15th Triennial World Congress, Barcelona, Spain

NUMERICAL ILLUSTRATIONS OF THE RELEVANCE OF DIRECT CONTINUOUS-TIME MODEL IDENTIFICATION

G.P. Rao * **and H. Garnier** **

* *Advisor, UNESCO-EOLSS Joint Committee,*
PO Box 2623, Abu Dhabi, United Arab Emirates,
gantirao@emirates.net.ae
** *Centre de Recherche en Automatique de Nancy, CNRS UMR 7039,*
Université Henri Poincaré, Nancy 1,
BP 239, F-54506 Vandœuvre-lès-Nancy Cedex, France,
hugues.garnier@cran.uhp-nancy.fr

Abstract: The aim of this paper is to establish a dependable approach to the identification of continuous-time models from sampled data. Some equation-error structure-based methods of the CONTSID (CONtinuous-Time System IDentification) Matlab toolbox that directly estimate continuous-time transfer function models from discrete-time data are compared with some classical discrete-time model identification techniques of the Matlab System IDentification (SID) toolbox. CONTSID is equipped with tools based on the developments over the past three decades and is noncommercial. The results of extensive numerical experiments presented in this paper suggest that direct approaches, that is, those in which continuous-time models are directly identified are superior to the indirect methods in which discrete-time models are first identified and then transformed into continuous-time models. It is also clear from this investigation that for identification problems in a wider context, that is, with choice between discrete-time and continuous-time models, it is desirable to have a set of tools, whose dependability is greatly enhanced by unifying all relevant approaches. *Copyright ©2002 IFAC*

Keywords: Continuous-time model, discrete-time model, Matlab toolbox, sampled data, system identification.

1. INTRODUCTION

Identification of continuous-time (CT) models is a problem of considerable importance in various disciplines such as economics, control, fault detection and signal processing. Early efforts in identifying CT linear time-invariant (LTI) systems began with CT models in their native continuous time domain. Subsequently, rapid developments in digital data systems and computers caused a major shift in the approaches with a go-completely-digital trend. Discrete-time (DT) became the working domain in the field of system identification and identification of DT models from sampled input/output data became the main approach. The identification techniques for DT models with discrete-time data are well documented (Ljung, 1999), (Söderström and Stoica, 1989) and widely applied. However, the last three decades have witnessed considerable development in CT approaches to system identification from sampled data (see (Pintelon *et al.*, 2000), (Söderström and Mossberg, 2000), (Bastogne *et al.*, 2001) for very recent references).

There are two ways to obtain a CT model. An indirect way, suggested by the DT model identification methodology is to estimate from sampled data a DT model first and then convert it into a CT model. A direct approach, suggested by the CT model identification methodology, consists in identifying directly a CT model from the discrete-time data. The basic problem of the direct approach is in handling of the non-measurable time-derivatives. Many methods to circumvent the need to reconstruct these time-derivatives have been devised. A comprehensive survey of these techniques has been first given by (Young, 1981) and then by (Unbehauen and Rao, 1987) and (Unbehauen and Rao, 1990). A book has also been devoted to these so-called 'direct' methods (Sinha and Rao, 1991). The CONtinuous-Time System IDentification (CONTSID) toolbox has been developed on the basis of these methods (Garnier and Mensler, 2000).

This paper examines the direct approaches included in the non-commercial Matlab CONTSID toolbox and indirect approaches available in the commercial Matlab System IDentification (SID) toolbox in the light of the results of an extensive simulation experiment in which each approach is subjected to a set of Monte Carlo simulations. The methods are assessed with reference to their performance in terms of accuracy and most importantly dependability. The latter criterion refers to 'stability rate' which means the number of runs that yield a stable model.

The paper is organized in the following way. Section 2 briefly reviews direct and indirect approaches available to estimate a continuous-time model from sampled data. Section 3 outlines the CT and DT model identification methods that are considered in this investigation. Section 4 describes the simulation conditions and section 5 summarizes and compares the results. Discussions and conclusions are given in section 6.

