Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Parameter Set Selection for Signal Transduction Pathway Models including Uncertainties

Wei Dai, Loveleena Bansal and Juergen Hahn*

*Department of Biomedical Engineering, Department of Chemical & Biological Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA (Tel: 518-276-2138; e-mail: hahnj@rpi.edu).

Abstract: It is common that only a subset of the parameters of dynamic models can be accurately estimated. One approach for identifying a subset of parameters for estimation is to perform clustering of the parameters into groups based upon their sensitivity vectors. However, this approach has the drawback that uncertainty cannot be directly incorporated as the sensitivity vectors are based upon the nominal values of the parameters. One technique to address this deficiency is to define sensitivity cones, where a sensitivity cone includes all possible sensitivity vectors of one parameter for different values resulting from the uncertainty. Parameter clustering can then be performed based upon the sensitivity cones, instead of the sensitivity vectors. This paper applies this new approach to a signal transduction pathway model with a large number of uncertain parameters.

1. INTRODUCTION

Mathematical models composed of ordinary differential equations (ODEs) are widely used to describe the dynamic behaviors of biological and biomedical systems. The prediction accuracy of these models not only depends on the structure of the model, but also relies on adjustable parameters of the model, many of which are either directly taken from the literature or estimated using experimental data. A number of studies have investigated various aspects of parameter estimation of dynamic systems (Anh *et al.*, 2006; Kravaris *et al.*, 2013; Poyton *et al.*, 2006). However, before any estimation is performed, it is important to determine if all of these parameters are numerically identifiable and, if not, then what subset of parameters can be accurately estimated (Chu *et al.*, 2007).

Sensitivity analysis is a powerful tool to study how parameter variations can qualitatively or quantitatively influence the model behavior. As such, a variety of methods for parameter set selection have been studied and are based on local or global sensitivity analysis. The main drawback of local sensitivity analysis is that the sensitivity vectors are dependent on the "true" parameter values that are not precisely known prior to estimation. This may result in identification of parameter subsets that are suboptimal, which can have a significant impact on the model's prediction accuracy if the parameter uncertainty is large. While global sensitivity analysis can incorporate the parameter uncertainty, the results from global sensitivity analysis are non-trivial to interpret as they do not rely on the concept of sensitivity vectors.

This paper uses a new approach to address challenges for parameter set selection under uncertainty by combining a hierarchical clustering method (Chu *et al.*, 2008) and a dynamic optimization technique so as to quantify the effect of the uncertainty in the parameter space on the sensitivity vectors. A sensitivity cone is computed for each parameter, where all sensitivity vectors for different possible parameter values are contained within the cone. The parameters can then be clustered based upon their sensitivity cones, similar to what was done for sensitivity vectors (Dai *et al.*, 2013), after a suitable measure has been defined. This approach is applied to a signal transduction pathway model as these models tend to contain a large number of parameters, many of which contain significant uncertainties and which cannot be estimated from available measurements.

The paper is structured as follows: preliminaries are presented in Section 2. Section 3 discusses in detail the problem formulation and the solution approach. A scheme for parameter selection is then introduced based upon hierarchical clustering of sensitivity cones. A case study involving a large-scale signal transduction pathway model is presented in Section 4, and conclusions can be found in Section 5.

2. PRELIMINARIES

2.1 Sensitivity equations

One form of dynamic systems containing n states, m parameters and l inputs can be represented as:

$$\frac{dx}{dt} = f(x, p, u) \tag{1}$$

where $x \in R^{n\times 1}$ is the state vector, $p \in R^{m\times 1}$ is the parameter vector, and $u \in R^{l\times 1}$ is the input vector. The sensitivity equation is derived by taking the derivative of Eq. (1) w.r.t. the parameters and by applying the chain rule. The resulting dynamic sensitivity equation for the state x_j and the parameter p_i is given by:

$$\frac{ds_{ij}}{dt} = \sum_{k=1}^{n}\left(\frac{\partial f_j}{\partial x_k}\cdot s_{ik}\right) + \frac{\partial f_j}{\partial p_i} \tag{2}$$

