KISSing in the Time of COVID-19: Some Lessons for Model Choice

Till Grüne-Yanoff

KTH Royal Institute of Technology, Stockholm

Abstract

I present and analyze the case of COVID-19 modeling at the *Public Health Agency of Sweden* (FoHM) between February 2020 and May 2021. The analysis casts the case as a decision problem: modelers choose from a strategically prepared menu that model which they have reasons to believe will best serve their current purpose. Specifically, I argue that the model choice at FoHM concerned a trade-off between model-target similarity and model simplicity. Five reasons for choosing to engage in such a trade-off are discussed: lack of information, avoiding overfitting, avoiding fuzzy modularity, maintaining good communication, and facilitating error avoidance and detection. I conclude that the case illustrates that model simplicity is an epistemically important principle.

Keywords: Modelling, Methodology, Similarity, Simplicity, Epistemic virtues.

1. Introduction

The epidemiological modelling toolbox has grown considerably over the last twenty years. The Public Health Agency of Sweden (FoHM) is a good illustration of that: it has systematically developed its menu of mathematical and computational modeling tools. But constructing a menu also forces a choice, and this is the focus of my case study in this paper. When Covid-19 came to Sweden, how did FoHM modelers choose their modeling tools from those menu options?

That is an interesting story in its own right, which I will sketch here, but my underlying interest is to identify the *reasons* for this choice. For this purpose, I first rehash in section 2 the discussion of two opposing desiderata, first of model-target similarity, and second of model simplicity, as exemplified by the KISS principle: "Keep It Simple Stupid". I argue that these are the relevant criteria for my case, by showing that the different options contained in the toolbox indeed differ mainly in how much they simplify and how many parameters they include (section 3.1). I then recount how the COVID-19 models at FoHM were chosen (section 3.2). In section 4, I discuss five epistemic reasons for choosing the simpler kind of model, in effect sacrificing a certain degree of potentially higher model-target similarity for the sake of more model simplicity. Section 5 concludes.

2. Modeling Between Similarity and Simplicity

At the heart of every methodological question about models lies a decision. The modeler pursues a certain goal—for example, predicting a future event based on current data, or explaining a current phenomenon with available theory—and then needs to choose which available model to employ for this purpose. This decision depends on other, previous choices—whether to model at all, and what models to make available for oneself—but in this paper I want to focus on the choice between available models.

Models can be assessed according to many different criteria, and some of these might trade off on each other (Levins 1966, Matthewson and Weisberg 2009, Elliott and McKaughan 2014). Even if there is no general trade-off, such desiderata can come into conflict under certain conditions. This is the case with similarity and simplicity, the two model virtues I focus on in this paper.

The similarity desideratum derives from the idea that models represent specific targets. Targets might either be actual or non-actual things; although scientists are interested in them, they represent them with models and then investigate the models in their stead. This might be because the target is not accessible or cannot be manipulated, be this for physical, economic or legal reasons. For example, cosmologists model black holes because they currently cannot be accessed or manipulated; pharmacologists experiment with rat models, because experimenting on humans is very restricted; economists use macroeconomic models because experimenting with the interest rate could have grave economic consequences.

This rationale for modelling implies two important consequences. First, if models are supposed to function as stand-ins or surrogates of targets for one of the above reasons, then it is important that they are similar to these targets. The less similar a model is to its target, so it seems from this perspective, the more difficult it is to justify its use in its target's stead. Similarity seems to emerge as a prime model desideratum from these considerations, and many authors have indeed defended such a view, with qualifications regarding background theory and modeler's purpose (Giere 1988, Weisberg 2013).[1]

Second, however, if models are employed because of one of the obstacles that the target poses for a direct investigation, then any successful model must differ from the target at least with respect to that obstacle. To be useful, a black hole model must be accessible and manipulable; it must be legal to experiment on the modal organism; and the manipulation of the Macro-model must not put millions into the poorhouse. Consequently, models must be to some degree different from their targets; demanding full similarity or identity between model and target would defeat the very rationale of using models in the first place.

Once one admits that much however, the question arises *how much* similarity to demand between model and target. Here views differ considerably. Some argue for similarity to a large extent: "Fruitful models […] share many, and do not fail to share too many, features [with their targets] that are thought to be salient by the scientific community" (Weisberg 2013: 155). Others, however, have argued for sacrificing some degree of similarity for the sake of simplicity (Paola and

[1] Similarity was once thought to be a vacuous concept (Goodman 1972). The more recent literature offers a number of operationalizations (e.g. Weisberg 2013), even if these remain controversial (Parker 2015).

Leeder 2011). These two desiderata trade off on each other if the target has a high degree of complexity: stressing similarity would then make the model complex, while stressing simplicity would keep its complexity comparatively lower, at the cost of reduced similarity.[2]

Model builders might take inspiration from engineers, who widely accept simplicity as a design principle, for at least two reasons. First, the more complex the device, the more difficult it is to control; keeping design simple thus allows, *ceteris paribus*, better control. Second, design defects are better remedied by basic redesign than by superficial modification, as this avoids conservative *ad hoc* maneuvers. Keeping devices simple forces such early redesign in the face of defects. In design, these considerations widely became known as the KISS principle (Rich 1995).

Mutatis mutandis, the KISS design principle has also been applied to modelling choice. In the computer simulation community, however, criticism of KISS grew with increasing computational power available to modelers. Why, these authors asked, should one sacrifice any degree of similarity if increased computational capacities allowed the construction and analysis of models of hitherto unreachable detail-richness? Some authors even formulated a counter principle to KISS, which they termed KIDS: "Keep It Descriptive Stupid", thus explicitly endorsing a maximization of similarity:

> The KIDS approach starts with a model which relates as strongly to the target domain as possible, but does not ensure that the models are "elegant". Before the advent of cheap computational power, it was only possible to get any results out of analytic (and hence relatively simple) models; this made the KIDS approach infeasible (Edmonds and Moss 2004: 142).

