Identification of a time-varying intracellular signalling model through data clustering and parameter selection: application to NF-κB signalling pathway induced by LPS in the presence of BFA

Dongheon Lee$^{1,2}$, Arul Jayaraman$^{1,3}$, Joseph Sang-II Kwon$^{1,2}$

$^1$Arte McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX 77843, USA
$^2$Texas A&M Energy Institute, Texas A&M University, College Station, TX 77843, USA
$^3$Department of Biomedical Engineering, Texas A&M University, College Station, TX 77843, USA
E-mail: kwonx075@tamu.edu

Abstract: Developing a model for a signalling pathway requires several iterations of experimentation and model refinement to obtain an accurate model. However, the implementation of such an approach to model a signalling pathway induced by a stimulus, which can be activated by many stimuli with different, usually unknown reaction mechanisms, is non-trivial. Specifically, as the initial model is likely to be erroneous, this approach may require a large number of iterations between experiments and model refinement to reach a relatively satisfying model. First, a number of different initial models, each of which corresponds to a different hypothesis about the underlying mechanisms and generate new hypotheses to be tested in new experiments. Previously, systems biology approach has been implemented to gain new insights into various signalling pathways such as the nuclear factor κB (NF-κB) signalling pathway [5, 6], Janus family of kinases – signal transducer and activator of transcription signalling pathway [7], and mitogen-activated protein kinase signalling pathway [8]. Ideally, one would like to develop a comprehensive signalling pathway model that can predict the signalling dynamics under various conditions. However, this can be a difficult task as a single signalling pathway can be activated by many stimuli with different, usually unknown, corresponding reaction mechanisms. For example, out of around 100 different stimuli of the NF-κB signalling pathway [9], only a handful, such as tumour necrosis factor-α (TNFα) and lipopolysaccharide (LPS), and their reaction mechanisms are well characterised. Consequently, investigating and modelling a signalling pathway induced by a stimulus, which has not been well studied, is non-trivial. Specifically, as the pathway is only partially known beforehand, a number of different model structures need to be formulated and discriminated, which can become very challenging. In the literature, different approaches have been proposed and implemented for solving this problem. First, extensive experiments are performed to characterise as many reactions as possible between intracellular molecules, which in practice is nearly impossible due to the large number of interactions to be studied. Second, the aforementioned iterative approach between modelling and experiments can be implemented to improve the model gradually [2, 10, 11]. However, as the initial model is likely to be erroneous, this approach may require a large number of iterations between experiments and model refinement to reach a relatively satisfying model. Third, a number of different initial models, each of which corresponds to a different hypothesis on the signalling pathway structure, are synthesised from the beginning, and the best model structure is selected by solving an optimisation problem against experimental observations [12–14]. Although the optimisation-based approach is promising, it has several numerical and algorithmic challenges such as efficiently finding a global optimal solution for a large system [12–16].

Alternatively, when an accurate model for describing a signalling pathway under one stimulus is available, we can modify that model to describe the same signalling pathway under a lesser-known stimulus [17, 18]. Hereafter, we refer to the model constructed for the well-studied stimulus as the nominal model. The rationale for using the nominal model is two-fold. First, the nominal model already contains a number of important pathway components as well as their interactions, which are likely to be important under the lesser-known stimulus as well. Second, this approach avoids the lengthy model selection procedure, which requires a number of different candidate models to be synthesised, calibrated, and compared [17]. On the other hand, the structure of the nominal model is likely to be insufficient to describe the signalling dynamics under the lesser-known stimulus due to unincorporated and unknown reactions and components specific to this stimulus [17]. Therefore, a poorly characterised signalling pathway induced by a lesser-known stimulus needs to be described by a data-driven approach to complement the inaccuracy of the
nominal model. Here, we choose to introduce time-varying parameters to the nominal model, which is usually time-invariant, based on the available experimental measurements [18, 19]. Through this data-driven approach, a more accurate model for the lesser-known stimulus can be derived based on the nominal model and the available data.

Motivated by the above considerations, we propose a numerical scheme to construct a time-varying model to simulate an intracellular signalling pathway with a lesser-known stimulus based on a nominal model. First, global sensitivity analysis (SA) is performed on the nominal model to identify a set of parameters that are identifiable given the model structure and experimental observations, and only these parameters are assumed to vary with time. Next, the temporal profiles of the model parameters are partitioned into several temporal subdomains whose boundaries are determined by clustering the experimental observations. And the parameters determined by the SA have fixed values in each temporal subdomain. Finally, a least-squares problem is solved to estimate the values of the parameters in each temporal subdomain by minimising the difference between the model predictions and the experimental data.

The paper is organised as follows: first, the motivation for formulating a time-varying signalling model is presented. Second, the proposed methodology that consists of the optimal temporal clustering and the global SA to construct a time-varying model is presented in details. Finally, the temporal profiles of the model parameters are implemented to develop the model for a time-varying NF-κB signalling pathway induced by LPS in the presence of brefeldin A (BFA) to assess the efficiency and accuracy of the proposed scheme.

2 Background

2.1 System description

Consider an intracellular signalling pathway initiated by an external stimulus, \( u \), which has been well characterised by the following model:

\[
\begin{align*}
\dot{x} &= f(x, \theta, u; t), \quad x(0) = x_0, \\
y &= g(x, \theta, u; t)
\end{align*}
\]

where \( x \in \mathbb{R}^n \) is the state vector, \( \theta \in \mathbb{R}^m \) is the parameter vector, \( x_0 \) is the initial value of the state vector \( x \), and \( y \in \mathbb{R}^p \) is the output vector.

When a lesser-known stimulus, \( u_a \), is added to a cell, the signalling dynamics deviate significantly from those predicted by (1). Due to the disparity in our understanding of the roles of \( u_a, u \), and their interplay in the signalling dynamics, little information on the signalling dynamics is available a priori. Consequently, the construction of a high-fidelity model, which faithfully simulates the signalling dynamics initiated by \( u_a \) and \( u \), requires iterative experimentation and model refinement, which can be an arduous and lengthy process [10, 20, 21].

A more viable alternative is to approximate the dynamics induced by \( u_a \) through introducing a time-varying model where \( \theta \) in (1) changes with time so that the well-defined model (1) can be used to describe the signalling dynamics under the two stimuli [19]. To this end, the temporal profile of \( \theta \) is described as piecewise constant functions. Under this representation, the entire temporal domain is partitioned into several temporal subdomains, of which has its own parameter values. Consequently, the following modified form of (1) is used to describe the signalling dynamics under two stimuli:

\[
\begin{align*}
\dot{x} &= f(x, \theta_{\sigma(t)}, u; t), \quad x(0) = x_0, \\
y &= g(x, \theta_{\sigma(t)}, u; t) \\
\sigma(t) &= i \text{ if } t \in T_i, \quad i = 1, \ldots, n_\theta
\end{align*}
\]

where \( \theta \in \mathbb{R}^m \), where \( i = 1, \ldots, n_\theta \), is the vector of parameter values used when the current time \( t \) belongs to the temporal subdomain, \( \sigma \), \( n_\theta \) is the number of temporal subdomains, and \( \sigma(t) \) is the discrete variable to denote which \( \theta \) is used at the time \( t \). Under this formulation, the overall temporal domain is partitioned into \( n_\theta \) subdomains, where different values of \( \theta \) are used. From here on, \( u_a \) is neglected since the use of \( \theta_{\sigma(t)} \) implies the presence of \( u_a \).

