10.1515/acsc-2017-0015

Archives of Control Sciences
Volume 27(LXIII), 2017
No. 2, *pages 239–250*

Sensitivity analysis of signaling pathway models based on discrete-time measurements

MALGORZATA KARDYNSKA and JAROSLAW SMIEJA

The paper is focused on sensitivity analysis of large-scale models of biological systems that describe dynamics of the so called signaling pathways. These systems are continuous in time but their models are based on discrete-time measurements. Therefore, if sensitivity analysis is used as a tool supporting model development and evaluation of its quality, it should take this fact into account. Such models are usually very complex and include many parameters difficult to estimate in an experimental way. Changes of many of those parameters have little effect on model dynamics, and therefore they are called sloppy. In contrast, other parameters, when changed, lead to substantial changes in model responses and these are called stiff parameters. While this is a well-known fact, and there are methods to discern sloppy parameters from the stiff ones, they have not been utilized, so far, to create parameter rankings and quantify the influence of single parameter changes on system time responses. These single parameter changes are particularly important in analysis of signalling pathways, because they may pinpoint parameters, associated with the processes to be targeted at the molecular level in laboratory experiments. In the paper we present a new, original method of creating parameter rankings, based on an Hessian of a cost function which describes the fit of the model to a discrete experimental data. Its application is explained with simple dynamical systems, representing two typical dynamics exhibited by the signaling pathways.

Key words: sensitivity analysis, signaling pathways, measurement uncertainty, discrete-time measurements.

1. Introduction

Signaling pathways (or regulatory pathways) are cascades of biochemical processes involving creation, degradation and modification of various molecules, specific for a given pathway, as well as their transport between cellular compartments such as cytoplasm, nucleus, mitochondrium, etc. These processes are activated by events taking place

The Authors are with Institute of Automatic Control, Silesian University of Technology, Akademicka str. 16, 44-100 Gliwice, Poland. E-mails: {malgorzata.kardynska, jaroslaw.smieja}@polsl.pl. M. Kardynska is the corresponding author.

The work has been supported by the NCN grant DEC-2013/11/B/ST7/01713 (JS, MK). Calculations were performed on the Ziemowit computational cluster (http://www.ziemowit.hpc.polsl.pl) created in the POIG.02.01.00-00-166/08 project (BIO-FARMA) and expanded in the POIG.02.03.01-00-040/13 project (Syscancer).

Received 11.12.2016. Revised 30.04.2017.

inside a cell (such as, e.g., DNA damage), changes in extracellular environment (e.g. in its chemical content or temperature), direct interactions with other cells (following their binding) or physical stresses (radiation, mechanical stress). They are regulated by positive and negative feedback loops that in many cases are not fully understood. Therefore, use of methods that have their origins in automatic control may help in learning these mechanisms, through formulating hypotheses about the structure of regulatory networks governing intracellular processes. Ultimately, such knowledge will support development of the protocols of external regulation of cell behavior.

There are many methods that can be used to describe the dynamics of signaling pathways. In this paper we are focused on deterministic models described by ordinary differential equations, in which variables denote concentrations of molecules involved in a given pathway. Each parameter (or, in case of Michaelis-Menten kinetics, a pair of parameters) correspond to a single biochemical process.

Due to rapid advances in experimental techniques, our knowledge about biochemical processes occurring in living cells is continuously expanding. In the literature there is a growing number of highly dimensional models, which describe the dynamics of signaling pathways components [1]. The more processes are taken into account, the more complex models arise with a large number of parameters. However, methods of measuring biochemical parameters are limited and often inaccurate [2]. Therefore, any mathematical model should be checked with respect to its sensitivity to parameter changes. In general, such model should be robust with respect to small parameter changes. Nevertheless, some parameters are always more important than others. The sensitivity analysis is the tool to be used to determine how a change of parameters influences the system behavior. It provides information about the most important parameters that have the greatest impact on the system output [3]. In the particular application that is considered in this paper, each parameter is associated with a particular biochemical process. Hence, sensitivity analysis may provide insights into how biological experiments should be planned to gain maximum information. Moreover, parameters with the highest ranking indicate prospective molecular drug targets affecting a pathway that is involved in a given disease.

