Simulation as a Scientific Instrument

Paul S. Andrews[1], Susan Stepney[1], and Jon Timmis[1,2]

[1] Department of Computer Science, University of York, UK
[2] Department of Electronics, University of York, UK

Abstract. Computer simulation approaches are starting to be used more extensively throughout scientific investigations. Some scientists, however, are skeptical about the benefits of simulation. We present computer simulation as a scientific instrument in order to explore issues of their construction and use, which we believe might increase their acceptance within science. We highlight the need to understand the model which the simulation implements, and examine the importance of calibrating simulations and presenting them in an open way to provide scientific reproducibility.

1 Introduction

The use of predictive simulation based approaches to facilitate research in a wide range of scientific disciplines is becoming ever more prevalent in the literature. The acceptance of this trend towards the use of simulation methods, however, is by no means universal with some scientists skeptical regarding the benefits of computer simulation to scientific understanding [6]. We believe it is the responsibility of the people engaged in constructing simulations to provide evidence demonstrating why simulation results can be used to provide real insight into scientific investigations.

In this paper we examine the analogy between computer simulators and a view of scientific instruments expressed by [4] to show how we can make simulation a tool more accessible to science. Computer simulations should be subject to the same rigour that goes into constructing other kinds of scientific instrument. They need to be calibrated to understand how the outputs relate to the system under study, and they should be presented in such a way that their findings can be reproduced.

2 Computer Simulation and Science

In general, Frigg and Reiss [2] state that the term simulation “refers to the entire process of constructing, using, and justifying a model that involves analytically intractable mathematics". Models do not, however, always involve analytically intractable mathematics. For example, simulation is often used where other approaches are intractable owing to ethical reasons (social experiments), cost, time, danger, or impossibility (galaxy formation or climate models). Often, computer simulation is the only way to greater insight into a system [4].

Simulators are built based on an underlying model that is used to represent the system or domain under investigation. Simulation runs (executing the simulator) then allow us to animate that model, exploring its temporal behaviour, hence a "simulation imitates a (usually real) process by another process" [1]. Models are used throughout science as surrogates to learn about the world, revealling features of the system the model represents. Learning takes place during both construction and manipulation of the model that underlies the simulation. In the simplest scenario, model construction results in computer code (the simulator), and model manipulation takes the form of in silico experimentation. During the former we learn about the system and gain an idea of the questions we wish to ask of it; during the latter we explore these questions and enhance our understanding of the model upon which the simulator is based.

Broadly speaking computational methods serve two purposes for scientists: brute-force/informatics approaches deal with large amounts of data or numerical calculations, for example genome sequencing; and predictive simulation aims to to explain observed natural phenomena by capturing the underlying behavioural processes. This paper is concerned with the latter, whereby simulation aims to explain real-world phenomena rather than describe it, and the model underlying the simulation provides a theory for how the components of the system interact to produce a particular outcome [8]. The results of computer simulation can be used for many complementary purposes, for example: to inform real-world experimentation on the system being investigated; to validate such experimentation; or simply to explore both concrete and abstract hypotheses.

One of the main advantages of computer simulation approaches is the complete control of the elements and parameters that make up the simulation. This allows us to explore various aspects that relate to a real-world system under study, which are otherwise difficult or even impossible to achieve. The flexibility of computational approaches, however, can have negative consequences. For instance, the increased access to computational power and reliance on computer simulations may lead to reduced levels of more expensive (but more informative) laboratory or field experiments [4], or overly complex and heavily parameterised models containing poorly understood assumptions [1]. Humphreys [4] also warns of the problems inherent in exploratory agent-based models that aim to show how simple rules can account for complex behaviour. It is possible to use simple models to produce patterns that have little connection to the actual underlying mechanism. In such cases, care must be taken not to blindly accept the explanation given by the computer simulation.

The outputs of simulations will depend on the way in which they have been constructed. This makes understanding the construction process key to interpreting and presenting what the simulation shows. We need to understand the details of the model that the simulation implements, and show how that model then relates to the simulation outputs in order that the outputs can be properly interpreted. This issue is true of most computational approaches in science; for example, Nyce [7] describes how to some radiographers, traditional x-ray images are not representations of the underlying biological structures, but are "very much the same kind of 'thing' " that do not need to be interpreted. This contrasts with their relationship with digital techniques, where a level of mistrust exists because the resultant images are produced via unknown machine operations. This produces a perceived distance between the digital images and the thing they represent.

