BIOINFORMATICS

Vol. 00 no. 00 2005
Pages 1–6

Supplement

Hengl, S.*, Kreutz, C., Timmer, J., Maiwald, T.

Physics Institute, University of Freiburg, Hermann Herder Strasse 3, 79104 Freiburg i.Br., Germany

The supplement is divided into sections entitled in accordance with the sections in the main text. First, the threshold values of the test function are calculated analytically. We then discuss the dependence of optimal transformations on the level of noise and apply the test function to a number of representative examples. The algorithm proposed is depicted in form of a flowchart. The biological motivation of the main example is given in the last section.

METHODS

Determination of Threshold Values

In the following we calculate the expectation value of the proposed test function Equation (8) for a parameter, which is either

- independent of the response variable, or
- functionally related with the response variable

These two cases correspond to the two types of curves depicted in the inset of Figure 2. We show, that the dependence of the test function on the number of bootstrap samples can be stated for the case of an error-free estimation of optimal transformations.

Suppose a parameter with no functional relation with the current response variable; thus its estimated optimal transformation equals gaussian noise which is smoothed by the kernel smoother in the ACE algorithm. Hence the estimated optimal transformation is similar to Figure 1(c), and it can be described by

$$\Phi_k(p_{kr}) = \sum_{r'=r-w_l}^{r'=r+w_l} \epsilon_{r'} \tag{12}$$

Hence, the test function H_k is

$$\begin{aligned} H_k &= \widehat{\mathrm{var}}\left[\frac{1}{NoB}\sum_{i=1}^{NoB}\alpha_k^i(p_{kr})\right] \\ &= \frac{1}{N}\sum_{r=1}^{N}\cdot\left(\frac{1}{NoB}\sum_{i=1}^{NoB}\alpha_k^i(p_{kr})\right)^2 \\ &\quad -\left(\frac{1}{N}\sum_{r=1}^{N}\frac{1}{NoB}\sum_{i=1}^{NoB}\alpha_k^i(p_{kr})\right)^2 \qquad (13)\\ &= \frac{1}{NoB^2}\left(\frac{1}{N}\sum_{r=1}^{N}\beta_k(p_{kr})^2 - \frac{1}{N^2}\left(\sum_{r=1}^{N}\beta_k(p_{kr})\right)^2\right), \end{aligned}$$

with $\beta_k(p_{kr}) = \sum_{i=1}^{NoB}\alpha_k^i(p_{kr})$. Thus, its expectation value $E[H_k]$ is

$$E[H_k] = \frac{1}{NoB^2}\left(\frac{1}{N}\sum_{r=1}^{N}E\left[\beta_k(p_{kr})^2\right] - \frac{1}{N^2}E\left[\left(\sum_{r=1}^{N}\beta_k(p_{kr})\right)^2\right]\right) \tag{14}$$

In the following, we separately calculate the two expectation values occurring in Equation (14).

$$\begin{aligned} E\left[\beta_k(p_{kr})^2\right] &= \sum_{i,j=1}^{NoB}E\left[\alpha_k^i(p_{kr})\alpha_k^j(p_{kr})\right] \\ &= \sum_{i=1}^{NoB}E\left[\alpha_k^i(p_{kr})\alpha_k^i(p_{kr})\right] \\ &\quad + \sum_{i,j=1,i\neq j}^{NoB}E\left[\alpha_k^i(p_{kr})\alpha_k^j(p_{kr})\right] \\ &= \sum_{i=1}^{NoB}E\left[\alpha_k^i(p_{kr})\alpha_k^i(p_{kr})\right] \\ &\quad + \sum_{i,j=1,i\neq j}^{NoB}E\left[\alpha_k^i(p_{kr})\right]E\left[\alpha_k^j(p_{kr})\right] \qquad (15)\\ &= NoB\cdot\gamma + (NoB)(NoB-1)\cdot\frac{1}{4} \end{aligned}$$

with

$$\begin{aligned} \gamma &= E[\alpha_k^i(p_{kr})\alpha_k^i(p_{kr})] \\ &= \sum_{r=1}^{N}\left(\frac{r}{N}\right)^2\cdot\frac{1}{N} \qquad (16)\\ &= \frac{1}{N^3}\sum_{r=1}^{N}r^2 \\ &= \frac{1}{N^3}\left(N(N+1)(2N+1)/6\right) \end{aligned}$$

