

NONUNIFORM SAMPLING FOR GLOBAL OPTIMIZATION OF KINETIC RATE CONSTANTS IN BIOLOGICAL PATHWAYS

Steven H. Kleinstein

Department of Pathology
Yale University School of Medicine
New Haven, CT, 06520 U.S.A.

Dean Bottino

Anna Georgieva
Ramesh Sarangapani

Modeling & Simulation
Novartis Pharmaceuticals Corporation
East Hanover, NJ, 07936 U.S.A.

G. Scott Lett

The BioAnalytics Group, LLC
Jamesburg, NJ, 08831 U.S.A.

ABSTRACT

Global optimization has proven to be a powerful tool for solving parameter estimation problems in biological applications, such as the estimation of kinetic rate constants in pathway models. These optimization algorithms sometimes suffer from slow convergence, stagnation or misconvergence to a non-optimal local minimum. Here we show that a nonuniform sampling method (implemented by running the optimization in a transformed space) can improve convergence and robustness for evolutionary-type algorithms, specifically Differential Evolution and Evolutionary Strategies. Results are shown from two case studies exemplifying the common problems of stagnation and misconvergence.

1 INTRODUCTION

Estimation of kinetic rate parameters is an important and often rate-limiting step in systems biology (Periwal and Szallasi 2002; Schoeberl et al. 2002; Voit 2002; Tsai and Wang 2005). The parameter estimation problems that arise from fitting complex models to data are typically nonlinear and may have multiple local minima. Standard local optimization algorithms used to solve these problems can fail for various reasons (Mendes and Kell 1998), and global optimization has become an increasingly important tool.

The popularity of global parameter estimation is due primarily to the increased robustness of these methods, and although convergence can be slow, global methods may be more efficient overall than restarted local methods (Moles et al. 2003). Some of the most promising methods have been Differential Evolution (Storn and Price 1996) and Evolutionary Strategy (Runarsson and Yao 2000). For example, a recent study found Evolutionary Strategies far superior to others for estimating metabolic pathway parameters (Moles et al. 2003). While these methods are more robust than local algorithms, convergence failure can occur

(Lampinen and Zelinka 2000). Failure is either due to *stagnation*, in which convergence slows or stops before an optimum is found, or *misconvergence*, in which the algorithm converges to a suboptimal local minimum. A possible source of these problems is the uniformly distributed samples used to generate the initial population and to generate mutations for subsequent generations. It is demonstrated here that nonuniform sampling methods can improve both the speed and robustness of Stochastic Ranking Evolutionary Strategy (SRES) and Differential Evolution (DE).

2 SYSTEMS AND METHODS

The general optimization problem we consider is to find the vector \mathbf{p} that globally minimizes the function $f(\mathbf{p})$ subject to $\mathbf{p} \in \Omega$, where $\Omega = \{\mathbf{p}: LB_j \leq p_j \leq UB_j, \forall j\}$ defines a rectangular feasible region. Nonuniform sampling is implemented by defining a new function to minimize $f^*(\mathbf{p})=f(\mathbf{p}^*)$ where

$$p_j^* = 10^{p_j} - \text{offset}_j,$$

$$\text{offset}_j = \max(\epsilon, -2*LB_j),$$

and ϵ is a small number (e.g., `eps` in Matlab). The offset is used to properly handle feasible regions that include zero. In addition, the lower and upper bounds on the parameters are transformed so that:

$$LB_j^* = \log_{10}(\text{offset}_j + LB_j) \text{ and } UB_j^* = \log_{10}(\text{offset}_j + UB_j).$$

We present results from two case studies exemplifying the common problems of stagnation and misconvergence. The first case study is a synthetic metabolic pathway used in prior published work comparing global optimization methods (Moles et al. 2003). The model is comprised of

eight ordinary differential equations with 36 kinetic parameters. We applied the same basic procedure described in (Moles et al. 2003), and attempted to fit all 36 kinetic parameters of the model. Lower and upper bounds were chosen as in (Moles et al. 2003): [0.1, 10] for the Hill coefficients and $[10^{-12}, 10^{12}]$ for all other parameters. The error function is a weighted sum-of-squares comparing observed and predicted values.