2. RELEVANCE OF CONTSID TOOLBOX METHODS

Much has been written in the literature on the relevance of CT models, see for instance (Unbehauen and Rao, 1990), (Rao and Sinha, 1991). As previously mentioned, the existing methods to determine a CT model from sampled data are usually classified into two broad categories: direct and indirect approaches.

In indirect approaches, a DT model is first identified. The desired CT model parameters are then obtained by transferring the DT model back to the CT. A clear advantage of the indirect methods is that well understood identification methods can be applied (Söderström and Stoica, 1989), (Ljung, 1999). Examples of such methods are the maximum likelihood and prediction error methods, which are known to give consistent and statistically efficient estimates under very general conditions. However, these approaches often require computationally costly minimization algorithms, without even guaranteeing convergence. Moreover, it is well known that, with rapidly sampled data, the DT model identification methods encounter problems due to numerical ill-conditioning, as the poles of the DT model cluster around the point $z = 1$ (Sinha and Rao, 1991). This also shows that conventional DT methods are not in harmony with the CT spirit, as in the limit of reduced sampling period they do not converge to the corresponding CT model. A way to partly overcome this problem is to use, for example, the δ-operator (Middleton and Goodwin, 1990). Further, the question of parameter translation between a DT description and a CT representation is non-trivial. The zeros of the DT model are first not as easily translatable to CT equivalents as the poles (Aström *et al.*, 1984). Secondly, due to the discrete nature of the measurements they do not contain all the information about the CT signals. To describe the signals between the sampling instants some additional assumptions have to be made, for example, assuming that the excitation signal is constant between the sampling intervals (zero-order hold assumption). Violation of these assumptions may lead to severe estimation errors (Schoukens *et al.*, 1994). And finally, non-uniformly sampled data cannot be handled directly.

On the other hand, direct CT model identification approaches available in the CONTSID toolbox are particularly well suited in the case of:

- multi-scale systems;
- fast sampled data;
- non-uniformly sampled data.

Two additional advantages can also be mentioned:

- they have an implicit pre-filtering stage. Data pre-filtering is indeed known to be an important prerequisite and can be considered as implicit in the direct approaches.
- they are not sensitive to the input type (ZOH or bandlimited assumption).

Some of these advantages of direct approaches, which can be considered as weaknesses for indirect methods, will be illustrated in the numerical simulation study presented below.

3. COMPARED METHODS

3.1 *Direct methods included in the CONTSID toolbox*

These methods usually work in two steps. The first step arises out of the input/output time-derivative measurement problem. The need to generate these time-derivatives is eliminated by applying linear transformations to the sampled input/output data. In the second step, the CT model parameters are estimated using some parameter estimation scheme. The pre-filtering feature may be characterized by three types of approaches : the methods of linear filters, the methods of modulating functions and the integral methods. Three linear transformations belonging to each class have been selected here, namely the Generalized Poisson Moment Functionals (GPMF) approach, the Fourier Modulating Function (FMF) method and the Linear Integral Filter (LIF) technique. The three chosen linear transformations are coupled with an instrumental variable (IV) method using an auxiliary model (Young, 1981). The routines denoted as IVGPMF, IVFMF and IVLIF, are available in the Matlab CONTSID toolbox (Garnier and Mensler, 2000).

3.2 *Indirect methods considered from the SID toolbox*

Version 5 of the SID toolbox was used. The chosen CONTSID toolbox techniques will be compared with the indirect approach using four classical DT model identification techniques of the Matlab System Identification toolbox: IV4, N4SID, OE and PEM . Note that in the case of the PEM algorithm, an ARMA noise model of order 2 is considered.

Note also that the SID toolbox includes the possibility of estimating directly CT model from sampled data. The algorithm is restricted to the PEM routine which can be used to determine a CT canonical state-space model. In the case of band-limited input signal, the inter-sample behavior can be set to 'FOH', which can improve the model estimation (see IDDEMO11 in the SID toolbox demonstration program for further information). Routines are also available in the CONTSID toolbox to estimate directly state-space models (see IDCDEMO demonstration program).

