An Agent-Based Model of COVID-19 Diffusion to Plan and Evaluate Intervention Policies

Gianpiero Pescarmona, Pietro Terna, Alberto Acquadro, Paolo Pescarmona, Giuseppe Russo, Emilio Sulis, and Stefano Terna

Abstract A model of interacting agents, following plausible behavioral rules into a world where the Covid-19 epidemic is affecting the actions of everyone. The model works with (i) infected agents categorized as symptomatic or asymptomatic and (ii) the places of contagion specified in a detailed way. The infection transmission is related to three factors: the characteristics of both the infected person and the susceptible one, plus those of the space in which contact occurs. The model includes the structural data of Piedmont, an Italian region, but we can easily calibrate it for other areas. The micro-based structure of the model allows factual, counterfactual, and conditional simulations to investigate both the spontaneous or controlled development of the epidemic.

The model is generative of complex epidemic dynamics emerging from the consequences of agents' actions and interactions, with high variability in outcomes and stunning realistic reproduction of the successive contagion waves in the reference region. There is also an inverse generative side of the model, coming from the idea

G. Pescarmona
University of Torino, Italy, e-mail: gianpiero.pescarmona@unito.it

P. Terna
University of Torino, Italy; Fondazione Collegio Carlo Alberto, Italy e-mail: pietro.terna@unito.it

A. Acquadro
University of Torino, Italy e-mail: alberto.acquadro@unito.it

P. Pescarmona
University of Groningen, The Netherlands e-mail: p.p.pescarmona@rug.nl

G. Russo
Centro Einaudi, Torino, Italy e-mail: russo@centroeinaudi.it

E. Sulis
University of Torino, Italy; e-mail: emilio.sulis@unito.it

S. Terna
tomorrowdata.io e-mail: stefano.terna@tomorrowdata.io

of using genetic algorithms to construct a meta-agent to optimize the vaccine distribution. This agent takes into account groups' characteristics—by age, fragility, work conditions—to minimize the number of symptomatic people.

1 A quick introduction to our agent-based epidemic model

The starting point is a compartmental model with Susceptible, Infected, and Recovered people (S.I.R.), but adding both a more detailed breakdown of the subjects involved in the contagion process [1] and a multi-scale framework to account for the interaction at different dimensional, and spatial levels [2]. From the virus micro-level, we move to individuals and up to the collective behavior of the population.

Following [3], we know that the analysis based on the assumption of heterogeneity strongly differs from S.I.R. compartmental structures modeled by differential equations. That work authors ponder when it is better to use agent-based models and when it would be better to use differential equation models. Differential equation models assume homogeneity and perfect mixing of characteristics within compartments, while agent-based models can capture heterogeneity in agent attributes and the structure of their interactions. We follow the second approach.

- Our model takes into consideration:

  i infected agents categorized as symptomatic or asymptomatic and
  ii the places of contagion specified in a detailed way, thanks to agent-based modeling capabilities.

- The infection transmission is related to three factors: the infected person's characteristics and those of the susceptible one, plus those of the space in which a contact occurs.

Finally, we subscribe the call of [4] to «cover the full behavioural and social complexity of societies under pandemic crisis» and we work arguing that «the study of collective behavior must rise to a "crisis disciplin" just as medicine, conservation, and climate science have, with a focus on providing actionable insight to policymakers and regulators for the stewardship of social systems», as in [5].

Now a look at the structure of the whole presentation.

In Section 1.1, we discuss models and specifically agent-based models; in Section 1.2, the biochemical support to agents' intrinsic susceptibility construction; in Section 1.3, the structure of the model, with the daily sequence of the agents' actions. Section 2 introduces a detailed description of the internal model mechanisms, with: conditional actions in Section 2.1, parameters in Section 2.2 and agents' interaction in Section 2.3.

A technique for contagion representation is introduced in Section 3. Then we explore simulation cases in Section 4, building several batches of runs and comparing extreme situations in Sections 4.1.

Section 4.2 reports the actual epidemic data in the reference region. With those data, we verify factual and counterfactual analyses in Section 5. Considering the possibility of calculating infection indicators without delays (Section 5.4), we experiment with the effect of adopting the control measure with 20 days of anticipation (Section 5.5). In Section 5.6 we verify another counterfactual policy, that of concentrating the efforts uniquely in defense of fragile persons. Section recap 6.1 summarizes these results.

The final application of the model is dedicated to a planning exercise on vaccination campaigns (Section 7). We introduce an analysis of the vaccine mechanism in the perspective of our model (Section 7.1), using both planned strategies (Sections 7.4, 7.5) and genetic algorithms (Section 7.6). The GAs goal is to optimize the behavior of a meta-agent, deciding the sequence of the vaccinations.

1.1 Why model? Why agents? Why another model?

Why another model, and most of all, why model? With [6]:

> The choice (. . .) is not whether to build models; it's whether to build explicit ones. In explicit models, assumptions are laid out in detail, so we can study exactly what they entail. On these assumptions, this sort of thing happens. When you alter the assumptions that is what happens. By writing explicit models, you let others replicate your results.

With even more strength:

> I am always amused when these same people challenge me with the question,"Can you validate your model" The appropriate retort, of course, is,"Can you validate yours?" At least I can write mine down so that it can, in principle, be calibrated to data, if that is what you mean by "validate" a term I assiduously avoid.

To reply to "why agents?", with [7] we define in short what an agent-based model is:

> An agent-based model consists of individual agents, commonly implemented in software as objects. Agent objects have states and rules of behavior. Running such a model simply amounts to instantiating an agent population, letting the agents interact, and monitoring what happens. That is, executing the model—spinning it forward in time—is all that is necessary in order to "solve" it.

More in detail:

> There are, ostensibly, several advantages of agent-based computational modeling over conventional mathematical theorizing. First, [. . .] it is easy to limit agent rationality in agent-based computational models. Second, even if one wishes to use completely rational agents, it is a trivial matter to make agents heterogeneous in agent-based models. One simply instantiates a population having some distribution of initial states, e.g., preferences. That is, there is no need to appeal to representative agents. [. . .] Finally, in most social processes either physical space or social networks matter. These are difficult to account for mathematically except in highly stylized ways. However, in agent-based models it is usually quite easy to have the agent interactions mediated by space or networks or both.

And now, "why another?" As a commitment to our creativity, using our knowledge to understand what is happening. Indeed, with arbitrariness: it is up to others and time to judge.

As any model, also this one is based on assumptions: time will tell whether these were reasonable hypotheses. Modeling the Covid-19 pandemic requires a scenario and the actors. As in a theater play, the author defines the roles of the actors and the environment. The characters are not real, they are prebuilt by the author, and they act according to their peculiar constraints. If the play is successful, it will run for a long time, even centuries. If not, we will rapidly forget it. Shakespeare's Hamlet is still playing after centuries, even if the characters and the plot are entirely imaginary. The same holds for our simulations: we are the authors, we arbitrarily define the characters, we force them to act again and again in different scenarios. However, in our model, the micro-micro assumptions are not arbitrary but based on scientific hypotheses at the molecular level, the micro agents' behaviors are modeled in an explicit and realistic way. In both plays and simulations, we compress the time: a whole life to two or three hours on the stage. In a few seconds, we run the Covid-19 pandemic spread in a given regional area.

1.2 Here a Gianpiero's section on the biochemical of the contagion

1.3 Our model

With our model, we move from a macro compartmental vision to a meso and micro-analysis capability. Its main characteristics are:

- scalability: we take into account the interactions between virus and molecules inside the host, determining individual susceptibility; the interactions between individuals in more or less restricted contexts; the movement between different environments (home, school, workplace, open spaces, shops); the movements occur in different parts of the daily life, as in [8];
  in detail, the scales are:
  - *micro*, with the internal biochemical mechanism involved in reacting to the virus, as in [9], from where we derive the critical importance assigned to an individual attribute of intrinsic susceptibility related to the age and previous morbidity episodes; the model indeed incorporates the medical insights and consistent perspectives of one of its co-authors, former full professor of clinical biochemistry, signing also the quoted article; a comment on Lancet [10] consistently signals the syndemic character of the current event: «Two categories of disease are interacting within specific populations—infection with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and an array of non-communicable diseases (NCDs)»;
  - *meso*, with the open and closed contexts where the agents behave, as reported above;

  - *macro*, with the emergent effects of the actions of the agents;

- granularity: at any level, the interactions are partially random and therefore the final results will always reflect the sum of the randomness at the different levels; changing the constraints at different levels and running multiple simulations should allow the identification of the most critical points, where to focus the intervention.

Summing up, S.I.s.a.R. (https://terna.to.it/simul/SIsaR.html) is an agent-based model designed to reproduce the diffusion of the COVID-19 using agent-based modeling in NetLogo [11]. We have Susceptible, Infected, symptomatic, asymptomatic, and Recovered people: hence the name S.I.s.a.R. The model works on the structural data of Piedmont, an Italian region, but we can quite easily calibrate it for other areas. It reproduces the events following a realistic calendar (national or local government decisions, as in Section 2.2), via its script interpreter. At the above address, it is also possible to run the code without installation. Into the *Info* sheet of the model, we have more than 20 pages of Supporting Information about both the structure and the calibration of the model.

The micro-based structure of the model allows factual, counterfactual, and conditional simulations. Examples of counterfactual situations are those considering:

i different timing in the adoption of the non-pharmaceutical containment measures;
ii an alternative strategy, focusing exclusively on the defense of fragile people.

The model generates complex epidemic dynamics, emerging from the consequences of agents' actions and interactions, with high variability in outcomes, and with a stunning realistic reproduction of the contagion waves that occurred in the reference region.

We take charge of the variability of the epidemic paths within the simulation, running batches of executions with 10,000 occurrences for each experiment.

Following [12], the AI and inverse generative side of the model comes from constructing a meta-agent optimizing the vaccine distribution among people groups—characterized by age, fragility, work conditions—to minimize the number of symptomatic people (as deceased persons come from there).

We can characterize the action of the planner both:

i introducing ex-ante rules following "plain" or "wise" strategies that we imagine as observers or
ii evolving those strategies via the application of a genetic algorithm, where the genome is a matrix of vaccination quotas by people groups, with their time range of adoption.

