8th International IFAC Symposium on
Dynamics and Control of Process Systems
Preprints Vol.3, June 6-8, 2007, Cancún, Mexico

A GREY-BOX MODELING APPROACH FOR THE REDUCTION OF NONLINEAR SYSTEMS [1]

Reinout Romijn * **Leyla Özkan** ** **Siep Weiland** ***
Jobert Ludlage ** **Wolfgang Marquardt** *

* *Process Systems Engineering, RWTH Aachen University, Turmstrasse 46, 52056 Aachen, Germany*
** *IPCOS, Bosscheweg 135b, 5282 WV Boxtel, The Netherlands*
*** *Eindhoven Univ. of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands*

Abstract: A novel model reduction methodology is proposed to approximate large-scale nonlinear dynamical systems. The methodology amounts to finding computationally efficient substitute models for an uncertain nonlinear system. Model uncertainty is incorporated by viewing the system as a grey-box or *hybrid model* with a mechanistic (first-principle) component and an empirical (black-box) component. The mechanistic part is approximated using proper orthogonal decomposition. Subsequently, the empirical part is identified by parameter estimation using the reduced order mechanistic part. As a consequence, the parameter estimation is computationally more efficient. An example with a distributed parameter system is provided.

Keywords: Grey-box modeling, Hybrid Modeling, Model Reduction, Parameter Estimation, Proper Orthogonal Decomposition, Distributed Parameter Systems

1. INTRODUCTION

Advances in computation and modeling tools have enabled the development of detailed complex mathematical models that yield reasonable and accurate predictions of the behavior of any type of (process) system. The intrinsic complexity of processes yields models that require a considerable computational effort to obtain solutions. As a consequence, the use of such models for on-line or real-time applications, such as model-based control, dynamic optimization or parameter estimation is computationally costly. Significant research efforts have been dedicated to the development and implementation of model reduction techniques, for an overview refer to (Marquardt, 2002; Antoulas and Sorensen, 2001), for reduction of distributed parameter systems refer to (Gay and Ray, 1995; Shvartsman and Kevrekidis, 1998; Mahadevan and Hoo, 2000; Hoo and Zheng, 2001). The method of Proper Orthogonal Decompositions (POD) is particularly popular in the fluid dynamics community and of considerable interest for the reduction and simplification of this type of systems.

There are a number of key disadvantages to most model reduction techniques, including the method of POD. Firstly, state reductions do not necessarily lead to computationally more efficient models (Schlegel *et al.*, 2002; Rathinam and Petzold, 2003; Astrid, 2004; van den Berg, 2005). Secondly, few techniques are able to cope with nonlinear uncertainty in the model. Thirdly, most reduction techniques do not allow to extract relevant features for the specific purpose for which the model is meant. This paper is motivated by the

[1] This work has been supported by the European Union within the Marie-Curie Training Network PROMATCH under the grant number MRTN-CT-2004-512441.

question to perform model approximation with the explicit aim to improve computational efficiency while keeping desirable model properties intact. We introduce a new methodology that utilizes a grey-box modeling approach to obtain a reduced model that is computationally more efficient. For arbitrary nonlinear systems a grey-box model defines an interconnection where the mechanistic part represents the known and the empirical part captures the unknown or uncertain parts of the system. At least three distinctive issues motivate combining grey-box model structures with model reduction:

(1) Model uncertainties and un-modeled dynamics can be represented as interconnections of a known system component with a (partly) unknown component. By tearing the uncertain part out of the model, the remaining part can be subject to model reduction without damaging the properties of the uncertain model part or the interconnection structure.

(2) Finite element discretization of the spatial geometry of nonlinear distributed parameter systems typically leads to performing nonlinear function evaluations in each of the mesh elements. As a consequence, the computational load per nonlinear function evaluation is critical. By *separating* computationally intensive functions in a model *before* any kind of model reduction is performed, a structure is created in which computationally intensive functions can be substituted by simpler ones which leads to more efficient approximate models.

