Dynamic Optimization with Particle Swarms (DOPS): A meta-heuristic for parameter estimation in biochemical models

Adithya Sagar, Christine Shoemaker[†] and Jeffrey D. Varner[*]

School of Chemical and Biomolecular Engineering

[†]School of Civil and Environmental Engineering

Cornell University, Ithaca NY 14853

Running Title: Parameter estimation in biochemical models

To be submitted: *PLoS ONE*

[*]Corresponding author:

Jeffrey D. Varner,

Associate Professor, School of Chemical and Biomolecular Engineering,

244 Olin Hall, Cornell University, Ithaca NY, 14853

Email: jdv27@cornell.edu

Phone: (607) 255 - 4258

Fax: (607) 255 - 9166

Abstract

Mathematical modeling is a powerful tool to analyze, and ultimately design biochemical networks. However, estimation of biochemical model parameters is a significant challenge. Parameter estimation in biochemical models generally involves expensive function evaluations, making it difficult to quickly obtain optimal solutions. Additionally, biochemical models often have many local minima or maxima which further complicates the parameter estimation problem. In this study, we developed Dynamic Optimization with Particle Swarms (DOPS), a novel meta-heuristic that combined features of multi-swarm particle swarm optimization with dynamically dimensioned search (DDS). DOPS uses a multi-swarm particle swarm optimization technique to generate candidate solution vectors of which the best one is greedily updated using dynamically dimensioned search. We first tested the performance of DOPS on a model of human coagulation cascade. We performed 25 trials with 4000 function evaluations per trial, and compared the performance of DOPS with other commonly used meta-heuristic approaches such as differential evolution (DE), simulated annealing (SA) and dynamically dimensioned search (DDS). We further tested the predictive power of the coagulation model parameters against data not used in training, and found good agreement between model simulations and experimental measurements. Thereafter we tested the performance of DOPS on commonly used test functions for global optimization (Ackley and Rosenbrock) and on published biochemical parameter estimation benchmark problems. For the wide range of problems that we considered, DOPS outperformed other meta-heuristic approaches despite a limited number of function evaluations. Taken together, DOPS is a promising meta-heuristic approach for the estimation of biochemical model parameters in relatively few function evaluations.

Keywords: Parameter identification, Meta-heuristic optimization, Biochemical modeling

Introduction

Cells process nutrients and respond to changes in their environment using complex biochemical networks. These networks contain thousands of components interconnected through nonlinear enzyme catalyzed reactions. Mathematical modeling has evolved as a powerful paradigm to analyze, and ultimately design these complex networks [**? ? ? ? ?**]. Mathematical modeling of biochemical networks is often an iterative process. First, models are formulated from biochemical knowledge, and then model parameters are estimated using experimental data [**? ? ?**]. Parameter estimation is typically framed as a non-linear optimization problem wherein the residual (or objective function) between experimental data and model simulations is minimized using an optimization strategy [**?**]. The optimal parameters obtained from training are then used to validate the model on hitherto unseen experimental data. If validation fails, model construction and calibration are repeated iteratively until satisfactory results are obtained.

Parameter estimation is a major challenge in the development of large biochemical models. Although parameter estimation has been a well studied problem in engineering for decades [**? ? ? ?**], the complex dynamics of large biological systems and noisy, often incomplete experimental data pose a unique estimation challenge. Most of the optimization problems pertaining to biological systems are non-linear and multi-modal i.e. most of them have multiple local minima or maxima [**? ?**]. The non-linearity of the problem coupled with multi-modality generally renders the local optimization techniques like pattern search method [**?**], Nelder-Mead simplex method [**?**], steepest descent method or Levenberg-Marquardt method [**?**] incapable of reliably obtaining an optimal solution since they generally stop at the first local minimum. Though deterministic global optimization techniques (for example algorithms based on branch and bound framework) can handle non-linearity and multi-modality [**? ?**], the absence of derivative information, discontinuity of the objective functions, non-smooth regions or the lack of any knowledge

about the objective function severely hamper the use of these techniques.

Meta-heuristic stochastic optimization approaches like Genetic Algorithms (GAs), Simulated Annealing (SA) [**?**], Evolutionary Programming and population based searches like Differential Evolution (DE) [**?**] have shown promise in this regard [**?**]. They do not make any assumptions about the structure of objective function or require any *a priori* information about the objective function. Though they do not guarantee strong convergence, these approaches are effective in finding near optimal solutions. Mendes et al. [**?**] used Simulated Annealing to estimate rate constants for the irreversible inhibition of HIV proteinase, Modchang et al. [**?**] used Genetic Algorithms to estimate parameters for a mathematical model of signal transduction, Differential Evolution based approaches have been an effective approach on various systems [**? ? ?**]. Tashkova et al. [**?**] compared different meta-heuristics for parameter estimation on a dynamic model of endocytosis and showed that DE was the most effective. Banga and co-workers have applied scatter-search based methods [**? ? ?**] to estimate parameters on non-linear biological processes. Hybrid approaches that combine a meta-heuristic with a local optimization search, wherein a near globally optimal solution that is obtained with a meta-heuristic is further refined using a local search have also become popular. Villaverde et al. [**?**] combined scatter search with local search methods for parameter estimation in large scale systems biology models. Fan et al. recently showed that population based meta-heuristics along with decomposition based methods can be used to model gene circuits from mRNA data [**?**]. Despite these successes, a major drawback with most of these metaheuristic approaches is the vast number of objective function evaluations they take in arriving at good search regions. As the models grow in size and complexity and the number of dimensions of parameter vector increases, evaluation of the objective function becomes computationally expensive. Thus performing a large number of evaluations is not computationally feasible.

In many of these high dimensional problems approaching an exact solution may not be necessary. Gutenkust et al. [**?**] showed that a number of systems biology models are 'sloppy'. Sloppy systems have specific parameter combinations that largely define the dynamics of the system. Large perturbations to the rest of the parameters does not greatly impact the system dynamics. Ensemble approaches [**?** **?**] have exploited this aspect to describe the dynamics of biological systems including coagulation which can be described using only a set of key species or parameters [**?**]. Tolson and Shoemaker [**?**] showed through Dynamically Dimensioned Search (DDS) that high-dimensional watershed models can be calibrated quickly by perturbing only a subset of dimensions.

In this study, we developed Dynamic Optimization with Particle Swarms (DOPS), a novel meta-heuristic that combines the global search capability of multi-swarm particle swarm optimization and dynamically dimensioned search (DDS). The objective of DOPS is to obtain near optimal parameter estimates for large biochemical models within a relatively few function evaluations. DOPS uses a multi-swarm particle swarm optimization technique to generate candidate solution vectors which are then greedily updated using dynamically dimensioned search. We first considered a model of human coagulation cascade to test the performance of DOPS. Coagulation is a large, complex biochemical network that is tightly regulated with several feedback loops. We then tested the performance of DOPS on commonly used test functions for global optimization (Ackley and Rosenbrock), published biochemical parameter estimation benchmark problems [**?**]. DOPS outperformed common meta-heuristic approaches like Differential Evolution (DE), Simulated Annealing (SA) and dynamically dimensioned search (DDS) on the test functions and the coagulation model. It also performed very well on the benchmark problems where it outperformed enhanced scatter search (eSS) and recovered the nominal parameters with only 4000 function evaluations across all the benchmark problems considered.