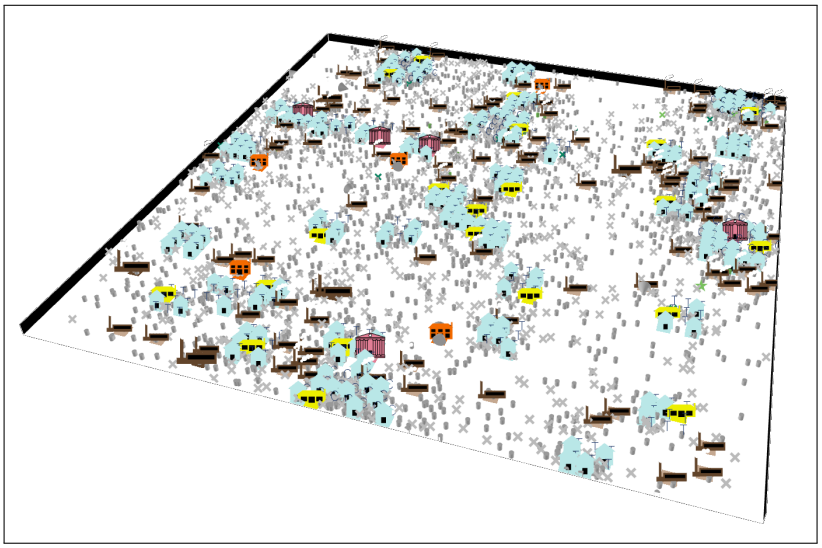


Fig. 1: A live 3D picture of the model world

2 How S.I.s.a.R. works

We have two initial infected individuals in a population of 4350 individuals, on a scale of 1:1000 with Piedmont. The size of the initial infected group is out of scale: it is the smallest number ensuring the epidemic's activation in a substantial number of cases. Initial infected people bypass the incubation period. For plausibility reasons, we never choose initial infected people among persons in nursing homes or hospitals. The presence of agents in close spaces—such as classrooms, factories, homes, hospitals, nursing homes—is set with realistic numbers, out of scale: e.g., a classroom contains 25 students, a home two persons, large factories up to 150 employees, small ones up to 15, etc.; the movements occur in different parts of the daily life, as in [8].

In Fig. 1 we have a 3D representation of the model world, with one of the possible random maps that the simulation generates. Persons are in gray, houses in cyan, nursing homes in orange, hospitals in pink, schools in yellow, factories (with shops and offices) in brown. Persons have a cylinder as shape, if regular or robust (young); a capital X if fragile; temporary their colors can be: red, if symptomatic; violet, if asymptomatic; turquoise, if symptomatic recovered; green, if asymptomatic recovered.

Doing the batches of 10,000 repetitions of the simulation, we use random maps to have a neutral effect of the structure of the space.

We can set:

- min and max duration of the individual infection;
- the length of the incubation interval;

- the critical distance, i.e., the radius of the possibility of infection in open air, with a given probability;
- the corrections of that probability, due to the personal characteristics of both active and the passive agents;
  - active agents can be symptomatic or asymptomatic, with different spreading characteristics (see ii in Section 2.2);
  - passive agents, as receivers, can be robust (young), regular, fragile, and extra fragile.

We have two main types of contagion: (a) within a radius, for people moving around, temporary in a house/factory/nursing home/hospital; (b) in a given space (room or apartment) for people resident in their home or in a hospital or in a nursing home or being in school or in a working environment.

People in hospitals and nursing homes can be infected in ways (a) and (b). Instead, while people are at school, they can only receive the disease from people in the same classroom, where only teachers and students are present, so this is a third infection mechanism (c). In all cases, the personal characteristics of the recipients are decisive.

We remark that workplaces are open to all persons, as clients, vendors, suppliers, external workers can go there. In contrast, schools are reserved for students and school operators.

All agents have their home, inside a city, or a town. The agents also have usual places (UPs) where they act and interact, moving around. These positions can be interpreted as free time elective places. When we activate the schools, students and teachers have both UPs and schools; healthcare operators have both UPs and hospitals or nursing homes; finally, workers have both UPs and working places. In each day (or tick of the model), we simulated full sequences of actions.

Fig. 2 describes what happens during every *day* in our simulated world, with the daily sequences of actions.

2.1 Conditional actions

Agents' movements in space, to go to work, school, and other UPs are subject to two interrelated general conditions.

I Symptomatic persons are at home or in a hospital or a nursing home and do not move.

II People not constrained by *condition I* can move if (primary rule) there are no general limitations (e.g., lockdown) *OR* if one of the following sub-conditions applies:

  a. agents who are hospital healthcare operators or nursing home healthcare operators;
  b. all people, according to the probability of moving of the whole non-symptomatic agents (2.2, iv);

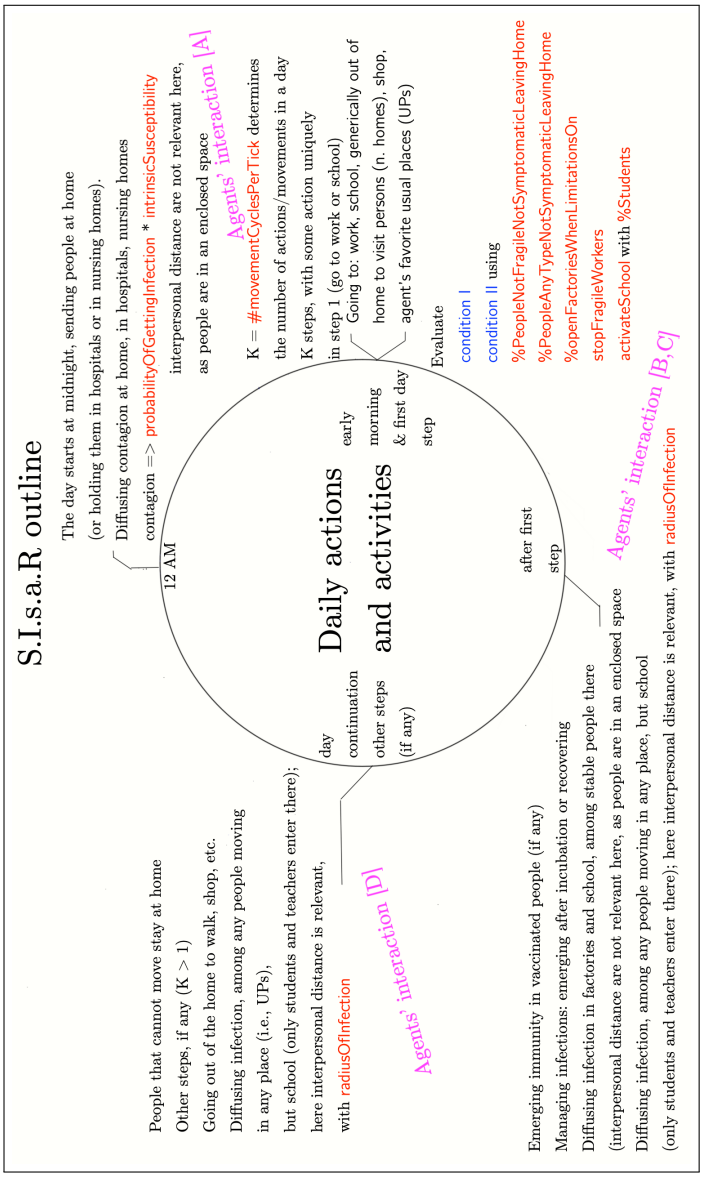


Fig. 2: A day in the simulation, with N repetitions where N is the duration of a given outbreak; look at: Section 2.1 for the rules of the conditional actions; Section 2.2 for the parameter definitions; and Section 2.3 for details on the agent interactions

c. regular people, according to the probability of moving of the regular non-symptomatic agents (2.2, v);

d. workers, if all the factories are open or it is open their own workplace (2.2, vi);
e. teachers, if the schools are open (2.2, vii);
f. students, if the schools are open, but with a possible quota limitation (2.2, viii).

2.2 Parameter definition

We define the parameters of Fig. 2, also with their short names used in program scripts, in round brackets. The values of the parameters are reported in detail in Appendix 1 - Parameter values (Section 9).

i *probabilityOfGettingInfection* (`prob`) is the base probability of getting infected, to be multiplied by the *intrinsicSusceptibility* factor (iii); it is activated if the subject is within a circle of radius (ix) with an infected person; values at (9, i);

ii *D%*, without the short name, is the percent increasing or decreasing factor of the contagion spread of an asymptomatic subject, compared to that of a symptomatic one, value at (9, ii);

iii the *intrinsicSusceptibility* in defined in Eq. 1

$$intrinsicSusceptibility = intrinsicSusceptibilityFactor^{groupFragility} \quad (1)$$

with *intrinsicSusceptibilityFactor* set to 5 , and $groupFragility$ exponent set to:

1 for extra-fragile persons,
0 for fragile persons,
-1 for regular persons,
-2 young people from 0 to 24 years old;

iv *%PeopleAnyTypeNotSymptomaticLeavingHome* (`%PeopleAny`) determines, in a probabilistic way, the number of people of any kind going around in case of limitations/lockdown; the limitations operate only if the lockdown is on (into our simulated world, from day 20); values at (9, iv);

v *%PeopleNotFragileNotSymptomaticLeavingHome* (`%PeopleNot`) determines, in a probabilistic way, the number of regular people going around in case of limitations/lockdown; as above, the limitations operate only if the lockdown is on (into our simulated world, from day 20); values at (9, v);
we try to reproduce the uncertainty of the decisions in the real world into the model via frequent changes of the parameters iv and v;
NB, the parameters iv and v produce independent effects, as in the following examples: (a) the activation of *%PeopleAny at 31, 0* and, simultaneously, of *%PeopleNot at 31, 80*, means that people had to stay home on that day, but people specifically not fragile could go out in 80% of the cases; (b) *%PeopleAny at 339, 80* and, simultaneously, *%PeopleNot at 339, 100* means that fragile and

not fragile persons cannot always go around, but only in the 80% of the cases; instead, considering uniquely non-fragile persons they are free to go out; the construction is an attempt to reproduce a fuzzy situation;
in future versions of the model, we will define the quotas straightforwardly:

- `%FragilePeopleNotSymptomaticLeavingHome`;
- `%NotFragilePeopleNotSymptomaticLeavingHome`;

vi *%openFactoriesWhenLimitationsOn* (`%Fac`) determines, in a probabilistic way, the factories (small and large industries, commercial surfaces, private and government offices) that are open when limitations are on; if the factory of a worker is open, the subject can go to work, not considering the restrictions (but uniquely in the first step of activity of each day); values at (9, vi);

vii *stopFragileWorkers* (`sFW`) is *off* (set to 0) by default; if *on* (set to 1), fragile workers (i.e., people fragile due to prior illnesses) can move out of their homes following the iv and v parameters, but cannot go to work; in the *off* case, workers (fragile or regular) can go to their factory (if open) also when limitations are on; values at (9, vii);
alternatively, we also have the `fragileWorkersAtHome` parameter; if *on* (set to 1) the total of the workers is unchanged, but the workers are all regular; we can activate this counterfactual operation uniquely at the beginning of the simulation;

viii when *activateSchools* (`aSch`) is *on* (set to 1), teachers and students go to school avoiding restrictions (but uniquely in the first step of activity of each day); *%Students* (`%St`) sets the quota of the students moving to school; the residual part is following the lessons from home; values at (9, viii);

ix following *radiusOfInfection* (`radius`), the effect of the contagion—outside enclosed spaces, or there, but for temporary presences—is possible within that distance; values at (9, ix).

