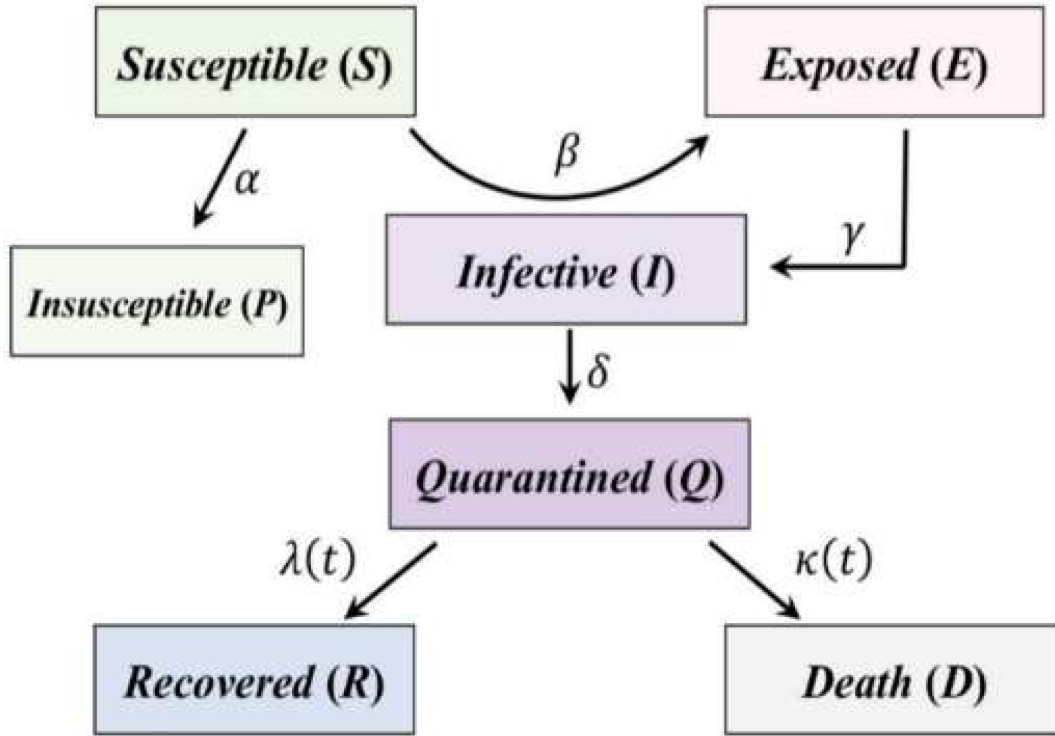


Figure 4: Esempio di modello SEIR preso dall'articolo [18]

Il motivo per cui viene utilizzato il modello SEIR come base è perchè permette di modellare una caratteristica intrinseca di una malattia infettiva, ovvero il periodo di latenza che un individuo appena infettato ha prima di diventare infettivo a sua volta e mostrare i sintomi di infezione. Questo permette di osservare quanto le misure di sicurezza e prevenzione non farmaceutiche sono efficaci sulla popolazione tenendo in considerazione un tempo di ritardo intrinseco nel feedback tra l'attuamento delle misure di prevenzione e i risultati positivi di queste ultime.

Una delle modifiche più utilizzate a questo modello è quella di avere un sistema ciclico, ovvero in cui gli individui che entrano nello stato R non diventano immuni alla malattia a tempo indefinito, ma perdono questa loro caratteristica di immunità dopo un dato periodo di tempo. Questo permette di modellare con più accuratezza le malattie infettive stagionali come ad esempio la comune influenza o il raffreddore, oppure mostrare l'andamento ad ondate di altre malattie che hanno la caratteristica di mutare molto velocemente, come è stato per il COVID-19 e le sue innumerevoli varianti.

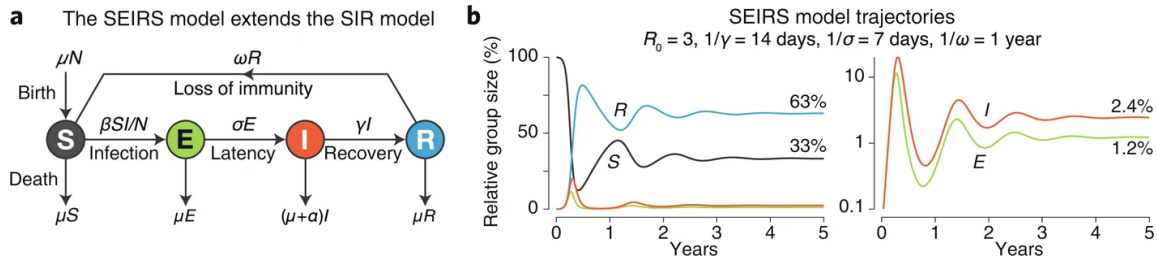


Figure 5: Modello SEIRS preso dall'articolo [7]

Questa variante denominata SEIRS permette invece di osservare, non solo l'andamento e l'efficacia delle contromisure non farmaceutiche, ma anche di quelle farmaceutiche, come ad esempio i vaccini; o più in generale l'andamento della così detta immunità di gregge [7].

Rimanendo sull'idea di voler analizzare l'efficacia di un vaccino, una modifica comune al modello SEIR è quella legata all'aggiunta dello stato V, Vaccinated, come stato esplicito oppure oppure implicito al modello. Questa variazione permette di modellare con più attenzione l'efficacia di un vaccino una volta introdotto all'interno della popolazione, ma più in generale permette di osservare l'efficacia di una politica di vaccinazione in relazione al numero di vaccinazioni effettuate in un determinato periodo di tempo. Questo viene solitamente affiancato con un modello ciclico, così da poter osservare come bisogna modificare le proprie politiche vaccinali in vista di ondate cicliche più o meno intense di infezioni.

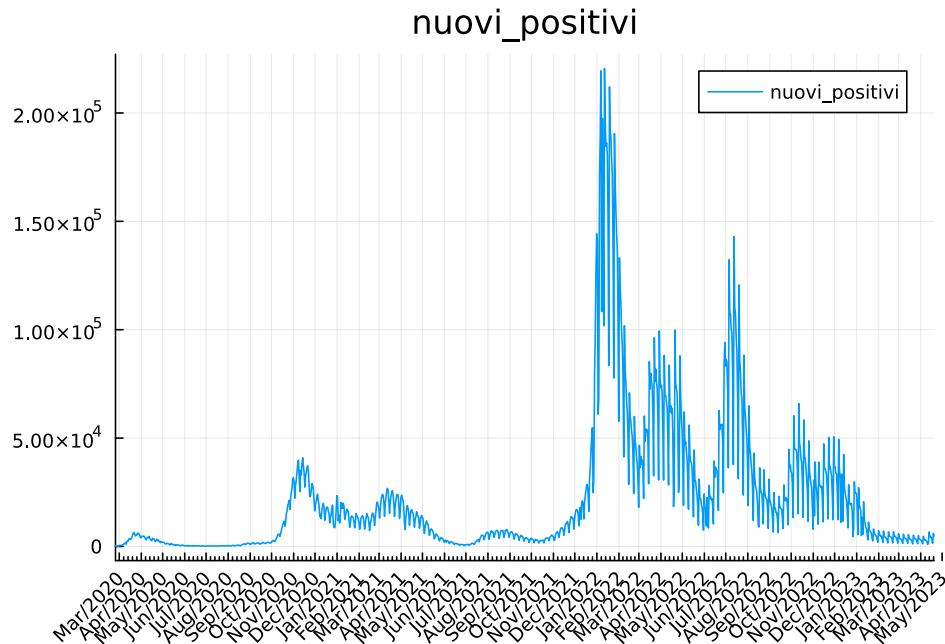


Figure 6: Esempio di ondate di infettività. Dati del Dipartimento di Protezione Civile Italiana

Ai fini pratici di una simulazione avere uno stato esplicitamente definito oppure ricavabile dalle probabilità di transizione degli altri stati è pressochè indifferente, e potrebbe essere richiesta una differenziazione solamente in caso in cui si avrebbe una differenza sostanziale tra lo stato R e V, ad esempio in termini di protezione dalla malattia, durata immunità etc... .

