Computers and Chemical Engineering 60 (2014) 86–101

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Review

Hybrid semi-parametric modeling in process systems engineering: Past, present and future

Moritz von Stosch[b], Rui Oliveira[b], Joana Peres[a], Sebastião Feyo de Azevedo[a,*]

[a] LEPAE, Departamento de Engenharia Quimica, Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias s/n, 4200-465 Porto, Portugal
[b] REQUIMTE, Departamento de Quimica, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

ARTICLE INFO

Article history:
Received 18 March 2013
Received in revised form 9 August 2013
Accepted 14 August 2013
Available online 31 August 2013

Keywords:
Hybrid modeling
Hybrid neural modeling
Semi-mechanistic modeling
Hybrid grey-box modeling
Hybrid semi-parametric modeling
Process operation/design

ABSTRACT

Hybrid semi-parametric models consist of model structures that combine parametric and nonparametric submodels based on different knowledge sources. The development of a hybrid semi-parametric model can offer several advantages over traditional mechanistic or data-driven modeling, as reviewed in this paper. These advantages, such as broader knowledge base, transparency of the modeling approach and cost-effective model development, have been widely recognized, not only in academia but also in the industry.

In this paper, the most common hybrid semi-parametric modeling and parameter identification techniques are revisited. Applications in the areas of (bio)chemical engineering for process monitoring, control, optimization, scale-up and model-reduction are reviewed. It is outlined that the application of hybrid semi-parametric techniques does not automatically lead into better results but that rational knowledge integration has potential to significantly improve model-based process operation and design.

1. Introduction

In process systems engineering process modeling takes a central role (Cameron & Hangos, 2001). In its essence, process modeling is an exercise of translation of knowledge about the process into an abstract mathematical representation (Cameron & Hangos, 2001). The nature of knowledge is diverse and thus modeling methods can naturally be segmented according to the nature of the knowledge. First-Principles, mechanistic or phenomenological models represent a broad class of more transparent (white box) models. In relation thereto, data-driven modeling represents a less transparent (black-box) modeling framework based exclusively on process data.

A closely related mathematical classification can be done with respect to the form of model parameterization. Parametric models are determined a priori on the basis of knowledge about the process (Thompson & Kramer, 1994; Walter, Pronzato, & Norton, 1997). Their number of parameters is fixed and they might have a physical or empirical interpretation depending on the level of knowledge sophistication. White-box models naturally fall in the category of parametric models. On the contrary, nonparametric models are determined exclusively from data (Haerdle, Mueller, Sperlich, & Werwatz, 2004; Thompson & Kramer, 1994). The term nonparametric is not meant to imply that these models completely lack parameters but that the number and nature of the parameters are flexible and not fixed in advance by knowledge. In between these two extremes lies hybrid semi-parametric modeling, which is the focus of this review (Fig. 1).

Hybrid semi-parametric models can thus be defined as model structures that combine parametric and nonparametric submodels (Thompson & Kramer, 1994). Their application to process modeling has evolved from the field of neural networks, being first reported in 1992 by Psichogios and Ungar (1992), Kramer, Thompson, and Bhagat (1992), Johansen and Foss (1992a), and Su, Bhat, Minderman, and McAvoy (1992). The central idea was to *a priori* structure the neural network model through the use of first-principle knowledge. The result was that, when trained with the same amount of process data, the hybrid semi-parametric model was capable to predict the process states better, was able to interpolate and extrapolate mostly more accurately and was easier to interpret than models based solely on neural networks.

Several other modeling methods exist that combine different types of knowledge or/and submodels. The term grey-box modeling appeared in the 1990-s in systems and control theory describing the incorporation of prior information (mainly structural information derived from first-principles, i.e. white-box models) into empirical (black-box) models (Bohlin & Graebe, 1995; Jorgensen & Hangos, 1995; Tulleken, 1993). According to Braake, van Can,

* Corresponding author. Tel.: +351 225 08 16 94.
E-mail address: diretor@fe.up.pt (S. Feyo de Azevedo).

0098-1354/$ – see front matter
http://dx.doi.org/10.1016/j.compchemeng.2013.08.008

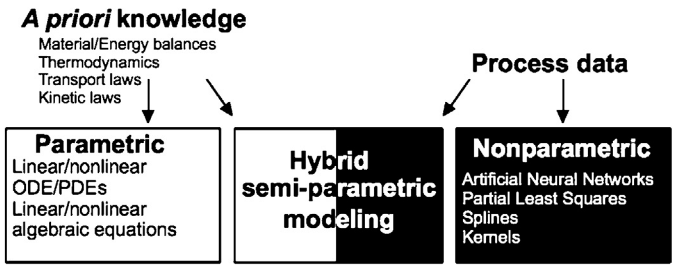

Fig. 1. Parametric, nonparametric and hybrid semi-parametric modeling and the types of knowledge they are based on.