(3) Reduction methods generally do not preserve model sparsity. It will be shown that by employing a grey-box model structure sparsity can be reintroduced in a projected low order model. Furthermore, a low order sparse nonlinear model part can be maintained when a full order model is reconstructed from an identified low order model.

The paper presents some preliminaries on grey-box modeling and POD reduction in section 2. Section 3 describes the methodology of approximate modeling which is subsequently applied to a distributed dynamical system in sections 4. Conclusions are deferred to section 5.

2. PRELIMINARIES

2.1 Grey-box modeling

A grey-box model, also referred to in literature as a *hybrid model*, consists of a combination of a mechanistic (*first principle*) and an empirical (*black-box*) model. Several grey-box model structures have been proposed. Psichogios and Ungar (1992) proposed a serial structure, Thompson and Kramer (1994) described a parallel structure. A more general structure has been formulated (Marquardt, 2002; Abonyi *et al.*, 2002) in a form based on the ordinary differential equation

$$\dot{x} = f_{\mathrm{FP}}\left(x, u, f_{\mathrm{EM}}\left(x, u\right)\right) \tag{1}$$

which contains first principle equations f_{FP} which describe the interaction of the model states x, inputs u and the outputs of an empirical model f_{EM}. Such grey-box models have been applied for modeling a variety of process systems (Marquardt, 2002). They are usually derived from conservation laws and balance equations but are, without exception, reduced to lumped models of low order. A distributed grey-box model that is governed by a partial differential equation (PDE) including an empirical term without lumping the states has not been investigated before to the knowledge of the authors. In this work the general partial differential equation

$$\frac{\partial T}{\partial t} = \mathcal{A}(T) + \mathcal{B}(u) + \mathcal{F}(T, u) \tag{2}$$

is considered. Here $T(x,t)$ denotes the state variable at position x in some spatial geometry Ω and at time t, $u(x,t)$ denotes the input. $\mathcal{A}$ is a linear operator, $\mathcal{B}$ denotes the input operator and $\mathcal{F}$ represents nonlinear terms and model mismatch. A separation of linear and nonlinear terms for a different class of systems as a basis for model reduction has been proposed in (Yousefi *et al.*, 2004), where the nonlinearities are not reduced however. In our work, the system (2) is separated into two parts

$$\frac{\partial T}{\partial t} = \mathcal{A}(T) + \mathcal{B}(u) + q \tag{3a}$$

$$q = \mathcal{F}(T, u) \tag{3b}$$

to facilitate the interpretation of the nonlinear function $\mathcal{F}(T,u)$ as the (known or unknown) empirical part of the model. A parametrization $\mathcal{E}(T,u,\theta)$ of the empirical part is proposed to replace $\mathcal{F}$ in (3b) by $\mathcal{E}(\cdot,\cdot,\theta)$ for a suitable parameter $\theta \in \Theta$. In this way, a hybrid structure is defined in which the estimation of $\theta \in \Theta$ will be combined with the reduction of the mechanistic part (3a).

2.2 Proper Orthogonal Decompositions

One of the most promising and significant techniques for an efficient reduction of large-scale nonlinear systems in fluid dynamics is the method of Proper Orthogonal Decompositions (POD) also known as the Karhunen-Loève method. The method is based on the observation that flow characteristics reveal coherent structures or *patterns* in many processes in fluid dynamics. This has led to the idea that the solutions of model equations may be approximated by considering

a small number of dominant coherent structures (called *modes*) that are inferred in an *empirical* manner from measurements or simulated data.

