Data Fitting in Partial Differential Algebraic Equations: Some Academic and Industrial Applications

K. Schittkowski

Department of Mathematics, University of Bayreuth, D - 95440 Bayreuth, Germany

Abstract

The paper introduces a numerical method to estimate parameters in systems of one-dimensional partial differential algebraic equations. Proceeding from given experimental data, i.e., observation times and measurements, the minimum least squares distance of measured data from a fitting criterion is computed, which depends on the solution of the dynamical system. We present a typical black box approach that is easily implemented proceeding from some standard numerical analysis tools. Main emphasis of the paper is to present a couple of practical applications from industry and academia, to give an impression on the complexity of real life systems of partial differential equations. The domains of application are pharmaceutics, geology, mechanical engineering, chemical engineering, food engineering, and electrical engineering.

Key words: parameter estimation, data fitting, least squares optimization, partial differential algebraic equations, method of lines, applications, software

1 Introduction

Parameter estimation or data fitting, respectively, is applied in practical situations, where a mathematical model is available to simulate and predict the dynamical structure of the system. The idea is compute unknown model parameters by minimizing the distance of a fitting function from experimentally observed data.

In this paper, we consider one-dimensional partial differential algebraic equations (PDAE) with optional flux functions, coupled ordinary differential algebraic equations, arbitrary fitting criteria, disjoint spatial integration areas, switching times, and dynamical constraints. Despite of the restriction that only
one-dimensional spatial variables are considered, the mathematical model is fairly general and broad from the practical point of view. But even if the underlying mathematical formulation requires a two- or three-dimensional simulation, it is often reasonable to estimate first some unknown model parameters by a least squares fit for a simplified one-dimensional variant.

The underlying data fitting problem is of the form

\[
\min \sum_{i=1}^{l} f_i(p)^2 \\
g_j(p) = 0, \quad j = 1,\ldots,m_e, \\
p \in \mathbb{R}^n : g_j(p) \geq 0, \quad j = m_e + 1,\ldots,m, \\
p_l \leq p \leq p_u.
\]  

(1)

We assume that the parameter vector \( p \) is \( n \)-dimensional and that all nonlinear functions are continuously differentiable with respect to \( p \). Upper and lower bounds are treated independently from the remaining constraints. The assumption that all problem functions must be smooth is essential. All efficient numerical algorithms are based more or less on the Gauss-Newton method and require first derivatives. The fitting criteria \( f_i(p) \) are supposed to depend also on the solution of a system of one-dimensional partial differential equations.

The main emphasis of the paper is to introduce a couple of real life case studies to get an impression on the complexity of practically relevant mathematical models. The considered applications are

- transdermal application of drugs,
- groundwater flow,
- cooling of a hot strip mill,
- drying maltodextrin in a convection oven,
- acetylene reactor,
- fluid dynamics of hydro systems,
- MCFC fuel cells,
- horn radiators for satellite communication.

The dynamical equations are outlined and numerical simulation results for a typical case study are presented. Besides the last one, all models are implemented in a special modeling language called PCOMP, which allows automatic differentiation of nonlinear functions, see Dobmann et al. [18] or Liepelt and Schittkowski [42]. The corresponding codes are available as part of the test problem collection of the data fitting program EASY-FIT, see Schittkowski [54,55] and numerical test runs can be repeated.

The data fitting model, alternative phrases are parameter estimation, nonlinear regression, or system identification, is outlined in Section 2. Possible
extensions of the general formulation are discussed to cover a broad domain of possible applications, for example flux functions, coupled ordinary differential algebraic equations, arbitrary fitting criteria, disjoint spatial integration areas, switching times, and dynamical constraints. Section 3 contains a brief summary of the numerical procedures that are involved to discretize the partial differential equations by the method of lines, to integrate the resulting system of ordinary differential algebraic equations, and to solve the constrained least squares problem efficiently. Only some basic features of the underlying ideas are presented. More details are found in the references and especially in Schittkowski [55]. The real life case studies are listed in Section 4, i.e., the mathematical equations and fitting criteria as detailed as possible, together with some numerical results.

2 The Dynamical Model

We proceed from \( r \) experimental data sets \((t_i, y_i^k), i = 1, \ldots, l_t, k = 1, \ldots, l, \) where \( l_t \) time values and \( l = l_t r \) corresponding measurement values are defined. Then the objective function to be minimized is

\[
\sum_{k=1}^{r} \sum_{i=1}^{l_t} (w_i^k (h_k(p, t_i) - y_i^k))^2 .
\]

The fitting criteria \( h_k(p, t) \) depend now on so-called state variables, i.e., the solution of an implicitly defined system of dynamic equations.

In its simplest form, a time-dependent one-dimensional system of \( n_p \) partial differential equations is given by

\[
u_t = F(p, u, u_x, u_{xx}, x, t) ,
\]

where \( u = (u_1, \ldots, u_{n_p})^T \) are the state variables. We denote the solution of (3) by \( u(p, x, t) \), since it depends on the time value \( t \), the space value \( x \), and the actual parameter value \( p \).

Flux functions are useful in two situations. First, they facilitate the declaration of highly complex model functions given by flux formulations. In these cases, it is often difficult or impossible to get the spatial derivatives in analytical form, and one possibility is to apply a first-order discretization scheme to the entire flux function. Another reason for using flux functions is to design special upwind formulae in case of hyperbolic partial differential equations, when usual approximation schemes break down. A typical reason is the propagation of shocks over the integration interval, enforced by non-continuous initial and boundary conditions. In many situations, advection or transport equations of
are considered with an additional source term on the right-hand side.

Now we permit additional algebraic equations as in the case of ordinary differential algebraic equations, and get the dynamic system

\[
\frac{\partial u_d}{\partial t} = F_d(p, u, u_x, u_{xx}, x, t), \quad 0 = F_a(p, u, u_x, u_{xx}, x, t),
\]

where \( x \in \mathbb{R} \) is the spatial variable with \( x_L \leq x \leq x_R \), and \( 0 < t \leq T \). But now the state variables are divided into \( n_d \) differential \( u_d = (u_1, \ldots, u_{n_d})^T \) and \( n_a \) algebraic variables \( u_a = (u_{n_d+1}, \ldots, u_{n_d+n_a})^T \), where the number of algebraic variables is identical to the number of algebraic equations summarized by the vector \( F_a \).

Initial values and boundary conditions may depend on the parameter vector to be estimated. Since the starting time is assumed to be zero, initial values have the form \( u(p, x, 0) = u_0(p, x) \) and are defined for all \( x \in [x_L, x_R] \). For both end points \( x_L \) and \( x_R \) we allow Dirichlet or Neumann boundary values

\[
\begin{align*}
  u(p, x_L, t) &= u^L(p, t), & u(p, x_R, t) &= u^R(p, t), \\
  u_x(p, x_L, t) &= \hat{u}^L(p, t), & u_x(p, x_R, t) &= \hat{u}^R(p, t)
\end{align*}
\]

for \( 0 < t \leq T \), where \( T \) is the final integration time, for example the last experimental time value \( t_l \). The availability of all boundary functions is of course not required. Their particular choice depends on the structure of the PDE model, for example whether second partial derivatives exist on the right-hand side or not.

In case of algebraic equations, we must treat initial and boundary conditions with more care. We have to guarantee that at least existing boundary conditions satisfy the algebraic equations. If initial values \( u_0(p, x) \) violate the algebraic equations after inserting corresponding approximations for spatial derivatives, it is tried to solve the corresponding system of nonlinear equations numerically. Thus, we have to assume that the resulting DAE is an index-1-system unless it is guaranteed that consistent initial values for the discretized DAE are available, see for example Caracotsios and Stewart [8] for a similar approach.

