15th Triennial World Congress, Barcelona, Spain

CONTINUOUS-TIME MODEL IDENTIFICATION OF REAL-LIFE PROCESSES WITH THE CONTSID TOOLBOX

H. Garnier

Centre de Recherche en Automatique de Nancy (CRAN), CNRS UMR 7039
Université Henri Poincaré, Nancy 1, BP 239, F-54506 Vandœuvre-lès-Nancy Cedex, France,
tel: (33) 3 83 91 20 69 - fax: (33) 3 83 91 20 30
e-mail : `hugues.garnier@cran.uhp-nancy.fr`

Abstract: The CONtinuous-Time System IDentification (CONTSID) toolbox methods are first applied to three real-life SISO and MIMO processes selected from thermal and electro-mechanical fields to illustrate their possibilities. The application results to an industrial distillation column demonstrate in a second part that continuous-time model based system identification techniques included in the CONTSID Matlab toolbox are now mature enough to be applied to industrial processes.

Keywords: Continuous-time model, distillation column, industrial application, Matlab toolbox, system identification.

1. INTRODUCTION

Interest in continuous-time (CT) approaches to system identification has been growing in the very recent years. A large amount of publications reflects the intensive effort devoted to the development of theory for such techniques (Johansson *et al.*, 1999), (Garnier *et al.*, 2000), (Pintelon *et al.*, 2000), (Söderström and Mossberg, 2000), (Bastogne *et al.*, 2001).

However although the interest of the system identification community has been increasing, no software tool to perform model identification directly in the CT domain was available until the very recent development of a Matlab toolbox called the CONTSID (CONtinuous-Time System IDentification) toolbox (Garnier and Mensler, 2000). It contains most of the identification methods found in the literature that make it the parameter estimation of CT models for linear time-invariant (LTI) systems possible, directly from discrete-time (DT) data. The toolbox has been used recently to compare the performances of most of the implemented methods (Mensler *et al.*, 2000).

The purpose of the current paper is firstly to present the application results of some of the CONTSID toolbox methods on real-life SISO and MIMO processes selected from thermal and electro-mechanical fields. This first part illustrates the available functions that identify directly continuous-time transfer function and state-space models from sampled data. Secondly, the results of the practical application to an industrial binary distillation column are described.

This paper is organized as follows. A brief overview of the contents of the CONTSID toolbox is given in section 2. Section 3 is devoted to the presentation of the application results obtained on three practical examples. In section 4, the identification of a continuous-time linear model of an industrial distillation column is described.

2. OVERVIEW OF THE CONTSID TOOLBOX

The CONTSID Matlab toolbox contains time-domain identification methods of CT parametric models for LTI SISO and MIMO systems operating in open-loop from sampled data (Garnier and Mensler, 2000). It is freely available for academic researchers and can be downloaded from:

http://www.cran.uhp-nancy.fr/cran/i2s/contsid/contsid.html

The general scheme for direct CT model identification can be divided into two distinct stages:

- the first stage is specific to CT model identification. It consists in applying to the input/output data a linear transformation (LT) in order to avoid the differential issue.
- the second stage concerns the parameter estimation where most algorithms developed for DT model identification can be used.

There are a multitude of choice for the LT required in the primary stage. The toolbox contains most of the LT methods developed over the last thirty years. From the comparative studies recently presented (Mensler *et al.*, 2000), methods which can be considered as those having the best performances are based on linear filtering: the Generalized Poisson Moment Functionals (GPMF) and State-Variable Filter (SVF) approaches, on Fourier and Hartley modulating functions (FMF and HMF), and on the two particular types of integral methods: the Linear Integral Filter (LIF) and Reinitialized Partial Moment (RPM) techniques. Parameter estimation techniques implemented in the toolbox can be subdivided into the following two families:

(1) Transfer function model estimation schemes to identify SISO or MISO systems.
(2) State-space model estimation schemes to identify MIMO systems.

2.1 *Schemes for transfer function identification*

The implemented schemes consider both Equation Error (EE) and Output Error (OE) model structure-based methods.

