

Coupling micro and macro dynamics models on networks: Application to disease spread

Arnaud Banos, Nathalie Corson, Benoit Gaudou, Vincent Laperrière, and Sébastien Rey Coyrehourcq

UMR Géographie-cités, CNRS, Paris, France

LMAH, University of Le Havre, France

UMR 5505 IRIT, CNRS, University of Toulouse, France

UMR 7300 ESPACE, CNRS, University Nice Sophia Antipolis, France

Contact author: benoit.gaudou@ut-capitole.fr

Abstract. A hybrid model coupling an aggregated equation-based model and an agent-based model is presented in this article. It is applied to the simulation of a disease spread in a city network. We focus here on the evaluation of our hybrid model by comparing it with a simple aggregated model. We progressively introduce heterogeneities in the model and measure their impact on three indicators: the maximum intensity of the epidemic, its duration and the time of the epidemic peak. Finally we present how to integrate mitigation strategies in the model and the benefits we can get from our hybrid approach over single paradigm models.

Keywords: hybrid model, ODE, metapopulation, network, disease spread

1 Introduction

The modelling of socio-environmental processes often requires to couple processes defined at distinct temporal and spatial scales. Models involving a single paradigm to describe all the processes, such as either an aggregate approach at the macroscopic scale (with a system dynamic approach) or a totally disaggregated microscopic approach (with an agent-based approach), fail to describe multi-scale phenomena. The aim of this article is thus to introduce a hybrid model coupling micro and macro dynamics models and to assess step-by-step the impact of the heterogeneity and the stochasticity introduced by microscopic model.

As a case study, we consider a disease spread in a network of cities. The two main dynamics we take into account are the disease spread within each city and its spread between cities through air traffic. The former dynamics involves a large number of people in each city (millions of people) and its temporal scale order is of several weeks. In contrarily the latter one involves few (hundreds) of people in each airplane, and a flight temporal scale order is of several hours.

The heterogeneities that will be studied in the model are related to the initial population distribution, heterogeneity in the network (we will experiment various network topologies) and heterogeneity in flight duration to take into account the space size.

The classical approach to model such phenomena is named *metapopulation approach*, comes from ecology [8] and has been applied to various fields and in particular to disease spread in large-scale networks [1,11,3]. The main idea is that each node of the network has a dynamic described using a system dynamics approach (often described using an Ordinary Differential Equation (ODE) system). In addition edges represent migration (modelled as instantaneous streams) between nodes. Such an approach allows modellers to study disease spread conditions and to test various spread mitigation strategies [11].

We argue that this approach is too limited to take into account strategies dealing with individual behaviours. In this article we thus go one step further by introducing individual and possibly heterogeneous passengers in order to better take into account intentionality, reflexivity and adaptability of human beings [4,5]. Due to space limitation, we cannot present this in details; we can only give an insight of the benefits of our approach. The main contribution is to design, build and evaluate a frame model that will be extend in the future with more complex individual behaviors.

The article is organized as follows. Section 2 introduces the various modelling approach used in this article. Section 3 presents the model and Section 4 the method we use to evaluate the hybrid model. Section 5 presents the results and Sections 6 and 7 introduces mitigation strategies in the model and conclude.

2 State of the art

2.1 System dynamics modeling of epidemic

One of the traditional ways to describe the dynamic of systems is the *system dynamics* approach [6], *i.e.* the description of the evolution of macroscopic variables using a set of equations, often an ODE system. Generally it is not possible to get an exact solution of such system, and a numerical integration method (*e.g.* Runge Kutta [13]) is used to approximate it. Nevertheless we can prove analytical properties on equilibrium points and on their stability. In epidemiology, the most famous equation-based model is the SIR model [9]. It considers the global population of N people as a whole and describes the evolution of the 3 stocks of Susceptible (S), Infected (I) and Recovered (R) people using the system presented in Fig. 1 ($N = S + I + R$ in this system). The numerical integration using the Runge Kutta 4 (RK4) method is plotted in Fig. 2.

The parameter β is the infection rate of the disease in case of contact between a Susceptible and an Infected individual, *i.e.* it describes the transition between Susceptible and Infected stocks. Similarly α describes the transition between the Infected and the Recovered states, *i.e.* it represents the recovery rate. This system can be graphically drawn using a Forrester diagram [6] as in Fig. 3.

$$\begin{cases} \frac{dS}{dt} = -\frac{\beta IS}{N} \\ \frac{dI}{dt} = \frac{\beta IS}{N} - \alpha I \\ \frac{dR}{dt} = \alpha I \end{cases} \quad (1)$$

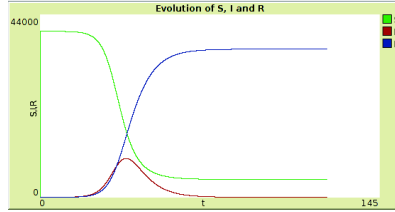


Fig. 1. ODE system for a SIR model

Fig. 2. Plot of the evolution of S, I and R for a SIR model. We choose: $\alpha = 0.2$, $\frac{\beta}{N} = 0.5$.



