BIOINFORMATICS ORIGINAL PAPER
Vol. 24 no. 8 2008, pages 1085–1092
doi:10.1093/bioinformatics/btn075

Systems biology

Inference of biochemical network models in S-system using multiobjective optimization approach

Pang-Kai Liu and Feng-Sheng Wang*
Department of Chemical Engineering, National Chung Cheng University, Chiayi 621-02, Taiwan, ROC

Received on November 20, 2007; revised on February 5, 2008; accepted on February 22, 2008
Advance Access publication March 5, 2008
Associate Editor: John Quackenbush

ABSTRACT

Motivation: The inference of biochemical networks, such as gene regulatory networks, protein–protein interaction networks, and metabolic pathway networks, from time-course data is one of the main challenges in systems biology. The ultimate goal of inferred modeling is to obtain expressions that quantitatively understand every detail and principle of biological systems. To infer a realizable S-system structure, most articles have applied sums of magnitude of kinetic orders as a penalty term in the fitness evaluation. How to tune a penalty weight to yield a realizable model structure is the main issue for the inverse problem. No guideline has been published for tuning a suitable penalty weight to infer a suitable model structure of biochemical networks.

Results: We introduce an interactive inference algorithm to infer a realizable S-system structure for biochemical networks. The inference problem is formulated as a multiobjective optimization problem to minimize simultaneously the concentration error, slope error and interaction measure in order to find a suitable S-system model structure and its corresponding model parameters. The multiobjective optimization problem is solved by the ε-constraint method to minimize the interaction measure subject to the expectation constraints for the concentration and slope error criteria. The theorems serve to guarantee the minimum solution for the ε-constrained problem to achieve the minimum interaction network for the inference problem. The approach could avoid assigning a penalty weight for sums of magnitude of kinetic orders.

Contact: chmfsw@ccu.edu.tw

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 INTRODUCTION

The rapid development of systems biology over the past few years has been driven by the advances in experimental methods that generate *in vivo* time-course data characterizing biochemical network interactions. In recent years, researchers intend to use such data for inferring a model structure and its parameters in order to examine intracellular dynamic behaviors on a systemic level. The ultimate goal of inferred modeling is to obtain expressions that quantitatively understand every detail and principle of biological systems (Chang *et al.*, 2005). How to select a suitable model structure and to estimate the parameter values involved is the main issue for mathematical modeling (Maki *et al.*, 2001; Mendes and Kell, 1998; Tsai and Wang, 2005).

Given a model structure, parameter estimation remains the limiting step in the modeling of biological systems. There exists, however, no unique method for estimating model parameters for nonlinear dynamic models. Most of the traditional nonlinear regression algorithms involving gradient methods have the possibility of getting trapped at local optima, depending upon the degree of system nonlinearity and the initial starting point (Mendes and Kell, 1998). Alternating regression (Chou *et al.*, 2006) dissects the nonlinear inverse problem of estimating parameter values into iterative steps of linear regression. The branch and bound algorithm (Polisetty *et al.*, 2006) is employed to convert the inverse problem of generalized mass action (GMA) or S-system into a convex optimization problem in order to obtain a global solution. Stochastic optimization methods, such as genetic algorithms, evolution strategy and simulated annealing (Edwards *et al.*, 1998; Gonzalez *et al.*, 2007; Moles *et al.*, 2003), are applied for parameter estimation in order to find a global solution. Many techniques have been employed to alleviate numerical integration burden. Voit and Almeida (2004) utilized a decoupling scheme to estimate the slopes of the dynamic processes. Tsai and Wang (2005) used the modified collocation method to approximate dynamic profiles at sampling points. The decomposing method (Kimura *et al.*, 2005; Maki *et al.*, 2002) is employed to convert the large network inference problem into subproblems. Such approximation techniques can be easily incorporated into an optimization method to avoid the computationally expensive numerical integrations for fitness evaluations.

To infer a realizable S-system structure, most articles have applied sums of magnitude of kinetic orders as a penalty term in the fitness evaluation (Ho *et al.*, 2005; Kikuchi *et al.*, 2003; Kimura *et al.*, 2004, 2005; Noman and Iba, 2005). A weighting factor in the penalty term needs to be carefully tuned in order to infer a realizable S-system model structure. The weighting factor in general depends on the problem of interest. An improper weighting factor should make to yield a wrong structure. According to our knowledge, no guideline has been published for tuning a suitable penalty weight to infer model structures of biochemical networks. In this study, we introduce the multiple-objective optimization approach to inferring a

*To whom correspondence should be addressed.

realizable S-system structure for biochemical networks. Such an approach can avoid assigning a weighting factor for sums of magnitude of kinetic orders. One dry-lab and one wet-lab case studies are made to illustrate the performance of the proposed approach.

2 METHODS

A biochemical network system can be modeled as a set of S-system canonical forms (Voit, 2000):

$$\dot{X}_i = f_i(\mathbf{X}, \mathbf{p}) = \alpha_i \prod_{j=1}^{n+m} X_j^{g_{ij}} - \beta_i \prod_{j=1}^{n+m} X_j^{h_{ij}}, i = 1, \ldots, n \tag{1}$$

where X_i is the ith component or pool in the network, the model parameter vector $\mathbf{p}$ consists of rate constants, α_j and β_j, and kinetic orders, g_{ij} and h_{ij}. f_i is the net rate equation, which consists of both influx and efflux. The m-dimensional independent variables in the S-system equations are expressed as X_{n+j}, $j = 1, \ldots, m$. The parameter estimation is to determine model parameters, rate constants and kinetic orders, so that the dynamic profiles fit satisfactorily the measured observation.