2.3 Agents' interaction

We underline that our simulation tool is not based on micro-simulation sequences, calculating the contagion agent by agent, on the base of their characteristics and ex-ante probabilities. It implements a true agent-based simulation, with the agents acting and, most of all, interacting. The effect is that of generating continuously contagion situations.

Each run creates a population with expected characteristics, but also with random specifications, to assure the heterogeneity in agents. The daily choices of the agents are partially randomized, to reproduce real-life variability.

Contagions arise from agents' interactions, in four time phases, as specified in Fig. 2:

A - in houses (at night), hospitals, nursing homes;
B - in schools and workplaces in general, among people stable there;

C - in the places above (excluding schools) by people temporary there and in open spaces (UPs above);
D - interactions mainly in open spaces (UPs above).

3 Contagion representation

We introduce a tool that allows analyzing the contagions' sequences in simulated epidemics and identifying the places where they occur.

- We represent each infected agent as a horizontal segment (from the starting date to the final date of the infection) with vertical connections to other agents if they receive the disease.
- We represent the new infected agents via further segments at an upper level.
- We display multiple information using three elements.
  - Colors in horizontal segments (areas of the infections): black for unknown places, gray for open spaces, cyan for houses, orange for nursing homes, pink for hospitals, yellow for schools, brown for factories, with shops and offices.
  - Vertical connecting segments keep the same color of the horizontal generating one.
  - Line thickness; proportional to fragility.
  - Styles: dotted lines for incubation, dashed lines for asymptomatic subjects, solid lines of symptomatic ones.
- This graphical presentation enables understanding at a glance how an epidemic episode is developing. In this way, it is easier to reason about countermeasures and, thus, to develop intervention policies.

At https://github.com/terna/contagionSequence we have the calculation tool, *sequentialRecords.ipynb*, generating these sequences.

Fig. 3 is useful as an example. We start with two agents from the outside, with black as the color code (unknown place). The first one is young—as reported by the thickness of the segment, starting at day 0 and finishing at day 22—and asymptomatic (dashed line); it infects no one. The second one—regular, as reported by the thickness of the segment, starting at day 0 and finishing at day 15—is asymptomatic (dashed line) and infects four agents on day 2. All the four infected agents received the infection at work (brown color) and turn to be asymptomatic after the days of incubation (dotted line); the first and the fourth are regular agents; the second and the third are fragile ones.

Continuing the analysis: on day 3, the second agent infects three other agents (at home, at work, at work) [. . .]; on day 13, agent number five infects seven regular agents at work and an extra-fragile one in a nursing home (orange color), etc.

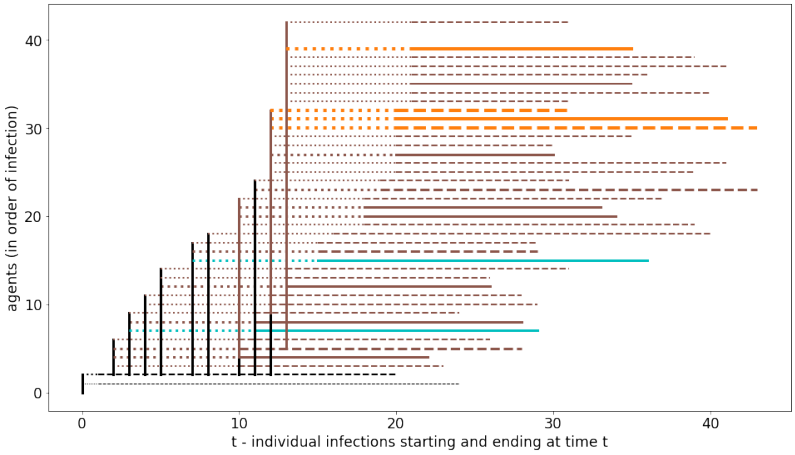


Fig. 3: A case with containment measures, first 40 infections: workplaces (brown) and nursing homes (orange) interweaving

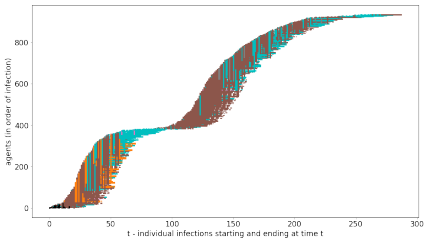


Fig. 4: A Case with containment measures, the whole epidemics: workplaces (brown) and nursing homes (orange) and then houses (cyan), with a bridge connecting two waves

If a vertical segment changes its color, we have an agent in an upper infecting someone on the same day of the infection transmitted by an agent in a lower row, so we lose some information.

In Fig. 4 we see an example of an epidemic where non-pharmaceutical containment measures are in adoption: a first wave shows an interlaced effect of contagions at home, in nursing homes, and at work. After a phase in which contagions develop mainly at home, a skinny bridge connects the first wave to a second one, restarting from workplaces. The thickness of the *snake* of the contagions measures the stock on infects agents on a given date; the slope reports the speediness of the epidemic development; the vertical coordinate reports the cumulative number of infected people.

In Appendix 2 - A gallery of contagion sequences (Section 10) we have several examples of contagion sequences.

4 Exploring scenarios with simulation batches

The sequences described in Section 3 offer mainly suggestions and interpretations. To explore systematically the introduction of factual, counterfactual, and prospective interventions to control the spread of the contagions, we need to analyze batches of simulations. In this perspective, each simulation run—whose length coincides with the disappearance of symptomatic or asymptomatic contagion cases—is a datum in a scenario of time and outcome variability. Consequently, we need to represent compactly the results emerging from batches of repetitions to compare the consequences of each batch's basic assumptions.

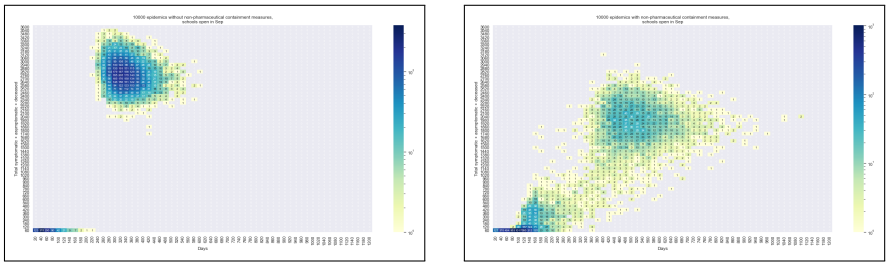
For this purpose, we used blocks of ten thousand repetitions. Besides summarizing the results with the usual statistical indicators, we adopt the technique of the heat-maps. With [13], we develop a search for comparative analyzes, not for forecasting. This consideration is consistent with the enormous standard deviation values that are intrinsic to the specific reality.

At https://github.com/terna/readSIsaR_BatchResults we have the codes producing the maps of the batches. A heat-map is a double histogram: in our application, it displays each simulated epidemic's duration in the x axis and the count of the symptomatic, asymptomatic, and deceased agents in the y axis (on a scale of 1:1000). The cells contain the number of epidemics of each specific duration and outcome, reported via a logarithmic color scale.

4.1 Epidemics without and with control

As a starting point, we compare the situations represented in Figs. 5a and 5b. In Fig. 5a the heat-map reports the distribution in duration and infection causation of 10,000 simulated outbreaks left to spread without any control; coherently, with the school always open. The numbers in Table 1 are scary. The concentration of the cases in the heat-map shows that, except a few instances spontaneously concluding in a short period (left bottom corner), produces a heavy *cloud* of cases lasting one year or one year and a half, hitting (as symptomatic, asymptomatic and deceased) from 2.000 to 3,500 persons on a total of 4,350 in the region (scale of 1:1000).

In Fig. 5b and the related Table 2 we report a similar simulation batch of 10,000 runs of the model, but with the adoption of the basic non-pharmaceutical containment measures, registered in the values of the parameters in Appendix 1 - Parameter values (Section 9). A calendar is at https://terna.to.it/simul/calendario092.pdf, and the model—version 0.9.6— is updated until April 2021. The results are dramatically different, showing the efficacy of the containment measures.



(a) Outbreaks without non-pharmaceutical containment measures

(b) Outbreaks with non-pharmaceutical containment measures

Fig. 5: Starting our analyses: 10,000 epidemics in Piedmont

(000)	symptomatic	totalInfected&Deceased	duration
mean	969.46	2500.45	303.10
std	308.80	802.88	93.50

Table 1: Mean values and standard deviations in Fig. 5a cases

(000)	symptomatic	totalInfected&Deceased	duration
mean	344.22	851.64	277.93
std	368.49	916.41	213.48

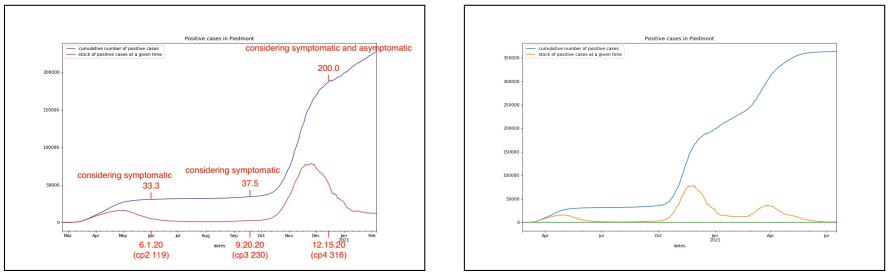
Table 2: Mean values and standard deviations in Fig. 5b cases

4.2 Actual data

The critical points for our experiments with the simulation model of the epidemic in Piedmont are Summer and Fall 2020, as in Fig. 6a, where we have the time series of the first part of Piedmont's actual epidemic. The blue line represents the cumulative number of infected persons. Initially, only symptomatic cases were accounted for, but after the 2020 Summer, with more generalized tests also asymptomatic patients are included.