Fig. 3. SIR compartments

This system is very simple but can be complexified by introducing additional stocks or streams between stocks. For example if people that have recovered from the disease can be infected again, a stream between the Recovered and Susceptible stocks can be added (we get the SIRS model). If the disease is characterized by a time period when the individual is exposed to the disease but not yet seek, an additional stock (Exposed) can be added between Susceptible and Infected, we thus get the SEIR model.

2.2 Metapopulation models

One of the strongest limitation of the system dynamic approach is that it considers the population as a whole. It has been extended in ecology by the so called *metapopulation* approach [8]: the aim is to represent several populations and the migration relationships between them. The whole population is thus split into several populations that are the nodes of a graph, the edges representing the possible migrations. The dynamic inside each node (population) is described by an ODE model whereas migrations are managed as instantaneous streams between nodes. This approach has also been applied in epidemiology [11,3] to deal with the spread of a disease over a network of cities.

The metapopulation approach allows modellers to study how to control an epidemic by testing some mitigation strategies at the city level, such as the quarantine (a city is closed for arrivals and departures), avoidance (the traffic avoids a city, but airplanes can still leave it) or at the individual level, such as the risk culture (people being aware they are infected can choose to postpone their travel in order to avoid to infect other persons). Nevertheless such strategies remain limited as they can not take into account heterogeneous individual travels: individuals are not and cannot be represented in such models. This is the main feature of the agent-based modelling approach.

2.3 Agent-based models

The main idea of the agent-based models is to describe individual entities of a system and to let emerge the macroscopic expected behavior thanks to the interactions among entities. In epidemiology, it has also been successfully applied in studying the diffusion of mumps in Portugal [14] or plague in Madagascar [10].

Works such as [12] or [2] have already investigated the link between agent-based and equation-based models, but their point was to compare the two approaches in their representation of the same phenomenon. They highlight the difference of paradigm, scale of representation and way of thinking the phenomenon. But as far as we are aware, few articles investigate deeply the coupling of these two approaches into one single model as we do in the sequel.

3 A micro-macro coupled model: the MicMac model

In this section, we present the MicMac coupled model. It has been developed using the Netlogo platform [16] (*c.f.* Fig. 4) dynamically coupled with a Scala extension we designed for Runge-Kutta 4 numerical integration method. Due to space limitation, we cannot present it using the full ODD protocol [7], but we keep only the main parts of it.

3.1 Overview of the model

The main purpose of this approach is to model the disease spread in a network of cities using a coupled model composed of an aggregated equation-based approach (SIR) for the epidemic dynamic within cities and an agent-based approach for the epidemic spread between cities (air traffic). The two main processes we take into account are the local spread of the epidemic in each city and the spread through the network (we consider here only the air traffic).

3.2 Entities, state variables and scales

The model is composed of two kinds of agents: the *cities* (that are the *nodes* of the network) and the *airplanes* (that will carry passenger from one city to another one).

A city agent (defined in the *node* breed) is dedicated to describe the epidemic evolution of the whole population of a city. As this evolution will be described using the SIR model, the population of the city is represented by three state variables (*S_Node*, *I_Node*, *R_Node*) representing respectively the number of inhabitants in the Susceptible, Infected and Recovered states. It is important to notice that these state variables will contain float values (due to the integration of the SIR system using the Runge-Kutta 4 method). A city agent is also characterized by a mobility rate (*mobility-rate-node*) that will be used to compute the number of inhabitants that will travel. The number of people that should travel is stored in the variable *stock-to-flight*. To deal with mitigation strategies, two

additional state variables are introduced (*in-airport* and *out-airport*): they take boolean values and represent the fact that the city accepts (or not) incoming flights and emits (or not) flights.

The airplane agents (defined in the *MobileGroup* breed) are characterized by their *speed* (that will be calibrated given the chosen disease and the spacial scale), their target city (*Next-Node*) and their population (*i.e.* the number of susceptible *S_Group*, infected *I_Group* and recovered *R_Group* people in the plane).

An edge of the network is only characterized by the two cities it connects. We consider it represents an air route without stopover. As a consequence, in this simple version of the model, we do not consider stopovers in the trips.

