**Mimicking Relations of Simulations in Empirical Science**

**Abstract** The discuss­­ion in this paper concerns the epistemic issues modelers face when they generate simulated data to solve problems with existing empirical datasets, research tools, or experiments. We argue that to count as epistemically justified and evidentially relevant, a simulation model does not have to mimic the target system of an empirical investigation, nor do simulated data need to be generated by mimicking the processes for generating empirical data about the system. Simulated data may successfully mimic the target system even if the model does not. In such cases simulated data typically improve the representational relation between empirical data and the target system. We argue that the epistemic evaluation of simulated data should start with recognizing six different mimicking relations, and we show how to employ them with a number of examples from different sciences. We also discuss difficulties in defining simulation and simulated data. We suggest that Monte Carlo modelling counts as simulation if one relaxes the requirement that a computational *model* must mimic something: it is sufficient if the data produced by the model mimic something.

**1 Introduction**

Data are typically defined as something given (lat. datum) to us in observation, or as that which registers on a measurement or recording device in a form which is accessible to humans (e.g., Woodward 1989, p. 394). Another way to characterize data is that they are the particular values of the variables we are interested in. Simulated data, however, are not what nature ‘gives’ us because they are the result of calculating the consequences of what is in the computer code. Hence, ‘simulated data’ is an oxymoron. Yet, in this paper we argue that the use of simulated data is widespread in empirical science, and that evaluating whether they are adequate to their methodological functions requires its own epistemology.

There is an extensive philosophical literature that compares computer simulation to experimentation in order to assess whether the former can be epistemically on a par with the latter (see Morrison 2015, Lenhard 2016, and Imbert 2017 for overviews). Authors who argue for the epistemic superiority of experiments appeal to materiality (e.g., Guala 2002, Morgan 2003), the possibility of being confounded (Morgan 2003, Giere 2009), or the less demanding informational requirements of experiments (Roush 2018). According to Parker (2009) and Winsberg (2009), the materiality thesis is faulty because what is epistemically important is relevant similarity rather than materiality. Several recent contributions posit that classifying a case as a simulation or as an experiment does not convey the relevant epistemic difference (Parke 2014, Massimi & Bhimji 2015, Morrison 2015). We agree, simply because we can see cases in which simulation is used even when one could conduct an experiment. However, we do not aim to solve issues in the debate on the epistemic priority of experiments. Instead, we argue that although using simulated data have various different methodological functions, there is always some mimicking relation that is relevant for their epistemic evaluation. An epistemological investigation should thus begin by asking: what mimics what, and is the mimicking successful?

A simulation model may aim to mimic the Data Generating Process (DGP) responsible for empirical data about some system. Alternatively, it may aim to mimic the system itself, without aiming to mimic the DGP to which it gives rise. A simulation may also be intended to mimic several possible DGPs or target systems, and the plural is important here. One could also forgo such attempts altogether, and instead aim to produce a dataset that represents the system better than empirical data does. In some such cases, the simulated data mimic the system even though the model does not. We believe that this kind of heterogeneity gives us good reason to elevate the study of simulated data to the status of an independent research topic, which one could discuss without concern for comparison to experiments.

We will show that the set of items among which the relevant similarity relation is taken to hold is larger than has been acknowledged hitherto. The similarity relation is generally assumed to hold between the simulation model and the target of the investigation. Most authors writing about ‘targets’ think of them as selected features of the real-world systems a simulation model aims to represent. However, given that a target can be anything a model aims to represent, a data model (Suppes 1962) could be taken to represent data, and a dataset is the target in some simulation models.

Why, then, would modelers aim to mimic empirical data or several DGPs or systems, or to generate data by means of simulation about some system without aiming to represent the system or its DGP? It is because empirical data and the methods of analyzing them are problematic in various ways. Depending on the case at hand, researchers may need to *correct* empirical data with respect to known biases, *filter* the signal from noise, *add* data because statistical procedures or a larger model cannot run with existing data, *augment* data when there is some but it is insufficient for the purposes of the researchers, *aggregate* data when researchers endeavor to obtain an overall view of some empirical literature, or *enhance interpretability* when empirical data is too ambiguous. Simulated data may be used in all such cases.

The aim of this paper, then, is to provide a framework for understanding the use of computer simulation in dealing with problems related to existing empirical datasets, research tools, or experiments. We discuss cases from several different disciplines. The purpose of having a large number of examples is to show that the reasons for using simulated data and the relevant mimicking relations are different in different examples. Another aim is to provide a framework for understanding the functions of computer simulations and simulated data in addressing problems related to the use of empirical data. We therefore analyze various data-to-phenomena inferences (Woodward 1989, 2000, 2011) in which simulated data play a prominent role.

The paper is structured as follows. Section 2 gives an account of mimicking relations, and in Section 3 we develop a definition of computer simulations that encompasses Monte Carlo methods. Section 4 presents a number of examples on the different mimicking relations. We present our conclusions in Section 5.

**2 Mimicking Relations**

As Alisa Bokulich (2018) argues, it is not the ‘pureness’ of data or the data model that matters, but their ‘fidelity’ in representing the relevant features of the world (see also Humphreys 2014). In fact, simulation is often used to enhance fidelity. However, we use the term *mimicking* instead of fidelity because it does not limit us to the end result of an endeavor, and we can talk about the aim as well. There are several different kinds of mimicking relations between simulations, data, data models, and the world. In other words, the epistemically relevant mimicking relation depends on the function of the simulation and the case at hand. We thus distinguish between the different mimicking aims that simulations or simulated data may have.

A Data Generating Process (DGP) consists of all the processes that are responsible for producing empirical data about a real-world target system. In some cases it is not necessary to distinguish between the system and the DGP because the DGP may be known to be particularly simple and reliable, and only one variable is of interest to researchers. However, in many circumstances in which simulated data are employed the DGP is problematic. A DGP always relates to a given variable, and the reason for distinguishing between DGPs and target systems is that computer simulations often mimic a DGP for a variable without mimicking the target system that is ultimately responsible for the values of the said variable. It is possible to do this because the DGP may be known to be responsible for various problems in the data, such as bias, gaps, and inhomogeneity. When a simulation mimics a DGP, the DGP rather than the system constitutes the target. The target system also generates data for other DGPs.

There may be several DGPs for a given variable if this variable is measured from outputs from interactions with different processes. For example, the DGP for temperature data from satellites is different from that for temperature data from surface stations because the measuring techniques are different. A DGP may also concern the data-collection methods and the ways in which they generate data.

Insofar as simulations produce data, they do so by means of the *Simulated* Data-Generating Process (‘SDGP’ for short). A SDGP is thus whatever generates data in a simulation model, irrespective of whether or not it aims to mimic a DGP or DGPs. It is always known because the modeler coded it into the model.