Next, we study the “waterfall problem,” which has a single global minimum at $x=x^*$, $y=y^*$.

Minimize:

$$f(x,y) = 20 + a^2 + b^2 - 10 \{ \cos(2\pi a) + \cos(2\pi b) \},$$

where

$$\begin{aligned} a &= (x-x^*), \\ b &= (y-y^*), \end{aligned}$$

subject to $x, y, x^*, y^* \in [0, 10^{10}]$.

All of the optimization methods and simulations were implemented in Matlab or in a Matlab interface to external compiled code. While DE was run using its default settings, SRES used the settings applied in (Moles et al. 2003): $\lambda = 350$, $\mu = 30$, $pf = 0.450$, $\text{varphi} = 1$. Integration of the synthetic pathway model was performed with the LSODA algorithm (Hindmarsh 1983) with algorithmic enhancements and Matlab interface developed by The Bio-Analytics Group LLC.

3 RESULTS

3.1 Nonuniform Sampling Speeds Convergence

Stagnation (or very slow convergence) of optimization algorithms such as Differential Evolution is a well-known phenomenon (Kelley 2000, Lampinen and Zelinka 2000). This was also the observation of (Moles 2003) on a synthetic biochemical pathway model. Figures 1A and 1B show the convergence of DE and SRES on this problem. In the figures, DE-log and SRES-log refer to the nonuniform sampling versions of DE and SRES, respectively. SRES-log-initial includes nonuniform sampling only for generating the initial population, while SRES-log-iterations includes it only during subsequent iterations. These results can be compared with Figure 2 in (Moles et al. 2003). However, while they plot convergence as a function of CPU time, we graph convergence as a function of the number of simulation runs (i.e., function evaluations). This allows for easy comparison across multiple computer architectures. Our results confirm that SRES converges more quickly than DE. However, over the number of function evaluations we considered, they both reached a point where convergence slowed to an unacceptable rate. In fur-

ther agreement with (Moles et al. 2003), several runs of a local optimizer (FMINCON) also failed to converge.

Nonuniform (logarithmic) sampling dramatically improves the performance of both DE and SRES on this problem (Figures 1A and 1B). SRES with nonuniform sampling performs the best, consistently converging to the ‘true’ parameter values. This improvement is not due to the initial population sampling since a modified version of SRES including only this aspect does not show improved performance. However, uniform sampling of the initial population followed by nonuniform sampling during the iterations performs equivalently to nonuniform sampling throughout.

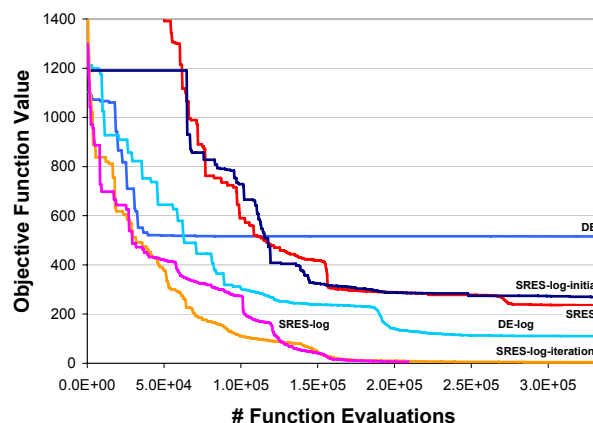


Figure 1A: Representative Convergence Plots for the Synthetic Metabolic Pathway Case

