New Zealand DEFENCE FORCE
Te Ope Kātua O Aotearoa

DEFENCE TECHNOLOGY AGENCY

DTA Report 464
NR 1768

A Novel Agent-based Model for COVID-19 Disease Spread

Gregory C McIntosh
Michael K Lauren
November 2021

DTA Report 464
NR 1768

A NOVEL AGENT-BASED MODEL FOR COVID-19 DISEASE SPREAD

G C McIntosh and M K Lauren

November 2021

ABSTRACT

Defence Technology Agency (DTA) has developed a novel agent-based model (ABM) for investigating the spread of the pre-Delta variant of the SARS-CoV-2 virus through a population of several tens of thousands of people. People are represented in the model by abstracted entities, the so-called agents, who reside at home locations defined on a grid topology. The mechanism for virus spread is via healthy or infected agents visiting various venues on a daily basis where other infected agents may or may not be present. We adopt a distillation approach whereby detailed social interactions are deliberately omitted from the model. This keeps the model manageable while still providing insights into the general dynamics of disease spread through the population. Our ABM is intended as an analysis tool for exploring pressure points in controlling the spread of COVID-19, and to provide visualizations of the virus spread for educational purposes. Effects considered include non-pharmaceutical interventions such as limited gathering sizes, restricting movement within the population, and contact tracing with isolation. The spatial component of how the virus spreads through a community can be understood and the time evolution of the reproduction number R_0 can be monitored as a scenario unfolds. This, in turn, is useful for highlighting different phases of the disease spread. The model demonstrates the way in which R_0 does not remain constant with time, even without interventions. Moreover, geometry appears to play a role in determining R_0, which leads to the model falling into different modes of evolution. Finally, by explicitly representing agent movement in terms of travel between ‘venues’, the model has the potential to explore the trade-off between reducing infection rates via interventions and the adverse effects this might have on economic activity or other social aspects such as education.

DTA Report
ISSN 1175-6594 (Print)
ISSN 2253-4849 (Online)

Published by
Defence Technology Agency
Private Bag 32901
Devonport
Auckland 0744
New Zealand

EXECUTIVE SUMMARY

BACKGROUND

Based on previous experience at Defence Technology Agency (DTA) in developing agent-based models (ABMs), a model has been developed for COVID-19 to demonstrate the potential benefit of using this method to simulate the behaviour of disease spread through a population and the effectiveness of different types of intervention. This contrasts with more conventional modelling approaches to epidemiology involving SIR (Susceptible-Infected-Recovered) equations, which are similar to predator-prey models. In that case, interventions are modelled more coarsely by simply adjusting the basic virus reproduction number, R_0. DTA's model was designed to allow a relatively large population to be represented with minimal computational expense. Importantly, the model is intended to explicitly represent the effects of non-pharmaceutical interventions on the spread of the SARS-CoV-2 virus. The behaviour of individual agents in the model can be tailored and the effects of population stratification can be explored. This allows a more nuanced way of examining interventions than SIR models allow, and could be useful, for example, in understanding the likely consequences of reopening a country's borders once its vaccination programme has been completed. More generally, as a prototype, this work is intended to promote a broader range of COVID-19 modelling approaches, as well as support allied partners through international collaborations such as the USINDOPACOM S&T Advisor Board.

AIM

To develop an agent-based epidemiological model which could assist decision makers in understanding different COVID-19 scenarios and options to mitigate disease spread.

RESULTS

For scenarios presented in this report, the ABM was typically set up with a population size of 40,000. The mechanism for virus spread is via healthy or infected agents visiting various locations on a daily basis where other infected agents may or may not be present. We deliberately adopt a distillation approach whereby detailed social interactions are omitted from the model. This keeps the model manageable while still providing insights into the general dynamics of disease spread through the population. By default, agents reside at their home venues, and rulesets can be applied as to whether agents are allowed to visit other venues on a day-by-day basis. This provides a mechanism by which different strategies can be explored for limiting disease spread. Model scenarios can be run within a few minutes on a standalone PC and the dynamics of the virus spreading through the population can be followed on the ABM's display; which uses a colour representation for the status of each agent. In order to quantify the effectiveness of different mitigation strategies, scenarios can be run multiple times to obtain average results along with other statistical measures.

Our model has been calibrated for a pre-Delta variant of the SARS-CoV-2 virus, applicable at the start of 2021. Therefore, with no interventions applied, the

reproduction number for the virus has been set to $R_0 = 2.0 – 2.5$, yielding doubling times for the virus in our model which agree with observations of real populations during the early phases of the COVID-19 pandemic.

The model has generated several insights as follows, some of which have been observed in real populations. On the other hand, the results listed below should be taken with the caveat that they are a reflection of the behaviour of the model, and the degree to which such behaviour would occur in the real world remains to be validated.

- Limiting agents' movement to their local neighbourhood is found to be effective at slowing (but not stopping) disease spread, with a resulting reproduction number of $R_0 \leq 1$. The disease can still eventually sweep through the entire population, but the lower resulting case numbers will have positive implications for managing a healthcare system.
- The above-mentioned localisation strategy must be adhered to strictly to be effective. Our model showed that as little as 1% of agents not following this rule could make it easy for new virus clusters to emerge elsewhere in the population.
- A 'lockdown' in our model has been defined as allowing only one person from each household out into the community each day. This was extremely effective at halting disease spread.
- Contact tracing along with self-isolation has been implemented in our model with the following rule set: if any person at a venue has developed symptoms then all attendees at that venue are requested to remain home until all have recovered. Likewise, all fellow home residents of those attendees are contacted and requested to remain home. This was found to be very effective at halting the disease spread, even with delays imposed on the process. In practice, contact tracing may not be feasible once very large numbers of people have become infected, due to the stress that this places on resources.
- Delays in contact tracing can be tolerated depending on other restrictions that have been applied but, beyond a certain threshold, will not stop the spread of the virus. For the scenarios examined in our model, delays of up to five days would still allow the virus spread to be brought under control when combined with other measures. This is not likely to be the case for the more recent Delta variant of COVID-19.
- Our model indicates minimal tolerance to people not complying with the above-mentioned lockdown rule. Above a small percentage of non-compliance, a transition occurs to a regime where this intervention entirely loses its effectiveness.
- Limiting venue sizes (social gatherings) is found to be very effective at slowing (but not stopping) disease spread, provided other rules have been applied; in particular, limiting agents' movement to their local neighbourhood.
- The virus reproduction number, R_0 was found to be an extrinsic measure of the virus's behaviour and can change dynamically as a scenario unfolds. Hence, this measure is useful for identifying different phases of disease spread. For example, R_0 tends to vary with changing conditions such as the topology of the disease spread through different regions of the population.
- The goal of reducing R_0 to 1 or slightly less is a weak condition for controlling disease spread. If already widespread in the population, the virus will continue to

propagate over many reproduction cycles and could still end up infecting the entire population.

This ABM can be made available to other epidemiological researchers and the source code provided so that the model can be customized. The model is quite extensible and other rule sets or features could be added, for example, to study the behaviour of virus propagation during a vaccine rollout.

SPONSOR

DDTA in support of VCDF request.

CONTENTS

1. INTRODUCTION

At the start of 2021, the SARS-CoV-2 virus, which causes COVID-19, had created a worldwide pandemic with approximately 100 million cases having emerged across the globe and around 2 million reported deaths [1]. In the absence of vaccines, a variety of strategies were adopted by different nations for curbing the spread of the virus. These ranged from strict lockdowns to more relaxed approaches which aimed to keep a country's economy afloat. Before the emergence of the Delta variant, a small number of countries had remarkable success at containing and/or eliminating the virus: New Zealand (NZ), Australia, South Korea, Taiwan and Vietnam. Their strategies involved a combination of lockdowns, strict border controls, effective contact tracing and hygiene practices.

At first glance, one might naively compare SARS-CoV-2 to influenza (flu). Both viruses give rise to respiratory symptoms and can lead to life-threatening pneumonia. However, a major difference is that people infected with the flu tend to exhibit symptoms straightaway and have the opportunity to isolate themselves before infecting others. On the other hand, people infected with SARS-CoV-2 can be contagious for several days before symptoms appear. This gives rise to the insidious nature of the disease whereby a large swathe of the population can become infected before anyone is aware that the virus is present. This factor alone leads to a long tail on the distribution of active cases in a country when attempting quarantine measures to eliminate the disease.

This report focuses on the pre-Delta variant of the SARS-CoV-2 virus and before vaccination programmes were put in place. Several facts regarding this virus had been established during the earlier phases of the pandemic [2 – 4]:

- Following infection, there is an incubation time of roughly 5 days before symptoms appear. During this time, people carrying the virus can be contagious and spread the virus unknowingly.
- The entire timeframe from first becoming infected to displaying symptoms and then recovering is approximately 14 days.
- The reproduction time of the virus tends to follow a Weibull distribution with a mean of about 5 days and a standard deviation slightly less than 2 days [3].
- The R_0-value for SARS-CoV-2 is in the range of 2 to 2.5, where R_0 is defined as the average number of new people infected by an active carrier while he/she remains infectious. This provides a measure of the infectiousness which allows comparison to other diseases.

These facts are sufficient to build mathematical models of COVID-19 disease spread and obtain a broad overview of how a population might be affected, given an initial small number of infected individuals. Various approaches can include the well-known Susceptible-Infectious-Recovered (SIR) model (or stochastic differential equations based on these), tree-branching models and agent-based models. Some of these approaches have been successful in highlighting the degree to which good contact tracing and quarantining procedures can help contain the spread of SARS-CoV-2 through a population [2 – 4]. In this paper, we demonstrate the utility of using an agent-based model (ABM) to study the behaviour of disease spread, building on Defence Technology Agency's (DTA's) previous experience with this technique [5, 6].

The idea of applying ABMs to epidemic modelling is not new, as other research teams have also used this approach for examining the spread of COVID-19 (e.g. [7 – 9]). ABMs represent a pandemic by explicitly modelling people individually as 'agents'. This contrasts with more traditional SIR (Susceptible Infectious Recovered) models, which represent a population in terms of aggregate properties/parameters (although some approaches split the population up into coarse subgroups). Agent-based models have the advantage that they can consider aspects of the virus propagating through a heterogenous population with, for example, differing daily schedules of activity. A final point is that ABMs are also highly suited to stochastic approaches that consider the randomness and uncertainty of the transmission of SARS-CoV-2. As such, the results from these models are usually expressed in terms of probability distributions.

In the following section we describe our novel ABM, which has been designed to allow the modelling of relatively large populations of agents (up to 100,000) without excessive computational expense. A baseline scenario has first been set up to yield an infection rate corresponding to the SARS-CoV-2 virus before the Delta variant emerged. Having established a baseline, we then explore the effect of various rulesets intended to limit the spread of the virus. These include lockdowns, restricting people's movement and simple forms of contact tracing. Finally, we consider real-world aspects in our model such as the distribution of venue sizes in a given population and inefficiencies in applying mitigating strategies.

2. AGENT-BASED MODEL FOR DISEASE SPREAD

ABMs are often – though not necessarily – designed to represent key aspects of a scenario while purposefully leaving out unnecessary detail which might slow down the model's computations [5]. Agents in the ABM are abstract entities which live on a grid of cells and can change state according to the state of nearby agents or local environmental conditions defined in the model. In the case of disease spread, agents can be assigned various attributes to represent people becoming infected with SARS-CoV-2 while visiting different venues according to a defined schedule.

The ABM we have developed for COVID-19 is illustrated in Figure 1. This shows a screenshot for a scenario involving a population size of 40,000 agents placed on a 200 × 200 grid.[1] It should be noted that the grid does not necessarily correlate to the geography of the scenario being modelled but, rather, provides a visualization of each agent's state and set of indices for keeping track of that agent. The venue scheme by which agents can become infected in our model is described below.

Agents can have one of four disease states depending on whether they have become infected, are displaying symptoms or have recovered, as listed in Table 1. An agent's state is indicated by its colour on the display and this allows the progress of the modelled disease spread to be visually monitored.[2] We note that the model could be extended to include other states such as the vaccination status of each agent so that vaccine rollouts could be modelled. The current model focuses on the pre-Delta variant of COVID-19 and before worldwide vaccination efforts had commenced.

[1] This is typical of the scenarios used throughout this report and is found to be large enough to allow complex behaviours to emerge that mimic those of real life, while still being computationally manageable.

[2] The colours in Table 1 have been selected due to ease of visibility between different agent states over adjacent grid squares.

State	Properties	Colour on Grid
Healthy/Uninfected	Able to be infected	Green
Infected	Not yet displaying symptoms	Red
Symptoms	Infectious, remain home	Blue
Recovered	No longer able to be infected in the model.	Dark green

Table 1: Possible agent states.

In our variant of the ABM, agents are assigned to fixed grid points and interact by notionally appearing at different venues according to a daily schedule. As mentioned, the grid does not necessarily correlate to the geography of the scenario being modelled but, rather, provides a visualization of each agent's state and a set of indices for keeping track of each agent. Agents only become infected by virtue of who they share a home with and which venues they travel to, rather than being in direct contact across grid locations. However, depending on the rulesets imposed, it is more likely that agents with indices close to each other will infect one another. In this regard, a distance metric between the agents comes into play after we introduce venues into our model.

The system of venues we use is illustrated schematically in Figure 2. Different venues could be considered, such as homes, workplaces, schools, cafes, shops, movie theatres, and so forth. These can be assigned sizes and visiting schedules representative of a typical population. In order to establish a baseline behaviour for our model, we first consider only two types of venue: (i) homes of four inhabitants and (ii) generic venues of size 100. For the purpose of this demonstration, agents have been assigned a daily schedule of 9 am to 5 pm to attend non-home venues.[3]

Scenarios begin with agents at their home venues in state = Healthy, as given by Table 1. They then appear at different venues each day where they can become infected if there are other infected agents at the same venue. They can also be infected at their home venues if other members of the household become infected.

The probability of infection per agent per unit time within a home or outside venue is given by:

$$p(infection) = \beta' \frac{N_I}{N_T} \quad (1),$$

where

- N_I = number of infected agents at the venue,
- N_T = total number of agents at the venue, and
- β' = scaling factor for overall infectiousness of a venue per unit time step.[4]

[3] The time step in our model is 1 hour.
[4] This parameter is referred to as β' so as to avoid confusion with β in the SIR model.

