

Sensitivity Analysis of Stochastic Models of Bistable Biochemical Reactions

Andrea Degasperi¹ and Stephen Gilmore²

¹ Department of Computing Science, University of Glasgow

² Laboratory for Foundations of Computer Science, University of Edinburgh

Abstract. Sensitivity Analysis (SA) provides techniques which can be used to identify the parameters which have the greatest influence on the results obtained from a model. Classical SA methods apply to deterministic simulations of ODE models. We extend these to stochastic simulations and consider the analysis of models with bifurcation points and bistable behaviour. We consider local, global and screening SA methods applied to multiple runs of Gillespie's Stochastic Simulation Algorithm (SSA). We present an example of stochastic sensitivity analysis of a real pathway, the MAPK signalling pathway.

1 Introduction

Reaction-based biochemical models use input parameters such as concentrations and kinetic rate constants to predict the time evolution of a biochemical system. The chemical species involved in the reactions have the role of the output variables of the model. Fig. 1 shows an example with four species.

Sensitivity Analysis (SA) studies the relationships between the inputs and the outputs of models. When we wish to perform SA we choose a time point at which to read the output values. In the case of an ODE model, a selected output (species) has a precise value at a given time. Changing one or more parameters of the model may alter this. In the case of stochastic simulation [1] the output of a selected species at a selected time can be considered to be the collection of the values given by the individual simulation runs. If it is sufficiently large, this set of values will reveal the distribution of the output.

One of the basic SA operations is to compute the difference between the output of a model and the output of the same model with one or more parameters perturbed. This is simple to do with ODE models but not so straightforward when facing stochastic simulation. One simple approach is to take as output the mean of the values coming from the simulations. However, this can lead to a loss of information: by taking the mean we are assuming a normal distribution and we are even neglecting the variance. Another possibility is the use of a distribution distance or *histogram distance* which, with sufficient simulation runs, is able to precisely describe the difference. In [2] this is used to quantify how well an approximate SSA emulates the exact SSA. We use it here with SA to quantify the effect of perturbation of the parameters of a stochastic model.

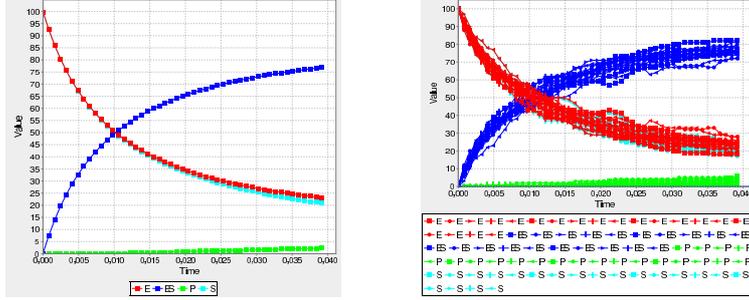


Fig. 1. Examples of time evolution of a biochemical model computed with ODE (on the left) and with SSA (on the right)

Histogram distance is computed as follows:

$$D_k(X, Y) = \sum_{i=1}^k \left| \frac{\sum_{j=1}^{|X|} \chi(x_j, I_i)}{|X|} - \frac{\sum_{j=1}^{|Y|} \chi(y_j, I_i)}{|Y|} \right|$$

where X and Y are two sets of numbers, k is the number of histogram columns or intervals which divide the range of the output variable, $|X|$ is the cardinality of the set X (resp. $|Y|$ is the cardinality of the set Y), x_j and y_j are elements of the sets X and Y respectively and the function χ returns 1 if the element x_j belongs to the interval I_i , 0 otherwise. I_i is the i -th interval in the range, which runs from $x_{min} + \frac{(i-1)L}{k}$ to $x_{min} + \frac{iL}{k}$, where $L = x_{max} - x_{min}$.

An interesting measure is then the *self distance*, given by $D_k(X, X')$. This runs the same experiment twice, with the same parameters, and then computes the histogram distance between the results. Perturbations in the parameters which generate values of distances less than or very close to the self distance will be considered not to have an influence, or, at least, we can say that we cannot distinguish any effect arising from this perturbation.

2 Sensitivity Analysis Classifications

According to [3], sensitivity analysis (SA) techniques can be classified as follows.

Local Methods: These concentrate the analysis around a particular point in the parameter space. For example, *local one at a time* and *elementary one at a time* approaches belong to this class.

Screening Methods: These are used to select the most important parameters when the complexity of the model is problematic or the number of parameters intractable. The main idea of these methods is that they should be computationally inexpensive and give the idea of which parameters can be fixed (low importance), even if the information that can be achieved is poor.

They are a tradeoff between information and algorithm complexity. Once the most influential parameters have been identified, it is then possible to apply a more informative and computationally expensive technique.

Global Methods: These techniques try to explore the entire space of the parameters or, at least, explore the subspace that is believed to contain the real value of the parameters and that represents their uncertainty. Usually these are the most computationally expensive, but also the most informative.

2.1 One-At-a-Time Methods

The classical and most widely used SA is the *one-at-a-time* (OAT) approach: a parameter is perturbed (usually by 1%) and the changes in the output measured. Alternatively it is possible to compute the derivative of the output with respect to each parameter to obtain its sensitivity coefficient:

$$S_{ij} = \frac{\delta y_j(\mathbf{p})}{\delta p_i}$$

where $y_j(\mathbf{p})$ is the j -th output of the model which depends on the parameters and p_i is the i -th parameter.

In the study of biochemical systems, OAT methods represent the prevalent practice when analysing ODE models. Other more complex and informative analysis has been proposed [4]. However, none of these are directly applicable to stochastic models whose output is defined as a *probability density function* (pdf) over the number of molecules for each species. The need to consider the entire pdf is very clear in the analysis of bistable systems. These present at a certain time a pdf which is not normal, but instead presents two distinct peaks of likelihood. In this particular context an analysis cannot make any assumptions about the pdf resulting from the model. For this reason SA of stochastic systems is an engaging research question [5] and here we are using histogram distance to quantify the change in the output value:

$$S_i = D(X_n, X_{p_i})$$

where X_n is a *random variable* (r.v.) with nominal pdf = $f(\mathbf{x}, \mathbf{p})$ and X_{p_i} is a r.v. with perturbed pdf = $f(\mathbf{x}, p_1, \dots, p_i + \Delta p_i, \dots, p_k)$. This distance can instead be divided by Δp_i , leading to a correspondent derivative-based approach.