This perspective on the similarity-simplicity trade-off will be central for this paper. Why, at a time of newly-won and still increasing technical feasibility, would modelers want to give up any potential similarity between model and target for the sake of keeping the model simple? With this question I turn to the case of COVID-19 modeling at the Public Health Agency of Sweden in 2020.

3. The Case: COVID-19 Modeling in Sweden

The Public Health Agency of Sweden (*Folkhälsomyndigheten*, FoHM) is a Swedish government agency with national responsibility for public health. It was formed in 2014 by a merger of the Swedish National Institute of Public Health (*Folkhälsoinstitutet*) and the Swedish Institute for Communicable Disease Control (*Smittskyddsinstitutet*, SMI). It has about 600 employees in six departments. Its task is to produce and disseminate scientifically sound knowledge that promotes health and prevents diseases and injuries. Its target groups are the national government, other state authorities, regions, and various interest groups (FoHM 2021c).

Epidemiological modeling at FoHM is performed at the Analysis Unit, which is part of the Department of Public Health Analysis and Data

---

[2] There are multiple notions of simplicity relevant for model choice. Rochefort-Maranda (2016), for example, distinguishes between parametric, theoretical, computational, epistemic, and dimensional simplicity. I will focus on parametric simplicity in this article.

Management. Since its inception, the unit has been headed by Dr. Lisa Brouwers, with a staff of 20, out of which 4-5 work with epidemiological modelling. Lisa received her PhD in Computer and Systems Sciences from Stockholm University in 2005. The title of her thesis was *Microsimulation Models for Disaster Policy Making*. Since 2004, she had been associated with the SMI, one of the predecessors of FoHM.

According to Brouwers, mathematical modelling—beyond statistical regression analysis—had not been practiced at FoHM until the early 2000s. This reflects the then-state of epidemiology more broadly: many epidemiologists in the early years of the millennium did not consider analytical models as part of their toolkit (Chubb and Jacobsen 2010; for a survey amongst epidemiologists about mathematical modelling, see Hejblum et al. 2011). Brouwers was hired at SMI into a project that aimed changing that.[3]

> I have been lucky to have managers who were interested and saw the relevance of modeling, so I have had the opportunity to over the years strengthen and form the modeling capacity within SMI, and then later on at FoHM (Brouwers interview 2021: 2).

Brouwers was tasked with implementing a long-term modelling strategy at FoHM: "it has been my responsibility to make sure we have had competence in modeling" (Brouwers interview 2021: 2). In particular, this involved constructing and maintaining a number of distinct model frameworks for epidemic modelling—a modeling toolbox—and developing staff competences in maintaining and applying them:

> What we had in mind was to have different types of models available or quite ready to deploy when we need them. The *MicroSim* model was one option, the *SEIR* models or variations of them, and then more statistical models. But maybe most of all we made sure that we had staff—competence—who can program such models, as well as have the ability to decide which models to start working with and when (Brouwers interview 2021: 2).

FoHM thus pursued a strategy of systematically developing and cultivating a menu of modelling tools, from which model choice for different epidemiological purposes could be made. Before describing the choice itself, it is worthwhile detailing what this menu actually consists of.

3.1 The Modelling Toolbox at FoHM

Compartmental models are some of the most commonly used models in infectious disease epidemiology, and various versions of these are also part of FoHM's toolbox. Its most popular version, the *SIR* model, consists of three compartments: *S* (susceptible) for individuals at risk of infection, *I* (infectious) for individuals currently infected, and *R* (recovered) for individuals who recovered from the infection and have immunity. Every individual in a population is assigned to one

---

[3] The project was headed by Johan Giesecke, state epidemiologist of Sweden from 1995 to 2005. During the pandemic, FoHM contracted again Giesecke, by then professor emeritus, to "support the Unit of analysis in their modelling and analysis of COVID-19, at a maximum of 800 h in 2020" (Karlsten 2020, my translation).

of the three compartments. Within each compartment, individuals are assumed to have the same properties and act in the same way. Individuals may progress between compartments according to predefined flow patterns. In the *SIR* model, for example, individuals progress from *S* to *I* to R. To model a specific epidemic in a particular population, one quantifies the proportion of the population located in each compartment at a specific time and assigns values to the rates of flow between compartments. The *SIR* model is commonly run with ordinary differential equations, which are deterministic. Alternatively, the parameters specifying the flow rates can be expressed as probability distributions to better capture the uncertainty of the estimates.

The *SIR* model can be varied by changing the number of compartments— either by expanding it (e.g. adding an 'exposed' compartment in a *SEIR* models, or connecting *R* back to *S* in a *SIRS* model, where immunity lasts only for a short period of time) or by contracting it (e.g. to a simple logistic *SI* model, or a *SIS* model where there is no immunity). Furthermore, compartmental models have occasionally included seasonally dependent flow rates, diffusion constants to model spatial distribution of the infected, vital statistics like births and deaths, age distributions and vaccination status. What all these variants share is the assumption that the transition rate between *S* and *I* is determined by the average number of contacts per person per time unit, in conjunction with the probability of disease transmission in a contact between a susceptible and an infectious subject. Individuals are assumed to mix homogeneously: their contact rates are assumed to be independent of their individual identities. Typical uses of such models include the prediction of disease spread, total number infected, epidemic duration or the infection's peak.