Several parameter estimation algorithms using all implemented LT are available for the EE approach. First, conventional LS-based LT algorithms have been completed. In order to overcome the bias problem associated with LS-based estimation in presence of noisy data, a bootstrap estimator of Instrumental Variable (IV) type where the IV are constructed using an auxiliary model has also been coupled with all available LT. A recently implemented approach concerns the iterative Simplified Refined Instrumental Variable for Continuous-time model identification (SRIVC) proposed by (Young and Jakeman, 1980). The SRIVC method was first available in the CAPTAIN Matlab toolbox (see http://www.es.lancs.ac.uk/cres/captain/). Its application results to a SISO thermal process are described in section 3.1.

In the case of OE model structure-based approach, a function to estimate the parameters of MISO models has been implemented using the Levenberg-Marquardt algorithm via sensitivity functions. The application results of the latter algorithm to an industrial binary distillation column are presented in section 4.

2.2 *Schemes for state-space model identification*

A first methodology is based on the *a priori* knowledge of structural indices, and considers the estimation of CT canonical state-space models. It consists first in transforming the canonical state-space model into an equivalent input/output polynomial description which is linear-in-its-parameters and therefore more suitable for the parameter estimation problem. A LT method may then be used to convert the differential equation into a set of linear algebraic equations. The unknown model parameters can finally be estimated by LS or IV-based algorithms (Garnier *et al.*, 1995). The application results to a SIMO pilot crane process are presented in section 3.2.

A second class of multivariable system identification schemes is based on subspace estimation methods. The most commonly known subspace methods were developed for DT model identification (Van Overschee and De Moor, 1996). The association of the more efficient LT methods with subspace methods for CT model identification has been recently developed (Johansson *et al.*, 1999), (Bastogne *et al.*, 2001). The successful application of one of these algorithms to a winding process is described in section 3.3.

3. LABORATORY PROCESS IDENTIFICATION RESULTS

In this section, identification results for three laboratory processes selected from thermal and electro-mechanical fields are summarized. They illustrate the use of the CONTSID toolbox routines to identify both transfer function and state-space models. To ensure maximal reproducibility, the data files have been included as parts of the demonstration program of the CONTSID toolbox (see *idcdemo.m* program).

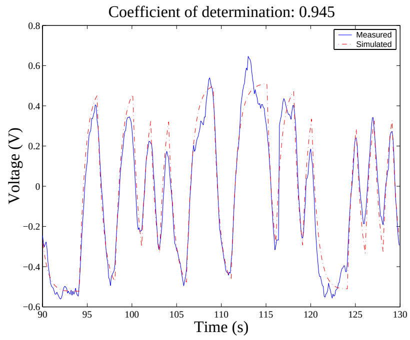

Fig. 1. Cross-validation results for the dryer

3.1 *Dryer identification*

Process description. This SISO laboratory set-up is a bench-scale hot air-flow device. It has been used many times to illustrate the performances of other identification methods. Air is pulled by a fan into a 30 cm tube through a valve and heated by a mesh of resistor wires at the inlet. The output is the voltage delivered by a thermocouple proportional to the air temperature at the outlet of the tube. The input is the voltage over the heating device. The relevant files are *idcdemo8.m* and *dryer.mat*.

Experiment design. The input signal was chosen to be a Pseudo Random Binary Signal (PRBS) of maximum length. The sampling period was set to 100 ms. Two data sets of 1905 measurements collected in the same conditions were used to perform the model estimation and validation. Mean and linear trend of the signals were removed.

Model structure selection. The CONTSID toolbox includes a routine which allows the user to automatically search over a range of different orders (Young, 2002). A simple first order model plus time-delay was selected (see *idcdemo8.m*).

Identification results. The process identification is performed with the SRIVC algorithm *srivc.m* (Young and Jakeman, 1980), (Young, 2002). The identification result is given as a transfer function model. To evaluate the model quality, the coefficient of determination R_T^2 will be considered as a performance index in all application results:

$$R_T^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_{\text{meas}}(i) - y_{\text{sim}}(i)\right)^2}{\sum_{i=1}^{N}\left(y_{\text{meas}}(i) - \bar{y}_{\text{meas}}\right)^2} \qquad (1)$$

where y_{meas} and y_{sim} represent respectively the measured and simulated outputs, $\bar{y}_{\text{meas}}$ is the mean of y_{meas}, N is the number of data.

