The effects of dynamic coarse-graining on simulation speed and consistency regarding SARS-CoV-2 simulation models

by

Max van Eck

Student number: 4468724

to obtain the degree of Master of Science

at the Delft University of Technology

to be defended publicly on 15-03-2022

Graduation committee

Chair	Dr.ir. J.H. (Jan) Kwakkel	Policy Analysis
First supervisor	Dr.ir. W.L. (Willem) Auping	Policy Analysis
Second supervisor	Dr. C.N. (Natalie) van der Wal	Systems Engineering

Executive summary

The SARS-CoV-2 virus, more commonly known as the coronavirus, is arguably responsible for the biggest global crisis in recent history. In an attempt to effectively deal with this crisis, politicians around the globe have been using simulation models in their policy development. There are multiple types of modelling techniques used for epidemiological transmission modelling (Alsharhan, 2021; Anastassopoulou, Russo, Tsakris, & Siettos, 2020). Each different method used has different advantages and disadvantages related to them.

One of the used simulation methods for transmission modelling is the Agent-Based Modelling (ABM) technique. The technique is a bottom-up approach, meaning it focusses on the behaviour of individuals to gather knowledge about the resulting emergent overall system behaviour. This technique specifically excels at developing early epidemic growth profiles, however, to gain this feature it needs to process a large amount of data. This data is not always readily available. Even if it is available, the amount of data that needs to be processed combined with the focus on individual behaviour, requires a lot of computational power for simulations making a robust uncertainty analysis very time consuming.

An alternative technique is equation-based modelling, with System Dynamics (SD) being an instance of it with additional benefits regarding communication. This is more of a top-down approach, focussing on the overall mechanics of the systems instead of the behaviour of the individuals in the system. Working with aggregate values for most if not all variables to create increased understanding in the system behaviour under different circumstances. Because this technique works with these aggregate values, there is a less of a computational strain when simulating the model. However, this comes at the costs of being able to generate accurate early epidemic growth profiles, as this technique is not capable of fully incorporating key concepts for transmission models, such as heterogeneity of agents, spatial effects, and stochasticity.

These two aforementioned modelling techniques, ABM and SD, have characteristic that lean themselves well to cover for each other's weaknesses. Utilising a technique that dynamically switches between the two modelling methods depending on the state of the model, could incorporate the strengths both models have, this is called dynamic coarse-graining. These strengths are the accuracy and incorporation of key concepts for the ABM side of the model, combined with the simulation speed of the SD side of the model. This dynamic coarse-graining is still in its infancy, as research related to it is very limited. The goal of this research is to examine what dynamically coarse-graining an ABM model to a SD model would mean for the simulation speed of the model, and whether the results will stay consistent with the more accurate ABM method.

The current research that has been performed on this topic has been on behaviourally stable models (Bobashev et al., 2007; Gray and Wotherspoon, 2012), meaning there are no changes to the behavioural mode of the model during the dynamic coarse-graining process. In this research two epidemiological transmission models are analysed. The first model will be a relatively simple model that is similarly incapable of exhibiting behavioural changes during the switching process. This Simple model is used as a test-case, for a more extensive SARS-CoV-2 specific model. This Extensive model will include the option of behavioural change during the switching process itself, meaning both the dynamic switching condition and agents' behaviour is dependent on disease state. As an added benefit, by comparing the results of the two different models, the gained insights are more generalisable.

By analysing the results of the dynamically coarse-grained models insights in regard to the consistency and simulation speed were generated. To be able to analyse the consistency of the models a metric of similarity, the normalised complexity invariant distance (NCID), has been proposed, based on the already existing Complexity Invariant Distance metric (Batista et al., 2013). The NCID metric combines the Euclidian distance between to data sets with the complexity of those datasets. The incorporation of the complexity aspect makes it well suited for analysing data that exhibits oscillatory behaviour, which is generally the case when examining transmission models. This metric allows for objective comparison of dynamically coarse-grained results, compared to the

results of the original ABM model it is supposed to replicate. Additionally, the outcomes were also analysed in conjunction with the uncertainty space to ascertain whether any uncertainty values predicted worse outcomes in terms of consistency.

The main results in term of simulation speed were, that both the Simple and Extensive model increased the simulation speed with 73,5% and 2,9% respectively. This increased simulation speed did come at the cost of consistency of the coarse-grained model, as the NCID metric indicated that respectively 66.7% and 57.5% were actually consistent with their ABM counterpart.

The difference in simulation speed is most likely explained by the computational burden created by the switching process itself. Due to the simplicity of the Simple model the switching process was very straightforward, however when switching the Extensive model significant amount of relevant information is lost. This information loss needs to be compensated, resulting in a more complex switching procedure, which has a significant computational burden associated with it. However, to be able to conclusively state this as the cause for the limited speed increase in the Extensive model, additional data on the speed of the different aspects of the simulation is required.

The outcome analysis of the consistency revealed that the inconsistencies in the Simple model can primarily be attributed to the method of modelling the disease duration. In the ABM model the agents will stay sick for exactly the given duration, whereas the agents in the SD model will stay sick for that duration on average, resulting in faster spill over into the immune group. The longer this main disease duration is, the bigger the impact on the overall model behaviour.

The main indicator for poor consistency results in the Extensive model, was the switching condition itself. The later the switch takes place from the ABM model to the SD model the more consistent the results are on average. This can mean that the stochastic components of the ABM model still have to much influence on the overall behaviour under the lower switching conditions. Alternatively, this could be the direct result of making the switch later in the modelling system, resulting in an increased percentage of the overall simulation being run in the ABM part of the coarse-grained model, and thus being more consistent with the ABM outcome. For the overall model this means there is no uncertainty space that specifically caused the inconsistencies, but the overall structure of the models seems to be lacking. The structure of the agents coming in contact, is the most probable cause of the problem as this structure was difficult to replicate in SD, and has big impact on the overall model behaviour.

From a more practical point of view, three distinct applications for dynamically coarse-grained models have been identified. Models that will be used for a sufficient amount of time, resulting in a raw time save in the long term, like Integrated Assessment Models (Weyant, et al., 1996). Models that can be created in advance and used in times of crisis, like epidemiological transmission models. Models working in real-time where speed is vital, for instance in case of self-driving cars.

The scientific contributions of this paper are the incorporation of an open exploration and behavioural changes with the dynamic coarse-graining method, the concretisation of the translation of some complex ABM properties to the SD paradigm, and lastly the introduction of NCID metric for consistency comparison of dynamically coarse-grained models.

It can be concluded that dynamic coarse-graining is a promising technique, yet only useful in specific niches. In order to gain the full benefits of this technique, the resulting model either needs to be used for near decades in case of Integrated Assessment Models or can be developed beforehand and used under extreme time pressures, in case of crises situations or self-driving cars. Once dynamic coarse-graining models have been fully developed, they do have the potential be an improvement on the currently utilised model, and dependent on their use even to help in saving lives.

Contents

List of Figures

List of Tables

1. Introduction

The coronavirus is arguably responsible for the biggest global crisis in recent history, having tremendous impact on, among other things, the economy, the healthcare systems, and peoples' personal lives (Søreide et al., 2020; Wójcik & Ioannou, 2020; Xiong et al., 2020). In order to find the best strategies in mitigating the consequences of the corona pandemic, governments have been using simulation models (Currie et al., 2020). There are several different simulation modelling methods, that all have their own advantages and disadvantages. The choice for a specific method is mainly based on the characteristics and type of the problem and the requirements the decision makers have (Chahal & Eldabi, 2010). In order to support the decision makers in dealing with the coronavirus, potential issues with the currently used modelling methods will be explored, and potential ways to mitigate these issues will be investigated.

Epidemiological modelling

Epidemiology is the study of the behaviour and determinants of disease in a defined population (Rothman, 2012). Since it would be unethical and dangerous to release diseases into non-laboratory populations, epidemiology is defined as an observational science (House, 2012). These restrictions make the use of mathematical and simulation models for research in this area almost a necessity. Simulation models have become more prevalent over the years to help policy-makers prepare for and deal with the increasing risk of pandemics, due to globalisation (Saker, Lee, Cannito, Gilmore, & Campbell-Lendrum, 2004). The most used methods of simulating epidemiology are Agent-Based Modelling (ABM), and System Dynamics (SD) (Gethmann et al., 2019; Huang, Lin, Chen, Huang, & Wu, 2013; Vonk Noordegraaf et al., 1998; Forrester, 1961; Sterman, 2000; Epstein & Axtell, 1996). This is also the case for the current modelling of SARS-CoV-2 (Alsharhan, 2021; Anastassopoulou, Russo, Tsakris, & Siettos, 2020). The main difference between the two paradigms is that ABM models have a stochastic nature versus the more deterministic nature of SD models (Railsback & Grimm, 2019; Jackson 1991). According to Epstein (2009), ABM models are best-suited for the simulation of a virus, as they can capture the global scale, complex social networks, and irrational behaviour, which is essential in confronting a virus. In accordance with this van Kleef, Robotham, Jit, Deeny, & Edmunds (2013) found that the prevalence of stochastic transmission models significantly increased relative to deterministic models over the years. However, SD is still recurrently used to research health policies and disease dynamics (Thompson & Tebbens, 2008), as they are among other reasons deemed to be easier to create (Maugeri, Barchitta, Battiato, & Agodi, 2020).

ABM & SD epidemiological models

When deciding between the two methods there are a few key characteristics that should be kept in mind. SD is a differential equation-based method, where every formula is pre-defined resulting in its more deterministic nature. This way of modelling results in implicit homogenous mixing assumption, as well as ignoring spatial effects of disease transmission (Merler et al., 2015; Di Stefano, Fuks, & Lawniczak, 2000). However, since people are likely to change their contact behaviour as a result of a deadly epidemic, it is important to model the heterogeneity of people and their adaptive behaviour (Epstein, Parker, Cummings, & Hammond, 2008). This lack of spatial effects generally results in overestimates of the very relevant R_0 value, which is an indication of the current reproductive number of the disease, making a precise definition of R_0 required for knowing vaccination thresholds and planning a vaccination strategy (Keeling & Grenfell, 2000). ABM models are able to tackle most of these issues. For one it is able to model of heterogeneous population and spatial factors, even tracking every individual's mobility and associated dynamic contacts, which is important for modelling realistic early epidemic growth profiles (Perez & Dragicevic, 2009; Chowell, Sattenspiel, Bansal, & Viboud, 2016). This is partly due to being able to incorporate so-called "super-spreading" events, situations in which a single individual directly infects a large number of others, which can have a large influence in the early course of the epidemic, especially in the case of a coronavirus (Lipsitch, 2003; Riley, 2003). The stochasticity that allows this inclusion, also allows the exploration of

the effects of various treatment strategies and model parameters on R_0 (Granich, Gilks, Dye, De Cock, & Williams, 2009). The combination of these capabilities makes ABM the state-of-the-art for simulating complex epidemic systems, especially the capacity of incorporating very detailed information is crucial (Siettos & Russo, 2013). The high complexity present in these ABM models comes at a cost, as very complex simulation models take a longer time to run (Chwif, Barretto, & Paul, 2000). To fully utilise the capabilities of ABM there is also a need for very detailed input data, which can be hard to obtain, whereas SD is less detailed, which makes it more scalable, increasing the spatial and temporal ranges that can be analysed (Ajelli et al., 2010). Ajelli et al. (2010) also suggest the possibilities of combining the methods, as the complexity present in agent-based models is primarily necessary to simulate the beginning stages of a pandemic more accurately when there are only a few infected citizens, making the behavioural impact a singular citizen or agent can have on the system as a whole significantly larger. When the number of infected citizens is getting larger, the Law of Large Numbers takes effect (Hsu & Robbins, 1947), meaning the agents could potentially get aggregated without significant loss of information. SD is a method that takes advantage of this, by allowing for a focus on aggregates (Tesfamariam & Lindberg, 2005). By modelling aggregates instead of every single individual, the computational power required is reduced significantly. When the purpose is to build a robust statistical portrait comparable to epidemic data, a model has to be run thousands of times (Epstein, 2009). This fact combined with the long run-time of complex ABM models can result in serious issues. Therefore, methods of increasing simulation speed should be looked into further.

Methods of increasing simulation speed

In principle two methods of speeding up simulation exist; one can either increase computing power, via methods like Distributed Simulation (Taylor, 2019), or one can try to reduce the required computing power, via methods like dynamic coarse-graining. Dynamic coarse-graining essentially looks for parts of the system that could potentially be reduced to a less complex form, without losing essential information. Simulating this part of the system in its reduced form would theoretically reduce the required computational power, resulting in a faster simulation (Xue, Ludovice, & Grover, 2011).

Coarse-graining itself, in regard to modelling, finds its origin in molecular modelling. Doing mathematical calculations using proteins is incredibly difficult, as even very small proteins have around 200 degrees of freedom. To improve the efficiency of these calculations Levitt and Warshel (1975) simplified the representation of these proteins, by averaging over the fine details. This process would later become known as coarse-graining. Since then, coarse-graining has become a fast developing methodology, yet primarily bounded to the field of biomolecular simulation (Saunders & Voth, 2013).

Biomolecular modelling was not the only modelling niche, that struggled with computational power issues when attempting to simulate very complex models (Railsback, Lytinen, & Jackson, 2006; Wang & Chatwin, 2004). While Biomolecular modelling focussed primarily on coarse-graining as a method to improve simulation speed, other areas focussed on parallel and distributed simulation (Fujimoto, 2015). Over the years parallel and distributed simulation methods have established themselves as very capable of speeding up the simulation time of models (R. M. Fujimoto et al., 2003; Mustafee & Taylor, 2009). In spite of this benefit, distributed simulation standards like the high-level architecture (HLA) are scarcely used in practice (Boer, de Bruin, & Verbraeck, 2009). According to Boer, de Bruin, and Verbraeck (2009) there are multiple reasons for this, ranging from performance issues to unfavourable cost-benefit ratios. Since then, there has been a lot of progress in the field of distributed simulation as demonstrated by Taylor (2019). However, issues like user-accessibility still persist. It is clear that the future of parallel and distributed simulation is very promising, but researchers and policy-analysts might not always be able to access the necessary computational power this still requires. Adaptation of coarse-graining techniques by the wider modelling community could therefore be a very interesting proposition, especially since the two techniques are not mutually exclusive (Ford, Weitzner, & Bahl, 2019).

Hybrid ABM-SD models

Integrating System Dynamics models with ABM models has the potential to utilise the strengths of both techniques while covering for each other's weaknesses (Nasirzadeh, Khanzadi, & Mir, 2017). Specially, the combination of them more accurate yet slower ABM models combined with the less accurate but faster SD models, to result in a both accurate and fast model. The potential for this combination was already identified by Scholl in his 2001 paper. He argued that even though the epistemological and ontological standpoints may differ, the practical benefits of working together are undeniable (Scholl, 2001). Since this paper there have been a number of successful collaborations between the two disciplines (Łatuszyńska, 2019; Monasterolo & Raberto, 2016; Swinerd & McNaught, 2014). However, these collaborations are still few and far between. If we look specifically at dynamic coarse-graining an ABM model into a SD model - also called IB-SD models swapping (Vincenot et al., 2011) - there are even fewer collaborations. Wallentin and Neuwirth (2017) used this method in regard to a simple predator-prey model. They found that it only sped the simulation up in some cases, this could probably be attributed to the added computational power needed for switching paradigms, which may not weigh up to the increase in modelling speed for simple models. Bobashev, Goedecke, Feng Yu, and Epstein (2007) used an approach similar to this to study epidemiological processes, however they did not use a System Dynamics approach specifically. Their results show that coarse-graining these models can save significant computational times but can also come at the expense of the desired accuracy. As mentioned in their discussion, they used a relatively simple agent-based model, not accounting for behavioural changes during an epidemic. In light of the Corona crisis, where behavioural changes play a very significant role, the spread of the virus and effectiveness of the mitigation measures (Luo, Yao, Zhou, Yuan, & Zhong, 2020), it is clear that additional research into the potential benefits of dynamic coarse-graining is required.

The potential benefits related coarse-graining can also be relevant for decision-makers. Since, decision-makers currently dealing with the corona pandemic, having access to faster yet accurate simulation models will allow them to test and develop new policies faster. The increased speed of development of dynamic policy pathways, has the potential to positively influence the amount of lives saved, the economy, and many other aspects of society currently suffering from the pandemic.

1.1. Research objective

Based upon the knowledge gap identified in the introduction, the following research question has been formulated:

"What is the effect of dynamically coarse-graining an Agent-Based SARS-CoV-2 model into a System Dynamics model on the simulation speed and the consistency of results?"

In order to adequately tackle this question, several sub questions have been developed. These sub questions have been devised in such a way that the intermediary objectives will be fully resolved, leading to an answer for the main research question. The devised sub questions are shown below:

1. What are different methods, and their associated advantages and disadvantages, in regard to modelling SARS-CoV-2 transmission in populations?
2. What set of rules should be used for dynamic coarse-graining an agent-based SARS-CoV-2 transmission model into a system dynamics model?
3. What would dynamic coarse-graining an agent-based SARS-CoV-2 transmission model mean in terms of simulation speed?
4. What are the differences and similarities between the coarse-grained and the original agent-based SARS-CoV-2 transmission model in terms of consistency of results?

1.2. Report structure

This research consists of five main chapters, that build on each other to come to a well substantiated conclusion. This chapter serves as an introduction to the topic of dynamic coarse-graining. In chapter 2 the literature regarding epidemiological models will be discussed. Based on this knowledge two dynamic coarse-graining models will be developed and discussed in chapter 3. Subsequently, the results of the simulation study will be showcased in chapter 4, with a discussion about the model and the conclusions to all sub-questions in chapter 5.

2. Methods of epidemiological modelling

This chapter will focus on the methods used for researching epidemiological transmission systems. First the most common type of transmission models will be discussed; "Susceptible-Exposed-Infectious-Removed" (SEIR) transmission models. Sections 2.2 and 2.3 will discuss modelling paradigms that are used most when utilising a SEIR model. The Agent Based Modelling (ABM) will be discussed followed by Equation-Based techniques, and System Dynamics (SD) in particular. These sections will primarily focus on the advantages and disadvantages in regard to modelling SARS-CoV-2. In Section 2.4 it will be discussed how a combination of these methods can be used to potentially gain additional benefits compared to using the methods in a vacuum. The combined knowledge gathered in these sections will be used to answer the second sub-question; "*What are different methods, and their associated advantages and disadvantages, in regard to modelling SARS-CoV-2 transmission in population?*". The fifth Section will describe how the knowledge gained in the aforementioned sections will be utilised for experimentation. This Section will be followed by a Section that will discuss how the results of the coarse-grained model will be compared to the results of the ABM model. Lastly in Section 2.7 the process of the outcome analysis is described in detail.

2.1. SEIR Transmission models

Epidemiological transmission simulation models have long been used to study the effects of potentially dangerous disease outbreaks in populations (Hethcote et al., 1981). The main structure of these models is based on the different stages of a disease in individuals of a population, as indicated by their more common name of SEIR models, where SEIR stands for Susceptible-Exposed-Infectious-Removed (H. A. Biswas et al., 2014). The "Susceptible" population is everyone who is currently not infected by the disease yet is able to get infected by the disease. Once someone gets infected, they will first go through some incubation period before they are able to infect others (Sartwell, 1966), individuals in this stage are called "Exposed". After the pathogen is sufficiently incubated, the exposed individuals will become able to spread the infection, moving them to the "Infectious" stage. For the dynamics of transmission models, it is generally important to differentiate between these two stages, however some more simpler models opt to combine these two stages into a singular "Infected" stage. The last stage, "Removed", is in its most basic form a combination of the individuals who succumbed to the disease and the individuals who recovered from it, and thus gained immunity. These groups are both categorised as removed, since they both do not play any role in further transmissions.

Besides these main stages of disease progression sometimes extended versions of the generic SEIR models are used. These models include additional progression stages that can have an influence on the overall spread of the disease, especially when modelling human populations this should be considered (Franco, 2021). Some extensions we will discuss are the inclusions of hospitalised individuals, the uncoupling of deceased and recovered populations, and differentiating between asymptomatic, symptomatic, and pre-symptomatic infectious individuals.

Including hospitalised individuals has two advantages, the pressure on the healthcare system can be ascertained and it influences the spread of the disease. People who are hospitalised are basically removed from the general population, as they do not partake in societal activities anymore. This significantly reduces the chance of them infecting susceptible people. Secondly, by tracking the amount of people that are hospitalised, the influence of potential mitigation measures on the healthcare system can be evaluated.

The inclusion of the different symptomatic types can greatly influence the disease dynamics, this is due to the big impact asymptomatic infectious individuals can have (Shao & Shan, 2020). When an individual gets symptoms of a disease, they are usually less likely to continue spreading the disease because of two reasons. For one the symptoms could make the infected individual bedridden or too sick to make them attend to social gatherings. Secondly the knowledge of having a transmissible disease can prevent them from attending social events based on moral grounds or have

them take precautions to reduce the chance of spreading the disease any further. However, when someone is either asymptomatically, or pre-symptomatically infectious they are often not aware of their infectious status, as they do not experience any (severe) symptoms of the disease. This difference has a great influence of the behaviour of the infectious populations and thus also influencing the spread of the disease.

The last extension we will discuss is the uncoupling of the "Removed" stage into deceased and either recovered or immune. Depending on the time the immunity is supposed to last for and the temporal scope of the model it could be relevant to include the loss of immunity in the model, replenishing the susceptible population.

Now the standard SEIR model and some potential extension to the model have been discussed, we can take a look at what modelling paradigm is most helpful for modelling SARS-CoV-2. Even though both techniques will be using the same basics in terms of SEIR, the actual implementation can result in quite different outcomes.

2.2. Agent-Based Modelling

Agent-Based Modelling (ABM) is a technique that can best be categorised as a bottom-up approach (Boulain et al., 2007), starting off will a collection of simple interactions and rules, that occur within the system of interest. The combination of all these different "simple" interactions can lead to emergent complex behavioural patterns. Section 2.2.1. will delve deeper in the basics of ABM. Section 2.2.2. will evaluate the capabilities of ABM for modelling transmission models.

2.2.1. ABM in general

The technique of Agent-Based modelling (ABM) was originally developed by Epstein and Axtell (1996) in an attempt to showcase, that the so-called soft sciences are just as decomposable into small subprocesses, as processes studied in subjects like physics. They argue that the aggregate behaviour we can all observe is a result of small processes going on everywhere. By modelling parts of systems, and the specific behaviours people may or may not preform, it would become easier to say whether these small behaviours actually lead to a particular property within a society, as performing controlled experiments to test such a hypothesis, is extremely difficult if not impossible.

ABM models can be best described by using the ODD protocol (Overview, Design concepts, Details) originally developed by Grimm et al. in 2006 and further updated by Grimm et al. in 2010 and by Grimm et al. in 2020. The seven main elements of this protocol form the core for every ABM model. An overview of these elements can be seen in figure 1.

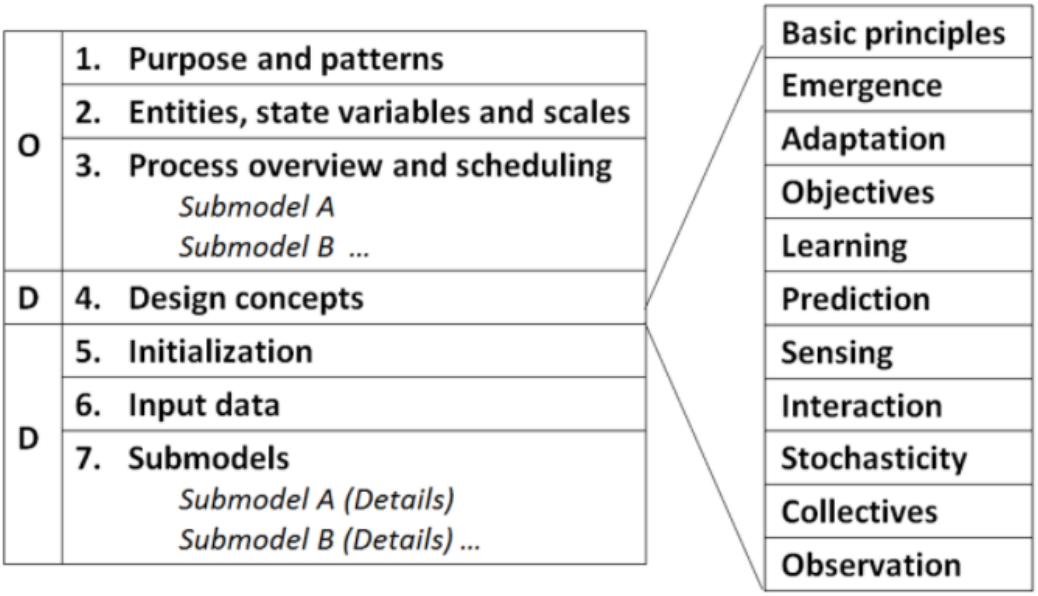

Figure 1: Elements of ODD protocol (Grimm et al. 2020)

The first element is in regard to the purpose of the model, and its expected patterns. The purpose of the model indicates why the model is created in the first place, and in extension also what should be and should not be included in the model. This follows from the principle of parsimony, also known as Ockham's razor, meaning only parts that serve the purpose of the model should be included in the model (Rodriguez-Fernández, 1999). To ensure the models suitability for its purpose, some expected behavioural patterns, that can be used to evaluate the model, should be pre-defined.

The second element of the ODD protocol is related to entities, state variables and scales, which can be considered the basic building blocks of the actual model itself. An entity is described as "a distinct or separate object or actor that behaves as a unit" (Grimm et all. 2020). These entities can be subdivided in four categories: agents, collectives, spatial units, and the environment. Agents are individual units that all have their own state variables and behaviours, creating the opportunity for heterogeneity within the model. Collectives are groups of agents that have certain attributes in common. Spatial units are used to define the conditions of the local space an agent exists in, most often this regards to geographical terrain. Lastly the environment holds state variables that are not space dependent. All these variables also have a predefined scale, the spatial units, or grids, have a certain size for instance.

The third element is about the processes that take place in the model, and on what schedule these processes take place. This primarily refers to the entities in the model, and how and when they interact with each other.

When movement options and interactions or "rules" between agents, based on these properties, are introduced the behaviour can become very complex. This gained complexity from "simple" behavioural patterns is called emergence, which is one of the design concepts within ABM. The design concepts can be interpreted as ABM specific characteristics. Not all design concepts all present in every model, as the use also depends on the purpose of the model. We will discuss a few design concepts most relevant for epidemiological transmission models.

Adaptation: This design concept encompasses the behavioural changes agents exhibit as a result of changes in their own state variables or the environment. In transmission models this is utilised when agents move through the different disease stages, as agents will for instance stop roaming around when they get ill, or change their behaviour based on a change in the active restrictions.

Emergence; As mentioned before, emergence can be seen as complex behavioural patterns that are results of simple behaviours or adaptive traits exhibited by agents. In transmission models this can for instance be seen in oscillating amount of infected individuals in persistent diseases.

Interactions: This characteristic is about how agents can affect other agents in the model. This is very relevant for transmission models, as the infections between agents are basically interactions. Infected agents effectively change the state variables of susceptible agents.

Stochasticity; or in Layman's terms randomness. This is one of the most important aspects for transmission models, as the inclusion of stochasticity is required to enable several key characteristics. By making both movement and infection chance stochastic, phenomenon like super-spreading events and local die-outs are able to occur, potentially having tremendous impact on the overall spread of the disease.

Collectives: are groups of individuals, that have certain state variables in common. In transmission models one could see the people in a certain disease phase as a collective, like the susceptible people. Another relevant collective is the age groups of the agents, as people are more likely to meet other people from the same age group, which influences the transmission mechanics.

The fifth part of an ODD is about the initialisation of the model, meaning the parameter values at the beginning of the simulation. The initial conditions of a model can have a big influence on the overall outcome of a model, making it very relevant for replication of results. This is most relevant when researching very specific regions, where the parameter values need to exactly match reality. However, when exploratory modelling is utilised, this is less relevant, as most if not all initial parameter values will be randomised.

The sixth aspect is in regard to external data sources, that are used during the simulation (not initial values). These external sources can be used to incorporate time series data of environmental variables, for instance the rain fall throughout a year or, more relevant for transmission models, the amount of incoming tourists.

Now that the basics of ABMs have been thoroughly discussed, the use of these models for the purpose of researching SARS-CoV-2 can be explored.

2.2.2. ABM SARS-CoV-2 models

Ever since the SARS-CoV-2 virus started to become a global pandemic a lot of scientific literature has been written about utilising ABM to assess the properties of the virus and the effectiveness of different mitigation methods on the spread of the virus (Alsharhan, 2021; Bossert et al., 2021; Chang, Harding, Zachreson, Cliff, & Prokopenko, 2020; Dignum et al., 2020; D'Orazio, Bernardini, & Quagliarini, 2021; Gopalan & Tyagi, 2020; Hinch et al., 2021; Hoertel et al., 2020; Kai, Goldstein, Morgunov, Nangalia, & Rotkirch, 2020; Rockett et al., 2020; Shamil, Farheen, Ibtehaz, Khan, & Rahman, 2021; Silva et al., 2020).

This wide utilisation is due to the capability of ABM models to incorporate some of the key characteristics of epidemiological transmission models (Silva et al., 2020): Heterogeneity, Movement, Contract tracing and superspreading events. One of these characteristics is heterogeneity of the population. For example, factors regarding health, like obesity and age (Alberca, Oliveira, Branco, Pereira, & Sato, 2020), are important to take note of, as they play a major role in determining the amount of SARS-CoV-2 patients in the ICU. Because ABM allows the setting of different parameters for every individual (Kai, Goldstein, Morgunov, Nangalia, & Rotkirch, 2020), it is possible to capture this heterogeneity.

A second important characteristic, that can be captured by ABM models is the movement of people. This movement is relevant as it determines whether infected people actually come in contact with susceptible people. By combining a detailed travel diary component with a classical SEIR model, ABM simulations can depict the heterogeneity between people and be able to model the sophisticated dynamic connections among them, including infection rate that is subject to peoples detailed personal activities, their methods of movement and their physical distance between one another (Shi, Wu, & Ben-Arieh, 2014).

