*Appl. Statist.* (2004)
**53**, *Part* 4, *pp.* 557–568

# Tests for cycling in a signalling pathway

T. G. Müller, D. Faller and J. Timmer

*Zentrum für Datenanalyse und Modellbildung, Freiburg, Germany*

and I. Swameye, O. Sandra and U. Klingmüller

*Max-Planck-Institut für Immunbiologie, Freiburg, Germany*

[Received April 2002. Final revision September 2003]

**Summary.** Cellular signalling pathways, mediating receptor activity to nuclear gene activation, are generally regarded as feed forward cascades. We analyse measured data of a partially observed signalling pathway and address the question of possible feed-back cycling of involved biochemical components between the nucleus and cytoplasm. First we address the question of cycling in general, starting from basic assumptions about the system. We reformulate the problem as a statistical test leading to likelihood ratio tests under non-standard conditions. We find that the modelling approach without cycling is rejected. Afterwards, to differentiate two different transport mechanisms within the nucleus, we derive the appropriate dynamical models which lead to two systems of ordinary differential equations. To compare both models we apply a statistical testing procedure that is based on bootstrap distributions. We find that one of both transport mechanisms leads to a dynamical model which is rejected whereas the other model is satisfactory.

*Keywords*: Bootstrap testing procedure; Differential equations; Model selection; Non-nested non-linear models; Signalling pathway

## 1. Introduction

Cells respond to their environment through complex signal transduction pathways that frequently begin at the cell membrane. In response to binding of specific ligands the receptor's cytoplasmic domain transmits the signal across the membrane and triggers a cascade of events in the cell. This can often be described by a series of biochemical reactions which may lead to gene expression.

Current research on signalling pathways and metabolic networks is still mainly dedicated to answering questions concerning the qualitative behaviour of biological systems, e.g. showing the effect of certain biochemical components on a system or investigating interdependences between variables (Bhalla and Iyengar, 1999; Fussenegger *et al.*, 2000). Although leading to a broader understanding in general, to comprehend these biochemical networks fully it is necessary to investigate their dynamical behaviour quantitatively (Koshland, 1998; Campbell, 1999; Zheng and Flavel, 2000; Endy and Brent, 2001; Downward, 2001). This inevitably leads to hypothesis testing to compare different models with statistical tests which are mostly applied under non-standard conditions in the finite sample case.

In this paper we address the question of modelling the dynamical behaviour of a specific signalling pathway representing a wide class of pathways. Starting with only a few assumptions

*Address for correspondence*: J. Timmer, Zentrum für Datenanalyse und Modellbildung, Universität Freiburg, Eckerstrasse 1, 79104 Freiburg, Germany.
E-mail: jeti@fdm.uni-freiburg.de

0035–9254/04/53557

about the underlying dynamics we first resolve the question of cycling, i.e. whether the biochemical reactant which activates the target genes may leave the nucleus to take part again in the signalling process. Afterwards we address the question of modelling the cycling behaviour with statistical tests, comparing two different models which are derived from two different biochemical hypotheses about the exact transport mechanism. In both cases mathematical models of the dynamical behaviour are tested for compatibility with the measured data.

The paper is organized as follows. In Section 2 we describe the biochemical background and formulate the mathematical properties on the basis of *a priori* knowledge. Furthermore we describe the experimental set-up, focusing on questions related to partial observability of the system and on measurement noise. In Section 3 we reformulate the mathematical problem and resolve the question of cycling as a testing problem. Moreover we describe how to deal with special non-standard conditions and present the results from this investigation. In Section 4 we present two candidates of global parametric models which comply with the results from the previous section and test for compatibility with the data.

## 2. The biological model and the experimental set-up

Signalling pathways play a dominant role in information processing within the cell. On binding of messenger molecules to cell surface receptors, signal transmission is initiated within the cell, leading to a cascade of biochemical reactions. In many cases this leads to migration of specific components into the nucleus where target genes are stimulated. A signalling pathway that has been studied in great detail is the Janus kinase–signal transducer and activator of transcription (STAT) pathway (Darnell, 1997; Pellegrini and Dusanter-Fourt, 1997) which can be activated by several receptors. On binding of erythropoietin (Epo) to the Epo receptor (EpoR) several biochemical reactions take place (Klingmüller *et al.*, 1996). In the first step the unphosphorylated STAT5 component is phosphorylated. In the next step, phosphorylated STAT5 molecules form dimers which can then enter the nucleus and stimulate activation of target genes; Fig. 1 (Haspel *et al.*, 1996). Hence, this signalling pathway can be described as a dynamical system consisting of one activation function, the time course of the activated EpoR, and four dynamical variables: unphosphorylated STAT5, phosphorylated STAT5, dimerized STAT5 and STAT5 in the nucleus. Although it is known that STAT5 leaves the nucleus and can be used in future signalling cascades, the timescale of this release is unknown and hence the question of cycling of STAT5 during one signalling cascade arises (Haspel and Darnell, 1999). This would mean that STAT5 leaving the nucleus can be rephosphorylated by the EpoR, thus leading to a cycling behaviour of the STAT5 component.