To indicate that the fitting criteria \( h_k(p, t) \) depend on the solution of the dynamical equation, where \( k \) denotes the index of a measurement set, we use
the notation
\[ h_k(p, t) = h_k(p, u(p, x_k, t), u_x(p, x_k, t), u_{xx}(p, x_k, t), t) \] . (7)

Each set of experimental data is assigned a spatial variable value \( x_k \in [x_L, x_R] \), \( k = 1, \ldots, r \), where \( r \) denotes the total number of measurement sets.

Coupled ordinary differential equations can be used to define a fitting criterion, for example if the in- or outflow of a system is investigated. Algebraic differential equations are highly useful in case of implicit boundary conditions, since coupling positions may coincide with boundary points. The coupled system of ordinary differential algebraic equations is given in the form
\[
\frac{\partial v_j}{\partial t} = G_j(p, u(p, x_j, t), u_x(p, x_j, t), u_{xx}(p, x_j, t), v, t) ,
\]
\[
0 = G_j(p, u(p, x_j, t), u_x(p, x_j, t), u_{xx}(p, x_j, t), v, t) ,
\] (8)

where \( u(p, x, t) \) is the state variable of the partial differential equation and \( x_j \) is any \( x \)-coordinate value where the corresponding ordinary differential or algebraic equation is to be coupled to the partial one, \( j = 1, \ldots, n_c \).

Now we extend the model structure to allow different integration intervals in the \( x \)-direction. A possible application is the diffusion of a substrate through different media, where we want to describe the transition from one area to the next by special conditions. Since these transition conditions may become non-continuous, we need a more general formulation and have to adapt the discretization procedure.

The general model is defined by a system of \( n_d \) one-dimensional partial differential equations and \( n_a \) algebraic equations in one or more spatial intervals, see also Schittkowski [51, 55]. These intervals are given by the outer boundary values \( x_L \) and \( x_R \) that define the total integration interval for the space variable \( x \), and optionally some additional internal transition points \( x^a_1, \ldots, x^a_{m_a-1} \). Thus, we get a sequence of \( m_a + 1 \) boundary and transition points
\[
x^a_0 = x_L < x^a_1 < \cdots < x^a_{m_a-1} < x^a_{m_a} = x_R .
\]

For each integration interval, we get a system of partial differential algebraic equations (5), eventually also with additional flux functions and coupled ordinary differential algebraic equations. A solution depends on the spatial variable \( x \), the time variable \( t \), the parameter vector \( p \), the corresponding integration interval, and is therefore written in the form \( v^i(p, t) \) and \( u^i(p, x, t) \) for \( i = 1, \ldots, m_a \).

Transition conditions between different integration areas are allowed at transition points in Dirichlet or Neumann form depending on the time variable, the parameters to be estimated, and the solution of the neighboring area. More
complex implicit boundary and transition conditions can be defined in the form of coupled algebraic equations.

There are many practical situations where model equations change during the integration over time and where corresponding initial values at the switching points must be adopted. A typical example is a pharmacokinetic application with an initial infusion and subsequent application of drug doses by injection. In this case, it is even possible that the solution becomes non-continuous at a switching point. Thus, we suppose that \( n_b \) break or switching points with

\[
\tau_0 = 0 < \tau_1 < \ldots < \tau_{n_b} < \tau_{n_b+1} = T
\]

are given, where \( T \) is the last experimental time value.

For the first integration interval, the same initial, boundary, and transition values are given as before. For all subsequent intervals, however, the integration subject to the time variable is to be restarted at a switching point and new function values can be provided that may depend now also on the solution of the previous section. Initial values at a switching point are evaluated from

\[
\begin{align*}
    u^i(p, x, \tau_k) &= b^i_k(p, u^i_-(p, x, \tau_k), v_-(p, \tau_k), x), \\
    v(p, \tau_k) &= \tilde{b}(p, v_-(p, \tau_k))
\end{align*}
\] (9)

for \( i = 1, \ldots, m_a \) and \( k = 1, \ldots, n_b \), where \( u^i_-(p, x, \tau_k) \) and \( v_-(p, \tau_k) \) denote the solution of the coupled PDAE system in the previous time interval at \( t = \tau_k \).

Since the right-hand side of the partial differential equation (5) and also the corresponding boundary and transition functions depend on the time variable, they may change from one interval to the next. Particularly non-continuous transitions at switching points are allowed. It is even possible that break points become variables to be adapted during the optimization process.

Additional nonlinear equality and inequality constraints of general form are allowed, see (1). Restrictions are useful to describe certain limitations on the choice of parameter values, for example monotonicity. It is often reasonable to define dynamical constraints, where the restriction functions depend on the solution of the partial differential equation and its first and second spatial derivatives at predetermined time and spatial values, and the solution of coupled ordinary differential equations,

\[
g_j(p) = g_j(p, u^{ij}(p, x_j, t_{k_j}), u_x^{ij}(p, x_j, t_{k_j}), u_{xx}^{ij}(p, x_j, t_{k_j}), v(p, t_{k_j}), t_{k_j})
\] (10)

for \( j = m_e + 1, \ldots, m_r \). Here the index \( i_j \) denotes the corresponding integration area that contains the spatial parameter \( x_j \), and \( k_j \) the corresponding experimental time where a restriction is to be formulated.
A widely used idea is to transform partial differential equations into a system of ordinary differential algebraic equations by discretizing the model functions subject to the spatial variable $x$. This approach is known as the numerical method of lines, see for example Schiesser [48]. For the $i$-th integration interval of the spatial variable, we define a uniform grid and get a discretization of the whole space interval from $x_L$ to $x_R$. To approximate the first and second partial derivatives of $u(x, t, p)$ subject to the spatial variable at a given point $x_k$, $k = 1, \ldots, n_g$, several different alternatives have been implemented in the code PDEFIT, see Schittkowski [53] for more details, which is applied for the numerical tests of this paper:

a) Difference Formulae: Derivatives are approximated by difference formulae at 3 and 5 grid points for first derivatives, that are applied recursively to get also second derivatives. Alternatively, a 5-point difference formula for second derivatives is implemented as well. The difference formulae are adapted at the boundary to accept given function and gradient values. Moreover, first derivatives can be approximated by simple forward and backward differences. They are particularly useful in a situation, where an upwind formula is desirable, but the right-hand side of the PDE is not given in flux form, see below. These difference formulae can be combined and applied individually to the spatial derivatives of the state variables under consideration.

b) Upwind Formulae for Hyperbolic Equations: In case of a scalar hyperbolic equation

$$u_i^t = f_i^x(x, t, u_i^t, p)$$

with a so-called flux function $f$, approximation by polynomials or difference formulae might become unstable, especially if non-continuous boundary conditions are supplied to describe for example the propagation of shocks, see Schiesser [48] for some numerical examples. Thus, special difference formulae have been derived in the past to overcome the instabilities, based, e.g., on so-called upwind formulae. They consist of simple forward or backward differences, if the flux direction is known in advance, or of more advanced higher order TVD formulae. For more information, see the original literature, for example Chakravarthy and Osher[10], Sweby [60], Wang and Richards [64], and Yang and Przekwas [68].

c) Systems of Advection-Diffusion Equations: Systems of non-homogeneous, nonlinear hyperbolic or advection equations depending on an area index $i$, $i = 1, \ldots, n_a$ and $u_i^t \in \mathbb{R}^{n_p}$, $n_p \geq 1$, can be solved by essentially non-oscillatory
(ENO) schemes, see Harten et al [29], Harten [30], or Walsteijn [63]. The equations are of the form

\[ u_t^i + f_x^i(u^i, p) = g^i(x, t, u^i, u_x^i, u_{xx}^i, p) \]  

(12)

High order polynomials are applied to approximate a primitive function, which is supposed to represent the flux function at intermediate spatial grid points. The choice of the corresponding stencil depends on the magnitude of divided differences, to direct the stencil away from discontinuities. To solve also systems of hyperbolic equations, a full eigenvalue-eigenvector decomposition of the Jacobian of the flux function subject to \( u^i \) is performed, and the scalar ENO method is applied to coefficient functions after a suitable transformation. The resulting system of ordinary differential equations is solved either by implicit ODE solvers as before, or by a special Runge-Kutta method with fixed stepsize to satisfy the CLF condition.