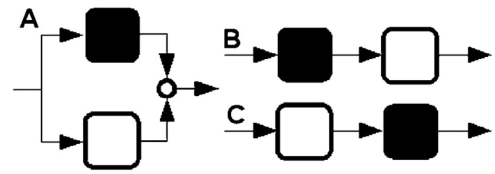

Fig. 2. Schematic sketch of the three ways to combine two models (represented by a white and a black box). **A** shows a parallel configuration. **B** and **C** present serial structures.

and Verbruggen (1998) a grey-box model is based on the same unstructured nature as a black-box model. The term has however evolved to designate all types of models that combine white-box and black-box submodels. For instance, the grey-box models in Akkari, Chevallier, and Boillereaux (2005), Estrada-Flores, Merts, De Ketelaere, and Lammertyn (2006), and Worden et al. (2007) combine first-principle based "white-box" models and empirical "black-box" models. In these cases both, white- and black-box models, are parametric models. According to the definition in this paper they are therefore not hybrid semi-parametric models. Hybrid semi-parametric models may however be viewed as a class of grey-box models in that the parametric and nonparametric submodels have different levels of transparency.

Block oriented models are another class of separable models consisting of linear dynamic and static nonlinear elements (Haber & Keviczky, 1999; Pearson & Pottmann, 2000), e.g. the Hammerstein model or the Wiener model (Haber & Keviczky, 1999). Block oriented models have some resemblance with hybrid semi-parametric models in that the nonlinear element could e.g. be represented by a neural network with a standard linear time invariant form as the linear dynamic model, such as in Su and McAvoy (1993). However, block oriented models are not necessarily hybrid semi-parametric models if the blocks do not explicitly combine parametric and nonparametric submodels.

Multiscale models are also compositions of two or more submodels that describe phenomena at different scales (Ingram, Cameron, & Hangos, 2004). In the vast majority of cases multiscale models are mechanistic parametric models (Ingram et al., 2004). However, multiscale models could also be hybrid semi-parametric models if parametric and nonparametric submodels target different scales (Teixeira, Alves, Alves, Carrondo, & Oliveira, 2007).

It should be noted that the term "hybrid modeling" has been frequently used as an equivalent to "hybrid semi-parametric modeling" in the literature, which is however a rather ambiguous definition as it can embrace many other types of modeling methods such as grey-box, block-oriented or multiscale modeling approaches referred to above. For coherence, we keep the term "hybrid semi-parametric" throughout this review.

1.1. Why hybrid semi-parametric modeling? What is the gain?

Mechanistic modeling and data-driven modeling constitute two approaches which are different in their traits. While the development of a mechanistic model is many times cumbersome/laborious and requires detailed knowledge about the process, data-driven approaches are rather quickly applicable and require less knowledge. In comparison to mechanistic models, more data are necessary for the derivation of data-driven models and its descriptive quality is good only in the vicinity to those regions for which it was derived. Hybrid semi-parametric modeling can balance the advantages and disadvantages of strictly mechanistic and nonparametric modeling. In relation to those approaches it can award with several benefits, such as higher estimation/prediction accuracy, better calibration properties, enhanced extrapolation properties, more efficient model development or better interpretability (for details see – supplementary material – section 1). The main advantage is a higher benefit/cost ratio to solve complex problems, which is a key factor for process systems engineering.

Problems in the application of hybrid semi-parametric models mostly concern the model implementation and especially the implementation of the algorithms for the parameter identification is error prone and laborious. However, once a general hybrid semi-parametric modeling tool is implemented, it can easily be reused. It should also be noted that the limitations of mechanistic or nonparametric models may pertain if the hybrid semi-parametric model is not carefully developed or/and the experiments are not carefully designed.

2. Hybrid semi-parametric modeling: the framework

Hybrid semi-parametric models combine nonparametric and parametric models that are based on different types of knowledge. Questions about model configuration, integration of various knowledge types, representation of unknown parts and their identification, best model set-up and requirements on experimental data will be addressed in detail below.

2.1. How to arrange the models? Hybrid semi-parametric model structures

Two models can be arranged in three ways, see Fig. 2, where structure **A** is referred to as parallel and structures **B** and **C** are called serial, sequential, cascade or consecutive. These structures are theoretically addressed in Agarwal (1997) considering that the white box would represent mechanistic information, and the black box consists of a nonparametric model. However, in the serial case, the order of the black and the white model might not be interchangeable. This is for instance the case when the white box in the serial structure **B** represents a material balance equation in which the kinetic rate term is an input variable that is always computed first by a nonparametric model (e.g. Psichogios & Ungar, 1992). However, the information flow between two serial connected submodels can be bidirectional.

2.1.1. The parallel structure

The parallel structure **A** usually finds application if a process model (white box) is available, but its performance, due to whatever reasons (e.g. unmodeled effects, nonlinearities, dynamic behavior) is limited. The parallel arrangement of a nonparametric model can lead to significantly improved estimations. Of course the prediction power of the nonparametric model remains poor on input constellations that have not been trained. The parallel approach is especially interesting if certain effects in the system can be uncoupled (e.g. a static nonlinear and dynamic linear behavior as in block-oriented models) and thus each effect can be represented by a separate model (Abonyi, Chovan, Nagy, & Szeifert, 1999; Chen, Hontoir, Huang, Zhang, & Morris, 2004; Klimasauskas, 1998; Masri,