Parameter Estimation in Computational Biology by Approximate Bayesian Computation (ABC) coupled with Sensitivity Analysis (SA)

Xin Liu[1] and Mahesan Niranjan[1]

xl1v09@soton.ac.uk, mn@soton.ac.uk

[1]School of Electronic and Computer Science, University of Southampton, UK

Abstract

We address the problem of parameter estimation in models of systems biology from noisy observations of model outputs. The models we consider are characterized by simultaneous deterministic nonlinear differential equations whose parameters are either taken from *in vitro* experiments, or are hand-tuned during the model development process to reproduces observations. We consider the family of algorithms coming under the Bayesian formulation of Approximate Bayesian Computation (ABC) and show that sensitivity analysis could be deployed to quantify the relative roles of different parameters in the system. Parameters to which a system is relatively less sensitive (known as sloppy parameters) need not be estimated to high precision, while the values of parameters that are more critical (stiff parameters) need to be determined with care.

The difficulty in estimating problem in high dimensions suggests a systematic re-allocation of computational effort, and we propose a three stage strategy in which sloppy parameters of a model are estimated in a coarse search followed by re-estimation of the stiff parameters to tighter error tolerances. We demonstrate the effectiveness of the proposed method on three oscillatory models and one transient response model taken from the systems biology literature.

Motivation

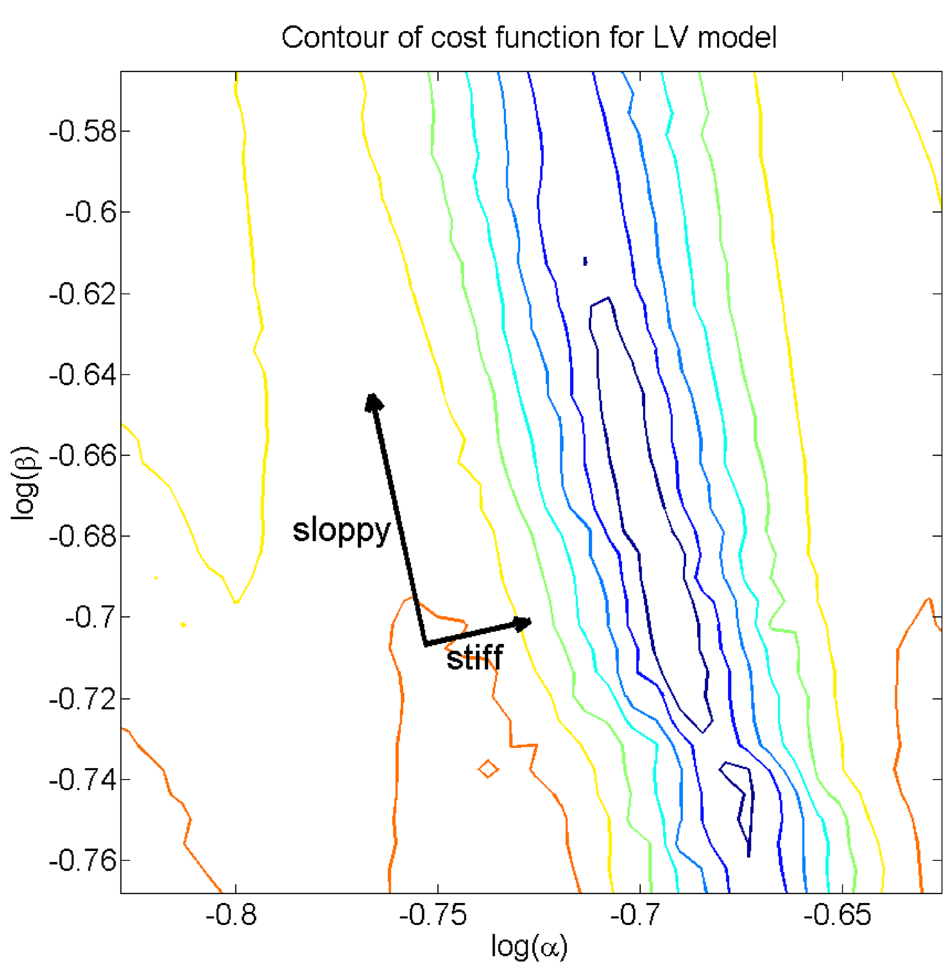
1) Estimate problem in high dimensions

- MCMC methods can tackle the problem in high dimensions, but they also have the slow chain mixing issue.
- Strike a balance between accuracy and efficiency.

