Parameter Identification and Model Verification in Systems of Partial Differential Equations Applied to Transdermal Drug Delivery

Klaus Schittkowski

September 27, 2007

Abstract

The purpose of this paper is to present some numerical tools which facilitate the interpretation of simulation or data fitting results and which allow to compute optimal experimental designs. They help to validate mathematical models describing the dynamical behavior of a biological, chemical, or pharmaceutical system, without requiring a priory knowledge about the physical or chemical background. Although the ideas are quite general, we will concentrate our attention to systems of one-dimensional partial differential equations and coupled ordinary differential equations. A special application model serves as a case study and is outlined in detail. We consider the diffusion of a substrate through cutaneous tissue, where metabolic reactions are included in form of Michaelis-Menten kinetics The goal is to simulate transdermal drug delivery, where it is supposed that experimental data are available for substrate and metabolic fluxes. Numerical results are included based on laboratory data to show typical steps of a model validation procedure, i.e., the interpretation of confidence intervals, the compliance with physical laws, the identification and elimination of redundant model parameters, the computation of optimum experimental designs and the identifiability of parameters by determining weight distributions.

Keywords: parameter estimation, data fitting, least squares optimization, statistical analysis, confidence interval, optimum experimental design, partial differential equation, method of lines, substrate diffusion, cutaneous tissue, transdermal application

1 Introduction

The numerical identification of unknown parameters of a dynamical model is extremely important in many technical, natural science, and medical disciplines. In this paper, we proceed from a system of one-dimensional partial differential equations. The goal is to minimize distances of appropriate fitting criteria depending on a set of model parameters and the solution of the partial differential equation from experimental data taken at predetermined time and spatial values. The underlying optimization problem consists of minimizing a sum of squared functions given in the form

$$s(p) := \sum_{i=1}^{l} (h(t_i, p) - y_i)^2 \tag{1}$$

with experimental time values t_i and measured data y_i, $i = 1, \ldots, l$.

We assume that a model function $h(t,p)$ is available depending on a vector $p \in I\!R^n$, the model parameters to be identified, and a time variable t. By varying p, the mathematical model is to be fitted to a set of measurements by a least squares estimate.

The fitting criterion $h(t,p)$ depends on the solution of a system of one-dimensional, time-dependent partial differential equations,

$$u_t = f(x,t,u,u_x,u_{xx},p) \;, \tag{2}$$

i.e., $h(t,p)$ depends on the time t, the parameter vector p, and the solution $u(x,t,p)$ of (2) at fixed spatial values. In addition, there are initial values for the solution $u(x,0,p)$ typically formulated for $t=0$ and boundary values for specifying function values or partial derivatives. As usual, the lower index t denotes the partial derivative subject to t, the lower index x the first partial derivative subject to x, and xx the second partial derivative subject to x.

More extensions of the mathematical model are possible, for example disjoint spatial integration areas with non-continuous transitions, coupled ordinary and algebraic equations, arbitrary boundary conditions, algebraic partial differential equations, dependencies of the fitting criterion also on higher derivatives, break or switching points, dynamic, i.e., time-dependent constraints, algebraic equality or inequality constraints, etc., see Schittkowski [28]. All these additional equations may depend on parameters to be estimated.

A standard discretization procedure is known as the method of lines, see Schiesser [22]. We fix the spatial variable x at a set of equidistant points, the *lines*, and approximate the corresponding first and second derivatives subject to x by an appropriate difference formula. Initial, boundary, and transition conditions between different integration areas must be taken into account. The resulting large and eventually stiff system of ordinary differential equations can be integrated by numerical standard techniques, see, e.g., the implicit methods developed by Hairer and Wanner [15]. For the iterative solution of the least squares problem, we propose any of the available Gauss-Newton and related methods, e.g., the codes DN2GB of Dennis et al. [7], NLSNIP of Lindström [17], or DFNLP of Schittkowski [24, 30].

A special model is investigated in detail in form of a case study, that describes the diffusion of a substrate (Ala-MNA) through cutaneous tissue, see Boderke et al. [3] or Wolf [42]. Two different layers are considered, the metabolically active tissue and a passive porous membrane. Metabolic reactions are taken into account by a Michaelis-Menten kinetics, leading to a system of two partial differential equations in two different integration areas. Non-continuous transition conditions between both areas are formulated. The underlying in-vitro experiment simulates transdermal drug delivery, where experimental data are known for substrate and metabolic fluxes. To be able to compute the total amounts of a drug entering and leaving the diffusion area, ordinary differential equations are coupled to the partial ones at both boundary points.

Whenever a complex dynamical system is applied for simulating a real-life phenomenon, the main question is whether the mathematical model predicts the observed behavior in a correct way. Typical questions are:

1. Can model parameters be identified from given experimental data?
2. Are there redundant parameters and how can we identify them?
3. How do measurement errors in experimental data influence the parameter values?
4. How can we interpret confidence intervals?
5. What do we know about the steady-state?
6. Are fundamental physical or other laws satisfied, e.g., the mass balance equation?

7. How can we predetermine optimal experimental design parameters?

Our intention is to outline a complete model validation process in form of a case study , i.e., we show how numerical tests can be designed in a practical situation to answer these questions. The procedures are general and can be applied to any other dynamical processes as well, see Schittkowski [32] for dynamic models based on systems of ordinary differential algebraic equations.

An alternative approach for parameter estimation and optimum experimental design of flow reactors is discussed in Carraro [6]. Investigation of the sensitivity matrix for reducing parameter correlations is proposed by Thomasethi and Cobelli [38] for input-output models based an ordinary differential equations. A statistical procedure for model selection is found in Bortz and Nelson [4], and a maximum-likelihood estimator for testing parameter sensitivities is derived by Gorfine and Freedman [12], in both cases applied to specific biological ordinary differential equation models.

Some statistical tool to analyze these questions are summarized in Section 2. In many cases, data fitting models are overdetermined, lead to too large confidence intervals, and prevent a precise identification of model parameters. Thus, we first present a brief outline of confidence intervals in Section 2.1. The notation is introduced and some formulas are presented for determining confidence regions needed subsequently. A numerical procedure is introduced in Section 2.2 based on successive elimination of the least important parameter until a given significance level is reached. Once we have a reliable model, we might be interested in the question how to compute the *best*, i.e., optimal set of design parameters based on the available mathematical equations. Section 2.3 contains a brief outline of optimum experimental design and some mathematical details, e.g., how to compute derivatives.

A pharmaceutical diffusion process serves as a case study, the transdermal application of a substrate including generation of a metabolite. The mathematical model is completely described in Section 3.

A brief outline of the numerical algorithms used to discretize the partial differential equation and to solve data fitting and optimum experimental design problems is given in Section 4.

Numerical test results are summarized in Section 5. Proceeding from available laboratory data, a typical first step is to estimate given parameters, to compute a least squares fit and confidence intervals, and to plot the results. The priority analysis allows to figure out those parameters, which are considered as redundant and which can be eliminated, see Section 5.1. By analyzing plots visually, the steady-state situation and a simple physical plausibility argument based on the mass conservation law help to validate the model, see Section 5.2. After identifying design parameters, numerical values for their optimal choice are computed following the procedure of Section 2.3, confer Section 5.3. Additional optimization of weights leads to the observation, that the number of non-zero weights is identical to the number of parameters, a further hint that the parameters we obtained from the priority analysis are uniquely identifiable.

2 Statistical Analysis Tools

2.1 Confidence Intervals

We proceed from a general nonlinear model in its simplest form

$$\eta = h(p,t) + \epsilon \quad . \tag{3}$$

$h(p,t)$ is our model function depending on a parameter vector $p \in I\!R^n$ and $t \in I\!R$ is the independent variable, also called explanatory or regression variable, in most cases the time. The function $h(p,t)$ is supposed to be differentiable subject to $p \in I\!R^n$, and at least continuous with respect to t. It is assumed

that there is a true parameter value $p^\star$, which is unknown and which is to be estimated by a least squares fit. The response $\eta \in I\!R$ is the dependent model variable.