This is also relevant in regard to the third key characteristic, the opportunity for contact tracing. Epidemiological transmission models of SARS-CoV-2 have indicated that the virus spread can be contained by fast recognition and quarantine of infected individuals and their recent contacts (Anderson, Heesterbeek, Klinkenberg, & Hollingsworth, 2020). Due to the ability of ABM to simulate every individual and their movement patterns explicitly, it is better at simulating contract tracing accurately relative to models based on ordinary differential equations (ODEs) (Arino & Portet, 2020). This tracking of the agent movement and associated contacts allows for the occurrence of so-called super-spreading events, where an infectious individual infects a significantly larger amount of susceptible people than would usually happen. The inclusion of these events is especially important, as it has a big impact on early epidemic growth profiles of the spread of the virus. (Lipsitch, 2003; Riley, 2003).

The use of ABM models to model SARS-CoV-2 has some caveats as well, mainly relating to availability of data and computational burden. These models and their parameterisation are often very reliant on the availability data, to be able to be correctly calibrated (Hoertel et al., 2020). The more detailed a model becomes, the more data is needed and also the more computationally intensive it becomes to actually simulate the model (Shamil, Farheen, Ibtehaz, Khan, & Rahman, 2021). This is especially an issue for ABM models as these models are very computationally intensive as is. This computational intensity does not only increase the time that is needed to fully simulate the model but can also limit the ability to explore a wide range of parameter combinations (Hinch et al., 2021). When new policies are not tested for a wide range of parameter combinations the outcomes can be disastrous, as was shown in The Netherlands in the beginning of July 2021. The amount of

confirmed corona-cases grew explosive due to some recent easing of the corona-mitigating measures (RIVM, 2021). Multiple members of the Dutch Outbreak Management Team (OMT) declared that this growth was above all expectations shown in the models (NU.nl, 2021a, 2021b), indicating that the uncertainties in the model were not explored to the fullest extent. Stochastic ABMs have been recognized as a robust method for discovering the impacts of heterogenous mitigation measures in varied pandemic and epidemic situations (Balcan et al., 2010; Barrett, Bisset, Leidig, Marathe, & Marathe, 2010; Chang, Harding, Zachreson, Cliff, & Prokopenko, 2020; Chao, Halloran, Obenchain, & Longini, 2010; Eubank et al., 2004; Ferguson et al., 2005; Germann, Kadau, Longini, & Macken, 2006; Halloran, 2002; Longini, 2004, 2005). This indicates that a stochastic ABM model should have been able to explore such an outcome, but that the high computationally burden that comes with it either prevented full exploration of the uncertainties or pushed the researchers towards a different modelling paradigm, that is less accurate when it comes to transmission models.

2.3. Equation-based modelling

Equation-based modelling is another technique that has widely been used to model epidemiological transmission systems. Utilising the equation-based paradigm instead of the agent-based method to model these types of systems has several advantages, and also some disadvantages. These will be discussed in Section 2.3.2. Section 2.3.1 will first focus on the characteristics of the technique itself, which is important to understand to be able fully grasps the differences between the two techniques.

2.3.1. Equation-based techniques in general

Equation-based simulation models are a collective of modelling techniques that most often use deterministic or ordinary differential equations (ODE) to analyse a system of interest. In this research we utilise a specific instance of those techniques, System Dynamics. The primary advantage of utilising this technique over the others is its ability to clearly communicate model structure to people not adapt at reading complicated mathematical equations, which is most of the population.

System Dynamics (SD) is a modelling technique developed by Forrester (1961) to increase understanding in the behaviour of complex dynamic systems. These kinds of systems are inherently difficult to grasp, since they are characterised by the occurrence of feedback structures, accumulations, and delays. Mental simulation of systems that incorporate these concepts is quite complicated, evidence suggests that people cannot mentally simulate any but the simplest of mental models without error (Doyle & Ford, 1998; Sterman, 1989). Utilising SD does not only support mental simulation of models, but also prompts the modeller to evaluate the direct and indirect dependencies between variables within the system. SD can also be viewed as a top-down approach, as the focus is mostly on the aggregate effects different parts of the system have on each other, which is opposite to the ABM approach (M'hamdi & Nemiche, 2018).

The SD technique utilises stock-flow diagrams to quantify its properties (Liu et al., 2021). The stocks are used for accumulation of mostly tangible objects, like people, but can also be used for intangible variables, like confidence levels (A. Ford, 1999). These objects or variables can enter or leave stocks via respectively inflows and outflows. Besides these two main variable types there are also constants and auxiliary variables. The constants are used to determine both the initial values for other variables, like stocks or delays, and the constant characteristics of the system. The auxiliary variables are used to combine the other variables with each other, via the use of mathematical equations, that are solved via the use differential equations (Ossimitz & Mrotzek, 2008). These equations are all deterministic from nature (Jackson, 1994), meaning the model will always result in the same outcome when it has the same initial starting conditions.

Besides utilising stock-flow diagrams for quantifying the model properties, it is also an excellent method of communicating the basic structure of models (Lane, 2008), as all non-stock or flow variables are also interconnected via the use of arrows, which indicate direct causal relationships. The combination of these variables and interconnections will results in a rough image

of how the system works, without needing to know any associated mathematical equations. This allows for effective communication of the model structure.

2.3.2. Equation based SARS-CoV-2 models

Equation based simulation methods have arguably been the most used technique of researching the SARS-CoV-2 pandemic (Abdo, Shah, Wahash, & Panchal, 2020; Anastassopoulou, Russo, Tsakris, & Siettos, 2020; Arino & Portet, 2020; Barlow & Weinstein, 2020; Chatterjee, Chatterjee, Kumar, & Shankar, 2020; Choi & Ki, 2020; Fanelli & Piazza, 2020; Ibarra-Vega, 2020; Ivorra, Ferrández, Vela-Pérez, & Ramos, 2020; Kim, Kim, Peck, & Jung, 2020; Kuniya, 2020; Li, Song, Yang, Gao, & Gao, 2020; Manchein, Brugnago, da Silva, Mendes, & Beims, 2020; Maugeri, Barchitta, Battiato, & Agodi, 2020; Tang et al., 2020; Tuite, Fisman, & Greer, 2020; H. Wang et al., 2020).

Diekmann, Heesterbeek, and Britton (2013), argue that the choice for a deterministic model is preferable as a first approach for analysing a transmission system, as such models have a few advantages compared to their stochastic counterpart, when not much is known about the pathogen. The primary advantage of using an equation-based method is the relative ease of creating such a model, the low computational burden it has, and its ability to give a clear overview of the feedback mechanisms at work.

The relative ease of building an equation-based simulation model is showcased by a few recent SARS-CoV-2 papers, where there were only a very few equations required to create a relatively adequately working model (Chatterjee, Chatterjee, Kumar, & Shankar, 2020; Maugeri, Barchitta, Battiato, & Agodi, 2020). These models comprised of only a few equations are not as accurate as more detailed/extensive models, since any model involves trade-offs between simplicity and realism (Tuite, Fisman, & Greer, 2020).

By creating models comprised of solely differential equations focussed on aggregates, instead of needing to model individual agents, the computational burden is decreased significantly (Parry & Evans, 2008). The smaller computational burden will result in faster simulations and the opportunity for a more extensive calibration of model parameters and exploration of the uncertainty space (Ivorra, Ferrández, Vela-Pérez, & Ramos, 2020).

The last big advantage of using ODEs is its ability to give clear insights in the (feedback) structure of the model. In their equation form these models will be able to communicate this to people very familiar with calculus, however when they are translated into a Stock-Flow diagram (SFD) they can be very effective in communicating with the general public (Lane, 2008).

Utilising a deterministic equation-based method to model epidemiological transmission systems also has some severe downsides, as it requires some unrealistic assumptions about key characteristics of these models (as mentioned in Section 2.2.2.).

The biggest issue is the implicit assumption of homogeneity present in ODEs, as examining pandemics while only considering the averages and aggregates of the population does not exhibit the complex emerging dynamics accurately (Großmann, Backenköhler, & Wolf, 2021). The lack of heterogeneity within these models, also dampers the inclusion of hubs (people with an unusually large amount of contacts) and super-spreading events (Shen, Taleb, & Bar-Yam, 2020). The importance of inclusion of these events, has already been mentioned in Section 2.2.2., and has also been indicated by several studies covering SARS-CoV-2 (Cave, 2020; Hasan et al., 2020; Riou & Althaus, 2020).

The use of aggregates and differential equations also fails to take spatial distribution and realistic movement patterns of individuals into account (Mammeri, 2020). Due to this is lack of actual movement within the model, it is difficult to ascertain results regarding testing and contact tracing, which would be possible when utilising a different modelling paradigm like ABM (Arino & Portet, 2020). Arino & Portret (2020), also mention this would be less of an issue when considering larger populations, as aggregate values will more accurately resemble reality in that case.

2.4. Multi-Modelling

An often discussed yet seldom performed method to model transmission systems is by coupling SD and ABM models (Scholl, 2001). The properties of these techniques are quite opposite to each other, which could potentially make them difficult to design. However, the techniques cover each other's weaknesses, while creating an opportunity to benefit from the advantages, resulting in a best of both worlds scenario (Banos et al., 2017). Specifically, the accuracy of which ABM models can simulate the beginning stages of virus transmissions (Perez & Dragicevic, 2009; Chowell, Sattenspiel, Bansal, & Viboud, 2016) and especially its ability to incorporate super-spreading events in these stages (Lipsitch, 2003; Riley, 2003), can be combined with the high speed the equation-based technique is able to simulate its models (Parry & Evans, 2008).

In general, there are four different types of model coupling that can be utilised, depending on when and how the different modelling techniques need to communicate (Vincenot et al., 2011). These four coupling methods are: Individual interacting with a single SD model, SD submodels embedded in individuals, individuals interacting with a space made of SD models, and model swapping. These different coupling methods will be shortly explored, to determine the best method to use for this research.

Individuals interacting with a single SD model relates to a situation where an ABM and SD model are run in unison. The agents behave in the same manner as they would in a solely ABM model, however some environmental variables are determined by the SD model. This SD model can also be influenced by the behaviour of the agents in the ABM model. This can be useful, as some environmental factor can more accurately be determined by a SD model.

SD submodels embedded in individuals, is similar to aforementioned technique. However, instead of utilising a single SD model to determine some global environmental variables, the agents utilise the SD models to dynamically determine some of their properties. The main use of this coupling method is to simulate the internal properties of agents more accurately.

Individuals interacting with a space made of SD models, utilises SD models to determine the environment of the agents again. However, instead of utilising a singular SD model to determine some global aspects of the environment, multiple SD models are used that each determine some specific space. This creates the ability to include heterogeneous spatial features in the modelling.

Model swapping, this method of model coupling is proposed with the express purpose of combining models that exhibit differences in computation time and accuracy based on the chosen paradigm. The concept relates to only having one of the two models running at any point in time, with a threshold value of certain event triggering a swap between the models used.

In this case, where to focus is on utilising the ABM paradigm during the beginning stages of an epidemic and the equation-based paradigm during the peak infections, the most logical choice is to utilise coarse-graining or model-swapping. Meaning the two models each work completely separate from each other, and only communicate when the epidemic in question swaps stages, leading to a swap in the model to use. This concept is illustrated in figure 2.

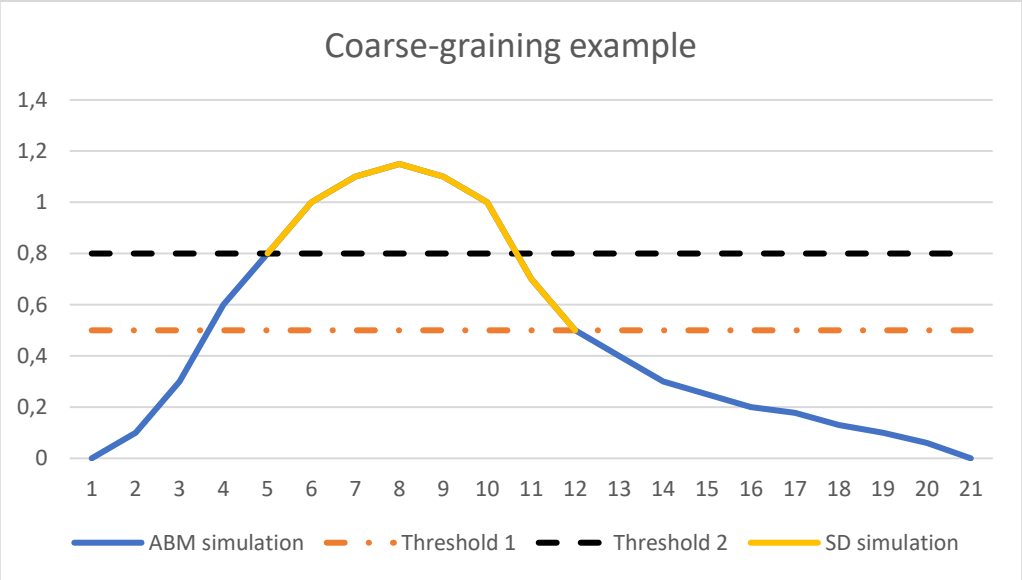

Figure 2: Coarse-graining concept

This swapping process itself will decrease the simulation speed somewhat, as some extra computations need to be performed, however this loss should be offset by a considerable margin due the speed increase gained from utilising the equation-based paradigm. This potential speed increase is only relevant, if meanwhile the results stay consistent, as there would be no point in utilising this type of multi-modelling if it would result in significantly less accurate results. Model-swapping, or dynamic coarse-graining, is a fairly new area of research, resulting in a limited amount of previous work, which will be discussed in the next section.

2.4.1. Previous coarse-graining work

One of the first attempts to actually combine the advantages of agent-based and equation-based approaches was done in 2007 by Bobashev et al.. They combined and ABM model and Equation based model based on a threshold switch, that was based on a stabilisation of a disease transmission parameter. This approach yielded good results, in terms of consistency, but required a fairly simplified model without any behavioural changes during the epidemic. Behavioural changes would alter the transmission parameter, and thus prevent the stabilisation. They did not look into the effects of the technique on the simulation speed.

A paper by Gray and Wotherspoon (2012), introduced the coarse-graining concept to contaminants in marine life models. They highlight the issues that transitioning between models can cause especially regarding fine-scale position data, that is lost. They argue that a plausible position for each individual can be reconstructed using the population distribution function, as there is no behavioural change associated with contaminant load. They found that coarse-graining the model significantly increases computational speed. However, their population's values suggest that a few variables behave quite differently under the coarse-graining. This "error" is not further investigated, but already showcases the difficulties of utilising this technique.

In 2015 Gray and Wotherspoon, wrote another paper regarding adaptive hybrid modelling. One of the key issues they discuss in the paper is regarding a suitable mathematical representation which allows for comparison of model configurations. They do not introduce a concrete method for this comparison, but do coin the term fidelity, to indicate the quality of a model's performance. Fidelity can be seen as a measure of the difference between real world results and the model's results in terms of distribution. The issue with using a metric that can only compare the results to reality, is that only models of real specific systems can be evaluated.

Wallentin and Neuwirth (2017) utilised the dynamic coarse-graining technique, when simulating the dynamics between fish and plankton. They specifically compared different types of models within this hybridised simulation. Like Bobashev et al. (2007) they also utilised emergence-

based aggregation, as the switching condition. Among their findings was that simulation time does not necessarily linearly correlate with aggregation level. With some less aggregated models outperforming more aggregated models of the same system in terms of speed. They also argue the emergence-based switches can connect spatial and hierarchical scale levels, while minimising loss of relevant information.

Based on these papers it can be concluded that the dynamic coarse-graining principle certainly has the capabilities to increase the simulation speed and also to stay consistent under specific circumstances. However, the combination of creating a fast, consistent and complex coarse-grained model has yet to be achieved. Since the studies that found significant speed increases, lacked in the consistency department, and the studies that found consistency restricted behavioural changes during the switching process. Lastly, there also seems to be a lack of a concrete method to determine the actual consistency between model results.

2.5. Experimental setup

In this study we will attempt to dynamically coarse-grain two separate epidemiological transmission models. First a very basic model will be utilised. The basic or simple ABM model, as we will call it, can be viewed as a test-case. The main benefit of first attempting to coarse-grain this Simple model is related to the identification of potential complications. The potential complications of the coarse-graining process are easier to be identified, as these complications are less likely to be buried beneath the complexity of the model. Additionally, the result of the Simple model can potentially lead to more insights when they are analysed in conjunction with a more complex model. To ensure the simplicity of the model it is important, that it cannot exhibit behavioural changes during the switching process, similarly to the model used by Bobashev et al. (2007).

Besides the Simple model there will also be an attempt made to coarse-grain a more extensive and complex ABM model. This Extensive model will be a SARS-CoV-2 specific model, include more complex concept, and exhibit potential behavioural changes during the switching process. The inclusion of these properties is necessary in order to be able to examine the ability of SARS-CoV-2 models to be coarse-grained. For both ABM models the Netlogo software will be utilised (Wilensky, 1999).

Once these models have been selected on the aforementioned criteria they first need to be translated. For this translation a rule set of sorts will be developed. These rules will primarily be based on a paper by Borshev & Filippov (2004). This paper describes how a SD model can be translated towards an ABM model, and its primary aim is to function as a practical reference for those who have prior experience with SD, and potentially want to supplement their models with an ABM add-on. Even though, this paper is focussed on creating an ABM model from an existing SD model, the principles it introduced should also be able to be used the other way around. This reverse engineering of their concepts slightly complicates the translation, however there is no paper available that concretely explains how an ABM model can be translated to the SD paradigm. Once this rule set has been completed the ABM Netlogo models will be translated into SD models using the Vensim software (Ventana Systems Inc., 2011).

The two versions of both models should each be capable of running completely autonomously, however in order to actually coarse-grain and be able to take advantage of both the accuracy of the ABM model and the speed of the SD model they will need to be connected. This connection only needs to take place during the switching process itself, as there is no need for communication between the models during the simulation itself. This connection will be established with use of the python programming language (Rossum, 1995).

Before starting to use the models for experimentation it is important to ensure the models are valid, meaning they are fit for their purpose. Every model will have its own limitations, as they are always simplifications of reality (Watson, Doherty, & Christensen, 2013), so the question is not whether these limitations exist, but whether the model is “good enough” for the intended purpose. These validations tests will focus primarily on the structure of the models, as a coherent structure is

vital for the coarse-graining process. This is not a trivial matter, as a valid ABM transmission model will most likely contain structures, which would be very demanding to accurately replicate in the equation-based paradigm. Additionally, the models' ability to exhibit the expected disease dynamics will be examined.

Once the models are deemed valid, the actual experimentation can commence. For this experimentation a form of exploratory modelling will be utilised (Kwakkel, 2017). The basic principle of exploratory modelling is to vary the values of uncertain variables, in such a range that policies can be tested for almost all possible scenarios, enabling testing the robustness of such policies. Or in case of this research to ensure the model functions as intended under all circumstances, instead of only under a pre-defined specific circumstance. Utilising uniform distributed Monte Carlo sampling a large amount of computational experiments can be generated (Shapiro, 2003). The advantage of using Monte Carlo sampling, over for instance Latin Hypercube sampling (Stein, 1987), is that the computational experiments are independent from each other. This allows for an expansion of computational experiments when required. Additionally, every computational experiment will be replicated multiple times, only saving the average results of these replications, to mitigate the stochasticity present in the ABM models.

Lastly, all the outcomes will be analysed and discussed. The analysis will focus on two parts of the outcome. First the increase in simulation speed will be examined by comparing the time it took for all the experiments to complete between both the ABM and the coarse-grained model for both the Simple and Extensive models. Secondly, the consistency of the coarse-graining results will be analysed. This will be done by quantifying the differences in the time series data for the Key Performance Indicators (KPIs) of both models. This quantified data will then be used to determine the percentage of computational experiments that result in consistent outcome, indicating the overall consistency of the coarse-graining process. Additionally, it can also be used to identify any potential uncertain variable values that results in less consistent result and are thus problematic when trying to coarse-grain a transmission model. An overview of all steps taken in this research can be found in figure 3.

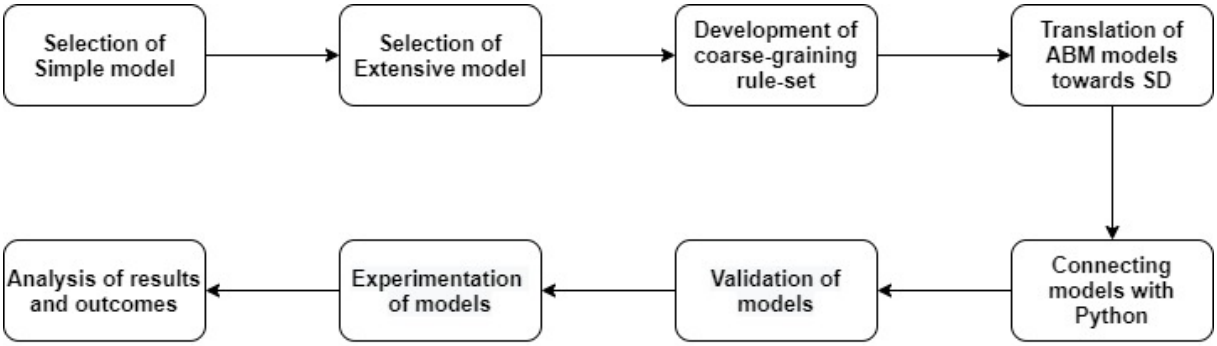

Figure 3: Overview of research steps

2.6. Comparing dynamically coarse-grained models

After the coarse-grained model has been developed, the model performance has to be evaluated. This evaluation will be based on the two main factors of interest due to the coarse-graining: the speed of the simulation and the accuracy of the simulation.

The main reason of utilising the coarse-graining technique is to improve the simulation speed of the model, as indicated by the third sub-question. This aspect can relatively easily be compared, by running both models for a large number of replications, utilising the same hardware, and tracking the time it took for the simulation to complete. By assessing the time, it took to fully complete the simulations it can be concluded whether the switch of paradigm will actually save time. However, if it turns out the coarse-gained model is actually slower than the original ABM model, it does not necessarily mean coarse-graining cannot be faster. This could, for instance, also be caused by inefficient coding. In that case the time it takes for specific parts of the simulation to run can be

analysed. That data can then be used to extrapolate, whether the simulation would potentially be faster on more extensive models. This could be the case when, for instance, the time saved by decreasing the computational burden does not weigh up to the time cost of making the switch itself.

Regardless of whether the coarse-grained model simulates faster or not, the coarse-graining process would be quite obsolete if it were to mean that the results are inconsistent with the original model. In order to adequately tackle sub-question 4, it is important to know how results acquired via different modelling paradigms can be compared to each other.

In order to determine the best method of comparing results between paradigms, the nature of the results must first be established. In the case of both models the data will be in the form of time series data, showcasing the evolution of the virus spread over time.

In a 2013 study performed by Figueredo, Siebers, and Aickelin they compared the results of ordinary differential equation (ODE) models, to an equivalent ABM model. They use the Wilcoxon rank-sum test (Rosner, Glynn, & Ting Lee, 2003). to determine whether the output of both models is statistically the same or not. They argue the Wilcoxon test is robust when the populations are not normally distributed (Rosner et al., 2003), which is the case for results gathered by using ABM and ODE models, whereas tests like the t-test could result in imprecise findings, as their performance suffers when the distribution does not follow the normal distribution. However, it is entirely possible that the Wilcoxon test will reject the similarity due to some paradigm differences. For instance, due to the continuous nature of ODE models most stocks will probably never fully reach zero once they have gone above it, in contrast to the ABM model, which does not exhibit this problem, as it is discrete in nature.

Auping, Pruyt and Kwakkel (2014), attempted to compare the results of three SD models of the global copper system, but with some structural differences between the models. They use absolute differences between trajectories of the runs to estimate the average difference between the runs. The main issue with this method of comparison is related potential oscillatory desynchronisation, which could result in counter phase behaviour. When relatively similar oscillatory behavioural modes are then compared, it might result in large differences. Since, epidemiological transmission models are characterised by oscillatory behavioural patterns Bjørnstad et al. (2021), this issue can be quite problematic.

Another method of assessing the similarity of time series data is by utilising the Complex Invariant Distance (CID) clustering algorithm (Batista et al., 2013; Steinmann et al., 2020). This algorithm utilises a complexity correction factor combined with the Euclidean distance between two time series to assess the similarity of the two time-series. The correction factor is calculated by comparing the length of the two time-series after they have been stretched into a straight line. If the stretched-out lines are equal in terms of length, the lines are equally complex, resulting in a correction factor of 1. This principle is illustrated in figure 4.

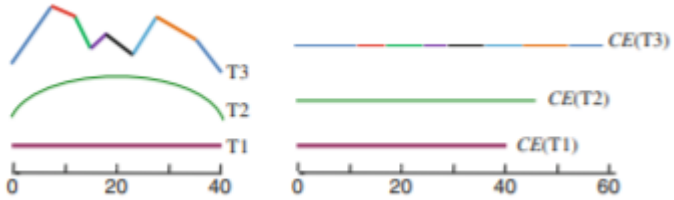

Figure 4: Complexity differences (Batista et al., 2013)

By comparing these complexity differences the issue related to oscillatory desynchronisation is largely resolved, since these desynchronised lines will be very similar in terms of overall complexity. Due to the importance of this property, the CID metric is deemed the best fit for this particular research. This does not mean other properties related to lines are irrelevant in regard to this subject, but since most of these properties, like for instance steepness and curvature, also influence the complexity of the lines and the Euclidean distance between the lines, these properties will not be examined explicitly.

The CID calculation will result in a distance measure, which indicates the similarity of two time series data. This measure cannot be interpreted directly, as it is also dependent on both the scale of time series data and the length of the data. Therefore, the data will be normalised by dividing it by both the average value of the dataset and the amount of data points in the dataset.

It is important that the coarse-graining process works under as many instances as possible, and not only for one pre-defined combination of variables, as a good scientist can draw an elephant with only three parameters (Seidenberg, 1993). Therefore, the model needs to be analysed for a wide range of uncertainties, ensuring its usability under all circumstances.

This normalised data can be used in several analysis. Firstly, it can be used to ascertain the percentage of runs above a pre-defined similarity threshold. This will indicate the overall performance of the coarse-graining process. Secondly, it can also be combined with the uncertainty data in order to determine under what conditions the dynamically coarse-grained model is potentially not able to reach the same results as the ABM model. This information can then be used to either make further improvement to the model or describe situations where the coarse-graining cannot be utilised in a satisfactory manner.

2.7. Outcome analysis

In order to analyse the results of the open exploration a PRIM inspired outcome analysis will be performed (Friedman & Fisher, 1999). The goal of this analysis is to find variables values that more often result in poor NCID values. Indicating either the existence of a bug in the program somewhere, or a variable that is not easily translated between paradigms, and therefore more difficult to correctly analyse using the dynamic coarse-graining technique.

The analysis works by first dividing all non-categorical variables up into 5 categories. Since the computational experiments are created by using a uniform distribution, the subdivision is made based on their value, as this will roughly results in categories of equal sizes. For instance, if the uncertainty range for a variable is between 0 and 100, roughly 20% of values should be between the values of 80-100. The NCID values are not a result of a uniform distribution, therefore these variables will be categorised in quintiles, based on their values. The analysis will therefore be based on relativeness of results instead of absolute results. The disadvantage of this is that values can vary a lot in absolute terms. For instance, if all NCID values are very good except 5%, the worst quantile will include some very good NCID values, but still the worst very good NCID values.

Once all variables have been categorised, they will get a score based on the amount of times they result in certain NCID quintiles. The score alterations related to the different categories of NCID values can be seen in table 1. Having both positive and negative score alterations is to account for variables that have less than five categories, the variables that are originally categorical. These variables will have more value instances per category, and would therefore always result in higher scores if this was not accounted for.

Table 1. Score alterations based on NCID category

Categorical NCID values	**Score alteration**
0 (best 20% of values)	-2
1	-1
2	0
3	+1
4 (worst 20% of values)	+2

By comparing the final score, the variables and associated values that cause the most inconsistency can be identified. Additionally, potential linearity between scores of categories associated with the same variable, can help indicate whether there is a trend, or a potential high-score is more likely to have happened due to randomness. If a certain variable category reaches a high score this is an

indication of the dynamic coarse-graining resulting in different outcomes, in regards to the ABM model. The reason behind the inconsistency can then be attempted to be ascertained manually. By removing the computational experiments that include the problematic variable category from the analysis. Additional problematic variable categories can potentially be identified. If there are no problematic variables in the analysis the scores of all categorised variable values should be around 0.

3. Model

This chapter focusses on the development of the coarse-grained SARS-CoV-2 transmission model. In order to ensure a correct and usable translation of the complex SARS-CoV-2 transmission model, the Extensive model, there will first be an attempt to coarse-grain a relatively simple ABM transmission model, the Simple model which is discussed in Section 3.1. This will make it easier to identify potential complications, as these complications are less likely to be buried beneath the complexity of the model. Finding these complications in the Simple model, will then create an opportunity for anticipation, making it easier to cope with them when coarse-graining the Extensive model, discussed in Section 3.2. In regard to the Extensive model, the basic structure of the ABM model will be discussed first. Subsequently paragraph 3.3 will focus on how this structure will be translated into a System Dynamics structure. Sections 3.4 and 3.5 of this chapter will discuss the translated System dynamics models of the Simple and Extensive model respectively. The Section 3.6 will put the focus on how these two separate models communicate with each other, to essentially form one big coarse-grained model. The validation of the models will be discussed in Section 3.7, and the practical experimental setup will be discussed in Section 3.8. All used models, data and code are also openly available on https://github.com/MaxvEck17/Dynamic_coarse_graining.

3.1. Simple ABM transmission model

The Simple model used in this research is part of NetLogo's (Wilensky, 1999) standard library of models, and can be found under sample models, Biology, Virus (Wilensky, 1998). The basics of the model will be shortly discussed, for a more extensive explanation of the model the official documentation can be consulted.

The ABM model can be considered a simplified version of the classical SEIR models. The exposed and infectious groups are taken together, resulting in three sperate possible disease states for the agents, which will also proxy as the Key Performance Indicators (KPI) of this model. Agents can be susceptible (S), sick (E + I), and immune (R). The progression of the disease states is visualised in figure 5.

