CaliBayes: Integration of GRID-based simulation and data resources for Bayesian calibration of biological models

W.E. Wolski, D.A. Henderson, R.J. Boys, T.B.L. Kirkwood and D.J. Wilkinson

Newcastle University, UK

The discipline of systems biology is greatly concerned with understanding complex non-linear biological processes through fitting dynamic and predictive computer models to post-genomic experimental data sources. If this is successful, scientists will be able to conduct *in silico* biological experiments that would be impossible to carry out *in vivo* and these could, in turn, transform our understanding of molecular biology. In order to use these computer models for prediction they first need to be calibrated to observed experimental data; this involves estimating the unknown internal parameters of the computer model.

The primary aim of the CaliBayes project is to develop Bayesian statistical software to assist biologists in calibrating their dynamic deterministic and stochastic systems biology models to available experimental data and expert opinion. Bayesian calibration involves determining plausible values for model parameters together with some idea of the uncertainty in these parameter values.

The Bayesian inference tasks are implemented via computationally intensive Markov chain Monte Carlo (MCMC) methods and so benefit from GRID-based computing and simulation resources such as CONDOR (`www.cs.wisc.edu/condor/`) and BASIS (`www.basis.ncl.ac.uk`). The inference engine developed by the CaliBayes project is implemented in the programming language C, and provides an application programming interface (API) to the statistical computing language R, as well as to the programming languages C++, java and python. This makes the CaliBayes API highly suited to be extended with GRID-based technologies such as Web Services Description Language (WSDL) in order to implement a Service Oriented Architecture (SOA).

The first version of the CaliBayes API provides Bayesian methods for the analysis of fast deterministic models. We present here the results of applying the CaliBayes API version 1 to the analysis of data on the cell cycle in frog eggs.