Cross-validation results are displayed on figure 1 where it may be noticed that the simulated output matches quite well to the measured one.

3.2 *Pilot crane identification*

Pilot description. The pilot is a simplified version of a real trolley crane since hoisting is not considered. It consists of a trolley which can be moved along a metal guiding bar. A pendulum rod with a weight at its end is fixed to the trolley. The trolley is driven via a flexible transmission belt by a current controlled DC-motor. The system input is the voltage of the DC-motor. The measured outputs are the trolley velocity and the load angle. The process is described in more detail in (Garnier *et al.*, 1996). The relevant Matlab files are *idcdemo5.m* and *crane.mat*.

Experiment design. The input was chosen to be a PRBS repeatedly sent five times to the system. The sampling interval is 10 ms. From a careful data analysis, data corresponding to the third and fourth PRBS response were selected for building a model while data corresponding to the fifth PRBS response were chosen for model validation purpose. Mean values of the signals were removed.

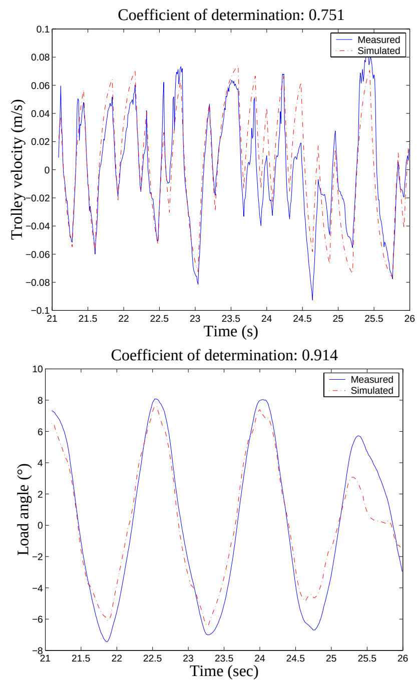

Fig. 2. Pilot crane cross-validation results

Model structure selection. The physical modeling indicates that the global order of the process is 3. The structural index for the trolley velocity is 1 while that for the load angle is equal to 2.

Identification results. The process identification is performed with the State-Space IV-based GPMF algorithm *ssivgpmf.m* (Garnier *et al.*, 1995). The identification result is given as a CT canonical state-space model. Cross-validation results are displayed on figure 2 which shows that despite the presence of friction that introduces non-linearities, there is a relatively good agreement between the measured and the simulated model outputs.