2.1 Estimation criteria

The canonical biochemical network inference problem is formulated as a function optimization problem to minimize an objective function that measures the goodness-of-fit of the model with respect to a given experimental time-course dataset. The least-squared error criterion is a commonly used objective function and is expressed as

$$J_1 = \frac{1}{nN_s} \sum_{i=1}^{n} \sum_{j=1}^{N_s} \frac{\left(X_{e_i}(t_j) - X_i(t_j)\right)^2}{X_{e_i \max}^2} \tag{2}$$

where $X_{e_i}(t_j)$ denotes the measured data for the ith component at the sampling time t_j, $X_i(t_j)$ is the computed concentration for the ith component at the sampling time t_j, and $X_{e_i \max}$ is the maximum measured concentration of the ith component. Here, N_s denotes the number of sampled data points. The dynamic profiles $X_i(t_j)$ are in general obtained by applying a numerical integration method to solve the differential Equation (1). Numerical integration for parameter estimation is time consuming and may cause a run-time error during the computation progress. Wang (2000) has used the modified collocation method to convert ordinary differential equations into algebraic equations. The piecewise linear Lagrange polynomial is the simplest shape polynomial for obtaining the approximate dynamic profiles. The approximate equations are expressed as:

$$X_i(t_l) \cong X_i(t_{l-1}) + 0.5\eta_l\left\{f_i(\mathbf{X}(t_l), \mathbf{p}) + f_i(\mathbf{X}(t_{l-1}), \mathbf{p})\right\}, i = 1, \ldots, n \tag{3}$$

Tsai and Wang (2005) have employed the algebraic equations to generate approximate profiles for parameter estimation in order to avoid solving the equations recursively. Such an approximation not only reduces computation time, but also converts the coupled algebraic equations into a set of uncoupled equations so that parallel computation can be straightforwardly applied for the parameter estimation.

The least-squared error criterion (2) is to directly employ concentration profiles of the system for evaluating fitness of the estimation. This error criterion refers to the concentration error in this study. An alternative error criterion is to use the slope information for evaluating fitness of the estimation (Voit and Almeida, 2004). The slope error criterion is therefore expressed as:

$$J_2 = \frac{1}{nN_s} \sum_{i=1}^{n} \sum_{j=1}^{N_s} \frac{\left(\dot{X}_{e_i}(t_j) - \dot{X}_i(t_j)\right)^2}{\dot{X}_{e_i \max}^2} \tag{4}$$

where $\dot{X}_{e_i}(t_j)$ is the approximate experimental slope for the ith component at the sampling time t_j, $\dot{X}_i(t_j)$ is the computed slope for the ith component at t_j, and $\dot{X}_{e_i \max}$ is the maximum slope of the ith component. Using the model slope to compute the error criterion can avoid the numerical integration of differential equations so it alleviates the computational burden. However, a smoothing filter, such as artificial neural network or spline smoothing, has to be utilized to smooth the measured data in order to generate the approximate experimental slopes for each variable (Almeida and Voit, 2003).

Inference of regulatory interactions in a biochemical system provides fundamental biological knowledge and significant efforts. Several network inference algorithms estimate all of the S-system parameters from time-course data. The estimation for a large-scale S-system often causes bottlenecks, and fitting the model to experimentally observed data is not simple. The decoupling approach, such as modified collocation method (Tsai and Wang, 2005), slope approximation (Voit and Almeida, 2004) and decomposition method (Kimura *et al.*, 2004), enables us to infer S-system models of genetic networks of a larger scale. To detect a suitable model structure for a large-scale S-system, the sum of magnitude of kinetic orders can be employed as a criterion to pruning a skeletal structure, and is expressed as (Kikuchi *et al.*, 2003; Kimura *et al.*, 2004; Voit and Almeida, 2003):

$$J_3 = \sum_{i=1}^{n} \sum_{j=1}^{I} \left(|g_{ij}| + |h_{ij}|\right) \tag{5}$$

where I is a set of cardinal numbers indicating which kinetic orders should be pruned. The kinetic orders g_{ij} and h_{ij} for S-systems quantify the regulation effect of X_j on the production or degradation of X_i so that less interaction means that the magnitude of the corresponding kinetic order is small. When there is no interaction between X_j and X_i, the S-system parameter values corresponding to the interaction (g_{ij} or h_{ij}) are zero. The criterion (5) referred to as the interaction measure (or sensitivity) therefore serves as the evaluation factor to prune smaller interactive kinetic orders. If the decoupling approach is applied to each subproblem, the concentration error criterion, slope error criterion and interaction measure are evaluated for each component only.

2.2 Multiobjective optimization approach

The aim of this study is to minimize simultaneously the concentration error, slope error and interaction measure in order to find a suitable S-system model structure and its corresponding model parameters. The multiobjective parameter estimation problem is therefore expressed as:

$$\min_{\mathbf{p} \in \Omega} \{J_1, J_2, J_3\} \tag{6}$$

where the feasible region $\mathbf{\Omega}$ is a set of all admissible model parameters $\mathbf{p}$ that satisfy the corresponding S-system model Equation (1).

Multiobjective optimization is a natural extension of the traditional optimization of a single-objective function. Typically, the objectives are incommensurable and often (partially or wholly) in conflict (Handl *et al.*, 2007). The incommensurability between multiobjective functions gives rise to the distinguishing difference between multiobjective optimization and traditional single-objective optimization. This fact leads to the lack of a complete order for multiobjective optimization problems. Concept of Pareto optimality or noninferiority is therefore employed to characterize a solution to multiobjective optimization problems. The definition of Pareto optimal solution is introduced as follows:

Definition. *A vector of the model parameters, $\mathbf{p}^*$, is the Pareto optimal point if and only if there does not exist $\mathbf{p} \in \mathbf{\Omega}$ such that*

$$J_j(\mathbf{p}) \le J_j(\mathbf{p}^*),\ j = 1, 2, 3$$

$$J_k(\mathbf{p}) < J_k(\mathbf{p}^*),\ \text{for some } k$$

The image of the Pareto optimal point is the Pareto optimal solution. The Pareto optimal point means that it is impossible to improve in any objective without a simultaneous worsening in some other objectives.

