Quantitative Methods in Systems Biology
Part IV: Sensitivity Analysis

Stephen Gilmore
LFCS, University of Edinburgh

SFM:08-Bio Summerschool
Bertinoro, Italy

2nd June 2008

Outline

Acknowledgements

A. Degasperi and S. Gilmore.

Sensitivity Analysis of Stochastic Models of Bistable Biochemical Reactions.

SFM-08:Bio,

Bertinoro, Italy, 2008.

Outline

Background

- Reaction-based biochemical models use input parameters such as concentrations and kinetic rate constants to predict the time evolution of a biochemical system.

Background

- 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.

Sensitivity Analysis

- *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.

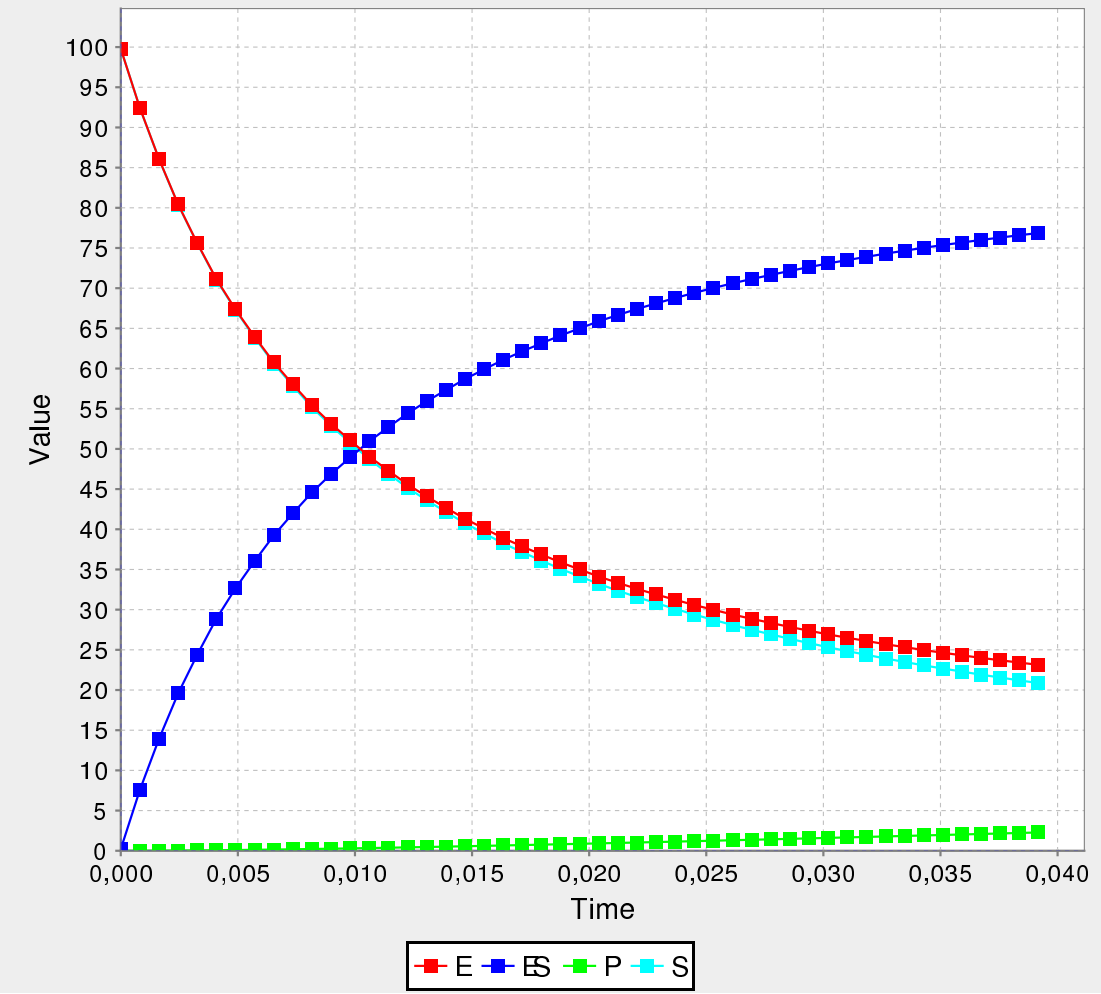
Sensitivity Analysis

- *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.

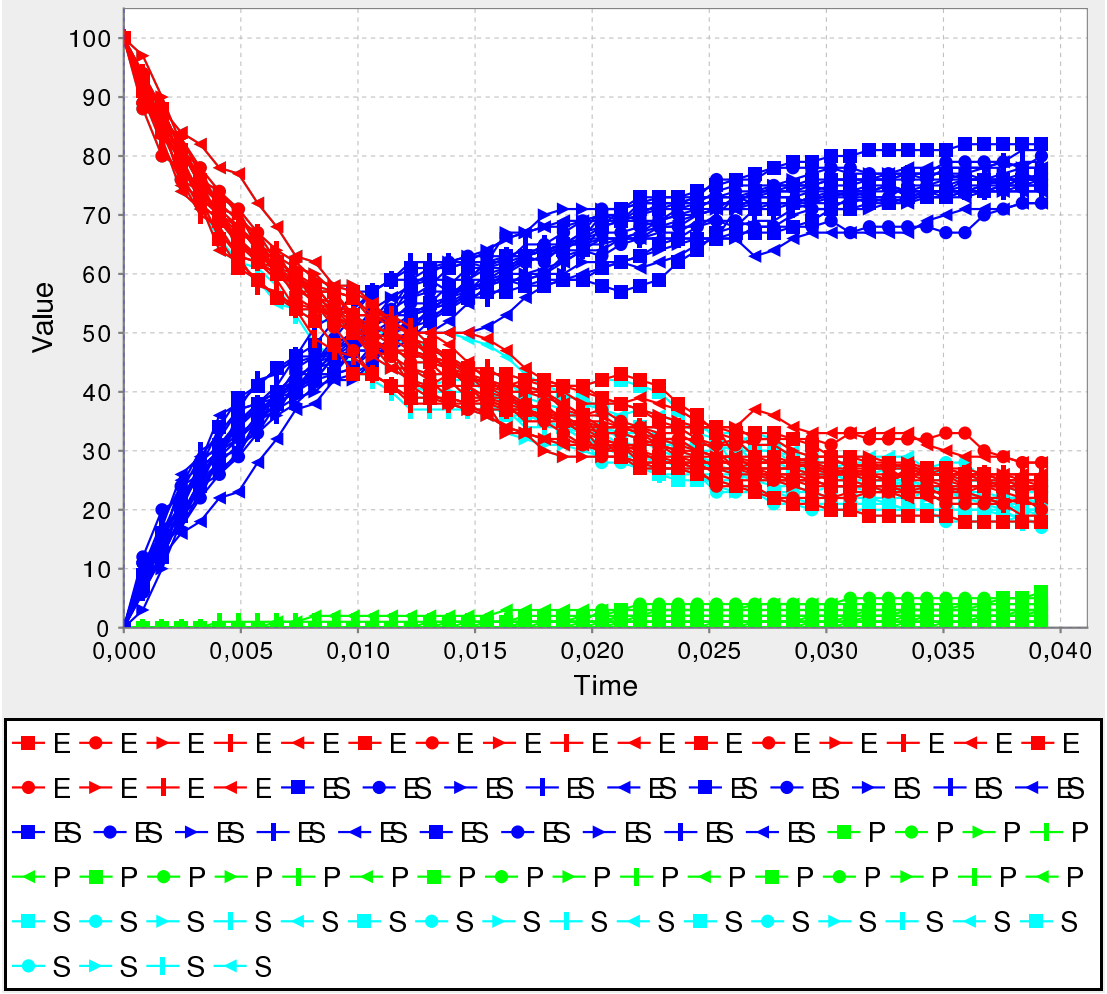
Sensitivity Analysis

- *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 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.

Modelling with ODEs



Modelling with Stochastic Simulation



Model perturbation

- 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.

Model perturbation

- 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.

Model perturbation

- 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.

Comparing simulation runs

- Another possibility is the use of a distribution distance or *histogram distance* which, with sufficient simulation runs, is able to precisely describe the difference.

Comparing simulation runs

- Another possibility is the use of a distribution distance or *histogram distance* which, with sufficient simulation runs, is able to precisely describe the difference.
- This has been used to quantify how well an approximate SSA emulates the exact SSA.

Comparing simulation runs

- Another possibility is the use of a distribution distance or *histogram distance* which, with sufficient simulation runs, is able to precisely describe the difference.
- This has been used to quantify how well an approximate SSA emulates the exact SSA.

Cao, Y., Petzold, L.:
Accuracy limitations and the measurement of errors in the stochastic simulation of chemically reacting systems.
J. Comput. Phys. **212**(1) (2006) 6–24

Comparing simulation runs

- Another possibility is the use of a distribution distance or *histogram distance* which, with sufficient simulation runs, is able to precisely describe the difference.
- This has been used to quantify how well an approximate SSA emulates the exact SSA.

Cao, Y., Petzold, L.:
Accuracy limitations and the measurement of errors in the stochastic simulation of chemically reacting systems.
J. Comput. Phys. **212**(1) (2006) 6–24

- We use it here with SA to quantify the effect of perturbation of the parameters of a stochastic model.

Histogram Distance

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 $L = x_{max} - x_{min}$.

Self-distance

- An interesting measure is then the *self distance*, given by $D_k(X, X')$.

Self-distance

- 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.

Self-distance

- 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.

Outline

Sensitivity Analysis Classifications

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.

Sensitivity Analysis Classifications

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.

Sensitivity Analysis Classifications

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.

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.

Sensitivity Analysis in Practice

- In the study of biochemical systems, OAT methods represent the prevalent practice when analysing ODE models.

Sensitivity Analysis in Practice

- In the study of biochemical systems, OAT methods represent the prevalent practice when analysing ODE models.
- However, they are not directly applicable to stochastic models whose output is defined as a *probability density function* (pdf) over the number of molecules for each species.

Sensitivity Analysis in Practice

- In the study of biochemical systems, OAT methods represent the prevalent practice when analysing ODE models.
- However, they are not 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.

Sensitivity Analysis in Practice

- In the study of biochemical systems, OAT methods represent the prevalent practice when analysing ODE models.
- However, they are not 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.

Sensitivity Analysis and Histogram Distance

SA of stochastic systems is an engaging research question 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 $= f(\mathbf{x}, p_1, ..., p_i + \Delta p_i, ..., p_k)$. This distance can instead be divided by Δp_i, leading to a correspondent derivative-based approach.

Local one-at-a-time methods

- 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.

Local one-at-a-time methods

- 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.

Local one-at-a-time methods

- 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.

Morris' method

- *Morris' method* 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' method

- *Morris' method* 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 $\mathbf{x}_1, ..., \mathbf{x}_r$ in the parameter space, and then taking their average.

Morris' method

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

1. negligible effects,
2. linear and additive effects, or
3. non-linear interaction effects.

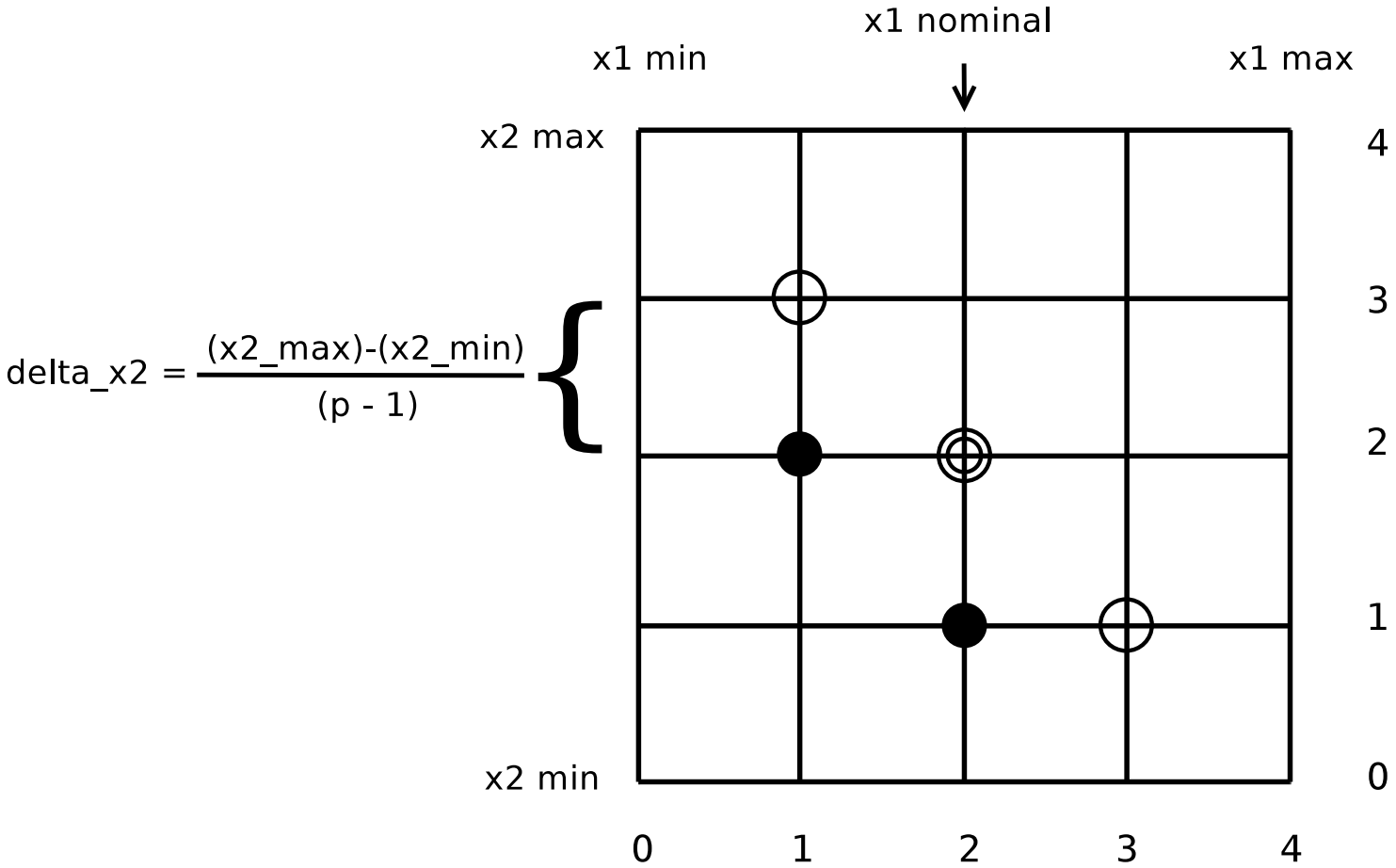
This will help to apply later the most appropriate global sensitivity analysis only on the relevant parameters.