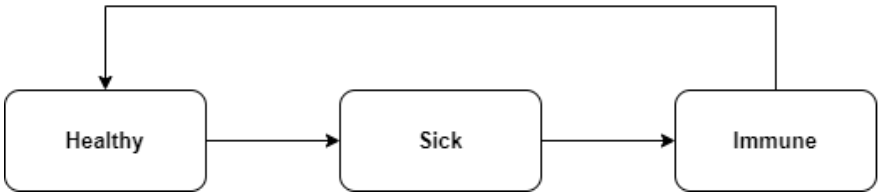

Figure 5: Disease progression in Simple ABM model

Additional to these main states the agents are also equipped with an age, which indicates when the agent will die, and internal timers that keep track for how many more ticks, which represents weeks, the agent should stay sick or immune. Infections have a pre-defined chance to take place between the sick and susceptible agents when they come in contact. Coming in contact means the spatial distance between two agents is smaller than a certain threshold. As the agents all move randomly through the model, the amount of infections will vary. Both the movement and infection chance add a certain amount of stochasticity to the model. The agent can die from either the virus or from reaching their total lifespan. Non-sick agents have a chance to reproduce as long as the total carrying capacity has not been reached yet. Depending on the settings used for the constants there are different modes of behaviour patterns possible. The virus has the potential to die out, however the most common behavioural mode is oscillating behaviour in regard to the amount of sick people. In figure 6 the flowchart of this behaviour is shown.

This model is deemed useable for the purpose of the Simple model as it exhibits the characteristics of SEIR models in their most basic forms. Furthermore, the model does not include

behavioural changes during different parts of the disease cycle, which is an important feature of the Simple model, as mentioned in Section 2.5.

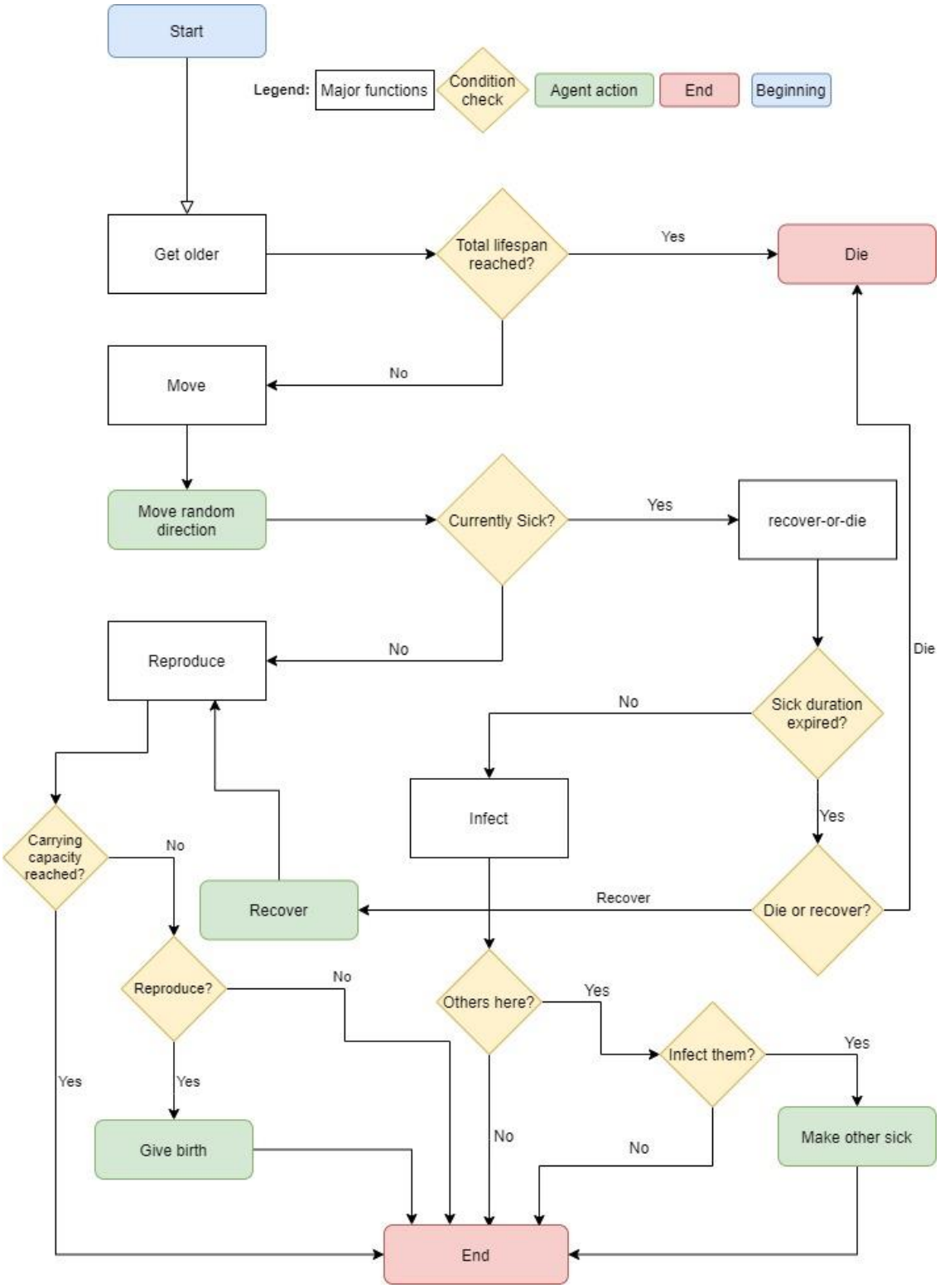

Figure 6: Flowchart Simple ABM model

3.2. Extensive Agent-Based model

The Agent-Based modelling (ABM) model has been developed by J. Badham (2021) and is primarily focussed on social, non-pharmaceutical, interventions and their influence on both transmission and

disease progression. Similar to the Simple model, the basics of the model will be shortly discussed, for a more extensive explanation of the model the official documentation can be consulted.

The basic structure consists of an extended person to person SEIR model, which excludes all other methods of infection. The agents, or people, in the model can have seven different epidemiological statuses. People start off as being susceptible (S), have a chance of becoming exposed (E) after coming in contact with an Infectious (I) individual, and become infectious themselves after a certain amount of time passes. While being infectious people can infect others and may experience symptoms. If these symptoms become severe people might need to be Hospitalised and in the most severe cases be moved to Critical care. All people that are in one of these three different infectious stages also have the possibility of recovery, making them immune, or in the worst case dying (R). Lastly, immune people have a chance of becoming susceptible again after a certain amount of time. This disease progression is visualised in figure 7.

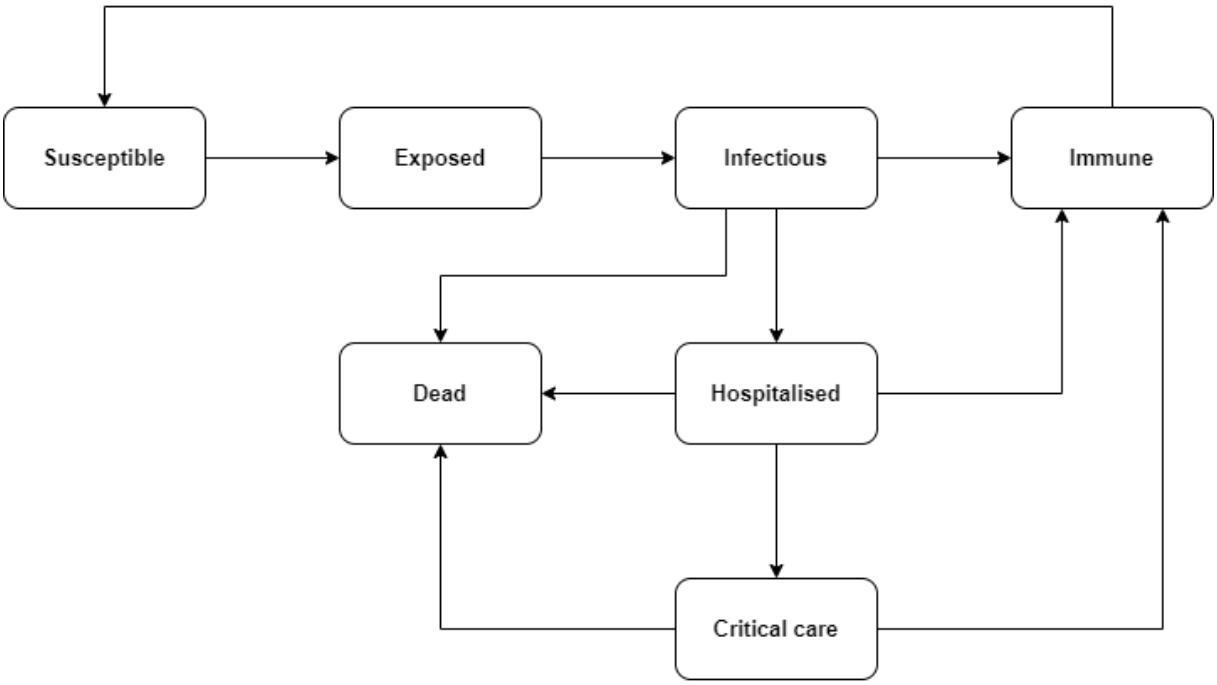

Figure 7: Disease progression in Extensive ABM model

The KPIs of this model are also related to the different stages of disease progression in the model. The three KPIs are: *Incidence, prevalence-I,* and *prevalence-all*. With *incidence* indicating the percentage of people that became exposed during the last tick, which represents a day. *Prevalence-I* is the percentage of people who are currently able to infect others, thus being either Infectious, Hospitalised or in Critical care. Lastly, *prevalence-all* indicates the percentage of agents who are currently infected with SARS-CoV-2, thus exposed, infectious, hospitalised or in critical care.

Transmissions are modelled by combining a chance of infection with spatial proximity. If a transmitter, an infectious person who is not isolated, comes in contact with a susceptible person a chance of transmission will be calculated based on a number of factors; like the age-groups the agents belong to, or possible social distancing interventions. When agents transition stage, both the time they will stay in the state, and their next state will be determined by a weighted draw. On onset of symptoms after infection they also have a probability to go into isolation, and potentially report their infectious state to the agents they already have infected. The movement that leads to these contacts is based on random draws, as there are no actual activities included in the model, which introduces a considerable stochastic component. A simplified version of how an agent goes through these interactions can be seen in the flowchart in figure 8.

Besides actions agents can take there are also global variables that affect most of these actions. These global variables can be seen as governmental interventions. Interventions can either

be activated for the entire duration of a simulation run or be trigger dependent. Interventions that run for the entire duration are the in essence the less extreme interventions, which are feasible to be maintained for a longer time period in a real population. An example of this would be to self-isolate oneself when starting to experience symptoms of the virus, or when being informed by an infected contact. The more intrusive, and trigger dependent, policies are mostly focussed on reducing the movement of agents. For instance, the high-risk-shielding intervention, puts every agent that has the high-risk state in isolation, regardless of current infection status. The trigger itself is activated if one of three variables; the amount of days passed since start of the outbreak, the amount of hospitalised people, or the amount of active cases, goes past a pre-defined threshold.

The model also includes an option of vaccinating the agents. This vaccination is not modelled in a gradual way, but instead as a button that needs to be pressed by the person who runs the simulation, to vaccinate a specified percentage of the agents. Since this is not an integrated process, there will not be any vaccinations during the coarse-graining process. However, this could be useful to determine the behaviour of SARS-CoV-2 when a large part of the population is already vaccinated at the start of a potential wave.

This model is deemed useable for the purpose of the Extensive model as it exhibits the characteristics of SEIR models in an extended form. It also includes all the required aforementioned characteristics as mentioned in Section 2.5. These characteristics include: being a SARS-CoV-2 specific model, including complex to translate concepts like individual tracking, and being able to exhibit behavioural changes during the switching process, which can be found in interventions that can be activated and are able to change the behavioural patterns.

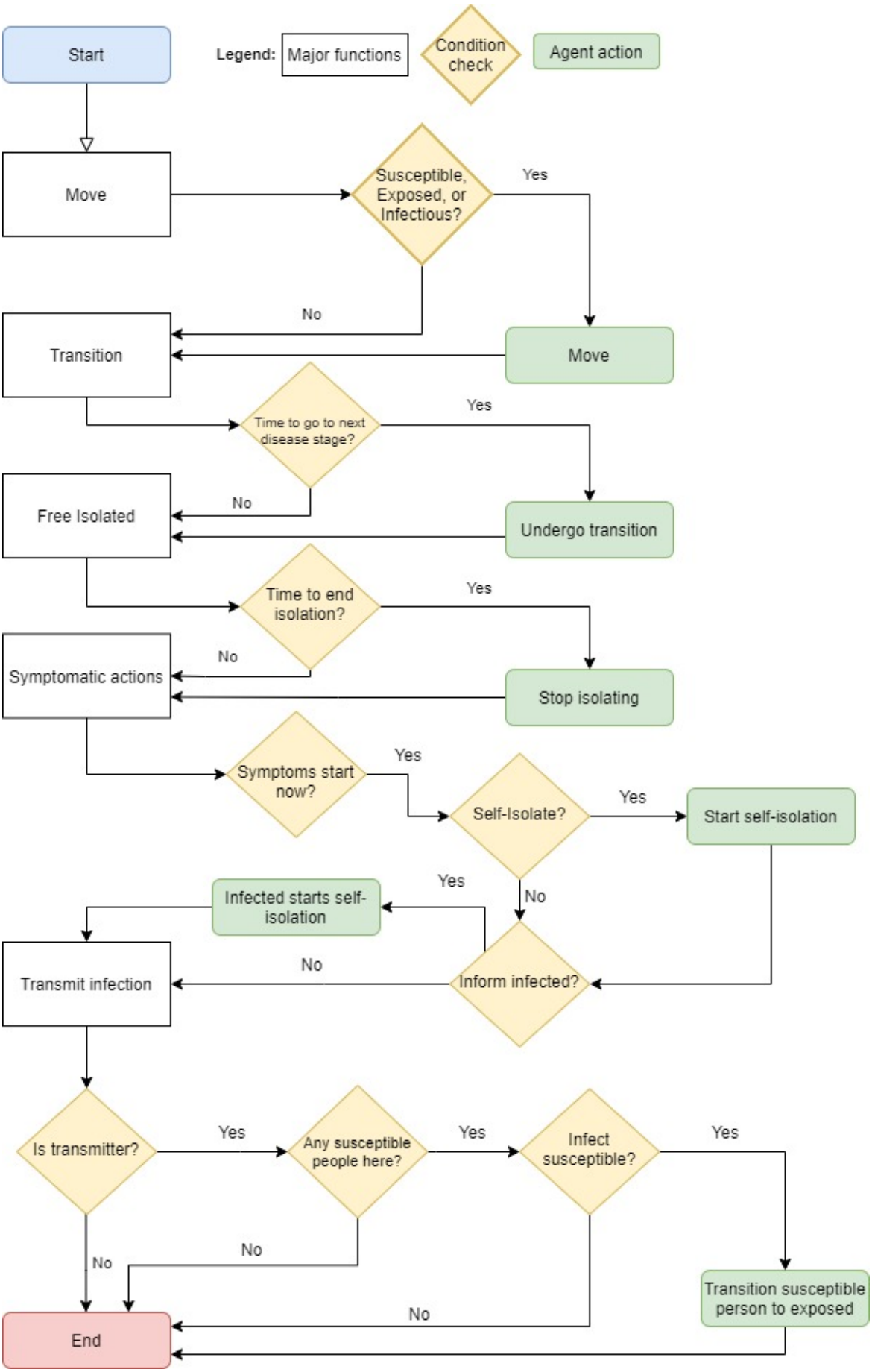

Figure 8: Simplified flowchart of Extensive ABM model

3.3. Coarse-grained modelling rule set

When combining the ABM and SD paradigm there are some issues that can occur, as the techniques were not developed with interconnection in mind. Making the translation of certain parts of the model more difficult than others. This could potentially be remedied somewhat if the original model, either ABM or SD, is developed with the other paradigm in mind. Processes in modelling almost never have one correct way to be modelled, but one way of modelling a certain process will allow it to be translated to the other paradigm more easily than an alternative way. By considering ways, it would be possible to translate it into the different paradigm, while making the original model a lot of issues could potentially be avoided. In this case, however, the original ABM models were not developed with a translation to SD in mind. Therefore, there is a need for a rough rule set of translating common, yet complex processes from ABM to SD. The goal of this section will hence be to resolve the second sub-question: "*What set of rules should be used for dynamic coarse-graining two SARS-CoV-2 transmission models in different modelling paradigms?"*

3.3.1. Basics of translating ABM models towards SD models

The basics of translating a model in the ABM paradigm to the SD paradigm will be based on the paper by Borshev & Filippov (2004). Their focus was on translating a SD model towards the ABM paradigm. Meaning their findings will essentially have to be reversed. The starting point will be the aggregation of agents and their different states into stocks. Because agents can theoretically have an infinite amount of states it is important to think about which states need their own stock, which states can be combined into the same stock, and which states are better of being subscripted. The flows between these different stocks can be seen as the transition between states agents might experience. It is also good practice to include every single variable that is included in the other model, assuming they are all used. Using the exact same variable name will also make the connection easier to make and help clarify what the original variables are and what the new variables are.

3.3.2. Translating complex concepts from ABM towards SD

Besides the easy to convert aspects of ABM models, there are some characteristics of ABM that SD does not have, which introduces some complications. The main issues are related to the lack of stochasticity, heterogeneity, and spatial effects in the SD paradigm. The most common occurrences of these aspects in ABM will be discussed. An overview of these concepts, potential solutions and chosen solution is shown in table 2.

Stochasticity in ABM is very common, nevertheless most uses can be generalised. The most common usage of stochastic aspects is in chance calculations on a nominal scale. Since SD is non-discrete those chances can be converted directly to distributions. If an agent has a 50% chance to recover from a disease or die alternatively, instead of choosing one of the two options SD can let half of the agent recover and half of the agent die, due to its continuous nature. An extension on this principle is drawing options based on a chance distribution on an interval or ratio scale. Theoretically speaking this could be resolved in the same way as the nominal chance, however the amount of nominal options are usually limited, whereas interval or ratio numbers can be scaled up much easier. For instance, when asking an agent to perform an action after a certain amount of time, the different time options are essentially limitless. Therefore, it is impractical to create separate flows for every possibility, in this case it is more practical to take the average value of the distribution. However, when these distributions are used in delay functions, most distributions are not fully captured correctly. Borshev and Filippov (2004) note that exponential distributions are the only correct way to transform a delay, in extension this means that delays are only capable of perfectly mimicking exponential distributions, being slightly off on other kinds of distributions. In spite of this, there is no clear alternative method of modelling this, so we will stick to the average outcomes of these distributions.

A considerable reason to analyse epidemiological transmission systems with the ABM paradigm is for the inclusion of special effects that are an important factor in the spread of viruses.

These factors are very difficult to incorporate in a SD model, forming a substantial obstacle in the translation of these models. Depending on the complexity of the movement options of the agents, it can be difficult to almost impossible to mathematically determine the related variables. In transmission models the main method of utilising spatial effects is to model the infections, as agents need to be close to each other to be able to infect each other. Mathematical attempts to calculate what the odds are for a susceptible agent to encounter another an infectious agent are complicated. Even if these odds can be determined the heterogeneity of the agents put another spoke in the wheel. If there exists a constant that explains this behaviour, it is very complicated to determine it. It needs to be controlled for just the right variables, and a statistical analysis will hardly help in finding these variables due to the many feedback mechanisms in transmission models. For instance, the amount of infected people determines how likely it is for a susceptible person to encounter an infected person, but this chance also influences the amount of infected people there will be. A solution to this problem would be to empirically estimate a variable, chance of contact for instance, controlled for some variable likely to affect it, and create a look-up function dependent on the control variables. This is not ideal, as it will always stay an estimate, but can function adequately enough. However, this method will not work when alterations to variables underlying the movement do not scale this look-up function in an easy to identify fashion.

A last difficult to translate concept is the use of individual tracking. In the Extensive ABM model this is for instance used to track the agents that got infected by a pre-symptomatic agent. This allows them to get informed of their infection once the symptoms for their infector start. In SD there are essentially no individuals, but only groups, or aggregates of individuals, this prevents any and all forms of individual tracking. This should be able to be remedied by using the fraction of pre-symptomatic people to determine what fraction of exposed people should get informed of their infection. However, there can also be structures in the model that cause agents in certain age groups to have an altered chance of infecting agents in other age groups, such a structure can be found in the Extensive model in the form of the contact-age-homophily variable. Structures like this make the correct implementation of individual tracking, as they are easy to overlook when constructing the SD model.

Table 2. Overview of complex to translate concepts and solutions

Complex concept	**Potential solutions**	**Chosen solution**
Stochasticity	• Create separate stocks (or subscripts) for every possible outcome and let the flows follow the pre-defined distribution in a continuous manner. • Aggregate the stochasticity to its average outcome	Nominal/Ordinal distributions: create separate stocks Ratio/Interval distributions: use average value
Spatial movement of agents	• Calculate the chance of agents coming in contact with each other based on the movement patterns, and variables like the amount of agents, and the size of the space the agents live in. • Empirically estimate the amount of contacts occurring in the model, controlled for variables that are likely to influence this amount	Simple model: Calculate the chance, but combined with a slight empirically estimated correction factor as the calculated result, did not exactly match the original outcome Extensive model: Empirically estimate controlled for the amount of susceptible and infectious agents.
Individual tracking of agents	• Combine relevant fractions to mathematically calculate the results this tracking should have, without explicit modelling of the tracking itself	No explicit individual tracking, but calculations utilising the relevant variables to gain similar outcomes.

3.4. Simple SD transmission model

Using the coarse-graining ruleset, developed in Section 3.3, the Simple ABM model, described in 3.1.1, has been translated into a SD model, utilising the Vensim Software (Ventana Systems Inc., 2011). The stock-structure in the SD model, is built upon the different states the agents can take. This can be seen in the conceptual Stock Flow diagram in figure 9.

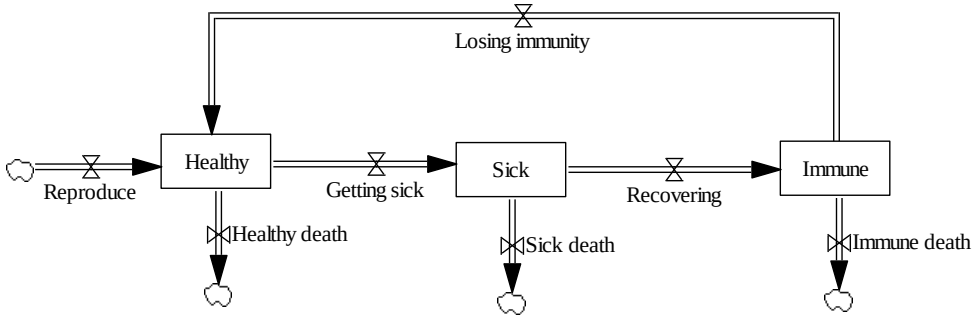

Figure 9: Conceptual Stock Flow diagram of Simple SD model

This diagram showcases the overlap with the agent states in the ABM model, as shown in Section 3.1.1, only adding outflows for deaths and an inflow for new-borns. Agents flow though these states by a third order delay of the average time they should spent in these stocks. However, being Healthy is not time dependent, but contact dependent. This creates the first issue, as the chance of coming in contact with another agent is purely spatial dependent and has a heavy stochastic component. This value should approach the amount of agents present in the model, thus the *Total population*, divided by the amount of space they live in, which can be interpreted as the total amount of patches present in the ABM model. A patch being the unit of one block of space. However, probably due to the heterogeneity of agents, this value was still off, making the use of an empirically found correction factor necessary.

A second issue also relates to the flows in the model. Because these flows are mostly modelled using third order delays and averages, any oscillatory behaviour is quickly lost. A delay function delays its inputs using an exponential distribution with a specified order and using the delay time as its average. By changing exact delay time to approximate delays time, the peaks will be smoothed out more, reducing the oscillatory behaviour. The main cause of this issue is the absence of stochasticity in the SD model. So, any singular replication performed in Netlogo instead of Vensim, will have drastically different results. Taking the average of multiple Netlogo replications will result in similar results to a singular replication of Vensim, meaning this is not necessarily something that needs changing.

The rest of the model was quite straightforward in terms of modelling. Most variables were able to be copied directly into the SD paradigm. All ratio and interval stochastic processes were changed to their average outcomes, and internal timers were translated in global chances to die or get past a certain state. A figure of the full model can be found in Appendix III.

3.5. Extensive System Dynamics model

Using the coarse-graining ruleset, developed in Section 3.3, the Extensive ABM model, described in Section 3.2, has been translated into a SD model, utilising the Vensim Software (Ventana Systems Inc., 2011). The basic concepts of the SD paradigm are described in Section 2.3.1. The stock-structure in the SD model, is mainly built on the different disease states agents can take. This can be seen in the conceptual model in figure 10.

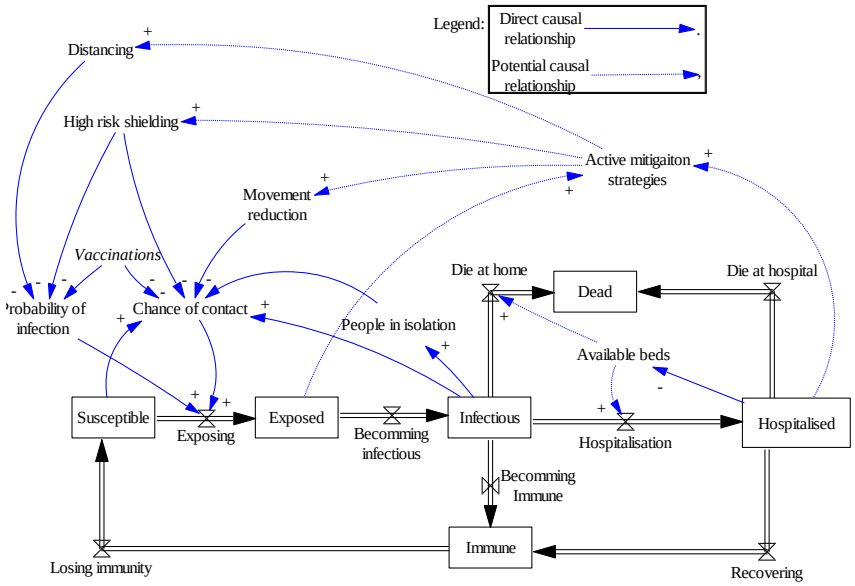

Figure 10: Conceptual Stock Flow diagram of Extensive SD model

The conceptual model is a very simplified version of the actual model, where the disease stages are broken up even further, including asymptomatic, pre-symptomatic and critical care. Besides the main disease stages, the main mitigation strategies are also included in the conceptual model. These strategies will either try to reduce the chance of a susceptible person to come in contact with an infectious person or reduce the probability of infection when they come in contact. The dashed lines indicate causal relationships, that can be turned on or off.

The other important agent states, present in the ABM model, are made into subscripts. Subscripts can be viewed as variables within other variables, which among other things helps in keeping some agent heterogeneity in the model. The subscripts used are; age-groups, risk-type, vaccination-type, and isolation status, creating 24 unique states within every disease state.

The flows between the stocks follow the same logic as in the ABM model. The time it takes to move on from one stock to the next is translated, from a weighted draw to the average outcome, and combined with a third order delay. As mentioned in Section 3.3.2. this is not a perfect translation of a weighted draw, as this changes the chance distribution into an exponential distribution.

One of the flows is not time dependent, and that is the *Exposing* flow between the *Susceptible* and *Exposed* stocks. This flow is dependent on the variable that was most difficult to quantify, the chance of contact between a susceptible person and a transmitter. This crucial variable is influenced by both the heterogeneity and the spatial aspects of the ABM model, making it hard to find a constant. When the variable was controlled for both the movement variables, the amount of susceptible people and the amount of transmitters, it still was far from constant in the ABM model. An attempt at a statistical analysis of the variable, also did not yield any useable results, due to the many feedback mechanisms present in this part of the model. Therefore, the decision to empirically estimate the variable was made. This was possible as the variable did consistently show about the same results, when also taking the stochasticity associated with this variable into account. This solution came with additional issues, as the resulting look-up function did not scale in an easily identifiable manner with the movement variables it should be dependent upon.

Another change that was made compared to the ABM model is related to the isolation of the agents. In the ABM model agents remember which agents they have infected, and more importantly stay isolated for a prespecified amount of time, independent from disease state. This makes it very difficult to make sure the people move through the disease model in a correct manner when they go in isolation halfway through. To remedy this, there was decided to couple the processes, meaning

people will only and always leave isolation when they either die, become immune or get admitted to a hospital. This change has the potential to result in an underestimate of the infections, as people that get notified of their infection in an early stage, for instance when they have just been exposed, could have already left their isolation before they have actually become immune. This also means the isolation duration will not be able to vary between runs, as the SD model is not capable of having shorter or longer isolation periods.

The last simplification that was made is in regard to the people who lose their immunity. In the Netlogo model the amount of episodes every agent experiences is recorded. However, since SD does not allow for tracking of individuals it is unable to say whether an individual has gone through the whole process multiple times. Therefore, it only tracks the amount of susceptible people who never have been infected. This will make it possible to determine the total impact the virus has had on the population. Every other part of the original model has been faithfully recreated, which should result in similar outputs for the two models. Especially for the key performance indicators (KPIs) of the model, which in this case are; The total prevalence of the virus (*prevalence-all*), the prevalence of infectious people (*prevalence-I*), and the amount of new people who get infected every day (*incidence*).

3.6. Interconnection of models

The two models are each capable of running completely autonomously, however in order to actually coarse-grain and be able to take advantage of both the accuracy of the ABM model and the speed of the SD model they will need to be connected. This connection only needs to take place during the switching process itself, as there is no need for communication between the models during the simulation itself. To accomplish this connection use of the python programming language will be made. The coarse-graining process will always start in the ABM model, as this is the best method of subdividing the agents over all stages. We can subdivide the swapping process in two separate processes, from ABM to SD and from SD to ABM, that each have their own set of steps to be taken. In order to account for the swapping process, there also need to be made a few alterations to the models themselves. In Section 3.6.1. the interconnection of the Simple model will be discussed, and in Section 3.6.2. the interconnection of the Extensive model. An overview of the complicated issues related to the interconnection of both models can be found in table 3, below Section 3.6.2.