The literature on multiobjective optimization is abundant (Sakawa, 1993), and we cannot hope to mention all the techniques that have been employed to generate a Pareto solution; however, one method is pervasive in multiobjective optimization literature. This technique is the weighted sum method for converting a multiobjective optimization problem such as (6) into a single-objective function problem. Such an approach is equivalent to introducing a penalty term to join with the concentration error criterion or slope error criterion as discussed in the literature (Kikuchi *et al.*, 2003; Kimura *et al.*, 2004; Tsai and Wang, 2005; Voit and Almeida, 2004). The penalty problem is therefore expressed as:

$$\min_{\mathbf{p}\in\Omega} J_1(\text{or } J_2) + \varpi \sum_{i=1}^{n}\sum_{j=1}^{I}\left(|g_{ij}| + |h_{ij}|\right) \tag{7}$$

The optimal estimates to the penalty problem (7) depend on the chosen weighting factor ϖ. The weighting factor needs to be carefully tuned in order to infer a realizable S-system model structure. No guideline has been published for tuning a suitable penalty weight to infer model structures of regulatory networks. In this study, we introduce the ε-constraint method to overcome such a drawback.

The ε-constraint method for characterizing the Pareto optimal estimates is to solve the following constraint problem formulated by taking one criterion as the objective function and letting all other criteria be inequality constraints (Sakawa, 1993). The first goal of this study is to find a suitable S-system structure so the constraint problem is formulated as:

$$\min_{\mathbf{p}\in\Omega} \frac{J_3(\mathbf{p})}{J_3^E} \tag{8}$$

subject to

$$C_1(\mathbf{p}) = \frac{J_1(\mathbf{p})}{J_1^E} - 1 \le 0 \tag{9}$$

$$C_2(\mathbf{p}) = \frac{J_2(\mathbf{p})}{J_2^E} - 1 \le 0 \tag{10}$$

where $J_i^E, i = 1, 2, 3$ are the expected values for the concentration criterion, slope criterion and interaction measure, respectively. In the following section, we will introduce an interactive computational algorithm to rationally provide these expected values. The relationships between the optimal solution $\mathbf{p}^*$ to the constraint problem and the Pareto optimal concept of the primal multiobjective parameter estimation problem (6) can be characterized by the following theorems.

Theorem 1. *If $\mathbf{p}^*\in\mathbf{\Omega}$ is a unique optimal solution of the constraint problem for some J_1^E and J_2^E, then $\mathbf{p}^*$ is the Pareto optimal solution to the multiobjective parameter estimation problem (6).*

Theorem 2. *If $\mathbf{p}^*\in\mathbf{\Omega}$ is a Pareto optimal solution of the multiobjective parameter estimation problem (6), then $\mathbf{p}^*$ is the optimal solution of the constraint problem for some J_1^E and J_2^E.*

Both theorems can be immediately proved from the definition of the Pareto optimality by making use of contradictory arguments following the similar procedures discussed in the textbook (Sakawa, 1993), and are expressed in the Supplementary 1. This fact indicates that a Pareto optimal estimate for the multiobjective parameter estimation problem (6) can be obtained by solving the converted constraint problems (8)–(10) using a global optimization method. Several constrained optimization methods can be employed to solve the converted constraint problem. In this study, the popular penalty-function method is introduced to solve the constraint problem. The inference problem is therefore expressed as:

$$\min_{\mathbf{p}\in\Omega} J = \frac{J_3(\mathbf{p})}{J_3^E} + \omega\left\langle\max\left\{C_1(\mathbf{p}), C_2(\mathbf{p})\right\}\right\rangle_+^2 \tag{11}$$

where the bracket operation in (11) is defined as $\langle C(\mathbf{p})\rangle_+ = \max\{C(\mathbf{p}), 0\}$. The second term in (11) indicates that a penalty is desired only if the point $\mathbf{p}$ is not feasible. If any or both C_1 and C_2 are positive, the worst value is employed to compute the penalty. If both inequality constraints are feasible, i.e. $\max\{C_1, C_2\} \le 0$, then the penalty is zero, i.e. $\langle C(\mathbf{p})\rangle_+ = 0$. This situation indicates that no penalty is incurred. From this result, it is clear that we can get arbitrarily close to the optimal interaction measure value of the constraint problem (8)–(10) by computing the inference problem (11) for a sufficiently large $\boldsymbol{\omega}$. To search easily a feasible point, the penalty parameter can be provided to be greater than the inverse for the minimum of the expected values, i.e. $\omega > 1/\max\{J_1^E, J_2^E\}$. Theorem 3 that serves to guarantee the minimum solution for the inference problem is also the optimal estimate for the constraint problems (8)–(10).

Theorem 3. *If $\langle C(\mathbf{p}_\omega)\rangle_+ = 0$ and $\mathbf{p}_\omega \in \Omega$ for some ω then $\mathbf{p}_\omega$ is a minimum solution to the constraint problems (8)–(10).*

Theorem 3 can be immediately proved following the similar procedures discussed in the textbook (Bazaraa and Shetty, 1979), and are expressed in the Supplementary 1. The aim of the theorem is to determine a feasible point $\mathbf{p}$ to the inference problem (11), i.e. both concentration and slope error criteria are less than their expected values, such that the interaction measure is minimized. Using this theorem, we introduce an interactive algorithm as shown in Table 1 for inferring biochemical regulatory networks. In Steps 1 and 3 of Table 1, we used the hybrid differential evolution (HDE) to minimize each corresponding objective function toward obtaining a global optimal solution. HDE enables a smaller population to be used for finding a global solution (Chiou and Wang, 1999) and has succeeded in solving several biochemical optimization problems (Wang and Sheu, 2000).