Morris' method

The k-dimensional factor vector $\mathbf{x}$ has components x_i that have p values in the set $\{0, 1/(p-1), 2/(p-1), ..., 1\}$. The region of experimentation Ω is then a k-dimensional p-level *grid*.

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.

Example of a grid in the Morris method



Variance-based methods

- Variance-based methods use the *variance of the conditional expectation* (VCE) as a measure of the importance of the input factors.

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.

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).

Variance decomposition

Probability theory states that:

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

The term $V_{\mathbf{x}}\big[E[Y|\mathbf{x}]\big]$ 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.

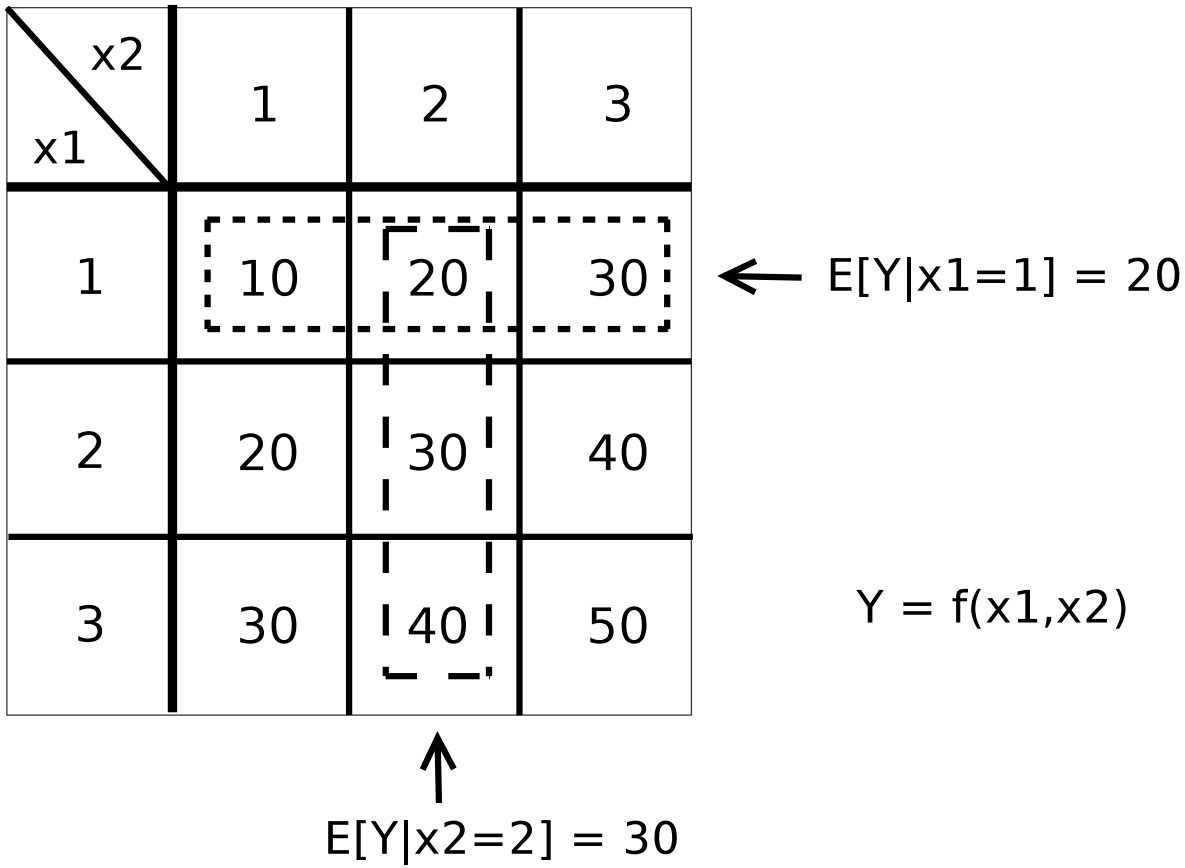
Variance of the Conditional Expectation

The variance of the conditional expectation is given by:

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

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 $\mathbf{x}$.

VCE example of a deterministic model



Outline

Sensitivity Analysis and Stochastic Simulation

In this section we compare SSA and ODE results for two bistable models.

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.

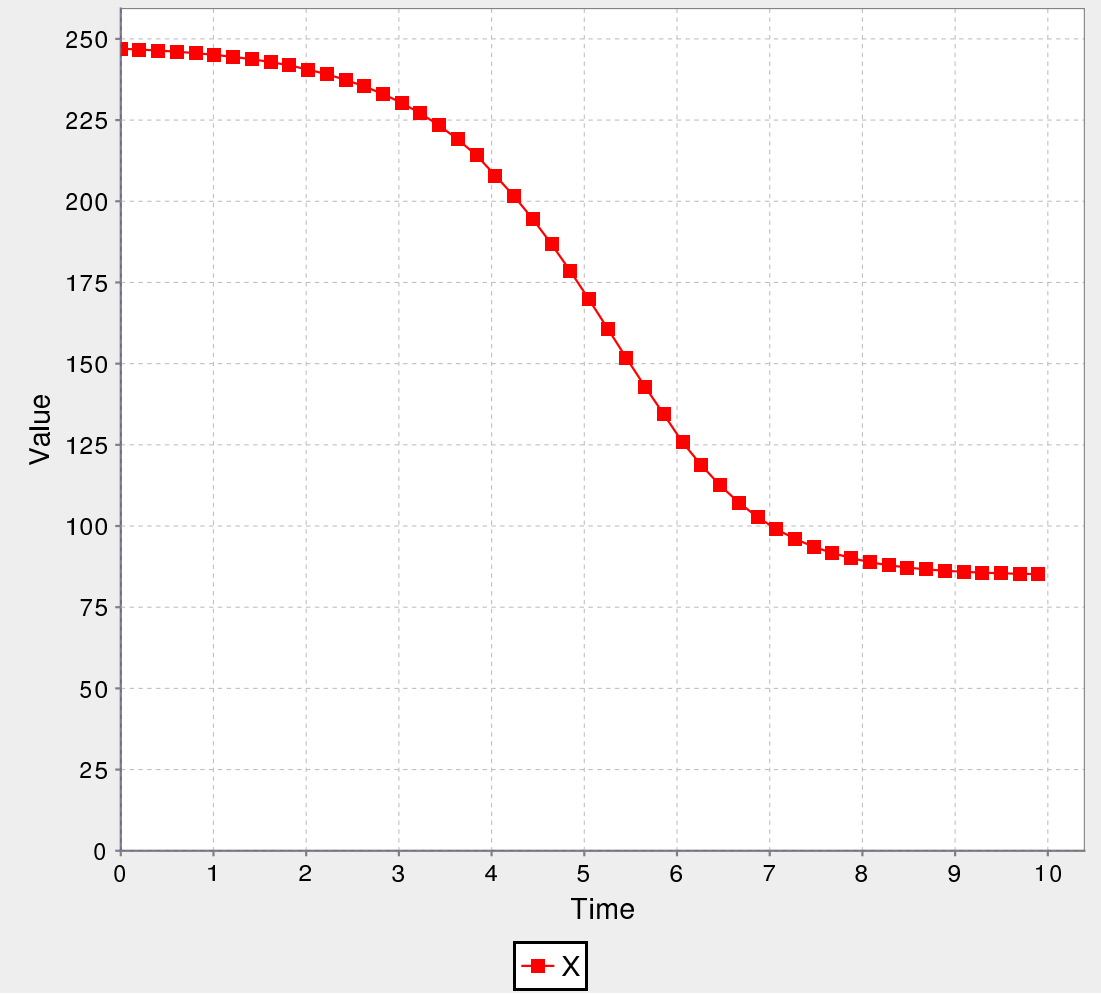
The Schlögl model

The Schlögl model 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:

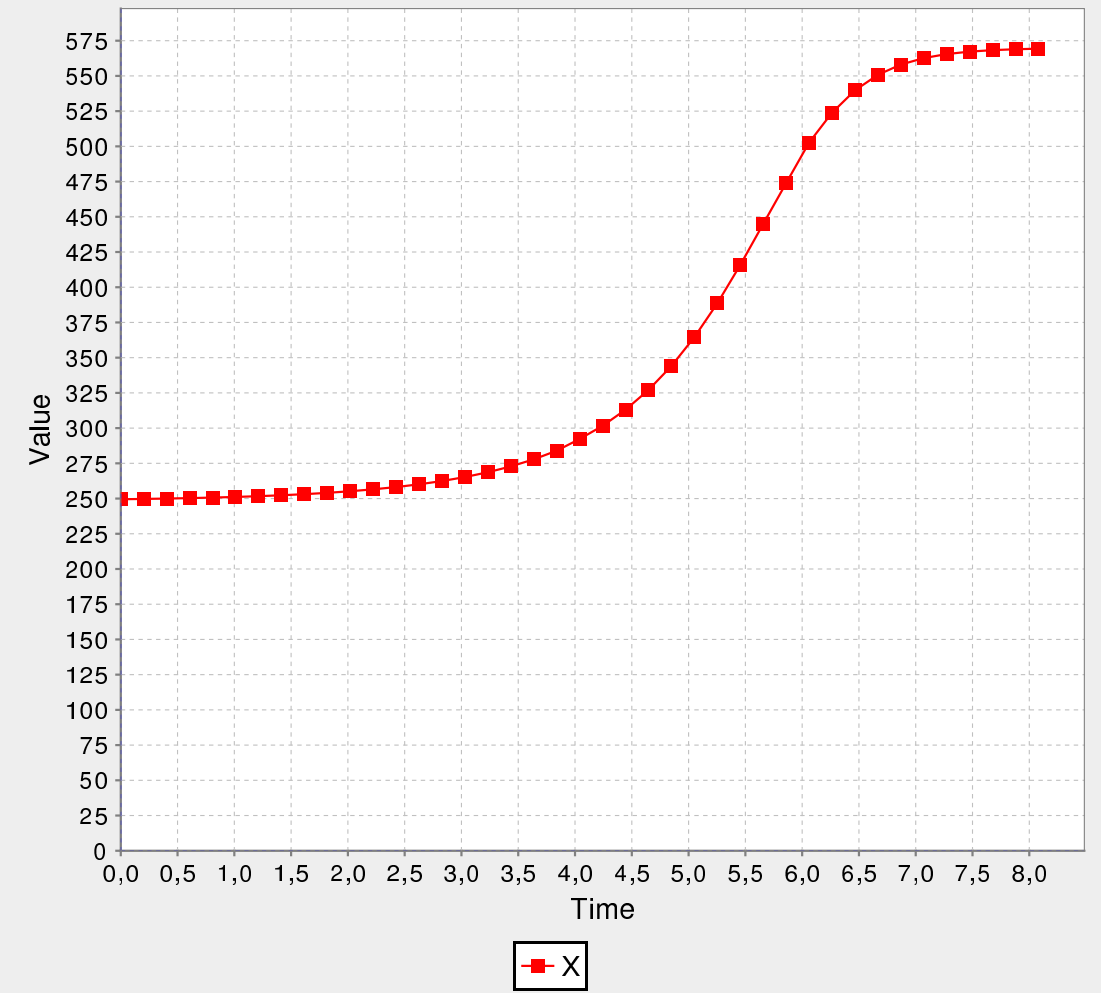
Reaction channels	**Propensity functions**	**Stochastic constants**	**Molecular populations**
$A + 2X \xrightarrow{a_1} 3X$	$a_1 = k_1 AX(X-1)/2$	$k_1 = 3 \cdot 10^{-7}$	$X_0 = 247$
$3X \xrightarrow{a_2} A + 2X$	$a_2 = k_2 X(X-1)(X-2)/6$	$k_2 = 1 \cdot 10^{-4}$	$A = 1 \cdot 10^5$
$B \xrightarrow{a_3}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

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.