Explanation:

- from http://www.protezionecivile.it/web/guest/department, the Italian Civil Protection Department web site, we find at https://github.com/pcm-dpc/COVID-1, i.e., the repository of regional data;
- initially, only symptomatic cases were accounted for, but after the 2020 Summer, with more generalized tests, asymptomatic patients are included;
- we had data about symptomatic infected people in the first wave, but from October 2020, data are mixed: in the above *git* repository, in October and November, we had "Positive cases emerged from clinical activity", unfortunately then re-



(a) Critical points in epidemic dynamic in Summer and Fall 2020 in Piedmont

(b) Data in Piedmont until July 2021, reporting three waves

Fig. 6: Actual data

ported as "No longer populated" (from the end of November, our observation) and "Positive cases emerging from surveys and tests, planned at national or regional level", again "No longer populated" (from the end of November, our observation);
- as a consequence, the subdivision between symptomatic and asymptomatic cases is no longer possible.

Considering the dynamic of the data in Fig. 6a, we search within the simulation batch for cases with both:

i numbers of infected persons quite similar at `cp2` and at `cp3`; besides, numbers not too different from those of the figure; (with *cp*, we indicate the internal check points of the simulation program. In Fig. 6a we also report the number of days from the beginning of the epidemic for each check point);
ii the number of infected persons at `cp4` has to be significantly greater than those at the previous check point.

In a lot of cases, epidemics satisfying condition (i) fail to match condition (ii); both the situations happen only in less than the 1.5% of the instances in a batch of ten thousand epidemic. We can conclude that the second wave registered in Piedmont after the Summer stability is due to new infected agents coming from outside and restarting the contagion process.

Other critical points in our analysis are the day on which the vaccination campaign starts, 373 of the simulation (Feb. 12th, 2021), and the day of the effectiveness of the initial vaccinations, 40 days later, day 413 (Mar. 22nd, 2021). At those dates, within the simulations, we can find either the presence of many infected agents or of few ones, as effectively was the situation in Piedmont.

As documentation, in Fig. 6b the time series covering the whole period.

5 Factual and counterfactual analyses

In Fig. 7 we collect the heat-maps of the experiments:

- observing the emergence of spontaneous second waves, in the absence of specific control measures (Sections 5.1);
- causing the emergence of the second wave through infections from outside, again in the absence of specific control measures (Section 5.2);
- causing the emergence of the second wave through infections from outside, in the presence of specific control measures (Section 5.3);
- reproducing the case of Section 5.3, anticipating by twenty days the start and end of all control measures (Section 5.5);
- reproducing the case of Section 5.3, limiting the control measures the fragile workers and fragile people in general (Section 5.6).

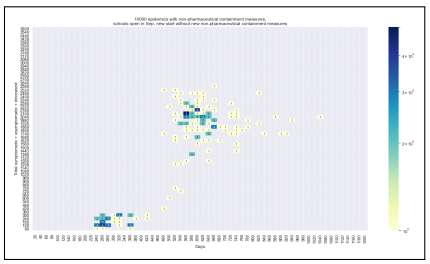
5.1 Spontaneous second wave, without specific containment measures

In an initial plain batch of 10.000 runs of the Piedmont related model, we count only 140 cases of epidemics with both the absence of new contagions in Summer 2020 and their explosion in Fall, as in Fig. 6a.

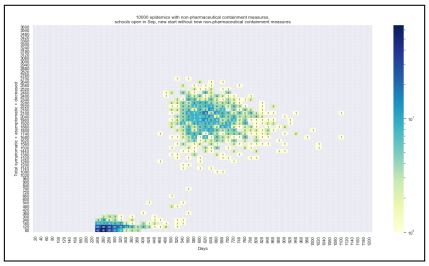
We select, first of all, 170 cases in the following way. Considering June 1st, we choose the epidemics that have a number of symptomatic agents in the (10, 70] interval (their mean: 37.9) and, on September 20th, a number of symptomatic agents in the (20, 90] interval (their mean: 60.4). Due to the lack of data described in Section 4.2, to compare December 15th and September 20th situations, we use symptomatic plus asymptomatic agents' count. We observe the existence of 140 outbreaks with the required characteristics; the December mean of the infected agents is 648.7, sensibly larger than the actual value: ≈200.0. We overestimate the reality being the long-lasting simulated outbreaks, the larger ones, and, most of all, having no containment measures operating in the simulations.

Fig. 7a and Table 3 show the outbreaks with similar cumulative numbers before and after the Summer 2020 pause (170 cases), with the second wave of 140 cases developing in the absence of containment measures.

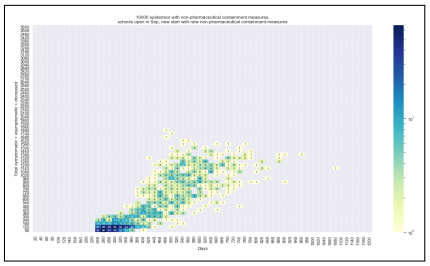
The adverse cases are 140 out of 10,000, a very light ratio as the second wave occurred in the Fall. The transition to the third wave that we see in Fig. 6b is easy to explain, as the second wave never completely ended.



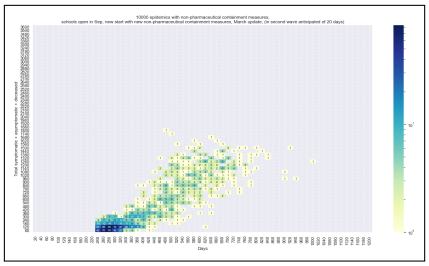
(a) First wave with non-pharmaceutical containment measures, spontaneous second wave, without specific measures



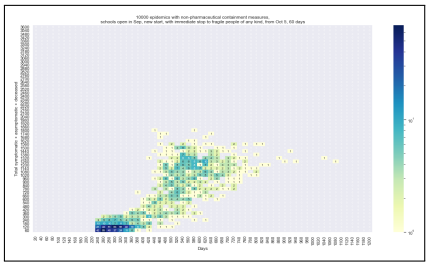
(b) First wave with non-pharmaceutical containment measures, forcing the second wave, without specific measures



(c) First wave with non-pharmaceutical containment measures, forcing the second wave, *with new specific non-ph. containment measures*



(d) First wave with non-pharmaceutical cont. meas,, forcing the second w., *with new specific non-ph. cont. meas., acting 20 days in advance*



(e) First wave with non-ph. cont. meas., forcing the sec. wave; *in sec. wave, uniquely stopping fragile people, including fragile workers*

Fig. 7: Heat-maps of the factual and counterfactual analyses

5.2 Second wave, new infections from outside, without specific containment measures

To generate a framework consistent with the presence of a second wave after a period of substantial inactivity of the epidemic, we introduced the occurrence of two cases of infected persons coming back from outside after Summer vacancies, conventionally on September 1st, 2020.

(1000)	Jun 1, 20		Sep 20, 20		Dec 15, 20		Feb 1, 21		May 1, 21		Dec 15, 20 to end		
cum. v.	sym.	all	sympt.	totalInf.	sympt.	totalInf.	sympt.	totalInf.	sympt.	totalInf.	sympt.	totalInf.	days
count	170.0	170.0	170.0	170.0	140.0	140.0	131.0	131.0	128.0	128.0	140.0	140.0	140.0
mean	37.9	100.2	60.4	159.3	248.4	648.7	432.2	1109.5	656.3	1655.5	701.1	1757.9	594.2
std	16.4	61.0	19.6	71.7	167.4	424.3	220.4	538.4	215.4	513.3	246.4	599.7	118.9

Table 3: Spontaneous second wave, without specific measures

(1000)	Jun 1, 20		Sep 20, 20		Dec 15, 20		Feb 1, 21		May 1, 21		Dec 15, 20 to end		
cum. v.	sym.	all	sympt.	totalInf.	sympt.	totalInf.	sympt.	totalInf.	sympt.	totalInf.	sympt.	totalInf.	days
count	1407.0	1407.0	1407.0	1407.0	1044.0	1044.0	1005.0	1005.0	980.0	980.0	1044.0	1044.0	1044.0
mean	35.6	72.7	40.0	84.1	180.4	462.1	354.1	900.4	623.8	1563.3	726.6	1810.9	620.9
std	14.1	42.6	16.7	52.8	134.6	354.6	213.8	535.4	217.9	527.0	221.9	544.0	110.8

Table 4: Second wave, new infections from outside, without specific measures

Now we have 1407 cases of epidemics with: (i) on June 1st, a number of symptomatic agents in the (10, 70] interval (their mean: 35.6) and (ii), on September 20th, a number of symptomatic agents in the (20, 90] interval (their mean: 40.0). As in Section 5.1, to compare December 15th and September 20th situations we use symptomatic plus asymptomatic agents' count. We observe the existence of 1044 outbreaks with the required characteristics; the December mean of the infected agents is 462.1, again sensibly larger than the actual value: ≈200.0. We overestimate the reality being the simulations run without the adoption of containment measures.

Both Fig. 7b and Table 4 show the outbreaks with similar cumulative numbers before and after the Summer 2020 pause (1407 cases), with the second wave of 1044 cases producing, in the absence of containment measures, a heavy cloud as that of Fig. 5a, with infected people of any kind in a range approximately from 1,500 to 2,800 realizations, with an equivalence, to the Piedmont scale, to 1,5-2,8 millions of subjects.

The number of cases is now sufficient to evaluate the effects of factual (Section 5.3 and counterfactual (Sections 5.5) and 5.6) simulation experiments.

5.3 Second wave, new infections from outside, with new specific containment measures

As in Section 5.2 we have 1407 cases of epidemics with: (i) on June 1st, a number of symptomatic agents in the (10, 70] interval (their mean: 35.6) and (ii), on September 20th, a number of symptomatic agents in the (20, 90] interval (their mean: 40.0). We use symptomatic plus asymptomatic agents' count December 15th and September 20th situations, recording 874 outbreaks with the required characteristics. The December mean of the infected agents is 340.6, closer to the actual value (≈200.0) due to the introduction into the simulation of specific control measures for the second

(1000)	Jun 1, 20	Sep 20, 20		Dec 15, 20		Feb 1, 21		May 1, 21		Dec 15, 20 to end		
cum. v.	sym.	all	sympt.	totalInf.	sympt.	totalInf.	sympt.	totalInf.	sympt.	totalInf.	sympt.	totalInf.	days
count	1407.0	1407.0	1407.0	1407.0	874.0	874.0	719.0	719.0	523.0	523.0	874.0	874.0	874.0
mean	35.6	72.7	40.0	84.1	*130.0*	*340.6*	*194.4*	*512.8*	*295.7*	*791.2*	252.7	666.4	494.1
std	14.1	42.6	16.7	52.8	83.9	232.6	104.1	276.9	119.1	300.6	156.8	416.4	122.7

Table 5: Second wave, new infections from outside, with new specific measures

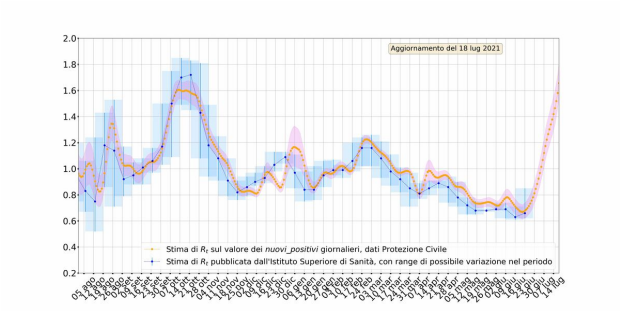


Fig. 8: In blue, the R_t values as reported by the Istituto Superiore di Sanità and in red the anticipated calculation published regularly, from the end of November 2020, at https://mondoeconomico.eu by Stefano Terna

wave. We always overestimate the reality because the surviving epidemics are the larger ones.