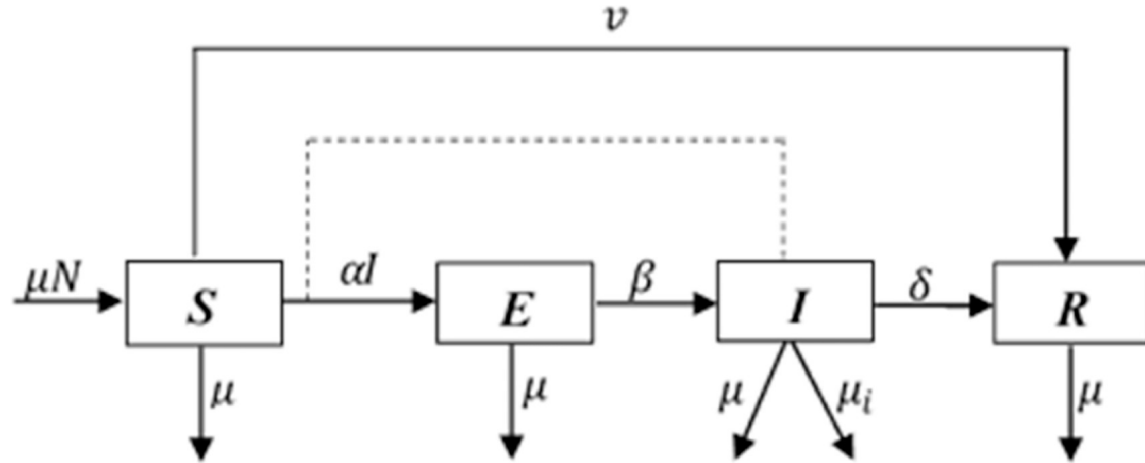


Figure 7: Esempio di modello SEIRV con stato implicito per la condizione V

Non essendoci un numero massimo di stati, e quindi di equazioni, utilizzabili all'interno del modello, ogni individuo è libero di definire un numero di equazioni arbitrario che rispecchia la sua idea di modellazione del sistema. Ne è un esempio il modello riportato in [17].

Come precedentemente introdotto esistono due grandi famiglie di modelli per la simulazione, e sono rispettivamente la famiglia di modelli deterministici e quella di modelli stocastici.

### 2.1.1 Modelli Deterministici

I modelli deterministici vengono principalmente utilizzati per la loro immediatezza e riproducibilità. Infatti un modello deterministico, una volta impostati i parametri necessari riprodurrà sempre lo stesso risultato. Questo tipo di approccio, seppur utilizzato in larga scala come ad esempio da [7] [26], [17] si basa su delle assunzioni molto forti che non sempre rispecchiano la realtà.

Infatti i modelli deterministici hanno il grosso problema di essere affidabili solamente nel caso in cui vi siano dati sufficientemente grandi, cosa che non sempre è possibile avere [42]. Essendo modelli di tipo deterministico, avendo dei parametri di infettività maggiori di zero, con un numero di individui infetti anch'esso maggiore di zero, si tenderà ad avere nel lungo periodo un andamento di equilibrio endemico derivato dalle equazioni e dal modello utilizzato. Questo comportamento però non sempre rispetta la realtà, ma come precedentemente accennato, in casi in cui si hanno grandi quantità di dati legati principalmente alla popolazione, questi modelli si comportano in maniera affidabile.

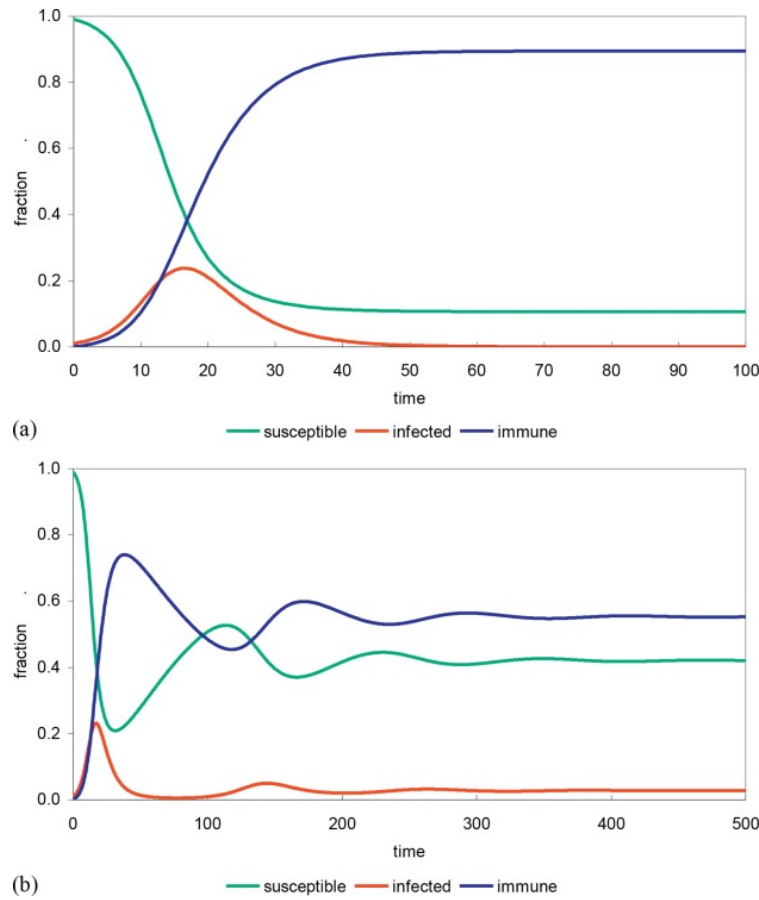


Figure 8: Esempio equilibrio endemico

### 2.1.2 Modelli Stocastici

I modelli stocastici, seppur più complessi e non determinabili a priori, permettono una modellazione più veritiera e simile alla realtà in quanto tengono in considerazione variazioni randomiche che possono capitare durante il decorso di una pandemia. Tuttavia questa loro caoticità richiede che per ottenere risultati robusti debbano essere eseguiti e computati molteplici volte, e la media dei loro risultati è il valore vero da tenere in considerazione. Questi modelli sono stati applicati durante la pandemia da COVID-19 come ad esempio da [18].