Together, these approaches can be classified as *local one-at-a-time* (LOAT) SA and are applicable if we assume that varying one parameter at a time affects the output of the model in a proportional way. However, that assumption is often not valid for biological systems making LOAT incapable of giving a complete view of the relationships between parameters and output and also between the parameters themselves. LOAT methods are useful mainly because they can give a first impression of sensitivity indices and because they are computationally inexpensive – an important consideration when dealing with thousands of stochastic simulations.

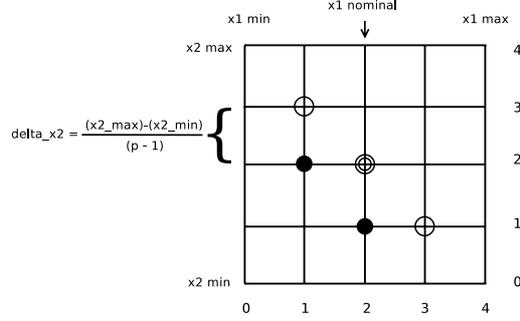


Fig. 2. Example of a grid in the Morris method. In this case we have two parameters ($k = 2$) and a grid level of five ($p = 5$), so the maximum possible combinations are $5^2 = 25$. The black dots are two possible random points, while the circles are other points computed during the algorithm iterations. An efficient implementation should not recompute the point circled twice.

2.2 Morris' Method

Morris' method [3] can be classified as one at a time, because it uses as a basic step the local OAT approach, and global, because the experiment covers the entire space over which the factors are believed to vary. Morris estimates the main effect of a factor by computing a number r of local measures, at different random points $\mathbf{x}_1, \dots, \mathbf{x}_r$ in the parameter space, and then taking their average.

When applying this method, a computationally expensive model is assumed, or a model with a large number of factors. The goal is to determine which factors have (a) negligible effects, (b) linear and additive effects, or (c) non-linear interaction effects. This will help to apply later the most appropriate global sensitivity analysis only on the relevant parameters.

The k -dimensional factor vector \mathbf{x} has components x_i that have p values in the set $\{0, 1/(p-1), 2/(p-1), \dots, 1\}$. The region of experimentation Ω is then a k -dimensional p -level *grid* (Fig. 2). In practice, the values sampled in Ω are then rescaled to generate the actual values of the parameters as sampled from a specific parameter range. Let Δ be a predetermined multiple of $1/(p-1)$. Then Morris defines the *elementary effect* of the i th factor at a given point \mathbf{x} as:

$$d_i(\mathbf{x}) = \frac{y(x_1, \dots, x_i + \Delta, \dots, x_k) - y(\mathbf{x})}{\Delta}$$

where \mathbf{x} is any value in Ω selected *such that the perturbed point $\mathbf{x} + \Delta$ is still in Ω* . After sampling r times, the result will be a distribution F_i of elementary effects. The characterisation of this distribution through its mean μ and standard deviation σ gives useful information about the influence of the i th input on the output.

$x_2 \backslash x_1$	1	2	3
1	10	20	30
2	20	30	40
3	30	40	50

$Y = f(x_1, x_2)$

\uparrow
 $E[Y|x_2=2] = 30$

$\leftarrow E[Y|x_1=1] = 20$

Fig. 3. An example of a possible deterministic model $f(x_1, x_2) = Y$ which depends on the factors x_1 and x_2 , together with examples of conditional expectations

2.3 Variance-Based Methods

Variance-based methods use the *variance of the conditional expectation* (VCE) as a measure of the importance of the input factors. The goal in these methods is to estimate the VCE by exploring the space made by all the possible values of the parameters. Applied to ODE-based models, the most well-known techniques are *correlation ratio*, *Sobol'*, and *Fourier amplitude sensitivity test* (FAST) [3,4,6]. Probability theory states that:

$$V[Y] = V_{\mathbf{x}}[E[Y|\mathbf{x}]] + E_{\mathbf{x}}[V[Y|\mathbf{x}]]. \quad (1)$$

The term $V_{\mathbf{x}}[E[Y|\mathbf{x}]]$ is the variance of the conditional expectation of Y , conditioned on \mathbf{x} . This is a suitable measure of the importance of \mathbf{x} , identifying the part of the variance of Y due to \mathbf{x} . If the variance of Y is matched by the VCE of \mathbf{x} we can say that \mathbf{x} is the only parameter (or set of parameters) which influences Y .

The variance of the conditional expectation is given by:

$$V_{\mathbf{x}}[E[Y|\mathbf{x}]] = \int (E[Y|\mathbf{x}] - E[Y])^2 p_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}$$

where $E[Y|\mathbf{x}] = \int y p_{Y|\mathbf{x}}(y) dy$. Here the integral is substituted with the sum over all the possible values of \mathbf{x} sampled from the range of \mathbf{x} . A simple example of a deterministic model is shown in Fig. 3.

The parameter space is sampled through the use of a grid. After having collected all the results, the conditional expectations are estimated by fixing a parameter to its possible values in the grid. A complete analysis of the influence of the parameters on the output and on the other parameters is provided but, as can be expected, the algorithm complexity increases exponentially with the grid level and the number of parameters.

Let $S_{\mathbf{x}}$ be the n -th order sensitivity index, with $\mathbf{x} \in \mathbb{N}^n$. This corresponds to the VCE fixing the factors in \mathbf{x} minus the sensitivity indices relative to all the possible combinations of the factors in \mathbf{x} . For example, S_{12} is given by $VCE_{12} - S_1 - S_2$ and S_{123} is given by $VCE_{123} - S_{12} - S_{13} - S_{23} - S_1 - S_2 - S_3$. The VCE relative to \mathbf{x} , where \mathbf{x} contains all the factors, is nothing but $V[Y]$.

Following [6] the sensitivity measure which is the most suitable to determine the influence of a parameter on the output of the model is the *Total Sensitivity Index* (TSI) or simply TS_i . This is defined as the sum of all the sensitivity indices that contain i in \mathbf{x} . For example, TS_1 is given by $S_1 + S_{12}$.

3 Sensitivity Analysis of Stochastic Simulations of Biochemical Reactions

In this section we introduce two new sensitivity measures and present them as variants of Morris' method and the variance-based approach respectively. When doing this, we compare these new techniques with their original versions.

From now on, when we refer to results obtained with ODE or deterministic methods, we implicitly intend that they are obtained using a 5/4 Dormand-Prince ODE solver with adaptive step-size. When we refer to results obtained with stochastic simulations, we implicitly intend that we used the original SSA [1], if not otherwise stated.