If a boundary or transition condition is given in Dirichlet-form, we insert the value of the boundary function directly to interpolate or approximate \( u(x, t, p) \) as described above. 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 and algebraic equations are added to the discretized system without any further modification. Since arbitrary coupling points are allowed, they are rounded to the nearest line of the discretized system. In the same way fitting criteria can be defined at arbitrary values of the spatial variable.

In case of algebraic partial or ordinary differential equations, boundary conditions have to satisfy the algebraic equations. Consistent initial values are computed internally, where some data must be given to serve as starting parameters for the nonlinear programming algorithm. Consequently, we allow only index-1-systems unless it is guaranteed, that consistent initial values for the discretized DAE are available.

It is possible that the right-hand side of a PDAE becomes non-continuous subject to integration time. Thus, it is necessary to supply time values and corresponding initial values depending on the solution of the previous interval, where the integration of the DAE is to be restarted with initial tolerances, for example with the initially given stepsize. The integration in the proceeding interval is stopped at the time value given minus a relative error in the order of the machine precision. Break or switching points are either constant or optimization variables to be adapted by the optimization code.

It can be shown that the resulting large system of ordinary differential equa-
tions becomes stiff in some situations, when discretization accuracy increases. Thus, the usage of implicit solvers is recommended. Since the Jacobian of the discretized right-hand side has a band structure, it is essential that the selected method is capable to exploit sparsity efficiently.

Finally, the resulting least squares problem (1) can be solved by any of the available standard solvers, see for example Dennis et al. [17], Gill et al. [21], or Schittkowski [49]. The algorithms discussed in these references, are based on the Gauss-Newton method and require first derivatives of the fitting criteria subject to the parameters to be estimated.

4 Applications and Numerical Simulation Results

Parameter estimation in dynamic systems is a widely used technique. Only a few examples are listed to present a brief review of some industrial and academic applications. The intention is to give an impression on the complexity of real life mathematical models.

The numerical test results are obtained by a code called PDEFIT, which is designed to solve parameter estimation problems based in one-dimensional, time-dependent partial differential algebraic equations with all the extensions discussed in Section 2, see Schittkowski [53]. Differential algebraic equations are integrated by the implicit code RADAU5, see Hairer and Wanner [28]. The underlying least squares problem is solved by the code DFNLP of Schittkowski [50]. By transforming the original problem into a general nonlinear programming problem in a special way, typical features of a Gauss-Newton and quasi-Newton least squares method are retained. The resulting optimization problem is solved by a standard sequential quadratic programming code NLPQL of Schittkowski [49].

In most cases, model functions are either interpreted and evaluated symbolically by a program called PCOMP, see Dobman et al. [18] that allows in addition automatic differentiation of nonlinear model functions. The source codes for the first six models are contained in the database of the interactive software system EASY-FIT, see Schittkowski [55], among the 1,000 real life and academic data fitting problems published there.

4.1 Transdermal Application of Drugs

We consider the permeation of substrates through cutaneous tissue with simultaneous metabolism. Possible practical background is the analysis and
numerical simulation of transdermal processes for developing new drugs and application devices, see for example Guy and Hadgraft [26]. Especially the effects 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 so-called donor contains a given volume and concentration of a substrate that penetrates through a thin layer of a cutaneous tissue of given size. At certain times, the concentration of the substrate is measured at both sides of the membrane.

The situation is a bit more complicated in the sense that at the same time, another substrate is generated, the so-called metabolite, caused by enzymatic interactions, see Hotchkiss [36] or Hadgraft [27]. The concentration of this new substrate is also measured on both sides of the membrane. It is assumed that metabolism can be described by Michaelis-Menten kinetics, see Pratt and Taylor [47], that the distribution of metabolically active enzymes in the layer is homogeneous, and that the mass transport is one-dimensional along the $x$-axis.

The dynamical variables we need to describe the process are substrate and metabolite concentration, $(u^{s/m}(x, t))$, substrate and metabolite mass at donor, $x = 0 (v^{s/m}(t))$, and substrate and metabolite mass at receiver, $x = l (w^{s/m}(t))$. Two partial differential equations describe the diffusion through the skin based on the Michaelis-Menten effect, see also Steinstr"asser [59] or Boderke et al. [6] or Crank [13] for a more rigorous mathematical treatment of diffusion processes. They are given by

$$
\begin{align*}
    u^{s}_t (x, t) & = D^{s}_T u^{s}_{xx}(x, t) - \frac{V_{\text{max}} u^{s}(x, t)}{K_{m} + u^{s}(x, t)}, \\
    u^{m}_t (x, t) & = D^{m}_T u^{m}_{xx}(x, t) + \frac{V_{\text{max}} u^{s}(x, t)}{K_{m} + u^{s}(x, t)},
\end{align*}
$$

(13)

$0 < x < l_T$ and $t > 0$, where $K_{m}$ is the Michaelis-Menten constant. Diffusion of substrate and metabolite through a membrane are described by the equations

$$
\begin{align*}
    u^{s}_t (x, t) & = D^{s}_M u^{s}_{xx}(x, t) , \\
    u^{m}_t (x, t) & = D^{m}_M u^{m}_{xx}(x, t) ,
\end{align*}
$$

(14)

$l_T < x < l$ and $t > 0$. $D^{s/m}_T$ is the diffusion coefficient of substrate and metabolite in the tissue, $D^{s/m}_M$ the diffusion coefficient of substrate and metabolite in the membrane. Corresponding initial values are $u^{s}(x, 0) = 0$ and $u^{m}(x, 0) = 0$ for all $x$ with $0 \leq x \leq l$.

Accumulation of mass flux at both sides of the two layers leads to four ordinary
differential equations coupled to (13) and (14),
\begin{align}
\dot{v}^s(t) &= F_a D_T^s u_x^s(0,t), \quad \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), \quad \dot{w}^m(t) = -F_a D_M^m u_x^m(l,t)
\end{align}

for $t > 0$. They describe the behaviour of substrate and metabolite at both end points, the mass flux 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 by
\begin{align*}
V_a u^s(0, t) &= P^s v^s(t), & V_a u^m(0, t) &= P^m v^m(t), \\
V_a u^s(l, t) &= P^s w^s(t), & V_a u^m(l, t) &= P^m w^m(t)
\end{align*}

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 at $t = l_T$ leading to
\begin{align}
&u^s(l_T^-, t) = T^s u^s(l_T^+, t), \quad u^m(l_T^-, t) = T^m u^m(l_T^+, t), \\
&D_T^s u_x^s(l_T^-, t) = D_T^m u_x^m(l_T^-, t), \quad D_T^m u_x^m(l_T^-, t) = D_M^m u_x^m(l_T^-, t)
\end{align}

These conditions are to be valid for all $t$ with $t > 0$. The remaining constants are donor volume $V_a$, membrane surface $F_a$, metabolization rate $V_{max}$, and transition parameter $T^{s/m}$ of substrate and metabolite. Parameters to be estimated are $D_T^{s/m}$, $D_M^{s/m}$, distribution coefficients $P^{s/m}$, and initial substrate mass $Y_0$. It is supposed that substrate and metabolite mass are measured at donor and receiver sides for different times leading to the fitting functions $v^s(t)$, $v^m(t)$, $w^s(t)$, and $w^m(t)$.

Fitting criteria and experimental data of a typical experiment are plotted in Figures 1 to 4, see Schittkowski [55]. The time horizon is extended to 1,000 min to analyze the steady state case.
4.2 Groundwater Flow

The mathematical model describes a tracer experiment that was conducted at the lake Gårdsjön in Sweden, see Andersson and Olsson [2], to investigate acidification of groundwater pollution. To conduct the experiment, a catchment of 1,000 m² was covered by a roof. A tracer impulse consisting of lithium-bromide was applied with steady state flow conditions. Tensiometer measurements of the tracer concentration were documented in a distance of 40 m from the center of the covered area.