Equations (9) and (10) are the inequality constraints for the constraint problems (8)–(10). All parameters those satisfy both inequality constraints make up a feasible set, i.e. $\Sigma = \left\{\mathbf{p} \in \Omega : C_1(\mathbf{p}) \le 0 \text{ and } C_2(\mathbf{p}) \le 0\right\}$. The expected values, J_1^E and J_2^E, are individually obtained from solving the corresponding single objective parameter estimation problems (2) and (4). Therefore, the aim of the constraint problem (8)–(10) is to determine the optimal parameters within the feasible set such that the interaction measure is minimized. In other words, given a super-structure, each single objective parameter estimation problem is first solved in order to yield its expected value. The multiobjective parameter estimation is then applied to find a compromised solution. The solution must be less than the expected value, and makes the state variables of the super-structure among the smallest impact. Many constrained optimization methods can be employed to solve the constraint problems (8)–(10). In this study, the problem is converted to the inference problem (11) through a penalty-function method. HDE is then applied to solve the inference problem. A large punishment should be added to the inference problem (11) in order to move the searching parameters into the feasible set, when the searching parameters violate the constraints (9) and (10) in the solution process. When the searching parameters satisfy the constraints (9) and (10), no penalty is incurred. The inference problem can continuously minimize the interaction measure toward obtaining a suitable structure.

3 RESULTS

In this article, we show two case studies, an artificial genetic network and a wet-lab system, for inferring a suitable interaction network. The detail of the computational results and three additional case studies, a 30 gene network (Kimura

Table 1. Interactive inference algorithm for biochemical regulatory networks using hybrid differential evolution

1. Calculate the expected values, J_1^E and J_2^E, for the concentration error and slope error using hybrid differential evolution to minimize its single-objective parameter estimation problem (2) and (4), respectively. Let each expected value be its corresponding minimum error criterion.
2. Compute the sum of the magnitude of kinetic orders for each single-objective parameter estimation problem. Let each expected value of the interaction measure for each single-objective parameter estimation problem be J_{31}^E and J_{32}^E. The expected value, J_3^E, for the interaction measure is set as $J_3^E = \max\{J_{31}^E, J_{32}^E\}$
3. Let the parameter be $\omega > \max\{J_1^E, J_2^E\}^{-1}$, and solve the inference problem (11) using hybrid differential evolution. Let the minimum solution be $\mathbf{p}^*$.
4. If $\langle C(\mathbf{p}^*)\rangle_+$ is smaller than the tolerance, e.g. 1.0E−6, then go to Step 5; otherwise, stop the interactive inference algorithm.
5. Sort the kinetic orders, g_{ij} and h_{ij}, for the synthesis and degradation terms using the score $|g_{ij}|/\max\{|g_{ij}|,|h_{ij}|\}$ and $|h_{ij}|/\max\{|g_{ij}|,|h_{ij}|\}$, respectively.
6. Delete the smaller kinetic orders with scores less than the assigned value, e.g. 1.0E−2, and then repeat the interactive procedures to infer the pruned model.

et al., 2004), a circadian oscillations of period protein in Drosophila (Ingalls, 2004) and an embryonic gene regulatory network in zebrafish (Huang *et al.*, 2006), are also provided in the Supplementary 2 to illustrate the effectiveness of the proposed algorithm. All computations were carried out on a Pentium IV computer using Microsoft Windows XP. The interactive inference algorithm is implemented in Compaq Visual Fortran. HDE serves as a minimization solver in the interactive algorithm, and has to provide four setting factors by the user. The setting factors used for all runs in the case studies are listed as follows: The crossover factor is set to be 0.5. Two tolerances used in the migration are set to be 0.05. The population size of 5 is used in the computation.

3.1 Case I: small-scale gene network

The dry-lab case study is a two-gene regulatory network shown in Hlavacek and Savageau (1996). The 'true' system is described in the S-system equations as follows:

$$\begin{aligned}\dot{X}_1 &= 5X_3X_5^{-1}X_6 - 10X_1^2\\ \dot{X}_2 &= 10X_1^2X_7 - 10X_2^2\\ \dot{X}_3 &= 10X_2^{-1}X_8 - 10X_2^{-1}X_3^2\\ \dot{X}_4 &= 8X_3^2X_5^{-1}X_6 - 10X_4^2\\ \dot{X}_5 &= 10X_4^2X_7 - 10X_5^2\end{aligned} \tag{12}$$

where X_1 is an mRNA produced from gene 1, X_2 is an enzyme protein it produces and X_3 is an inducer protein catalyzed by X_2. X_4 is an mRNA produced from gene 4 and X_5 is a regulator protein it produces. Positive feedback from the inducer protein X_3 and negative feedback from the regulator protein X_5 are assumed in the mRNA production processes of genes 1 and 4. X_6, X_7 and X_8 denote a pool of nucleic acid, amino acid and substrate, respectively, and are considered as independent variables in the system.