3.1 The Schlögl Model

The Schlögl model [5,2] is a suitable model to show the differences between usual Local OAT approaches and the one based on histogram distance. It is defined as follows:

Reaction channels	Propensity functions	Stochastic constants	Molecular populations
$A + 2X \xrightarrow{a_1} 3X$	$a_1 = k_1AX(X-1)/2$	$k_1 = 3 \cdot 10^{-7}$	$X_0 = 247$
$3X \xrightarrow{a_2} A + 2X$	$a_2 = k_2X(X-1)(X-2)/6$	$k_2 = 1 \cdot 10^{-4}$	$A = 1 \cdot 10^5$
$B \xrightarrow{a_3} X$	$a_3 = k_3B$	$k_3 = 1 \cdot 10^{-3}$	$B = 2 \cdot 10^5$
$X \xrightarrow{a_4} B$	$a_4 = k_4X$	$k_4 = 3.5$	

where A and B are kept constant. That is, they are available in sufficient supply that we do not model changes to their molecular populations. The parameter values are set close to a bifurcation point, where a small perturbation in them can lead to completely different results in the ODE time evolution, as can be seen in Fig. 4 (left and centre).

From a single set of parameters the time evolution of the stochastic simulations will follow either one of two possible behaviours, as can be observed in Fig. 4 (right). With the goal of describing the behaviour of this system, ODE models, or the simple average of X from different stochastic simulations could be inappropriate if not misleading. The use of estimated distributions can be considered a more suitable choice.

3.2 Local Methods

Three local one-at-a-time sensitivity analyses have been applied to the Schlögl model: *LOAT (ODE)*, *LOAT (Gillespie Average)* and *LOAT (Gillespie Density)*.

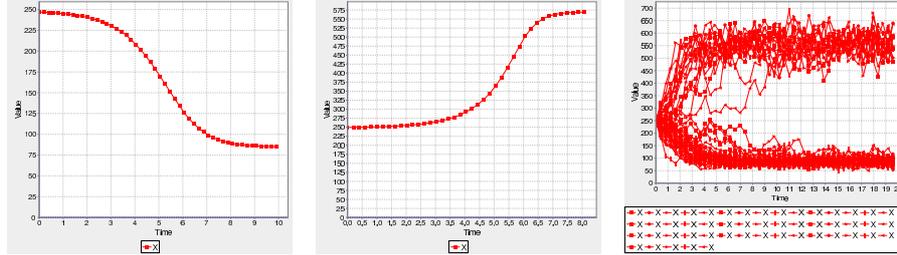


Fig. 4. Left, the time evolution of the output variable X of the Schlögl ODE model obtained with the value of the parameters stated in the text. Centre, the time evolution of X changing only the initial number of molecules of X from 247 to 250: the behaviour seems to completely change. Right, 50 runs of the SSA on the Schlögl model shows the behaviour in a more informative way.

These differ in the way in which the distance is calculated and in the method used to compute the time evolution of the system.

LOAT (ODE): The difference is computed from the output resulting from the ODE models. Performing the analysis more than once will lead to the same result, due to the deterministic nature of the ODEs.

LOAT (Gillespie Average): Many exact SSA simulations are computed here, so the result may change from analysis to analysis, reducing its variation if the number of stochastic simulations increases. The average of the simulations output is used.

LOAT (Gillespie Density): Also in this case, the exact SSA (Gillespie's Direct Method) is used to compute the evolution of the system. In this analysis the histogram distance is used instead of the simple difference of the averages.

Given the difference in the order of magnitude of the parameters of the Schlögl model, we may be more interested in the *relative* perturbation. For this reason we consider the simple output difference a more interesting sensitivity index than the derivative and we will discuss that first.

In Fig. 5 the first significant observation is that the ODE and Gillespie Density procedures share common results. They both show that k_1 produces the same variation as A and that k_3 produces the same variation as B . Indeed, we know that k_1 and A are related, because they could have been considered a single parameter (consider the propensity functions), and this fact has been captured by the analysis. The same reasoning holds for k_3 and B . On the other hand, an important and expected difference appears in the influence of X_0 . With ODEs, the output variation induced by the perturbation of X_0 is similar to that of k_1 and k_3 , showing high sensitivity. This is due to crossing the bifurcation point. The Gillespie Density method shows instead a low value of histogram distance for the same perturbation, revealing it to be far less influential than k_1 . (This latter method can easily be proved to be the correct one by considering Fig. 6,

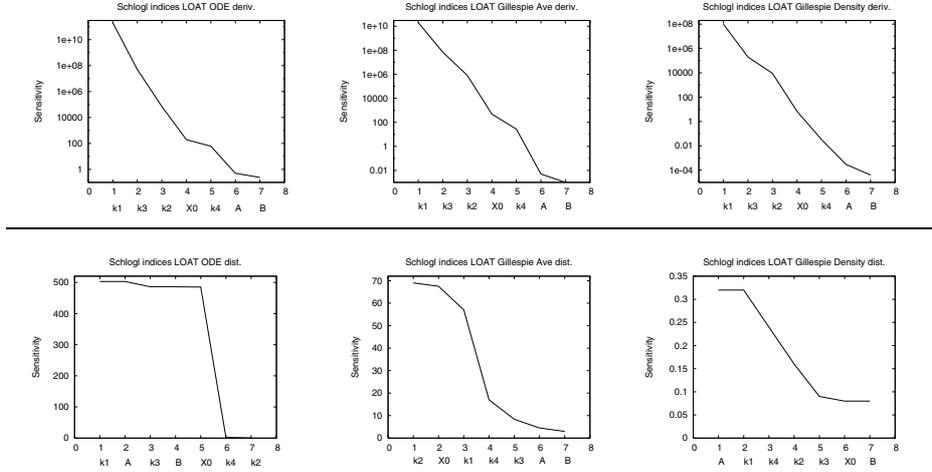


Fig. 5. Results of the three LOAT SA described in the text. The time of the analysis is 20 seconds with a perturbation of 1%. Above the line all results are computed using derivatives (distance divided by the perturbation). Below the line all results are computed using elementary OAT (just distance). The histogram distance is computed with 50 histogram columns and 5000 runs. The histogram self distance for X is 0.068. ODE fractional value 0.001. Results obtained using the simulator Dizzy [7].

where the histograms of the distribution of X at time 20, generated with nominal and perturbed parameter values, nearly coincide. In Fig. 6 can also be observed how the perturbation of k_1 influences the outcome of the stochastic simulations.) Parameter k_3 , along with B , has been discovered to be not particularly influential, with a histogram distance close to the self distance.

The Gillespie Average approach seems instead inconsistent, particularly when it shows k_1 and A to have different sensitivities.

To conclude the discussion of the results we can notice how the derivative approach (Fig. 5, above the line) attributes the same order of importance to the parameters in all three cases. However, this is due mainly to their different orders of magnitude and not significant with regard to the sensitivity of the system. It is clear that, at least in this context, a parameter that is estimated to be of the order of 10^{-5} and a parameter that is estimated to be of the order of 10^3 are not directly comparable.