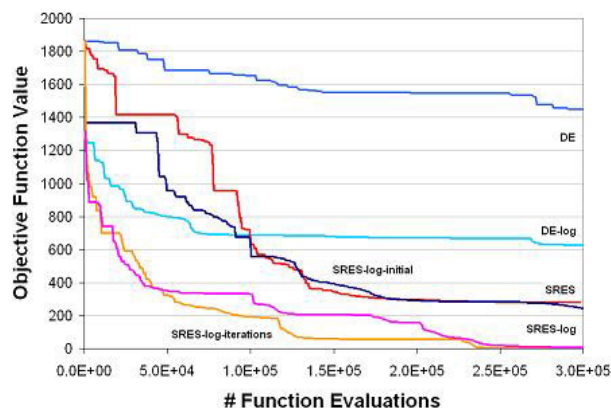


Figure 1B: Representative Convergence Plots for the Synthetic Metabolic Pathway Case

3.2 Convergence to Correct Local Minimum

There is a well-known tradeoff between the speed of convergence and the probability to misconverge (Storn and Price 1996). To explore the impact of nonuniform sampling on the potential to misconverge, we applied our algorithms to the waterfall problem, which has a large number of local minima, but a single global one (Figure 2). Specifically, we analyzed the relative performance of the un-

modified and modified versions of DE and SRES on the waterfall problem for various locations of the global minimum. At each location of the global minimum (i.e., every combination of $\log(x^*) = [-10, -9, \dots, 9, 10]$ and $\log(y^*) = [-10, -9, \dots, 9, 10]$), 100 optimizations were run for 100 iterations each, and the difference in the fraction of runs that converged (nonuniform - uniform sampling) was calculated. Convergence was defined as reaching a function value below 10^{-4} . In all cases, the actual function value at the global minimum is zero.

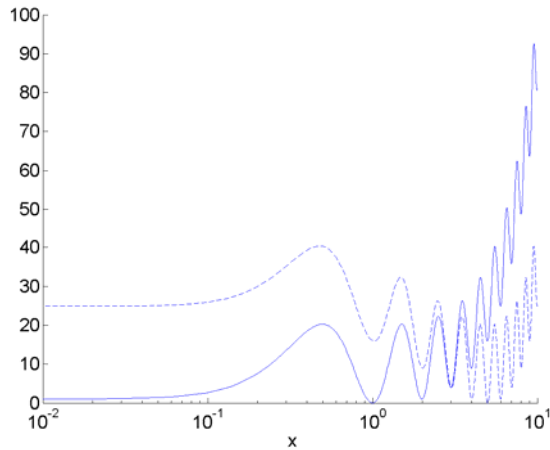


Figure 2: The Waterfall Function with Global Minimum at $(x^*=1, y^*=1)$, Solid Line, or at $(x^*=5, y^*=5)$, Dashed Line, with the y Value Set to the Global Minimum y^*

DE often misconverges on this problem. For all tested values of the global minimum, the fraction of runs that converged averaged 0.08 ± 0.03 . In most cases, nonuniform sampling greatly improves the performance, increasing the fraction converged by up to 0.97 (Figure 3). A notable exception occurs when the x or y value of the global minimum has a value of one (or zero in log space). In these cases, nearby local minima occur at x or y equal to zero. The log transform expands the basin of this minimum and, since only the global minimum has a lower value, many points end up here (Figure 2). It is a far jump from a value near zero (as small as -16 in transformed space) to a value near one (0 in transformed space) so leaving this basin is difficult and the optimization often stagnates here resulting in misconvergence. Although this “worst-case” scenario reveals one potential pitfall of nonuniform sampling based on the log transform, this appears to be a peculiarity of the waterfall problem. We have not yet encountered this issue in real-world parameter estimation problems.

The unmodified version of SRES already performs well on this problem (with an average fraction converged of 0.90 ± 0.08). Nonuniform sampling is still useful since it reduces the average number of iterations until convergence from 57 ± 1 to 28 ± 18 , among optimizations that converged. As was the case for stagnation in the previous sec-

tion, this improvement is mainly due to the impact of non-uniform sampling during the iterations. Nonuniform sampling of the initial population distribution only shows an improvement when the global minimum is very near the lower boundary. Like DE, the modified version of SRES performs poorly when the global minimum has x or y equal to one (for the reasons discussed above). However, its performance does not degrade when the global minimum is near the upper boundary of the feasible region (data not shown).