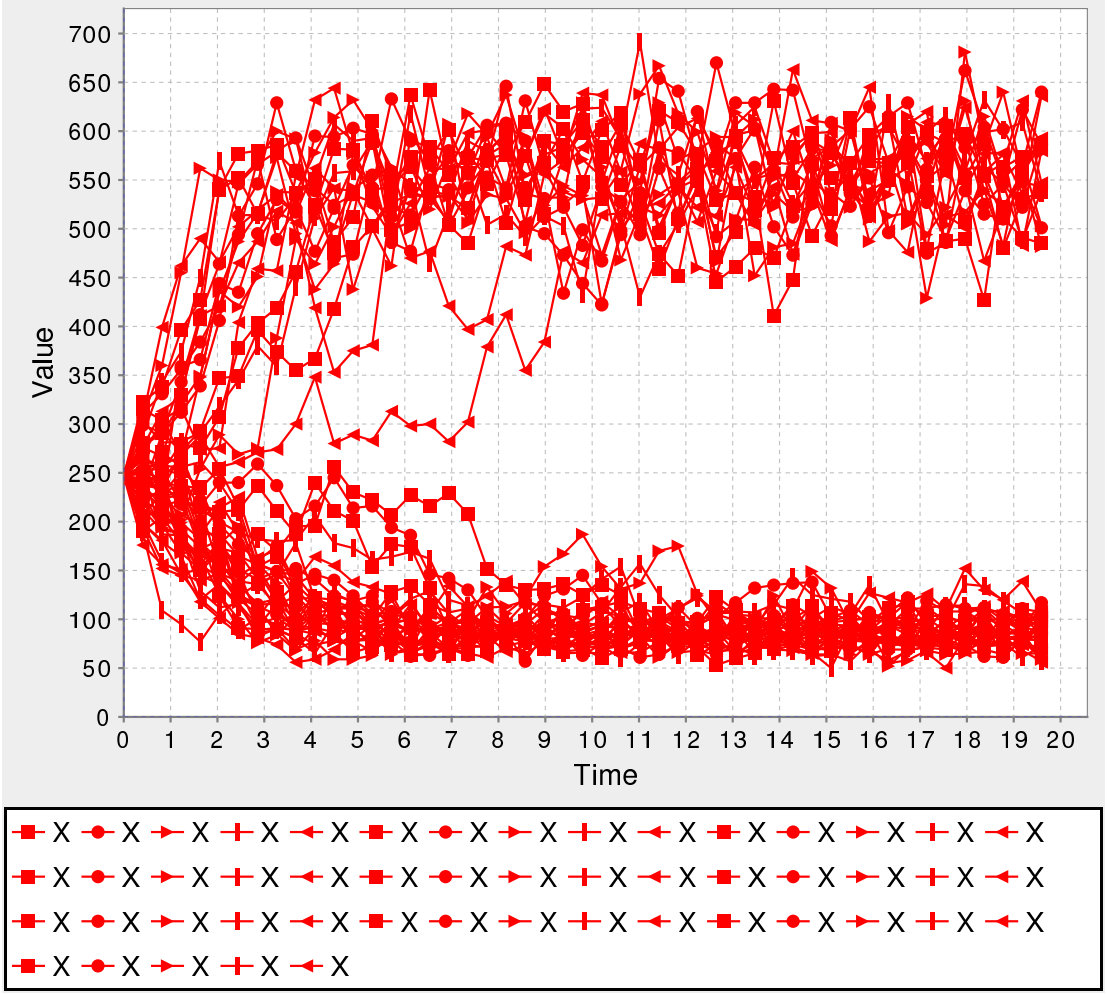
ODE analysis of the Schlögl model ($X_0 = 247$)



ODE analysis of the Schlögl model ($X_0 = 250$)



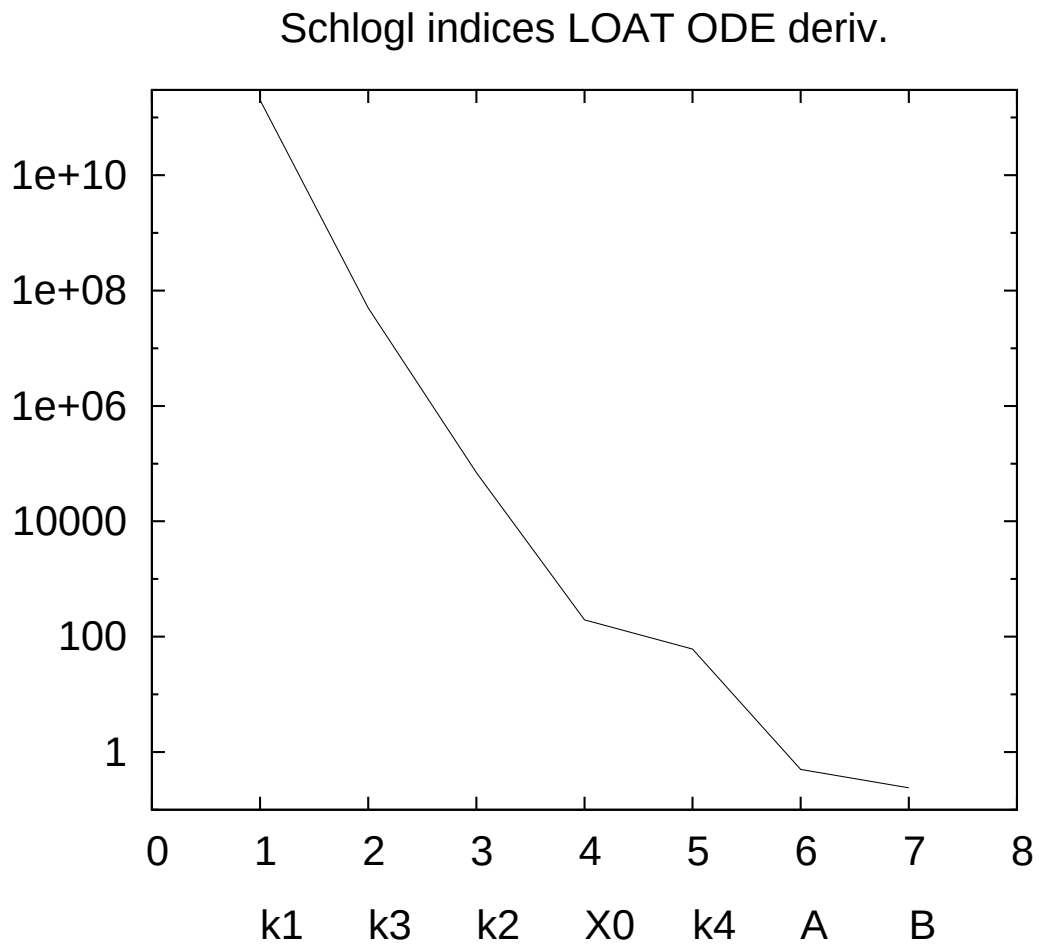
50 simulation runs of the Schlögl model



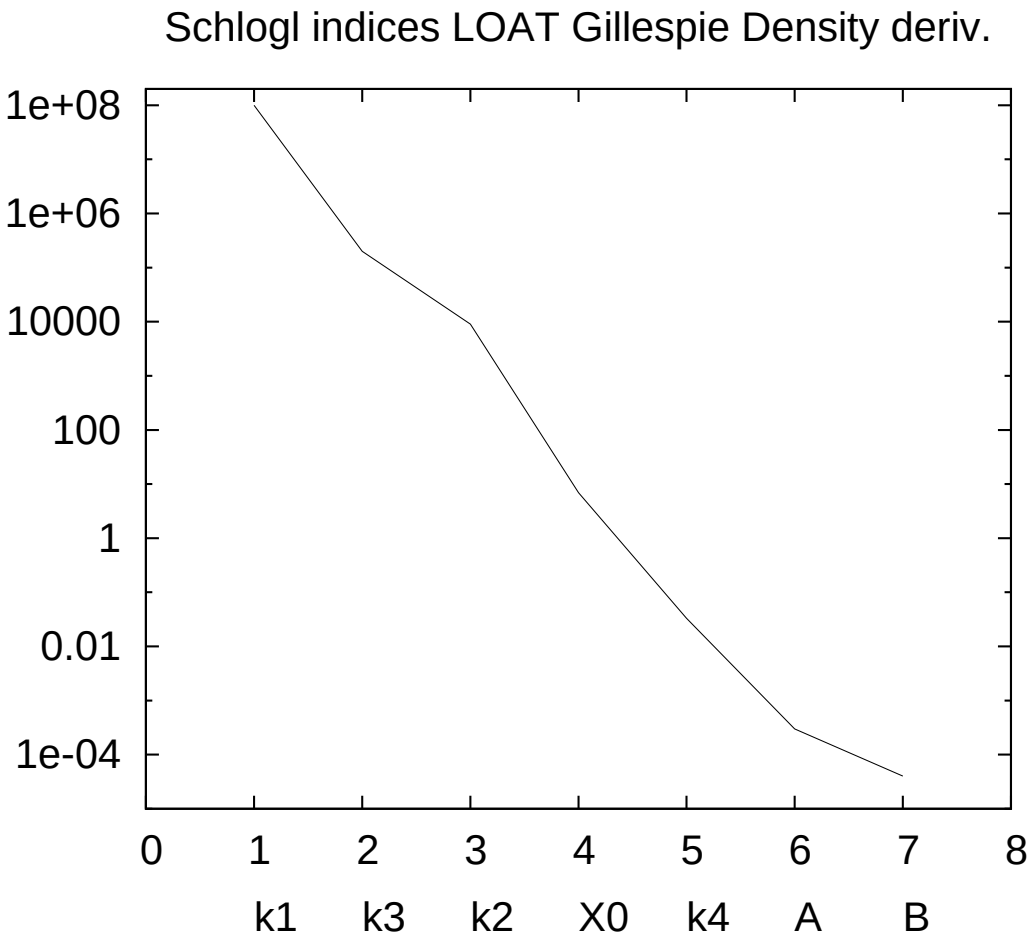
Bistable systems

From a single set of parameters the time evolution of the stochastic simulations will follow either one of two possible behaviours. 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.

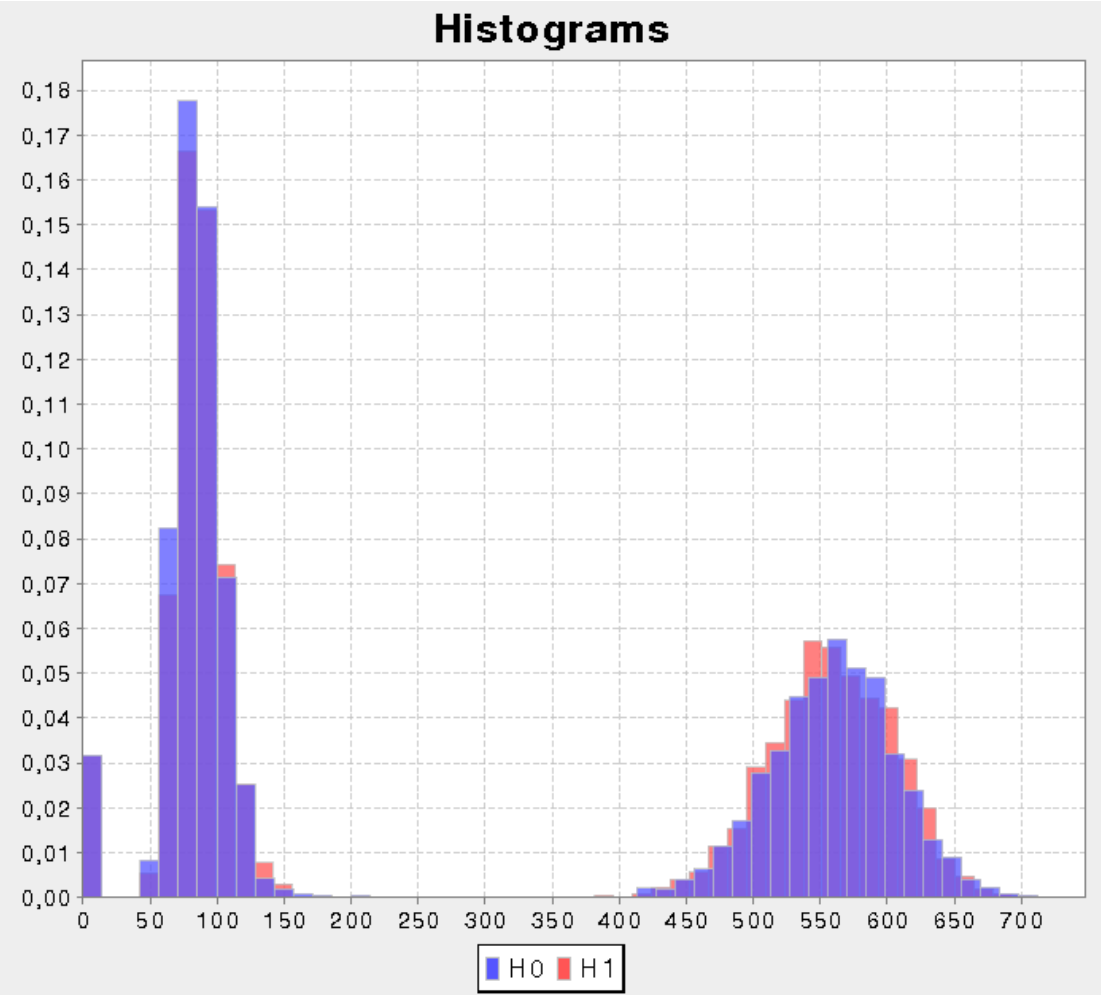
LOAT analysis (ODEs)