According to the results of this first study we will from now on prefer the simple distance, specifying the relative perturbation in percentage.

3.3 Screening Methods

In this section we apply Morris' method in two different versions: firstly an adapted version of the original algorithm which makes use of the output of ODEs; and secondly a novel modification which uses the information captured by sets of stochastic simulations.

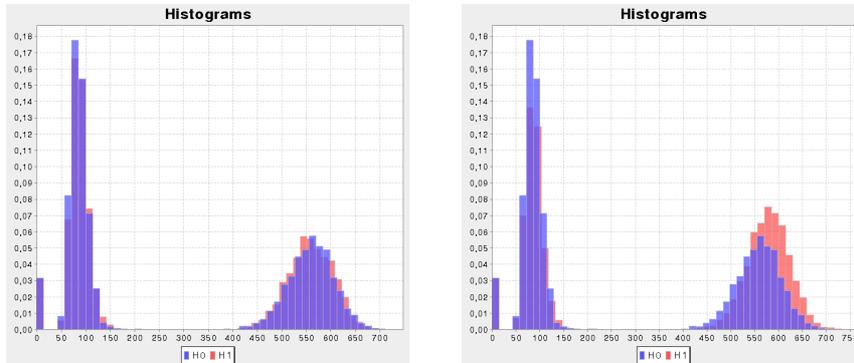


Fig. 6. Values of output X of the stochastic simulations at time 20. On the left with nominal parameter values (H0) and X_0 perturbed by 1% (H1). On the right with nominal parameter values (H0) and k_1 perturbed by 1% (H1). Each histogram is obtained from 5000 samples grouped in 50 columns.

Given our previous experience with LOAT SA we make use of two different *elementary effects*: one being the simple difference of the outputs of ODE models; the other the histogram distance of outputs of stochastic simulations. Moreover we consider the possibility of having multiple outputs:

$$d_{ij}(\mathbf{x}) = y_j(x_1, \dots, x_i + \Delta, \dots, x_k) - y_j(\mathbf{x})$$

$$d_{ij}(\mathbf{x}) = D(Y_j, Y'_j)$$

where in general d_{ij} is the local influence of the i th input on the j th output of the model. Considering a certain fixed time t when the analysis is performed, y_j is the outcome of the output j at that time and \mathbf{x} is the vector of parameters. Y_j is the *random variable* (r.v.) of the outcome of the j th output at time t distributed following the pdf $f(y_j, \mathbf{x})$ and Y'_j is the r.v. of the outcome of the j th output at the same time distributed following the pdf $f(y_j, x_1, \dots, x_i + \Delta, \dots, x_k)$. $D(Y_j, Y'_j)$ is the histogram distance between Y_j and Y'_j .

In order for these two measures to have meaning, we modified slightly the method to generate perturbations which are always comparable. The ranges are chosen as displacement from a nominal value which is proportional ($\pm 10\%$) to that value. In the p -level grid we allow only unitary perturbations (not multiples of $1/(p-1)$ but exactly $1/(p-1)$ every time). This way, every difference corresponds to the same percentage in perturbation with respect to the parameter nominal value which is central in the grid (Fig. 2).

Morris' Methods on the Schlögl Model. The two screening methods have been applied to the Schlögl model. Fig. 7 shows the outcome of the analysis with ODEs used to determine the time evolution of the system. The average elementary effect has the role of ordering the parameters from the most to the least influential. However, the elevated standard deviation of all the parameters

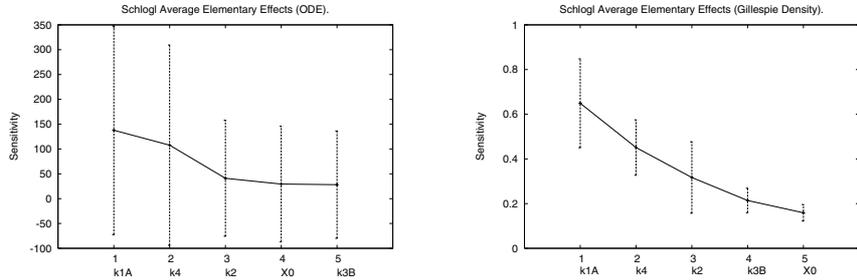


Fig. 7. Result of Morris' method on the Schögl model. Time evolution is computed with ODEs (left) and with stochastic simulations (right). The adopted parameters are 1000 random points for ODE r (40 for stochastic simulations), grid level 5 p , time of the analysis 10s, $\pm 5\%$ from nominal value, ODE fractional value 0.001. Number of stochastic simulations 1000 and number of histogram columns 50. The average histogram self distance of the random points was 0.141 with std. dev. 0.025.

sensitivities makes this classification difficult and reveals that the model is likely to be nonlinear with respect to the parameters and strong dependency between the parameters is also likely to exist.

Fig. 7 shows also Morris' method applied using histogram distance. It is important to bear in mind that all the conclusions are up to the level of precision that is given by the average self distance. Observing Fig. 7 (right) we can at this point say that, with the current approximations, the initial number of molecules of the species X is a factor that appears not to be influencing the value of the species X at time 10 seconds. We can also see that the product k_3B has a weak influence and that this influence does not change particularly as other parameters change (relatively low standard deviation). The other three parameters, k_2 , k_4 and k_1A show instead that they have a significant influence, particularly k_1A , and their relative larger standard deviation implies non-linearity and correlations. The reduction of the relative standard deviation in the novel Morris' method helps us to be more confident when stating which factors are the most important and which require to be further analysed.

3.4 Global Methods

We applied variance-based analysis to the Schögl model, both with the ODE and the stochastic simulation approach. The analysis has been performed considering a subset of three parameters, selected as the most important factors arising from a previous analysis with Morris' method (Fig. 7 on the right). The factors are k_4 , k_1A and k_2 .

The results of the analysis of the two variance-based approaches are shown in Table 1. We notice that the order of importance of the three parameters is the same, according to the total sensitivity indices. Differences in the first and second order sensitivity indices may be due to the relative weaker importance

Table 1. Variance-based sensitivity analysis of Schlögl model. First-order sensitivity indices relative to the factors k_4 (1), k_1A (2) and k_2 (3) and other combined effects are shown. The last three rows show the total sensitivity indices. Time of the analysis 10s, grid level 5, fractional step size of ODE method 0.001, number of stochastic simulations 1000, number of histogram columns 50.

index	VCE with histogram distance		VCE with ODEs	
	sensitivity	rank	sensitivity	rank
S_1	0.244	2	21278	2
S_2	0.325	1	30366	1
S_3	0.064	5	4028	5
S_{12}	0.086	4	10033	3
S_{13}	0.008	7	299	7
S_{23}	0.086	3	456	6
S_{123}	0.054	6	8969	4
TS_1	0.392	2	40580	2
TS_2	0.551	1	49826	1
TS_3	0.213	3	13752	3

that k_2 seems to have in the classical analysis. Indeed, sensitivities involving k_2 , like S_{13} or S_{23} are weaker in the classical analysis.