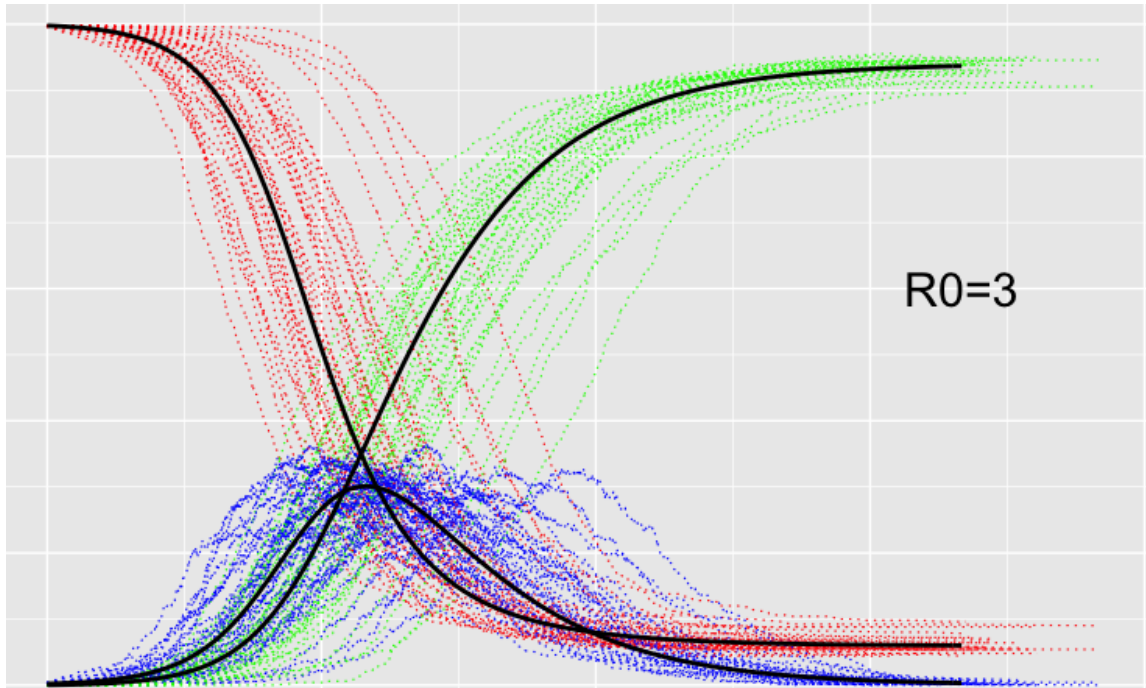


Figure 9: Modello SIR stocastico

E' immediato notare come il comportamento delle curve sia imprevedibile se preso singolarmente, e che non sempre esiste uno stato di equilibrio endemico chiaro e definito come quello ottenibile da un modello deterministico. Ciò nonostante effettuando molte simulazioni è possibile vedere come il comportamento generale del modello sia comunque simile a quello di un modello deterministico.

## 2.2 Modelli ad Agente

Un modello ad agente e' un modello computazionale per la simulazione delle azioni e interazioni di un insieme di agenti autonomi, siano essi individui o gruppi di individui, con l'obiettivo di comprendere il comportamento del sistema e la relazione che vige con i suoi risultati [40] [39].

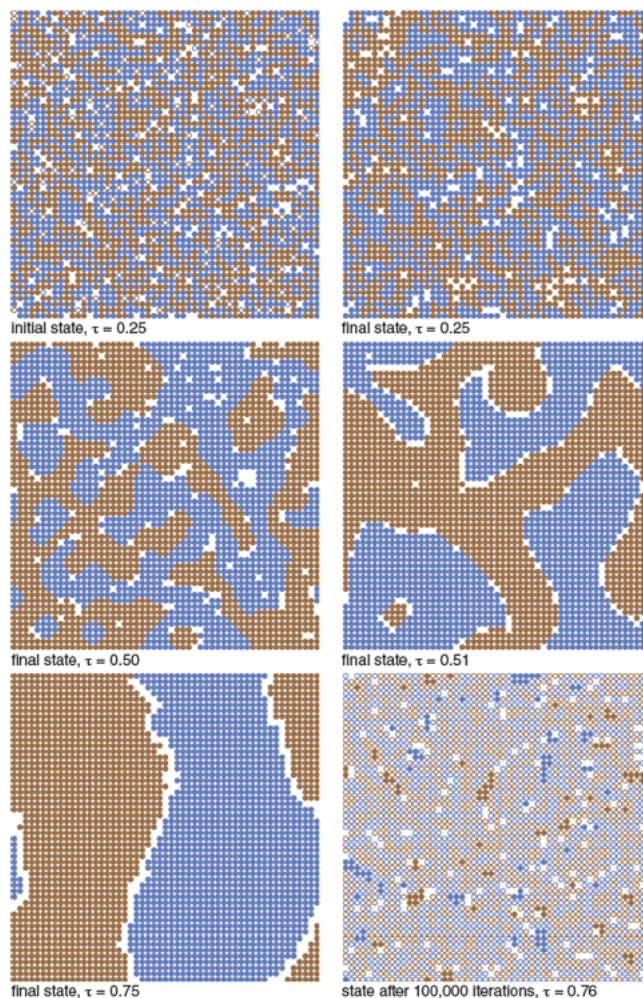


Figure 10: Schelling's model of segregation simulated via ABM

L'utilizzo di modelli ad agente in epidemiologia e' una tecnica nota che da piu' di un decennio viene impiegata per simulare e comprendere le situazioni e i problemi piu' disparati, tutti pero' generalmente accomunati dal fatto che come parametro portante vi sia il comportamento umano [19] [14] [38] [6]. Uno dei parametri piu' caratterizzanti che da sempre sono stati tenuti in considerazione e assunti come ristretti ad una piccola cerchia, sono le interazioni sociali tra individui. Questo sovrainsieme di parametri racchiude molteplici sottoinsiemi di parametri che descrivono delle in-

terazioni piu' specifiche ma che possono essere raggruppate come macro categoria se si scende a compromessi.

Questa specifica e' importante in quanto uno delle sfide piu' grandi che il mondo della simulazione, e quindi quello epidemiologico devono affrontare e' proprio quello di trovare un modo efficiente e soprattutto realistico di simulare le interazioni sociali tra individui, in quanto queste possono influenzare notevolmente i risultati di una simulazione definendola utile oppure inutile [37]. Non soltanto, un'altra sfida e' il modo con cui si decide di rappresentare lo spazio (e il tempo) all'interno della simulazione. In base al tipo di discretizzazione effettuata una simulazione potrebbe essere utile in un campo ma totalmente inutile in un altro [23].