We first consider noise-free time-course data for evaluating the penalty problem (7) and the constraint problems (8)–(10) for comparison. The eight sets of training data generated by Tsai and Wang (2005) were employed to infer an S-system model structure. In this example, we set the kinetic orders $g_{ii} = 0$, which precluded the direct effect of a variable on its own production and required the kinetic orders h_{ii} to be greater than zero, indicating that the degradation of compounds almost depends always on the concentration. The search ranges used for the regression were α_i and $\beta_i \in [0, 20]$, g_{ij} and $h_{ij} \in [-4, 4]$, $i \neq j$ and $h_{ii} \in [0, 4]$. The HDE algorithm was employed to solve the penalty problem (7) with the weighting factor of 10^{-3}, 10^{-4} and 10^{-6}, respectively. The computational results were shown in Supplementary 2. For the cases using the weighting factor of 10^{-3} and 10^{-6}, we cannot infer a convergent structure from the penalty problem (7) using several trials for HDE. For the weighting factor of 10^{-4}, the inferred structure is identical to the 'true system' after four iterations. We next applied the proposed interactive inference algorithm, as shown in Table 1, except using a different penalty parameter in Step 3, to solve the inference problem (11). In this computation, we used, respectively, the penalty parameters of 1, 10^2 and 10^4 in Step 3 to solve each inference problem (11). The proposed algorithm enabled us to infer the identical S-system structure to the 'true' system although the assigned penalty parameters were widely different. The computation time for each case was about the same. 38.8 min and two iterations required on a single-CPU Pentium IV 3.0 GHz. The proposed algorithm requires one-fifth CPU times that of the result solved by Tsai and Wang (2005) because it uses decoupling computation to solve each subsystem. Table 2 summarizes the comparison between the proposed algorithm and the reported methods for this inference problem (Kikuchi *et al.*, 2003; Kimura *et al.*, 2004; Tsai and Wang, 2005).

Next, we test the performance of the proposed method in a real-world situation by conducting the experiment with the sets of noisy time-course data. To imitate real profiles, 10% random noises are added into the eight sets of 'true' time-course data. The rational method in the curve-fitting toolbox for MATLAB is then employed to smooth the measured data in order to yield the noise-free time-course profiles for evaluating concentration error criterion and slope error criterion. In this work, the proposed interactive inference algorithm not only can infer the S-system model structure for dependent variables, as discussed in the previous run, but can also infer interaction relations between dependent and independent variables. The superstructure for the S-system is therefore expressed as:

$$\dot{X}_i = \alpha_i \prod_{j=1}^{5+3} X_j^{g_{ij}} - \beta_i \prod_{j=1}^{5+3} X_j^{h_{ij}}, i = 1, \ldots, 5 \tag{13}$$

Table 2. Comparison between the proposed interactive inference algorithm and three reported methods

Method	ODE solver	CPU time	Estimated result
This work	Modified collocation and slope approximation for each subsystem	38.8 min and two iterations required on a single-CPU Pentium IV 3.0 GHz	The parameter values are almost identical to the true values.
Modified collocation approximation	Modified collocation for the whole system	2.84 h required for all computation on a single-CPU Pentium IV 2.4GHz	The parameter values are almost identical to the true value.
PEACE 1	Numerical integration	10 h for one loop on PC cluster with 1040 CPUs and Pentium III 933 MHz	The structure is not completely identical to the true system. There exists h_{53}.
GLSDC	Decomposed numerical integration	58.8 min for minimizing each subsystem on a single-CPU Pentium III 1 GHz	This method could not estimate the parameter values with perfect precision.

Modified collocation approximation was reported by Tsai and Wang (2005). PEACE 1 was reported by Kikuchi *et al.* (2003). GLSDC was reported by Kimura *et al.* (2004).

In this example, three independent variables, a pool of nucleic acid, amino acid and substrate, are also included in the superstructure so there are 90 S-system parameters to be estimated. However, using the decoupling approach, parallel computation can be employed to infer each subsystem that includes 18 parameters only.

In the first run, the HDE algorithm was employed to solve each single-objective parameter estimation problem. The expected values for the first subsystem were ($J_1^E = 2.746\text{E}-3, J_{31}^E = 1.746\text{E}+1$) and ($J_2^E = 7.514\text{E}-2, J_{32}^E = 2.491\text{E}+1$), respectively. These expected values were then provided for the inference problem (11) in Step 2 of the interactive inference algorithm in Table 1. The optimal estimates for each subsystem, as shown in the Supplementary 2, were feasible. The optimal concentration error and slope error were $J_1^* = 2.506\text{E}-3$ and $J_2^* = 4.678\text{E}-2$, respectively, so the inequality constraints in (9) and (10) were less than zero, i.e. $\langle C(\mathbf{p}^*)\rangle_+ = 0$. Many kinetic orders g_{ij} and h_{ij} were very small. We then deleted those smaller kinetic orders with scores <0.01. The interactive inference algorithm was then repeated to refit each pruned subsystem. For the second iteration, the inferred regulatory structure for independent variables approached to the 'true' system. One iteration is enough to achieve the minimum connective network for the first subsystem, three iterations for the second, fourth and fifth subsystem. For the third subsystem, after the fifth iteration, the concentration error criterion and slope error criterion were 5.326E−3 and 1.449E−2, respectively, both of which were almost equal to their expected values. Both inequality constraints for the third subsystem were less than zero. The score of the kinetic order g_{35} was smaller than 1.0E−2, after deleting this parameter; the inferred S-system structure was essentially identical to the 'true' system. The estimated parameter values were employed to evaluate an extra test-experiment in order to validate the model. The initial condition for the test-experiment is beyond the training dataset. The 'true' dynamic profiles are shown as the data-points in Figure 1. The model profiles shown as solid curves are capable of predicting the dynamic responses under the condition. In order to yield more accurate estimates, the solution obtained by the proposed algorithm is employed as the initial starting point for a gradient-based method, a subroutine BCONF in IMSL Math/Library, to solve the parameter estimation problem. The local search procedure employs Runge–Kutta pairs of various orders, a subroutine IVMRK in IMSL Math/Library, to solve differential equations towards obtaining time-course profiles of the system. The refined estimates are then employed to evaluate the extra test-experiment to validate the model. The model profiles shown as dashed curves in Figure 1 can satisfactorily fit the test-experiments.

