Strengthening the Reporting of Empirical Simulation Studies (STRESS)

FACS: A geospatial agent-based simulator for analysing COVID-19 spread and public health measures on local regions

Imran Mahmood, Anastasia Anagnostou, Simon J E Taylor, David Bell, Derek Groen

Department of Computer Science, College of Engineering Design and Physical Sciences,7Brunel University, Kingston Lane, Uxbridge, London, UB8 3PH, UK

Section/Subsection	Item	Recommendation
1. Objectives		
Purpose of the model	1.1	Explain the background and rationale for the model. *** The proposed framework (i) Incorporates geospatial data, epidemiological data, disease parameters, population dynamics and demography for a given region (ii) Models the transmission of COVID-19 (or other viruses as defined in the input) (iii) Simulates the COVID-19 outbreak in a given region using computational disease dynamics at a local (e.g., city or borough) scale (iv) Forecasts the number of infectious cases spread and identifies spatial hot spots across the region, and estimates number of arrivals of the infected patients at the hospitals (v) Evaluates different lock down scenarios and analyses counterfactuals (vi) Generates early warnings for rapid emergency response management. ***
Model Outputs	1.2	State the qualitative or quantitative system level outputs that emerge from agent interactions within the ABS. Define the quantitative performance measures that the ABS model will produce, using equations where necessary. Specify how and when they are calculated during the model run along with how any measures of error such as confidence intervals are calculated *** The outputs of the model are the number of Susceptible, exposed, infectious, recovered, dead and immune population on each timestep (day) of the simulation. It also outputs: number of daily infections, number of daily ICU admissions and daily ICU bed occupancy. An ensemble of 25 runs are analysed with a confidence interval of 95%. ***
Experimentation Aims	1.3r o	If the model has been used for experimentation, state the research questions that it was used to answer.

a.) Theory driven analysis. – Provide details and reference the theories that are tested within the model.

b.) Scenario based analysis – Provide a name and description for each scenario, including a rationale for the choice of scenarios and ensure that item 2.3 (below) is completed.

c.) Design of experiments – Provide details of the overall design of the experiments with reference to performance measures and their parameters (provide further details in *data* below).

d.) Simulation Optimisation – (if appropriate) Provide full details of what is to be optimised, the parameters that were included and the algorithm(s) that was be used. Where possible provide a citation of the algorithm(s).

***

(A) Which establishments and areas are the main drivers of the spread of the epidemic within a single region?

(B) How much of an increase in admissions to hospital (ICU) can we expect when specific individual lock down measures are lifted or applied?

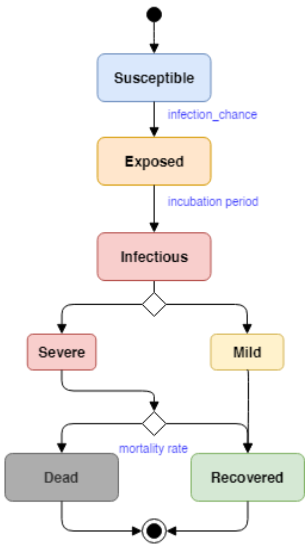
***

2. Logic

Base model overview diagram | 2.1 | Provide one or more of: state chart, process flow or equivalent diagrams to describe the basic logic of the base model to readers. Avoid complicated diagrams in the main text.

***

The state chart of the disease transmission model is shown below:



***

Base model logic | 2.2 | Give details of the base model logic. This could be text to explain the overview diagram along with extra details including ABS product and process patterns. Include details of all intermediate calculations.

***

The model starts an initial number of infected and susceptible agents. All the agents move while susceptible agents are exposed to the infection through a direct/indirect contact of infected agents during every simulation time unit (day). Infected agents recover after a period of time if the illness is mild and