Given an ensemble of K measurements $\mathbf{T}^k(\cdot)$, $k = 1, \ldots, K$ with each measurement defined on some spatial domain Ω, the POD method amounts to assuming that each observation $\mathbf{T}^k$ belongs to a Hilbert space $\mathcal{H}$ of functions defined on Ω. With the inner product defined on $\mathcal{H}$, it then makes sense to call a collection $\{\varphi_j\}_{j=1}^{\infty}$ an *orthonormal basis* of $\mathcal{H}$ if any element, say $\mathbf{T} \in \mathcal{H}$, admits a representation

$$\mathbf{T}(x) = \sum_{j=1}^{\infty} a_j \varphi_j(x), \quad x \in \Omega. \tag{4}$$

Here, the a_j's are referred to as the *coefficients* and the φ_j's are the *modes* of the expansion. The truncated expansion

$$\mathbf{T}_n(x) = \sum_{j=1}^{n} a_j \varphi_j(x), \quad x \in \Omega \tag{5}$$

causes an approximation error $\|\mathbf{T} - \mathbf{T}_n\|$ in the norm of the Hilbert space. We will call $\{\varphi_j\}_{j=1}^{\infty}$ a *POD basis* of $\mathcal{H}$ whenever it is an orthonormal basis of $\mathcal{H}$ for which the *total approximation error* $\sum_{k=1}^{K} \|\mathbf{T}^k - \mathbf{T}_n^k\|$ is minimal for all truncation levels n. This is an *empirical basis* in the sense that every POD basis depends on the data ensemble.

Using variational calculus, the solution to this optimization problem amounts to finding the normalized eigenfunctions φ_j of a positive semi-definite operator $R : \mathcal{H} \to \mathcal{H}$ that is defined as

$$\langle \psi_1, R\psi_2 \rangle := \frac{1}{K} \sum_{k=1}^{K} \langle \psi_1, \mathbf{T}^k \rangle \cdot \langle \psi_2, \mathbf{T}^k \rangle \tag{6}$$

with $\psi_1, \psi_2 \in \mathcal{H}$. R is well defined in this manner and corresponds to a positive semi-definite matrix whenever $\mathcal{H}$ is finite dimensional. In that case, a POD basis is obtained from the normalized eigenvectors of R (Astrid, 2004; Cazemier, 1997).

Subsequently, a Galerkin projection is used to obtain the reduced model as follows. Suppose that the system is governed by a PDE of the form (2). Let $\mathcal{H}_n$ denote an n dimensional subspace of $\mathcal{H}$ and let $P_n : \mathcal{H} \to \mathcal{H}_n$ and $I_n : \mathcal{H}_n \to \mathcal{H}$ denote the canonical projection and canonical injection maps. The reduced model is then given by

$$P_n \frac{\partial T_n}{\partial t} = P_n \mathcal{A}(T_n) + P_n \mathcal{B}(u) + P_n \mathcal{F}(T_n, u) \tag{7}$$

where $T_n(\cdot, t) = \mathbf{T}_n(t)$ belongs to $\mathcal{H}_n = P_n \mathcal{H}$ for all t. In the specific case of a POD basis, the finite dimensional subspace $\mathcal{H}_n = \text{span}(\varphi_1, \ldots, \varphi_n)$ where the φ_j's denote POD basis functions. In that case, (7) becomes an *ordinary differential equation* in the coefficients $a_j(t)$ in the expansion of T_n.

3. METHODOLOGY

A conventional construction of a grey-box model which approximates the model (3) without any form of model order reduction consists of the replacement of the relation (3b) by the empirical model

$$q = \mathcal{E}(T, u, \theta) \tag{8}$$

where $\mathcal{E} : \mathcal{H} \times \mathcal{U} \times \Theta \to \mathcal{H}$ defines a parametrization of $\mathcal{F}$. In a grey-box *modeling* framework $\mathcal{F}$ is an unknown process part that is approximated by an empirical model $\mathcal{E}$. In a grey-box *model reduction* framework $\mathcal{F}$ is a possibly highly complex, computationally expensive nonlinear term that is approximated by a computationally less expensive term $\mathcal{E}$.