So far, the inference problem is to minimize simultaneously the concentration error, slope error and interaction measure in order to find the S-system model structure and its corresponding model parameters. The problem can be solved by minimizing two objective functions. That is to minimize simultaneously the concentration error and interaction measure or the slope error and interaction measure, respectively. The bi-objective minimization problem is therefore expressed as $\min\{J_1, J_3\}$ or $\min\{J_2, J_3\}$. The interactive inference algorithm is then employed to solve both problems, respectively. For noise-free time-course data, both bi-objective minimization problems can achieve the exact model structure.

However, for noisy time-course data, we cannot obtain the exact structure for minimizing $\{J_1, J_3\}$ after four iterations and $\{J_2, J_3\}$ after five iterations, respectively. The optimal estimates for each subsystem are shown in the Supplementary 2. The optimal estimates from $\min\{J_1, J_3\}$ and $\min\{J_2, J_3\}$ are then employed to evaluate the extra test-experiment to validate the model, respectively. Although both structures are different from the 'true' system, the model profiles shown as dashed-dot curves ($\min\{J_1, J_3\}$) and dashed-dot-dot curves ($\min\{J_2, J_3\}$) in Figure 1 can suitably fit the test-experiments.

3.2 Case studies II: kinetics model of ethanol fermentation

In this wet-lab case study, we analyze a batch fermentation process discussed by Wang *et al.* (2001). The fermentation

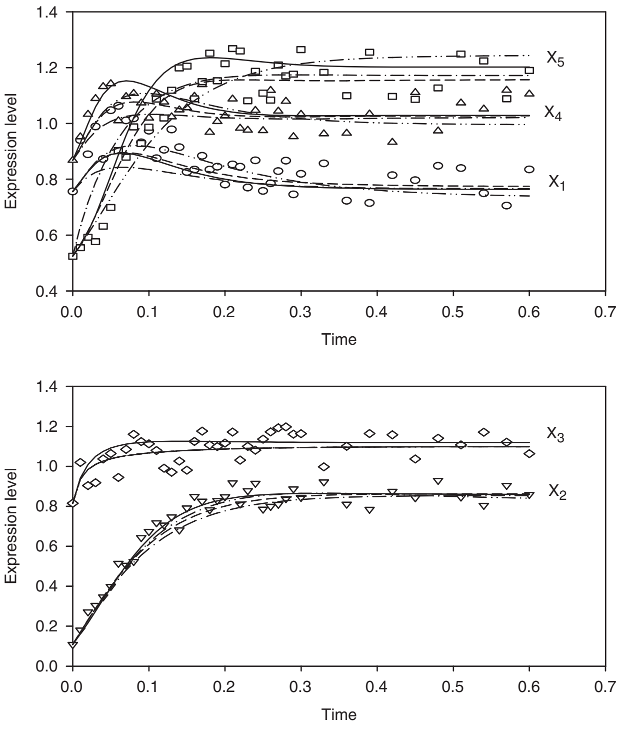

Fig. 1. Model validated using various estimated parameters. The symbolic data points are the 'true' dynamic profiles with 10% random noise using the initial condition [0.56, 0.08, 0.57, 0.54, 0.34, 0.8, 0.8, 0.8]. The solid curves are the computed profiles using the optimal estimates from min{J_1, J_2, J_3}. The dashed curves are the computed profiles using the refined estimates. The dashed-dot curves are the computed profiles using the optimal estimates from min{J_1, J_3}. The dashed-dot-dot curves are the computed profiles using the optimal estimates from min{J_2, J_3}.

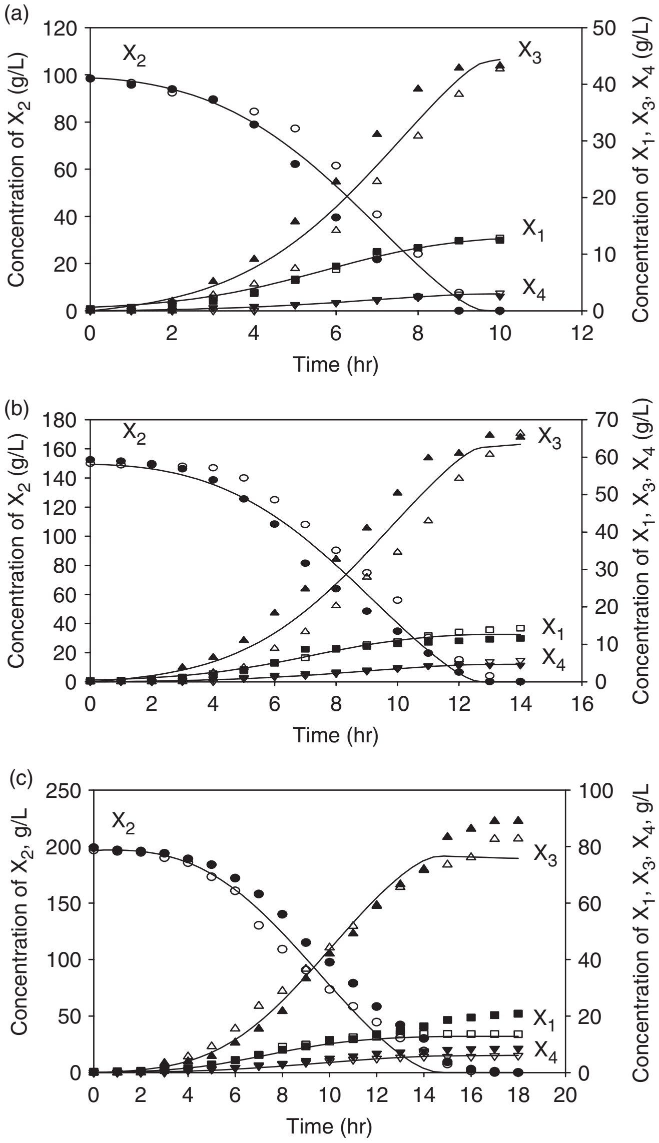

Fig. 2. Two batch fermentations using the initial glucose concentrations of (**a**) 100 g/L and (**b**) 150 g/L for inferring an S-system model, and one batch fermentation using the initial concentration of (**c**) 200 g/L to validate the inferred mode.