		become recovered with a level of immunity. If the illness is of high severity the infected agent may recover or die (with a given mortality rate). ***
Scenario logic	2.3	Give details of any difference in the model logic between the base case model and scenarios. This could be incorporated as text or, where differences are substantial, could be incorporated in the same manner as 2.1. *** Following scenarios are simulated: No Transition Scenario Description 1 NO-MEASURES No interventions at all. 2 EXTEND-LOCKDOWN Extend London lockdown as of 15-05-2020 infinitely. 3 DYNAMIC-LOCKDOWN 50% of the population is sent to work (and associated facilities opened), when fewer than 100 ICU beds are occupied. This is checked once a week. 4 PERIODIC-LOCKDOWN Switch between the 15-05-2020 lock down situation, and 50% sent to work every 90 days. 5 OPEN-ALL Lift the lockdown entirely at a pre-set date. 6 OPEN-SCHOOLS Open all schools from a pre-set date. 7 OPEN-SHOPPING Open all shops from a pre-set date. 8 OPEN-LEISURE Open all leisure locations from a pre-set date. 9 WORK50 50% sent to work, 40% of the shops are opened. 10 WORK75 75% sent to work, schools and shops are open, 50% of the leisure locations are open. 11 WORK100 All measures lifted, except for social distancing, isolation of infected cases, and household quarantine. ***
Algorithms	2.4	Provide further detail on any algorithms in the model that (for example) mimic complex or manual processes in the real world (i.e. scheduling of arrivals/appointments/operations/maintenance, operation of a conveyor system, machine breakdowns, etc.). Sufficient detail should be included (or referred to in other published work) for the algorithms to be reproducible. Pseudo-code may be used to describe an algorithm. *** Three algorithms are proposed: In Algorithm 1, the system initializes the Ecosystem of the Simulation Core, using the input data, where each iteration evolves the ecosystem on a day basis. Algorithm 2 presents the daily evolution of the simulation. If the day reaches a transition day, the corresponding transition scenario will take into effect. Algorithm 2, present a hierarchical loop to iterate each house in the region, for each household in the house and for each person in the household. It accumulates a list of visits planned during the day as per the agent's needs. Then it processes the condition of the agent based on its current state and the time elapsed. Algorithm 3 shows the process of agent movement and possible exposure during the visit, where the person's condition is progressed in time according to the state chart mentioned above. ***

Components	2.5	2.5.1. Environment	Describe the environment agents interact within, indicating its structure, and how it is generated. For example, are agents bound within a homogeneous grid, or do they have continuous movement through a detailed landscape incorporating geographic or environmental information? *** (1) GIS based Location Graph at borough level of the city of London with GIS coordinates of different types of locations including: 'hospital', 'house', 'office', 'park', 'leisure', 'school', 'supermarket', 'shopping'. ***

		2.5.2. Agents	List all agents and agent groups within the simulation. Include a description of their role in the model, their possible states, state transitions, and all their attributes.
			***
			Initial population:
			{Infected=20, Susceptible=Borough population, Recovered=0}
			States and transitions are described in the state-chart mentioned above.
			***
			Describe all decision-making rules that agents follow in either algorithmic or equation form. Where relevant authors should report:
			
			• Disease parameters for Covid-19:
			  ○ Infection rate = 0.07
			    ▪ When one infectious and one susceptible person spend 24 hours in a 4 m^3 area, the likelihood of the susceptible person to become infected is equal to the infection rate (0.07, or 7%).
			  ○ 5 day incubation period.
			    ▪ Patients are not infectious in the first 3.5 days of the incubation period (they are in the last 1.5 days).
			    ▪ We use a Poisson distribution to mimic the variety
			  ○ Two variations:
			    ▪ Mild, no hospitalisation, 14 days to recovery
			    ▪ Non-mild, hospitalisation, 20 days to recovery (if surviving)
			  ○ 12 day average period to hospitalisation
			  ○ Full details are in Appendix 2.
			  ○ Mortality rate is age-dependent

2.5.3. Interaction Topology		Describe how agents and agent groupings are connected with each other in the model define: • with whom agents can interact, • Interacting agents: infected and susceptible • how recipients of interactions are selected • When one infectious and one susceptible person spend 24 hours in a 4 m3 area, the likelihood of the susceptible person to become infected is equal to the infection rate • what frequency interaction occurs. • Every simulation time unit • How agents handle and assign priorities to concurrent events • No priorities (random execution of actions scheduled for the same time unit) It is recommended that interactions are described using a combination of equations pseudo-code and logic diagrams. Report how interactions are affected by agent states and the environment state
2.5.4 Entry / Exit		Where relevant, define how agents are created and destroyed in the model. *** All agents are created at initialisation. They are not destroyed however they change state. ***