In order to allow the numerical integration of the SIR systems, we chose that the simulation step corresponds to the integration step. We thus need to synchronize the air traffic process on this time discretisation. The time and space scales are defined as parameters of the simulation. Given a disease (characterized by its *alpha* and *beta* parameters) and an initial population, the modeller can choose the size of the environment (*TerritorySize-km* in km) and the integration step (*h*). Then a calibration phase is done at the initialization to compute other parameters (*e.g.* speed of airplanes).

3.3 Process overview and scheduling

The 2 main processes in the model are (i) the epidemic evolution and (ii) the air traffic. In terms of scheduling, at each simulation step, first the model computes the disease spread in each city and each airplane and then each city computes whether it should create a new airplane and chooses its target. Finally each airplane moves toward its target and when it reaches it, its population is integrated into the city population.

Epidemic evolution. In each city and airplane, the epidemic dynamics is driven by a SIR ODE system (*c.f.* Fig. 1). We consider only one disease and we do the hypothesis that the same system (with the same disease-related parameters) can be applied to each city and each airplane. So taken as input the number of people of a city in each state, we update these numbers by discretising the system using the Runge Kutta 4 numerical integration method¹. This computes the new number of people in each state. It is important to notice that the number of people in each state computed using the chosen integration method is a float value.

The equation-based approach has been chosen to describe the epidemic evolution in each city because this approach is dedicated to describe dynamics in a huge and homogeneous population, which is implicitly the case when we consider a city as one agent. In addition, it allows us to increase the population on each node without lose in terms of computation time.

¹ The numerical integration is done using an external plugin computing each RK4 integration step.

Air traffic. Each city has a variable *stock-to-flight* that represents the number of people that should leave the city and travel to another one. At each simulation step, each city computes (thanks to its *mobility-rate-node* attribute and its population) the number of people that should leave the city and add it to the *stock-to-flight* variable. If this number is greater than an airplane capacity, airplanes are created in order to empty this stock. Given the number of people that take an airplane, we extract a set of people that is representative of the whole city population by using a proportional random draw: for each individual to be added in the airplane we randomly choose his epidemic state. As a consequence, the number of Susceptible, Infected and Recovered people are initialized with an integer value: as the population in airplane is small we need to extract and manage individually each people.

Once an airplane has been populated, its target is randomly chosen among all the adjacent cities.

3.4 Initialization

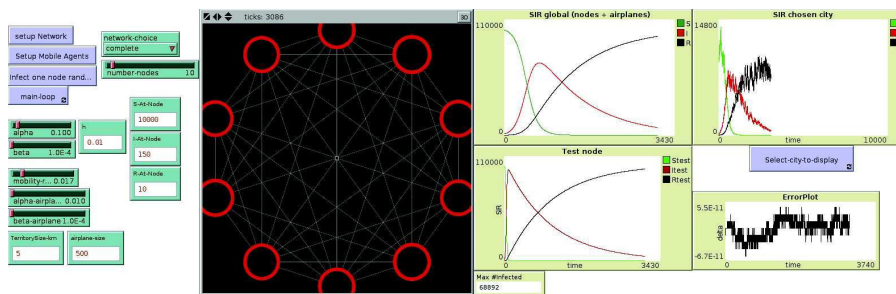


Fig. 4. Interface of the MicMac model. The modeller can choose the topology of the network (among complete, random, small-world...), the mobility rate and the initial population. The interface allows the modeller to observe the epidemic dynamics of one city but also of the whole network. The difference between the current and the initial population (error) is computed at each step, in order to ensure that the population remains constant during the simulation (a divergence could appear as we both manipulate float and integer values for the population).

Calibration. The duration of flights depends on the distance between the source and the target and is calibrated on the integration step. Therefore, the simulation setup contains a calibration step between the integration step, the flight duration and the distance. Given a particular disease (characterized by its *alpha* and *beta* (parameters) and its observed duration, the calibration is done on an additional reference node containing the population of the whole network: a SIR dynamics is applied until the epidemic end (*i.e.* until the step where the number of infected people is lower than a given threshold). This gives us the

simulation step corresponding to the end of the epidemic and thus the duration corresponding to an integration step h . Given the size of the environment, we can thus adjust the flight speed.

As an example, if we consider a disease characterized by $\alpha = 0.2$, $\beta/N = 0.5$ and a duration of 60 days and an initial population of 39999 susceptible and 1 infected individuals in each of the 10 cities of the network, we get a step duration of 0.1 days. So 1 simulation step duration is about 2.4 hours in this case, which justifies the need to take into account flights that are not instantaneous.

Parameters values. Parameters of the MicMac model are based on the U.S. domestic flights data². The 10 cities with the biggest airports have 57 709 474 inhabitants and about 21 420 000 passengers per month. The mobility rate is thus 0.37 per month and 0.012 per day. In addition, the area of the U.S. is about 9600000 km² and the mean size of airplanes is 80 passengers.

