British Journal of Mathematics & Computer Science
4(7): 952-968, 2014

SCIENCEDOMAIN *international*
www.sciencedomain.org

Constrained Stochastic Space Search Method for Parameter Estimation in Biological Networks

Jimmy Omony[1*]

[1]*University of Groningen, Molecular Genetics Department, P.O. Box 11103, 9700 CC Groningen, The Netherlands.*

Original Research Article

Received: 11 November 2013
Accepted: 18 January 2014
Published: 12 February 2014

ABSTRACT

Parameter estimation is an important part of computational systems biology – especially in studies on biological networks. Numerous stochastic search methods have been applied in parameter estimation in biological networks. In this paper, a constrained stochastic space search (CSSS) method for parameter estimation is proposed and evaluated for estimating the parameters of a genetic network described by differential equations. Both linear and nonlinear model formalisms were used for the data evaluation. The performance of the CSSS method was compared to the Integrated Controlled Random Search for Dynamic Systems (ICRS/DS) stochastic optimization algorithm. Compared to the ICRS/DS, the CSSS algorithm is faster with at least a 7-fold shorter convergence time. Independent replicates were run and identification performed. For the same initialization conditions prior to optimization, the CSSS had on average smaller relative mean errors than the ICRS/DS.

Keywords: Biological network; differential equations; optimization; parameter estimation.

1 Introduction

There is growing interest in parameter estimation and understanding the dynamics of biological networks. Numerous methods have been used so far to gain insight into various aspects of biological networks, e.g. (probabilistic) Bayesian networks, regression methods, Boolean methods, and differential equations. The ordinary differential equation (ODE) formalism is a popular method to model biological networks [1,2,3,4,5]. This formalism is used to model the regulatory status of genes, hence, rendering ODEs suitable for describing gene regulation activity [6]. Generally, many parameters are required in ODEs to describe the dynamics and regulatory roles of various components of a biological network. In principle, the number of parameters increases with the level of details required to describe the regulatory mechanisms involved in a particular biological network. There is currently numerous global parameter estimation methods proposed in computational systems biology [7,8,9,10,11]. There are many space search methods that address parameter estimation problems in biological networks and the number keeps rising [12].

**Corresponding author: jimmy.omony@gmail.com;*

Although much progress has been made in developing powerful and accurate parameter estimation methods, such as that in [13,14], there is still need for improvement in the computational robustness and speed of existing methods. Many genetic network inverse modeling problems require robust parameter estimation routines to ensure that the parameters are obtained with high precision. The classical way to do this is by least squares estimation or maximum likelihood estimation is used [15]. Similarly, stochastic search based methods can be used – in which case a specified set of parameter constraints and heuristics for initializations on some parameters for optimization are required [7,11,14]. The optimization is based on minimization of goal function to fit a given model to a dataset.

The challenge of accurately describing the dynamics and change in concentration of molecular species, e.g. messenger Ribose Nucleic Acid (mRNA) in the cell. In Computational Systems Biology, Bioinformatics and Biotechnology, parameter estimation remains a big challenge – especially in the study of biological networks. To help address this challenge, various methods have been proposed in literature; e.g., Englezos and Kalogerakis [16] provided insight into optimization methods used for parameter estimation. Currently, there are numerous approaches to estimating parameters from data. Fundamentally, the mathematical principles and problem formulation vary. Additionally, there exist differences between the most methods, for instance: (i) some methods are computational more expensive, (ii) variation in levels of accuracy and precision in prediction of state variables and/or models parameters, and (iii) the scalability to larger dimensional networks (or compartmental systems). Even small functional modules in networks require a large number of parameters to describe the underlying kinetics and regulatory mechanisms between the molecular units (genes and/or proteins). This large number of parameters is required for accurate network reconstruction irrespective of the model formalism.

Here, the constrained stochastic space search (CSSS) method is proposed and its performance is compared to the Integrated Controlled Random Search for Dynamic Systems (ICRS/DS) which can be found in [17]. The ICRS/DS is a modification of the ICRS algorithm [18]. Much as the ICRS/DS is robust, it comes with weaknesses, e.g.: (i) it requires making heuristic guesses for the direction search tuning parameters, and (ii) it takes a long time to converge to an optimum solution, if a solution exists in the search space and the optimization tolerance level is sufficiently stringent. The CSSS method addresses these issues by using a technique of variance scaling on the parameters during optimization. The performances of these methods were validated on two test case networks through *in silico* experiments and optimization.

Other stochastic search algorithms such as the Genetic Algorithms and Particle (Swarm) Optimization have also been applied to solve parameter estimation in biochemical systems (see Bosezzi et al. [19] and Yang et al. [20]. Unlike the CSSS method, many of the stochastic algorithms in literature are quite complex and require curation of initial parameter values prior to optimization. Extensive insight into the performance of commonly used algorithms for parameter estimation in Systems Biology is found in the work of Ashyraliyev et al. [21]. The CSSS algorithm aims to get the best parameter estimates by optimizing the sum of squared errors (SSE) goal function. Just like the ICRS/DS, the CSSS algorithm requires pre-specification of constraints on the parameter for initialization. The parameter precision and accuracy levels were found to be influenced by network topology and the existence of a feedback and/or feed-forward loop. For instance, the existence of edges (which represents the regulatory strength between a target gene and its transcription factor) in a network being significantly stronger than others may cause a drastic shift in the system condition number - ultimately affecting how accurately parameter are estimated.

2 Methods

2.1 Network Representation

Consider a genetic network system that is represented by the differential equation

$$\dot{X} = f(X, u; \Omega); \qquad Y = g(X, \varepsilon) \tag{1}$$

where $X = [x_1, \dots, x_n]^T \in \Re^{N \times n}$ represents the state of the system which details the gene activities. Here $y_i = [y_{i1}, \dots, y_{iN}]^T \in \Re^{N \times 1}$ is the measured activity of gene i in time, N - number of time points; the matrix of measured data is $Y = [y_1, \dots, y_n]^T \in \Re^{N \times n}$, Ω - parameter space, ε - Gaussian measurement error term and u - external perturbation vector. The transition function $f(\cdot)$ captures the network dynamics and $g(\cdot)$ is the output or mRNA measurement function which can be obtained by transcriptomics experiments.

