

# Model parameter estimation of simplified linear models for a continuous paper pulp digester

Limei Ding <sup>\*</sup>, Thomas Gustafsson, Andreas Johansson

*Division of Systems and Interaction, Luleå University of Technology, SE-97187, Luleå, Sweden*

Received 18 October 2005; received in revised form 30 April 2006; accepted 19 September 2006

## Abstract

A physical model of a continuous paper pulp digester is simplified and two subprocesses selected from the digester are modelled by coupled linear partial differential equations. This study focuses on the parameter identification of the simplified linear models. Finite-dimensional approximation of the model is made and a software package developed for identification of distributed parameter processes is applied. This identification system is developed for flexibility to allow identification for different choices of subprocesses and process variables. Unknown parameters of the subprocess models are estimated and the results are illustrated by process simulation and model validation.

© 2006 Elsevier Ltd. All rights reserved.

*Keywords:* Identification; Continuous digester; Distributed parameter process; Model reduction; Model validation

## 1. Introduction

A continuous paper pulp digester is a complex process and model identification of digesters has been paid attention in the past. Funkquist [1,2] used a grey-box identification approach for model parameter estimation of a continuous digester. Wisniewski and Doyle [3] applied an identification procedure based on normalized moments of an impulse response to identify a set of linear models used for model predictive control of a continuous digester. Alexandridis [4] applied the partial least squares method in model identification based on a Purdue digester model proposed by Wisniewski et al. [5]. Amirthalingam and Lee [6,7] presented subspace identification based inferential control of a continuous pulp digester.

The current study focuses on the estimation of unknown parameters in physical models of two subprocesses in a continuous paper pulp digester, and aims to obtain low order lumped models with the parameters that retain their

physical meanings. From a control and monitoring point of view, the parameters in a physical model bear certain physical meanings, which can be very useful in the controller design or diagnostic procedure. On the other hand, a process model with lower order and less complexity will be easier to implement and can save computation time.

The price to pay for having a low complexity model is often reduced model accuracy. In recent years, control relevant identification (or robust identification) has been frequently studied [8–10]. In this research area, definitions of model error and estimates of a model error model provide tools for robustness considerations in control system design.

In this paper, physical models of two subprocesses in the digester are constructed by coupled partial differential equations (PDEs). Model simplification and approximation are performed and a software package developed for identification of distributed parameter physical models is applied. Although the black-box identification can not directly be applied for the estimation of physical models, it will be applied to construct high-order models both for a subprocess modelling and its model error modelling, and the

<sup>\*</sup> Corresponding author. Tel.: +46 920 491788.  
E-mail address: [limei.ding@ltu.se](mailto:limei.ding@ltu.se) (L. Ding).

high-order models will be utilized for our physical nominal model validation.

This paper is organized as the follows: Next section introduces a simplified physical model and then this infinite-dimensional model is lumped using an approximation algorithms that belongs to a class of approaches called method of weighted residuals. In Section 3, a flexible identification system is briefly introduced, then subprocess choice and variable selection for model identification are made. Section 4 deals with parameter identification of two subprocesses. The simplified model is applied for both subprocesses. In Section 5, model verification is demonstrated and discussed. The conclusions are given in Section 6. In addition, some notations of the physical model are listed in Appendix A, and some detailed mathematical descriptions dealing with the model reduction and the model improvement are supplemented in Appendix B and C, respectively.

## 2. A simplified physical model

### 2.1. Model selection and simplification

A paper pulp digester is a chemical reactor in which a mixture of wood chips and liquor is heated by steam and the chemical reactions called delignification take place. In this process, the lignin, a substance that line the cell walls of wood fibers and keep the wood intact, is dissolved and the cellulose fibers, product of the process, are separated.

A continuous paper pulp digester at Husum pulp mill, M-real, Sweden is shown in Fig. 1. Two subprocesses are selected from this digester for the parameter identification study.

In the last few decades, physical model studies of a continuous paper pulp digester have been frequently reported [5,11–17]. A fundamental model proposed by Bhartiya

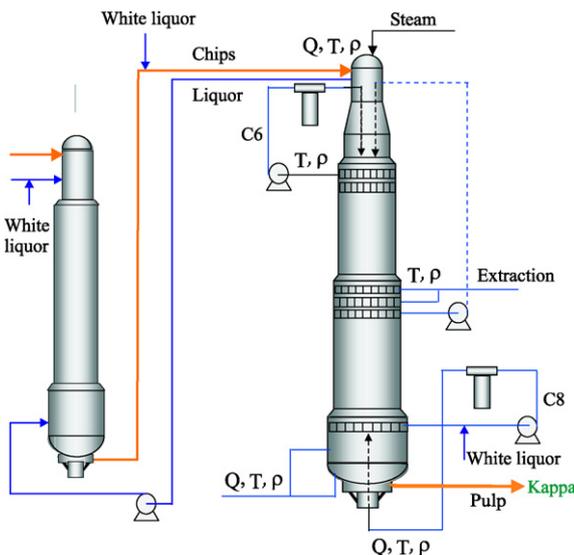


Fig. 1. A continuous paper pulp digester.

[17], 2003, is chosen as process model of the studied digester.

Each volume in a digester is assumed to contain three phases: solid phase, entrapped liquor phase and free liquor phase. The combined solid phase and entrapped liquor phase is referred to as chip phase. The solid phase is assumed to consist of five components: high-reactive lignin ( $s_1$ ); low-reactive lignin ( $s_2$ ); cellulose ( $s_3$ ); araboxylan ( $s_4$ ) and galactoglucomannan ( $s_5$ ). The entrapped liquor is an aqueous solution with six components: active effective alkali (EA) ( $e_1$ ); passive EA ( $e_2$ ); active hydrosulfide (HS) ( $e_3$ ); passive HS ( $e_4$ ); dissolved lignin ( $e_5$ ); and dissolved carbohydrates ( $e_6$ ). The free liquor phase is an aqueous solution of the same six components, as in the entrapped liquor phase.

The subscripts e, f and c denote entrapped liquor, free liquor and chip phase, respectively. The subscript  $i$  denotes the components in solid phase or in liquor phase. In the process the axial distribution of transport properties (along a spatial variable  $z$ ) occurs and the process is modelled as the following partial differential equations.<sup>1</sup>

Free liquor phase thermal energy balances:

$$C_{pf}M_f \frac{\partial T_f}{\partial t} = -T_f \frac{\partial(C_{pf}M_f)}{\partial t} - v_f \frac{\partial(C_{pf}M_f T_f)}{\partial z} - U \frac{1-\eta}{\eta} (T_f - T_c) - \frac{D\epsilon(1-\eta)D_E}{\eta} \pm \frac{\dot{V}_{ext}M_{ext}T_{ext}}{\Delta V_f} \quad (1)$$

Chip phase thermal energy balances:

$$(C_{ps}M_s + C_{pe}M_e\epsilon) \frac{\partial T_c}{\partial t} = -T_c \left( C_{ps} \frac{\partial M_s}{\partial t} + C_{pe} \frac{\partial M_e\epsilon}{\partial t} \right) - v_c \frac{\partial}{\partial z} [(C_{ps}M_s + C_{pe}M_e\epsilon)T_c] + \Delta H_R \sum_{i=1}^5 R_{s,i} + U(T_f - T_c) + D\epsilon D_E \quad (2)$$

Free liquor phase mass continuity:

$$\frac{\partial \rho_{fi}}{\partial t} = -v_c \frac{\partial \rho_{fi}}{\partial z} - D\epsilon \frac{(1-\eta)}{\eta} (\rho_{fi} - \rho_{ei}) \pm \rho_{fi,ext} \frac{\dot{V}_{ext}}{\Delta V_f} \quad (3)$$

Entrapped liquor phase mass continuity:

$$\frac{\partial \rho_{ei}}{\partial t} = -\frac{\rho_{ei}}{\epsilon} \frac{\partial \epsilon}{\partial t} - \frac{v_c}{\epsilon} \frac{\partial(\epsilon \rho_{ei})}{\partial z} + D(\rho_{fi} - \rho_{ei}) + R_{ei} \quad (4)$$

Solid phase mass continuity:

$$\frac{\partial \rho_{si}}{\partial t} = -v_c \frac{\partial \rho_{si}}{\partial z} + R_{si} \quad (5)$$

The rate of mass consumption of solid  $i$  per chip volume can be given based on the kinetic model developed by Christensen et al. [14]:

<sup>1</sup> The notations of the model are listed in Appendix A.

$$R_{si} = -e_f \left( k_{1i} \rho_{c1} + k_{2i} \rho_{c1}^{1/2} \rho_{c3}^{1/2} \right) (\rho_{si} - \rho_{si}^\infty) \quad (6)$$

where  $e_f$  is the reaction rate effectiveness factor. The reaction rate constants  $k_{1i}$  and  $k_{2i}$  are determined using the Arrhenius law which is a nonlinear function of the temperature.