Sensitivity analysis have been developed for over half a century, initially for applications in engineering [4, 5, 6]. While sensitivity methods proved to be helpful in analysis of various pathways [7, 8, 9, 10, 11], they were focused on simulation results whose units were clearly determined (as concentration units). However, in most cases biological experiments provide data about the fold increase of the number of molecules or of the concentration, while their absolute values are not known. In that case the same methods of sensitivity analysis may lead to false conclusions [12]. Furthermore, measurements are discrete in time with irregular and sparse sampling periods. Even if live microscopy is used, data is collected every couple of minutes. For these reasons it is necessary to develop methods which take into account specific properties of biological systems and experimental data.

In this paper we present a new method for creating parameter rankings. It is based on the Hessian of a cost function describing the fit of the model to discrete measurement data [13]. The rankings allow not only to identify the most important parameters of the

model, which is facilitated by the methods developed in earlier works [14], but allows to quantify and rank the importance of single parameters. These parameters should be determined with the highest accuracy when developing a biologically relevant model and may provide hints to indicate prospective molecular targets for new drugs [15]. In addition, the proposed method takes into account the measurement uncertainty at specific sampling times, which is also an improvement over standard sensitivity analysis methods. Most of these methods assume that the model perfectly describes the biological process, which is not true - mathematical models are built on the basis of experimental measurements that are subject to uncertainty and in the case of biological experiments may be very high.

In the following section the concept of stiff and sloppy parameters and the method to find them are introduced, followed by the definition of the proposed ranking. Then, the rankings obtained with the proposed method are shown and compared with those given by standard sensitivity analysis for two models, each representing particular dynamics exhibited by signaling pathways. Both examples show that the presented method allows to create reliable parameter rankings that helps to identify parameters substantially affecting the fit of the model to experimental data.

2. Stiff and sloppy parameters

Sensitivity analysis is usually divided into two major categories: local and global. Local sensitivity analysis describes how the system output changes when parameters deviate in a close neighborhood of their nominal values. Global sensitivity, in turn, describes how the system output changes when multiple parameters are allowed to change in a relatively wide range [16, 17]. The method presented in this work may be classified as a local one.

Let the model be described by the state equation:

$$\frac{d\mathbf{X}}{dt} = F(\mathbf{X}, \mathbf{U}, \theta), \quad \mathbf{X}(t_0) = \mathbf{X}_0, \tag{1}$$

where $\mathbf{X}$ is a vector of state variables, representing concentrations or the average number of molecules of proteins, enzymes or transcripts involved in the signaling pathway, $\mathbf{U}$ denotes control vector, which is usually a scalar in biological systems, θ is a vector of parameters. A solution of the described model is defined by:

$$\mathbf{X}(\theta, t). \tag{2}$$

Model quality can be evaluated using the least-squares cost function [13] that describes the difference between the variable values obtained from simulation and quantified experimental data:

$$C_s(\theta) = \sum_n \frac{1}{2} \frac{(x_s(\theta, t) - x_{s,n})^2}{\sigma_{s,n}^2} = \sum_n \frac{1}{2} r_{s,n}^2, \tag{3}$$

where $x_{s,n}$ is the value of s-th state variable in the n-th sample measured with the uncertainty $\sigma_{s,n}$, $x_s(\theta,t)$ is a solution of the model at corresponding time t, while $r_{s,i}$ is the residual describing the deviation of a dynamical variable $x_s(\theta,t)$ from its measured values. If the model perfectly fits to measurement data the cost function C_s is equal to 0, and the vector of parameters θ, giving a perfect fit, is denoted as θ^*.

To analyze model sensitivity to parameter variation, let us consider the Hessian matrix corresponding to the cost function C_s calculated at θ^*. Since the value of one biochemical parameter may vary in the range of several orders or more from another, to eliminate the impact of relative changes in parameter values the derivatives with respect to $\log\theta$ are taken [14]:

$$H^{C_s}_{j,k} = \frac{\partial^2 C_s}{\partial \log\theta_j \partial \log\theta_k}, \tag{4}$$

where j and k denotes $j-$th and k-th parameter, respectively.