LOAT analysis (Simulation)

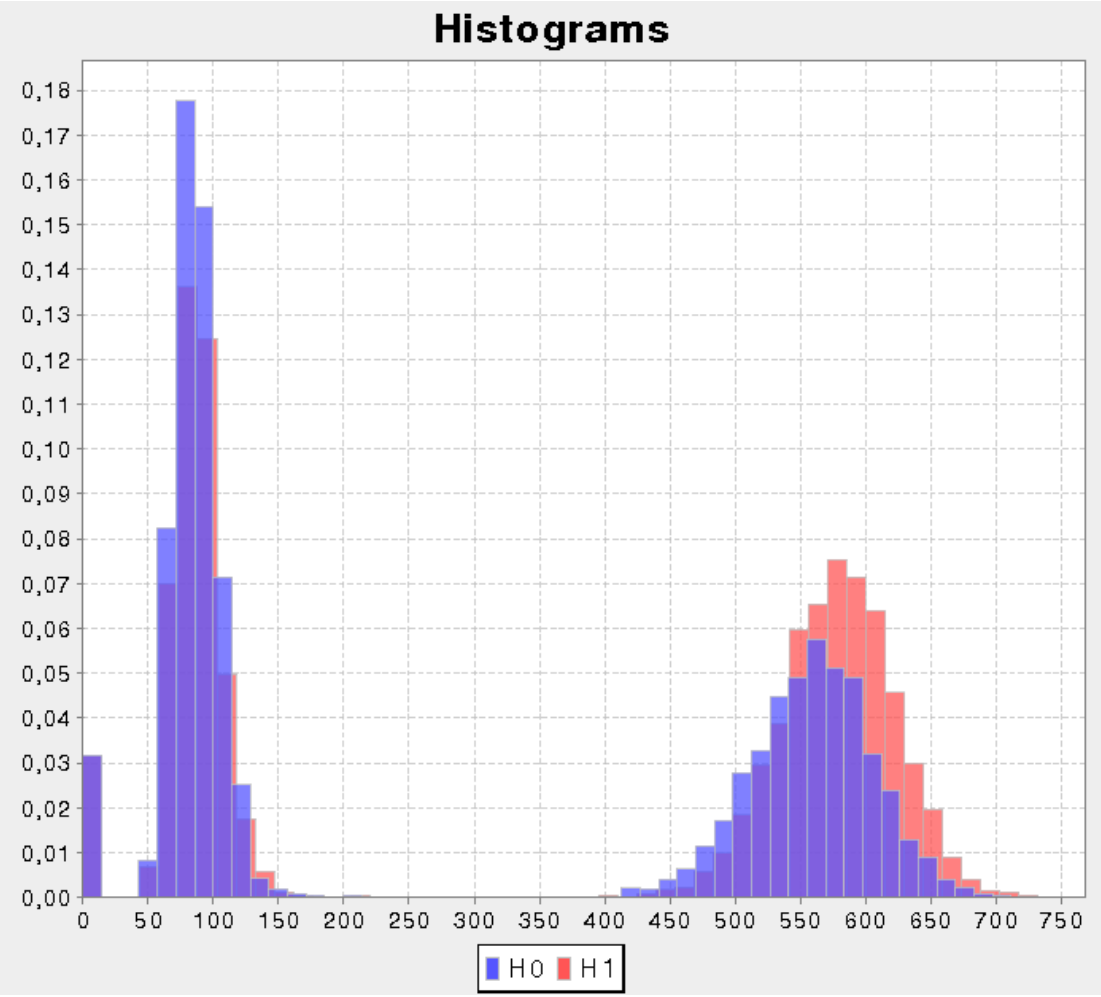


Values of X at $t = 20$ (from 500 simulations)



Here we have the model with nominal parameter values (H0) and X_0 perturbed by 1% (H1).

Values of X at $t = 20$ (from 500 simulations)



Here we have the model with nominal parameter values (H0) and k_1 perturbed by 1% (H1).

Mitogen-activated protein kinase (MAPK) cascades

- Mitogen-activated protein kinase (MAPK) cascades 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.

Mitogen-activated protein kinase (MAPK) cascades

- Mitogen-activated protein kinase (MAPK) cascades 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.

Mitogen-activated protein kinase (MAPK) cascades

- Mitogen-activated protein kinase (MAPK) cascades 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.
- The MAPK protein that triggers the cell response usually needs to be activated through a two-site phosphorylation.

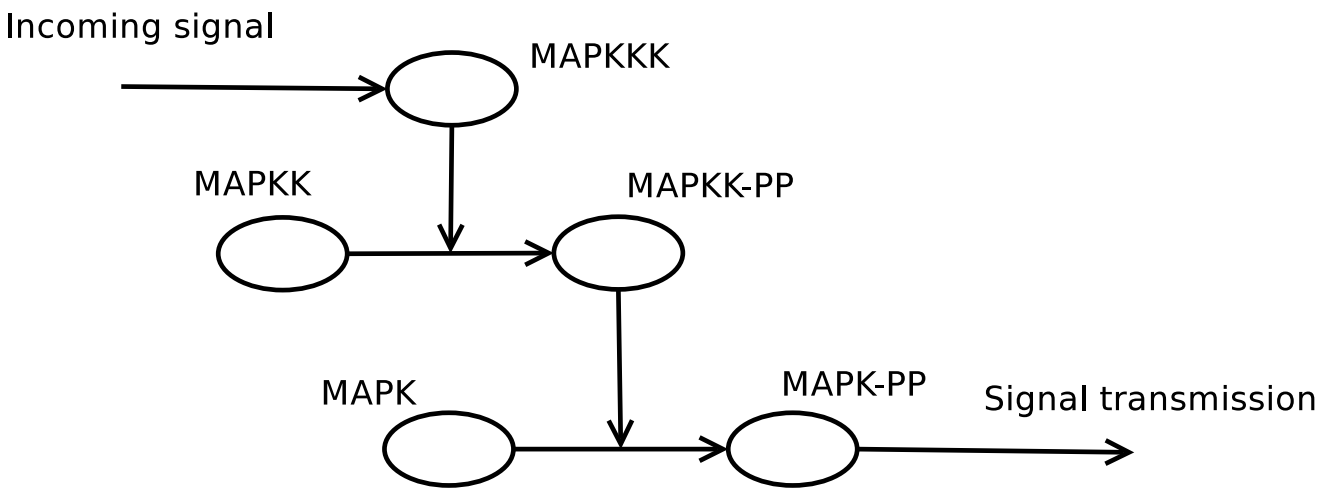
Mitogen-activated protein kinase (MAPK) cascades

- 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.

Mitogen-activated protein kinase (MAPK) cascades

- 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.

Mitogen-activated protein kinase (MAPK) cascade



Mitogen-activated protein kinase (MAPK) model

- 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 theronine.

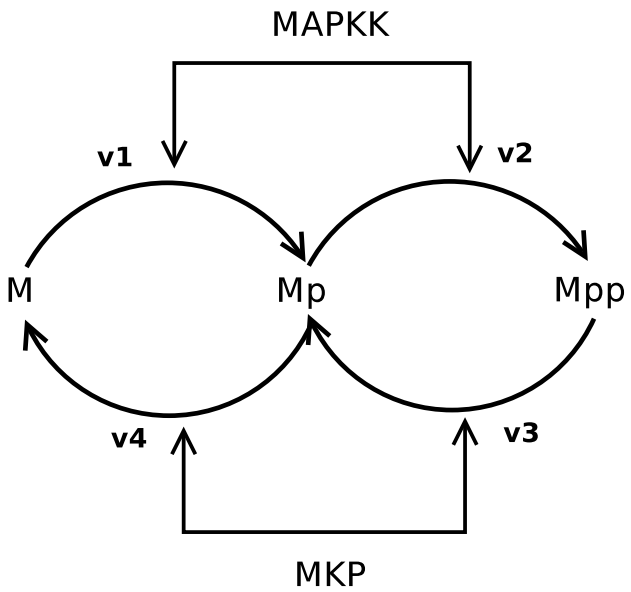
Mitogen-activated protein kinase (MAPK) model

- 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 theronine.
- The model consists of a two step double phosphorylation.

Mitogen-activated protein kinase (MAPK) model

- 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 theronine.
- The model consists of a two step double phosphorylation.
- When speaking about this level of the MAPK cascade, we use M, Mp and Mpp as the unphosphorylated, monophosphorylated and biphosphorylated forms of MAPK.

Mitogen-activated protein kinase (MAPK) model



Mitogen-activated protein kinase (MAPK) reactions

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.

$$M \xrightarrow{v_1} Mp \qquad v_1 = \frac{k_1^{cat} \cdot MAPKK \cdot M/K_{m1}}{(1 + M/K_{m1} + Mp/K_{m2})}$$

$$Mp \xrightarrow{v_1} Mpp \qquad v_2 = \frac{k_2^{cat} \cdot MAPKK \cdot Mp/K_{m2}}{(1 + M/K_{m1} + Mp/K_{m2})}$$

$$Mpp \xrightarrow{v_3} Mp \qquad v_3 = \frac{k_3^{cat} \cdot MKP3 \cdot Mpp/K_{m3}}{(1 + Mpp/K_{m3} + Mp/K_{m4} + M/K_{m5})}$$

$$Mp \xrightarrow{v_4} M \qquad v_4 = \frac{k_4^{cat} \cdot MKP3 \cdot Mp/K_{m4}}{(1 + Mpp/K_{m3} + Mp/K_{m4} + M/K_{m5})}$$

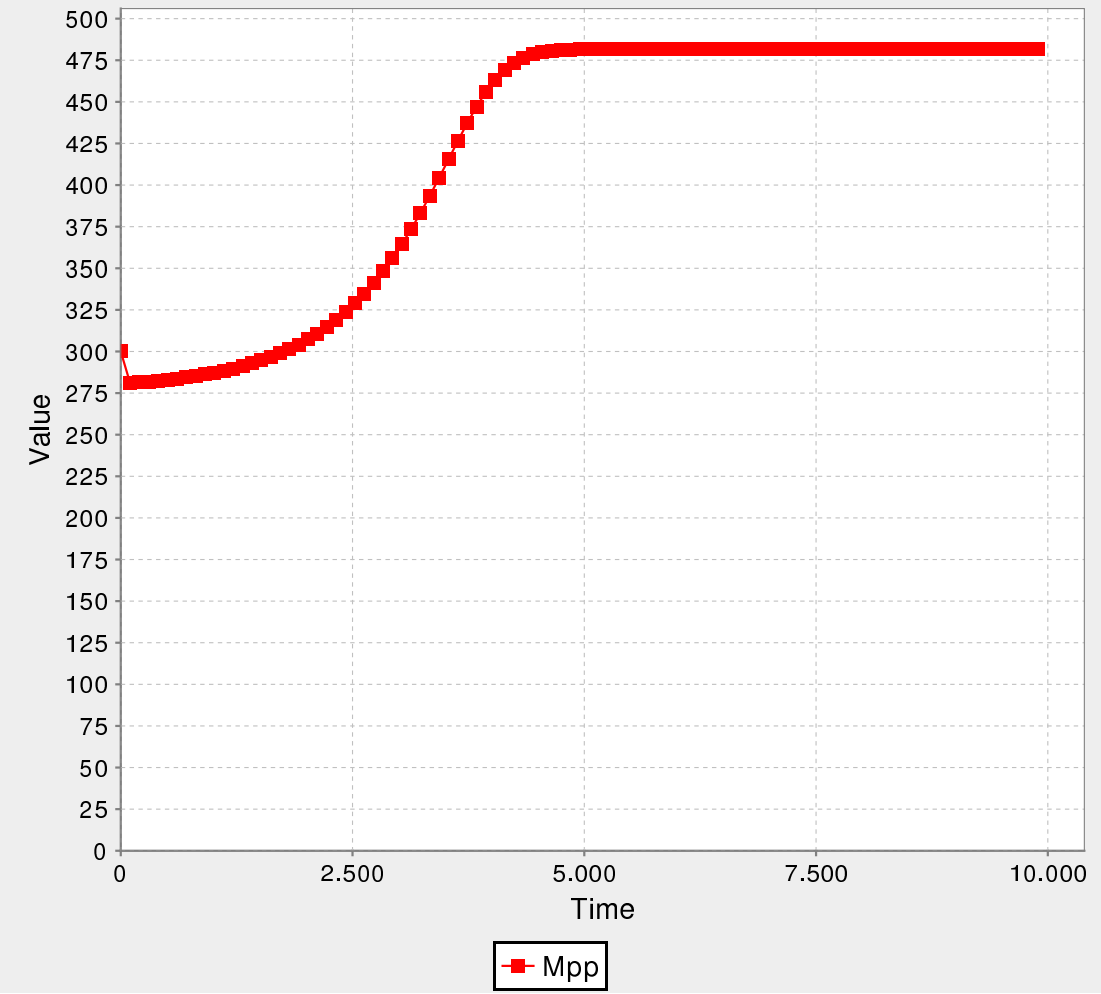
Parameters and relationship with kinetics

$$k_1^{cat} = k_2 = 0.01$$
$$k_2^{cat} = k_4 = 15$$
$$k_3^{cat} = h_2/(1 + h_2/h_3) = 0.084$$
$$k_4^{cat} = h_5 \cdot (1 + h_5/h_6 + h_{-3} \cdot (h_{-4} + h_5)/(h_3 \cdot h_4))^{-1} = 0.06$$