The second term in Equation (15) holds, because ranks of parameter estimates at t in different bootstrap samples are independent. In Equation (16) we used the explicit form of the estimated optimal transformation (cf. Equation (12)).

$$\begin{aligned} &E\left[\left(\sum_{r=1}^{N}\beta_k(p_{kr})\right)^2\right] \\ &= E\left[\sum_{r=1}^{N}\beta_k(p_{kr})\sum_{r'=1}^{N}\beta_{r'}\right] = \sum_{r,r'=1}^{N}E\left[\beta_k(p_{kr})\beta_k(p_{kr'})\right] \\ &= \sum_{r=1}^{N}E\left[\beta_k(p_{kr})^2\right] + \sum_{r,r'=1,r\neq r'}^{N}E\left[\beta_k(p_{kr})\beta_k(p_{kr'})\right] \\ &= N\cdot E\left[\beta_k^2(p_{kr})\right] + \sum_{r,r'=1,r\neq r'}^{N}E\left[\beta_k(p_{kr})\beta_k(p_{kr'})\right] \qquad (17) \end{aligned}$$

*to whom correspondence should be addressed

$$
\begin{aligned}
&E\left[\beta_k(p_{kr})\beta_k(p_{kr'})\right] \\
&= \sum_{i=1}^{NoB} E\left[\alpha_k^i(p_{kr})\alpha_k^i(p_{kr'})\right] + \sum_{i,j=1,i\neq j}^{NoB} E\left[\alpha_k^i(p_{kr})\alpha_k^j(p_{kr'})\right] \\
&= \sum_{i=1}^{NoB} E\left[\alpha_k^i(p_{kr})\alpha_k^i(p_{kr'})\right] + NoB(NoB-1)\cdot\frac{1}{4} \\
&= NoB\cdot E\left[\alpha_k^i(p_{kr})\alpha_k^i(p_{kr'})\right] + NoB(NoB-1)\cdot\frac{1}{4}
\end{aligned}
$$

$$
\begin{aligned}
&E\left[\alpha_k^i(p_{kr})\alpha_k^i(p_{kr'})\right] \\
&\quad= \sum_{i=1}^{NoB} E\left[(\alpha_k^i(p_{kr})-\mu_r+\mu_r)\cdot(\alpha_k^i(p_{kr'})-\mu_r+\mu_r)\right] \\
&\quad= E\left[\left(\alpha_k^i(p_{kr})-\mu_r\right)\cdot\left(\alpha_k^i(p_{kr'})-\mu_r\right)\right]+\mu_r^2 \\
&\quad= E\left[(\alpha_k^i(p_{kr})-\mu_r)\cdot(\alpha_k^i(p_{kr'})-\mu_r)\right]+\frac{1}{4} \qquad (18)
\end{aligned}
$$

We now calculate an upper bound for the correlation term in Equation (18). Consider Equation (17), thus

$$
\begin{aligned}
&\sum_{r,r'=1,r\neq r'} E\left[(\alpha_k^i(p_{kr})-\mu_r)\cdot(\alpha_k^i(p_{kr'})-\mu_r)\right] \\
&\quad= \sum_{r=1}^{N}\sum_{r=r'-w_l}^{r'+w_l} E\left[(\alpha_k^i(p_{kr})-\mu_r)\cdot(\alpha_k^i(p_{kr'})-\mu_r)\right] \\
&\quad= N\cdot\sum_{r=r'-w_l}^{r'+w_l} E\left[(\alpha_k^i(p_{kr})-\mu_r)\cdot(\alpha_k^i(p_{kr'})-\mu_r)\right] \\
&\quad\leq N\cdot 2w_l E\left[(\alpha_k^i(p_{kr})-\mu_r)\cdot(\alpha_k^i(p_{kr'})-\mu_r)\right] \\
&\quad\leq N\cdot 2w_l\frac{N/2}{N}\cdot\frac{N/2-1}{N} \\
&\quad= w_l\left(\frac{N}{2}-1\right)
\end{aligned}
$$