Instead of (4), the Hessian approximation H^{C_s} can be used - the so called Fisher information matrix $J^T J$ [18]:

$$J = \frac{\partial r_{s,n}}{\partial \log\theta}. \tag{5}$$

The Hessian matrix H^{C_s} is positive definite and symmetric, so it has real eigenvalues λ and eigenvectors v [19]. It describes the surface of deviations of the model from measured data. For a model with N_p parameters the surface is an N_p-dimensional ellipsoid in parameter space. The principal axes of the ellipsoids are the eigenvectors of H^{C_s}, while the width (denoted by d_i) of the ellipsoids along each principal axis is given by:

$$d_i = \frac{1}{\sqrt{\lambda_i}}. \tag{6}$$

The narrowest axes are called *stiff* and they define directions in the parameter space, leading to large changes in the model response. The broadest axes, called *sloppy*, represent the directions along which parameters changes, even in a wide range, do not result in a worse fitting of the model to experimental data [20]. The meaning of eigenvalues and eigenvectors of Hessian H^{C_s} is illustrated with a simple example of a hypothetical model with two parameters: θ_1 and θ_2 (Fig. 1), where Hessian describes an ellipse in the θ_1/θ_2 parameter space, d_1 and d_2 denote the width of the ellipse along each principal axis, corresponding to eigenvalues λ_1 and λ_2, respectively, while v_1 and v_2 denote the eigenvectors of H^{C_s} and define the position of the ellipse. In this example, the eigenvectors v_1 and v_2 are associated with a large value d_1 and small value d_2, respectively. Since v_2 eigenvector depends mostly on θ_1, this parameter is *stiff*, i.e. the change in its value leads to much greater change in system response than in the case θ_2 was changed.

It should be stressed, however, that models of signaling pathways are much more complex and often include dozens of parameters. Therefore graphical presentation of results illustrating deviations of the model from measured data is not possible. In this

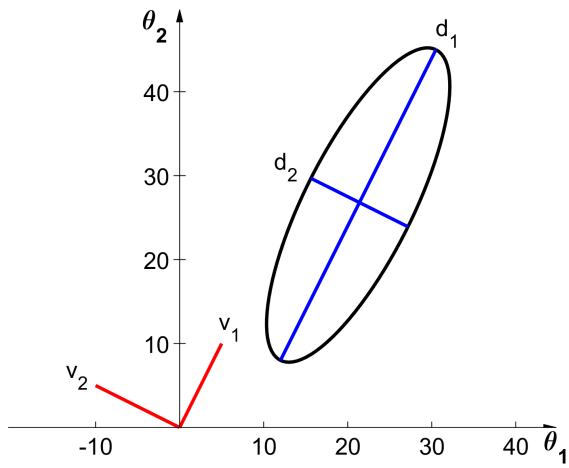


Figure 1: An ellipse illustrating deviations of the model from measured data in a 2-dimensional parameter space.

case, plots showing parameter rankings are the preferred way of presenting the results. The method of creating parameter rankings, proposed in this paper, is based on eigenvectors and eigenvalues of the Hessian H^{C_s}. R_j, denoting the ranking value for the j-th parameter, can be defined as:

$$R_j = \sum_i \left| \frac{v_{j,i}}{d_i} \right|, \tag{7}$$

where d_i is the width of the ellipsoid along i-th principal axis, and $v_{j,i}$ is the element of the i-th eigenvector corresponding to the j-th parameter.

3. Ranking examples

To show applicability of the proposed method, two examples are presented in this section. Both represent typical dynamics exhibited by signaling pathways - with and without oscillations. For each model its step response have been simulated and the measurements have been sampled for arbitrarily chosen time points, shown in the figures, with 10% uncertainty.

The results obtained are compared to standard rankings based on the area under the curve (AUC) of sensitivity functions with L^1 norm [21] as a metric. While other methods can be found in the literature [22], sensitivity functions constitute the most often used

base for the rankings in local analysis of signaling pathways [7, 8, 10, 11, 23], unless it is qualitative behavior of the pathway that is under consideration [24].