2) Implicit expression of measurement noise

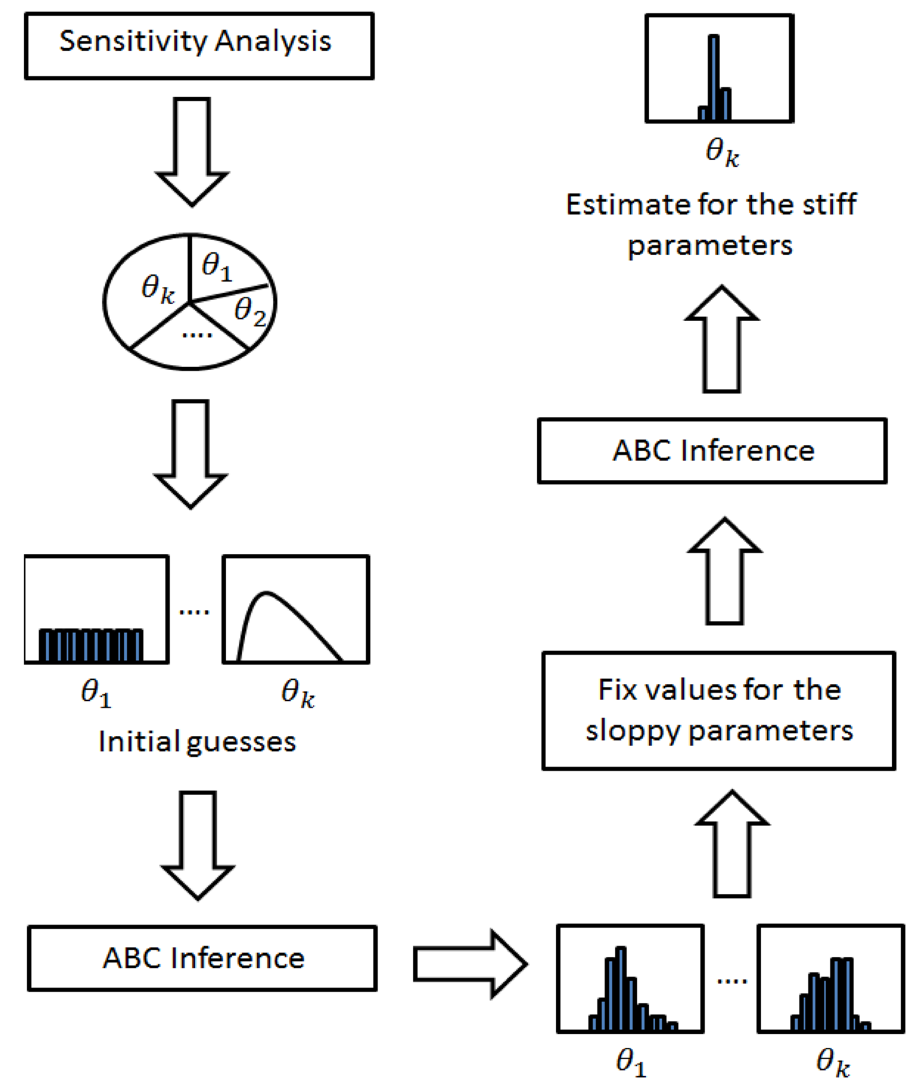
- Intractable likelihood evaluation causes the use of state-of-art inference methods e.g. Expectation-Maximization (EP), Maximization-Likelihood (ML) and Particle Filter (PF) becomes ineligible.

3) Sensitivity of parameters



ABC-SMC coupled with SA

1. Differentiate sloppy/stiff parameters.
2. Run ABC-SMC with coarse criterion.
3. Assign inferred values to sloppy parameters and reduce dimensions.
4. Run ABC-SMC with tight acceptance condition.



