Automated calibration of agent-based immunological simulations

Mark Read[1,*], Magnus Tripp[2], Hannah Leonova[3], Louis M. Rose[2] and Jon Timmis[1,2]

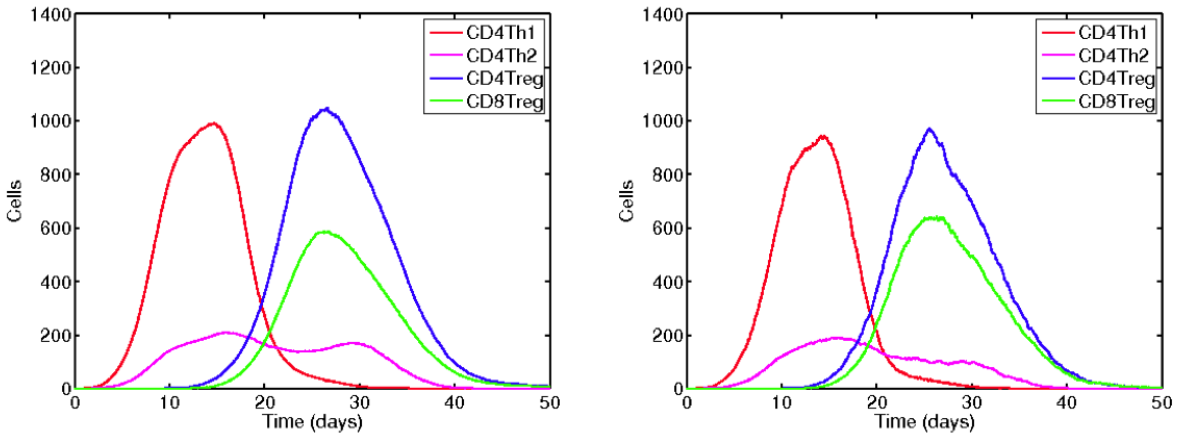
[1]Department of Electronics, the University of York, UK.
[2]Department of Computer Science, the University of York, UK.
[3]Department of Chemical and Biological Engineering, the University of Sheffield, UK.
HL is now at the University of Sheffield; her contribution to this work was conducted at York.
*mark.read@york.ac.uk

The emerging field of computational immunology shows great promise to advance immunological research. Simulations of immunological systems provide a platform for *in silico* experimentation, facilitating the formulation and evaluation of hypotheses.

A major challenge in this field, however, lies in parameterization, particularly in agent-based simulations. These simulations contain many parameters (>50 is plausible), many pertaining to aspects of immunology that either have not or cannot be examined with current wet-lab technologies. Curve-fitting (e.g. linear regression) based-calibration is tractable only for relatively simple simulations with few parameters, and will not necessarily lead to biologically-plausible parameter values (e.g., if the model is a bad representation of the biology). For larger systems the current state of the art is to calibrate by hand/eye, with some values based on wet-lab data or expert opinion, and the rest on trial and error. Furthermore, it is typical to calibrate simulations against data from only a single wet-lab experiment. Although these data may comprise observations of multiple cells/molecules/disease scores (termed *responses*), given that a simulation is likely to be used to perform *multiple* novel experiments that have not been attempted in the wet-lab, it still constitutes calibration against a single datapoint (single experiment). Put another way, with so many degrees of freedom there may be multiple points in parameter space for which a simulation re-creates data from a single experiment; a simulation calibrated in such a manner will not necessarily be representative of the biology when used for a different experiment. We propose that to have genuine trust that they reliably capture the biology, immunological simulations should be calibrated against multiple wet-lab experiments.

Performing calibration of this magnitude by hand is intractable, and as such we are investigating alternatives based on automated multi-objective meta-heuristic search techniques. Each wet-lab experiment used in calibration is performed also in simulation. Each response from each experiment is treated as an individual objective that the search algorithm must align simulation behaviour with. The search algorithm searches for parameter values that satisfy all these constraints. We are developing this methodology by calibrating ARTIMMUS, an existing simulation of the murine autoimmune disease Experimental Autoimmune Encephalomyelitis (Read, 2011), using NSGA-II (K Deb, 2002), a multi-objective search technique. ARTIMMUSs hand-calibrated parameters have been shown to reflect the *in vivo* disease dynamics of EAE (Read, 2011; Mark Read, 2012). ARTIMMUS comprises around 70 parameters, representing a very large search space. In demonstrating proof of principle of this technique we first calibrate a restricted set of 8 key parameters (shown in figure 1c; all other parameters retain their previously calibrated values, see Read (2011)), gradually increasing the range of permitted values over which the search process may operate, and the number of objectives.