Here, s_{ij} is defined as the sensitivity of the state x_j w.r.t. the parameter p_i. The total number of sensitivity equations is $m\times n$ since $i \in \{1,\ldots,m\}$ and $j \in \{1,\ldots,n\}$. The sensitivities can be calculated by integrating all of the ODEs simultaneously using an ODE solver. However, considering

the fact that the original model is independent of the sensitivity equations and that the sensitivity equations in Eq. (2) are independent for different p_i, only $2n$ equations need be integrated at a time for one particular p_i. After calculating the sensitivity variables using numerical integration, the sensitivity vector can be obtained by sampling at uniformly spaced time points $\boldsymbol{t} = (t_1,t_2,\ldots,t_h)^T$, where h is the time step. The i-th column of the sensitivity matrix $\mathbf{S}$ represents the sensitivity vector of the output x_j w.r.t. to the parameter p_i.

$$\mathbf{S} = \begin{bmatrix} s_{1j}(t_1) & \cdots & s_{mj}(t_1) \\ \vdots & \ddots & \vdots \\ s_{1j}(t_h) & \cdots & s_{mj}(t_h) \end{bmatrix} \tag{3}$$

The sensitivity matrix shown in Eq. (3) is normalized by dividing each sensitivity vector by (x_j^{ss}/p_i^0), where x_j^{ss} is the steady state value of the corresponding output and p_i^0 is the nominal value of the corresponding parameter.

2.2 Hierarchical clustering

Hierarchical clustering is a technique to group data based upon pairwise similarities of statistical properties (Karypis *et al.*, 1999). Results from hierarchical clustering are commonly presented in a dendrogram. For parameter set selection, hierarchical clustering is used to reduce the parameter number for estimation by determining several groups of parameters that are pairwise indistinguishable (i.e. they cannot be uniquely estimated for a reasonable level of noise in the measurements). It is then possible to only consider one parameter per group for estimation. Hierarchical clustering is implemented using the following steps:

1) Calculate the pairwise distance of the objects. For a data set containing m objects, there are $m(m\text{-}1)/2$ pairs.
2) Group the objects into a dendrogram based on the distance.
3) Determine a threshold cut-off value to partition the objects into different clusters.

In step one, cosine similarity is usually used to measure the angle between vectors (Van der Laan *et al.*, 2003). A larger angle implies a larger distance (less similarity) between two sensitivity vectors, and vice versa. As explained in Section 2.1, each column of the sensitivity matrix represents the sensitivity vector for one corresponding parameter, and the sign of the direction of the sensitivity vector has no influence on parameter set selection. The distance between two sensitivity vectors is defined by modifying the cosine similarity, shown in the following equation, where w and v represent two different sensitivity vectors, and d is the defined cosine distance.

$$d = 1 - \left|\cos\theta\right| = 1 - \frac{\left|w \cdot v\right|}{\left\|w\right\| \cdot \left\|v\right\|} \tag{4}$$

In step two, the data are clustered into a dendrogram based on the distance from step one. Commonly used linkage criteria between two clusters include complete linkage clustering, where the longest distance between two clusters is found, and single linkage clustering, where the shortest distance is found. For parameter set selection, parameters need to be clustered into several groups such that the parameters within a group are pairwise indistinguishable. Therefore, the complete linkage criterion is preferred.

A threshold cut-off value is determined in the last step. This value directly affects the number of clusters that the sensitivity vectors are partitioned into. As a rule of thumb, the cut-off value is chosen to be small enough so that any pair of sensitivity vectors within a cluster is sufficiently similar.

2.3. Dynamic optimization

Dynamic optimization refers to optimization problems that address time-varying systems. Generally, these problems seek to optimize an objective function by determining a group of input profiles that may change over time. It is worth noting that, for parameter estimation applications, the objective function is to minimize the fitting error and the parameters are assumed to be constant over time. The dynamic optimization problem is formulated as follows:

$$\max_{u(t),\,p} \quad y = y\left(x, t, p\right) \tag{5}$$

$$s.t. \quad \frac{dx}{dt} = f\left(x, p, u\left(t\right)\right), \; x\left(0\right) = x_0 \tag{6}$$

$$\Psi\left(x, u\left(t\right)\right) \le 0, \; \Omega\left(x\left(t_f\right)\right) \le 0 \tag{7}$$

Here, y is the objective function, $u(t)$ is the input profile, p are the model parameters. Eq. (6) represents the dynamic system with states x and the known initial values x_0. Path constraints and terminal constraints are denoted as Ψ and Ω, respectively.

A simultaneous approach is commonly used to solve dynamic optimization problems (Dai *et al.*, 2013; Vassiliadis *et al.*, 1994). The simultaneous approach parameterizes the input variables and also discretizes the dynamic system. The discretized system is then included as algebraic constraints in a nonlinear programming problem.