The reaction rates  $R_{ei}$  and  $R_{sj}$  are related via stoichiometric coefficients  $b_{ij}$

$$R_{ei} = \sum_{j=1}^5 b_{ij} R_{sj} \quad (7)$$

For more detail of the model please refer to the original paper [17].

Parameter identification of a complex physical model such as (1)–(7) is a troublesome task. Therefore, we have simplified the model using the following assumptions:

- A1: The dynamic variations of chip porosity, the mass variations of solid and liquor, and the change in external volume flow rate can be neglected.
- A2: The heat released by the exothermic reactions and energy transfer due to diffusion of components between the entrapped liquor and free liquor can be ignored.
- A3: The process works at an equilibrium point and only small variations around this work point are considered.
- A4: Considering the most important components in the chemical reactions taken place in the digester, only lignin in solid phase and effective alkali in liquor phase are taken into account in the reaction rate equations.
- A5: A further simplification is made by considering only temperature behavior. Due to assumption A2, neglecting the reaction heat makes the temperature variables independent of the concentration variables.

The temperature variables are functions of both vertical position  $z$  and time  $t$ . Notation  $T_f(z, t)$  denotes the temperature of free liquor phase, and  $T_c(z, t)$  denotes the temperature of chip phase.

The simplified model is given as the follows:

$$\frac{\partial T_f(z, t)}{\partial t} = -v_f \frac{\partial T_f(z, t)}{\partial z} + d_f (T_c(z, t) - T_f(z, t)) \quad (8)$$

$$\frac{\partial T_c(z, t)}{\partial t} = -v_c \frac{\partial T_c(z, t)}{\partial z} + d_c (T_f(z, t) - T_c(z, t)) \quad (9)$$

where  $d_c = U$ , and  $d_f = U \frac{1-\eta}{\eta}$ . The temperature Eqs. (8) and (9) are the same as that in the model derived by Lundquist [18].

## 2.2. A finite-dimensional approximation of the simplified model

The orthogonal collocation (OC) method is one of the methods of Weighted Residuals [19]. It is commonly used

to obtain a lumped model of a distributed parameter system for process identification, simulation and control [1,2,20], and is widely used and accepted in chemical engineering, especially for the reduction of dynamical models of tubular reactors [20,21]. Many advantages of applying the OC method have been reported in the literature: it may substantially reduce the number of required ordinary differential equations (ODEs) [22]; its formulation and implementation is fairly easy [21,23]; the nature and the dimension of state variables remain unchanged after the reduction procedure [20,21], etc. Furthermore, there is potential to improve the application result by modifying the original OC method, e.g., using the orthogonal collocation on finite elements and the spline collocation technique [24–26]. For the purpose of the current study presented in Section 1, it is well worth to test the OC approximation for direct identification of model parameters based on the physical model, so as to obtain a model with simplicity and accuracy suited for process control and monitoring.

For the simplified model (8) and (9), the following notation is introduced:

$$\frac{\partial x(z, t)}{\partial t} = f \left( x(z, t), \frac{\partial x(z, t)}{\partial z} \right), \quad z \in [0, 1] \quad (10)$$

where

$$x(z, t) = [T_f(z, t) \quad T_c(z, t)]^T = [x_1(z, t) \quad x_2(z, t)]^T \quad (11)$$

and the boundary condition is given as

$$x(0, t) = u(t) \quad (12)$$

where  $u(t)$  is a vector of process inputs.

Note that the spatial variable  $z$  used in this subsection does not refer to the real physical length, although the notation used here is the same as that used in the preceding subsection. Here we define  $z \in [0, 1]$ . For all reduced models applying the OC approximation, that discussed in the rest of this paper, we have  $z \in [0, 1]$ . No matter what the real length is, it can be normalized via a linear variable transformation.

A trial solution of  $x(z, t)$  is chosen as the polynomial:

$$x_T(z, t) = \sum_{j=0}^{N+1} x(z_j, t) l_j(z) \quad (13)$$

where the basis functions  $l_j(z)$  are Lagrange interpolation polynomials with the interpolation points chosen as the interior collocation points  $z_1, z_2, \dots, z_N$ , and the boundary points  $z_0 = 0$  and  $z_{N+1} = 1$ .

The interior collocation points for the interval  $z \in (0, 1)$  should be chosen as the zeros of a Jacobi polynomial  $P_N^{(\alpha, \beta)}(z)$  and this guarantees accuracy of the OC approximation [19].

The first-order spatial derivative can then be approximated as

$$\frac{\partial x(z, t)}{\partial z} \approx \frac{\partial x_T(z, t)}{\partial z} \quad (14)$$

In next section two subprocesses of the digester are selected and the OC method will be applied in the model reduction based on the simplified digester model (8) and (9).

### 3. An identification tool and subprocess selection

Depending on model application purpose, physical models of a distributed parameter system, such as a continuous paper pulp digester, can be built up with more or less complexity. Furthermore, many subprocesses in a complex system need to be modelled appropriately due to the requirements of process control and monitoring in different levels. In practice, we need a flexible identification tool with a rich collection of physical models and model reduction methods, to be able to deal with the identification problems. This work has been carried out and our first stage application result will be presented in this paper.

#### 3.1. An identification tool

A software package is developed for model identification of distributed parameter processes. Fig. 2 demonstrates the structure of the identification tool that contains a flexible simulator. The software package consists of program routines encoded in MATLAB 7.0.1, and the function of the programs can be divided into three categories:

1. *Model implementation:* a collection of physical models of distributed parameter systems works as a model library, and a collection of finite-dimensional approximation

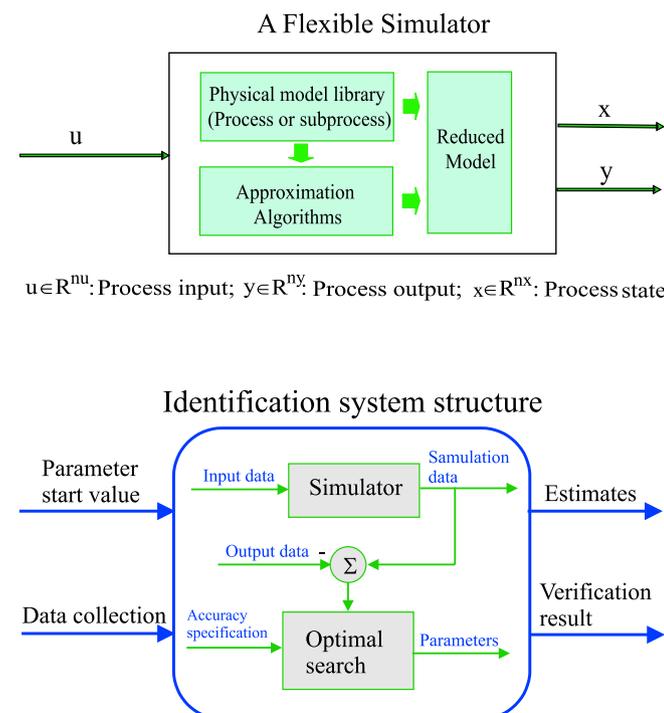


Fig. 2. Upper: a flexible simulator containing in the identification system; lower: identification system structure.

methods as well as linearization methods supports model reductions.

2. *Process simulation:* process variables are selected according to simulation and identification objectives, and the variable selection should be suited with model selection; in addition measured data are treated and supplied appropriately.
3. *Parameter identification:* vectors of unknown parameters and a loss function are defined, estimation start points are given. The data fitting problems are solved in the least-squares sense; finally, the identified model is validated using different methods.

#### 3.2. Subprocess selection

In the current paper, we only discuss the model identification of two subprocesses that is shown in Fig. 3. Subprocess  $G_a$  is selected in the co-current zone from the outlet position of circulation  $c_6$  to the upper extraction. The entire co-current zone is selected as another subprocess denoted  $G_b$ . The temperatures of free liquor and chips are considered as the process variables. The model of the two subprocesses consists of Eqs. (8) and (9).