3.6.1. Interconnection of the Simple model

As one of the goals of this research is to see what speed improvement can result from utilising the coarse-graining technique, it is important the interconnection of the two models does not take up to much extra computational power. To this end it is important there is no constant connection between the models. In order to achieve this, there needs to be a variable stopping condition added to the ABM model. This function is implemented by creating a "*go_while*" procedure, which will run the model until either the switching condition is met, the amount of sick turtles has become zero or the pre-defined runtime is exceeded.

Once the ABM model has stopped running, it will return the relevant data to python (Rossum, 1995). This data will subsequently be used to specify all constants and initial values in the SD model. To reduce the chance of an error occurring, all constants are overwritten to the most recent values in the ABM model, even those that are not supposed to vary.

At this point the SD model will start running, starting from the point the ABM model stopped. Similarly, to the ABM model, the SD model also has a build-in stopping condition, preventing the waste of time by simulating the model longer than necessary. The stopping condition in the SD model will always be half of the stopping condition in the ABM model. This prevents rapid switching between the models if the model is closely oscillating around the switching condition, as multiple unnecessary switches can quickly increase the required simulation time significantly.

After the model falls below the switching condition, it will return the relevant data to python and prepare the ABM model for taking over. This is a bit more complicated, due to the paradigm

differences. There are three major issues that need to be addressed. Firstly, the continuous values need to be discretised. This is only an issue when switching form SD towards ABM, as all discrete values lay on the continuous spectrum, but not all continuous values lay on the discrete spectrum. For the Simple model this issue is not too difficult to solve, as the only values that need to be transferred are the amount of Sick, Immune and Total people. Which can just be rounded to their nearest number. However, this will create a sudden yet subtle change of values in the data, which could potentially be seen in graphs of the data.

The second issue relates to the spatial aspect of ABM models. The spatial information is largely lost when the swap from the ABM towards the SD is made. When returning to ABM, all agents need to have a spatial position that results in reasonable behaviour, in regard to in which phase of the transmission cycle the model exists. If, for instance, all agents would be put on the same place, the amount of infections that occur will be significantly higher than could be reasonably expected. In this Simple model the movement of the agents is completely random, independent from any variables and large enough to get through the entire "map". The combination of these factors make it possible to put the agents back on the map in a completely random fashion. This will be done in the same fashion as at the start of the simulation by using the setup button.

The last issue is related to the internal timers the agents have. These internal timers are used to determine when agents are supposed to perform a certain action. In the case of the Simple model internal timers ascertain whether turtles should change their status from Sick to Immune/Dead or from Immune to Susceptible and when they should die. The easiest way of resolving this issue is by setting their internal timers for a random time period between zero and the maximum time. This solution can possibly create issues when spatial effects and heterogeneity play an important part, as it disregards whether it is more likely, that the agent was recently infected, or has been infected for a long time already. Due to the simplistic nature of this model, this does not seem to affect the results in any noticeable manner.

These two models will keep switching between each other until one of them reaches the total run-time, or the amount of sick people is reduced to zero. In this latter case the SD model will run until completion. Once the simulations are completed the partial runs from both models are combined and saved for potential analyses. An overview of the pseudocode underlying this process can be seen in figure 11.

```
def coarse_graining():
    set constants to ABM_model
    run ABM_model until (runtime, switchcondition, virus died out)
    Save ABM data
    if runtime is reached --> report saved data
    else set variables to SD_Model based on saved data
    run SD_model until (runtime, switchcondition)
    Save SD Data
    Combine Data
    if runtime is reached --> report saved data
    else: while runtime is not reached
        set variables to ABM_model based on saved data
        Run ABM model until (runtime, switchcondition, virus died out)
        save ABM data
        Combine Data
        if runtime is reached --> report saved data
        else set variables to SD_Model based on saved data
        run SD_model until (runtime, switchcondition)
        Save SD Data
        Combine Data
        if runtime is reached --> report saved data
```

Figure 11: Pseudo code of coarse-graining process for both models.

3.6.2. Interconnection of the Extensive model

Similar to the Simple model, both the ABM and SD model need to have a built-in switching condition to minimise the computational power required for the interconnection of the models. This will make the ABM model simulate until either, the amount of infectious people reaches a certain threshold, the full run-time is completed or there are no more infected people in the model. Additionally, to these requirements, the model will also continue running when new mitigation measures have just been activated or the simulation has almost reached its full run-time. The respective reasons are that the coarse-graining preforms poorly when switching shortly after activation of mitigation measures, and if the simulation is almost completed it will take more time to actually switch models than to have the ABM model run its course.

Once the ABM model has first reached its switching condition all relevant data is saved, and all variables are copied and set to the Vensim model. Some variable values need to be translated due to the paradigm switch, as ABM allows for variables to hold String values and Booleans, where SD only utilises numbers. Additionally, percentages are used differently which means they sometimes need to be divided by a hundred when making the transition. Besides all real variables, that need to be set there is also a correction factor present in the model. This variable is used to adjust the look-up movement variable to be more consistent with the ABM model. It calculates ratio between the last amount of contacts per transmitter per susceptible person and the look-up of the same variable and multiplies the look-up variable by this ratio. This has empirically shown to yield more similar results than solely utilising the look-up function.

After all variables in the SD model have been set to their correct values, its simulation can commence. This simulation will last until "*Prevalence-I*" falls below the 0.013 threshold, and the model has to at least be running for a prespecified amount of days, 7 in this case. This prevents the coarse-graining process from oscillating between the two models. When the simulation is completed all its data is saved in a run file.

The data in this run file is used to make the switch back to the ABM paradigm. Swapping back is a fair bit more complicated as information not present in the SD model needs to put into the ABM model. Unlike the Simple model it is crucial the Extensive ABM model is not reloaded, setup again, or changed in any way shape or form before switching back. By keeping the original ABM model as is, during the Vensim simulation, the information still contained in the model can be used in the switching process. This is especially useful, as the amount of agents in the model stays constant during the run, and most agent states, like age-group, also do not change.

Keeping the previous model will thus save the locality and heterogeneity of agents. This is important, as when all agents are put back in a homogenous or random manner, the amount of infections will be greatly overestimated after the switch. An attempt to preserve the heterogeneity present in the original model is made by utilising the initial infected agent. Due to the spatial manner of infecting other agents, the agents closest to the original agent should have gotten infected first and be the furthest in the infection cycle. Therefore, agents will gain a state based on their own original constant attributes (risk-type, age, and vaccination-status), and their proximity to the original agent.

The three states, the closest agents are theoretically most likely to have, are; to be immune, dead, or susceptible again. To accommodate for this, the amount of agents in these groups are tagged based on their proximity to the original agent. The tagged agents will then be made dead, immune, or susceptible, based on their permanent states and random chance. Once these three groups have been put back in the model, the other disease stages will also be put back based in a similar manner. The order in which they will be returned is; critical, hospitalised, severe symptomatic, mild symptomatic, asymptomatic, pre-symptomatic, exposed, and finally susceptible. This will result in a circular pattern around the initial infected agent.

This will not perfectly mimic the original situation, but probably be a close approximation. The only part that is disregarded, is that not every agent in the model will have gone through the virus process. So in between the sick and recovered agents, there should be some agents scattered who

managed to dodge the virus. To accommodate for this all susceptible agents who have not gone through the virus yet will be placed first. Fifty percent of these agents will first take a random spot of a random agents that belongs to fifty percent of closed agents with the other fifty percent being returned completely random. These agents will be locked away from being able to get infected during the transformation process. A rough estimation of how this should look can be seen in figure 12.

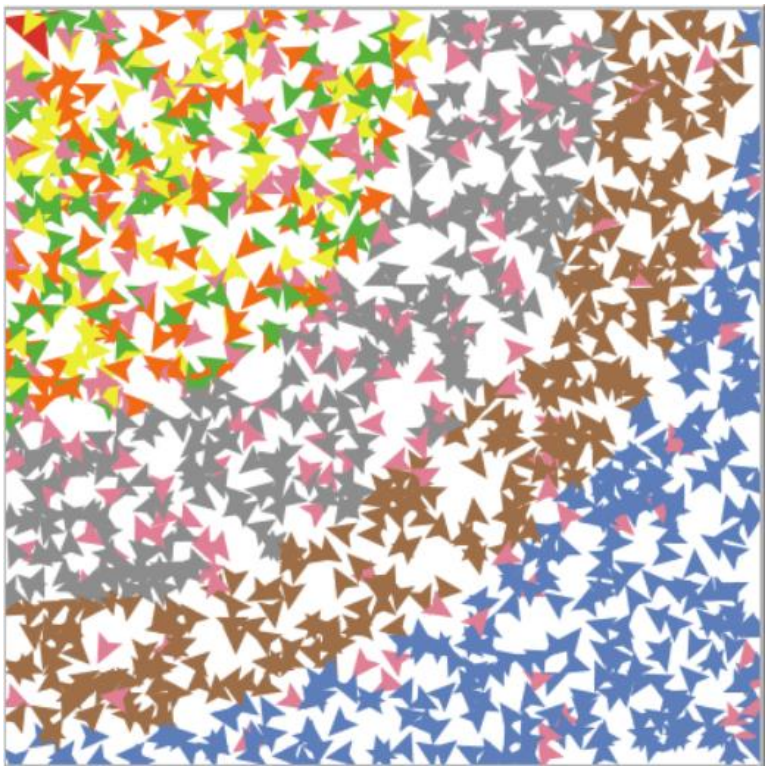

Figure 12: Rough estimation of ABM transformation result

Another issue with the transformation process of the ABM model is the discretisation of values. Since the amount of agents with specific attributes is constant, and thus unchanging, situations can occur where there are not enough agents left to fit a disease state. This happens when the disease states that were put back first were mostly rounded-up. This will slightly alter the composition of the agents in terms of disease state, however due to the large amount of agents present in the model, the impact should be minimal.

Dependent on which disease stage the agents get during the transformation process they need to have different set of timers and attributes set. To accomplish this, there are distinct procedures established. The only state not able to be somewhat reproduced is "*my- infected*", which holds the data on what agents each transmitter has infected. This is only relevant for the pre-symptomatic agents, as they could tell their infected agents to go in isolation once they become symptomatic. However, since agents only stay pre-symptomatic for a very short time, this should not have a big impact on the overall model behaviour. Especially as the model will be in a downwards trend when making the switch, resulting in a very low proportion of pre-symptomatic agents.

Table 3. Overview of complications and solutions related to interconnection of models

Issue	**Solution**
Ensure minimal computational requirements during switching process	Embed switching conditions within the models themselves, to reduce communication during simulation.
Ensure variables keep their correct value	Overwrite every variable every switch including constants as a failsafe.

Oscillatory switching between models	Set the switching condition for switching back lower than other the switch condition. Prevent model from switching back the first few time steps after switching.
Continuous variables need to be discretised	Round continuous variable to the closest discrete value.
Loss of spatial information in model	**Simple model**: re-setup the model with correct amount of agents and alter attributes randomly. **Extensive model**: save original model and only change their attributes related to their disease state. Combined with an algorithm that determines what agents need to change based on their proximity to the original infected agent.
Internal timers of agents	Set the time to a random time between the maximum value of the timer and 0.
Switching after activation of intervention	Prevent switching at this time as it results in poorer results.
Loss of agent tracking information	Tracking is only relevant for pre-symptomatic agents, which are a very small group at the time of switching back, thus ignoring this has only a very minor impact on the model.

3.7. Verification and validation

Before starting to use the models for experimentation it is important to ensure the models are valid, meaning they are fit for their purpose. Every model will have its own limitations, as they are always simplifications of reality (Watson, Doherty, & Christensen, 2013), so the question is not whether these limitations exist, but whether the model is "good enough" for the intended purpose. For instance, you should not judge a fish on its ability fly. So, depending on what the model is intended to be used for, the set of validation tests will vary. In Section 3.7.1. the purpose of the four models will first be discussed, followed by the validation tests, and the conclusion regarding their validity.

3.7.1. Purpose of the models

In order to validate any model, it is paramount that the purpose of the model is clearly defined. To this end the purpose of this research needs to be analysed. In this case the research is mainly of a methodological nature, as the models are not intended to give any direct policy advice. This means the models will need to be able to function in such a way the use of the new methodology can be analysed. To this end, the purpose of the SD models is purely whether they able to mimic the results of their related ABM model under certain conditions. So, it could be argued that the validity of the SD models is the heart of the research. Therefore, the validity of these models should theoretically only be based on its similarity to the ABM model, which behavioural components will be discussed further in chapters 4 and 5. However, these models are in essence also still epidemiological transmission models, so they will also be judged upon their ability to exhibit the standard disease mechanics expected from such models.

The same holds true for the ABM model to some extent. When looking at the research in a vacuum, for the coarse-graining it is not really relevant whether the behaviour exhibited by the ABM models is a good representation of real disease dynamics. The primary purpose of both models is to include concepts that are either simple or complex to translate to the SD paradigm. If these concepts, like stochasticity for instance, are included in the ABM model and the SD model is able to replicate it, the coarse-graining was successful, regardless of the behaviour that is actually exhibited. However, since the goal of the research is to specifically look at coarse-graining SARS-CoV-2 models, it could be argued that the behaviour displayed by the ABM models is relevant to this end. Therefore, the ABM

models will also be validated upon their ability to exhibit behaviour expected from such epidemiological transmission models.

3.7.2. Validation of Simple model

Based on the purpose of the both the ABM and SD version of the Simple model explained in Section 3.7.1, a selection of behavioural and structural tests will be performed. Starting with the ABM model it needs to include concepts deemed not very complex to translate to the SD paradigm, the basic structure of epidemiological transmission models and exhibit basic disease dynamics. The SD model structure will be validated based in unison with the ABM model, as they should be perfectly aligned due to the coarse-graining. The behaviour of the SD model will also separately be tested on its ability to exhibit basic disease dynamics.

3.7.2.1. Parameter verification test of Simple model

The parameter verification test entails the comparison of the different variables in the model to their real world counter-part. Ideally every variable should have a distinctive meaning and a logical value attached to them. If there are a significant amount of variables without such meaning, this could be an indication of overfitting a model, meaning the model is less likely to be an accurate representation of reality. The extended execution of this test is located in Appendix I**.** It can be concluded that, even though not every single variable has an established real-world counterpart, the parameters make enough sense for the purpose of this research.

3.7.2.2. Structure verification test

Verifying the structure of the Simple models will be the most important test in regard to the purpose of the model. Since this model is purely utilised to serve as a test-case for the Extensive model, the most important thing is that the basic structure follows the structure of standard transmission models. If this is not the case, it could be argued that using this model as the test-case would be senseless. In Appendix I is discussed how the Simple ABM model incorporates every attribute to make it a viable option in regard to the methodological goal.

3.7.2.3. Dimensional-consistency

This test will help to build confidence in the internal consistency of the SD model. Since most variables in the Netlogo model do not have dimensions attached to them, this test will only be performed on the SD side of the Simple model. Consistency of the dimensions in the Vensim model also implicates, that the dimensions of the Netlogo model variables are consistent, as the functions are essentially the same. Dimensions failing to add up implicates a structural mistake in the model and should therefore be analysed carefully. This test is especially useful in conjunction with the parameter verification test, as this eliminates the variables used to fix the dimensions without having any real-life meaning. Based on the information provided in Appendix I it can be concluded that both models are dimensionally consistent, building a significant amount of confidence in the underlying structure of the model.

3.7.2.4. Extreme conditions test

To ensure the models behaviour does not exhibit any impossibilities, the model is tested under conditions that are most likely to break the model. For instance, when putting the carrying-capacity to zero the amount of agents in the model should never increase and always decline. More obvious mistakes, like variable going negative, are also indicators of failing this test. By passing this test, the likelihood of the basic structure functioning correctly is increased. The full implementation of the extreme conditions test, for both Simple models, can be found in Appendix I. Based on these results it can be seen the Simple ABM model functions perfectly under all circumstances. The same cannot be said for the SD model, as its capacity to work correctly is contingent on the duration of both the

immunity and sickness being at least one week, and the Lifespan of the people being not smaller than either of these variables. If these specific conditions are kept in mind when experimenting on the model, it does function well enough for its purpose of being a test-case.

3.7.2.5. Basic behavioural test

Once it has been established that the main structure of the Simple models functions well enough for the purpose of this research, the actual behaviour can be analysed. This behaviour should roughly replicate know disease dynamics. According to Bjørnstad et al. (2021) SIR models, and their extensions, should exhibit periodicity when $R_0 > 1$ and there is recruitment into the susceptible population. This behaviour is visualised in figure 13.

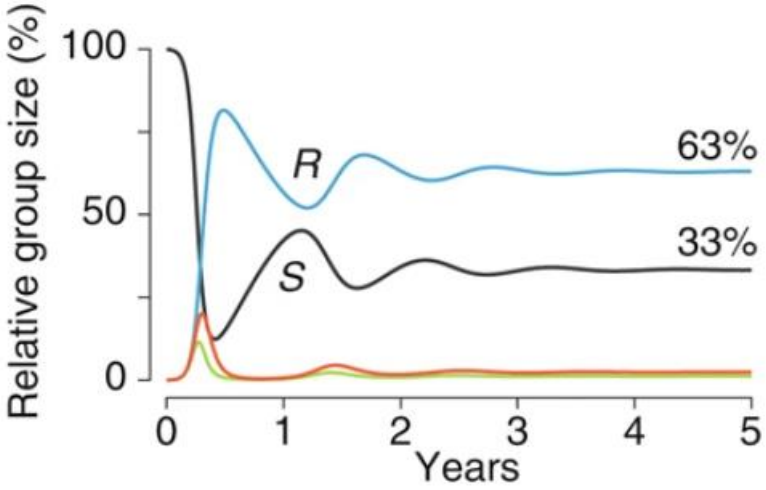

Figure 13: Disease dynamics when R0 > 1, Bjørnstad et al. (2021)

The ability of the models to exhibit these behavioural patterns is analysed in Appendix I**.** Based on the findings it can be concluded that the models are able to replicate these disease dynamics. Therefore, the confidence in the model to result in known disease dynamics is increased, and the model's behaviour is deemed good enough for the purpose of this research.

3.7.2.6. Conclusion Simple model

Based on the five validation tests performed it can be concluded that the Simple models satisfy all requirements for them to serve as a test case for the Extensive model. Not only do their structures seem very solid under most conditions, all relevant structures for preparing to coarse-grain a more complex model are also included. One can never be a hundred percent certain of a model's validity but passing the proposed validations does create enough confidence to move forward with these models.

3.7.3. Validation of Extensive model

Similar to the Simple model, one of the goals of utilising the ABM model is to find out if it is methodologically possible to perform coarse-graining on an ABM model, that includes more complex concepts. In order to be able to successfully achieve this goal, the only requirement of the model is that it is structurally valid. To this end, the same structural validation tests as the Simple model went through will be performed: the parameter verification, structure verification, and the dimensional-consistency test. For the Extensive model there will not be an extreme conditions test, as there are too many constant variables, and the test will be performed indirectly during the exploratory modelling. In contrast with the Simple model, the Extensive model also serves to model a specific pathogen. Therefore, validation tests regarding behaviour should be performed, to check whether the behaviour of the SARS-CoV-2 virus is mimicked. However, due to the many mitigation measure countries have deployed against this virus, it will not be possible to perform an accurate historical

validation test. This means the correct implantation of the SARS-CoV-2 virus can merely be tested via parameter verification. Therefore, the only real test of behaviour will be testing the basic disease dynamics. The structural tests will be performed in unison of the models, and the behaviour test will be performed separately. All these tests are fully performed in Appendix II.

3.7.3.1. Parameter verification of Extensive model

All model parameters are fully analysed to ensure they have a real world meaning and also have values correspondent with this real world meaning. A full discussion of the parameter verification can be found in Appendix II. In this discussion it was found that the variables related to movement, lack adequate real world meaning, making the model unusable for exploration of movement reduction policies. Additionally, a few variables lack real world meanings, this lack can either be explained by the simplifications made in the model, making these variables necessary or will be further explored in the dimensional consistency test. Overall, almost all parameters present in the model are valid, in both real world meaning and value, increasing confidence in the model's ability to be used for the purpose of this research.

3.7.3.2. Structure verification of Extensive model

For the structural verification there are three structural elements that are studied in more detail. These are general epidemiological transmission structure, realistic structural constructions, and complex structural concepts. The full analysis of these structural elements can be found in Appendix II. The only issues that have been identified are related to realistic structural constructs. Especially the structures related to isolation mechanics are a bit unrealistic. This makes the model unable to accurately represent concrete dynamics and mitigation measures in the model. However, for the primary methodological objectives of this research the model can be considered sufficient.

3.7.3.3. Dimensional consistency of Extensive model

By analysing the dimensional consistency of the models, any potential structural mistakes can be identified. Because the structure of both models is almost identical, any dimensional issue in one of the models will also exists in the other. The analysis will therefore only be performed on the SD version of the model. In Appendix II the full test has been discussed. It was found that only one variable had dimensional consistency issues, but because that variable is purely supplementary in nature it does not affect the rest of the model. So, in conclusion the model is almost completely dimensionally consistent, which increases the confidence in the structural foundation of the models.

3.7.3.4. Basic behaviour Extensive model

For the basic behaviour of the Extensive model an attempted will be made to replicate the expected behaviour as shown by Bjørnstad et al. (2021) in figure 13. The basic concept is that there should occur when both $R_0 > 1$ and there is recruitment into the susceptible population. In the model the R_0 value is above 1, however the recruitment into the susceptible population generally happens only after the disease has already died out. Therefore, the immunity duration will be decreased and the percentage of people who will lose their immunity will be increased. The full results can be found in Appendix II Based on the findings it can be concluded that the models are able to replicate these disease dynamics. Therefore, the confidence in the model to result in known disease dynamics is increased, and the model's behaviour is deemed good enough for the purpose of this research.

3.7.4. Conclusion of Extensive model validation

Based on the structure and behaviour tests, fully performed in Appendix II. there is a significant amount of confidence gained in the validity of both the basic structure of the model and the behaviour the model performs under different conditions. There are some caveats to the validity of the model, for instance the movement options are not easily translated to the real world. The closest

thing would be within city movements as short movement and between city movement for long movements, however this interpretation would require some city like structures in the model, which are also not present. Nevertheless, the model is deemed valid enough for use.

3.8. Experimental setup

The goal of the experimentation will primarily be to discover whether the coarse-grained model generates statistically similar results to the ABM model, and under what circumstances it is or is not capable of reproducing accurate results. It is important that the coarse-graining process works under as many instances as possible, and not only for one pre-defined combination of variables, as a good scientist can draw an elephant with only three parameters (Seidenberg, 1993). Therefore, the model needs to be analysed for a wide range of uncertainties, ensuring its usability under all circumstances. The more computational experiments that are performed the better, as both the amount of analyses that can be performed, and the potential to gain knowledge will increase. However, running these computational experiments is also very time consuming. Due to these time constraints, a total of 200 computational experiments will be performed on all models. But since the computational experiments are constructed with the use of Monte-Carlo sampling, further research can simply add more experiments to the already existing ones.

When experimenting with a stochastic model, it is also important to replicate the model multiple times for the same combination of uncertainties, as the results of unreplicated simulation models cannot be trusted (Edmonds & Hales, 2003). According to M.E. Warnier (personal communication, December 2, 2019) the amount of replications you need per computational experiment is more of an art than a science. The reason for this is that it is highly dependent on the impact of the stochastic processes on the overall performance of the model. If the model is strongly path dependent or complex you would need more replications, as the variation of the outcomes will be greater. Some combinations of uncertainties have been shown to result in behavioural modes, that occur in less than 5% of the replications, meaning those runs will most likely only occur in one of the models, if not enough replications are ran. To fully eliminate this issue, the amount of replications per computational experiment can be further increased, however for this research we will perform 20 replications of every computational experiment to balance between simulation time and accuracy of the results.

Besides establishing whether the models generate statistically similar behaviour, it is also important to look at the switching condition. In order to obtain the full benefits of coarse-graining process, the switch should not be made too early, as the outcomes of the results will potentially start to vary too much. However, the switch should also not be made too late since this will be detrimental to the simulation speed. During the simulation of all computational experiments, the time it takes the simulation to fully complete for the two models will be measured.

To ensure reproducibility the software programs utilised in this research can be found in table 4. Since this software is subject to change the used versions are also listed.

Table 4. Software used for research

Software	**Version**	**Reason**
NetLogo	6.1.0.	Software for ABM modelling
Vensim DSS	9.0.1.	Software for SD modelling
Python	3.8.3.	Programming language used
Jupyter Notebook	6.4.6	Interface for using Python
Ema-workbench	2.0	Interconnection of models

3.8.1. Experimental setup of the Simple model

In order to find out whether it is possible to coarse-grain the Simple model, while keeping the results consistent, the model needs to be run for a large amount of computational experiments. Each

computational experiment exists out of values for seven variables that have been created via Monte-Carlo sampling with an uniform distribution. The variables that were varied, and their value ranges, can been found in table 5**.**

Table 5. Simple model uncertainties

Variable	**Value range**	**Extra information**
Carrying capacity	50-1000	
Infectiousness	0%-100%	
Immunity-duration	1-150	Minimum of 1, as it is the SD model limit
Lifespan	50-10000	Minimum of 50, to decrease chance of being lower than either immunity-duration or Sickness-duration
Chance-recover	0%-100%	
Chance-reproduce	0%-10%	
Sickness-duration	1-100	Minimum of 1, as it is the SD model limit

Due to the stochasticity present in the Netlogo model, each computational experiment will be replicated twenty times, using the aggregate results of these replications. The amount of runs, and thus unique computational experiments, used to analyse the Simple model will be 200. Every unique run will be saved, as the time series data is needed for the NCID analysis. This will ensure, that any combination of factors that result in inconsistencies between the two models can be identified and analysed individually. All variables not mentioned in table 5. are set to their base value, as can be found in the full model, which is available on GitHub.

3.8.2. Experimental setup of the Extensive model

The Extensive model will be tested for a total of 31 variables, of which 19 are uncertainties and 12 are policy levers. The values of these variables are all obtained by Monte-Carlo sampling using uniform distribution. The variables that are analysed, their potential ranges and potential extra information can be seen in table 6.

Table 6. Extensive model uncertainties and levers

Variable	**Value range**	**Extra information**
Transmission-parameter	0-015	Based on Netlogo-range
Beds-H	0-100	
Beds-C	0-10	
Death-no-bed	0-1	
Distancing-reduction	0-1	
HR-Shield-duration	1-52	
Isolation-efficacy	0.8-1	Based on Netlogo-range
Mild-asymptomatic	0-1	
Self-isolators	0-1	
Informers	0-1	
Found-and-isolate	0-1	
Prob-InfDeath	0-0.02	Based on Netlogo-range
Prob-InfHosp	0-0.25	Based on Netlogo-range
Prob-HospDeath	0-1	Prob-HospDeath + prob-HospCrit cannot exceed 1
Prob-HospCrit	0-1	Prob-HospDeath + prob-HospCrit cannot exceed 1
Prob-CritDeath	0-1	

Immune-mild	0-1	
Immune-severe	0-1	
Immune-loss-when	4-26	Based on Netlogo-range
When-symptoms-if-I	0.05-2.5	Must be a bit bigger than the Time Step used in Vensim
Prop-high-risk	0-0.2	Based on Netlogo-range
Relative-risk	1-15	Based on Netlogo-range
Intervention-duration	1-52	
Blocked-bed-effect	0,1	Boolean variable
Distancing-option	0,1,2,3	Categorical variable
Isolate-inform?	0,1	Boolean variable
High-risk-shielding	0,1	Boolean variable
Use-age-mixing?	0,1	Boolean variable
Trigger-type	1,2,3	Categorical variable, but 0 is not possible, as that requires interactivity with the modeller during the simulation
Trigger-level	X	Value depends on the trigger-type
Trigger-level(days)	5-200	
Trigger-level(hospital)	0-1	Percentage of total available beds
Trigger-level(cases)	50-2000	
Switch-level	1000,1500,2000,2500	Categorical variable

It is important to note that four relatively important variables will not be varied during the open exploration of the coarse-grained model (see Appendix IV). These variables are all related to the movement of the agents. As mentioned in Section 3.7.3.1, the movement of the agents is not implemented in a structurally realistic manner, this combined with difficulties in scaling the related look-up variable made the outcomes very inconsistent. For a more in-depth explanation of the issues related the movement variables Appendix IV can be examined.

Due to the stochasticity present in the model, it will be ran twenty times for each computational experiment, comparing the aggregate results of computational experiments. As there are a large number of variables that need to vary, there is also a need for a large amount of replications, to ensure the correct behaviour of the model under all circumstances. Both models will be run for 200 different computational experiments. This should give a good representation of most potential behaviour there can occur in the model.

4. Results

In this chapter the outcomes of the experiments will be evaluated and discussed. The experimental setup, which is used for the exploration is described in chapter 3 paragraph 8. First the results of simulation speed will be discussed for both models. In Section 4.2 the consistency of the Simple model will be discussed followed by the Extensive model in Section 4.3.

4.1. Simulation speeds results

The main goal of the coarse-graining process is to increase the speed at which the results for simulations are gained. To this end the simulation speed of both models for the 200 computational experiments under 20 replications has been measured. The observed time it took for these simulations to complete can be found in table 7.

Table 7. Simulation speed of the different models.

Model	Simple simulation speed	Extensive simulation speed
ABM	00:15:46.86	18.00:19.15
Coarse-grained	00:09:05.66	17:29:58.18
Speed increase	73,5%	2,9%

The difference in simulation speed is significant for both models, although much higher for the Simple model. The formula for the speed increase is as follows: $\frac{(ABM\ speed - Coarsegrained\ speed)}{Coarsegrained\ speed}$. A doubling of the simulation speed would thus be indicated by a speed increase of 100%. This is probably due to the computational burden of the switching process itself. In earlier versions of the Extensive coarse-grained model, running the coarse-grained model took even longer than running the ABM model, this time was upset by some coding optimisations.

These results indicate the promise the coarse-graining process holds. When the coarse-graining process gets even more streamlined and efficiently coded, the speed increase can also be further improved. However, improving the simulation speed only serves a purpose under the condition that the results stay consistent. So, to this end the results themselves need to be analysed further.