Problem Formulation

The problem of parameter estimation in dynamic biological systems consists of finding an optimal parameter vector that results in the best fit between the model predictions and observed experimental data.The cost function or objective that measures the goodness of fit is generally the Euclidean norm which is minimized under certain constraints. This can be mathematically expressed as follows.

$$\begin{aligned} \text{minimize} \quad & K(\mathbf{p}) = (\sum_{i=1}^{N}(g_i(t_i, \mathbf{x}, \mathbf{p}, \mathbf{u}) - y_i)^2) \\ \text{subject to} \quad & \dot{x} = f(t, \mathbf{x}(\mathbf{t}, \mathbf{p}), \mathbf{p}, \mathbf{u}(\mathbf{t})) \\ & x(t_0) = x_0 \\ & c(t, \mathbf{x}, \mathbf{p}, \mathbf{u}) \geqslant 0 \\ & \mathbf{p}^L \leqslant \mathbf{p} \leqslant \mathbf{p}^U \end{aligned} \tag{1}$$

where t is time, x is the state variable with an initial state x_0, u is the input vector p is the parameter vector and f is the system of equations (like differential equations or set of algebraic constraints), c in the set of linear and non-linear constraints, g is the set of observables and p^L and p^U are the lower and upper bounds. Thus the problem eventually is to find p^* where

$$\mathbf{p}^* = \arg\min_p K(\mathbf{p}) \tag{2}$$

Results

Dynamic Optimization with Particle Swarms (DOPS). DOPS is a novel meta-heuristic that combines multi-swarm based particle swarm methods with dynamically dimensioned search (Fig.**??**). The goal of this approach is to obtain optimal or near optimal parameters for high-dimensional complex biological systems within a pre-specified number of

function evaluations. We randomly initialized a swarm of $\mathcal{K}$-dimensional particles (represented as x_i), wherein each of these particles corresponds to a $\mathcal{K}$-dimensional parameter vector. After initialization, the particles were grouped into different sub-swarms randomly. Thereafter within each sub-swarm S_k, particles were updated according to the following rule.

$$\mathbf{x}_{i,j} = \theta_{1,j-1}\mathbf{x}_{i,j-1} + \theta_2\mathbf{r}_1\left(\mathcal{L}_i - \mathbf{x}_{i,j-1}\right) + \theta_3\mathbf{r}_2\left(\mathcal{GL}_k - \mathbf{x}_{i,j-1}\right) \tag{3}$$

where $(\theta_1, \theta_2, \theta_3)$ are adjustable parameters, $\mathcal{L}_i$ denotes the best solution found by particle i within the sub-swarm till function evaluation j-1, and $\mathcal{GL}_k$ denotes the best solution found over the population of all particles within the sub-swarm S_k. The quantities r_1 and r_2 denote uniform random vectors with the same dimension as the number of unknown model parameters ($\mathcal{K} \times 1$). Equation (3) is suggested by the general form of particle swarm equations. However this equation does not contain the velocity terms that are generally associated with particle swarm optimization. In our algorithm the parameter $\theta_{1,j-1}$ depends on the function evaluations and is controlled according to the following equation

$$\theta_{1,j} = ((\mathbf{N} - j) * (\mathbf{w}_{max} - \mathbf{w}_{min}))/(\mathbf{N} - 1) + \mathbf{w}_{min} \tag{4}$$

where $\mathbf{N}$ represents the total number of function evaluations, $\mathbf{w}_{max}$ and $\mathbf{w}_{min}$ are the maximum and minimum inertia weights respectively. While updating the particle, we made sure all dimensions of the solution represented by the particle were within bounds using a set of reflection boundary conditions shown below.

(5)

if $x_{i,j}^{old} < x_i^{min}$ **then**
  $x_{i,j}^{new} = x_{i,j}^{old} + (x_i^{min} - x_{i,j}^{old})$ **if** $x_{i,j}^{new} > x_i^{max}$ **then**
    $x_{i,j}^{new} = x_i^{max}$
  end
end
if $x_{i,j}^{old} > x_i^{max}$ **then**
  $x_{i,j}^{new} = x_{i,j}^{old} + (x_{i,j}^{old} - x_i^{max})$ **if** $x_{i,j}^{new} < x_i^{min}$ **then**
    $x_{i,j}^{new} = x_i^{min}$
  end
end

After every g function evaluations, the particles within all sub-swarms were mixed and then randomly redistributed to a new sub-swarm. The particles were then again updated according Eq.1. This process continued till $\mathcal{FR} * \mathbf{N}$ number of functions evaluations, where $\mathcal{FR}$ represents the fraction of evaluations with the multi-swarms. At the end of these function evaluations, we froze all the solutions represented by various particles and chose the particle with best solution among $\mathcal{GL}_1 \cdots \mathcal{GL}_{NS}$ as the initial candidate vector $\mathcal{G}$ for the remaining $(1 - \mathcal{FR}) * \mathbf{N}$ number of function evaluations.

This particle was then updated according to the following rule

$$\mathcal{G}_{new}(\mathbf{J}) \quad = \quad \mathcal{G}(\mathbf{J}) + \mathbf{r}_{normal}(\mathbf{J})\sigma(\mathbf{J}) \tag{6}$$

where $\mathbf{J}$ represents the set containing the specific dimensions being perturbed, r_{normal} denotes a normal random vector of the same dimensions as $\mathcal{G}$. σ is the amplitude of perturbation given by following equation:

$$\sigma \quad = \quad \mathbf{R}(\mathcal{MAX} - \mathcal{MIN}) \tag{7}$$

where $\mathbf{R}$ is the scalar perturbation size parameter, $\mathcal{MAX}$ and $\mathcal{MIN}$ are $(\mathcal{K} \times 1)$ vectors that represent the maximum and minimum bounds on each dimension. The set $\mathbf{J}$

was constructed using a probability function $\mathcal{P}_j$ that represented the probability whether a specific dimension j was perturbed or not. This function is a monotonically decreasing function that decreases with the number of function evaluations. $\mathcal{P}_j$ can be any monotonically decreasing function, in our approach we used the following function:

$$\mathcal{P}_j \quad = \quad 1 - \log(j/((1 - \mathcal{FR}) * \mathbf{N})) \tag{8}$$

Thus the number of dimensions of the candidate vector that are updated or perturbed decreases with the as the number of function evaluations increase. These updates are greedy in nature, so $\mathcal{G}_{new}$ becomes the new solution vector only if it is better than the old one $\mathcal{G}$. The decrease of weight function in equation 4, reflection boundary conditions in equation 5, the update function described in equations 6 and 7, the selection probability in equation 8 are the means by which DDS ideas are incorporated into DOPS. The fraction of evaluations $\mathcal{FR}$ within the swarm phase and DDS phase is based on a switching strategy wherein the switch from swarm phase to DDS phase happens when the error due to the best solution does not drop more than 1% of the original error, continuously for more than a prescribed number of function evaluations. This allows the solution to quickly jump out of local optima and avoid any convergence issues that are generally associated with swarm based searches.

Performance of DOPS on a model of the human coagulation cascade. We compared the performance of DOPS on a model of blood coagulation against commonly used meta-heuristics like simulated annealing (SA), differential evolution (DE), and dynamically dimensioned search (DDS). Coagulation is an archetype biochemical network that is highly interconnected and tightly regulated with multiple positive and negative feedback loops (Fig. **??**). The biochemistry underlying coagulation, though quite complex has been well studied [**? ? ? ? ? ? ?**], and reliable experimental coagulation models have been

developed [**?** **?** **?** **?**]. This makes it an ideal system for mathematical modeling and parameter estimation. Coagulation is regulated by a set of serine proteases also known as coagulation factors and blood platelets.