2.1.1 Linear representation

Let us consider the linear system representation:

$$\dot{\mathrm{X}} = AX + bu \tag{2}$$

Here, A is the network connectivity matrix with typical attributes of real biological networks like activation, repression, auto-regulation (which is also often referred to as self-regulation) and feedback. The network is triggered using an external perturbation function:

$$u(t) = u(0)[1/(\beta + e^{Kt})] \tag{3}$$

β and K are nonnegative constants and $u(0)$ is the initial concentration of the triggering compound. The nature of $u(t)$ is based on many physical and biological systems in which the concentrations of an inducing compound is consumed [22,23].

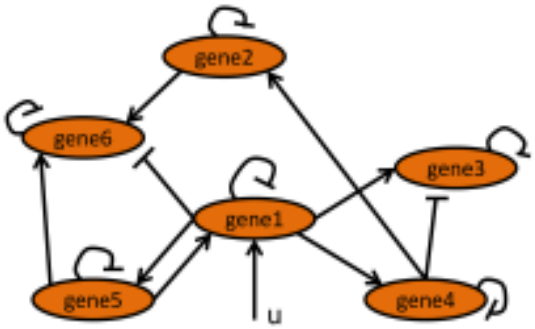

Fig. 1. A synthetic genetic network. The sharp headed (→) and flat headed (⟞) arrows respectively represent activation and repression of transcription. The nodes represent genes and u is the perturbation (input) signal, which can be nutrient uptake, glucose uptake etc depending the biological network under consideration

Consider a system of ODEs represented with the connectivity matrix constructed from the network in Fig. 1. The fully connected matrix A $\in \Re^{n\times n}$ that corresponds to this network has a total of $n^2 = 36$ parameters. However, not all the parameters need to be considered given that sparseness is a common attribute of biological networks [24,25]. Let $\mathrm{A}(\Omega)$ denote the parameterized connectivity matrix which, in this case, is a sparse network with 16 parameters – including the input vector coefficient b_1. The diagonal entries θ_{ii} in (4) are the auto–regulation effects. The parameterized connectivity matrix is

$$\mathrm{A}(\Omega) = \begin{pmatrix} -\theta_{11} & 0 & 0 & 0 & \theta_{15} & 0 \\ 0 & -\theta_{22} & 0 & \theta_{24} & 0 & 0 \\ \theta_{31} & 0 & -\theta_{33} & -\theta_{34} & 0 & 0 \\ \theta_{41} & 0 & 0 & -\theta_{44} & 0 & 0 \\ \theta_{51} & 0 & 0 & 0 & -\theta_{55} & 0 \\ -\theta_{61} & \theta_{62} & 0 & 0 & \theta_{65} & -\theta_{66} \end{pmatrix} \tag{4}$$

and the input vector

$$b = [b_1, 0,0,0,0,0]^T \tag{5}$$

The network is triggered by the signaling function in (3)as an external perturbation, therefore, enhancing the gene activities. For the ease of notations, let the parameter from (4) are represented by:

$$\omega = [\theta_{11}, \theta_{15}, \theta_{22}, \dots, \theta_{66}, b_1]^T \tag{6}$$

Several sets with different parameter values were assigned to the connectivity matrix to enable synthetic data generation. The initial parameters are depicted as:

$$\omega^{(0)} = \left[\theta_{11}^{(0)}, \theta_{15}^{(0)}, \theta_{22}^{(0)}, \dots, \theta_{66}^{(0)}, b_1^{(0)}\right]^T \tag{7}$$

The vector dimension $dim\left(\omega^{(0)}\right) = m$ depends on the number of unknown parameters.

2.1.2 Boundary constraints on parameters

The restrictions imposed on the lower (L) and upper (U) parameter bounds are:

$$\Omega_i = \begin{cases} \theta_{ij}^{\mathrm{L}} \le \theta_{ij} \le \theta_{ij}^{\mathrm{U}} \quad \text{for all} \quad i,j \\ \qquad b_1^{\mathrm{L}} \le b_1 \le b_1^{\mathrm{U}} \end{cases} \tag{8}$$

These constraints ensure that nonrealistic parameter values are avoided [26,27]. Once the model structure is known and the constraints on the parameters in (8) specified, then optimization of the goal function is initiated. The parameter values are updated during optimization until convergence at some optimal goal function value is obtained.

2.2 Nonlinear Network Representation

Consider the nonlinear model representation of a genetic network of three genes as in [28]. With a slight change of notation, the model is given by

$$\Gamma_{\text{nls}}(\Omega_i) = \begin{cases} \frac{dx_1}{dt} = k_{1s}\frac{1}{1+k_{13}x_3} - k_{1d}x_1 \\ \frac{dx_2}{dt} = k_{2s}\frac{k_{21}x_1}{1+k_{21}x_1} - k_{2d}x_2 \\ \frac{dx_3}{dt} = k_{3s}\frac{k_{31}x_1}{1+k_{31}x_1}\frac{k_{32}x_2}{1+k_{32}x_2} - k_{3d}x_3 \end{cases} \tag{9}$$

The term x_i represents the state for gene i. The equations in (9) describe a simple regulatory network in which gene 3 (x_3) represses transcription of gene 1 (x_1), gene 1 activates transcription of gene 2 (x_2), gene 1 (x_1) and gene 2 (x_2) collectively activate gene 3 (x_3). A popular form of function used in modeling biological networks is the Hill function [29,30]. The parameter vector for this nonlinear system (nls) representation is given by

$$\omega_{\text{nls}} = [k_{1s}, k_{13}, k_{1d}, k_{2s}, k_{21}, k_{2d}, k_{3s}, k_{31}, k_{32}, k_{3d}]^T \tag{10}$$

In (9), the mRNA degradation is modeled as a first-order reaction and is represented by the constants k_{1d}, k_{2d} and k_{3d}; The terms k_{1s}, k_{2s} and k_{3s} are the synthesis parameters and k_{ij} is the effective affinity constant for gene j activating i. The expression for the measured system outputs are as described in subsection 2.1. The constraints on the parameters in (9) were specified as:

$$\Omega_{i,\text{nls}} = k_{ij}^{\text{L}} \leq k_{ij} < k_{ij}^{\text{U}} \tag{11}$$

The dots in the parameter subscripts in (11) denote the corresponding indices in (10). This restriction and the initial guess $\omega_{\text{nls}}^{(0)}$ were used for the optimization.

