Is Computational Biology a Reliable Science?

Hans Ekkehard Plesser[1,2,3,*], Eilen Nordlie[1], and Marc-Oliver Gewaltig[4]

[1] Department of Mathematical Sciences and Technology, Norwegian University of Life Sciences, Ås, Norway

[2] Center for Biomedical Computing, Simula Research Laboratory, Lysaker, Norway

[3] RIKEN Brain Science Institute, Wako-shi, Saitama, Japan

[4] Honda Research Institute Europe GmbH, Offenbach, Germany

[*] Corresponding author; e-mail: hans.ekkehard.plesser@umb.no.

October 10, 2009

Systems biology, neuroscience and other disciplines in biology rely increasingly on computer simulations to study the properties of models of biological systems, ranging from molecular processes via whole organs to the complete human physiome. This raises the issue of how reliable these computer simulations are as a scientific tool. In our own field, computational neuroscience, most colleagues agree that published results based on simulations cannot be replicated in many cases, often not even by their original authors. Scientists in other fields have come to similar conclusions: "The vast body of results being generated by current computational science practice suffer a large and growing credibility gap: it is impossible to verify most of the computational results shown in conferences and papers" [1]. In at least one case, and innocent software error forced the retraction of several high-profile publications [2], and several authors have pointed out that many computational scientists lack even basic software engineering skills [3, 4].

In response to this reliability crisis, several groups have proposed rather technical solutions, namely that all scientist in a field should use the same, thoroughly tested software [1], and create systems for the seamless documentation of scientific workflow from the script specifying a simulation model to the final figure appearing in a journal, and should share the source code of their simulations [5].

In our view, this addresses only the *replicability* of simulations [6]. While important, reforms that address replicability only, fall short of the thorough appraisal of models and simulations that computational science needs in our view. We believe that computational biologists need to focus on *reproducibility* as least as much as on replicability [6]. *Replicability* is the ability to repeatedly—also after a long period of time—generate identical results from a simulation. Trivial as this may sound, it is challenging in practice, as simulations scripts and software are under constant development, and few scientists diligently record and preserve *all* versions of their code.

Reproducibility, in contrast, entails that Bob can read Alice's paper, build his own mental model, implement it independently, and obtain results in agreement with Alice's. Only if Bob successfully—and independently—reproduces Alice's result, can the scientific community be assured that the model described in Alice's paper is indeed complete, in the sense that the description contains all information required to yield the reported results. We are aware of at least one rather well-known paper in computational neuroscience in which the results depended crucially on what the authors apparently considered an insignificant implementation detail. This fact went unnoticed for close to a decade.

Computational biology will benefit in at least three ways from an increased focus on independent reproducibility: (i) It will force authors to describe their models more precisely [7] and to reflect more carefully about the borders between properties of their scientific models and details of the computer implementation. This will make models more meaningful. (ii) Scientist will need to engineer, test, and validate their simulation and data-analysis software more carefully than today, where even irreplicability is shrugged off as one of the dirty secrets of the community. This in turn will make simulation results more reliable. (iii) Independent implementations of the same model will almost necessarily yield different results, as different numerical methods, random number sources, etc, are used. Scientists will thus have to reflect carefully upon the requirements for stating that two implementations of a model are in agreement (or are not). This reflection may lead to a deeper understanding of the predictive power of the model, and might even lead to new discoveries: Erwin Schrödinger first developed the theory of first-passage times in order to resolve conflicting evidence on Millikan's oil-drop experiment for measuring the elementary charge [8].

References

[1] D. L. Donoho, A. Maleki, I. U. Rahman, M. Shahram, and V. Stodden. 15 years of reproducible research in computational harmonic analysis. *Computing in Science & Engineering*, 11(1):8–18, 2009. doi: 10.1109/MCSE.2009.15.

[2] Greg Miller. Scientific publishing. a scientist's nightmare: software problem leads to five retractions. *Science*, 314(5807):1856–1857, Dec 2006. doi: 10.1126/science.314.5807.1856.

[3] Gregory V. Wilson. Where's the real bottleneck in scientific computing? *American Scientist*, 94(1):5–6, 2006. doi: 10.1511/2006.1.5.

[4] Jo Erskine Hannay, Hans Petter Langtangen, Dietmar Pfahl, Janice Singer, and Greg Wilson. A survey of how scientists develop software. In *Proceedings of ESEM2009*, 2009.

[5] Patrick Vandewalle, Jelena Kovačević, and Martin Vetterli. Reproducible research in signal processing. *IEEE Signal Proc Mag*, 26(3):37–47, 2009. doi: 10.1109/MSP.2009.932122.

[6] Chris Drummond. Replicability is not reproducibility: Nor is it good science. In *Proceedings of the Evaluation Methods for Machine Learning Workshop at the* 26^{th} *ICML*, Montreal, CA, 2009. URL `http://www.site.uottawa.ca/~cdrummon/pubs/ICMLws09.pdf`.

[7] Eilen Nordlie, Marc-Oliver Gewaltig, and Hans Ekkehard Plesser. Towards reproducible descriptions of neuronal network models. *PLoS Comput Biol*, 5(8):e1000456, Aug 2009. doi: 10.1371/journal.pcbi.1000456.

[8] E. Schrödinger. Zur Theorie der Fall- und Steigversuche an Teilchen mit Brownscher Bewegung. *Physikalische Zeitschrift*, 16:289–295, 1915.