Unfortunately, no measurement technique is currently available which permits the direct measurement of any system variable. This leads to unobserved components, which is a general problem in the analysis of biochemical systems. Only the activation function of EpoR,

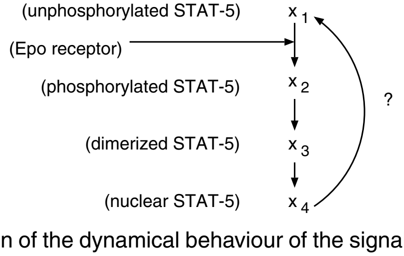

**Fig. 1.** Schematic representation of the dynamical behaviour of the signalling pathway: the presence of the cycling $x_4 \longrightarrow x_1$ is investigated

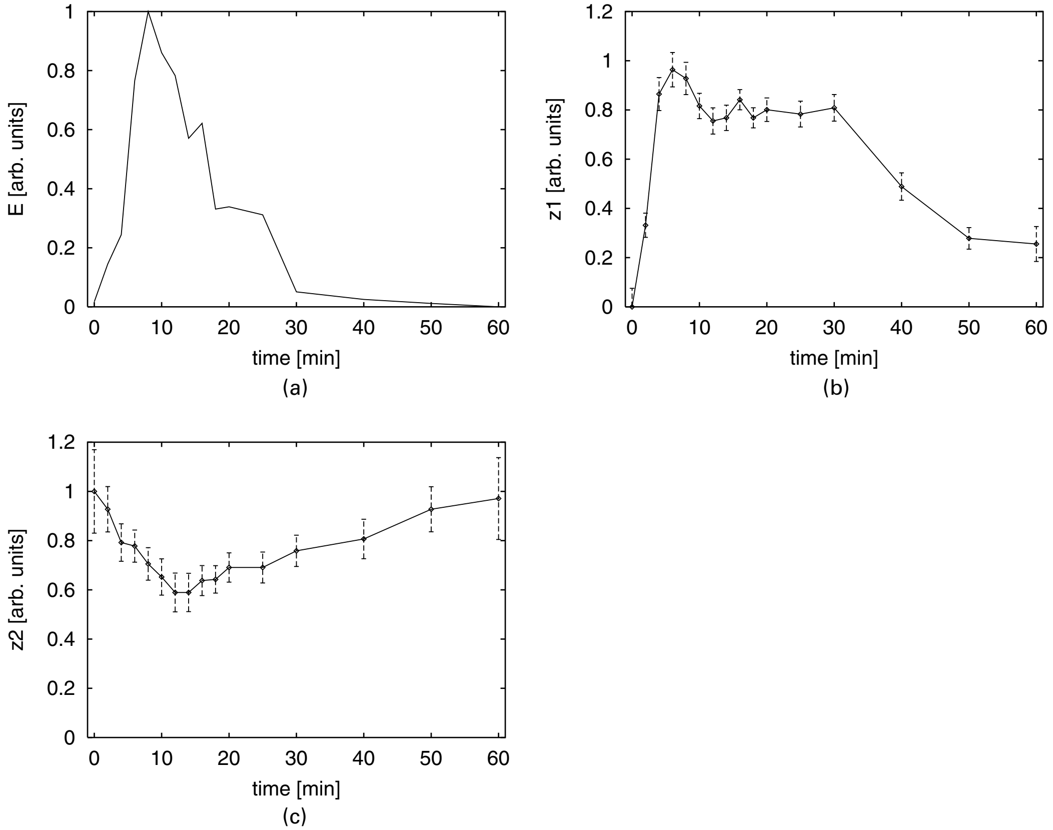

**Fig. 2.** Measured time series of a typical data set: (a) amount of activated EpoR, (b) amount of phosphorylated STAT5 and (c) amount of total STAT5 (for visualization purposes, we connected the data points with lines)

the total amount of STAT5 in the cytoplasm and the amount of phosphorylated STAT5 in the cytoplasm can be measured over time. Even worse, all measurements are only proportional to the true amount, introducing additional unknown nuisance scaling parameters. Additionally, measurements are corrupted by observational noise. Hence, we deal with partially observed noisy data. A detailed description of the experimental measurement techniques that are used can be found in Swameye *et al.* (2003). Basically, six biological samples are taken at 16 different time points, and the amounts of STAT5, phosphorylated STAT5 and EpoR are measured by immunoblotting followed by chemiluminescence measurements. Since this measurement technique involves a counting process with high intensity, the error distribution can be approximated by a Gaussian distribution. In Fig. 2 we display a typical measurement of all three observation functions. Error bars in Fig. 2(a) have been omitted since EpoR activation serves as an input function to the system and is not modelled in a parametric way.

To increase the power of the test it is possible to use independent experiments. Then, the dynamical parameters governing the biochemical reactions of STAT5 are the same for every experiment whereas the nuisance parameters differ.

To simplify the mathematical modelling we shall use the following notation: $E(t)$ is the time course of the activated EpoR at time $t$, serving as the input function to the system. $x_1(t)$, $x_2(t)$

and $x_3(t)$ are the amounts of unphosphorylated, phosphorylated and dimerized STAT5 in the cytoplasm and $x_4(t)$ is the amount of STAT5 in the nucleus.

With *a priori* knowledge from biochemistry some properties of the system are known and we can derive mathematical descriptions of some biochemical reactions.