4. SIMULATION CONDITIONS

4.1 *System used for the simulation study*

The system considered is a linear fourth-order non-minimum phase system with complex poles whose Laplace transfer function is

$$G_o(s) = \frac{K(-Ts+1)}{\left(\frac{s^2}{\omega_{n,1}^2} + \frac{2\zeta_1 s}{\omega_{n,1}} + 1\right)\left(\frac{s^2}{\omega_{n,2}^2} + \frac{2\zeta_2 s}{\omega_{n,2}} + 1\right)}, \quad (1)$$

with $K = 1$, $T = 4$s, $\omega_{n,1} = 20$ rad/s, $\zeta_1 = 0.1$, $\omega_{n,2} = 2rad/s$, $\zeta_2 = 0.25$.

This is an interesting system from two points of view. It first can be considered as a multi-scale system since it has one fast oscillatory mode with relative damping 0.25 and one slow oscillatory mode with relative damping 0.1. Secondly, the system has a zero in the right half plane. The Bode plot of the system is displayed in figure (1).

4.2 *Configuration of the trials*

In order to study the sensitivity of direct and indirect approaches to the sampling period T_s, noise level, and input type, three different parameters were used:

$$T_s \in \{\text{over-sampling, "normal"}\}$$
$$\text{noise} \in \{\text{deterministic, white } (10 \text{ dB})\}$$
$$\text{input} \in \{\text{PRBS, Multi-sine}\}$$

This leads to several trials whose features are summarized in table 1.

4.3 *Sampling period*

Assuming that the frequency bandwidth of interest is limited to $\omega_{n,1}$, two different sampling periods may be considered. The choice of $T_s = 50$ ms corresponds to what has been called a "normal sampling" choice for DT model identification which represents $\left(\frac{1}{\pi}\right)$ of the Shannon maximum sampling period $\left(\frac{\pi}{\omega_{n,1}}\right)$. In the over-sampling case recommended for CT model identification, the sampling time is set to $T_s = 10$ ms, which represents $\left(\frac{1}{5\pi}\right)$ of the Shannon maximum sampling period.

4.4 *Input signals*

In order to investigate the sensitivity to the input type of CT and DT model identification methods, two signals are considered: a PRBS (Pseudo-Random Binary Signal) of maximum length, respecting the zero-order hold assumption and multi-sine signals, respecting the band-limited assumption. These signals are generated in order to excite the system in its frequency bandwidth.

Multi-sine. The input signal is chosen as the following sum of five sinusoidal signals:

$$u(t) = sin(t) + sin(1.9t) + sin(2.1t) + sin(18t) + sin(22t). \quad (2)$$

The observation time is set to $T = 75s$. Because of the two sampling periods, two multi-sine signals are generated of 7500 and 1250 points respectively.

PRBS of maximum length. The characteristics of the signal whose amplitude switches between -1 and $+1$, are the following: the number of stages of the shift register is set to $ns = 10$, p= 7, (where p is such that the PRBS signal is constant over intervals of length p), which makes a number of points $N = 7161$ in the case of over-sampling while $p = 3$ and $ns = 9$ which makes a number of points $N = 1533$ in the case of "normal" sampling.

4.5 *Stochastic cases*

Most of the runs have been conducted using noisy output by adding a zero-mean white gaussian noise signal $v(t_k)$ to the noise-free output $y_u(t_k)$ (with $t_k = kT_s$) such that

	T_s	input	noise	name
Influence of T_s	over-sampling	Multi-sine	noise-free	trial1
			10 dB	trial2
	"normal" sampling		noise-free	trial3
			10 dB	trial4
Influence of the input type	over-sampling	PRBS	noise-free	trial8
			10 dB	trial9
		Multi-sine	noise-free	trial1
			10 dB	trial2
	"normal" sampling	PRBS	noise-free	trial10
			10 dB	trial11
		Multi-sine	noise-free	trial3
			10 dB	trial4

Table 1. Features of the trials

$$y(t_k) = y_u(t_k) + v(t_k). \tag{3}$$

Monte Carlo simulations with 100 experiments are used for a Signal-to-Noise Ratio (SNR in dB) equals to 10 dB. The SNR is defined as

$$SNR = 10log\frac{P_{y_u}}{P_v}, \tag{4}$$

where P_v and P_{y_u} represent the average power of the additive noise on the system output and of the noise-free output respectively.