process uses high-ethanol tolerance yeast, *Saccharomyces diastaticus* LORRE 316, to produce ethanol. The experimental materials and methods were illustrated in Wang *et al.* (2001). The yeast can utilize glucose to produce ethanol and glycerol so the super-structure of S-system is described as

$$\dot{X}_i = \alpha_i \prod_{j=1}^{4} X_j^{g_{ij}} - \beta_i \prod_{j=1}^{4} X_j^{h_{ij}}, i = 1, \ldots, 4 \tag{14}$$

where X_1, X_2, X_3 and X_4 denotes, respectively, the concentrations of the biomass, glucose, ethanol and glycerol, respectively. The interactive inference algorithm is then employed to determine a suitable S-system structure and its parameter values.

The time-course data with 1 h sampling were obtained from two batch fermentations using the initial glucose concentrations of 100 and 150 g/l. Each fermentation was carried out in two experiments giving data points shown in Figure 2a and b. These repeated time-course data were noisy so the curve-fitting toolbox in MATLAB software was first employed to smooth the observed data for evaluating concentration error criterion and slope error criterion. By following similar procedures as discussed in the previous dry-lab case study, the inference problem is decoupled into four subsystems. The search ranges for each parameter used in the interactive inference algorithm were α_i and $\beta_i \in [0, 5]$, and g_{ij} and $h_{ij} \in [-3, 3]$.

Table 3 shows the expected values for each iteration. The Pareto optimal error criteria for the concentration and slope are also listed in Table 3. Two iterations are enough to achieve the minimum connective network for the first, third and fourth subsystem, three iterations for the second subsystem. The rate constant of degradation for glycerol was almost zero so the degradation term was neglected. The inferred model and its parameters were then provided as the initial starting point for a gradient-based method, a subroutine BCONF in IMSL Math/Library, to yield the more accurate solution.

Table 3. The expected values for each iteration obtained by each single objective parameter estimation and the Pareto optimal values obtained by multiobjective optimization problem

Iteration	Variable	Expected values		Pareto optimal values	
		J_1^E	J_2^E	J_1^*	J_2^*
1	1	2.36E−3	1.39E−2	1.03E−3	1.39E−2
	2	2.22E−4	6.23E−3	2.22E−4	3.82E−3
	3	4.55E−4	3.32E−3	3.60E−4	3.32E−3
	4	1.17E−3	6.52E−3	2.98E−4	6.42E−3
2	1	3.31E−3	4.14E−2	3.05E−3	4.14E−2
	2	1.98E−3	6.49E−3	5.70E−4	6.49E−3
	3	4.64E−4	1.01E−2	4.63E−4	4.71E−3
	4	4.22E−4	1.62E−2	3.70E−4	1.20E−2
3	1	–	–	–	–
	2	4.78E−3	9.20E−2	1.99E−3	3.34E−2
	3	–	–	–	–
	4	–	–	–	–

The optimal model is obtained as follows:

$$
\begin{aligned}
\dot{X}_1 &= 3.1125\ X_1^{0.8993} X_2^{-0.2771} - 0.4777\ X_1^{1.6345} X_2^{-0.2768} \\
\dot{X}_2 &= 3.3067 - 1.8574\ X_1^{0.5929} X_2^{0.2960} \\
\dot{X}_3 &= 0.5433\ X_1^{0.8782} X_2^{0.2282} - 0.0735\ X_3^{0.8352} \\
\dot{X}_4 &= 0.3568\ X_1^{0.6600} X_2^{0.3187}
\end{aligned}
\tag{15}
$$

The optimal computed profiles for initial glucose concentration of 100 g/l and 150 g/l are shown as the solid curves in Figure 2a and b. To validate the inferred model, an additional experiment with the initial glucose of 200 g/l, which is 33% greater than the training data, are employed to predict the dynamic behavior. Figure 2c shows the computed results and experimental data. As seen from this figure, the inferred S-system model is suitable for describing the dynamic behaviors of the batch fermentation process.

4 CONCLUSION

To infer a suitable interaction network for biological systems from time-course data poses many challenges. Numerical integration for differential equations and finding global parameter values are two major problems. Modified collocation and slope approximation can be employed to alleviate the computation burden. Hybrid differential evolution is utilized to obtain a global estimate. However, when inferring a minimum interaction network, sums of magnitude of kinetic orders serve as the penalty term to evaluate the fitness for the inference problem. How to tune a penalty weight to yield a realizable model structure is also a challenging problem. No guideline has been published for tuning a suitable penalty weight to infer a suitable model structure of biochemical networks. The multiobjective optimization approach could avoid assigning a penalty weight for sums of magnitude of kinetic orders. We have proved that the approach could guarantee the minimum solution for the constrained problem to achieve the minimum interaction network for the inference problem.

A multiobjective optimization can consider many goals at the same time. This study is to investigate the concentration error criterion, the slope error criterion and interaction measure. The concentration error criterion is employed to measure the goodness-of-fit of the model with respect to a given experimental time-course dataset. The slope error criterion is used to judge the accuracy of the dynamic function, i.e. the net rate equation in (1). Each kinetic order, g_{ij} or h_{ij}, is applied to quantify the effect of X_j variable on the production or degradation of X_i. A smaller parameter value means less interaction between state variables, X_j and X_i. The interaction measure sums up magnitude of kinetic orders which serves as an index to prune a skeletal structure of S-systems. Such a pruning strategy may be not suited for inferring whether genetic interactions are fragile or robust. Since the fragile interaction has higher sensitivity, a slight change in parameter value of this interaction, g_{ij} or h_{ij}, should cause a big difference of dynamic behaviors of gene expression. Under such circumstances, the interaction measure may be unable to infer such high-sensitive systems. An additional goal, such as dynamic sensitivities of state variables with respect to g_{ij} and h_{ij}, should be considered in the multiobjective parameter estimation problem toward inferring a suitable network structure for a high-sensitive gene-regulatory system.