2.2 Experimental measurements

In order to train and validate (2), \( y \) is measured experimentally under different conditions. Here, \( n_\theta \) different values of \( u \) with a fixed value of \( u_a \) are used. Due to the technical and economic constraints in a biology experiment, \( y \) can be measured at only a few sampling time instants, \( t_i, i = 1, \ldots, N_t \), where \( N_t \) is the number of sampling instants [22]. Also, it should be noted that commonly used biochemical assays such as Western blots, flow cytometry, or microarrays typically give qualitative or semi-quantitative datasets, which measure relative but not absolute concentrations of biomolecules [23]. In other words, the measured output is defined as

\[
z^c_i(t) = c_i \cdot \bar{y}(u; t) + v_i
\]

where \( z^c_i, i = 1, \ldots, n_\theta \), is the relative measured output that is corrupted with measurement noise under the input \( u' \), \( s = 1, \ldots, n_\theta \), \( y_i \) is the output in absolute concentration that is not directly measurable in the experiments, \( c_i \) is the proportional constant relating \( z^c_i \) and \( y_i \), and \( v_i \) is the measurement noise. Here, it is assumed that the mean value of \( v_i \) can be inferred during the equipment calibration procedure [24].

Since the values of \( e \) are not usually known beforehand, an alternative quantity is computed to facilitate the comparison between the model and the experimental measurements. Specifically, fold changes in the measurements are calculated as follows [21]:

\[
\begin{align*}
\hat{y}_i(t) &= \frac{z^c_i(t) - \hat{y}_i}{z^c_i(t) - \hat{y}_i} \\
&= \frac{\hat{y}_i(u'; t)}{\hat{y}_i(u'; t)}
\end{align*}
\]

where \( \hat{y}_i(t) \) is fold change of \( z^c_i \) at the time instant at \( t_i \), \( t_i \) is the first sampling instant (usually \( t_i = 0 \)), and \( \hat{y}_i \) is the average measurement noise that can be obtained by performing a negative control measurement without reagents.

2.3 Problem statement

In this study, we seek to construct a time-varying model (2) by estimating the temporal dynamics of \( \theta \), and this can be achieved by addressing the following two problems:

**Problem 1**: Given the model (2) and the experimental measurements (3) and (4), determine the number of temporal subdomains, \( n_\theta \), as well as the temporal subdomains, \( T_i, \forall i = 1, \ldots, n_\theta \),

Since \( n_\theta \) and \( T_i \) are not known a priori, the experimental measurements are clustered to estimate the value of \( n_\theta \) and the temporal subdomains, \( T_i \).

**Problem 2**: Given the model (2), the experimental measurements (3), and the temporal subdomains of the parameters \( (T_1, \ldots, T_{n_\theta}) \), estimate the values of parameters, \( \theta_i, i = 1, \ldots, n_\theta \), in each temporal subdomain.

For many intracellular signalling pathways, only a small subset of \( \theta \) is identifiable from the experimental measurements [22, 25]. As a result, the additional parameters introduced in the time-varying model (2), which increases the size of the parameter space by \( n_\theta \)-fold, are likely to be even more undetachable. Hence, a sequential parameter selection methodology is implemented to identify the most important parameters in \( \theta \), and the values of these
parameters in each temporal subdomain are estimated. The resultant model then can be used to investigate the system dynamics and design the optimal experiments for future studies to advance our understanding.

3 Temporal clustering

Since the intracellular signalling dynamics are described by the time-varying model (2) with the piecewise constant \( \theta \), the value of \( n_a \) and all the temporal subdomains, \( \mathbb{T}_i \), need to be determined. In this work, they are inferred by clustering the experimental measurements into several temporal subdomains in a way that the data points contained in each subdomain exhibit similar temporal behaviours [26]. This inference assumes that the time-invariant parameters in one temporal subdomain, \( \mathbb{T}_i \), result in the relatively uniform dynamics in \( y \).

For the given experimental measurements \( D \in \mathbb{R}^{N_y \times N_t} \), where \( N_t = n_a \cdot n_u \),

\[
\begin{bmatrix}
\bar{y}_1(t_1) & \cdots & \bar{y}_1(t_{N_i}) \\
\vdots & \ddots & \vdots \\
\bar{y}_{N_y}(t_1) & \cdots & \bar{y}_{N_y}(t_{N_i})
\end{bmatrix}
\]

a clustering algorithm will assign \( N_i \) column vectors of \( D \) into \( n_a \) different temporal subdomains by minimising the distance between vectors in a subdomain and the centre of the subdomain, which is measured by the following intra-cluster error sum [26]:

\[
\Lambda = \sum_{i=1}^{N_i} \sum_{k=1}^{n_a} z_{ik} \| D_k - c_i \|^2;
\]

where \( D_k \) is the \( k^{th} \) column of \( D \), \( z_{ik} \) is a binary variable indicating whether \( D_k \) is in the \( k^{th} \) subdomain, and \( c_i \in \mathbb{R}^{N_y} \) is the centre of the \( k^{th} \) cluster.

Since a value of \( n_a \) is not known a priori, a clustering method is implemented with all possible number of subdomains \( (1, \ldots, N_i) \) to find an optimal \( n_a \) by computing and comparing the values of \( \Lambda \) as well as the inter-cluster error sum, \( \Gamma \), which is defined as follows [26]:

\[
\Gamma = \sum_{i=1}^{N_i} \| c' - c_i \|^2
\]

where \( c' \in \mathbb{R}^{N_y} \) is the global cluster centre, which is defined as

\[
c'_j = \frac{1}{N_i} \sum_{i=1}^{N_i} D_{ji}
\]

where \( c'_j \) is the \( j^{th} \) element of \( c' \). When an optimal clustering configuration is achieved, the value of \( \Lambda \) is minimised while the value of \( \Gamma \) is maximised to achieve the maximum intra-cluster similarity and inter-cluster dissimilarity [26, 27]. Mathematically, this is quantified by the clustering balance, \( \epsilon \), which was proposed in [27], as follows:

\[
\epsilon = 0.5 \Gamma + 0.5 \Lambda
\]

where 0.5 in front of \( \Gamma \) and \( \Lambda \) is a weight coefficient, which can be adjusted based on the problem [28]. As \( \Lambda \) and \( \Gamma \) are expected to decrease and increase, respectively, with the increase in the number of subdomains, a turning point in the value of \( \epsilon \) determines the optimal value of \( n_a \) [26]. Once the value of \( n_a \) is determined, all the \( \mathbb{T}_i \) can also be determined by clustering \( D \) into \( n_a \) temporal subdomains.

