The limitations of model-based experimental design in sloppy systems

Andrew White[1], Malachi Tolman[1], Mark K. Transtrum[1]

Complex models in systems biology can include a large number of unknown parameters. Many such models are "sloppy", i.e., exhibit an extreme insensitivity to coordinated changes in many parameter combinations. Because of this extreme insensitivity experimental design methods have been developed to optimally select those experiments that allow accurate estimates of all the parameters. These methods typically assume that the model is a complete representation of the system. In practice, however, this is assumption is almost never true--models always involve simplifying approximations. We explore the effects of these approximations on model-based experimental design methods. We conduct several numerical experiments in which data is generated from a complex model (acting as a surrogate for the actual system) but experiments are selected based on an approximate model. We find that although the simple model is able to fit data generated by the complex model for many potential experiments, it is unable to fit data for those experiments selected as "optimal" as determined by experimental design methods. This is because the "optimal" experiments are those most likely to make microscopic details more important, including those omitted from the model.

MODELS of complex biological systems can involve a large number of unknown parameters. Considerable attention has been given to the problem of parameter inference in systems biology. Many models are "sloppy," i.e., exhibit an extreme insensitivity to coordinated changes in the parameters. Because of the near-universal appearance of sloppiness among systems biology models it was suggested that sloppiness was an inherent feature of such models and that accurate parameter inference would be practically impossible [1]. Subsequently, it was shown that model-based experimental design could be used to identify a collection of experiments that would enable accurate parameter estimates. The idea was that although the model of each experiment would be sloppy individually, complementary experiments could be identified that would allow the accurate estimates of all the parameters [2].

There has been considerable interest in experimental design techniques for parameter inference in systems biology [2, 3, 4, 5]. However, nearly all of these methods assume that the model is a complete representation of the system. In practice, however, this assumption is almost never true. Models always employ simplifying. Indeed, it would be hard to imagine a "complete" model of systems biology. Any model will have rates and binding affinities that will be altered by the surrounding complex stew of proteins, ions, lipids, and cellular substructures. Furthermore, in such systems there is no clear distinction between which parameters are important and which are not.

We consider the model of EGFR signaling due to Brown et al. [6] for which optimal experiments were later identified by Apgar et al., [2]. We generate data for the Apgar et al. experiments using a model similar to that of Brown et al. but with Michaelis-Menten reaction replaced by the more accurate mass-action reactions. We find that although both models can fit the data for the experiments in Brown et al. [6], the Michaelis-Menten model is unable to fit data generated by the mass action model for the experiments proposed by Apgar et al.

We argue that this result will be generic for systems in which there is no clear separation between important and unimportant system features. We describe such systems as "sloppy," a natural extension of "sloppy" models. Optimal experiments are those that highlight features of the system that were unimportant for other experiments. This includes those components of the system that were omitted from the model. When these experiments are carried out, the model will typically be unable to fit the resulting data. Our results suggest that more careful uncertainty quantification is necessary when modeling and selecting experiments for such systems.

REFERENCES

[1] Gutenkunst RN, et al. (2007) Universally Sloppy Parameter Sensitivities in Systems Biology Models. *PLoS Comput Biol* 3(10): e189. doi:10.1371/journal.pcbi.0030189

[2] Apgar, Joshua F., et al. Sloppy models, parameter uncertainty, and the role of experimental design. *Molecular BioSystems* 6.10 (2010): 1890-1900.

[3] Chachra, Ricky, Mark K. Transtrum, and James P. Sethna. "Comment on "Sloppy models, parameter uncertainty, and the role of experimental design"."*Molecular BioSystems* 7.8 (2011): 2522-2522.

[4] Balsa-Canto, Eva, Antonio A. Alonso, and Julio R. Banga. "Computational procedures for optimal experimental design in biological systems." *IET systems biology* 2.4 (2008): 163-172.

[5] Transtrum, Mark K., and Peng Qiu. "Optimal experiment selection for parameter estimation in biological differential equation models." *BMC bioinformatics* 13.1 (2012): 181.

[6] Brown, Kevin S., et al. "The statistical mechanics of complex signaling networks: nerve growth factor signaling." *Physical biology* 1.3 (2004): 184.

[1]Department of Physics & Astronomy, Brigham Young University, Provo Utah. E-mail: mktranstrum@byu.edu