We chose thus parameters proportional to these values.

4 Evaluation method

In order to assess the impact of the agent-based model coupled with SIR models, we will progressively study several heterogeneities introduced by the individual airplane transportations and compare them with a reference model. This model considers only the epidemic dynamic in the cities: it is thus the integration of several SIR nodes³. This model is also equivalent to a homogeneous meta-population model on a complete graph.

On this model we will consider three particular values (illustrated on the Fig. 5) as indicators:

- the maximum value of the number of infected people: MaxI ;
- the time when the number of infected people is maximum: TimeofMaxI ;
- the duration of the epidemic: Duration .

These three indicators will be denoted $MaxI_{nSIR}$, $TimeofMaxI_{nSIR}$ and $Duration_{nSIR}$ for the reference model and $MaxI_{MicMac}$, $TimeofMaxI_{MicMac}$ and $Duration_{MicMac}$ for the hybrid MicMac model.

Whereas the two last indicators are computed by (numerical) simulation, this model is interesting because the first one can be computed analytically. If we have n nodes with on each an initial population of I_{init} infected, S_{init} susceptible and R_{init} recovered ($N = S_{init} + I_{init} + R_{init}$ being the total number of people), the value of MaxI is:

$$MaxI_{nSIR} = n \left(I_{init} + S_{init} + \frac{N\alpha}{\beta} \left(-1 + \ln \left(\frac{N\alpha}{\beta} \right) - \ln (S_{init}) \right) \right)$$

² http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/press_releases/airline_traffic_data.html

³ This is not equivalent to 1 node containing the whole population.

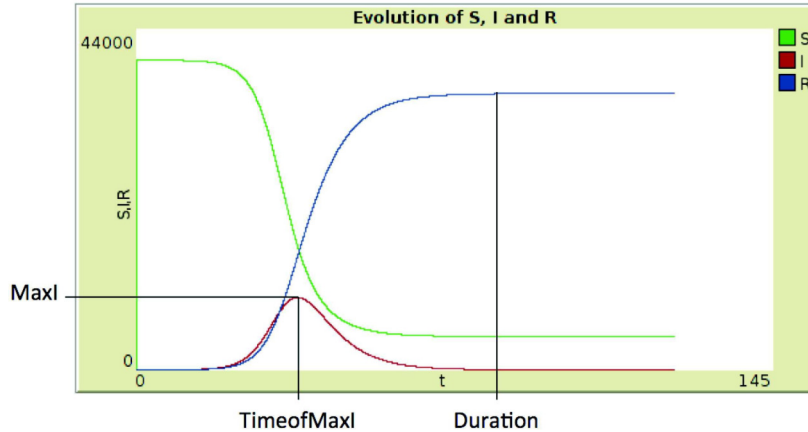


Fig. 5. Evolution of the number of susceptible, infected and recovered people over time, computed with $\alpha = 0.2$, $\beta/N = 0.5$ and $S_{init} = 1000$, $I_{init} = 10$, $R_{init} = 1$. The numerical integration method was Runge Kutta 4 with an integration step $h = 10^{-3}$.

5 Comparison

From a mean field model fitting with the reference model (Section 5.1), we progressively introduce heterogeneity on various dimensions: (i) initial population distribution (Section 5.2), (ii) time of flights (Section 5.3) and (iii) network (Section 5.4) and evaluate their impact.

5.1 Reference case: equivalence of both models

In this model, we remove all heterogeneity and the effect of the spatial component on the model:

- the mobility rates and initial population are the same for every node (no heterogeneity among cities).
- flights are instantaneous (the model is aspatial);
- the network is a complete network (no heterogeneity due to the network);

We compute (by simulation) the various indicators and we get the following results:

- $MaxI_{nSIR} = MaxI_{MicMac}$;
- $TimeofMaxI_{nSIR} = TimeofMaxI_{MicMac}$;
- $Duration_{nSIR} = Duration_{MicMac}$.

This first experiment shows that the MicMac model fits the reference model: we have thus built an agent-based model able to reproduce a set of equation-based models under mean field assumptions. This initial step was necessary to carefully assess the following models. From this we can go deeper in the exploration of the impact of heterogeneity in our agent-based model.

5.2 Impact of initial conditions (populations)

Instead of having a homogeneous population in each node, we introduce heterogeneity among nodes by creating initially all the infected people in only one city (this represents a more realistic situation where an epidemic starts in one location before spreading). Nevertheless the global population remains the same in the two models. Other hypotheses made in the previous section remain true. We get the following results:

- $MaxI_{nSIR} > MaxI_{MicMac}$;
- $TimeofMaxI_{nSIR} < TimeofMaxI_{MicMac}$;
- $Duration_{nSIR} < Duration_{MicMac}$.