It is possible to mimic a target system without mimicking a DGP related to it because a simulation model may represent the causal relationships in a target system without specifically representing the processes from which empirical data are collected. Furthermore, given that a simulation model may generate data all by itself, there may be no need to specify a particular DGP in the simulation model. In other cases, the simulation model aims to mimic the DGP without aiming to mimic the target system from which the data derive. A simulation model may also mimic a diverse set of target systems or their DGPs. Thus, although in many cases mimicking the DGP amounts to mimicking the target system, there are also cases in which the target system is mimicked without mimicking any of its DGPs, and cases in which a DGP is mimicked, but not the target system.

A *simulation* *model* may aim at mimicking

1. a specific target system, or
2. a specific DGP producing empirical data , or
3. a diverse set of DGPs producing empirical data, or
4. a diverse set of target systems.

It is also possible that the simulation *model* itself does not need to mimic the target systems(s) or the DGP(s) for empirical data: it is rather the *data* produced by the simulation that it is intended to mimic

(e) specific properties in existing empirical data, or

(f) a specific system.

In the latter cases, mimicking may happen without the model aiming to represent the system or the DGP for empirical data. Only the end result of the simulation, the data, mimics something. One might wonder why (f) constitutes a separate category. The idea here is that the data generated by the simulation mimic the system *without* mimicking in sense (e) and without the simulation model mimicking in senses (a), (b), (c), or (d). This case is most relevant when the empirical data are known to be problematic. There is a mimicking relation between the empirical data and the target system, but researchers know the respects in which this relation is less than perfect. If empirical data were not incomplete, unambiguous, or tainted by biases, and if they had a high signal-to-noise ratio, one would say they were ‘ideal’. Ideal data could be described as data that would be available in non-actual, ideal circumstances. In case (f), then, one generates data that are ideal in the sense that they better represent a system than the available empirical data. One must know the particular respect in which the empirical data are non-ideal. The simulated data are ideal compared to empirical data in at least one respect.

Note that in cases (a) and (b), whether the model successfully mimics the target system or the DGP is the crucial epistemic relation. However, if this mimicking succeeds, then the simulated *data* will also successfully mimic the target system or the DGP. However, such mimicking is merely an epiphenomenon that carries no independent epistemic weight in cases (a) and (b). Similar remarks apply to case (e). In contrast, this latter mimicking is the primary epistemic relation in case (f).

Generating data that are ideal in all possible respects is to all intents and purposes a utopian event. When simulated data mimics a specific system (f), the aim is to change empirical data by means of simulation so as to represent the target system better. As we will show, however, although it is necessary to know how the real DGP yields erroneous data, methods of generating more ideal data do not purport to mimic a DGP but are intended to generate data that takes the error into account. This is why case (f) is different from cases (a) and (b). It is thus not necessary to mimic the error for correcting the data, but we will show that testing the *tools* with which one *estimates* the error does require mimicking it correctly.

In addition to using ‘pure simulated data,’ one can combine simulated data with empirical data, the aim being to obtain a corrected, enhanced, or unbiased dataset as a whole. Such ‘hybrid datasets’ are then used instead of empirical data for different evidential and epistemic purposes. The distinction between *hybrid* and *pure* simulated data is not epistemically important because both categories are internally heterogeneous. As we will show, different simulations that end up with hybrid data mimic different things, and pure simulated data are generated for several different purposes and with different mimicking relations in mind.

**3 What Are Simulated Data?**

For the current purposes, let us assume that there are two kinds of data. The origin of empirical data lies in reality such that its production requires causal interaction with measuring or detection devices, or with human sensory capacities, whereas the origin of simulated data lies in a computer simulation model. A simulation is not in causal interaction with the system the simulated data is taken to represent, nor with the empirical data if that is what the simulation aims to represent. Stating that the origin of simulated data is a computer simulation model does not give a complete definition insofar as it does not clarify the notion of a computer simulation. In this section, we provide an account of such definitional issues. Hartmann’s (1996) definition of simulation refers to mimicking a process with another process, whereas Humphreys (2004, p. 110) takes into account the possibility that simulation may mimic an object rather than a process, and requires that it provides solutions to a computational model. This computer simulation model is a program that runs on a computer (see also Winsberg 2015).

It would seem to follow from the above that the simulation model must mimic the *target* system of the investigation, the target system being interpreted as a spatiotemporal object or process (cf. Peschard forthcoming). Hence, it would seem to follow that a simulation must mimic the target system (b) or the DGP responsible for the empirical data about that system (a). However, some simulations do not purport to mimic them. Monte Carlo simulations are a common case in point because the randomness on which the method is based is not normally meant to be a claim about the object or process of interest (Beisbart & Norton 2012).

By a Monte Carlo method, we mean any method that exploits computer-generated randomness in calculating some results. The randomness is typically the result of using a pseudorandom number generator. The generators are algorithms that give rise to complicated deterministic processes that produce data that mimic those provided by genuine random processes. Although some philosophers (e.g., Hartmann 1996, Lenhard 2016) and many scientists refer to Monte Carlo methods as ‘simulations,’ the above definitions do not do justice to Monte Carlo models (Winsberg 2015). Moreover, Grüne-Yanoff and Weirich (2010) argue that Monte Carlo methods lack the mimicking aspect and therefore should be counted as calculations, not simulations. We propose, however, that the best way to account for why Monte Carlo methods are often counted as simulations is with reference to their mimicking aspects.

Although the data generated by random number generators always mimic the corresponding analytical probability distributions, we do not appeal to this mimicking relation in arguing that some Monte Carlo methods count as simulation. It is epistemically relevant only if it fails, and if the failure affects the relevant results calculated with pseudorandom numbers.[[1]](#footnote-1) The relevant mimicking relation does not hold between the data generated by the pseudorandom number generators and the analytical distributions, but rather between the simulation model that *embeds* the pseudorandom generator, and something else, or alternatively, between the data generated by the simulation model (instead of the pseudorandom generator) and something else.

What, then, could this ‘something else’ be? Some Monte Carlo models mimic the target system or the DGP that generates the empirical data (see Galison 1996), whereas others mimic the empirical data without mimicking the DGP or the system. If a target refers to whatever a model represents, then the target of such simulation models is a dataset. However, the randomness of the Monte Carlo model may well not mimic the data, either. In such cases it is usually the *data* produced by the Monte Carlo simulation that *mimics* the empirical data.