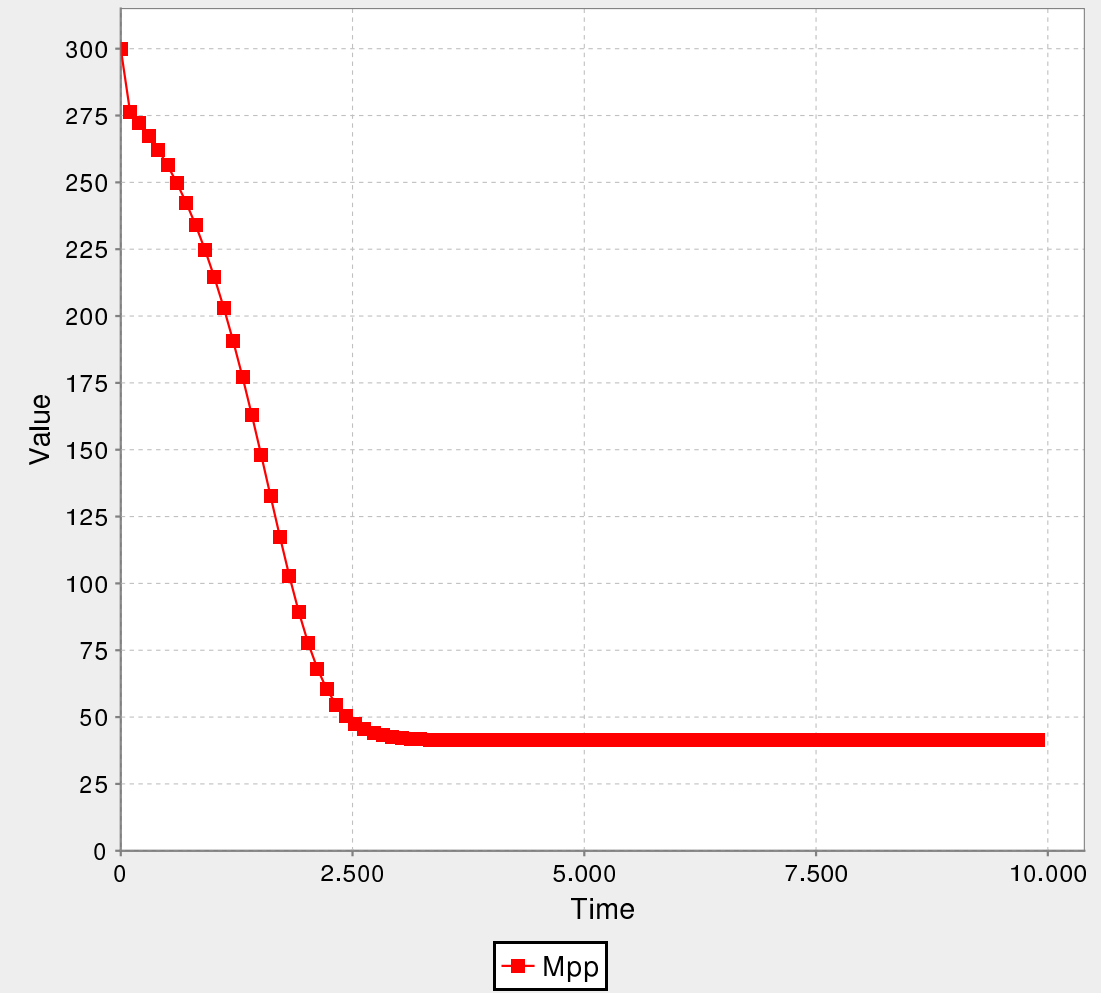
$$M_0 = 200$$
$$Mp_0 = 0$$
$$Mpp_0 = 300$$
$$MAPKK_0 = 50$$
$$MKP3_0 = 100$$

$$K_{m1} = (k_{-1} + k_2)/k_1 = 50$$
$$K_{m2} = (k_{-3} + k_4)/k_3 = 500$$
$$K_{m3} = (h_{-1} + h_2)/(h_1 + h_1 \cdot h_2/h_3) = 22$$
$$K_{m4} = (h_{-4} + h_5) \cdot (h_4 \cdot (1 + h_5/h_6 + h_{-3} \cdot (h_{-4} + h_5)/(h_3 \cdot h_4)))^{-1} = 18$$
$$K_{m5} = (h_6/h_{-6}) = 78$$

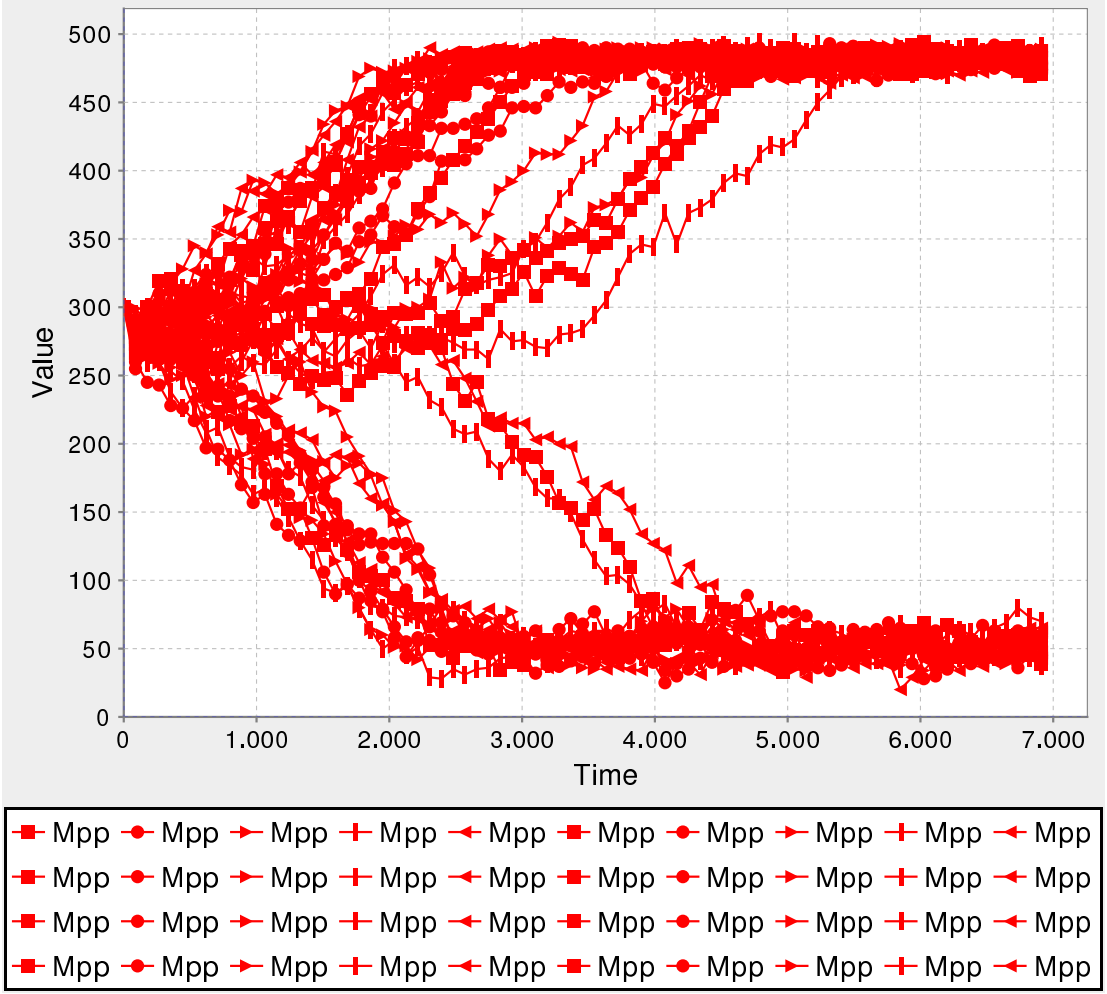
Mpp molecules using nominal parameter values (ODEs)



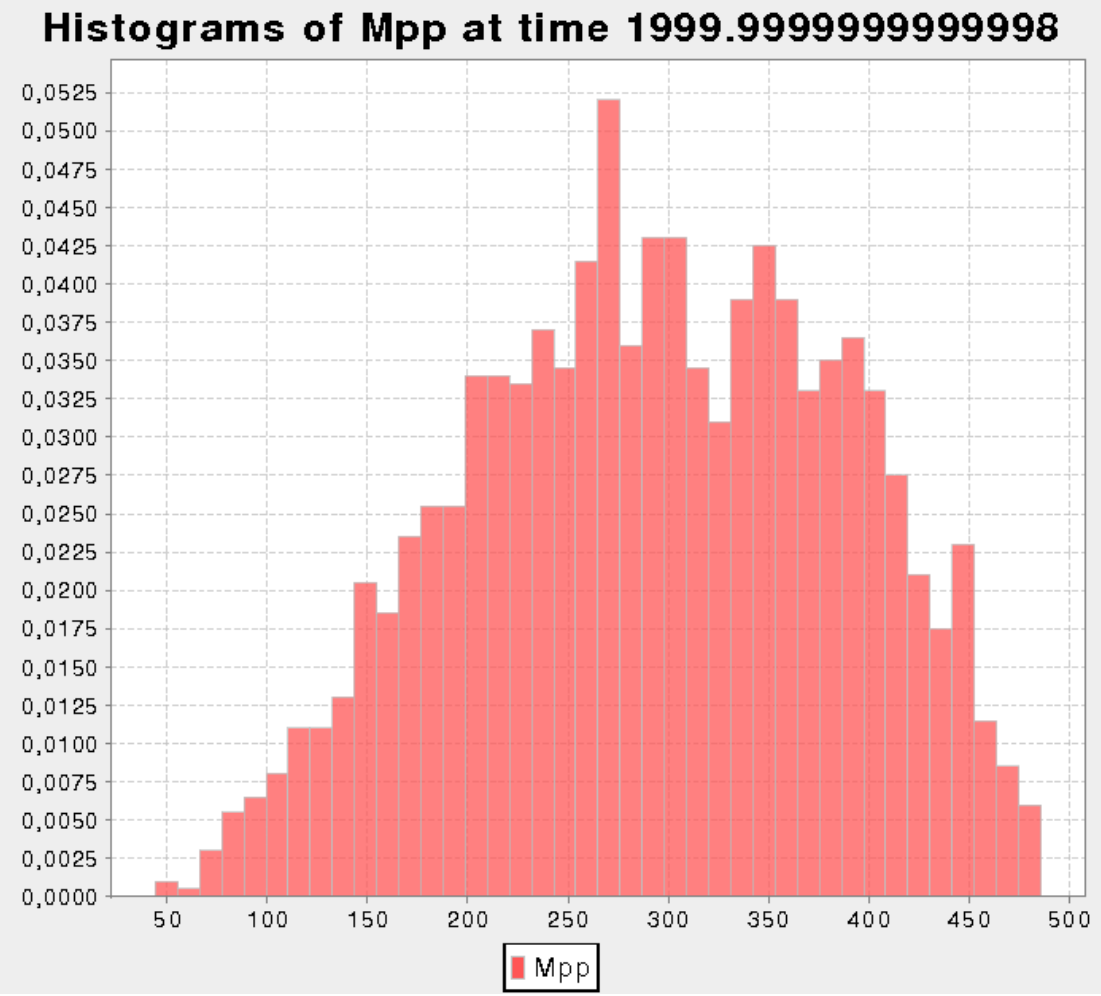
Mpp molecules with MKP3 incremented by 5% (ODEs)