2.3 Data Generation and Parameter Estimation

The synthetic data was generated prior to parameter estimation. A comparison of the "true" and the estimated parameters was done. In the simulation, measurement noise was added to the data. The expression for gene i at time t is $y_i(t) = x_i(t) + \varepsilon_i(t)$, where $\varepsilon_i = \alpha\|x_i\|_1$ is the relative data measurement error, α is the measurement noise. A value of $\alpha = 0.05$ was considered at sampling moments in the *in silico* data simulations. The notation $\|\cdot\|_1$ represents the *City Block Distance* in the L_1 space. The SSE goal function is expressed as

$$J(\widehat{\Omega}_i) = \arg min\left(\xi_i^{\text{T}}(t|\widehat{\Omega}_i)\xi_i(t|\widehat{\Omega}_i)\right) \text{ for } i = 1, \ldots, n \tag{12}$$

where $\xi_i(t|\widehat{\Omega}_i) = y_i^{\text{obs}}(t) - y_i^{\text{est}}(t|\widehat{\Omega}_i)$ is the residual vector, and $y_i^{\text{est}}(t|\widehat{\Omega}_i)$ is the estimated expression for gene i conditioned on the parameter space.

2.4 Performance Evaluation

A precision measure, relative errors were used on the individual parameters. The expression for computing the relative error on a parameter is given by

$$R_{\text{error},l} = |\omega_l^{\text{true}} - \widehat{\omega}_l|/|\omega_l^{\text{true}}| \tag{13}$$

where ω_l^{true} and $\widehat{\omega}_l$ represent the considerably "true" and estimated parameter values, respectively. Overall, two goodness-of-fit measures on parameter estimates were used, namely: the goal function and the relative error in parameter estimates. The mean and standard deviations of the relative errors in the vector $\widehat{\omega}$ are given by (14) and (15), respectively.

$$\hat{\mu}_{R_{\text{error}}} = \frac{1}{p}\sum_{l=1}^{p} R_{\text{error},l}(\widehat{\omega}) \tag{14}$$

$$\hat{\sigma}_{R_{\text{error}}} = \sqrt{\frac{1}{p-1}\sum_{l=1}^{p}\left(R_{\text{error},l}(\widehat{\omega}) - \mu_{R_{\text{error}}}\right)^2} \tag{15}$$

Here p is the number of independent identifications performed. Twenty different *in silico* experiments were performed (i.e. $p = 20$). Firstly, performing different experiments involved making a nonnegative change $\delta\omega_l$ in the parameter vector (i.e. set $\omega_l \coloneqq \omega_l \pm \delta\omega_l$) and then simulating the data – this was repeated p times. Secondly, independent identifications were performed using both the CSSS and ICRS/DS methods. In each *in silico* experiment, a random number of parameters was kept constant while the others were varied around at most 20% of their starting values – this was repeated p times. The above different cases in data simulation and identification help to specify reliability of the optimization methods with respect to $\widehat{\omega}$. Low values of $\bar{\mu}_{R_{\text{error}}}$ are associated with a good performance, and low values of the mean relative standard errors $\hat{\sigma}_{R_{\text{error}}}$ represent a high consistency in estimating ω. To assess the performance of the method over the parameter space, the average values, $\bar{\mu}_{R_{\text{error}}}$ were computed using

$$\hat{\mu}_{\text{Total}} = \frac{1}{m}\sum_{k=1}^{m} \hat{\mu}_{R_{\text{error},k}} \tag{16}$$

This statistic is an unbiased estimator since $\bar{\mu}_{\text{Total}} \approx 0$. Similarly, the total variations in the estimate $\hat{\sigma}_{\text{Total}}$ were calculated using the expression

$$\hat{\sigma}_{\text{Total}} = \sqrt{\frac{1}{m-1}\sum_{k=1}^{m}\left(\hat{\sigma}_{R_{\text{error},k}} - \text{E}\left(\hat{\sigma}_{R_{\text{error}}}\right)_k\right)^2} \tag{17}$$

The mean standard deviation was calculated from

$$\hat{\sigma}_{\text{mean}} = \frac{1}{m}\sum_{k=1}^{m} \hat{\sigma}_{R_{\text{error},k}} \tag{18}$$

where $\text{E}(\cdot)$ is the expectation operator. Equation (16) and (18) provide additional information on the accuracy and consistency of the computed parameter estimates. A small confidence bound in (18) indicates replicable experiments with respect to the identification method used. The closer (16) gets to zero, the better the identification.

2.5 CSSS Algorithm Pseudo-code

The specifications of the subsequent steps in the CSSS algorithm are given in Table 1. The implementation and systems identification was done in MATLAB version 7.4.0. Setting an infinitesimal tolerance value $\delta_\varepsilon = 1e-5$ (a relative deviation in goal functions) reduces the likelihood of the optimization getting stuck at a local minimum. This low tolerance value ensures that the error in the goal function is negligible, hence, increasing the accuracy in parameter estimation.