(a) The phosphorylation process of unphosphorylated STAT5 $x_1(t)$ which depends on the activated EpoR leads to a reduction in unphosphorylated STAT5 which is described by $[p_1 x_1 E]$.
(b) Release of STAT5 from the nucleus is modelled by $[p_4 x_4]$. This is a first-order approximation for many other transport mechanisms and therefore a rather general approach.
(c) The observation function of the total amount of phosphorylated STAT5 in the cytoplasm is $z_1 = p_5(x_2 + x_3)$.
(d) The observation function of the total amount of STAT5 in the cytoplasm is $z_2 = p_6(x_1 + x_2 + x_3)$.
(e) The observation function of activated EpoR is $z_3 = p_7 E$.
(f) The initial conditions are $x_2(0) = x_3(0) = x_4(0) = 0$, whereas $x_1(0)$ must be estimated from the data.
(g) Conservation of mass requires $\Sigma_i x_i = C =$ constant.
(h) All parameters are positive.

These properties are quite general and apply to many signalling pathways. Hence, the following analysis is applicable to other systems.

It is important to note that in this first investigation we start from basic assumptions since we do not assume any knowledge about a mathematical model of the dimerization process ($x_2 \to x_3$) or the transport of dimers into the nucleus ($x_3 \to x_4$).

## 3. Mathematical modelling and the formulation of the testing problem

Using *a priori* knowledge from Section 2, we obtain the following ordinary differential equations and observation functions:

$$\begin{aligned} \dot{x}_1 &= -p_1 x_1 E + p_4 x_4, \\ z_1 &= p_5(x_2 + x_3), \end{aligned} \tag{1}$$

$$\begin{aligned} \dot{x}_2 &= p_1 x_1 E - f(x_2), \\ z_2 &= p_6(x_1 + x_2 + x_3), \end{aligned} \tag{2}$$

$$\begin{aligned} \dot{x}_3 &= f(x_2) - g(x_3), \\ z_3 &= p_7 E =: D, \end{aligned} \tag{3}$$

$$\dot{x}_4 = g(x_3) - p_4 x_4, \tag{4}$$

with $p_1$ and $p_4$ unknown dynamical parameters, $p_5, p_6$ and $p_7$ unknown nuisance parameters and $f(x_2)$ and $g(x_3)$ unknown functions describing the biochemical reactions of dimerization and transport into the nucleus. Parameters $p_2$ and $p_3$ will later be used to parameterize functions $f$ and $g$.

Using conservation of mass and the starting conditions we obtain

$$x_1(t) = \frac{z_2(t)}{p_6} - \frac{z_1(t)}{p_5}, \tag{5}$$

$$\dot{x}_1(t) = \frac{\dot{z}_2(t)}{p_6} - \frac{\dot{z}_1(t)}{p_5}, \tag{6}$$

$$x_4(t) = \frac{1}{p_6}\{z_2(0) - z_2(t)\}, \tag{7}$$

and differential equation (1) can be rewritten as

$$\frac{\dot{z}_2(t)}{p_6} - \frac{\dot{z}_1(t)}{p_5} = -p_1\left\{\frac{z_2(t)}{p_6} - \frac{z_1(t)}{p_5}\right\}\frac{z_3(t)}{p_7} + \frac{p_4}{p_6}\{z_2(0) - z_2(t)\}. \tag{8}$$

Testing for cycling now amounts to the question whether parameter $p_4$ is different from 0. We interpret the measured time series as realizations of random variables and reformulate this question as a problem of model selection between two different linear regression models A and B defined as follows: model A,

$$\begin{aligned} Y_i &= f_{\mathrm{A}}(X_{1i}, X_{2i}, X_{3i}) \qquad (9)\\ &= K_1 X_{1i} - K_2 X_{2i} + K_1 K_2 X_{3i}; \qquad (10) \end{aligned}$$

model B,

$$\begin{aligned} Y_i &= f_{\mathrm{B}}(X_{1i}, X_{2i}, X_{3i}, X_{4i}) \qquad (11)\\ &= K_1 X_{1i} - K_2 X_{2i} + K_1 K_2 X_{3i} - K_3 X_{4i} + K_4 \qquad (12) \end{aligned}$$

with

$$Y_i = \dot{z}_2(t_i), \tag{13}$$

$$\left.\begin{aligned} X_{1i} &= \dot{z}_1(t_i),\\ X_{2i} &= z_2(t_i)\,E(t_i),\\ X_{3i} &= z_1(t_i)\,E(t_i),\\ X_{4i} &= z_2(t_i) \end{aligned}\right\} \tag{14}$$

and parameters

$$\left.\begin{aligned} K_1 &= p_6/p_5,\\ K_2 &= p_1,\\ K_3 &= p_4,\\ K_4 &= p_4 p_6\, x_1(0). \end{aligned}\right\} \tag{15}$$

To discriminate between models A and B, we would like to apply the test statistic $L = 2(L_{\mathrm{B}} - L_{\mathrm{A}})$ originating from the likelihood ratio test with $L_{\mathrm{A}}$ and $L_{\mathrm{B}}$ being defined as

$$L_{\mathrm{A}} = \sum_i -\frac{1}{2}\log(2\pi) - \log(\sigma_i) - \frac{\{Y_i - f_{\mathrm{A}}(X_{1i}, X_{2i}, X_{3i})\}^2}{2\sigma_i^2}, \tag{16}$$

$$L_{\mathrm{B}} = \sum_i -\frac{1}{2}\log(2\pi) - \log(\sigma_i) - \frac{\{Y_i - f_{\mathrm{B}}(X_{1i}, X_{2i}, X_{3i}, X_{4i})\}^2}{2\sigma_i^2}, \tag{17}$$