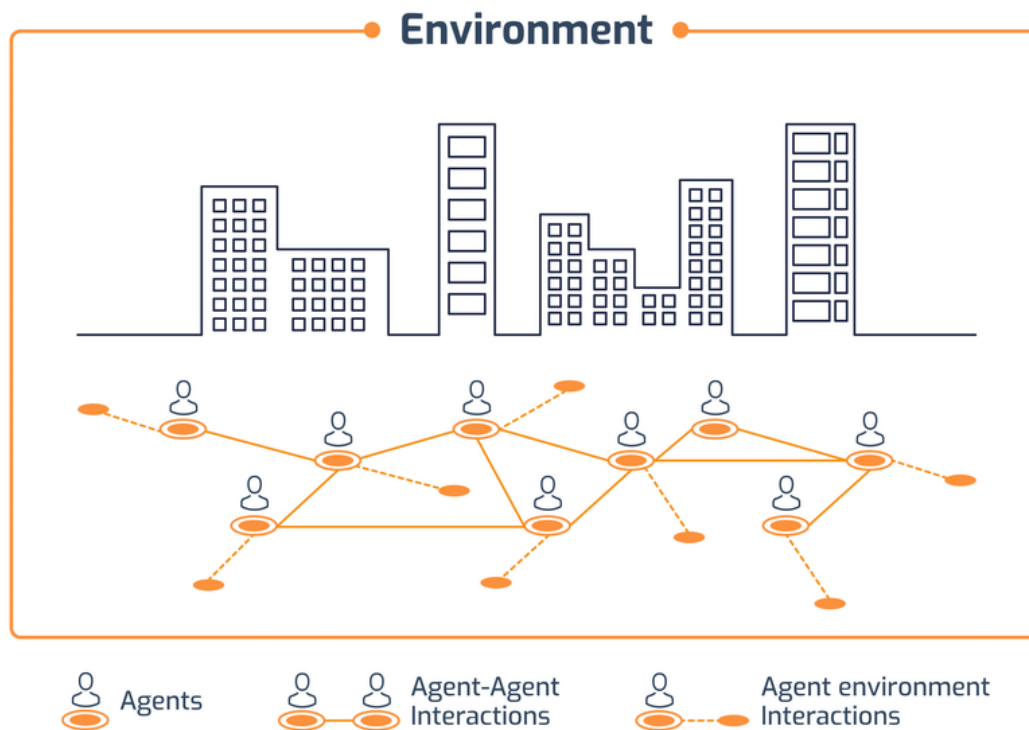


Figure 11: Rappresentazione schematica di un modello ad agente

Con l'arrivo della pandemia da COVID-19 molti ricercatori hanno focalizzato la propria attenzione sull'idea di sviluppare un modello ad agente puro oppure ibrido [25] (ibridato con delle ODE oppure delle Partial DE), con l'obiettivo di trovare un modello di simulazione in grado di simulare in maniera affidabile il decorso di una pandemia tenendo in considerazione le variabili piu' stocastiche e imprevedibili come il comportamento umano. L'ambiente di lavoro simulato era generalmente un ambiente controllato che poteva essere una citta' come in [32],

Un'altro obiettivo e' quello di osservare l'impatto degli interventi non farmaceutici sulla popolazione



come riporta [29] [22]. Altri ancora invece utilizzano l'approccio simulativo tramite modelli ad agente per estrapolare delle ODE tramite l'analisi del modello, come riportato da [27].

### 2.2.1 Discretizzazione

La tematica della discretizzazione e' una delle proprieta' fondamentali e al contempo uno dei problemi atavici della simulazione. Il mondo in cui viviamo e' un mondo continuo, ma gli strumenti che attualmente abbiamo per simularlo sono discreti, per cui ogni qualvolta che vogliamo simulare un evento dobbiamo decidere in che modo adattare la realta' alla simulazione, andando inequivocabilmente a perdere informazioni nel processo [23].

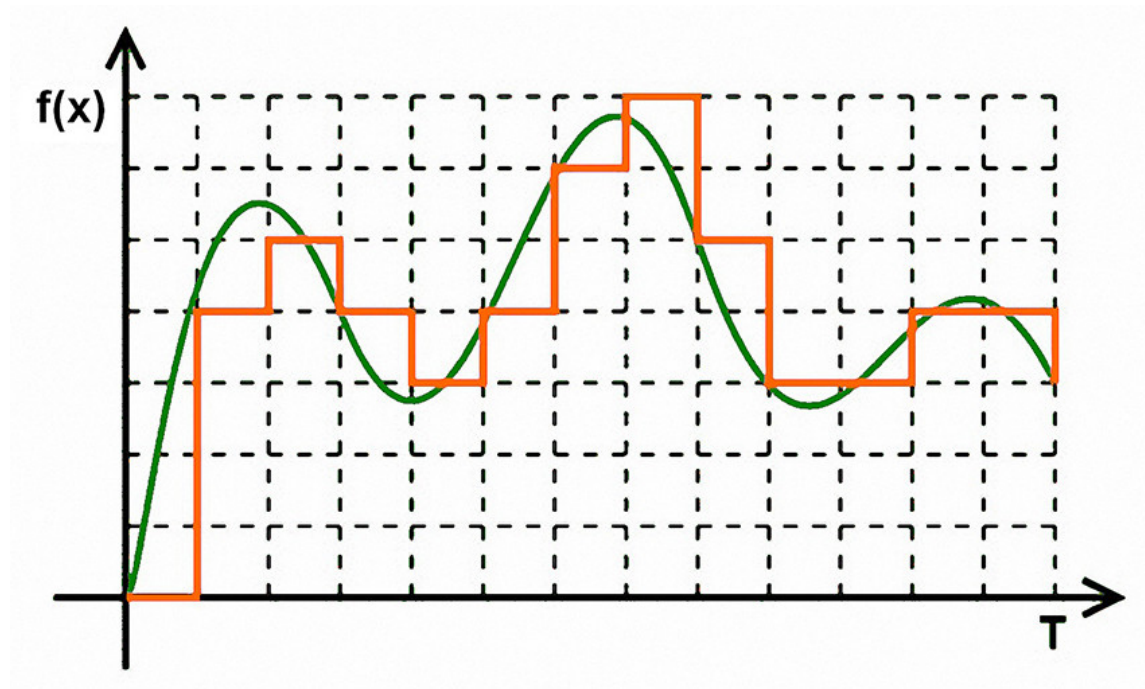


Figure 12: Esempio di discretizzazione dei dati

Il processo di discretizzazione in una simulazione puo' prendere principalmente due macro aree che sono: lo spazio e il tempo. Molti framework di simulazione implementano al loro interno diversi trucchetti per simulare in maniera quanto piu' accurata un insieme di dati continuo, permettendo all'utente di utilizzare parole chiave come ad esempio ContinuousSpace. Quello che praticamente viene fatto e' creare un ambiente per il modello che e' il piu' preciso e granulare possibile, per cui servono allo scopo fintanto che quest'ultimo non richiede una maggiore precisione.

Il processo di discretizzazione comunque non e' sempre negativo, in quanto alcuni problemi possono essere simulati in maniera estremamente fedele anche effettuando questi accorgimenti, e anzi alle volte non e' perfino necessario avere una precisione troppo alta per la simulazione di determinati eventi.

Se ad esempio si volesse simulare tramite un agente una partita a Risiko, non e' necessario richiedere uno spazio e un tempo continuo della simulazione, in quanto questi possono essere rimpiazzati dalla loro controparte discreta che sono caselle e turni.



Figure 13: Plancia di gioco di Risiko

Se si volesse simulare il traffico aereo si potrebbe utilizzare uno spazio a grafo dove ogni nodo rappresenta un aeroporto e gli archi rappresentano la tratta in partenza o in arrivo. Anche in questo caso l'applicazione della discretizzazione non sarebbe qualcosa di problematico.

## 2.3 Epidemiologia

L'epidemiologia è la disciplina biomedica che studia la distribuzione e la frequenza delle malattie ed eventi di rilevanza sanitaria nella popolazione [44]. Si occupa di analizzare le cause, il decorso e le conseguenze delle malattie.

[16] [31]

### 2.3.1 Causalità in Epidemiologia

L'epidemiologia è una disciplina molto pratica, che visto l'obiettivo che si pone, quello di trovare le cause relative ad un dato effetto, non può esentarsi dagli svariati problemi che gravitano e definiscono questo nobile obiettivo, primo tra tutti: cosa vuol dire che un evento è causa di un altro? e come definisco questo tipo di rapporto in maniera inequivocabile?

Questi interrogativi possono sembrare banali in quanto come specie ci siamo evoluti per trovare una correlazione di causalità tra eventi anche quando questi non ne hanno. Ad esempio se fossimo in un bosco, al buio e soli con l'unico rumore ad accompagnarci che è quello di una piacevole brezza estiva, se percepissimo un rumore tra i cespugli, molto probabilmente penseremmo che c'è qualcosa che non va, che ci sia un pericolo in agguato, anche se magari la motivazione è la suddetta brezza.

Questo adattamento evolutivo ci ha permesso di sopravvivere in situazioni di pericolo, ma sfortunatamente quando si parla di scienza e di dati, non sempre l'istinto è qualcosa a cui affidarsi.