Identification of the model part $q = \mathcal{E}(T, u, \theta^*)$ takes place after discretization of the spatial domain Ω into N points. By this, the variable $\mathbf{T}(t)$ becomes $\mathbf{T}(t) := \text{col}(T(x_1, t), ..., T(x_N, t))$ which is an element of a finite dimensional Hilbert space $\mathcal{H}$. Introducing expression $\mathcal{M}(T, u) = \mathcal{A}(T) + \mathcal{B}(u)$ for the mechanistic model part, the identified *full order* grey-box model is given by

$$\frac{\partial T}{\partial t} = \mathcal{M}(T, u) + \mathcal{E}(T, u, \theta^*) \tag{9}$$

3.1 3-step method of Grey-box model identification

The reduction methodology that we propose in this contribution is illustrated in Figure 1. The method consists of the following steps:

(1) *Reduction of the mechanistic part.* The grey-box model formulation as in equation (3) allows a separate treatment of the separate model parts. Following (7), a Galerkin projection is performed on the mechanistic part (3a) to yield a reduced order model

$$P_n \frac{\partial T_n}{\partial t} = P_n \mathcal{M}(T_n, u) + q_n \tag{10}$$

where $T_n = P_n T$, and $q_n = P_n q$ and where $P_n : \mathcal{H} \to \mathcal{H}_n$ is the projection matrix defined in Section 2.2. This step reduces the order of the model to n.

(2) *Parametrization and estimation of empirical part.* After reduction of (3a), the reduced model

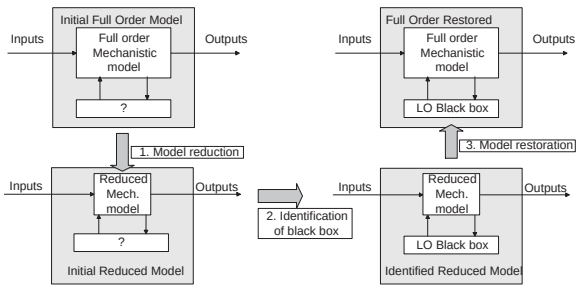

Fig. 1. Scheme of the proposed methodology

(10) is extended with the relation $\mathcal{E}_n(T,u,\theta) = \mathrm{col}(\mathrm{E}_1(T,u,\theta),\ldots,\mathrm{E}_n(T,u,\theta))$ where $\mathcal{E}_n : \mathcal{H}_n \times \mathcal{U} \times \Theta \to \mathcal{H}_n$ defines a parametrization of the empirical part. An optimal parameter $\theta^* \in \Theta$ is identified, such that a criterion function $J : \Theta \to \mathbb{R}$ on the interconnection (10)-(11) is minimized. This results in:

$$q_n = \mathcal{E}_n(T_n, u, \theta^*) \tag{11}$$

In this step a low order model with a low number of nonlinear function evaluations has been obtained by the interconnection of (10) and (11) which we refer to as the *reduced model*.

(3) *Model restoration.* In order to obtain solutions in terms of the variable $T(x,t)$, three possibilities exist. Firstly, the solution trajectories of the *reduced model* can be injected to the full order space. Secondly, the *reduced model* equations (10, 11) as well as the variable T_n can be injected to the full order space which results in the *injected model* structure:

$$\frac{\partial T}{\partial t} = I_n P_n \mathcal{M}(I_n P_n T, u) + I_n \mathcal{E}_n(P_n T, u, \theta^*) \tag{12}$$

Thirdly, the identified low order empirical part can be injected to the full order space and coupled to the original mechanistic part, by which the *coupled model* structure is obtained:

$$\frac{\partial T}{\partial t} = \mathcal{M}(T,u) + I_n \mathcal{E}_n(P_n T, u, \theta^*) \tag{13}$$

With both the restored *injected* and *coupled* models structures are obtained in which fewer nonlinear functions have to be evaluated than in (9).