The diffusion equations proposed by Van Genuchten and Wierenga [61] are chosen by Hoch [35] to analyze the diffusion process and to get a simulation model. A two-domain approach was selected in the form of two equations, to model the mobile and the immobile part of the system. The first one describes the diffusion of the flow through soil by convection and dispersion, the second one the so-called immobile part, the mass transfer orthogonal to the flow direction,

\[
\theta_m \frac{\partial c_m}{\partial t} (x,t) + \theta_{im} \frac{\partial c_{im}}{\partial t} (x,t) = \theta_m D_m \frac{\partial^2 c_m}{\partial x^2} (x,t) - \theta_m V_m \frac{\partial c_m}{\partial x} (x,t) .
\]

\[
\theta_{im} \frac{\partial c_{im}}{\partial t} (x,t) = \alpha (c_m(x,t) - c_{im}(x,t))
\]
for \( t > 0 \) and \( 0 < x < L \). Boundary conditions are
\[
c_m(0, t) - \frac{D_m \partial c_m}{V_m} (0, t) = \begin{cases} 
a, & \text{if } t < t_0 \\
0, & \text{otherwise} \end{cases}
\]
\( c_m(0, t) + \frac{D_m \partial c_m}{V_m} (L, t) = 0 \)
\( c_m(L, t) + \frac{D_m \partial c_m}{V_m} (L, t) = 0 \)
for \( t > 0 \), and initial values are given by \( c_m(x, 0) = 0 \) and \( c_m(x, 0) = 0 \) for
\( 0 < x < L \). Then we evaluate the fitting function
\[
h(t) = c_m(\frac{1}{2}L, t) - \frac{D_m \partial c_m}{V_m} \left( \frac{1}{2}L, t \right)
\]
defined for \( t > 0 \). In the above equations, \( c_m(x, t) \) and \( c_m(x, t) \) denote the
tracer concentrations, \( \theta_m \) and \( \theta_m \) the corresponding water contents, \( D_m \) the
dispersion coefficient, and \( \alpha \) the mass transfer coefficient.

Experimental measurements are given inside the spatial area at \( x = 40 \). Parameters to be estimated are \( p_m, p_m, \) and \( D_m \), whereas \( t_0, V_m, \) and \( a \) are
considered as constants. Figure 5 shows all experimental data and the fitting
criterion. The corresponding surface plots for mobile and immobile parts are
found in Figures 6 and 7. For more details, see Hoch [35] and Schittkowski [55].

4.3 Cooling a Hot Strip Mill

We consider a mathematical model for cooling a thin metal plate of thickness
\( L \) in a rolling mill by water at one side and by surrounding air at the other side.
For simplicity, only one cooling section is considered and we suppose that the
temperature can be measured at both sides of the plate. Moreover, we assume
constant speed and neglect heat transfer orthogonal to the move direction.
Then we are able to apply the standard one-dimensional heat equation

\[ C_p(T(z,t)) \rho(T(z,t)) \frac{\partial T(z,t)}{\partial t} = \frac{\partial}{\partial z} \left( \lambda(T(z,t)) \frac{\partial T(z,t)}{\partial z} \right), \]

where \( T(z,t) \) denotes the temperature at time \( t \) and the spatial position \( z \) orthogonal to the plate surface. The density \( \rho(T) \) and the heat transfer coefficient \( \lambda(T) \) are given by

\[ \rho(T) = k_0^p + k_1^p T, \quad \lambda(T) = k_0^\lambda + k_1^\lambda T. \]

The specific heat capacity \( C_p(T) \) known from piecewise linear interpolation of tabulated data, see also Groch [23] or Kopp and Philipp [39].

Boundary conditions are formulated for air and water, cooling both sides with surrounding temperatures \( T_a \) and \( T_w \), respectively, depending on the operational conditions of the mill. Corresponding heat transfer constants are \( \alpha_a \) in the first and \( \alpha_w \) in the second case. At \( z = 0 \), the plate is cooled by water only between a time interval from \( t = t_{1w} \) to \( t = t_{2w} \), which depends on the speed of the plate. Neumann boundary conditions are obtained by combining Newton and Stefan-Boltzmann laws leading to

\[ \lambda(T(0,t)) \frac{\partial T(0,t)}{\partial z} = \alpha_w(T(0,t) - T_w) + \epsilon(T(0,t)) C (T^4(0,t) - T_w^4) \]

for \( 0 < t_{1w} < t < t_{2w} \) and

\[ \lambda(T(0,t)) \frac{\partial T(0,t)}{\partial z} = \alpha_a(T(0,t) - T_a) + \epsilon(T(0,t)) C (T^4(0,t) - T_a^4) \]

otherwise. In a similar way, the right boundary condition is given for air cooling only, see Seredynski [56]. Initial temperature is set to \( T(z,0) = T_0 \).

For a typical data fitting test run, we consider \( \alpha_w \) and \( \alpha_a \) as unknown parameters to be identified. The specific heat capacity \( C_p(T) \) is defined by linear interpolation of some data. The integration is restarted at the two switching
4.4 Drying Maltodextrin in a Convection Oven

Hot-air drying is a common food preservation process, characterized by the removal of water using air as heat delivering and water removal agent, see for example Hayashi [32]. To model the degradation kinetics during drying, the fundamental laws that govern the drying process are Fick’s second law of diffusion and the energy conservation law. The model was first investigated by Mishkin [44,45], and has been used in the past for the optimization of drying processes, for example by Banga et al. [3,4]. The overall diffusion process is described by

$$\frac{\partial}{\partial t} M_w(x,t) = \frac{\partial}{\partial x} \left( D(M_w(L,t), T(t)) \frac{\partial}{\partial x} M_w(x,t) \right)$$

(21)

defined for all $x$ with $0 \leq x \leq L$ and $t \geq 0$, where the diffusion coefficient $D(M_w, T)$ is a function of temperature $T$ and moisture content $M_w$ at the
right boundary. The thickness \( L \) depends on the actual experiment and is measured.

Initial condition is \( M_w(x,0) = M_{w,0} \) with an initial moisture content \( M_{w,0} \).
There exists a left homogeneous Neumann boundary condition and a right Dirichlet boundary condition \( M_w(L,t) = M_{w,eq}(t), \ t \geq 0 \). The moisture-temperature dependence of polymer solutions viscosity is modeled using the William, Landel and Ferry [66] equation. Assumed is a linear chemical structure and that the polymer is in the rubbery state, see Nelson and Labuza [46] and William et al. [66],

\[
D(M_w, T) = D_r 10^{(C_2 - (T_g - T_r))/(C_2 + (T_g - T_r))}.
\]  

(22)

Coefficients \( D_r, C_2 \), and reference temperature \( T_r \) are unknown parameters that are to be estimated by a least squares fit, and \( C_1 \) is known a priori.