4 Parameter estimation

The aim of the parameter estimation is to quantitatively calibrate a model so that it can make an accurate and robust prediction of the system, which then can be used to analyse underlying mechanisms and design optimal experiments [4, 20]. We can formulate a parameter estimation for (2) as a least-squares problem to minimise the difference between model predictions and measurements as follows:

\[
\min_{\epsilon, \theta_1, \ldots, \theta_{n_u}} \sum_{i=1}^{n_a} \sum_{j=1}^{N_i} \| y(t_j; \theta) - \bar{y}(t_j) \|^2 \quad (10a)
\]

\[
s.t. \quad \hat{x} = f(x, \theta, u, t_i), \quad x(0) = x_0 \quad (10b)
\]

\[
y = g(x, \theta, u, t_i) \quad (10c)
\]

\[
\sigma(t_i) = k \quad \text{if } t_i \in \mathbb{T}_i, \quad k = \{1, \ldots, n_a\} \quad (10d)
\]

\[
\theta^b \leq \theta \leq \theta^u \quad (10e)
\]

where \( \theta^b \) and \( \theta^u \) are lower and upper bounds for the values of the model parameters, respectively.

It should be noted that the parameter estimation (10) is often ill-conditioned and results in a non-unique solution [29]. This is especially problematic for calibrating biological models since biological systems are often partially observable and over-parameterised (i.e. \( n_a \ll n_{\theta} \)) [25]. As the time-dependency of the model parameters is introduced, the issue of the non-uniqueness in the parameter estimation exacerbates since the number of parameters increases by \( n_a \cdot n_{\theta} \) fold. In order to handle this issue, we assume that only identifiable parameters, which is a subset of \( \theta \), vary with time while the remaining parameters are time-invariant and fixed at their nominal values. Consequently, this study carries out the parameter selection methodology before the parameter estimation to determine the identifiable parameters and estimate their values in each \( \mathbb{T}_i \) by solving the least-squares problem.

4.1 Parameter selection

The objective of the parameter selection procedure is to determine the identifiable parameters that will be estimated in the parameter estimation step. In this study, two global SA techniques are implemented to determine which parameters are identifiable.

4.1.1 Sensitivity analysis: In the literature, several analytical methods have been proposed to determine the parameter identifiability, including Taylor series expansion [30], differential algebra [31], or similarity transformation [32]. But these methods require symbolic manipulation and thus only applicable to a relatively small system (for \( n_a + n_{\theta} \leq 10 \)) due to the computational requirements of these methods [33].

Alternatively, the parameter identifiability can be assessed by SA, which evaluates the importance of the model parameters by quantifying changes in model outputs due to changes in model parameters. A common method is the local SA method that is based on the direct differentiation of a system model with respect to its parameters. However, the evaluation of the system model as well as its derivatives with respect to its parameters depends on the values of the model parameters, which are unknown before the parameter estimation. Therefore, a result of the local SA method is local in nature and likely to be unreliable, particularly when the parameter values are largely uncertain [34, 35].
In this study, two global SA methods, Morris method [36] and Sobol’ method [37], are implemented sequentially to determine the most important parameters. Even though the parameters take different values in each temporal subdomain, the model structure remains the same. Hence, the results of global SA on the time-invariant model will be valid for the time-varying one because the global SA computes the importance of model parameters over the entire parametric domain. Therefore, all the analysis in the following sections is conducted based on the time-invariant model (1).

### 4.1.2 Morris method

The Morris method computes the average sensitivity of a model parameter by calculating the average change in model outputs due to changes in its value. Specifically, the value of a parameter \( \theta_j \) at a time instant \( t_i \), \( j = 1, \ldots, N_p \), is perturbed by \( \Delta_j \) to compute its effect on an output \( y \), which is quantified as follows [38]:

\[
d_j(a^t; t_i) = \frac{y(x, \theta^t, \ldots, \theta_j + \Delta_j, \ldots, \theta_{N_p}, u^t; t_i) - y(x, \theta, u^t; t_i)}{\Delta_j}
\]

where \( d_j(a^t; t_i) \) is called the elementary effect of \( \theta_j \) on \( y \) at a time instant \( t_i \), \( j = 1, \ldots, N_p \). By calculating \( N_m \) different \( d_{ij} \) with \( N_m \) different values of \( \Delta_j \), the average sensitivity measure of the parameter \( \theta_j \), which is denoted as \( s_j \), is computed as follows:

\[
s_j(a^t; t_i) = \frac{1}{N_m} \frac{\theta_j}{y(x, \theta^t, u^t; t_i)} \sum_{j=1}^{N_m} \left[ d_{ij}(a^t; t_i) \right]
\]

where \( \Delta_j^k \) is the \( k \)th perturbation applied to the parameter \( \theta_j \). Here, the term \( d_{ij}(a^t; t_i) \) is normalised by \( \theta_j/y(x, \theta, u^t; t_i) \) to eliminate possible scaling effects [38]. And, the suggested value for \( N_m \) is \( r(N_0 + 1) \), where \( r \) is usually around six [39].

Then, the final scaled sensitivity of all the model outputs with respect to a parameter across all the time instants is defined as follows:

\[
S_j = \frac{1}{n_s} \frac{n_s}{n_p} \left[ s_j(a^t_1; t_1) \ldots s_j(a^t_{n_s}; t_1) \ldots s_j(a^t_{n_p}; t_1) \right]
\]

where \( s_j(a^t; t_i) \) is the average sensitivity computed under an input \( u_i \), \( s = 1, \ldots, n_s \), at a time instant \( t_i \).

Although the Morris method is conceptually simple and easy to be implemented, it has a limited capability in capturing the nonlinear output behaviour and the dependency among parameters [38]. Therefore, this study utilises the Morris method as a screening tool to reduce the number of parameters to be analysed by the Sobol’ method, which overcomes the problems of the Morris method but is computationally more expensive.

### 4.1.3 Sobol’ method

Once the Morris method screens out less important parameters from \( \theta \), the importance of the remaining parameters, which are denoted as \( \theta \in \mathbb{R}^{N_p} \), \( N_p < n_0 \), is analysed via the Sobol’ method. Different from the local SA method or the Morris method, the Sobol’ method is a variance-based method. Specifically, the sensitivity of a parameter is computed by quantifying how much each parameter contributes to the output variance. A brief overview on the Sobol’ SA method is presented below, and further details can be found elsewhere [37, 39, 40].

The main idea of the Sobol’ method is the decomposition of the model output into summands of increasing dimensionality. Specifically, a model output \( y \) can be decomposed as follows:

\[
y(\theta) = y_0 + \sum_{i=1}^{n_p} yi(\bar{\theta}) + \sum_{i=1}^{n_p-1} \sum_{j=i+1}^{n_p} yi\cdot ji(\theta, \bar{\theta}) + \sum_{i=1}^{n_p-1} \sum_{j>i} \sum_{k=1}^{j-1} \sum_{l=k+1}^{j-1} \sum_{m=l+1}^{j-1} \cdots \sum_{p=q+1}^{j-1} yi\cdot ji\cdot kj\cdot l\cdot m\cdots p(\theta) + \cdots + y_{1\ldots N_p}(\theta)
\]

where \( s \), where \( 3 \leq s \leq n_p \), is the number of parameters involved in a summand \( y_{i_1\ldots i_s}(\theta_{i_1}, \ldots, \theta_{i_s}) \), and \( y_i \) is defined as follows:

\[
y_i = \int y(\theta) d\theta
\]

Here, we assume \( y_0 \) is a constant and the integrals of every summand over any of its variables are zero, i.e.

\[
\int y_{i_1\ldots i_s}(\theta_{i_1}, \ldots, \theta_{i_s}) d\theta_k = 0
\]

in order for the decomposition of \( y \) as (14) to hold [22, 37].

