Journal of Process Control 24 (2014) 272–285

Contents lists available at ScienceDirect

Journal of Process Control

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

Review

# A review on data-driven linear parameter-varying modeling approaches: A high-purity distillation column case study☆

A.A. Bachnas[a], R. Tóth[a,*], J.H.A. Ludlage[a], A. Mesbah[b]

[a] Control Systems Group, Electrical Engineering Department, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
[b] Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

ARTICLE INFO

*Article history:*
Received 22 May 2013
Received in revised form
24 December 2013
Accepted 27 January 2014

*Keywords:*
System identification
Linear parameter-varying systems
Support vector machine
Orthonormal basis functions
Prediction error minimization
High-purity distillation column

ABSTRACT

Model-based control strategies are widely used for optimal operation of chemical processes to respond to the increasing performance demands in the chemical industry. Yet, obtaining accurate models to describe the inherently nonlinear, time-varying dynamics of chemical processes remains a challenge in most model-based control applications. This paper reviews data-driven, Linear Parameter-Varying (LPV) modeling approaches for process systems by exploring and comparing various identification methods on a high-purity distillation column case study. Several LPV identification methods that utilize input–output and series expansion model structures are explored. Two LPV identification perspectives are adopted: (i) the local approach, which corresponds to the interpolation of Linear Time-Invariant (LTI) models identified at different steady-state operating points of the system and (ii) the global approach, where a parametrized LPV model structure is identified directly using a global data set with varying operating points. For the local approach, various model interpolation schemes are studied under an Output Error (OE) noise setting, whereas in the global case, a polynomial parametrization based OE prediction error minimization approach, an Orthonormal Basis Functions (OBFs) based model estimator and a Least-Square Support Vector Machine (LS-SVM) based non-parametric approach are investigated. Through extensive simulation studies, the aforementioned LPV identification approaches are analyzed in terms of the attainable model accuracy and local frequency response behavior of the obtained models. Recommendations are provided to achieve adequate choice between the methods for a particular process system at hand.



**Contents**



☆ This work was supported by the European Union's 7th Framework Program (FP7/2007-2013) under the grant Autoprofit (Grant No.: 257059); (www.fp7-autoprofit.eu).
* Corresponding author. Tel.: +31 638312635.
*E-mail addresses:* a.a.bachnas@tue.nl (A.A. Bachnas), r.toth@tue.nl, tothrola@ieee.org (R. Tóth), j.ludlage@tue.nl (J.H.A. Ludlage), amesbah@mit.edu (A. Mesbah).

0959-1524/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jprocont.2014.01.015

## 1. Introduction

Chemical processes often exhibit a significant nonlinear behavior when operated over a wide range of operating conditions. Despite the advances in model-based control technologies, it remains a challenge to realize high-performance operation of nonlinear chemical processes (in terms of product quality and process productivity) using a single *linear time-invariant* (LTI) model-based controller. For example, when a chemical process is operated under transient conditions (e.g., set-point changes and start-up procedures), the nonlinear behavior of the process becomes more dominant, which makes the use of a single LTI model inadequate to describe the system dynamics over the entire operating window. Distillation columns are a representative example of process systems that exhibit nonlinear dynamics and gain directionality when operated in the high-purity region [1–3].

To meet with the increasing performance demands in the process industry, model-based control strategies are commonly utilized for optimal process operation (e.g., see [4–8] and the citations therein for model-based control of distillation columns). These control approaches require accurate dynamic models to obtain satisfactory closed-loop performance and robustness with respect to process uncertainties. Describing the dynamics of chemical processes using first-principles laws often leads to complex, nonlinear models consisting of a large number of ordinary differential equations, which are typically computationally intractable for real-time control applications. For example, the dynamics of a distillation column can be described by a set of nonlinear, *differential algebraic equations* (DAEs), whose numerical complexity (in terms of the number of equations) can increase significantly with the number of theoretical trays, which can be in the order of hundreds in the high-purity case. Another common complication in first-principles modeling of chemical processes can arise from the need for rigorous descriptions of physicochemical phenomena governing the system dynamics (e.g., thermodynamics driving vapor–liquid equilibria in a distillation column) that may not be obtained straightforwardly.

Data-driven modeling (system identification) methods offer an alternative modeling approach to address the inherent challenges of first-principles modeling of process systems. These approaches enable arriving at a relatively simple, control-relevant description of the system dynamics. To achieve this objective, the primary question is which (possibly nonlinear) model structure(s) should be selected to capture the behavior of the process with a low complexity description.

In this paper, the concept of *linear parameter-varying* (LPV) identification [9] is exploited to describe the nonlinear dynamics of process systems over a wide range of operating conditions. As a generalization of the classical concept of *gain scheduling*, the LPV framework is able to describe nonlinear process dynamics in terms of a time-varying linear dynamical relation between the system input and output signals. The variations of the latter dynamical relation depend on a so-called *scheduling signal*, which represents the changes in the operating conditions of the process. Thus, LPV models preserve the advantageous properties of LTI models, while being able to represent a large class of nonlinear systems [10]. LPV model-based control has been applied in many application areas (e.g., aerospace engineering, automotive applications, high-tech systems) as it benefits from the extension of successful LTI control design strategies, such as *proportional-integral-derivative* (PID) control [11], *model predictive control* (MPC) [12], optimal control [13], and robust control [14] (see [15–25,10] and the citations therein).

This work explores the applicability of various LPV identification approaches for data-driven modeling of chemical processes using a high-purity distillation column case study. To give an overview of the available approaches, several methods that use *input–output* (IO) and *series-expansion* model structures are reviewed and compared. In view of these approaches, two different LPV identification perspectives are explored to describe the process dynamics. The first perspective, called the *local approach*, relies on the identification of multiple LTI models at several operating points of the process, which is followed by the interpolation of the resulting models over the entire operating range. Several widely used interpolation schemes and methods (see [19,26–30,21,17]) are studied. The other investigated LPV identification approach, called the *global approach*, is based on a single data set that spans over a large range of operating conditions. The data set is used to identify the functional dependencies of a linear model structure on the scheduling variable. To explore the performance of the global approach, a classical prediction error minimization approach using an input–output polynomial model structure with a priori chosen parametrization of the scheduling dependencies [31,20] is compared with an *orthonormal basis functions* (OBF's) based model estimator [17] and a *least-square support vector machine* (LS-SVM) based non-parametric approach [32]. The OBF method, which uses a series expansion based model structure, has been shown to exhibit high performance in the modeling of a tank reactor