To estimate $p^\star$ from given experimental data t_i and y_i, $i = 1, \ldots, l$, we minimize the least squares function (1) over all $p \in I\!R^n$. Let $\hat{p}$ denote the solution of this data fitting problem. The question we are interested in is how far away $\hat{p}$ is from the true parameter $p^\star$.

It is assumed that the time values t_i are given a priori without errors, and that ϵ_i denotes the statistical error of the measurements or the response variable, respectively. As usual, we suppose that the errors $\epsilon_i = h(p^\star, t_i) - y_i$ are independent and normally distributed with mean value zero and known constant variance σ^2, i.e., $\epsilon_i \sim N(0, \sigma^2)$ for $i = 1, \ldots, l$.

The basic idea is to linearize the nonlinear model

$$\overline{h}(p) := (h(p, t_1), \ldots, h(p, t_l)^T$$

in a neighborhood of $p^\star$ and to apply linear regression analysis, since linear models are very well understood, see Seber [35]. Under additional regularity assumptions, $\hat{p}$ and

$$\hat{\sigma}^2 := s(\hat{p})/(l - n)$$

are consistent estimates of $p^\star$ and σ^2, respectively. They converge with probability one to the true values, and are asymptotically normally distributed as l goes to infinity. Moreover, we know that due to the normal distribution of the errors, $\hat{p}$ is also a maximum likelihood estimator.

The error in parameters, $\hat{p} - p^\star$, is approximately normally distributed with mean value 0 and covariance matrix $I^{\star -1}$, where

$$I^\star := \frac{1}{\sigma^2} \nabla \overline{h}(p^\star) \nabla \overline{h}(p^\star)^T \quad ,$$

see Seber and Wild [36]. An approximate $100(1 - \alpha)\%$ confidence region for $p^\star$ is given by the set

$$\{p : (p - \hat{p})^T \hat{I} (p - \hat{p}) \leq n F^\alpha_{n,l-n}\} \quad , \tag{4}$$

where $F^\alpha_{n,l-n}$ denotes the F-distribution with $(n, l - n)$ degrees of freedom within the linearization error, and where

$$\hat{I} := \frac{1}{\hat{\sigma}^2} \nabla \overline{h}(\hat{p}) \nabla \overline{h}(\hat{p})^T \tag{5}$$

estimates $I^\star$. This result is very similar to the corresponding confidence region for linear models.

For a numerical implementation, however, (4) is inconvenient. $100(1 - \alpha)\%$ confidence intervals for the parameters are given by

$$\left[\hat{p}_i - t^{\alpha/2}_{l-n} \sqrt{\hat{d}_{ii}} \ , \ \hat{p}_i + t^{\alpha/2}_{l-n} \sqrt{\hat{d}_{ii}} \right] \tag{6}$$

for the i-th individual model parameter value $p^\star_i$, $i = 1, \ldots, n$. In this case, $\hat{p}_i$ is the i-th coefficient of $\hat{p}$ and $\hat{d}_{ii}$ the i-th diagonal element of $\hat{I}^{-1}$, see also Gallant [10] or Donaldson and Schnabel [9]. t_{l-n} denotes the t-distribution with $l - n$ degrees of freedom.

However, (6) is valid only approximately depending on the quality of the linearization or the curvature of $h(p, t)$, respectively. Donaldson and Schnabel [9] present some examples, where the confidence intervals are very poor. Thus, we have to be very careful when computing (6) without additional linearization checks.

2.2 Significance Levels by Eigenvalue/-vector Analysis of the Fisher Information Matrix

Proceeding from a parameter estimation model, corresponding data, and a successful least squares fit, significance levels of the estimated parameters are to be evaluated. If a model seems to be overdetermined, i.e., contains too many parameters compared to the number of equations, the levels give an impression of the significance of parameters and help to decide upon questions like

- which parameters can be identified,
- which parameters can be treated as constants,
- whether additional experimental should be added or not.

Moreover, overdetermined data fitting problems almost always lead to unstable and slow convergence of Gauss-Newton-type least squares algorithms with a large number of iterations until termination tolerances are satisfied.

We have seen in the previous section that $\hat{I}^{-1}$ can be considered as an approximation of the covariance matrix $I^{\star -1}$ and that $\hat{p}$ is a least squares estimate for the true, but unknown parameter $p^{\star}$. Assumptions are independent and normally distributed errors in the measurements with mean value 0 and variance σ^2.

A more rigorous analysis based on the maximum-likelihood function leads to the theorem of Cramér and Rao, which states that the inverse of the so-called Fisher information matrix (5) is a lower bound for the covariance matrix of the parameter errors. For a precise definition of this matrix and a proof see, e.g., Goodwin and Payne [11].

Since all induced matrix norms are greater than the spectral radius of a matrix, we apply the L_2-norm, i.e.,

$$\|\hat{I}^{-1}\|_2^2 = |\lambda_{max}(\hat{I}^{-1})| = \frac{1}{|\lambda_{min}(\hat{I})|} \quad . \tag{7}$$

λ_{max} and λ_{min} denote the largest and smallest eigenvalues of a matrix, respectively. Since small eigenvalues of $\hat{I}$ enforce large entries of the covariance matrix, we try to reduce them by successive elimination of parameters corresponding to large eigenvector coefficients. The order by which the variables are eliminated, can be considered as an indication about their relative significance. The highest level reflects the highest priority.

We proceed from a given significance tolerance $\gamma > 0$, known time values $t_1, \ldots, t_l$, and an optimal solution $\hat{p}$ of the corresponding least squares data fitting problem. We try to satisfy

$$\|\hat{I}^{-1}\|_2^2 = \frac{1}{|\lambda_{min}(\hat{I})|} < \gamma^2 \quad . \tag{8}$$

Assuming a sufficiently accurate approximation of $p^{\star}$, the true parameter vector, we hope to get sufficiently small variances.

Note that very small or zero eigenvalues lead to the conclusion that some parameters cannot be estimated at all by the underlying model and the available data, or that there are combinations of highly correlated parameters, see Caracotsis ans Stewart [5]. To detect the significant parameters on the one hand and the redundant or dependent parameters on the other we apply a procedure described in Schittkowski [32] in more detail, see also Schneider, Posten, and Munack [34] or Majer [19]. The idea is to successively eliminate parameters until (8) is satisfied, which corresponds to the largest absolute value of the eigenvector of the smallest eigenvalue λ_{min}. The cycle is terminated in one of the following situations:

1. The smallest eigenvalue of the Fisher information matrix is smaller than a threshhold value, see (8).

2. The parameter correlations are significantly reduced, e.g., by 25 %.

3. None of the above termination reasons is met and all parameters have been eliminated.

After termination, significance levels of the parameters are available. Level 1 corresponds to the first eliminated variable, level 2 to the second, etc. The final level can be assigned to several parameters indicating a group of identifiable parameters. Possible conclusions are to add more experimental data or to fix some parameters for subsequent evaluations. Thus, the determination of significance levels is part of the experimental design process to validate a parameter estimation model. It is recommended to scale the parameters to the value one before starting the priority analysis. Note that the priority analysis does not depend on any experimental data other than a set of time values.

It should be noted, however, that the proposed heuristic procedure is by no means justified by a mathematically rigorous theory. For alternative approaches to identify system parameters, see for example Bard [2], Ljung [18], van den Bosch and van der Klauw [39], Walter and Pronzato [40], or Banks and Kunisch [1]. In these monographs, numerous methods are presented proceeding either from a statistical analysis, theoretical model building, or more special model structures.