Then, by assuming that \( y \) is square integrable, its variance can be expressed as follows:

\[
V = \int (y^2(\theta) - y^2) d\theta
\]

\[
= \int \left( \sum_{i=1}^{n_p} y_i^2(\bar{\theta}) + \sum_{i=1}^{n_p-1} \sum_{j=i+1}^{n_p} y_{i\cdot j}^2(\theta, \bar{\theta}) + \cdots + y_{1\ldots N_p}^2(\theta) \right) d\theta
\]

\[
= \sum_{i=1}^{n_p} V_i + \sum_{i=1}^{n_p-1} \sum_{j=1}^{n_p} V_{i\cdot j} + \sum_{i=1}^{n_p-1} \sum_{j>i} \sum_{k=1}^{j-1} \sum_{l=k+1}^{j-1} \cdots \sum_{p=q+1}^{j-1} V_{i\cdot j\cdot k\cdot l\cdot m\cdots p}
\]

where \( V \) is the total variance of a model output \( y \), and \( V_{i\ldots j} \) is the partial variance of the output due to the parameters \( \theta_{i}, \ldots, \theta_{j} \). Based on the \( V \) and \( V_{i\ldots j} \), the importance of a parameter, \( \theta_i \), can be quantified by the first order and total sensitivity indices, which are defined as follows [41]:

\[
S_{i} = \frac{V_i}{V} \quad S_{i} = \frac{V_i + \sum_{k=i+1}^{n_p} V_{i\cdot k} + \cdots + V_{i\ldots n_p}}{V - 1}
\]

where \( S_{i} \) and \( S_{i} \) are the first-order and total sensitivity indices, respectively, of the model parameter \( \theta_i \). \( V_j \) is the partial variance of a model output due to \( \theta_j \), and \( V_{i\ldots j} \) is the partial variance of a model output due to join effects of the model parameters \( \theta \) except \( \theta_j \). Here, \( S_{i} \) refers to the main effect of the parameter \( \theta_i \), and \( S_{i} \) measures the importance of a parameter \( \theta_i \) by taking into account the direct effect \( \{S_{i}\} \) as well as its joint effects with other parameters. It should be noted that the difference between \( S_{i} \) and \( S_{i} \) indicates how much \( \theta_i \) is involved in interactions with other parameters in terms of changing the model output [40].

In this study, a Monte Carlo method proposed by Homma and Saltelli [41] is implemented to estimate the total sensitivity indices. First of all, two matrices \( (A \in \mathbb{R}^{N_0 \times n_0}) \) and \( (B \in \mathbb{R}^{N_0 \times n_0}) \) are generated randomly from the parameter space via a Sobol’ sequence to produce parameter samples without overlapping [37, 42]. Here, \( N_0 \) is the sample size for the Monte Carlo estimation, which is typically around a few hundreds to thousands [39]. Then, another set of matrices \( C_j \in \mathbb{R}^{N_0 \times n_0} \), \( V_j = 1, \ldots, n_p \), can be defined for
every parameter, \( \theta_j \), by replacing the \( j \)th column of \( B \) with the \( j \)th column of \( A \). Next, the model outputs can be computed for all the sampled parameter values in the matrices \( A, B, \) and \( C_p \). Finally, the first-order and total sensitivity indices in (18) can be approximated as follows:

\[
S_i(u^t; t) \approx \frac{1}{N_s} \sum_{i=1}^{N_s} \left[ \gamma(u^t, a^{i0}; t) \cdot \gamma(u^t, e^{i0}; t) - f_j^i(u^t; t) \right],
\]

\[
S_T(u^t; t) \approx 1 - \frac{1}{N_s} \sum_{i=1}^{N_s} \left[ \gamma(u^t, b^0; t) \cdot \gamma(u^t, e^{i0}; t) - f_j^i(u^t; t) \right],
\]

where \( a^{i0}, b^0, \) and \( e^{i0} \) are \( i \)th rows of \( A, B, \) and \( C_p \) respectively, and \( f_j^i(u^t; t) \) is defined as follows:

\[
f_j^i(u^t; t) = \left( \frac{1}{N_s} \sum_{i=1}^{N_s} \gamma(u^t, a^{i0}; t) \right)^2
\]

Since the proposed model (2) has \( n_s \) outputs obtained under \( n_s \) values for \( u \) sampled at \( N_t \) time instants, a lumped sensitivity metric for the total sensitivity index is defined to ease the parameter selection process:

\[
ST_{ij} = \frac{1}{N_t} \sum_{t=1}^{N_t} S_i(u^t; t)
\]

where \( ST_{ij} \) is the sensitivity of \( y_i \) with respect to \( \theta_j \), which will be used to select the most influential parameters, and \( S_T \) is the total sensitivity index, \( ST \), computed for an output, \( y_i, i = 1, \ldots, n_s \). Similarly, a lumped sensitivity metric for the first-order sensitivity index is defined as follows:

\[
SS_{ij} = \frac{1}{N_t} \sum_{t=1}^{N_t} S_i(u^t; t)
\]

where \( SS_{ij} \) is the sensitivity of \( y_i \) with respect to \( \theta_j \), which will be used to select the most influential parameters, and \( S_S \) is the first-order sensitivity index, \( SS \), computed for an output, \( y_i, i = 1, \ldots, n_s \).

Based on the values of \( ST_{ij} \) and \( SS_{ij} \), a set of identifiable parameters, \( \Theta \in \mathbb{R}^{n_p} \), where \( n_p \leq n_p \), can be identified from \( \hat{\Theta} \). And the final parameter estimation problem (4) is reformulated as follows:

\[
\min_{\theta^{\text{min}} \leq \theta \leq \theta^{\text{max}}} \sum_{s = 1}^{N_s} \sum_{i = 1}^{N_i} (\gamma(u^{i0}; t) - \gamma(t)) \quad \text{s.t.} \quad \tilde{x} = f(x, \Theta_{\text{model}}; u^t; t), \quad x(0) = x_0
\]

\[
y = g(x, \Theta_{\text{model}}; u^t; t)
\]

\[
\sigma(t) = k \quad \text{if} \quad t \in T_k, \quad k \in \{1, \ldots, n_k\}
\]

\[
\theta^{\text{lb}} \leq \Theta_{\text{model}} \leq \theta^{\text{ub}}
\]

The computational time required for the SA depends on the sample size, the number of parameters, and the time for running a model. For the Morris and Sobol’ methods, the number of simulations required to compute the sensitivity indices are \( n_p \times N_t \times n_s \) and \((n_p + 2)N_t \times n_s \) respectively, which shows that the computational cost will increase linearly. Moreover, the computational cost of solving (23) depends on the time required for running a model, the number of the model parameters, the number of the temporal subdomains, and the number of different initial guesses to solve (23).

