Modelling Practices and Practices of Modelling

Fotia, Giorgio

Modelling Practices and Practices of Modelling

Giorgio Fotia *

Abstract: *»Modellierungspraktiken und Praktiken der Modellierung«*. Modelling represents a core method of investigation in the sciences. Relying on a number of case studies, I want to explore the main concepts that denote the practice of modeling in pure and applied sciences. I argue that these concepts could be seen as metaphors to reflect upon when exploring how the practices of modeling are characterised across different disciplines.

Keywords: Mathematical models, numerical models, computational science, scientific discovery, data-science.

1. Introduction

Modelling is pervasive in the sciences, where it represents a core method of investigation as well as a subject of research *per se*. This paper considers some concepts that characterise the practice of modeling in pure and applied science. I argue that these concepts could be useful metaphors in trying to understand how models are used to investigate or represent reality.

2. Topics in Modelling

There are many examples of how mathematical models are exploited in science and engineering. Effectively, J. T. Oden points out that mathematical models "provide the vehicle with which precision is given to theory and to the mental processes used to establish and perpetuate what is known in science and engineering" (Oden 2002, 13).

Use cases include, among others, fluid dynamics and turbulence, wave propagation in complex layered media, and the prediction of the behaviour of complex engineered systems. With reference to this latter application area, in Fig. 1 some results of an integrity analysis simulation for a component of the Large Hadron Collider (LHC) at CERN in Geneva are shown. The numerical

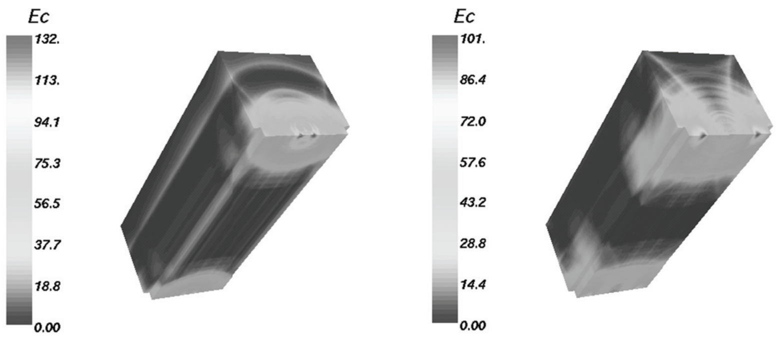
* Giorgio Fotia, CRS4 Biosciences, CRS4, Center for Advanced Studies, Research and Development in Sardinia, Polaris Edificio 1, 09010 Pula (Cagliari), Italy; Giorgio.Fotia@crs4.it.

Historical Social Research Supplement 31 (2018), 147-153 │ published by GESIS
DOI: 10.12759/hsr.suppl.31.2018.147-153

computations account for the coupled thermal and structural responses of the component to the loads due to high energy proton beam deposition. The results of numerous simulations of the baseline design (not shown here) demonstrated a possible lack of structural integrity under specific conditions in some of the components. For the sake of optimisation, the design was revised accordingly — a new material was selected that would guarantee their structural integrity where necessary, and modifications of the geometry were adopted to optimise the performance of the whole system.

From this example and from many others as well[1], it is apparent that the use of mathematical models for the simulation of complex physical phenomena as opposed to the use of experiments has nowadays emerged as a novel part of the scientific method in addition to theory and observation.

Figure 1: Computer Simulation of the Structural Thermoelastic Response of an LHC Component Subject to Extreme Loads



Computed distribution of kinetic energy at time t = 12μs (left) and t = 16μ (right). Calculation carried out at CRS4, (L. Bruno, G. Fotia, Y. Kadi, F. Maggio, L. Massidda, and F. Mura, 1998–2004).

Yet numerical (or computational) models are necessarily tied to experimental data, and this suggests that we should explore how these models are validated and in some cases improved by the use of experimental data. For example, A. Quarteroni notes that

> I am not aware of any numerical method which can produce valid and accurate results on a variety of complex problems without exploiting to a certain extent information which are provided by experimental measurements (Quarteroni 2004, 5).

[1] See Stein, de Borst, and Hughes (2017) for a comprehensive presentation of the subject and an exhaustive overview of the recent developments in this area of research.

This remark suggests that the concept of reliability of predictions is essential to the practice of computational modelling. In fact, it is generally agreed that the reliability of a computer simulation (i.e the measure of confidence that can be assigned to it)

> depends upon the assessment and control of errors inevitable in the computational process – modeling errors due to the impossibility of capturing all of nature with mathematical abstraction and approximation error due to the impossibility of solving exactly the mathematical models. (Oden 2002, 15)

Another concept that is worth further methodological investigations is the use of computing as an heuristic tool (Lax 1999, 24-8 passim). In a recent interview, P. D. Lax reports that John Von Neumann already realised in 1945 that computing, that is the practice of running a numerical model on a computer,

> [...] gives us those hints without which any progress is unattainable, what the phenomena are that we are looking for." In other words, computing may be used not only for solving concrete problems "but rather to explore which way science should be developed (Lax 2004, 3).