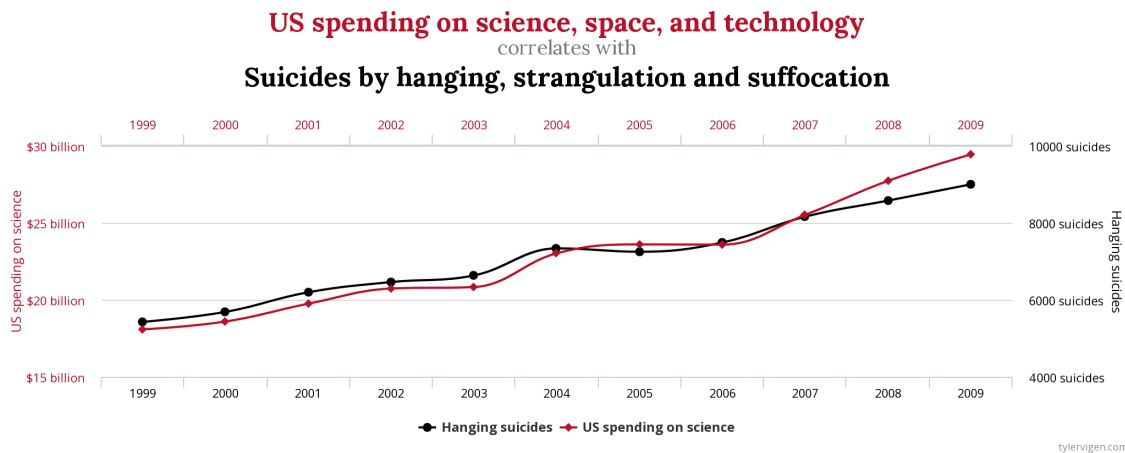


Figure 14: Esempio di correlazione spuria da <https://www.tylervigen.com/spurious-correlations>

Affidandoci solamente al grafico e ai dati riportati in figura, ci verrebbe da pensare che le due categorie siano in qualche modo correlate, e che il governo americano debba essere fermato, in quanto responsabile di incitamento al suicidio. Ebbene non è così, questo è un caso di relazione spuria [52], ovvero che due o più variabili sono associate ma non causalmente correlate.

E' chiaro quindi che non sia così semplice comprendere le cause di un determinato effetto, o insieme di effetti. Alle volte le cause sono completamente contro intuitive e auto alimentanti [50], il che rende

ancora più difficile la loro determinazione. Altre volte invece le cause dipendono dalle interazioni che hanno gli individui tra loro e per cui sono mutevoli in base al comportamento degli individui, rendendo pressochè impossibile determinare con precisione la causa ma al più è possibile restringere la propria ricerca ad un insieme minimale di esse [16].

Il problema della causalità quindi non è da prendere sotto gamba, ed è uno dei problemi cardine da comprendere quando si vogliono determinare e applicare degli interventi all'interno di una popolazione per cercare di mitigare la diffusione di un agente patogeno [31]. Conoscere l'agente patogeno, o quanto meno la sua natura può aiutare, ma non sempre è sufficiente. l'utilizzo di modelli di Machine Learning per l'estrapolazione di dati, di correlazioni e successivamente per la definizione di policy di intervento può risultare in un rischio non indifferente se dovesse succedere di confondere una relazione spuria con una relazione causale.

Questa confusione sembra impossibile, ma avendo sotto mano solamente un grande insieme di dati, e una formula che ne descrive il comportamento, definire con certezza la relazione o la non relazione che esiste tra questi è certamente un compito arduo e delicato. Tuttavia avere un aiutante in grado di analizzare ed estrapolare relazioni da una mole molto grande di dati in poco tempo può essere l'aiuto necessario per comprendere più velocemente e reattivamente le relazioni reali che ci sono tra i dati, arrivando quindi a definire delle policy di intervento che sono, al meglio delle proprie conoscenze attuali, prive di correlazioni spurie.

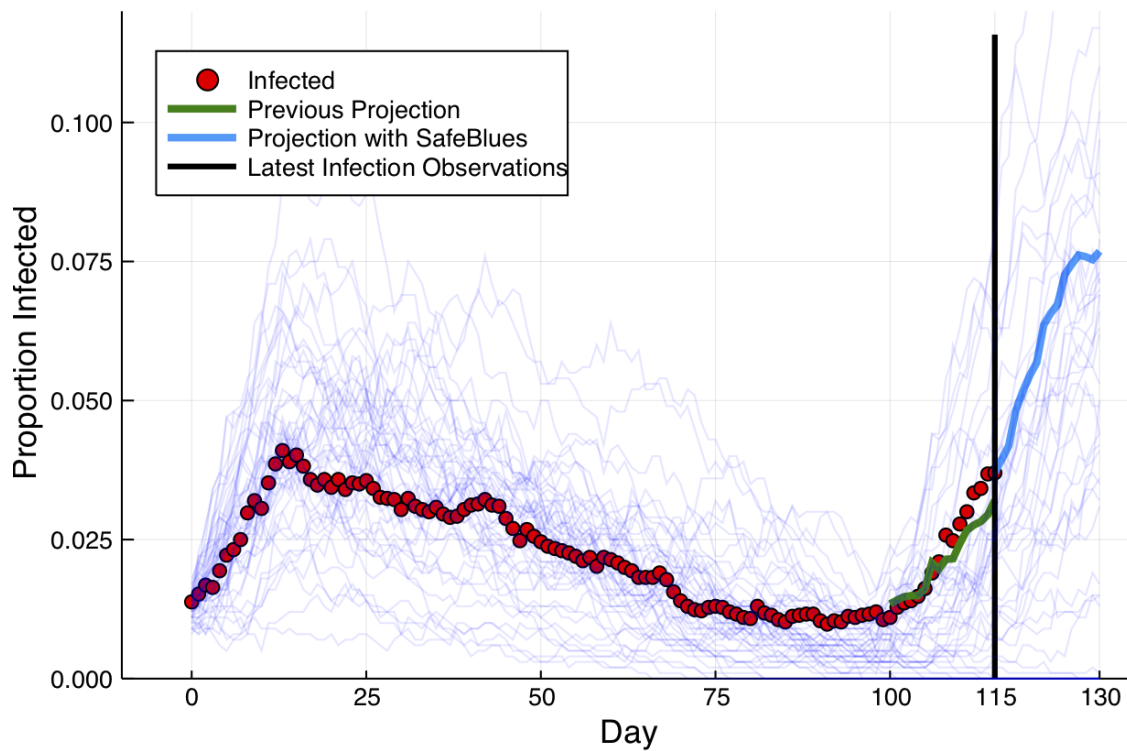


Figure 15: Esempio di proiezione e previsione dati tramite SafeBlues

## 2.4 Julia

Julia è un linguaggio di programmazione ad alto livello, multi-paradigma e open-source ideato per compiere analisi numerica ed effettuare operazioni di computer science in maniera rapida e stabile. Julia è nato ufficialmente come linguaggio di programmazione nell'anno 2012 con lo scopo di fornire uno strumento potente, robusto e veloce tanto se non più dei linguaggi considerati in questo ambito lo stato dell'arte, ovvero C e Fortran; ma anche facile da approcciare, al contrario dei linguaggi sopra citati. Julia è un linguaggio di programmazione scritto in C++ e Scheme, ma gran parte della sua composizione è scritta in Julia stesso [46].