3. PARAMETER SET SELECTION FOR DYNAMIC SYSTEMS UNDER UNCERTAINTY

3.1 Visualization of effect of parameter uncertainty on sensitivity vectors

As discussed in Section 2, sensitivity vectors can be clustered into different groups in a dendrogram based on the pairwise cosine distance. However, if parameter uncertainty is considered, the sensitivity vector generated from the sensitivity equation for each parameter will not be a single fixed vector but a group of vectors distributed around the nominal vector (shown in Fig. 1.a). A sensitivity cone can be defined which contains all possible sensitivity vectors corresponding to one parameter for different values of all parameters. Any vectors inside the sensitivity cone cannot be distinguished from each other due to uncertainty in the parameter values. Furthermore, vectors associated with different parameters may have sensitivity cones that overlap due to uncertainty (shown in Fig. 1.b). One inference is that

parameters whose sensitivity cones overlap should to be grouped into the same cluster. In order to capture the largest uncertainty for each parameter p_i, $\theta_i^{\max}$ is defined in the following equation:

$$\theta_i^{\max} = \max \cos^{-1}\left(\frac{\left|\mathbf{s}_i(p)\cdot\mathbf{s}_i(p_0)\right|}{\left\|\mathbf{s}_i(p)\right\|\cdot\left\|\mathbf{s}_i(p_0)\right\|}\right) \mid \forall\ i\in\{1,\ldots,m\} \quad (8)$$

Here, $\boldsymbol{s}_i(p_0)$ and $\boldsymbol{s}_i(p)$ are the sensitivity vectors of the i-th parameter at the nominal values p_0, and at other values p, chosen from the uncertainty range of all parameters, respectively.

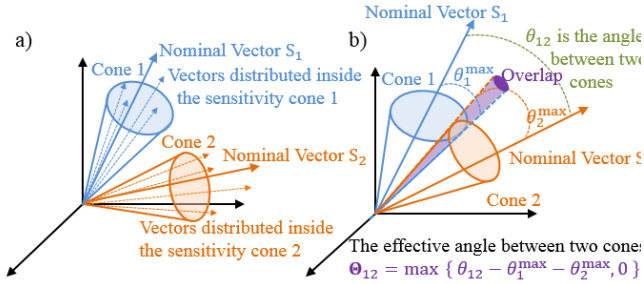

Fig. 1. Visualization of effect of uncertainty on the sensitivity vectors

It can easily be seen that two sensitivity cones will not overlap when $\theta_{12} > \theta_1^{\max} + \theta_2^{\max}$, thus the corresponding parameters p_1 and p_2 can be distinguished as long as the cut-off value is sufficiently small. In other words, if $\theta_{12} \le \theta_1^{\max} + \theta_2^{\max}$, then the two sensitivity vectors may be colinear, and the two parameters cannot be distinguished no matter how small the cut-off value is. It should be noted that this condition is conservative and accounts for the worst possible situation. As the sum of $\theta_i^{\max}$ and $\theta_j^{\max}$ will always be used to compare with θ_{ij} for different pair of parameters p_i and p_j, an effective angle Θ_{ij} is defined in Eq. (9). The corresponding cosine distance between two sensitivity cones becomes 1-cos Θ_{ij}. Hierarchical clustering is then performed based on the cosine distance of the effective angle between each pair of sensitivity cones.

$$\Theta_{ij} = \max\left\{\theta_{ij} - \theta_i^{\max} - \theta_j^{\max}, 0\right\} \quad (9)$$

This approach incorporates the parameter uncertainty of the parameter values into the parameter set selection procedure while retaining the existing methods for generating a local sensitivity matrix and performing hierarchical clustering. The next subsection will describe the numerical implementation of this approach.

3.2 Problem formulation and computational solution

For each specific parameter p_{i^*} (i^* is a fixed index here), the cone angle $\theta_{i^*}^{\max}$ can be calculated by solving the following dynamic optimization problem:

$$\min_p \quad y = \cos^2(\theta_{i^*}) = \frac{\left[\mathbf{s}_{i^*j}(p)\cdot\mathbf{s}_{i^*j}(p_0)\right]^2}{\left\|\mathbf{s}_{i^*j}(p)\right\|^2\cdot\left\|\mathbf{s}_{i^*j}(p_0)\right\|^2} \quad (10)$$

$$s.t. \quad \frac{dx_j}{dt} = f(x,p,u), \quad x_j(t_0) = x_j^0, \quad (11)$$

$$\frac{ds_{i^*j}}{dt} = \sum_{k=1}^{n}\left(\frac{\partial f_j}{\partial x_k}\cdot s_{i^*k}\right) + \frac{\partial f_j}{\partial p_{i^*}}, \quad s_{i^*j}(t_0) = 0, \quad (12)$$

$$\mathbf{s}_{i^*j}(p) = \left[s_{i^*j}(t_1) \quad \cdots \quad s_{i^*j}(t_h)\right]^T \quad (13)$$