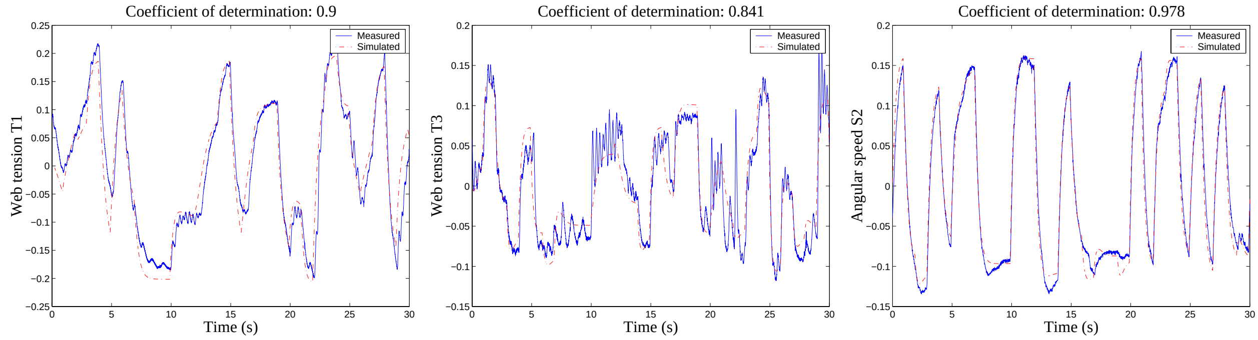

Fig. 3. Cross-validation results for the winding process

3.3 *Winding process identification*

Pilot description. The main part of this MIMO pilot plant is a winding process composed of a plastic web and three reels. Each reel is coupled with a direct-current motor via gear reduction. The angular speed of each reel (S_1, S_2, S_3) is measured by a tachometer while the tensions between the reels (T_1, T_3) are measured by tension meters. At a second level, each motor is driven by a local controller. Two PI control loops adjust the motor currents (I_1) and (I_3) and a double PI control loop drives the angular speed (S_2). The set-points of the local controllers (I_1^*, S_2^*, I_3^*) constitute the manipulated inputs of the winding system $u(t) = \left[I_1^*(t)\ S_2^*(t)\ I_3^*(t)\right]^T$. Driving a winding process essentially comes down to controlling the web linear velocity and the web tensions (T_1) and (T_3) around a given operating point. Consequently, the output variables of the winding system are $y(t) = \left[T_1(t)\ T_3(t)\ S_2(t)\right]^T$. The process is described in more detail in (Bastogne *et al.*, 2001). The relevant Matlab files are *idcdemo7.m* and *winding.mat*.

Experiment design. Discrete-time internal binary sequences were used as excitation signals. The sampling period is set to 10 ms. Mean and linear trend of the signals were removed.

Model structure selection. The system order has been previously estimated and set to $n = 3$. Note however that the algorithm makes it possible to estimate the system order along with the model parameters if it is not known a priori.

S*Identification results.* The process identification is performed with the 4SID-based GPMF algorithm *sidgpmf.m*. The identification result is given as a CT state-space model which can be used to obtain the simulated outputs.

Cross-validation results are plotted on figure 3 where it may be observed that there is a very good agreement with quite high values for the coefficient of determination.

4. INDUSTRIAL DISTILLATION COLUMN IDENTIFICATION

4.1 *Column description*

Figure 4 shows a schematic description of the industrial binary distillation column. It is equipped with 48 trays, a steam-heated reboiler and a total condenser. The column is fed in at the 18^{th} tray with a binary mixture of carbonate components. The separation of components takes place under controlled pressure. The objectives are to control the impurity of the top product or distillate X_t and the impurity of the bottom product or residue X_b with respect to changes on reflux flow F_r and heating power Q while preventing influence of changes on feed flow F_f and feed composition. The distillate and residue X_t and X_b are measured by means of analyzers and expressed in volume per million (vpm). The process is described in more detail in (Defranoux *et al.*, 2000).

4.2 *Experiment design*

Two kinds of experiment were carried out while respecting constraints imposed by the industrial company. These constraints were first to not perturb the production, since the top composition is a finished product; and secondly to manipulate the inputs separately for security and productivity reasons. This latter constraint has imposed that the inputs were perturbed separately and that MISO identification was utilized. The sampling time was set to 10 s. The two experiments, therefore, consisted of manipulating separately the set-points of the reflux flow, F_r, and of the temperature of tray 40, T_{40}, around their normal operating point; the other variables being locally controlled. The experiment lasted between 5 and 17 hours. The manipulated variables were chosen as zero-mean Random Binary Signals (RBS). Two RBS with a magnitude of 0.3 t/h and of 1.5°C were separately applied to the reflux flow F_r and to the temperature T_{40}, respectively, as illustrated in Figure 5. Before executing the estimation pro-

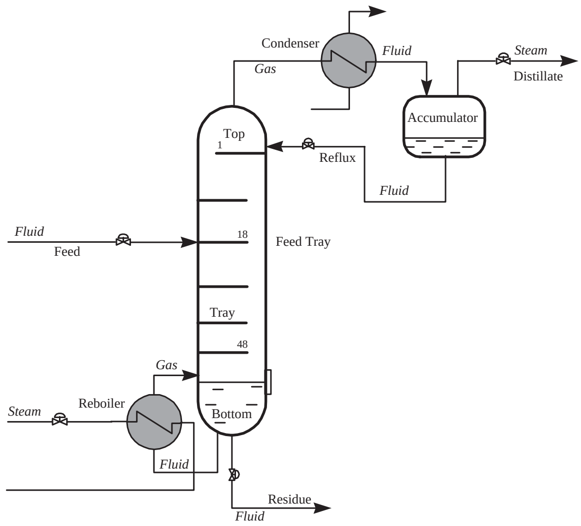

Fig. 4. Schematic description of the column

cedure, a classical data pre-processing was carried out on the raw data sets.