Figure 16: Logo del linguaggio di programmazione Julia

le caratteristiche principali di questo linguaggio sono principalmente:

- **Alte performance:** lo scopo per cui Julia è nato è stato quello di offrire un linguaggio estremamente performante con la capacità di poter compilare programmi in codice nativo per molteplici piattaforme grazie all'utilizzo di LLVM
- **Dinamico:** la scelta di rendere Julia un linguaggio dinamicamente tipizzato lo rende di facile utilizzo in quanto rende molto più semplice il suo approccio anche a chi non ha una base solida di programmazione, in quanto ritorna la stessa sensazione di immediatezza di un linguaggio di scripting. Inoltre questo permette un alto supporto per l'uso interattivo
- **Ambiente riproducibile:** lo scopo del linguaggio è quello di poter permettere all'utente di

ricreare le stesse condizioni ogni volta su ogni macchina su cui un programma viene eseguito. Questo può essere ottenuto tramite l'utilizzo di file binari pre compilati

- Componibile: Julia utilizza l'approccio multiple dispatch come paradigma, permettendo una grande flessibilità nell'esprimere una elevata quantità di pattern di programmazione, dall'object-oriented al funzionale
- General Purpose: lo scopo del linguaggio è quello di creare un ecosistema in grado di poter soddisfare qualsiasi esigenza di un utente, permettendo la creazione di applicativi e microservizi senza dover ricorrere ad integrazioni con codice non nativo Julia
- Open source: Julia abbraccia la filosofia open source, e il codice sorgente dell'intero linguaggio, così come di tutte le librerie è disponibile sulla piattaforma GitHub sotto la licenza MIT. Questo permette una crescita eterogenea grazie al contributo di più di 1000 utenti che si impegnano a migliorare il linguaggio

#### 2.4.1 Agents.jl

Seguendo la filosofia propria del linguaggio di programmazione in cui è sviluppata, la libreria Agents.jl [12] viene sviluppata con l'obiettivo di essere facile da imparare e usare ed estendibile, con forte attenzione sulla creazione ed evoluzione di modelli veloci e soprattutto scalabili. Molteplici esempi comparativi sono stati effettuati mostrando come il framework sviluppato permetta di avere un notevole guadagno prestazionale rispetto ai maggiori competitor attualmente presenti sul mercato (Mesa, Netlogo, MASON) [1].

La facilità di interazione con questa libreria non è da confondersi con una mancanza di opzioni durante lo sviluppo, in quanto nativamente Agents.jl permette l'integrazione con altre librerie che in maniera altrettanto semplice e veloce offrono all'utente la possibilità di addentrarsi nel mondo del machine learning, in particolar modo il mondo del Scientific Machine Learning [35], branca che soprattutto grazie alla pandemia da Covid-19 ha visto un enorme interesse per l'analisi di dati per lo sviluppo di policy di prevenzione e contenimento della pandemia.



Figure 17: Logo framework Agents.jl

### 2.4.2 SciML.ai

SciML è una collezione di librerie dedite al calcolo scientifico e non solo al machine learning. Questo framework permette di avere tutti gli strumenti per poter utilizzare facilmente, velocemente e in maniera robusta tecniche di analisi numerica molto avanzata, così da poter sviluppare applicazioni semplicemente senza essere banali [35] [33] [34].



Figure 18: Logo SciML.ai

Durante la pandemia da Covid-19 questo framework è stato utilizzato per sviluppare applicazioni le quali grazie all'utilizzo di tecniche di scientific machine learning riuscivano sia a prevedere in maniera molto accurata l'andamento dell'epidemia seppur in presenza di una scarsa quantità di dati (uno dei grandi problemi dei modelli di machine learning e di artificial intelligence sono gli enormi dataset necessari per addestrare le reti in maniera robusta) e le stesse presentavano misure di contenimento e prevenzione che si sono dimostrate essere efficaci [5] [11].

### 2.4.3 SafeBlues

Un esempio di un modello di scientific machine learning può essere il modello denominato SafeBlues [5] [11] il quale simulando una rete bluetooth in cui gli individui potevano venire infettati da un virus e poi infettare a loro volta tutti gli individui nella rete con una certa probabilità, aveva riprodotto fedelmente l'andamento della pandemia da Covid-19. In aggiunta questa soluzione, aveva mostrato come l'applicazione di policy per il contenimento del virus bluetooth erano perfettamente applicabili anche al caso reale della pandemia.

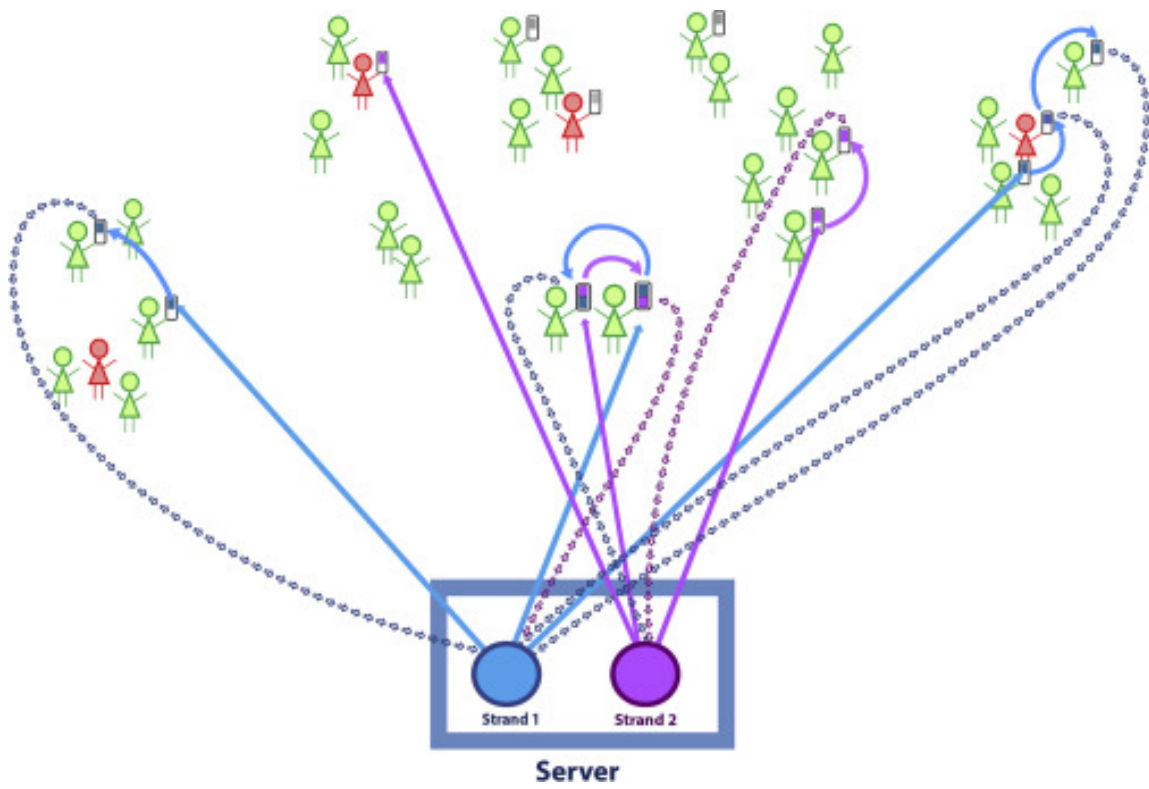
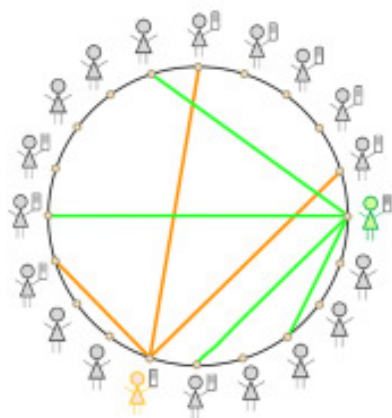
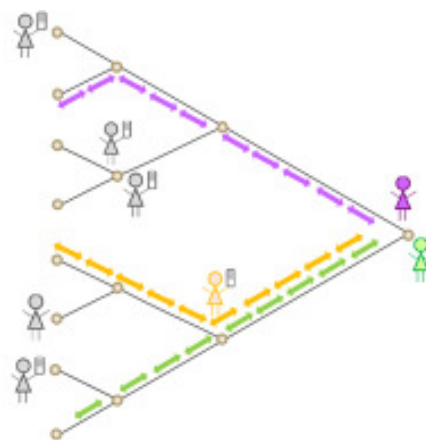