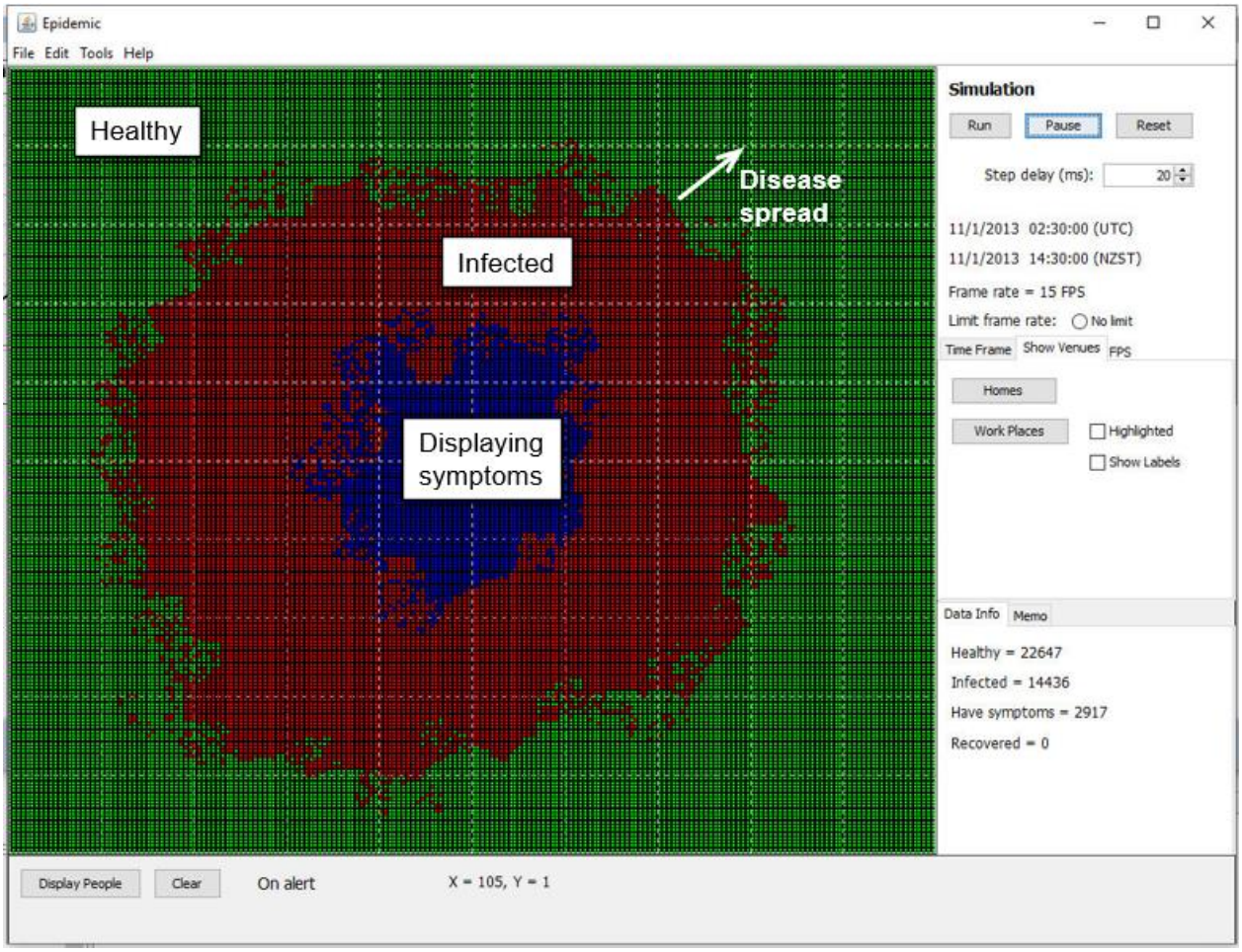


Figure 1: An agent-based model for disease spread.

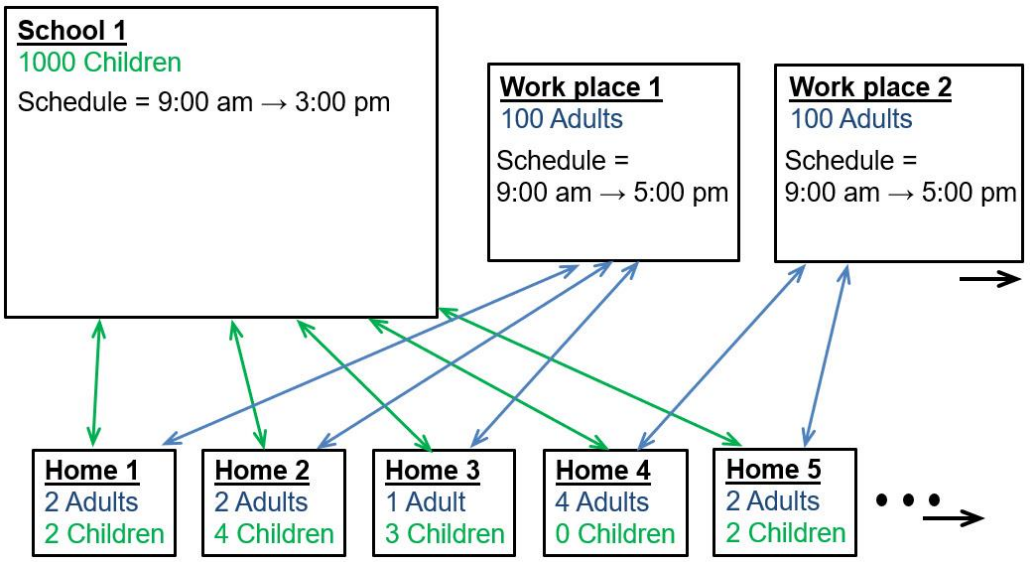


Figure 2: System of venues used in the model. Agents switch venues each day depending on a prescribed schedule. By default, agents reside at home venues. Note that all agents in the population are assigned to homes and schools or workplaces until these venues are at full capacity. For example, a population of 40,000 adults implies 400 workplace venues in the population, each with capacity of 100. This is assuming that nobody works from home.

Equation (1) is based on the infection rate formula used in the SIR model (see Eq. (2) on Page 8). It is calculated for all agents at every time step in the model. In our case, the model time step is set to 1 hour, which gives a reasonable compromise between having the model run quickly and the required time resolution for agents' schedules.[5] (We note that the parameter β' scales according to the model time step, Δt.)

Equation (1) provides a simple but effective representation for infection between people. It has the following characteristics:

- For reasonable-sized venues such as large workplaces or schools, there is a low probability of becoming infected to begin with if there are only one or two other infected people at the same venue.
- If most other people at a venue are already infected, there is a high probability of any remaining people becoming infected.
- Households, with their small number of occupants, are almost guaranteed to infect any healthy people within the same household.
- The scaling factor, β', allows scenarios to be calibrated to yield R_0-values consistent with COVID-19. Through the calibration scheme explained in Section 3, we find that a value of $\beta' = 0.015$ gives $R_0 \sim 2.2$ for unrestrained exponential growth in the disease spread. (At first glance, this value of β' appears small but, unlike the reproduction number, it is a probability of transmission for a unit time interval, rather than for the entire period of contagiousness.)

Our baseline scenario has the following specifications, and is based on the pre-Delta variant of the SARS-CoV-2 virus:

- Population size = 40,000.
- $\beta' = 0.015$ ($R_0 \sim 2.2$).
- Each agent attends a pre-designated venue every day. There are no restrictions on which venue each agent can be assigned to. For example, this represents agents' freedom to visit any venue within the population. (In Section 3, we explore scenarios where agents are restricted in their movement.) All agents in the population are pre-assigned to a home or outside venue to the extent that those venues will be at full capacity. For example, a population of 40,000 adults assigned to workplaces of size 100 implies 400 workplace venues in the population. This is a simplifying assumption compared to the work life of a real population, while still allowing for interesting dynamics of the virus spread to play out in the model.
- At the start of each scenario, one infected agent is placed at the centre of the population grid.
- After becoming infected, agents are infectious for 5 days before displaying symptoms. During this time, they continue to travel to their prescribed venues each day and infect other agents. After symptoms appear, agents remain at their home venue until they recover.
- Recovery time is 9 days after symptoms appear. Deaths are not included in our model, since the main focus is on the dynamics of the disease spread. The death rate could be inferred as a percentage of infected case numbers.
- We assume a flat probability distribution for infecting other people over the time frame for which people are infectious. Our preliminary investigation has revealed that including a more realistic probability distribution, such as a Weibull distribution, does not qualitatively change the nature of our results, and essentially involves a

[5] The model time step could be adjusted to give finer detail in agents' schedules if required.

recalibration of the β' value to achieve the same required reproduction number of $R_0 \sim 2$.

- Asymptomatic cases have also been introduced into the model. These agents represent people who contract the virus, are infectious but do not display symptoms. We have assigned 30% of the population to be asymptomatic, which is of similar order to the rate that has been estimated in real populations for the pre-Delta variant of COVID-19 [10]. As it stands, we have applied the simplifying assumption that asymptomatic agents are equally as infectious as symptomatic agents, whereas, in reality they are likely to be less infectious. The effects of lowering their infectiousness could be explored as this research is further progressed, but would undoubtedly make the disease less aggressive.

Further details of how our ABM operates are illustrated by the pseudocode in Figures 3 and 4. Figure 3 shows tests which are carried out at every model time step for an agent travelling from a home venue to an outside venue. These tests are carried out for every agent in the population grid and involve a sequence of logical tests of whether or not an agent is eligible to travel outside its home venue. Examples include whether an agent is displaying symptoms, has been told to self-isolate after contact tracing or is assigned to work from home. Also, the scheduled travel time is included as a further test. If an agent meets all criteria, then its state is changed from ‘at home’ to ‘at work’ and the attendance records of the agent’s home and outside venues are adjusted accordingly. The module for testing whether or not agents return home from an outside venue follows similarly. However, this is simpler because agents remain at their home venue by default under most contingencies and fewer tests are required.

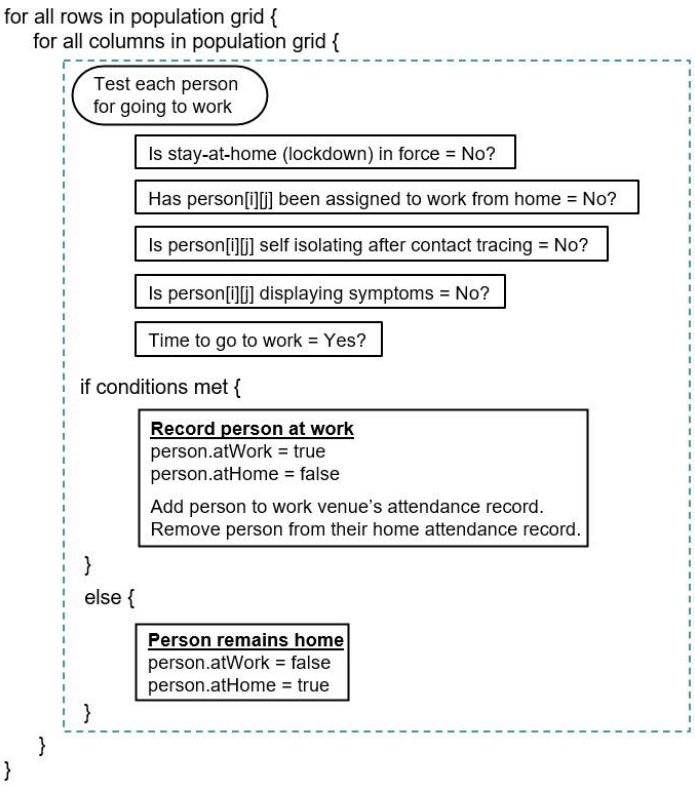


Figure 3: Pseudo code for testing agents travelling from their home to outside venues at each time step.

Figure 4 shows a module for calculating whether or not an agent becomes infected while attending a particular venue, and makes use of Eq. (1). Once again, all agents in the population are considered and this is carried out for every model time step. Other modules, such as deciding whether an agent has become symptomatic or has recovered, based on elapsed time, also appear similarly. However, these modules only need to be calculated once every 24 hours in scenario time, thus, reducing computational load.

```
for each workplace venue {
    Test for attendees becoming infected
        Count number of attendees, N_T
        Count number of attendees who are infected, N_I
        Calculate probability of infection at the venue:
        p(infection) = β' N_T/N_I, (see Eq. (1)).
        for each healthy attendee at the venue {
            if (Random number > p(infection)) {
                attendee.status = Infected
                attendee.timeOfInfection = Simulation time
            }
            else {
                attendee remains healthy
            }
        }
}
```

Figure 4: Pseudo code for calculating infections at each time step.

Results of running the baseline scenario are given in Figure 5. This shows the number of agents in each health state as listed in Table 1. The initial exponential increase in infected cases is apparent and the disease sweeps through the entire population in 3 to 4 months of scenario time. The delay between agents becoming infectious and their symptoms appearing is clear from the graphs.

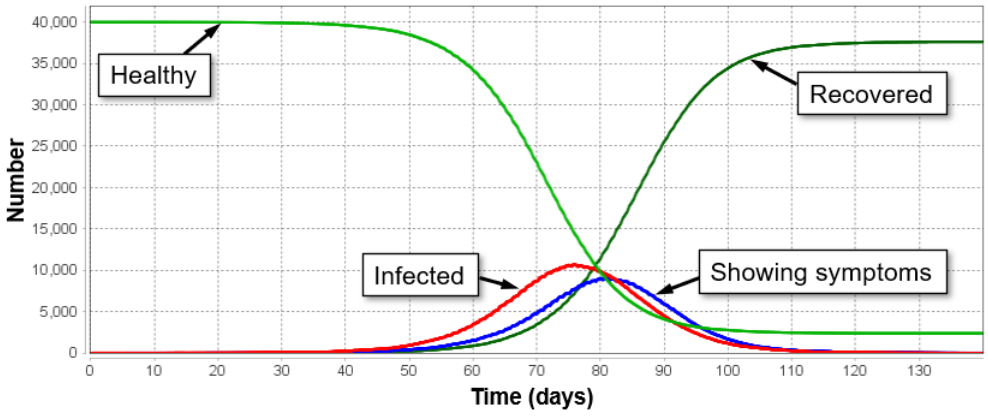


Figure 5: HISR (Healthy-Infected-Symptoms-Recovered) graph obtained from our model with 40,000 agents and calibrated with a reproduction number of R_0 ~ 2.

The results seen in Figure 5 have some similarity with graphs obtained from simpler SIR models for disease spread, defined by the following differential equations:

$$\frac{dS}{dt} = -\beta\frac{IS}{N} \qquad \frac{dI}{dt} = \beta\frac{IS}{N} - \gamma I \qquad \frac{dR}{dt} = -\gamma I \qquad (2),$$

where

- N is the population size.
- S = number of susceptible people.
- I = number of infected people.
- R = number of people who have recovered.
- β = infection rate for the disease.
- γ = recovery rate.

Our scaling factor β', used in Eq. (1), is loosely related to the infection rate, β, in the SIR model. To illustrate, solutions of Eqs. (2) are plotted in Figure 6 for $\beta = 0.015$ and γ adjusted to give a similar number of peak infected cases as our ABM in Figure 5.

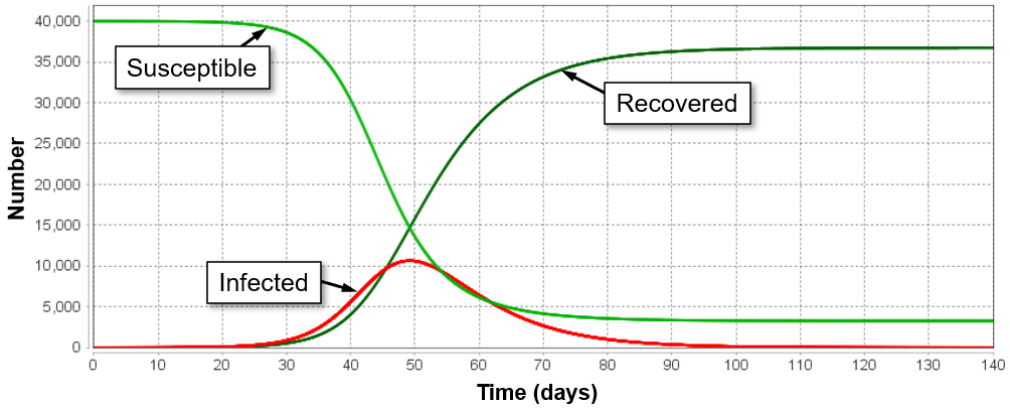


Figure 6: Results from an SIR (Susceptible-Infected-Recovered) model with β = 0.015, γ = 0.0045 and initial susceptible population of 40,000.

(In both cases, there was initially one infected person in the population, $I(0) = 1$.) It can be seen that peak infections occur sooner in the SIR model. This is because all entities are in direct contact in the SIR model, while in the ABM interactions only occur among a limited number of agents at each time step. Furthermore, since infection only occurs at local venues in our model, β' is not trivially related to R_0.

In addition to the HSIR graphs shown in Figure 5, our model is capable of producing other graphs such as shown in Figures 7 and 8. These include the dynamic R_0 versus time and the total number of confirmed cases with time. Since we have control of individual agents in our model, R_0 can be measured at each time step by applying a counter to each agent after infecting another agent and averaging these counts over the timeframe for which the infectious agents were active in the community. For example, Figure 7 shows daily R_0 values measured from a particular scenario and the way in which R_0 can change as a scenario unfolds. In this scenario, agent movements were restricted to venues within the neighbourhood of their assigned homes. This restriction causes the infection to spread out as a wave front, similar to that seen in Figure 1, and results in the average R_0 value becoming less than the initial calibrated baseline value of ~ 2. (The reduced R_0 value is due to the supply of healthy agents able to be infected being limited to the forward direction of the wave front.)