4.3 *Model structure selection*

From an acute data analysis, it turns out that the temperature measurement of the sensitive tray T_{12} could be considered as a continuous image of the distillate X_t that reacts quickly towards changes. This sensitive tray temperature T_{12} has, therefore, been considered as an output variable instead of the distillate. No temperature tray could however represent a continuous image of the residue X_b which constitutes the second output of the model. Classically the reflux flow F_r and heating power represented by the controlled temperature T_{40} are used as input variables for the system. The most important disturbance entering this distillation column is a change in the feed flow rate. The feed flow rate F_f being measured, it has therefore been included as a third input variable for the model. The multivariable coupling in the process can be then be described by the following model:

$$\begin{pmatrix} T_{12}(s) \\ X_b(s) \end{pmatrix} = \begin{pmatrix} H_{11}(s) & H_{12}(s) & H_{13}(s) \\ H_{21}(s) & H_{22}(s) & H_{23}(s) \end{pmatrix} \begin{pmatrix} F_r(s) \\ T_{40}(s) \\ F_f(s) \end{pmatrix} \quad (2)$$

where s denotes the Laplace variable.

4.4 *Identification results*

The data used here are real measurements on a real industrial commercial production plant. Unfortunately we were not allowed to include the data sets in the CONTSID toolbox. The process identification is performed with the OE structure-based algorithm named *coe.m*.

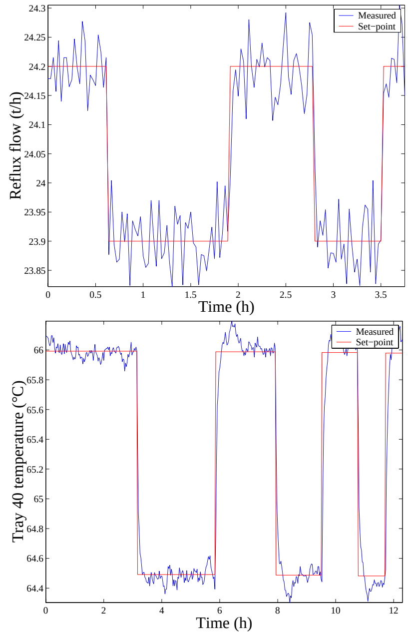

Fig. 5. RBS excitation signals

The measured reflux flow and tray 40 temperature and not the set-points for these variables were considered in the identification procedure. Time delays from input to output variables were previously estimated from step responses. During the experiment, the feed flow changes did not disturb explicitly the bottom product composition of the column. The distant position of the feed tray (figure 4) in respect to the bottom of the column probably explains this phenomenon. The transfer function H_{23} was not, therefore, considered in the estimation procedure and was set to zero. Furthermore, no coupling between the reflux flow F_r and the residue X_b could be demonstrated. Consequently, the transfer function H_{21} was also set to zero and was not considered in the estimation procedure. This explains why there is no cross-validation plot for X_b in the case of excitation on the reflux flow set-point. Cross-validation results are presented in figures 6 and 7. They are of identical quality to those obtained by using an indirect approach consisting first of estimating a discrete-time model by a prediction error method and then converting it into a continuous-time one (Defranoux *et al.*, 2000).

5. CONCLUSION

This paper first presents a brief overview of continuous-time approaches to system identification implemented in the CONTSID toolbox. The toolbox contains a set of Matlab functions which implement most of the continuous-time model identification techniques from discrete-time data. These Matlab functions are easy to use and enhance the understanding as well as the applicability of the algorithms. The paper then demons-