2.3 Experimental Design

Mathematical models describe the dynamical behavior of a system with the goal to allow numerical evaluation of model parameters which cannot be measured directly. These parameters identify the system under consideration, and must be verified by experiments. Experimental design depends on parameters which have to be set in advance to be able to measure the output data of an experiment. Examples are initial concentration of substrates, input feeds of a chemical reactor, temperature distributions, etc.

To determine the optimal experimental design parameters, we first have to find a suitable guess for the model variables either from the literature or some preliminary experiments. We have seen in the previous sections that the Fisher information matrix or the covariance matrix, respectively, determines the confidence region (4) of the model parameters. Since we have additional freedom to setup an experiment a priori, we can use the design parameters to minimize the volume of the corresponding ellipsoid based on a suitable criterion.

To formalize the approach, we denote the model parameters by $p \in \mathbb{R}^{n_p}$ and the design parameters by $q \in \mathbb{R}^{n_q}$. Our model function is now extended by the design parameters, $h(p, q, t)$, and we assume that we know a set of experimental time values t_i, $i = 1, \ldots, l$. Moreover, we let

$$\overline{h}(p, q) := (h(p, q, t_1), \ldots, h(p, q, t_l)^T$$

and denote by $F(p, q) = \nabla_p \overline{h}(p, q)$ the Jacobian matrix of $\overline{h}(p, q)$ subject to $p \in \mathbb{R}^{n_p}$, where $q \in \mathbb{R}^{n_q}$ is fixed. For simplicity, we assume that $F(p, q)$ has full rank for all p and q.

A formal performance measure is available based on the covariance matrix $C(p, q) = I(p, q)^{-1}$, where $I(p, q) = F(p, q)F(p, q)^T$ denotes an approximation of the Fisher information matrix, and where we omit a guess for the error variances of measurements to simplify the notation. In other words, we assume that all experimental data are measured with constant error.

The volume of a confidence region for a given model parameter $p \in \mathbb{R}^{n_p}$ is that of an an ellipsoid, see (4), and the goal is to minimize its volume on the one hand, but on the other to prevent also degenerate situations where the maximum and minimum eigenvalues drift away, i.e., where the ellipsoid is *compressed*. This is to be achieved by adapting the design parameter q for a given model parameter p,

which is obtained either from a preliminary experiment, literature, or a reasonable guess. Possible criteria are available either for $C(p,q)$ or $I(p,q)$, respectively, depending on the procedure how to measure or estimate the volume and the structure of the ellipsoid. The most popular ones are

$$\begin{array}{lcl} D & : & \det(C(p,q)) \\ A & : & \text{trace}(C(p,q)) \\ A^\star & : & \text{-trace}(I(p,q)) \\ E & : & \lambda_{min}(I(p,q)) \\ E^\star \text{ or } C & : & \lambda_{min}(I(p,q))/\lambda_{max}(I(p,q)) \end{array}$$

Here $\lambda_{min}(I(p,q))$ and $\lambda_{max}(I(p,q))$ denote the minimum and maximum eigenvalues of $I(p,q)$. For a more detailed discussion, see, e.g., Winer, Brown and Michels [41] or Ryan [21].

For our numerical implementation, we use the A-criterion, since the computationally attractive confidence intervals by which the size of the ellipsoid is estimated, take only the diagonal elements of the covariance matrix into account, see (6). This leads for each $p \in I\!R^{n_p}$ to the optimization problem

$$q \in I\!R^{n_q} : \quad \begin{array}{l} \min \ \text{trace}(C(p,q)) \\ q_l \le q \le q_u \ , \end{array} \tag{9}$$

where we add additional bounds for the variable q.

There remains the question how to compute the derivatives of the objective function

$$\phi(q) := \text{trace}(C(p,q)) \tag{10}$$

subject to q in an efficient way. Numerical differentiation of $\phi(q)$ subject to q by a difference formula based on a previous numerical differentiation of $h(p,q,t)$ subject to p by another difference formula for computing $F(p,q)$ is unstable because of accumulation of truncation errors. It is assumed that second order analytical partial or mixed-partial derivatives are not available. Thus, we try to find a reasonable compromise which nevertheless leads to sufficiently stable procedure.

As shown by Schittkowski [32], partial differentiation of the objective function of (9) subject to q_k, $1 \le k \le n_q$, gives

$$\begin{array}{rcl} \dfrac{\partial}{\partial q_k}\phi(q) & = & -\text{trace}\left(I(p,q)^{-1} \ \left(\dfrac{\partial}{\partial q_k} F(p,q) F(p,q)^T \right.\right. \\ & & \left.\left. + \ F(p,q) \dfrac{\partial}{\partial q_k} F(p,q)^T \right) \ I(p,q)^{-1} \right) \ . \end{array} \tag{11}$$

The mixed partial derivatives of $\overline{h}(p,q)$ subject to p and q are approximated by forward differences

$$\begin{array}{rcl} \dfrac{\partial^2}{\partial q_k \partial p_j} h(p,q,t_i) & \approx & \dfrac{1}{\epsilon_k \epsilon_j} \left((h(p+\epsilon_j e_j, q+\epsilon_k e_k, t_i) + h(p,q,t_i)) \right. \\ & & \left. - \ (h(p, q+\epsilon_k e_k, t_i) + h(p+\epsilon_j e_j, q, t_i)) \right) \end{array} \tag{12}$$

for $i = 1, \ldots, l$ and $j = 1, \ldots, n_p$. Here, $e_j \in I\!R^{n_p}$ and $e_k \in I\!R^{n_q}$ are the j-th and k-th unit vectors, respectively, and ϵ_j, ϵ_k are suitable perturbation tolerances, e.g., chosen by $\epsilon_j = \max(\nu, |p_j|)\epsilon$ and $\epsilon_k = \max(\nu, |q_k|)\epsilon$ with tolerances $\epsilon > 0$ and $\nu > 0$, which must be selected very carefully.

The perturbation tolerance ϵ should not be too small. Depending on the condition number of the information matrix, large values like $\epsilon = 0.01$ or even $\epsilon = 0.1$ are applicable and can lead to stable solution processes subject to a surprisingly small optimality criterion.

The experimental design approach introduced in the previous section assumes that the time values are known in advance. However, there are very many situations where one would like to know in advance their approximate number and also their optimal locations, to improve the confidence intervals of the parameters to be estimated, and to reduce the number of time-consuming or expensive experiments.

Our idea is to proceed from a given set of artificial time values which could be large and dense, and to formulate an experimental design optimization problem as before by introducing weights w_i, $i = 1, \ldots, l$. Thus, we replace the model function $h(p, q, t_i)$ by $w_i h(p, q, t_i)$ with additional weight factors w_i, $i = 1, \ldots, l$, which are to be treated as optimization variables in our optimum design problem (9) and which becomes

$$q \in I\!R^{n_q}, w \in I\!R^l : \quad \begin{array}{ll} \min \ \mathrm{trace}(C(w,p,q)) & \\ \sum_{i=1}^{l} w_i = 1 \ , & \\ q_l \le q \le q_u \ , & \\ \tau \le w_i \le 1 \ , & i = 1, \ldots, l \ . \end{array} \tag{13}$$

The covariance matrix depends now on additional weights, i.e., on

$$\begin{aligned} F(w,p,q) &:= \nabla_p \overline{h}(w,p,q) \ , \\ \overline{h}(w,p,q) &:= (w_1 h(p,q,t_1), \ldots, w_l h(p,q,t_l))^T \ . \end{aligned}$$

Note that for stability reasons, a small lower bound $\tau > 0$ is introduced.