Together with the factor $\frac{1}{N^2}$ in the second term of Equation (18) and with w_l increasing less than N, e.g. $w_l \sim \sqrt{N}$, we conclude that the overall contribution of the correlation term to $E\left[H_k\right]$ decreases at least as fast as $\frac{1}{\sqrt{N}}$, and hence,

$$
\begin{aligned}
E\left[H_k\right] &= \frac{1}{NoB^2}\left\{\left[NoB(\frac{1}{3}+\frac{1}{2N}+\frac{1}{6N^2})+(NoB^2-NoB)\frac{1}{4}\right]\right. \\
&\quad -\frac{1}{N}\left[NoB(\frac{1}{3}+\frac{1}{2N}+\frac{1}{6N^2})+(NoB^2-NoB)\frac{1}{4}\right] \\
&\quad -\frac{1}{N^2}\left[N(N-1)\left(NoB\cdot\frac{1}{4}+(NoB^2-NoB)\frac{1}{4}\right)\right] \\
&\quad \left.-NoB\cdot\Xi(N)\right\}, \qquad (19)
\end{aligned}
$$

where $\Xi(N)\in O(\frac{1}{\sqrt{N}})$ denotes the overall contribution of the correlation term to expectation value. If we neglect all terms of order $\frac{1}{N}$ and smaller, we yield

$$
E[H_k] = \left(\frac{1}{12}-\Xi(N)\right)\frac{1}{NoB} \qquad (20)
$$

We now calculate H_k for the case that the estimated transformation is a monotone function in the estimated parameter replicates. Since we rank transform estimated transformation all monotone transformations are mapped on a straight line starting from zero with slope one. For a strong functional relation the straight line is the same for all bootstrap samples. We can calculate the value of the variance analytically.

$$
\begin{aligned}
H_k &= var\left[\frac{1}{NoB}\sum_{i=1}^{NoB}\alpha_k^i(p_{kr})\right] \\
&= \frac{1}{N}\sum_{r=1}^{N}\left(\frac{1}{NoB}\sum_{i=1}^{NoB}\alpha_k^i(p_{kr})\right)^2 \\
&\quad -\frac{1}{N^2}\left(\sum_{r=1}^{N}\frac{1}{NoB}\sum_{i=1}^{NoB}\alpha_k^i(p_{kr})\right)^2 \qquad (21) \\
&= \frac{1}{N}\sum_{r=1}^{N}\left(\alpha_k^i(p_{kr})\right)^2-\frac{1}{N^2}\left(\sum_{r=1}^{N}\alpha_k^i(p_{kr})\right)^2 \qquad (22) \\
&= \frac{1}{N}\sum_{r=1}^{N}\left(\frac{r}{N}\right)^2-\frac{1}{N^2}\left(\frac{1}{N}\cdot\frac{N(N+1)}{2}\right)^2 \\
&= \frac{1}{N^3}N(N+1)(2N+1)/6-\frac{1}{N^2}\frac{(N+1)^2}{4} \\
&= \frac{1}{12}(1-\frac{1}{N^2}) \qquad (23)
\end{aligned}
$$

Equation (22) holds, because $\alpha_k^i(p_{kr})$ is independent of i. The expectation value of a real number is the number itself, thus

$$
E[H_k] = \frac{1}{12}(1-\frac{1}{N^2}) \qquad (24)
$$

Identifiability of Identifiability

After having determined which parameters of a given set have a functional relation, the explicit analytical expression could be of interest. In theory, transformation function of non-identifiabilities often can be expressed by simple functions like the exponential or logarithmic function. Thus, analytic expression can be found by fitting subsequently a set of standard functions to the empirically determined transformation functions and selecting the *best* fitting one. In a different setting this has been done by Wang and Murphy (2005), based on the Bayesian Information Criterion (BIC) in order to compare the different fits. In practice, however, problems may arise from the non-uniqueness of determined optimal transformations, as we will demonstrate with an example suggested by these authors.
Consider

$$
\log(y) = \sqrt{x}+\epsilon, \qquad (25)
$$

where ϵ denotes gaussian noise. The transforms of y and x are linearly related, but same is true for

$$
y = \exp(\sqrt{x}+\epsilon), \qquad (26)
$$

where we just applied the exponential. In Figure 7 we show that the estimated transformations depend on the level of noise. In general, optimal transformations estimated for a two parameter dependency are unique only up to nonlinear transformations; this renders impossible to define a finite set of standard functions. Note that this variability is largely compensated in our algorithm. First, the bootstrap approach ensures that sources of noise are equal for each bootstrap sample. Second, the estimated optimal transformations are ranked within the test function before the mean transformation is calculated.