In the noise-free case, the three linear transformations are coupled with a simple least-squares algorithm while in the noisy case, they are associated with an IV algorithm based on an auxiliary model.

4.6 *Criteria*

The criteria selected for the performance evaluation are the mean average square error (MSE) of the output and the empirical standard deviation (σ_{SE}) of the average square error (SE) of the output defined by

$$MSE = \frac{1}{N_{exp}} \sum_{i=1}^{N_{exp}} SE(i), \tag{5}$$

$$\sigma_{SE}^2 = \frac{1}{N_{exp}-1} \sum_{i=1}^{N_{exp}} (SE(i) - MSE)^2, \tag{6}$$

where

$$SE = \frac{1}{N} \sum_{k=1}^{N} \varepsilon^2(t_k), \tag{7}$$

$$\varepsilon(t_k) = y_u(t_k) - \hat{y}_u(t_k), \tag{8}$$

where y_u and $\hat{y}_u$ represent the noise-free output of the system and the simulated output of the estimated model respectively. Simulation results, with these two indices being representative of the quality of performance, will be presented here. A third performance index will also be considered. It represents the "stability rate" of the methods, that is, the number of stable models N_{stable} estimated during the Monte Carlo simulations. Note that estimated unstable models were not used to compute the above performance indices.

4.7 *CT model structure selection*

Whatever the input type, a CT model of the following form was considered (p is the differential operator):

$$G(p) = \frac{b_0 + b_1 p}{a_0 + a_1 p + a_2 p^2 + a_3 p^3 + p^4}. \tag{9}$$

4.8 *DT model structure selection*

In the case of the multi-sine input signal which is a band-limited excitation signal, a DT model of the following form was considered (q^{-1} is the backward shift operator):

$$G_{BL}(q) = \frac{b_0^d + b_1^d q^{-1} + b_2^d q^{-2} + b_3^d q^{-3} + b_4^d q^{-4}}{1 + a_1^d q^{-1} + a_2^d q^{-2} + a_3^d q^{-3} + a_4^d q^{-4}}. \tag{10}$$

In the case of the PRBS input signal respecting the ZOH assumption, the form of the DT model to be estimated is the following:

$$G_{ZOH}(q) = \frac{b_1^d q^{-1} + b_2^d q^{-2} + b_3^d q^{-3} + b_4^d q^{-4}}{1 + a_1^d q^{-1} + a_2^d q^{-2} + a_3^d q^{-3} + a_4^d q^{-4}}. \tag{11}$$

Note also that in the case of the four DT model identification methods, the MSE of the output has been computed from the estimated DT model and not from the CT model converted from the estimated DT model.

4.9 *CT method user parameter choice*

The choice of the user parameters for the three CT model identification methods has to be exercised a priori. To analyse its influence, Monte Carlo simulations of 100 experiments were carried out in which the value of the design parameters were varied in the case of trial2 and trial4 simulation conditions. The value which minimized the MSE for the output for each technique was retained. For the minimal-order GPMF approach, the user parameter λ which is the cut-off frequency of the identical first order filter element $\frac{\beta}{s+\lambda}$, has been chosen to 12.5 rad/s (with $\beta = \lambda$). The user parameter of the FMF method has been set to $M\omega_o = 18 rad/s$. The user parameter of the LIF method, which depends upon the sampling interval of the data, has been chosen to $l = 18$ when $T_s = 0.01$s and $l = 4$ when $T_s = 0.05$s.

5. COMPARATIVE SIMULATION RESULTS

5.1 *Performance evaluation in the time-domain*

The simulation results are displayed in table 3. It may be noticed from the table that in a noise-free context (trials 1, 3, 8 and 10), all methods accurately identify the system, even if the SE are higher in the case of CONTSID toolbox approaches. This is due to errors introduced in the implementation stage of all direct approaches which need numerical approximations of either integral or analog filtering. These errors can be however controlled with the sampling time.