We find that plotting changes in R_0 as a scenario progresses is useful for identifying different phases of the disease spread, as indicated in Figure 7. For example, at the start of the scenario a peak typically occurs in R_0 after the first infected agent begins infecting other healthy agents in their local vicinity. During this phase, disease spread tends to be unrestrained, and the growth is exponential. This yields an R_0 value close to 2, for which the model was originally calibrated. Furthermore, there is usually a peak in R_0 within the first day of infection because households are naturally very infectious in our model and the first infected agent tends to infect all fellow household members straightaway. After a scenario gets underway, R_0 settles down to a constant value representative of the population dynamics and any restrictions that have been imposed. This value of R_0 provides a useful signature of the propagation mechanism and the degree to which geometry might be limiting disease spread.

Figure 8(a) shows a cumulative graph of confirmed cases for the population. This was obtained for a scenario with a minimum of restrictions applied. Such graphs are useful for identifying the exponential phase of disease spread after the population has first become infected, and where an intrinsic value of R_0 can be obtained. This is highlighted in Figure 8(b) where the graph of confirmed cases has been re-plotted on a log-linear scale. The exponential phase of the disease spread is clearly identified by a linear section of the graph. The slope of this section can be used as another measure of infectiousness in the population and from which the doubling time of the disease can be extracted.

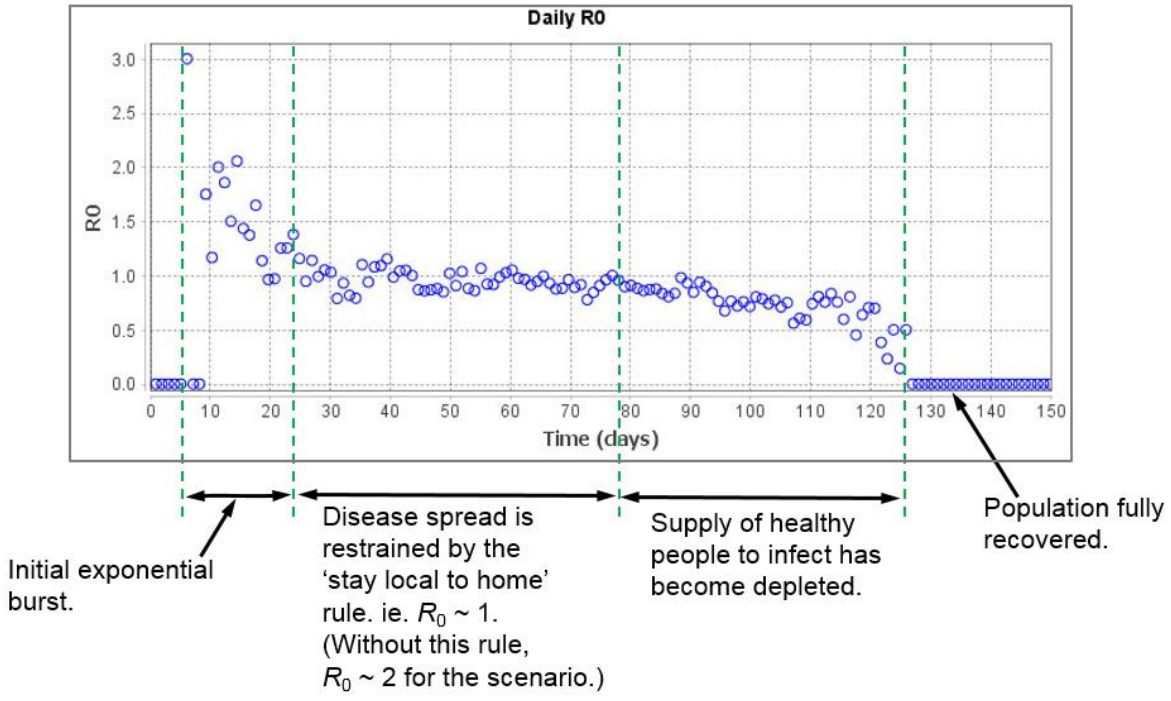


Figure 7: Example of daily R_0 values obtained from a single scenario run.

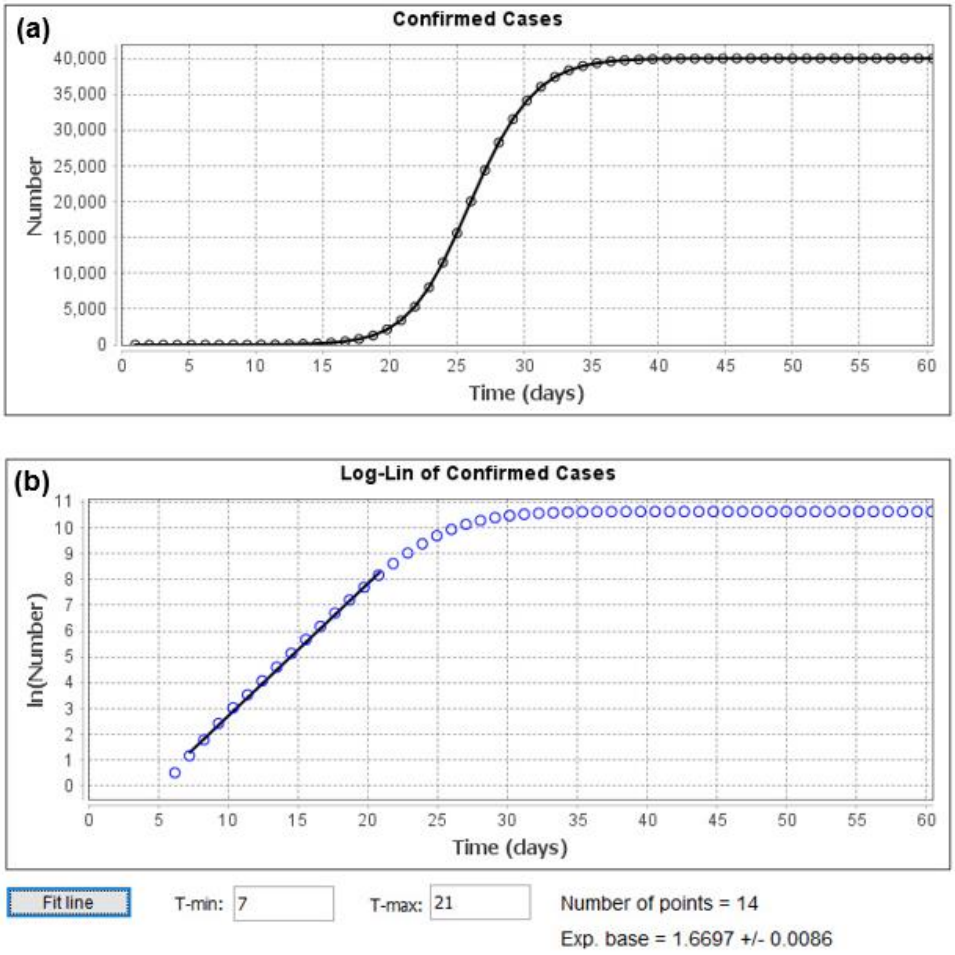


Figure 8: Sample output graphs from the ABM: cumulative confirmed cases. (a) Plotted on a linear scale in both axes. (b) Plotted on a log-linear scale.

3. RESULTS

In this section, we present results from our baseline scenario and demonstrate the general behaviour of the model. To reiterate, the baseline scenario has the following settings:

- Population size = 40,000 agents, each referenced to a point on a 200 × 200 grid.
- Home venue sizes = 4 agents.
- Daily venue sizes = 100 agents. This is intended to represent medium-sized work premises or small schools.[6]
- Agents are assigned to venues such that they will be at full capacity each day. For example, there are 400 such outside venues in this scenario.
- Agents have no restrictions on which venue they can be assigned to within the population.

This initial setup allows the virus to spread freely through the population and leads to the number of infected agents increasing exponentially with time until the supply of healthy agents to infect becomes depleted.

[6] Note that a venue size of 100 would represent a larger workplace environment in a New Zealand context. See Section 4.1.

As mentioned, this baseline scenario is calibrated so as to have an 'intrinsic' R_0 value close to 2. This was achieved by running the scenario with different values of β' and measuring R_0 during the exponential phase of disease growth. The required β' value can depend on assumptions placed into the model to begin with. Results of experimenting with two types of calibration scheme are shown in Figure 9:

- Figure 9(a): Infected agents do not self-isolate after symptoms appear, and continue to infect other healthy agents in the population for the full 14-day timeframe that they remain infectious. This is deemed to yield a R_0 value in line with an uncontrolled COVID-19 outbreak and corresponds to a 'natural' value of R_0 whereby familiarity with the virus has not yet been established and mitigating strategies have not been put in place. This is the calibration scheme used in our scenarios, and produces doubling times in agreement with those found in real populations.
- Figure 9(b): Infected agents self-isolate after 5 days when symptoms appear. This calibration scheme produces more aggressive virus spread than seen in real populations and is useful for illustrative purposes. It is unclear to what degree people would 'naturally' self-isolate in the absence of mandated mitigations, and this represents the other extreme case from Figure 9(a). Although this alternative case is useful for explorative purposes, it probably requires β' to be higher than it would be in reality. (Note also that agents who have self-isolated will still infect fellow home residents in the model.)

For each data point in Figure 9, an average value of R_0 was measured during the initial exponential phase of the population becoming infected. This timeframe can be determined by inspecting graphs of total confirmed cases such as in Figure 8 (here, between days 6 and 21). Using the calibration scheme in Figure 9(a), we have selected a value of β' = 0.015 to be used in Eq. (1) for all scenarios and this yielded R_0 = 2.2 for the baseline scenario. It is interesting to note that this β' value is relatively small and not too far above the regime for which virus spread can no longer be sustained within the model population. In this regard, the infectiousness of the virus is somewhat weak within our model and it is the inclusion of asymptomatic cases which give the virus more potency.

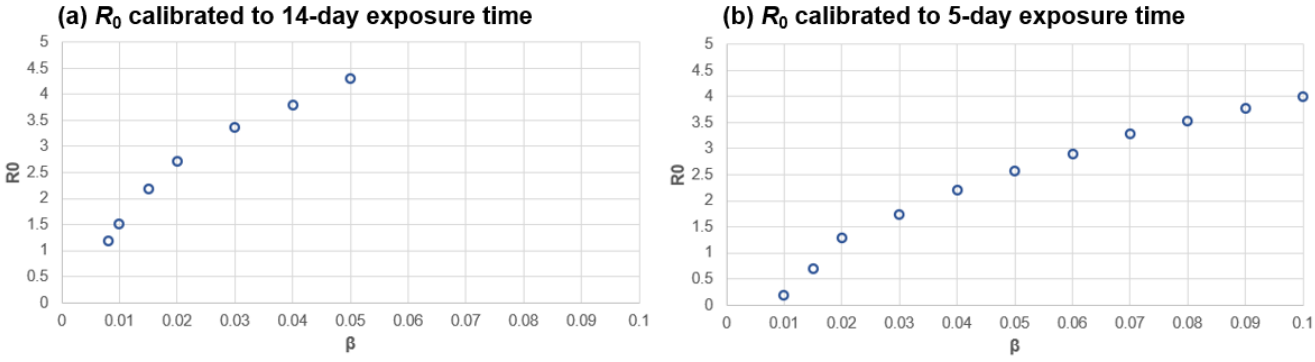


Figure 9: Calibration graphs for R_0 versus β' for two cases: (a) agents remain infectious for 14 days while spreading the virus and (b) agents remain infectious for 5 days while spreading the virus before self-isolating.

In the following sections, we present outcomes from our model with different mitigation rulesets applied.

3.1. Base-line Scenario: Full Freedom of Movement

Results from running our scenario with no restrictions on agents' movement are illustrated in Figures 10 and 11. Figure 10 shows a screenshot during the early phase of the disease spread. Despite the first infection appearing at the centre of the grid, the virus spreads quickly to other parts of the population and then sweeps through the entire population. This is further demonstrated from the corresponding R_0 graph in Figure 11 where it can be seen that the disease has swept through the entire population within about 3 months. The value of R_0 in the graph remains above 2 during the exponential phase of the disease spread, in agreement with our calibration scheme.

Tables 2 and 3 summarize overall results from our scenario if different rulesets are applied to represent different mitigation strategies. Table 2 gives the total number of agents that were infected and the peak in numbers that occurred during the virus spread. Table 3 gives the time of peak infections and the total number of days from the first infection until the final case disappears. The following list gives a description of the different interventions that were considered in our model and the rules used to implement them:

- No Intervention: This is the default scenario with a minimum of restrictions applied. The only rule applied is that any agents displaying symptoms will self-manage and remain home until fully recovered.
- Lockdown = Only One Person per Household Allowed out: A lockdown in our model is defined as only one healthy person from each household allowed out each day. If any person in the population displays symptoms, then this is applied. This continues until the number of infections returns to zero in the population. Although this is not a 100% lockdown, it represents a more realistic situation where household members would still leave the house for shopping or to work in essential services.
- Contact Tracing with Self-Isolation: If any person at a venue has developed symptoms, then all attendees at that venue are requested to remain home until all have recovered. Likewise, all fellow home residents of those attendees are contacted and requested to remain home. An assumption is that this process occurs without delay. The effect of delays in this simple form of contact tracing is explored in Section 4.2.

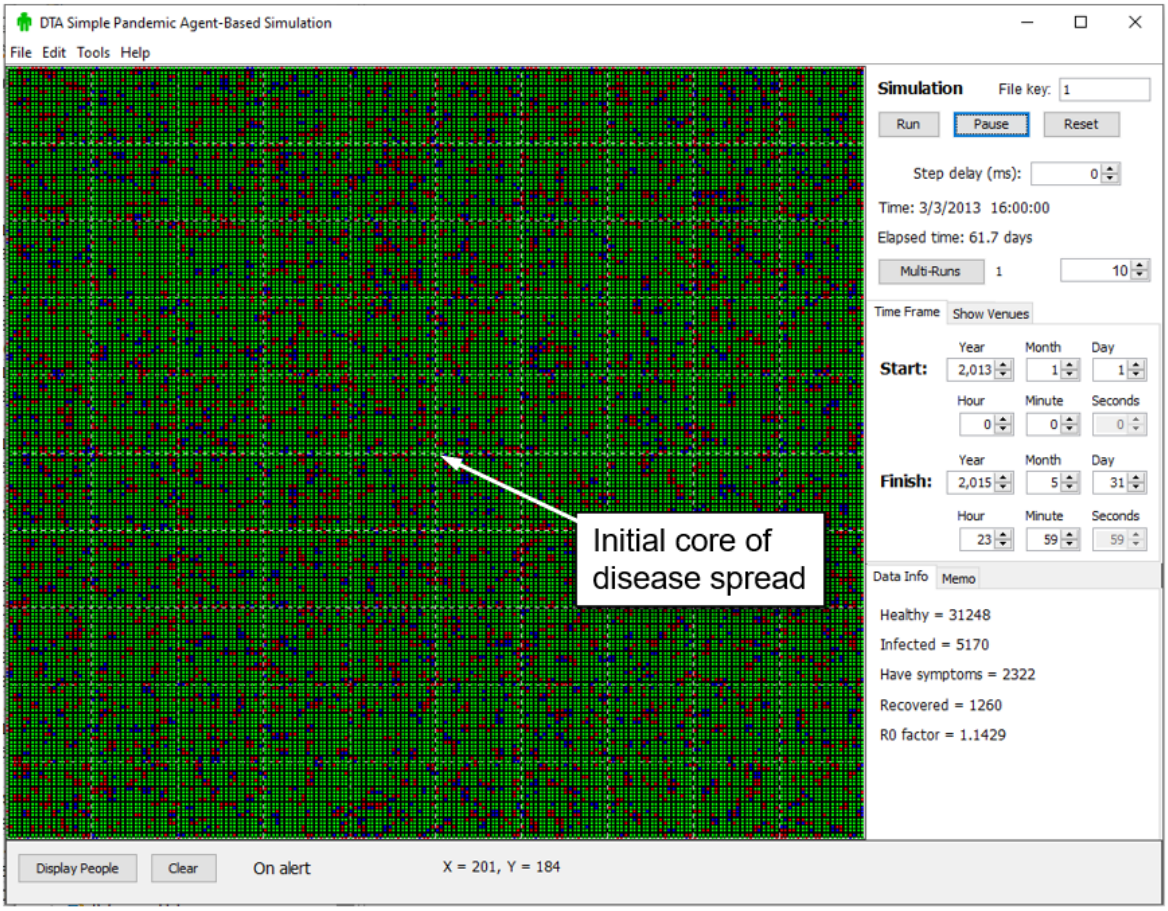


Figure 10: Initial disease spread when full freedom of movement is allowed in the population. The initial infection appears at the centre of the grid.