The coagulation factors are generally in an inactive state and are known as zymogens. These zymogens are activated through certain triggers. Trigger events like injury or trauma or sepsis expose factors like collagen, tissue factor and von Willebrand factor (vWF) to blood. The exposure of these factors to blood kick-starts a series of convergent cascades that lead to conversion of zymogen prothrombinase to thrombin. For example when coagulation is initiated through tissue factor pathway, tissue factor and activated factor VIIa (FVIIa) form a complex that activates factors X (fX) and IX (fIX). Activated factor X (fXa) thereafter activates downstream factors VIII and IX. The initial activation leads to the production of picomolar amounts of thrombin (fIIa) which activates platelets and amplifies its own production through the formation of a prothrombinase complex (FXa-FVa) on the surface of the activated platelets. Thrombin also downregulates its own production by forming a complex with thrombomodulin which then activates protein C (PC). PC inhibits the formation of prothrombinase complex. In addition, Tissue factor pathway inhibitor (TFPI) downregulates FXa formation and activity by sequestering free FXa and TF-FVIIa in a FXa-dependent manner. Antithrombin III (ATIII) inhibits all proteases. surface protein thrombomodulin (TM). Luan et al. modeled coagulation using coupled non-linear ordinary differential equations with 148 reactions and 92 species [**?**] and validated model using 21 published datasets.

To train the model parameters, we used data sets from TF-VIIa initiated coagulation with no anticoagulants. The objective function was a weighted linear combination of two different error functions that used data sets representing coagulation initiated with different concentrations of TF-VIIa (5pM, 5nM) [**?**]. This choice of using a linear combination of two different error functions was motivated by poor validation results while using a

single error function. We restricted the number of function evaluations to 4000 for each algorithm and performed 25 trials of this experiment. DOPS exhibits a much faster rate of error convergence and has a much lower final error as compared to the other algorithms (Figure 3).

Within the first 1000 function evaluations of DOPS there is a very rapid drop in error. Approximately between 500 -1000 function evaluations a switch to dynamically dimensioned search phase happens (this switch varies from trial to trial since the switch is based on the error from the swarm phase). Overall at the end of 4000 function evaluations DOPS minimizes the error (final objective error is 90456.9) to a much greater extent than any of the other algorithms. Using the parameters obtained at the end of 4000 function evaluations we examined the 'fits' between models predictions and experimental data (Figure 4). The solid lines represent the mean value of prediction over 25 trials and the shaded region represents the 99% confidence interval. Subsequently we compared the model predictions against completely 'unseen' or untrained experimental where coagulation was initiated with 500pM,50pM,10pM concentrations of TF-VIIa respectively (Figure 5).

Performance of DOPS on benchmark problems and test functions. Villaverde and co-workers recently published a set of benchmark biochemical problems to evaluate parameter estimation methods [**?**]. From a computational cost perspective problems they categorized the problems as most expensive, intermediate and least expensive. We evaluated the performance of our algorithms on a problem from the most expensive and least expensive categories. The first problem is a genome wide kinetic model of *Saccharomyces cerevisiae* with 1759 parameters (henceforth referred to as problem B1). The second problem (henceforth referred to as problem B4) is a metabolic model of Chinese Hamster Ovary (CHO) with 35 metabolites, 32 reactions and 117 parameters. In both cases pseudo time series data was generated. For problem B1, the time series data

consisted of 44 observables and for problem B4 the data corresponded to 13 different metabolites. We fixed the number of function evaluations at 4000 for DOPS and trained both the models against the pseudo experimental data. We recaptured the 'nominal' parameters (Figure S7 A and B) in both cases within 4000 evaluations. The final objective function value (Table 1) is an order smaller for B1 possibly due to overfitting. In Table 1 we summarize the results obtained on the coagulation model and the benchmark problems.

We compared the performance of DdPSO with algorithms on commonly used test functions for global optimization, as well as benchmark biochemical problems from literature (Table **??**). We used a 300 dimensional Rastrigin function and 300 dimensional Ackley function. Both Ackley and Rastrigin have multiple local minima and maxima and attain a global minimum value of 0. We tested the performance of the 5 different heuristic approaches on these two functions. In each experiment we again fixed the number of function evaluations at 4000 and ran 25 experiments. In both cases (Figure S6 A and B) we see that the error convergence rate for DDSMLSPSO is much faster as compared to the other four s and it finds the global minimum of 0 in both cases.

Discussion

Our study presents a novel approach for high-dimensional parameter estimation in complex biological systems with relatively few function evaluations. In this approach we combined a variant of a well known meta heuristic particle swarm optimization with Dynamically Dimensioned Search (DDS). We tested our approach on an ODE model of coagulation with 148 parameters and 92 species. Coagulation is an ideal system to test our approach since the biology is well known and complex, with multiple feed back loops that are tightly regulated. We used experimental data under different conditions to obtain optimal parameters and used these parameters to make predictions against unseen experimental data. We obtained good fits and made sufficiently accurate enough predictions using parameters obtained from 4000 function evaluations. Further, we also used high-dimensional forms of commonly used test functions of global optimization and showed that we were able to find the global minimum for 300 dimensional Ackley and Rastrigin functions faster than other meta-heuristics. We also considered two recently published benchmark problems to test parameter optimization approaches and showed that we were able to retrieve the nominal parameter vector within 4000 function evaluations. Meta-heuristic approaches are generally effective in finding close to optimum solutions of complex, multimodal functions. In addition, they generally obviate the need for any *a priori* knowledge (like function derivative). However they take an exorbitant number of objective function evaluations to come close to an optimum. When the objective function evaluations tend to become expensive it is infeasible to take up a large number of evaluations. As the dimensionality of parameter space increases, the search region gets widened and thus the problem becomes more challenging. In addition, most of these approaches require optimization of 'algorithm parameters' before the actual optimization and also involve computationally expensive update operations. Tolson and Shoemaker, through DDS, showed that randomly perturbing a subset of dimensions in high dimensional parameter space

is an effective way to obtain near optimal solutions with few function evaluations. However their approach was based on a single solution. Single solutions based methods may prove ineffective in complex systems when they start from a 'bad' region of search or from a local minimum. In our approach, we tried to eliminate the probability of starting from a bad region of search using a variant of particle swarm optimization. Particle Swarm Optimization (PSO) is a population based meta heuristic that is known for rapid convergence without any complex operations like recombination, mutation or selection that are associated with other population based meta-heuristics like Differential Evolution (DE) or Genetic Algorithm (GA). Several particle swarm variants have been proposed to improve the search ability and rate of convergence, that involve different neighborhood structures, multi-swarms or adaptive parameters. Multi-swarm PSO with small particle neighborhoods have been shown to better in searching on complex multi-modal solutions [**?**]. Multi swarm methods, in addition, avoid rapid convergence to a local optimum or stable point and are able to generate diverse solutions. Generation of diverse solutions in the early stage gives a better exploratory capability and thus converge of upon multiple optima.