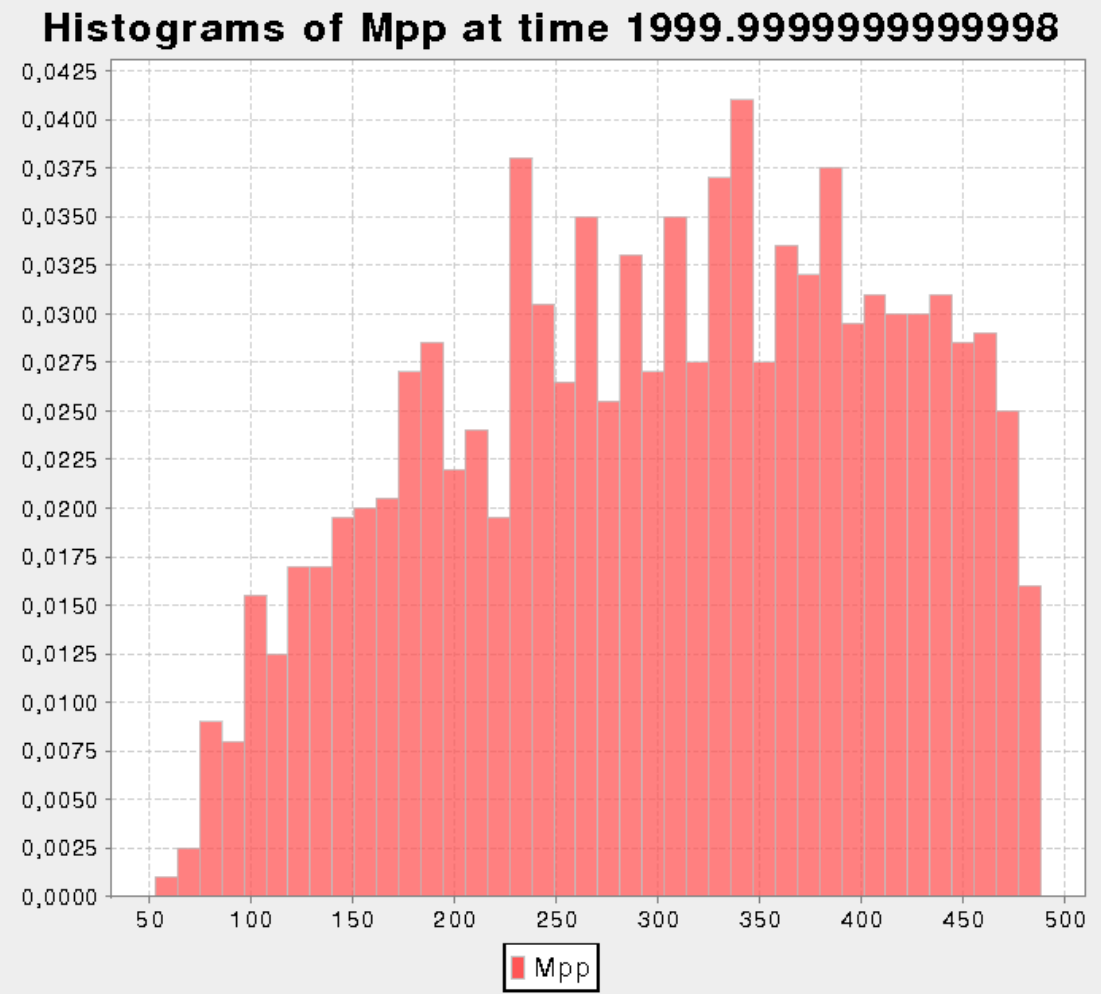
Mpp molecules over 40 runs of the SSA



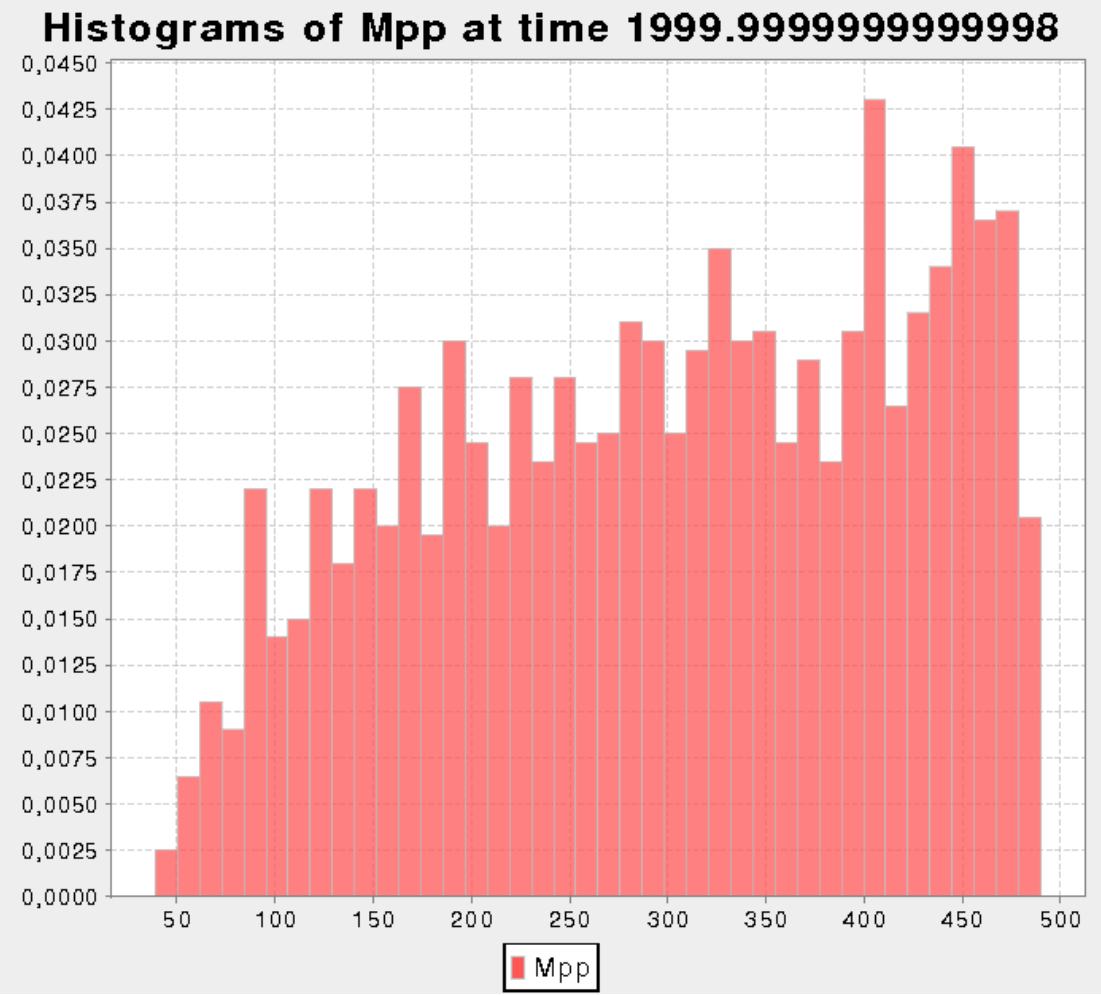
Mpp molecules (2000 SSA runs, enzymes decreased 20%)



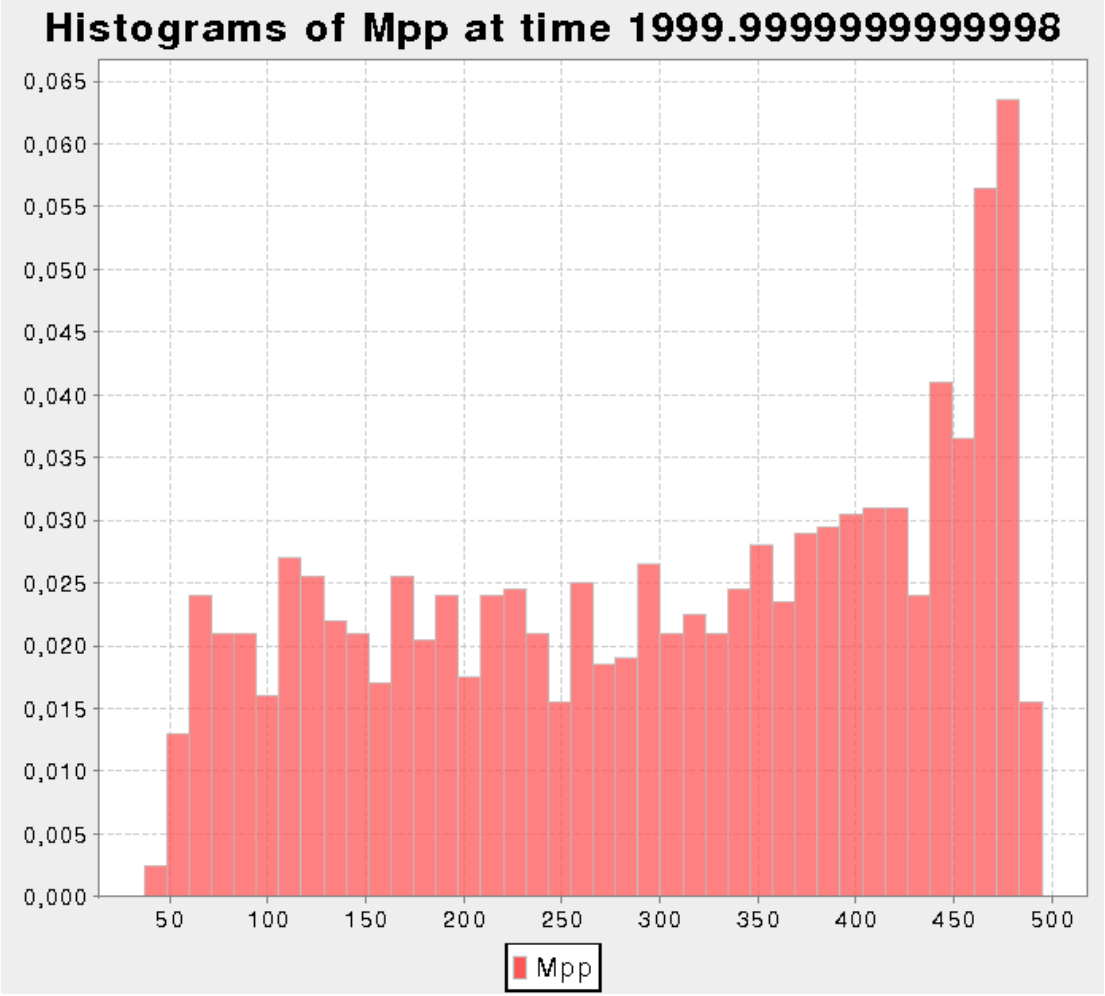
Mpp molecules (2000 SSA runs, enzymes decreased 10%)



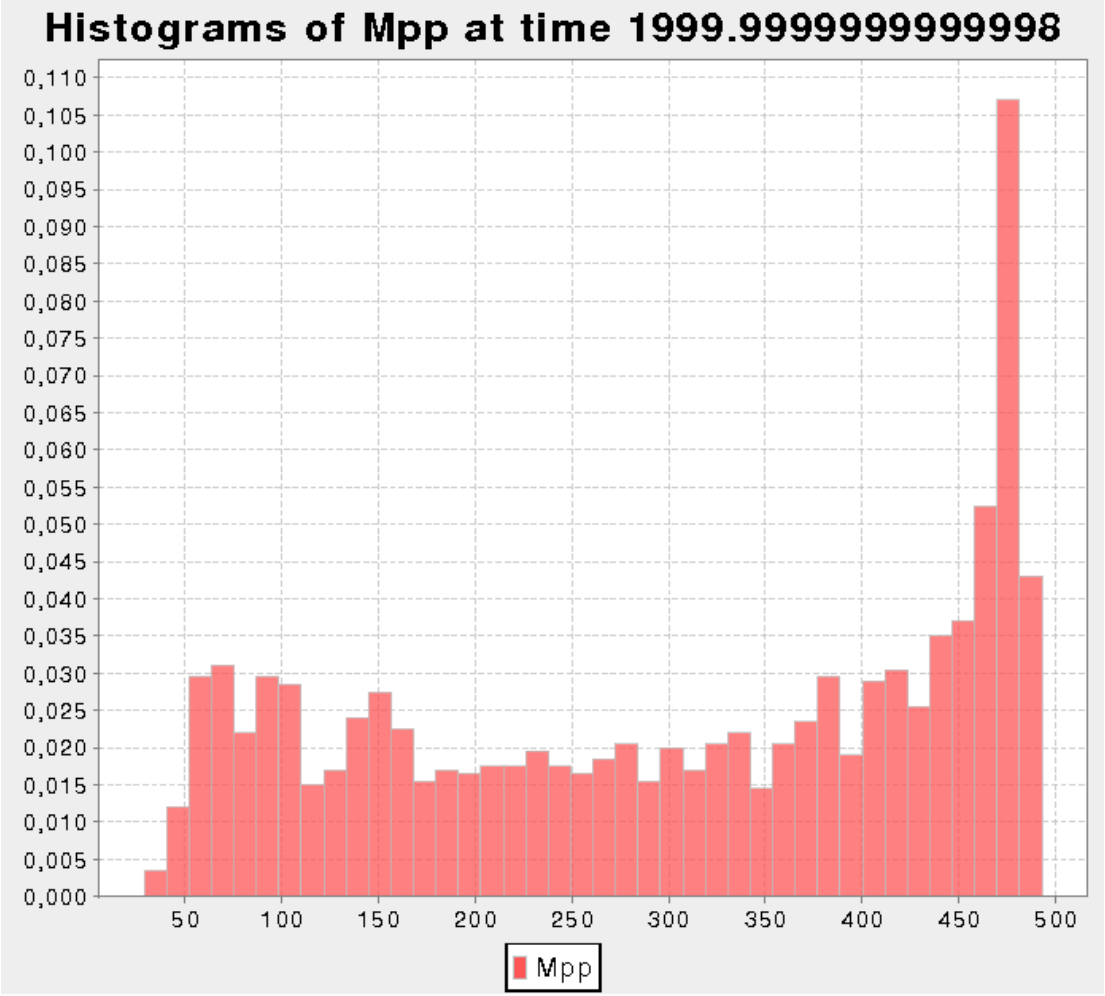
Mpp molecules (2000 SSA runs, enzymes unchanged)