4 Sensitivity Analysis of the Mitogen-Activated Protein Kinase (MAPK) Cascades

Mitogen-activated protein kinase (MAPK) cascades [8,9] are signalling pathways which share a particular common structure consisting usually of three levels, where the signal is transmitted from one level to another through the phosphorylation of a kinase. Once activated this phosphorylates the kinase at the next level down the cascade (Fig. 8, left). The MAPK protein that triggers the cell response usually needs to be activated through a two-site phosphorylation. The catalyst for this reaction is a MAPKK (MAPK kinase) molecule and, at the upper level, the same role belongs to a MAPKKK (MAPKK kinase) molecule. The last molecule in this model is the MKP (MAP kinase phosphatase) which dephosphorylates, and so deactivates, the MAPK molecule.

We consider a single level of the MAPK cascade with only one MAPK kinase and without making any distinction between MAPK phosphorylated on tyrosine or threonine. The model consists of a two step double phosphorylation (Fig. 8 on the right). When speaking about this level of the MAPK cascade, we use M , Mp and Mpp as the unphosphorylated, monophosphorylated and biphosphorylated forms of MAPK.

The model of MAPK which we use in this section has been presented in [8] as a system of ODEs which describe the evolution of the concentration of M , Mp and Mpp in time. The rate at which these concentrations change is obtained using assumptions from the Michaelis-Menten kinetics. We use the same set of

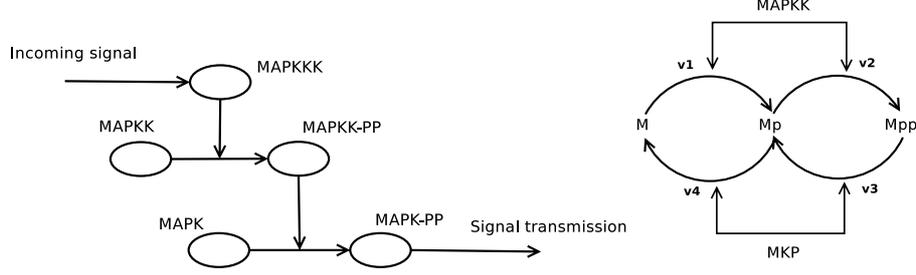
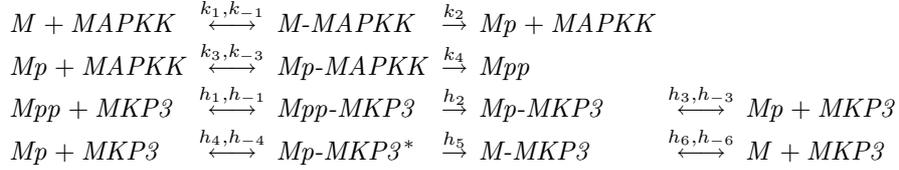


Fig. 8. On the left: structure of a MAPK cascade. At each level, the enzyme that catalyzes the reaction in the next level is activated by a two-site phosphorylation. On the right: model of a level of the MAPK cascade.

equations, but with number of molecules instead of concentrations. This procedure is correct if we assume that the product of the cell volume and the Avogadro number is equal to 1.

The system in Fig. 8 (on the right) is defined by the following enzymatic reactions. Notice how, in the first two lines, phosphorylation and product dissociation are considered a single step, while, in the last two lines, dephosphorylation and product release are two distinct steps.



This system can be reduced to only four reactions, under the assumptions of constant number of ATP/ADP molecules and protein-protein complexes at steady-state. These are the resulting reactions and rate equations.

$$\begin{aligned}
 M &\xrightarrow{v_1} Mp & v_1 &= \frac{k_1^{cat} \cdot MAPKK \cdot M / K_{m1}}{(1 + M / K_{m1} + Mp / K_{m2})} \\
 Mp &\xrightarrow{v_2} Mpp & v_2 &= \frac{k_2^{cat} \cdot MAPKK \cdot Mp / K_{m2}}{(1 + M / K_{m1} + Mp / K_{m2})} \\
 Mpp &\xrightarrow{v_3} Mp & v_3 &= \frac{k_3^{cat} \cdot MKP3 \cdot Mpp / K_{m3}}{(1 + Mpp / K_{m3} + Mp / K_{m4} + M / K_{m5})} \\
 Mp &\xrightarrow{v_4} M & v_4 &= \frac{k_4^{cat} \cdot MKP3 \cdot Mp / K_{m4}}{(1 + Mpp / K_{m3} + Mp / K_{m4} + M / K_{m5})}
 \end{aligned}$$

In these expressions $MAPKK$ and $MKP3$ are the total amount of molecules of the two enzymes and are considered constant through time. The nominal

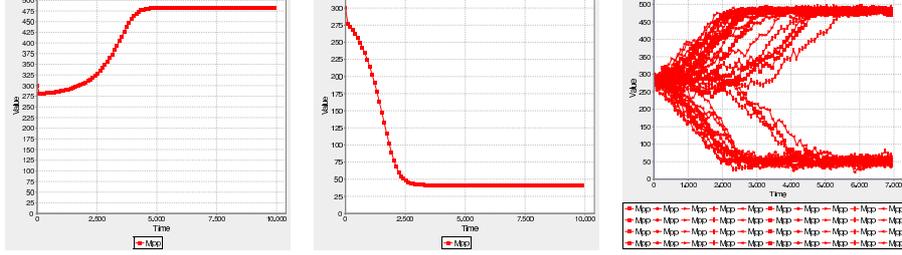


Fig. 9. Left, time evolution of the Mpp molecule of the MAPK model computed with ODEs using the nominal parameter values described in the text. Centre, the initial number of molecules of the phosphatase MKP3 is incremented by 5%. Right, 40 runs of the SSA with the nominal parameter values show how the evolution of the system may lead to two different stable systems.

values of the parameters and the relationship with the kinetics of the elementary enzymatic reactions are given below.