3.2 Sparsity structure

Figure 2 illustrates the sparsity structure of the full order model and the reduced model. In the full order model $\mathcal{M}$ and $\mathcal{E}$ are solved on each (of the N) grid points or mesh elements. In the reduced model, the same structure is present whereby $\mathcal{M}$ and $\mathcal{E}$ are only to be solved n times. The restored coupled model has a structure where $\mathcal{E}$ is coupled to $\mathcal{M}$ via the projection and injection matrices. In case the input is a distributed variable $u(x,t)$, the input hast to be projected on a suitable subspace as well in order to maintain a low order mapping $\mathcal{E}_n : \mathcal{H}_n \times \mathcal{U} \times \Theta \to \mathcal{H}_n$.

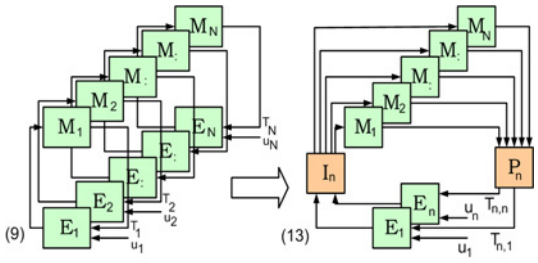

Fig. 2. Sparse structure of models (9) and (13)

4. APPLICATION TO HEAT DIFFUSION

4.1 Model description

This section considers a simple one-dimensional model for heat conduction and convection in melted glass. The model is given by the PDE

$$\frac{\partial T(x,t)}{\partial t} = -v\frac{\partial T(x,t)}{\partial x} + \alpha\frac{\partial^2 T(x,t)}{\partial x^2} + Q(x,t) \tag{14}$$

where $T(x,t)$ is the temperature at position x and time t, v is the convective flow velocity, α is the thermal diffusivity coefficient and

$$Q(x,t) = \frac{\sigma}{\rho c_p}\left(c_1 T_{\mathrm{fl}}^4(x,t) - c_2 T^4(x,t)\right) \tag{15}$$

is a spatially distributed heat input term depending on T and the heating flames T_{fl} above the glass. The parameter values of the model are listed in Table 1.

As a first step, the spatial domain is discretized in N equidistant grid points at intervals d_x. Let $\mathbf{T}(t) = col(T(x_1,t),\ldots,T(x_N,t))$ denote the temperature vector at the grid points. After the spatial derivatives of T have been approximated by central finite differences the resulting model assumes the form

$$\dot{\mathbf{T}} = A\mathbf{T} + BT_b + Q(\mathbf{T}, \mathbf{T}_{\mathrm{fl}}) \tag{16}$$

where T_b is the temperature at the boundary at the upstream side. The gradient of the glass temperature at the downstream boundary is assumed to be zero.

A hybrid structure is employed for (16) where the nonlinearity $Q(\mathbf{T}, \mathbf{T}_{\mathrm{fl}})$ is replaced by an empirical function $\mathcal{E}(\mathbf{T}, \mathbf{T}_{\mathrm{fl}}, \theta)$ defined by a *feed-forward neural network* in the parameter vector $\theta \in \mathbb{R}^{13}$. Specifically, let $\mathcal{E} = col(E_1,\ldots,E_N)$ denote the nonlinearity in each gridpoint $x_1\ldots,x_N$, the ith output E_i of $\mathcal{E}$ is defined by

$$\begin{cases} \begin{bmatrix} a_{i,1} \\ a_{i,2} \\ a_{i,3} \end{bmatrix} = \tanh\left(\begin{bmatrix} \theta_1 & \theta_3 & \theta_5 \\ \theta_2 & \theta_4 & \theta_6 \\ \theta_7 & \theta_8 & \theta_9 \end{bmatrix} \begin{bmatrix} T_i \\ T_{fl,i} \\ 1 \end{bmatrix} \right) \\ \mathrm{E}_i = \begin{bmatrix} \theta_{10} & \theta_{11} & \theta_{12} & \theta_{13} \end{bmatrix} \begin{bmatrix} a_{i,1} & a_{i,2} & a_{i,3} & 1 \end{bmatrix}^\top \end{cases} \tag{17}$$