All these indicators show that we now have a disease spread behaving as a classical diffusion phenomenon: the maximum of the epidemic is lower in the MicMac model because all the cities are not synchronized anymore and the TimeofMaxI is postponed (which induces a longer epidemic). We have to notice that the network topology has no effect here (as it is complete), but the diffusion is not instantaneous because of the fact that an airplane without infected people can leave an infected city (due to the random filling of the airplanes).

5.3 Impact of the time in flight

From the MicMac model presented above (with initial heterogeneous population distribution), we now release the instantaneity of flights. The size of the area is now taken into account and thus travels now take several simulation steps: the temporal coupling between travel time and integration step is made at initialization during the calibration step (*c.f.* Section 3.4).

We have explored the impact of the size of the area over the three indicators and the results are presented in Fig. 6. We observe that the size of the area does not have a significant impact on the results (for all the indicators).

This result is quite surprising as it shows that the distance has no influence on results in the current MicMac model. We can imagine two explanations:

- the total flying population is constant (*i.e.* independent on the travel distance and time);
- the contagion model in the airplanes is not adapted (the population in airplanes is small, so out of the scope of the classical SIR model).

It could be interesting to investigate the second explanation. A way to do could be to replace the ODE model by a Gillespie algorithm that is dedicated to small population (everything else in the model remaining unchanged).

5.4 Impact of the network

In order to study the impact of the network topology on results, we consider the same model initialized with various networks. In every case, we consider

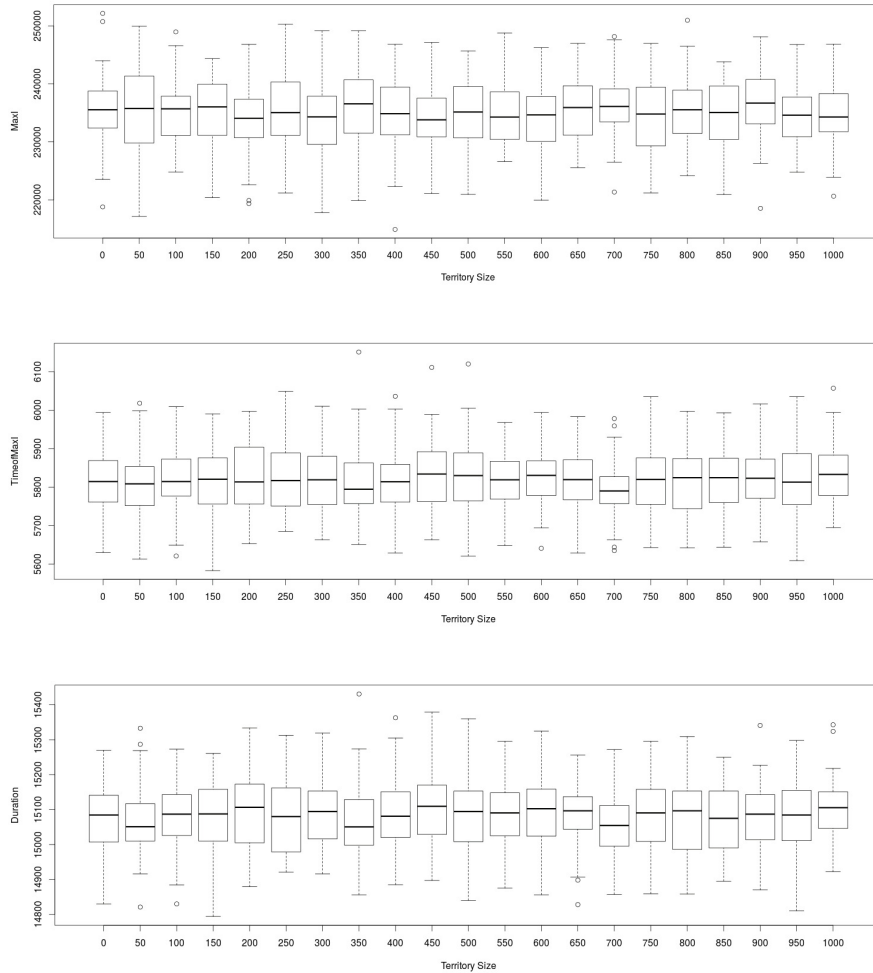


Fig. 6. Influence of the area size on the indicators MaxI , TimeofMaxI and Duration

a network with 100 nodes (cities). First we use regular networks with 4 to 99 neighbours for each node (the last one is a complete network). Then we use a small-world network. To produce small-world networks, we use the Watts & Strogatz algorithm [15]: we start with a regular network with 4 neighbors for each node. Then we rewire some of the nodes. To evaluate the impact of various small-world network, we consider rewiring probabilities from 0.2 to 1. Results are summarized in Fig. 7 and Fig. 8 .