3.1. Transcription-translation pathway

As the first example, let us consider a simple pathway, in which gene transcription is activated, leading to production of mRNA and, subsequently, protein, whose concentrations are denoted by x_m and k_p, respectively. These molecules are degraded in a first-order process. In the simplest case such system is described by linear state equations:

$$\frac{dx_m}{dt} = k_m u - k_{dm} x_m, \tag{8}$$

$$\frac{dx_p}{dt} = k_p x_m - k_{dp} x_p, \tag{9}$$

where u represents the system input (induction of transcription), k_m, k_p, k_{dm} and k_{dp} are mRNA and protein production and degradation rates, respectively.

Let us also assume that only the protein levels are observed in the experiment, i.e. x_p is the output variable.

Then, the system may be alternatively represented by the transfer function

$$K(s) = \frac{X(s)}{U(s)} = \frac{k}{(1+sT_1)(1+sT_2)}, \tag{10}$$

whose parameters have been arbitrarily chosen as $T_1 = 1/k_{dm} = 0.1$, $T_2 = 1/k_{dp} = 1$ and $k = k_p k_m/(k_{dm} k_{dp}) = 1$ (in this system only three parameters are identifiable).

The model step response with sampled measurements is shown in Fig. 2.

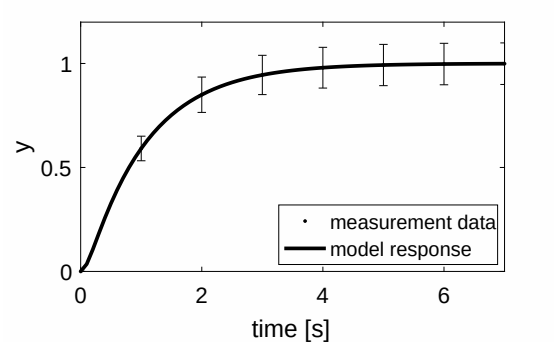
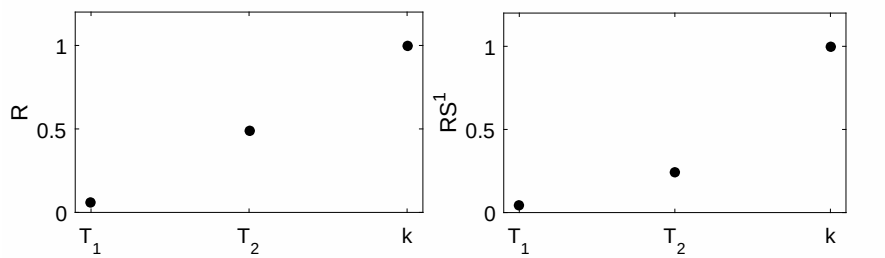


Figure 2: The step response of a second-order inertial system with parameters $T_1 = 0.1$, $T_2 = 1$ and $k = 1$ and the standard deviation of measurement data.

Parameter rankings, obtained with the procedure described in the previous section are shown in Fig. 3a.



a *b*

Figure 3: Parameter rankings for the second-order system based on: (a) the proposed method and (b) sensitivity functions.

The order of parameters importance is the same for both rankings. However, in the ranking based on sensitivity functions, the change in system response caused by change of T_2 value is similar to the one caused by T_1 change. The ranking created with the method proposed in this paper assigns much greater importance to T_2, which is a greater time constant and determines transient system response. These differences result from taking into account the measurement uncertainty in the cost function C_s (3). As mentioned before, uncertainty of measurement, defined as the standard deviation σ, has been assumed to be 10% of measurement value. Fig. 2 shows that the absolute standard deviation of measurements made at time-points: 1s and 2s (before reaching the steady state) is smaller than the standard deviation of measurements in steady state. As a result, deviation of the model response at points with a lower σ will be treated as more important than deviation at points with a higher σ. This explains the higher position of parameter T_2 in our ranking, compared to the parameter k, which is responsible for the steady state system response and according to Fig. 2 cannot be precisely determined due to measurement uncertainty.