4.2. Consistency of results

Two hundred unique computational experiments have been run with twenty replications per set for both the coarse-grained and ABM model and for both the Simple and Extensive model. During the simulation these twenty replications were combined into a singular run for every computational experiment. Additionally, all singular runs have also been combined with each other to form one aggregate run per model. In order to find out whether the results stayed consistent when using the coarse-graining method, the runs were compared using a normalised Complex Invariant Distance (NCID) metric, as discussed in Section 2.6. The overall results of the consistency will be discussed first, followed by a more in-depth outcome analysis. The Simple model will be analysed first in Section 4.2.1. followed by the Extensive model in Section 4.2.2.

It is first important to note, that the NCID values of the computational experiments can be seen as an equality measure. However, there is no predefined value, at which the data can be considered equal “enough”. For the purpose of this research a NCID value of 0.05 will be considered the cut-off point for equality. This value is chosen arbitrarily but based on the standard p-value cut-off point used in statistics.

4.2.1. Consistency of Simple model

Based on the NCID cut-off value of 0.05, 200 of the 600 variable outcomes are deemed unequal. All these runs can be seen in the ensemble shown in figure 14. The orange lines indicate lines with a NCID value below the cut-off value, with the blue lines having higher NCID values.

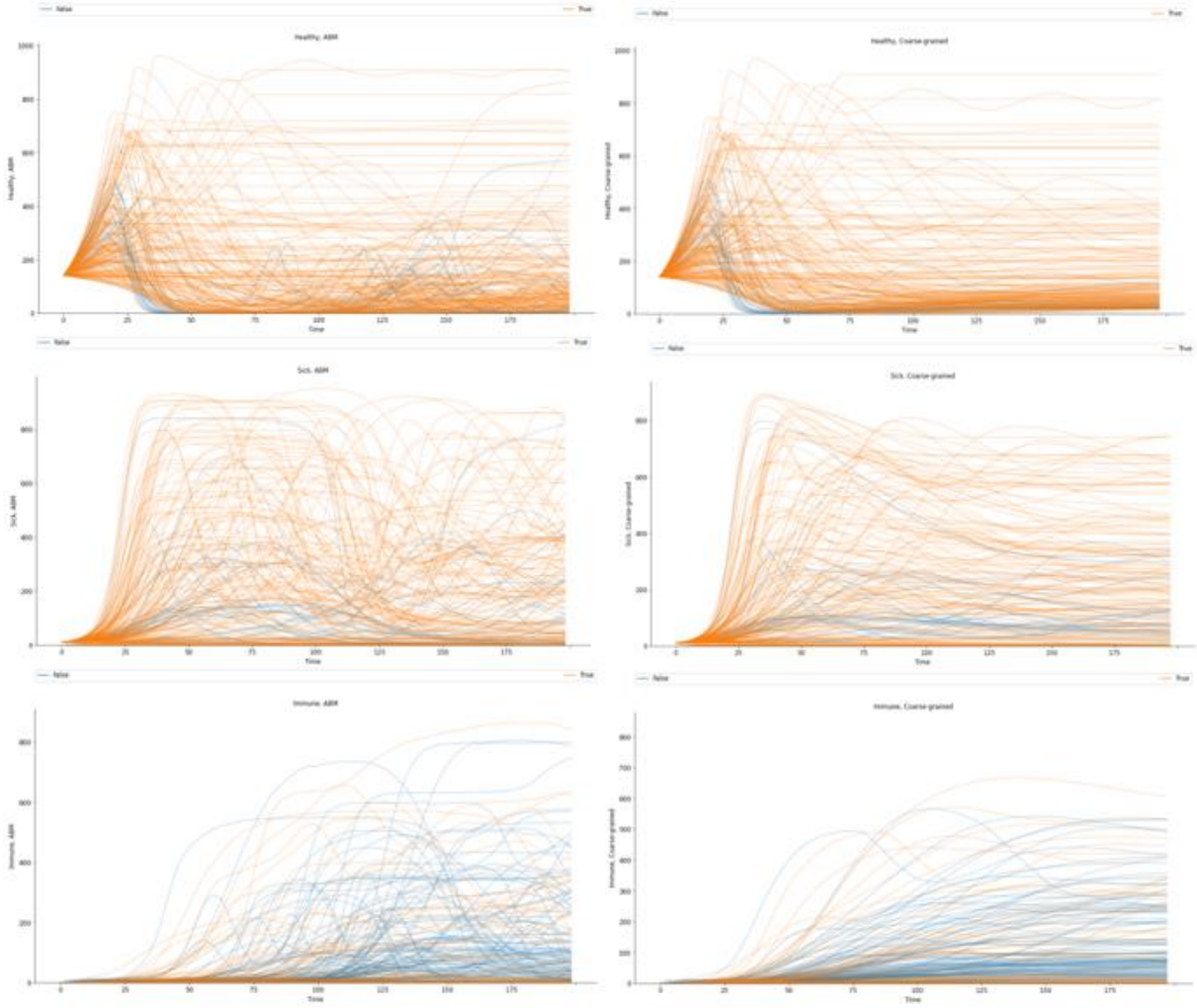

Figure 14: Ensemble of runs Simple model

This figure indicate that the coarse-grained Simple model relatively accurately models the healthy population, but the Immune populations seem to differ a lot. Additionally it can be seen that there is a big difference in oscillatory behaviour between the models, with a loss of oscillation when utilising the coarse-grained model.

However, perhaps more importantly the overall aggregate results have values significantly below the 0.05 threshold. The results of the aggregate run can be seen in figure 15. The aggregate Healthy, Sick, and Immune variables have NCID values of 0.0041, 0.0097 and 0.0162 respectively, as also shown in table 8.

Table 8. Aggregate NCID values of KPIs Simple model

Variable	**NCID value**
Healthy	0.0041
Sick	0.0097
Immune	0.0162

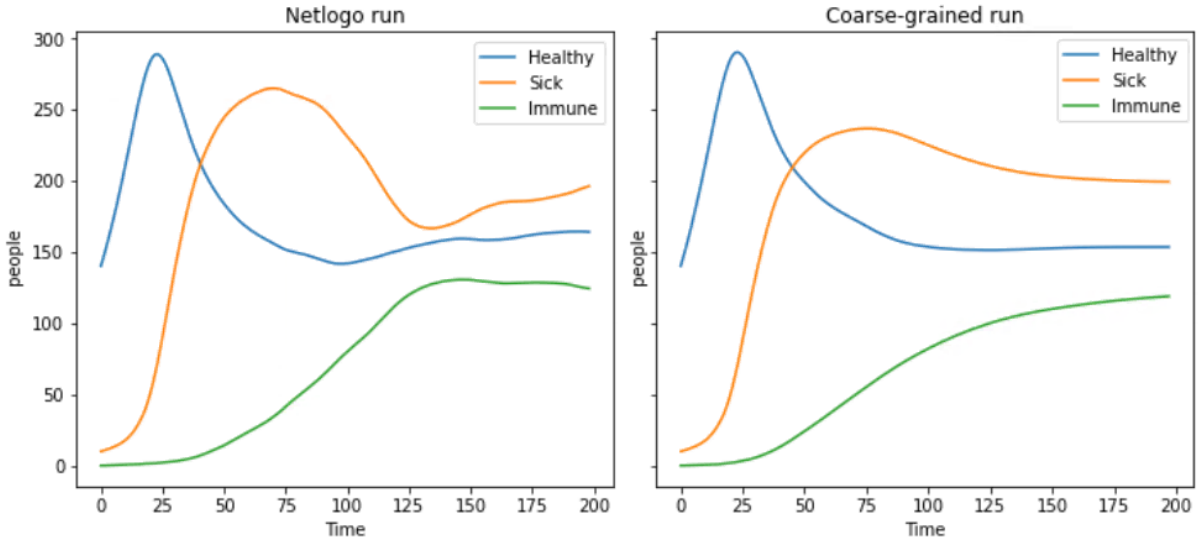

Figure 15: Aggregate run of Simple model

The aggregate run shows that the two models start very similar, than diverge to converge in the end. This behaviour can best be explained by the fact both models start out in ABM. The coarse-grained model loses its oscillatory behaviour quickly, but in the long run both models lose the oscillations. Based on this hypothesis it is likely that the NCID values will decrease as the run-time of the models increases.

Besides knowing what percentage of runs exhibit the same behaviour it is also important to gain knowledge regarding parameter values that result in unequal runs. To this end an output analysis has been performed the full explanation of this method can be found in Section 2.7**.** In short, this method categorises all uncertainty values as well as the NCID outcomes. The more often a parameter value in a certain category leads to unfavourable outcomes (High NCID values), the score of that category will be increased. Each time a certain parameter value leads to a more favourable outcome the score will also be decreased. The final score will then give a rough estimation of the performance of the variable during the coarse-graining process. As there are three Key Performance Indicators (KPIs) in the Simple model, there are also three NCID scores per computational experiment. This can be used to differentiate between the impact of the parameter value on the different KPIs. For this analysis the combined NCID value of the three KPIs will be used.

This analysis found that the variable *Duration* with values higher than 80 ended up with a score of 55, which is an incredibly high score. The most likely explanation for this occurrence, is related to paradigm differences. In the ABM model, people will stay sick for exactly the duration of the disease. Whereas on the SD side, people will stay sick for the duration of the disease on average. This will mean people will start to spill over into either immunity or death far earlier. This behaviour is also showcased in figure 16.

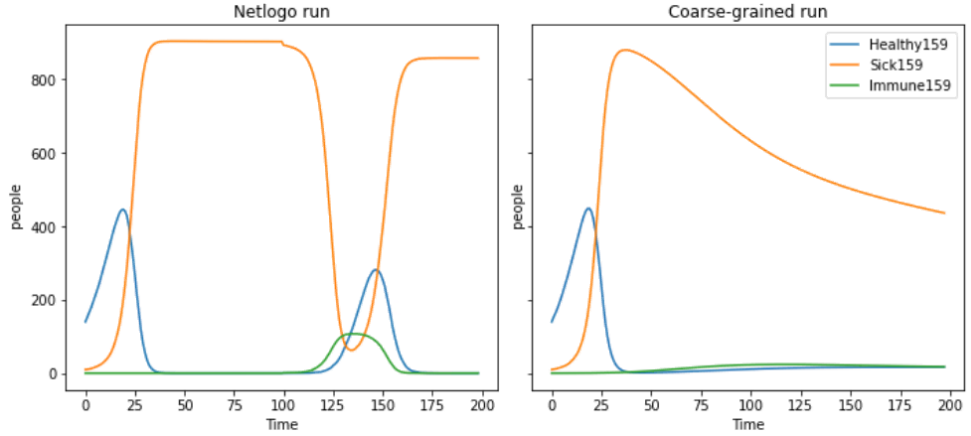

Figure 16: Effect of high duration on coarse-graining Simple model

When removing the computational experiments that include durations with values above 80 the outcome analysis resulted in a similar outcome. This instance showed durations above the value of 60 performed worst. By further removing those computational experiments the next worst score was negative two. This generally indicates that the variable value more often leads to a run with a low NCID, than a high NCID. This result is slightly skewed, as the remaining computational experiments contain significantly more low NCIDs than high NCIDs. Overall, it does not seem that there is another variable that causes a significant amount of poor results, when dynamically coarse-graining the Simple model.

4.2.2. Consistency of Extensive model

Using the cut-off value of 0.05 to determine the consistency of the models, it can be seen that the Extensive model is less consistent than the Simple model, as here 255 of the 603 variable outcomes have higher NCID values. This increased inconsistency can also be seen when looking at the ensemble of runs shown in figure 17.

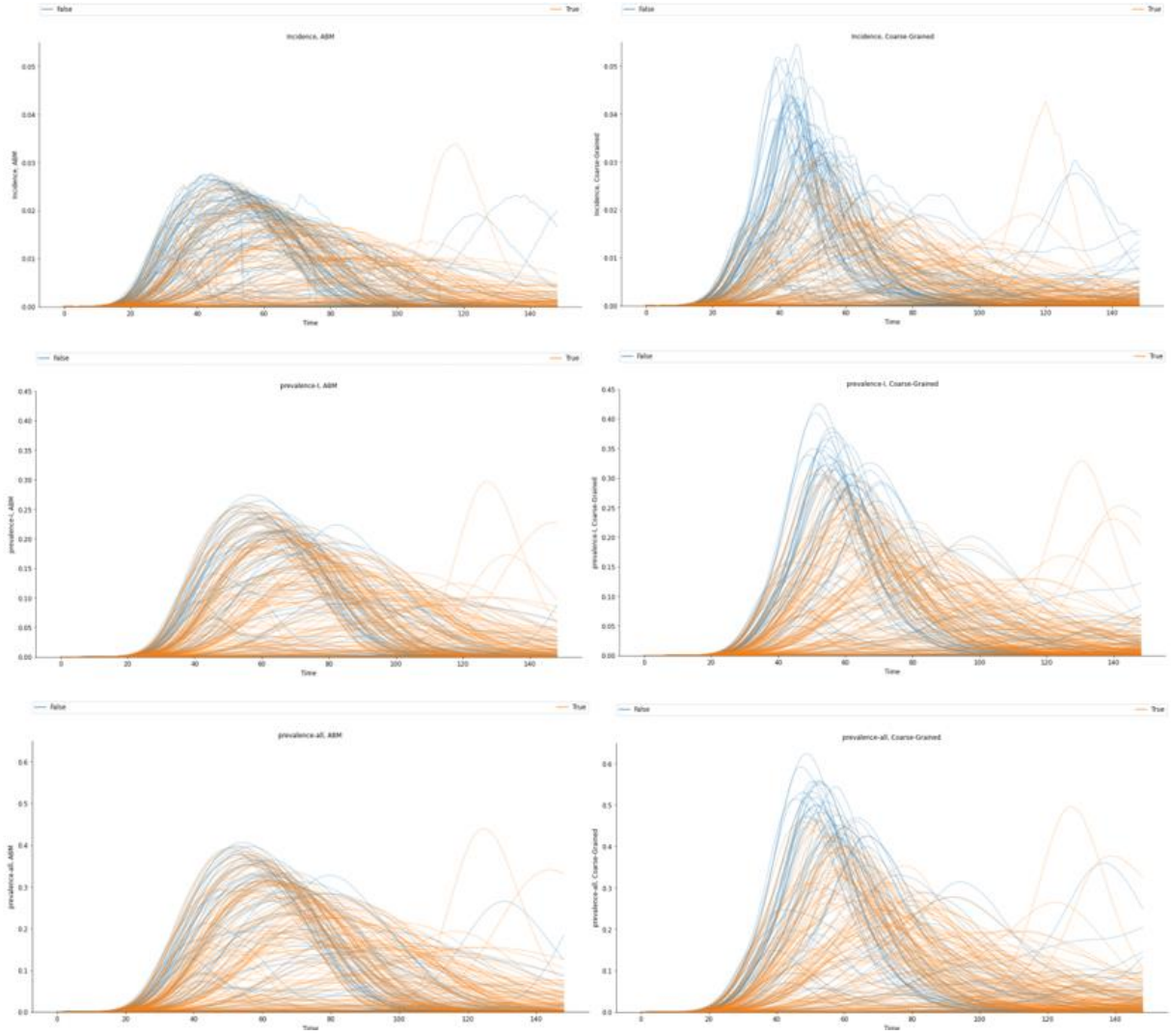

Figure 17: Ensemble of runs Extensive model

This ensemble shows that the coarse-grained model has a tendency to overestimate infections at the beginning of the simulation leading to higher peaks. When looking at the ABM side of the model it does not seem that there is a certain behavioural mode that specifically leads to this overestimation. This lower consistency compared to the Simple model can also be seen in the NCID values for the aggregate run for the KPIs; *Incidence, Prevalence-I,* and *Prevalence-all*. Their aggregate NCID values

are respectively 0.0258, 0.0158, and 0.0161, as also shown in table 9. Even though they have comparatively higher NCID values, these values are still below the cut-off point, this could be interpreted as the model on average being consistent. The results of these aggregate runs can be seen in figure 18.

Table 9. Aggregate NCID values of KPIs Extensive model

Variable	**NCID value**
Incidence	0.0258
Prevalence-I	0.0158
Prevalence-all	0.0161

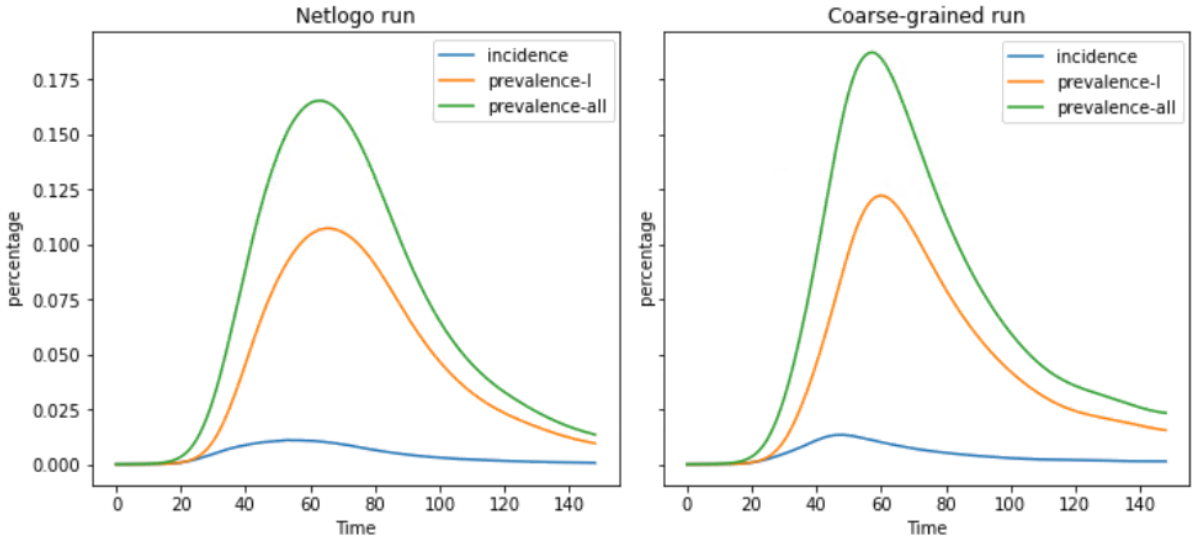

Figure 18: Aggregate results of Extensive model

These aggregate results showcase that the biggest difference between the models can be found at the peak, as the coarse-graining results in a significantly higher peak. To potentially determine the cause of this difference a closer look will be taken at the variables that lead to the worst outcomes, using an outcome analysis, explained in Section 2.7.

The variable that is most indicative of inconsistent results is the switching level. This variable indicates when the model makes the switch form the ABM model to the SD model. In this case both switching levels of 1000 of 1500, resulted in significantly more inconsistent runs. With the score of switching level 1000, being more than twice as high as the 1500 switching level. This can be seen as a positive, as this indicates that the model is less accurate when switching too early. However, this does not necessarily mean the model actually functions better under an even higher switching level. By heightening the switch-level it can occur, that the switching level never gets reached, resulting in a purely ABM simulation. As can also be seen in figure 18. a *Prevalence-I* of 0.1 (which is comparable to a switch level of 2000), is at the very peak of the graph. Due to computational experiments that die out immediately after the start this peak is a bit lower than the real peak could reach, so switching at the level of both 2000 and 2500 does still happen, yet less often than compared to the 1000 and 1500 levels, as can also be seen in figure 17. This has the potential to skew the results. The failure to identify any real parameters that cause the behavioural differences, indicate that the differences are most likely due to structural issues, with the *movement look-up* variable being the most likely issue. A theoretical solution would be to artificially lower this look-up variable, however this will also impact the runs that are consistent at the current moment. Potentially under-estimating the amount of infections occurring, which has an additional detrimental effect of resulting in switches occurring more often and in extension a decreased simulation speed.

Another issue that can be identified by closely examining figure 18, is that the decrease at the end of the simulation of the coarse-grained model drops off, compared to its ABM counterpart. This is due to a second wave of infections that is build-up in the coarse-grained models after a switch-back occurred. The complexity related to the transformation of the ABM model, makes the model significantly less often die out. This divergence at the end of the simulation indicates, that the NCIDs will keep increasing as the run-time is increased, which is opposite to the Simple model. Additionally, to the less consistent results the simulation speed of the coarse-grained model will potentially also suffer. When a second wave of infections occur, there will also be additional switches made between the two models, which could increase the computational power required. Based on these results it can be concluded that the principle of coarse-graining works, but there is still ample room for improvement.

5. Discussion and conclusions

In this chapter, the results will be discussed. More specifically, Section 5.1. discusses the main results, while Section 5.2 addresses the strengths and weaknesses. Next, Section 5.3. explains the potential real-world applications dynamic coarse-graining, followed by the scientific contribution of this work in Section 5.4. Finally, recommendations for future research are given in Section 5.5, followed by the main conclusions presented in Section 5.6.

5.1. Discussion of main results

The main research question was: *"What is the effect of dynamically coarse-graining an agent-based SARS-CoV-2 model into a System Dynamics model on the simulation speed and the consistency of results?"*. This question was divided in four sub-questions which will be answered in Sections 5.1.1 (sub-question one and two), 5.1.2 (sub-question three), and 5.1.3 (sub-question four). Together, they answer the main research question.

5.1.1. Method and rules for dynamic coarse-graining of an agent-based SARS-CoV-2 transmission model.

In this section, the answers to the first two sub-questions will be summarised.

- *What are different methods, and their associated advantages and disadvantages, in regard to modelling SARS-CoV-2 transmission in populations?*
- *What set of rules should be used for dynamic coarse-graining an agent-based SARS-CoV-2 transmission model into a system dynamics model?*

These sub-questions were answered in Chapters 2 and 3 with a literature review. The main results for the different methods of modelling of transmission models there are that there are in general two main simulation techniques utilised, and thus also in regard to SARS-CoV-2 modelling, that follow the same basic structure. This basic structure is based on the general disease dynamics of virtually all known transmissible diseases, the SEIR structure (H. A. Biswas et al., 2014). The two methods most commonly used are Agent-Based Modelling (ABM) and Equation-Based Modelling.

ABM models are generally more computationally intensive, but also more accurate as it allows for inclusion of more important disease dynamics, especially at the early stages of an epidemic. Equation-based models are very fast to run, as they can exists of only a very limited amount of differential equations, but in exchange for this speed they lack some key characteristics for modelling an accurate epidemiological transmission model.

To take advantages of both the accuracy of the ABM paradigm and the speed of the equation-based paradigm, a hybrid structure is proposed focussed on dynamically switching between the two models during simulation.

When answering the second sub-question, it is relevant to note that the most important aspect during the dynamic coarse-graining is the consistency of results between the models. In order to achieve this consistency, the main structure of the two models need to be identical. Borshev & Filippov (2004), described how an equation-based model could be translated to the ABM paradigm. Their proposed methods were used as the main guidelines to follow in translation process. For the complete method of translating all concepts, Section 3.3 and specifically table 2 can be consulted.

5.1.2. Coarse-graining and simulation speed

The third research sub-question was:

- *What would dynamic coarse-graining an agent-based SARS-CoV-2 transmission model mean in terms of simulation speed.*

Based on the results presented in chapter 4, it can be seen that dynamic coarse-graining has the potential to improve simulation speed quite considerable. Two coarse-grained models were created, a simple and extensive model differing in both size and complexity. The Simple model and Extensive model had an increased simulation speed of 73.5% and 2.9% respectively, compared to their respective original ABM model. This showcases the potential of speed increase coarse-graining can have. However, it also indicates that the benefits can be very limited when the switching procedure is too computationally intensive. As, when the extensive SD and ABM model are run independently, the simulation time of the SD model is significantly shorter compared to the ABM model.

As showcased in chapter 4, the coarse-graining technique does improve the simulation speed of the model, just as theorised in chapter 2, and in coherence with the increased simulation speed of a coarse-grained model in Gray and Wotherspoon (2012). This improved speed of simulation does come at the costs of simulation accuracy, with a between 33.3% and 42.5% of the runs not reaching the predefined consistency level of 0.05 NCID between the models. In the case of the Simple model this was largely the case due to the simplistic and also unrealistic manner the sickness duration was implemented in the original ABM model (Wilensky, 1998). If this duration was used in a more realistic manner, preferably with the use of an exponential distribution the coarse-grained model output would come really close to the ABM model output.

The bigger issue at hand has to do with the Extensive model, which failed to find variables responsible for the inconsistencies. This indicates the inconsistency was a result of a structural issue within the model. As most of the model was duplicated in an exact manner, the issue is most likely in the look-up variable for the movement options. This variable was constructed fairly crude, as the coherence between the movement variables, the susceptible people and transmitters were too difficult to find. There most likely does exist a more mathematical solution to the problem, which could in turn significantly improve the coarse-graining consistency.

The second big issue in the Extensive model is related to the swapping between models, specifically going from the SD model to the ABM model. A lot of information goes lost when the initial switch towards the SD model is made, which was also noted as a problem by Wallentin and Neuwirth (2017). To counteract this the agents are kept in place during the SD simulation. However, when switching back, an estimate of which agents should be in which disease state needs to be made. The choice was made to make this dependent on proximity to the first agent infector. With agents that are in an early disease state, like exposed, being positioned far away from that agents' positions, and people who are at the end of the dynamics, like immune, being close to the agent. This can be seen as realistic, however that is only the case under a specific condition. That is the impact of the virus at hand was high. If the disease dies down or out because of mitigation measures put in place, when only a small portion of agents have been impacted the spatial situation is very different. Therefore, it would probably better to have the switching method be dependent on the total impact of the virus, instead of only having one method of transforming the ABM model.

5.1.3. Comparison of a dynamically coarse-grained model with the original agent-based SARS-CoV-2 transmission model.

The fourth sub-question was:

- *What are the differences and similarities between the dynamically coarse-grained and the original agent-based SARS-CoV-2 transmission model in terms of consistency of results?*

In order to determine the similarity of the two models in terms of consistency of results, there first needs to be an objective metric for the similarity between the original and dynamically coarse-grained model. We presented a normalised CID metric, to assess this similarity. Based on this metric, and the devised cut-off point for similarity of 0.05, the Simple and Extensive models were inconsistent for 33.3% and 42.5% of the computational experiments respectively. Perhaps more

importantly, the aggregate run of both models were deemed consistent, indicating the models are consistent on average.

Besides the issues that are related to the coarse-graining process itself, there are also issues related to analysis of the outcomes of the coarse-graining process. Firstly, the normalised CID, is not a very consistent metric. The original CID measure is in concept a measure of distance (Batista et al., 2013). However, when looking at distance measures the concept of relativity is very important. When a measurement has a 10 centimetres error margin this can be a significant margin when considering a small distance or only a minor margin when considering a measurement of several kilometres. When talking about normal distance measurements, one could express this margin as a percentage difference, which is a good indication of the actual accuracy (Rabinovich, 2017). This concept was attempted to be utilised, however time series data does not have a singular value that can be used for this comparison. When using the highest value in the data, the results for data that includes very few large numbers, and a lot of small numbers would be skewed. Therefore, the average value of the time series data and the length of the time series data were used to "normalise" the CID value and make it viable for general comparison. The resulted NCID values are unfortunately not very consistent. Some series had relatively high NCID values, even though the resulting graphs look very similar, and did not even reach the switching conditioning. Meaning they should have been almost identical after twenty replications, but the discreteness of the model combined with a little stochasticity thus resulted in inconsistencies. Since the main variable that was used for further analysis was a bit inconsistent, the subsequent analysis was also made more inconsistent.

This subsequent analysis could be considered as a PRIM inspired analysis (Friedman & Fisher, 1999). The need for an inspired analysis was necessary as the amount of data needed for a real PRIM analysis was too time consuming to collect. This simplified version required less data, in exchange for more crude results. By clustering the continuous uncertainty values to five groups per uncertainty, it becomes easier to see what group of inputs leads to what group of outputs. This process is dependent on the grouping to be done in a correct manner, as when part of the same group results in high NCID values and another part of the same group results in low NCID values the only result that is shown, is no significant influence on NCID values. This problem does not occur when working with a regression analysis instead of clusters (Dazard et al., 2015).

Assuming the clusters of uncertainty values were valid, every time a cluster and the corresponding computational experiments get removed from the analysis two issues start to develop. For one a big chunk of "bad" NCID values get removed from the analysis. Therefore, the remaining clusters will most likely result in relatively "good" NCIDs just from the virtue of still being in the analysis. After two or three rounds of removing clusters, this analysis can therefore not accurately depict the remaining problematic clusters (if any). This can be solved by either re-clustering the NCID values every time a cluster gets removed, or even better running the entire simulation again without the removed cluster. This last option will take a considerable amount of extra time but will ensure there are always enough datapoints available for the analysis.

A second issue with this analysis is the lack of multivariate testing. It is entirely possible that a specific cluster only results in undesirable outcomes when it is combined with a specific other cluster. The outcome analysis is not equipped to identify these reinforcing effects, whereas a real PRIM is able to actually find those scenarios if and when they occur (Guivarch et al., 2016).

It must also be noted that the outcome analysis utilising the NCID values is far from obsolete, despite the many issues associated with it. The outcome analysis was able to identify several errors in the coarse-graining process, which helped to improve it. Therefore, it can be concluded that even though the outcome analysis is far from perfect it is definitely valuable.

5.2. Strengths and weaknesses

This study appears to be the first study to research modelling concepts in the relatively new field of dynamic coarse-graining, or model swapping (Vincenot et al., 2011). These modelling concepts are an open exploration and allowing for behavioural changes to occur during the dynamic switching

process. This study can then function as the basis for future studies looking into a combination of these concepts. More specifically on how an open exploration of the uncertainty space can help in evaluating the consistency of coarse-grained models, and on what the effects of behavioural mode switches during the coarse-graining process are on both the consistency and simulation speed of this technique.

The second key strength of this study is also related to the evaluation of consistency and simulation speed, and that the evaluation of multiple models that differ on most scales but are related to the same system (Gray and Wotherspoon, 2012). By first evaluating the Simple model as a case-study, additional insights in the potential generalisability of the found results. By comparing the outcomes of the two models it was for instance found that dynamic coarse-graining does have the potential to significantly increase simulation speed, but only under certain circumstances.