Here, y is the square of the cosine of the angle between the sensitivity vector and the corresponding nominal vector. The square is used instead of the absolute value so as to guarantee that the objective function will be continuously differentiable (a common requirement for NLP solvers). Eq. (11) is the original dynamic system and Eq. (12) represents the sensitivity equations of the output x_j w.r.t. parameter p_{i^*}. In this problem, the input u is fixed and treated as a known parameter, while all the parameters p are perturbed to determine the maximum angle $\theta_{i^*}^{\max}$. The perturbation range of p can be specified to reasonable values according to its physical meaning or prior knowledge. For m different parameters, there will be m different maximum angles $\theta_{i^*}^{\max}$.

The problem is formulated by discretizing the dynamic system with a 3-point Radau collocation method (Kameswaran *et al.*, 2008) in AMPL and is solved with the interior-point nonlinear solver IPOPT (Wächter *et al.*, 2006).

3.3 Schematic for parameter set selection under uncertainty

Table 1. Algorithm for Parameter Set Selection with Uncertainty

Step 1.	Calculate the normalized local sensitivity matrix at the nominal parameter values.
Step 2.	Fix parameters whose nominal sensitivity vectors have small lengths (e.g., less than 5% of the largest one) at their nominal values.
Step 3.	Calculate the angle $\theta_i^{\max}$ of each sensitivity cone associated with parameter p_i that is not fixed at its nominal value in Step 2.
Step 4.	Cluster the parameters into a dendrogram by performing hierarchical clustering on the basis of the pairwise cosine distance between two cones 1-cosΘ_{ij}, where $\Theta_{ij} = \max\{\theta_{ij} - \theta_i^{\max} - \theta_j^{\max}\}$.
Step 5.	Choose a cut-off value (usually smaller than 0.05) to partition the parameters into n different clusters.
Step 6.	Select the parameters with the largest nominal sensitivity vectors of each of the n clusters as representatives from these clusters. These n parameters form the subset that needs to be estimated.

Now that all the preliminaries have been introduced, and the technique for quantifying the effect of uncertainty on the sensitivity vectors has been discussed, the steps of the parameter set selection algorithm under uncertainty are summarized in Table 1. More details about the approach can be found in the literature (Dai *et al.*, 2013).

In Table 1, step 2 represents an optional preliminary screening procedure for reducing the parameter set, as parameters with small sensitivity vector lengths are unlikely to be chosen for estimation. Step 3 solves an optimization problem to compute the sensitivity cones of all parameters. Step 4 performs clustering on the basis of the cosine distance

of the effective angle between two cones which is defined in Eq. (9). An appropriately small cut-off value is chosen for partitioning the clusters in Step 5. In Step 6, the parameters with the largest nominal sensitivity vectors of each group are chosen as the representatives for these groups. These selected parameters form the parameter subset that should be estimated.

4. CASE STUDY: TNF-α SIGNALING PATHWAY MODEL

Modeling and analysis of intracellular signaling networks is an important area in systems biology. Signaling pathways initiate essential processes for regulating cell growth, division, apoptosis, or responses to environmental stimuli. These pathways include a large number of components, which detect, amplify, and integrate diverse external signals to generate responses, such as changes in enzyme activity or gene expression. It is infeasible to measure all the components in these pathways which limits the number of parameters that can be estimated. Therefore, the values of most of the kinetic parameters are taken directly from the literature and often contain a significant level of uncertainty.

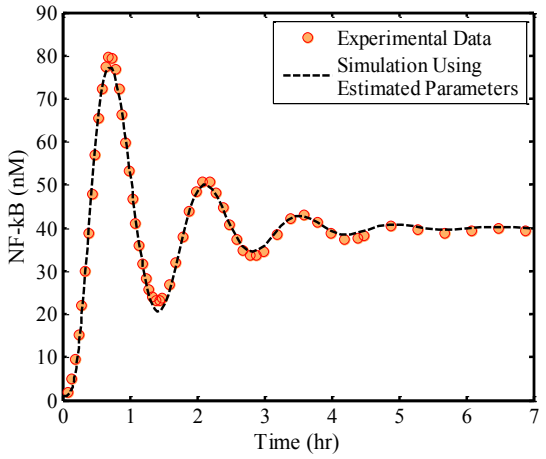

Fig. 2. Simulation result of TNF-α signaling pathway model using estimated parameters determined by the presented approach.