In Fig. 7c we see that the heavy cloud of the previous figure dissolved, and in Table 5 the figures in italic emphasize the effect of the containment interventions on the cases of epidemic continuation, in any case, reduced for the same motivation.

5.4 Here a Stefano's section on calculating the infection indicator without delays

See Fig. 8 about . . .

Methodology:

https://github.com/tomorrowdata/COVID-19/blob/main/notebooks/Rt_on_italian_national_data.ipynb

5.5 Second wave, new infections from outside, with new specific containment measures, introduced 20 days in advance

The counterfactual situation is related to the start and end dates of the actions of containment, both occurring 20 days in advance, with a natural barrier set on October

(1000)	Jun 1, 20		Sep 20, 20		Dec 15, 20		Feb 1, 21		May 1, 21		Dec 15, 20 to end		
cum. v.	sym.	all	sympt.	totalInf.	sympt.	totalInf.	sympt.	totalInf.	sympt.	totalInf.	sympt.	totalInf.	days
count	1407.0	1407.0	1407.0	1407.0	769.0	769.0	637.0	637.0	471.0	471.0	769.0	769.0	769.0
mean	35.6	72.7	40.0	84.1	**112.2**	**294.2**	*172.0*	*467.9*	*276.5*	*748.6*	248.9	663.4	499.3
std	14.1	42.6	16.7	52.8	66.8	188.4	91.5	251.3	112.9	286.9	158.0	417.5	124.1

Table 6: Second wave, new infections from outside, with new specific measure anticipation of -20 days

5th, 2020. Before that date, the conditions to start second wave new policies were missing.

As in the last two sections, we have 1407 cases of epidemics alive at the critical dates of June 1st and September 20th, after a Summer interval characterized by a quiet phase. Considering December 15th and September 20th situations, the second wave epidemics are 769, again decreasing because the anticipated actions have eliminated some other cases. The December mean of the infected agents is 294.2, higher than the actual value (≈200.0). We always overestimate the mean of the epidemic effects, being the surviving epidemics the larger ones.

Comparing Fig. 7d and Fig. 7c the difference is not evident; instead, the italic figures, and must of all, the red bold ones, in Table 6 reports clearly the comparative advantage of this counterfactual experiment with respect to the values of Table 5.

5.6 Second wave, new infections from outside, with a unique intervention measure: stopping fragile people for 60 days

This second counterfactual experiment is based on an immediate stop to the circulation of fragile persons and specifically of fragile workers, plus isolating nursing homes and hospitals. Schools are always open in this experiment. The decision is activated on October 5th, 2020, as the second wave is becoming evident. In [14] we have important consideration suggesting the importance of taking into account fragility in a long-term fighting perspective against this kind of epidemics.

As in the last three sections, we have 1407 cases of epidemics alive at the critical dates of June 1st and September 20th, after a Summer interval characterized by a quiet phase. Considering December 15th and September 20th situations, the second wave epidemics are 886, lightly above the values of Section 5.3 and Section 5.5, but without locking the economic and the society as a whole. The December mean of the infected agents is 326.3, higher than the actual value (≈200.0) for the explained overestimation bias.

Comparing Fig. 7e and Fig. 7c the difference is not evident; instead, the italic figures, and must of all, the violet bold ones, in Table 7 reports the similarity of the effects of this counterfactual experiment with respect to the values of Table 5.

(1000)	Jun 1, 20	Sep 20, 20		Dec 15, 20		Feb 1, 21		May 1, 21		Dec 15, 20 to end		
cum. v.	sym.	all	sympt.	totalInf.	sympt.	totalInf.	sympt.	totalInf.	sympt.	totalInf.	sympt.	totalInf.	days
count	1407.0	1407.0	1407.0	1407.0	886.0	886.0	761.0	761.0	637.0	637.0	886.0	886.0	886.0
mean	35.6	72.7	40.0	84.1	**128.1**	**326.3**	*211.0*	*555.1*	*323.3*	*862.1*	301.1	792.3	515.5
std	14.1	42.6	16.7	52.8	89.6	234.2	118.1	306.7	126.4	315.9	170.7	450.2	116.9

Table 7: Second wave, new infections from outs., stop fragile people. 60 days from Oct. 5

6 Cost-benefit analysis of the of interventions [Beppe's contribution]

draft, draft, draft, draft, draft

Considering the different interventions of Section 5.3 and 5.6, we can evaluate . . .

First of all, we can evaluate the costs for the attribution of sick pay to workers who are compulsorily absent from work, if they are not able to carry out remotely their work activities . . .

We have also relevant social benefits, e.g., schooling, and economic benefits, as activities do not stop . . .

An interesting reference is [15], where . . . WORKS IN PROGRESS.

6.1 To recap

Write a nice global comment here

All waves, in second wave and a the end, with day count

Fig. 8

7 Planning vaccination campaigns

7.1 Some notes on vaccines (Alberto's section)

Vaccines are biological products made from killed or attenuated microorganisms, from viruses or from some of their components (antigens), or from substances they produce made safe by chemical (e.g.: formaldehyde) or heat treatment, while maintaining their immunogenic properties (https://www.who.int/vaccines); today, vaccines can be composed of proteins obtained by recombinant DNA techniques using genetic engineering approaches.

They usually contain, in addition to the antigenic fraction, sterile water (or a saline-based physiological solution), adjuvants, preservatives and stabilisers. Adju-

Scenarios		Dec 15, 20		Dec 15, 20 to end		
		sympt.	totalInf.	sympt.	totalInf.	days
no containment	count	140.0	140.0	140.0	140.0	140.0
in spontaneous	mean	248.4	648.7	701.1	1757.9	594.2
second wave	std	167.4	424.3	246.4	599.7	118.9
no containment	count	1044.0	1044.0	1044.0	1044.0	1044.0
in forced	mean	180.4	462.1	726.6	1810.9	620.9
second wave	std	134.6	354.6	221.9	544.0	110.8
basic containment	count	874.0	874.0	874.0	874.0	874.0
in forced	mean	*130.0*	*340.6*	252.7	666.4	494.1
second wave	std	83.9	232.6	156.8	416.4	122.7
-20 days cont.	count	769.0	769.0	769.0	769.0	769.0
in forced	mean	**112.2**	**294.2**	248.9	663.4	499.3
second wave	std	66.8	188.4	158.0	417.5	124.1
frag. subj. &						
workers control	count	886.0	886.0	886.0	886.0	886.0
in forced	mean	**128.1**	**326.3**	301.1	792.3	515.5
second wave	std	89.6	234.2	170.7	450.2	116.9

Table 8: Report of the key results, with count, mean, and std

vants are included in the vaccine in order to enhance the immune system response; preservatives are added to prevent contamination of the preparate by bacteria; stabilisers are introduced to increase the shelf life of the product and to maintain the properties of the vaccine during storage.

How vaccines work: a step back in the 18th century

Although early forms of empirical immunization appear to have been present in different cultures (India and China; [16]), the creation of the first vaccine (for smallpox immunization) dates back to 1798 by Edward Jenner, an English physician. Jenner had noticed that milkmaids who became infected with cowpox (Vaccinia Virus), a virus that causes similar symptoms to human smallpox (Variola Virus or Smallpox virus) but not fatal, did not subsequently develop the disease [17]. This suggested that Cowpox inoculation could protect against Smallpox. Jenner decided to test his theory by inoculating an eight-year-old boy, the son of his gardener (sic!), James Phipps, with material taken from the cowpox lesions of a local milkmaid. As expected, James developed few local lesions and a modest fever. Two months later Jenner inoculated James with variolous matter from a case of human smallpox, without a sensible effect was produced: Jenner had proved that the boy had been immunized. By definition, all subsequent immunizations would be called vaccinations as in 1881 Louis Pasteur proposed it as a general term for the new protective inoculations, in honor of Jenner.

Once inoculated, vaccines (all of them), mimicking the first contact between man and pathogen, are able to stimulate an immunological response (humoral and

cellular) as if this occurred through a natural contagion, although not leading to disease and without giving the associated complications. The rationale behind this phenomenon is immunological memory: the body/immune system that has already experienced a pathogenic microorganism, treasures the experience by responding rapidly to the same microorganism (the absence of immunological memory is the reason why Covid-19 emerged as a problem for humans). For some vaccines it is necessary to make recalls at a distance of time. Normally our body reacts to an unwanted host, but it can take up to two weeks to produce a sufficient amount of antibodies versus the pathogen. In the absence of vaccination, in this interval of time a pathogen can create damage to the body and even lead to death.

Types of vaccines

A long way has been covered since 1798, and technologies have steadily improved to arrive at hi-tech vaccines such as those we are using today to fight Covid. The types of vaccine that exist today are:

- live attenuated vaccines (e.g.: measles and tuberculosis): these are produced from infectious agents that have been rendered non-pathogenic;
- inactivated vaccines (e.g.: poxvirus): these are produced using infectious agents killed by heat or chemicals;
- purified antigen vaccines (e.g.: anti-meningococcal): these are produced by purifying specific components (bacterial or viral);
- anatoxin vaccines (e.g.: tetanus): these are produced using molecules from the infectious agent, which are not capable of causing the disease on their own, but which can stimulate/activate the immune defenses of the vaccinee;
- recombinant protein vaccines (e.g.: hepatitis B): these are produced using recombinant DNA technology, which involves inserting genetic material coding for the antigen (a protein/peptide) into microorganisms capable of producing the antigen specifically, allowing it to be purified;
- recombinant mRNA vaccines (e.g., Pfizer/BioNTec and Moderna): these are produced using an mRNA coding for a target gene encapsulated nanoparticle made with lipid bilayers; this information is able to drive the synthesis of an antigenic protein, through the cell machinery, and the triggering of the immunitary system;
- recombinant viral vector vaccines (e.g., Astrazeneca/Oxford): these are produced using an DNA coding for a target gene carried within a defective adenovirus (from human or from chimpanzee), able to vehicle the gene within the cell nucleus. This information is able to drive the synthesis of an antigenic protein, through the cell machinery, and the triggering of the immunitary system.