Figure 19: Esempio funzionamento SafeBlues

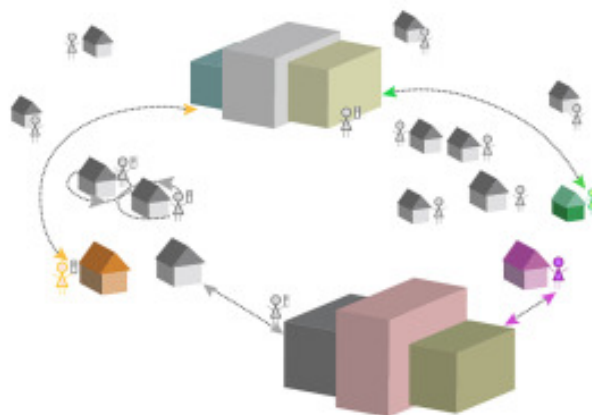




Model I



Model II



Model III

Figure 20: Esempio modelli SafeBlues

## **3 Metodi e Modelli**

### **3.1 COVID-19**

#### **3.1.1 Modelli SEIR**

#### **3.1.2 Modelli ad Agente**

## **3.2 Altre epidemie**

### **3.2.1 Ebola**

### **3.2.2 Aviaria**

### **3.2.3 Influenza**

### **3.3 Monitoraggio e Intervento**

## 4 Risultati Ottenuti

## 5 Conclusioni

## References

- [1] Sameera Abar, Georgios K. Theodoropoulos, Pierre Lemarini r, and Gregory M.P. O’Hare. Agent based modelling and simulation tools: A review of the state-of-art software. *Computer Science Review*, 24:13–33, 2017.
- [2] Linda J. S. Allen. *An Introduction to Stochastic Epidemic Models*, pages 81–130. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.
- [3] Naomi Altman and Martin Krzywinski. Association, correlation and causation. *Nature Methods*, 12(10):899–900, Oct 2015.
- [4] Justin Angevaare, Zeny Feng, and Rob Deardon. Pathogen.jl: Infectious disease transmission network modeling with julia. *Journal of Statistical Software*, 104(4):1–30, 2022.
- [5] Azam Asanjarani, Aminath Shausan, Keng Chew, Thomas Graham, Shane G. Henderson, Hermanus M. Jansen, Kirsty R. Short, Peter G. Taylor, Aapeli Vuorinen, Yuvraj Yadav, Ilze Ziedins, and Yoni Nazarathy. Emulation of epidemics via bluetooth-based virtual safe virus spread: Experimental setup, software, and data. *PLOS Digital Health*, 1(12):1–23, 12 2022.
- [6] Keith R. Bissett, Jose Cadena, Maleq Khan, and Chris J. Kuhlman. Agent-based computational epidemiological modeling. *Journal of the Indian Institute of Science*, 101(3):303–327, Jul 2021.
- [7] Ottar N. Bj rnstad, Katriona Shea, Martin Krzywinski, and Naomi Altman. The seirs model for infectious disease dynamics. *Nature Methods*, 17(6):557–558, Jun 2020.
- [8] Matthew H Bonds, Donald C Keenan, Pejman Rohani, and Jeffrey D Sachs. Poverty trap formed by the ecology of infectious diseases. *Proc Biol Sci*, 277(1685):1185–1192, December 2009.
- [9] Fred Brauer. *Compartmental Models in Epidemiology*, pages 19–79. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.
- [10] P. Ciunkiewicz, W. Brooke, M. Rogers, and S. Yanushkevich. Agent-based epidemiological modeling of covid-19 in localized environments. *Computers in Biology and Medicine*, 144:105396, 2022.
- [11] Raj Dandekar, Shane G. Henderson, Hermanus M. Jansen, Joshua McDonald, Sarat Moka, Yoni Nazarathy, Christopher Rackauckas, Peter G. Taylor, and Aapeli Vuorinen. Safe blues: The case for virtual safe virus spread in the long-term fight against epidemics. *Patterns*, 2(3):100220, 2021.
- [12] George Datseris, Ali R. Vahdati, and Timothy C. DuBois. Agents.jl: a performant and feature-full agent-based modeling software of minimal code complexity. *SIMULATION*, 0(0):003754972110688, January 2022.
- [13] Bernhard Ebbinghaus, Lukas Lehner, and Elias Naumann. Welfare state support during the covid-19 pandemic: Change and continuity in public attitudes towards social policies in germany. *European Policy Analysis*, 8(3):297–311, 2022.

- [14] Abdulrahman M El-Sayed, Peter Scarborough, Lars Seemann, and Sandro Galea. Social network analysis and agent-based modeling in social epidemiology. *Epidemiol Perspect Innov*, 9(1):1, February 2012.
- [15] Mathilde Frérot, Annick Lefebvre, Simon Aho, Patrick Callier, Karine Astruc, and Ludwig Serge Aho Glélé. What is epidemiology? changing definitions of epidemiology 1978-2017. *PLOS ONE*, 13(12):1–27, 12 2018.
- [16] Sandro Galea, Matthew Riddle, and George A Kaplan. Causal thinking and complex system approaches in epidemiology. *Int J Epidemiol*, 39(1):97–106, October 2009.
- [17] Giulia Giordano, Franco Blanchini, Raffaele Bruno, Patrizio Colaneri, Alessandro Di Filippo, Angela Di Matteo, and Marta Colaneri. Modelling the covid-19 epidemic and implementation of population-wide interventions in italy. *Nature Medicine*, 26(6):855–860, Jun 2020.
- [18] Alberto Godio, Francesca Pace, and Andrea Vergnano. Seir modeling of the italian epidemic of sars-cov-2 using computational swarm intelligence. *International Journal of Environmental Research and Public Health*, 17(10), 2020.
- [19] Elizabeth R. Groff, Shane D. Johnson, and Amy Thornton. State of the art in agent-based modeling of urban crime: An overview. *Journal of Quantitative Criminology*, 35(1):155–193, Mar 2019.
- [20] World Bank Group. Wdr 2022 chapter 1. introduction, Mar 2023.
- [21] Masoud Jalayer, Carlotta Orsenigo, and Carlo Vercellis. Cov-abm: A stochastic discrete-event agent-based framework to simulate spatiotemporal dynamics of covid-19, 2020.
- [22] Cliff C. Kerr, Robyn M. Stuart, Dina Mistry, Romesh G. Abeysuriya, Katherine Rosenfeld, Gregory R. Hart, Rafael C. Núñez, Jamie A. Cohen, Prashanth Selvaraj, Brittany Hagedorn, Lauren George, Michał Jastrzębski, Amanda S. Izzo, Greer Fowler, Anna Palmer, Dominic Delport, Nick Scott, Sherrie L. Kelly, Caroline S. Bennette, Bradley G. Wagner, Stewart T. Chang, Assaf P. Oron, Edward A. Wenger, Jasmina Panovska-Griffiths, Michael Famulare, and Daniel J. Klein. Covasim: An agent-based model of covid-19 dynamics and interventions. *PLOS Computational Biology*, 17(7):1–32, 07 2021.
- [23] Garyfallos Konstantinoudis, Dominic Schuhmacher, Håvard Rue, and Ben D Spycher. Discrete versus continuous domain models for disease mapping. *Spatial and Spatio-temporal Epidemiology*, 32:100319, 2020.
- [24] James Ladyman, James Lambert, and Karoline Wiesner. What is a complex system? *European Journal for Philosophy of Science*, 3(1):33–67, Jan 2013.
- [25] Sadegh Marzban, Renji Han, Nóra Juhász, and Gergely Röst. A hybrid PDE-ABM model for viral dynamics with application to SARS-CoV-2 and influenza. *R Soc Open Sci*, 8(11):210787, November 2021.
- [26] Samuel Mwalili, Mark Kimathi, Viona Ojiambo, Duncan Gathungu, and Rachel Mbogo. Seir model for covid-19 dynamics incorporating the environment and social distancing. *BMC Research Notes*, 13(1):352, Jul 2020.