A finite-dimensional approximation of the model using the OC method is given by the following (where  $j = 1, 2, \dots, N + 1$ ):

$$\begin{aligned} \dot{T}_f(z_i, t) = & -v_f \sum_{j=1}^{N+1} \left. \frac{\partial I_j}{\partial z} \right|_{z=z_i} T_f(z_j, t) + d_f(T_c(z_i, t) - T_f(z_i, t)) \\ & - v_f \left. \frac{\partial I_0}{\partial z} \right|_{z=z_i} T_f(z_0, t) \end{aligned} \quad (15)$$

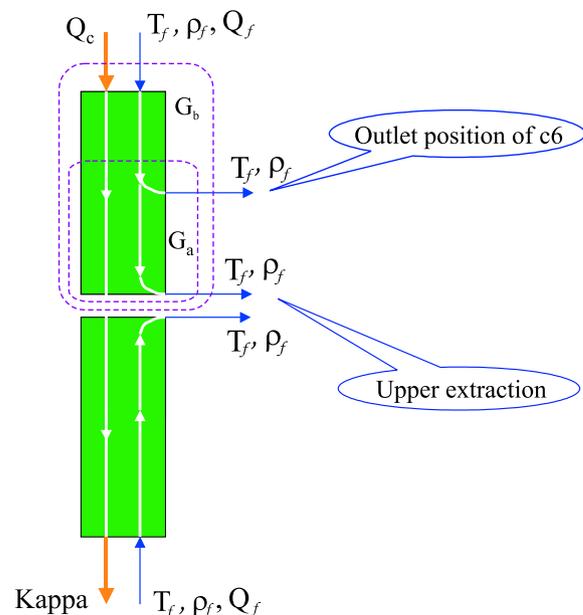


Fig. 3. A digester and its different subprocesses.

$$\begin{aligned} \dot{T}_c(z_i, t) = & -v_c \sum_{j=1}^{N+1} \left. \frac{\partial l_j}{\partial z} \right|_{z=z_i} T_c(z_j, t) + d_c(T_f(z_i, t) - T_c(z_i, t)) \\ & - v_c \left. \frac{\partial l_0}{\partial z} \right|_{z=z_i} T_c(z_0, t) \end{aligned} \quad (16)$$

This model approximation is a linear multi-input–multi-output state space description and it will be explained in more detail in the rest of this paper.

At the top of the digester, both the free liquor stream and the chip stream are heated by steam.

Then the input  $u$  and the output  $y$  for subprocess  $G_b$  are defined as

$$u(t) = T_s(z_{\text{top}}, t), \quad y(t) = T_f(z_{\text{ext}}, t) \quad (17)$$

where  $T_s(z_{\text{top}}, t)$  is the steam temperature at the top of the digester and the output variable  $T_f(z_{\text{ext}}, t)$  is the free liquor temperature at upper extraction.

For subprocess  $G_a$ , we assume that the free liquor temperature at the subprocess input point  $z_{c6}$  influences both the free liquor stream and the chip stream which move downwards from  $z_{c6}$ . Then the input and the output of  $G_a$  are defined as

$$u(t) = T_f(z_{c6}, t), \quad y(t) = T_f(z_{\text{ext}}, t) \quad (18)$$

where  $T_f(z_{c6}, t)$  is the free liquor temperature at the outlet position of c6.

#### 4. Parameter identification

In this section, the unknown parameter identification of  $G_a$  and  $G_b$  will be discussed. The physical model is given by (8) and (9) and their finite-dimensional approximation is obtained as (15) and (16). In common form the state space description can be given as the follows:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned} \quad (19)$$

where the matrices  $A$ ,  $B$ ,  $C$  and  $D$  are given by the following:

$$\begin{aligned} A &= \begin{bmatrix} -v_f A_1 - d_f I & d_f I \\ d_c I & -v_c A_1 - d_c I \end{bmatrix} \\ B &= \begin{bmatrix} -v_f B_1 \\ -v_c B_1 \end{bmatrix} \\ C &= [C_1 \quad 0] \\ D &= 0 \end{aligned} \quad (20)$$

Due to (13) and (14), matrices  $A_1$ ,  $B_1$  and  $C_1$  that are used for the model approximation are defined as the following<sup>2</sup>:

$$A_1 = \begin{bmatrix} \left. \frac{\partial l_1}{\partial z} \right|_{z=z_1} & \cdots & \left. \frac{\partial l_{N+1}}{\partial z} \right|_{z=z_1} \\ \vdots & \ddots & \vdots \\ \left. \frac{\partial l_1}{\partial z} \right|_{z=z_{N+1}} & \cdots & \left. \frac{\partial l_{N+1}}{\partial z} \right|_{z=z_{N+1}} \end{bmatrix} \quad (21)$$

$$B_1 = \begin{bmatrix} \left. \frac{\partial l_0}{\partial z} \right|_{z=z_1} & \cdots & \left. \frac{\partial l_0}{\partial z} \right|_{z=z_{N+1}} \end{bmatrix}^T \quad (22)$$

$$C_1 = [l_1(z_{N+1}) \quad \cdots \quad l_{N+1}(z_{N+1})] \quad (23)$$

The identity matrix “ $I$ ” and the zero matrix “ $0$ ” that appear in the matrices  $A$  and  $C$  have the same dimensions as matrices  $A_1$  and  $C_1$ , respectively.

Four unknown parameters need to be estimated, and the parameter vector is defined as

$$\theta = [v_f \quad d_f \quad v_c \quad d_c]^T \quad (24)$$

The order of a model approximated by the OC method depends on the number of collocation points  $N$ . The OC approximation with  $N = 3, 5, 7, 9$  are chosen and applied in the identification.

The model agreement resulted from the parameter estimation can be indicated using a loss function  $V(\theta)$  defined as

$$V(\theta) = \sum_{i=1}^n (y_s(i) - y_m(i))^2 \quad (25)$$

where  $y_s$  is the simulated output,  $y_m$  is the measured output and  $n$  is the number of data points.

Table 1 demonstrates the start value of  $V(\theta)$  and the final value of  $V(\theta)$  obtained during model identification of  $G_a$  and  $G_b$ . It can be seen from Table 1 that for both  $G_a$  and  $G_b$ , the estimate based on the model approximation with  $N = 5$  gives a smaller value of  $V(\theta)$  comparing to the results using other order approximations.

In Fig. 4, the upper plots demonstrate the frequency properties of the models estimated from the OC approximations with  $N = 3, 5, 7, 9$ , and the lower plots show the spectral of the process input and output data collected from the real plant. The bandwidth of the OC models increase slightly with the increase of the model order. The Bode plot of the model with  $N = 5$  has a noticeable magnitude peak within a frequency band around 0.1. Similarly, in data spectral plots, there are a set of signals within this frequency band have greater magnitude compared with

Table 1  
Identification of models of different order

Number of collocation points	$N = 3$	$N = 5$	$N = 7$	$N = 9$
$G_a$				
Number of iterations	121	144	77	135
Start value of $V(\theta)$	668.916	898.613	863.024	866.232
Final value of $V(\theta)$	255.419	246.662	249.131	251.038
$G_b$				
Number of iterations	18	32	319	271
Start value of $V(\theta)$	243.26	199.264	211.303	228.232
Final value of $V(\theta)$	235.073	186.864	209.177	208.645

<sup>2</sup> For details please see Appendix B, and refer to [20].

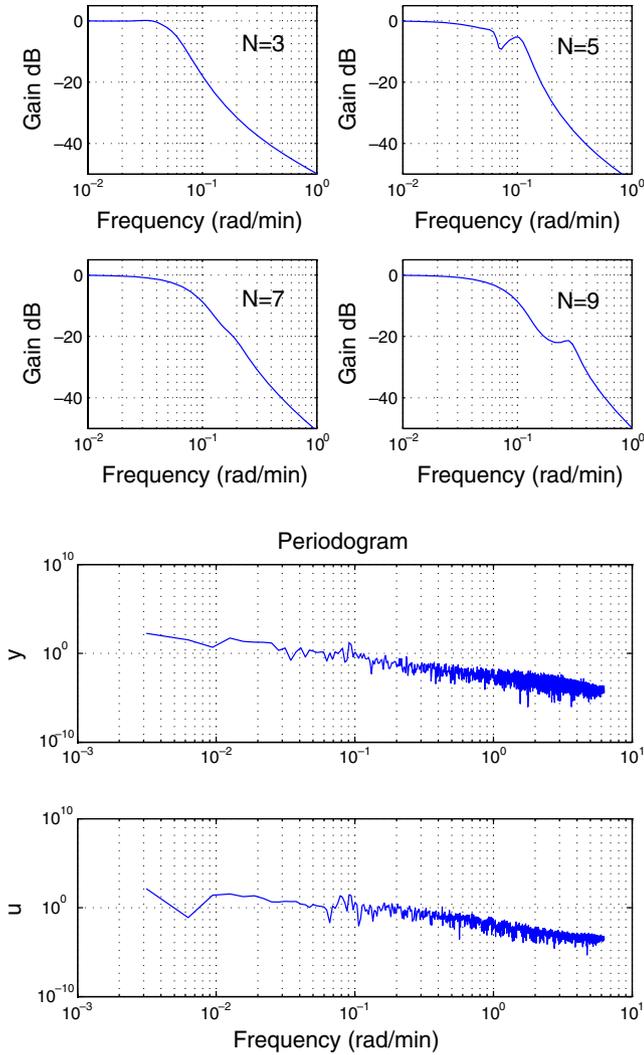


Fig. 4. Bode diagrams of models for  $G_b$  using final parameter estimates (the top steam temperature is used as a single input).

the signals in higher frequencies and the signals in an adjoining lower frequency area. It may be a reason that the OC model with  $N = 5$  results a smaller value of final  $V(\theta)$  in the identification.