The Gordon-Taylor equation has been widely used to describe the moisture dependence of the glass transition temperature for a binary mixture of biopolymer and plastisizer, see Nelson and Labuza [46],

\[
T_g(t) = \frac{kX_w(t)T_{gw} + T_{gM_D12}(1 - X_w(t))}{kX_w(t) + (1 - X_w(t))},
\]  

(23)

where the water-mass fraction is defined by \( X_w(t) = \frac{M_w(L,t)}{1+M_w(L,t)} \) and where the remaining constants \( k, T_{gw}, \) and \( T_{gM_D12} \) are given. An energy balance at the drying slab is used to model the temperature during drying according to Mishkin [45],

\[
\frac{d}{dt}T(t) = \frac{h_A(t)}{\alpha(t)} (T_{db}(t) - T(t)) + \frac{\lambda_w(t)m_s}{\alpha(t)} \frac{d}{dt} M_w(t)
\]  

(24)

for \( t > 0 \) with initial condition \( T(0) = T_0 \), where \( \alpha(t) = m_s C_p^s + m_s M_w(t) C_p^p \), \( h_A(t) = A_0 P_1 M_w(t) \), \( \lambda_w(t) = \lambda^0_w - \lambda^1_w (T(t) + 273) \). \( P_1 \) is an unknown parameter we want to estimate and \( C_p^s, C_p^p, \lambda^0_w \), and \( \lambda^1_w \) are known constants. Solid mass \( m_s \) and initial temperature \( T_0 \) are determined by the experiment. The circular drying area \( A_0 \) is computed from the measured diameter. The average moisture content is defined by

\[
\bar{M}_w(t) = \frac{1}{L} \int_0^L M_w(\zeta,t) d\zeta.
\]  

(25)

The dry bulb temperature \( T_{db}(t) \) is continuously monitored for each experiment. The raw temperature data are smoothed by B-splines, and values for the sampling times are stored. The equilibrium moisture content is determined by weight difference from the state and from the known initial moisture content.
In order to model the temperature dependence of maltodextrin, a modified GAB model equation is used,

\[ M_{w,eq}(t) = \frac{C(t)K(t)W_m(t)a_w}{(1 - K(t)a_w)(1 - K(t)a_w + C(t)K(t)a_w)}, \quad (26) \]

where \( a_w \) denotes the measured water activity. The model parameters \( C(t) \), \( K(t) \) and \( W_m(t) \) are fitted to an Arrhenius temperature dependence and the data are fitted in a single least squares regression yielding the parameters

\[ C(t) = 0.04e^{\frac{1257.14}{T(t)+273}}, \quad K(t) = 0.65e^{\frac{144.57}{T(t)+273}}, \quad W_m(t) = 0.05e^{\frac{-99.27}{T(t)+273}}, \]

see also Frias et al. [20] for a more detailed outline of the model equations and the experimental design.

To give an impression on a typical data fitting result, we select one of the experimental data sets of Frias et al. [20]. Measured data for moisture content and temperature at different time values are to be fitted to (25) by adapting the model parameters \( D_r \), \( P_1 \), \( C_2 \), \( T_r \), and \( M_{w,0} \). The corresponding data and surface plots in Figures 11, 12, and 13. Because of lack of a sufficiently large number of experimental data, the parameter estimates are poor, highly correlated, and possess large confidence intervals. More accurate results are obtained by Frias et al. [20] by simultaneously taking 18 experimental data sets into account differing only by the initial moisture contents.

4.5 Acetylene Reactor

The computation of optimal feed controls for chemical reactors, especially for tubular reactors, is a well-known technique, see Edgar and Himmelblau [19] or Buzzi-Ferraris et al. [7]. The mathematical model is given as a distributed parameter system consisting of a set of first-order partial differential equations in one space dimension. The chemical reactions and the temperature depend on the spatial variable, whereas the dynamical decrease of the cross-sectional
area caused by coke deposition is time-dependent. In both cases, we know initial values either in the form of time-dependent feed control functions or a constant tube diameter. Alternative approaches to compute optimal reactor feed rates are discussed in Birk et al. [5], see also Liepelt and Schittkowski [41].

We consider a chemical reactor producing acetylene (C$_2$H$_2$), reacting methane (CH$_4$) in natural gas with oxygen. This reaction requires less oxygen compared with complete combustion. The products are quickly quenched to keep the acetylene from being converted entirely to coke, see Wansbrough [65]. During the reaction process, a small part of the carbon is deposited in the reactor as coke. The quantity and its distribution in the reactor depend on the reaction equations. Since it is impossible to measure the cross-sectional area directly, we need a mathematical model that describes the functional dependence of the cross-sectional area upon other system parameters. If the deposition of coke reaches a certain limit, the reactor must be stopped and the tube cleaned.

There are six reactions to be taken into account, see Birk et al. [5] for details, which can be described by a system of eight ordinary differential equations. Let $C_i$ denote the molar concentration of the $i$-th component, $i = 1, \ldots, 8$, with initial values

$$C_1(0,t) = C_{1}^{0}(t) \quad , \quad C_2(0,t) = C_{2}^{0}(t) \quad , \quad C_i(0,t) = 0 \quad , \quad i = 3,\ldots,8 \ .$$

The eight material balance equations depend on the rates of the various reactions,

$$r_i(x,t) = k_i \exp\left(-\frac{E_i}{R}\left(1/T(x,t) - 1/T_r\right)\right)c_i(x,t)$$

for $i = 1, \ldots, 5$, where $c_1(x,t) = C_1^{a1}(x,t)$, $c_2(x,t) = C_1(x,t)C_2^{a2}(x,t)$, $c_3(x,t) = C_6(x,t)C_2^{a5}(x,t)$, $c_4(x,t) = C_3^{a4}(x,t)$, $c_5(x,t) = C_5(x,t)C_2^{a5}(x,t)$, and on the velocity of the mixture in the reactor, because this speed deter-
mines the time that the components spent in the reactor,

\[ v(x,t) \frac{\partial}{\partial x} C_1(x,t) = -r_1(x,t) - r_2(x,t), \]
\[ v(x,t) \frac{\partial}{\partial x} C_2(x,t) = -r_2(x,t) - \frac{1}{2} r_3(x,t) - \frac{1}{2} r_5(x,t), \]
\[ v(x,t) \frac{\partial}{\partial x} C_3(x,t) = \frac{1}{2} r_1(x,t) - r_4(x,t), \]
\[ v(x,t) \frac{\partial}{\partial x} C_4(x,t) = r_3(x,t), \]
\[ v(x,t) \frac{\partial}{\partial x} C_5'(x,t) = \frac{3}{2} r_1(x,t) + r_2(x,t) + r_4(x,t) - r_5(x,t) \]
\[\quad - n(1 - \epsilon) r_4(x,t), \]  
\[ v(x,t) \frac{\partial}{\partial x} C_6(x,t) = r_2(x,t) - r_3(x,t), \]
\[ v(x,t) \frac{\partial}{\partial x} C_7(x,t) = r_2(x,t) + r_5(x,t), \]
\[ v(x,t) \frac{\partial}{\partial x} C_8(x,t) = 2(1 - \epsilon) r_4(x,t). \]  

The reaction constants are \( k_1, \ldots, k_5 \) with five activation energies \( E_1, \ldots, E_5 \), and three reaction orders \( a_1, a_2, \) and \( a_4 \). For the smaller and less important reactions, the stoichiometric order can be used as an estimate for the reaction order. For the other reactions, these parameters have to be derived from the real reactor that is going to be examined. The average temperature \( T_r \) is used to scale the exponential functions and \( R \) denotes the gas constant.

Velocity and density of the mixture are given by

\[ v(x,t) = \frac{\dot{m}(t)}{\rho(x,t) A(x,t)}, \quad \rho(x,t) = \sum_{j=1}^{8} C_j(x,t) M_j, \]  

where the total mass flow \( \dot{m}(t) = \dot{m}_1(t) + \dot{m}_2(t) \) is the sum of the two input flows. \( M_j \) denotes the molar weight of the \( j \)-th component. Since the acetylene reactor is controlled by the feeds of natural gas and oxygen, these are the only components with non-vanishing initial values. Finally, the temperature in the reactor can be described by an ordinary differential equation,

\[ \frac{\partial}{\partial x} T(x,t) = \frac{1}{\rho(x,t) v(x,t) c_p(x,t)} \sum_{i=1}^{5} r_i(x,t) \Delta H_i \]  

with the initial condition \( T(0,t) = T_0 \). The incremental change of the temperature is determined by the rate of heat release for all reactions, which depends
on the total heat capacity

\[ c_p(x, t) = \sum_{j=1}^{8} c_{pj} M_j C_j(x, t) / \sum_{j=1}^{8} M_j C_j(x, t) \] .

The individual heat capacities \( c_{pj} \) are considered to be constant, and the parameters \( \Delta H_i \) denote the known heats of reaction. Coke deposition is modeled by the time-dependent differential equation

\[ \frac{\partial}{\partial t} A(x, t) = -\beta r_4(x, t) \] (30)

with initial condition \( A(x, 0) = A_0 \) and reaction parameter \( \beta \).