3. Data

Data sources	3.1	List and detail all data sources. Sources may include: • Interviews with stakeholders • samples of routinely collected data, • prospectively collected samples for the purpose of the simulation study, • public domain data published in either academic or organisational literature. Provide, where possible, the link and doi to the data or reference to published literature. All data source descriptions should include details of the sample size, date ranges and use within the study. *** Three main data sources are: 1. **Location Graph using Open Street Map data** Description: This dataset is used to initialize locations of different types including: hospital, house, leisure, office, park, school, shopping, supermarket within a borough. Sample size: ~150,000 – 300,000 locations Date range: N/A 2. **Demographic data from the Office for National Statistics** Description: Used for the input of different demographic patterns e.g., age distribution of the borough population, age wise

		movement patterns. Sample size: Age distribution between 0-90 Date: N/A 3. **COVID-19 case-study validation data from National Health Services (NHS), North West London:** Description: The validation data encompasses the borough of Brent Sample size: 60 days Date: Data collected for the month of March2020 and April2020. ***
	3.2	Provide details of any data manipulation or filtering that has taken place before its use in the simulation, e.g. interpolation to account for missing data, removal of outliers or filtering of large-scale data. *** None. ***
Input parameters	3.3	List all input parameters in the model, providing a description of each parameter and the values used. For stochastic inputs provide details of any continuous, discrete or empirical distributions used along with all associated parameters. Where applicable define the time/spatial dependence of parameters and any correlation structure. Clearly state: • Base case inputs • Inputs used in experimentation, where different from the base case. • Where optimisation or design of experiments has been used, state the range of values that parameters can take. Where theoretical distributions are used, state how, these were selected and prioritised above other candidate distributions. *** Model parameters: (a) Infection rate = 0.07 (b) Incubation Period = 3.5 days (c) Illness Duration: Mild, no hospitalisation, 14 days to recovery Severe, hospitalisation, 20 days to recovery (if surviving) (d) Time period to hospitalisation = 12 days (average) (e) Mortality rate = 1.6% (f) Mortality per Age = ~ 0 - 0.134 (g) Mortality period: 8 days (h) Recovery period: 8 days (i) Mild recovery period: = 8.5 days (7-10 days) (j) Period to hospitalisation: 12 days ***
Assumptions	3.4	Where data or knowledge of the real system is unavailable, state and justify the assumptions used to set input parameter values and distributions; agent

		interactions or behaviour; or model logic. *** When an susceptible person resides at a location where infectious persons also reside on the same day, then the probability of the infection is given by $\mathbb{P}_{inf} = \left(\frac{LSs}{OD_{loc}} \times \frac{LSi}{A_{con}}\right) \times \frac{CR}{360} \times M$ ***
4. Experimentation		
Initialisation	4.1	State if a warm-up period has been used, its length and the analysis method used to select it. 30 days warm-up period. This length is selected to provide ample time for initial serious cases to conclude, and reduce the bias in disease stages at the start of the actual simulation. State what if any initial agent and environmental conditions have been included. For example, the initial agent population size, agent states and attributes, initial agent network structure(s), and resources within the environment. Report whether initialisation of these variables is deterministic or stochastic. *** Deterministic initialization of the agent population is based on location graph and demographic data ***
Run length	4.2	Detail the run length of the simulation model and time units. *** Typical runs are 180 days to 2 years. ***
Estimation approach	4.3	State if the model is deterministic or stochastic. If the model is stochastic, state the number of replications that have been used. If an alternative estimation method has been used (e.g. batch means), provide full details. *** Stochastic with 25 replications. ***
5. Implementation		
Software or programming language	5.1	State the operating system and version and build number. State the name, version and build number of commercial or open source ABS software that the model is implemented in. State the name and version of general-purpose programming languages used (e.g. Python 3.5.2). Where packages, frameworks and libraries have been used provide all detailed including version numbers. *** It can run on any OS FACS ABM code, distributed under BSD 3-clause license, version 1.0. Python 3.8 + FabSim3 + FabCovid19 for remote supercomputer ***
Random sampling	5.2	State the algorithm or package used to generate random samples within the software/programming language used e.g. Mersenne Twister or Java.Random version x.y *** Python/NumPy, both using Mersenne Twister.

		***
Model execution	5.3	If the ABS model has a time component, describe how time is modelled (e.g. fixed time steps or discrete-event). State the order of variable updating within the model. In time-stepped execution state how concurrent events are resolved. If the model is parallel, distributed and/or use grid or cloud computing, etc., state and preferably reference the technology used. For parallel and distributed simulations the time management algorithms used. If the HLA is used then state the version of the standard, which run-time infrastructure (and version), and any supporting documents (FOMs, etc.) *** Fixed time steps Random execution of agent actions ***
System Specification	5.4	State the model run time and specification of hardware used. This is particularly important for large scale models that require substantial computing power. For parallel, distributed and/or use grid or cloud computing, etc. state the details of all systems used in the implementation (processors, network, etc.) *** Runtime: ~45 min ***