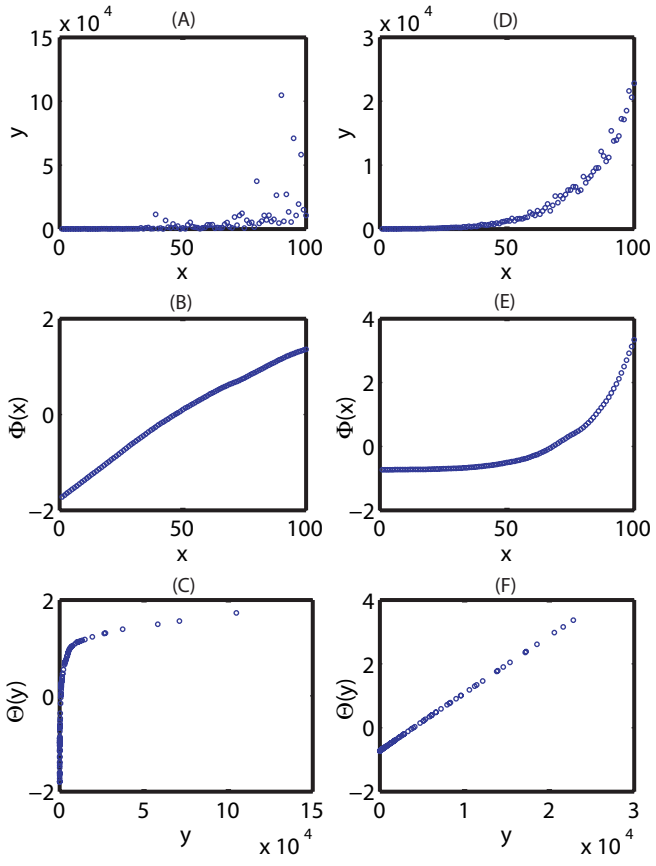

Fig. 7. Optimal transformations depend on strength of noise, here demonstrated with the example $\log(y) = \sqrt{x} + \epsilon$ as discussed by Wang and Murphy (2005). A)(D) x versus y, (B)(E) estimated transformation of x, $\Phi(x)$, versus x and (C)(F) estimated transformation of y $\Theta(y)$ versus y. On the left (A-C) $\epsilon \in N(0,1)$. On the right (D-F) $\epsilon \in 0.1 \cdot N(0,1)$. Comparing left and right hand side shows that estimated transformations strongly depend on the level of noise and are in general nonlinear transforms of each other. In (B), the estimated optimal transformation of x, $\Phi(x)$, can best be described by fitting $\sqrt{x}$ to data, because it yields the lowest BIC score. Thus the left column (A-C) corresponds to Eg. (25). In (F), the estimated optimal transformation of y in the low noise case, $\Theta(x)$, is estimated to be of the form $a \cdot y + b$, thus the right column corresponds to Equation (26). This renders impossible to define a finite set of standard functions which can be compared to estimated transformations, because it depends on the noise, which representative of the equivalence class of possible optimal transformations is estimated. Note, we chose to exemplify the problem with a very low noise example, but it also occurs with considerable higher noise $\epsilon \in 0.5 \cdot N(0,1)$.

For more than two parmaters, the freedom of nonlinear transformations is reduced to linear transformations. However, in practice, the ACE-algorithm which estimates optimal transformations (Breiman and Friedman, 1985) restricts the freedom of possible linear transformations by setting $\Theta(p_k) = \frac{p_k}{||p_k||}$ as initial value. Hence, estimated optimal transformations for more than two parameters may only vary in scale and location. Thus, due to the rank transformation, all possible optimal transformations are mapped onto the same value of the test function.