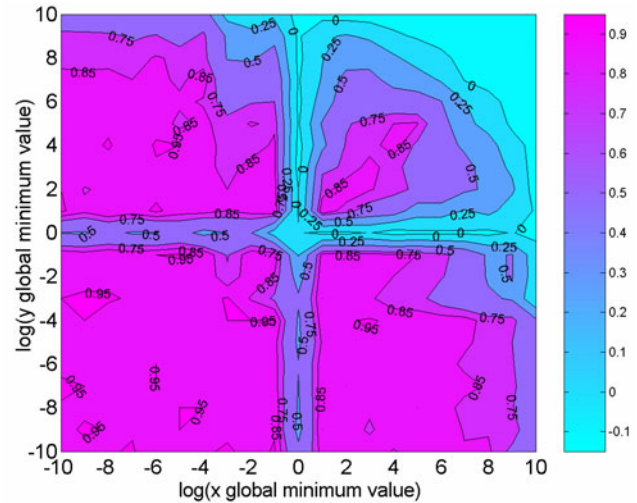


Figure 3: Difference in Performance of the Unmodified and Modified Versions of DE on the Waterfall Problem

The benefit of nonuniform sampling, as well as the performance of these optimization algorithms in general, will depend not only on the location of the optimum, but also on the precise bounds for the feasible region. Although we have not investigated this issue in detail, a limited analysis suggests that the results presented here are not overly sensitive to the particular choice of upper bound. For example, choosing upper bounds for x and y anywhere between 10^{-6} and 10^{20} does not significantly alter the performance of DE with nonuniform sampling at either good or poor performing optima, $(10^{-6}, 10^{-6})$ and $(10^6, 10^6)$ respectively, so long as the upper bound is greater than the optimal value (data not shown). On the other hand, increasing the lower bound for x and y (e.g., up to 10^{-6}) can dramatically improve robustness at the otherwise poorly performing optimum $(10^6, 10^6)$. This is not surprising since raising the lower bound will decrease the search space for the optimization.

3.3 Test Cases from the “Real-world”

In addition to the two synthetic case-studies described above, we have extensive experience applying these global optimization approaches to dozens of industrial and research problems. Here we briefly highlight two real-world

examples from the pharmaceutical industry. Specifically, we have considered a PBPK model and a biologically-based pathway model as test cases to evaluate the optimization methods. In both cases we tested SRES, DE, and DE with nonuniform sampling. SRES with nonuniform sampling was not tested on these problems since it had not been implemented when these projects were ongoing.

3.3.1 Physiologically-based Pharmacokinetic Modeling

A human physiologically-based pharmacokinetic (PBPK) model for a liver-active drug was developed to describe the *in vivo* clinical PK in humans (Figure 4). This model consists of multiple compartments representing the gut, liver, blood, poorly perfused and richly perfused tissues. The model compartments are connected by systemic circulation and are assumed to be homogenous and well-mixed. The venous effluents of the various compartments combine to yield a flow-averaged venous drug concentration. The venous blood exiting each compartment was equilibrated with the mixed-mean tissue concentration of the drug in the respective compartments. The orally administered drug was assumed to be cleared from the three compartment gut lumen by a first-order elimination process into the feces and another first-order uptake process into the liver via the portal blood flow. Derivation of the governing mass balance equations in each compartment have been described elsewhere (Sarangapani et al. 2003).

All the physiological parameters, such as tissue volume and regional tissue blood perfusion rates, and physicochemical parameters, such as partition coefficient and plasma protein binding, were obtained from either the literature or *in vitro* experimentation. Phase I clinical data from 32 healthy subjects was used as the basis to estimate the remaining eleven unknown kinetic parameters by global optimization.