Below we provide examples of three kinds of mimicking relations in Monte Carlo simulations. First, they may *mimic* *the error or the bias* that is known to distort the DGP (b). What we mean by bias is that the DGP is known to produce data that systematically distorts some characteristics of the system. Second, the data they produce may *mimic* the *target system or process* (better than the empirical data) (f). Third, if there are some empirical data, but not sufficient for the purpose at hand, the data produced by a Monte Carlo simulation may mimic *the relevant properties* *of existing empirical data (e).*

In the first case, the simulation model represents both the error mechanism and how the DGP would operate if it were not perturbed by error. The mimicking relation thus holds between the simulation model and the DGP. Given that the DGP is modeled, this means that these simulation models contain a description of a spatiotemporal target or a description of an experiment or measurement device. In the second and the third cases, however, the simulation model does no such thing, and it is not required to. It is only the data resulting from the simulation that has to mimic something.

Let us now consider our source-based definitions of data in more detail. Our notion of empirical data is tantamount to Barberousse et al.’s (2009) ‘dataE’, which they define as being of ‘empirical origin, namely produced by physical interactions with measuring or detection devices’ (p. 560). The problem with source-based definitions is that they do not always allow clear differentiation between simulation and computation. In particular, it is difficult to distinguish between a simulation that uses empirical data as input, and a measurement that involves the computational refinement of empirical data (Arnold 2013). Clearly, calculating an average of empirical data with a computer does not change its empirical status. Nevertheless, such data may be processed and modified much more extensively, and such processing could even change what the data are about (Humphreys 2013).

Consider ‘data imputation,’ understood as a general class of methods aimed at filling in the missing values of empirical data with artificial values. The filling-in may be carried out randomly, or by means of interpolation, mean substitution or maximum likelihood, for example. It would seem natural to call the data *artificial* if imputation methods add data points into the datasets where empirical data are missing. Given that the use of computers and statistical software facilitates and automates the production of such artificial data, does this mean that they are always also simulated? Let us bear in mind that the origin of simulated data must be a computer simulation model. The answer to this question depends on what it means to say that the source of data is empirical or simulated.

Lusk (2016; see also Barberousse & Vorms 2013) proposes that data produced by a computer simulation model are always empirical as long as the model has any empirical data as input, but he does not limit the transformations to data in any way. Arnold (2013), in contrast, proposes that the data are empirical if the output data from a simulation are interpreted to concern the same variables that are responsible for the empirical input data. Although Arnold’s proposal is appealing, neither of them is satisfactory. Lusk’s proposal provides misleading information about what is epistemically relevant about data. Calibrating a simulation model with empirical data typically occurs in *modelling* the data-generating process or the target system, whereas when empirical data are filtered or corrected, they typically retain their status as *evidence* about a target system. Lusk’s proposal forces one to ignore this epistemic difference. It is also somewhat implausible in that the weakest possible empirical input seems to be sufficient to render the outgoing data empirical. Lusk accepts data produced in a simulation as empirical even on the basis of a mere qualitative fit with the target system. In fact, any simulation has some qualitative fit to some system insofar as it successfully mimics something. Thus, if any qualitative fit is acceptable, it becomes well-nigh impossible for a simulation not to yield empirical data. Finally, climate models may yield data about the future climate, and such data cannot, in principle, be empirical, even if the simulation model is calibrated with historical empirical data.

Arnold’s proposal fails because data assimilation for climate models takes empirical data about some variables as input to set up the model, and then replaces the values of the same variables with simulated values (see e.g., Parker 2017; Werndl 2019). Even though the input data and the output data concern exactly the same variables[[2]](#footnote-2), the input data are empirical whereas the output data are clearly simulated.

We agree with Arnold (2013) that it is difficult to distinguish cases in which a model produces simulated data due to being calibrated with empirical parameter values from the mere computational refinement of empirical input data, but we are obliged to solve such classification problems only if their solution solves major epistemic problems. Although there is often a significant epistemic difference between these two uses of empirical data, it does not lie in whether the hybrid method is to be conceptualized as empirical or simulated.

Even though we do not agree with the proposals discussed thus far, we do not provide a solution that would solve the problems of clearly delineating when data are simulated, artificial, or empirical. We readily admit that if one is interested in being able to categorize every method into either simulation or experiment, we have no answers to offer. As a matter of fact, we doubt whether a universally applicable distinction can be made. Indeed, if we are right about this, the conclusion spells trouble for those who believe there is a significant epistemic difference between the different kinds of data. Yet, there are clear cases of all such kinds of data. Indeed, given that our examples include cases in which simulated data are superior to artificial and/or empirical data, the interesting question is not where exactly to draw the line between empirical and simulated data, but rather what mimics what, and whether the mimicking is successful and thereby succeeds in solving epistemic problems.[[3]](#footnote-3)

We do need a clear way of distinguishing between computation and simulation, however, in order to properly delimit the range of cases that involve simulated data. We propose to do so as follows (cf. Lehtinen & Kuorikoski 2007). Calculating an average of empirical data counts as computation because the epistemic credibility of a computation does not depend on any mimicking relation. It depends, instead, on whether or not the mathematical calculation is correctly carried out. For similar reasons, some Monte Carlo exercises should be classified as *computations* rather than simulations.

The epistemic credibility of a simulation always depends on whether it, or the data it generates, successfully mimics something in the empirical world in the sense of (a)-(f) discussed above. In contrast, if a model is computational but not a simulation, it is so because the computational operation does not (aim to) mimic anything empirical, and because the resulting data do not mimic anything empirical, either. For example, calculating the value of pi using Monte Carlo methods does not count as a simulation because the thus-calculated value does not need to be compared with anything in the empirical world, in that it can be compared with the result of an analytical mathematical calculation. This kind of Monte Carlo calculation computes a mathematical object. Recall that the pseudorandom number generator always mimics another mathematical object, namely a genuine random distribution. We did not claim that Monte Carlo methods are simulations in virtue of this mimicking relation because what is mimicked is not empirical. However, Monte Carlo methods are also often used to calculate the values that various functions or equations would yield as data. If these data have an empirical interpretation, in other words if they aim to represent some empirical target, then the data yielded by the Monte Carlo method are simulated. In other words, if the Monte Carlo method calculates the abstract mathematical form of a function or an equation, then it counts as computation, but if it calculates the values such a function or equation would have if it were instantiated with empirical values, then it counts as a simulation. This way of distinguishing between computation and simulation does not hinge on whether the computer model takes empirical data as input or produces them as output, but rather on whether or not the epistemic evaluation of the computer model or the data it produces requires considering some mimicking relation.

Before embarking on a discussion of our examples of simulated data, let us explain the nature of our enterprise. Our primary interest lies in epistemically evaluating different uses of simulated data. We do not aim to show that such data constitutes a uniform category. On the contrary, there are many different kinds of simulated data in that they are generated for various different reasons (correcting, filtering, and so on), they can be generated using different simulation techniques (such as discretization, Monte Carlo, and IBMs), they may form hybrid datasets with empirical data or replace empirical data, and they may derive from a model that uses calibrated empirical data or from a purely theoretical enterprise. Although all these aspects and considerations may have systematic epistemic consequences and even systematic relations with different mimicking relations, we are not able to discuss every possible kind of simulated data. The point of the examples is rather to show that it is fruitful to organize the epistemic evaluation of simulated data around the six mimicking relations that we introduced in Section 2: identifying the relevant mimicking relation indicates where to look for the epistemically relevant considerations. For expositional reasons, we present the examples in the order a, b, d, e, and then finally f and c together in one subsection.