ACKNOWLEDGEMENT

The financial support from the National Science Council, Taiwan, ROC (Grant NSC96-2627-B-194-001), is highly appreciated.

Conflict of Interest: none declared.

REFERENCES

Almeida,J.S. and Voit,E.O. (2003) Neural-network-based parameter estimation in S-system models of biological networks. *Genome Inform.*, **14**, 114–123.

Bazaraa,M.S. and Shetty,C.M. (1979) *Nonlinear Programming. Theory and Algorithms*. John Wiley and Sons, pp. 336–340.

Chang,W.C. *et al*. (2005) Quantitative inference of dynamic regulatory pathways via microarray data. *BMC Bioinformatics*, **6**, 1–19.

Chiou,J.P. and Wang,F.S. (1999) Hybrid method of evolution algorithms for static and dynamic optimization problems with application to a fedbatch fermentation process. *Comput. Chem. Eng.*, **23**, 1277–1291.

Chou,I.C. *et al*. (2006) Parameter estimation in biochemical systems models with alternating regression. *Theor. Biol. Med. Model.*, **3**, 1–25.

Edwards,K. *et al*. (1998) Kinetic model reduction using genetic algorithms. *Comput. Chem. Eng.*, **22**, 239–246.

Gonzalez,O.R. *et al*. (2007) Parameter estimation using simulated annealing for S-system models of biochemical networks. *Bioinformatics*, **23**, 480–486.

Handl,J. *et al*. (2007) Multiobjective optimization in bioinformatics and computational biology. *IEEE/ACM Trans. Comput. Biol. Bioinform.*, **4**, 279–292.

Hlavacek,W.S. and Savageau,M.A. (1996) Rules for coupled expression of regulator and effector genes in inducible circuits. *J. Mol. Biol.*, **255**, 121–139.

Ho,S.Y. *et al*. (2005) Evolutionary divide-and-conquer approach to inferring S-system models of genetic networks. *Proceedings of the 2005 IEEE Congress on Evolutionary Computation.*, **1**, 691–698.

Huang,W.H. *et al*. (2006) Reverse engineering for embryonic gene regulatory network in zebrafish via evolutionary optimization with data collocation. In *Proceeding of 7th International conference on systems biology*, pp. 9–13. http://www.icsb-2006.org/program/postersF.htm.

Ingalls,B.P. (2004) Autonomously oscillating biochemical systems: parametric sensi-tivities of extrema and period. *IEE Syst. Biol.*, **1**, 62–70.

Kikuchi,S. *et al*. (2003) Dynamic modeling of genetic algorithm and S-system. *Bioinformatics*, **19**, 643–650.

Kimura,S. *et al*. (2004) Inference of S-system models of genetic networks from noisy time-series data. *Chem-BioInformatics J*., **4**, 1–14.

Kimura,S. *et al*. (2005) Inference of S-system models of genetic networks using a cooperative coevolutionary algorithm. *Bioinformatics*, **21**, 1154–1163.

Maki,Y. *et al*. (2001) Development of a system for the inference of large scale genetic networks. *Pac. Symp. Biocomput*., **6**, 446–458.

Maki,Y. *et al*. (2002) Inference of genetic network using the expression profile time course data of mouse P19 cells. *Chem-BioInformatics J*., **13**, 382–383.

Mendes,P. and Kell,D.B. (1998) Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation, *Bioinformatics*, **14**, 869–883.

Moles,C.G. *et al*. (2003) Parameter estimation in biochemical pathways: a comparison of global optimization methods. *Genome Res*., **13**, 2467–2474.

Noman,N. and Iba,H. (2005) Reverse engineering genetic networks using evolutionary computation. *Genome Inform*., **16**, 205–214.

Polisetty,P.K. *et al*. (2006) Identification of metabolic system parameters using global optimization methods,. *Theor. Biol. Med. Model*., **3**, 1–15.

Sakawa,M. (1993) *Fuzzy Sets and Interactive Multiobjective Optimization*. Plenum Press, New York, pp. 91–116.

Tsai,K.Y. and Wang,F.S. (2005) Evolutionary optimization with data collocation for reverse engineering of biological networks,. *Bioinformatics*, **21**, 1180–1188.

Voit,E.O. (2000) *Computational analysis of biochemical systems*. Cambridge University Press, pp. 37–75.

Voit,E.O. and Almeida,J.S. (2003) Dynamic profiling and canonical modeling: powerful partners in metabolic pathway identification. In Goodacre,R. and Harrigan,G.G. (eds.) *Metabolite Profiling: Its Role in Biomarker Discovery and Gene Function Analysis*. Kluwer Academic Publishing, Dordrecht, pp. 125–139.

Voit,E.O. and Almeida,J.S. (2004) Decoupling dynamic systems for pathway identification from metabolic profiles. *Bioinformatics*, **20**, 1670–1681.

Wang,F.S. (2000) A modified collocation method for solving differential-algebraic equations. *Appl. Math. Comput*., **116**, 257–278.

Wang,F.S. and Sheu,J.W. (2000) Multiobjective parameter estimation problems of fermentation processes using a high ethanol tolerance yeast. *Chem. Eng. Sci*., **55**, 3685–3695.

Wang,F.S. *et al*. (2001) Hybrid differential evolution for problems of kinetic parameter estimation and dynamic optimization of an ethanol fermentation process. *Chem. Eng. Sci*., **40**, 2876–2885.