Table 1. Pseudo-code for the CSSS algorithm. Here ℓ, l, s are the iteration indices, r: number of successful iterations satisfying the specified criterion, $\Re$: a definite integral domain, $\|\cdot\|_1$: the city block distance (or L_1 norm), $\mathrm{E}(\cdot)$: is the expectation operator. $J^{(0)}$ – initial value of goal function, $[y_1^{\mathrm{est}},\cdots,y_n^{\mathrm{est}}]^{(\ell)}$ and $[y_1^{\mathrm{est}},\cdots,y_n^{\mathrm{est}}]^{(\mathrm{cg})}$ are the vectors of mRNA measurements (or model output for individual genes) during iteration and upon convergence, respectively. The superscripts "L" and "U" in ω^{L} and ω^{U} represent the lower and upper bounds, respectively. The other variables and parameters are as defined elsewhere in the manuscript

Require: $\left\{\omega^{(0)}, J^{(0)}, \Gamma\left(\Omega_i \mid \mathrm{Y}_{N\times n}\right)\right\}$
Ensure: $\Gamma\left(\Omega_i\right)$ holds
$\omega^{(0)}, J^{(0)}, \Gamma^{(0)}\left(\Omega_i\right);\ \ell \leftarrow 1$
$J_\ell\left(\Omega_i\right), \left[y_1^{\mathrm{est}},\cdots,y_n^{\mathrm{est}}\right]^{(\ell)} \leftarrow \int_{\Re_\Omega\subset\Re} f(\cdot)\, d\Re_\Omega$
$s \leftarrow 0$
while $\left\{\left|J_{\ell+1}\left(\Omega_i\right) - J_\ell\left(\Omega_i\right)\right| \Big/ \left|J_\ell\left(\Omega_i\right)\right| > \delta_\varepsilon\right\}$ do
  $\omega^{(\ell+1)} \leftarrow \omega^{(\ell)};\ s+1 \leftarrow s$
  $\sigma^{2(\ell)} \leftarrow \min\left[\mathrm{E}\left(\omega^{\mathrm{U}} - \omega^{(\ell)}\right)^2, \mathrm{E}\left(\omega^{(\ell)} - \omega^{\mathrm{L}}\right)^2\right]$
  $l \leftarrow 0$
  while $s < 1$ do
    $r \leftarrow 0$
    while $(r < 1) == 1$ do
      $\omega^{(\ell+1)} \leftarrow \omega^{(\ell)} + \sigma^{2(\ell)} / \ell + 1;\ \sigma^{2(\ell)} \square\ \aleph\left(0, \mathrm{I}_m\right)$
      if $\left\|\omega^{\mathrm{U}} - \omega^{(\ell)}\right\|_1 + \left\|\omega^{(\ell)} - \omega^{\mathrm{L}}\right\|_1 \le 2N$ then
        $r = 1;\ \omega^{(\ell+1)} \leftarrow \omega^{(\ell)}$
      end if
    end while
    $J_{\ell+1}\left(\Omega_i\right) \leftarrow \int_{\Re_\Omega\subset\Re} f(\cdot)\, d\Re_\Omega$
    if $J_{\ell+1}\left(\Omega_i\right) < J_\ell\left(\Omega_i\right)$ then
      $\ell + 1 \leftarrow \ell;\ \omega^{(\ell+1)} \leftarrow \omega^{(\ell)}$
    end if
    $\omega^{(cg)} \leftarrow \omega^{(\ell+1)}$
  end while
end while
return $\omega^{(cg)}, J^{(cg)}, \left[y_1^{\mathrm{est}},\cdots,y_n^{\mathrm{est}}\right]^{(cg)}, \ell + 1$

2.6 Pre-conditioner Verification

The condition number, just like correlation coefficients, can be used to explain any anomalies in parameter estimates. If the condition number is large (i.e. deviates significantly from the value 1), then the gene expression values have similar time trajectories. This is an attribute that may lead to biased parameter estimates. Let the measured gene expression (mRNA levels) data be $Y_{N\times n}$, then the condition number for this matrix is defined as $\kappa(Y) \coloneqq \|Y\|\|Y^{-1}\|$if $Y_{[N=n]}$ is a nonsingular matrix. If Y is a non-square matrix, then the analogous definition for the condition number is:

$$\kappa(Y^T Y) \coloneqq \kappa(Y^2) = \|(Y^{\mathrm{T}}Y)\|\|(Y^T Y)^{-1}\| \tag{19}$$

3Results and Discussions

3.1 Model Goodness-of-fit to Data

On performing the identification, the results for the linear, the non-linear both for the CSSS and ICRD/DS methods all showed good model fits to the data. See, for instance Fig. 2, which concerns results of the CSSS method for the non-linear case which involves three genes and a total of 10 parameters to be estimated. Individual data points such as those in Fig. 2 represent averaged data values from duplicates, triplicates or more measurements – depending on the costs. The condition number for the linear and nonlinear network cases was found to be in the order of magnitude $\kappa(\cdot) = 1.7e + 4$ and $\kappa_{\mathrm{nls}}(\cdot) = 1.2e + 4$, respectively. This order of magnitude in this number may compromise the accuracy of parameter estimates. It should be noted that by invoking a network by a single trigger most of the gene transcription profiles turn out to be correlated.

3.2 Identification from the Linear Model Formalism

Parameter estimation with the CSSS and ICRS/DS algorithms is done for a 20 sets of parameters for the linear model formalism. The SSE (or goal function value) using the ICRS/DS algorithm was found to be $J^{(\mathrm{cg})}_{\mathrm{ICRS/DS}} = 0.0191$, and $J^{(\mathrm{cg})}_{\mathrm{CSSS}} = 0.0170$ for the CSSS method. These values indicate the lowest achievable goal function resulting from the optimization. It basically shows a marginally better performance for CSSS compared to ICR/DS (i.e. $J^{(\mathrm{cg})}_{\mathrm{CSSS}} < J^{(\mathrm{cg})}_{\mathrm{ICRS/DS}}$). Table 2 provides an example of the parameter estimates and the true parameters of one of the sets. For both methods there is a deviation from the true values. The main reason is that a large number of parameters are estimated from a dataset with correlated response variables resulting in a high condition number. Identification with reduced system dimension (e.g. with gene 5 and 6 knocked out) shows that the CSSS method still yields results with lower variances on the relative errors. The overall relative error rates for the reduced system (with 9 parameters) were also lowered.