Extended Fourier amplitude sensitivity test (eFAST)

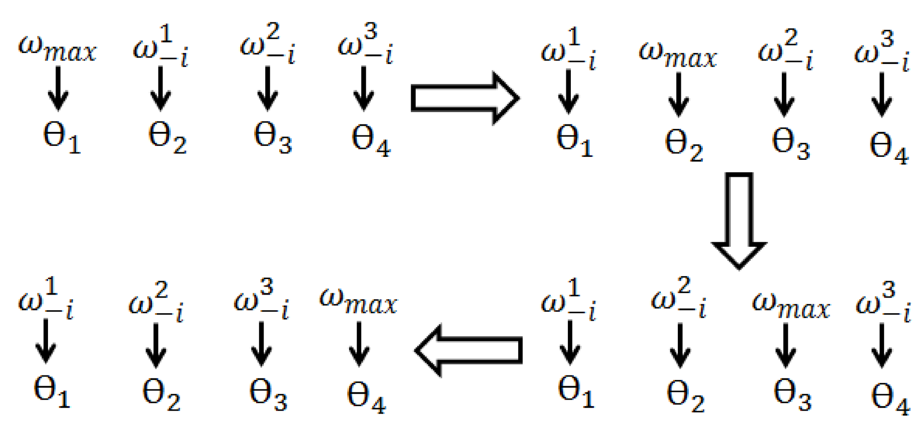
1) Variance decomposition method

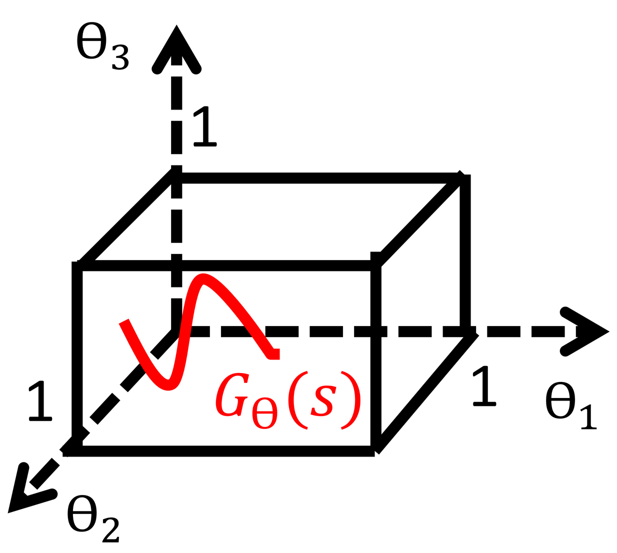
- eFAST apportions variance of output to individuals caused by the variations of parameters:

$$\mathrm{SI} = \frac{\mathrm{Var}_{\Theta_i}[\mathrm{E}(\mathrm{X}^*|\Theta_i)]}{\mathrm{Var}(\mathrm{X}^*)} = \frac{V_i}{V}$$

- Sinusoidal function for generating samples:

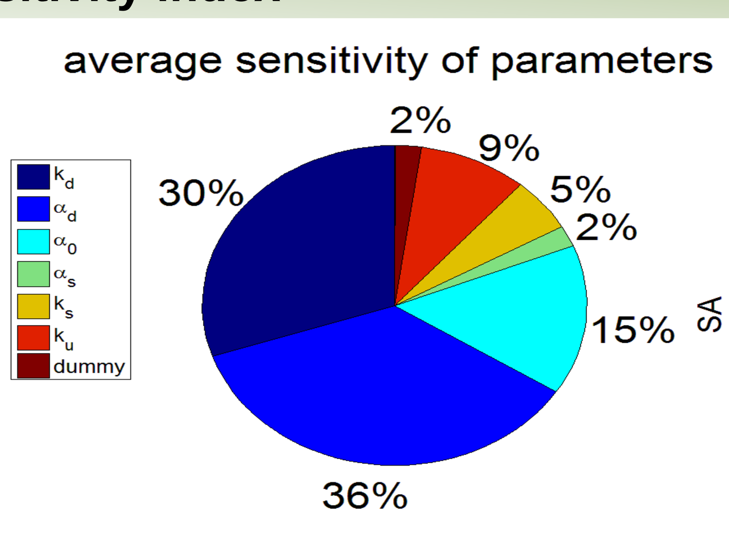
$$\Theta = G_\Theta(s) = \frac{1}{2} + \frac{1}{\pi}\arcsin(\sin(\omega s + \varphi))$$