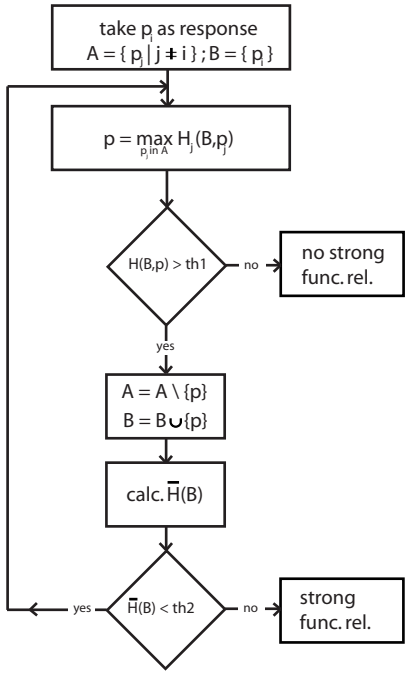

Fig. 8. Simplified flowchart of proposed algorithm. We start with parameter p_i as response. Let B denote the set of all parameters assigned to be functionally related with p_i, and A the set of all other parameters. Thus, at first, $B = \{p_i\}$. Every parameter of A is tested if it is functionally related with all parameters in B, and if so, it is added to B. Successively, more and more parameters are added to B until a strong functional relation is assigned. We stop adding parameters to B either if there is no parameter in A supposed to have a strong functional relation with the response p_i, or if there are no parameters left.

The Algorithm

Figure 8 shows a simplified flowchart of the proposed algorithm. For clarity, control steps and break conditions that are not necessary for the understanding of the algorithm are omitted.

Test Settings

Below we present a short list of representative examples which were generated as follows. A $(n \times k)$ matrix K, uniformly distributed on the interval $[0\ \ 5]$, was drawn and functional relations were introduced as indicated.

$$\begin{aligned} p_1 &= p_2 + 0.2 \cdot p_3 \cdot p_4 \\ p_6 &= p_7 \\ p_8 &= 5 \cdot sin(p_9) + p_{10}^2 \end{aligned} \tag{27}$$

$$\begin{aligned} p_4^2 &= 1/5 \cdot (\exp(p_1) + p_2^3) \\ p_6 &= 5 \cdot (p_3 + p_7) \\ p_8 &= 1/p_5 + 0.3 \cdot p_9 \end{aligned} \tag{28}$$

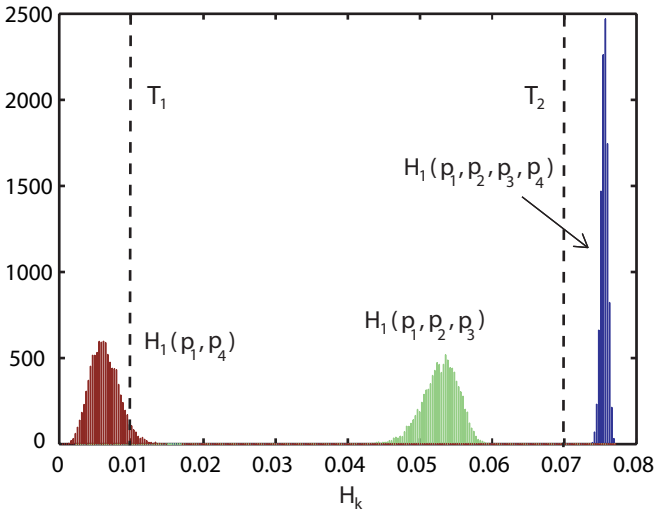

Fig. 9. Histogram showing the distribution of H_k for Equations (27) for three different combinations of parameters; p_1 was taken as response in all three cases. Values of H_k below T_1 indicate that no functional relation exists. Values in the intermediate region between T_1 and T_2 indicate that additional parameters have to be added to establish a complete functional relation. The test-function as well as the proposed algorithm is robust concerning a violation of the assumption of an additive functional relation.

$$\begin{aligned} p_1 &= log(p_2) + 0.1 \cdot p_3 \\ p_5 &= (1 - exp(-p_4)) - 0.1 \cdot p_6 + 100 \end{aligned} \quad (29)$$

$$\begin{aligned} p_1 &= p_2 \cdot p_3 \cdot p_4 \\ p_5 &= \frac{1}{5} \cdot (3 \cdot p_6 + 3 \cdot p_7 + p_8 + p_9) \end{aligned} \quad (30)$$