We utilized this capacity of multi swarms to generate diverse candidate solutions which can be used as initial solutions for a DDS like search. Thus the solution vectors obtained at the end of multi swarm search phase have a better propensity to not get stuck in bad regions in DDS phase as opposed to starting with a totally random initial solution. Given a fixed number of function evaluations, we have been able to show that we were able to obtain better solutions for the coagulation model and the test functions faster that other commonly known meta-heuristics and also DDS alone. Choosing the number of function evaluations is largely a function of cost and complexity of the objective function. Traditionally the stopping condition for a parameter optimization problem can be the number of function evaluations, percentage of initial error achieved or an absolute error threshold.

However in case of complex, expensive functions where we desire a value within a certain period of time, the number of evaluations are used as a stopping criterion. In our current study we used a value of 4000 which we based upon the time taken (approximately 8-10 seconds) for a single objective function evaluation in the coagulation case. We used the same value of 4000 for benchmarks published by Villaverde and co-workers [**?**]. Quite surprisingly we took a couple of orders lesser number of function evaluations (1e3) to obtain the optimal parameter vector as compared to the enhanced Scatter Search (eSS) with a local optimizer, which took around the order of 1e5 number of evaluations. The amount of CPU time taken (on an Intel Xeon processor 2.4 GHz) is lesser as compared to eSS on a similar architecture.

A surprisingly remarkable aspect about our algorithm has been that we did not 'pre optimize' any parameters of the algorithm to suit a specific problem. In the swarm phase this includes the number of particles, number of sub-swarms, acceleration constants or the number of generations after which the particles are redistributed and the neighborhood perturbation parameter in DDS phase. We used the same parameters for all the problems. The same rule was applied to the rest of the meta-heuristics barring Simulated Annealing. For SA, we optimized the cooling schedule for the coagulation model. Thus, in this approach any overhead that usually comes with additional function evaluations to optimize parameters was avoided. The only parameter that we adjusted was the fraction of evaluations within swarm search. For the coagulation model it was set to be 0.4 and for the benchmarks it is at 0.6. This parameter was set based on empirical observations rather than any definitive analysis. When the error convergence rates stagnated in the swarm phase we switched to DDS phase. A further improvement to our approach could be to incorporate a meta intelligent way of switching in and switching out of both the phases.

The performance of our approach seems impressive given that it performed well on

different complex systems with no specific conditioning being required. We comfortably outperformed common meta-heuristics and were also able to find minima of high dimensional global optimization test functions. Thus this approach may be well suited to large scale global optimization. In addition, surprisingly, we were able to obtain optimal parameter vectors for two different large scale systems biology models with a couple of orders fewer number of function evaluations as compared a hybrid approach. However it is quite possible that highly optimized versions of common meta heuristics may outperform us on these systems. This aspect is currently beyond the scope of this study. Our approach can also be combined with local derivative based searches to improve upon the accuracy of the solutions.

Acknowledgements

This study was supported by an award from the Army Research Office (ARO #59155-LS).

References

1. Assmus HE, Herwig R, Cho KH, Wolkenhauer O (2006) Dynamics of biological systems: role of systems biology in medical research. Expert Review of Molecular Diagnostics .

2. van Riel NAW (2006) Dynamic modelling and analysis of biochemical networks: mechanism-based models and model-based experiments. Briefings in Bioinformatics 7: 364–374.

   KEY: Riel:2006aa

   ANNOTATION: 10.1093/bib/bbl040

3. Jaqaman K, Danuser G (2006) Linking data to models: data regression. Nat Rev Mol Cell Biol 7: 813–819.

   KEY: Jaqaman:2006aa

   ANNOTATION: 10.1038/nrm2030

4. Kitano H (2002) Systems biology: a brief overview. Science 295: 1662–1664.

5. Hood L, Heath JR, Phelps ME, Lin B (2004) Systems biology and new technologies enable predictive and preventative medicine. Science 306: 640–643.

6. Aldridge BB, Burke JM, Lauffenburger DA, Sorger PK (2006) Physicochemical modelling of cell signalling pathways. Nat Cell Biol 8: 1195–1203.

   KEY: Aldridge:2006aa

   ANNOTATION: 10.1038/ncb1497

7. Banga JR (2008) Optimization in computational systems biology. BMC systems biology 2: 47.

8. Ashyraliyev M, Fomekong-Nanfack Y, Kaandorp JA, Blom JG (2009) Systems biology: parameter estimation for biochemical models. Febs Journal 276: 886–902.

9. Moles CG, Mendes P, Banga JR (2003) Parameter estimation in biochemical pathways: a comparison of global optimization methods. Genome research 13: 2467–2474.
10. Nieman R, Fisher D, Seborg D (1971) A review of process identification and parameter estimation techniques†. International Journal of Control 13: 209–264.
11. Beck JV, Arnold KJ (1977) Parameter estimation in engineering and science. James Beck.
12. Young P (1981) Parameter estimation for continuous-time models—a survey. Automatica 17: 23–39.
13. Beck JV, Woodbury KA (1998) Inverse problems and parameter estimation: integration of measurements and analysis. Measurement Science and Technology 9: 839.
14. Hooke R, Jeeves TA (1961) "direct search"solution of numerical and statistical problems. Journal of the ACM (JACM) 8: 212–229.
15. Nelder JA, Mead R (1965) A simplex method for function minimization. The computer journal 7: 308–313.
16. Moré JJ (1978) The levenberg-marquardt algorithm: implementation and theory. In: Numerical analysis, Springer. pp. 105–116.
17. Esposito WR, Floudas CA (2000) Deterministic global optimization in nonlinear optimal control problems. Journal of Global Optimization 17: 97–126.
18. Horst R, Tuy H (2013) Global optimization: Deterministic approaches. Springer Science & Business Media.
19. Kirkpatrick S, Gelatt CD, Vecchi MP, et al. (1983) Optimization by simulated annealing. science 220: 671–680.
20. Storn R, Price K (1997) Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. Journal of global optimization 11: 341–359.
21. Sun J, Garibaldi JM, Hodgman C (2012) Parameter estimation using metaheuristics in

systems biology: a comprehensive review. Computational Biology and Bioinformatics, IEEE/ACM Transactions on 9: 185–202.

22. Mendes P, Kell D (1998) Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation. Bioinformatics 14: 869–883.
23. Modchang C, Triampo W, Lenbury Y (2008) Mathematical modeling and application of genetic algorithm to parameter estimation in signal transduction: Trafficking and promiscuous coupling of g-protein coupled receptors. Computers in Biology and Medicine 38: 574–582.
24. Tsai KY, Wang FS (2005) Evolutionary optimization with data collocation for reverse engineering of biological networks. Bioinformatics 21: 1180–1188.
25. Wang FS, Su TL, Jang HJ (2001) Hybrid differential evolution for problems of kinetic parameter estimation and dynamic optimization of an ethanol fermentation process. Industrial & engineering chemistry research 40: 2876–2885.
26. Noman N, Iba H (2007) Inferring gene regulatory networks using differential evolution with local search heuristics. IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB) 4: 634–647.
27. Tashkova K, Korošec P, Šilc J, Todorovski L, Džeroski S (2011) Parameter estimation with bio-inspired meta-heuristic optimization: modeling the dynamics of endocytosis. BMC systems biology 5: 159.
28. Villaverde AF, Egea JA, Banga JR (2012) A cooperative strategy for parameter estimation in large scale systems biology models. BMC systems biology 6: 75.
29. Rodriguez-Fernandez M, Egea JA, Banga JR (2006) Novel metaheuristic for parameter estimation in nonlinear dynamic biological systems. BMC bioinformatics 7: 483.
30. Egea JA, Rodríguez-Fernández M, Banga JR, Martí R (2007) Scatter search for chemical and bio-process optimization. Journal of Global Optimization 37: 481–503.