$i = 1, \ldots, 16$. In our setting the observation error $\varepsilon_i$ is assumed to be distributed according to

$$\varepsilon_i = N(0, \sigma_i^2), \qquad i = 1, \ldots, 16, \tag{18}$$

and the distribution of the test statistic is known analytically. However, if the distribution of the test statistic is not known analytically, it can be simulated by a parametric bootstrap, as will be shown in what follows. We assume that all observations are corrupted by white noise and obtain

$$\sigma_i^2 = \sigma_{Y_i}^2 + \sum_j \left( \frac{\partial f}{\partial X_j} \right)^2 \sigma_{X_{ji}}^2. \tag{19}$$

The distribution of the test statistic $L$ critically depends on the amount of information that can be drawn from the system. Before we deal with the actual setting of the recorded experimental data, we shall investigate other realistic experimental settings and their influence on the distribution of the test statistic. First we assume that $X_{ji}$ are realizations of independent random variables and that only realizations of $Y$ are corrupted by observational noise.

Models A and B originate from the same model class and are nested. Therefore, under standard conditions, we would assume that $L = 2(L_{\mathrm{B}} - L_{\mathrm{A}})$ is distributed as a $\chi^2$-distribution with $\Delta\mathrm{df} = \mathrm{df}_{\mathrm{B}} - \mathrm{df}_{\mathrm{A}}$ degrees of freedom where $\Delta\mathrm{df}$ is the difference in degrees of freedom and $\mathrm{df}_M$ denotes the degrees of freedom of model $M$. For the case with one experiment we obtain $\Delta\mathrm{df} = 2$ since model B has additional parameters $K_3$ and $K_4$; see Fig. 3, curve A. In fact, owing to the additional constraints for all parameters, $K_j \geqslant 0$, this distribution is different from the standard distribution (Self and Liang, 1987). The constraints lead to a distribution consisting of a mixture of $\chi^2$-distributions so that

$$L \sim \tfrac{1}{4}(\chi_0^2 + 2\chi_1^2 + \chi_2^2); \tag{20}$$

see Fig. 3, curve B. The distribution incorporating the inequality constraints leads to a new 95% quantile $q_{0.95} = 4.23$. It is noteworthy to recall the 95% quantile under standard conditions, which is $q_{0.95} = 5.99$. Hence the constraints are important to consider to avoid a loss of power of the test.

In the increasingly more realistic settings that are investigated in what follows, the error structure of the random variables $X_j$ entering the test statistic exhibits an unknown dependence structure. This error structure is increasingly complicated as the model assumptions become increasingly more realistic. To the best of our knowledge, it is then not possible to describe the distribution of $L$ analytically, since the probability model underlying the standard likelihood as a test statistic does not hold. Nevertheless this likelihood ratio test statistic remains a reasonable test statistic, if its distribution is calculated by using a parametric bootstrap. Hence, a parametric bootstrap procedure was applied to obtain the distribution of the likelihood ratio test statistic under these conditions. For this, the parameter values based on estimates with

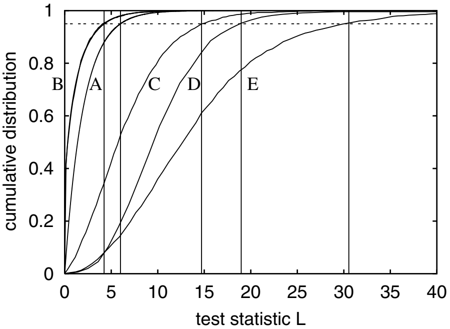

**Fig. 3.** Cumulative distributions of the test statistic for various assumptions about the experimental conditions: curve A corresponds to the standard likelihood ratio test, curve B includes the effects of positivity constraints on the parameters, the implications for the test statistic resulting from the dynamical nature of the underlying system assuming $X_{ji}$ can be observed is given by curve C, curve D shows the resulting distribution of the test statistic if only $z_i$ and their temporal derivatives are accessible and curve E finally shows the resulting distribution if only the variables $z_i$ are observed (included in the figure is the 0.95 significance level and the corresponding critical values)

the smaller model were used to simulate the system dynamics. Then, at each of the original 16 time points, Gaussian observational noise was added to the simulated system dynamics. These bootstrap samples were then used to simulate the distribution of the test statistic.

In the first setting, we investigate the test statistic, assuming that the variables $X_{jl}$, equation (14), can be observed directly. This leads to the error-in-variables problem. The result is displayed in Fig. 3, curve C. The new conditions have a huge effect on the test statistic with the new 95% quantile being $q_{0.95} = 14.7$. In the next setting, we assume that only time series $z_k(t)$ and their temporal derivatives are directly accessible. Then, equations (13) and (14) may be used to compute the realizations of the $X_j$. This leads to a new distribution of the test statistic due to the non-linear transformation, equation (14). Again, we simulate this distribution, which is displayed in Fig. 3, curve D. We obtain a value of $q_{0.95} = 19.0$. Finally, we investigate the experimental setting under which the data analysed were recorded, i.e. observations are restricted to $z_1$, $z_2$ and $z_3$ in equations (1)–(3). In this case temporal derivatives are not accessible by measurement and must be estimated from the data. Here, we use a method that is based on splines (Hanke and Scherzer, 2001) which estimates derivatives based on the coefficients of the local polynomials of the fitted spline. In this case, the errors of random variables $Y, X_1, \ldots, X_4$ follow complex dependence structures and beforehand it is difficult to predict the effect on the test statistic. Again, we use a simulation to compute the distribution of the test statistic $L$ which is displayed in Fig. 3, curve E, and yields $q_{0.95} = 30.5$.