The latter types of vaccine (mRNA and Adenoviral vector-based) were today adopted mainly because they can be manufactured very quickly, and being their production process highly standardized. As a matter of fact, they are the fastest way to create a vaccine in the middle of a pandemic.

mRNA and adenovirus-based vaccines

Let us take a step back. The CoV-SARS-2019 virus has on its surface a protein called Spike (S protein), that it uses to enter a human cell via binding to the ACE2 receptor (Pescarmona figure?). The S-protein has therefore been chosen as the specific target to produce a vaccine since it is exposed in large quantities on the surface of the virus. mRNA-based and adenovirus-based vaccine for Covid target the S-protein through the production of a RNA messengers (mRNAs), the classical molecule that routinely instructs all the cells what to build.6 Once the S-protein is produced within the body and presented the the immune system, it is considered an antigen, and the body starts producing antibodies against it. The same thing can be done by using a pre-made protein and injecting it... But its production, testing and approval is longer (years to decades) and more expensive. In a mRNA-based vaccine (e.g., Pfizer/Moderna) the mRNA coding for the Cov-Sars-2019 spike (S protein) is encapsulated in lipid nanoparticles. This preparate is then injected (usually in the deltoid muscle). After that, the nanoparticles fuse with the cell membranes and mRNA is released into the cell cytoplasm, without entering in the nucleus nor getting incorporated into the genomic DNA. In a adenovirus-based vaccine (e.g.: AstraZeneca) the gene coding for the spike protein is inglobated as DNA in a defective Chimpanzee adenovirus which is not able to proliferate, alone. This virus once injected latch on the host cell and released DNA (carrying the spike protein gene) in the cell cytoplasm. DNA then migrates to the nucleus where it is transcribed into mRNA, which will migrate to the cytoplasm. In both the vaccines, at this particular stage the mRNA coding for the Spike uses the cellular machinery (e.g.: ribosomes) for being translated into protein, imitating virus-infection-like humoral immunity and cellular immunity ([18]. Both mRNA-based and adenovirus-based vaccines are able to increase the host's anti-virus effects by increasing T cells' antigen reactiveness [19]. Normally, are these white blood cells, as the first defenses of the body, to "detect" the presence of the pathogen and to organize a protection by generating specific antibodies to combat it through B-lymphocytes, as particular blood cells deputed to antibody production [20] These antibodies cover the virus and prevent it from attacking our body. The immune system memory can be compared to a human's memory. Once it encounters an unwanted visitor, it will remember it and will be able to recognize it in the future. This process typically takes a few weeks for the body to produce the antibodies, but these cells will be there to guard the body for a long time.

7.2 Planning a vaccination campaign using genetic algorithms, with non-pharmaceutical containment measures in action

We compare the effect of choosing the vaccination quotas via genetic algorithms (GAs) with two predetermined strategies. Our model considers three hypotheses: vaccinated people still spread the contagion; they do not spread the contagion; they do it in the 50% of the case. We show here only the results of the first case, the worst.

Important dates:

- in the internal calendar of the model, day 373 is February 12th, 2021; it is effectively the starting point of the vaccinations in Piedmont;
- the day of the effectiveness of the initial vaccinations, 40 days later, is day 413 (March 22nd, 2021).

A technical detail: we simulate the vaccination campaigns with the GAs using the BehaviorSearch program, https://www.behaviorsearch.org, which is strictly related to NetLogo.

7.2.1 Vaccination groups

We take into consideration seven groups in order of decreasing fragility and considering the exposure to contagions:

g1 extra fragile people with three components;

- due to intrinsic characteristics: people in living in nursing homes;
- due to risk exposure:
  - nursing homes operators;
  - healthcare operators;

g2 teachers;
g3 workers with medical fragility;
g4 regular workers;
g5 fragile people without special characteristics;
g6 regular people, not young, not worker, and not teacher;
g7 young people excluding special activity cases (a limited number in *g1*).

7.3 A specific realistic case

The description of the vaccination effects on an outbreak is quite lengthy. Considering the collection at https://terna.to.it/simul/GAresultPresentation.pdf, we concentrate on a unique case: the experiment I reported there, maintaining the reference to I in the titles of the figures here. Considering the adoption of the government non-pharmaceutical measures, we search—in the batch of the 10,000 outbreaks of Section 5.3—for realizations of sequences similar to the actual events that occurred in Piedmont. As we see in Fig. 9 and the related Fig. 10, the artificial case that we adopt for the GAs exploration has the following critical characteristics:

(i) numbers of infected persons quite similar at `cp2` and at `cp3` in Fig. 6a; besides, numbers not too different from those of the same figure;
(ii) number of infected persons at `cp4` significantly greater than those at the previous checkpoint.

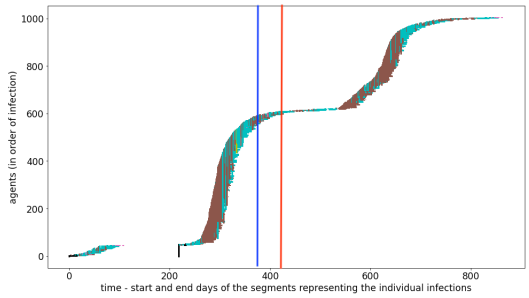


Fig. 9: Crucial dates: blue line for the starting point of the vaccination campaign and red line for the start of the effectiveness of the initial vaccinations; all the situations without vaccination

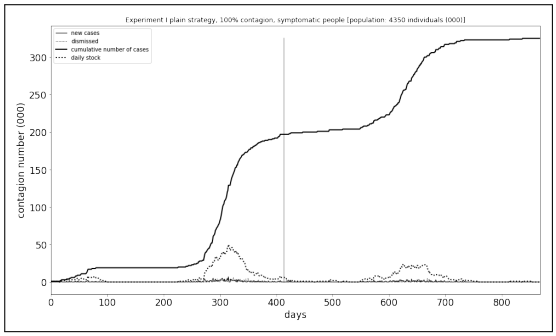


Fig. 10: Base symptomatic series; the vertical line at day 413 is not relevant here

In Fig. 10, with all the situation without vaccinations, we have the first wave in Spring 2020, a larger one in Fall 2020, a limited one between the end of 2020 and the beginning of 2021; then, a relatively quiet interval and successively, just while we write these notes, some restarting signals; finally, currently in the future, a fourth wave. Very realistic with Piedmont's actual situation, the limited thickness of the *snake* of Fig. 9 when vaccinations start and when their effectiveness develops. The hole in the series identifies a period of quasi-extinct epidemics. Then it restarts with the arrival of infected persons from outside.

Here and in the following sections, we analyze the count of symptomatic persons, being the goal of our simulated vaccination campaign exactly that of decreasing the number of symptomatic people, as deceased persons come from there.

From day	Q. of vaccines (000)	*g1*	*g2*	*g3*	*g4*	*g5*	*g6*	*g7*
373	5	0.1	0.1	0.1	0.1	0.1	0.1	0.1
433	10	0.1	0.1	0.1	0.1	0.1	0.1	0.1
493	10	0.1	0.1	0.1	0.1	0.1	0.1	0.1
553	10	0.1	0.1	0.1	0.1	0.1	0.1	0.1
613	20	0.1	0.1	0.1	0.1	0.1	0.1	0.1
738	end							

Table 9: From the day of the first column, considering the quantity of the second column (000), the vaccination of each group follows the quota of the related columns

(000)	*g1*	*g2*	*g3*	*g4*	*g5*	*g6*	*g7*
Susc. at t = 0	133	84	240	1560	1179	254	900
Susc. when vacc. starts	124	81	162	1234	1032	245	891

Table 10: Susceptible persons at the beginning of the simulation and when the vaccination campaign starts, day 373, Feb. 12th, 2021

7.4 Vaccination quotas, *plain* strategy

The vaccination plans are related to the first dose; the second dose is automatically scheduled, with an independent supply. The vaccinated person is supposed to benefit from immunity 40 days after the first dose.

Considering a *plain* option as that adopted in Table 9 with, in each day, the quantities of doses of the first column, we will primarily vaccinate the left column groups to move gradually to people of the other columns, as those on the left have already received the vaccine. The order is (*g1*) extra fragile people, (*g2*) teachers, (*g3*) fragile workers, (*g4*) regular workers, (*g5*) fragile people, (*g6*) regular people, (*g7*) young people. In Table 10 we have numbers both of persons in each category at the beginning of this experiment (and in the following ones) and when the vaccination campaign starts.

Some of the coefficients in Table 9, and all the successive similar ones, are not used in two situations:

i when the persons of a group have fully vaccinated, the quotas in the rows below that day are not relevant;
ii when the people in the columns to the left of a given column completely absorb the available doses of vaccine on that day, the quotas in that column have unimportant values.

We anticipate that the GAs procedure does not optimize the coefficients of cases i and ii.

From day	Q. of vaccines (000)	*g1*	*g2*	*g3*	*g4*	*g5*	*g6*	*g7*
373	5	0.1	0.1	0.1	0.0	0.1	0.0	0.0
433	10	0.1	0.1	0.1	0.0	0.1	0.0	0.0
493	10	0.1	0.1	0.1	0.1	0.1	0.1	0.1
553	10	0.1	0.1	0.1	0.1	0.1	0.1	0.1
613	20	0.1	0.1	0.1	0.1	0.1	0.1	0.1
738	end							

Table 11: From the day of the first column, considering the quantity of the second column (000), the vaccination of each group follows the quota of the related columns

The series that we introduce hereafter are most significant from day 413, March 22nd, when the initial vaccinations' effectiveness—if any—begins, after 40 days from initial vaccinations.

In Fig. 11a we have the effects of the vaccination plan as the number of vaccinated persons by groups. In Fig. 11b we have the most important outcome: the no vaccination test-bed is that of Fig. 10. We note the waves after the vertical line—when vaccinations start to operate—are less relevant than in the test plot, but anyway, those further waves are there.

7.5 Vaccination quotas, *wise* strategy

Considering now a *wise* option, as an attempt to mimic the actual (and complex) vaccine distribution of the vaccination in the region, we use the quotas of Table 11, with the exact mechanism of the previous section. We primarily vaccinate the left column groups to move gradually to other columns, but postponing group *g4* (regular workers), *g6* (regular people), and *g7* (young people). In Table 10 we have numbers both of persons in each category at the beginning of this experiment (and in the following ones) and when the vaccination campaign starts. The considerations sub i and ii in Section 7.4 apply also here.

In Fig. 11c we have the effects of the vaccination plan as the number of vaccinated persons by groups. In Fig. 11d we have the experiment outcome: the no vaccination test-bed is that of Fig. 10. We note the waves after the vertical line—when vaccinations start to operate—are less relevant than in the test plot, but we have significant further waves in this case too.