Corresponding partial derivatives of the objective function $\phi(w, q)$ subject to a weight w_i are obtained from

$$\begin{aligned} \frac{\partial}{\partial w_i}\phi(w,q) &= -\mathrm{trace}\Big(I(w,p,q)^{-1} \ \Big(\frac{\partial}{\partial w_i} F(w,p,q) F(w,p,q)^T \\ &\quad + F(w,p,q) \frac{\partial}{\partial w_i} F(w,p,q)^T \Big) \ I(w,p,q)^{-1} \Big) \ . \end{aligned} \tag{14}$$

see (11), and from

$$\frac{\partial}{\partial w_i} F(w,p,q) = \left(\frac{\partial}{\partial p_j} h(p,q,t_i) \right)_{j=1,n_p;i=1,l} \ . \tag{15}$$

Thus, we get the weight derivatives more or less for free, since the partial derivatives subject to the model parameters are known from the computation of the objective function.

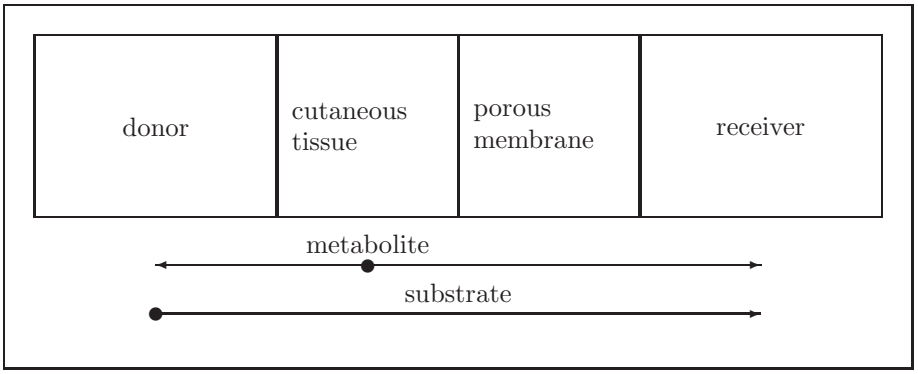
3 The Transdermal Diffusion Model

We consider the permeation of a substrate through cutaneous tissue with simultaneous metabolism and through a subsequent passive porous membrane. Goal is the analysis and numerical simulation of transdermal processes for developing new drugs and application devices, see, e.g., Guy and Hadgraft [13]. Especially the affects of various parameters influencing diffusion and metabolism can be studied in detail, once a reliable model validated by experimental data is available.

A special laboratory experiment is to be modeled, where a donor contains a given volume and concentration of a substrate, e.g., Ala-MNA, that penetrates through a thin layer of a cutaneous tissue with

thickness l_T and through an underlying membrane with thickness l_M with known porosity. $l = l_T + l_M$ is the total thickness of the diffusion area under consideration. From time to time, the concentration of the substrate is measured on both sides of the membrane.

At the same time, a metabolite is generated because of enzymatic interactions, see Hotchkiss [16], Hadgraft [14], or Guy and Hadgraft [13]. Also the concentration of this new substrate is measured on both sides of the membrane. It is assumed that metabolism can be described by Michaelis-Menten kinetics, see Pratt [20], that the distribution of metabolically active enzyms in the layer is homogeneous, and that the mass transport is one-dimensional along the x-axis.



To describe the mathematical model of the experiment as outlined above, we use the notation

l_T	-	thickness of tissue,
l_M	-	thickness of membrane,
l	-	total thickness of diffusion area, i.e., $l = l_T + l_M$,
K_m	-	Michaelis-Menten constant $[nmol/mm^3]$,
V_a	-	volume of donor and receiver $[mm^3]$,
F_a	-	surface of the membrane $[mm^2]$,
Y_0	-	initial mass of substrate $[nmol]$

for design parameters and constants. Model parameters $p = (p_1, \ldots, p_9)^T$ to be estimated, are

D_T^s	-	diffusion coefficient of substrate in tissue,
D_M^s	-	diffusion coefficient of substrate in membrane,
D_T^m	-	diffusion coefficient of metabolite in tissue,
D_M^m	-	diffusion coefficient of metabolite in membrane,
P^s	-	distribution coefficient of substrate,
P^m	-	distribution coefficient of metabolite,
$V_{\max}$	-	metabolization rate,
T^s	-	transition parameter of substrate,
T^m	-	transition parameter of metabolite.

The dynamic variables we need to describe the process, are

$u^s(x,t)$	-	substrate concentration at x and t $[nmol/mm^3]$,
$u^m(x,t)$	-	metabolite concentration at x and t $[nmol/mm^3]$,
$v^s(t)$	-	substrate mass at donor, i.e., for $x = 0$ $[nmol]$,
$w^s(t)$	-	substrate mass at receiver, i.e., for $x = l$ $[nmol]$,
$v^m(t)$	-	metabolite mass at donor, i.e., for $x = 0$ $[nmol]$,
$w^m(t)$	-	metabolite mass at receiver, i.e., for $x = l$ $[nmol]$.

Two partial differential equations describe the diffusion through the skin and take the Michaelis-Menten effect into account, see also Wolf [42], Steinsträsser [37], or Boderke et al. [3],

$$\begin{aligned} u_t^s(x,t) &= D_T^s\, u_{xx}^s(x,t) - \frac{V_{\max} u^s(x,t)}{K_m + u^s(x,t)} \ , \\ u_t^m(x,t) &= D_T^m\, u_{xx}^m(x,t) + \frac{V_{\max} u^s(x,t)}{K_m + u^s(x,t)} \end{aligned} \tag{16}$$

for $0 < x < l_T$ and $t > 0$. The corresponding initial values are $u^s(x,0) = 0$ and $u^m(x,0) = 0$ for all $x > 0$, and $u^s(0,0) = Y_0 P^s / V_a$, $u^m(0,0) = 0$. Diffusion of substrate and metabolite through membrane are described by the equations

$$\begin{aligned} u_t^s(x,t) &= D_M^s\, u_{xx}^s(x,t) \ , \\ u_t^m(x,t) &= D_M^m\, u_{xx}^m(x,t) \end{aligned} \tag{17}$$

defined for $l_T < x < l$ and $t > 0$. Corresponding initial values are $u^s(x,0) = 0$ and $u^m(x,0) = 0$ for all x.

Accumulation of mass fluxes at both sides of the two layers leads to the integrals

$$\begin{aligned} v^s(t) &= F_a D_T^s \int_0^t u_x^s(0,\tau)d\tau \ , \\ v^m(t) &= F_a D_T^m \int_0^t u_x^m(0,\tau)d\tau \ , \\ w^s(t) &= -F_a D_M^s \int_0^t u_x^s(l,\tau)d\tau \ , \\ w^m(t) &= -F_a D_M^m \int_0^t u_x^m(l,\tau)d\tau \ . \end{aligned} \tag{18}$$

From these equations we get four ordinary differential equations that are coupled to (16) and (17) to obtain a dynamic system in standard form, i.e.,

$$\begin{aligned} \dot{v}^s(t) &= F_a D_T^s \ u_x^s(0,t) \ , \\ \dot{v}^m(t) &= F_a D_T^m \ u_x^m(0,t) \ , \\ \dot{w}^s(t) &= -F_a D_M^s \ u_x^s(l,t) \ , \\ \dot{w}^m(t) &= -F_a D_M^m \ u_x^m(l,t) \end{aligned} \tag{19}$$

for $t > 0$. They describe the behavior of substrate and metabolite at both end points, i.e., the mass fluxes into and out of tissue and membrane. Initial conditions are $v^s(0) = Y_0$, $v^m(0) = 0$, $w^s(0) = 0$, and $w^m(0) = 0$.