**4 Examples of different mimicking relations**

**4a) Mimicking the target system: Data Assimilation in Climate Science**

Empirical raw data about the climate are problematic in various different ways. There are temporal and spatial gaps in climate datasets, they suffer from discontinuities resulting from changes in the calibration of instruments or in their surroundings, and some of the data are systematically biased, for example. *Data assimilation* refers to the whole process of integrating various sets of data with global simulation models (see Edwards 2010, Parker 2017).

Before they started using simulations climate modelers employed various more-or-less ad-hoc methods to fill the gaps with artificial data. These included spatial-interpolation techniques, meaning that they inserted the missing values into the missing cells in the grid by imposing a smoothly changing average from the edge points in the existing data.

The justification for employing simulation in data assimilation is that it allows for using physical theory in determining the properties of the missing data rather than adding artificial data points ad hoc or based on statistical methods. Constructing the reanalysis data with the help of global simulation models is an attempt to mimic the real processes in the target system that generate the needed missing climate data: in other words, the simulation *model purports to mimic the target system* (a), the climate system. The simulation does not purport to mimic the real DGPs that produce empirical data about the climate, given that the *DGPs* are known to produce erroneous and incomplete data. Instead, the purpose of using simulation is to apply knowledge about the physical relationships governing the climate system in order to correct, fill and smooth the erroneous empirical data produced by the DGPs. Hence, the epistemic evaluation of models that generate reanalysis data hinges on whether they are able to mimic the target system successfully (see e.g., Edwards 2010, Ch. 12 for an account of how failures of such models affect the reanalysis data).

As a matter of fact, using simulated data has brought about a clear improvement in data quality. Artificial data from interpolation is not based on solid, theoretical-background knowledge of the physical processes: conversely, using general circulation model (GCM) simulation to fill the gaps in the dataset constitutes an attempt to better mimic the target system (a). As Parker (2016) notes, reanalysis data is not self-evidently less reliable than empirical data. However, evaluating the reliability of such data is complex because it depends on the overall performance of the global model as well as on any success in correcting for errors in the empirical data. This explains why some climate modelers lean on the predictive success of global simulation models in justifying the use of reanalysis data (e.g., Dee et al. 2011, p. 555). Given that it is not possible to analyze improvements in climate models attributable to reanalysis as opposed to the empirical data separately from the rest of the model, modelers appeal to the overall predictive improvement it generates.

The overall purpose of the reanalysis data is to provide a better description of the state of the climate than that produced from empirical data deriving from the DGPs. In this sense, then, climate scientists interpret reanalysis *data* as intending to mimic the target system (f). However, in this case, even if the *model* successfully mimicked the target system (a), this would not be sufficient for the data successfully to do so (f). This is because the reanalysis derives from a huge assortment of methods used to manipulate the empirical data over and above simulations of the target system. Without even trying to describe these complexities, let us merely note that many additional kinds of mimicking relations are involved in generating reanalysis data.

**4b) Mimicking a specific DGP producing empirical data**

We do not provide a case study on this kind of simulation because other philosophers have already discussed several examples. Tal (2011), for example, describes how computer simulations were used in modeling the detection procedures used to study superfluid-to-Mott-insulator phase transitions. Here the simulation model mimics a specific DGP (b) in an experiment, and simulated data are used to clarify the status of empirical data as evidence (see also Massimi and Bhimji (2015), and Lusk (forthcoming)).

**4d) Mimicking a diverse set of possible target systems: Metapopulation Models and Parameter Estimation Methods**

Metapopulation models describe the colonization and extinction dynamics of a species over a spatial area. The species is described as one metapopulation. It consists of subpopulations living in separate areas. The spatial area consists of habitat patches and the habitat matrix, where subpopulations can and cannot persist, respectively. Subpopulations can migrate through it to other habitat patches and colonize them. The extinction and colonization probabilities of a metapopulation depend on the size of the habitat patches, the distances and connections between them, and species-specific parameters concerning survival and migration ability.

The predictive accuracy of metapopulation models depend on the data required for estimating their parameter values. However, the data required to estimate parameter values are often biased, inaccurate, and non-projectable in space or time. Thus, additional tools –parameter estimation methods – are needed to provide more accurate estimations of parameter values for models based on empirical data. One of the advantages of “incidence function” models (Hanski 1994) over earlier models is that the data they require for parameter estimation is more accurately measurable. In addition, Hanski (1994) introduced a new nonlinear regression method for estimating the values of a model’s parameters when fitted to data.

Although the incidence-function framework displaced previous modelling frameworks, critics pointed out problems in Hanski’s parameter-estimation method. The method required that the metapopulation be in a state of colonization and extinction equilibrium, for instance. The problem is that this does not hold in many metapopulation scenarios, hence the method provides biased parameter estimates in many cases. Hanski’s method seemed to lack robustness.

Given that Hanski’s (1994) method gave biased estimates in some metapopulation scenarios Moilanen (1999) developed an improved Monte-Carlo-based parameter-estimation method. Determined to show proof that his method outperformed Hanski’s method, he used simulated data to provide the known “truth” to which the performance of different methods could accurately and reliably be compared.

Moilanen (1999) created a diverse set of simulated metapopulation data. He varied the sizes of the habitat patches and the distances between them, the migration abilities and extinction probabilities of metapopulations, and the minimum patch-size area required for subpopulations to persist. The different simulations created a diverse set of data concerning patch presences/absences and extinction as well as colonization turnover events in metapopulations of different generic species with their varying extinction and colonization dynamics. These simulations thus *mimicked a diverse set of possible target systems* (d).

Moilanen (1999) tested the predictive accuracy and performance of parameter-estimation methods on the above-mentioned simulated data. If a method captures a diverse set of simulated data produced by different simulated systems, then its performance is robust. The results clearly indicated that the new Monte Carlo method not only produced more accurate parameter estimates of metapopulation data than Hanski’s regression method, it was also more robust in its performance under diverse metapopulation simulated datasets.

We now consider cases in which simulated data have mimicking relations that have not previously been recognized. The next two cases are different from the previous ones in that the simulation model does not mimic a target or a DPG, nor a set of potential DGPs or target systems. Instead, the simulated data mimics the specific properties of empirical data (e).