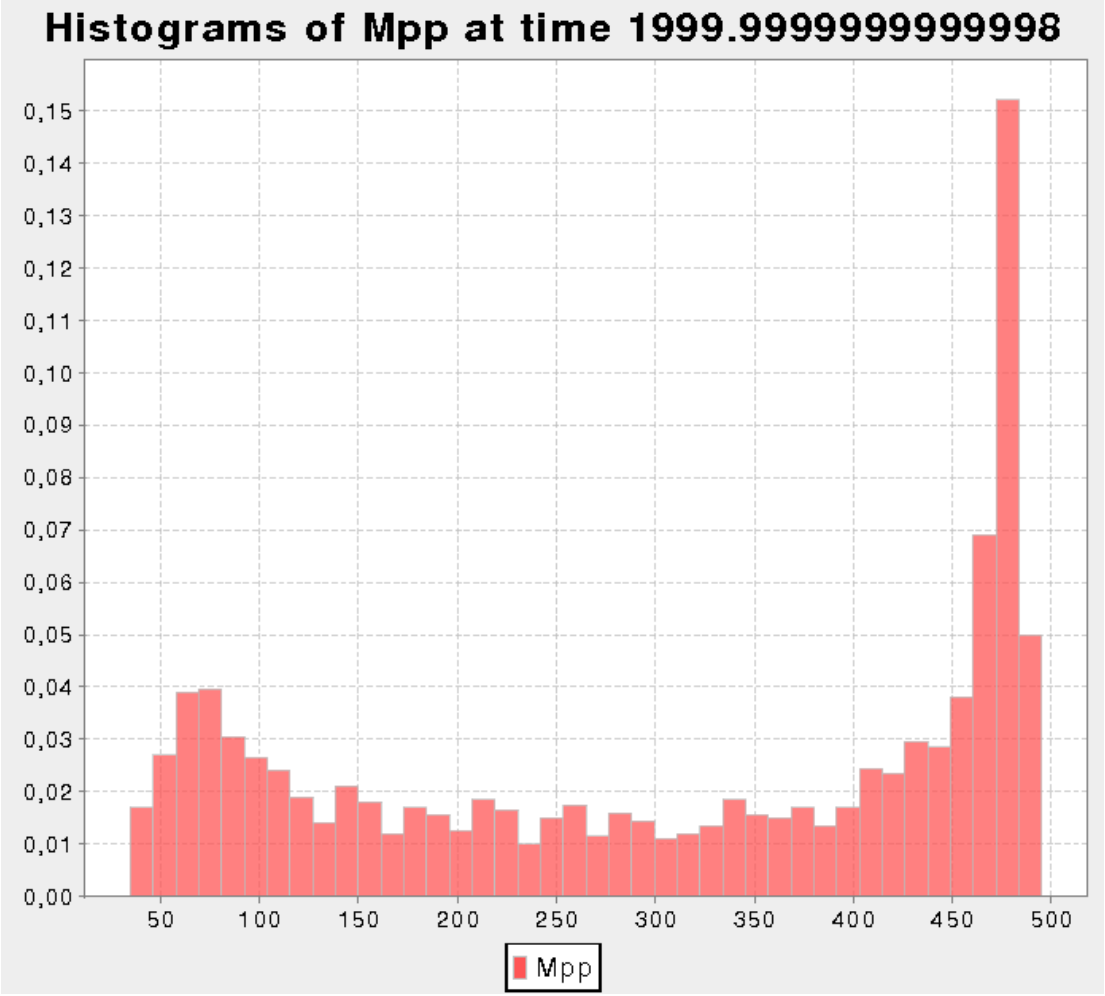
Mpp molecules (2000 SSA runs, enzymes increased 10%)



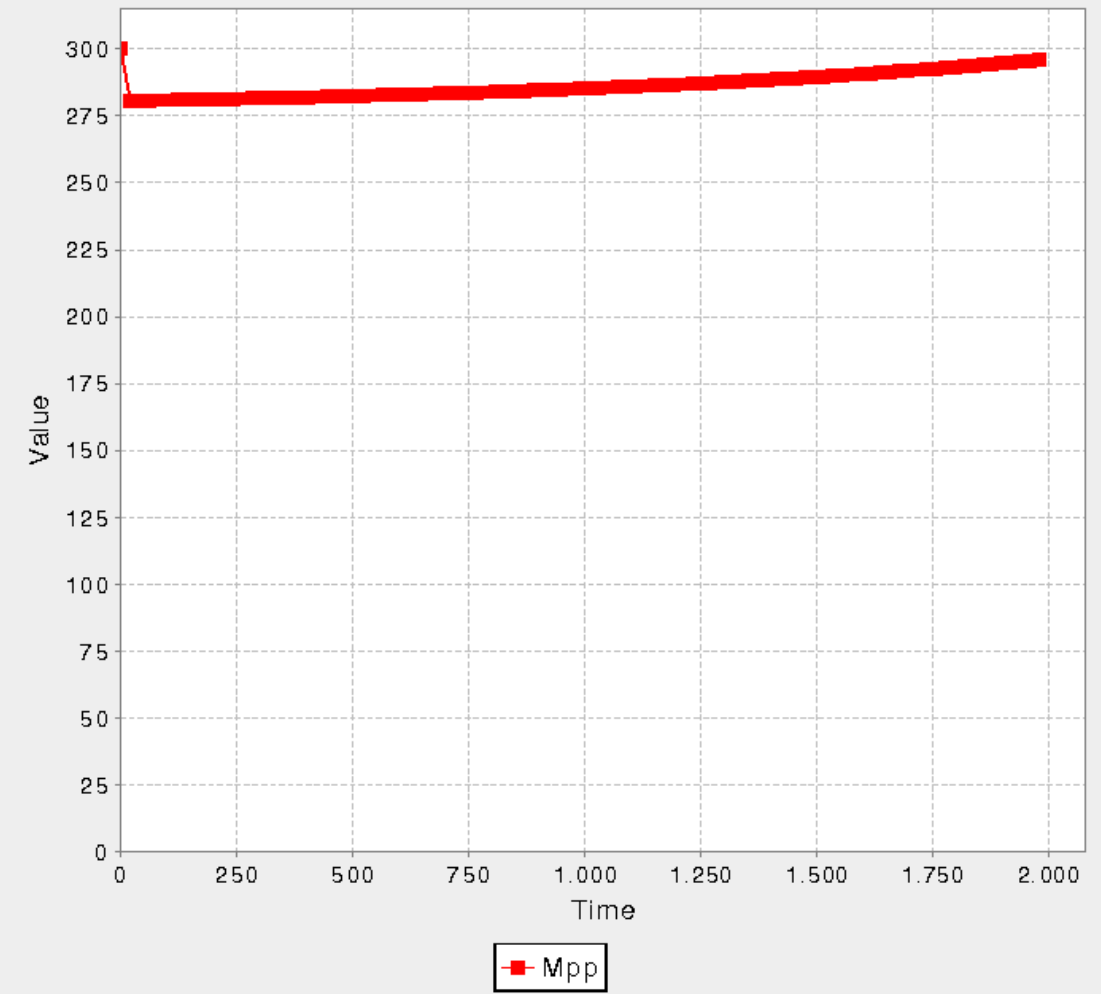
Mpp molecules (2000 SSA runs, enzymes increased 20%)



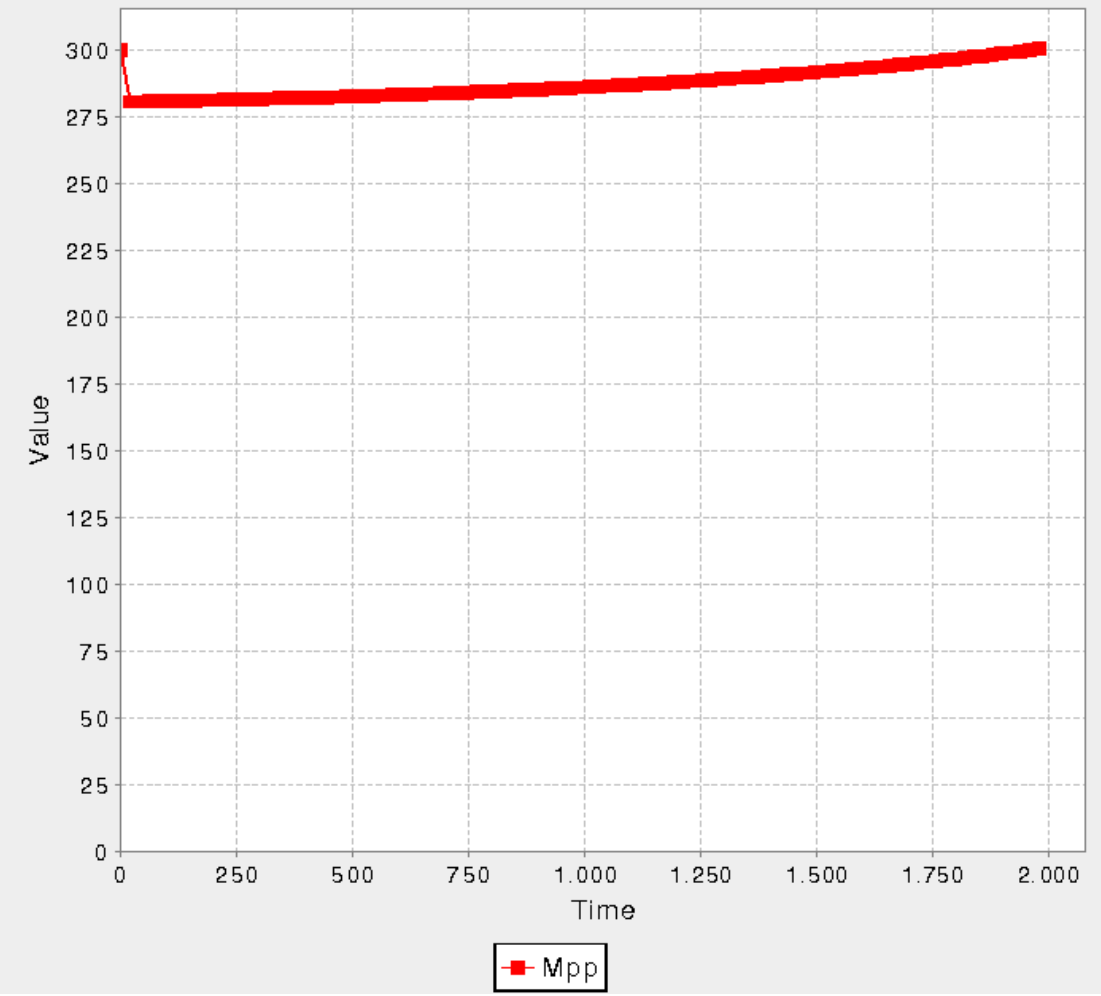
Mpp molecules (2000 SSA runs, enzymes increased 30%)