3.2. A closed-loop regulatory module

As a second example a model of p53/Mdm2 regulatory module has been considered. It is one of the simplest oscillatory systems that can be fitted to experimental data, described by the following equations [25]:

$$\frac{d(p53)}{dt} = ms_1 - k_{d1} \cdot (p53) \cdot (Mdm2_{nuc})^2, \tag{11}$$

$$\frac{d(Mdm2_{cyt})}{dt} = n \cdot \left(s_2 + \frac{s_3 \cdot (p53)^3}{s_4^3 + (p53)^3} \right) - k_1 k_2 \cdot \frac{MdM2_{cyt}}{k_2 + (p53)}, \tag{12}$$

$$\frac{d(Mdm2_{nuc})}{dt} = \frac{k_1 k_2 \cdot (Mdm2_{cyt})}{k_2 + (p53)} - k_{d2} \cdot (Mdm2_{nuc}), \tag{13}$$

where the variables $p53$, $Mdm2_{cyt}$ and $Mdm2_{nuc}$ denote concentrations of total p53 protein, cytoplasmic Mdm2 and nuclear Mdm2, respectively. It is a minimal model reflect-

ing oscillatory response of p53 protein to system excitation, e.g., DNA damage. The parameters m and n are the numbers of p53 and Mdm2 gene copies, respectively, s_1, s_2, s_3 and s_4 are the production rates per gene copy, k_{d1} and k_{d2} are p53 and Mdm2 degradation rates and k_1, k_2 are Mdm2-mediated nuclear import rates. Nominal parameter values are given in Table 1.

Table 1: p53/Mdm2 model parameters.

Parameter	s_1	s_2	s_3	s_4	k_{d1}
Value	16	8	80	$1 \cdot 10^5$	$1 \cdot 10^{-13}$
Unit	s^{-1}	s^{-1}	s^{-1}	–	s^{-1}
Parameter	k_{d2}	k_1	k_2	m	n
Value	$2.2 \cdot 10^{-4}$	$3.5 \cdot 10^{-3}$	$2.3 \cdot 10^3$	2	2
Unit	s^{-1}	s^{-1}	–	–	–

Let the system output be the $p53$ concentration, measured, e.g., with an experimental technique called *Western Blotting*. A characteristic feature of biological measurements, including the one mentioned in the preceding sentence, is their large uncertainty, affected by many different factors. As a result, the standard deviation of measurements may vary significantly, as shown in Figure 4 (measurement points are marked for illustration only - their values have been calculated in simulation, with the random error superimposed on the results).

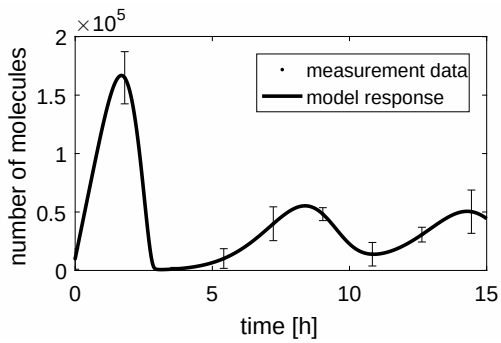


Figure 4: p53 model responses against the measurement data.

As in the previous example, two different sensitivity rankings have been calculated. The results are shown in Fig. 5.

Methods used for rankings calculation produced different results. In the case of signaling pathways, it may be more difficult to determine the impact of individual param-

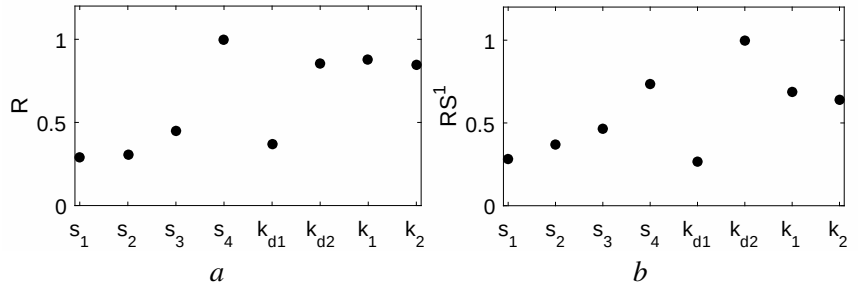


Figure 5: Parameter rankings for p53/Mdm2 model based on: our method (a) and sensitivity functions (b).