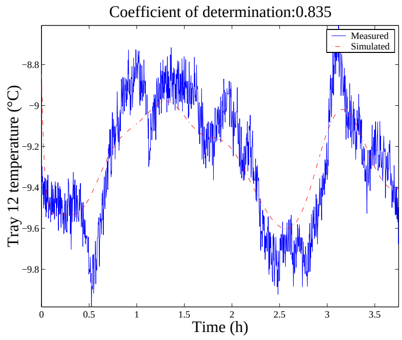

Fig. 6. Cross-validation results for RBS excitation on the reflux flow set-point

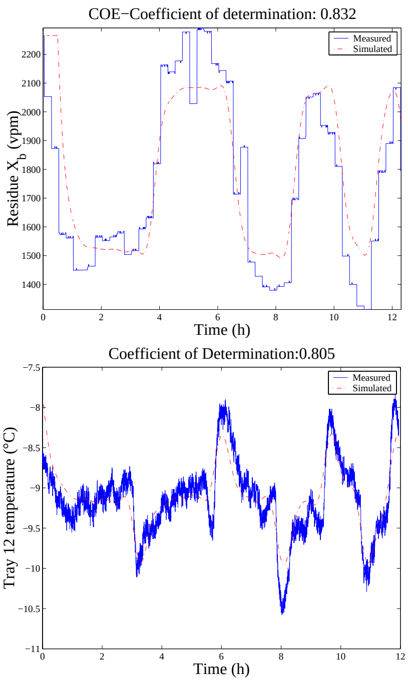

Fig. 7. Cross-validation results for RBS excitation on the tray 40 temperature set-point

trates the efficiency of the CONTSID identification algorithms on three sets of laboratory process data. These practical examples can be easily reproduced by running the demonstration programs available in the toolbox. One application of the algorithms to an industrial binary distillation column has been further presented. This latter application of the CONTSID toolbox techniques demonstrates that continuous-time models based system identification techniques are now mature enough to be applied to industrial processes.

Acknowledgement

The author is grateful to many colleagues, who directly or indirectly, have contributed to the practical identification results presented in this paper: T. Bastogne, C. Defranoux, M. Gilson, E. Huselstein, M. Mensler, P. Sibille, A. Richard.

6. REFERENCES

Bastogne, T., H. Garnier and P. Sibille (2001). A PMF-based subspace method for continuous-time model identification. Application to a multivariable winding process. *International Journal of Control* **74**(2), 118–132.

Defranoux, C., H. Garnier and P. Sibille (2000). Identification and control of an industrial binary distillation column: a case study. *Chemical Engineering and Technology* **23**(9), 745–750.

Garnier, H. and M. Mensler (2000). The CONTSID toolbox: a Matlab toolbox for CONtinuous-Time System IDentification. In: *12th IFAC Symposium on System Identification*. Santa Barbara (USA).

Garnier, H., M. Gilson and W.X. Zheng (2000). A bias-eliminated least-squares method for continuous-time model identification of closed-loop systems. *International Journal of Control* **73**(1), 38–48.

Garnier, H., P. Sibille and A. Richard (1995). Continuous-time canonical state-space model identification via poisson moment functionals. In: *34th IEEE Conference on Decision and Control*. Vol. 2. New Orleans - USA. pp. 3004–3009.

Garnier, H., P. Sibille, M. Mensler and A. Richard (1996). Pilot crane identification and control in presence of friction. In: *13th Triennial World Congress*. San Francisco (USA). pp. 107–112.

Johansson, R., M. Verhaegen and C.T. Chou (1999). Stochastic theory of continuous-time state-space identification. *IEEE Trans. on Signal Processing* **47**(1), 41–50.

Mensler, M., H. Garnier and E. Huselstein (2000). Experimental comparison of continuous-time model identification methods on a thermal process. In: *12th IFAC Symposium on System Identification*. Santa Barbara (USA).

Pintelon, R., J. Schoukens and Y. Rolain (2000). Box-Jenkins continuous-time modeling. *Automatica* **36**, 983–991.

Söderström, T. and M. Mossberg (2000). Performance evaluation of methods for identifying continuous-time autoregressive processes. *Automatica* **36**, 53–59.

Van Overschee, P. and B. De Moor (1996). *Subspace Identification for Linear Systems - Theory, Implementation, Applications*. Kluwer Academic Publishers.

Young, P.C. (2002). Optimal IV identification and estimation of continuous-time TF models. In: *15th Triennial World Congress on Automatic Control*. Barcelona (Spain).

Young, P.C. and A.J. Jakeman (1980). Refined instrumental variable methods of time-series analysis: Part III, extensions. *Int. J. Control* **31**, 741–764.