The full order grey-box model then becomes

$$\dot{\mathbf{T}} = A\mathbf{T} + BT_b + \mathcal{E}(\mathbf{T}, \mathbf{T}_{\mathrm{fl}}, \theta). \tag{18}$$

Table 1. Parameter values

Parameter		value	unit
flow	v	5E-4	ms^{-1}
thermal diffusivity	α	5E-3	Wm^{-2}
density	ρ	2500	kgm^{-3}
specific heat	c_p	400	$\frac{J}{kgK}$
transmission coeff.	c_1, c_2	0.20, 0.33	
Stefan-Boltzmann const.	σ	5.67E-8	$\frac{W}{m^2K^4}$
boundary temperature	T_b	1773	K
glass bath length	L	10	m

4.2 Reduction of mechanistic part

The method of POD is employed to reduce the mechanistic part of the model. A POD basis $\{\varphi_k\}_{k=1}^N$ for $\mathcal{H} = \mathbb{R}^N$ is computed from a collection of temperature measurements $\mathbf{T}(t_k) \in \mathcal{H}$. Here, $N = 100$ and we define the $n = 5$ dimensional projection space $\mathcal{H}_5$ by the column-span of the matrix $\Phi := [\varphi_1, \ldots, \varphi_5]$ that consists of 5 dominant POD basis functions. Since the basis is orthonormal, the canonical projection $P_5 = \Phi^\top$ and the canonical injection matrix $I_5 = \Phi$. A Galerkin projection of the mechanistic part then yields the model which we call the *projected model*. $\mathcal{E}$ is still of dimension N:

$$\dot{\mathbf{z}} = A_r\mathbf{z} + B_r T_b + \Phi^\top \mathcal{E}\left(\Phi\mathbf{z}, \mathbf{T}_{\text{fl}}, \theta^*\right). \qquad (19)$$

where $\Phi\mathbf{z} = \mathbf{T}$, $A_r = \Phi^\top A \Phi$ and $B_r = \Phi^\top B$. A gain in computational speed is noticeable for the mechanistic model part (see Table 3).

4.3 Estimation of empirical part

To improve the computational speed, the empirical part in (19) is replaced by a neural network $\mathcal{E}_n(z, z_{\text{fl}}, \theta)$ with the same sparse structure as $\mathcal{E}$ in (17) but with $n = 5$ dimensional input vectors. This simplifies (19) to

$$\dot{\mathbf{z}} = A_r\mathbf{z} + B_r T_b + \mathcal{E}_n\left(\mathbf{z}, \mathbf{z}_{\text{fl}}, \theta^*\right) \qquad (20)$$

where $\mathbf{z}_{\text{fl}} = \Phi^\top \mathbf{T}_{\text{fl}}$ denotes the projection of $\mathbf{T}_{\text{fl}}$ on the dominant basis functions. This projection which is required because of the enforced sparse structure as mentioned in section 3.2, is not optimal because T_{fl} has not been taken into account when constructing Φ. However, since the arguments of the nonlinear terms in $\mathcal{E}$ are affine combinations of $\mathbf{T}$ and $\mathbf{T}_{\text{fl}}$ (eq. 17), the exact sparse structure between the states and input variables is preserved:

$$\Phi^\top\left(\theta_1\mathbf{T} + \theta_2\mathbf{T}_{\text{fl}}\right) + \theta_3 = \theta_1\Phi^\top\mathbf{T} + \theta_2\Phi^\top\mathbf{T}_{\text{fl}} + \theta_3. \qquad (21)$$