**4e) Mimicking the distributional properties of the existing empirical data**

**Amplifying data in voting theory**

The aim of normative voting theory is to determine which voting rules are the best in different circumstances concerning the numbers of voters and candidates, the range of options, and the institutional setting. There are several criteria governing good performance in the case of voting rules, but they all reflect the idea that the rules should select a clearly best candidate when there is one, and they should not exhibit various sorts of paradoxical changes in selected candidates when the electorate changes slightly.

Until recently it was very rare to use empirical data on real elections to test the performance of voting rules, and most approaches were based on analytical mathematical models or computer simulations. The standard assumption in the simulations was of an *impartial anonymous culture* (IAC), meaning that each preference ordering is equally likely in a randomly generated dataset. The reason why *simulations* were used to generate data on voting preferences instead of proving theorems about distributions of voters, as in analytical modelling, had to do with tractability: it was easier to simulate than to compute the probabilities analytically. IAC fails to mimic the empirical data (see e.g., Gehrlein 2006, p. 104), and this matters if the frequency with which various problems emerge is different in IAC simulations and real elections.

Plassmann and Tideman (2014) sought to evaluate how often various voting rules generated problematic outcomes but, given the rarity of many of these problems, they needed to study millions of elections. However, empirical data are available for only a few thousand elections. Their solution was to use a statistical model that was calibrated with empirical data on votes. This allowed them to construct an *amplified* simulated dataset of a million simulated elections that mimicked empirical data in that it preserved and replicated relevant aspects, especially the distributional properties of the empirical data on voting.[[4]](#footnote-4) Their point in introducing a model calibrated with empirical data was thus to provide information about real rather than theoretical problems. Plassmann and Tideman (2014) evaluated how frequently 14 different voting rules encountered 10 voting paradoxes. Plassmann and Tideman’s approach and IAC are both examples of Monte Carlo methods, but the distributions in their datasets look very different from those simulated under the IAC assumption: the former tend to have fewer candidates with a genuine chance of winning than the latter.

The amplified datasets are similar to each other (and to empirical data) in terms of their distributional properties, but they differ in the details. Given that Plassmann and Tideman evaluate how often it is reasonable to expect various problems to occur under different voting rules, it is important that the amplified data reflect the relevant distributional properties of empirical data.

Mimicking the real DGP for empirical data on votes (b) is considered to be difficult if not impossible. If Plassmann and Tideman were interested in modelling the DGP for a single election they could have studied various socioeconomic factors that explain why different voters vote for different kinds of candidates or policies.[[5]](#footnote-5) However, the DGP for the next election implies another society, another voting rule, and so on, and the relevant data must concern thousands of elections. Hence, in some sense there is no DGP for the kind of data that studying the properties of voting rules requires.

Rather than mimicking a specific DGP for an election or elections, therefore, Plassmann and Tideman *sought to mimic the distributional properties of the existing empirical data* on voting (e). Constructing such simulated datasets yields information about how slightly different yet realistic datasets might affect the performance of various voting rules. In other words, they aimed at mimicking empirical data. Note that although the simulation model includes a simulated DGP, this SDGP is not meant to mimic the real one. It does not do so because it represents the DGP as if it were random, even though the real DGP is known not to be random in this sense. In this context, the relevant distributional properties of the data refer to the following kinds of issues: the proportion of elections in which there is one clear winner, two close contestants, or three genuine competitors; and how often is there a consensus candidate who is not the best choice of many voters but who is the second-best for many.

**Other cases of mimicking empirical data**

We believe that mimicking empirical data without mimicking the DGP or a target system is widespread. Let us thus briefly consider two more examples, from climate science and macroeconomics, respectively: Santer et al. (2008) and Hoover & Salyer (1997). While Plassmann and Tideman would perhaps try to mimic the DGP for elections if it were possible, these contributions are different. The point of these simulations is to provide an artificial DGP which does not aim to mimic the real DGP or the target system.[[6]](#footnote-6) The starting point here is that other authors claim to find a relationship X on the basis of empirical data. The point of using simulated data is to show that data-analysis techniques T used to arrive at this conclusion are faulty and hence that X has not been shown to be found. In both cases a simulated dataset aims to mimic the statistical properties of empirical data, but it is generated in such a way that X is not part of the SDGP. The simulations then show that T finds X in the simulated data even though X is not there by design. Hence the failure of T is demonstrated.

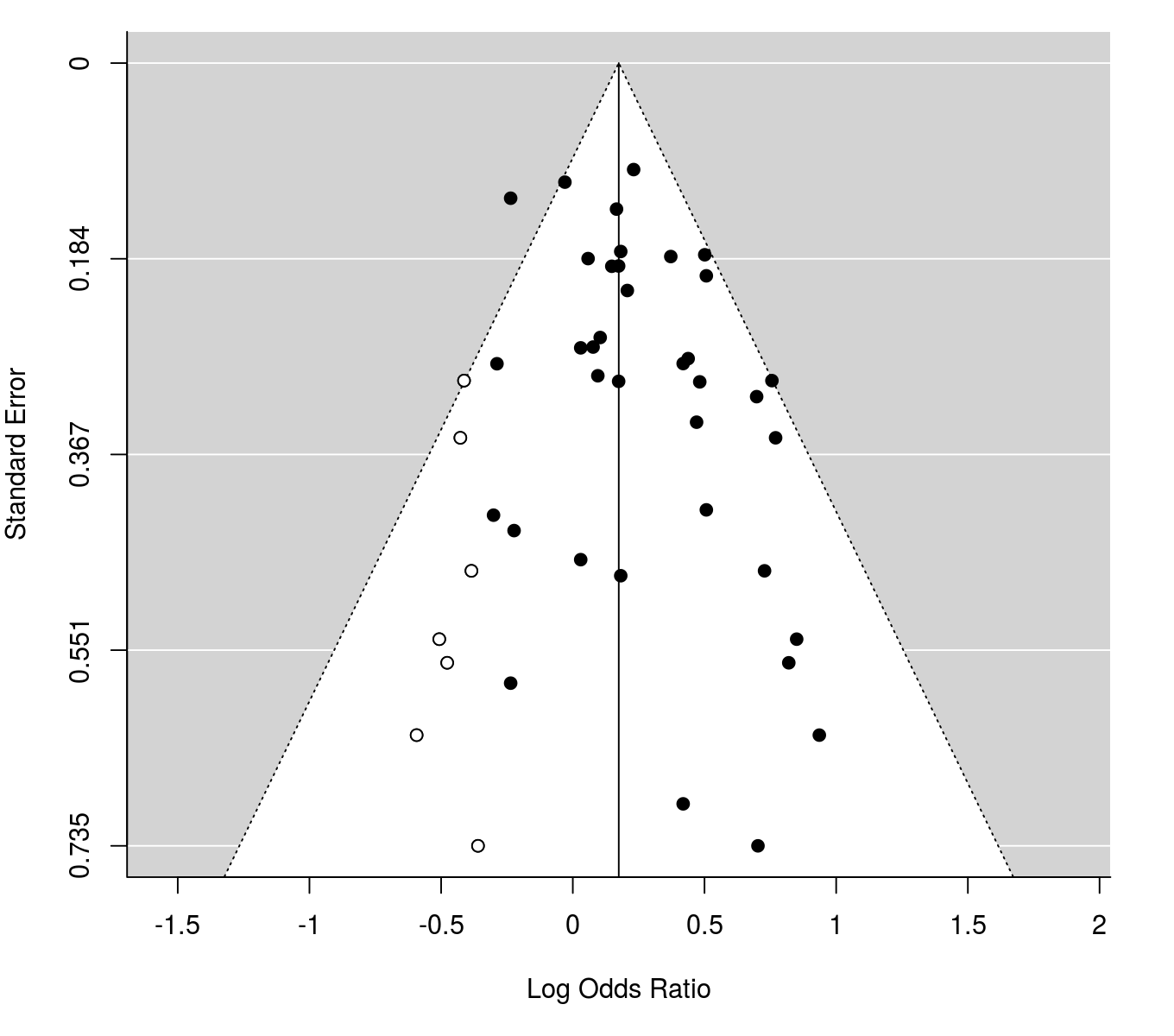
**4f) Generating simulated data that mimic the target system better than the empirical data and 4c) mimicking a diverse set of DGPs producing empirical data: Correcting for Publication Bias in Meta-analyses**