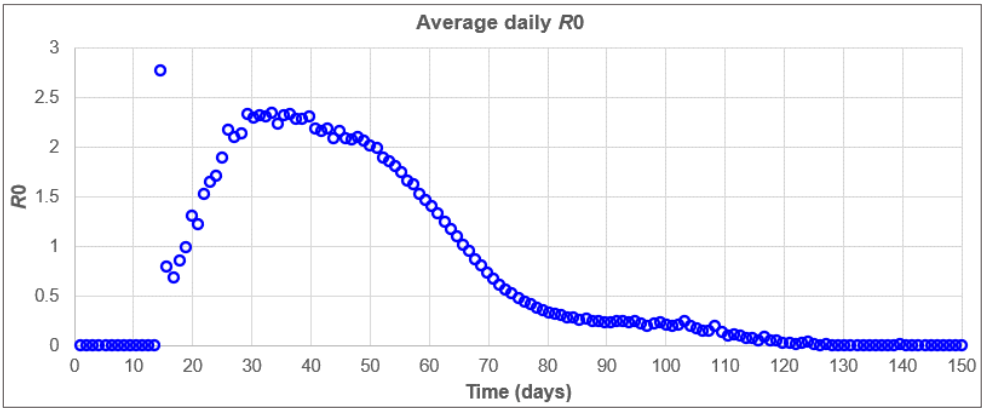


Figure 11: Average daily R_0 graph obtained when agents have full freedom of movement.

Tables 2 and 3 show average results from multiple runs of this scenario for each rule set. In each case the scenario was run 100 times, which provides reasonable statistical accuracy without excessive computing times. Also included are the standard deviation and minimum and maximum values in the results. We find that the standard deviation, minimum and maximum are useful measures for the type of distribution arising from different scenario outcomes. For example, if the standard deviation is relatively small and the minimum and maximum values are close to the average then the results tend to be normally distributed. In these cases, the scenario will most likely give similar

results each time it is run. On the other hand, cases where the standard deviation is large and of similar magnitude to the average (and where there is a large difference between the minimum and maximum values) indicate situations which are much less predictable. In these cases, there can be a wide variety of different outcomes each time a scenario is run. For example, as seen in Table 2, a lockdown might halt disease spread almost immediately with only a single household becoming infected. Other times, several hundred people might become infected before the disease spread finally comes to a halt.

	Total infected				Peak at any one time			
Rule Set	**Average**	**Std**	**Minimum**	**Maximum**	**Average**	**Std**	**Minimum**	**Maximum**
No intervention	37750	80	37584	37968	18980	220	18525	19530
Lockdown	29	35	3	170	16	15	3	75
Contact tracing	16	14	2	76	13	11	2	59
Stay local	37362	180	36867	37672	2855	250	2264	3722
Stay local + Lockdown	23	25	3	139	13	11	2	67
Stay local + Contact tracing	13	9	3	46	11	7	2	35
Stay local + Leakage = 1 %	37461	140	37015	37741	7190	520	6218	8388
Stay local + Small gatherings	17388	11000	2	29481	501	320	2	1125
Stay local + Large gatherings	37774	170	36762	38022	6994	690	4965	8458

Table 2: Results from applying different rulesets to control the spread of the virus: total number of agents in the population infected and peak number of infections during the outbreak.

	Days to Peak Infections				Days to Zero Infections			
Rule Set	**Average**	**Std**	**Minimum**	**Maximum**	**Average**	**Std**	**Minimum**	**Maximum**
No intervention	77	10	58	112	126	10	111	157
Lockdown	20	15	5	78	23	17	6	93
Contact tracing	13	4	4	32	15	6	4	35
Stay local	271	30	200	347	435	40	374	545
Stay local + Lockdown	19	13	3	85	21	15	3	88
Stay local + Contact tracing	14	6	6	50	15	6	6	50
Stay local + Leakage = 1 %	162	20	134	219	246	20	211	304
Stay local + Small gatherings	495	340	6	2270	860	540	6	2287
Stay local + Large gatherings	137	20	105	185	214	20	190	265

Table 3: As for Table 2 but recording the number of days until peak number of infections occur and number of days until no new infections occur and the virus has been eliminated.

These aspects are illustrated further in Figures 12 and 13 which show the results from Tables 2 and 3 in a graphical format. Here, the results are plotted as horizontal bar charts where the extent of the bars is given by the minimum and the maximum values from Tables 2 and 3, and average values are plotted as vertical lines. Mitigation strategies that could result in a wide variety of outcomes are made evident by the bars which extend well beyond their average values.

Referenced against our baseline scenario, several trends can be discerned from Tables 2 and 3:

- Disease spread is relatively aggressive if no rule sets are applied. The entire population becomes infected in about 3 months of scenario time. There is a huge surge in infected agents during the exponential growth phase and almost half the population is infected during the peak in this phase.
- The lockdown intervention was extremely effective in this model. On average, only about 30 agents became infected and the population is clear of the virus after about three weeks. As mentioned above, there can be a huge variation in scenario outcomes whereby, in the worst-case scenario, a few hundred agents might have been infected. We note that our lockdown has been implemented after just one person in the population has exhibited symptoms, whereas, in real populations a lockdown might be implemented after numerous people have been identified with symptoms.

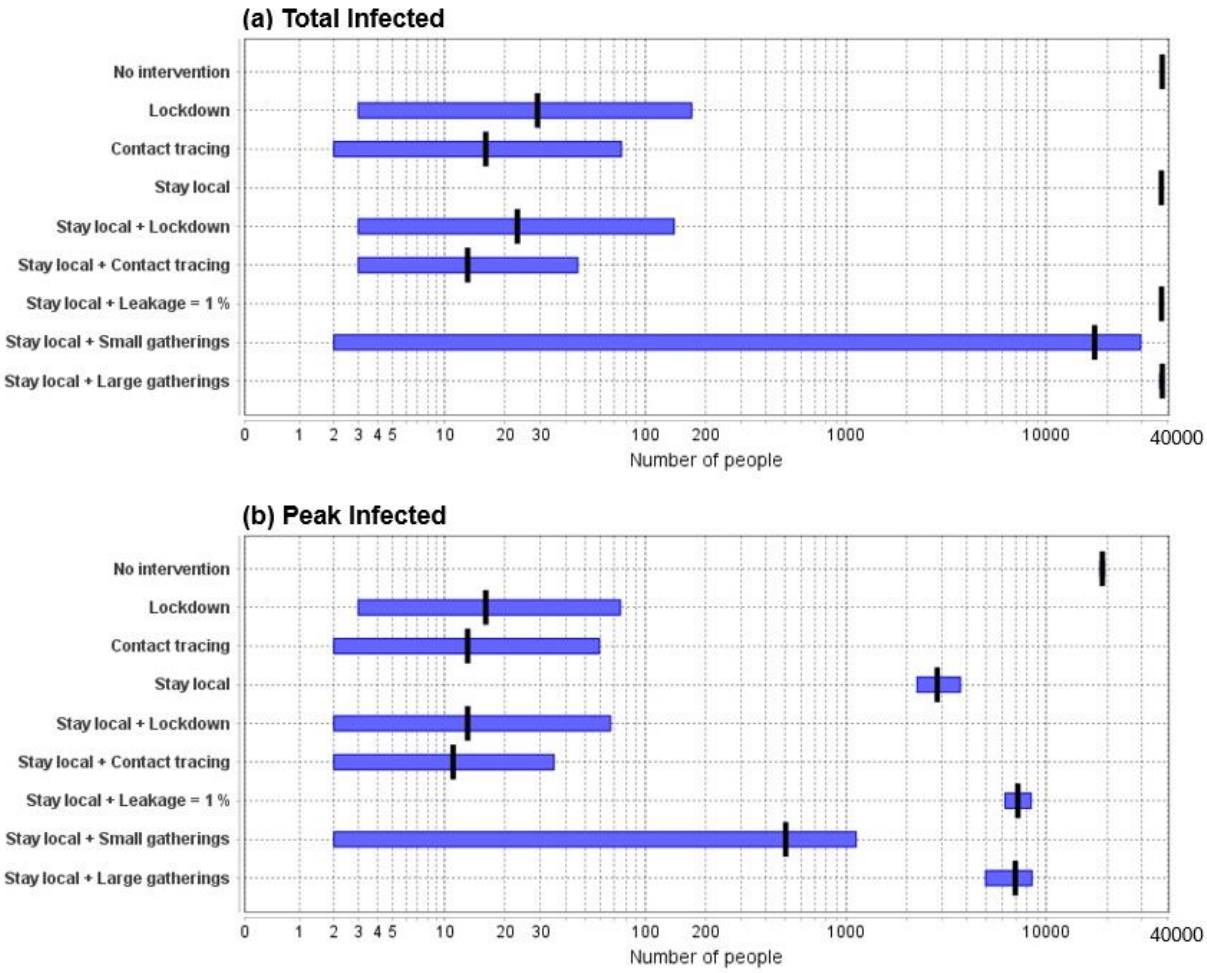


Figure 12: Graphical view of results given in Table 2: numbers of infected.

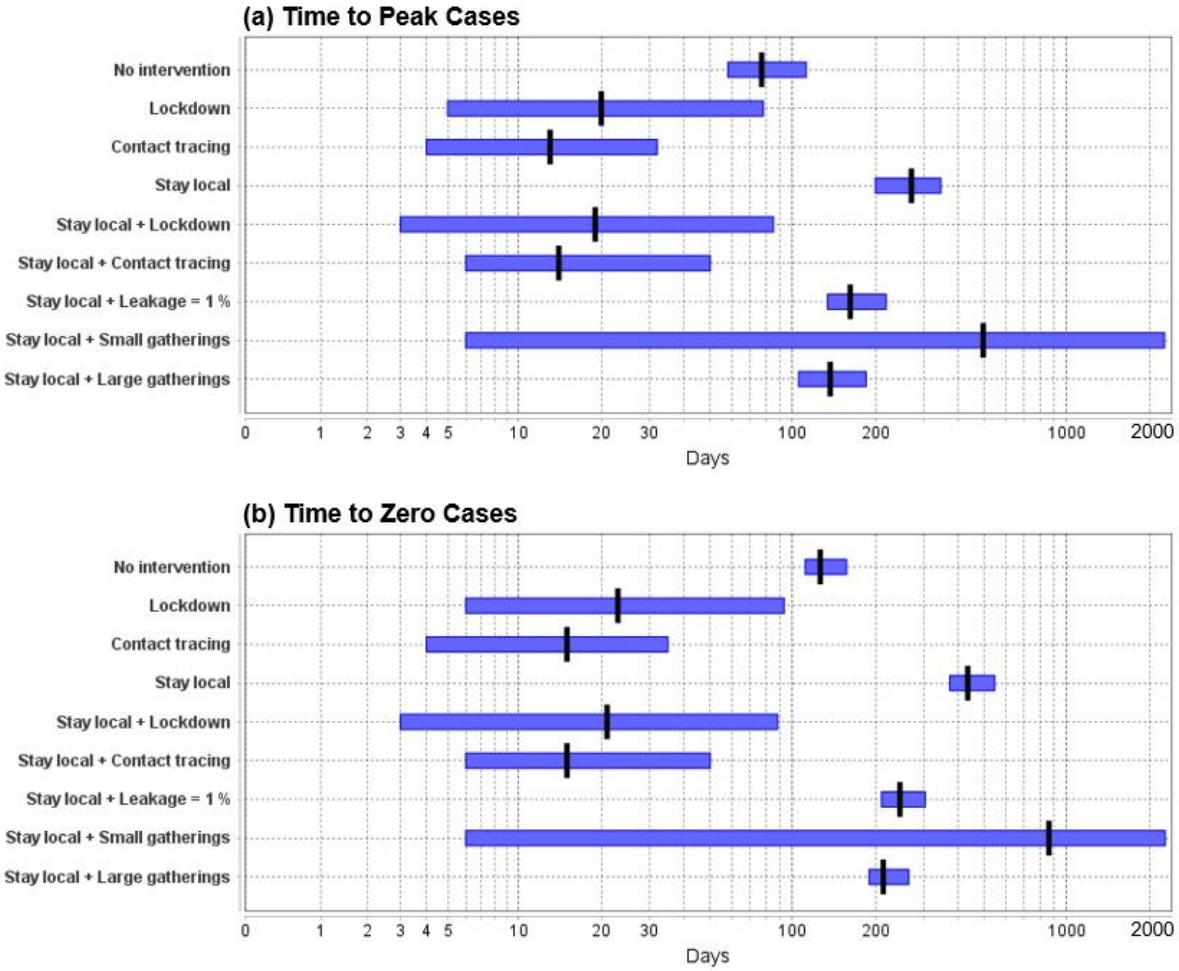


Figure 13: Graphical view of results given in Table 3: time of peak cases and zero new cases.

- Contact tracing with self-isolation was effective and achieved similar results to a lockdown. However, in our model we effectively assume that contact tracing works perfectly; that is, there are infinite resources available for this and there is no delay in tracing contacts.

The trends described above are also evident in scenario excursions explored in the following sections.

3.2. Mitigating Virus Spread via Local Movement Restrictions

In this scenario, an additional restriction, less severe than a full lockdown, is implemented in which agents' movement is restricted to venues closest to their homes, called the 'stay local' rule. The initial disease spread is shown in Figure 14 and has considerably different character to the situation seen in Figure 10 without this rule. The virus spreads out as a wave from the initial infection at the centre of the grid. At first glance, it might appear that the disease is propagating by direct contact between adjacent grid cells. However, the actual propagation mechanism is via cross-contamination between agents from different homes appearing at nearby venues together.

Corresponding measured R_0 values are shown in Figure 15 and highlight the slower spread of the disease for this scenario. The entire population still becomes infected,

but it takes more than one year of model time for the disease to sweep through. Concomitantly, the peak in number of cases is significantly reduced.

Figure 15 reveals a lower effective value of R_0 throughout most of the disease spread when compared to Figure 11. This is due to the disease spreading out as a wavefront and the supply of healthy agents being limited to those forward of the wavefront. Once the wavefront is established, $R_0 < 1$ for most of the disease spread. Established knowledge tends to consider reducing R_0 to 1 or below as a goal for controlling disease spread. On the other hand, Figure 15 highlights a case where the virus continues to spread under this condition, albeit slowly, until the entire supply of healthy agents in the population has been exhausted.

Overall results of applying these local movement restrictions, as well as lockdown rules and contact tracing, are given in Tables 2 and 3. As mentioned above, our model suggests that this 'stay local' rule will not prevent most of the population from eventually becoming infected, but the rate of disease spread is much slower compared to the baseline scenario. Similar to the baseline scenario, lockdowns and contact tracing with self-isolation are extremely effective at halting the disease spread for this case.