Fig. 3 summarizes the effect of different simplifying assumptions in deriving the test statistic. It is important to note that inference based on all except the final full procedure is prone to produce false positive results. When using the data set to calculate the value of the test statistic, we obtain a value of $L = 78.7$ which corresponds to a $p$-value of $p < 10^{-4}$. Therefore we conclude that model A is not sufficient to describe the data. In Figs 4(a) and 4(b), a typical fit for both models for the data set is shown. Model A cannot fit the data, whereas model B only misrepresents some time points at the beginning.

Model selection between models A and B may be also resolved by approximating the probability distribution of estimated parameter $p_4$ with a bootstrap procedure (Efron, 1982; Hinkley, 1988; Efron and Tibshirani, 1993). The Gaussianity of the error distribution, as motivated from the measurement process, applies only to the observed variables, $z_1$, $z_2$ and $z_3$. Variables $Y$, $X_1$, $X_2$ and $X_3$ are computed by multiplication or temporal differentiation of the time series of $z_1$, $z_2$ and $z_3$ and the distribution of these random variables is no longer Gaussian. To the best

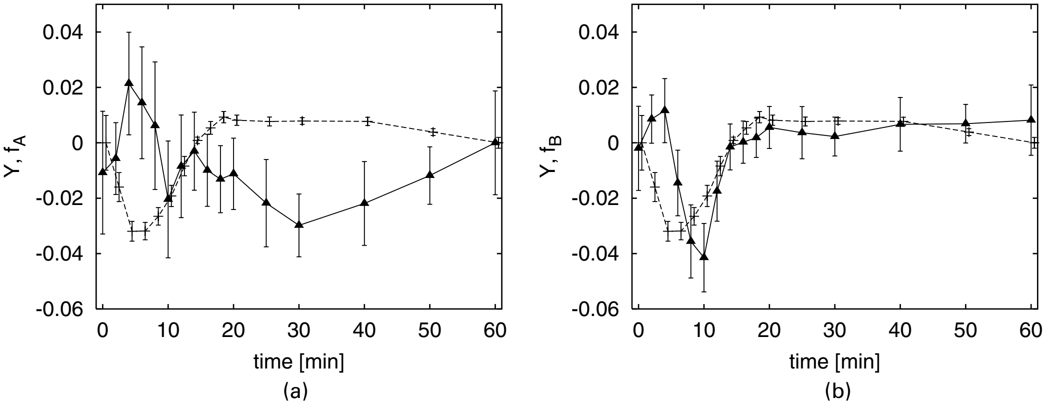

**Fig. 4.** Best fit for two different linear regression approaches $f_A$ and $f_B$ of equations (10) and (11) (———) for data $Y$ (-------): (a) smaller model without cycling; (b) larger model with cycling

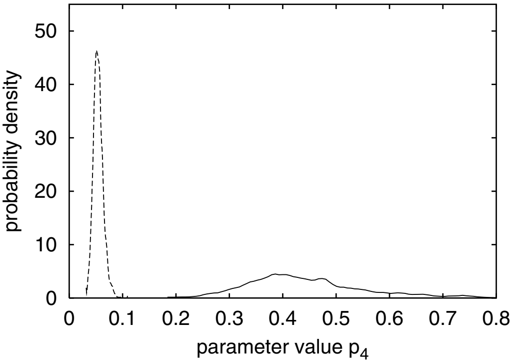

**Fig. 5.** Approximated probability density for the estimated dynamical parameter $p_4$: -------, under the null hypothesis; ———, with a bootstrap procedure with $10^4$ bootstrap samples

of our knowledge, there is no possibility of constructing a maximum likelihood estimator for parameters $K_1, \ldots, K_4$ owing to the unknown distribution of the random variables $Y$, $X_1$, $X_2$ and $X_3$. Nevertheless, the distribution of the parameters is estimated by minimizing $L_\mathrm{A}$ or $L_\mathrm{B}$. To investigate the distribution of the parameters that are computed in this way under the null hypothesis, we use a parametric bootstrap as described above.

Drawing $10^4$ different bootstrap samples from the original data set and computing parameter estimates for every bootstrap sample, we obtain an approximation of the probability distribution for $p_4$ as displayed in the full curve in Fig. 5. It is important to note that using this distribution to test whether $p_4$ is significantly greater than 0 does produce false positive results. Our simulation reveals that the estimate of $p_4$ under the null hypothesis is biased; see the broken curve in Fig. 5. Therefore, when applying bootstrap procedures in such cases, it is necessary to simulate the distribution of parameter estimates under the null hypothesis explicitly. Nevertheless, if we compare both distributions of $p_4$, under the null hypothesis and with the actual data set, in our case we may conclude that parameter $p_4$ is significantly greater than the expected value under the null hypothesis. This confirms our results from the previous analysis of the test statistic $L$.

## 4. Dynamical modelling

To understand the dynamical behaviour of the signalling pathway in general, it is necessary to find an appropriate mathematical description for the temporal evolution of all dynamical variables. Hence, we look for a dynamical model in the form of a deterministic differential equation. Such a model is helpful for gaining insight into the hidden mechanisms of the pathway and for designing future experiments since it can be used to predict the behaviour of the system under perturbations. Possible alterations can be analysed theoretically and numerically in a fast and cheap manner.