31. Villaverde AF, Henriques D, Smallbone K, Bongard S, Schmid J, et al. (2015) Biopredyn-bench: a suite of benchmark problems for dynamic modelling in systems biology. BMC systems biology 9: 8.
32. Fan M, Kuwahara H, Wang X, Wang S, Gao X (2015) Parameter estimation methods for gene circuit modeling from time-series mrna data: a comparative study. Briefings in bioinformatics : bbv015.
33. Gutenkunst RN, Waterfall JJ, Casey FP, Brown KS, Myers CR, et al. (2007) Universally sloppy parameter sensitivities in systems biology models. PLoS computational biology 3: e189.
34. Song SO, Chakrabarti A, Varner JD (2010) Ensembles of signal transduction models using pareto optimal ensemble techniques (poets). Biotechnology journal 5: 768–780.
35. Luan D, Szlam F, Tanaka KA, Barie PS, Varner JD (2010) Ensembles of uncertain mathematical models can identify network response to therapeutic interventions. Molecular bioSystems 6: 2272–2286.
36. Sagar A, Varner JD (2015) Dynamic modeling of the human coagulation cascade using reduced order effective kinetic models. Processes 3: 178–203.
37. Tolson BA, Shoemaker CA (2007) Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. Water Resources Research 43.
38. Mann KG, Butenas S, Brummel K (2003) The dynamics of thrombin formation. Arteriosclerosis, thrombosis, and vascular biology 23: 17–25.
39. Mann K, Brummel K, Butenas S (2003) What is all that thrombin for? Journal of Thrombosis and Haemostasis 1: 1504–1514.
40. Mann KG (2003) Thrombin formation. CHEST Journal 124: 4S–10S.
41. Vogler EA, Siedlecki CA (2009) Contact activation of blood-plasma coagulation. Bio-

materials 30: 1857–1869.

42. Diamond SL (2013) Systems biology of coagulation. Journal of Thrombosis and Haemostasis 11: 224–232.

43. Fogelson AL, Tania N (2005) Coagulation under flow: the influence of flow-mediated transport on the initiation and inhibition of coagulation. Pathophysiology of haemostasis and thrombosis 34: 91–108.

44. Anand M, Rajagopal K, Rajagopal K (2003) A model incorporating some of the mechanical and biochemical factors underlying clot formation and dissolution in flowing blood: review article. Journal of Theoretical Medicine 5: 183–218.

45. Hockin MF, Jones KC, Everse SJ, Mann KG (2002) A model for the stoichiometric regulation of blood coagulation. Journal of Biological Chemistry 277: 18322–18333.

46. Chatterjee MS, Denney WS, Jing H, Diamond SL (2010) Systems biology of coagulation initiation: kinetics of thrombin generation in resting and activated human blood. PLoS computational biology .

47. Mann KG, Brummel-Ziedins K, Orfeo T, Butenas S (2006) Models of blood coagulation. Blood Cells, Molecules, and Diseases 36: 108–117.

48. Luan D, Zai M, Varner JD (2007) Computationally derived points of fragility of a human cascade are consistent with current therapeutic strategies. PLoS computational biology 3: e142.

49. Zhao SZ, Liang JJ, Suganthan PN, Tasgetiren MF (2008) Dynamic multi-swarm particle swarm optimizer with local search for large scale global optimization. In: Evolutionary Computation, 2008. CEC 2008.(IEEE World Congress on Computational Intelligence). IEEE Congress on. IEEE, pp. 3845–3852.

Table 1: Error Analysis.

Simulation	Concentration	Normalized Standard Error	R-Squared Value
FVIIa-TF level	5nM	0.0367	0.7680
	500pM	0.0636	0.8939
	50pM	0.1204	0.8384
	10pM	0.0813	0.9492
	5pM	0.0574	0.9794

Table 2: Table with optimization settings and results for coagulation and different benchmarks using DdPSO.

	Coagulation	Benchmarks			
	Experimental Data Set	B1	B2	B4	B5
Ackley 300D	Rastrigin 300D				
Lower Bound	0.001.pnom	5.pnom	10.pnom	5.pnom	varying
30	5.12				
Upper Bound	1000.pnom	0.2.pnom	0.1.pnom	0.2.pnom	varying
-15	-5.12				
CPU Time	10.0835 hours	65.26 hours	17.9 minutes	6.2 minutes	32.686 minutes
2.8 seconds	2.6 seconds				
Function Evaluations	4000	4000	4000	4000	4000
4000	4000				
Initial Objective Value	1.11837e+07	1.3824e+07	3.8848e+03	1.8536e+07	1.4288e+04
21.12	99.985				
Final Objective Value	90456.9	3.5348e+04	324.4182	38.9375	3.4495e+03
0	0				
Nominal Objective Value	4.7785e+06	1.0986e+06	-	39.0676	4.2737e+03
0	0				