5 Application to NF-κB signalling

In this section, we applied the proposed methodology to model the NF-κB signalling dynamics in RAW murine macrophages induced by LPS in the presence of BFA.

5.1 NF-κB signalling pathway

NF-κB is an important regulator of inflammation and immune responses in various immune cells such as macrophages [43]. Under homeostatic conditions, the activity of NF-κB is minimal because it is sequestered by isomers of IκB (inhibitors of κB) proteins such as IκB-α, -β and -ε [43]. In the classical NF-κB activation pathway, an external stimulus (e.g. LPS) activates IκB kinase (IKK), which leads to degradation of IκB and thus activates NF-κB [43]. Then, the derepressed NF-κB protein translocates to the nucleus and upregulates the expression of various target genes such as IκB, and pro-inflammatory cytokines such as TNFα, which propagates the inflammatory signals to adjacent cells and tissues [44, 45].

As a component in gram-negative bacteria’s outer membranes, LPS is a potent activator of the NF-κB signalling pathway in macrophages through Toll-like receptor 4 (TLR4) [46]. By forming a complex with LPS, TLR4 and its accessory molecules activate NF-κB signalling through the classical activation pathway as described earlier. In contrast, BFA activates NF-κB through an alternate signalling pathway [47, 48]. Since exposure to BFA leads to the Golgi apparatus fusing with the endoplasmic reticulum (ER), normal intracellular trafficking is disrupted, which leads to the accumulation of proteins in the ER. This, in turn, initiates the ER-stress pathway and leads to the activation of NF-κB [21, 47, 49].

Although several mechanisms have been proposed to explain how NF-κB activity is induced by the ER-stress pathway, mechanistic details have not been fully elucidated yet due to the complexity of the ER-stress signalling pathway [50, 51]. Furthermore, recent studies demonstrated that interactions between the ER-stress and NF-κB signalling pathways are bidirectional, which further complicates the system analysis (see [48] and references therein). To unravel the complexity of the ER-stress signalling pathway, several computational models [50–52] have been proposed; however, they have not been validated thoroughly under various physiological conditions, whereas the NF-κB signalling pathway model has been continuously tested and improved since the early 2000s [53–55]. Furthermore, few studies have attempted to model the crosstalk between the ER-stress and NF-κB signalling pathways. Consequently, this study chose to use the time-varying model to represent the LPS-induced NF-κB signalling dynamics in the presence of BFA because the detailed model structure is still not known fully. The proposed model can be used to design future experiments that can help elucidate the underlying molecular interactions in future studies.

Motivated by the above considerations, we considered the LPS-induced NF-κB signalling model as the well-characterised model (1) while the model for the NF-κB signalling dynamics induced by LPS in the presence of BFA is considered as the unknown high-fidelity model, which would be approximated by the LPS-induced signalling model with time-varying parameters.

5.2 Dynamic model of LPS-induced NF-κB signalling

The schematic diagram for the NF-κB signalling pathway and the TNF-α production induced by LPS in the presence of BFA is shown in Fig. 1. The starting point of the model is the LPS-induced NF-κB signalling model developed by Hoffmann et al. [56, 58], where the LPS-NF-κB signalling pathway model was adapted from Caldwell et al. [58], and a model describing the regulation of the TNF-α production by internalised LPS-TLR4 complexes was adopted from Junkin et al. [56]. Lee et al. [21] further updated the model by incorporating a new role for A20 protein as an inhibitor of LPS-induced signalling. Also, the well-known effect of a BFA addition on the collapse of Golgi apparatus was taken into account by introducing time-dependent decays in rate constants associated.
with protein secretion and protein translocation to the membrane [21]. The model outputs are the dynamics of IkBa protein and intracellular TNF-α-protein (i.e. \( n_\alpha = 2 \)), and the updated model contains 49 states and 146 parameters (i.e. \( n_\alpha = 49 \) and \( n_\beta = 146 \), and see [21] for the details on the model).

The datasets obtained through flow cytometry in our previous study [21] were used to perform the temporal clustering as well as clustering the experimental datasets into three temporal subdomains, which are shown in Fig. 4. Specifically, the first, second and third subdomains contain the data points spanning from 0 to 60 min, 120 min, and 240 to 360 min, respectively. Each number of subdomains.

Since the value of \( N_t \) is eight, the maximum number of possible subdomains is eight in this work. Fig. 2 shows the changes in the intra- and inter-cluster error sums (\( \Lambda \) and \( \Gamma \), respectively) for all possible number of subdomains. As expected, the value of \( \Lambda \) decreases with the number of subdomains, while the value of \( \Gamma \) increases. Based on these two values, the cluster balance (\( \epsilon \)) defined in (9) can be computed for each number of subdomains and plotted in Fig. 3. As described earlier, a turning point in Fig. 3 is used to determine the optimal value of \( n_\sigma \), which is found to be three.

Based on \( n_\sigma = 3 \), each temporal subdomain can be determined by clustering the experimental datasets into three temporal subdomains, which are shown in Fig. 4. Specifically, the first, second and third subdomains contain the data points spanning from 0 to 60 min, 120 min, and 240 to 360 min, respectively. Each temporal subdomain can be interpreted to represent a different phase of the NF-κB signalling pathway induced by LPS in the presence of BFA. The first subdomain shows the early phase of the NF-κB signalling, where IkBa is quickly degraded while TNFα has not been synthesised. The second subdomain corresponds to the transition from the late phase of the LPS-induced NF-κB signalling pathway, where the rate of TNF-α synthesis accelerates and the IkBα is being re-synthesised, to the BFA-dominated signalling. The last subdomain can be seen as BFA-induced NF-κB dynamics, where the IkBα concentration is sustained at a low level due to the inhibition of its translation by the BFA [21].

5.4 SA result
The Morris and Sobol’ sensitivity methods were implemented as described above, and all the sensitivity computation was performed in parallel in the ADA supercomputing cluster at Texas A&M University. The result of the SA via the Morris method is shown in Table 1. For each parameter, six different values were randomly sampled from its parameter domain ranging from 10 to 1000% of its nominal value, and the average sensitivity of each parameter with respect to the two outputs was computed (13). Table 1 only lists the parameters whose sensitivity measures were at least 1% of that of the most important parameter. Interestingly enough, parameters whose normalised \( S_i \) values are at least 0.1 are the ones directly involved in the TNF-α dynamics such as synthesis rate concentrations, so the fold changes were computed based on (4). Therefore, the model output functions \( y = g(x, \theta_{opt}, u, t) \) are also defined as the fold change of the two states with respect to their initial conditions as follows:

\[
y_i(t_i) = \frac{I_kB_{tot}(t_i)}{I_kB_{tot}(0)}
\]

\[
y_j(t_i) = \frac{TNF_{tot}(t_i)}{TNF_{tot}(0)}
\]

\( y_i(t_i) \) and \( y_j(t_i) \) are the predicted fold changes of IkBa and intracellular TNFα concentrations, respectively, at the time \( t_i \) and TNFα and IkBa are the predicted IkBa and intracellular TNF-α concentrations, respectively, by the model.