- [27] John T Nardini, Ruth E Baker, Matthew J Simpson, and Kevin B Flores. Learning differential equation models from stochastic agent-based model simulations. *J R Soc Interface*, 18(176):20200987, March 2021.
- [28] Hiroshi Nishiura. Correcting the actual reproduction number: a simple method to estimate  $r(0)$  from early epidemic growth data. *Int J Environ Res Public Health*, 7(1):291–302, January 2010.
- [29] Aleksandar Novakovic and Adele H. Marshall. The cp-abm approach for modelling covid-19 infection dynamics and quantifying the effects of non-pharmaceutical interventions. *Pattern Recognition*, 130:108790, 2022.
- [30] World Health Organization. Who coronavirus (covid-19) dashboard.
- [31] M Parascandola and D L Weed. Causation in epidemiology. *J Epidemiol Community Health*, 55(12):905–912, December 2001.
- [32] Sunny Prakash Prajapati, Rahul Bhaumik, and Tarun Kumar. An intelligent abm-based framework for developing pandemic-resilient urban spaces in post-covid smart cities. *Procedia Computer Science*, 218:2299–2308, 2023. International Conference on Machine Learning and Data Engineering.
- [33] Chris Rackauckas, Mike Innes, Yingbo Ma, Jesse Bettencourt, Lyndon White, and Vaibhav Dixit. Diffeqflux.jl—a julia library for neural differential equations. *arXiv preprint arXiv:1902.02376*, 2019.
- [34] Christopher Rackauckas, Yingbo Ma, Julius Martensen, Collin Warner, Kirill Zubov, Rohit Supekar, Dominic Skinner, and Ali Ramadhan. Universal differential equations for scientific machine learning. *arXiv preprint arXiv:2001.04385*, 2020.
- [35] Christopher Rackauckas and Qing Nie. Differentialequations.jl—a performant and feature-rich ecosystem for solving differential equations in julia. *Journal of Open Research Software*, 5(1):15, 2017.
- [36] Constantinos I. Siettos and Lucia Russo. Mathematical modeling of infectious disease dynamics. *Virulence*, 4(4):295–306, 2013. PMID: 23552814.
- [37] Eric Silverman, Umberto Gostoli, Stefano Picascia, Jonatan Almagor, Mark McCann, Richard Shaw, and Claudio Angione. Situating agent-based modelling in population health research. *Emerging Themes in Epidemiology*, 18(1):10, Jul 2021.
- [38] Melissa Tracy, Magdalena Cerdá, and Katherine M Keyes. Agent-Based modeling in public health: Current applications and future directions. *Annu Rev Public Health*, 39:77–94, January 2018.
- [39] Christopher W. Weimer, J. O. Miller, and Raymond R. Hill. Agent-based modeling: An introduction and primer. In *2016 Winter Simulation Conference (WSC)*, pages 65–79, 2016.
- [40] Wikipedia. Agent-based model — Wikipedia, the free encyclopedia. <http://en.wikipedia.org/w/index.php?title=Agent-based%20model&oldid=1148935023>, 2023. [Online; accessed 30-April-2023].

- [41] Wikipedia. Causality — Wikipedia, the free encyclopedia. <http://en.wikipedia.org/w/index.php?title=Causality&oldid=1150225747>, 2023. [Online; accessed 27-April-2023].
- [42] Wikipedia. Compartmental models in epidemiology — Wikipedia, the free encyclopedia. <http://en.wikipedia.org/w/index.php?title=Compartmental%20models%20in%20epidemiology&oldid=1150020798>, 2023. [Online; accessed 27-April-2023].
- [43] Wikipedia. Cycle of poverty — Wikipedia, the free encyclopedia. <http://en.wikipedia.org/w/index.php?title=Cycle%20of%20poverty&oldid=1151766525>, 2023. [Online; accessed 27-April-2023].
- [44] Wikipedia. Epidemiologia — Wikipedia, the free encyclopedia. <http://it.wikipedia.org/w/index.php?title=Epidemiologia&oldid=124857046>, 2023. [Online; accessed 27-April-2023].
- [45] Wikipedia. Influenza spagnola — Wikipedia, the free encyclopedia. <http://it.wikipedia.org/w/index.php?title=Influenza%20spagnola&oldid=132751127>, 2023. [Online; accessed 27-April-2023].
- [46] Wikipedia. Julia (programming language) — Wikipedia, the free encyclopedia. [http://en.wikipedia.org/w/index.php?title=Julia%20\(programming%20language\)&oldid=1151847776](http://en.wikipedia.org/w/index.php?title=Julia%20(programming%20language)&oldid=1151847776), 2023. [Online; accessed 27-April-2023].
- [47] Wikipedia. Pandemia — Wikipedia, the free encyclopedia. <http://it.wikipedia.org/w/index.php?title=Pandemia&oldid=132968684>, 2023. [Online; accessed 27-April-2023].
- [48] Wikipedia. Pandemia di COVID-19 — Wikipedia, the free encyclopedia. <http://it.wikipedia.org/w/index.php?title=Pandemia%20di%20COVID-19&oldid=133076664>, 2023. [Online; accessed 27-April-2023].
- [49] Wikipedia. Peste nera — Wikipedia, the free encyclopedia. <http://it.wikipedia.org/w/index.php?title=Peste%20nera&oldid=132822978>, 2023. [Online; accessed 27-April-2023].
- [50] Wikipedia. Positive feedback — Wikipedia, the free encyclopedia. <http://en.wikipedia.org/w/index.php?title=Positive%20feedback&oldid=1147282594>, 2023. [Online; accessed 27-April-2023].
- [51] Wikipedia. Simulation software — Wikipedia, the free encyclopedia. <http://en.wikipedia.org/w/index.php?title=Simulation%20software&oldid=1149041989>, 2023. [Online; accessed 27-April-2023].
- [52] Wikipedia. Spurious relationship — Wikipedia, the free encyclopedia. <http://en.wikipedia.org/w/index.php?title=Spurious%20relationship&oldid=1139468569>, 2023. [Online; accessed 30-April-2023].
- [53] Wikipedia. Storia della pandemia di AIDS — Wikipedia, the free encyclopedia. <http://it.wikipedia.org/w/index.php?title=Storia%20della%20pandemia%20di%20AIDS&oldid=132423486>, 2023. [Online; accessed 27-April-2023].
- [54] Wikipedia. Tifo esantematico — Wikipedia, the free encyclopedia. <http://it.wikipedia.org/w/index.php?title=Tifo%20esantematico&oldid=132730114>, 2023. [Online; accessed 27-April-2023].

- [55] Hualei Xin, Yu Li, Peng Wu, Zhili Li, Eric H Y Lau, Ying Qin, Liping Wang, Benjamin J Cowling, Tim K Tsang, and Zhongjie Li. Estimating the Latent Period of Coronavirus Disease 2019 (COVID-19). *Clinical Infectious Diseases*, 74(9):1678–1681, 09 2021.
- [56] J H Zaccai. How to assess epidemiological studies. *Postgrad Med J*, 80(941):140–147, March 2004.