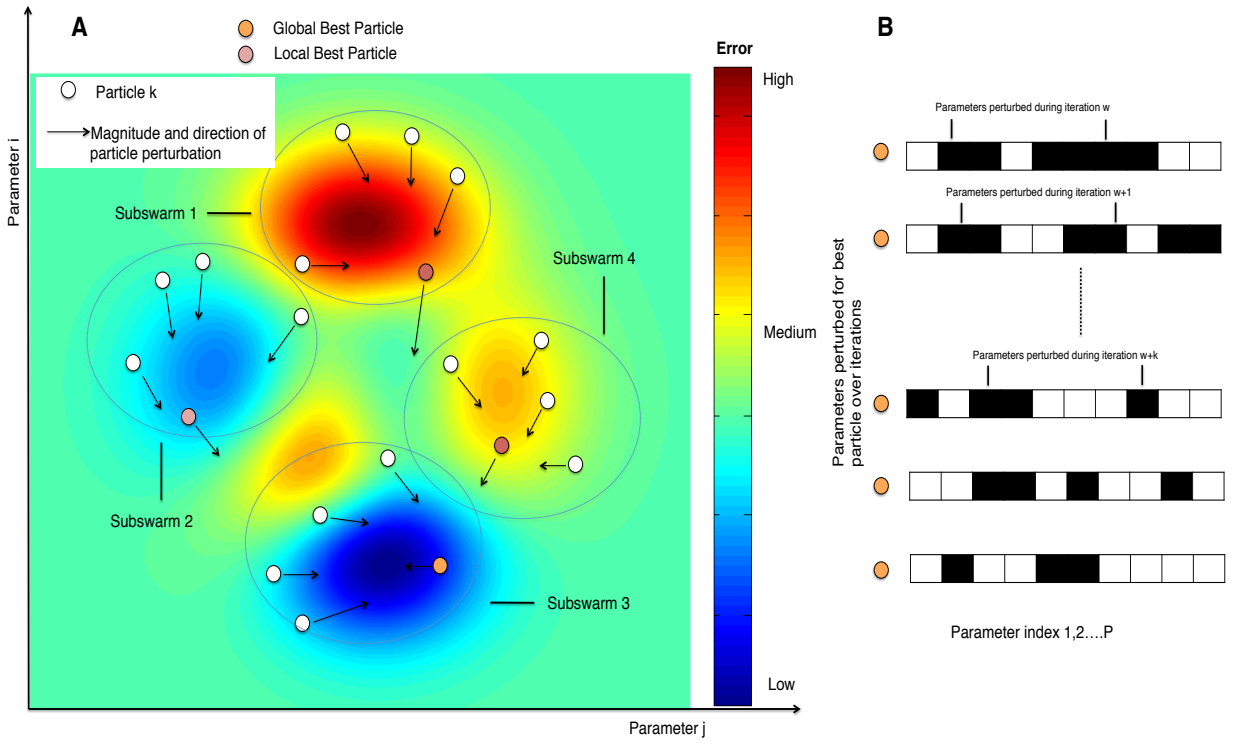

Fig. 1: Multi Swarm Particle Swarm Optimization with Dynamically Dimensioned Search. **A**: Each particle represents an N dimensional parameter vector. Particles are randomly initialized and grouped into different sub-swarms. Within each swarm the magnitude and direction of the movement a particle is influenced by the position of the best particle and also by its own experience. After every g number of function evaluations the particles are mixed and randomly assigned to different swarms. At the end of the evaluations ($\mathcal{FR}$) assigned to swarm search, the global best particle (orange color) amongst all sub-swarms is chosen as the candidate parameter vector for Dynamically Dimensioned Search **B**: The candidate vector does a greedy search in a dynamic neighborhood. The search neighborhood is dynamically adjusted by varying the number of dimensions that are perturbed (in black) in each evaluation step. The probability that a dimension is perturbed decreases as the number of function evaluations increase. Thus as the evaluations increase the optimality of the solution is preserved.

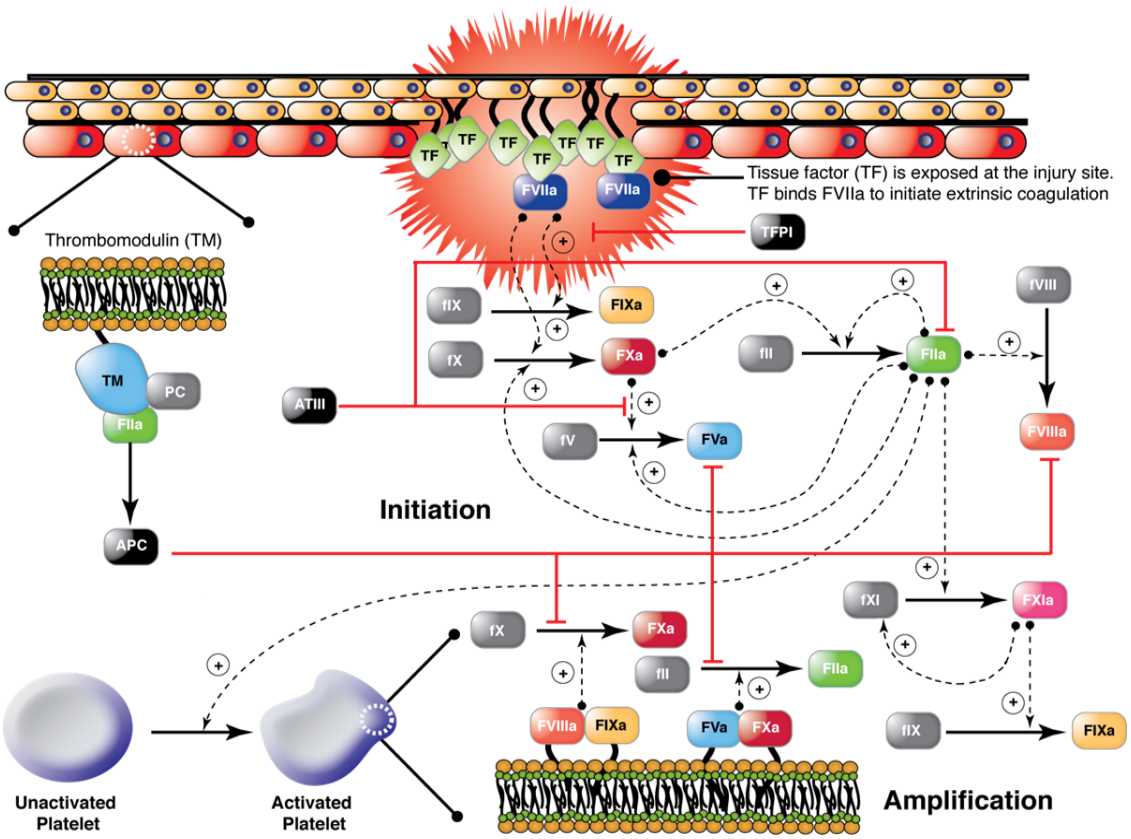

Fig. 2: Schematic of the extrinsic and intrinsic coagulation cascade. Inactive zymogens upstream (grey) are activated by exposure to tissue factor (TF) following vessel injury. Tissue factor and activated factor VIIa (FVIIa) form a complex that activates factor X (fX) and IX (fIX). FXa activates downstream factors including factor VIII (fVIII) and fIX. Factor V (fV) is primarily activated by thrombin (FIIa). In addition, we included a secondary fV activation route involving FXa. FXa and FVa form a complex (prothrombinase) on activated platelets that converts prothrombin (fII) to FIIa. FIXa and FVIIIa can also form a complex (tenase) on activated platelets which catalyzes FXa formation. Thrombin also activates upstream coagulation factors, forming a strong positive feedback ensuring rapid activation. Tissue factor pathway inhibitor (TFPI) downregulates FXa formation and activity by sequestering free FXa and TF-FVIIa in a FXa-dependent manner. Antithrombin III (ATIII) inhibits all proteases. Thrombin inhibits itself binding the surface protein thrombomodulin (TM). The IIa-TM complex catalyzes the conversion of protein C (PC) to activated protein C (APC), which attenuates the coagulation response by the proteolytic cleavage of fV/FVa and fVIII/FVIIIa.