Boundary conditions are formulated to couple the partial differential equations and the ordinary differential equations, i.e.,

$$\begin{aligned} u^s(0,t) &= \frac{P^s}{V_a} \ v^s(t) \ , \\ u^m(0,t) &= \frac{P^m}{V_a} \ v^m(t) \ , \\ u^s(l,t) &= \frac{P^s}{V_a} \ w^s(t) \ , \\ u^m(l,t) &= \frac{P^m}{V_a} \ w^m(t) \end{aligned} \tag{20}$$

for all $t > 0$.

Between both integration areas we allow non-continuous conditions for the transition of substrate and metabolite from tissue to membrane and vice versa, but require continuous mass flows, i.e.,

$$\begin{aligned} u^s(l_T^-,t) &= T^s \ u^s(l_T^+,t) \ , \\ u^m(l_T^-,t) &= T^m \ u^m(l_T^+,t) \ , \\ u_x^s(l_T^+,t) &= \frac{D_M^s}{D_T^s} \ u_x^s(l_T^-,t) \ , \\ u_x^m(l_T^+,t) &= \frac{D_M^m}{D_T^m} \ u_x^m(l_T^-,t) \ . \end{aligned} \tag{21}$$

These conditions are to be valid for all t with $t > 0$. Here $u^s(l_T^-, t)$ and $u^m(l_T^-, t)$ denote the concentrations of substrate and metabolite at the right tissue boundary and $u^s(l_T^+, t)$ or $u^m(l_T^+, t)$ the corresponding values at left membrane boundary.

It is supposed that substrate and metabolite mass are measured at donor and receiver sides at different times leading to the fitting functions $v^s(t)$, $v^m(t)$, $w^s(t)$, and $w^m(t)$.

4 Algorithms and Test Environment

In our case, partial differential equations are discretized by the method of lines, see Schiesser [22], which is outlined briefly. More details about the approach chosen are found in Schittkowski [25, 26, 28, 33]. For the first integration area we define 31 equidistant grid points and for the second 11 equidistant lines, i.e., the resulting system of ODEs consists of 82 equations, where we take the structure of the boundary and transition conditions into account.

To approximate the first and second partial derivatives of $u(x, t, p)$ at these points subject to the spatial variable x, different alternatives are available. For our numerical tests, a 5-point difference formula for first and second derivatives is used. The difference formula is adapted at the boundary to accept given function and gradient values.

Whenever a boundary or transition condition is given in Dirichlet-form, we know the value of the boundary function and use it to interpolate or approximate the function $u(x, t, p)$ as described above. In other words the corresponding function value of the right-hand side of the discretized system is replaced by the value given. Alternatively a boundary condition may appear in Neumann-form. In this case the derivative values at the boundary are replaced by the given ones before evaluating the second order spatial derivative approximations.

Ordinary differential equations are added to the discretized system without any further modifications. The resulting system of ordinary differential equations is integrated by the implicit ODE solver RADAU5 of Hairer and Wanner [15]. The absolute and relative error tolerance is set to 10^{-7}.

The resulting least squares problem is solved by the SQP-Gauss-Newton code DFNLP of Schittkowski [24, 30]. The original problem is transformed into a general nonlinear programming problem by introducing additional variables and constraints. It can be shown that typical features of a Gauss-Newton and quasi-Newton method are retained, see Schittkowski [24, 28]. The resulting nonlinear programming problem is solved by a standard SQP code called NLPQLP, cf. Schittkowski [23, 31]. Stopping tolerance for checking the KKT optimality conditions is 10^{-10}.

Experimental design requires the numerical solution of complex optimization problems, where mixed partial derivatives of second order are required to compute the gradient of the objective function, see (11) and (14). As outlined in Section 2.3, forward differences are applied to approximate mixed partial derivatives of $h(p, q, t)$. Since the Fisher information matrix $I(p, q)$ is often ill-conditioned or even rank-deficient, the generalized inverse is evaluated by the Lapack routine DGELSS which applies a singular value decomposition (SVD).

The numerical modules are all combined within one Fortran code called PDEFIT, see Schittkowski [26, 33], which is executed from an interactive data fitting system EASY-FIT, see Schittkowski [27, 28]. The whole package can be downloaded from the home page of the author[1] and the numerical tests of the subsequent section can be repeated. Model functions are provided in form of a special modeling language called PCOMP, see Dobmann et al. [8] or Schittkowski [29], which is based on a Fortran-similar syntax. Since model functions are interpreted during runtime, compilation and linking is prevented, and new model variants are tested rapidly.

[1] `http://www.klaus-schittkowski.de/`

i	t_i	y_i^1	y_i^2	y_i^3	y_i^4
1	0	323.80	-	-	-
2	2	315.79	2.35	1.7	0.78
3	5	306.17	6.56	4.11	2.34
4	7	297.34	9.1	5.67	3.15
5	10	292.15	13.51	7.39	5.32
6	20	265.91	26.66	12.1	11.79
7	30	242.96	39.26	15.06	18.46

Table 1: Experimental Data

Numerical tests are performed on a Dual Core AMD Opteron(tm) processor running with 1.81 GHz under Windows XP64.

5 Numerical Results

5.1 Step 1: Data Fitting and Significance Analysis

The laboratory experiment uses the following design parameters by which experiments are conducted,

$$\begin{aligned} l_T &= 0.005\, mm, \\ l_M &= 0.005\, mm, \\ V_a &= 3,000\, mm^3, \\ F_a &= 63.6\, mm^2, \\ Y_0 &= 324.0\, nmol, \end{aligned}$$

The Michaelis-Menten constant is set to $K_m = 7.0 \cdot 10^{-3}$ as found in the literature, see Boderke et al. [3] for details. Experimental times and measured substrate concentrations are shown in Table 1 for four measurement sets, seven time values, and 27 experimental data to be fitted over a time horizon of 30 *min*.

The laboratory conditions enforce that only a limited number of data can be measured within a very short period of 30 minutes. Corresponding model parameters to be estimated and starting values are found in Table 2. The initial concentration Y_0 is added to the list of model parameters to take possible measurement errors into account.

The least squares code DFNLP computes a solution subject to the termination tolerance of 10^{-10} after 361 iterations and reduces the residual scaled by the sum of scaled squared measurement values from 0.86 to 0.00034. The relatively large number of iterations is one of the difficulties we encounter in case of overdetermined parameter estimation models. Obviously, the area around an optimal solution is very flat and makes it extremely difficult to approximate the true solution exactly within the required accuracy. In our case, the KKT optimality value as reported by DFNLP, is less than 10^{-7} after 49 iterations with residual value 0.00069 and less than 10^{-9} after 65 iterations with residual value 0.00051. We will never know by sure how close this value is to the true solution, and have to expect further changes of model parameter values if we would continue the process, which is numerically impossible in our case.

p	p_0	$\hat{p}$	$\hat{\sigma}$	p_l
D_M^s	0.02	0.0017	0.018	7
D_M^m	0.1	5.32	7.75	1
D_T^s	1.0	26.1	0.0044	2
D_T^m	8.0	9.22	4.48	4
P^s	0.01	0.741	8.45	3
P^m	0.05	0.0051	0.0029	7
V_{max}	30.0	8.02	19.7	7
T^s	1.0	0.376	0.11	5
T^m	1.0	0.53	0.041	6
Y_0	324.0	318.8	1.98	7

Table 2: Parameter Identification

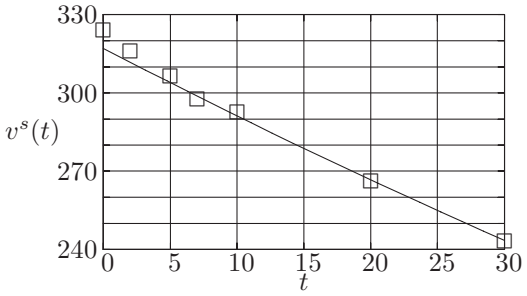


Figure 1: Substrate at Donor

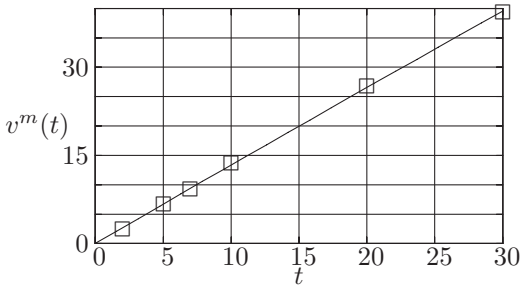


Figure 2: Metabolite at Donor

Figures 1 to 4 show the corresponding plots of the experimental data and the fitting functions $v^s(t)$, $v^m(t)$, $w^s(t)$, and $w^m(t)$. Concentration profiles of substrate and metabolite are found in Figures 5 and 6, where the time axis is extended to $t = 400$.