Meta-analyses provide quantitative overviews of empirical literature, and explain the differences in primary studies. A meta-estimate is a number that purports to describe the quantitative content of a collection of empirical studies on some topic. Publication bias is a widely recognized problem in meta-analysis, which arises when studies fail to be published either because they do not provide statistically significant estimates of effects or because the estimates turn out to be inconsistent with a widely recognized theory. Studies that have produced estimates of the “right” sign, and that have passed the five-percent-significance threshold are typically published. The task of the meta-analyst is to develop tools for reliably recognizing publication bias and correcting for it. The traditional way of dealing with this was to construct a *funnel graph*[[7]](#footnote-7), inspect it visually to verify whether or not it was asymmetric, and if so, to propose some statistical or ad-hoc correction methods based on estimations of bias in empirical data. Researchers then came up with computer-based approaches aimed at automatic correction, which produce artificial or simulated data that supplement existing meta-analyzed data.

Meta-analysts concerned with publication bias may try to construct a dataset with properties that empirical data would have were it not thus tainted. The properties of this dataset depend crucially on the bias *estimators* that are applied to empirical data before a hybrid dataset is constructed.

Simulated data is used for two different functions in meta-analysis. The meta-analysts first construct an estimator of publication bias. They then test the performance of the estimator with simulated data on a diverse set of DGPs (c). Second, if the estimator provides satisfactory results with simulated data, it is then used to correct and supplement the empirical data so as to better mimic the target system (f). Below we give two examples of this methodology, the “trim and fill” method described by Duval and Tweedie (2000a, 2000b), and the Monte Carlo method of Bowden et al. (2006).

Having tested the applicability and credibility of their trim-and-fill method, as an estimator of publication bias, with simulated data, Duval and Tweedie (2000a, 2000b) go on to use the method to derive estimates from an empirical dataset. They first estimate the degree of bias by calculating the number of studies (with certain characteristics) that would have to be added to the empirical dataset so as to remove it. An iterative algorithm adds data points into the dataset until the funnel graph is sufficiently symmetric, such that they have the same standard deviations as the data points in the empirical data, and the absolute value of the divergence from the mean (the vertical line in the figure below) is the same but the points are on opposite sides of it: see the figure below.



The added data points are depicted in the figure as empty circles, and the empirical data points as black dots. Thus, the funnel plot is trimmed by filling in the empty circles.

Bowden et al. (2006), propose a Monte Carlo method to deal with publication bias in meta-analysis data, generating simulated data to test their modification of the method. Having thus tested its performance with simulated data, they apply it to an empirical dataset. Bowden et al.’s method adds data points randomly to empirical data. They provide an example in which the Monte Carlo model generates data that mimic the distributional properties that empirical data would have in the absence of bias. This amounts to *generating simulated data that mimic the target system better than the empirical data* (f). Nevertheless, the Monte Carlo model does not represent the target system or the DGP, and thus cannot be said to mimic them.

Note that Duval and Tweedie’s trim-and-fill method is different from Bowden et al.’s Monte Carlo method: Duval and Tweedie rather use an iterative mathematical algorithm in adding data to an empirical meta-analysis dataset. Given that their algorithm is not based on randomness, one could take their supplementary data to be artificial but not simulated. However, according to our definition, the resulting data are simulated in virtue of the fact that they aim to generate a dataset that better mimics the target system. Whether or not one chooses to call the data simulated has no epistemic significance, however. What is crucial is that the methods are able to generate data that mimic the target system correctly (f), and this means getting the distributional properties right.

To better understand why this is so, note that the data-generating process for a meta-analysis (MADGP) consists in the inclusion rules for selecting the published primary studies, the studies themselves, and the process that is responsible for publication bias. Although taking the bias into account means better modelling the empirical MADGP, the simulation *models* that generate hybrid datasets do not even purport to mimic the MADGP, in contrast to meta-analytical simulations aimed at testing the performance of bias estimators.

We mention these considerations because the meta-analytical methods of generating hybrid datasets do not seem to be problematic *if* the publication-bias estimators are reliable. If they are, just about any way of fixing the bias will do because the usual aim of meta-analysis is to generate a correct meta-estimate about some empirical literature. Nor is there any epistemic difference between tweaking the empirical data by hand in some ad-hoc way and using computerized methods. The real problem lies in correctly estimating the bias. Perhaps this is a more general feature of cases in which empirical data are known to be biased: as long as the direction and the magnitude of the bias are reliably known, it does not really matter how it is corrected.

In the case of publication bias, testing the performance of estimators with simulated data is clearly the only method that yields information about whether the estimators are sufficiently reliable. It is thus not surprising that there is a growing body of simulations that test the performance of various estimators when the data may be tainted by publication bias (Callot & Paldam 2010, Ringquist 2013, Reed 2015, Stanley & Doucouliagos 2016). The credibility of these simulations hinges on whether the possibly relevant causal factors that generate and bias empirical data are correctly modeled. Because these factors concern the DGP, *meta-analytical simulations that test the performance of research tools aim to mimic a diverse set of DGPs producing empirical data* (c).

**7 Conclusions**

This paper began with an account of the difficulties in defining simulation and simulated data. We suggest that one could acknowledge Monte Carlo modelling as simulation if one were to relax the requirement that a computational *model* must mimic something: it is sufficient if the data produced by such a model mimic something. Having identified six different mimicking relations that are relevant for evaluating the epistemic credentials of various simulation methods and the data being produced, we considered various cases in which a simulation model or simulated data do mimic something. The common denominator in the cases is that the available empirical data, research tools, or experiments are acknowledged to be deficient in some way, and simulation methods are then employed to make them more reliable or relevant.