FoHM developed such models. Specifically, at the beginning of COVID-19, it used a deterministic *SEIR* model with compartments *E* for exposed and a distinction between I_r for reported and I_o for unreported infections. The flow pattern of this model was: $S \to E \to \{I_r, I_o\} \to R$ (FoHM 2020b). This model was later expanded to include two R-compartments: R_1 in which an individual still can test positive on a PCR test, and a second compartment R_2, in which an individual no longer tests positive on a PCR test (FoHM 2020d). In a third analysis, called *VirSim*, the initial *SEIR* model was modified to contain three separate age cohorts (0-19, 20-69, 70+) (FoHM 2020c). This model had already been developed at SMI ten years earlier, under participation of Brouwers (Fasth et al 2010). Until June 2021, all COVID-19 analyses published by FoHM relied on some variant of these compartment models.

However, these models were not the only ones in FoHM's toolbox. To the contrary, Lisa Brouwers' early research work concentrated on another kind of model, in which individuals and their contacts are represented explicitly and heterogeneously. Such models are often called agent-based models or microsimulations. Brouwers developed her first microsimulation model, *MicroPox*, as part of her PhD (Brouwers 2005). The model is a microsimulation model, representing all 8,861,393 Swedes (the size of the Swedish population when the data set was collected). A unique feature of the model is that it uses government census data on where each person works, who the person works with, and who the person lives with. This makes it possible to extract a network of contacts that shows the professional and family contacts. These contacts are depicted deterministically. A day in the simulation model is divided into day and night. In the first hour of the day, people with a job go to work. If people are unemployed or retired, they stay

at home. Since school and kindergarten data were not available, the model uses a proxy based on age and physical distance. Everyone returns home after work and sleeps there with their families. The model gives transmission probabilities for each of these locations (Brouwers 2005, Brouwers and Liljeros 2005).

Brouwers illustrated the model use at the hand of a number of simulated smallpox epidemics. This was motivated by the knowledge that smallpox spreads mainly through close contacts, and that therefore the contact network is of greater importance than it would be for a highly contagious disease like measles. The purpose of these modelling efforts was to create a tool for testing the effects of intervention policies, including mass vaccination, targeted vaccination, isolation and social distancing. Initially focused on smallpox, the model, renamed *MicroSim*, was modified to support simulations of pandemic influenza in 2006 (Brouwers et al. 2009a). Specifically, *MicroSim* was used to estimate the economic consequences of reduced absenteeism through a sufficiently strong H1N1 influence vaccination campaign (Brouwers et al. 2009b).

Models like *MicroPox* or *MicroSim* are costly to maintain, both in terms of the regularly needed census data updates, as well as in terms of the staff competences and worktime required for their maintenance.[4] Despite these costs, FoHM kept these models in their toolbox:

> They are kind of, or at least they were before COVID started, maintained up to date. I have had persons working with what we call *MicroSim* [...] *MicroSim* is quite maintained, documented, and possible to run. In other words, yes, it would be possibly without too much trouble to get it started again (Brouwers interview 2021: 1).

Not all agent-based models are as complex, detail-rich and data-intensive as *MicroPox* or *MicroSim*, however. Other agent-based models might still represent individuals and their contacts explicitly and heterogeneously, but focus on a smaller population or avoid reliance on census data altogether.

As an example, consider Burke et al. (2006), who simulated a single initial infected person attack on a town network of either 6,000 or 50,000 people. Town networks either consist of one town, a ring of six towns, or a 'hub' with four 'spokes.' Each town consists of households of up to seven persons, one workplace, and one school. All towns share a hospital. Each space is represented as a grid, so that each cell in the grid has eight neighbors. Agents are distinguished by type (child, health care worker, commuter) by family ID and by infectious status. Each 'day,' agents visit spaces according to their type, and then return home. On the first day of the simulation, the position in schools and workplaces is randomly assigned, but after that, agents remember their positions. During the day, agents interact with all of their immediate neighbors: 10 times at home, 7 times at work, and 15 times in the hospital. After each interaction, they move positions to the first free cell in their neighborhood. Homogeneous mixing is thus completely eschewed; instead, agents interact in a number of dynamic neighborhoods. These models represent some recognizable "town-properties" without representing any

---

[4] "I think that to maintain a microsimulation or agent-based model you either need to have a large group of modelers so you can have one or two modelers working with it part-time every year, or you need a specific interest to drive the project yourself as head of the modeling group, or maybe some other collaborations with a university" (Brouwers interview 2021: 3).

actual town or drawing on any data from such actual towns (For a more detailed discussion, see Grüne-Yanoff 2021).

These intermediate agent-based models are also part of FoHM's toolbox.

> We use these kinds of models as well. […] it is more of a tool to study network effects of different phenomena. […] When we implement our *SEIR* models in different age groups they become so complex that it is easier to implement the models as agents, i.e. each agent or each version of the *SEIR* structure is an implementation of an agent. So, we are kind of using it but not conceptually as an agent-based model. Nevertheless, we are considering the whole spectra of different kinds of models (Brouwers interview 2021: 3).

To summarize, FoHM for almost 20 years has developed a modelling toolbox, consisting of "the whole spectra of models" to suit various modelling purposes. This toolbox included both compartmental models, census-based agent-based models and more abstract agent-based models. Given this menu of available tools, it is interesting to see how FoHM actually chose its models for investigating COVID-19, when the epidemic came to Sweden.

3.2 Choosing the Covid-19 Model February 2020-May 2021

In February 2020, Brouwers recalls, her boss first suggested that they start looking at their models and get prepared to assist the Swedish regions in making predictions in terms of hospital and care needs. "Out of curiosity", Brouwers and her team also tried to see if they could fit *SEIR* models to the very early data available from Wuhan. However,

> quite early we realized that we could not do it because the *SEIR* models overshoot by predicting huge outbreaks. We of course understood that this happened because there must be an unreported fraction of infected persons [whose] size we did not know. Also, we thought that it is probably not a totally homogenous spread. In other words, there was still so much we did not know about how the virus was spread—how infectious it was. Simply taking the data from Wuhan and implement it in models for Sweden would render enormous outbreaks that were not realistic, because otherwise we would have seen the local outbreaks spread faster, probably, from Wuhan to the rest of China at that time (Brouwers interview 2021: 4).