Fig. 5 demonstrates the measured data (solid line) and process simulation (dashed line) using the obtained parameter estimates. The upper figure and the lower figure show the data of  $G_a$  and  $G_b$ , respectively, where the number of collocation points is chosen as  $N = 5$ .

It can be seen from the figure that the simulation of  $G_b$  agrees with the measured data better than the simulation of  $G_a$ . In addition, the identification of  $G_b$  resulted smaller final  $V(\theta)$  than the identification of  $G_a$ , and this also indicates that the identification of  $G_b$  provides better parameter estimates. Since the temperature of the chip stream in  $G_a$  at c6 outlet position is influenced not only by the liquor temperature at this position, but also the top steam temperature, the identification of  $G_a$  can be improved by modifying its model structure. We assume that the input of the subprocess is a combination of liquor temperature

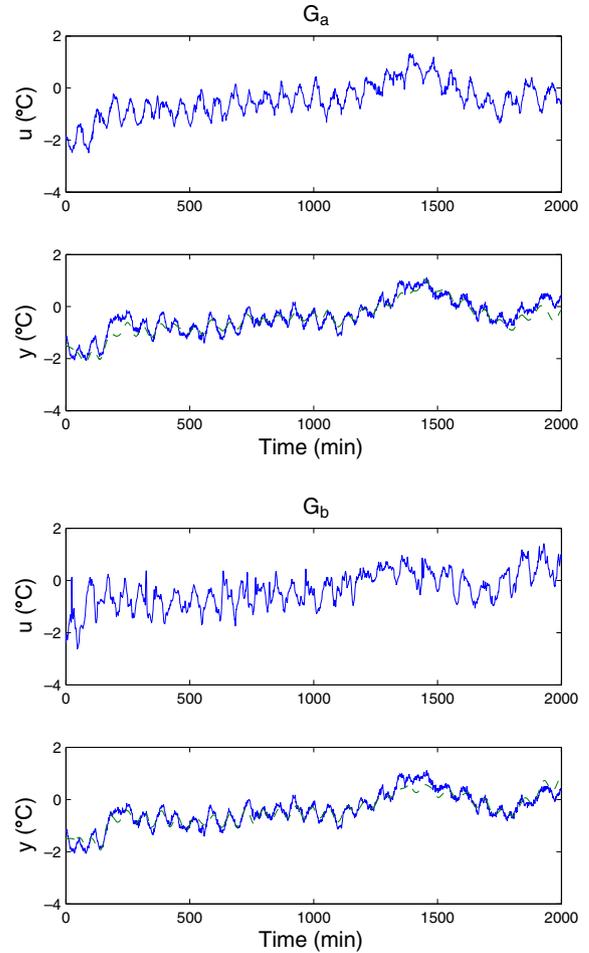


Fig. 5. Measured data and process simulation of  $G_a$  (upper) and  $G_b$  (lower).

at the outlet point of c6 and a delayed steam temperature, then the process input and the output can be given as following:

$$\begin{aligned} u(t) &= [T_s(z_{\text{top}}, t - \tau) \quad T_f(z_{\text{c6,out}}, t)]^T \\ y(t) &= T_f(z_{\text{ext}}, t) \end{aligned} \quad (26)$$

where  $\tau$  is a pure time delay from the top of the digester to the outlet point of c6.

The process model is modified by adding an additional partial differential equation given below<sup>3</sup>:

$$\frac{\partial T_d(z_d, t)}{\partial t} + \frac{1}{\tau} \frac{\partial T_d(z_d, t)}{\partial z_d} = 0 \quad (27)$$

and (27) represents a time delay. The boundary condition of the equation is

$$T_d(0, t) = T_s(z_{\text{top}}, t) \quad (28)$$

where  $T_d(z_d, t)$  denotes a delayed steam temperature over the length from the digester top to the outlet position of c6. The vertical spatial variable  $z_d \in [0, 1]$  is defined from

<sup>3</sup> In Appendix C the equation is explained in detail.

a linear transformation of variables  $z_d = (Z_d - z_{c6}) / (z_{top} - z_{c6})$ , where  $Z_d \in [z_{top}, z_{c6}]$ . The parameter vector of the modified model is then defined as

$$\theta = [v_f \quad d_f \quad v_c \quad d_c \quad \tau]^T \quad (29)$$

The matrices of the improved state space model are given below:

$$A = \begin{bmatrix} -\frac{1}{\tau}A_1 & 0 & 0 \\ 0 & -v_f A_1 - d_f I & d_f I \\ \tilde{0} & -v_c B_1 & d_c I & -v_c A_1 - d_c I \end{bmatrix}$$

$$B = \begin{bmatrix} -\frac{1}{\tau}B_1 & 0 \\ 0 & -v_f B_1 \\ 0 & 0 \end{bmatrix} \quad (30)$$

$$C = [0 \quad C_1 \quad 0]$$

$$D = 0$$

where “ $I$ ” and “ $0$ ” in  $A$  have the same dimension as  $A_1$ , and “ $0$ ” in  $B$  and  $C$  have the same dimensions as  $B_1$  and  $C_1$ , respectively, and “ $\tilde{0}$ ” in  $A$  is a zero matrix by  $(N + 1) \times N$ .

The measured and simulated data of  $G_a$  using the improved model structure is demonstrated in Fig. 6. Comparing it with the old result which is shown by the upper figure in Fig. 5, the agreement between the model simulation and the measured data is obviously improved.

The parameter estimates for  $G_a$  using model (30) and for  $G_b$  using model (20) are shown as the follows:

$$\hat{\theta}_{G_a} = [\hat{v}_f \quad \hat{d}_f \quad \hat{v}_c \quad \hat{d}_c \quad \hat{\tau}]^T \quad (31)$$

$$= [0.0453 \quad 0.2262 \quad 0.0133 \quad 0.046 \quad 29.7478]^T$$

$$\hat{\theta}_{G_b} = [\hat{v}_f \quad \hat{d}_f \quad \hat{v}_c \quad \hat{d}_c]^T \quad (32)$$

$$= [0.0131 \quad 0.0093 \quad 0.0082 \quad 0]^T$$

Due to the normalization of vertical lengths and the different model structures, the estimates for the two subprocesses  $G_a$  and  $G_b$  can not be compared directly.

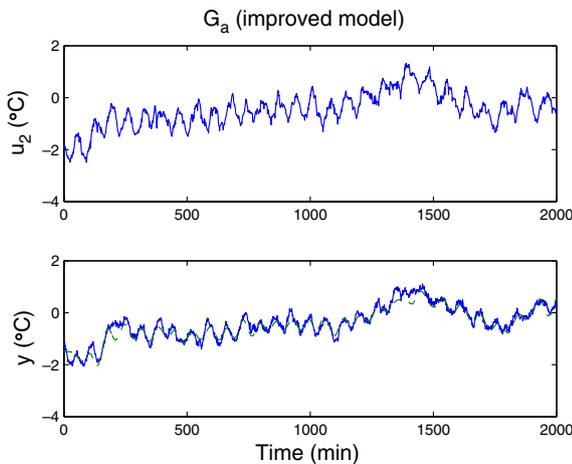


Fig. 6. Measured and simulated data of  $G_a$ , the improved model is used in simulation.