Fig. 3 gives the mean relative error (MRE) for each parameter as estimated by both methods. Overall, the MRE for the CSSS method ($\bar{\mu}_{\mathrm{Total}} = 0.4072$) is below that of the ICRS/DS ($\bar{\mu}_{\mathrm{Total}} = 0.6001$). Moreover, an important property of the CSSS method is that the MRE for an individual parameter is nearly constant while the MRE for the ICRS/DS shows a large variation (compare corresponding error bars). In other words the CSSS method is consistent in its results. The lower variation is a result of the variance scaling in step 11 of the algorithm. To estimate ωin the linear

model set up, parameter estimation a choice of $\omega_l^{(0)} \in \{-1, +1\}$ was made. Hereby, -1 was chosen for the auto-regulation effects, -1 for the repressing and $+1$ for the activating effects of the transcription factors to a particular target gene. With this choice of initial parameter, the CSSS algorithm needs a significantly lower amount of iterations to achieve the convergence than ICRS/DS (Fig. 3). If the initial parameters were not chosen according the above strategy, the convergence rate of both algorithms slightly worsens, but the performance of both algorithms remain invariantly comparable with CSSS marginally outperforming the ICRS/DS.

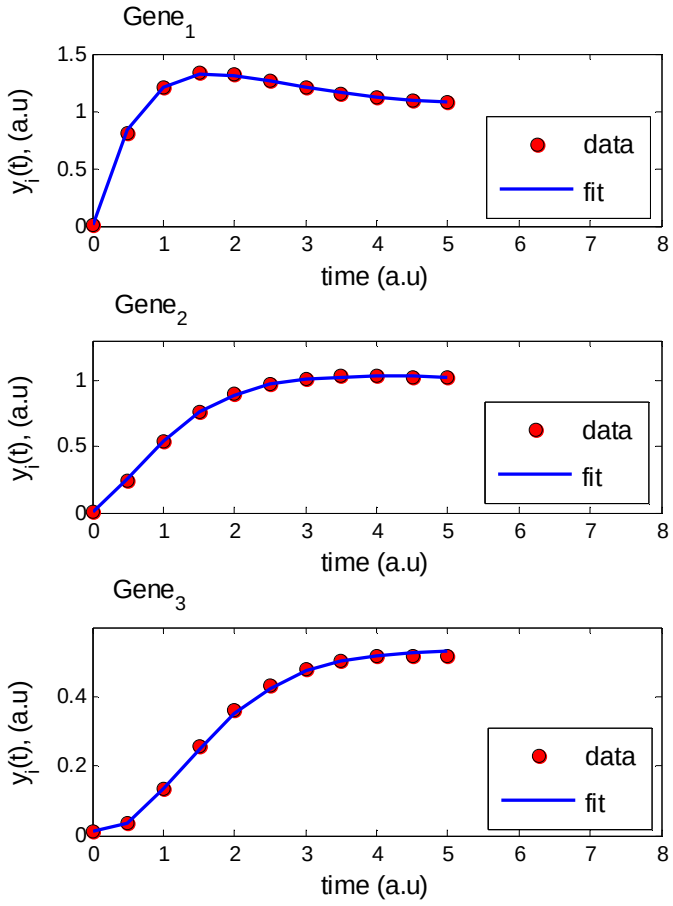

Fig. 2. Model fit to data using the CSSS method to the nonlinear example in section 2.2. The data points are indicated by points and the model fit by lines (measurement noise $\alpha = 0.05$, $N = 11$ data points and sampling step size of $\Delta t = 0.5$ units).

Table 2. Parameter estimates obtained by using the CSSS and ICRS/DS algorithm. Estimations were done using time course data from the linear network formalism (subsection 2.1) with measurement noise level of $\alpha = 0.05$ and $N = 20$ data points

ω_{nls}	"true" value	ICRS/DS	CSSS
		$\hat{\omega}_{\text{nls},i}^{(\text{cg})}$	$\hat{\omega}_{\text{nls},i}^{(\text{cg})}$
θ_{11}	2.0	0.6922	1.6589
θ_{15}	1.0	0.1398	0.6760
θ_{22}	2.1	2.0944	1.6817
θ_{24}	1.2	1.1965	0.9581
θ_{31}	2.5	2.7175	1.0710
θ_{33}	2.6	2.8466	1.1614
θ_{34}	1.6	1.7097	0.6087
θ_{41}	1.4	1.1515	0.6307
θ_{44}	2.9	2.3845	1.2947
θ_{51}	1.2	1.2235	0.9866
θ_{55}	1.1	1.1193	0.8993
θ_{61}	1.0	0.2683	0.3777
θ_{62}	2.6	2.4020	1.1511
θ_{65}	2.1	0.9770	1.6007
θ_{66}	2.1	1.5236	1.7599
b_1	1.0	0.5854	1.0011

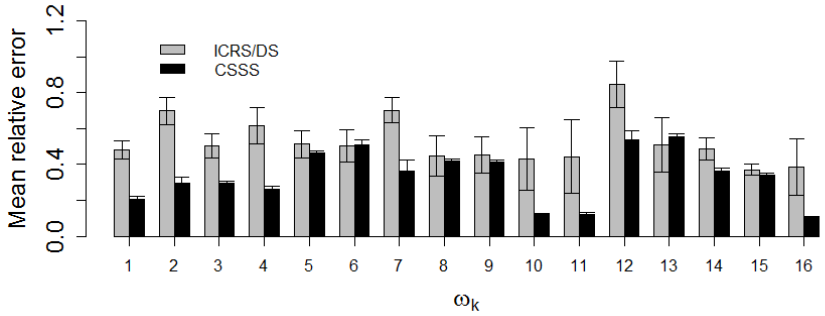

Fig. 3. Mean relative error (MRE) for the linear model formalism by the ICRS/DS and CSSS method, respectively. Analysis was done using dataset in which $\alpha = 0.05$. The error bars are the standard deviations ($\hat{\sigma}_{R_{\text{error},k}}$) to the MRE on parameter ω_k

3.3 Identification from the Nonlinear Model Formalism

The ODE nonlinear model formalism is known to be prone to having correlated parameters. This is problematic for network identification since it leads to reduced accuracy and precision in parameter estimates. Parameters were estimated for 20 different datasets. For the two methods being compared, the SSE values for the goal functions using the nonlinear identification problem were $J_{\text{ICRS/DS}}^{(\text{cg})} = 0.0017$ and $J_{\text{CSSS}}^{(\text{cg})} = 0.0014$. The parameters estimated for one of the datasets are shown in Table 3. Despite the correlation between parameters, for all these sets the parameters

estimated are much closer to the true values compared to the linear model formalism (see also error estimates in Fig.5). As expected, for the linear case with reduced number of parameters, better parameter estimates were observed.