This failure at replicating the Wuhan data with compartmental models led to two decisions. First, in order to estimate the burden for hospitals in the regions, FoHM would initially not try to replicate the data or model the dynamic, even in an *SEIR* model, for a prediction of the outbreak in Sweden. Instead, they used prototypical epi-curves where they beforehand decided how many would be infected, in order to sketch a realistic worst-case scenario for the regions, without modelling any transmission. This led to the first report on March 20th (FoHM 2020a).

> We thought that this could be used to help the regions in Sweden by providing an answer [to] the following question: If it in each region of Sweden would become as bad as it was in Wuhan, then what would the need of hospital beds be? Divided into ICU-beds and ordinary hospital beds. So we constructed outbreaks with a clinical attack rate of the same magnitude as in Wuhan, 1%, for each region in Sweden, using a simple *SIR*-model (Brouwers interview 2021: 4).

The second decision was that the explicit modelling of transmission with any model should wait until relevant data was available about the pandemic in Sweden—in particular on the fraction of unreported cases. In the so-called *Gloria*-studies, FoHM in collaboration with the Swedish army tested population samples in different regions for COVID-19 and from this concluded that 98.7% of all infections go unreported (FoHM 2020e).

> At that time, we had that piece of puzzle we missed previously, so with this modelling report we switched from just doing some prototypical to try to make a realistic representation of the dynamics in the region where we had that information we were previously lacking (Brouwers interview 2021: 7).

With that information they built the first *SEIR* model with separate compartments I_r for reported and I_o for unreported infections. Because the infection rate of unreported cases might differ from the reported ones, they modelled three scenarios with different infection rates—one identical to, another at 55%, and one at 11% of those in I_r. (FoHM 2020b). The purpose of this model was to estimate when the infection would peak in Stockholm, and how many infected were to be expected until the end of April.

When the results of this study were presented at a press conference on April 21st, deputy state epidemiologist Anders Wallensten explained the model's assumption with an illustration, saying that "there is about one confirmed case of COVID-19 out of 1000 cases in total". The figure caused confusion. As a journalist pointed out, at that time about 6,000 infections were confirmed in Stockholm alone—would this imply that almost six million were actually infected (Stockholm region has less than 2,5 million inhabitants). Brouwers and her team stated that the figure was incorrect—a fact that the press reported probably more than any other results from FoHM's analysis unit.

Nevertheless, Brouwers two days later could also see a positive side of this hiccup: "It's almost lucky that the mistake was so obvious, it could have been much more subtle" and hard to find, she said in an interview with *Aftonbladet*.[5] Before the publicly accessible code was taken down, just after the 21/4 press conference, five members of the public had already written to the programmer about the same error. The model was reprogrammed, and the report adjusted accordingly.

FoHM has performed a large number of model studies of Covid-19 since, including projections of rising infection rates due to increased summer travel (FoHM 2020c), estimates of infections rates in some of the Swedish regions until early 2021 (FoHM 2020d) and scenarios for future developments of the pandemic into summer 2021 (FoHM 2020f, FoHM 2021a and 2021b). All of these reports are based on compartmental models, the later ones specifically on FoHM's *VirSim*. Not a single study was performed with agent-based tools, specifically not *MicroSim*.

The main reason for that was a lack of relevant data, especially about the basic epidemiology and the medical features of the disease. Comparing it to the influenza simulation they did with *MicroSim* before, Brouwers argued that

---

[5] "Det var nästan tur att felet vår så uppenbart, det kunde ha varit klart mer subtilt" (Karlsson 2020).

> When we model the flu, we have quite a lot of data, but we still have to make a lot of assumptions regarding how many days after infection individuals' peak their infectiousness, how do people act when they are asymptomatically infected, etc. But for COVID we had no such information at all, and therefore we decided that moving to a *MicroSim* model with so much uncertainty is not worth it. It would not have been wise. Further, no one pushed for such a *MicroSim* model either (Brouwers interview 2021: 6).

This does not exclude that FoHM will soon begin making use of *MicroSim*, however:

> After the summer [2021] it could be the case that we start doing more network and maybe agent-based modeling to look at the spread within certain groups, or between certain groups, and the rest of the population where you have a quite heterogeneous vaccination coverage in the population—pockets with the risk of infection put among young people and also some groups that have a lower vaccination coverage. We are discussing this, but we have not started this modeling yet (Brouwers interview 2021: 5).

There are legitimate purposes for which agent-based simulations might be used in modelling COVID-19; yet these purposes have not been the immediate pressing ones during the early and current stages of the epidemic—at least not under the given circumstances. How such considerations of purpose and circumstance constitute reasons for choosing certain options from the modelling toolbox will be my focus in the next section.

4. KISS Despite Technical Feasibility: Reasons for Model Choice

Modelers at FoHM systematically build a toolbox comprising the whole spectrum of epidemiological model types. They did so because they believed that each of these model types had their own advantages that made them a best choice for certain purposes, under certain conditions. The long-term strategy was to provide the technical means to quickly apply the best model to a host of possible eventualities. When faced with the COVID-19 outbreak in Sweden, modelers at FoHM chose compartmental models over agent-based models, even though the latter were technically feasible and available. In this section I investigate the reasons for this choice—both those explicitly considered by FoHM staff, as well as implicit ones that could justify such trade-off decisions.

This choice is philosophically interesting, because it exemplifies the trade-off between simplicity and similarity sketched in section 2. The compartmental models weren't simply better than the agent-based models; to the contrary, the FoHM modelers explicitly acknowledge that the compartmental models are probably too simplifying to get a sufficiently accurate representation of the COVID-19 dynamic:

> Covid-19 is primarily transmitted through droplet infection, which indicates that the social contact structure in the population is important for the dynamics of infection. The compartmental model does not take into account variation in contacts between people, which occur in a society where few individuals could have many contacts and the majority have fewer contacts. This simplification in the model, i.e. a homogenous contact structure, usually results in a somewhat faster growth

of an epidemic than if heterogeneity is included in the model. The model, therefore, runs the risk of overestimating the speed of the outbreak in the Stockholm region. This is not included in the specified confidence intervals, as a confidence interval cannot report such uncertainties (FoHM 2020b: 23).