From day	Q. of vaccines (000)	*g1*	*g2*	*g3*	*g4*	*g5*	*g6*	*g7*
373	5	0.01	0	0	0.79	0.18	0.38	0.19
433	10	0.94	0.06	0.32	0.54	0.19	0.83	0.5
493	10	0.97	0.97	0.74	0.79	0.2	0.14	0.52
553	10	0.98	0.83	0.02	0.39	0.99	0.04	0.48
613	20	0.52	0.01	0.83	0.6	1	0.27	0.9
738	end							

Table 12: GAs best strategy with *vaccinated people still spreading the infection*: from the day of the first column, considering the quantity of the second column, the vaccination of each group follows the quota of the related columns

7.6 GAs quotas in the experiment, with vaccinated people spreading the infection

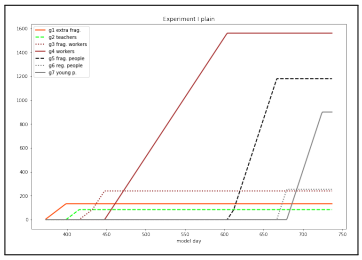
Finally, the objective of this whole section: the use of GAs, evolving populations of models repeating this same experiment, but choosing the parameters to decide daily vaccination on a random basis initially and successively considering them as a genetic chromosome of the model re-productively crossed with those of other models searching for the best fitness related to the goal of reducing the number of symptomatic persons. [21], also quoted at https://www.behaviorsearch.org, is a helpful introduction to the methodology; the sources of the GAs used here are at https://github.com/terna/GAs. The GAs action, determining the vaccination quotas, optimizes the behavior of a meta-agent, in a sort of *inverse generative social science perspective* [22].

With the GAs option, we use the quotas of Table 12, with the exact mechanism of the previous section. The considerations sub i and ii in Section 7.4 also apply here. Still, we underline that the GAs procedure does not optimize the coefficients of the two cases because they do not affect the fitness related to the minimum number of symptomatic subjects.

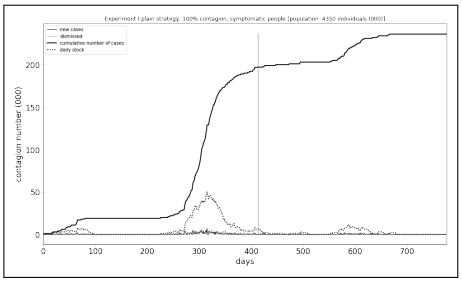
In Table 10 we have numbers both of persons in each category at the beginning of this experiment (and in the following ones) and when the vaccination campaign starts.

In Fig. 11e we have the effects of the vaccination plan as the number of vaccinated persons by groups. The main attention of the GAs is initially is related to the groups: *g4* (workers), *g1* (extra-fragile persons), *g3* (fragile workers), *g2* (teachers). Then *g5* (fragile people) finally *g6* (regular people) and *g7* (young people). The priority is for highly circulating persons (workers and teachers), then for fragile persons.

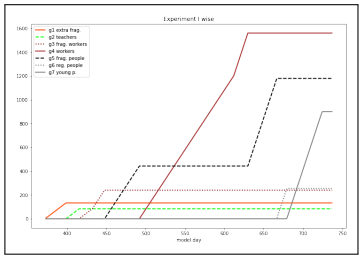
In Fig. 11f we have the crucial result of this experiment: the no vaccination test-bed is always that of Fig. 10. With GAs' choices, the waves after the vertical line—when vaccinations start to operate—disappear, and the whole outbreak is a lot shorter.



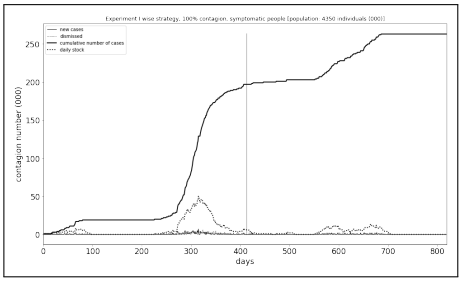
(a) *Plain* vaccination sequence; on the y axis the number of vaccinated subjects of each group (if vaccination is complete, the line is horizontal)



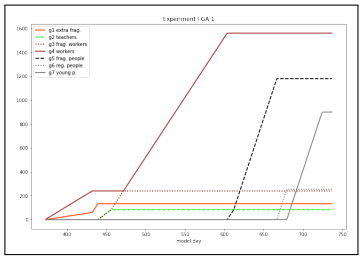
(b) *Plain* vaccination symptomatic series; the vertical line is at day 413, when the effectiveness of first vaccination starts



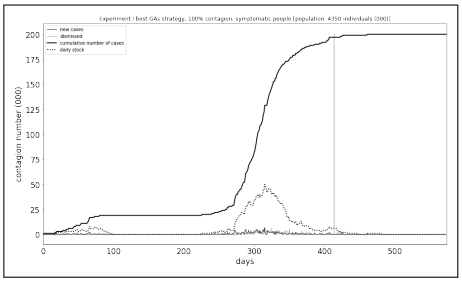
(c) *Wise* vaccination sequence; on the y axis the number of vaccinated subjects of each group (if vaccination is complete, the line is horizontal)



(d) *Wise* vaccination symptomatic series; the vertical line is at day 413, when the effectiveness of first vaccination starts



(e) GA vaccination sequence; on the y axis the number of vaccinated subjects of each group (if vaccination is complete, the line is horizontal)



(f) GAs vaccination symptomatic series; the vertical line is at day 413, when the effectiveness of first vaccination starts

Fig. 11: Vaccination sequences and time series

8 A new model and future developments

Using SLAPP, https://terna.github.io/SLAPP/ a second model is under development, with a ratio of 1:100 to the Piedmont population, so 43,500 agents. It will contain the same items as the current one, plus transportation, aggregation places as happy hours, nightlife, sports, stadiums, discotheques, etc. We will also consider networks as fam-

ily networks, professional networks, high-contact individual networks [23]. Finally, we will take into consideration the socioeconomic conditions of the individuals.

As seen, the S.I.s.a.R. model is a tool for comparative analyses, not for forecasting, mainly due to the enormous standard deviation values intrinsic to the problem. The model is highly parametric, and more it will be, precisely in the comparative perspective. It also represents a small step in using artificial intelligence tools and the inverse generative perspective in agent-based models.

9 Appendix 1—Parameter values

We report here the values of parameters of Fig. 2, also with the short names used in program scripts, in round brackets. Look at Section 2.2 for the definition. Day numbering is related to actual dates via the Table 13. Day 1 is fixed at Feb 4th, 2020.

The values adopted in the experiments reported in this work are the following.

i The values of *probabilityOfGettingInfection* (`prob`) are: 0.05 (starting phase); 0.02 at day 49 (adoption of non-pharmaceutical measures); 0.035 at day 149 (some relaxation in compliance); 0.02 at day 266 (again, compliance to rules).

ii The value of *D%* is -50 in all the runs;

iii *intrinsicSusceptibility* is set discussing Eq. 1 in (2.2, iii).

iv The values of *%PeopleAnyTypeNotSymptomaticLeavingHome* (`%PeopleAny`) are: at (day) 20, 90; at 28, 80; at 31, 0; at 106, 80; at 110, 95; at112, 85; at 117, 95; at 121, 90; at 259, 90; at 266, 80; at 277, 50; at 302, 70; at 320, 90; at 325, 50; at 329, 80; at 332, 50; at 336, 80; at 337, 50; at 339, 80;

v The values of *%PeopleNotFragileNotSymptomaticLeavingHome* (`%PeopleNot`) are: at (day) 31, 80; at 35, 70; at 36, 65; at 38, 15; at 42, 25; at 84, 30; at 106, 0; at 302, 90; at 325, 50; at 332, 50; at 337, 50; at 339, 100; at 349, 90;

vi The values of *%openFactoriesWhenLimitationsOn* (`%Fac`) are: at (day) 38, value4 0; at 49, 20; at 84, 70; at 106, 100; at 266, 90; at 277, 70; at 302, 80; at 320, 90; at 325, 30; at 329, 90; at 332, 30; 336, 90; at 337, 30; at 339, 100;

vii *stopFragileWorkers* (`sFW`): by default, 0; in one of the experiments we used `sFW` with set to 1 (on) at day 245 and to 0 (off) at day 275;

viii The values of *activateSchools* (`aSch`) are: at (day) 1, on; at 17, off; at 225, on; at 325, off; at 339, on;
the values of *%Students* (`%St`) are: at (day) 0, 100; at 277, 50; at 339, 50; at 350, 50 (repeated values are not relevant for the model, but for the use of the programmer-author);

ix The value of *radiusOfInfection* (`radius`) is 0.2; in the model, space is missing of a scale, but forcing the area to be in the scale of a region as Piedmont, 0.2 is equivalent to 20 meters; we have to better calibrate this measure with movements and probabilities; this is a critical step in future developments of the model.

Day	Date	Day	Date	Day	Date	Day	Date
25	28- 2-2020	200	21- 8-2020	375	12- 2-2021	550	6- 8-2021
50	24- 3-2020	225	15- 9-2020	400	9- 3-2021	575	31- 8-2021
75	18- 4-2020	250	10-10-2020	425	3- 4-2021	600	25- 9-2021
100	13- 5-2020	275	4-11-2020	450	28- 4-2021	625	20-10-2021
125	7- 6-2020	300	29-11-2020	475	23- 5-2021	650	14-11-2021
150	2- 7-2020	325	24-12-2020	500	17- 6-2021	675	9-12-2021
175	27- 7-2020	350	18- 1-2021	525	12- 7-2021	700	3- 1-2022

Table 13: The days of the simulation and their equivalent dates in the calendar

10 Appendix 2—A gallery of contagion sequences

The gallery of contagion sequences, reported in Table 14, shows the vast variety of situations generated by the agent-based simulations, producing actions and interactions of the agents.

- 1a An outbreak without containment measures, with a unique wave, but very heavy: contagions are in nursing homes (orange), workplaces (brown), homes (cyan), hospitals (pink).
- 1b This is the previous epidemic without containment measures, considering the first 200 infections, with the main contribution of nursing homes (orange) and workplaces (brown).
- 1c Another outbreak, always without containment measures: nursing homes (orange) as a starter.
- 2a The [1c] epidemic, without containment measures, first 200 infections: nursing homes (orange) as a starter; around day 70, a unique contagion at home makes the epidemic continue.
- 2b Another case without containment measures showing the initial action of contagions in workplaces (brown) and homes (cyan).
- 2c Here we see the first 200 infections showing that the initial profound effects of contagions in workplaces (brown) and homes are due, in the beginning, to the fragile persons, also asymptomatic,
- 3a An outbreak with containment measures, where we see another influential contribution of workplaces (brown) and homes (cyan) to the epidemic diffusion.
- 3b Here the first 200 infections: after day 100, we observe many significant cases of fragile workers diffusing the infection.
- 3c In this outbreak, with containment measures, the infections arise from workplaces (brown), nursing homes (orange), and homes (cyan), but also hospitals (pink).
- 4a Here we explore the first 200 infections of [3c]: in the beginning, workplaces (brown), hospitals (pink), nursing homes (orange), and homes (cyan) are interweaving.