A TNF-α signaling pathway model (Huang *et al.*, 2008) is used here to illustrate the selection of a subset of uncertain parameters for estimation using limited experimental data. This model consists of 37 state variables and 60 parameters. The TNF-α concentration is the input and the concentration of the transcription factor, NF-κB in the nucleus, is the only measured output. The mathematic model, the initial value of the state variables and the nominal values of the unknown parameters can be found in the literature (Huang *et al.*, 2008).

First, sensitivity equations are generated based on Eq. (2). Then, an extended model is formulated by combining the original signaling pathway model and the sensitivity equations. According to the experimental data (shown in Fig. 2), the system reaches steady state after approximately 6 hours, so the sensitivity vector is generated by sampling the sensitivity variables from 0 to 6 hours with a step size of 1 minute. The sensitivity vectors corresponding to different parameters compose the sensitivity matrix of the form in Eq. (3). After applying step 2 in Table 1, 18 parameters are left as the candidates for parameter clustering: p_1, p_3, p_5, p_6, p_7, p_9, p_{10}, p_{14}, p_{15}, p_{19}, p_{21}, p_{22}, p_{25}, p_{27}, p_{32}, p_{40}, p_{46}, and p_{60}.

Assuming that there is no uncertainty, the parameters are grouped in a dendrogram, as shown in Fig. 3.a), using the complete linkage clustering. The parameters can be grouped into different numbers of clusters, based upon the cut-off value of the cosine distance. For example, if a cut-off value of

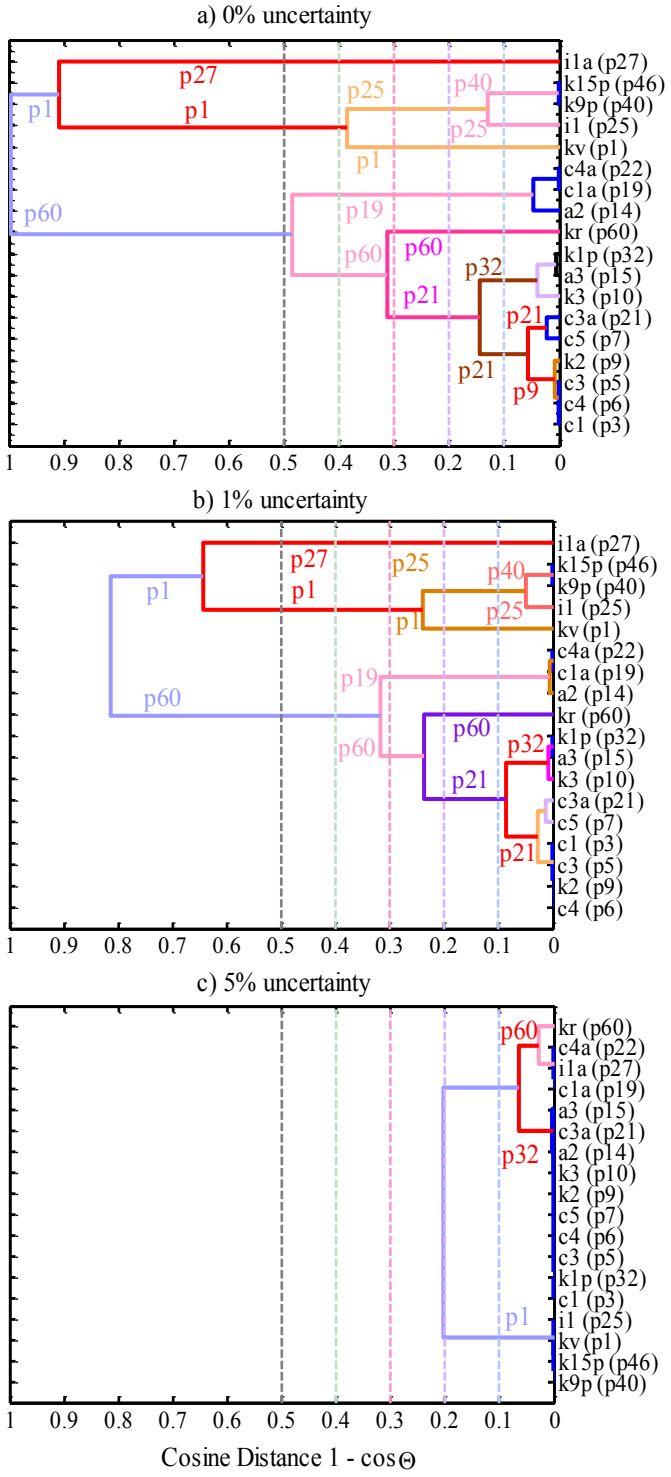

Fig. 3. Hierarchical clustering of parameters of the TNF-α signaling pathway model including uncertainty.