Results are promising: figure 1a depicts previously calibrated simulation behaviour (left), and NSGA-IIs best attempt at recreating it (right) (Tripp, 2013). The absolute difference between target (hand-calibrated) values and those obtained by NSGA-II are shown in figure 1b. This experiment represents the most difficult problem setup that was attempted. NSGA-II operated on 6 objectives, attempting to match the peak numbers of CD4Th1, CD4Treg, CD8Treg, the times at which these peaks occurred for CD4Th1 and CD8Treg, and the number of CD4Th1 cells remaining at 40 days. The difficulty of this search problem must be emphasized, the 8 parameters being calibrated constitutes a substantial search space, and 6 objectives is a lot for a multi-objective optimization algorithm such as this. Nonetheless, the results are promising and further investigation, in particular incorporating dynamics from a second experiment, is warranted. Interestingly, the search process highlighted how disparate areas of search space could provide seemingly well-aligned behaviours (termed *local optima*) when calibrating against this one experiment. This highlights the importance of more powerful calibration techniques: the existence of multiple local optima in immunological simulations when calibrating against single experiments is problematic for reasons outlined above. Furthermore, if state-of-



(a) Effector T cell population sizes within the simulation over time. Left, the hand-calibrated simulation dynamics. Right, the result of NSGA-II calibrating ARTIMMUS parameters and recreating hand-calibrated dynamics

response	CD4Th1P	CD4TregP	CD8TregP	CD4Th1T	CD8TregT	CD4Th140
abs difference	47.0	76.0	55.0	5.0	0.0	0.0

(b) The absolute difference in values between hand calibrated response values, and those resulting from NSGA-II. Response names ending in 'P' denote peak population sizes, 'T' denotes the time at which these peaks occur.

Parameter	hand	NSGA-II (183)	range	NSGA-LA (261)
CD4Th	40	30	0 - 100	43
CD4Treg	30	38	0 - 90	55
CD8Treg	30	47	0 - 90	50
Neurons	500	527	440 - 560	460
Microglia	75	106	15 - 135	117
DCs in LN	10	13	0 - 70	43
DCs in CNS	40	59	0 - 100	40
DCs in Spleen	100	72	40 - 160	109

(c) The parameters over which NGSA-II performs optimization. Hand calibrated and NSGA-II parameter values are given, as are the ranges of values over which NSGA-II operated. The best solution found is labelled 'NSGA-II'. DC, dendritic cells; LN, lymph node; CNS, central nervous system. A sub-optimal, but good, result from NSGA-II (labelled 'NSGA-LA') is also given. Fitnesses are shown in parentheses, and represent the sum of absolute differences between hand calibrated and NSGA-optimised response values. Lower fitnesses represent better solutions.

Figure 1: The result from NSGA-II that most closely calibrated ARTIMMUS parameter values against the hand-calibrated target values.

the-art automated methods with access to considerable computational power are challenged by this calibration problem, then those calibrating by eye and hand with trial and error aught to be cautious.

Acknowledgements

MR and JT are supported by the EC CoCoRo Project, GA 270382. JT is also supported by the Royal Society.

References

K Deb, A Pratap, S. A. T. M. (2002). A fast and elitist multiobjective genetic algorithm. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197.

Mark Read, Paul S. Andrews, J. T. V. K. (2012). Techniques for grounding agent-based simulations in the real domain: a case study in experimental autoimmune encephalomyelitis. *Mathematical and Computer Modelling of Dynamical Systems*, 18(1):67–86.

Read, M. (2011). *Statistical and Modelling Techniques to Build Confidence in the Investigation of Immunology through Agent-Based Simulation.* PhD thesis, The University of York, UK.

Tripp, M. (2013). Automated calibration of agent-based simulation for an autoimmune disease. Master's thesis, The University of York, UK.