5.3 Temporal clustering
In the flow cytometry experiments described in our previous study [21], the sampling time instants were 0, 10, 20, 30, 60, 120, 240, 360 min (i.e. \( N_t = 8 \)) after LPS and BFA were added to the cell culture. Two LPS concentrations (10 and 250 ng/ml) and one concentration of BFA (1 µg/ml) were used to obtain the experimental data sets (i.e. \( n_\sigma = 2 \)). Then, the temporal clustering methodology described in the preceding section was implemented to partition the measurement data sets to determine the optimal value of \( n_\sigma \) as well as the corresponding temporal subdomains.

\[
\text{Cluster balance value } \epsilon = \frac{\sum_{l=1}^{N_t} \left| I_kB_{tot}(t_l) \right| - \left| I_kB_{tot}(t_l) \right|}{\sum_{l=1}^{N_t} \left| I_kB_{tot}(t_l) \right|}
\]

\( \text{Cluster balance value } \epsilon \) is normalised by its maximum value.

The parameter estimation required in the proposed methodology. As discussed earlier, the datasets obtained through flow cytometry are relative data, which will not give the measurements in absolute number of subdomains (\( n_\sigma \)). The cluster balance value is normalised by its maximum value.

\[
\text{Error sum } \Lambda = \frac{\sum_{l=1}^{N_t} \left| I_kB_{tot}(t_l) \right| - \left| I_kB_{tot}(t_l) \right|}{\sum_{l=1}^{N_t} \left| I_kB_{tot}(t_l) \right|}
\]

\( \text{Error sum } \Lambda \) is normalised by its maximum value.
the Morris method, the values of these eleven parameters were measured under 10 and 250 ng/ml of LPS, respectively, in the presence of 1 μg/ml of BFA, and three different temporal subdomains are separated by blue dash lines.

Out of 22 parameters selected from the Morris method, the first 11 parameters were further analysed by the subsequent SA through the Sobol’ method [38]. Here, the parameters after the 11th parameter with respect to each output separately through the Sobol’ method were selected for the subsequent parameter estimation. This means that the parameters are highly important to at least one output and that the parameters are highly significant to the overall IκBα dynamics. Therefore, only one parameter, the Hill coefficient for TNF-α transcription, was selected for the parameter estimation as it has the highest sensitivity measures. In summary, five parameters were selected to vary with time, and their values in each temporal subdomain were determined in the following parameter estimation step (Table 3).

5.5 Parameter estimation

With the results from the temporal clustering and SA, the parameter estimation problem (23) was solved to obtain the values of these five parameters in each temporal subdomain (Table 3). Here, the model validation and the parameter estimation were performed via MATLAB built-in functions, ode15s and fmincon, and the multistart function available in MATLAB was used to solve the optimisation problem multiple times with different initial values.

Figs. 5 and 6 show the predicted dynamics of TNF-α and IκBα after the parameter estimation. In order to show the improvement of the prediction accuracy, the predicted dynamics after the parameter estimation were compared with the experimental measurements [21] and those predicted before the estimation. The prediction accuracy for the dynamics of the proteins was significantly improved. In particular, the model was able to track the TNF-α dynamics very accurately under both conditions (Fig. 5). Although there was some discrepancy between the model prediction and the experimental measurements for the IκBα dynamics under 10 ng/ml LPS, the overall prediction was improved.

In order to further validate the resulted model, the prediction accuracy of the resulted model was assessed with the experimental dataset, which was not used to train the model. In Fig. 7, the TNF-α and IκBα dynamics predicted by the model under 50 ng/ml LPS in the presence of BFA were plotted and compared with the corresponding experimental dataset. As shown in Fig. 7, the resulted model was able to accurately predict the TNF-α and IκBα dynamics.

Table 1 Result of Morris SA

| Rank | Parameter                                      | Normalised Sj |
|------|-----------------------------------------------|---------------|
| 1    | Hill coefficient for TNF-α transcription      | 1.00          |
| 2    | constant for TRIF-α-induced                   | 0.99          |
| 3    | TNF-α production enhancement (κα)             | 0.99          |
| 4    | constant for TRIF-α-induced                   | 0.99          |
| 5    | TNF-α production enhancement (κα)             | 0.99          |
| 6    | TNF-α protein degradation rate constant       | 0.99          |
| 7    | maximum degradation rate for TNF-α transcript | 0.99          |
| 8    | IKK-α-mediated degradation rate constant for IκBα | 0.99          |
| 9    | IKK-α-mediated IKK activation rate constant   | 0.99          |
| 10   | constitutive IKKK activation rate constant    | 0.99          |
| 11   | IκBα transcript degradation rate constant     | 0.99          |
| 12   | IκBα translation rate constant                | 0.99          |
| 13   | IκBα degradation rate in nucleus              | 0.99          |
| 14   | IκBα degradation rate in cytoplasm            | 0.99          |
| 15   | EC50 constant for TNF-α transcription         | 0.99          |
| 16   | constitutive IκBα degradation rate constant   | 0.99          |
| 17   | hill coefficient for IκBα transcription       | 0.99          |
| 18   | NF-αB-induced TNF-α transcription rate constant | 0.99          |
| 19   | constitutive IKKK deactivation rate constant  | 0.99          |
| 20   | rate constant for IκBα and NF-αB association  | 0.99          |
|      | in nucleus                                     | 0.99          |
| 21   | constitutive IKKK inactivation                | 0.99          |
| 22   | constitutive IκBα inactivation                | 0.99          |
by comparing their temporal dynamics under two different LPS not used in the model calibration, which demonstrated the robustness of the calibrated model and thus validates the proposed methodology. The result of the parameter estimation for the presented model is 1.75 while reproducing the Iκκ-β-mediated degradation rate constant for IkBα in NFκB-IκκB complexes, /min.

It should be noted that the development of the previous model went around 100 stimuli that can trigger the NFκB signalling pathway since the underlying signalling network structures. For example, it has been found that there are multiple stimuli for one signalling pathway, and the number of stimuli is likely to be higher for those with highly complex network structures. For example, it has been found that there are around 100 stimuli that can trigger the NFκB signalling pathway. Moreover, the dynamics of one signalling pathway induced by different stimuli can be very different since these stimuli activate the intracellular signalling pathway through different mechanisms. Again, with the NFκB signalling pathway as an example, TNFα receptor and TLR4, respectively, resulting in the distinctive signalling dynamics.

Therefore, the comprehensive characterisation of an intracellular signalling pathway is non-trivial since each stimulus of the signalling pathway has its own distinct activation mechanism and corresponding dynamics. Under this circumstance, a model-based approach can be implemented to facilitate the study. However, this model-based approach is often feasible only for a handful of well-characterised stimuli such as TNFα and LPS. For the NFκB signalling pathway since the underlying signalling mechanisms induced by these stimuli are relatively well studied. Motivated by the above considerations, the current study proposes a methodology to construct a data-driven mechanistic model for through the iterative implementation of experiments and modelling, which can be time-consuming. However, through the proposed approach, one can get a model with a reasonable prediction accuracy in a shorter amount of time.