4b An outbreak with containment measures where the effect of the contagions in workplaces (brown), nursing homes (orange), and homes (cyan) is evident.
4c In the first 200 infections of [4b], workplaces (brown) and nursing homes (orange) are strictly interweaving.
5a An outbreak with containment measures where the effect of nursing homes (orange) is prevalent.
5b An outbreak with containment measures with a highly significant effect from workplaces (brown).
5c Stopping fragile workers at day 20 in the previous case, we obtain a beneficial effect, but home contagions (cyan) keep alive the pandemic, which explodes again in workplaces (brown).
6a Exploring the first 200 infections of the case [5c], we have evidence of the event around day 110 with the new phase due to a unique asymptomatic worker.
6b Finally, the same epidemic stopping fragile workers and any fragility at day 15 case and isolating nursing homes.
6c An outbreak with containment measures spontaneously stopping in a short period.

Acknowledgements

Many thanks to Simone Landini, Nizar Mansour, Matteo Morini, Fabio Pammolli, Enrico Scalas, and Federico Tedeschi for precious discussions, insights, and critics. The usual disclaimer applies.

To run our model, we tremendously benefit from the use of HPC facilities provided by Matteo Morini and Sarah de Nigris, http://tomorrowdata.io, and the HP4AI (https://hpc4ai.unito.it/) and C3S centers at University of Torino.

References

1. A. Scala, A. Flori, A. Spelta, E. Brugnoli, M. Cinelli, W. Quattrociocchi, F. Pammolli, Scientific Reports **10**(1), 13764 (2020). DOI 10.1038/s41598-020-70631-9. URL https://doi.org/10.1038/s41598-020-70631-9
2. N. Bellomo, R. Bingham, M.A.J. Chaplain, G. Dosi, G. Forni, D.A. Knopoff, J. Lowengrub, R. Twarock, M.E. Virgillito, Mathematical Models and Methods in Applied Sciences **30**(08), 1591 (2020). URL https://doi.org/10.1142/S0218202520500323
3. H. Rahmandad, J. Sterman, Management Science **54**(5), 998 (2008)
4. F. Squazzoni, J.G. Polhill, B. Edmonds, P. Ahrweiler, P. Antosz, G. Scholz, E. Chappin, M. Borit, H. Verhagen, F. Giardini, N. Gilbert, Journal of Artificial Societies and Social Simulation **23**(2), 10 (2020). DOI 10.18564/jasss.4298. URL http://jasss.soc.surrey.ac.uk/23/2/10.html
5. J.B. Bak-Coleman, M. Alfano, W. Barfuss, C.T. Bergstrom, M.A. Centeno, I.D. Couzin, J.F. Donges, M. Galesic, A.S. Gersick, J. Jacquet, A.B. Kao, R.E. Moran, P. Romanczuk, D.I. Rubenstein, K.J. Tombak, J.J. Van Bavel, E.U. Weber, Proceedings of the National Academy

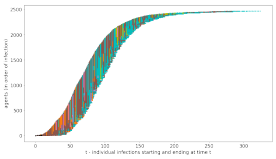
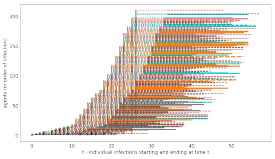
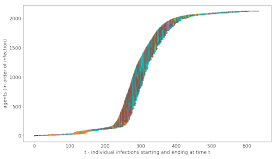
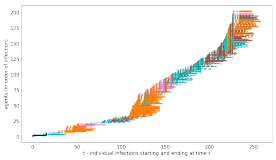
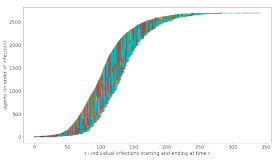
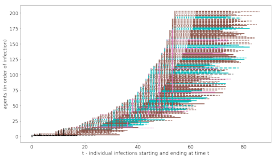
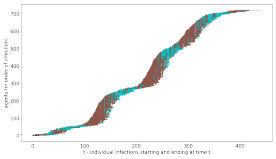
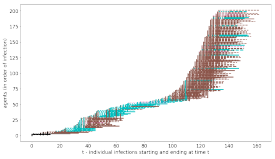
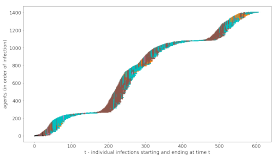
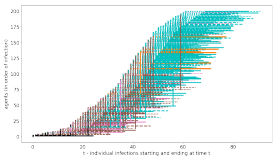
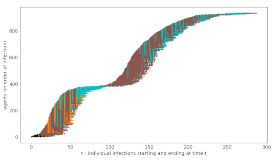
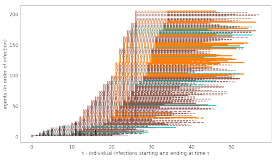
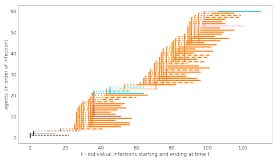
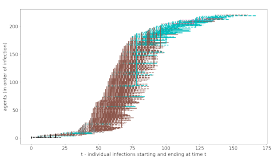
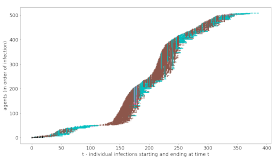
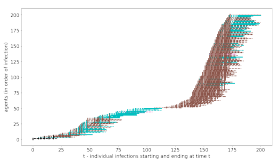
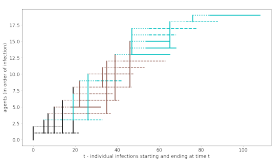
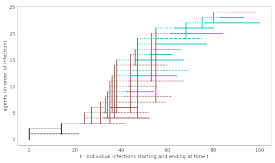
Nr.	a	b	c
1			
2			
3			
4			
5			
6			

Table 14: Gallery of sequences, symptomatic and asymptomatic agents

of Sciences **118**(27) (2021). DOI 10.1073/pnas.2025764118. URL https://www.pnas.org/content/118/27/e2025764118

6. J. Epstein, Journal of Artificial Societies and Social Simulation **11**(4), 12 (2008). URL http://jasss.soc.surrey.ac.uk/11/4/12.html
7. R. Axtell, Why agents? On the varied motivations for agent computing in the social sciences. Tech. rep., Center on Social and Economic Dynamics Brookings Institution (2000)
8. A. Ghorbani, F. Lorig, B. de Bruin, P. Davidsson, F. Dignum, V. Dignum, M. van der Hurk, M. Jensen, C. Kammler, K. Kreulen, L.G. Ludescher, A. Melchior, R. Mellema, C. Păstrăv, L. Vanhée, H. Verhagen, Review of Artificial Societies and Social Simulation (2020). URL https://rofasss.org/2020/04/25/the-assocc-simulation-model/

9. F. Silvagno, A. Vernone, G.P. Pescarmona, Antioxidants **9**(7), 624 (2020). DOI 10.3390/antiox9070624. URL http://dx.doi.org/10.3390/antiox9070624
10. R. Horton, Lancet (London, England) **396**(10255), 874 (2020). URL https://www.thelancet.com/action/showPdf?pii=S0140-6736%2820%2932000-6
11. U. Wilensky. Netlogo (1999). URL http://ccl.northwestern.edu/netlogo/
12. T.M. Vu, C. Probst, J.M. Epstein, A. Brennan, M. Strong, R.C. Purshouse, Genetic and Evolutionary Computation Conference : [proceedings]. Genetic and Evolutionary Computation Conference **2019**, 1356 (2019). DOI 10.1145/3321707.3321840. URL https://pubmed.ncbi.nlm.nih.gov/33083795
13. P. Steinmann, J.R. Wang, G.A. van Voorn, J.H. Kwakkel, Review of Artificial Societies and Social Simulation **17** (2020). URL https://rofasss.org/2020/04/17/deep-uncertainty/
14. N. Phillips, Nature **590**(7846), 382 (2021). URL https://www.nature.com/articles/d41586-021-00396-2
15. D. Miles, M. Stedman, A.H. Heald, International Journal of Clinical Practice **n/a**(n/a), e13674 (2020). DOI 10.1111/ijcp.13674. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ijcp.13674
16. A. Boylston, Journal of the Royal Society of Medicine **105**(7), 309 (2012). URL https://doi.org/10.1258/jrsm.2012.12k044
17. E. Jenner, *An inquiry into the causes and effects of the variolae vaccinae, a disease discovered in some of the western counties of England, particularly Gloucestershire, and known by the name of the cow pox* (Springfield [Mass.] : Re-printed for Dr. Samuel Cooley, by Ashley & Brewer, 1802, 1800). URL http://resource.nlm.nih.gov/2559001R
18. M.A. Monslow, S. Elbashir, N.L. Sullivan, D.S. Thiriot, P. Ahl, J. Smith, E. Miller, J. Cook, S. Cosmi, E. Thoryk, et al., Vaccine **38**(36), 5793 (2020). URL https://www.sciencedirect.com/science/article/pii/S0264410X20308483
19. Y. Wang, Z. Zhang, J. Luo, X. Han, Y. Wei, X. Wei, Molecular Cancer **20**(1), 33 (2021). DOI 10.1186/s12943-021-01311-z. URL https://doi.org/10.1186/s12943-021-01311-z
20. W. Ratajczak, P. Niedźwiedzka-Rystwej, B. Tokarz-Deptuła, W. Deptuła, Central-European journal of immunology **43**(2), 194 (2018). URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6102609/
21. J.H. Miller, Management Science **44**(6), 820 (1998). URL https://www.jstor.org/stable/pdf/2634650.pdf
22. T.M. Vu, C. Probst, J.M. Epstein, A. Brennan, M. Strong, R.C. Purshouse, in *Proceedings of the Genetic and Evolutionary Computation Conference* (2019), pp. 1356–1363. URL https://dl.acm.org/doi/abs/10.1145/3321707.3321840
23. G. Manzo, A. van de Rijt, Journal of Artificial Societies and Social Simulation **23**(4), 10 (2020). DOI 10.18564/jasss.4435. URL http://jasss.soc.surrey.ac.uk/23/4/10.html