Acknowledging that a model is overly simplifying to fully satisfy a certain purpose and yet choosing it over less simplifying alternatives indicates that the chosen model has other advantages that are more important for one's purpose and under prevailing conditions than the sacrifice in similarity. In the following, I will discuss five epistemic reasons that are all connected to simplicity considerations—thus arguing that FoHM's model choice indeed was motivated by a similarity-simplicity trade-off.

4.1 Lack of Reliable Data

The first reason for choosing compartmental models over agent-based ones is the lack of information needed for specifying some of the agent-based parameters. The latter contain a much larger number of parameters than the former for at least two reasons. First, agent-based models like *MicroPox* or *MicroSim* contain a lot of individual and institutional structures—e.g. demographic data and potential meeting places like dwellings, hospitals, offices, public transport—left implicit in compartmental models (Brouwers 2005, table 1). Second, these structures can take heterogeneous values for e.g. transmission probabilities at different places, individual probabilities of visiting the emergency when feeling ill or the propensity to travel (ibid.). This provides agent-based models with a much higher *potential* to represent social contact networks and individual heterogeneities. But this modelling potential also imposes high demands on measurement and data provision. Only if sufficiently reliable information is available can the model potential be actualized into a useful model:

> The [agent-based models'] usefulness would be huge if you had more knowledge about individual differences, like susceptibility, immunity, etc. But we don't have that. Hence no reason to use those models! (Brouwers IFFS talk 2020)

> The more information we have, the more possibilities we have, but we are still not at the point where we see that it is useful to switch to agent-based models because there is still so much that we do not know. (Brouwers interview 2021: 5).

At the beginning of the pandemic, the lack of data about the specifics of COVID-19 and its viral SARS-CoV-2 agent was particularly acute and led FoHM to choose models that did not explicitly require this information, instead replacing them with simple random-mixing assumptions. However, it is quite common for agent-based models to suffer from such data deficits, even outside of emergency situations. Take for example Brouwers' earlier *MicroPox* model:

> The data set contains […] no information about which school a child is enrolled in. Therefore, we must generate a proxy for this connection in the *MicroPox* model [We use] proxies for relations that are important to include in the social network, but for which we have no real data (Brouwers 2005: 73).

Instead of specifying these free parameters with empirically well-founded information, they are filled through plausibility considerations that *might* be correct, although the modeler has little reason that they actually are (for further discussion of these proxies, see Grüne-Yanoff 2021).

For obvious reasons, it is more difficult to obtain reliable information of infection rates at different locations, instead of obtaining an average over the whole population, as used e.g. in the *SEIR* model. To explicitly include additional parameters that cannot be specified based on reliable evidence then constitutes a source of uncertainty and error that the FoHM modelers sought to avoid:

what would be the use of using an agent-based model if you don't have that information? [One might] overestimate the risk that [agents] are part of large families or they work in certain places, and [...] what you get out from the model would be based on that misassumption. So I would say: [...] it's the risk of introducing more errors when using individual-based models, when you don't have additional information (Brouwers interview 2021: 10-11).

Model choice thus was motivated by weighing potential errors from different modeling strategies. Clearly, the FoHM modelers saw the simplifications of the compartmental models as a threat to relevant similarity between model and target, and thus as potential errors (as expressed in the quotation at the end of section 4). But the richer and more flexible structure of the agent-based models only offered the *potential* of building more similar models; this potential similarity could however only be realized with sufficiently good data (Grüne-Yanoff 2021). As such data—e.g. about individual differences in susceptibility and immunity or individuum-based social contact networks—was not available, the demands of high parameter specificity posed its own danger of generating errors. The modelers apparently consider the latter the graver threat, and thus chose simpler models over potentially more similar ones.

4.2 Avoiding Overfitting

The above problem concerns the unavailability of data for an independent determination of certain model parameters—e.g. the determination of the proportion of recorded and non-recorded cases based on the *Gloria* studies (section 3.2). However, many model parameters are not determined that way, but by estimation. For example, the early *SEIR* model estimated infectivity rate parameters θ, δ, ε and b_t from recorded infections and the measured proportion of recorded and unrecorded cases, using least-squares regression analysis (FoHM 2020b: 12-14).

For such model estimates, a too large number of parameters poses an additional source of error. It starts from the fact that data is to some degree always contaminated by random measurement error. Random error can in principle be reduced through increased sample size, but sample sizes are often limited and the size of random error is not known for many data-generating processes. A concern for the modeler therefore is to not *overfit* the model to the data set—i.e. to not fit the parameters in such a way that the model begins to describe the random error in the data rather than the relationships between variables. Such a result is more likely the higher the number of free parameters is in the model (Zucchini 2000). Overfitting has at least two negative consequences. First, an overfitted model describes the relationships between variables less accurately than an ideally fitted

one. Second, overfitting fits a model to a specific set of data, including all its idiosyncrasies, thus deteriorating its abilities to predict future data. In both cases, increasing the number of parameters, and thus decreasing simplicity, makes the model sensitive to additional error (Grüne-Yanoff 2021).

Overfitting also illustrates that the trade-off in model choice is between simplicity and *potential* for similarity between model and target. A model that contains more parameters in principle can of course be fitted better to a (complex) target than a model with less parameters. But the practices of model estimation and calibration put a limit to how far this ideal can be reached. Not only is the data sample limited, it is also contaminated with noise. Therefore, the potential of a model with many parameters can rarely be fully realized. The modeler choosing a simpler model thus does not trade off genuine similarity, but rather only the (often unrealizable) promise of potential similarity (Zucchini 2000: 45).