Naturally, all deterministic differential equations to model the dynamical behaviour are only approximations to the truth. However, an analysis of the system with *a priori* knowledge, e.g. from biochemistry, normally leaves only a few candidate models which can be compared with model selection procedures.

From the previous investigations the model class is already reduced to models with cyclic behaviour. In what follows we use *a priori* biochemical knowledge to model the dimerization process and the transport into the nucleus mathematically. This will enable us later to focus on comparing different models of cycling. The dimerization process ($x_2 \rightarrow x_3$) is modelled as the reduction of phosphorylated STAT5 ($x_2$) and the production of dimer ($x_3$) by $[p_2x_2^2]$. The transport into the nucleus ($x_3 \to x_4$) is described as the reduction of dimer ($x_3$) and production of nuclear STAT5 ($x_4$) by $[p_3x_3]$. However, there are two possible transport mechanisms governing the export to the cytoplasm ($x_4 \to x_1$) about which no reliable information is available.

(a) A compartmental approach: reduction of nuclear STAT5 ($x_4$) and production of unphosphorylated STAT5 ($x_1$) are proportional to the amount of nuclear STAT5 and can be modelled by $[p_4x_4]$.

(b) A model with delay: reduction of nuclear STAT5 ($x_4$) and production of unphosphorylated STAT5 ($x_1$) are exactly the amount of dimerized STAT5 ($x_3$) that entered the nucleus at time $t-\tau$. This mechanism can be modelled by $[p_3x_3(t-\tau)]$ and represents the idea that dimerized STAT5 binds to the deoxyribonucleic acid for a certain time, is dephosphorylated quickly and may then enter the cytoplasm again.

It is important to note that the discrimination between these parametric models of the transport mechanisms (a) or (b) with the help of the measured data offers the possibility to gain information about a mechanism which currently cannot be investigated directly by measurements. In what follows, terms from model $M_i$ ($i=1,2$) will be denoted as $[\,]_i$ in the mathematical description and the differential equations and observation equations are

$$\begin{aligned} \dot{x}_1 &= -p_1x_1E[+p_4x_4]_1[+p_3x_3(t-\tau)]_2, \\ z_1 &= p_5(x_2+x_3), \end{aligned} \tag{21}$$

$$\begin{aligned} \dot{x}_2 &= p_1x_1E - p_2x_2^2, \\ z_2 &= p_6(x_1+x_2+x_3), \end{aligned} \tag{22}$$

$$\begin{aligned} \dot{x}_3 &= -p_3x_3 + p_2x_2^2, \\ z_3 &= p_7E := D, \end{aligned} \tag{23}$$

$$\dot{x}_4 = p_3x_3[-p_4x_4]_1[-p_3x_3(t-\tau)]_2. \tag{24}$$

Owing to the observation function of the system which allows only limited measurements of the dynamical behaviour, not all dynamical and nuisance parameters as well as the starting value $x_1(0)$ can be extracted from the data. Therefore it is not possible to identify all the parameters of the system. This non-identifiability often occurs in dynamical models with unobserved components (Chappell *et al.*, 1990; Müller *et al.*, 2002). It is possible to obtain a dynamical system with fewer parameters consisting of the same observation function. Therefore it is necessary to apply a transformation of the differential equation to ensure that all transformed parameters are identifiable.

Using the transformation

$$\left.\begin{aligned} v_i &= p_2x_i, \\ r_1 &= p_1/p_7, \\ r_3 &= p_3, \end{aligned}\right\} \tag{25}$$

$$\left.\begin{aligned} r_4 &= p_4, \\ r_5 &= p_5/p_2, \\ r_6 &= p_6/p_2, \end{aligned}\right\} \tag{26}$$

we obtain the following differential equations:

$$\begin{aligned} \dot{v}_1 &= -r_1v_1D[+r_4v_4]_1[+r_3v_3(t-\tau)]_2, \\ y_1 &= r_5(v_2+v_3), \end{aligned} \tag{27}$$

$$\begin{aligned}\dot{v}_2 &= r_1 v_1 D - v_2^2,\\ y_2 &= r_6(v_1 + v_2 + v_3),\end{aligned} \tag{28}$$

$$\begin{aligned}\dot{v}_3 &= -r_3 v_3 + v_2^2,\\ y_3 &= D,\end{aligned} \tag{29}$$

$$\dot{v}_4 = r_3 v_3[-r_4 v_4]_1[-r_3 v_3(t-\tau)]_2. \tag{30}$$

For all parameters $r_j$ it must hold that $r_j \geqslant 0$.

Assuming Gaussian errors, maximum likelihood estimates of all unknowns are computed by minimizing

$$\chi^2\{\mathbf{r}, v_1(0)\} = \sum_{i=1}^{N} \sum_{j=1}^{M} \frac{[y_j^D(t_i) - y_j^M\{t_i, \mathbf{r}, v_1(0)\}]^2}{\sigma_{ij}^2} \tag{31}$$

where $N$ is the number of time points and $M$ is the number of observed components. $y_j^D(t_i)$ are the data points at time $t_i$ and $y_j^M\{t_i, \mathbf{r}, v_1(0)\}$ is the solution of the ordinary differential equation for model $M$ with dynamical parameters $\mathbf{r}$ and starting value $v_1(0)$. For details of the minimization technique that is used see Bock (1981), Schittkowski (1995) and Timmer *et al.* (1998).