Fig. 4 gives the MRE for the individual parameters. The average MRE for the CSSS is 0.1670 with a deviation of $1.4e-4$ and for the ICRS/DS method 0.2274 with a deviation of $1.5e-3$. Again the CSSS is more consistent in finding the parameters. In contrast to the linear case no rules could be applied for the parameters initial guess, therefore initial values were randomly chosen from the interval $[0,5]$.

Table 3. Parameter estimates obtained by using the CSSS and ICRS/DS algorithm. The estimations were performed using with data having measurement noise $\alpha = 0.05$, $N = 11$ data points and sampling step size of $\Delta t = 0.5$ units

$\omega_{\mathbf{nls}}$	"true" value	ICRS/DS	CSSS
		$\hat{\omega}_{\mathbf{nls},i}^{(\mathbf{cg})}$	$\hat{\omega}_{\mathbf{nls},i}^{(\mathbf{cg})}$
k_{1s}	2.0	2.0221	2.2348
k_{13}	2.0	1.4011	0.9060
k_{1d}	1.0	1.1755	1.5165
k_{2s}	2.0	1.4067	1.8152
k_{21}	1.0	1.8981	2.0154
k_{2d}	1.0	1.7361	1.8558
k_{3s}	2.0	3.1916	2.0154
k_{31}	1.0	0.2739	0.9287
k_{32}	1.0	4.0543	1.6019
k_{3d}	1.0	0.8022	0.7271

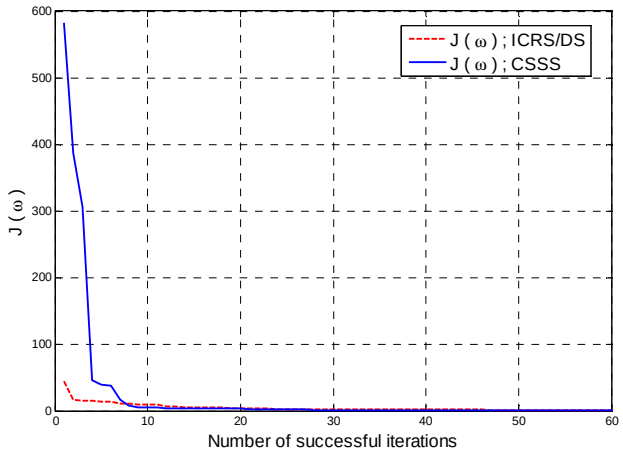

Fig. 4. Plot of goal function convergence values using the CSSS and ICRS/DS algorithms. The figure axis narrowed for easy visibility – the final successful iteration number at which the goal function converges is about 150 and 900 for the ICRS/DS and CSSS, respectively

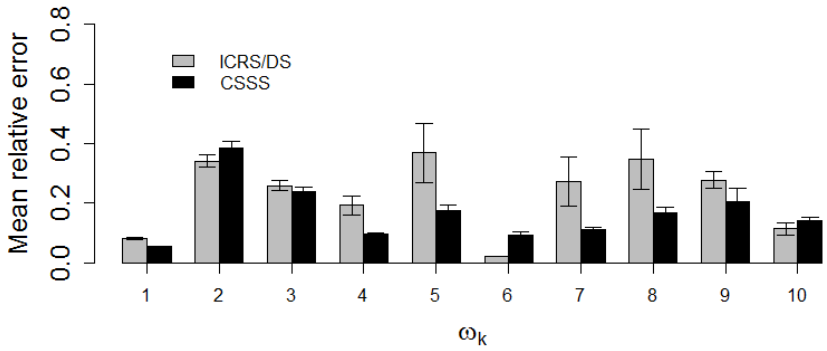

Fig. 5. Mean relative error (MRE) for the non-linear model formalism by the ICRS/DS and CSSS method, respectively. Analysis using data with a $\alpha = 0.05$noise level. The error bars are the standard deviations ($\hat{\sigma}_{R_{\text{error},k}}$) to the MRE on parameter ω_k

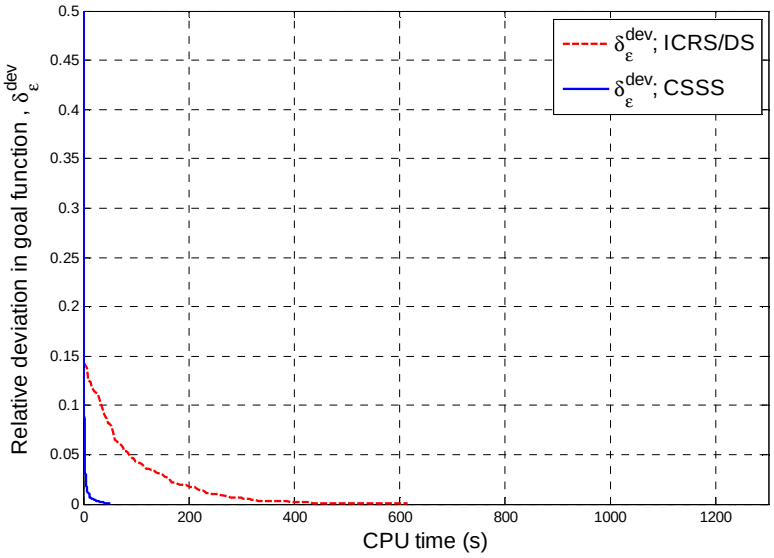

Fig. 6. Comparison of relative deviations in goal functions $\delta_{\varepsilon}^{\text{dev}}$ to convergence time for the CSSS algorithm versus with the ICRS/DS algorithm for the linear model formalism

3.4 Comparison of Algorithm Performance: ICRS/DS Versus CSSS

To assess the performance of the convergence times for the two identification methods under consideration, the optimization times to convergence were computed. For the same model and dataset, the CSSS algorithm converges faster than the ICRS/DS algorithm (Figs. 7 and 8). This holds true even for a highly connected network (with $m = 16$ parameters) used in the linear network as opposed to the nonlinear network with only three genes (with $m = 10$). During the optimization process, an iteration run was either rejected as unsuccessful or accepted as successful

by checking if $\omega^{(\ell)}$ and $J^{(\ell)}$ satisfies the conditions in steps 12 and 17, respectively (Table 1). A distinguishing feature of stochastic optimization methods is that much computational time is wasted on unsuccessful iterations. Compared to the ICRS/DS; steps 12 and 17 in the CSSS method ensure that the number of unsuccessful iterations is reduced.