### 5. Model validation and discussion

The models of the two subprocesses with the obtained parameter estimates are verified using validation data measured from the real process. Fig. 7 demonstrates the validation results of  $G_a$  and  $G_b$ , where the measured data are plotted using solid lines and the simulated data are plotted using dashed lines. Comparing Fig. 7 with the results using the identification data (shown in Figs. 5 and 6), there is no

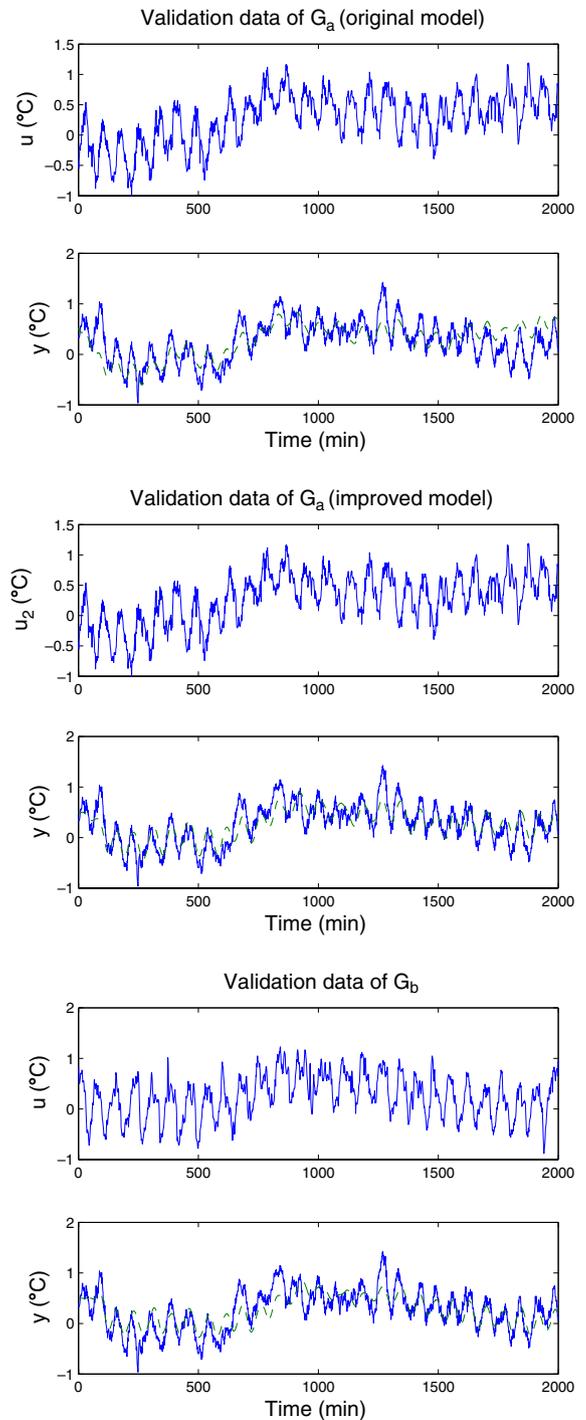


Fig. 7. Validation data and model simulation of  $G_a$  and  $G_b$  (solid line: measurement; dashed line: simulation).

obvious deterioration of the data agreement. In addition, it can be seen from the results using both the validation data and the identification data, that the model simulation of  $G_a$  with the parameter estimates from the improved model are obviously better than that from the original model.

The mean squares error (MSE) of the residual  $\varepsilon$  is defined as

$$\text{MSE} = \frac{1}{n} \left( \sum_{i=1}^n \varepsilon^2(i) \right)^{\frac{1}{2}} \quad (33)$$

where

$$\varepsilon(i) = y_s(i) - y_m(i) \quad (34)$$

and  $n$  is the number of data points. The MSE of the identification results using both identification data and validation data are shown in Table 2.

The quality of the models with parameter estimates can be observed in Fig. 7 and Table 2. Firstly we compare the identification results of  $G_a$  and  $G_b$  using the original model presented by (8) and (9). In Table 2, the MSE values of  $G_b$  are less than those of  $G_a$ . In Fig. 7, the model simulation of  $G_b$  agrees with the real process data better than  $G_a$ . This implies that the original model is more appropriate to be applied for subprocess  $G_b$  than for subprocess  $G_a$ .

Then the identification results of  $G_a$  using the original model and the improved model will be compared. In Section 4 the original model consisting of (8) and (9) is improved for subprocess  $G_a$  by adding an additional partial differential equation that represents a time delay. In Table 2, the MSE values of  $G_a$  using the improved model are much less than the results using the original model, and they are even less than the results of  $G_b$ . In Fig. 7, the data agreement shown in the middle figure is much better than that shown in the upper figure. This implies that the improved model gives a better representation for the process dynamics of  $G_a$ .

Furthermore, the deteriorations of MSE using the validation data compared with that using the identification data are limited in an acceptable degree.

We express a real plant as follows:

$$y = Gu + v \quad (35)$$

where  $y$  is the observed output variable of the plant,  $G$  is the true process and  $v$  is the system noise. Eq. (35) can be a continuous time description or a discrete time description, since both continuous time models and discrete time models will be discussed in the rest of this section.

Table 2  
Mean squares error of identification results

Data	Ident. data	Valid. data
$G_a$ Model (20)	0.0042	0.0047
$G_a$ (improved) Model (30)	0.0030	0.0036
$G_b$ Model (20)	0.0034	0.0039

In Section 4, the frequency properties of the identified OC models illustrated in Fig. 4 shows that the amplification of the models at high frequencies decrease fast. Thus the OC models are not able to represent high frequency dynamics well. In order to construct a model that can represent the subprocesses better in the high frequency region, a high-order black-box model of  $G_b$  is estimated using real

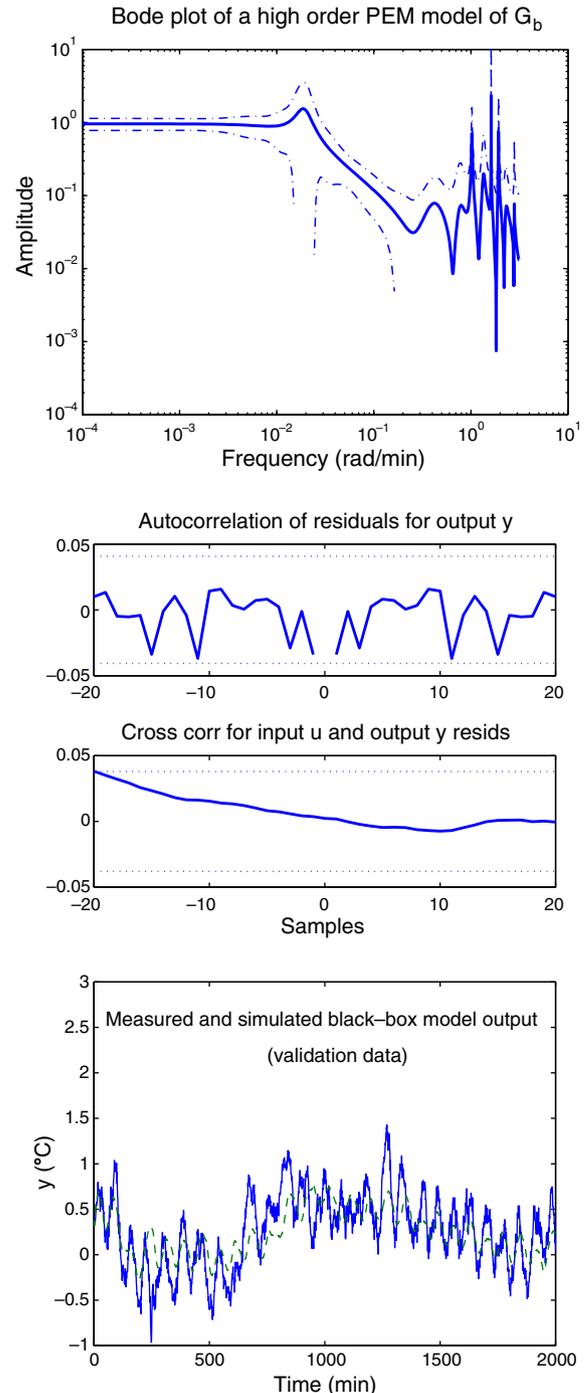


Fig. 8. Upper: Bode plot of a high-order black-box PEM model of  $G_b$  with 99% confidence interval; middle: correlation functions of the PEM model (applying validation data); lower: validation output data (solid line: measurement, dashed line: simulated black-box model output).

plant data with the help of a black-box identification tool [27]. The model structure is

$$\hat{y}(t) = \frac{b_0 + b_1q^{-1} + \dots + b_nq^{-n}}{1 + f_1q^{-1} + \dots + f_nq^{-n}}u(t) + \frac{1 + c_1q^{-1} + \dots + c_mq^{-m}}{1 + d_1q^{-1} + \dots + d_mq^{-m}}e(t) \quad (36)$$

where  $u(t)$  is the measured process input and  $e(t)$  is assumed to be white noise. For the experiments in the sequel, we choose  $n = 20$  and  $m = 5$ . The frequency properties of the OC models will be compared to the high-order black-box model.

In Fig. 8, the upper part demonstrates the Bode plot of the high-order black-box model of  $G_b$ . Its 99% confidence region is illustrated using dash-dotted lines. The model is estimated by applying the prediction error method (PEM). The middle part represents the correlation functions of the PEM model and shows that the model passes the correlation test. The second term on the right side of (36) is a prediction of  $v$  given in (35), and  $v$  is obviously not white due to its model structure. The lower part plots the measured and simulated output data.