Overfitting might not have been a concern for FoHM at the beginning of the pandemic. I could not find evidence that they specifically worried about random error in the data, nor whether they explicitly compared the number of free parameters in the *SEIR* and the agent-based models (it isn't even obvious how many of the latter's parameters would have to estimated). Yet the quality of the data, and the worry about parameter uncertainty played an important role, as I showed in the previous section; and these are considerations that also raise overfitting worries.

4.3 Easier Communicability

FoHM has well-specified client groups: parliament, the national government, other state authorities, regions, district councils and municipalities, district administrative councils and various interest groups (FoHM 2021c). FoHM not only provides these clients with facts, but also provides them with information about methods, so that clients trust the results and can explain to the public how they came about.

> [A]s soon as we publish results for one region, or a forecast or scenarios for a specific region, they will get questions from journalists, and they need to be able to answer those questions, and they need to understand what the model is showing and how it came about. So we have had lot of meetings [where we are] pedagogically going through how the modelling is performed, what data is used? What is a *SEIR* model? How has it been calibrated to real-world data? And what are the different scenarios? [...] they would be confident that what is in the report, the modeling, is something they kind of trust themselves. [...] And they could say [...] yeah, we have been part of this process (Brouwers interview 2021: 9).

To be able to achieve understanding and epistemic trust in their clients is a good reason to keep models simple. Simplicity here again concerns in the first place the number of parameters included in the model. The more parameters, the more computational steps are required to obtain a model result. With sufficiently many computational steps, a model can only be solved by a machine; and, in the more extreme cases, humans cannot even grasp how the machine arrives at a solution. Agent-based models, specifically those based on large data sets like *MicroSim*, are very much located at this extreme end. Such strong forms of *epistemic opacity* (Humphreys 2004) prevent the kind of shared understanding and epistemic trust that is part of FoHM's mission.

One might reply that ease of communication to clients, while important, does not concern the creation of knowledge and therefore does not constitute an epistemic value. The trade-off sketched in section 2, however, is about competing epistemic virtues of models; hence the above considerations should not count as relevant for that debate. Fair enough.

But ease of communication also concerns interaction within the knowledge-building team, and here it *does* constitute an epistemic value:

> being able to communicate within the group, but also with the managements at the FoHM, so we have a constant dialogue and they understand what we are doing in the model. Even Johan [Carlsson, the Managing director] understands, Anders [Tegnell, the state epidemiologist] understands what we are doing in the model: what are the drawbacks, what are the positive aspects of this model etcetera? […] So within the modeling group we have epidemiologists working tightly together with the modelers, like Jerker [Jonsson] for instance, who is an infectious disease doctor, who can directly advise on how we should implement the risk for hospitalization etcetera. What does it mean? How should we interpret what this data from Wuhan—what does it say? How can we think about this in terms of Sweden? […] So have a multidisciplinary working [group] together with the modelers, not only in the last stage but early, from the beginning (Brouwers interview 2021: 12-3).

The basic argument here is that the multidisciplinary team *as a whole* produces the relevant knowledge—neither modelers, computer scientists, epidemiologists or infectious disease doctors alone can produce it. This requires that each team member understands what the others are working with, and this crucially includes a basic understanding of the model. The more parameters and computational steps the model has, however, the less likely non-modelers will achieve this understanding, even if within-team communication is optimal. Thus, in a multidisciplinary team, there are good epistemic reasons to sacrifice some degree of (potential) similarity for the sake of understanding-facilitating simplicity.

4.4 Avoiding Fuzzy Modularity

There is a more specific kind of opacity, beyond that of generic parameter count, which arises in massive agent-based models like *MicroSim*. The more complex a model is, the more sub-components it has. In addition, when simulating a complex model, these model components run together and in parallel. Not each one, however, contributes to the model result independently. Rather, during a simulation, the components often exchange the results of intermediary calculations among one another—so that the contribution of each component to the model result is in turn affected by all the components that have interacted with it. Due to this interactivity, such agent-based models cannot be divided into separately manageable parts. Instead, these models represent a form of "fuzzy modularity" that makes understanding difficult (Lenhard and Winsberg 2010).

First of all, this is a problem for the explanatory power of agent-based models. Even if such a model could generate the explanandum quite accurately, it would be difficult to determine which of the modeled mechanisms contributed to the generated result. If understanding consists in identifying the mechanisms that created the explanandum, the fuzzy modularity of a model undermines improvements in understanding.

Now, policy makers perhaps do not have to worry about understanding. So why would fuzzy modularity be a problem for them? Due to it, users of agent-based models do not know how individual mechanisms contribute to the generation of a relevant effect. Knowing how individual mechanisms contribute is, however, of great importance for both (i) design and (ii) justification of interventions. First, without knowing how individual mechanisms contribute, the designer does not know where to intervene, because an intervention in a contributing cause can have several effects—through several mechanisms—that can reinforce or interrupt each other (Grüne-Yanoff 2021). Furthermore, they do not know whether the relationship between intervention and effect can be transferred to other contexts where some of the parallel mechanisms may work differently. This does not apply yet to FoHM's modeling, as they so far have modelled only few interventions.[6] But in later stages of modeling the epidemic and various interventions in it, where Brouwers saw some potential for employing *MicroSim* (section 3.2), fuzzy modularity might become an important argument against its use and for a continued sacrifice of (potential) similarity for the sake of simplicity.

4.5 Easier Error Detection

A final reason for trading off similarity for simplicity is that simplicity facilitates error avoidance and detection. All models used at FoHM today are computer-based; implementing models thus means programming them, and programming inevitably brings with it programming errors. Even though there might not be a correlation between complexity of code and number of bugs, there is a correlation with volume (e.g. "lines of code") and bugs (Fenton and Ohlsson 2000). Because simpler models have less code than more complex ones, on average simpler models contain less errors. Keeping computer-based models simple is thus a strategy for programming error avoidance.