To sum up, we get a distributed system in the form of ten first-order partial differential equations. We can select either the time or the spatial variable to transform the equations into an explicit system of partial differential algebraic equations.

To find out whether a given input feed can be identified correctly, a data fitting problem is generated proceeding from \( \dot{m}_k(t) = a_k \) for \( k = 1, 2 \), where \( a_1 \) and \( a_2 \) are the parameters to be fitted, see Birk et al. [5] for details. Corresponding surface plots for temperature and cross sectional area are shown in Figures 14 and 15.

4.6 Fluid Dynamics of Hydro Systems

We consider the flow process in an open rectangular channel. Suppose that \( Q(x, t) \) denotes the discharge of the incompressible fluid, say water, also expressed in the form \( Q(x, t) = U(x, t) A(x, t) \), where \( U(x, t) \) is the velocity averaged over the cross sectional area \( A(x, t) \). It is assumed that there is no
lateral inflow or outflow, and that \( A(x, t) = b \, H(x, t) \) because of the rectangular geometry of the channel, where \( b \) is the constant width and \( H(x, t) \) the height or water level.

Flow dynamics can be modeled by the equations of Saint-Venant [15], see also Cunge and Holly [14], Graf [22], or Abbott and Minns [1]. The first partial differential equation results from the continuity equation to model the flux of the momentum,

\[
\frac{\partial}{\partial t} Q(x, t) + \frac{\partial}{\partial x} \left( \frac{1}{2} g b H(x, t)^2 + \frac{\beta}{A(x, t)} Q(x, t)^2 \right) = -g A(x, t) (F_r(x, t) - F_l(x))
\]

(31)

where \( g \) is the gravitational constant and \( \beta \) the Boussinesq velocity distribution coefficient, \( F_l(x) \) the bed slope of the channel, and \( F_r(x, t) \) the friction slope. A second differential equation is needed for the conservation of mass

\[
\frac{\partial}{\partial t} A(x, t) + \frac{\partial}{\partial x} Q(x, t) = 0 \quad .
\]

(32)

Together with the Manning number \( c \) and the fraction of cross sectional area versus wetted boundary \( R(x, t) \), we use the so-called Chezy formula for computing the friction slope

\[
F_r(x, t) = \frac{c^2 \sqrt{Q(x, t)|Q(x, t)|}}{R(x, t)^{4/3} A(x, t)^2} \quad , \quad R(x, t) = \frac{A(x, t)}{b + 2 H(x, t)}
\]

Equations (31) and (32) define a system of hyperbolic partial differential equations in flux formulation, an important assumption for applying special discretization techniques discussed in Section 3. Initial values describe the discharge and water level distribution at \( t = 0 \), \( Q(x, 0) = Q_0(x) \) and \( H(x, 0) = H_0(x) \). Boundary values are chosen to model a specific situation, where a time dependent inflow at one side is given, say at \( x = 0 \), with an input function \( s(t) \), and an outflow controlled by an underflow gate opening subject to a control function \( u(t) \), leading to

\[
Q(0, t) = s(t) \quad , \quad Q(L, t) = \alpha u(t) \sqrt{g(H(L, t) - H_f)} \quad ,
\]

see Graf [22], where \( H_f \) is the right water level outside the reach and \( \alpha > 0 \) a constant, see Figure (16) for a series of two gates. Since the underlying system is of order 1, we need only two boundary conditions to get a well-defined model.

In many practical situations, channels are connected and form complex networks with various topologies, see Gugat et al. [24] and Gugat [25]. To give a simple example, consider the two serial channels of Figure (16) with the same width for simplicity, which are connected at \( x = L_1 \). Water height can
be controlled by two underflow gates, one between the two channels, one at
the right boundary $x = L_2$. Both are controllable by time-dependent, smooth
functions $u_1(t)$ and $u_2(t)$. The dynamical equation (31) is defined in two dif-
ferent integration areas from $x = 0$ to $x = L_1$, then again from $x = L_1$ to
$x = L_2$, leading to boundary and transition conditions of the form

$$
Q(0, t) = s(t) , \quad Q^+(L_1, t) = \alpha u_1(t) \sqrt{g(H^-(L_1, t) - H^+(L_1, t))} ,
Q^-(L, t) = Q^+(L, t) , \quad Q(L_2, t) = \alpha u_2(t) \sqrt{g(H(L_2, t) - H_f)} .
$$

Here the minus and plus signs denote the corresponding limits from the left
and right side at transition point $x = L_1$.

We consider two serial channels shown above. The inflow $s(t)$ is given by linear
interpolation of some data. We are interested in the question, whether the flow
in the channel can be controlled at the two underflow gates subject to given
water levels. Controlled are the openings of the gates by simple polynomial
control functions, see Gugat et al. [24] or Schittkowski [55] for details. Corre-
sponding discharge and water level surface plots are shown in Figures 17 and
18.
Molten carbonate fuel cells (MCFC) convert chemical energy contained in fuel and oxidizer to electrical energy by a complex electro-chemical reaction. Typically, a mixture of natural gas and steam is fed to an anode channel, where internal reforming and oxidation at an electrode take place. The gas exhausting the anode, moves through a catalytic combustion chamber with additional air supply, and enters the cathode channel from the reverse side leading to a reduction reaction. Carbonate ions are transferred between the electrodes through a solid, an electrolyte, see Figure 19. Meanwhile, fuel cells are commercially available and get increasing interest because of the potential to replace traditional coal or gas burning power systems.

One of the key questions is how to stabilize temperature within the cells to improve performance and operation time, and we are interested in a mathematical model to develop a control strategy. Most available models consider the steady-state case, see for example Koh et al. [38]. Transient models for MCFCs and related techniques are discussed in He and Chen [33], Lukas et al. [43], and Simoglou et al. [58]. The presented approach is based on the equations developed by Heidebrecht and Sundmacher [34], see also Chudej et al. [11].

First, we normalize all quantities and let $t_b$ et $h_b$ et $i\text{ and } z$ be the spatial variables, $0 \leq t \leq 1, 0 \leq z \leq 1$. If $x_{a/c} = (x_{a/c}^1, \ldots, x_{a/c}^n)^T$ denote the mol fractions of $n$ gas components under consideration in anode $(a)$ and cathode $(c)$ channels, the gas phase balance equations can be derived in the form

$$
\frac{1}{T_a(z,t)} \frac{\partial}{\partial z} x_a^i(z,t) = -g_a(z,t) \frac{\partial}{\partial z} x_a^i(z,t) + \sum_{k=1}^m \left( \nu_{k,i}^{ox} - x_a^i(z,t) \nu_{b,k}^{ox} \right) D_k^{ox} R_k^{ox}(z,t) \\
+ \sum_{k=1}^m \left( \nu_{k,i}^{ref} - x_a^i(z,t) \nu_{b,k}^{ref} \right) D_k^{ref} R_k^{ref}(z,t)
$$

$$
\frac{1}{T_c(z,t)} \frac{\partial}{\partial z} x_c^i(z,t) = g_c(z,t) \frac{\partial}{\partial z} x_c^i(z,t) + \left( \nu_{b,1}^{ref} - x_c^i(z,t) \nu_{b,1}^{ref} \right) D_1^{ref} R_1^{ref}(z,t)
$$

for $i = 1, \ldots, n$, where $m$ is the number of all reforming reactions in the
gas phase, $D^\text{ox/ref/red}_k$ the Damköhler number, $T^{a/s/c}(z,t)$ the temperature in anode, solid and cathode, and $g^{a/c}(z,t)$ the convective molar flows in anode and cathode. The remaining constants are not discussed, see Heidebrecht and Sundmacher [34] for details. The reaction rates are computed by