Simulated data are used for various evidential, epistemic, and methodological purposes. We discuss cases in which simulated data are used as a testing benchmark, for filling gaps in and correcting data, for example. However, we do not believe that these methodological uses exhaust the possibilities. Furthermore, the relevant mimicking relations do not map neatly onto an account of the different uses of simulated data. For these reasons, we have not attempted to give an exhaustive account of different kinds of simulated data. Our aim instead was to shed light on the epistemic interpretation of using such data: figuring out the relevant mimicking relation helps to enhance understanding of whether or not the use of simulated data is justified. Table 1 summarizes our cases and the relevant mimicking relations.

**Table 1: Mimicking relations**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| What is mimicked | One specific system  (a) | One specific DGP  (b) | Several DGPs and the error mechanism  (c) | Several possible target systems  (d) | Existing empirical data  (e) | One specific system  (f) |
| Cases | Climate reanalysis | Superfluids | Meta-analysis  Estimators | Meta-population  Models | Voting rules | Meta-analysis studies |
| What does the mimicking | Simulated data | Simulation model | Simulation model | Simulation model | Simulated data | Simulated data |
| Purpose of simulation | Correct, smooth, and fill gaps in climate data | Interpret empirical data | Test MA estimators | Test MP estimators | Test voting rules | Correct biases in empirical data |

The use of simulated data is ubiquitous in empirical science. What is it that facilitates or enables its use? One explanation is that in the cases we discuss, individual data points in empirical data do not have much epistemic relevance: if they did, the purity of empirical data would be much more relevant. What usually matters are the distributional properties of data rather than individual data points per se. Our addition to Bogen and Woodward’s (1988) account is thus that the extent to which simulated data are used in empirical science provides evidence about the extent to which observable data points are epistemically irrelevant. This feature of empirical data makes it possible to use simulations or simulated data to filter and augment empirical data, correct for bias, and fill in missing empirical data points, for example.

Given that there is no account that studies the variety of mimicking relations, it seems to us that philosophers of science have taken for granted that simulation models – and scientific models in general – have one privileged mimicking relation. Typically, they have focused on the idea that simulation models should mimic correctly their specific target system (a). We have shown that simulation models have other mimicking relations and that there is no privileged mimicking relation, More importantly, it is not always the *model* that should do the mimicking. In cases (e) and (f), it is not the simulation model but rather the *data* it produced that is intended to mimic certain specific properties of existing empirical data (e) or the target system (f). It is impossible to understand the epistemology of these cases without paying attention to the special features of simulated data. Recognizing the relevant mimicking relation is the starting point for evaluating the evidential and epistemic relevance of the use of simulated data in empirical science.

**References**

Arnold, E. (2013) Experiments and Simulations, Do They Fuse? In Durán, J.M. & Arnold, E. (eds.): *Computer Simulations and the Changing Face of Scientific Experimentation*. Newcastle, Cambridge Scholars Publishing, 46-75

Barberousse, A., Franceschelli, S. & Imbert, C. (2009) Computer Simulations as Experiments. *Synthese* 169: 557–574

Barberousse, A. & Vorms, M. (2013) Computer Simulations and Empirical Data. In Durán, J.M. & Arnold, E. (eds.): *Computer Simulations and the Changing Face of Scientific Experimentation*. Newcastle, Cambridge Scholars Publishing, 29-45

Beisbart, C. & Norton, J. D. (2012) Why Monte Carlo Simulations Are Inferences and Not Experiments. *International Studies in the Philosophy of Science* 26: 403-422

Bogen, J. & Woodward, J. (1988) Saving the Phenomena, *The Philosophical Review* 97: 303-352.

Bokulich, A. (2018) Using Models to Correct Data: Paleodiversity and the Fossil Record. *Synthese*  https://doi.org/10.1007/s11229-018-1820-x

Bowden, J., Thompson, J. R. & Burton, P. (2006) Using Pseudo-Data to Correct for Publication Bias in Meta-Analysis. *Statistics in medicine* 25: 3798-3813

Bueno, O. (2014) Computer Simulations: An Inferential Conception. *Monist* 97: 378-398

Callot, L. & Paldam, M. (2011) Natural Funnel Asymmetries; A Simulation Analysis of the Three Basic Tools in Meta-Analysis. *Research Synthesis Methods* 2(2): 84-102

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. -J., Park, B. -K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. -N. & Vitart, F. (2011) The ERA-Interim Reanalysis: Configuration and Performance of the Data Assimilation System. *Quarterly Journal of the Royal Meteorological Society* 137: 553-597

Duval, S. & Tweedie, R. (2000a) A Nonparametric "Trim and Fill" Method of Accounting for Publication Bias in Meta-Analysis. *Journal of the American Statistical Association* 95: 89-98

Duval, S. & Tweedie, R. (2000b) Trim and Fill: A Simple Funnel-Plot-Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. *Biometrics* 56: 455-463

Edwards, P. N. (2010): *A vast machine: computer models, climate data, and the politics of global warming*, MIT Press, Cambridge, Mass.; London

Galison, P. (1996) Computer Simulations and Trading Zone. In Galison, P. & Stump, D. (eds.): *Disunity of Science: Boundaries, Context, and Power*. Stanford, Stanford University Press, 118-157

Gehrlein, W. V. (2006) *Condorcet's Paradox.* Heidelberg, Springer, 292 p.

Giere, R. N. (2009) Is Computer Simulation Changing the Face of Experimentation? *Philosophical Studies* 143(1): 59-62.

Grüne-Yanoff, T. & Weirich, P. (2010) The Philosophy and Epistemology of Simulation: A Review. *Simulation & Gaming* 41: 20-50

Guala, F. (2002) Models, Simulations, and Experiments. In Magnani, L. & Nersessian, N. (eds.): *Model-Based Reasoning.* Boston, Springer, 59-74

Hanski, I. (1994) A Practical Model of Metapopulation Dynamics. *Journal of Animal Ecology* 63: 151-162

Hartmann, S. (1996) The World as a Process: Simulations in the Natural and Social Sciences. In Hegselmann, R., Mueller, U. & Troitzsch, K. (eds.): *Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View*. Dordrecht, Kluwer, 77-100

Hoover, K. & Salyer, K. (1998) Technology Shocks or Coloured Noise? Why Real-Business-Cycle Models Cannot Explain Actual Business Cycles, *Review of Political Economy* 10: 299-327

Humphreys, P. (2004). *Extending Ourselves: Computational Science, Empiricism, and Scientific Method.* New York, Oxford University Press, 192 p.