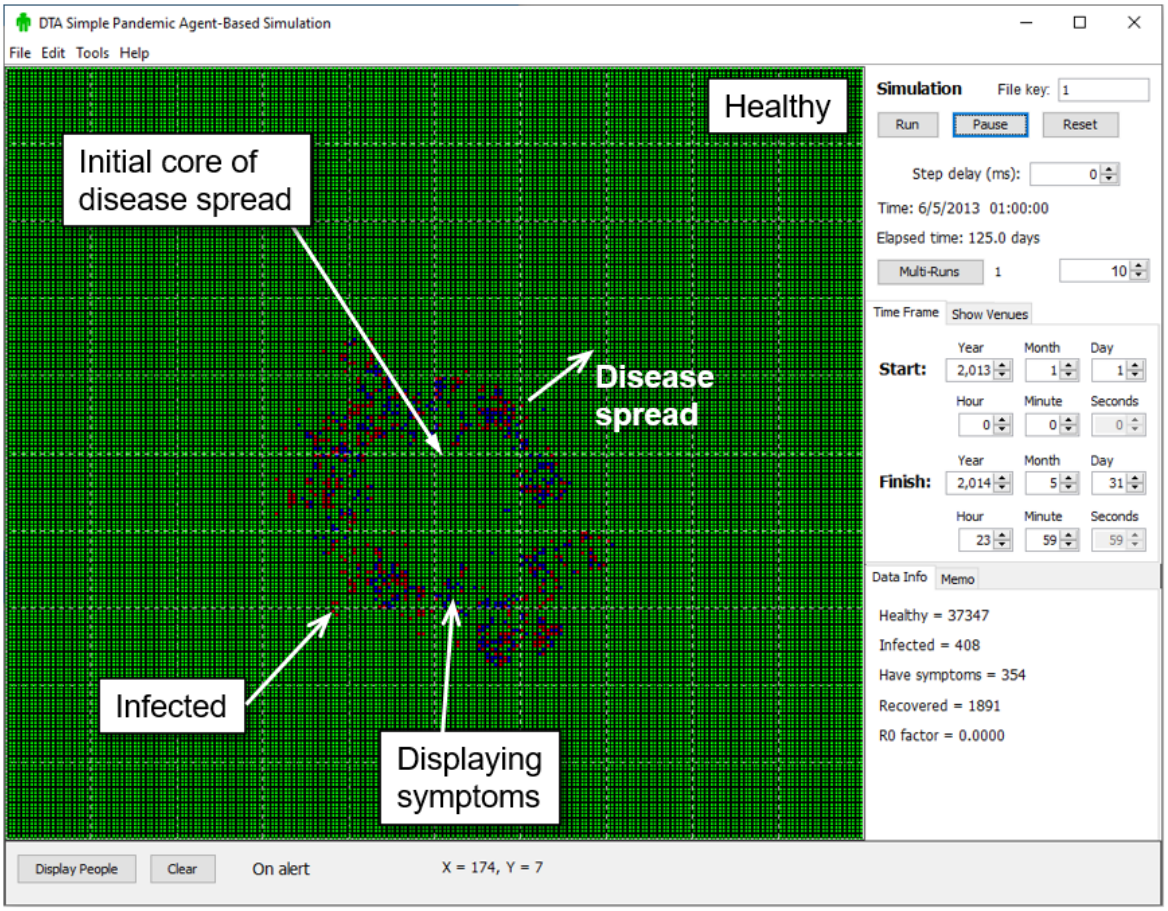


Figure 14: Initial disease spread if agents are restricted to only attending venues closest to home.

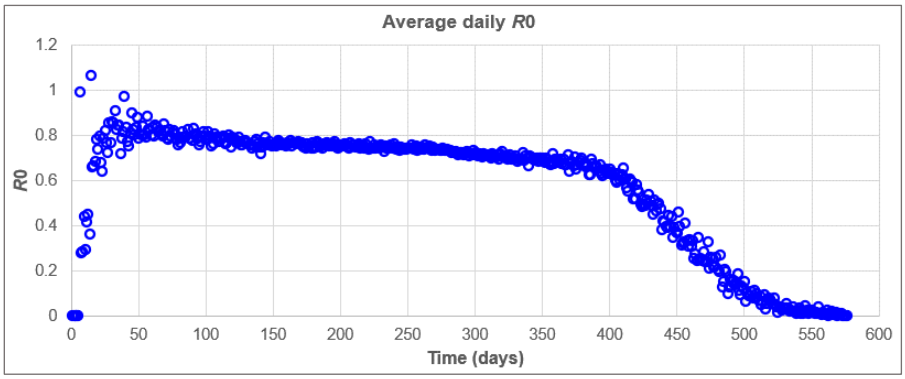


Figure 15: Average daily R0 values obtained if agents are restricted to attending venues closest to home.

3.3. Effect of Non-Compliance in Movement Restrictions

In this section we explore the effect of relaxing the rule whereby agents must only attend venues closest to home. In Figure 16, the rule was removed for 1 % of randomly selected agents in the population. Despite this low number, there has been a significant effect on the scenario outcome. The peak in infections at any one time has more than doubled and the disease now sweeps through the population in less than a year. Visually, one can identify new clusters of outbreaks appearing in Figure 16 which is not dissimilar to what one might expect to see in a real population.

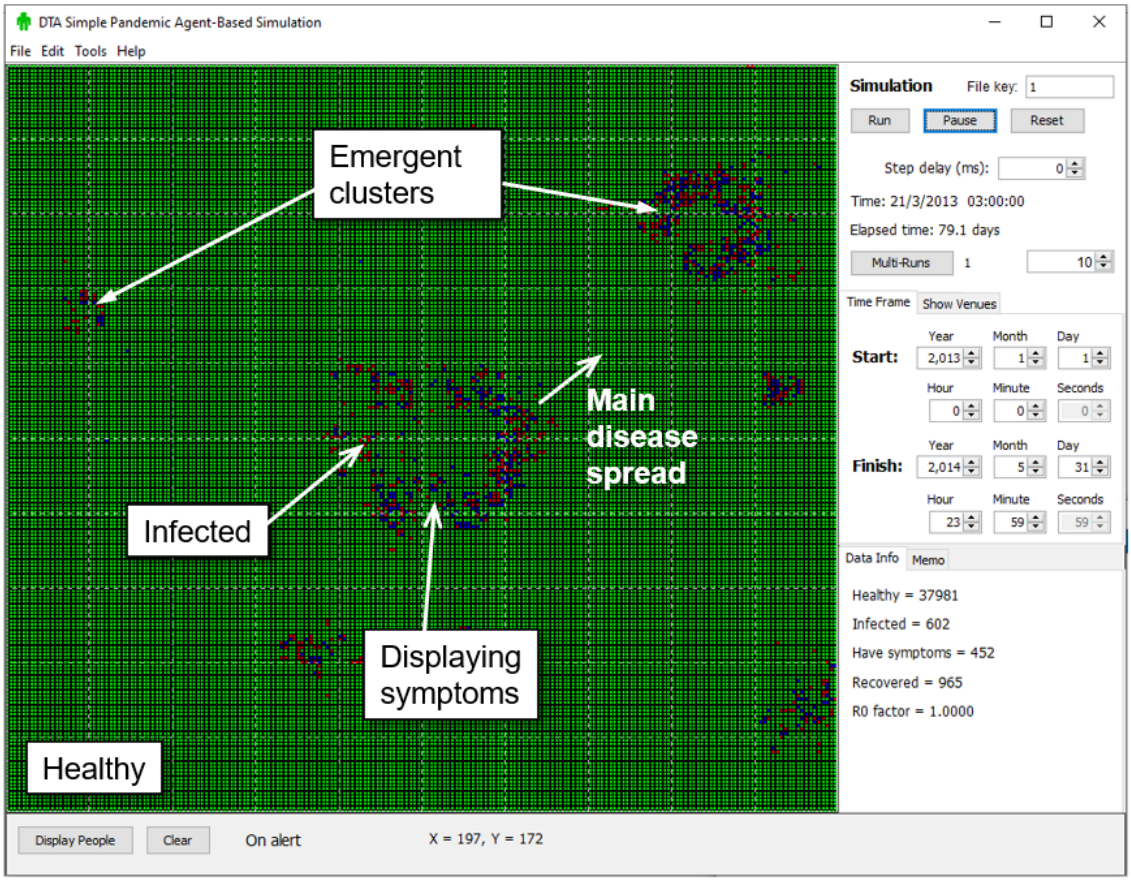


Figure 16: Initial disease spread if 1 % of agents disobey rule 'only attend venues closest to home'.

3.4. Effects of Gathering Sizes

Our baseline scenario used venue sizes of 100 agents. In this section, we briefly explore the effects of changing this number to smaller venues (size = 20) or larger venues (size = 1000) for the case that the 'stay local' rule has been applied. Results are summarised in Tables 2 and 3 in the final two rows. We note that agents are assigned to venues to their full capacity. For example, 40,000 agents assigned to venues of size 20 implies 2000 such venues in the population.

It can be seen that smaller venues (size = 20) slow the spread of the virus significantly. It takes over two years for the disease to sweep through the population and, on average, only about 500 people are infected at any one time. Only one third of the population ends up becoming infected overall, and a type of herd immunity has come about due to the lower transmission rate. By contrast, larger venue sizes speed up the virus transmission with most of the population becoming infected within 7 – 8 months.

Figure 17 shows a snapshot of the virus spread with venue sizes of 20. This illustrates the virus spreading through the population as small self-sustaining clusters. Under the given calibration scheme, propagation of the virus cannot be sustained through the population if venues are limited to a capacity of 10 agents or less.

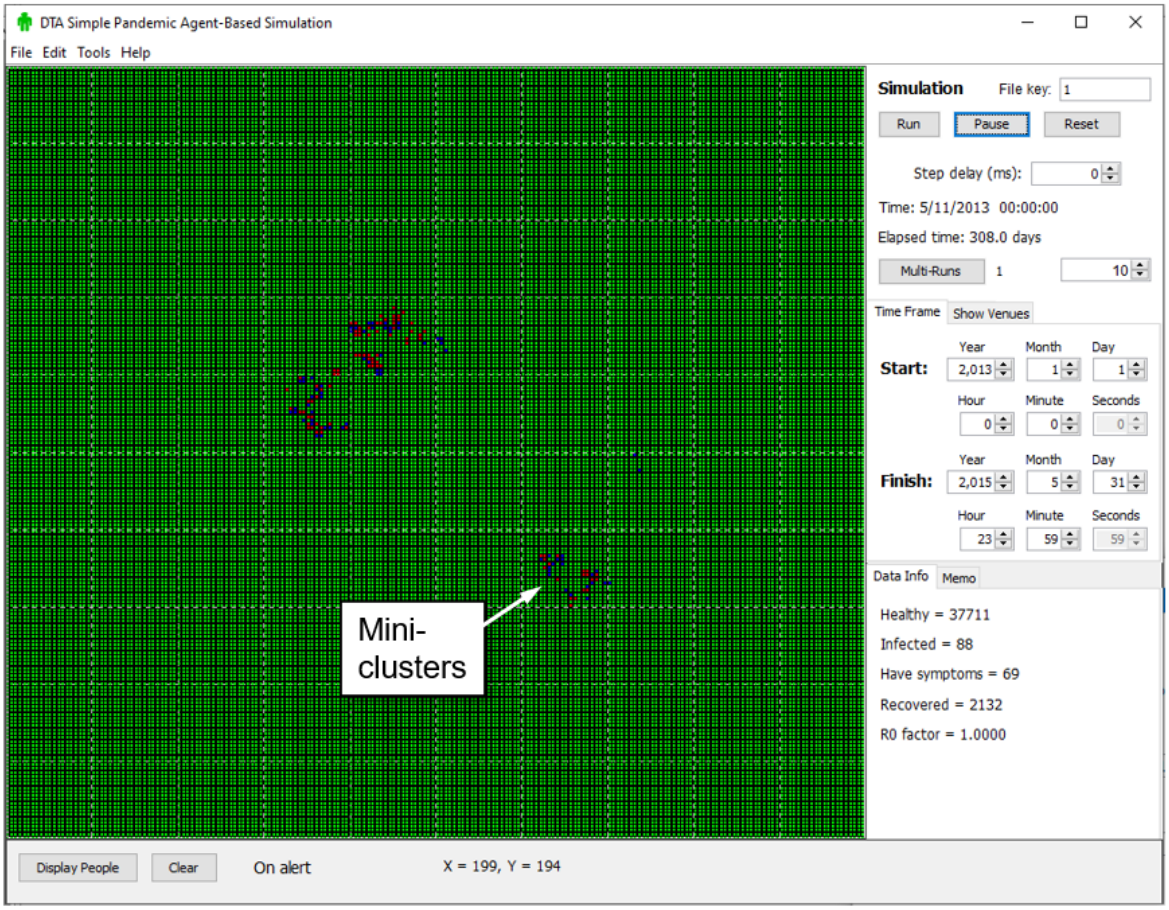


Figure 17: Disease spread with reduced venue sizes.

4. REAL-WORLD CONSIDERATIONS

4.1. Venues

In the previous section we explored the effect of smaller venues (size = 20) or larger venues (size = 1000). This revealed the extent to which larger gatherings accelerated spread, while limiting venue sizes/gatherings could be used to slow the disease.

In this section we briefly explore the effect of a distribution of venue sizes based on more realistic demographics. Various venues we could consider include homes, workplaces, schools, cafes, retail stores, movie theatres, amongst others. Our model has already been designed to include homes and workplaces. For the third type, schools have been selected as an example. This has been implemented by splitting the population into two types of agent at each home venue: (i) two adults attending a workplace and (ii) two children attending a school. This is a simple approximation to what would be a distribution of family units in a real population, but simplifies bookkeeping in our software and allows the model to run faster.[7] Furthermore, we only consider weekdays and exclude other types of venue that agents might visit related to retail or leisure. A simplifying assumption of our model is that visits to retail stores, cafes or leisure facilities can be wrapped into the venues that are already being visited by agents each day.

To obtain a distribution of venue sizes for schools, we have used the city of Christchurch, NZ as an example, with a population of approximately 380,000. The distribution of school sizes for this city is shown in Figure 18(a). It can be seen that most schools have sizes of 200 – 500 pupils and there are a few schools with over 1000 pupils. These larger sizes tend to be associated with secondary schools. This distribution can be drawn from to produce a selection of school sizes for our model.

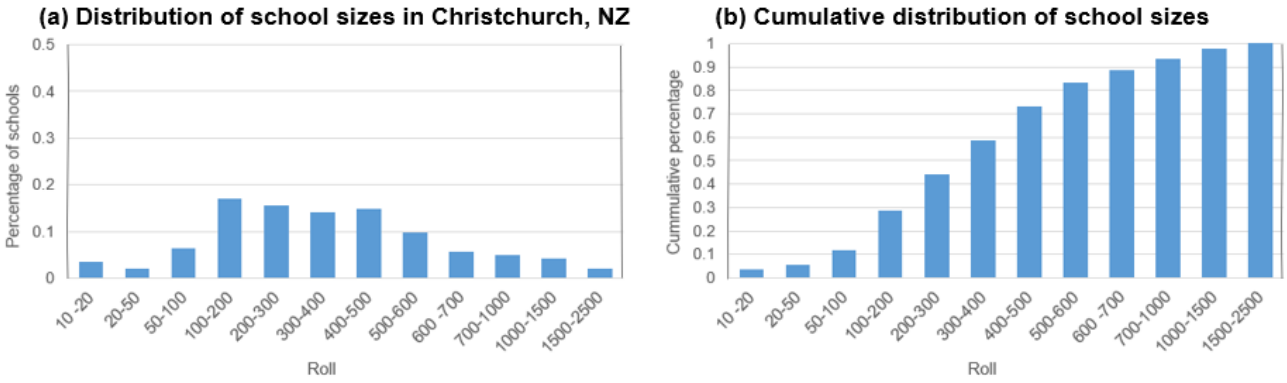


Figure 18: School size distribution for Christchurch, New Zealand. Derived from the website: www.educationcounts.govt.nz.

Figure 19 shows the overall distribution of workplace sizes in NZ, obtained from the Statistics NZ website for 2018[8]. A plot of this distribution is given in Figure 19(b). We note that a large portion of workplace venues have zero employees (~65%) and that these represent self-employed businesses. However, as a fraction of the entire working population, they only represent 17% of the work force and we do not consider these venues in our model.

[7] The model could be extended to include distributions of family units based on a real population.
[8] www.stats.govt.nz/information-releases/new-zealand-business-demography-statistics-at-february-2018

Figure 20(a) shows the distribution of non-zero workplace sizes on a log-log plot. This is found to be approximately linear, indicating that the distribution of workplace sizes follows a power law. The upshot of this is that there are many more small workplaces than large workplaces. The existence of this power law distribution is not a surprise and is a typical characteristic of complex self-organizing systems (i.e. this distribution was not 'planned', but is the result of self-organizing behaviour within the system). An index which defines the degree to which the number of venues scales with venue size can be obtained from the linear fit in Figure 20(a). This, in turn, can be used to generate our own distribution of workplace sizes, as shown in Figure 20(b).