Since intracellular signalling pathways regulate various cellular behaviours, their dynamics and outcomes bear great importance for studying and predicting the tissue-level responses in vivo. One important factor dictating the signalling pathways is different stimuli that initiate the pathways. As discussed in the paper, there can be multiple stimuli for one signalling pathway, and the number of stimuli is likely to be higher for those with highly complex network structures. For example, it has been found that there are around 100 stimuli that can trigger the NFκB signalling pathway.

Table 2: Result of SA by the Sobol’ method

| Rank | Parameter | ST_{ij} | SS_{ij} |
|------|------------|---------|---------|
| 1 | IKKγ-mediated degradation rate constant for IkBα in NFκB-IκκB complexes | 0.49 | 0.30 |
| 2 | IKKα-mediated IKK activation rate constant | 0.36 | 0.19 |
| 3 | constitutive IKK activation rate constant | 0.31 | 0.10 |
| 4 | IkBα transcript degradation rate constant | 0.30 | 0.15 |
| 5 | Kα | 0.075 | 0.02 |
| 6 | TNF-α protein degradation rate constant | 0.075 | 0.02 |
| 7 | maximum degradation rate constant for TNF-α transcript | 0.075 | 0.02 |
| 8 | TNF-α nascent mRNA processing rate constant | 0.075 | 0.02 |
| 9 | TNF-α protein synthesis rate constant | 0.075 | 0.02 |
| 10 | Hill coefficient for TNF-α transcription | 0.075 | 0.02 |
| 11 | Kα | 0.075 | 0.02 |

Table 3: Result of the parameter estimation

| Parameter | Parameter values in each temporal subdomain |
|-----------|---------------------------------------------|
| T1 | T2 | T3 |
| IKKγ-mediated degradation rate constant for IkBα in NFκB-IκκB complexes | 2.59 | 0.23 | 0.04 |
| IKKα-mediated IKK activation rate constant, /μM-min | 5200 | 52 | 4230 |
| constitutive IKK activation rate constant, /min | 5x10^{-6} | 1.3x10^{-7} | 4.9x10^{-6} |
| IkBα transcript degradation rate constant, /min | 0.33 | 0.18 | 0.12 |
| Hill coefficient for TNF-α transcription | 3.73 | 1.96 | 2.02 |

Fig. 5: Result of parameter estimation. The predicted dynamics of TNFa before (dash line) and after (solid line) the parameter estimation were compared with the experimental observations under (a) 10 ng/ml and, (b) 250 ng/ml of LPS in the presence of BFA.
those less-studied stimuli, whose corresponding signalling dynamics are less characterised. This is feasible since the mechanisms of the signalling pathway induced by different stimuli overlap with each other. For example, the NF-κB signalling pathway network induced by TNF-α and LPS will converge at the IKK level, which frees the NF-κB proteins from their inhibitors. Therefore, for a new or less-studied stimulus of an intracellular signalling pathway, its corresponding signalling dynamics can be described by modifying the nominal model into a time-varying one as discussed in the paper. Then, the constructed model can be used for the optimal experimental design to enhance our understanding.

It should be noted that the proposed methodology is a semi-data-driven approach, where the model construction is guided by both the available experimental data and the mechanistic model. Specifically, based on the experimental data, the temporal profiles of the model parameters are inferred to complement the model mismatch due to the use of a nominal model. As a result, the resultant model can provide relatively accurate predictions in spite of the incomplete knowledge of the underlying system. At the same time, the use of the mechanistic model allows the resultant model to be used in the detailed analysis of the underlying mechanisms, which is difficult to be performed through a data-driven model.

Additionally, the proposed time-varying model was able to robustly predict the dynamics of IκBα and TNF-α proteins, which are the core components in the NF-κB signalling pathway, under the various conditions although the detailed ER-stress signalling mechanisms were not incorporated into the model. Due to the accuracy and robustness of the model, it can be used in future studies to design optimal experiments to enhance our understandings on how BFA can activate the NF-κB signalling pathway.

Although the proposed methodology can be used to obtain a more accurate and predictive model as described above, it has the following limitations. First, the increase in the number of parameters to be estimated due to the temporal partitioning the parameters may exacerbate the unidentifiability issue in the model calibration. This can be a severe issue since a signalling pathway model is often over-parameterised while the available experimental measurements are limited. Second, the identified model may not reflect the true mechanisms associated with the less-studied stimulus. Specifically, the proposed method relies on the global SA to identify which parameters are time-varying, but it does not consider any biological significance while selecting the parameters. Therefore, the identified temporal profiles of the parameters may not have the biological relevance, which will constrain the process analysis based on the resultant model. It should be noted that this limitation can be mitigated by adding additional constraints into the minimisation problem (23) so that the resultant parameters retain their biological significance.

6 Conclusion

In this work, we presented a methodology for constructing a time-varying model for an intracellular signalling pathway when its reaction network is not fully known a priori. First, experimental data were clustered through the k-mean clustering algorithm to determine the temporal subdomains for the model parameters, where the parameters have different values in each temporal subdomain. Next, the global SA, which uses the Morris and Sobol’ methods in sequence, was carried out to identify the most important parameters with respect to the model outputs. And only
these parameters were determined to be time-varying while the remaining parameters were fixed at their nominal values. Finally, the least-squares problem was solved to estimate the values of five parameters in each temporal subdomain to construct an accurate time-varying model. The proposed methodology was implemented to model the NF-κB signalling pathway induced by LPS in the presence of BFA to predict the dynamics of IκBα and TNF-α proteins. The prediction accuracy of the resulted model was comparable to that of a more detailed model proposed by Lee et al. [21], which demonstrated the performance of the proposed methodology. In summary, the proposed methodology speeds up the overall model development process without losing the prediction accuracy by avoiding the time-consuming procedure of experimentation and literature survey for developing a high-fidelity model.

7 Acknowledgments

The authors gratefully acknowledge financial support from Artie McFerrin Department of Chemical Engineering and the Texas A&M Energy Institute and the supercomputer resources from the Texas A&M University High Performance Computer Center. Also, support from the Nesbitt Chair Endowment to A.J. and flow cytometry data provided by Dr. Yufang Ding (Texas A&M University) are gratefully acknowledged.