The $\chi^2$ goodness-of-fit tests for models $M_1$ and $M_2$ were at the border of significance. When comparing models $M_1$ and $M_2$ directly to determine the mechanism of cycling, we deal with a model selection problem, which is rather difficult since we compare non-nested non-linear models in the finite sample case. Unfortunately, strict results only exist for the asymptotic case (Cox, 1962; Pesaran and Deaton, 1978; Vuong, 1989). Therefore, we use a bootstrap procedure to deal with this problem. In the first attempt we apply a procedure that is based on Wahrendorf *et al.* (1987) and Hall and Wilson (1991). We first compute the log-likelihood of both models and obtain the difference $\hat{L}_{12} = [\hat{L}_1 - \hat{L}_2]$. Afterwards, for each of the two model fits, bootstrap samples $\mathbf{Y}^*$ of the data $\mathbf{Y}$ are generated. These bootstrap samples are then used to obtain the distribution of $\hat{L}_{12}^*$ by computing the log-likelihood for each model given the data from the corresponding bootstrap. The bootstrap distribution $S^* = [\hat{L}_{12}^* - \hat{L}_{12}]$ now serves as an approximation of the distribution of $[\hat{L}_{12} - L_0]$ where $L_0$ is the true value. Thus, it is possible to test the null hypothesis $H_0$ that $M_1$ or $M_2$ is the true model and therefore $L_0 \leqslant 0$ or $L_0 \geqslant 0$ respectively. Applying this procedure, we reject neither model $M_1$ nor $M_2$. This is consistent with the results of Hall and Wilson (1991), who showed that this testing procedure has little power. Therefore, to increase the power, we apply a procedure that was suggested by Hall and Wilson (1991).

In this approach, a bootstrap procedure is used to simulate data according to the null hypothesis. Fitting one of the two models with the experimental data yields residuals which are used in what follows to construct bootstrap samples of the data under the hypothesis that the corresponding model is true.

We first simulate separately under fitted model $M_1$ to obtain bootstrap samples $\mathbf{Y}^{*1}$. Afterwards, similarly to the previous procedure, we fit both models $M_1$ and $M_2$ using these bootstrap data $\mathbf{Y}^{*1}$ and compute the likelihood ratio $\hat{L}_{12}^{*1} = [\hat{L}_1^{*1} - \hat{L}_2^{*1}]$. Again we can obtain the bootstrap distribution of $S^{*1} = [\hat{L}_{12}^{*1} - \hat{L}_{12}]$ and as before $S^{*1}$ serves as an approximation to the distribution of $[\hat{L}_{12} - L_0]$. The same procedure is done under fitted model $M_2$, leading to bootstrap samples $\mathbf{Y}^{*2}$. Again we fit both models $M_1$ and $M_2$ and obtain the bootstrap distribution $S^{*2} = [\hat{L}_{12}^{*2} - \hat{L}_{12}]$. With these approximations we may compute significance levels in the case of each statistic and we obtain $\alpha = 0.007$ for model $M_1$ and $\alpha = 0.26$ for $M_2$. Provided that our test procedure has enough power, this indicates that $M_2$ is compatible with the data whereas $M_1$ is not.

The main difference between the two bootstrap procedures is the construction of the bootstrap samples. In the first approach, both models are used to fit the experimental data. Then the

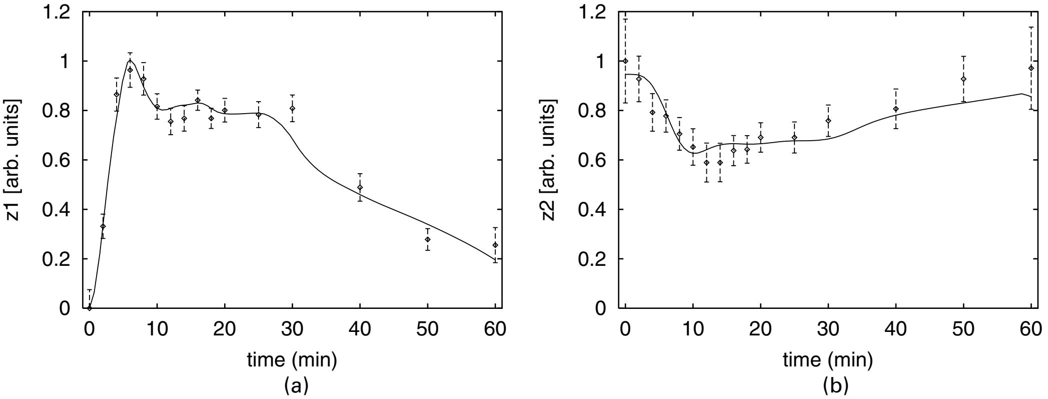

**Fig. 6.** Fit of the selected global model $M_2$ for the experimental data set

residuals of both fits are used to construct distinct bootstrap samples for each of the two models. The second approach uses the fit of one model, e.g. $M_1$, to construct a common bootstrap sample which is then used to compute the likelihood for both models. Hence, the second approach explicitly incorporates the null hypothesis in the construction of the bootstrap samples. In Fig. 6, we display the resulting fit for selected model $M_2$.