The output of  $G_b$  predicted by the OC models,  $y_s = \hat{G}_u$ , are different from the observed output  $y$ , and the difference can be represented using a model residual  $\varepsilon$ :

$$y = \hat{G}u + \varepsilon \quad (37)$$

Considering the OC model structure (19), (20) and (30), and comparing them with (35) and (36), one sees that the OC models have a lack in noise model representation,

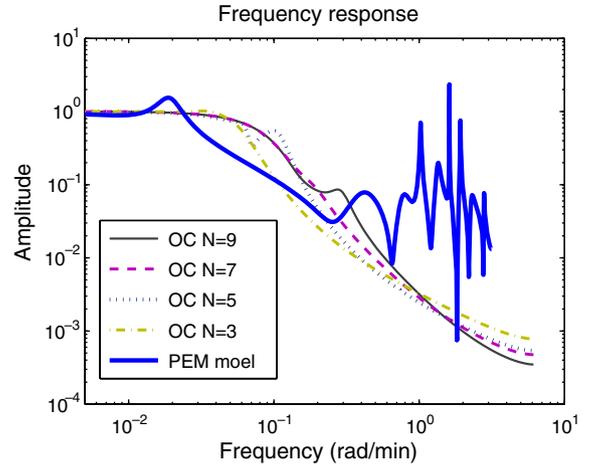


Fig. 9. Comparing frequency properties of the high-order black-box model and the OC physical models.

and the model residual  $\varepsilon$  may contain the model error caused by this lack. The estimated black-box model of  $G_b$  shows that the process noise can be represented using a fifth order model. We suppose that the residual  $\varepsilon$  of  $\hat{G}_b$  is not white, and the OC models of  $G_b$  will fail the residual correlation test. Furthermore, Fig. 9 demonstrates the frequency properties of the black-box model and the OC physical models of  $G_b$ . The magnitude of all the OC models decrease faster than the PEM process model in frequencies higher than  $\omega \approx 0.2$ . This implies that the high-order black-box model represents the subprocess better in the high frequency region compared to the OC models.

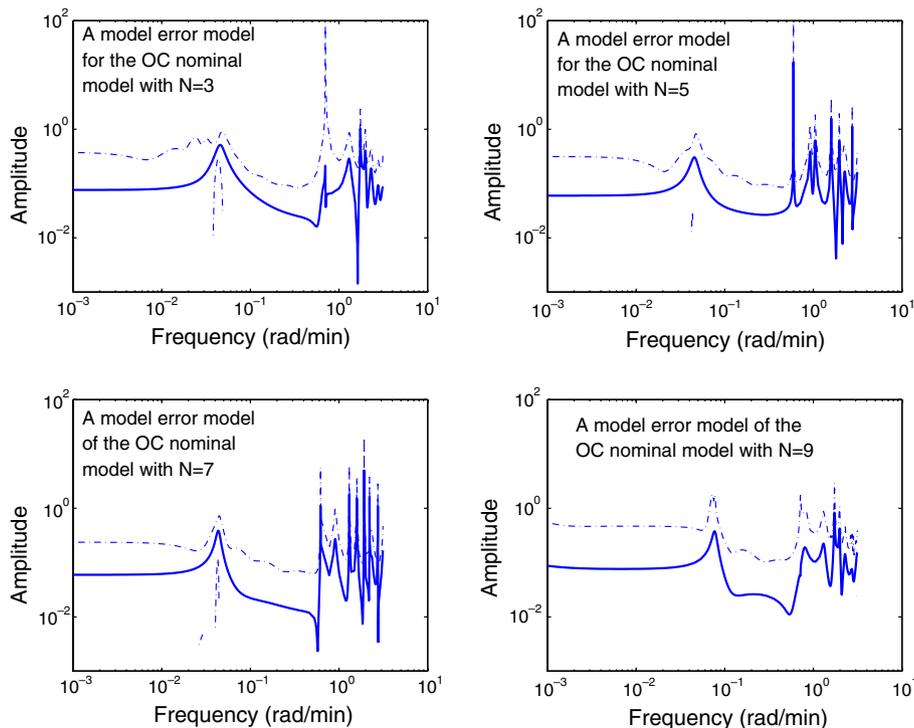


Fig. 10. Frequency responses of the high-order black-box model error models (applying validation data).

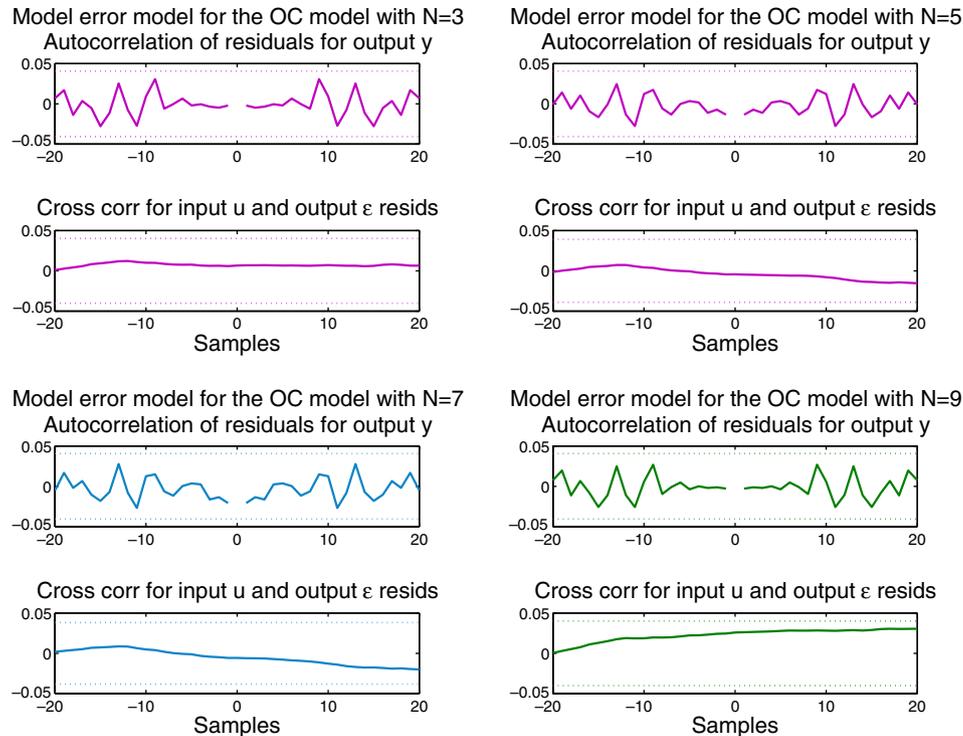


Fig. 11. Correlation functions of the high-order black-box model error models (applying validation data).

From a point of view of control oriented model validation [8], the model residual  $\varepsilon$  can be separated into model error  $\Delta$  and disturbances  $w$ :

$$\varepsilon = \Delta u + w \quad (38)$$

where no further requirement is put on  $w$ . A nominal model could fail in a traditional correlation test, but pass a control oriented model validation test if the norms of its model error model  $\Delta$  and disturbance  $w$  are limited in an accepted level [8]. In addition, an advantage with the model error model concept is that the nominal model still can be used for control design, even if it is falsified by  $\Delta$ . This is the case, for example, if the falsification takes place in “unimportant” frequency regions [9]. Thus it is necessary to construct model error models for the OC nominal models obtained in our study.

A high-order black-box model can be estimated using  $\varepsilon$  as the output data and  $u$  as the input data. The model structure is chosen to be the same as the black-box process model given by (36) with  $n = 20$  and  $m = 5$ . Fig. 10 demonstrates the frequency responses of the model error models for the OC nominal models. The 99% confidence regions of the model error models are illustrated by dash-dotted lines. The corresponding correlation functions of the model error models are shown in Fig. 11, and all the estimated models pass the correlation test.

The magnitudes of the model error models can be easily compared using Fig. 12. The model error model for the OC model with  $N = 5$  has smaller magnitude at the frequencies lower than  $\omega \approx 0.02$  compared to the OC models in other

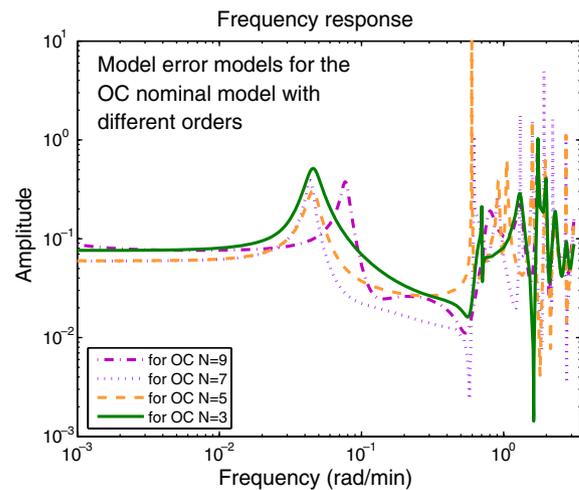


Fig. 12. Comparing magnitudes of frequency responses of the high-order black-box model error models.

orders, but it has greater magnitude at the frequencies higher than  $\omega \approx 0.2$ .