Furthermore, once errors are in the code, it is easier to detect them in simpler models. The programming error detected at the April 21st FoHM press conference (see section 3.2) is a good example. By the time the press conference was over, five members of the public had already contacted Brouwers' collaborator to point out the same mistake. This required that the code was sufficiently simple so that educated laypeople could understand it and parse possible bugs. It would be difficult to imagine that something like this could happen with a massive agent-based model like *MicroSim*.

Some might reply that ease of error detection is not a primarily epistemic value, but rather only of pragmatic relevance. I disagree. Modelling tools' susceptibility to error, like that of any tool through which we hope to acquire knowledge, is of direct epistemic interest. Philosophers of science have accepted as much when they discuss strategies to avoid measurement error or ways of controlling background factors in experiments. They should treat strategies for avoiding programming error in the same vein. If indeed keeping one's model simple is an effective strategy in this regard, then modelers might well have a good epistemic reason for trading off some potential similarity for model simplicity.

---

[6] Exceptions are the introduction of vaccine compartments in some of the more recent scenario studies (FoHM 2020f, FoHM 2021a and 2021b) and Camitz' (unpublished) work on house quarantine.

5. Conclusion

The case study I presented in this paper illustrates that modeling methodology is a decision problem: modelers choose from a strategically prepared menu that model which they have reasons to believe will serve best their current purpose, under current conditions.

The examination of the menu developed at FoHM showed that these modelling alternatives differed mainly with respect to their simplicity—the number of parameters they contained—and therefore also with respect to the potential complexity of the target that they could represent. I argued that these differences are connected to the ongoing methodological discussion about whether modelers should trade off model-target similarity for the sake of increasing model simplicity—and thus about the validity of the KISS principle.

An analysis of the case study provided five reasons for choosing to engage in such a trade-off: lack of information, avoiding overfitting, avoiding fuzzy modularity, maintaining good communication, and allowing for error avoidance and detection. In addition, I argued for two observations: first, the purported trade-off is really between potential (not realized) model-target similarity; and second, each of these are indeed epistemic reasons. I conclude from these arguments that KISS, even in the time of COVID-19, is an epistemically important principle.[7]

References

Brouwers, L. 2005, "MicroPox: A Large-Scale and Spatially Explicit Microsimulation Model for Smallpox Planning", in Ingalls, V. (ed.), *The Proceedings of the 15th International Conference on Health Sciences Simulation*, San Diego: Society for Modeling and Simulation International (SCS), 70-66.

Brouwers IFFS talk 2020, Personal notes from a talk by Lisa Brouwers, Institute for Future Studies, Stockholm, 10/9/2020 (unpublished material).

Brouwers interview 2021, Transcript of interview with Lisa Brouwers, Stockholm, 28/5/2021 (unpublished material).

Brouwers, L. and Liljeros, F. 2005, "The Functional Form of an Epidemic in a Real-World Contact Network", Stockholm University working paper.

Brouwers, L., Camitz, M., Cakici, B., Mäkilä, K., and Saretok, P. 2009a, "MicroSim: modeling the Swedish population", preprint arXiv: 0902.0901.

Brouwers, L., Cakici, B., Camitz, M., Tegnell, A., and Boman, M. 2009b, "Economic Consequences to Society of Pandemic H1N1 Influenza 2009-Preliminary Results for Sweden", *Eurosurveillance*, 14, 37, 19333.

Burke D.S., Epstein J.M., Cummings D.A., Parker J.I., Cline K.C., Singa R.M., and Chakravarty S. 2006, "Individual-Based Computational Modeling of Smallpox Epidemic Control Strategies", *Academic Emergency Medicine*, 13, 11, 1142-49.

Chubb, M.C. and Jacobsen, K.H. 2010, "Mathematical Modeling and The Epidemiological Research Process", *European Journal of Epidemiology*, 25, 1, 13-19.

Edmonds, B. and Moss, S. 2004, "From KISS to KIDS—An 'Anti-Simplistic' modelling Approach", in Davidsson, P., Logan, B., and Takadama, K. (eds.), *Multi-Agent and Multi-Agent-Based Simulation: Joint Workshop MABS 2004*, Berlin-Heidelberg: Springer, 130-44.

Elliott, K.C. and McKaughan, D.J. 2014, "Nonepistemic Values and The Multiple Goals of Science", *Philosophy of Science*, 81, 1, 1-21.

Fasth, T., Ihlar, M., and Brouwers, L. 2010, "VirSim: A Model to Support Pandemic Policy Making", *PLoS currents*, 2.

Fenton, N.E. and Ohlsson, N. 2000, "Quantitative Analysis of Faults and Failures in a Complex Software System", *IEEE Transactions on Software Engineering*, 26, 8, 797-814.

FoHM. 2020a, "Skattning av behov av slutenvårdsplatser Covid-19", technical report, Public Health Agency of Sweden, release date 20/3/2020, revised 27/3/2020, https://www.folkhalsomyndigheten.se/contentassets/1887947af0524fd8b2c6fa7 1e0332a87/skattning-av-vardplatsbehov-folkhalsomyndigheten.pdf (last accessed 01/06/2021).

FoHM. 2020b, "Estimates of the Peak-Day and the Number of Infected Individuals during the COVID-19 Outbreak in the Stockholm Region, Sweden February—April 2020", technical report, Public Health Agency of Sweden. Artikelnummer 20055, revidering 1, release date 5/5/2020, https://www.folkhalsomyndigheten. se/contentassets/e1c3b83fa24f4d019e4842053ffd8300/estimates-peak-day-infect ed-during-covid-19-outbreak-stockholm-feb-apr-2020.pdf (last accessed 5/6/2021).