In the case of regular networks, we can observe that the MaxI value increases and the TimeofMaxI and Duration decrease when the number of edges increases.

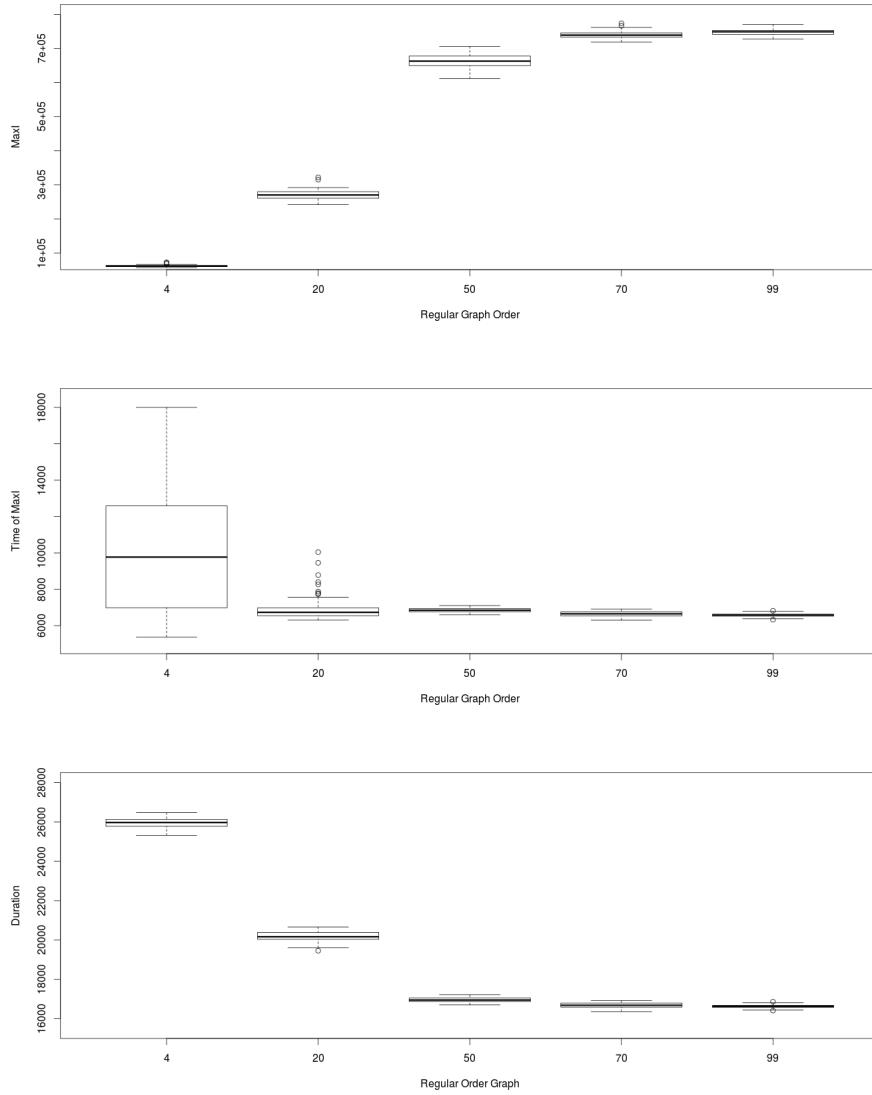


Fig. 7. Indicators values for regular networks with order among $\{0.04, 0.2, 0.5, 0.7, 0.99\}$

It is due to the fact that it is easier to infect other nodes when it is easier to access them via the network (which is allowed by increasing the number of edges). In the case of the small-world networks, we have similar observation when we increase the rewiring probability. The rewiring indeed creates shortcut in the