0.1 is chosen, the parameters can be grouped into eight distinguishable clusters (from top to bottom) containing $\{p_{27}\}$,

$\{p_{46}, p_{40}\}$, $\{p_{25}\}$, $\{p_1\}$, $\{p_{22}, p_{19}, p_{14}\}$, $\{p_{60}\}$, $\{p_{32}, p_{15}, p_{10}\}$, and $\{p_{21}, p_7, p_9, p_5, p_6, p_3\}$. The corresponding parameter subset that should be estimated is $p_{27}, p_{40}, p_{25}, p_1, p_{19}, p_{60}, p_{32}$, and p_{21}, since these are the parameters with the largest norms of the nominal sensitivity vectors in each of these eight clusters. If the cut-off value is increase to 0.5, the parameters can now only be grouped into three clusters containing $\{p_{27}\}$, $\{p_{46}, p_{40}, p_{25}, p_1\}$, and $\{p_{22}, p_{19}, p_{14}, p_{60}, p_{32}, p_{15}, p_{10}, p_{21}, p_7, p_9, p_5, p_6, p_3\}$, where p_{27}, p_1, and p_{60} should be estimated.

Table. 2. The largest angle (°) of the sensitivity cones for each parameter for two different uncertainty ranges.

	p_1	p_3	p_5	p_6	p_7	p_9	p_{10}	p_{14}	p_{15}
1%	6.6	4.3	5.2	4.8	1.0	5.0	3.6	7.3	4.8
5%	29.9	23.0	25.3	24.9	24.5	24.9	19.2	34.0	25.0
	p_{19}	p_{21}	p_{22}	p_{25}	p_{27}	p_{32}	p_{40}	p_{46}	p_{60}
1%	4.3	2.2	4.9	6.3	9.1	5.0	5.0	5.0	1.4
5%	22.8	15.1	23.2	29.4	45.4	26.0	24.1	24.1	6.1

The cluster number is also influenced by the uncertainty of the parameters. The larger the uncertainty, the shorter will be the cosine distance between pairs of sensitivity cones, thus fewer clusters can be partitioned using the same cut-off value. Furthermore, once two sensitivity cones overlap, the corresponding two parameters cannot be distinguished no matter how small the cut-off value is chosen. In this example, the largest angle of each sensitivity cone for two different uncertainty ranges is listed in Table 2, the cosine distance between each pair of sensitivity cones is calculated based on Eq. (4) and Eq. (9), and the dendrograms for different uncertainty ranges are shown in Fig. 3.b) and 3.c). It can be seen from Fig. 3 that an increase of the uncertainty results in the content of the dendrogram being pushed towards zero and the sensitivity cones start to overlap, resulting in clusters of indistinguishable parameters. For example, if 5% uncertainty is considered, the cosine distance between the sensitivity cones of p_{25}, p_1, p_{46}, and p_{40} are 0. Thus, no matter how small the cut-off value is, these parameters cannot be partitioned into different groups.

The D-optimality criterion, which is the most popular experimental design criterion, is used in this example to verify the performance of the parameter set selection. This criterion minimizes the volume of the confidence ellipsoid with an arbitrary fixed confidence level for a least-squares estimator. If no uncertainty is considered, then the criterion value is obtained by calculating the nominal sensitivity matrix. However, if a reasonable uncertainty range is considered, e.g., 1% or 5%, a Monte Carlo simulation is performed where 10,000 parameter sets are randomly chosen where the parameters are sampled to have values within the uncertainty range. The average criterion value is obtained from calculating the sensitivity matrix for these different parameter values. Therefore, the magnitude of the criterion value reflects the performance of the parameter selection for estimation. Since 18 parameters are considered for parameter set selection, there are a total number of $2^{18}-1$ different combinations of parameter subsets. It is obviously infeasible to compare all these possible combinations here.