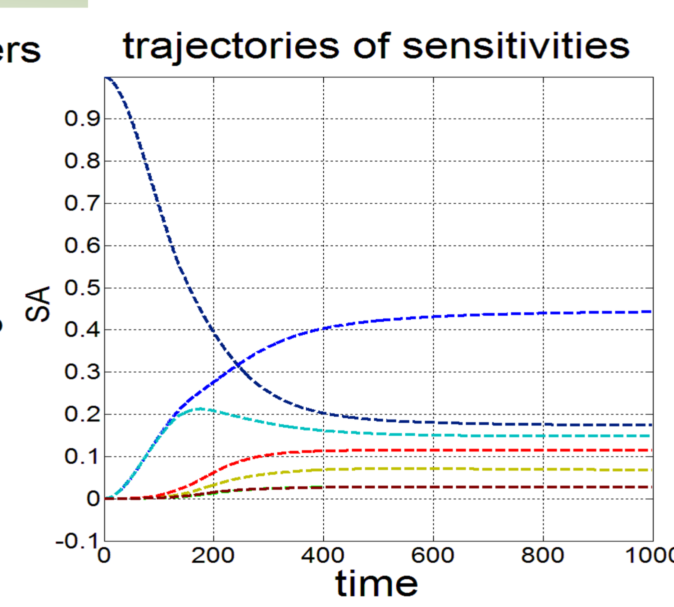




- System output is approximated by the Fourier coefficients, due to the Parseval's theorem.
- If the underling Θ_i is strong influence on system output, then oscillations of system output at frequency ω_i and it harmonics $M\omega_i$ have the high amplitude.