Not surprisingly, computational models are now well established as a tool for theoretical investigation in science, (see e.g. McCurdy et al. 2002). For example, massive computation of turbulence – performed by solving the exact equations of hydrodynamic turbulence – have provided new quantitative data and enhanced the understanding of this area (see e.g. Yeung, Zhai, and Sreenivasan 2015). This suggests to us that it is worthwhile to further investigate the concept of computing as an instrument for discovery in the sciences and to understand how computational models are used in this endeavour.

Further opportunities to reflect upon the way to describe the practice of modeling in the sciences are provided by examining how these practices are viewed from the perspective of the now emerging data-driven science.

In recent years, in fact, data driven science has emerged as a novel framework, due to recent developments in the technology of experimentation and measurement. This trend has forced scientists to change their attitudes toward data, and data methods are leading to transformative changes across the engineering, physical and biological sciences as well as the social sciences. Striking examples include, among many others, data-intensive computing systems (Mattmann 2014), genomics and systems biology (Stephens et al. 2015), medicine and health (Rotmensch 2017), urban informatics (Ota et al. 2016, Zhao et al. 2016), political and social sciences (Alvarez 2016), and social media and computer-mediated communication (Olshannikova et al. 2017, Barberá 2015).

Indeed, in most cases the term *data-driven modeling* seems to have relevance as opposed to a-posteriori validation (or optimisation) of models (Efron 2016, Hastie 2015). But this is not necessarily true. Interestingly enough, the use of such rich data sets has been recently proposed in conjunction with more traditional analysis, modeling, and computation. In fact, on one hand, numerical simulation of large complex systems can easily strain available computational resources. Similarly, experiments can generate overwhelming amounts of data. On the other hand, recent advances in data-driven methods are allowing for significant advances in the prediction and control of highly complex, often networked, systems.

This approach, known as Dynamic Data-Driven Applications Systems (DDAS), represents an emerging paradigm in computational science, in which "simulations and experiments (or field data) interact in real time to dramatically improve the fidelity of the simulation tool, its accuracy, and its reliability" (NSF 2006, 37). Homeland security, control of hazardous materials, environmental remediation, manufacturing processes, and vehicle flight control are just a few of the recent applications of this technology (Darema 2004). As recently pointed out (Kuske at al. 2017), it is expected that the combination of these approaches may provide a transformative mathematical framework for modeling the behaviour of complex systems. This interesting remark suggests that it might be useful to investigate how, and in what sense, these different modelling techniques are interconnected.

To sum up, there are a number of concepts that characterise the practice of modeling in the sciences. What emerges is a practice-based overview of what modelling in the context of different domains of applied sciences means in operational terms and a glimpse of what could be entailed if the practice of modelling is analysed from this viewpoint.

3. Conclusion

In this paper, I attempted to explore the concept of modeling in the context of different domains of applied sciences from the perspective of modeling as a practice. However, there are a number of issues that would need further analysis. For instance, it would be useful to further explore the relationship between theory and practice that emerges as a consequence of resorting to computations for discovery in the pure sciences.

Further opportunities to reflect upon the way to describe the practice of modeling in the sciences can be elicited by examining these concepts from the perspective of the emerging domain of data-driven science. To this end it would be important, for example, to compare a number of use cases across different disciplines.

Exploring how practices of modeling are characterised across disciplines seems to me a promising way to examine how, and in what sense, practices of modeling are interconnected, and whether and how the concept of modelling in the sciences can be appropriately (re)defined.

4. Discussion

Paul Fishwick' questions
Paul Fishwick was the respondent to my talk. While agreeing that computation is undoubtedly a pillar of contemporary science, Paul pointed out that models can also be seen as ways of physically encoding information using a specific technology, with associated analogies and metaphors. As such, they can be considered to be informational representations of our world. He considers that, if one wants to characterise exhaustively the practice of modeling, diagrammatic and physical representations, and mathematical notation, should be considered as well, and he asked me to comment on this issue. Another question he raised was about the potential connections between the concepts and the practices I discussed and the arts and humanities.

My answer
I consider that while information representation may be part of the effort of building a computational model, whether their role is essential or not is strongly dependent on the particular goal of the model building process and of the application problem one may want to solve. However, I do agree with Paul in considering that these representations should be taken into account if one wants to unravel how modeling is used in practice. As far as the potential connections between the concepts and the practices I discussed and the arts and humanities, we both agreed that data science can provide the appropriate framework for non-traditional research and discovery in the humanities. In this same framework, we posit that the concept of computing as an instrument for discovery in the sciences I described can be a useful metaphor to reflect upon when trying to unify the description of the practices of modeling in many different domains, both in science and in the humanities.

References

Alvarez, Michael, ed. 2016 *Computational social science. Discovery and prediction*. New York, NY: Cambridge University Press.

Barberá, Pablo. 2015. Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data. *Political Analysis* 23 (1): 76-91. doi: 10.1093/pan/mpu011.