A third strength of this study is the attempt to resolve the issue regarding the metric of comparing modelling outcomes in terms of consistency. This issue has been mentioned in multiple papers that researched the comparison of models (Auping, Pruyt and Kwakkel, 2014; Gray and Wotherspoon, 2015). To this end the NCID metric was introduced, which is especially useful for comparison of oscillatory behaviour, while also resolving the issue of relativity mentioned by Auping, Pruyt and Kwakkel (2014).

Besides its strengths the study also has some limitations, which need to be kept in mind when evaluating the contents of this paper. However, as the great Dutch football coach Johan Cruijff once said, every disadvantage has its advantage (RTL, 1997), so do these limitations, as they can form the basis for future research.

A methodological weakness of this research is related to the limited amount of runs performed on the Extensive model, namely 200 whereas it would have been more advisable to perform a couple thousand runs. The limited amount of performed runs has two detrimental effects on the research as a whole. For one it could have prevented some modes of behaviour from occurring as they might only occur under very specific circumstances, that would normally be found when fully exploring the uncertainty space. When attempting to develop robust policies using open exploration this can be a big issue, as it is essential to test the robustness under all circumstances. However, in this case the goal was not to create a robust policy, but to see under what circumstances the coarse-grained model is able to stay consistent to its ABM counterpart. Therefore, in the worst-case scenario these undiscovered behavioural modes would also result in inconsistent modelling results, and therefore slightly lower the percentage of consistent runs.

Another related limitation is the impact it has on the methods of analysis. Due to the lack of runs there is also a lack of available data that can be used for analysis. This for instance prevented the use of a PRIM analysis (Friedman & Fisher, 1999), and forced the use of outcome analysis as described in Section 2.7. The detrimental effects of using this analysis instead of the PRIM are described in more detail in Section 5.2.2.

A second limitation is related to the evaluation of the simulation speed. To get a good grasp of the different aspects influencing the speed differences between the models it would be better to gain more data related to the simulation besides the total simulation time. By broadening the scope of the time analysis to include data about individual runs or even individual aspects of runs, like for instance the time the switching process itself takes, a better understanding could have been gained regarding the possibilities of the speed increase, and the causes of any potential lag.

The last major limitation is in regard to the setup of the Extensive model. The Extensive ABM model is SARS-CoV-2 specific and includes some complex to translate properties but is in its current form not useable to give actual policy advice. If a model, that was actually used for the creation of policies in a country, was coarse-grained and analysed a more concrete overlook of the specific advantages for SARS-CoV-2 could be given. The analysis would in turn be able to both give insights in the coarse-graining technique as well as insights relating SARS-CoV-2 mitigation policies.

5.3. Practical applications for coarse-graining

What the new insights gained during this research actually mean for the viability for coarse-graining simulation models moving forward will be discussed now. The perhaps most important lesson is the difficulty involved in developing a coarse-grained model. As discussed in Section 5.2.1 there are still a plethora of both small and big issues in the coarse-grained model, especially in the Extensive model. Before a coarse-grained model is actual able to be used for practical applications it is paramount these kinds of issues are all fixed first. The more complex a model becomes, the complexity of coarse-graining that model grows exponentially. Creating a good valid model from scratch, already takes a significant amount of time. Making two of them in different paradigms that flawlessly fit together will take even more time. All this time loss needs to be offset by the resulting coarse-grained model otherwise it would not be practical to go to the effort of creating the model in the first place. So, the question becomes when would it be viable to utilise the dynamic coarse-graining technique?

There are essentially three distinct applications for coarse-grained models.

1. Models that will be used for a sufficient amount of time that the time investment will be gained back by the increased simulation speed
2. Models that can be created in advance, and used in times of crisis
3. Models working in real-time, where speed is vital

The first option mentioned here, relates to models that will need to be used for simulation on a regular basis for years or even decades. A model type that encompasses these attributes are the Integrated Assessment Models (IAMs) (Weyant, et al., 1996). These IAMS are primarily used to model the impact of policies on global climate change (van Vuuren et al., 2009). Since there are many uncertainties regarding climate science, an extensive uncertainty analysis would be necessary to evaluate such a model, whereas most only utilised a few sensitivity analyses to this end (Stanton et al., 2009). Utilising more realistic open exploration methods like Monte-Carlo analyses showed to have a major impact on model outcomes (Dietz and Hope, 2007). In one stochastic IAM 2000 parameters values were sampled from a probability distribution (Dowlatabadi, 1998). Performing an open exploration on such a model would be incredibly computationally intensive. This is a good example of where coarse-graining to save simulation time could have significant benefit. When looking at the results displayed in Section 4.1, it clearly shows that coarse-graining increases the speed at which simulation can be completed. The speed increase is very dependent on the computational burden the switching process between the models encompasses. It is not unthinkable, that a more streamlined coarse-graining process could potentially half the simulation time needed for models. However, these IAMs will probably also have quite a bit of switching lag, due to the size of these model. So, the increased simulation speed will probably lie somewhere in between the speed increases seen in table 7**.** Keeping the long and difficult development process in mind, these models will likely need to be ran for a very long time to see a return on time investment.

The second option relates to crises like the current SARS-CoV-2 crisis. When a good coarse-grained model can be developed before a virus reaches the country of interest for instance, the time sink is not relevant if it makes the ability to find robust policies faster at that time, when more specifications of the virus become known. In cases like this swift actions from governments have the potential to save lives of a lot of individuals. By supporting this decision-making process with a faster method of simulating, this can be an invaluable tool in savings lives.

The last real-world application and perhaps the most promising is related to real-time simulations, where speed is vital. A good example of such a situation is with self-driving cars. These self-driving cars need to perform an enormous amount of simulations while driving to ensure the safety of both the passenger and others in traffic (Chen et al., 2019). By coarse-graining parts of these simulations, for instance cars a bit further away, vital milliseconds can be saved, which in turn can be used to save lives. This is an especially promising avenue, as other solutions like distributed simulations are not viable alternatives, due to safety concerns.

5.4. Scientific contributions

In order to see the full scientific contributions of this research, the previous research containing dynamic coarse-grained simulation models must be examined. In Section 2.4.1 the previous works in this relatively new field were discussed. By comparing those works with this new research, the contributions of this research can be made explicit. Based on those paper (Bobashev et al, 2007; Gray and Wotherspoon, 2012; Gray and Wotherspoon, 2015; Wallentin and Neuwirth, 2017) this research seems to be the first research into dynamically coarse-graining populations while allowing for behavioural change during the dynamic coarse-graining process. This behavioural change, due to for instance implementation of mitigation measures, makes the switching back process to the ABM model significantly more complex, as relevant information needed for good spatial positions of the agents relative to each other is lost. By removing the confines of behavioural stability from the coarse-graining process its uses can be a lot more widespread.

A second addition to the current scientific literature is the open exploration of the uncertainty space in order to gain insights into overall behaviour of the coarse-grained model. Wallentin and Neuwirth (2017) also tested their coarse-grained model under multiple scenarios, but their scenarios focussed on the aggregation levels and moment of switching instead of the inherent uncertainty in the underlying model.

A third addition is the concretisation of the translation of some more complex ABM properties to the SD paradigm, which extends on the paper of Borshev & Filippov (2004), who primarily discussed the translation of basic properties between the two paradigms. This extension can serve to support researchers when translating different modelling, being another step towards and exhaustive concrete list regarding methods of translating these properties.

Lastly, this is also the first research that tries to mathematically analyse the similarity between the dynamic coarse-grained model and the ABM model. In this similarity comparison the starting point is to have the original ABM model as the benchmark instead of reality itself. This allows for the comparison of every coarse-grained model instead of only realistic models. The NCID term of consistency has been introduced to this end.

5.5. Future Research

Based on the limitations discussed in Section 5.2 there are a few avenues, that could benefit from further investigations. These avenues can be divided into two categories. Firstly, improvements made to this research specifically, and secondly standardisation of concepts for dynamic coarse-graining.

For the improvements made to this research specifically, the analysis can be expanded upon significantly. Utilising a real PRIM to analyse the results, can potentially result in new insights that were missed due to simplifications that were made. Since the current computational experiments are both openly available and sampled via the Monte-Carlo method, more experiments can be added upon the existing ones.

A second avenue of future research in context of improving the current analysis is related to the used NCID metric. This metric generally only focusses on Euclidean distance and the complexity of the analysed lines. Whereas perhaps different metrics relating to other attributes of lines might give a better insight in the similarity between the runs of the two different models, which in turn would help by any further analysis regarding the consistency of the dynamic coarse-graining process.

Besides improving the analysis of the current models, the models themselves could potentially also benefit from improvements. The concepts that created the most issues could primarily take advantage of improvements made to them. Firstly, by devising a more accurate way of creating a variable related to the chance of contact could significantly improve consistency results, as this would potentially prevent both over- and under estimations of infections happening. The second big issue, is related during the switch back from the SD model to the ABM model. During the switch back all agents need to get the correct attributes attached to them, to ensure the coarse-grained

ABM model will follow the same behavioural mode as the original ABM model. Creating an improved algorithm for this process or making the SD model more geographically specific could potentially solve this issue. By making these improvements to the models the limits of consistency for the dynamic coarse-graining can be explored further. Knowing these limits of consistency can be crucial to determine whether it is safe to use such a technique in self-driving cars for instance.

Lastly, by optimising the coding for the interconnection and additional time measurements, the knowledge regarding improvements in simulation speed can be increased. Optimisations would give better insights into the real potential speed increases of the techniques, as that might be hindered, by sluggish code. Secondly, additional time measurements of all different steps of the dynamic coarse-graining process can give extra insights into potential bottlenecks of the technique. By knowing these bottlenecks, or main reasons of time loss, a better extrapolation could be made regarding the potential time save of even bigger models, like the Integrated Assessments Models mentioned in Section 5.3.

More interesting is perhaps improvement to standardisation of concepts. The concepts that could benefit most from standardisation is the comparison metric used for dynamically hybrid models. The NCID proposed here gives an insight into the similarity of the time series data but does seem to be inconsistent for some cases. Just like how the introduction of the standardised metric system, reduced the amount of different metrics used greatly, and by extension improved both trade and scientific progress (Poirier, n.d.). Standardisation of the similarity metric would change the current status quo of almost every paper using a different metric to measure similarity. This can in turn support the analysis of the different papers on the subject, and extension improve scientific progress moving forwards.

The second standardisation, that can be looked into is regarding the translation of complex concepts form ABM models to an equation-based equivalent. Concepts like how random movement relates to chances of encountering, or an algorithm for gaining relevant information for switching between models, could be really helpful. Now every time someone attempts this the wheel has to be reinvented, whereas there is probably a more robust solution for these issues.

5.6. Conclusion

Due to the new SARS-CoV-2 pandemic, governments all over the world started to use simulation models in an afford to mitigate the effects on both the economy and human lives. In times of pandemics every day a mitigation measure gets activated sooner, more damage to society is prevented. To achieve this goal a new dynamic coarse-graining method of combining ABM and SD is proposed, where during the coarse-graining behavioural changes can occur. Two dynamically coarse-grained models (simple and extensive) were compared to their original ABM counterpart model. By utilising the newly proposed NCID metric, the consistency of this method was analysed. Using a NCID cut-off point of 0.05, it was found that between 33.3% and 42.5% of the runs were not consistent enough with the original ABM model. The simulation speed increase was considerable, ranging between a 73.5% and 2.9% increase. To achieve more consistent dynamically coarse-grained models a lot of additional development time will be required. Therefore, it can be concluded that dynamic coarse-graining is a promising technique, yet only useful in specific niches. In order to gain the full benefits of this technique, the resulting model either needs to be used for near decades in case of Integrated Assessment Models or can be developed beforehand and used under extreme time pressures, in case of crises situations or self-driving cars. Once dynamic coarse-graining models have been fully developed, they do have the potential be an improvement on the currently utilised model, and dependent on their use even to help in saving lives.

6. Literature list

Abdo, M. S., Shah, K., Wahash, H. A., & Panchal, S. K. (2020). On a comprehensive model of the novel coronavirus (COVID-19) under Mittag-Leffler derivative. *Chaos, Solitons & Fractals*, *135*, 109867. https://doi.org/10.1016/j.chaos.2020.109867

Ajelli, M., Gonçalves, B., Balcan, D., Colizza, V., Hu, H., Ramasco, J. J., Merler, S., & Vespignani, A. (2010). Comparing large-scale computational approaches to epidemic modeling: Agent-based versus structured metapopulation models. *BMC Infectious Diseases*, *10*(1), 190. https://doi.org/10.1186/1471-2334-10-190

Alberca, R. W., Oliveira, L. D. M., Branco, A. C. C. C., Pereira, N. Z., & Sato, M. N. (2020). Obesity as a risk factor for COVID-19: an overview. *Critical Reviews in Food Science and Nutrition*, 1–15. https://doi.org/10.1080/10408398.2020.1775546

Alsharhan, A. M. (2021). Survey of Agent-Based Simulations for Modelling COVID-19 Pandemic. *Advances in Science, Technology and Engineering Systems Journal*, *6*(2), 439–447. https://doi.org/10.25046/aj060250

Anastassopoulou, C., Russo, L., Tsakris, A., & Siettos, C. (2020). Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PLOS ONE*, *15*(3), e0230405. https://doi.org/10.1371/journal.pone.0230405

Anderson, R. M., Heesterbeek, H., Klinkenberg, D., & Hollingsworth, T. D. (2020). How will country-based mitigation measures influence the course of the COVID-19 epidemic? *The Lancet*, *395*(10228), 931–934. https://doi.org/10.1016/s0140-6736(20)30567-5

Arino, J., & Portet, S. (2020). A simple model for COVID-19. *Infectious Disease Modelling*, *5*, 309–315. https://doi.org/10.1016/j.idm.2020.04.002

Auping, W. L., Pruyt, E., & Kwakkel, J. H. (2014). Dealing with Multiple Models in System Dynamics: Perspectives on the Future of Copper. International Journal of System Dynamics Applications, 3(4), 17-35. doi:10.4018/ijsda.2014100102

Badham, J. (2021). *JuSt-Social COVID-19* (1.2.0) [Software]. CoMSES Computational Model Library. https://www.comses.net/codebases/1669ffd5-ed68-4495-96e9-7aa635365ce7/releases/1.2.0/

Balcan, D., Gonçalves, B., Hu, H., Ramasco, J. J., Colizza, V., & Vespignani, A. (2010). Modeling the spatial spread of infectious diseases: The GLobal Epidemic and Mobility computational model. *Journal of Computational Science*, *1*(3), 132–145. https://doi.org/10.1016/j.jocs.2010.07.002

Banos, A., Corson, N., Lang, C., Marilleau, N., & Taillandier, P. (2017). Multiscale Modeling: Application to Traffic Flow. *Agent-Based Spatial Simulation with NetLogo, Volume 2*, 37–62. https://doi.org/10.1016/b978-1-78548-157-4.50002-9

Barlow, N. S., & Weinstein, S. J. (2020). Accurate closed-form solution of the SIR epidemic model. *Physica D: Nonlinear Phenomena*, *408*, 132540. https://doi.org/10.1016/j.physd.2020.132540

Barrett, C., Bisset, K., Leidig, J., Marathe, A., & Marathe, M. V. (2010). An Integrated Modeling Environment to Study the Co-evolution of Networks, Individual Behavior and Epidemics. *AI Magazine*, *31*(1), 75. https://doi.org/10.1609/aimag.v31i1.2283

Batista, G. E. A. P. A., Keogh, E. J., Tataw, O. M., & de Souza, V. M. A. (2013). CID: an efficient complexity-invariant distance for time series. *Data Mining and Knowledge Discovery*, *28*(3), 634–669. https://doi.org/10.1007/s10618-013-0312-3

Bjørnstad, O. N., Shea, K., Krzywinski, M., & Altman, N. (2021). The SEIRS model for infectious disease dynamics. *Nature Methods*, *18*(3), 321. https://doi.org/10.1038/s41592-021-01079-6

Bobashev, G. V., Goedecke, D. M., Feng Yu, & Epstein, J. M. (2007). A Hybrid Epidemic Model: Combining The Advantages Of Agent-Based And Equation-Based Approaches. *2007 Winter Simulation Conference*, 1532–1537. https://doi.org/10.1109/wsc.2007.4419767

Boer, C. A., de Bruin, A., & Verbraeck, A. (2009). A survey on distributed simulation in industry. *Journal of Simulation*, *3*(1), 3–16. https://doi.org/10.1057/jos.2008.9

Borshchev, A., & Filippov, A. (2004). From System Dynamics and Discrete Event to Practical Agent Based Modeling: Reasons, Techniques, Tools. *Paper Presented at the Proceedings of the 22nd International Conference of the System Dynamics Society*, 1. https://proceedings.systemdynamics.org/2004/SDS_2004/PAPERS/381BORSH.pdf

Bossert, A., Kersting, M., Timme, M., Schröder, M., Feki, A., Coetzee, J., & Schlüter, J. (2021). Limited containment options of COVID-19 outbreak revealed by regional agent-based simulations for South Africa. *F1000Research*, *10*, 98. https://doi.org/10.12688/f1000research.28250.1

Boulain, N., Simioni, G., & Gignoux, J. (2007). Changing scale in ecological modelling: A bottom up approach with an individual based vegetation model. *Ecological Modelling*, *203*(3–4), 257–269. https://doi.org/10.1016/j.ecolmodel.2006.11.024

Cave, E. (2020). COVID-19 Super-spreaders: Definitional Quandaries and Implications. *Asian Bioethics Review*, *12*(2), 235–242. https://doi.org/10.1007/s41649-020-00118-2

Chahal, K., & Eldabi, T. (2010). A multi-perspective comparison for selection between system dynamics and discrete event simulation. *International Journal of Business Information Systems*, *6*(1), 4. https://doi.org/10.1504/ijbis.2010.034001

Chang, S. L., Harding, N., Zachreson, C., Cliff, O. M., & Prokopenko, M. (2020). Modelling transmission and control of the COVID-19 pandemic in Australia. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-19393-6

Chao, D. L., Halloran, M. E., Obenchain, V. J., & Longini, I. M. (2010). FluTE, a Publicly Available Stochastic Influenza Epidemic Simulation Model. *PLoS Computational Biology*, *6*(1), e1000656. https://doi.org/10.1371/journal.pcbi.1000656

Chatterjee, K., Chatterjee, K., Kumar, A., & Shankar, S. (2020). Healthcare impact of COVID-19 epidemic in India: A stochastic mathematical model. *Medical Journal Armed Forces India*, *76*(2), 147–155. https://doi.org/10.1016/j.mjafi.2020.03.022

Chen, S., Chen, Y., Zhang, S., & Zheng, N. (2019). A Novel Integrated Simulation and Testing Platform for Self-Driving Cars With Hardware in the Loop. *IEEE Transactions on Intelligent Vehicles*, *4*(3), 425–436. https://doi.org/10.1109/tiv.2019.2919470

Choi, S., & Ki, M. (2020). Estimating the reproductive number and the outbreak size of COVID-19 in Korea. *Epidemiology and Health*, *42*, e2020011. https://doi.org/10.4178/epih.e2020011

Chowell, G., Sattenspiel, L., Bansal, S., & Viboud, C. (2016). Mathematical models to characterize early epidemic growth: A review. *Physics of Life Reviews*, *18*, 66–97. https://doi.org/10.1016/j.plrev.2016.07.005

Chwif, L., Barretto, M. R. P., & Paul, R. J. (2000). On simulation model complexity. *2000 Winter Simulation Conference Proceedings (Cat. No.00CH37165)*, *1*, 449–455. https://doi.org/10.1109/wsc.2000.899751

Currie, C. S. M., Fowler, J. W., Kotiadis, K., Monks, T., Onggo, B. S., Robertson, D. A., & Tako, A. A. (2020). How simulation modelling can help reduce the impact of COVID-19. *Journal of Simulation*, *14*(2), 83–97. https://doi.org/10.1080/17477778.2020.1751570

Dazard, J. E., Choe, M., LeBlanc, M., & Rao, J. S. (2015). R package PRIMsrc: Bump Hunting by Patient Rule Induction Method for Survival, Regression and Classification. *Proceedings. American Statistical Association. Annual Meeting*, *2015*, 650–664. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4718587/

di Stefano, B., Fuks, H., & Lawniczak, A. T. (2000). Object-oriented implementation of CA/LGCA modelling applied to the spread of epidemics. *2000 Canadian Conference on Electrical and Computer Engineering. Conference Proceedings. Navigating to a New Era (Cat. No.00TH8492)*, 26–31. https://doi.org/10.1109/ccece.2000.849664

Diekmann, O., Heesterbeek, H., & Britton, T. (2013). *Mathematical Tools for Understanding Infectious Disease Dynamics*. Princeton University Press. http://www.jstor.org/stable/j.cttq9530

Dietz, S., & Hope, C. (2007). Reflections on the Stern Review. *World Economics*, *8*(1), 121-168. https://personal.lse.ac.uk/sternn/104NHS.pdf

Dignum, F., Dignum, V., Davidsson, P., Ghorbani, A., van der Hurk, M., Jensen, M., Kammler, C., Lorig, F., Ludescher, L. G., Melchior, A., Mellema, R., Pastrav, C., Vanhee, L., & Verhagen, H. (2020).

Analysing the Combined Health, Social and Economic Impacts of the Corovanvirus Pandemic Using Agent-Based Social Simulation. *Minds and Machines*, *30*(2), 177–194. https://doi.org/10.1007/s11023-020-09527-6

D'Orazio, M., Bernardini, G., & Quagliarini, E. (2021). A probabilistic model to evaluate the effectiveness of main solutions to COVID-19 spreading in university buildings according to proximity and time-based consolidated criteria. *Building Simulation*, *14*(6), 1795–1809. https://doi.org/10.1007/s12273-021-0770-2

Doyle, J. K., & Ford, D. N. (1998). Mental models concepts for system dynamics research. *System Dynamics Review*, *14*(1), 3–29. https://doi.org/10.1002/(sici)1099-1727(199821)

Edmonds, B., & Hales, D. (2003). Replication, Replication and Replication: Some Hard Lessons from Model Alignment. *Journal of Artificial Societies and Social Simulation*, *6*(4). http://jasss.soc.surrey.ac.uk/6/4/11.html

Epstein, J. M. (2009). Modelling to contain pandemics. *Nature*, *460*(7256), 687. https://doi.org/10.1038/460687a

Epstein, J. M., & Axtell, R. L. (1996). *Growing Artificial Societies*. Brookings Institution Press.

Epstein, J. M., Parker, J., Cummings, D., & Hammond, R. A. (2008). Coupled Contagion Dynamics of Fear and Disease: Mathematical and Computational Explorations. *PLoS ONE*, *3*(12), e3955. https://doi.org/10.1371/journal.pone.0003955

Eubank, S., Guclu, H., Anil Kumar, V. S., Marathe, M. V., Srinivasan, A., Toroczkai, Z., & Wang, N. (2004). Modelling disease outbreaks in realistic urban social networks. *Nature*, *429*(6988), 180–184. https://doi.org/10.1038/nature02541

Fanelli, D., & Piazza, F. (2020). Analysis and forecast of COVID-19 spreading in China, Italy and France. *Chaos, Solitons & Fractals*, *134*, 109761. https://doi.org/10.1016/j.chaos.2020.109761

Ferguson, N. M., Cummings, D. A., Cauchemez, S., Fraser, C., Riley, S., Meeyai, A., Iamsirithaworn, S., & Burke, D. S. (2005). Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature*, *437*(7056), 209–214. https://doi.org/10.1038/nature04017

Figueredo, G. P., Siebers, P. O., & Aickelin, U. (2013). Investigating mathematical models of immuno-interactions with early-stage cancer under an agent-based modelling perspective. *BMC Bioinformatics*, *14*(S6). https://doi.org/10.1186/1471-2105-14-s6-s6

Figueredo, G. P., Siebers, P. O., Owen, M. R., Reps, J., & Aickelin, U. (2014). Comparing Stochastic Differential Equations and Agent-Based Modelling and Simulation for Early-Stage Cancer. *PLoS ONE*, *9*(4), e95150. https://doi.org/10.1371/journal.pone.0095150

Ford, A. (1999). *Modeling the Environment: An Introduction To System Dynamics Modeling Of Environmental Systems* (1st Edition). Island Press.

Ford, A. S., Weitzner, B. D., & Bahl, C. D. (2019). Integration of the Rosetta suite with the python software stack via reproducible packaging and core programming interfaces for distributed simulation. *Protein Science*, *29*(1), 43–51. https://doi.org/10.1002/pro.3721

Forrester, J. W. (1961). *Industrial dynamics.*

Franco, N. (2021). COVID-19 Belgium: Extended SEIR-QD model with nursing homes and long-term scenarios-based forecasts. *Epidemics*, *37*, 100490. https://doi.org/10.1016/j.epidem.2021.100490

Friedman, J. H., & Fisher, N. I. (1999). Bump hunting in high-dimensional data. *Statistics and Computing*, *9*, 123–143. https://doi.org/10.1023/a:1008894516817

Fujimoto, R. (2015). Parallel and distributed simulation. *2015 Winter Simulation Conference (WSC)*, 45–59. https://doi.org/10.1109/wsc.2015.7408152

Fujimoto, R. M., Perumalla, K., Park, A., Wu, H., Ammar, M. H., & Riley, G. F. (2003). Large-scale network simulation: how big? how fast? *11th IEEE/ACM International Symposium on Modeling, Analysis and Simulation of Computer Telecommunications Systems, 2003. MASCOTS 2003.*, 116–123. https://doi.org/10.1109/mascot.2003.1240649

Germann, T. C., Kadau, K., Longini, I. M., & Macken, C. A. (2006). Mitigation strategies for pandemic influenza in the United States. *Proceedings of the National Academy of Sciences*, *103*(15), 5935–5940. https://doi.org/10.1073/pnas.0601266103

Gethmann, J., Probst, C., Bassett, J., Blunk, P., Hövel, P., & Conraths, F. J. (2019). An Epidemiological and Economic Simulation Model to Evaluate Strategies for the Control of Bovine Virus Diarrhea in Germany. *Frontiers in Veterinary Science*, *6*, 406. https://doi.org/10.3389/fvets.2019.00406

Gopalan, A., & Tyagi, H. (2020). How Reliable are Test Numbers for Revealing the COVID-19 Ground Truth and Applying Interventions? *Journal of the Indian Institute of Science*, *100*(4), 863–884. https://doi.org/10.1007/s41745-020-00210-4

Granich, R. M., Gilks, C. F., Dye, C., de Cock, K. M., & Williams, B. G. (2009). Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. *The Lancet*, *373*(9657), 48–57. https://doi.org/10.1016/s0140-6736(08)61697-9

Gray, R., & Wotherspoon, S. (2012). Increasing model efficiency by dynamically changing model representations. *Environmental Modelling & Software*, *30*, 115–122. https://doi.org/10.1016/j.envsoft.2011.08.012

Gray, R., & Wotherspoon, S. (2015). Adaptive submodel selection in hybrid models. *Frontiers in Environmental Science*, *3*. https://doi.org/10.3389/fenvs.2015.00058

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S. K., Huse, G., Huth, A., Jepsen, J. U., Jørgensen, C., Mooij, W. M., Müller, B., Pe'er, G., Piou, C., Railsback, S. F., Robbins, A. M., . . . DeAngelis, D. L. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, *198*(1–2), 115–126. https://doi.org/10.1016/j.ecolmodel.2006.04.023

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., & Railsback, S. F. (2010). The ODD protocol: A review and first update. *Ecological Modelling*, *221*(23), 2760–2768. https://doi.org/10.1016/j.ecolmodel.2010.08.019

Grimm, V., & Railsback, S. F. (2005). *Individual-based Modeling and Ecology: chapter 5*. Princeton University Press. https://doi.org/10.1515/9781400850624

Grimm, V., Railsback, S. F., Vincenot, C. E., Berger, U., Gallagher, C., DeAngelis, D. L., Edmonds, B., Ge, J., Giske, J., Groeneveld, J., Johnston, A. S., Milles, A., Nabe-Nielsen, J., Polhill, J. G., Radchuk, V., Rohwäder, M. S., Stillman, R. A., Thiele, J. C., & Ayllón, D. (2020). The ODD Protocol for Describing Agent-Based and Other Simulation Models: A Second Update to Improve Clarity, Replication, and Structural Realism. *Journal of Artificial Societies and Social Simulation*, *23*(2). https://doi.org/10.18564/jasss.4259

Großmann, G., Backenköhler, M., & Wolf, V. (in press). Why ODE models for COVID-19 fail: Heterogeneity shapes epidemic dynamics. *MedRxiv: Preprint*. https://doi.org/10.1101/2021.03.25.21254292

Guivarch, C., Rozenberg, J., & Schweizer, V. (2016). The diversity of socio-economic pathways and CO2 emissions scenarios: Insights from the investigation of a scenarios database. *Environmental Modelling & Software*, *80*, 336–353. https://doi.org/10.1016/j.envsoft.2016.03.006

Biswas, M. H. A., Paiva, L. T., & de Pinho, M. (2014). A SEIR model for control of infectious diseases with constraints. *Mathematical Biosciences and Engineering*, *11*(4), 761–784. https://doi.org/10.3934/mbe.2014.11.761

Halloran, M. E. (2002). Containing Bioterrorist Smallpox. *Science*, *298*(5597), 1428–1432. https://doi.org/10.1126/science.1074674

Hasan, A., Susanto, H., Kasim, M. F., Nuraini, N., Lestari, B., Triany, D., & Widyastuti, W. (2020). Superspreading in early transmissions of COVID-19 in Indonesia. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-79352-5

Hethcote, H. W., Stech, H. W., & van den Driessche, P. (1981). PERIODICITY AND STABILITY IN EPIDEMIC MODELS: A SURVEY. *Differential Equations and Applications in Ecology, Epidemics, and Population Problems*, 65–82. https://doi.org/10.1016/b978-0-12-148360-9.50011-1

Hinch, R., Probert, W. J. M., Nurtay, A., Kendall, M., Wymant, C., Hall, M., Lythgoe, K., Bulas Cruz, A., Zhao, L., Stewart, A., Ferretti, L., Montero, D., Warren, J., Mather, N., Abueg, M., Wu, N.,

Legat, O., Bentley, K., Mead, T., . . . Fraser, C. (2021). OpenABM-Covid19—An agent-based model for non-pharmaceutical interventions against COVID-19 including contact tracing. *PLOS Computational Biology*, *17*(7), e1009146. https://doi.org/10.1371/journal.pcbi.1009146

Hoertel, N., Blachier, M., Blanco, C., Olfson, M., Massetti, M., Limosin, F., & Leleu, H. (2020). Facing the COVID-19 epidemic in NYC: a stochastic agent-based model of various intervention strategies. *MedRxiv : The Preprint Server for Health Sciences*, *2020.04.23.20076885*. https://doi.org/10.1101/2020.04.23.20076885

House, T. (2012). Modelling epidemics on networks. *Contemporary Physics*, *53*(3), 213–225. https://doi.org/10.1080/00107514.2011.644443

Hsu, P. L., & Robbins, H. (1947). Complete Convergence and the Law of Large Numbers. *Proceedings of the National Academy of Sciences*, *33*(2), 25–31. https://doi.org/10.1073/pnas.33.2.25

Huang, S.-K., Lin, M.-T., Chen, H.-C., Huang, S.-C., & Wu, M.-H. (2013). Epidemiology of Kawasaki Disease: Prevalence from National Database and Future Trends Projection by System Dynamics Modeling. *The Journal of Pediatrics*, *163*(1), 126–131.e1. https://doi.org/10.1016/j.jpeds.2012.12.011

Ibarra-Vega, D. (2020). Lockdown, one, two, none, or smart. Modeling containing covid-19 infection. A conceptual model. *Science of The Total Environment*, *730*, 138917. https://doi.org/10.1016/j.scitotenv.2020.138917

Ivorra, B., Ferrández, M., Vela-Pérez, M., & Ramos, A. (2020). Mathematical modeling of the spread of the coronavirus disease 2019 (COVID-19) taking into account the undetected infections. The case of China. *Communications in Nonlinear Science and Numerical Simulation*, *88*, 105303. https://doi.org/10.1016/j.cnsns.2020.105303

Jackson, M. (1991). *Systems Methodology for the Management Sciences*. Springer Publishing.