The results of using these distributions to generate workplaces and school venues in our model are shown in Figure 21. In this case, we have used a population size of 40,000, with 20,000 agents designated as adults and the other 20,000 designated as children. We note that the scatter plots in Figure 21 appear to reflect the distributions in Figure 18(a) and Figure 20(b) reasonably well. For example, a large number of workplace venues have been generated with sizes of 10 – 20 and only a few workplace venues of size greater than 50 have been created. (Larger venue sizes would have appeared had we considered a population size greater than 40,000.) Similarly, a relatively large number of school venues have been generated with sizes of 200 – 500 and only a few school venues of size greater than 1000 have appeared. (In the case of Figure 21(b), this would represent two secondary schools in our population of 40,000.)

(a) Distribution of workplace sizes in NZ

Workplace Size	Number of Venues	Number of Employees	Inferred Average Venue Size
0	381,552	0	0
1 - 5	114,114	275,700	2.4
5 – 9	29,322	212,600	7.3
10 – 19	23,847	320,000	13.4
20 – 49	13,923	415,700	29.9
50 – 99	4,308	295,500	68.6
>100	2,835	718,400	253.4
Sum	569,901	2,237,900	

(b) Graph of workplace size distribution

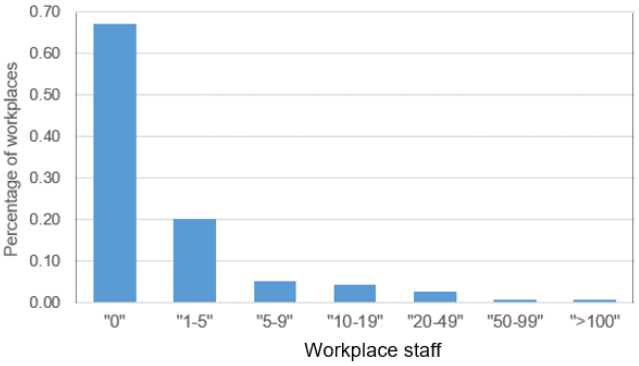
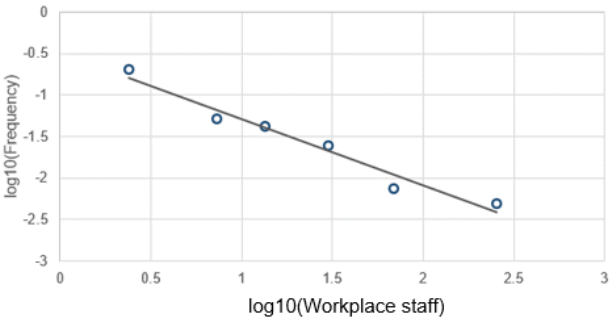


Figure 19: Distribution of workplace sizes in New Zealand.

(a) Distribution of workplace sizes on a log-log plot



(b) Distribution of workplace sizes used in the ABM

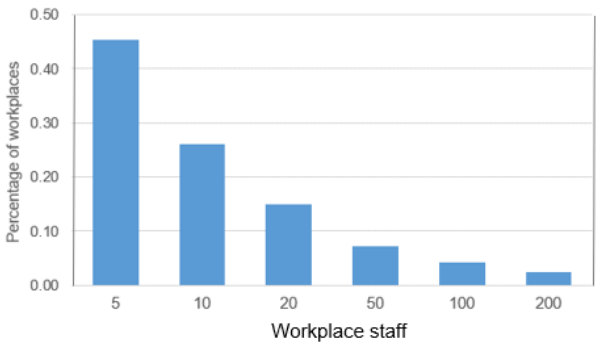


Figure 20: (a) Log-log plot indicating an approximate power law for the distribution of workplace sizes in New Zealand. (b) This power law has been applied to generate a distribution of workplace sizes for use in our model.

The results of running our model with these distributions for workplace and school venues are summarized in Tables 4 and 5. The model has first been run for the case of no movement restrictions applied to adults. Children have been defined as always travelling to school venues closest to their homes. Variants of the scenario have been run with adults restricted to only attending workplace venues closest to home and/or schools closed.

It can be seen that, without movement restrictions on adults, the disease spreads through the population very rapidly in a similar fashion to the baseline scenario presented in Section 3.1. Restricting adults to local workplaces slows the disease spread somewhat. However, the disease still spreads through the population faster than when the 'stay local' rule was applied in Section 3.2, despite the smaller average workplace sizes seen in Figure 21(a). This is because the relatively large school venues are acting as a conduit for disease spread. Alternatively, closing schools by itself does not significantly slow disease spread if adults are not restricted to their local neighbourhood.

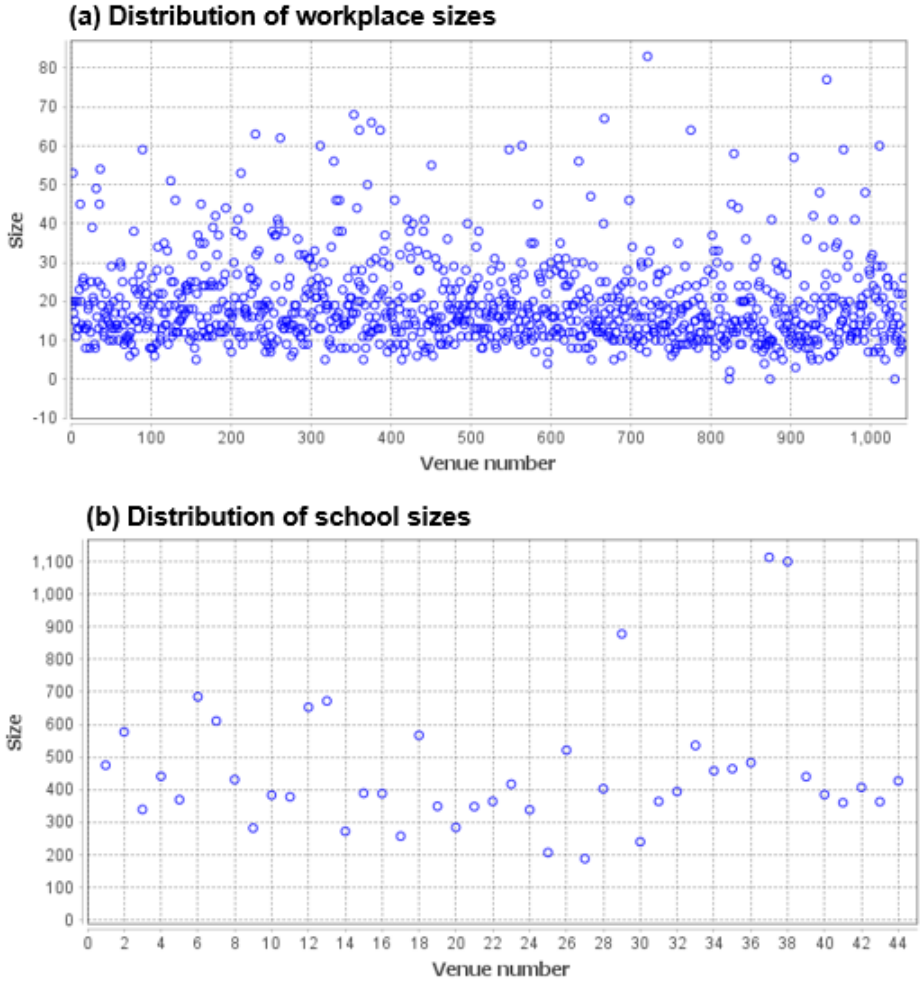


Figure 21: Resulting distribution of workplace sizes and school sizes generated from our model using the distributions in Figures 18 to 20.

	Total Infected				Peak Infected				Virus does not take hold (%)
Scenario	Average	Std	Minimum	Maximum	Average	Std	Minimum	Maximum	
Non-local work venues	37586	100	37574	37803	18057	250	17407	18706	17
Local work venues	37415	90	37184	37625	6514	680	4917	8487	12
Non-local work venues, schools closed	24824	730	22970	26061	4185	340	3413	5100	43
Local work venues, schools closed	46	60	3	283	18	15	3	64	N/A

Table 4: Results from using a distribution of workplace sizes and school sizes based on a real population: total number of agents in the population infected and peak number of infections during the outbreak.

	Days to Peak Infections				Days to Zero Infections				Virus does not take hold (%)
Scenario	Average	Std	Minimum	Maximum	Average	Std	Minimum	Maximum	
Non-local work venues	80	9	64	104	130	10	116	157	17
Local work venues	141	15	103	183	221	20	186	282	12
Non-local work venues, schools closed	177	30	134	268	295	30	240	391	43
Local work venues, schools closed	26	22	4	130	29	25	4	137	N/A

Table 5: As for Table 4 but recording the number of days until peak of infections is reached and number of days until no new infections occur.

A significant advantage comes from closing all schools and restricting adults to only attending workplace venues closest to their homes. Under the present calibration scheme, the disease propagates out as small clusters to begin with but cannot be sustained, given the small workplace venues involved on average. This suggests that assumptions in the model, such as only two types of external venues, are too simple and that other types of gathering or mechanisms of virus spread should be considered. On the other hand, this does confirm the importance of limiting gathering sizes. It should be noted that children agents are treated within the model as being as infectious as adult agents, which does not appear to have been the case for the pre-Delta variant of COVID-19 but may be applicable with new variants of concern.

A screenshot of the model running with a distribution of venue sizes as described above is given in Figure 22 for the case that adults have been restricted to venues in their local neighbourhood and schools remain open. The virus spreads out in the form of a wavefront, similar to results seen in Section 3.2. The effect of large school venues is clear in Figure 22, with schools acting as an accelerated conduit for disease spread.

In summary, the results presented in this section further emphasize the importance of small venue sizes in curtailing disease spread. The precise nature of each venue is less important in our model since we have not attempted to model different types of social interactions within either schools or workplaces. Furthermore, we have not considered any potential difference in infectiousness of children or adults. The effects of variation in infectiousness with age can be explored in due course as our model is further developed.

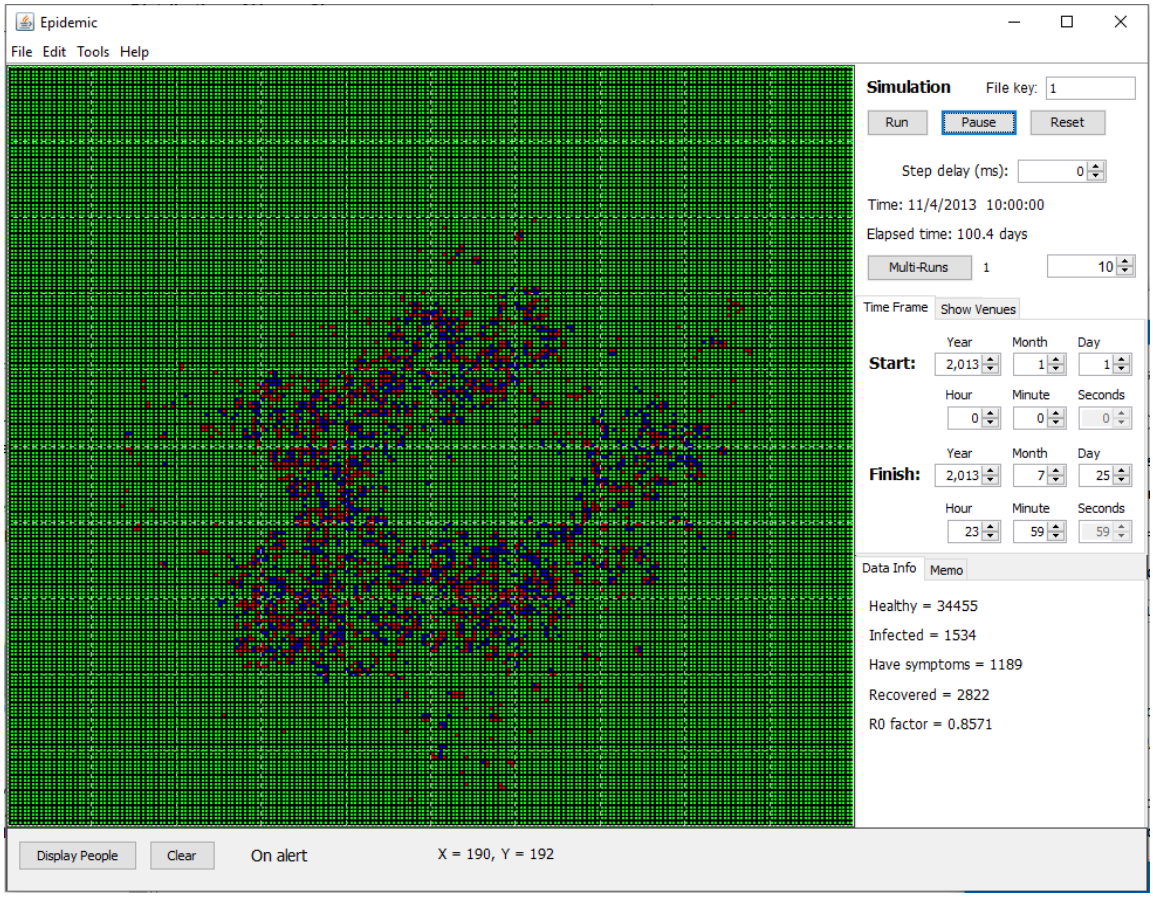


Figure 22: Example of disease spread associated with schools. The effect of relatively large school sizes is evident.

4.2. Contact Tracing with Self-Isolation

In Section 3 we implemented a simple form of contact tracing whereby all attendees at a workplace venue were told to stay home if any other attendees at the same venue had developed symptoms. Those attendees' fellow home residents were also directed to remain home. This was found to be effective at halting disease, but it was assumed that there would be zero delay between someone at a workplace developing symptoms and then all other attendees being notified.

In this section we briefly explore the effect of delays in this form of contact tracing for the case where agents have full freedom of movement in the population. Results are summarized in the graphs in Figure 23 where different delay times have been applied. (Corresponding numerical results are given in Appendix B.) These results confirm that contact tracing without delay is very effective at halting disease spread. The total number of infections in Figure 23 with zero delay is found to be < 100.

Figure 23 suggests some tolerance to delays in contact tracing and then self-isolation (this might also be related to our calibration scheme and simplified assumptions used in the model.) Delays of up to 5 days can be tolerated, after which average infections rise above 100 cases. There is a phase transition above a delay of 8 days where contact tracing fails to prevent the entire population becoming infected.

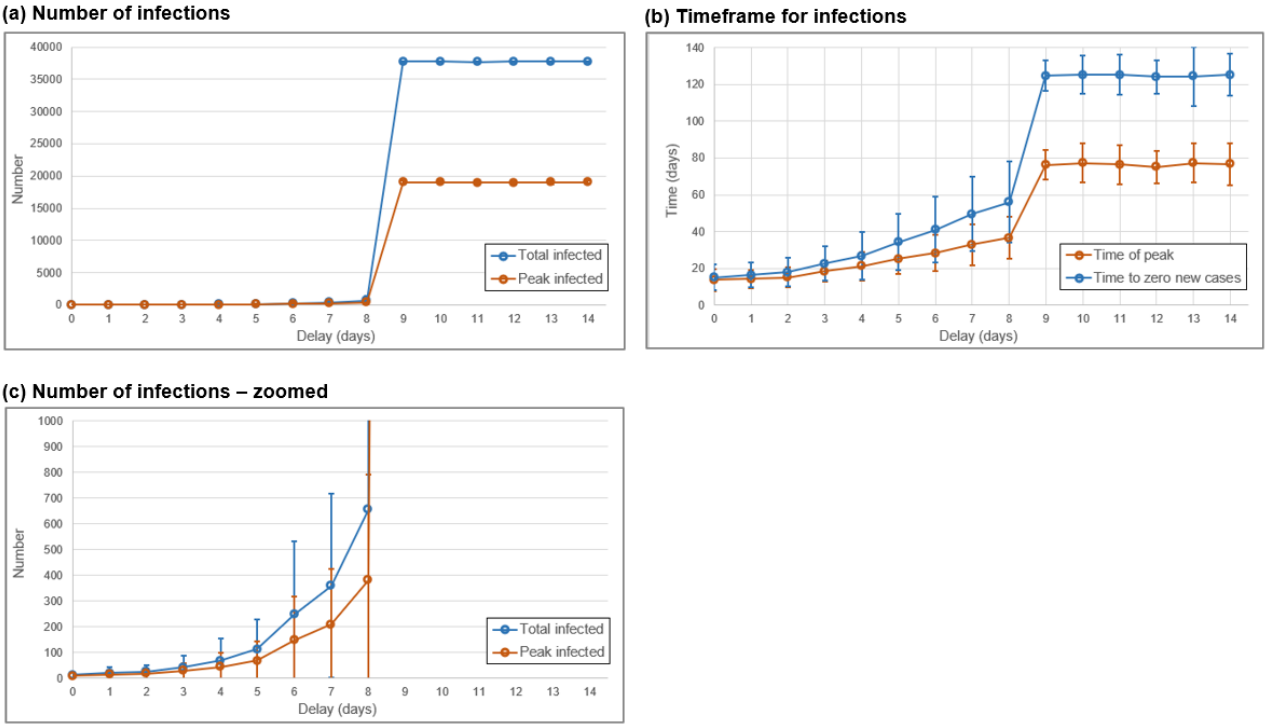


Figure 23: The effect of delays in contact tracing for pre-Delta COVID-19.