To increase trust and confidence in computer simulations further, their outputs need to be reproducible. Reproducibility is a key axiom of science, and Timmer [11] reports that the increased reliance on computational methods in most areas of science has lead to an inadvertent loss of scientific reproducibility. The examples given by Timmer [11] for the apparent loss of reproducibility focus on the more data intensive aspects of using computers to analyse large quantities of data. However, many of the issues raised are equally applicable to the simulation techniques used in the Alife community. These issues include: a complex mix of data from both public and internally generated; data that is often quick to become out of date; a complex pipeline of software programs used for analysis, etc; multiple sets of parameters for each piece of software; different software versions; and software bugs.

3 Simulation as a Scientific Instrument

We can consider computer simulation in the same terms as other tools or instruments used by scientists in their scientific endeavours. Given this, such simulations should be subject to the same rigour that goes into constructing other kinds of scientific instrument. They need to be calibrated to understand how the outputs relate to the system under study, and they should be presented in such a way that their findings can be reproduced. We present here a discussion that aims to show that computer simulations fall within a spectrum of instruments used every day by the scientific community.

3.1 Scientific Instruments

The description presented throughout this section has been drawn from the analysis of Humphreys [4], who examines the role of computational models (including simulation) in science.

The role of scientific instruments is to enhance the range of our natural human abilities, such as our perceptual and mathematical capacities. In the case of mathematics, computational devices can be used to move beyond what is accessible naturally to the human brain, such as the number of calculations performed, which is many orders of magnitude greater on a computer.

Instruments of all types have been used for hundreds of years throughout science. These range from everyday instruments such as bench microscopes and optical telescopes to specialised medical imaging equipment. Many modern day instruments incorporate explicit computational approaches in addition to physical detection to provide enhancements. For example, magnetic resonance imaging (MRI) and computerised axial tomography (CAT) scanners contain various physical devices to measure nuclear spin and radiation respectively and then use computer algorithms to transform these readings into two- and three-dimensional images.

One thing in common with all instruments is that they are calibrated to produce outputs that are directly accessible to the human observer. The process of calibration relies on correctly observing and reproducing the structure of known features measured by the instrument. This affords confidence in using the instrument, establishing the scope of its usage along with its accuracy, precision and resolution.

We can often take for granted familiar instruments such as microscopes and telescopes, which are the product of many years of testing, refinement and adjustment. Because of this, the user of such instruments does not need to know the precise details of the theory behind how it works as it has been deliberately designed to be used without need for that knowledge. However, when dealing with contemporary research-level instruments, the user needs to understand the instrument in much greater detail owing to their complexity and the lack of many years of refinement. This type of instrument may routinely malfunction and produce spurious data or need close attention to work in the desired way. Knowing how the instrument works should help reduce the occurrence of unwanted artefacts increasing the instrument's stability and highlighting situations in which malfunctions take place. Another related issue is knowing how to interpret the outputs of the instrument. It is these issues that we tackle when calibrating an instrument. The main benefit of knowing how instruments work is when they provide unexpected outputs, whether this be because something has gone wrong or not. Instrumental knowledge should tell us when we have gone outside the domain of application of the instrument, and how this can be corrected.

3.2 A Spectrum of Instruments

Based on the previous description of scientific instruments, we consider computer simulation as a technique that allows us to develop bespoke scientific instruments. Once engineered, scientific instruments are applied to some object/system of study, which we call the domain. Here we explore how computer simulations relate to their domain of study in the context of other types of scientific instrument. The purpose is to show how the inputs and outputs of a simulation instrument might be applied to understand its intended domain. Conceptually, an instrument takes some form of observation as input from a domain and transforms it based on a model of understanding that has been encoded into the instrument during its construction. The output of the instrument is then presented to a human observer, with the aim of extending the understanding of the domain. Consider three different examples of scientific instruments within the setting just described:

Optical telescope: light is passively received as a direct physical input from the domain (the object that the telescope is directed towards). Lenses are used to refract the light, which is emitted as the output so that the domain appears magnified to the human observer.