## 5. Discussion

We have shown the cycling behaviour of a signalling pathway with specific statistical tests. Since the methods that we used are quite general, they provide a common framework for analysing such biochemical systems. In the first analysis, we investigated cycling in general with the help of a model selection procedure in a linear regression setting. Because of the inequality constraints for dynamical parameters and variables which are a common requirement for signalling pathways, likelihood ratio tests under non-standard conditions are the appropriate technique when analysing these systems. We have shown that different approximating assumptions concerning the error structure and normality may have a huge effect on the distribution of the test statistic, in the present case leading to a too high actual level of the test, bearing the risk of false positive results. Moreover in most cases it is necessary to simulate this distribution under the null hypotheses since analytical results are not available.

In the second investigation, we compared two dynamical models which consist of different cycling mechanisms. This led to a model selection problem with non-nested non-linear models in the finite sample case. A specialized bootstrap technique turned out to have enough power to discriminate between both models.

Owing to an improvement in measurement techniques and the general importance of signalling pathways we expect interest in modelling these systems to grow rapidly in the near future. Therefore it is necessary to provide the mathematical means for model selection procedures to avoid false positive results. Especially in the non-limit case with few data points, statistical testing under non-standard conditions remains a crucial problem.

## References

Bhalla, U. and Iyengar, R. (1999) Emergent properties of networks of biological signaling pathways. *Science*, **283**, 381–387.

Bock, H. (1981) Numerical treatment of inverse problems in chemical reaction kinetics. In *Modelling of Chemical Reaction Systems* (eds K. Ebert, P. Deuflhard and W. Jäger), pp. 102–125. Berlin: Springer.

Campbell, P. (1999) Can physics deliver another biological revolution? *Nature*, **397**, 89.

Chappell, M., Godfrey, K. and Vajda, S. (1990) Global identifiability of the parameters of nonlinear systems with specified inputs: a comparison of methods. *Math. Biosci.*, **102**, 41–73.

Cox, D. R. (1962) Further results on tests of separate families of hypotheses. *J. R. Statist. Soc.* B, **24**, 406–424.

Darnell, J. (1997) STATs and gene regulation. *Science*, **277**, 1630–1635.

Downward, J. (2001) The ins and outs of signaling. *Nature*, **411**, 759–762.

Efron, B. (1982) *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia: Society for Industrial and Applied Mathematics.

Efron, B. and Tibshirani, R. (1993) *An Introduction to the Bootstrap*. New York: Chapman and Hall.

Endy, D. and Brent, R. (2001) Modelling cellular behaviour. *Nature*, **409**, 391–395.

Fussenegger, M., Bailey, J. and Varner, J. A. (2000) A mathematical model of caspase function in apoptosis. *Nat. Biotechnol.*, **18**, 768–774.

Hall, P. and Wilson, S. (1991) Two guidelines for bootstrap hypothesis testing. *Biometrics*, **47**, 757–762.

Hanke, M. and Scherzer, O. (2001) Inverse problems light: numerical differentiation. *Am. Math. Mnthly*, **108**, 512–521.

Haspel, R. and Darnell, J. (1999) A nuclear protein phosphatase is required for the inactivation of STAT1. *Proc. Natn. Acad. Sci. USA*, **96**, 10188–10193.

Haspel, R., Salditt-Georgieff, M. and Darnell, J. (1996) The rapid inactivation of nuclear tyrosine phosphorylated STAT1 depends upon a protein tyrosine phosphatase. *Embo J.*, **15**, 6262–6268.

Hinkley, D. V. (1988) Bootstrap methods. *J. R. Statist. Soc.* B, **50**, 321–337.

Klingmüller, U., Bergelson, S., Hsiao, J. and Lodish, H. (1996) Multiple tyrosine residues in the cytosolic domain of the erythropoietin receptor promote activation of STAT5. *Proc. Natn. Acad. Sci. USA*, **93**, 8324–8328.

Koshland, D. (1998) The era of pathway quantification. *Science*, **280**, 852.

Müller, T. G., Noykova, N., Gyllenberg, M. and Timmer, J. (2002) Parameter identification in dynamical models of anaerobic waste water treatment. *Math. Biosci.*, **177–178**, 147–160.

Pellegrini, S. and Dusanter-Fourt, I. (1997) The structure, regulation and function of the janus kinase (JAK) and the signal transducers and activators of transcription (STATs). *Eur. J. Biochem.*, **248**, 615–633.

Pesaran, M. H. and Deaton, A. S. (1978) Testing non-nested nonlinear regression models. *Econometrica*, **46**, 677–694.

Schittkowski, K. (1995) Parameter estimation in differential equations. In *Recent Trends in Optimization Theory and Applications* (ed. R. Agarwal). Singapore: World Scientific.

Self, S. and Liang, K. Y. (1987) Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *J. Am. Statist. Ass.*, **82**, 605–610.

Swameye, I., Müller, T., Timmer, J., Sandra, O. and Klingmüller, U. (2003) Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by data-based modeling. *Proc. Natn. Acad. Sci. USA*, **100**, 1028–1033.

Timmer, J., Müller, T. and Melzer, W. (1998) Numerical methods to determine calcium release flux from calcium transients in muscle cells. *Biophys. J.*, **74**, 1694–1707.

Vuong, Q. H. (1989) Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, **57**, 307–333.

Wahrendorf, J., Becher, H. and Brown, C. (1987) Bootstrap comparison of non-nested generalized linear models: applications in survival analysis and epidemiology. *J. Appl. Statist.*, **36**, 72–81.

Zheng, T. and Flavel, R. (2000) Death by numbers. *Nat. Biotechnol.*, **18**, 717–718.