Initial conditions are set to $\mathbf{z}(0) = \Phi^\top\mathbf{T}(0)$. The grey-box model has now been completely reduced to a system with both a reduced number of equations and a reduced number of nonlinear function evaluations. The parameter estimation problem of the nonlinear dynamic system that has been obtained can be formulated as follows:

$$\min_\theta \sum_{i=1}^{N_m}\sum_{j=1}^{n}\left(z_{i,j} - \hat{z}_{i,j}\right)^2$$

$$\text{subject to } \begin{cases} \dot{\mathbf{z}} = A_r\mathbf{z} + B_r T_b + \mathcal{E}_n\left(\mathbf{z}, \mathbf{z}_{\text{fl}}, \theta\right) \\ \mathcal{E}_n\left(\mathbf{z}, \mathbf{z}_{\text{fl}}, \theta\right) = \text{col}(\text{E}_1, \ldots, \text{E}_n) \\ \hat{\mathbf{z}} = \Phi^\top\hat{\mathbf{T}} \\ \mathbf{z}_0 = \Phi^\top\mathbf{T}_0 \end{cases} \qquad (22)$$

N_m is the number of measurements over time. The data from the snapshot matrix have been used as measurement data $\hat{\mathbf{T}}$. A standard Runge-Kutta scheme has been used to solve the model equations. The SQP optimization algorithm SNOPT (Gill *et al.*, 2005) has been used to solve the parameter estimation problem using finite differences to approximate the gradients.

A brief assessment of the computational load of the parameter estimation is given in Table 2 which gives the averaged results for 10 runs starting from random initial parameter values. The estimation time has been reduced with an order of magnitude. The exact reason for this phenomenon and how to exploit it will be subject to future research.

Table 2. Estimation results

model	function eval.	gradient eval.	Iter.	CPU time
Full order	1864	112	111	1401 s
Reduced	1680	105	104	145 s

4.4 Model restoration and validation

The computational load of the different models and the distribution over the empirical part and the mechanistic part as well as the cost of the actual evaluation of the nonlinear function are shown in detail in Table. 3. The total computational load including overhead in the model evaluation, has been reduced from 0.738 to 0.446 s, which is a reduction of about 40%. The computational load of the projected and restored models lie in between these values.

Table 3. Simulation results

model (eq.)	nonlin. func.	emp. part	mech. part	total
Original (16)	.871 s	.939 s	.107 s	1.236 s
Full order (18)	.247 s	.409 s	.192 s	.738 s
Projected (19)	.215 s	.355 s	.154 s	.650 s
Reduced (20)	.074 s	.188 s	.142 s	.446 s
Coupled (13)	.074 s	.189 s	.176 s	.506 s
Injected (12)	.086 s	.214 s	.252 s	.613 s

The reduced, injected and coupled model are validated by inspection of the quality of prediction. Figure 3 shows the nominal trajectory when the original model (16) is subjected to different inputs and initial conditions as well as the prediction errors of the reduced, the injected and the coupled model:

$$\left|\varepsilon\right|_{\text{val}} = \left|\frac{T_{\text{prediction}} - T_{\text{nominal}}}{\max\left(T_{\text{nominal}}\right) - \min\left(T_{\text{nominal}}\right)}\right|.$$

The figure shows the robustness of the coupled model structure. Inputs and initial conditions that have not been taken into account when constructing the POD basis functions could be interpreted as disturbances. Because the new inputs and initial conditions can not be projected sufficiently

well on the POD basis, the reduced model and the injected model show a discrepancy up to about 36% of the temperature variation. The coupled model is more robust against these disturbances and shows a maximum error of about 7%.

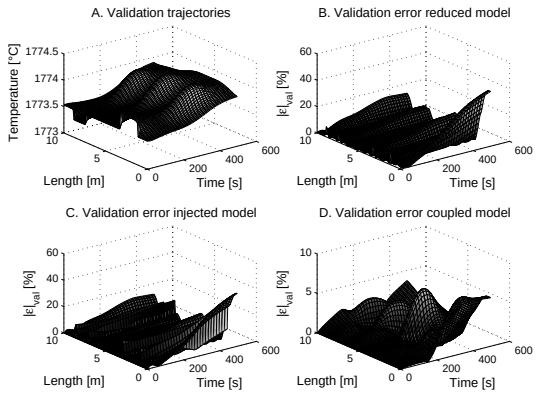

Fig. 3. Validation results.