One might assume that delays greater than 5 days would render contact tracing useless, since agents already self-isolate 5 days after becoming infected and symptoms appear. However, the 5-day delay relates to the first case showing symptoms at a venue. Other agents who have recently become infected at the venue, but have yet to display symptoms, will be made aware sooner. Furthermore, the delayed contact tracing is also capturing the asymptomatic cases that would otherwise have gone on to infect others in the population.

4.3. Effect of Non-Compliance

In this section we explore what happens if a small percentage of the population do not follow lockdown rules. Results for different percentages of agents not obeying lockdown rules are given in Figure 24. As was found in Section 3.1, lockdowns would be very effective if there is 100% compliance. (Note that our definition of a lockdown still allows 1 person from each household to attend outside venues each day. This could represent essential workers.) Figure 24 suggests that there is some tolerance to people not obeying lockdown rules and that a non-compliance rate of up to 5 – 10 % could be accepted. Above this level the virus is likely to defeat the lockdown and take hold in the population. Figure 24(c) shows the likelihood that the lockdown will work in spite of higher levels of non-compliance. It can be seen that there is still some chance that the lockdown will succeed but that this decreases with increasing disobedience. Below a 10 % disobedience level the lockdown tends to be fully effective.

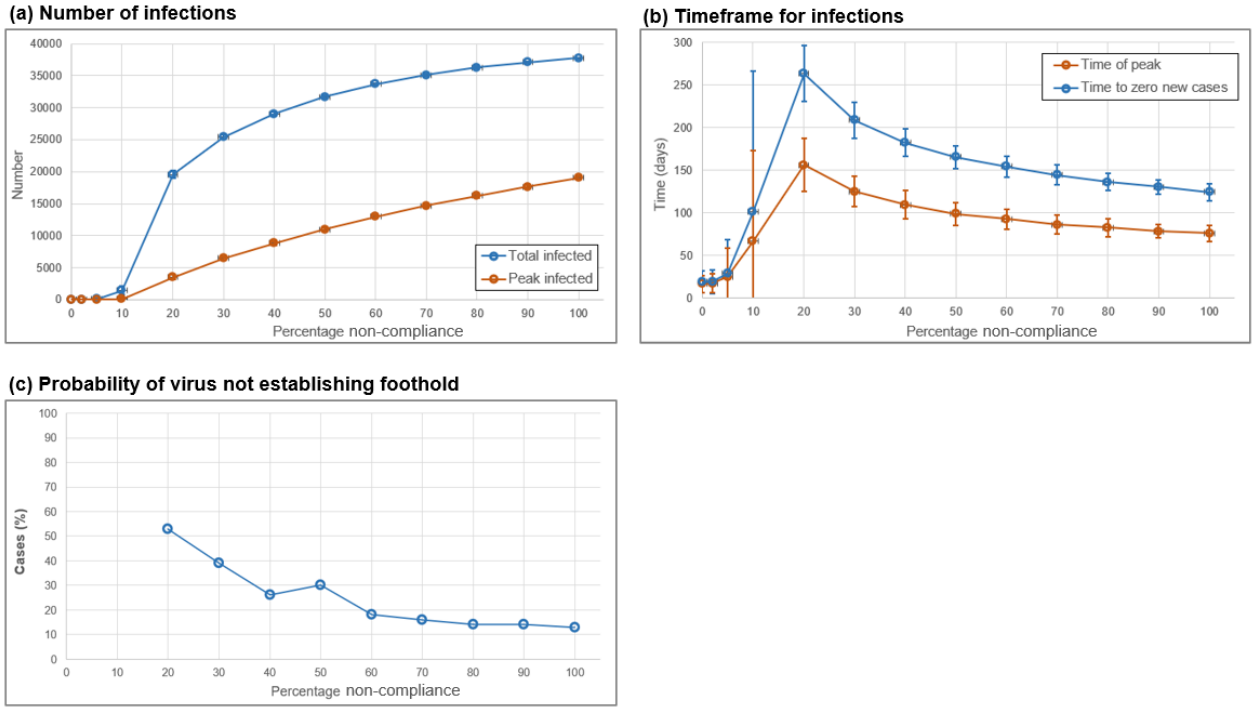


Figure 24: Effects of people disobeying lock-down requirements. (a) Number of infections. (b) Timeframe for infections. (c) Probability that the virus fails to establish itself in the population, despite the presence of non-compliant members in the community.

4.4. Working from Home versus Social Distancing

In this section we use our model to explore the trade-off between working from home versus being allowed out into the population while maintaining disease control interventions such as physical distancing and the wearing of face masks. Working from home is easily considered in our model by simply designating a percentage of agents as not visiting workplace venues each day. Effective disease-control practices, such as physical distancing or mask wearing, are more difficult to define in the model, but could be represented by changing our β' parameter for infectiousness at each venue. That is, good control/social practices could be modelled by reducing β' below its originally calibrated value of 0.015.

The effects of working from home versus good control/social practices, modelled by lowering β', are shown in Figure 25 as surface plots. These graphs plot the number of infections that occurred in the population during the outbreak, whereby each data point was obtained from 100 simulation runs. Figure 25(a) shows average results while Figures 25(b) and 25(c) show worst case results from each collection of 100 runs. All three graphs display similar trends.

Results from Figure 25 suggest that increasing the rate of people working from home does reduce infections in the population but that a significant fraction of the population must be confined to the home for this to take effect. By contrast, there is a steep drop-off in infections with decreasing β' whereby, below a certain value, propagation of the virus cannot be sustained in the population. In summary, Figure 25 suggests that good control/social practices (modelled by decreasing β') are potentially more effective than keeping some people at home, so long as there are mechanisms by which β' can be reduced in such a fashion.

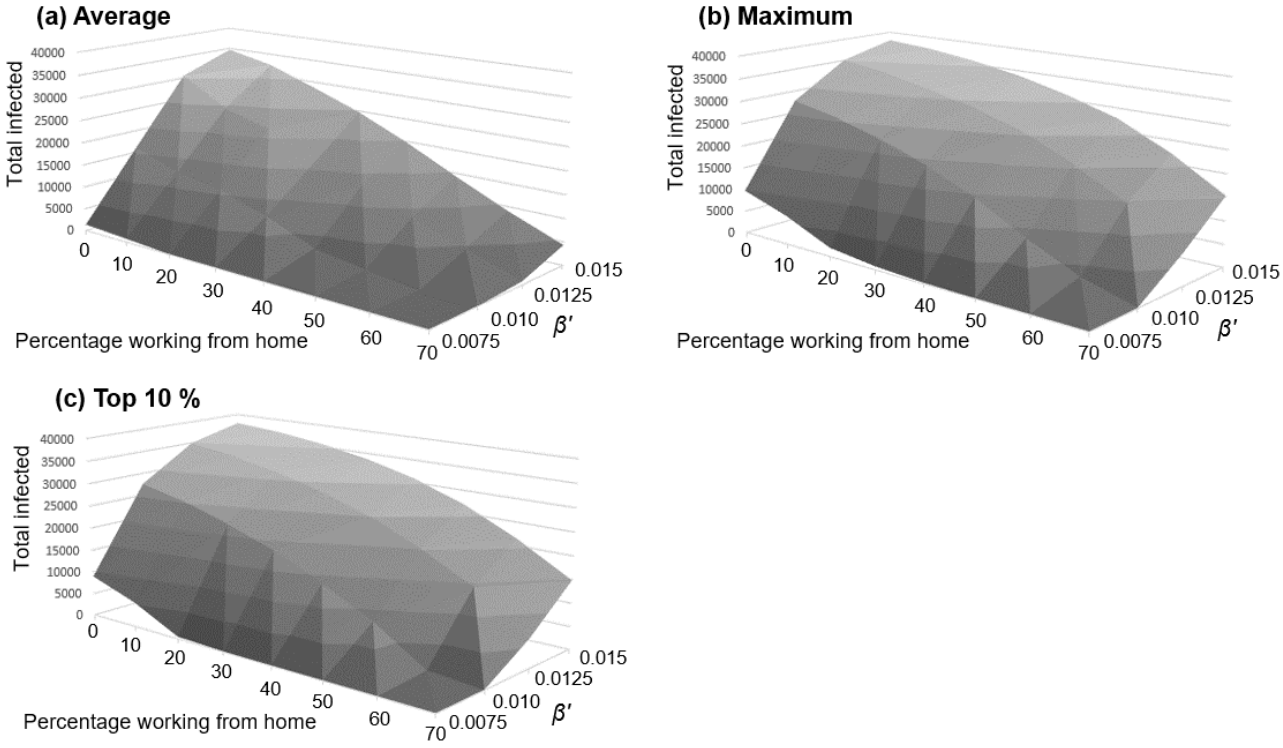


Figure 25: The relative benefits of confinement at home versus good social practices such as physical distancing and mask wearing. Good social practises are modelled coarsely by reducing the virus reproduction number β'. Each data point was obtained from 100 runs of the model: (a) average results, (b) worst case result from 100 runs and (b) top 10% worst case results. Note the relatively sharp drop-off with β', suggesting that good social practises can provide more benefit than limited confinement at home.

5. SUMMARY

We have developed a novel agent-based model (ABM) for studying COVID-19 disease spread which provides an alternative approach to standard SIR models typically employed for these types of study. The agents in our model are abstracted entities that can take on different attributes according to whether they are healthy, infected by COVID-19, and/or are displaying symptoms. The mechanism for disease spread is via healthy or infected agents appearing at prescribed venues each day where other healthy or infected agents may be present. We adopt a distillation approach whereby detailed social interactions are deliberately omitted from the model. This keeps the model manageable while still providing insights into the general dynamics of disease spread through the population. This study is based on the pre-Delta variant of COVID-19.

A key observation from the modelling was that the geometry limiting the way the population interacts appears to be a key in determining the evolution of the outbreak. This is despite the fact that the population is uniformly spread, i.e. this observation does not simply reflect some geographically unusual distribution of a human population, but rather is a property of agents interacting in two dimensional space. The advantage of our modelling approach is that such effects can be uncovered, which would not be the case for an SIR-type model.

Observing the evolution of R_0 over the timeframe for which each scenario played out was found to be useful for identifying different phases of disease spread. This further demonstrated the relationship between the shape of the outbreak and the rate of disease spread. For example, some scenarios had movement restrictions applied such that the virus propagated out as a wavefront, and these tended to have a measured R_0 close to 1.

Following from the previous point, our model suggests that the goal of reducing R_0 to 1 or slightly below is a weak condition for controlling the spread of the virus, particularly if it has already taken hold in a significant portion of the population. In various scenarios using our model, the virus was able to continue propagating well into the population under this condition. Such geometrical effects might be partly responsible for the multiple-surge character of the pandemic in real populations, and could be explored further.

Though the intention of the model is to explore the relative effectiveness of various interventions in the model – rather than to attempt to quantify their impact in the real world – we note the following findings:

- Restricting agents' movement is effective at slowing down disease spread. This does not prevent the disease from eventually infecting the entire population, but the slower infection rate could help manage the epidemic (i.e. 'flatten the curve'). In our case, a ruleset was applied that limited agents' daily movement to their local neighbourhoods.
- We found that limiting agents' movement to local areas in order to control the disease spread must be strictly adhered to. If as few as 1% of agents do not follow this rule, and travel to other parts of the population, then new clusters of infection will emerge and infected numbers in the population will increase significantly. This ignores, however, strategies that might allow some agents to travel more widely on the condition that they adopt more stringent hygiene approaches, such as strict social distancing and frequent testing.
- A lockdown (defined in our model as only one person per household allowed out each day) is extremely effective at halting disease spread. In our scenarios, less than 1 % of the population became infected if a lockdown was applied immediately after the first agent with symptoms appeared in the population.
- Contact tracing with self-isolation was also effective at halting disease spread in our model. This assumed no delay between any agent at a venue displaying symptoms and all other attendees being isolated at home.
- We further found from our model that, for the pre-Delta variant of COVID-19, a delay of up to 5 days in contact tracing could be tolerable depending on other restrictions that have been applied to the population. However, such delays are undesirable since, although the virus spread is eventually halted, many more individuals will become infected during the interim. We also did not consider the impact of resource limitations on contact tracing.
- We briefly explored the effect of some individuals not obeying lockdown rules. Depending on the severity of the lockdown our model indicates a tolerance to non-compliance of up to 5%. More than this, and the disease will take hold and sweep through the entire population.

- Limiting venue sizes (social gatherings) was effective at slowing disease spread (but not stopping it), provided other rules had been applied such as limiting agents' movement to their local neighbourhood. By contrast, large venue sizes (up to 1000 people in our model) acted as a conduit for the virus to sweep through the population extremely rapidly, to the extent that other restrictions lost their impact.

Broadly speaking, many of the above-mentioned effects have been observed in the COVID-19 response of various countries. Our model agrees that a winning strategy for halting COVID-19 disease spread, without being overly restrictive and in the absence of vaccines, would be: (i) limit gathering sizes, whether these are social gatherings or work related, (ii) limit peoples' movement to their local neighbourhoods and (iii) good contact tracing and self-isolation with minimum delays.

The work presented in this report is ongoing and suggestions for further developments are as follows:

- Our model was set up on a square grid which, for bookkeeping purposes and following the chain of infections between agents, made it easy to keep track of the venues attended by agents each day. The square grid assumption could be relaxed and geographical zones based on real cities or townships considered.
- The model could be applied to explore scenarios representing relaxation of the current border restrictions in New Zealand, most appropriately those for which some fraction of the population had been immunised.
- Our scenarios were limited to a maximum population size of about 100,000 agents due to computational resources. It would be worthwhile exploring parallel processing techniques to increase this population size. For example, a population of one million people would be suitable for studying COVID-19 spread in large cities.
- Our model was able to produce a rich set of behaviours based on relatively simple assumptions. For example, we assumed a flat infectiousness rate for all members of our population and fixed timeframes for becoming infected and then exhibiting symptoms. The extent to which statistical distributions for these parameters might change the outcomes from our model may be worth exploring. Our initial investigations have revealed that outcomes from the model are not changed qualitatively by including, for example, a Weibull distribution for the infectiousness of individuals over time, and that this merely entails a recalibration of the model in order to achieve a reproduction number for the virus of $R_0 \sim 2$. On the other hand, distributions for age variability would be worth further investigation since this could have implications for policy makers in terms of school closures.

ACKNOWLEDGEMENTS

The authors would like to thank John Kay, Brian Young, Andy Richardson from Defence Technology Agency and Pubudu Senanayake from Statistics New Zealand for helpful discussions. Also, helpful information was obtained from meetings with the USINDOPACOM FVEY COVID-19 Modelling Working Group.