$k_1^{cat} = k_2 = 0.01$ $k_2^{cat} = k_4 = 15$ $k_3^{cat} = h_2 / (1 + h_2 / h_3) = 0.084$ $k_4^{cat} = h_5 \cdot (1 + h_5 / h_6 + h_{-3} \cdot (h_{-4} + h_5) / (h_3 \cdot h_4))^{-1} = 0.06$	$M_0 = 200$ $Mp_0 = 0$ $Mpp_0 = 300$ $MAPKK_0 = 50$ $MKP3_0 = 100$
$K_{m1} = (k_{-1} + k_2) / k_1 = 50$ $K_{m2} = (k_{-3} + k_4) / k_3 = 500$ $K_{m3} = (h_{-1} + h_2) / (h_1 + h_1 \cdot h_2 / h_3) = 22$ $K_{m4} = (h_{-4} + h_5) \cdot (h_4 \cdot (1 + h_5 / h_6 + h_{-3} \cdot (h_{-4} + h_5) / (h_3 \cdot h_4)))^{-1} = 18$ $K_{m5} = (h_6 / h_{-6}) = 78$	

The particularity of these parameter values is that they are close to a bifurcation point. As can be seen in Fig. 9 (left and centre), a small perturbation of an ODE parameter value can lead to a radical change in the behaviour of the time evolution of the double phosphorylated MAPK (Mpp). As with the Schlögl model a set of runs of the SSA shows that the real behaviour of the system with the nominal parameters is a choice between two stable systems. Moreover, thanks to [8], we know that this system with the stated parameters presents three steady-states which we can consider to be three attractors for the stochastic simulations. This situation is confirmed by the graph of the time evolution of Mpp in Fig. 9 (right). Although the choice appears to be between two attractors, it is delayed in those runs which are influenced by a central attractor.

4.1 Sensitivity Analysis

In this section we apply both classical SA and the techniques which we developed earlier to the presented MAPK model. We will proceed with a comparison of the methods throughout the analysis. Our choice is to measure the influence of

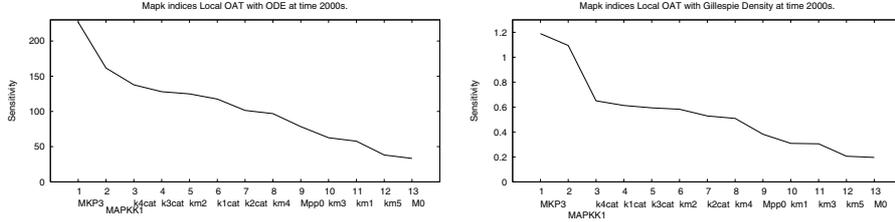


Fig. 10. LOAT sensitivity analysis of the MAPK model at time 2000 seconds. The result of classical analysis is on the left and the result of the analysis based on histogram distance is on the right. ODE time evolution is computed with fractional step size of 0.0001, while we used 10000 stochastic simulations and 50 histogram columns in the SSA runs. The perturbation of each parameter has been by 5%. The histogram self distance is 0.1.

the factors, kinetics and initial number of molecules, on the amount of double phosphorylated MAPK (Mpp). To do so, we choose the time of the analysis to be 2000 seconds. This time, as revealed in Fig. 9, is at the core of the choice between the two possible behaviours of the system and is within the limits of our possibilities in terms of computational power when using the SSA.

Local one-at-a-time Analysis. As a first step in the sensitivity analysis of the MAPK model, we performed a LOAT analysis. As we have seen, this consists of the perturbation of one of the factors at a time and in the measurement of the corresponding output change with respect to the original model. We used two different measures: the simple difference of the values of Mpp at time 2000 seconds generated using ODE-based results; and the histogram distance between the sets of values of Mpp at time 2000 seconds collected using stochastic simulations. With this first and computationally inexpensive analysis, we can have an idea of the relevance of the factors in the immediate surrounding of the factor nominal values. However, we have to bear in mind that without a global analysis we cannot be certain of the implications that may arise from perturbing more than one factor simultaneously. This last point cannot be neglected when trying to assert the influence of a factor on the model.

The results of the local one-at-a-time analysis are shown in Fig. 10. The thirteen factors are listed in the graphs from the most relevant to the least. We can notice that the relative order of importance is not particularly affected by the method used for the analysis. However, with the deterministic approach it appears that just the amount of phosphatase MKP3 is the most relevant factor, while with the stochastic approach, the intuition is that both the amount of kinase MAPKK and phosphatase MKP3 are the most relevant factors, above all the others.

This last statement can be defended, at least in this local analysis, through reference to the histograms generated using the results of the stochastic simulations of the perturbed models. Fig. 11 highlights that the initial amount of MAPKK and MKP3 are both the most influential factors. Moreover, they play

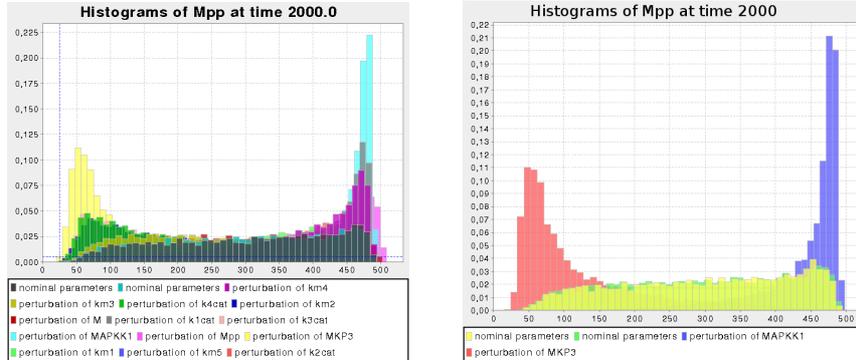


Fig. 11. Histograms which collect the values of Mpp obtained using 10000 stochastic simulations. Each histogram is divided into 50 columns. Left, all the histograms resulting from the one-at-a-time (OAT) analysis, one for each factor perturbed. Right, a second OAT analysis with only the histograms relative to the perturbations of the initial amount of MAPKK and MKP3. The histograms labelled with *nominal parameters* are those generated with the values of the parameters stated in the text.

the strongest role in the choice between the two possible stable systems. They have opposite roles, since increasing the amount of one of the two enzymes leads to opposite choices. It is indeed not surprising that the condition of bi-stability is guided by the right proportion in the amount of enzymes that catalyze the reactions.

Screening with Morris' Methods. Before we proceed to a more detailed analysis, we wish to use a screening method to identify and then exclude those factors that are clearly the least influential. Once we have isolated only a small part of most influential factors, we can proceed with the computationally expensive techniques which can provide the most detailed analysis. To do so, we use the techniques we developed earlier based on Morris' method. As we have seen, we consider a range of possible values for each factor and then we sample in the vector space generated by all the possible combinations of values of all the factors. This sampling is done randomly and through the use of a grid. We use here a grid level of five, meaning that each of the thirteen parameters can assume one of five possible values. For each random point selected in the grid of all the possible combinations of values, a LOAT analysis is performed. The indices resulting from that are the *elementary effects* which are local with respect to that random point. Averaging over all these local analyses reveals whether the degree of importance of a parameter is constant or changes when the other factors assume other values. The results of Morris' method applied to the MAPK model are shown in Fig. 12.