In Table 2, the starting values p_0, the obtained final solution values $\hat{p}$, and the standard deviations $\hat{\sigma}$ subject to an expected error of 1% of the measurements are reported. The last column, p_l, shows the priority levels as introduced in Section 5.2. Obviously, some parameters got extremely large confidence intervals indicating that a identification of this parameter is not possible. Moreover, 15 pairs of parameters got a correlation coefficient greater than 0.98. We conclude that in this specific situation, six parameters D_M^m, D_T^s, D_T^m, P_s, T^s, and T^m seem to be less relevant then the others and cannot be identified as accurately as needed despite of some small standard deviations.

After fixing these parameters and repeating the test run from the same starting values for the indentifiable parameters, we get the results of Table 3. Their estimations are significantly better than those obtained before, and the maximum correlation coefficient is 0.81. Now the least squares algorithm stops after 73 iterations.

5.2 Step 2: Physical Plausibility Analysis: Steady-State and Mass Balance

By applying simple physical arguments, it is often possible to check in advance whether the mathematical model violates some basic laws, for example mass balance or steady-state conditions. If a model fails to pass these tests, it can be rejected immediately.

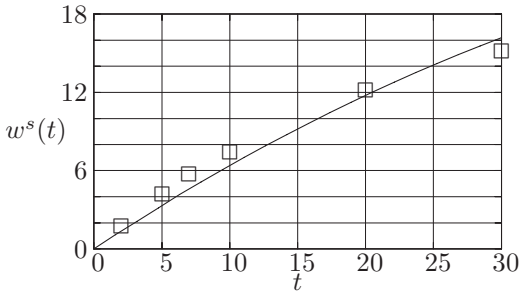


Figure 3: Substrate at Receiver

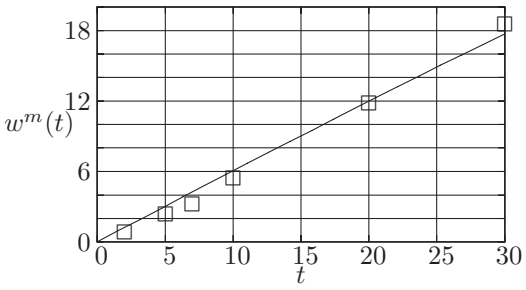


Figure 4: Metabolite at Receiver

p	p_0	$\hat{p}$	$\hat{\sigma}$
D_M^s	0.02	0.0017	0.000012
P^m	0.05	0.0051	0.00029
V_{max}	30.0	8.02	0.036
Y_0	324.0	318.8	1.65

Table 3: Parameter Identification after Priority Analysis

The underlying experiment is assumed to measure substrate and metabolic mass only within a relatively short time interval of 30 min. Longer measurement times are not possible because of additional chemical side effects. An important practical question is, whether the system achieves a reasonable and interpretable steady-state condition, i.e., a constant substrate and metabolite distribution at both sides of the two layers in equilibrium.

We insert the optimal parameters $\hat{p}$ of Table 3 and perform a simulation, i.e., we integrate the system equations from $t = 0$ to $t = 400$. The resulting function plots are shown in Figure 7 to 10. We observe that the system reaches its steady-state where the substrate is completely transformed into the metabolite. The initially given substrate mass permeates completely through both layers.

A further possibility to check the validity of the model, is to compute the mass balance

$$b(t) := \frac{1}{Y_0}\left(v^s(t) + v^m(t) + w^s(t) + w^m(t) + F_a\int_0^l u^s(x,t)dx + F_a\int_0^l u^m(x,t)dx - Y_0\right) . \qquad (22)$$

The mass distribution over a time horizon of 400 min is shown in Figure 11. The error is in the order of the discretization and numerical integration accuracy chosen.

5.3 Step 3: Optimum Experimental Design

The numerical results of the previous sections indicate that D_M^s, P^m, and V_{max} should be considered as model parameters to get estimated. Y_0 is now considered as an experimental design parameter, since the initial concentration at the donor can be predetermined in advance. In addition, we consider also V_a, the volume of the donor, and F_a, the surface of the membrane, as experimental design parameters. The question is whether proceeding from the mathematical model, optimal values can be found without requiring any further experimental data.