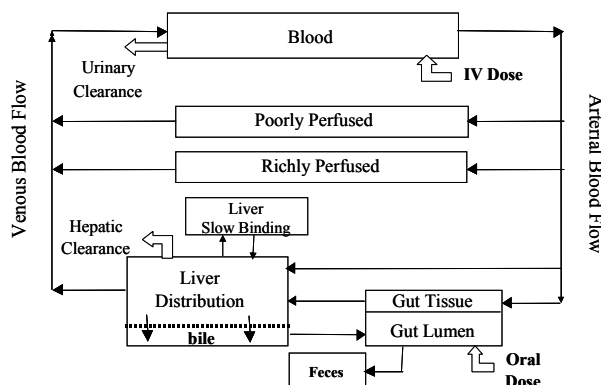


Figure 4: Schematic of the PBPK Model Structure

The model parameter estimation was conducted in two steps. In the first step, all eleven unknown parameters were estimated using the PK data by minimizing the sum-squared error (SSE) between the model predictions and the

measured data. All of the global optimization algorithms tested converged reasonably well on this problem. However, DE with nonuniform sampling achieved the fastest convergence with the smallest SSE (Figure 5). Several runs of a local optimizer (FMINSEARCH) failed to converge, including some with starting points very near the optimal value. Adaptive simulated annealing (Ingber 1989) was also tried on this problem with results that were better than SRES and DE, but not as good as DE with nonuniform sampling (data not shown). Following this initial optimization, we identified a reduced set of three highly observable parameters by exploring the SSE function in the neighborhood of the optimal estimates. We further optimized these parameters while the remaining kinetic parameters were assigned physiologically feasible values. Once again, DE with nonuniform sampling had the fastest convergence (data not shown).

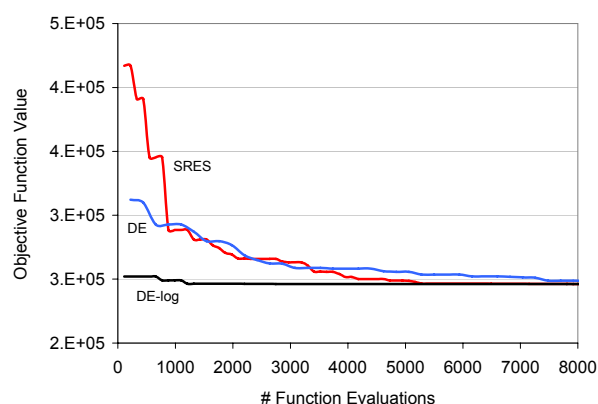


Figure 5: Convergence Plot with All Parameters Being Optimized for the PBPK Model Shown in Figure 4

3.3.2 Signal Transduction Pathway Modeling

We also tested the ability of nonuniform sampling to improve convergence on a model of the Wnt signal transduction pathway extended from (Lee et al. 2003). This model consists of a system of twenty ordinary differential equations describing the concentration of various protein and protein complexes in the pathway as a function of time. Each rate equation follows mass action kinetics and requires two parameters.

Several parameter values (including both rate constants and protein expression levels) were taken from the literature or measured directly. Time course kinetic information for a downstream entity in the pathway was then used as a basis to estimate the remaining twelve unknown parameters by global optimization. All approaches converged and yielded model predictions that were very close to the fitting data. Nevertheless, DE with nonuniform sampling converged 20% faster compared with the uniform

sampling version, and with a SSE that was 10% of that for uniform sampling.

4 CONCLUSIONS

Nonuniform sampling of parameters proved effective in improving speed of convergence and avoidance of stagnation for both DE and SRES. The results are mixed regarding misconvergence. Logarithmic transformation preferentially increases the density of samples near the region of parameter space where parameters have small magnitude. Consequently, the likelihood of misconvergence may increase if the solution is near the opposite end of the feasibility region. We found this was a problem for DE at the extreme upper boundary, but SRES performed well regardless of the location of the global minimum. The usual remedy for this problem in DE is to adjust the parameters to increase diversity in the sampled population (Storn and Price 1996), which may be an option. However, this motivates the search for a nonuniform sampling method (such as the logit transformation) that preserves samples in both ends of the feasibility intervals.