Humphreys, P. (2013) What Are Data About? In Durán, Juan & Arnold, Eckhart (eds.): *Computer Simulations and the Changing Face of Experimentation.* Cambridge, Cambridge scholars publishing, 12-28

Humphreys, P. (2014) X-Ray Data and Empirical Content. In Schroeder-Heister, P., Heinzmann, G., Hodges, W., Bour, P. (eds.): Logic, Methodology and Philosophy of Science - Proceedings of the 14th International Congress (Nancy). London, College Publications, 219-234

Imbert, C. (2017) Computer Simulations and Computational Models in Science. In **Magnani**, L. & **Bertolotti, T.** (eds.): *Springer Handbook of Model-Based Science.* Springer, Cham, 735-781

Lehtinen, A. & Kuorikoski, J. (2007) Computing the Perfect Model: Why do Economists Shun Simulation?, *Philosophy of Science* 74: 304-329

Lenhard, J. (2016) Computer Simulation. In Humphreys, P. (ed.): *The Oxford Handbook of Philosophy of Science*. Oxford, Oxford University Press, 717-737

Lusk, G. (2016) Computer Simulation and the Features of Novel Empirical Data. *Studies in History and Philosophy of Science* 56: 145-152

Lusk, G. (forthcoming): Saving the Data. *British Journal for Philosophy of Science*

Massimi, Michela & Bhimji, Wahid (2015) Computer Simulations and Experiments: The Case of Higgs Boson. *Studies in History and Philosophy of Modern Physics* 51: 71-81

Moilanen, A. (1999) Patch Occupancy Models of Metapopulation Dynamics: Efficient Parameter Estimation Using Implicit Statistical Inference. *Ecology* 80: 1031-1043

Morgan, M. (2003) Experiments without Material Intervention. Model Experiments, Virtual Experiments, and Virtually Experiments. In Radder, H. (ed.): *The Philosophy of Scientific Experimentation*. Pittsburgh, University of Pittsburgh Press, 216–233

Morrison, M. (2015) *Reconstructing Reality: Models, Mathematics, and Simulations.* New York, Oxford University Press, 344 p.

Parke, E. C. (2014) Experiments, Simulations, and Epistemic Privilege. *Philosophy of Science* 81: 516-536

Parker, W. S. (2009) Does Matter Really Matter? Computer Simulations, Experiments, and Materiality. *Synthese* 269: 483-496

Parker, W. S. (2016) Reanalyses and Observations: What’s the Difference? *Bulletin of the American Meteorological Society* 97: 1565-1572

Parker, W. S. (2017) Computer Simulation, Measurement, and Data Assimilation. *British Journal for the Philosophy of Science* 68: 273-304

Peschard, I. (forthcoming) Computer Simulation as Substitute for Experimentation? In Vaienti, S. (ed.): *Simulations and Networks*. Hermann, Paris.

Plassmann, F. & Tideman, N. (2014) How Frequently Do Different Voting Rules Encounter Voting Paradoxes in Three-Candidate Elections? *Social Choice and Welfare* 42: 31-75

Reed, W. R. (2015) A Monte Carlo Analysis of Alternative Meta-Analysis Estimators in the Presence of Publication Bias. Economics: The Open-Access, Open-Assessment E-Journal, 9: 1–40. [http://dx.doi.org/10.5018/economics-ejournal.ja.2015-30](http://dx.doi.org/10.5018/economics-ejournal.ja.2015-30%20)

Ringquist, E. J. (2013) *Meta-Analysis for Public Management and Policy.* San Francisco, Jossey-Bass

Roush, S. (2018) The Epistemic Superiority of Experiment to Simulation. *Synthese* 195: 4883-4906

Santer, B., Thorne, P., Haimberger, L., Taylor, K., Wigley, T., Lanzante, J., Solomon, S., Free, M., Gleckler, P., Jones, P., Karl, T., Klein, S., Mears, C., Nychka, D., Schmidt, G., Sherwood, S. & Wentz, F. (2008) Consistency of Modelled and Observed Temperature Trends in the Tropical Troposphere. *International Journal of Climatology* 28: 1703-1722

Stanley, T. D. & Doucouliagos, H. (2016) Neither Fixed nor Random: Weighted Least Squares Meta-Regression. *Research Synthesis Methods* 8: 19-42

Suppes, P. (1962). Models of Data. In Nagel, E., Suppes, P. & Tarski, A. (eds.): *Logic, Methodology and Philosophy of Science*. Stanford, Stanford University Press, 252-261

Tal, E. (2011) From Data to Phenomena and Back Again: Computer-Simulated Signatures. *Synthese* 182: 117-129

Tideman, N. & Plassmann, F. (2014) Which Voting Rule Is Most Likely to Choose the "Best" Candidate? *Public Choice* 158: 331-357

Werndl, C. (2019) Initial Conditions Dependence and Initial Conditions Uncertainty in Climate Science. *British Journal for the Philosophy of Science* 70: 953-976

Winsberg, E. (2009) A Tale of Two Methods. *Synthese* 169, 575-592.

Winsberg, E. (2015) Computer Simulations in Science. In Zalta, Edward (ed.): *The Stanford Encyclopedia of Philosophy.* <https://plato.stanford.edu/entries/simulations-science/> (accessed in March 2019)

Woodward, J. (1989) Data and Phenomena. *Synthese* 79: 393-472

Woodward, J. (2000) Data, Phenomena, and Reliability, *Philosophy of Science* 67: 163-179

Woodward, J. (2011). Data and Phenomena: A Restatement and Defense. *Synthese* 182: 165-179

1. Such problems are relevant in cases in which the difference between the pseudorandom and genuinely random numbers becomes compounded in repeated calculations. [↑](#footnote-ref-1)
2. The data points also concern the same spatiotemporal units. [↑](#footnote-ref-2)
3. This is why we do not aim to take part in the discussion about how to classify data produced by *hybrid methods* (see also Parker 2017, Bueno 2014, Morgan 2003). [↑](#footnote-ref-3)
4. We find it appropriate to refer to the data as *amplified* because the existing properties of the empirical data are replicated but with random variation due to sampling. Their calibration procedure is described in more detail in an unpublished paper: Plassmann and Tideman (2011): “How to Predict the Frequency of Voting Events in Actual Elections”, <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1911286>. [↑](#footnote-ref-4)
5. There is field of research purporting to do this, the study of voting behavior, but it has little to do with normative voting theory. [↑](#footnote-ref-5)
6. Hoover and Salyer (1997) are explicit about not mimicking the DGP: ‘It is not the case that we have created an artificial world in which there are artificial analogues to true technology shocks’ (p. 313). [↑](#footnote-ref-6)
7. A funnel graph typically plots estimates from a set of similar statistical studies such that the standard error (or sample size) and effect size are used as the two axes. [↑](#footnote-ref-7)