$$R_1^{\text{ref}}(z,t) = \exp \left( E_1^{\text{ref}} w^a(z,t) \right) \left( x_1^a(z,t)x_2^a(z,t) - \frac{x_4^a(z,t)x_3^a(z,t)^3}{K_1^{\text{ref}}} \right),$$

$$R_2^{\text{ref}}(z,t) = \exp \left( E_2^{\text{ref}} w^a(z,t) \right) \left( x_4^a(z,t)x_3^a(z,t) - \frac{x_5^a(z,t)x_3^a(z,t)}{K_2^{\text{ref}}} \right),$$

$$R_k^{\text{ox}}(z,t) = \frac{x_k^{a+2}(z,t)w_k^a(z,t)}{1 + D_k^{II,\text{ox}} w_k^a(z,t)},$$

$$R_1^{\text{red}}(z,t) = -x_5^c(z,t)\sqrt{x_6^c(z,t)} \frac{w_1^c(z,t)}{1 + D_1^{II,\text{red}} w_1^c(z,t)\sqrt{x_6^c(z,t)}}$$

with $w^a(z,t) = 1 - \frac{1}{T^a(z,t)}$, $w^s(z,t) = 1 - \frac{1}{T^s(z,t)}$, and

$$w_k^a(z,t) = \exp \left( E_1^{\text{ox}} w^s(z,t) \right) \exp \left( -\alpha_k^{\text{ox}} n_1^{\text{ox}} D_0^a / T^s(z,t) \right),$$

$$w_k^s(z,t) = \exp \left( E_2^{\text{ox}} w^s(z,t) \right) \exp \left( \alpha_k^{\text{ox}} n_1^{\text{ox}} D_0^a / T^s(z,t) \right),$$

$$w_1^c(z,t) = \exp \left( E_1^{\text{red}} w^s(z,t) \right) \exp \left( -\alpha_1^{\text{red}} n_1^{\text{red}} D_0^c / T^s(z,t) \right),$$

$k = 1,2$. Enthalpy balances lead to the temperature equations

$$\frac{c_p^a(z,t)}{T^a(z,t)} \frac{\partial}{\partial z} T^a(z,t) = -g^a(z,t)c_p^a(z,t) \frac{\partial}{\partial z} T^a(z,t)$$

$$+ \sum_{k=1}^{m} D_{ad,k}^{\text{ref}} D_k^{\text{ref}} R_k^{\text{ref}}(z,t) + q_{as}(z,t), \quad (33)$$

$$\frac{c_p^c(z,t)}{T^c(z,t)} \frac{\partial}{\partial z} T^c(z,t) = g^c(z,t)c_p^c(z,t) \frac{\partial}{\partial z} T^c(z,t) + q_{cs}(z,t)$$

and

$$c_p^s \frac{\partial}{\partial z} T^s(z,t) = \lambda^s \frac{\partial^2}{\partial z^2} T^s(z,t) - D_0^c i_{e}F_{1} + q_{sa}(z,t) + q_{sc}(z,t)$$

$$+ \sum_{k=1}^{m} (D_{ad,k}^{\text{ox}} + n_k^{\text{ox}} D_0^a) D_k^{\text{ox}} P_k^{\text{ox}}(z,t)$$

$$+(D_1^{\text{red}} + n_1^{\text{red}} D_0^c) D_1^{\text{red}} R_1^{\text{red}}(z,t) \quad (34)$$

24
with \( c_p^i(z, t) = c_p^i \sum_{i=1}^{n} x^i_a(z, t) \), \( c_p^i(z, t) = c_p^i \sum_{i=1}^{n} x^i_c(z, t) \). For computing \( q_{as}(z, t) \), \( q_{sa}(z, t) \), \( q_{sc}(z, t) \), and \( q_{cs}(z, t) \), somewhat more complex formulae are available. The ‘+’ sign indicates that only positive values are taken into account. Convecive molar flow densities lead to two further algebraic partial differential equations

\[
0 = \left( -\frac{\partial}{\partial z} g^a(z, t) + \sum_{k=1}^{m} \left( \nu_{b,k}^{ref} D_k^{ref} R_k^{ref} + \nu_{b,k}^{ox} D_k^{ox} R_k^{ox} \right) \right) T^a(z, t) c_p^a(z, t) - g^a(z, t) c_p^a(z, t) \frac{\partial}{\partial z} T^a(z, t) + q_{as}(z, t),
\]

\[
0 = \left( -\frac{\partial}{\partial z} g^c(z, t) + \nu_{b,1}^{red} D_1^{red} R_1^{red} \right) T^c(z, t) c_p^c(z, t) - g^c(z, t) c_p^c(z, t) \frac{\partial}{\partial z} T^c(z, t) + q_{cs}(z, t).
\]

State variables are \( x^i_a(z, t) \), \( x^i_c(z, t) \), \( T^a(z, t) \), \( T^c(z, t) \), \( g^a(z, t) \) and \( g^c(z, t) \), \( i = 1, \ldots, n \), where initial values are set to suitable constants. It is assumed that the system can be controlled at the left-hand side of the anode channel by input feeds for mol fractions, temperature and molar flows. Dirichlet boundary conditions for the right-hand side of the cathode describe the burner outlet concentrations related to the cathode inlet concentrations, where heat exchanges with the environment are taken into account. Boundary functions are completed by two additional homogeneous Neumann conditions for the temperature in the electrolyte solid, see Heidebrecht and Sundmacher [34] for details.

For a simple numerical test, we simulate a typical application by constant temperature and input feed for the molar flow at the anode. The question is whether the potential differences in anode, cathode, and electrolyte, \( D_{a/s/c} \phi \), can be identified correctly by assuming that experimental data for temperature and some mol fractions of gas components are available at anode and cathode outflows. Since the flow directions in anode and cathode channel are known, the corresponding temperature and molar flow spatial derivatives are approximated by backward and forward differences, respectively. The corresponding temperature distributions are shown in Figures 20 and 21.

4.8 Horn Radiators for Satellite Communication

Corrugated horns are frequently used as reflector feed sources for large space antennae, for example for INTELSAT satellites, see Wolf et al. [67]. The goal
is to achieve a given spatial energy distribution of the radio frequency (RF) waves, called the radiation or directional characteristic. The transmission quality of the information carried by the RF signals is strongly determined by the directional characteristics of the feeding horn as determined by its geometric structure.

The electromagnetic field theory is based on Maxwell’s equations consisting of four differential equations relating the electrical field \( E \), the magnetic field \( H \), the electrical displacement, and the magnetic induction to electrical charge density and current density, see Collin [12], Silver [57], or Waldron [62]. Under some basic assumptions, particularly homogeneous and isotropic media, Maxwell’s equations can be transformed into an equivalent system of two coupled equations.

By assuming that the surface of the wave guide has ideal conductivity, and that homogeneous Dirichlet and Neumann boundary conditions at the surface are applied, we get the eigenmodes or eigenwaves for the circular wave guide. Since they form a complete orthogonal system, electromagnetic field distribution in a circular wave guide can be expanded into an infinite series of eigenfunctions, and is completely described by the amplitudes of the modes.

For the discussed problem, only the transversal eigenfunctions of the wave guides need to be considered. The eigenfunctions of the circular wave guide are given in the form

\[
T_{np}^{H} = \frac{2 - \delta_{n0}}{(x'_{np}^2 - n^2)\pi} \frac{J_n(x'_{np}\rho/z)}{|J_n(x'_{np})|} \left\{ \begin{array}{l} \sin n\phi \\ \cos n\phi \end{array} \right\} .
\]  

(36)

Here \( J_n \) denotes the \( n \)-th Bessel function, \( x_{np} \) and \( x'_{np} \) the \( p \)-th root of the \( n \)-th Bessel function and the \( p \)-th root of the first derivative of the \( n \)-th Bessel function, respectively, \( n = 0, 1, 2, \ldots \), \( p = 0, 1, 2, \ldots \). For the last term in (36), either the upper or lower trigonometric function is used. Similarly, \( T_{np}^{E} \) is defined.
In principle, the radiated far field pattern of a horn is determined by the field distribution of the waves emitted from the aperture. On the other hand, the aperture field distribution itself is uniquely determined by the excitation in the feeding wave guide and by the interior geometry of the horn. Therefore, assuming a given excitation, the far field is mainly influenced by the design of the interior geometry of the horn.