REFERENCES

1. See World Health Organization COVID-19 Dashboard, website: covid19.who.int.
2. M. J. Plank, R. N. Binny, S. C. Hendy, A. Lustig, A. James and N. Steyn, A stochastic model for COVID-19 spread and the effects of Alert Level 4 in Aotearoa New Zealand, preprint, April 2020.
3. L. Ferretti, C. Wymant, M. Kendall, L. Zhao, A. Nurtay, L. Abeler-Dörner, M. Parker, D. Bonsall and C. Fraser, 'Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing', Science, Vol. 368, Issue 6491, May 2020.
4. J. Hellewell, S. Abbott, A. Gimma, N. I. Bosse, C. I. Jarvis, T. W. Russell, J. D. Munday, A. J. Kucharski and W. J. Edmunds, 'Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts', The Lancet Global Health, Vol. 8, April 2020.
5. M. K. Lauren and R. T. Stephen, 'Map-Aware Non-uniform Automata (MANA), a New Zealand Approach to Scenario Modelling', Journal of Battlefield Technology, 5(1), 2002.
6. D. P. Galligan, 'Modelling Shared Situational Awareness using the MANA Model', Journal of Battlefield Technology, 7(3), 2004.
7. E. Cuevas, 'An agent-based model to evaluate the COVID-19 transmission risks in facilities', Computers in Biology and Medicine 121 (2020) 103827.
8. P. C. L. Silva, P. V. C. Batista, H. S. Lima, M. A. Alves, F. G. Guimarães and R. C. P. Silva, 'COVID-ABS: An agent-based model of COVID-19 epidemic to simulate health and economic effects of social distancing interventions', Chaos, Solitons and Fractals 139 (2020) 110088.
9. Neil M Ferguson, Daniel Laydon, Gemma Nedjati-Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, Zulma Cucunubá, Gina Cuomo-Dannenburg, Amy Dighe, Ilaria Dorigatti, Han Fu, Katy Gaythorpe, Will Green, Arran Hamlet, Wes Hinsley, Lucy C Okell, Sabine van Elsland, Hayley Thompson, Robert Verity, Erik Volz, Haowei Wang, Yuanrong Wang, Patrick GT Walker, Caroline Walters, Peter Winskill, Charles Whittaker, Christl A Donnelly, Steven Riley, Azra C Ghani, 'Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand', Imperial College COVID-19 Response Team internal report, 16 March, 2020.
10. Pratha Sah, Meagan C. Fitzpatrick, Charlotte F. Zimmer, Elaheh Abdollahi, Lyndon Juden-Kelly, Seyed M. Moghadas, Burton H. Singer and Alison P. Galvani, 'Asymptomatic SARS-CoV-2 infection: A systematic review and meta-analysis', Proceedings of the National Academy of Sciences of the United States of America (PNAS) August 24, 2021 118 (34) e2109229118.

APPENDIX A: SETTING UP VENUES

Venues are an important component of our model since they provide the conduit through which disease is spread through the population, as defined by Eq. (1). Home venues have each been allocated 4 agents by default to represent a small household. This could be extended to represent a distribution of different households in a population. However, a scheme with 4 agents per household is advantageous on a square grid in terms of keeping track of which agents are infectious at each household. This, in turn, removes the need for search algorithms which would slow down our model execution due to the large number of households involved. Other types of venues in our model can have a distribution of sizes and are stored as container lists in software. Hence, these do require search algorithms to determine statistics such as the ratio of infected to healthy people at each venue.

Our model has a feature for visualising venues in the population according to their allocated attendance, as illustrated in Figure A.1. This shows a scheme where equal-sized venues were distributed evenly throughout the population in the form of a regular lattice. However, the periodic structure causes resonances to occur between homes and venues which tend to block disease spread. To overcome this problem, we have used an amorphous distribution of venues, as shown in Figure A.2. The amorphous distribution has been created by treating the venues as equal-sized patches in a simple 'molecular dynamics simulation'. The patches are allowed to shift their positions over several iterations until they are approximately equally spaced within the confines of the population grid. This process is carried out at the beginning of each scenario run so that randomisation can be achieved for venue distributions from one scenario run to the next.

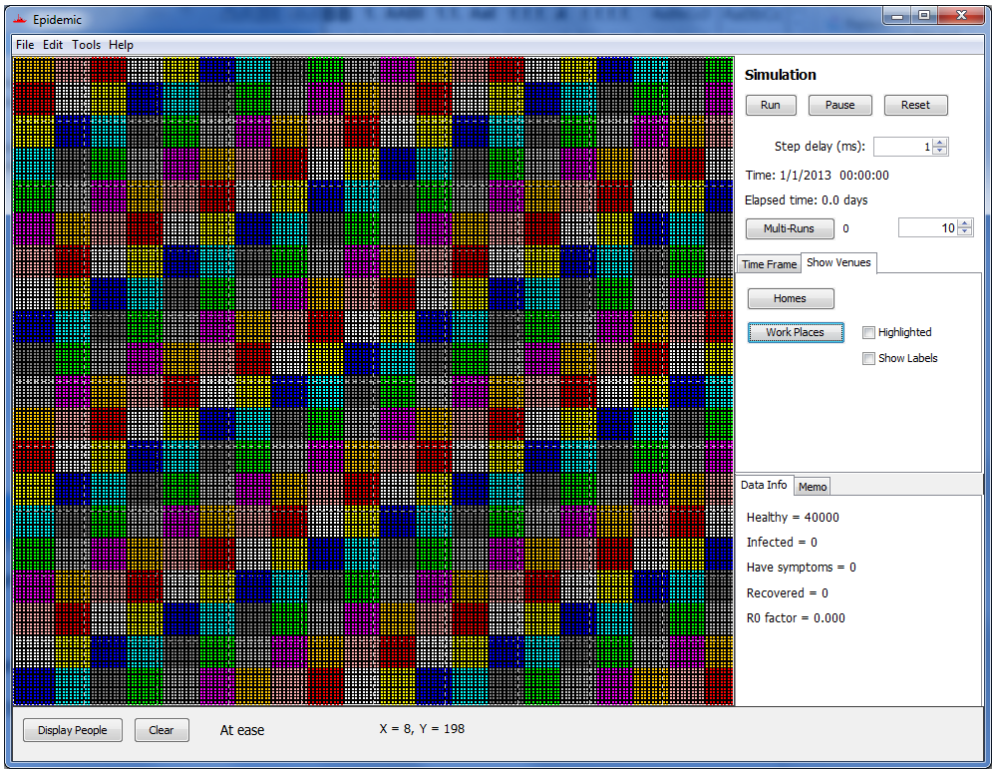


Figure A.1: Workplace venues set up on the population grid with a regular lattice distribution. In this case, each workplace has size = 100 agents.

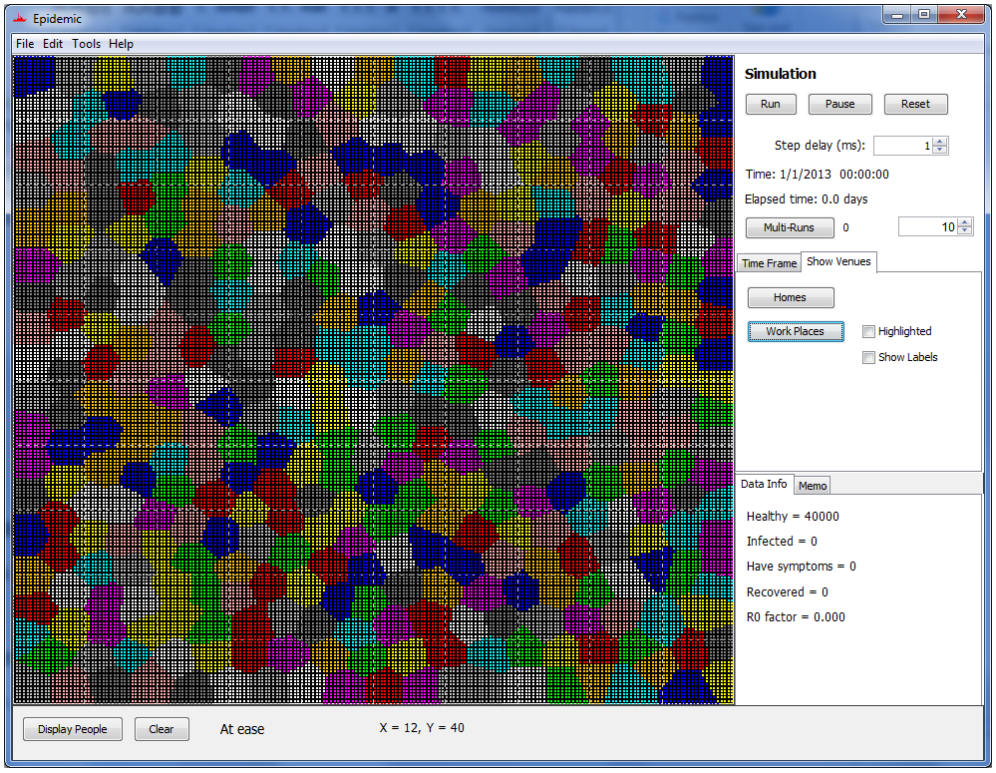


Figure A.2: As for Figure A.1 but with an amorphous distribution of workplace venues. In this case, average workplace size = 100 agents.

APPENDIX B: RESULTS – DELAYS IN CONTACT TRACING AND DEFECTORS

This appendix includes tables of numerical results used for the graphs in Section 4.2 and 4.3. Those sections explored the effect of hindrances in strategies that can limit virus spread, including delays in contact tracing and cases where a percentage of the population do not follow lockdown rules.

1. Delays in Contact Tracing

This section includes results obtained from exploring the effects of delays in contact tracing used in our model. Tables B.1 and B.2 show numerical results used in Figure 19 for the case where agents have full freedom of movement within the population.

	Total Infected				Peak Infected				Prob. virus not taking hold (%)
Delay (days)	Average	Std	Minimum	Maximum	Average	Std	Minimum	Maximum	
0	13	13	4	96	11	10	2	76	1
1	21	23	3	117	16	17	3	89	2
2	26	25	3	126	19	19	2	97	1
3	43	45	2	266	29	30	2	175	2
4	68	87	3	437	44	57	2	305	4
5	113	120	2	600	70	73	2	382	2
6	249	280	2	1637	148	170	2	980	5
7	360	360	2	1963	209	220	2	1227	4
8	655	690	2	2418	382	410	2	1412	3
9	37734	90	37526	37979	19016	240	18424	19558	16
10	37739	90	37543	38006	19045	220	18491	19601	15
11	37720	80	37498	37880	18984	260	18233	19583	9
12	37733	90	37513	37961	18969	240	18445	19451	5
13	37754	90	37552	37981	19011	260	18134	19581	6
14	37745	90	37526	37926	19041	250	18615	19673	11

Table B.1: Results from exploring the effect of delays in contact tracing. The scenario involves full freedom of movement of agents within the population.

	Days to Peak Infections				Days to Zero Infections				Prob. virus not taking hold (%)
Delay (days)	Average	Std	Minimum	Maximum	Average	Std	Minimum	Maximum	
0	14	6	4	43	15	7	4	48	1
1	14	5	3	33	17	7	3	39	2
2	15	5	4	27	18	8	4	39	1
3	19	6	2	32	23	9	2	46	2
4	21	8	3	36	27	13	3	56	4
5	25	8	5	43	34	15	5	69	2
6	28	10	2	44	41	18	2	73	5
7	33	11	1	53	50	20	1	76	4
8	37	12	4	52	56	22	4	88	3
9	76	8	60	94	124	10	107	143	16
10	77	11	60	112	125	10	108	156	15
11	76	11	61	119	125	10	107	165	9
12	75	9	63	108	124	10	106	156	5
13	77	11	61	125	124	20	107	159	6
14	77	11	63	107	125	10	107	161	11

Table B.2: As for Table B.1 but recording the number of days until peak of infections is reached and number of days until no new infections occur.

2. Defectors

This section includes results obtained from exploring the effects of different portions of the population not obeying lockdown rules. Tables B.3 and B.4 show numerical results used in Figure 20 for the case of a full lockdown.

	Total Infected				Peak Infected				Prob. virus not taking hold (%)
Defectors (%)	Average	Std	Minimum	Maximum	Average	Std	Minimum	Maximum	
0	30	43	2	272	16	18	2	119	1
2	27	40	3	208	14	15	3	75	1
5	74	190	4	1542	22	31	2	192	1
10	1454	3000	3	10305	123	250	3	915	1
20	19573	610	18109	21066	3495	360	2841	4188	53
30	25444	380	24653	26301	6431	360	5678	7109	39
40	29024	320	28238	29734	8797	340	8023	9605	26
50	31708	230	31177	32184	11010	280	10307	11557	30
60	33668	210	33205	34169	12953	240	12458	13537	18
70	35217	170	34683	35483	14660	270	14110	15348	16
80	36274	140	35929	36593	16235	270	15693	16810	14
90	37117	120	36810	37335	17666	260	17079	18500	14
100	37744	90	37549	37932	19020	220	18533	19494	13

Table B.3: Results from exploring the effect of agents not obeying full lockdown.

	Days to Peak Infections				Days to Zero Infections				Prob. virus not taking hold (%)
Defectors (%)	Average	Std	Minimum	Maximum	Average	Std	Minimum	Maximum	
0	17	10	2	62	19	12	2	62	1
2	17	11	4	59	19	14	4	72	1
5	25	31	4	288	29	40	4	320	1
10	67	110	6	487	101	170	6	569	1
20	156	30	119	285	263	40	210	396	53
30	125	20	86	169	209	20	165	262	39
40	110	20	87	171	182	20	151	240	26
50	99	15	73	136	165	20	140	204	30
60	93	15	68	127	154	20	129	184	18
70	86	15	68	137	144	20	123	196	16
80	83	10	65	113	136	10	117	167	14
90	78	10	65	98	130	10	112	160	14
100	76	10	63	108	124	10	109	158	13

Table B.4: As for Table B.3 but recording the number of days until peak of infections is reached and number of days until no new infections occur.

DOCUMENT CONTROL SHEET

1. ORIGINATING ACTIVITY Defence Technology Agency Auckland, New Zealand	2. RELEASE AUTHORISED BY: 
3. REPORT NUMBER DTA Report 464	4. CONTROL NUMBER NR 1768
5. DATE November 2021	6. NUMBER OF COPIES 2
7. SECURITY CLASSIFICATION UNCLASSIFIED	8. RELEASE LIMITATIONS UNLIMITED
9. TITLE A Novel Agent-Based Model for COVID-19 Disease Spread	
10. AUTHOR(S) G C McIntosh and M K Lauren	11. AUTOMATIC DOWNGRADING
12. KEYWORDS EJC THESAURUS TERMS	NON-THESAURUS TERMS

13. ABSTRACT

Defence Technology Agency (DTA) has developed a novel agent-based model (ABM) for investigating the spread of the pre-Delta variant of the SARS-CoV-2 virus through a population of several tens of thousands of people. People are represented in the model by abstracted entities, the so-called agents, who reside at home locations defined on a grid topology. The mechanism for virus spread is via healthy or infected agents visiting various venues on a daily basis where other infected agents may or may not be present. We adopt a distillation approach whereby detailed social interactions are deliberately omitted from the model. This keeps the model manageable while still providing insights into the general dynamics of disease spread through the population. Our ABM is intended as an analysis tool for exploring pressure points in controlling the spread of COVID-19, and to provide visualizations of the virus spread for educational purposes. Effects considered include non-pharmaceutical interventions such as limited gathering sizes, restricting movement within the population, and contact tracing with isolation. The spatial component of how the virus spreads through a community can be understood and the time evolution of the reproduction number R_0 can be monitored as a scenario unfolds. This, in turn, is useful for highlighting different phases of the disease spread. The model demonstrates the way in which R_0 does not remain constant with time, even without interventions. Moreover, geometry appears to play a role in determining R_0, which leads to the model falling into different modes of evolution. Finally, by explicitly representing agent movement in terms of travel between 'venues', the model has the potential to explore the trade-off between reducing infection rates via interventions and the adverse effects this might have on economic activity or other social aspects such as education.

INITIAL DISTRIBUTION:

	No. of Copies
NEW ZEALAND	
Director DTA	2

New Zealand
DEFENCE
FORCE
Te Ope Kātua O Aotearoa

DEFENCE TECHNOLOGY AGENCY

Devonport Naval Base.
Private Bag 32901.
Devonport, Auckland
New Zealand 0744

T +64 (0)9 445 5902
F +64 (0)9 445 5890
www.dta.mil.nz