Darema, Frederica. 2004. Dynamic Data Driven Applications Systems: A New Paradigm for Application Simulations and Measurements. In *Computational Science - ICCS 2004, Bubak, M., van Albada, G.D., Sloot, P.M.A., Dongarra, J.* Lecture Notes in Computer Science, vol. 3038. Springer, Berlin, Heidelberg. doi: 10.1007/978-3-540-24688-6_86

Efron, Bradley, Hastie, Trevor. 2016. *Computer Age Statistical Inference: Algorithms, Evidence, and Data Science*. Cambridge University Press.

Hastie, Trevor. 2015. Statistical Learning with Big Data [Slides]. Stanford, October 21. Retrieved from <https://web.stanford.edu/~hastie/lectures.htm> (Accessed March 23, 2018).

Kuske, Rachel, Nathan J. Kutz, and Karen Willcox. 2017. Workshop on Data-Driven Methods for Reduced-Order Modeling and Stochastic Partial Differential Equations; Banff International Research Station, Canada, January 29-February 3, 2017. <https://www.birs.ca/events/2017/5-day-workshops/17w5140> (Accessed March 23, 2016).

Lax, Peter D. 1999. Mathematics and Computing. In *Useful Knowledge: The American Philosophical Society Millennium Program*, ed. Alexander G. Bearn. Memoirs of the American Philosophical Society Held at Philadelphia for Promoting Useful Knowledge, vol. 234, 23-44. Philadelphia: American Philosophical Society.

Lax, Peter D. 2004. Oral history interview by Philip Colella, 14 November, 2003, UCLA and 26 April, 2004, Courant Institute. Society for Industrial and Applied Mathematics, Philadelphia, PA. Available at <http://history.siam.org/pdfs2/Lax_final.pdf> (Accessed December 12, 2016).

Mattmann, Chris A. 2014. *Cultivating a research agenda for data science*. J. Big Data 1, 6. doi:10.1186/2196-1115-1-6

NSF. 2006. *Simulation-based engineering science – revolutionizing engineering science through simulation.* A Report of the National Science Foundation Blue Ribbon Panel on Simulation-Based Engineering Science. Available at <https://www.nsf.gov/pubs/reports/sbes_final_report.pdf> (Accessed December 12, 2016).

McCurdy Bill, Robert Sugar, Marsha Berger, David Ceperly, Steve Koonin, Jeremiah Ostriker, Dan Reed, John Wilkins, and Paul Woodward. 2002. Computation as a Tool for Discovery in Physics. Report to the National Science Foundation by the Steering Committee on Computational Physics. Available at <https://www.nsf.gov/pubs/2002/nsf02176/start.htm> (Accessed December 12, 2016).

Oden, Tinsley J. 2002. The Promise of Computational Engineering & Science: Will it be kept? *IACM Expressions* 12 (02): 12-5.

Olshannikova, Ekaterina, Thomas Olsson, Jukka Huhtamäki, Hannu Kärkkäinen. 2017 Conceptualizing Big Social Data. J. Big Data 4, 3. doi:10.1186/s40537-017-0063-x.

Ota, Masayo, Huy Vo, Claudio Silva, and Juliana Freire. 2017. STaRS: Simulating Taxi Ride Sharing at Scale. In *IEEE Transactions on Big Data* 3 (3): 349-61, Sept. 1 2017. doi: 10.1109/TBDATA.2016.2627223.

Quarteroni, Alfio. 2004. Oral history interview by Philip Davis, 22 June. Brown University. Society for Industrial and Applied Mathematics, Philadelphia, PA. Available at <http://history.siam.org/pdfs2/Quarteroni_final.pdf> (Accessed December 12, 2016).

Rotmensch, Maya, Yoni Halpern, Abdulhakim Tlimat, Steven Horng, and David Sontag. 2017. Learning a Health Knowledge Graph from Electronic Medical Records, Nature Scientific Reports, July. doi: 10.1038/s41598-017-05778-z.

Stein, Erwin, René de Borst and Thomas J. R. Hughes, eds. 2017. *Encyclopedia of Computational Mechanics*, 2nd ed. vols. 1 & 2 (Fundamentals); vols 3 & 4 (Solids and Structures) vols. 5 & 6 (Fluids).

Stephens Zachary D., Skylar Y. Lee, Faraz Faghri, Roy H. Campbell, Chengxiang Zhai, Miles J. Efron, Ravishankar Iyer, and Michael C. Schatz 2015. *Big Data: Astronomical or Genomical? PLoS Biol13* (7): e1002195. doi: 10.1371/journal. pbio.1002195.

Yeung Pui-Kuen, Xiameng Zhai and Katepalli Sreenivasan. 2015. Extreme events in computational turbulence. *PNAS* 112 (41): 12633-18. doi: 10.1073/pnas. 1517368112.

Zhao, Khai, Denis Khryashchev, Juliana Freire, Claudia Silva, and Huy Vo. 2016. Predicting taxi demand at high spatial resolution: Approaching the limit of predictability. In *IEEE International Conference on Big Data (Big Data)*, Washington, DC, 833-42. doi: 10.1109/BigData.2016.7840676.

Historical Social Research
Historische Sozialforschung

All articles published in HSR Supplement 31 (2018): Models and Modelling between Digital Humanities – A Multidisciplinary Perspective