Additionally, Gaussian noise was added to account for imperfect data and to test robustness of our algorithm. With threshold values ($T_1 = 0.01, T_2 = 0.07$) all functional relations could be recovered with no false positive results. Note that usage of *ACE* implicitly assumes that the underlying functional relation can be written in an additive form, which is not always true, cf. Equation (27). Our studies confirmed that the proposed algorithm is robust concerning violations of the assumption of additivity, see Figure 9.

RESULTS

Guidelines

Here we provide a list of guidlines which summarize statements in the main text and mirror experience gathered with the algorithm and identifiability analysis applied to simulated and biological data:

1. **The model size should always reflect the amount and quality of data at disposal, as well as the current biological research focus.** A completely overparametrizied model results in a great fraction of non-identifiable parameters most of which would require to be fixed in order to render the model identifiable. Thus, one should start the analysis with the smallest model that explains all the data with a minimum amount of parameters.
2. **Distinguish between practically and structurally non-identifiable parameters.** The latter may be fixed at arbitrary values in parameter space without loss of flexibility in the model's dynamics. Practically non-identifiable parameters, however, require considerable care and may demand new experiments or values from the literature.
3. **Iterate identifiabilty analysis and fixation of parameters until all free parameter are rendered identifiable.**
4. **Check always if a practical non-identifiablity is relevant**, e.g., practically non-identifiable parameters may comprise a functional relation, but may nevertheless be determined within a small enough intervall.
5. **The number of fits necessary for the algorithm to work properly depends on the underlying functional relations**. In our simulations we were able to detected linear relations down to 15 points. We found that 50 points and more yield stable results with reasonable sensitivity.
6. **All parameter sets employed for identifiability analysis should in principle yield the same value of the cost-function**, e.g. χ^2-value. In practice, with only limited number of fits, it suffices to take parameter sets with comparable values of the cost function below a certain threshold.
7. **The choice of the cost-function may inherently introduce non-interpretable non-identifiabilities** which can not be eliminated even by perfect measurements. Von Dassow *et al.* (2000) propose and discuss a model for the segment polarity network in drosophila and introduce a cost function which includes an user defined threshold level above which genes are considered to be switched on. This cut-off criteria, though justified, projects a lot of different parameter sets on the same expression profile. The set of sources of non-identifiabilty might therefore be extended by the choice of the cost function.

Endocytosis Model

Endocytosis is the process of engulfing substances outside the cell with a membrane and transporting them into the cytoplasm. The endocytosis of the erythropoietin receptor (EPO receptor) is supposed to consist of a constitutive part and an EPO-induced part, which accelerates the overall degradation of the receptor. A fraction of EPO bound to the internalized receptor is recycled to the extracellular medium, the rest is degraded and then exported as well.
The model in the main text is an effective model supposed to include only the essential steps in the endocytosis of the EPO receptor. For example: The EPO receptor is a homodimer whose two domains dimerize upon stimulation of EPO. This reaction, however, is known to be fast and may thus be neglected on the timescale of endocytosis.

It is assumed, that the internalized EPO receptor is degraded after dissociation from the vesicle. The details of the degradation process are not known, however, it is assumed that degraded receptors do not interfere with the dynamic of the system. Additionally, with the employed measurement techniques (see below), dissociated receptors can not be observed, neither separately nor in combination with other species. Hence, the pool of degraded receptors may be omitted in the formulation of model. In fact, with respect to the available observables, the introduction of an additional rate constant for the degradation of the receptor would yield an over-parametrized

model. The non-identifiabilities as disccued in the main text are not affected by the effective modelling of the receptor degradation.

As we highlight in the main text, even the effective, highly condensed model exhibits non-identifiabilities under the given experimental conditions. Therefore, a further extension of the model would lead to a total overparametrization with respect to available data.