MRI scanner: a physical input from the domain is achieved through manipulation. A magnetic field is used to line up proton spins of hydrogen atoms of the intended domain and a radio signal used to disrupt this and measurements are made of the distortion. Computer algorithms are then used to transform the measurements to create an image of the domain that is displayed to the human observer.

Predictive computer simulation: no physical input is received directly from the domain, but a set of derived starting conditions for computational agents is represented within the simulation. These starting conditions then drive the dynamics of the computational model encoded within the simulation, and may be subject to further inputs. A representation of the model is output to a human observer over a period of time.

Whilst each of the three instruments just described attempts to help us understand the domain of study, the relationship with that domain differs. Even though all three instruments are based on models of domain understanding, the way in which these models are encoded differs. For example, it might be within the physical components of the instrument and/or within a computational model. Computer simulations are at the far end of this spectrum, based purely on a computational model with no direct domain understanding specifically encoded within their hardware.

The way in which inputs are received from the domain also differs with scientific instruments. As we move from instruments such as the optical telescope to an MRI scanner, the domain input changes from a passive observation to requiring a direct perturbation of the domain in order to measure its effects. When we move to computer simulation, this domain input becomes far more indirect with regard to space and time in the sense that no direct physical input is present. In this case, we rely more on logical connections to the domain rather than direct physical inputs. This results in an added layer of interpretation required to understand how the inputs of computer simulations map to the entities of the domain under investigation.

In summary, even though different instruments are fulfilling the same role of investigating a domain of study, the way in which instruments interact with that domain can differ immensely. Understanding the relationship between an instrument and the domain it measures is vital to interpreting its output. This is especially true of computer simulation instruments that fall at the far end of a spectrum of instruments, with a reliance on computational models and an indirect relationship to their domain of study.

4 Calibration

We have discussed in the previous section that computer simulations can be viewed as scientific instruments. It follows that simulations should be subject to the same rigorous process of construction as other scientific instruments, generating an understand of how the model upon which the simulation instrument is based relates its inputs to its outputs. In order to help achieve this, the simulation needs to be calibrated.

Construction of a computer simulation, and any other kind of scientific instrument, relies on both processes of science and engineering. Science is employed in the development of the model upon which the instrument will be based. Engineering is then used to implement this model resulting in the construction of an instrument suitable for the purpose of scientific investigation. Before the instrument can be used, however, it should be subjected to the process of calibration. As previously mentioned, calibration involves establishing the relationship between the output of an instrument (across a range of operating conditions) and the system under observation; it lets us interpret what the simulation is showing us.

For instruments such as microscopes and telescopes, Humphreys [4] tells us that the calibration process relies first on correctly observing and reproducing the structure of features that are already known. For computational devices, calibration standards often include reproducing analytically derived reference points. One problem is that due to the complexity of calculations, access to results independent of the simulation is often impossible, thus comparison with existing (instrumental) techniques is required [4].

It is often the case with predictive computer simulation that we do not have access to the types of data typically used to calibrate (for example a set of reference standards). In this case calibration can only be achieved by comparing simulation observations with predictions from a pre-existing and explicitly stated model that formed the basis for simulation construction. One example would be in the case of emergent properties such as flocking. Our model might predict that flocking is the result of a combination of certain agent behaviours. These behaviours would be encoded within the simulation, and then calibration would ascertain whether or not the flocking behaviours are perceived in the actual simulation. This emphasises that need to understand and identify what the underlying model of a simulation is actually a model of.

A further problem with calibration and computer simulation is that all simulators are essentially different specialised instruments that have been constructed to answer a specific question or set of questions. The purpose of these instruments is more often than not different, therefore calibration is going to be different for each individual simulator. When changes are made to the simulator (no matter how small), it may have to be re-calibrated depending on how the change affects the encoded behaviours.

5 Openness

Previously we highlighted the need for the results generated by computer simulations to be reproducible. Whilst proper calibration would be a first step towards this, we need to be open about further aspects of the simulation. Timmer [11] describes the beginning of a movement towards researchers adopting approaches to ensure that computational tools are in line with existing scientific methods. However, whilst there may be a recognition that nearly everyone doing science uses some form of computation, there are few who know what is needed to make sure that documentation of approaches is sufficient for reproducibility.

It is probably intractable to expect complete reproducibility of a piece of science performed using computer simulation without access to the exact piece of computer code and all the initialising variables (parameter settings, initial states of data). This gives support to the argument for complete openness of code, an emotive issue amongst many who develop computational tools. This can only be tackled by a sea-change within research communities to require this level of openness.