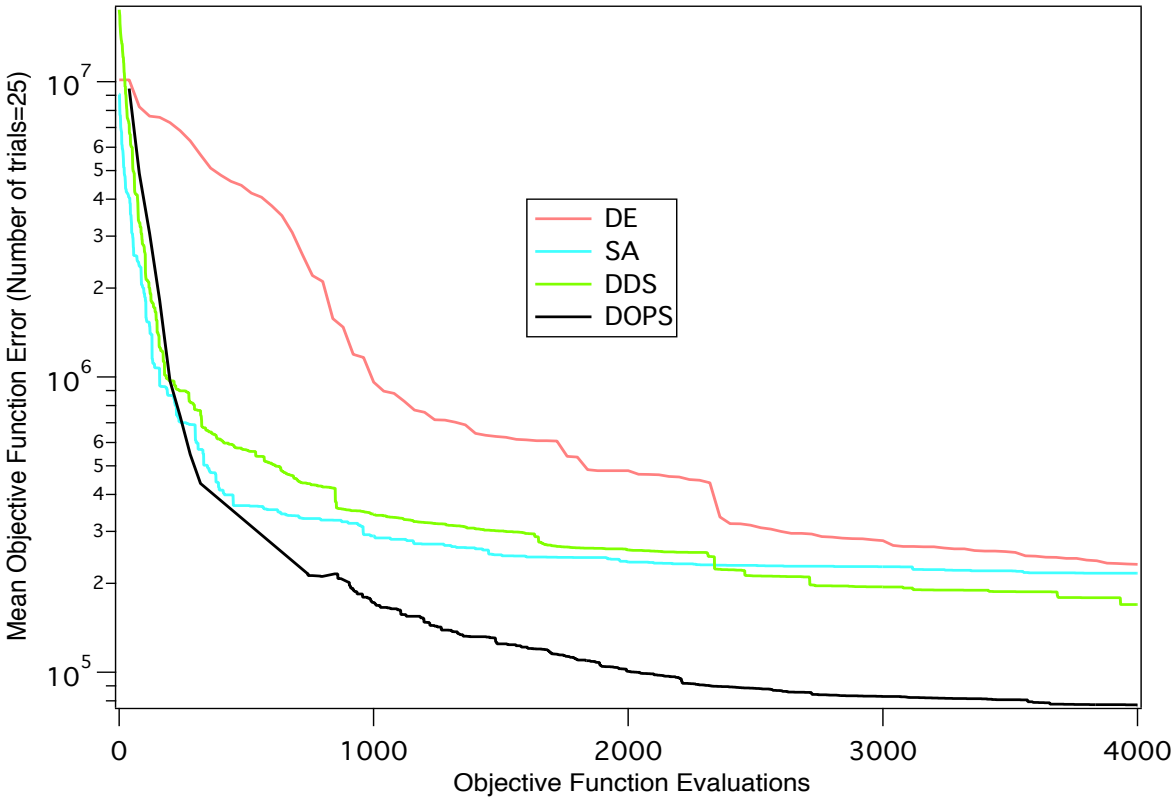

Fig. 3: Error convergence rates of the 5 different algorithms on the coagulation model. The objective error is the mean over N= 25 trials. DDSMLSPSO and DDS have the steepest drop in error during first 2000 function evaluations. Thereafter the error drop in DDS remains nearly constant whereas DDSMLSPSO drops further from around 2000 to 4000 function evaluations. At the end of 4000 function evaluations DDSMLSPSO attains the lowest error.

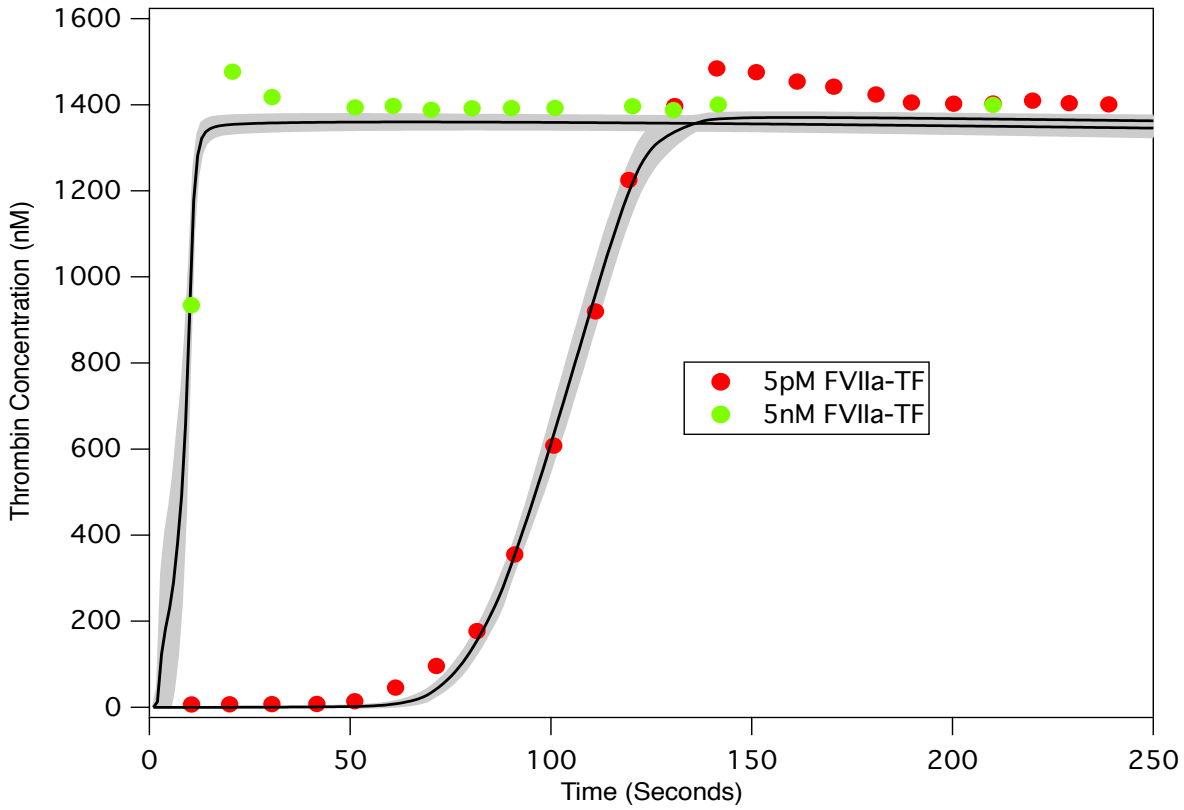

Fig. 4: Model fits on experimental data using DDSMLSPSO. The model parameters were estimated using DDSMLSPSO. Solid black lines indicate the simulated mean thrombin concentration using parameter vectors from N= 25 trials. The grey shaded region represents the 99% confidence estimate of the mean simulated thrombin concentration. The experimental data is reproduced from the synthetic plasma assays of Mann and co-workers. Thrombin generation is initiated by adding Factor VIIa-TF (5nM - Red and 5pM - Green) to synthetic plasma containing 200 μmol/L of phospholipid vesicles (PCPS) and a mixture of coagulation factors (II,V,VII,VIII,IX,X and XI) at their mean plasma concentrations.