In the following we motivate the two constants $B_{\max}$ and K_D included in the model: The two constants are derived from a Scatchard-analysis (Scatchard, 1949), i.e., we measure specificly bound EPO at various concentrations of EPO in the medium. The experiments are conducted at temperatures far below the physiological body temperature of 37 degrees Celsius, often at room temperature or $4^\circ C$. All cellular processes are therefore assumed to be stopped or at least to be slowed down considerably. We can thus neglect the process of constitutive and enforced endocytosis. At very high concentrations of EPO, all receptors are saturated and the curve of receptor-bound EPO in dependence of freely diffusing EPO is a hyperbola (Kitamura *et al.*, 1989):

$$[EPO\ EPOR] = \frac{B_{\max} \cdot EPO}{K_D + EPO}. \tag{31}$$

The parameters $B_{\max}$ and K_D may be determined by fitting a hyperbola to data. $B_{\max}$ may be interpreted as the total amount of EPO receptors at the cell surface without stimulation. The equation for the unstimulated EPO receptor without EPO in the medium is $\dot{EPOR} = -k_1 EPOR + k_1 B_{\max}$. Thus for the steady state, $\dot{EPOR} = 0$, we yield $EPOR = B_{\max}$.

In order to reveal how K_D is incorporated in the model, we rewrite Equation (31) like follows

$$\frac{[EPO\ EPOR]}{EPO} = -\frac{1}{K_D} \cdot [EPO\ EPOR] + \frac{B_{\max}}{K_D}. \tag{32}$$

If we consider the differential equation for *EPO* in the model and if we neglect the fraction of internalized *EPO*, we may write

$$\dot{EPO} = +k_2 K_D [EPO\ EPOR] - k_2 EPO\ EPOR.$$

Thus, for the steady state of EPO ($\dot{EPO} = 0$) we get

$$\begin{aligned} \frac{[EPO\ EPOR]}{EPO} &= \frac{1}{K_D} \cdot EPOR \\ &= \frac{1}{K_D} \cdot (B_{\max} - [EPO\ EPOR]) \\ &= -\frac{1}{K_D} \cdot [EPO\ EPOR] + \frac{B_{\max}}{KD}. \end{aligned}$$

We choose $EPOR(t = 0) = B_{\max} = 1000$ and $K_D = 100$, which corresponds approximately to the values measured by Sawada *et al.* (1988) with concentrations given in $pMol/L$.

The particular combination of species to observables depends on the measurement technique. We assume EPO to be radiolabeled with ^{125}I. The quantification is then achieved by measuring the emission rate of a probe. The three observables correspond to the three reaction compartments, i.e. the extracellular medium, the proteins bound on the surface of the membrane, and the intracellular medium. The separation of the three compartments is a standard procedure and is described, e.g, in (Broudy *et al.*, 1988).

REFERENCES

Breiman, L. and Friedman, J. (1985). Estimating optimal transformations for multiple regression and correlation. *Journal of the American Statistical Associtation*, **80(391)**, 580–598.

Broudy, V. C., Lin, N., Egrie, J., de Han, C., Weiss, T., Papayannopoulou, T., and Adamson, J. W. (1988). Identification of the receptor for erythropoietin on human and murine erythroleukemia cells and modulation by phorbol ester and dimethyl sulfoxide. *Proc. Natl. Acad. Sci. USA*, **85**(17), 6513–6517.

Kitamura, T., Tojo, A., Kuwaki, T., Chiba, S., Miyazono, K., Urabe, A., and Takaku, F. (1989). Identification and analysis of human erythropoietin receptors on a factor-dependent cell line, tf-1. *Blood*, **73**(2), 375–380.

Sawada, K., Krantz, S. B., Sawyer, S. T., and Civin, C. I. (1988). Quantitation of specific binding of erythropoietin to human erythroid colony-forming cells. *J Cell Physiol*, **137**(2), 337–345.

Scatchard, G. (1949). The attraction of proteins for small molecules and ions. *Ann. N.Y. Acad. Sci.*, **51**, 660–672.

Von Dassow, G., Meir, E., Munro, E. M., and Odell, G. M. (2000). The segment polarity network is a robust developmental module. *Nature*, **406**(6792), 188–192.

Wang, D. and Murphy, M. (2005). Identifying nonlinear relationships in regression using the ACE algorithm. *Journal of Applied Statistics*, **32(3)**, 243–258.