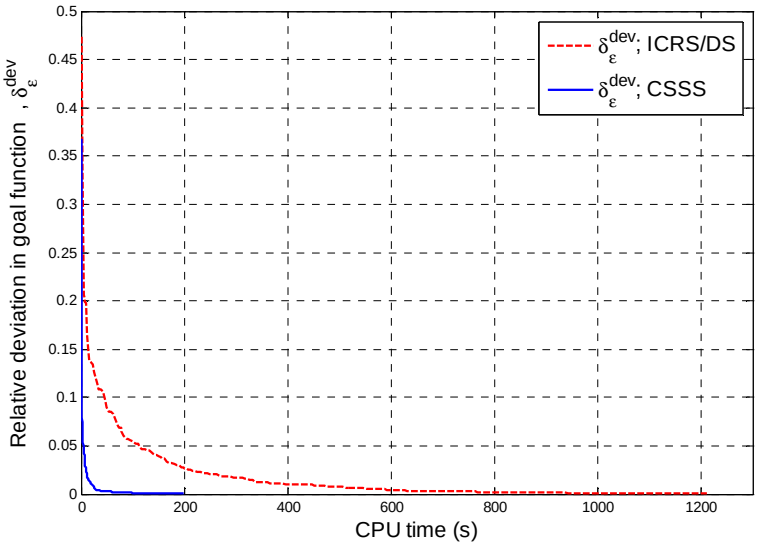

Fig. 7. Comparison of relative deviations in goal function δ_ε^{dev} to convergence time for the CSSS algorithm versus with the ICRS/DS algorithm for the non-linear model formalism

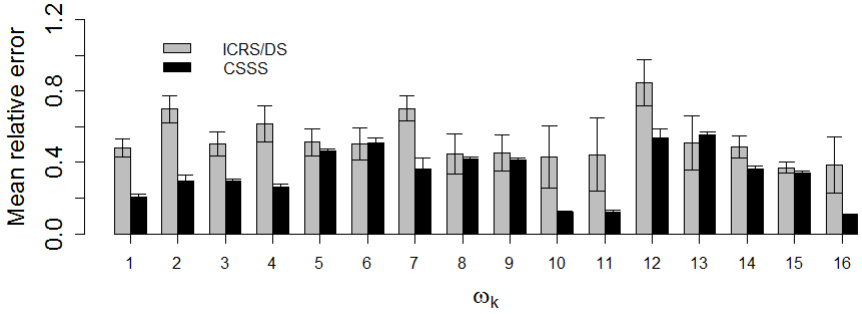

Fig. 8. Mean relative error (MRE) for the linear model formalism by respectively the ICRS/DS and CSSS method, analysis with the noise-free dataset in which $\alpha = 0.0$. The error bars are the standard deviations ($\widehat{\sigma}_{R_{error,k}}$) to the MRE on parameter ω_k

An assessment on the parameter accuracy was performed with five times the number of time points for all the cases considered. A total of 100 data points were used for both the linear and nonlinear network representation problems. The results do not deviate much from those observed earlier-on in which the CSSS method was shown to be superior to the ICRS/DS method. Relative error analysis was performed with noise-free data and the results indicated that the standard

deviations on the parameters are again lower for identification with the CSSS as compared to the ICRS/DS method. Therefore, irrespective of the measurement noise level in the dataset, the CSSS method is consistent and reliable in the parameter estimates – as seen from the low mean relative error and standard deviation values. This analysis indicates that the errors in the parameters are not caused by the noise level (Fig. 8) but are associated with the high condition number.

Using the CSSS method, the averaged relative errors on the estimated parameters were less than those obtained using the ICRS/DS method (Figs. 3 and 4). The CSSS method achieves the best accuracy results and is faster in convergence. The advantage of using the CSSS method over the ICRS/DS method is that the CSSS method does not require heuristic tuning of the parameters. In step 6: Table 1, the variance of the parameters is estimated and the minimum deviation from the two terms is used for updating the next parameter estimate. Scaling the variance term σ^2 by the iteration index $\ell + 1$ (step 11: Table 1) ensures that the variance decreases during optimization, leading to faster convergence (Figs. 6 and 7). This increased speed of convergence is associated with less function rejection during the space search in the optimization process. In the presence of correlated parameters in a model, the likelihood of obtaining good parameter estimates is reduced because a change in one parameter value proportionally causes a shift in the other parameter value – this is a challenge in systems identification. In essence, applying a method to a parameter estimation problem for larger network dimensions may yield a model with good fits to the data, but with fading degrees of parameter accuracy. Quicker convergence and better parameter estimates can be achieved by good parameter initialization since this increases the likelihood of obtaining optimal solutions [31]. This statement seems relevant for the ICRS/DS method, but this work shows that with the fast converging CSSS algorithm the initial choice is not that important.