Instead, twenty different parameter subsets are listed in Table 3 and compared to each other. Set #1 is selected from the literature (Huang *et al.*, 2008). Set #2, #3, #4, and #5 are selected when the parameters are clustered based on the nominal sensitivity vectors and the cut-off value is chosen to be 0.5, 0.4, 0.2 and 0.1 respectively (shown in Fig. 3.a)). When the 5% uncertainty is considered, the parameters are clustered based upon the sensitivity cones shown in Fig. 3.c). If the cut-off value is 0.05, then the parameters are only partitioned into three clusters and three parameters are selected, each one a representative of a cluster, as is given by set #6. In order to study the effect of correlation among the parameters on the performance of parameter selection, two representatives (i.e., p_{19}, p_{60} in set #7) from the same cluster, and three representatives (i.e., p_{19}, p_{27}, p_{60} in set #8) from the same cluster are selected. Set #9 and #10 are chosen as counter examples where the criterion values are large when there is no uncertainty considered, while the value decreases significantly even when only a small amount of uncertainty is considered. Set #10 to #20 are random combinations of four parameters selected from set #5.

Table 3. Different combinations of parameter subsets and the corresponding criterion values*

#	Parameter Subset	0%	5%	10%	20%	30%
1	$\{p_5, p_{32}, p_{60}\}$	-58.2	-59.3	-59.9	-60.3	-61.9
2	$\{p_1, p_{27}, p_{60}\}$	-23.5	-34.3	-34.8	-39.9	-42.4
3	$\{p_1, p_{19}, p_{27}, p_{60}\}$	-51.66	-66.65	-66.96	-72.53	-75.51
4	$\{p_1, p_{19}, p_{21}, p_{25}, p_{27}, p_{60}\}$	-57.7	-90.2	-90.7	-97.2	-100.5
5	$\{p_1, p_{19}, p_{21}, p_{25}, p_{27}, p_{32}, p_{40}, p_{60}\}$	-171.7	-172.5	-174.2	-180.4	-188.3
6	$\{p_1, p_{32}, p_{60}\}$	-39.2	-40.7	-40.9	-42.6	-46.2
7	$\{p_1, p_{19}, p_{60}\}$	-45.4	-42.7	-43.6	-49.1	-51.8
8	$\{p_{19}, p_{27}, p_{60}\}$	-38.8	-54.3	-55.3	-56.9	-59.0
9	$\{p_{25}, p_{27}, p_{60}\}$	-16.6	-46.1	-46.6	-47.0	-49.5
10	$\{p_{21}, p_{27}, p_{60}\}$	-24.1	-47.4	-47.8	-48.1	-50.54
11	$\{p_1, p_{27}, p_{32}, p_{40}\}$	-78.2	-79.4	-80.7	-81.5	-85.6
12	$\{p_1, p_{27}, p_{32}, p_{60}\}$	-59.3	-61.8	-62.8	-66.2	-69.8
13	$\{p_1, p_{27}, p_{40}, p_{60}\}$	-60.9	-61.1	-61.7	-65.6	-69.0
14	$\{p_{19}, p_{27}, p_{32}, p_{40}\}$	-97.7	-98.5	-99.7	-100.6	-102.5
15	$\{p_{19}, p_{27}, p_{32}, p_{60}\}$	-80.1	-82.2	-82.5	-83.2	-86.4
16	$\{p_{19}, p_{27}, p_{40}, p_{60}\}$	-81.3	-81.5	-81.6	-82.6	-85.6
17	$\{p_{25}, p_{27}, p_{32}, p_{40}\}$	-88.7	-90.8	-90.1	-91.5	-92.9
18	$\{p_{25}, p_{27}, p_{32}, p_{60}\}$	-73.0	-73.4	-73.9	-74.3	-76.8
19	$\{p_{25}, p_{32}, p_{40}, p_{60}\}$	-82.1	-83.0	-83.8	-84.5	-86.2
20	$\{p_{27}, p_{32}, p_{40}, p_{60}\}$	-82.6	-84.0	-84.2	-85.9	-86.5

* D-optimality criterion: max $\log_{10}(\det(\mathbf{S}^T\mathbf{S}))$, $\mathbf{S}$ is the normalized sensitivity matrix of the subset