eters since they may affect many characteristics of the system response. For example, both parameter rankings indicate the high position of parameter k_{d2}, which affects the amplitude of $p53$ oscillations and introduces a small phase shift. In the parameter ranking based on sensitivity functions the parameter k_{d2} is indicated as the most important, while in the ranking taking into account the measurement uncertainty its position is slightly lower. On the other hand, the parameter s_4 has been found to be more relevant according to our method. In order to determine which of these two parameters has more significance on the fit of the model to experimental data, we performed simulations with parameters k_{d2} and s_4 changed by 15%. On Fig. 6 we compared received time courses with the model response for the nominal parameter set and measurement data. When we analyze Fig. 6 we find that despite the significant change in the model response caused by change of the parameter k_{d2}, the model may still quite good fit to the experimental data due to measurements uncertainty. Therefore, in the ranking taking into account the measurement uncertainty the position of parameter k_{d2} is lower. Similarly, we can explain the differences for parameter s_4, which also significantly changes the model response, however, the model response after changing the parameter s_4 by 15% slightly worse fits to the experimental data.

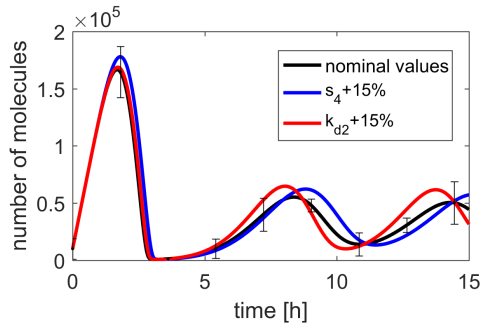


Figure 6: A comparison of p53 model responses after the change of selected parameters.

4. Conclusion

The new method for creating parameter rankings based on the known method called *sloppy / stiff* sensitivity analysis has been proposed. It facilitates taking into account the impact of measurement uncertainty, which is a major problem in analysis of any kind of biological experimental data.

Two simple examples have been used to show that the presented method allows to create reliable parameter rankings that helps to identify parameters substantially affecting the fit of the model to experimental data. This allows to choose parameters that should be determined with the highest accuracy in the experimental research and, as a consequence, the method can be used to plan biological experiments and helps to use the funds for experimental research in the most efficient way.

References

[1] R. WILLIAMS, J. TIMMIS and E.E. QWARNSTROM: Computational models of the NF-kB signalling pathway. *Computation*, **2**(4), (2014), 131-158.

[2] S.J. MAERKL and S.R. QUAKE: A systems approach to measuring the binding energy landscapes of transcription factors. *Science*, **315** (2007), 233-237.

[3] D.A. RAND: Mapping the global sensitivity of cellular network dynamics. *J. Royal Society Interface*, **5** (2008), 59-69.

[4] R.W. NEWCOMB: Linear Multiport Synthesis. McGraw-Hill, NY, 1966.

[5] J.J. CRUZ: Feedback Systems. McGraw-Hill, New York, 1972.

[6] R.K. BRAYTON and R. SPENCE: Sensitivity and Optimization. Elsevier, Amsterdam, 1980.

[7] M. BENTELE, I. LAVRIK, M. ULRICH, S. STOSSER, D.W. HEERMANN, H. KALTHO, P.H. KRAMMER and R. EILS: Mathematical modeling reveals threshold mechanism in cd95-induced apoptosis. *The Journal of Cell Biology*, **166**(6), (2004), 839-851.

[8] H. YUE, M. BROWN, J. KNOWLES, H. WANG, D.S. BROOMHEAD and D.B. KELL: Insights into the behaviour of systems biology models from dynamic sensitivity and identifiability analysis: a case study of an NF-κ B signalling pathway. *Molecular BioSystems*, **2**(12), (2006), 640-649.