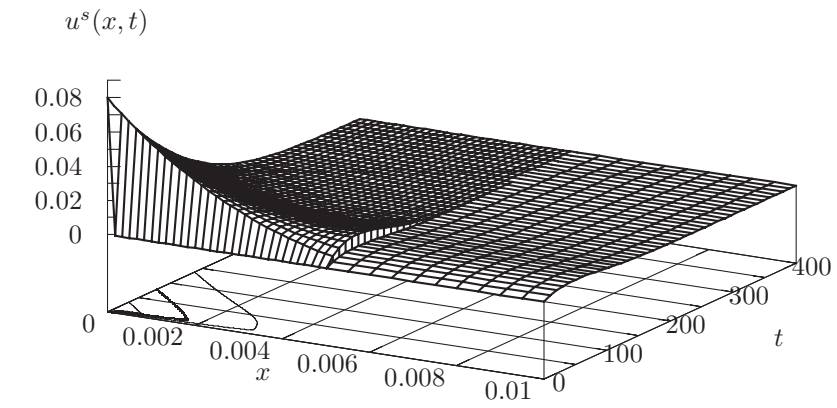


Figure 5: Substrate Concentration

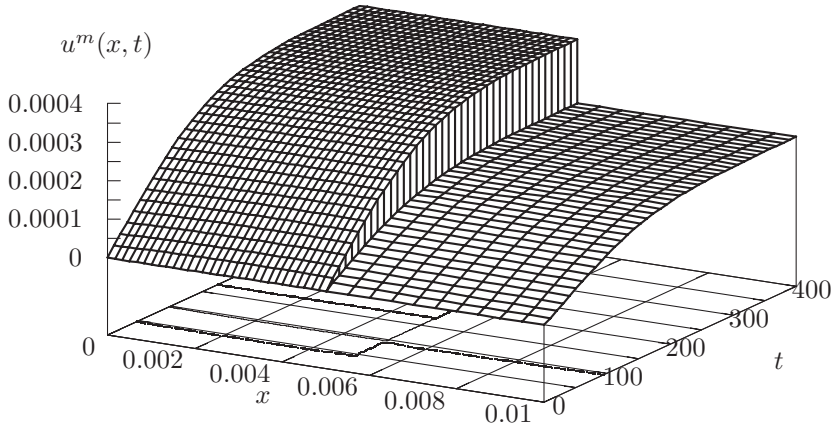


Figure 6: Metabolite Concentration

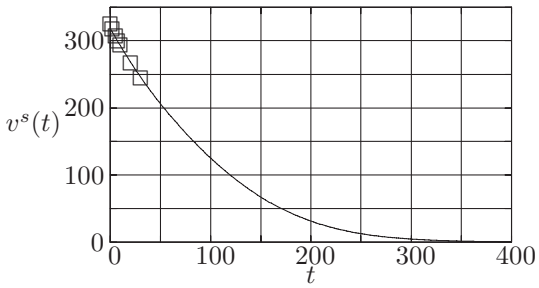


Figure 7: Substrate at Donor

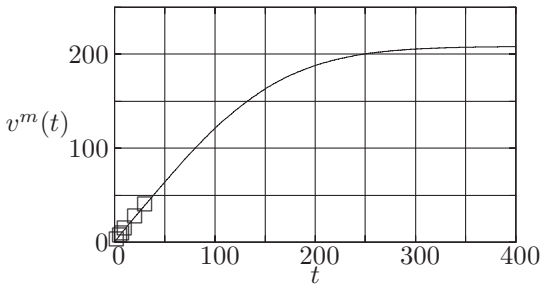


Figure 8: Metabolite at Donor

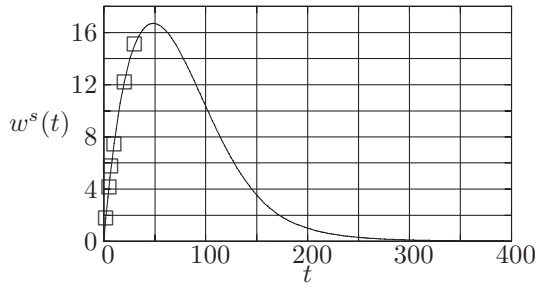


Figure 9: Substrate at Receiver

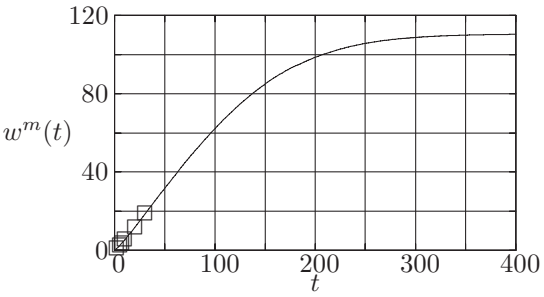


Figure 10: Metabolite at Receiver

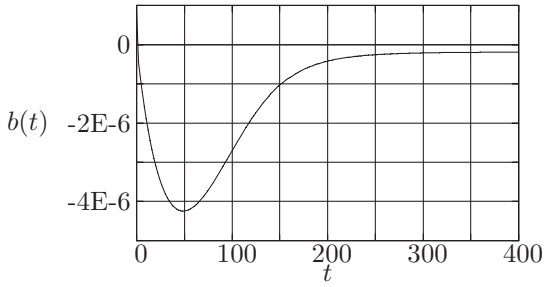


Figure 11: Mass Balance

q	q_0	$\hat{q}$
Y_0	300	1,000
F_a	100	6,211
V_a	1,000	89,517

Table 4: Optimal Design Parameters

p	$\hat{p}$	$\hat{\sigma}$ (original)	$\hat{\sigma}$ (optimum design)
D_M^s	1	0.0035	0.0028
P^m	1	0.098	0.036
V_{max}	1	0.0019	0.0027

Table 5: Parameter Identification before and after Optimum Design

To avoid perturbing scaling effects, model and design variables are scaled to one based on the numerical data achieved so far. Mixed partial derivatives are approximated by forward differences with a tolerance of 0.1 and the A-criterion is applied, see Section 2.3 for details.

After 30 iterations, the code NLPQLP stops at the design parameters denoted by $\hat{q}$ in Table 4. The scaled A-criterion is reduced from 1.0 to 0.0027, where the stopping criterion is set to 10^{-10}. Table 4 contains the initial design parameters q_0 and the optimal ones $\hat{q}$. Y_0 hits its upper bound 1,000.

The standard deviations of the model functions are computed by numerical simulation at the initial and the optimal parameter set by adding subsequently uniformly generated errors in the order of 1%, and evaluating the confidence intervals, see Table 5. The model parameters p are scaled to one. It seems that the larger the membrane surface, the donor volume, and the initial substrate concentration, the smaller are the standard deviations. A general recommendation could be to design these items as big as possible. A significant improvement of the estimate for the metabolite distribution coefficient P^m, the most difficult parameter, is obtained.

The experimental design problem studied so far, uses a constant weight one for each measurement. Now we are interested in the question whether it is possible to predict an optimal distribution of weights of the given time horizon of 30 *min*. By applying the approach introduced in Section 2.3, we add the scaled weights to the existing set of design parameters, and repeat the test run. We then get an optimization problem consisting of 28 variables and an additional linear equality constraint to enforce that the sum over all weights is one.

NLPQLP needs 47 iterations to reduce the scaled A-criterion from 1.0 to 0.014 subject to a termination accuracy of 10^{-10}. The interesting observation is that only three weights are different from zero corresponding to $v^m(10)$, $w^s(7)$, and $w^m(7)$ according to the number of model parameters, $n_p = 3$. Obviously, these parameters are identifiable, since any measurements at these time values lead to an exact fitting with zero residuals.

For illustration, we perform a simple experiment. We predetermine artificial experimental data *by hand* for the three criteria, each at the given time position $t = 7$ or $t = 10$, respectively, namely $y_1 = 300$ for $v^m(10)$, $y_2 = 50$ for $w^s(7)$, and $y_3 = 100$ for $w^s(7)$. The values are significantly different from the simulated curves discussed so far, and are exactly approximated. The least squares code DFNLP needs 9 iterations to compute a perfect fit with residual less than 10^{-12}.

6 Conclusions

We present an approach to compute unknown parameters in a dynamical model consisting of one-dimensional, time-dependent partial and coupled ordinary differential equations by a least squares data fit. The numerical discretization is based on the method of lines and difference formulae for first and second derivatives subject to the spatial variable. The resulting system of ordinary differential equations is integrated by a standard solver, and the least squares optimization problem is solved iteratively by a modified Gauss-Newton method.

A couple of practical tools are proposed to compute optimal designs, to analyze results of a data fitting run based on confidence intervals, and to get significance levels for the estimated parameters. Of particular interest is the possible reduction of experimental measurements to be retrieved, by weight optimization. The feasibility of this approach and the robustness of the numerical implementation are evaluated by a practical case study. The application model is outlined in detail, which is supposed to simulate the transdermal diffusion of drugs. Partial differential equations describe diffusion of a substrate through a metabolically active tissue, where generation of a metabolite is taken into account by the Michaelis-Menten kinetics. The mass fluxes at donor and receiver sides are modeled by coupled ordinary differential equations.

The transdermal diffusion problems serves as a typical case study to analyze a complex mathematical dynamic model in a practical situation. We show how to check which parameters can be identified, which not, and investigate steady-state and mass balance. The goal is to test the validity of our model by numerical experiments, and to compute optimal experimental designs.

A priority analysis based on successive elimination of less significant model parameters reduces their number from 9 to 3. The analysis is based on computing the absolutely largest coefficient of an eigenvector belonging to the smallest eigenvalue of the inverse of the Fisher information matrix, i.e., the inverse of the covariance matrix. A rigorous mathematical theory justifying this approach, is not known to the author. However, all numerical tests performed so far, see also Schittkowski [32] for some more systematic experiments, show its effectiveness.

Experimental design is a very well understood tool. Mixed second derivatives are required for computing the gradients of the design criterion, the main numerical bottleneck for practical applications. It is shown that a careful implementation based on a difference formula could lead to acceptable numerical results. But the stability of the solution still depends on a careful choice of the tolerance for approximating derivatives. More numerical tests and analytical investigations are necessary to find out guidelines how to choose the tolerance in advance in a proper way. Due to the complexity of the basic state equations, analytical derivatives of solutions of differential equations are not available in our case.

By considering weight factors as experimental design parameters, a further hint for the identifiability of the final three model parameters is obtained. It is shown that the total number of positive weights is reduced to three and that the curvature of the fitting criteria is completely identified by these parameters.