Jackson, M. C. (1994). Critical systems thinking: Beyond the fragments. *System Dynamics Review*, *10*(2–3), 213–229. https://doi.org/10.1002/sdr.4260100209

Kai, D., Goldstein, G. P., Morgunov, A., Nangalia, V., & Rotkirch, A. (2020). Universal Masking is Urgent in the COVID-19 Pandemic: SEIR and Agent Based Models, Empirical Validation, Policy Recommendations. *ArXiv Preprint ArXiv*, *2004.13553*. https://arxiv.org/abs/2004.13553

Keeling, M. J., & Grenfell, B. T. (2000). Individual-based Perspectives on R0. *Journal of Theoretical Biology*, *203*(1), 51–61. https://doi.org/10.1006/jtbi.1999.1064

Kim, S., Kim, Y. J., Peck, K. R., & Jung, E. (2020). School Opening Delay Effect on Transmission Dynamics of Coronavirus Disease 2019 in Korea: Based on Mathematical Modeling and Simulation Study. *Journal of Korean Medical Science*, *35*(13). https://doi.org/10.3346/jkms.2020.35.e143

Kreft, J. U., Picioreanu, C., van Loosdrecht, M. C. M., & Wimpenny, J. W. T. (2001). Individual-based modelling of biofilms. *Microbiology*, *147*(11), 2897–2912. https://doi.org/10.1099/00221287-147-11-2897

Kuniya, T. (2020). Prediction of the Epidemic Peak of Coronavirus Disease in Japan, 2020. *Journal of Clinical Medicine*, *9*(3), 789. https://doi.org/10.3390/jcm9030789

Kwakkel, J. H. (2017). The Exploratory Modeling Workbench: An open source toolkit for exploratory modeling, scenario discovery, and (multi-objective) robust decision making. *Environmental Modelling & Software*, *96*, 239–250. https://doi.org/10.1016/j.envsoft.2017.06.054

Lane, D. C. (2008). The emergence and use of diagramming in system dynamics: a critical account. *Systems Research and Behavioral Science*, *25*(1), 3–23. https://doi.org/10.1002/sres.826

ŁAtuszyńska, M. ł. (2019). Hybrid System Dynamics—Agent-Based Simulation for Research in Economics and Business. *Experimental and Quantitative Methods in Contemporary Economics*, 229–248. https://doi.org/10.1007/978-3-030-30251-1_17

Levitt, M., & Warshel, A. (1975). Computer simulation of protein folding. *Nature*, *253*, 694–698. https://doi.org/10.1038/253694a0

Li, S., Song, K., Yang, B., Gao, Y., & Gao, X. (2020). Preliminary Assessment of the COVID-19 Outbreak Using 3-Staged Model e-ISHR. *Journal of Shanghai Jiaotong University (Science)*, *25*(2), 157–164. https://doi.org/10.1007/s12204-020-2169-0

Lipsitch, M. (2003). Transmission Dynamics and Control of Severe Acute Respiratory Syndrome. *Science*, *300*(5627), 1966–1970. https://doi.org/10.1126/science.1086616

Liu, J., Zou, Y., Wang, W., Zhang, L., Liu, X., Ding, Q., Qin, Z., & ČEpin, M. (2021). Analysis of dependencies among performance shaping factors in human reliability analysis based on a system dynamics approach. *Reliability Engineering & System Safety*, *215*, 107890. https://doi.org/10.1016/j.ress.2021.107890

Longini, I. M. (2004). Containing Pandemic Influenza with Antiviral Agents. *American Journal of Epidemiology*, *159*(7), 623–633. https://doi.org/10.1093/aje/kwh092

Longini, I. M. (2005). Containing Pandemic Influenza at the Source. *Science*, *309*(5737), 1083–1087. https://doi.org/10.1126/science.1115717

Luo, Y., Yao, L., Zhou, L., Yuan, F., & Zhong, X. (2020). Factors influencing health behaviours during the coronavirus disease 2019 outbreak in China: an extended information-motivation-behaviour skills model. *Public Health*, *185*, 298–305. https://doi.org/10.1016/j.puhe.2020.06.057

Macal, C. M. (2010). To agent-based simulation from System Dynamics. *Proceedings of the 2010 Winter Simulation Conference*, 371–382. https://doi.org/10.1109/wsc.2010.5679148

Mammeri, Y. (2020). A reaction-diffusion system to better comprehend the unlockdown: Application of SEIR-type model with diffusion to the spatial spread of COVID-19 in France. *Computational and Mathematical Biophysics*, *8*(1), 102–113. https://doi.org/10.1515/cmb-2020-0104

Manchein, C., Brugnago, E. L., da Silva, R. M., Mendes, C. F. O., & Beims, M. W. (2020). Strong correlations between power-law growth of COVID-19 in four continents and the inefficiency of soft quarantine strategies. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *30*(4), 041102. https://doi.org/10.1063/5.0009454

Maugeri, A., Barchitta, M., Battiato, S., & Agodi, A. (2020). Estimation of Unreported Novel Coronavirus (SARS-CoV-2) Infections from Reported Deaths: A Susceptible–Exposed–Infectious–Recovered–Dead Model. *Journal of Clinical Medicine*, *9*(5), 1350. https://doi.org/10.3390/jcm9051350

Merler, S., Ajelli, M., Fumanelli, L., Gomes, M. F. C., Piontti, A. P., Rossi, L., Chao, D. L., Longini, I. M., Halloran, M. E., & Vespignani, A. (2015). Spatiotemporal spread of the 2014 outbreak of Ebola virus disease in Liberia and the effectiveness of non-pharmaceutical interventions: a computational modelling analysis. *The Lancet Infectious Diseases*, *15*(2), 204–211. https://doi.org/10.1016/s1473-3099(14)71074-6

M'hamdi, A., & Nemiche, M. (2018). Bottom-Up and Top-Down Approaches to Simulate Complex Social Phenomena. *International Journal of Applied Evolutionary Computation*, *9*(2), 1–16. https://doi.org/10.4018/ijaec.2018040101

Monasterolo, I., & Raberto, M. (2016). A Hybrid System Dynamics Agent Based Model to Assess the Role of Green Fiscal and Monetary Policies. *SSRN Electronic Journal*, 22. https://doi.org/10.2139/ssrn.2748266

Mustafee, N., & Taylor, S. J. E. (2009). Speeding up simulation applications using WinGrid. *Concurrency and Computation: Practice and Experience*, *21*(11), 1504–1523. https://doi.org/10.1002/cpe.1401

Nasirzadeh, F., Khanzadi, M., & Mir, M. (2017). A hybrid simulation framework for modelling construction projects using agent-based modelling and system dynamics: an application to model construction workers' safety behavior. *International Journal of Construction Management*, *18*(2), 132–143. https://doi.org/10.1080/15623599.2017.1285485

NU.nl. (2021a, July 11). *Gommers: 'Rutte en De Jonge hadden excuses moeten aanbieden'*. NU - Het laatste nieuws het eerst op NU.nl. Retrieved 14 September 2021, from https://www.nu.nl/coronavirus/6144953/gommers-rutte-en-de-jonge-hadden-excuses-moeten-aanbieden.html

NU.nl. (2021b, July 11). *OMT-lid vreest dat besmettingen 'doorsijpelen' naar kwetsbaren*. NU - Het laatste nieuws het eerst op NU.nl. Retrieved 14 September 2021, from https://www.nu.nl/coronavirus/6144895/omt-lid-vreest-dat-besmettingen-doorsijpelen-naar-kwetsbaren.html

Ossimitz, G., & Mrotzek, M. (2008, July). *The Basics of System Dynamics: Discrete vs. Continuous Modelling of Time*. In Proceedings of the 26th International Conference of the System Dynamics Society. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.408.4644&rep=rep1&type=pdf

Parry, H. R., & Evans, A. J. (2008). A comparative analysis of parallel processing and super-individual methods for improving the computational performance of a large individual-based model. *Ecological Modelling*, *214*(2–4), 141–152. https://doi.org/10.1016/j.ecolmodel.2008.02.002

Perez, L., & Dragicevic, S. (2009). An agent-based approach for modeling dynamics of contagious disease spread. *International Journal of Health Geographics*, *8*(1), 50. https://doi.org/10.1186/1476-072x-8-50

Poirier, J. (n.d.). *work of Lavoisier: Chapter 8*. Antoine-Laurent de Lavoisier (1743–1794) Life and Works. Retrieved 2 December 2021, from http://historyofscience.free.fr/Lavoisier-Friends/a_chap8_lavoisier.html#unification

Rabinovich, S. G. (2017). Evaluating Measurement Accuracy. *Springer Series in Measurement Science and Technology*. Published. https://doi.org/10.1007/978-3-319-60125-0

Railsback, S. F. (2001). Concepts from complex adaptive systems as a framework for individual-based modelling. *Ecological Modelling*, *139*(1), 47–62. https://doi.org/10.1016/s0304-3800(01)00228-9

Railsback, S. F., & Grimm, V. (2019). *Agent-Based and Individual-Based Modeling: A Practical Introduction, Second Edition* (2nd ed.). Princeton University Press.

Railsback, S. F., Lytinen, S. L., & Jackson, S. K. (2006). Agent-based Simulation Platforms: Review and Development Recommendations. *SIMULATION*, *82*(9), 609–623. https://doi.org/10.1177/0037549706073695

Riley, S. (2003). Transmission Dynamics of the Etiological Agent of SARS in Hong Kong: Impact of Public Health Interventions. *Science*, *300*(5627), 1961–1966. https://doi.org/10.1126/science.1086478

Riou, J., & Althaus, C. L. (2020). Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020. *Eurosurveillance*, *25*(4). https://doi.org/10.2807/1560-7917.es.2020.25.4.2000058

RIVM. (2021, September 14). *Positief geteste mensen*. Rijksoverheid. Retrieved 14 September 2021, from https://coronadashboard.rijksoverheid.nl/landelijk/positief-geteste-mensen

Rockett, R. J., Arnott, A., Lam, C., Sadsad, R., Timms, V., Gray, K. A., Eden, J. S., Chang, S., Gall, M., Draper, J., Sim, E. M., Bachmann, N. L., Carter, I., Basile, K., Byun, R., O'Sullivan, M. V., Chen, S. C. A., Maddocks, S., Sorrell, T. C., . . . Sintchenko, V. (2020). Revealing COVID-19 transmission in Australia by SARS-CoV-2 genome sequencing and agent-based modeling. *Nature Medicine*, *26*(9), 1398–1404. https://doi.org/10.1038/s41591-020-1000-7

Rodriguez-Fernández, J. (1999). Ockham's razor. *Endeavour*, *23*(3), 121–125. https://doi.org/10.1016/s0160-9327(99)01199-0

Rosner, B., Glynn, R. J., & Ting Lee, M. (2003). Incorporation of Clustering Effects for the Wilcoxon Rank Sum Test: A Large-Sample Approach. *Biometrics*, *59*(4), 1089–1098. https://doi.org/10.1111/j.0006-341x.2003.00125.x

Rossum, G. (1995). *Python reference manual*. CWI (Centre for Mathematics and Computer Science). https://ir.cwi.nl/pub/5008/05008D.pdf

Rothman, K. J. (2012). *Epidemiology: An Introduction* (2nd ed.). Oxford University Press.

RTL (1997, April 24). https://citaten.net/zoeken/citaten_van-johan_cruijff.html

Saker, L., Lee, K., Cannito, B., Gilmore, A., & Campbell-Lendrum, D. (2004). *Globalization and infectious diseases: A review of the linkages*. World Health Organisation. https://www.who.int/tdr/publications/documents/seb_topic3.pdf

Sartwell, P. E. (1966). THE INCUBATION PERIOD AND THE DYNAMICS OF INFECTIOUS DISEASE. *American Journal of Epidemiology*, *83*(2), 204–216. https://doi.org/10.1093/oxfordjournals.aje.a120576

Saunders, M. G., & Voth, G. A. (2013). Coarse-Graining Methods for Computational Biology. *Annual Review of Biophysics*, *42*(1), 73–93. https://doi.org/10.1146/annurev-biophys-083012-130348

Scholl, H. J. (2001). Agent-based and system dynamics modeling: a call for cross study and joint research. *Proceedings of the 34th Annual Hawaii International Conference on System Sciences*, 8. https://doi.org/10.1109/hicss.2001.926296

Seidenberg, M. S. (1993). Connectionist Models and Cognitive Theory. *Psychological Science*, *4*(4), 228–235. https://doi.org/10.1111/j.1467-9280.1993.tb00266.x

Shamil, M. S., Farheen, F., Ibtehaz, N., Khan, I. M., & Rahman, M. S. (2021). An Agent-Based Modeling of COVID-19: Validation, Analysis, and Recommendations. *Cognitive Computation*. Published. https://doi.org/10.1007/s12559-020-09801-w

Shao, P., & Shan, Y. (in press). Beware of asymptomatic transmission: Study on 2019-nCoV prevention and control measures based on extended SEIR model. *BioRxiv*. https://doi.org/10.1101/2020.01.28.923169

Shapiro, A. (2003). Monte Carlo Sampling Methods. *Handbooks in Operations Research and Management Science*, 353–425. https://doi.org/10.1016/s0927-0507(03)10006-0

Shen, C., Taleb, N. N., & Bar-Yam, Y. (2020). Review of Ferguson et al 'Impact of non-pharmaceutical interventions...' *New England Complex Systems Institute*. Published. https://www.academia.edu/download/62435084/ReviewOfFergussonf.pdf

Shi, Z. Z., Wu, C. H., & Ben-Arieh, D. (2014). Agent-Based Model: A Surging Tool to Simulate Infectious Diseases in the Immune System. *Open Journal of Modelling and Simulation*, *02*(01), 12–22. https://doi.org/10.4236/ojmsi.2014.21004

Siettos, C. I., & Russo, L. (2013). Mathematical modeling of infectious disease dynamics. *Virulence*, *4*(4), 295–306. https://doi.org/10.4161/viru.24041

Silva, P. C., Batista, P. V., Lima, H. S., Alves, M. A., Guimarães, F. G., & Silva, R. C. (2020). COVID-ABS: An agent-based model of COVID-19 epidemic to simulate health and economic effects of social distancing interventions. *Chaos, Solitons & Fractals*, *139*, 110088. https://doi.org/10.1016/j.chaos.2020.110088

Sklar, E. (2007). NetLogo, a Multi-agent Simulation Environment. *Artificial Life*, *13*(3), 303–311. https://doi.org/10.1162/artl.2007.13.3.303

Søreide, K., Hallet, J., Matthews, J. B., Schnitzbauer, A. A., Line, P. D., Lai, P. B. S., Otero, J., Callegaro, D., Warner, S. G., Baxter, N. N., Teh, C. S. C., Ng-Kamstra, J., Meara, J. G., Hagander, L., & Lorenzon, L. (2020). Immediate and long-term impact of the COVID-19 pandemic on delivery of surgical services. *British Journal of Surgery*, Advance online publication. https://doi.org/10.1002/bjs.11670

Stanton, E. A., Ackerman, F., & Kartha, S. (2009). Inside the integrated assessment models: Four issues in climate economics. *Climate and Development*, *1*(2), 166–184. https://doi.org/10.3763/cdev.2009.0015

Stein, M. (1987). Large Sample Properties of Simulations Using Latin Hypercube Sampling. *Technometrics*, *29*(2), 143–151. https://doi.org/10.1080/00401706.1987.10488205

Steinmann, P., Auping, W. L., & Kwakkel, J. H. (2020). Behavior-based scenario discovery using time series clustering. *Technological Forecasting and Social Change*, *156*, 120052. https://doi.org/10.1016/j.techfore.2020.120052

Sterman, J. D. (2000). *Business dynamics: Systems thinking and modeling for a complex world*.

Swinerd, C., & McNaught, K. R. (2014). Simulating the diffusion of technological innovation with an integrated hybrid agent-based system dynamics model. *Journal of Simulation*, *8*(3), 231–240. https://doi.org/10.1057/jos.2014.2

Tang, B., Xia, F., Tang, S., Bragazzi, N. L., Li, Q., Sun, X., Liang, J., Xiao, Y., & Wu, J. (2020). The effectiveness of quarantine and isolation determine the trend of the COVID-19 epidemic in the final phase of the current outbreak in China. *International Journal of Infectious Diseases*, *96*, 636–647. https://doi.org/10.1016/j.ijid.2020.05.113

Taylor, S. J. E. (2019). Distributed simulation: state-of-the-art and potential for operational research. *European Journal of Operational Research*, *273*(1), 1–19. https://doi.org/10.1016/j.ejor.2018.04.032

Tesfamariam, D., & Lindberg, B. (2005). Aggregate analysis of manufacturing systems using system dynamics and ANP. *Computers & Industrial Engineering*, *49*(1), 98–117. https://doi.org/10.1016/j.cie.2005.05.001

Thompson, K. M., & Tebbens, R. J. D. (2008). Using system dynamics to develop policies that matter: global management of poliomyelitis and beyond. *System Dynamics Review*, *24*(4), 433–449. https://doi.org/10.1002/sdr.419

Tuite, A. R., Fisman, D. N., & Greer, A. L. (2020). Mathematical modelling of COVID-19 transmission and mitigation strategies in the population of Ontario, Canada. *Canadian Medical Association Journal*, *192*(19), E497–E505. https://doi.org/10.1503/cmaj.200476

van Kleef, E., Robotham, J. V., Jit, M., Deeny, S. R., & Edmunds, W. J. (2013). Modelling the transmission of healthcare associated infections: a systematic review. *BMC Infectious Diseases*, *13*(1), 1. https://doi.org/10.1186/1471-2334-13-294

van Vuuren, D. P., Lowe, J., Stehfest, E., Gohar, L., Hof, A. F., Hope, C., Warren, R., Meinshausen, M., & Plattner, G. K. (2009). How well do integrated assessment models simulate climate change? *Climatic Change*, *104*(2), 255–285. https://doi.org/10.1007/s10584-009-9764-2

Ventana Systems Inc. (2011). *Vensim Reference Manual*.

Vincenot, C. E., Giannino, F., Rietkerk, M., Moriya, K., & Mazzoleni, S. (2011). Theoretical considerations on the combined use of System Dynamics and individual-based modeling in ecology. *Ecological Modelling*, *222*(1), 210–218. https://doi.org/10.1016/j.ecolmodel.2010.09.029

Vonk Noordegraaf, A., Buijtels, J. A. A. M., Dijkhuizen, A. A., Franken, P., Stegeman, J. A., & Verhoeff, J. (1998). An epidemiological and economic simulation model to evaluate the spread and control of infectious bovine rhinotracheitis in the Netherlands. *Preventive Veterinary Medicine*, *36*(3), 219–238. https://doi.org/10.1016/s0167-5877(98)00081-6

Wallentin, G., & Neuwirth, C. (2017). Dynamic hybrid modelling: Switching between AB and SD designs of a predator-prey model. *Ecological Modelling*, *345*, 165–175. https://doi.org/10.1016/j.ecolmodel.2016.11.007

Wang, H., Wang, Z., Dong, Y., Chang, R., Xu, C., Yu, X., Zhang, S., Tsamlag, L., Shang, M., Huang, J., Wang, Y., Xu, G., Shen, T., Zhang, X., & Cai, Y. (2020). Phase-adjusted estimation of the number of Coronavirus Disease 2019 cases in Wuhan, China. *Cell Discovery*, *6*(1). https://doi.org/10.1038/s41421-020-0148-0

Wang, Q., & Chatwin, C. R. (2004). Key issues and developments in modelling and simulation-based methodologies for manufacturing systems analysis, design and performance evaluation. *The International Journal of Advanced Manufacturing Technology*, *25*(11–12), 1254–1265. https://doi.org/10.1007/s00170-003-1957-7

Watson, T. A., Doherty, J. E., & Christensen, S. (2013). Parameter and predictive outcomes of model simplification. *Water Resources Research*, *49*(7), 3952–3977. https://doi.org/10.1002/wrcr.20145

Weyant, John & Davidson, Ocen & Dowlatabadi, Hadi & Edmonds, James & Grubb, Michael & Parson, Edward & Richels, Richard & Rotmans, Jan & Shukla, Priyadarshi & Tol, Richard & Cline, William & Fankhauser, Samuel. (1996). Integrated Assessment of Climate Change: An Overview and Comparison of Approaches and Results. https://www.researchgate.net/publication/221678860_Integrated_Assessment_of_Climate_Change_An_Overview_and_Comparison_of_Approaches_and_Results

Wójcik, D., & Ioannou, S. (2020). COVID-19 and Finance: Market Developments So Far and Potential Impacts on the Financial Sector and Centres. *Tijdschrift Voor Economische En Sociale Geografie*, *111*(3), 387–400. https://doi.org/10.1111/tesg.12434

Xiong, J., Lipsitz, O., Nasri, F., Lui, L. M. W., Gill, H., Phan, L., Chen-Li, D., Iacobucci, M., Ho, R., Majeed, A., & McIntyre, R. S. (2020). Impact of COVID-19 pandemic on mental health in the general

population: A systematic review. *Journal of Affective Disorders*, *277*, 55–64. https://doi.org/10.1016/j.jad.2020.08.001

Xue, Y., Ludovice, P. J., & Grover, M. A. (2011). Dynamic coarse graining in complex system simulation. *Proceedings of the 2011 American Control Conference*, 5031–5036. https://doi.org/10.1109/acc.2011.5990888

Appendix I: Validation of Simple model

In this Appendix all validation test performed on the Simple ABM and SD model will be extensively discussed. The resulting conclusions can be found in paragraph 3.7.2.

Parameter verification Simple model

The parameter verification test entails the comparison of the different variables in the model to their real world counter-part. Ideally every variable should have a distinctive meaning and a logical value attached to them. If there are a significant amount of variables without such meaning, this could be an indication of overfitting a model, meaning the model is less likely to be an accurate representation of reality. In regard to the Simple model the parameters will be primarily judged to their meaning and not to their actual value, as the model is highly simplified.

In the ABM model there are a few aspects that are not present in the SD model, these will be discussed first. The movement procedure is specified as a random turn right combined with a random turn left and a single movement forward. This does not relate to any real-world walking phenomenon, which is generally with intent in mind, instead of being random. However, due to the simplified nature of the model, in combination with the absence of a utilisation goal this is perfectly acceptable.

All the variables that are present in both models have a clear meaning, and their values relative to each other make sense. The only variable present in Vensim without having a real-world counterpart is the “correction factor” variable. This variable is not used to correct unit errors, but because the Vensim overestimates the amount of contacts between Sick and Susceptible people that take place. This is probably needed to account for the heterogeneity present in the Netlogo model. In the Vensim model every person has an equal amount of chance of encountering any other person. In the Netlogo model this is not the case, as recently infected people have an increased chance of encountering their infector, due to them being close to each other. It would be better if it was not necessary to use an arbitrary factor to solve this problem, but in lights of the purpose of this model it is not very relevant.

Structure verification test Simple model

Verifying the structure of the Simple models will be the most important test in regard to the purpose of the model. Since this model is purely utilised to serve as a test-case for the Extensive model, the most important thing is that the basic structure follows the structure of standard transmission models. If this is not the case, it could be argued that using this model as the test-case would be senseless. For agent-based transmission models there a number of key characteristics that need to be included for it to be considered representative. These are the standard SEIR structure, spatial effects, heterogeneity of the population and stochasticity. The first aspect relates primarily to transmission models as a concept, whereas the latter three constitute properties that have an important role in real world transmissions, while also being difficult to include in a SD model of transmission.

The SEIR structure is included in one of its most basic forms, reducing it to a SIR model. These three states are used to generate the oscillatory behaviour usually observed when analysing infectious diseases, like viruses. Spatial effects are included via the movement options combined with the proximity-based infections. Heterogeneity is a direct result of the combination of these two properties. Stochasticity is also seen included in the movement options, but is also found in odds of reproducing, infecting one another, and the chance of recovering. The inclusion of these properties, make this Simple ABM virus model a good choice in regard to the methodological purpose of this research.

Since the structure of the SD model is the same as the ABM model structure it also incorporates a reduced SEIR model as its foundation.

Dimensional-consistency Simple model

This test will help to build confidence in the internal consistency of the SD model. Since most variables in the Netlogo model do not have dimensions attached to them, this test will only be performed on the SD side of the Simple model. Consistency of the dimensions in the Vensim model would also implicate the dimensions of the Netlogo model variables are consistent, as the functions essentially the same. Dimensions failing to add up implicate a structural mistake in the model and should therefore be analysed carefully. This test is especially useful in conjunction with the parameter verification test, as this eliminates the variables used to fix the dimensions without having any real-life meaning.

As Netlogo models do not include units for most variables, this test will be performed using the SD version of the model. Since that model incorporates all variables present in the Netlogo model, this should also give a good indication of whether to original model was dimensionally consistent. In the parameter verification test it was already concluded that there are no variables included without meaning, with the purpose of altering the dimension of a variable, since removing the correction factor will keep the dimension of the related variable the same. Utilising the build-in unit check function of Vensim, it can be concluded that all included dimensions add up. This does not necessarily mean all dimensions are definitely correct for every variable but does build a significant amount of confidence in the base structure of the model.

Extreme conditions test Simple model

To ensure the models behaviour does not show any impossibilities, the model is tested under conditions that are most likely to break the model. For instance, when putting the carrying-capacity to zero the amount of agents in the model should never increase and always decline. More obvious mistakes, like variable going negative, are also indicators of failing this test. By passing this test, the likelihood of the basic structure functioning correctly is increased.

In table 10 below it is shown which variables are changed, their original value, and the value(s) used for the extreme conditions testing. All changes are tested both individually and in conjunction with each other. Due to the stochastic nature of the ABM model, the model will be ran five times. All these runs will be evaluated individually, as looking at the aggregate of the five might hide some issues.

Table 10. Variables and values used for extreme condition testing

Netlogo variable	Vensim variable	Original value	Lower value	Upper value
carrying-capacity	Carrying capacity	300	0	10000
Infectiousness	Chance of infection	65% - 0.65	0	100% - 1
Immunity-duration	Immunity duration	52	0	1000
Lifespan	Lifespan	2600	0	10000
Chance-recover	Recovery-rate	75% - 0.75	0	100% - 1
Chance-reproduce	Reproduction chance	1 – 0.01	0	100 – 1
Duration	Sickness duration	20	0	1000

Depending on the variable that underwent the change the most relevant KPI will be shown, from a choice of the Healthy people, Sick people, Immune people, or All people. Due to visibility concerns we will not show all KPIs for every run. The images on the right showcase the Netlogo outcome, with the left showing the outcome Vensim generates. All results of these tests can be seen in the figures below.

During the test, it became evident that the SD model breaks when immunity duration goes below 1, this is probably due to a combination of the Time Step used (0.03125) and the fact that it is used inside a delay function. For diseases that have a shorter immune period than 1 week this model does not function correctly. In that case it is advisable to switch to a model that utilises days every time step instead of weeks. The same holds true of the sickness duration. Additionally, the flows are

coupled to each other by using the fraction of sickness or immunity duration compared to lifespan. Due to this the model starts to fail when Lifespan becomes lower than either of these variables.

When looking at the combination of all high factors in figure 26, it can be seen that the behaviour of Immune people is different in the SD model. This is due to the method of disease modelling. In the SD model Turtles will stay sick for 1000 weeks on average, whereas turtles in the ABM model stay sick for exactly 1000 weeks. This results in a discrepancy in the immune population

Lastly, to ensure the model also functions under different combinations of these factors the sensitivity tool in Vensim is utilised. The result can be seen in figure 27. All variables are varied between the ranges depicted in table 10. It showcases the basic structure of the model functions as it should.