There is, however, a more subtle and no less important source of knowledge that should also be open. We have previously mentioned that all simulations encode a model. This model is often only explicitly expressed as the computer code and is the process of much work. It contains many different assumptions which are vital to understanding what the model represents. This is an issue of validation (see [10]): how do you know that you have built the right system to answer the questions you are exploring? This is hard to express as a yes/no answer and it typically expressed as a level of confidence. In some circumstances, for example where outputs of a simulation instrument have a high level of criticality, that a structured argument is required to express confidence in a computer simulation (see [9] for a more in depth discussion).

6 Conclusion

There are many reasons why scientific instruments based on computer simulation might not be accepted for use in scientific investigations. In this paper we have suggested that if we can show how simulations relate to more traditional scientific instruments, and highlight some important issues regarding how they might be constructed and presented, that we may stand a greater chance of simulation approaches becoming a useful instrument for science. It is also important to emphasise that simulation does not replace direct experimentation; simulation is a tool to assist more traditional approaches. Simulation should be part of the scientists toolkit and used where it is appropriate.

Humphreys [4] argues that when we use new instruments we need to understand how they are built. Over time they can become more generally accepted: in the case of simulation the ‘acceptance’ is going to be the acceptance of a class of predictive software artefacts and development strategies rather than of an instance of one. This general acceptance is only going to come from an increased number of instance acceptances.

In summary, we can consider general purpose computers as physical instruments that can be used to construct a wide variety of logical instruments in the form of simulations. However, as previously discussed, the benefit of computational approaches conceals its drawback, with many unknowns liable to populate the simulation. In the aftermath of 'Climategate' [3], there should be greater scrutiny on the way in which scientists use computational devices as part of their scientific process [5]. There have been calls for open-source code to enable repeatability. In our view, this is not enough: we need to perform calibration and present the results of this calibration to provide us with the knowledge to decode the output of simulation and interpret in the context of the real domain being modelled and simulated. We need to show how the simulation has been engineered and why it is a good instrument to enhance our domain knowledge.

7 Acknowledgements

This work is part of the CoSMoS project (www.cosmos-research.org), funded by EPSRC grants EP/E053505/1 and EP/E049419/1.

References

[1] R. Frigg and S. Hartmann. Models in science. In E. Zalta, editor, *Stanford Encyclopedia of Philosophy*. 2006.

[2] Roman Frigg and Julian Reiss. The philosophy of simulation: hot new issues or same old stew? *Synthese*, 169:593–613, 2009.

[3] Roger Harrabin. Climate science must be more open, say MPs. 2010. http://news.bbc.co.uk/1/hi/sci/tech/8595483.stm.

[4] Paul Humphreys. *Extending Ourselves: Computational Science, Empiricism, and Scientific Method*. Oxford University Press, New York, 2004.

[5] Darrel C. Ince, Leslie Hatton, and John Graham-Cumming. The case for open computer programs. *Nature*, (482):485–488, 2012.

[6] Lucas Laursen. Biological logic. *Nature*, 462(26):408–410, 2009.

[7] James M. Nyce. Artifice, interpretation and nature: Key categories in radiology work. In *Unconventional Computation, 8th International Conference, UC 2009*, volume 5715 of *Lecture Notes in Computer Science*, pages 11–15, 2009.

[8] Steven L. Peck. Simulation as experiment: a philosophical reassessment for biological modeling. *Trends in Ecology and Evolution*, 19(10):530–534, 2004.

[9] Fiona A.C. Polack, Tim Hoverd, Adam T. Sampson, Susan Stepney, and Jon Timmis. Complex systems models: engineering simulations. In S. Bullock, J. Noble, R. Watson, and M. A. Bedau, editors, *Eleventh International Conference on the Simulation and Synthesis of Living Systems*, pages 482–489. MIT Press, Cambridge, MA, 2008.

[10] R. G. Sargent. Verification and validation of simulation models. In *37th Winter Simulation Conference*, pages 130–143. ACM, 2005.
[11] John Timmer. Keeping computers from ending science's reproducibility. 2010. http://arstechnica.com/science/news/2010/01/keeping-computers-from-ending-sciences-reproducibility.ars.