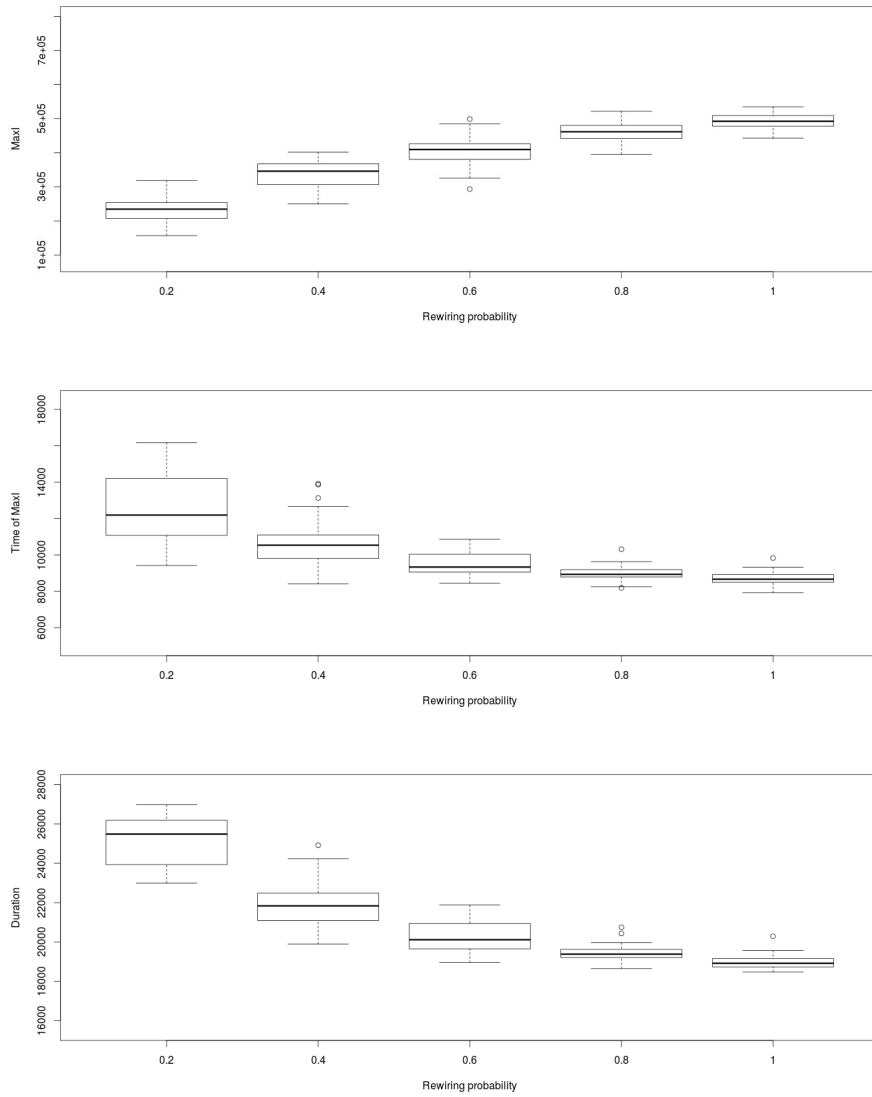


Fig. 8. Indicators values for small-world networks with rewiring probability among $\{0.2, 0.4, 0.6, 0.8, 1\}$

network. The more shortcuts the network has, the easier it is to infect other nodes.

In fact, increasing the rewiring probability (small-world networks) or the network order (regular networks) have the same effect: they decrease the diameter

and the average path length (APL) of the network (*c.f.* Fig. 9 and 10). In addition it is interesting to notice that the topology itself has not such an influence on the result, mainly the diameter (and the APL) of the network has the higher influence on results⁴.

Rewiring probability	0.2	0.4	0.6	0.8	1
APL	4.23	3.73	3.46	3.44	3.41

Fig. 9. Average Path Length for various small-worlds networks

Order	0.04	0.2	0.5	0.7	0.99
APL	12.88	2.98	1.5	1.22	1

Fig. 10. Average Path Length for various regular networks

6 Perspectives: Application of mitigation strategies

Once the model being designed, we can use it to evaluate several mitigation strategies. Due to space limitation, we give here only the method to implement these strategies. Several strategies have been proposed in the literature [11,3], such as the quarantine, avoidance of risk culture (*c.f.* Section 2.2).

The two first strategies are parameterized by a threshold: a city is put in quarantine or avoidance when its rate of infected people is greater than the given threshold. As an example, in the quarantine strategy, we consider that a city is put in quarantine when the ratio of infected people reached the threshold $\theta_{quarantine}$, *i.e.* when the following condition is fulfilled:

$$\frac{I_i}{S_i + I_i + R_i} > \theta_{quarantine}$$

We might consider that such strategy of containment is well defined at a macroscopic level and directly compatible with the metapopulation model. Therefore, adding a microscopic component would not add much. However, simulations run with both models suggest a different conclusion (Fig. 11). While the dynamics of the epidemic is largely driven by the SIR dynamics on nodes composed of large static populations (compared to small flying populations), the discretisation and stochasticity implied by the agent formalism has a significant impact on containment strategies, even if they rely to a macroscopic level. Indeed, as can be seen on Fig. 11, the impact of such a quarantine strategy on the epidemic, defined by the three indicators MaxI, TimeofMaxI and Duration, is much more progressive and realistic.