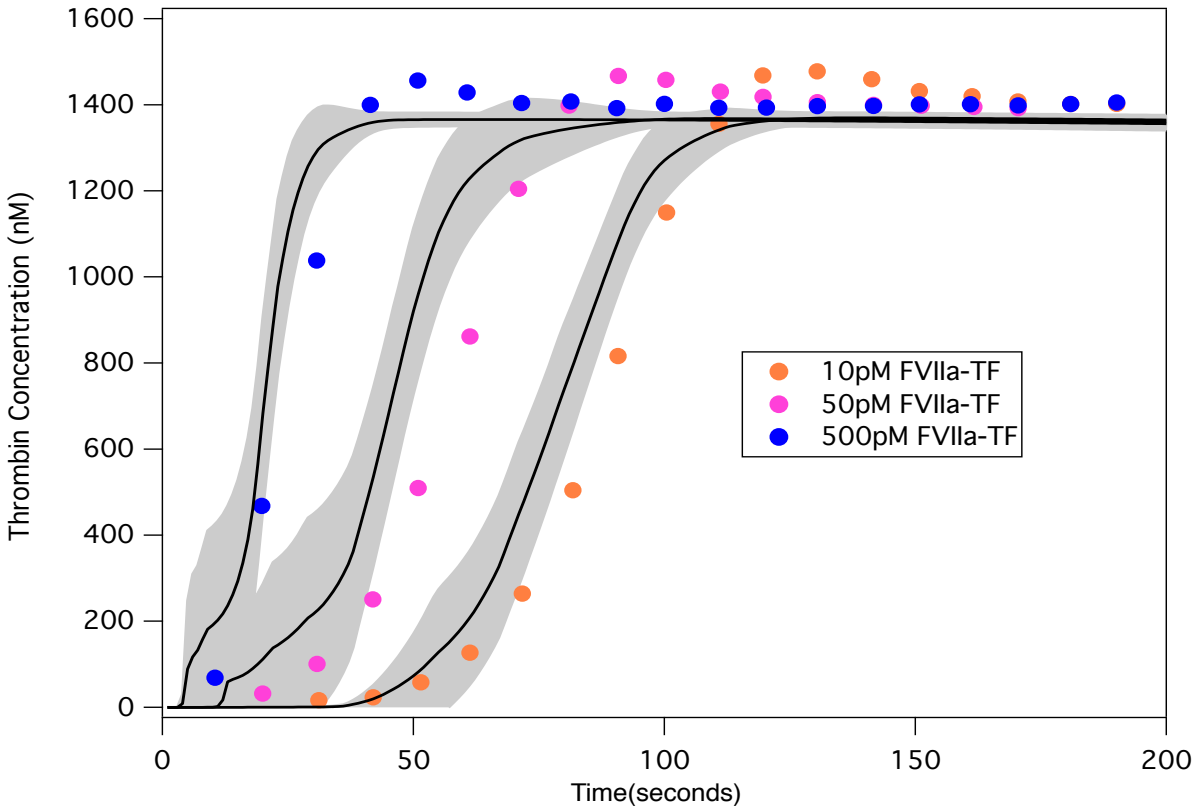

Fig. 5: Model predictions on unseen experimental data using parameters obtained from DDSMLSPSO. The parameter estimates that were obtained using DDSMLSPSO were tested against data that was not used in the model training. Solid black lines indicate the simulated mean thrombin concentration using parameter vectors from N= 25 trials. The grey shaded region represents the 99% confidence estimate of the mean simulated thrombin concentration. The experimental data is reproduced from the synthetic plasma assays of Mann and co-workers. Thrombin generation is initiated by adding Factor VIIa-TF (500pM - Blue, 50pM - Pink and 10pM - Orange respectively) to synthetic plasma containing 200 μmol/L of phospholipid vesicles (PCPS) and a mixture of coagulation factors (II,V,VII,VIII,IX,X and XI) at their mean plasma concentrations.

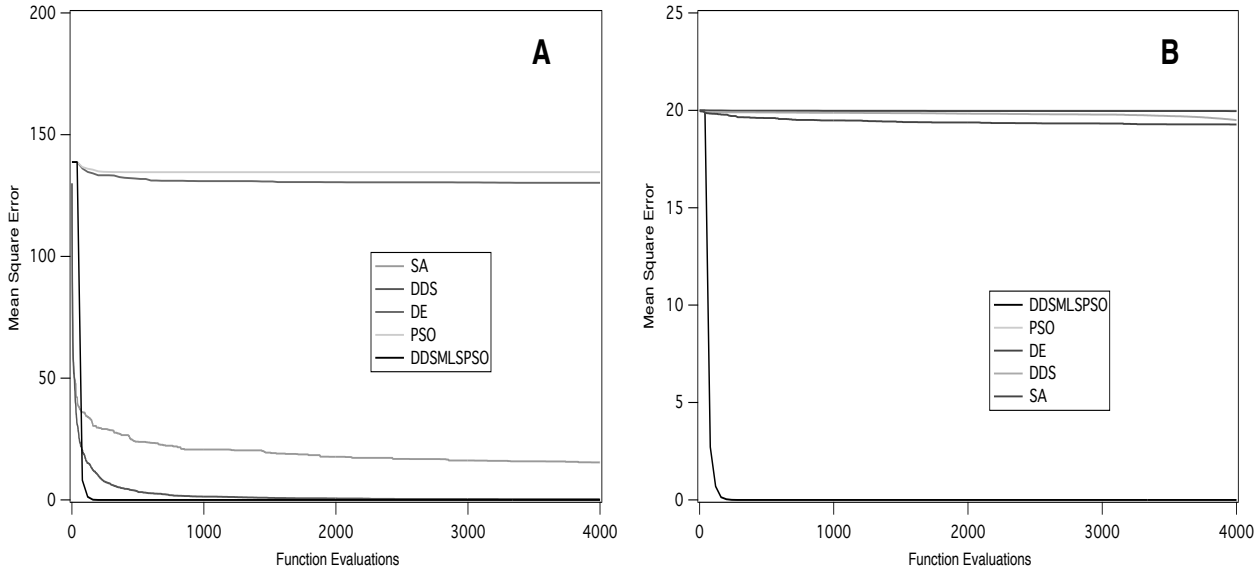

Fig. S1: Error convergence on 300 dimensional Ackley function and Rastrigin function. **(A)** Error convergence of 5 different meta-heuristics on 300 dimensional Ackley function **(B)** Error convergence of 5 different meta-heuristics on 300 dimensional Rastrigin function

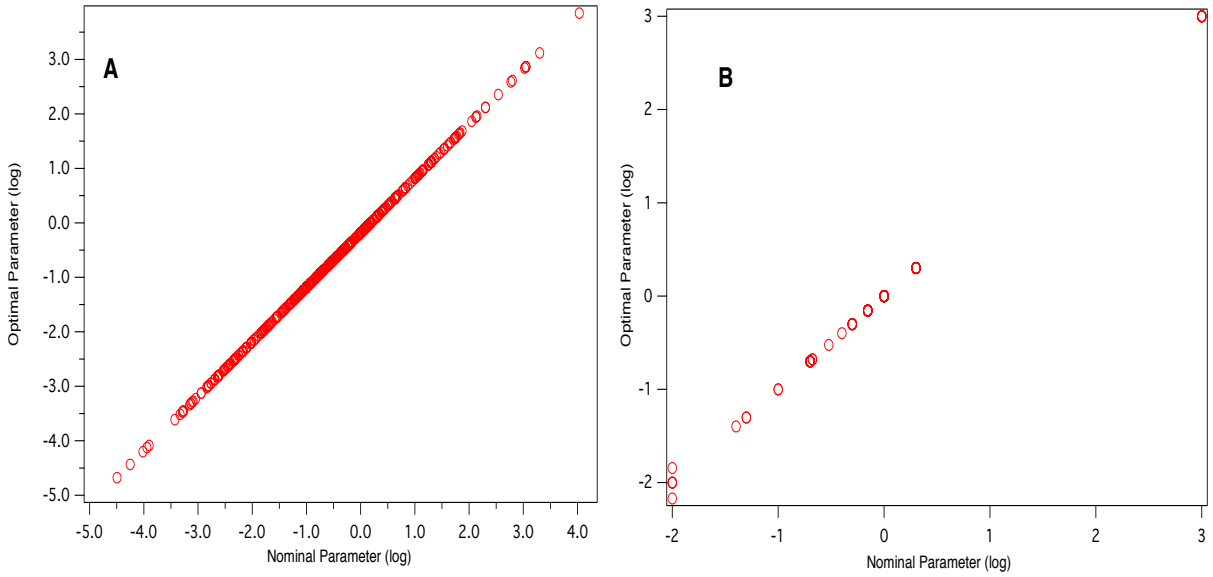

Fig. S2: Difference between optimal and nominal parameter vector values on benchmark problems. **(A)** Problem B1: Genome wide kinetic model of E.coli *S.cerevisiae* with 1759 unknown parameters. **(B)** Problem B4: Metabolic model of Chinese Hamster Ovary Cells (CHO) cells with 117 parameters.