Usually, the horn is excited by the $TE_{11}$ mode, which is the fundamental, i.e., the first solution of the wave equation in cylindrical coordinates. In order to obtain a rotational symmetric distribution of the energy density of the field in the horn aperture, a quasi-periodical corrugated wall structure according to Figure 22 is assumed, see Wolf et al. [67] and Johnson and Jasik [37]. To reduce the number of optimization parameters, the horn geometry is described by a set of envelope functions from which the actual geometric data for ridges and slots can be derived. Typically, a horn is subdivided into three sections, an input section, a conical section, and an aperture section. For the input and the aperture section, the interior and outer shape of slots and ridges is approximated by a second-order polynomial, while a linear function is used to describe the conical section. It is assumed that the envelope functions of ridges and slots are parallel in conical and aperture section. By this simple analytical approach, it is possible to approximate any reasonable geometry with sufficient accuracy by the design parameters shown in Figure 22.

![Envelope Function of a Circular Corrugated Horn](image)

Fig. 22. Envelope Functions of a Circular Corrugated Horn

A circular corrugated horn has a modular structure, where each module consists of a step transition between two circular wave guides with different diameters, see Figure 23.

From Maxwell’s equations, it follows that the tangential electrical and magnetic field components must be continuous at the interface between two wave guides. This continuity condition is exploited to compute a relation between the mode amplitudes of the excident $b^k_{E,j}$, $b^k_{H,j}$ and incident $a^k_{E,j}$, $a^k_{H,j}$ waves in each wave guide of a module, see Figure 23, $k = 1, 2$. Then voltage and
current coefficients are defined by

\[
U^k_{H,j} = \sqrt{z^k_{H,j}} (a^k_{H,j} + b^k_{H,j}) , \quad U^k_{E,j} = \sqrt{z^k_{E,j}} (a^k_{E,j} + b^k_{E,j}) ,
\]

where \( z^k_{H,j} \) is the \( j \)-th magnetic and \( z^k_{E,j} \) the \( j \)-th electric field impedance, \( k = 1, 2 \). \( I^k_{H,j} \) and \( I^k_{E,j} \) are given by the same formulae, where only the + sign is replaced by the − sign. From the eigenfunctions, the tangential fields in both areas are obtained,

\[
E_k = \sum_{j=1}^{\infty} \left( U^k_{H,j} e^k_{H,j} + U^k_{E,j} e^k_{E,j} \right) , \quad H_k = \sum_{j=1}^{\infty} \left( I^k_{H,j} h^k_{H,j} + I^k_{E,j} h^k_{E,j} \right) ,
\]

\( k = 1, 2 \), where the tangential field vectors in case of excitation by the \( TE_{11} \) mode are computed from (36). In a similar way, \( h^k_{H,j}(\rho, z, \phi) \), \( e^k_{E,j}(\rho, z, \phi) \), and \( h^k_{E,j}(\rho, z, \phi) \) are defined, see Hartwanger et al. [31] for details. The tangential fields must be continuous at the transition between two wave guides. Moreover, boundary conditions must be satisfied, \( E_2 = 0 \) for \( r_1 \leq r \leq r_2 \). Now only \( n_1 \) eigenwaves in region 1 and \( n_2 \) eigenwaves in region 2 are considered. The electric field in area 1 is expanded subject to the eigenfunctions in area 2 and the magnetic field in area 2 subject to the eigenfunctions in area 1. After some manipulations, in particular interchanging integrals and finite sums, the following relationship between voltage coefficients in region 1 and 2 can be formulated in matrix notation:

\[
\begin{pmatrix} U^2_E \\ U^2_H \end{pmatrix} = \begin{pmatrix} X_{EE} & X_{HE} \\ X_{EH} & X_{HH} \end{pmatrix} \begin{pmatrix} U^1_E \\ U^1_H \end{pmatrix} .
\]

(37)

Here \( U^k_E \) and \( U^k_H \) are vectors, consisting of the coefficients \( U^k_{E,j} \) and \( U^k_{H,j} \) for \( j = 1, \ldots, n_k \), respectively, \( k = 1, 2 \). The elements of the matrix \( X_{EE} \) are given by

\[
X_{EE}^{ij} = \int_0^{r_2} \int_0^{2\pi} e^2_{E,i}(\rho, z, \phi)^T e^1_{E,j}(\rho, z, \phi) \rho \, d\phi \, d\rho .
\]

(38)
In the same way, $X_{HE}$, $X_{EH}$, and $X_{EE}$ are defined. Moreover, matrix equations for the current coefficients are available. Next, the relationship between the mode amplitude vectors $b_k^E$ and $b_k^H$ of the incident waves $b_{E,j}^k$, $b_{H,j}^k$, and $a_k^E$ and $a_k^H$ of the incident waves $a_{E,j}^k$, $a_{H,j}^k$, $j = 1, \ldots, n_k$, $k = 1, 2$, are evaluated. After some algebraic manipulations of matrices, we obtain the total scattering matrix

$$
\begin{pmatrix}
 b_1(p) \\
 b_2(p)
\end{pmatrix}
 = 
\begin{pmatrix}
 S_{11}^* (p) & S_{12}^* (p) \\
 S_{21}^* (p) & S_{22}^* (p)
\end{pmatrix}
\begin{pmatrix}
 a_1 \\
 a_2
\end{pmatrix},
$$

(39)

see Kühn and Hombach [40] or Hartwanger et al. [31], relating the amplitudes at the feed input with those at the aperture. Now the vector $a_1$ describes the amplitudes of the modes exciting the horn, the $TE_{11}$ mode in our case. Thus, $a_1$ is the $2n_1$-dimensional unity vector. The vector $a_2$ contains the amplitudes of the reflected modes at the horn aperture, known from the evaluation of the far field. Only a simple matrix times vector computation is performed to get the modes of reflected waves $b_1(p)$ and $b_2(p)$, once the scattering matrix is known.

The main goal of the optimization procedure is to find an interior geometry $p$ of the horn so that the distances of $b_j^i(p)$ from given amplitudes $\bar{b}_j^2$ for $j = 1, \ldots, 2n_2$ become as small as possible. The first component of the vector $b_1(p)$ is a physically significant parameter, the so-called return loss, representing the power reflected at the throat of the horn. Obviously, this return loss should be minimized as well. The phase of the return loss and further components of $b_1(p)$ are not of interest.

From these considerations, the least squares optimization problem

$$
p \in \mathbb{R}^n : \min \sum_{j=1}^{2n_2} (b_j^i(p) - \bar{b}_j^2)^2 + \mu b_1^1(p)^2
$$

(40)

is obtained. The upper index $j$ denotes the $j$-th coefficient of the corresponding vector, $\mu$ a suitable weight, and $p_l$, $p_u$ lower and upper bounds for the parameters to be optimized. Note also that complex numbers are evaluated throughout this section, leading to a separate evaluation of the regression function of (40) for the real and imaginary parts of $b_2^i(p)$.

**Conclusions**

We show an approach to compute unknown parameters in a dynamical model consisting of partial differential algebraic equations by a least squares data
fit. The dynamical equations are discretized by the method of lines leading to large system of ODEs or DAEs, respectively. Ordinary differential equations and least squares problems can be solved by available standard codes.

The intention of the paper is to present a review on some techniques that are routinely used to estimate data in dynamical systems. The model structure is very flexible and covers a broad and realistic domain. Some real life applications are included, which reflect typical situations in industry and academia. The complexity of practical mathematical models is illustrated and numerical results are included. They show that unknown parameters of realistic dynamical systems can be estimated more or less routinely by available standard algorithms.

References