The idea of nonuniform distributions of parameters has been applied to other algorithms. Globalized Nelder-Mead uses a nonuniform distribution of parameters in restarted Nelder-Mead, but the method is applied to untransformed parameters (Luersen and Riche 2002). Thus we expect Globalized Nelder-Mead would suffer the same convergence problems seen in the benchmark cases in this paper. Thus, it may be worth investigating whether Globalized Nelder-Mead can be accelerated using nonuniform transformation of the parameters. In equation of state phase calculations in fluid modeling, the partition coefficient is often transformed using the logarithm to accelerate convergence of optimization methods to minimize Gibbs free energy. In this case, however, local derivative-based methods are used to solve the resulting optimization problem. Logarithmic transformation of parameters as used in this approach is not applicable to general optimization problems.

In global optimization, nonuniform distribution of initial guesses is a standard technique. TOMLAB© is a commercial optimization package that has options for nonuniform distributions of initial populations. A deterministic distribution of population members at corners and centroids of hypercubic feasible regions in bounds constraint problems, and a Latin hypercube probabilistic distribution is used to distribute initial populations in global methods. It would be of interest to combine the Latin hypercube experimental design with nonlinear transformation of parameters as described in this paper.

In summary, we have demonstrated that nonuniform sampling can improve both the speed and robustness of the evolutionary-type algorithms SRES and DE. Although there is no single optimization approach that is guaranteed

to work well for all problems (Wolpert and Macready 1997), we have found that SRES and DE with nonuniform sampling work well in practice for estimating kinetic rate constants in several types of biological pathway models.

Matlab code for the SRES and DE methods used here are available on the Web at the following URLs: <http://www.cs.princeton.edu/~stevenk/optimization> and http://www.bioanalyticsgroup.com/default_files/Page295.htm.

ACKNOWLEDGMENTS

Steven H. Kleinstejn was partially supported by NSF IGERT grant DGE-9972930.

REFERENCES

- Hindmarsh, A. C. 1983. ODEPACK, a systematized collection of ODE solvers. *Scientific Computing*. e.a. R.S. Stepleman. Amsterdam, North-Holland. 1: 55-64.
- Ingber, L. 1989. Very fast simulated re-annealing. *Mathematical and Computer Modelling* 12(8): 967-973.
- Lampinen, J., and I. Zelinka. 2000. On stagnation of the differential evolution algorithm. *MENDEL 2000, 6th International Mendel Conference on Soft Computing*, Brno, Czech Republic, Brno University of Technology.
- Lee, E., A. Salic, R. Kruger, R. Heinrich, and M. W. Kirschner. 2003. The roles of APC and Axin derived from experimental and theoretical analysis of the Wnt pathway. *PLoS Biology* 1(1): E10. Epub 2003 Oct 13.
- Luersen, M. A., and R. L. Riche. 2002. Globalized Nelder-Mead method for engineering optimization. In *Proceedings of the Third International Conference on Engineering Computational Technology*. Stirling, Scotland, Civil-Comp press: 165-166.
- Mendes, P., and D. Kell. 1998. Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation. *Bioinformatics* 14(10): 869-83.
- Moles, C. G., P. Mendes, and J. R. Banga. 2003. Parameter estimation in biochemical pathways: a comparison of global optimization methods. *Genome Research* 13(11): 2467-74.
- Periwal, V., and Z. Szallasi. 2002. Trading "wet-work" for network. *Nature Biotechnology* 20(4): 345-6.
- Runarsson, T. P., and X. Yao. 2000. Stochastic ranking for constrained evolutionary optimization. *IEEE Transactions on Evolutionary Computation* 4(3): 284-294.
- Sarangapani, R., J. Teeguarden, M. E. Andersen, R. H. Reitz, and K. P. Plotzke. (2003). Route-specific differences in distribution characteristics of octamethylcyclotetrasiloxane in rats: analysis using PBPK models. *Toxicological Sciences* 71(1): 41-52.