5. CONCLUSIONS

This paper is motivated by the observation that many model reduction techniques for large-scale dynamical systems yield lower order models that fail to be computationally efficient. To remedy this problem, a general framework is proposed for model reduction for nonlinear systems using a grey-box modeling approach. Grey-box models allow a separation of a mechanistic, first principle part and an empirical, black-box part. It is shown that by combining model reduction and parameter estimation on grey-box models, the computational complexity can be reduced. The method is illustrated on a distributed model for heat diffusion. The prediction error of the restored coupled grey-box model is reasonably small and robust. A considerable improvement in computational speed of the parameter estimation problem has been observed. As a result, a method that provides a faster way to identify a non-linear reduced model has been presented.

REFERENCES

Abonyi, J., J. Madar and F. Szeifert (2002). *Soft Computing in Industrial Applications - Recent Advances*. Chap. Combining first principles models and neural networks for generic model control, pp. 111–122. Springer Engineering Series.

Antoulas, A. C. and D.C. Sorensen (2001). Approximation of large-scale dynamical systems: an overview. *Int. J. Appl. Math. Comput. Sci.* **11**(5), 1093–1121.

Astrid, P. (2004). Reduction of process Simulation Models: A Proper Orthogonal Decomposition Approach. PhD thesis. Technical Unicversity of Eindhoven.

Cazemier, W. (1997). Proper Orthogonal Decompsition and Low Dimensional Models for Turbulent Flow. PhD thesis. State University of Groningen.

Gay, D. H. and W. H. Ray (1995). Identification and control of distributed parameter systems by means of singular value decompsotion. *Chemical Engineering Science* **50**, 1519–1539.

Gill, P.E., W. Murray and M.A. Saunders (2005). Snopt: An sqp algorithm for large-scale constrained optimization. *SIAM Review* **47**, 99 – 131.

Hoo, K.A. and D. Zheng (2001). Low-order control-relevant models for a class of distributed parameter systems. *Chemical Engineering Science* **56**, 6683 – 6710.

Mahadevan, N. and K. Hoo (2000). Wavelet-based model reduction of distributed parameter systems. *Chemical Engineering Science* **55**, 4271–4290.

Marquardt, W. (2002). Nonlinear model reduction for optimization based control of transient chemical processes. In: *Proceedings of the 6th International Conference on Chemical Process Control*. Vol. 98 of *AIChe Symp. Ser.326*. pp. 12–42.

Psichogios, D.C. and L.H. Ungar (1992). A hybrid neural network-first principles approach to process modeling. *AIChE Journal* **38**(10), 1499–1511.

Rathinam, M. and L.R. Petzold (2003). A new look at proper orthogonal decomposition. *SIAM J. Numer. Anal.* **41**, 1893 – 1925.

Schlegel, M., J. van den Berg, W. Marquardt and O.H. Bosgra (2002). Projection based model reduction for dynamic optimization. AIChE Annual Meeting, Indianapolis. http://www.lpt.rwth-aachen.de/Publication/abstract.php?Nummer=LPT-2002-22.

Shvartsman, S.Y. and I.G: Kevrekidis (1998). Nonlinear model reduction for control of distributed systems: a computer-assisted study. *AIChE Journal* **44**, 1579 – 1595.

Thompson, M.L. and M.A. Kramer (1994). Modeling chemical processes using prior knowledge and neural networks. *AIChE Journal* **40**(8), 1328–1340.

van den Berg, J. (2005). Model Reduction for Dynamic real Time Optimization of Chemical Processes. PhD thesis. Delft Technical University.

Yousefi, A., B. Lohmann, J. Lienemann and J.G. Korvink (2004). Nonlinear heat transfer modelling and reduction. In: *Proceedings of the 12th mediterranean conference on control and automation*.