8 References

[1] Klipp, E., Liebermeister, W.: ‘Mathematical modeling of intracellular signaling pathways’, BMC Bioinformat., 2006, 7, p. S10
[2] Kitano, H.: ‘System biology: a brief overview’, Science, 2002, 295, pp. 1662–1664
[3] Singh, A., Jayaraman, A., Hahn, J.: ‘Modeling regulatory mechanisms in IL-6 signal transduction in hepatocytes’, Biotechnol. Bioeng., 2006, 95, pp. 850–862
[4] Moya, C., Huang, Z., Cheng, P., et al.: ‘Investigation of IL-6 and IL-10 signalling via mathematical modelling’, IET Syst. Biol., 2011, 5, pp. 15–26
[5] Hoffmann, A., Levchenko, A., Scott, M.L., et al.: ‘The IκB-NF-κB signalling module: temporal control and selective gene activation’, Science, 2002, 298, pp. 1241–1245
[6] Lipniacki, T., Paszek, P., Brasier, A.R., et al.: ‘Mathematical model of NF-κB regulatory module’, J. Theor. Biol., 2004, 228, pp. 195–215
[7] Swameye, I., Müller, T.G., Timmer, J., et al.: ‘Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by database modeling’, Proc. Natl. Acad. Sci., 2003, 100, pp. 1028–1033
[8] Schoeberl, B., Eichler-Jonsson, C., Gilles, E.D., et al.: ‘Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors’, Nat. Biotechnol., 2002, 20, pp. 370–377
[9] Pahl, H.L.: ‘Activators and target genes of Rel/NF-κB transcription factors’, Oncogene, 1999, 18, pp. 6853–6866
[10] Gadkar, K.G., Gunawan, R., Doyle, F.J.: ‘Iterative approach to model identification of biological networks’, BMC Bioinform., 2005, 6, p. 155
[11] Balsa-Canto, E., Alonso, A.A., Banga, J.R.: ‘An iterative identification procedure for dynamic modeling of biochemical networks’, BMC Syst. Biol., 2010, 4, p. 11
[12] Rodriguez-Fernandez, M., Rehberg, M., Kremling, A., et al.: ‘Simultaneous model discrimination and parameter estimation in dynamic models of cellular systems’, BMC Syst. Biol., 2013, 7, p. 76
[13] Penas, D.R., Henrikis, D., Gonzalez, P., et al.: ‘A parallel metaheuristic for large mixed-integer dynamic optimization problems, with applications in computational biology’, PLoS ONE, 2017, 12, e0182186
[14] Maurya, M.R., Bornheimer, S.J., Venkatasubramanian, V., et al.: ‘Mixed-integer nonlinear optimisation approach to coarse-graining biochemical networks’, IET Syst. Biol., 2008, 3, pp. 24–39
[15] Verheijen, P.J.T.: ‘Model selection: an overview of practices in chemical engineering’, Comput. Aided Chem. Eng., 2003, 16, pp. 85–104

Fig. 8 Comparison between the model developed in this study and the model developed by Lee et al. [21]. The predicted dynamics of IκBα were compared with the experimental observations under (a) 10 ng/ml and, (b) 250 ng/ml LPS concentration in the presence of BFA

Fig. 9 Comparison between the model developed in this study and the model developed by Lee et al. [21]. The predicted dynamics of TNF-α were compared with the experimental observations under (a) 10 ng/ml and, (b) 250 ng/ml LPS concentration in the presence of BFA
Pohjanpalo, H.: ‘System identifiability based on the power series expansion of Ljung, L., Glad, T.: ‘On global identifiability for arbitrary model Vajda, S., Godfrey, K.R., Rabitz, H.: ‘Similarity transformation approach to the solution’, Math. Comput. Simul., 2001, 55, pp. 271–280.

Chu, Y., Jayaraman, A., Hahn, J.: ‘Parameter sensitivity analysis of IL-6 signalling pathways’, IET Syst. Biol., 2007, 1, pp. 342–352.

Saltelli, A., Ratto, M., Tarantola, S., et al.: ‘Sensitivity analysis for chemical models’, Chem. Rev. 2005, 105, pp. 2831–2860.

Saltelli, A., Gatelli, D., Campolongo, F., et al.: ‘Global sensitivity analysis: the primer’ (John Wiley & Sons, 2008)

Homer, T., Sattler, A.: ‘Improving measures in global sensitivity analysis of nonlinear models’, Reliab. Eng. Syst. Saf., 1996, 52, pp. 1–17

Kontoravdi, C., Asprey, S.P., Pistikopoulos, E.N., et al.: ‘Application of global sensitivity analysis to determine goals for design of experiments: an example study on antibody-producing cell cultures’, Biotechnol. Prog., 2005, 21, pp. 1128–1135

Ghosh, G., Wang, Y.Y., Huang, D., et al.: ‘NF-κB regulation: lessons from structures’, Immunol. Rev., 2012, 246, pp. 36–58

Parameswaran, N., Patial, S.: ‘Tumor necrosis factor-α signaling in macrophages’, Crit. Rev. Eukaryot. Gene Expr., 2010, 20, pp. 87–103

Maiti, S., Dai, W., Alainz, R.C., et al.: ‘Mathematical modeling of pro- and anti-inflammatory signaling in macrophages’, Processes, 2015, 3, pp. 1–18

Kawai, T., Akira, S.: ‘The role of pattern-recognition receptors in innate immunity: update on toll-like receptors’, Nat. Immunol., 2010, 11, pp. 373–384

Pahl, H.L., Baeuerle, P.A.: ‘A novel signal transduction pathway from the endoplasmic reticulum to the nucleus is mediated by transcription factor NF-κB’, EMBO J., 1995, 14, pp. 2580–2588

Smith, J.A.: ‘Regulation of cytokine production by the unfolded protein response; implications for infection and autoimmunity’, Front. Immunol., 2018, 9, p. 422

Chardin, P., McCormick, F.: ‘Brefeldin A: the advantage of being uncompetitive’, Cell, 1999, 97, pp. 153–155

Erguler, K., Pieri, M., Deltas, C.: ‘A mathematical model of the unfolded protein stress response reveals new mechanism for recovery, adaptation and apoptosis’, BMC Syst. Biol., 2013, 7, p. 16

Diedrichs, D.R., Gomez, J.A., Huang, C., et al.: ‘A data entrained computational model for testing the regulatory logic of the vertebrate unfolded protein response’, Mol. Biol. Cell, 2018, 29, pp. 1502–1517

Cho, H., Wu, M., Zhang, L., et al.: ‘Signaling dynamics of palmitate-induced ER stress responses mediated by ATF4 in HepG2 cells’, BMC Syst. Biol., 2013, 7, p. 9

Lipniacki, T., Kimmel, M.: ‘Deterministic and stochastic models of NF-κB pathway’, Cardiovasc. Toxicol., 2007, 7, pp. 215–234

Cheong, R., Hoffmann, A., Lechպenko, A.: ‘Understanding NF-κB signaling via mathematical modeling’, Mol. Biol. Cell, 2008, 20, pp. 191–209

Williams, R.A., Timmis, J., Qwarnstrom, E.E.: ‘Computational models of the NF-κB signalling pathway’, Computation, 2014, 2, (46), pp. 131–158

Junkin, M., Kaestle, A.J., Cheng, Z., et al.: ‘High-content quantification of single-cell immune dynamics’, Cell. Rep., 2016, 15, pp. 411–422

Werner, S.L., Kearns, J.D., Zadorozhnaya, V., et al.: ‘Encoding NF-κB temporal control in response to TNF: distinct roles for the negative regulators IκBα and A20’, Genes Dev., 2008, 22, pp. 2093–2101

Caldwell, A.B., Cheng, Z., Vargas, J.D., et al.: ‘Network dynamics determine the autocrine and paracrine signaling fuctions of TNF’, Genes Dev., 2014, 28, pp. 2120–2133