The model error models discussed above provide estimates of the magnitude of the model error in important frequency ranges. This will be helpful for designing a robust controller for the process.

## 6. Conclusions

An identification system for distributed parameter processes is applied and is shown to be able to identify

unknown parameters of physical models using data collected from the real process. The flexibility of the identification system allows different choices of subprocesses and variables.

Two subprocesses are selected for the parameter identification study. The model of the selected subprocesses consists of coupled PDEs and then the model is reduced by the OC method in linear state space form. The estimation results are improved by adjusting the model structure and the variable selection, etc. The obtained models are verified by simulation using validation data and mean square error analysis. The model outputs provide an acceptable agreement to the measured data. In addition, high-order black-box models are constructed for a subprocess and its model error. These black-box models are utilized for model validation of the OC nominal models. The frequency properties of the OC models and the black-box models are demonstrated and compared. The model error descriptions provide helpful information for controller design with robustness considerations.

In the future, the identification tool will be further developed and a richer physical model collection and approximation method collection will be available. More complex models and other subprocesses in the digester will be considered in the identification work.

### Acknowledgements

The author would like to thank Thore Lindgren who is working at Metso Panelboard, for the help from him when he worked at Eurocon. During the early stage of this research he did helpful work in process investigation and data collection of the full-scale digester. Author's gratitude goes also to Terje Fossum and Håkan Fridén at NPI, Mid Sweden University; Harry Forsgren, and Christer Svanhol at Eurocon; and Torbjörn Sjölund at More Research, for their support and helpful discussions.

### Appendix A

Notation of model (1)–(7) cited from [17]:

*Main variables:*

$T_f$	temperature of free liquor phase
$T_c$	temperature of chip phase
$\rho_{fi}$	concentration of $i$ th component of free liquor phase
$\rho_{ei}$	concentration of $i$ th component of entrapped liquor phase
$\rho_{si}$	concentration of $i$ th component of solid phase
$R_{si}$	reaction rate of solid component $i$
$R_{ei}$	reaction rate of entrapped liquor component $i$

*Rest notation:*

$C_{pf}$	specific heat capacity of free liquor phase
$M_f$	mass of free liquor phase

$v_f$	velocities of free liquor phase
$U$	interphase heat-transfer coefficient
$\eta$	fraction of control volume occupied by free liquor
$D$	coefficient of diffusion
$\epsilon$	fraction of chip phase occupied by entrapped liquor, or porosity
$D_E$	energy transported by diffusion into entrapped liquor
$\dot{V}_{\text{ext}}$	volumetric flow rate of external stream
$M_{\text{ext}}$	mass of external stream
$T_{\text{ext}}$	temperature of external stream
$\Delta V_f$	volumes of free liquor phase in control volume
$C_{ps}$	specific heat capacity of solid phase
$M_s$	mass of solid phase
$C_{pe}$	specific heat capacity of entrapped liquor phase
$M_e$	mass of entrapped liquor phase
$v_c$	velocities of chip phase
$\Delta H_R$	heat of reaction
$e_f$	reaction rate effectiveness factor
$k_{1i}$	the first kinetic rate constant for consumption of $i$ th solid component
$k_{2i}$	the second kinetic rate constant for consumption of $i$ th solid component
$\rho_{si}^\infty$	concentration of $i$ th component of solid phase which does not react
$b_{ij}$	stoichiometric coefficients for consumption of white liquor components

### Appendix B

Define an arbitrary state vector as

$$x(z, t) = [x_1(z, t) \quad x_2(z, t) \quad \cdots \quad x_{nx}(z, t)]^T \quad (39)$$

where  $nx$  is the number of the state variables. Then a single state variable in  $x(z, t) \in R^{nx}$  is denoted as  $x_m(z, t)$ , where  $m = 1, \dots, nx$ . The OC approximation of the spatial derivative  $\frac{\partial x(z, t)}{\partial z}$  and the OC approximation of a single state variable  $x_m(z, t)$  are presented here in detail. According to (13) and (14), we have

$$\frac{\partial x(z, t)}{\partial z} \approx \frac{\partial x_T(z, t)}{\partial z} = \frac{\partial}{\partial z} \sum_{j=0}^{N+1} x(z_j, t) l_j(z) \quad (40)$$

where  $x(z_j, t)$  is a vector and it is independent of  $z$ . Then,

$$\begin{aligned} \frac{\partial}{\partial z} \sum_{j=0}^{N+1} x(z_j, t) l_j(z) &= \sum_{j=0}^{N+1} \frac{\partial l_j(z)}{\partial z} x(z_j, t) \\ &= \sum_{j=0}^{N+1} \left[ \frac{\partial l_j(z)}{\partial z} x_1(z_j, t) \quad \cdots \quad \frac{\partial l_j(z)}{\partial z} x_{nx}(z_j, t) \right]^T \end{aligned} \quad (41)$$

Thus for each single state  $x_m$ , its spatial derivative at a collocation point  $z_i$  is

$$\begin{aligned}
\left. \frac{\partial x_{mT}}{\partial z} \right|_{z_i} &= \sum_{j=0}^{N+1} \left. \frac{\partial l_j(z)}{\partial z} \right|_{z_i} x_m(z_j, t) \\
&= \left. \frac{\partial l_0(z)}{\partial z} \right|_{z_i} x_m(z_0, t) + \left. \frac{\partial l_2(z)}{\partial z} \right|_{z_i} x_m(z_1, t) + \cdots + \left. \frac{\partial l_{N+1}(z)}{\partial z} \right|_{z_i} x_m(z_{N+1}, t) \\
&= \left. \frac{\partial l_0}{\partial z} \right|_{z_i} x_m(z_0, t) + \left[ \left. \frac{\partial l_1}{\partial z} \right|_{z_i} \cdots \left. \frac{\partial l_{N+1}}{\partial z} \right|_{z_i} \right] x_m(t)
\end{aligned} \quad (42)$$

where

$$x_m(t) = [x_m(z_1, t) \cdots x_m(z_{N+1}, t)]^T \quad (43)$$

For  $i = 1, 2, \dots, N+1$ , the spatial derivative of  $x_m$  at  $z_1, z_2, \dots, z_{N+1}$  is given as

$$\begin{aligned}
\begin{bmatrix} \left. \frac{\partial x_{mT}}{\partial z} \right|_{z_1} \\ \left. \frac{\partial x_{mT}}{\partial z} \right|_{z_2} \\ \vdots \\ \left. \frac{\partial x_{mT}}{\partial z} \right|_{z_{N+1}} \end{bmatrix} &= \begin{bmatrix} \left. \frac{\partial l_0}{\partial z} \right|_{z_1} \\ \left. \frac{\partial l_0}{\partial z} \right|_{z_2} \\ \vdots \\ \left. \frac{\partial l_0}{\partial z} \right|_{z_{N+1}} \end{bmatrix} x_m(z_0, t) \\
&+ \begin{bmatrix} \left. \frac{\partial l_1}{\partial z} \right|_{z_1} & \left. \frac{\partial l_2}{\partial z} \right|_{z_1} & \cdots & \left. \frac{\partial l_{N+1}}{\partial z} \right|_{z_1} \\ \left. \frac{\partial l_1}{\partial z} \right|_{z_2} & \left. \frac{\partial l_2}{\partial z} \right|_{z_2} & \cdots & \left. \frac{\partial l_{N+1}}{\partial z} \right|_{z_2} \\ \vdots & \vdots & \cdots & \vdots \\ \left. \frac{\partial l_1}{\partial z} \right|_{z_{N+1}} & \left. \frac{\partial l_2}{\partial z} \right|_{z_{N+1}} & \cdots & \left. \frac{\partial l_{N+1}}{\partial z} \right|_{z_{N+1}} \end{bmatrix} x_m(t) \\
&= B_1 x_m(z_0, t) + A_1 x_m(t)
\end{aligned} \quad (44)$$

For all states  $x_1, \dots, x_{n_x}$  we have

$$\frac{\partial x_T(z, t)}{\partial z} = (I_{n_x} \otimes A_1)x(t) + (I_{n_x} \otimes B_1)x(z_0, t) \quad (45)$$

where  $I_{n_x} \otimes A_1$  and  $I_{n_x} \otimes B_1$  are Kronecker products, and

$$x(t) = [x_1(t) \cdots x_{n_x}(t)]^T \quad (46)$$

$$x(z_0, t) = [x_1(z_0, t) \cdots x_{n_x}(z_0, t)]^T \quad (47)$$

Matrices  $A_1$  and  $B_1$  are used to approximate spatial derivatives of the state variables. At an arbitrary position  $z_{arb} \in [0, 1]$ , a process state  $x_m$  can be solved directly using (13):

$$\begin{aligned}
x_m(z_{arb}, t) &= \sum_{j=0}^{N+1} x_m(z_j, t) l_j(z_{arb}) = [l_1(z_{arb}) \cdots l_{N+1}(z_{arb})] x_m(t) \\
&+ l_0(z_{arb}) x_m(z_0, t)
\end{aligned} \quad (48)$$

thus at the boundary point  $z = z_{N+1}$ , we have

$$\begin{aligned}
x_m(z_{N+1}, t) &= [l_1(z_{N+1}) \cdots l_{N+1}(z_{N+1})] x_m(t) + l_0(z_{N+1}) x_m(z_0, t) \\
&= C_1 x_m(t) + l_0(z_{N+1}) x_m(z_0, t)
\end{aligned} \quad (49)$$

Due to the orthogonality:

$$l_j(z_i) = \begin{cases} 1 & j = i \\ 0 & j \neq i \end{cases} \quad (50)$$

we have

$$C_1 = [l_1(z_{N+1}) \cdots l_N(z_{N+1}) \quad l_{N+1}(z_{N+1})] = [0 \cdots 0 \quad 1] \quad (51)$$

and

$$l_0(z_{N+1}) = 0 \quad (52)$$

Then

$$x_m(z_{N+1}, t) = C_1 x_m(t) \quad (53)$$

Matrix  $C_1$  is used to approximate the state variables at boundary point  $z = z_{N+1}$ .