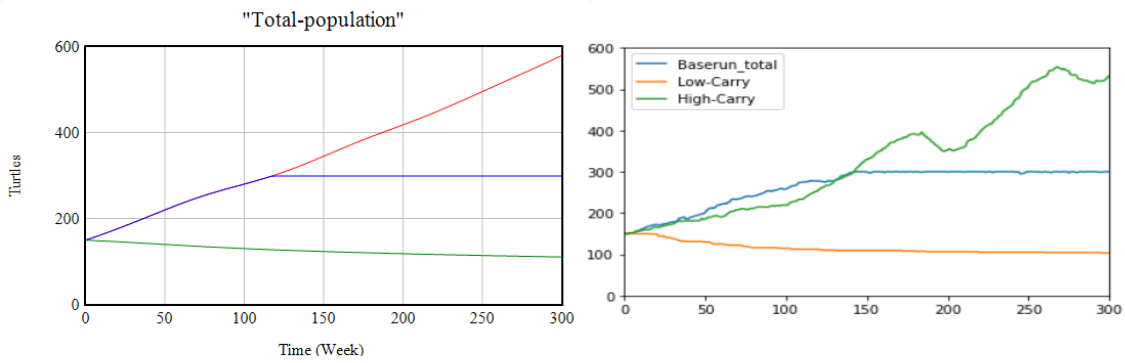

Figure 19: Extreme carrying capacities

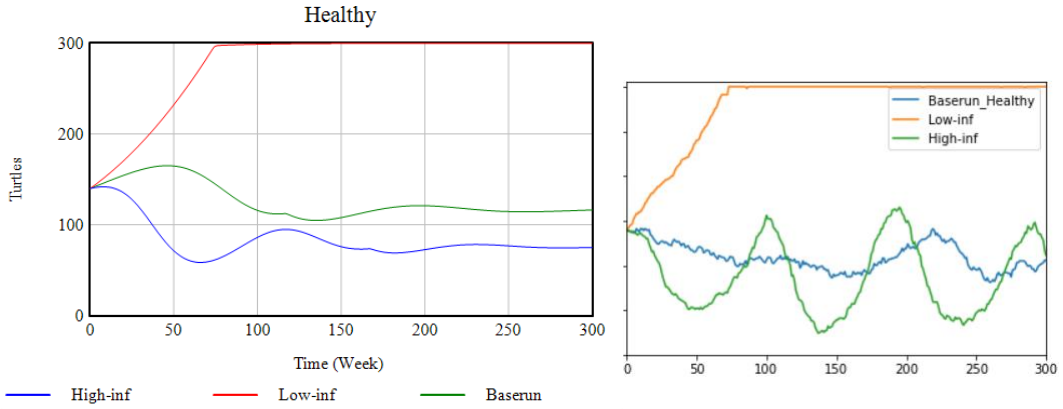

Figure 21: Extreme infectiousness

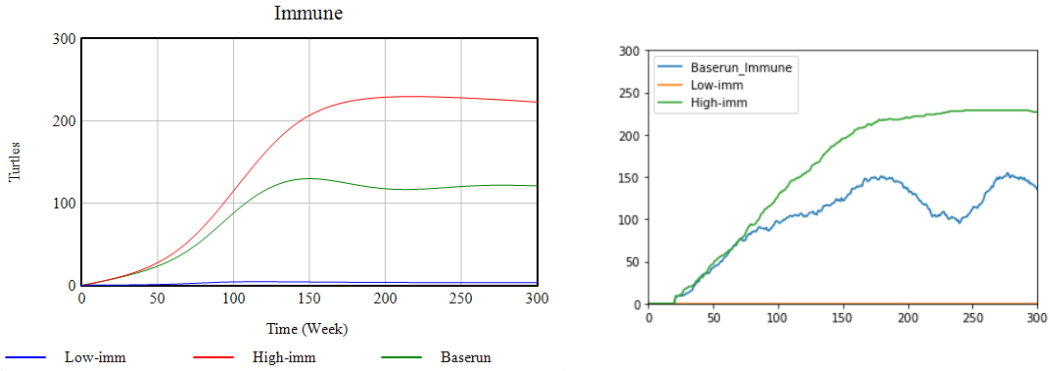

Figure 20: Extreme Immunity duration

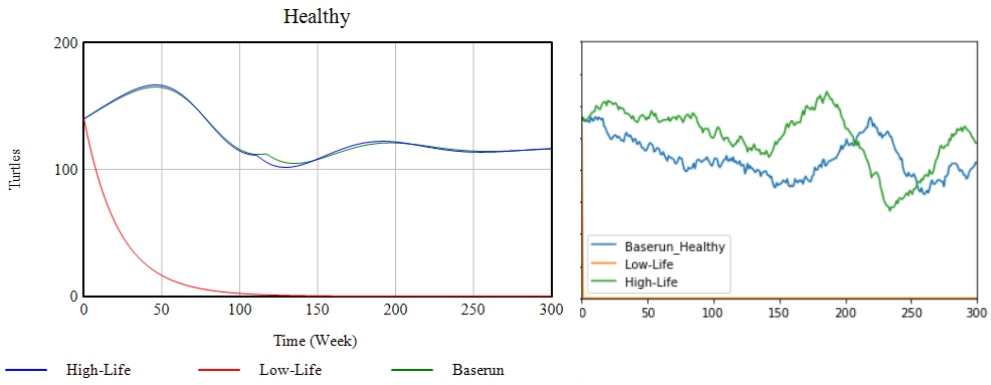

Figure 22: Extreme lifespan

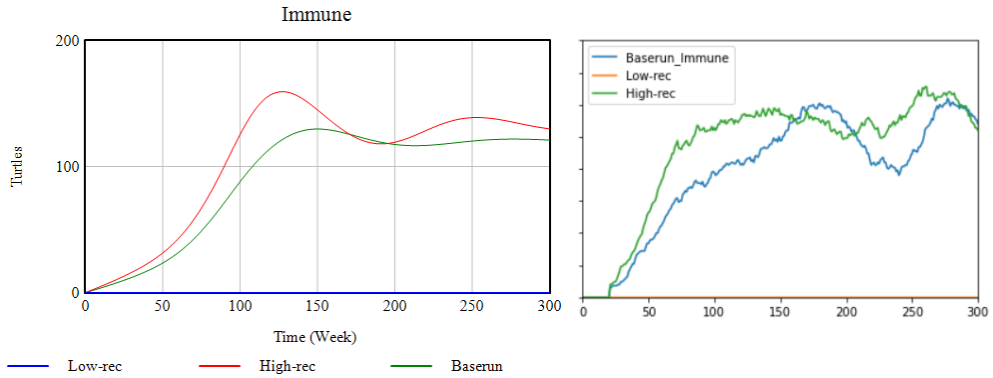

Figure 24: Extreme recovery-rate

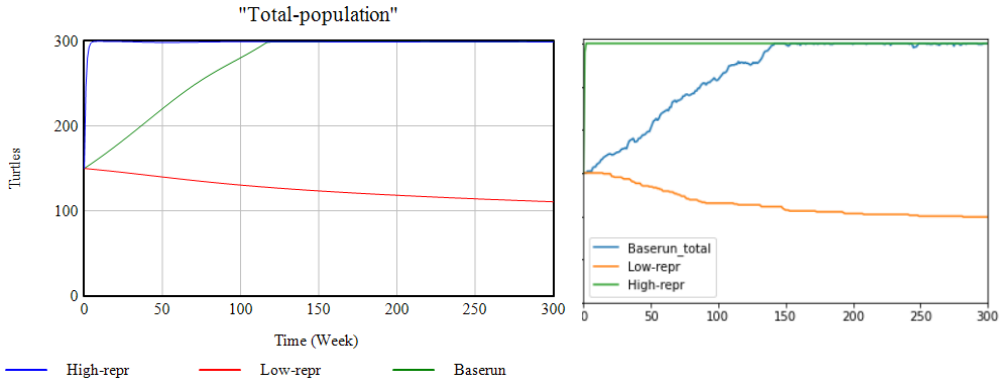

Figure 23: Extreme reproduction-rate

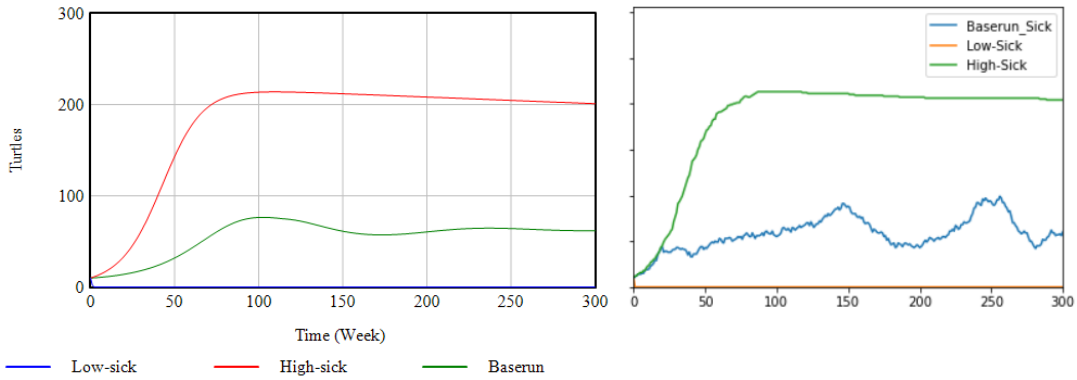

Figure 25: Extreme sickness duration

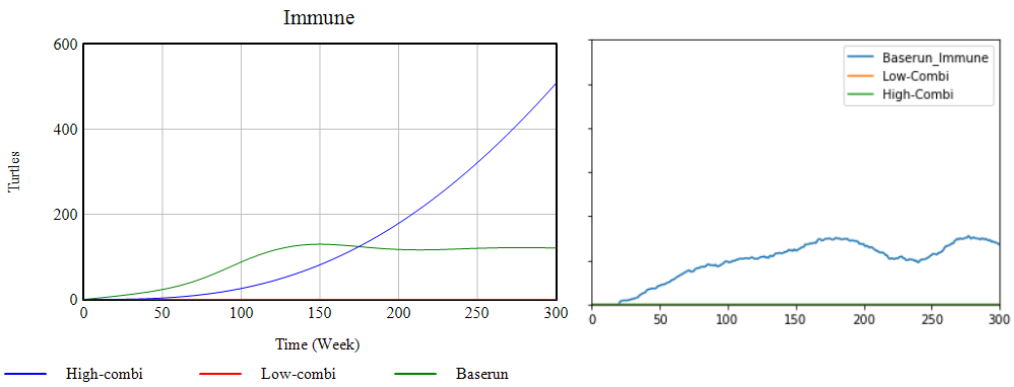

Figure 26: Extreme combination

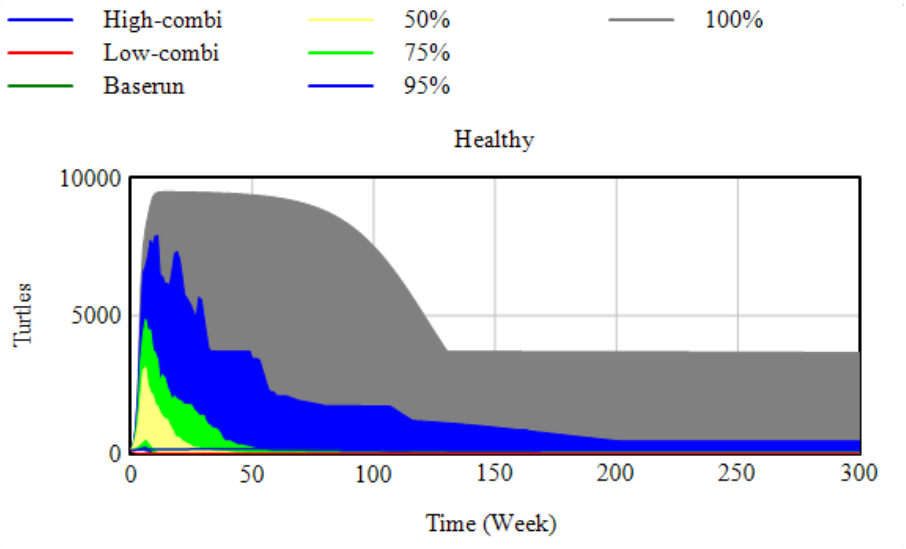

Figure 27: Extreme combination sensitivity

Basic behavioural test Simple model

Once it has been established that the main structure of the Simple models functions well enough for the purpose of this research, the actual behaviour can be analysed. This behaviour should roughly replicate know disease dynamics. According to Bjørnstad et al. (2021) SIR models, and their extensions, should exhibit periodicity when $R_0 > 1$ and there is recruitment into the susceptible population. This behaviour is visualised in figure 28.

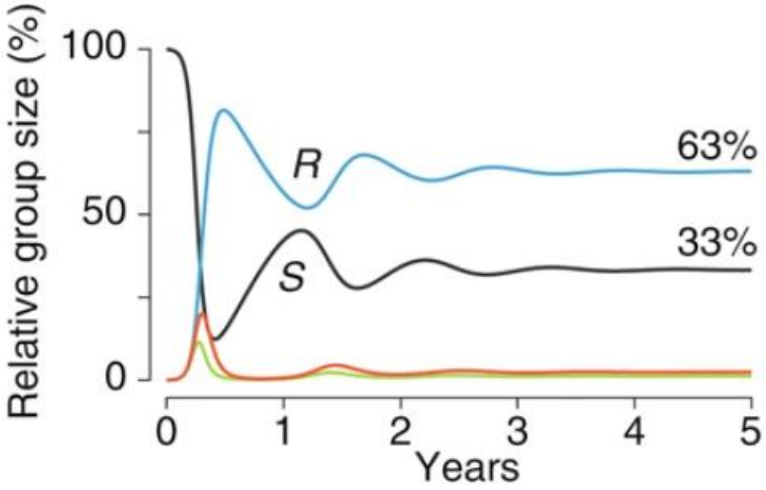

Figure 28: Disease dynamics when R_0>1 Bjørnstad et al. (2021)

The Simple model should roughly be able to replicate this behaviour by altering a few of its variables. By reducing the *carrying-capacity* to the amount of initial agents and increasing the *Lifespan* and *Recovery-rate* to a value that prevents any deaths from occurring, 260100 and 1 respectively, the amount of agents in the model should stay constant. Dependent on the infection and delay rates, between the different disease stages, behaviour akin to figure 28 should be exhibited if disease dynamics are implemented correctly.

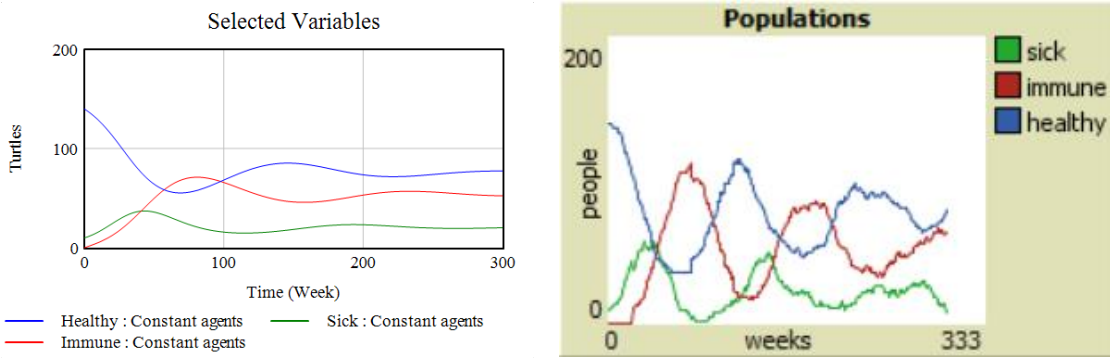

Figure 29: Basic disease behaviour Simple model

In figure 29 the results for both models under these conditions are showcased. Both models exhibit this oscillatory periodicity to some extent. In the SD model this periodicity is quite quickly lost but does showcase the long-term behaviour. In the ABM model the oscillations are significantly bigger, but also slowly decrease in size towards a constant amount.

Based on the behaviours displayed in figure 29, it can be conluded that both models follow the basic disease dynamics present in SIR models, and its extensions. Therefore, the confidence in the model to result in known disease dynamics is increased, and the model's behaviour is deemed good enough for the purpose of this research.

Appendix II; Validation of Extensive model

In this appendix the full behavioural and structural validation of the Extensive model will be performed. As mentioned in paragraph 3.5.2. the validation tests of model structure that will be performed are: the parameter verification test, structure verification test, dimensional-consistency test, and extreme conditions test. For the test of the behaviour of the model a behaviour reproduction test will be performed.

Parameter verification Extensive model

The parameter verification test entails the comparison of the different variables in the model to their real world counter-part. Ideally every variable should have a distinctive meaning and a logical value attached to them. If there are a significant amount of variables without such meaning, this could be an indication of overfitting a model, meaning the model is less likely to be an accurate representation of reality. As the model is supposed to be a representation of a system with the SARS-CoV-2 virus, all variables will also be judged on their similarity to known values associated with this virus. The variables will be discussed based on their use in the model.

Within the variables related to movement there are a few variables that have lacklustre real-life meaning. Starting with both *prop-move-long* and *prop-move-short*, these should indicate what the chance is of a person to move either 3 or 1 patch respectively. This is probably supposed to mimic the movement of people going either far from home or staying close. In reality people will almost always move from their in the beginning of a day and to their home at the end. This implementation of movement will thus result in people eventually meeting a random set of other people whereas people often meet the same people quite regularly (friends, family, colleagues, etc.) and strangers on more rare occasions (going to an event for instance). This makes the model quite unusable, when trying to analyse the impact of mitigation aimed at reducing the movement of people. Due to this stochastic element in the model the uncertainty in regards the chance of people meeting each other is quite considerable, to reduce this a *Fix variable* has been added to the model that has no real world meaning, but only helps to increase consistency between the two models.

The second group of variables is related to SARS-CoV-2 attributes. Almost all of these variables have been based on reports and studies publicised in late 2020. The figures used are therefore viewed as valid but can potentially be outdated. A few parameters do not seem to have any real variables associated, for instance the standard value for *Transmission-parameter* seems to be determined via method of trial-and-error. This variable can be explained as the chance of transmitting the disease when actually coming in contact with someone, however as coming in contact is loosely defined as being on the same patch, it would be unreasonable to expect a real life tested variable value.

All other groups of variables have logical meanings and values attached to them, except for the unit correction factors. These are used to alter the dimension of some variables, to make the model dimensionally consistent. These do not have any real-world counterpart, and their use should be closely examined in the dimensional consistency test. Overall, this test increases confidence that the model is adequate for the purpose of exploring the effects of SARS-CoV-2, except when movement reduction policies are explored.

Structure verification test Extensive model

The structure of the Extensive model will be tested for basic epidemiological transmission structure, realistic structural constructions, and inclusion of complex to translate ABM concepts. These aspects are important as they form the foundation for respectively, disease dynamics, interpretability of the results, and coarse-graining capabilities.

The models include an extended SEIR structure, including asymptomatic, pre-symptomatic and hospital states. The inclusion of these extra states makes the model able to represent the real disease dynamics more accurately, as the impact of these groups is quite significant (see Section 2.1).

Regarding the structural constructs, we will discuss the concepts which can be viewed as unrealistically implemented. The structure related to the availability of hospital and ICU beds can be viewed as a bit unrealistic. The amount of available beds is predefined and constant in the model. When all beds are in use people either get rejected by hospitals or the availability of beds does not play any factor in hospitalisation of people. However, in reality the amount of beds available for SARS-CoV-2 patients will be scaled up when necessary. When the choice is made to have the availability of beds not affect the hospitalisations, this will not be entirely accurate, but can be used to see at what point and under what circumstances this amount gets eclipsed.

A second structural issue is related to the isolation mechanics in the model. There are two main methods of going into isolation: on onset of symptoms of when getting informed by your infector. This means only people who have actually gotten infected by the virus will ever go in isolation. In reality the infector does not know which people they have come in contact with were actually infected by them, making them also warn uninfected susceptible people. These susceptible people would also go in isolation for some time. During the time they are isolated, they cannot become infected by any transmitters, which can thus have a considerable impact on the disease dynamics. A second structural issue related to isolation mechanics is solely present in the SD version of the model. In this model people will not leave isolation after a certain time period, but only when becoming immune, going to the hospital, or dying. This simplification was necessary as the correct flows would otherwise become too difficult to accurately determine. In reality people are able to leave isolation before they have lost their infectiousness, which could lead them to infect more susceptible people. These structural simplifications are not optimal; however, the model should still be able to be used to gain insight in the disease dynamics under these circumstances.

The last structural element that needs to be examined is related to complex structures in regard to coarse-graining. Besides the elements that were already present in the Simple model, like stochasticity, heterogeneity of agents, and spatiality, more complex aspects need to be included to gain confidence in the ability of coarse-graining to work for more complex models. The increased structural complexity in the model can be found in the inclusion of individual tracking and the increased heterogeneity. Especially the increased heterogeneity in conjunction with the spatiality made the coarse-graining significantly more complex.

Based on the analysis of the structural elements in the two models, it can be concluded that the main structure of the model is sufficient for the goals of this research. However, when trying to research concrete dynamics and mitigations measures present in the model, a distorted view can be generated, due to the simplifications in the model. This model should thus primarily be used with methodological goals in mind.

Dimensional-consistency Extensive model

The dimensional-consistency tests will only be performed on the SD version of the model, as it includes a tool for this purpose, and every variable has a specified unit. Due to the coarse-graining, the main structure and functions of the models are exactly the same, so consistency in the SD model also indicates consistency in the ABM version. Dimensions failing to add up implicate a structural mistake in the model and should therefore be analysed carefully. This test is especially useful in conjunction with the parameter verification test, as this eliminates the variables used to fix the dimensions without having any real-life meaning.

The parameter verification test identified a few variables that were solely used to alter the dimensions of other parameters. The use of these variables; *day unit correction* and *people unit correction*, will be discussed to identify whether their use is justified or hides structural issues.

d*ay unit correction*, is utilised four times in the model. Two uses are in conditional checks where several types of variables, are checked against the same condition. To this end, variables that are checked against this condition are changed to be dimensionless with use of this unit correction. A third use is when a look-up function uses time as its input. The final use is to calculate *Mix R0*. The need for a correction factor here seems to indicate a dimensional error, as the formula does not add

up without the correction. This could be caused by a forgotten factor with a value of 1 in the ABM model, or a structural mistake. However, since this is a purely supplementary variable (meaning it is not used any further in the model), the issue can be ignored.

people unit correction, is used twice in the model. The first use is in the same conditional check as mentioned for *day unit correction*. The second use is also related to the *Mix R0* variable. So, it can be concluded that this variable was not used to hide any relevant structural mistakes.

Overall, there are not any other issues with the dimensions present in either model. This does not indicate that all variables are correct, only that these variables are consistent with each other. This consistency increases the confidence in the correctness of the internal structure of the model.

Basic behavioural test Extensive model

Once it has been established that the main structure of the Extensive models functions in lights of the purpose of this research, the actual behaviour can be analysed. This behaviour should roughly replicate know disease dynamics. According to Bjørnstad et al. (2021) SIR models, and their extensions, should exhibit periodicity when $R_0 > 1$ and there is recruitment into the susceptible population. This behaviour is visualised in figure 30.

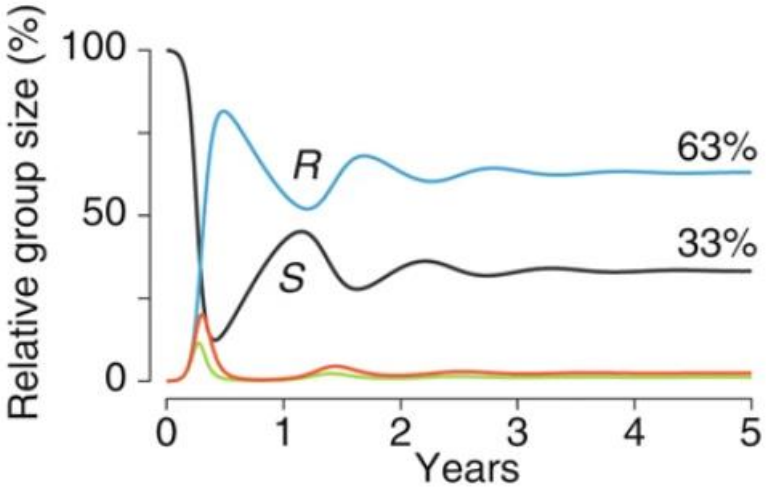

Figure 30: Disease dynamics when R0>1 Bjørnstad et al. (2021)

This behaviour should be replicated when the aforementioned conditions are in place. To ensure the recruitment into susceptible population and prevent the disease from dying out before this recruitment a few variables are altered. The time people are immune is decreased to 14 days on average by setting *Immune-loss-when* to 2. And by putting both *Immune-mild* and *Immune-severe* to 0, everybody can lose their immunity. The resulting behaviour for both models can be seen in figure 31.

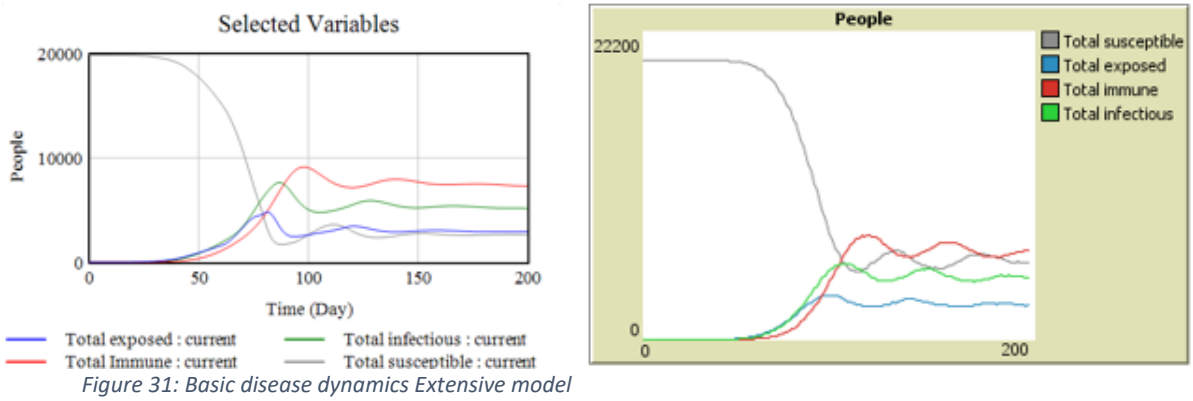

Figure 31: Basic disease dynamics Extensive model

In figure 31 the results for both models under these conditions are showcased. Both models exhibit this oscillatory periodicity, while slowly reaching an equilibrium near the end of the simulation. Both models clearly showcase the expected disease behaviour mentioned by Bjørnstad et al. (2021). Therefore, we can conclude that the main disease functions act appropriately and we can use these models to research the behaviour of SARS-CoV-2 in the population.

Appendix III: Simple SD model

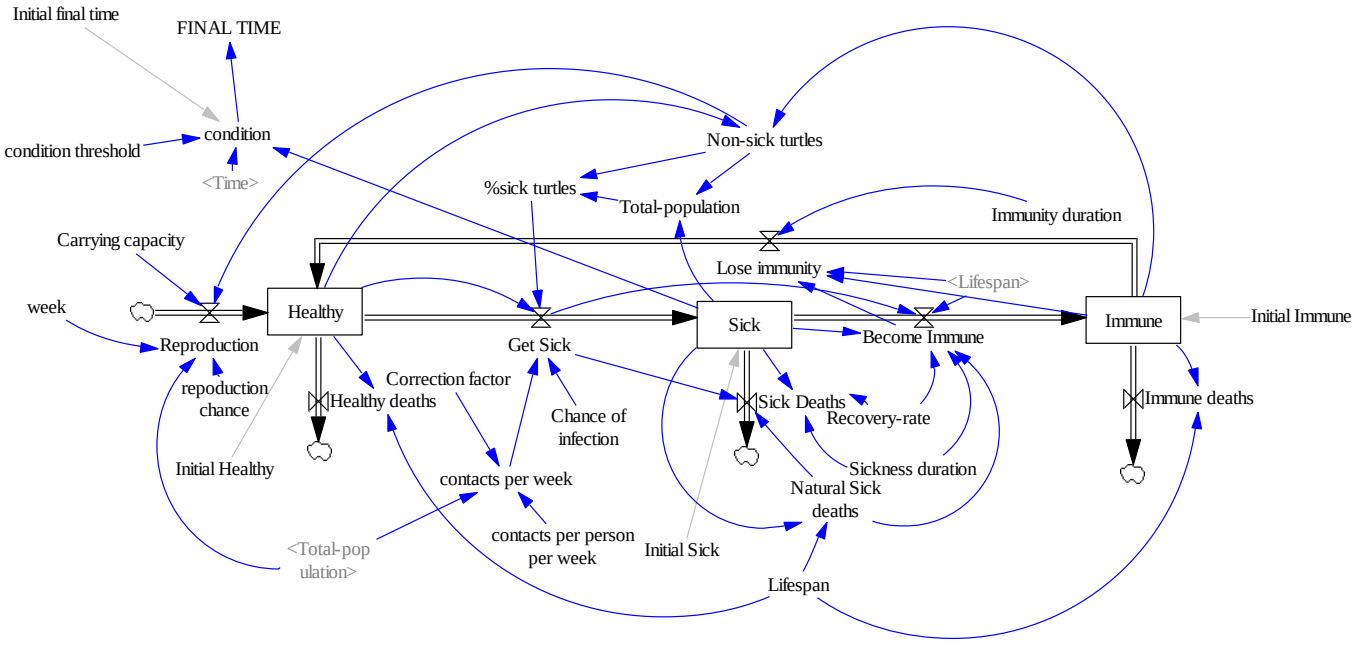

Figure 32: Full Simple SD model

Appendix IV: Explanation for exclusion of movement

It is important to note that four relatively important variables will not be varied during the open exploration of the coarse-grained model. These variables are all related to the movement of the agents. As mentioned in Section 3.5, the movement of the agents and related chance of running into each other has been translated into a look-up variable, which does not scale in an identifiable manner with these movement variables. Since the movement in the ABM model is completely random, alterations to the movement options work differently than they would in reality. In the ABM model the movement reductions have a compounding effect over time, as they will primarily reduce the chance of the virus spreading to faraway places in a short time. Locally these reductions have a very limited impact on reducing infections, and perhaps even detrimental by reducing the chance an agent near an outbreak has to dodge the virus. However, on the really long term the effects of these reductions will start decreasing as the virus will most likely still reach the far places, only at a slower pace. This combination of factors makes it borderline impossible to accurately represent this in a SD model. Not only is it not able to be translated easily, the effects these changes cause are not comparable to the real-life counterparts these changes should have. In real life people will generally always start and end up at the same place, their home. Therefore, there is no compounding effect related to reaching faraway places, as people are just as likely to visit those places at the onset of the reduction, compared to several weeks after the onset. Inclusion of these variables in the open exploration will only result in less consistent results, without gaining any useful insights.