There are several trends that can be clearly seen when analyzing the data form Table 3. From set #2 to #5, it can be seen that the more parameter are selected, the worse estimation performance will get, no matter how small the uncertainty. Additionally, the average performance of sets of three parameters is better than that of the sets composed of four parameters, which can be found from set #11 to #20. From set #6 to #8, it can be seen that if more than one parameter is selected from the same cluster and no parameter is selected from other clusters, the performance will get significantly worse. From set #9 and #10, an interesting conclusion can be drawn as some parameter subsets will

exhibit a good estimation performance when there is no uncertainty, however, even if only a small amount of uncertainty is considered, then the performance will drop significantly. This phenomenon directly shows the importance of considering uncertainty when selecting parameter subsets for estimation. Sets #2 and #6 are highlighted in the table since they are the best two candidates that are chosen for parameter estimation. The parameter estimation result using set #6 can be seen in Fig. 2. It is important to note that there is likely no perfect parameter set for estimation of uncertain systems, especially since determination of an appropriate cut-off value is still a topic of current research; however, it is important to determine one or more potential sets of parameters which are good candidates for estimation and differentiate those from other sets which would clearly result in worse prediction accuracy.

5. CONCLUSION

This work applies an approach for parameter set selection under uncertainty to a signal transduction pathway model. The technique is based upon an extension of existing methods that cluster the parameters according to their sensitivity vectors. The extension is made by realizing that the sensitivity vectors can vary due to the parameter uncertainty and that a sensitivity cone can be defined, where all possible sensitivity vectors of a parameter computed for different parameter values lie within one sensitivity cone. Computation of the sensitivity cones is non-trivial, and a dynamic optimization formulation is presented for the computation. The effective angle between two sensitivity cones can then be used for clustering the sensitivity cones. This technique is used in a case study where the parameter subset determined by this approach exhibits good fitting and prediction accuracy for the investigated TNF-α signaling pathway model.

ACKNOWLEDGMENT

The authors gratefully acknowledge partial financial support by the National Science Foundation (Grant CBET#0941313) and the American Chemical Society (ACS-PRF#50978-ND9).

REFERENCES

Anh, D. T., Bonnet, M.P., Vachaud, G., Van Minh, C., Prieur, N., Vu Duc, L., & Lan Anh, L. (2006). Biochemical modeling of the Nhue River (Hanoi, Vietnam): Practical identifiability analysis and parameters estimation. *Ecological Modelling, 193*(3-4), 182-204.

Chu, Y., & Hahn, J. (2007). Parameter set selection for estimation of nonlinear dynamic systems. *AIChE Journal, 53*(11), 2858-2870.

Chu, Y., & Hahn, J. (2008). Parameter set selection via clustering of parameters into pairwise indistinguishable groups of parameters. *Industrial & Engineering Chemistry Research, 48*(13), 6000-6009.

Dai, W., Bansal, L., Word, D., & Hahn, J. (2013). Parameter set selection for dynamic systems under uncertainty via dynamic optimization and hierarchical clustering. *AIChE Journal, 60*(1), 181-192.

Dai, W., Word, D., & Hahn, J. (2013). Modeling and dynamic optimization of fuel-grade ethanol fermentation using fed-batch process. *Control Engineering Practice, 22*, 231-241.

Huang, Z., Senocak, F., Jayaraman, A., & Hahn, J. (2008). Integrated modeling and experimental approach for determining transcription factor profiles from fluorescent reporter data. *BMC Systems Biology, 2*(1), 64.

Kameswaran, S., & Biegler, L.T. (2008). Convergence rates for direct transcription of optimal control problems using collocation at Radau points. *Computational Optimization and Applications, 41*(1), 81-126.

Karypis, G., Han, E.-H., & Kumar, V. (1999). Chameleon: Hierarchical clustering using dynamic modeling. *Computer, 32*(8), 68-75.

Kravaris, C., Hahn, J., & Chu, Y. (2013) Advances and Selected Recent Developments in State and Parameter Estimation. *Computers & Chemical Engineering,* 51, 111-123.

Poyton, A., Varziri, M.S., McAuley, K.B., McLellan, P., & Ramsay, J. O. (2006). Parameter estimation in continuous-time dynamic models using principal differential analysis. *Computers & Chemical Engineering, 30*(4), 698-708.

Van der Laan, M., & Pollard, K.S. (2003). A new algorithm for hybrid hierarchical clustering with visualization and the bootstrap. *Journal of Statistical Planning and Inference, 117*(2), 275-303.

Vassiliadis, V., Sargent, R., & Pantelides, C. (1994). Solution of a class of multistage dynamic optimization problems. 1. Problems without path constraints. *Industrial & Engineering Chemistry Research, 33*(9), 2111-2122.

Wächter, A., & Biegler, L.T. (2006). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming, 106*(1), 25-57.