FoHM. 2020c, "Effekt av ökade kontakter och ökat resande i Sverige sommaren 2020", technical report, Public Health Agency of Sweden. Artikelnummer 21035, release date 15/6/2020, https://www.folkhalsomyndigheten.se/contentassets/2 9b815266baa4b409905c096be773df5/effekter-okade-kontakter-okat-resande-sve rige-sommaren-2020.pdf (last accessed 5/6/2021).

FoHM. 2020d, "Estimates of the Number of Infected Individuals during the Covid-19 Outbreak in the Dalarna Region, Skåne Region, Stockholm Region, and Västra Götaland Region, Sweden", technical report, Public Health Agency of Sweden. Artikelnummer 200103, release date 1/7/2020, https://www.folkhalsomyn digheten.se/contentassets/e1702f53eea144cdb1ca2ef854b45c35/estimates-peak-day-infected-during-covid-19-outbreak-20103.pdf (last accessed 5/6/2021).

FoHM. 2020e, "Förekomsten av covid-19 i Sverige21-24 april och 25-28 maj 2020", technical report, Public Health Agency of Sweden. Artikelnummer 20105, release date 2/7/2020, https://www.folkhalsomyndigheten.se/contentassets/fb47e0345 3554372ba75ca3d3a6ba1e7/forekomstren-covid-19-sverige-21-24-april-25-28-maj-2020_2.pdf (last accessed 7/6/2021).

FoHM. 2020f, "Scenarier för fortsatt spridning—Delrapport 1", technical report, Public Health Agency of Sweden. Artikelnummer 20223, release date 21/12/2020, https://www.folkhalsomyndigheten.se/contentassets/fa087223be5c4bef8298ee2 943f099ca/scenario-fortsatt-spridning.pdf (last accessed 5/6/2021).

FoHM. 2021a, "Scenarier för fortsatt spridning—Delrapport 2", technical report, Public Health Agency of Sweden. Artikelnummer 21035, release date 2/3/2021, https://www.folkhalsomyndigheten.se/contentassets/cc4a46bdfba54df38b9507c 5d6242f77/scenarier-fortsatt-spridning-21035.pdf (last accessed 5/6/2021).

FoHM. 2021b, "Scenarier för fortsatt spridning—Delrapport 3", technical report, Public Health Agency of Sweden. Artikelnummer 21087, release date 6/6/2021, https://www.folkhalsomyndigheten.se/contentassets/21443db264d84f7fbf495003 726b89f5/scenarier-for-fortsatt-spridning--delrapport-3.pdf (last accessed 5/6/2021).

FoHM. 2021c, "Brief Facts and Organization", https://www.folkhalsomyndigheten.se/the-public-health-agency-of-sweden/about-us/brief-facts-and-organization/ (last accessed 01/06/2021).

Giere, R.N. 1988, *Explaining Science: A Cognitive Approach*, Chicago: University of Chicago Press.

Goodman, N. 1972, "Seven Strictures on Similarity", in *Problems and Projects*, Indianapolis-New York: Bobbs-Merrill, 437-46.

Grüne-Yanoff, T. 2021, "Choosing the Right Model for Policy Decision-Making: The Case of Smallpox Epidemiology", *Synthese*, 198, 2463-84.

Humphreys, P. 2004, *Extending Ourselves: Computational Science, Empiricism, and Scientific Method*, New York: Oxford University Press.

Hejblum, G., Setbon, M., Temime, L., Lesieur, S., and Valleron, A.J. 2011, "Modelers' Perception of Mathematical Modeling in Epidemiology: A Web-Based Survey", *PLoS One*, 6, 1, e16531.

Karlsson, E. 2020, "Chefen efter kaosdagarna: "Vi har kört nattarbete"", Aftonbladet, 23/4/2020, https://www.aftonbladet.se/nyheter/a/QobQ6J/chefen-efter-kaosd agarna-vi-har-kort-nattarbete (last accessed 01/06/2021).

Karlsten, E. 2020, "Johan Gieseckes kontrakt med Folkhälsomyndigheten—kan arbeta motsvarande en halvtid under 2020", https://emanuelkarlsten.se/johan-gieseckes-kontrakt-med-folkhalsomyndigheten-kan-arbeta-motsvarande-en-halvt id-under-2020/ (last accessed 01/06/2021).

Lenhard, J. and Winsberg, E. 2010, "Holism, Entrenchment, and the Future of Climate Model Pluralism", *Studies in History and Philosophy of Science*, *Part B: Studies in History and Philosophy of Modern Physics*, 41, 3, 253-62.

Levins, R. 1966, "The Strategy of Model Building in Population Biology", *American Scientist*, 54, 4, 421-31.

Matthewson, J. and Weisberg, M. 2009, "The Structure of Tradeoffs in Model Building", *Synthese*, 170, 1, 169-90.

Parker, W.S. 2015, "Getting (Even More) Serious About Similarity*", Biology and Philosophy*, 30, 2, 267-76.

Paola, C. and Leeder, M. 2011, "Simplicity versus Complexity", *Nature*, 469, 38-39.

Rich, B.R. 1995, *Clarence Leonard (Kelly) Johnson 1910-1990: A Biographical Memoir*, Washington, DC: National Academies Press, http://www.nasonline.org/publications/biographical-memoirs/memoir-pdfs/johnson-clarence.pdf

Rochefort-Maranda, G. 2016, "Simplicity and Model Selection", *European Journal for Philosophy of Science*, 6, 2, 261-79.

Weisberg, M. 2013, *Simulation and Similarity: Using Models to Understand the World*, Oxford-New York: Oxford University Press.

Zucchini, W. 2000, "An Introduction to Model Selection", *Journal of Mathematical Psychology*, 44, 1, 41-61.