- Schoeberl, B., C. Eichler-Jonsson, E. D. Gilles, and G. Muller. 2002. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nature Biotechnology* 20(4): 370-5.
- Storn, R., and K. Price. 1996. Minimizing the real functions of the ICEC'96 contest by differential evolution. In *Proceedings of IEEE International Conference on Evolutionary Computation*, 842-844.
- Tsai, K. Y., and F. S. Wang. 2005. Evolutionary optimization with data collocation for reverse engineering of biological networks. *Bioinformatics* 21(7): 1180-8.
- Voit, E. O. 2002. Metabolic modeling: a tool of drug discovery in the post-genomic era. *Drug Discovery Today* 7(11): 621-8.
- Wolpert, D. H., and W. G. Macready. 1997. No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation* 1(1): 67-82.

AUTHOR BIOGRAPHIES

STEVEN H. KLEINSTEIN is an Assistant Professor in the Department of Pathology at Yale University School of Medicine. He was previously a member of the research staff at Princeton University where he ran the Program in Integrative Information, Computer and Application Sciences (PICASSo). He holds a Ph.D. degree in Computer Science from Princeton University. His research focuses on the development and application of computational methods that leverage mathematical/statistical models and numerical simulations in order to improve understanding of experimental and clinical data. The immune response has been a particular focus of his work. He is a member of the Society for Mathematical Biology and the International Society for Computational Biology. His e-mail address is steven.kleinstein@yale.edu.

DEAN BOTTINO is an associate director in the Department of Modeling & Simulation, part of Clinical Development and Medical Affairs at Novartis Pharmaceuticals. Dean holds a Ph.D. in applied mathematics from Tulane University. He provides modeling support for compounds in both clinical and pre-clinical development to better understand mechanism of action and help in the discovery of markers of both efficacy and toxicity. In addition, he has worked on implementing and further developing different types of systems-level models, ranging from signal transduction pathways, tumor pathophysiology and cardiac electrophysiology models. His e-mail address is: dean.bottino@novartis.com.

ANNA GEORGIEVA, Ph.D. is an associate director in the Department of Modeling & Simulation, part of Clinical Development and Medical Affairs at Novartis Pharmaceuticals. She provides modeling support for compounds in

both clinical and pre-clinical development to better understand mechanism of action and help in the discovery of markers of both efficacy and toxicity. In addition, she has worked on implementing and further developing different types of systems-level models, ranging from signal transduction pathway models, network inference models and cardiac electrophysiology. Her e-mail address is anna.georgieva@novartis.com.

RAMESH SARANGAPANI is a senior associate in the Modeling & Simulation department at Novartis Pharmaceuticals. Prior to joining Novartis in 2002, he held multiple positions of increasing responsibility at ICF consulting. He holds a Ph.D. in bioengineering from the University of Delaware. He specializes in applying advanced pharmacostatistical modeling techniques to assess the efficacy and safety of new chemical entities. As a member of the modeling group at Novartis, he supports integration and optimal use of pre-clinical and clinical data in R&D. He is a member of the American Association of Pharmaceutical Scientists. His e-mail address is ramesh.sarangapani@novartis.com.

G. SCOTT LETT is the President of The BioAnalytics Group LLC, a scientific computing and modeling company in the medical research industry. His research focuses on the development of numerical and visualization algorithms and their application to life sciences research. He held senior positions in the spacecraft industry and the petroleum industry before joining Physiome Sciences in 1999 to focus on computational biology. He holds a Ph.D. in Applied Mathematics from the University of Colorado. He is a member of the Society of Industrial and Applied Mathematics, the Institute of Electrical and Electronics Engineers, the IEEE Computer Society, the IEEE Engineering in Medicine and Biology Society, and the American Heart Association. His e-mail address is slett@bioanalyticsgroup.com.