Also in this case we compare the results obtained with the deterministic method which uses the time evolution computed with ODE and the stochastic method which uses time evolution computed with SSA. The ODE-based

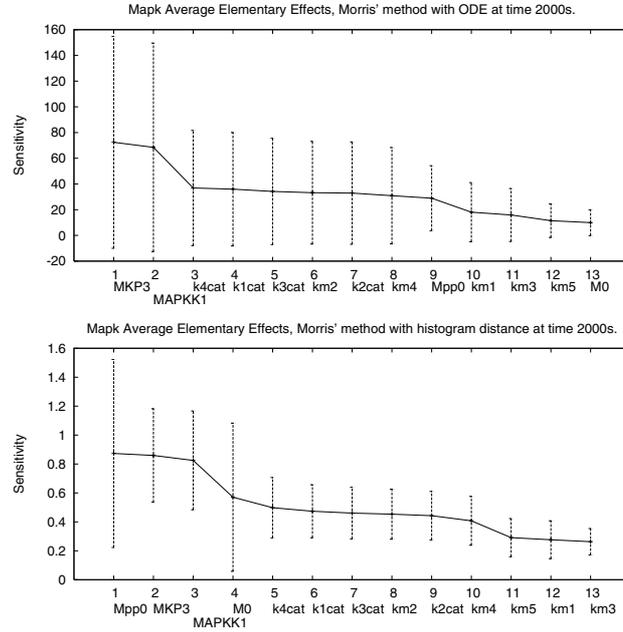


Fig. 12. The result of Morris' method applied to the MAPK model. ODE integration uses a fractional step size of 0.0001. Result obtained with a grid level of 5 and an average over 1000 random points. In the approach based on histogram distance, 1000 runs of SSA, 50 columns and 40 random points have been used and the average histogram distance is 0.15 with standard deviation of 0.061. The parameters vary within $\pm 10\%$ of their nominal value.

approach highlights that, although the most influential parameters are confirmed to be the initial amount of MAPKK and MKP3, the elementary effects of the factors are extremely variable. In this case it is difficult to say which factors we want to include in the detailed global analysis, if we exclude MAPKK and MKP3. The important standard deviation of the elementary effects is certainly due to a correlation between the factors and the non-linearity of the model output with respect to the parameters.

Before discussing the results obtained with Morris' method based on histogram distance, we need to point out that we were forced to limit the accuracy of the analysis, due to the demanding asymptotic complexity of the algorithms and the computational power available to us. Each experiment is made of 1000 stochastic simulations, number which leads to a relatively high histogram self distance of 0.150, and a standard deviation of 0.061. However, we have already seen in the LOAT analysis that the self distance can be considerable even with the greater precision of 10000 stochastic simulations (self distance of 0.1, see Fig. 10). Therefore, it appears that the point in time where we perform our analysis is particularly unstable, with high stochasticity and indecision from the single runs about which stable system to choose. We can then assume that we

have two factors that limit the accuracy of our results: a limited number of stochastic simulations and a strong stochasticity already present in the model.

The results of Morris' method based on histogram distance, shown in Fig. 12, confirm the strongly non-linear dependence in the model and the inconstant influence of the parameters on the amount of double phosphorylated MAPK at time 2000 seconds. On a more positive note, this method appears to achieve a more precise information than the ODE-based analysis. First of all, many parameters have reduced the standard deviation of their elementary effect. We can therefore be more confident when stating that some factors are less influential than others. Moreover, the strong influence which is attributed to the initial amount of the enzymes MAPKK and MKP3 is more clearly evident. Finally, this second analysis assigns a different role to the factors Mpp_0 and M_0 . Here, they appear to have a stronger average sensitivity, though this sensitivity may vary considerably (large standard deviation), showing a strong dependence on the value of the other parameters.

Global Analysis with Variance Decomposition. Thanks to the screening which we applied in the previous section, we can now apply a global and more informative method to a reduced set of parameters taken from the factors of the MAPK model. The factors that proved to be the most influential are the initial number of molecules of MAPKK and MKP3, so we investigate their influence as single parameters and their combined effect. For this purpose we used the techniques developed in Section 3.4. Again, a method based on differences of outputs of ODEs and one based on histogram distances of executions of SSA are compared. These measures consider the variance of the output: while the former focuses on the variance of the ODE output, the latter estimates the variance in the distribution approximated by histograms. In both cases, the quantity of the variance that is due to each parameter is identified (Table 2).

In both the approaches, the initial amount of MAPKK and MKP3 present the same level of importance, with the former that is slightly more influential. The difference lies in the importance that is given to the combined effect of the two factors. While with the first approach the combined effect is considerably less than the single effects, with the second approach it appears that the two

Table 2. First and second order sensitivity indices relative to the factors $MAPKK_0$ (1) and $MKP3_0$ (2) of the MAPK model and their combined effect (12), obtained computing the variance of the conditional expectation. The fractional step size used in the ODE integration is 0.0001, the number of stochastic simulations used is 5000 and the number of histogram columns is 50. The parameters vary within $\pm 10\%$ of their nominal value.

index	Variance-based with ODEs		Variance-based with histogram distance	
	sensitivity	rank	sensitivity	rank
S_1	15695.65	1	0.350	2
S_2	15308.66	2	0.332	3
S_{12}	5631.88	3	0.811	1

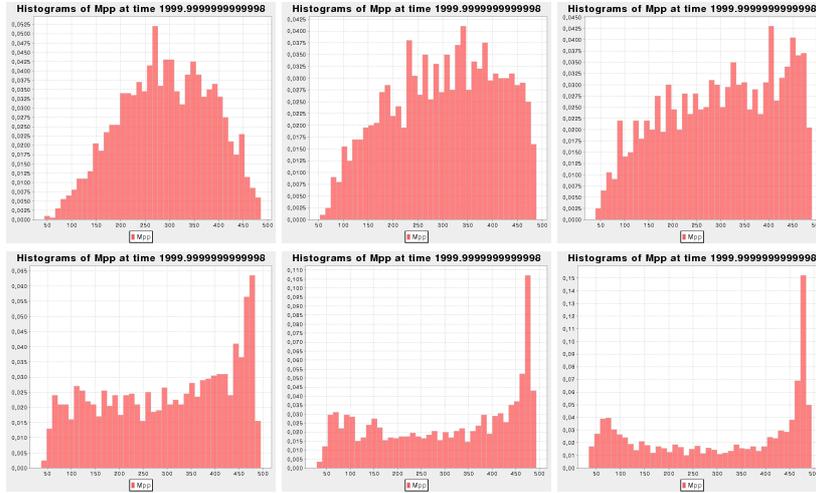


Fig. 13. Distribution of the values of Mpp from 2000 stochastic simulations at time 2000 seconds, with simultaneous perturbation of MAPKK and MKP3. Above, from left to right: number of molecules of the enzymes decreased by 20%, 10% and with their nominal value. Below, from left to right: number of molecules of the enzymes increased by 10%, 20% and 30%.

