Exploiting the exciting interface between mechanistic and statistical models

Matthew P. Adams [a,b]

[a] *School of Mathematical Sciences, ARC Centre of Excellence for Mathematical and Statistical Frontiers, and the Centre for Data Science, Queensland University of Technology, Brisbane QLD Australia*
[b] *School of Chemical Engineering, The University of Queensland, St Lucia QLD Australia*

Email: mp.adams@qut.edu.au

Abstract: When modelling real-world phenomena, there may be a flood (or a drought!) of data accompanied by a cloud of uncertainty. But where does this uncertainty come from? The uncertainty could be irreducible, and thus our best bet is to use the available data to devise a statistical model. We may, given time, be able to obtain process-based understanding of the system, and then our best bet might be to devise a (usually complicated!) mechanistic model.

However, the true pathway in many cases lies somewhere in between. This creates tension for mathematical modellers who may typically have been trained in one of two camps: (1) a process-based world that hinges on a physical understanding of the system, represented by some combination of ordinary differential equations, partial differential equations, and other physical formulas, or (2) a world built on correlations or formulas with little physical underpinnings but provide a solid foundation to produce meaningful predictions quickly. Research in the latter camp has exploded recently with the expansion of machine learning and artificial intelligence techniques, but is also beginning to introduce regulations that ensure quality control in the application of these techniques (e.g. responsible artificial intelligence). Mechanistic and statistical models can learn much from each other. For example, process-based models can be used to sense-check the predictions of statistical models, and statistical models can emulate process-based models for rapid gains in computational efficiency. Strategic model hybridization from these two different approaches may yield a balance between mechanistic insights and impactful outcomes.

This talk explores research work which bridges the divide between mechanistic and statistical models. Such work typically requiring cross-training at the postgraduate or postdoctoral level and is a potential route to take best advantage of the strengths of both modelling strategies. Analogous to the broader tension felt between fundamental and applied research, hybrid modelling strategies may also have the potential to balance the two scientific progress goals of mechanistic insights and practical predictions. I will present some of the lessons my colleagues and I have learned about navigating this mechanistic/statistical interface with examples of surprising and unexpected outcomes we have seen in the fields of biology, ecology, agriculture and elsewhere. These examples include applications where:

- Statistical techniques can inform strategic simplification of mechanistic models,
- Mechanistic and statistical models have equal predictive power, but each has a preferred usage depending on the model's goal,
- Emulation of a mechanistic model will be utterly essential to practically inform decision-making, and
- Statistical assumptions vastly alter conclusions about parameter identifiability in both model types.

Keywords: *Mechanistic models, modelling approaches, predictions, scientific impact, statistical models*