4. Conclusions

The results presented in this paper are based on two stochastic optimization methods addressing the parameter identification problem from time course datasets. The findings indicate that the performance of the CSSS is better than that of the ICRS/DS. The CSSS method shows lower variation on relative error on the parameter estimates. The mean relative errors on the parameters are also lower for the results obtained using the CSSS method as opposed to those obtained from the ICRS/DS method. The CSSS method can be used for parameter estimation in small sized networks. The CSSS method is computationally efficient and out-performs the ICRS/DS in the convergence speed. The performance of the proposed optimization method for both the linear and nonlinear model representation was better than that of the ICRS/DS. Achieving accurate parameter identification remains one that is crucial to address, especially with growing interests from understanding larger biological networks, thereby, up-scaling the challenge. However, by decomposing networks into sub-units, it is easier to estimate the parameters with much better precision.

Competing Interests

The author declares no competing interests.

References

[1] Goodwin BC. Temporal Organization In Cells; A Dynamic Theory Of Cellular Control Processes. Academic Press, New York; 1963.

[2] Hartemink AJ, Gifford DK, Jaakola TS, Young RA. Combining location and expression data for principled discovery of genetic regulatory network models. Proc. Pac. Symp. Biocomput. 2002;7:437–449.

[3] Zak D, Pearson RK, Vadigepalli R, Gonye G, Schwaber JS, Doyle III FJ. Continuous-time identification of gene expression models. Omics. 2004;7:373–386.

[4] Savageau MA. Biochemical systems analysis: a study of function and design in molecular biology. Addison Wesley Publishing Company; 1976.

[5] Kikuchi S, Tominaga D, Arita M, Takahashi K, Tomita M. Dynamic modeling of genetic networks using genetic algorithm and S-system. Bioinformatics. 2003;19:643-650.

[6] Smolen P, Baxter DA, Byrne JH. Modeling transcriptional control in gene networks: methods, recent results, and future directions. Bull. Math. Biol. 2000;62:247-292.

[7] Zwolak J, Tyson J, Watson L. Globally optimized parameters for a model of mitotic control in frog egg extracts. IEE Proceedings Systems Biology. 2005;152:81-92.

[8] Chaitankar VP, Perkins EJ, Gong P, Deng Y, Zhang C. A novel gene network inference algorithm using predictive minimum description length approach. BMC Syst. Biol. 2010;4:S7.

[9] Tsai KY, Wang FS. Evolutionary optimization with data collocation for reverse engineering of biological networks. Bioinformatics. 2005;21:1180-1188; doi:10.1093/bioinformatics/bti099

[10] Ashyraliyev M, Jaeger J, Blom JG. Parameter estimation and determinability analysis applied to *Drosophila* gap gene circuits. BMC Syst. Biol. 2008;2:83.

[11] Polisetty PK, Voit EO, Gatzke EP. Identification of metabolic system parameters using global optimization methods. Theoretical Biology and Medical Modelling. 2006;3:4.

[12] Torres NV, Voit EO. Pathway analysis and optimization in metabolic engineering. New York, Cambridge University Press. 2002.

[13] Espocito WR, Floudas CA. Global optimization for the parameter estimation of differential-algebraic systems. Ind. Eng. Chem. Res. 2000;39:1291-1310.

[14] Papamichial I, Adjiman CS. A rigorous global optimization algorithm for problems with ordinary differential equations. J. Global Optim. 2002;24:1-33.

[15] Li Z, Osborne MR, Prvan T. Parameter estimation of ordinary differential equations. IMA J. Numer. Anal. 2005;25:264-285.

[16] P E, N K. Applied Parameter Estimation for Chemical Engineers. 1st Edition ed. CRC Press: USA; 2000.

[17] Carrasco EF, Banga JR. Dynamic optimization of batch reactors using adaptive stochastic algorithms. Ind. Eng. Chem. Res. 1997;36:2252-2261.

[18] Banga JR, Casares JJ. Integrated Controlled Random Search: application to a wastewater treatment plant model. Inst. Chem. Eng. Symp. Ser. 1987;100:183.

[19] Besozzi D, Cazzaniga P, Mauri G, Pescini D, Vanneschi L. A Comparison of Genetic Algorithms and Particle Swarm Optimization for Parameter Estimation in Stochastic Biochemical Systems. Lecture Notes in Computer Science. 2009;5483:116-127.

[20] Yang F, Zhang C, Sun T. Comparison of Particle Swarm Optimization and Genetic Algorithm for HMM training. ICPR 2008 19th International Conference on Pattern Recognition IEEE Xplore Digital Library. 2008.

[21] Ashyraliyev M, Fomekong-Nanfack Y, Kaandorp JA, Blom JG. Systems biology: parameter estimation for biochemical models. FEBS J. 2009;276:886-902; doi:10.1111/j.1742-4658.2008.06844.x; 10.1111/j.1742-4658.2008.06844.x

[22] Eungdamrong NJ, Iyengar R. Modeling cell signaling networks. Biol. Cell. 2004;96:355-62.

[23] Omony J, de Graaff LH, van Straten G, van Boxtel AJB. Modeling and analysis of the dynamic behavior of the XlnR regulon in *Aspergillus niger*. BMC Syst. Biol. 2011;5:S14.

[24] Hoguland M, Frigyesi A, Mitelman K. A gene fusion network in human neoplasia. Oncogene. 2006;25:2674-8.

[25] Nacher JC, Ochiai T. Power-law distribution of gene expression fluctuations. Physics letter A. 2008;372:6202-6206 .

[26] Tucker W, V M. Parameter reconstruction for biochemical networks using interval analysis. Reliable Comput. 2006;12:389-402.

[27] Tucker W, Kutalik Z, V M. Estimating parameters for generalized mass action models using constraint propagation. Math. Biosci. 2007;208:607-620.

[28] Karleback G, Samir R. Modeling and analysis of gene regulatory networks. Molecular cell biology. 2008;9:770-780.

[29] Hill AV. The possible effect of the aggregation of the molecules of haemoglobin on its dissociation curves. J. Physiol. 1910;40(Suppl.):4-7.

[30] Polynikis A, Hogan SJ, di Bernardo M. Comparing different ODE modeling approaches of gene regulatory networks. J. Theor. Biol. 2009;261:511-530.

[31] Averick BM, Carter RG, Moré JJ. The MINPACK-2 test problem collection; 1991.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
www.sciencedomain.org/review-history.php?iid=410&id=6&aid=3611