Acknowledgment: The paper is dedicated to PD Dr. M. Wolf from Boehringer Ingelheim Pharma KG, who provided the transdermal diffusion model, who motivated the author to do some research in this area, who helped the author to understand practical issues of pharmaceutical laboratory work, and who passed away too early.

References

[1] Banks H.T., Kunisch K. (1989): *Estimation Techniques in Distributed Parameter Systems*, Birkhäuser

[2] Bard Y. (1974): *Nonlinear Parameter Estimation*, Academic Press

[3] Boderke P., Schittkowski K., Wolf M., Merkle H.P. (2000): *A mathematical model for diffusion and concurrent metabolism in metabolically active tissue*, Journal on Theoretical Biology, Vol. 204, No. 3, 393-407

[4] Bortz D.M., Nelson P.W. (2005): *Model selection and mixed-effects modeling in HIV infection dynamics*, Bulletin of Mathematical Biology, Vol. 68, 2005-2025

[5] Caracotsios M., Stewart W.E. (1985): *Sensitivity analysis of initial value problems with mixed ODE's and algebraic equations*, Computers and Chemical Engineering, Vol. 9, 359-365

[6] Carraro T. (2005): *Parameter estimation and optimal experimental design in flow reactors*, Dissertation, Faculty of Natural Sciences, University of Heidelberg

[7] Dennis J.E. jr., Gay D.M., Welsch R.E. (1993): *Algorithm 573: NL2SOL-An adaptive nonlinear least-squares algorithm*, ACM Transactions on Mathematical Software, Vol. 7, No. 3, 369-383

[8] Dobmann M., Liepelt M., Schittkowski K. (1994): *Algorithm 746: PCOMP: A FORTRAN code for automatic differentiation*, ACM Transactions on Mathematical Software, Vol. 21, No. 3, 233-266

[9] Donaldson J.R., Schnabel R.B. (1987): *Computational experience with confidence regions and confidence intervals for nonlinear least squares*, Technometrics, Vol. 29, 67-82

[10] Gallant A.R. (1975): *Nonlinear Regression*, American Statistics, Vol. 29, No. 2, 73-81

[11] Goodwin G.C., Payne R.L. (1977): *Dynamic System Identification: Experiment Design and Data Analysis*, Academic Press, New York

[12] Gorfine M., Freedman L., Shahaf G., Mehr R. (2005): *Maximum likelihood estimator and likelihood ratio test in complex models: An application to B lymphocyte development*, Bulletin of Mathematical Biology, Vol. 65, 1131-1139

[13] Guy R.H., Hadgraft J. (1988): *Physicochemical aspects of percutaneous penetration and its enhancement*, Pharmaceutical Research, Vol. 5, No. 12, 753-758

[14] Hadgraft J. (1979): *The epidermal reservoir, a theoretical approach*, International Journal of Pharmaceutics, Vol. 2, 265-274

[15] Hairer E., Wanner G. (1991): *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*, Series Computational Mathematics, Springer

[16] Hotchkiss S.A.M. (1992): *Skin as a xenobiotic metabolizing organ*, in: *Process in Drug Metabolism*, G.G. Gibson ed., Taylor and Francis, London, 217-262

[17] Lindström P. (1983): *A general purpose algorithm for nonlinear least squares problems with nonlinear constraints*, Report UMINF-103.83, Institute of Information Processing, University of Umea, Umea, Sweden

[18] Ljung L. (1987): *System Identification: Theory for the User*, Prentice Hall

[19] Majer C. (1998): *Parameterschätzung, Versuchsplanung und Trajektorienoptimierung für verfahrenstechnische Prozesse*, Fortschrittberichte VDI, Reihe 3, Nr. 538, VDI, Düsseldorf

[20] Pratt W.B., Taylor P. (1990): *Principles of Drug Action*, Churchill Livingstone, New York

[21] Ryan T.P. (2007): *Modern Experimental Design*, Wiley Series in Probability and Statistics

[22] Schiesser W.E. (1991): *The Numerical Method of Lines*, Academic Press, Inc., San Diego

[23] Schittkowski K. (1985): *NLPQL: A FORTRAN subroutine solving constrained nonlinear programming problems*, Annals of Operations Research, Vol. 5, pp. 485-500

[24] Schittkowski K. (1988): *Solving nonlinear least squares problems by a general purpose SQP-method*, in: *Trends in Mathematical Optimization*, K.-H. Hoffmann, J.-B. Hiriart-Urruty, C. Lemarechal, J. Zowe eds., International Series of Numerical Mathematics, Vol. 84, Birkhäuser, 295-309

[25] Schittkowski K. (1997): *Parameter estimation in one-dimensional time-dependent partial differential equations*, Optimization Methods and Software, Vol. 7, No. 3-4, 165-210

[26] Schittkowski K. (1999): *PDEFIT: A FORTRAN code for parameter estimation in partial differential equations*, Optimization Methods and Software, Vol. 10, 539-582

[27] Schittkowski K. (2002): *EASY-FIT: A software system for data fitting in dynamic systems*, Structural and Multidisciplinary Optimization, Vol. 23, No. 2, 153-169

[28] Schittkowski K. (2002): *Numerical Data Fitting in Dynamical Systems - A Practical Introduction with Applications and Software*, Kluwer Academic Publishers, Dordrecht, Boston, London

[29] Schittkowski K. (2004): *PCOMP: A modeling language for nonlinear programs with automatic differentiation*, in: *Modeling Languages in Mathematical Optimization*, J. Kallrath ed., Kluwer, Norwell, MA, 349-367

[30] Schittkowski K. (2005): *DFNLP: A Fortran Implementation of an SQP-Gauss-Newton Algorithm - user's Guide, version 2.0, Report*, Report, Department of Computer Science, University of Bayreuth

[31] Schittkowski K. (2006): *NLPQLP: A Fortran implementation of a sequential quadratic programming algorithm with distributed and non-monotone line search - user's guide, version 2.2*, Report, Department of Computer Science, University of Bayreuth

[32] Schittkowski K. (2007): *Experimental design tools for ordinary and algebraic differential equations*, to appear: Industrial and Engineering Chemistry Research

[33] Schittkowski K. (2007): *Parameter identification in one-dimensional partial differential algebraic equations*, GAMM-Mitt., Vol. 30, No. 2, 352-375

[34] Schneider R., Posten C., Munack A. (1992): *Application of linear balance equations in an online observation system for fermentation processes*, Proceedings of the IFAC *Modelling and Control of Biotechnical Processes*, Boulder, Colorado, 319-322

[35] Seber G.A.F. (1977): *Linear Regression Analysis*, John Wiley, New York

[36] Seber G.A.F., Wild C.J. (1989): *Nonlinear Regression*, John Wiley, New York

[37] Steinsträsser I. (1994): *The organized HaCaT cell culture sheet: A model approach to study epidermal peptide drug metabolism*, Dissertation, Pharmaceutical Institute, ETH Zürich, Switzerland.

[38] Thomaseth K., Cobelli C. (1999): *Generalized sensitivity functions in physiological system identification*, Annals of Biomedical Engineering, Vol. 27, 607-616

[39] van den Bosch P.P.J., van der Klauw A.C. (1994): *Modeling, Identification and Simulation of Dynamical Systems*, CRC Press

[40] Walter E., Pronzato L. (1997): *Identification of Parametric Models*, Springer

[41] Winer B.J., Brown D.R., Michels K.M. (1971): *Statistical Principles In Experimental Design*, McGraw-Hill

[42] Wolf M. (1994): *Mathematisch-physikalische Berechnungs- und Simulationsmodelle zur Beschreibung und Entwicklung therapeutischer Systeme*, Habilitation, Faculty of Mathematics and Natural Sciences, University of Bonn, Germany