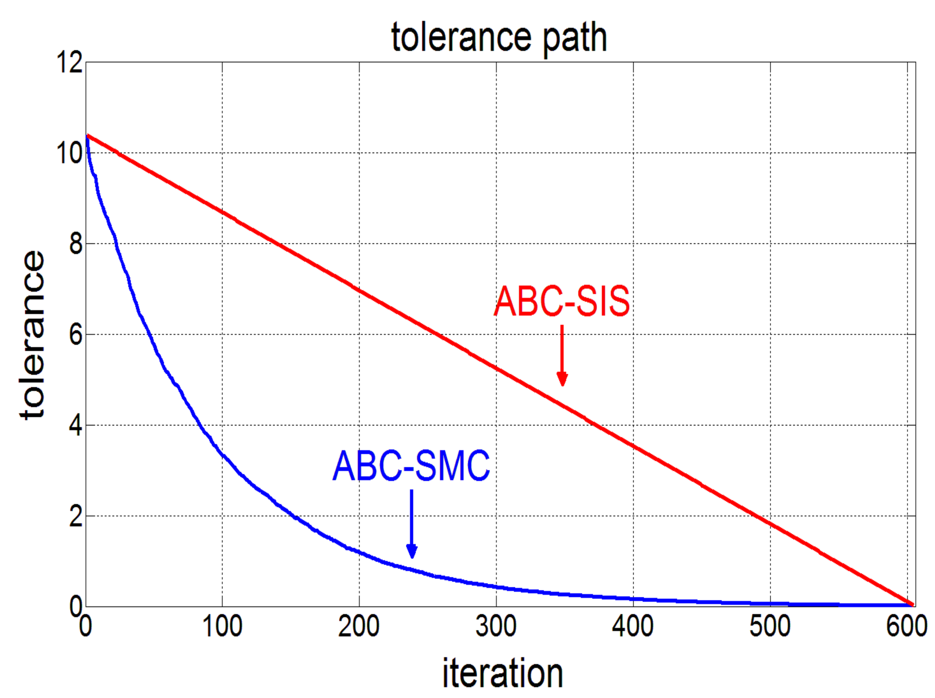
2) Sensitivity index

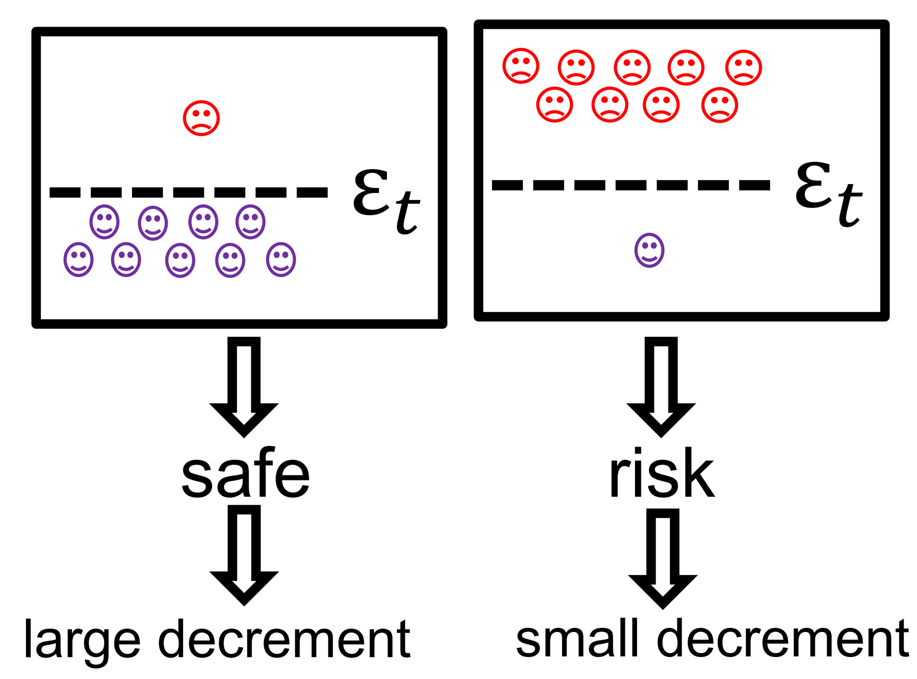




Adaptive ABC-SMC

1) Adaptive acceptance threshold reduction



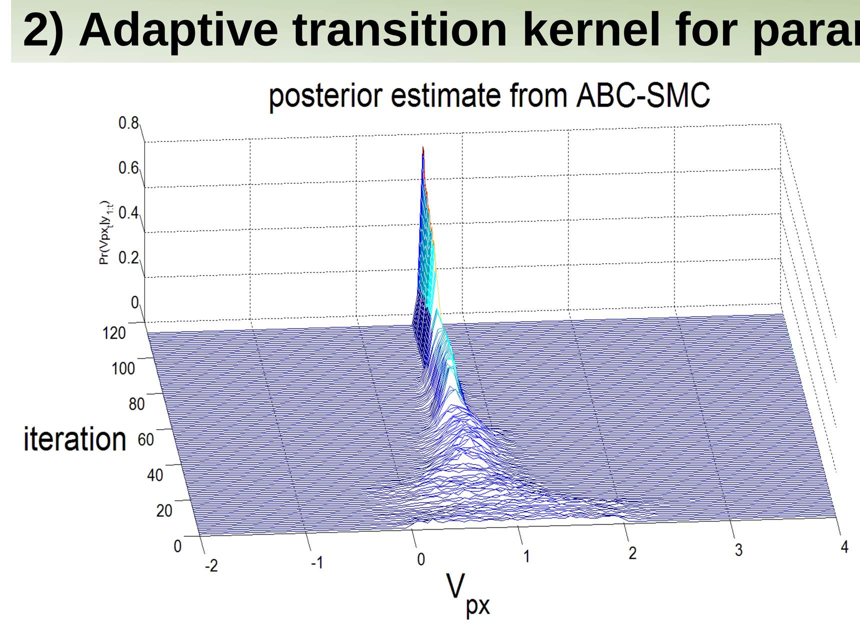


Proportion of 'alive' particles under the current threshold:

$$\mathrm{PA}(\mathrm{X}^*, \varepsilon_t) < \alpha \mathrm{PA}(\mathrm{X}^*, \varepsilon_{t+1})$$

$$\mathrm{PA}(\mathrm{X}^*, \varepsilon_t) = \sum_{m=1}^{M} \mathrm{I}_{\varepsilon_t}(\mathrm{X}^*_m, \mathrm{X})$$