## Appendix C

**Lemma.** Consider the PDE:

$$\frac{\partial T_d(z_d, t)}{\partial t} + \frac{1}{\tau} \frac{\partial T_d(z_d, t)}{\partial z_d} = 0, \quad z_d \in [0, 1] \quad (54)$$

Define

$$T_d(0, t) = u(t) \quad (55)$$

$$T_d(1, t) = y(t) \quad (56)$$

then

$$y(t) = u(t - \tau) \quad (57)$$

**Proof.** Taking Laplace transform for (54), we have

$$sT_d(z_d, s) + \frac{1}{\tau} \frac{\partial T_d(z_d, s)}{\partial z_d} = 0 \quad (58)$$

Rearranging (58) we have

$$\frac{\partial T_d(z_d, s)}{T_d(z_d, s)} = -s\tau dz_d \quad (59)$$

Integrating both sides of (59):

$$\ln T_d(z_d, s) = -s\tau z_d + c_1 \quad (60)$$

then

$$T_d(z_d, s) = C e^{-s\tau z_d} \quad (61)$$

where  $c_1$  and  $C$  are constants and  $C = e^{c_1}$ .

Applying (55) we have

$$T_d(z_d, s) = u(s) e^{-s\tau z_d} \quad (62)$$

Taking the inverse Laplace transform, the time function of  $T_d$  along  $z_d$  is obtained as

$$T_d(z_d, t) = u(t - \tau z_d), \quad 0 \leq z_d \leq 1 \quad (63)$$

The exact solution of the equation at  $z_d = 1$  is obtained as

$$T_d(1, t) = u(t - \tau) \quad (64)$$

Applying (56) we have

$$y(t) = u(t - \tau) \quad (65)$$

The Lemma is proved. Applying the Lemma into the improved model for  $G_a$ , that consists of PDEs (8), (9) and (27), where

$$T_d(0, t) = T_s(z_{top}, t) \quad (66)$$

and

$$T_d(1, t) = T_f(z_{c6}, t) \quad (67)$$

Then the delayed steam temperature at the outlet position of c6 is given as

$$T_d(z_{c6}, t) = T_s(z_{top}, t - \tau) \quad (68)$$

## References

- [1] J. Funkquist, Modeling and identification of a distributed parameter process: the continuous digester, PhD thesis, Automatic Control, Department of Signals, Sensors and Systems, Royal Institute of Technology, Sweden, 1995.
- [2] J. Funkquist, Grey-box identification of a continuous digester – a distributed-parameter process, *Control Engineering Practice* 5 (7) (1997) 919–930.
- [3] P.A. Wisnewski, F.J. Doyle III, Control structure selection and model predictive control of the Weyerhaeuser digester problem, *Journal of Process Control* 8 (5–6) (1998) 487–495.
- [4] A. Alexandridis, H. Sarimveis, A. Angelou, T. Retsina, G. Bafas, A model predictive control scheme for continuous pulp digesters based on the partial least square (PLS) modelling algorithm, in: *Proceedings of the 10th Control Systems conference*, June 3–5, 2002, Stockholm, Sweden, pp. 253–257.
- [5] P.A. Wisnewski, F.J. Doyle III, F. Kayihan, Fundamental continuous pulp digester model for simulation and control, *AIChE Journal* 43 (1997) 3175–3192.
- [6] R. Amirthalingam, J.H. Lee, Subspace identification based inferential control applied to a continuous pulp digester, *Journal of Process Control* 9 (1999) 397–406.
- [7] R. Amirthalingam, J.H. Lee, Subspace identification based inferential control of a continuous pulp digester, *Computers and Chemical Engineering* 21 (1997) 1143–1148.
- [8] L. Ljung, Model validation and model error modeling, in: B. Wittenmark, A. Rantzer (Eds.), *Proceedings of the Åström Symposium on Control*, Studentlitteratur, Lund, Sweden, 1999, pp. 15–42.
- [9] W. Reinelt, A. Garulli, L. Ljung, Comparing different approaches to model error modeling in robust identification, *Automatica* 38 (5) (2002) 787–803.
- [10] S.L. Quinn, T.J. Harris, D.W. Bacon, Accounting for uncertainty in control-relevant statistics, *Journal of Process Control* 15 (2005) 675–690.
- [11] K.E. Vroom, The *H* factor: a means of expressing cooking times and temperatures as a single variable, *Pulp and Paper Magazine of Canada* (1957) 228–231, Convention issue.
- [12] R.R. Gustafson, C.A. Sleicher, W.T. McKean, B.A. Finlayson, Theoretical model of a Kraft pulping process, *Industrial and Engineering Chemistry Process Design and Development* 22 (1983) 87–96.
- [13] A.C. Butler, T.J. Williams. A Description and User's Guide for the Purdue Kamyrdigester Model. Technical Report 152, Purdue University, PLAIC, Purdue Engineering, West Lafayette, IN 47907, December, 1988.
- [14] T. Christensen, L.F. Albright, T.J. Williams. A mathematical model of the Kraft pulping process, Technical Report 129, Purdue University, PLAIC, Purdue University, West Lafayette, N 47907, May 1982.
- [15] F.A. Michelsen. A dynamic mechanistic model and model-based analysis of continuous Kamyrdigester, PhD Thesis, 1995 Report no. 95-4-W, University of Trondheim, 1995.
- [16] F. Kayihan, M.S. Gelormino, E.M. Hanczyc, F.J. Doyle III, Y. Arkun, A Kamyrdigester model for identification and controller design, *Proc. IFAC World Cong. San Francisco*, Elsevier Science Publishers, New York, 1996, pp. M37-42.
- [17] S. Bhartiya, P. Dufour, F.J. Doyle III, Fundamental thermal-hydraulic pulp digester model with grade transition, *AIChE Journal* 49 (2) (2003) 411–425.
- [18] S.O. Lundqvist, Matematisk modell av kontinuerlig sulfatkokare, in: *SCAN FORSK*, vol. 72, Stockholm, Sweden, STFI, 1975 (in Swedish).
- [19] J. Villadsen, M. Michelsen, *Solution of Differential Equation Models by Polynomial Approximation*, Prentice-Hall, Englewood Cliffs, NJ, 1978.
- [20] D. Dochain, J.P. Babary, N. Tali-Maanar, Modeling and adaptive control of nonlinear distributed parameter bioreactors via orthogonal collocation, *Automatica* 28 (5) (1992) 873–883.
- [21] D. Dochain, N. Tali-Maanar, J.P. Babary, On modelling monitoring and control of fixed bed bioreactors, *Computers and Chemical Engineering* 21 (11) (1997) 1255–1266.
- [22] L. Lefèvre, D. Dochain, S. Feyo de Azevedo, A. Magnus, Optimal selection of orthogonal polynomials applied to the integration of chemical reactor equations by collocation methods, *Computers and Chemical Engineering* 24 (12) (2000) 2571–2588.
- [23] C.A.J. Fletcher, *Computational Galerkin Methods*, Springer-Verlag, New York, 1984.
- [24] K. Alhumaizi, R. Henda, M. Soliman, Numerical analysis of a reaction–diffusion–convection system, *Computers and Chemical Engineering* 27 (4) (2003) 579–594.
- [25] G.F. Carey, B.A. Finlayson, Orthogonal collocation on finite elements, *Chemical Engineering Science* 30 (1975) 587–596.
- [26] M.A. Soliman, A spline collocation method for the solution of diffusion–convection problems with chemical reactions, *Chemical Engineering Science* 47 (1992) 4209–4213.
- [27] MathWorks, *System Identification Toolbox*.