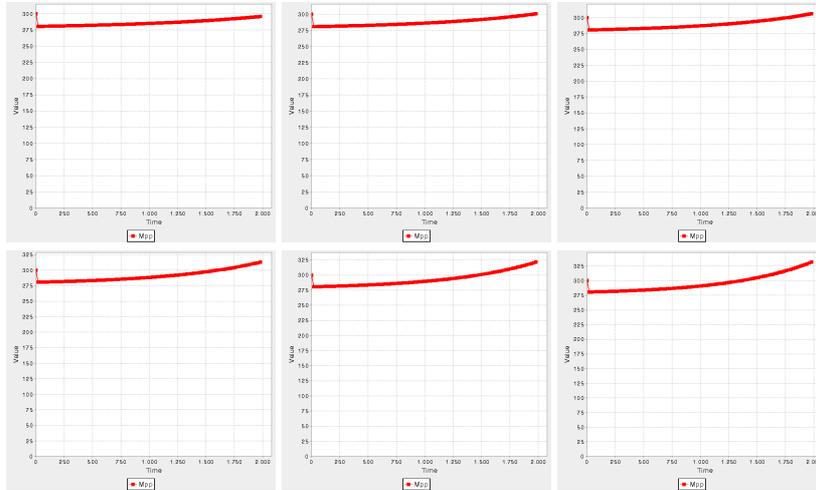


Fig. 14. Time evolution of the double phosphorylated MAPK (Mpp) with ODE for 2000 seconds, with simultaneous perturbation of MAPKK and MKP3. Above, from left to right: number of molecules of the enzymes decreased by 20%, 10% and with their nominal value. Below, from left to right: number of molecules of the enzymes increased by 10%, 20% and 30%.

parameters are more linked. Changing them together leads to a stronger influence with respect to a one-at-a-time change.

The visualisations in Fig. 13 help to prove the connection and reciprocal influence of the factors of this model. In this figure one can see that the combined perturbation of MAPKK and MKP3 leads to a variation of the distribution of the set of values obtained with the stochastic simulations. Although the mean of these values appears to be the same, the distributions seem to pass from a compact and largely Gaussian shaped (on the left) to a more irregular one, which begins to show the two peaks of the bi-stability. This observations can be interpreted as the simple fact that increasing the amount of enzymes accelerates the process, allowing the two stable choices to be reached sooner. Other interesting visualisations are those in Fig. 14, where we can observe that the ODE integration fails to interpret the high stochasticity and indecision present in the system at time 2000 seconds. However, also in this case, incrementing or decreasing the quantity of enzymes accelerates or slows the production of MAPK-PP (Mpp).

5 Conclusions

In this paper we have shown an example of how sensitivity analysis of a model of biochemical reactions can be performed using both deterministic and stochastic approaches. As a first result, we have shown how global analysis such as Morris' method first and the variance decomposition after, are necessary and must be used to identify the relationship between the factors. For example, if we had to rely only on a local analysis, we would just accept the order of importance given in Fig. 10. However, thanks to the further application of a global screening method (Fig. 12), we have been able to state that this order of importance may vary if we change the value of more than a single factor at once. This suggested, if not actually demanded, a further and more informative analysis concerning those factors that seemed the most influential and dependent on the others. In this case, we showed the intuitive relationship between the enzymes MAPKK and MKP3, whose simultaneous increment accelerates the system and whose proportions play the main role in the bi-stability of the system.

As a second but not less important result, the comparison between deterministic and stochastic approaches to sensitivity analysis highlighted how, when dealing with bistable systems near a bifurcation point, it becomes necessary to have a sensitivity analysis tool that takes into account the distribution behind a set of stochastic simulations. Although the analytical analysis of the ODEs is fundamental to identify the bifurcation points and the multiple steady-states, ODE integrations cannot model the uncertainty in the time evolution of the system close to those bifurcation points. In this situation of high stochasticity, a more suitable sensitivity analysis is one that takes into account the variations between sets of stochastic simulations rather than the simple output of a ODE integration. Here, for example, we have seen how a modified version of Morris' method, identified some properties that the deterministic method was not able to capture.

Acknowledgements. Stephen Gilmore is supported by the “Stochastic process algebra for biochemical signalling pathway analysis” (SIGNAL) project funded by the Engineering and Science Research Council (EPSRC) under grant number EP/E031439/1.

References

1. Gillespie, D.T.: Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry* 81(25), 2340–2361 (1977)
2. Cao, Y., Petzold, L.: Accuracy limitations and the measurement of errors in the stochastic simulation of chemically reacting systems. *J. Comput. Phys.* 212(1), 6–24 (2006)
3. Saltelli, A., Chan, K., Scott, E.M. (eds.): *Sensitivity Analysis*. Wiley, Chichester (2000)
4. Saltelli, A., Ratto, M., Stefano, T., Francesca, C.: Sensitivity analysis for chemical models. *Chem Rev* 105(7), 2811–2828 (2005)
5. Gunawan, R., Cao, Y., Petzold, L., Doyle, F.J.: Sensitivity analysis of discrete stochastic systems. *Biophys J* 88(4), 2530–2540 (2005)
6. Chan, K., Saltelli, A., Tarantola, S.: Sensitivity analysis of model output: Variance-based methods make the difference. In: *Proceedings of the 1997 Winter Simulation Conference*, pp. 261–268 (1997)
7. Ramsey, S., Orrell, D., Bolouri, H.: Dizzy: stochastic simulation of large-scale genetic regulatory networks. *J Bioinform Comput Biol* 3(2), 415–436 (2005)
8. Markevich, N.I., Hoek, J.B., Kholodenko, B.N.: Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades. *The Journal of Cell Biology* 164, 353–359 (2004)
9. Klipp, E., Herwig, R., Kowald, A., Wierling, C., Lehrach, H. (eds.): *Systems Biology in practice*. Wiley-Vch, Chichester (2005)