2) Adaptive transition kernel for parameters



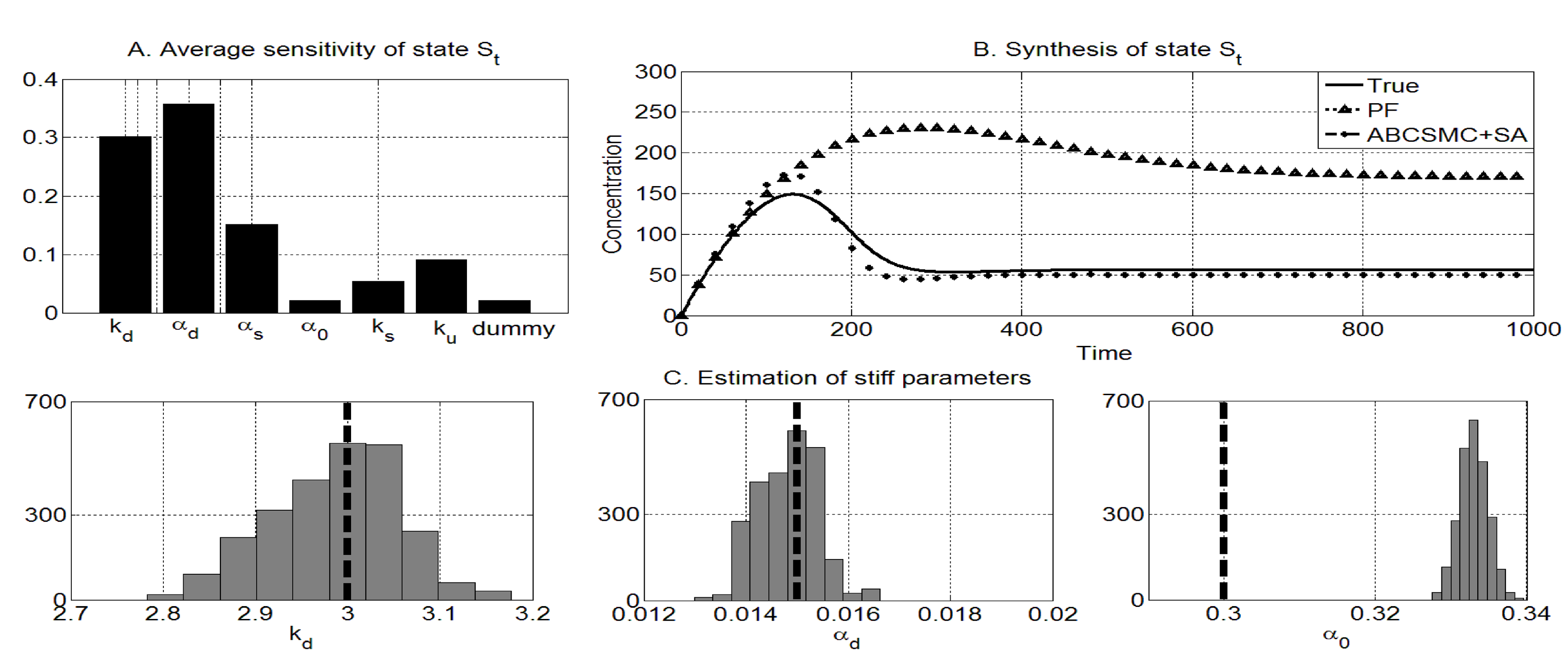
- moves along the mean of population.
- stepsize is determined by variance of population

$\Theta_{t+1} = \alpha_k \Theta_t + (1-\alpha_k)\overline{\Theta}_t + \omega_t^\Theta$

$\alpha_k = \frac{3\delta - 1}{2\delta}$, where δ is a discount factor and valued between 0.95 and 0.99.

$\omega_t^\Theta \sim N(0, h^2 V_t)$ where $h^2 = 1 - \alpha_k^2$ and V_t is the variance of the current population of particles.

Results of heat shock system



References

1. Gutenkunst, R.N., Waterfall, J. J., Casey, F. P., Brown, K. S., Myers, C. R. and Sethna, J. P. (2007). Universally sloppy parameter sensitivities in systems biology model. *PLoS Comput. Biol.*, **3**(10), 189-197.
2. Del Moral, P., Doucet, A. and Jasra, A. (2012). An adaptive sequential Monte Carlo method for approximate Bayesian computation. *Stat. Comput*, **22**(5),1009-1020.
3. Saltelli, A. , Tarantola, S. and Chan, K. P. –S. (1999). A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, **41**(1), 39-56.
4. Liu, X. and Niranjan, M. (2012). State and parameter estimation of the heat shock response system using Kalman and particle filters. *Bioinformatics*, **28**(11), 1501-1507.