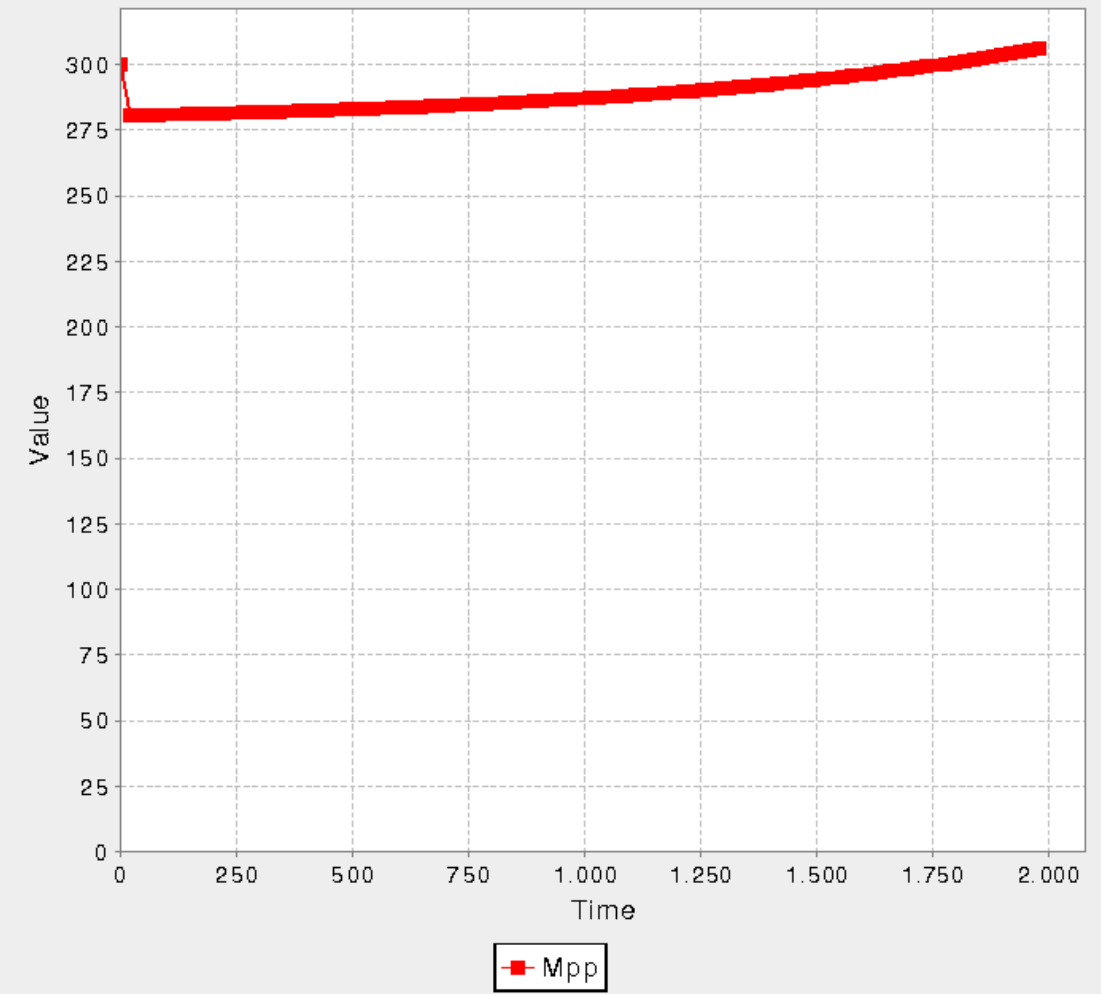
Mpp molecules (ODEs, enzymes decreased 20%)



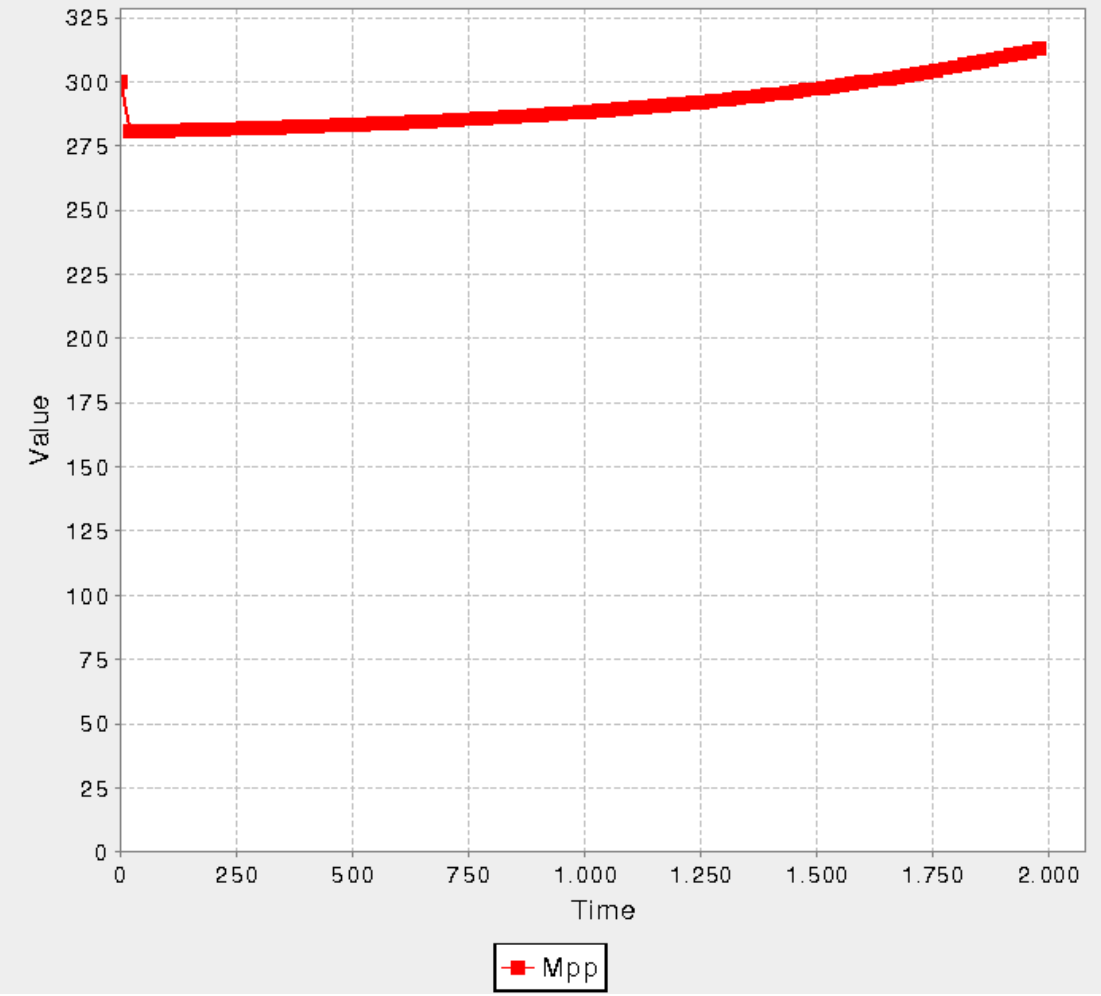
Mpp molecules (ODEs, enzymes decreased 10%)



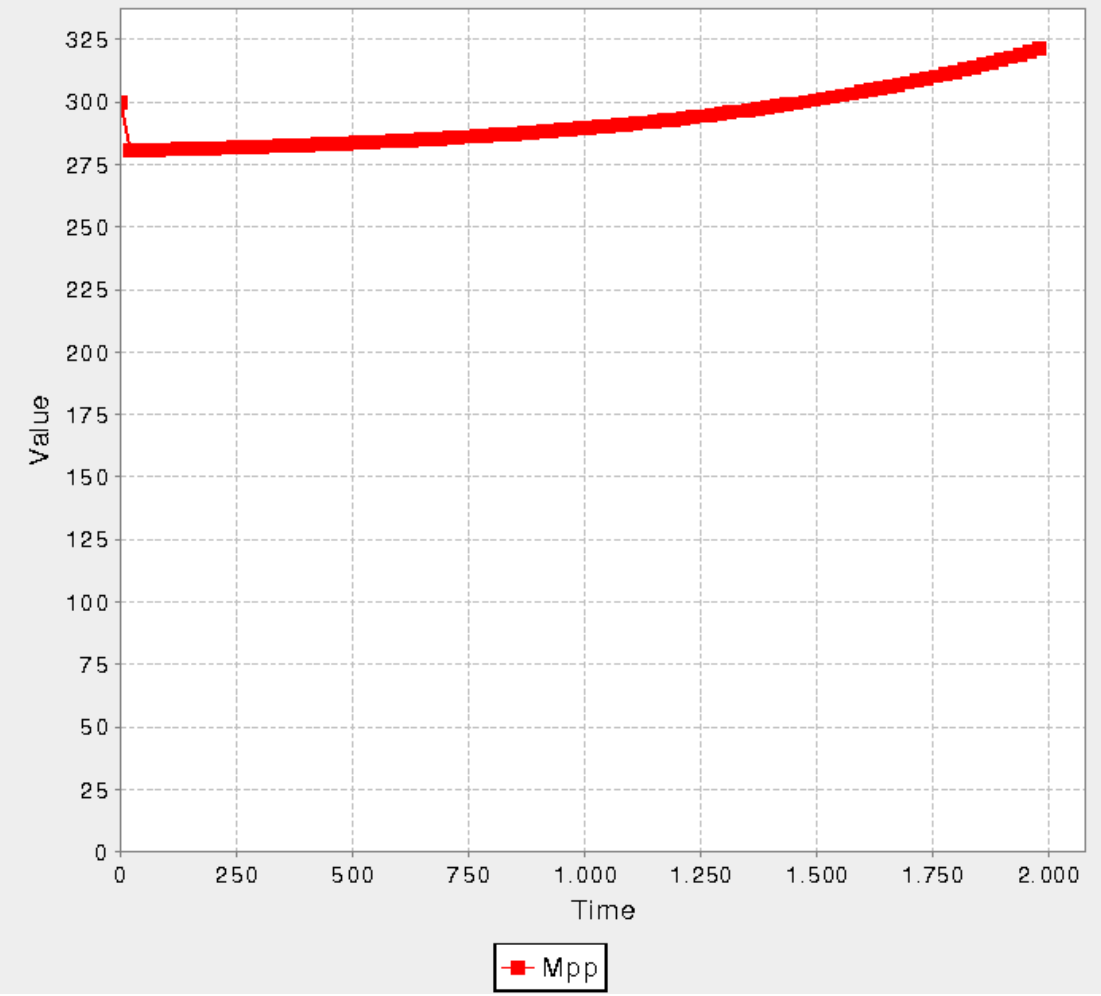
Mpp molecules (ODEs, enzymes unchanged)



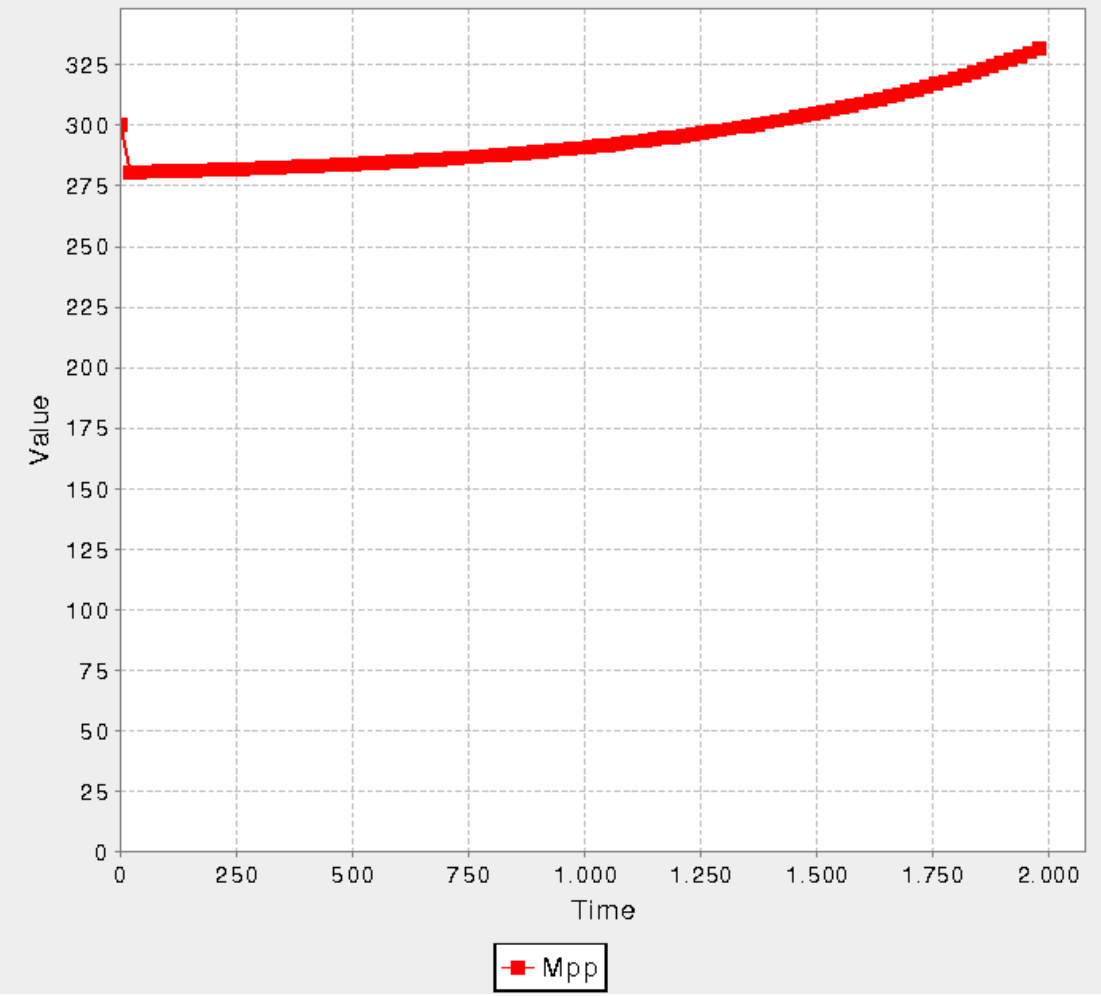
Mpp molecules (ODEs, enzymes increased 10%)



Mpp molecules (ODEs, enzymes increased 20%)



Mpp molecules (ODEs, enzymes increased 30%)



Conclusions

The comparison between deterministic and stochastic approaches to sensitivity analyis 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.