[9] E. BALSA-CANTO, A.A. ALONSO and J.R. BANGA: An iterative identification procedure for dynamic modeling of biochemical networks. *BMC Systems Biology*, **4**(11), (2010).

[10] V.RAIA, M. SCHILLING and M. BOEHM: Dynamic mathematical modeling of IL13-induced signaling in Hodgkin and Primary Mediastinal B-Cell Lymphoma Allows Prediction of Therapeutic Targets. *Cancer Research*, **71** (2011), 693-704.

[11] K. RATEITSCHAK, F. WINTER, F. LANGE, R. JASTER and O. WOLKENHAUER: Parameter identifiability and sensitivity analysis predict targets for enhancement of STAT1 activity in pancreatic cancer and stellate cells. *PLoS Computational Biology*, **8**(12), (2012), e1002815.

[12] J. SMIEJA, M. KARDYNSKA and A. JAMROZ: The meaning of sensitivity functions in signaling pathways analysis. *Discrete and Continuous Dynamical Systems – series B*, **10** (2014), 2697-2707.

[13] B.C. DANIELS, Y.J. CHEN, J.P. SETHNA, R.N. GUTENKUNST and C.R. MYERS: Sloppiness, robustness, and evolvability in systems biology. *Current Opinion in Biotechnology*, **19** (2008), 389-395.

[14] R.N. GUTENKUNST, J.J. WATEFALL, F.P. CASEY, K.S. BROWN, C.R. MYERS and J.P. SETHNA: Universally sloppy parameter sensitivities in systems biology models. *PLoS Computational Biology*, **3**(10), (2007), 1871-1878.

[15] A. MARIN-SANGUINO, S.K. GUPTA, E.O. VOIT and J. VERA: Biochemical pathway modeling tools for drug target detection in cancer and other complex diseases. *Methods in Enzymology*, **487** (2011), 319-369.

[16] J. LEIS and M. KRAMER: Sensitivity analysis of systems of differential and algebraic equations. *Computers & Chemical Engineering*, **9** (1985), 93-96.

[17] P. IGLESIAS and B. INGALLS (Ed): Control theory and systems biology. MIT Press, 2010.

[18] J.J. WATERFALL, F.P. CASEY, R.N. GUTENKUNST, K.S. BROWN, C.R. MYERS, P.W. BROUWER, V. ELSER and J.P. SETHNA: Sloppy-model universality class and the Vandermonde matrix. *Physical Review Letters*, **97**(15), (2006), 150601.

[19] R.A. HORN and C.R. JOHNSON: Matrix Analysis. Cambridge University Press, 1990.

[20] K.S. BROWN and J.P. SETHNA: Statistical mechanical approaches to models with many poorly known parameters. *Physical Review E*, **68** (2003), 021904.

[21] M. KARDYNSKA and J. SMIEJA: L-1 and L-2 norms in sensitivity analysis of signaling pathway models. *21st Int. Conf. on Methods and Models in Automation and Robotics*, Miedzyzdroje, Poland, (2016), 589-594.

[22] K. PUSZYNSKI, P. LACHOR, M. KARDYNSKA and J. SMIEJA: Sensitivity analysis of deterministic signaling pathways models. *Bulletin of the Polish Academy of Sciences Technical Sciences*, **60**(3), (2012), 471-479.

[23] K. A. KIM, S. L. SPENCER, J. G. ALBECK, J. M. BURKE, P. K. SORGER, S. GAUDET and H. KIM DO: Systematic calibration of a cell signaling network model. *BMC Bioinformatics*, **11** (2010), 202.

[24] M. KARDYNSKA and J. SMIEJA: Sensitivity analysis of signaling pathways in the frequency domain. *Information Technologies in Medicine*, **2** (2016), 275-285.

[25] B. HAT, K. PUSZYNSKI and T. LIPNIACKI: Exploring mechanisms of oscillations in p53 and nuclear factor-κB systems. *IET Systems Biology*, **3** (2009), 342-355.