⁴ This is only a first attempt of result: it should be verified on other network topologies. Other characteristics should also be tested.

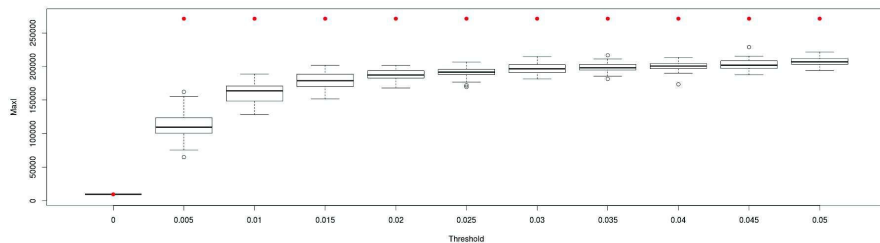


Fig. 11. Influence of the quarantine threshold on the indicator MaxI for the MetaPop (red plots) and the MicMac models.

7 Conclusion

We have presented the MicMac model that is a model coupling equation-based and agent-based models in a model representing the spread of a disease in a city network. We have studied carefully the influence of heterogeneities due to the agent-based model and compared results *w.r.t.* a reference model that does not contain the agent part. Finally we have presented how to integrate mitigation strategies inside.

As perspectives we plan to show the benefits of the hybrid approach (and in particular the benefits brought by the agent approach) by introducing much more complex diffusion processes: in particular, passengers will be able to do multi-stop travels, with possibility to be infected in the transit airport and passengers will be able to do round trips (they will be able to bring back disease to their home city).

Acknowledgements

This work has been partially founded by the CNRS through the PEPS HuMaIn (2013 and 2014) and by the French National Network of Complex Systems (RNSC) through the interdisciplinary network MAPS (<http://maps.csregistry.org/>).

References

1. J. Arino and P. Van den Driessche. A multi-city epidemic model. *Mathematical Population Studies*, 10(3):175–193, 2003.
2. F. Cecconi, M. Campenni, G. Andrighetto, and R. Conte. What do agent-based and equation-based modelling tell us about social conventions: The clash between ABM and EBM in a congestion game framework. *Journal of Artificial Societies and Social Simulation*, 13(1):6, 2010.

3. V. Colizza and A. Vespignani. Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: theory and simulations. *J. of Theoretical Biology*, 251:450–467, 2008.
4. R. Durrett and S. Levin. The importance of being discrete (and spatial). *Theoretical Population Biology*, 46(32):363–394, 1994.
5. L. Fahse, C. Wissel, and V. Grimm. Reconciling classical and IB approaches in theoretical population ecology : a protocol for extracting population parameters from IBM. *American Naturalist*, 152:838–852, 1998.
6. J. W. Forrester. *Industrial Dynamics*. MIT Press, 1961.
7. V. Grimm, U. Berger, D. L. DeAngelis, J. G. Polhill, J. Giske, and S. F. Railsback. The ODD protocol: A review and first update. *Ecological Modelling*, 221(23):2760–2768, 2010.
8. I. A. Hanski and M. E. Gilpin, editors. *Metapopulation Biology: Ecology, Genetics, and Evolution*. Academic Press, Waltham, Massachusetts, USA, 1997.
9. W. O. Kermack and A. G. McKendrick. Contributions to the mathematical theory of epidemics. *The Journal of hygiene*, 39(3):271–288, 1939.
10. V. Laperrière, D. Badaritti, A. Banos, and J.-P. Müller. Structural validation of an individual-based model for plague epidemics simulation. *Ecological Complexity*, 6(2):102–112, 2009.
11. S. Meloni, N. Perra, A. Arenas, S. Gomez, Y. Moreno, and A. Vespignani. Modeling human mobility responses to the large-scale spreading of infectious diseases. *Scientific Reports*, 1, 62(7):1–7, 2011.
12. H. V. D. Parunak, R. Savit, and R. L. Riolo. Agent-based modeling vs. equation-based modeling: A case study and users’ guide. In *First International Workshop on Multi-Agent Systems and Agent-Based Simulation (MABS)*, pages 10–25, 1998.
13. W. H. Press, S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical Recipes. The Art of Scientific Computing*. Cambridge University Press, 3rd edition, 2007.
14. J. Simoes. *An Agent-Based Approach to Spatial Epidemics through GIS*. PhD thesis, University College London, 2006.
15. D. J. Watts and S. H. Strogatz. Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684):440–442, 1998.
16. U. Wilensky and I. Evanston. Netlogo. center for connected learning and computer based modeling. Technical report, Northwestern University, 1999.