5.1.1. *Influence of the sampling period* Results obtained in case of trials 2 and 9 in table 3 can be used to study the effects of over-sampling. The study of the stability rate for the different techniques points out to the clear difference between direct and indirect methods in the case of over-sampling. Not only the DT methods more often lead to unstable models (this is especially true for IV4 and N4SID algorithms), but also when the identified model is stable, its behavior is significantly different from that of the the actual system (MSE is higher). These results confirm the fact that while CT model identification methods are very efficient here, indirect methods using DT model identification techniques may encounter problems due to numerical ill-conditioning effects in the case of rapidly sampled data.

5.1.2. *Influence of the input type* In case of multi-sine input (trials 2 and 4), as expected, CT model identification methods perform much better than DT model identification techniques. When the conditions of experiments become more favourable to DT model identification techniques (that is for a PRBS input and $Ts = 0.05s$ - see results for trial11), the difference in the performance between CT and DT model identification methods become less significant; the CT methods (and more particularly the IVGPMF method) having nevertheless the best global performances in the case of the considered example.

5.2 *Performance evaluation in the frequency-domain*

To evaluate the quality of the estimated models in the frequency-domain, Bode plots of stable estimated models are plotted on figures (1) and (2). Due to lack of place, only Bode diagrams for the IVGPMF and OE methods in the case of trial2 are plotted. These plots confirm the superiority of the direct CT model identification techniques.

Note again that in the case of the DT model identification method, the Bode plots have been computed from the estimated DT models and not from the CT models converted from the estimated DT models.

	IVGPMF	IVFMF	IVLIF	IV4	N4SID	OE	PEM
Trial2	0.5	0.3	0.9	0.6	3.6	13.8	23.4
Trial11	0.15	0.22	0.08	0.11	0.65	0.76	0.97

Table 2. Computational time for the different methods

5.3 *Computational efficiency*

Here, computational efficiency of the different methods is investigated in terms of the computational time to run each algorithm. The computational time for all methods in the case of trial2 and trial11 are given in table 2. All simulations have been executed on a 800 Mhz Pentium computer with 192 Mhz of internal memory with Matlab 6.1. The main conclusion is that there is a significant difference between the direct CT methods which are least-squares based and DT methods except the IV4 routine. The three CT methods require approximately the same amount of time to produce the estimates. They are computationally very efficient compared with PEM and OE routines which are much more time-consuming as it is well-known.

6. DISCUSSION AND CONCLUSION

In this paper, direct methods of the CONTSID toolbox and indirect methods of the Matlab SID toolbox to estimate CT models from sampled data are compared in various conditions of simulation. The results of this study clearly show that the indirect route to CT model identification is not fully dependable. Although DT model based methods have proved to be highly successful and useful for many purposes, it is desirable not to use them as an intermediate step on the path towards CT models. This suggests that system identification tools deserve to be enhanced in their capacity; they should offer wider choice of both models and methods. It is appropriate to have a system of tools unifying the various approaches so that it becomes dependable in a situation characterized by a diversity of needs. The SID and CONTSID toolboxes are complementary. With an appropriate unification arrangement, they can form a unique system of tools for system identification that will be comprehensive, more dependable and effective.

7. REFERENCES

Aström, K.J., P. Hagander and J. Sternby (1984). Zeros of sampled systems. *Automatica* **20**(1), 31–38.

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.

Garnier, H. and M. Mensler (2000). The CONTSID toolbox: a Matlab toolbox for CONtinuous-Time

name	T_s	input	noise	criterion	IVGPMF	IVFMF	IVLIF	IV4	N4SID	OE	PEM
trial1	10ms	multi-sine	noise-free	SE	7.8e-005	7.2e-005	4.7e-003	2.2e-004	1.8e-006	3.5e-021	3.5e-021
trial2	10ms	multi-sine	white 10dB	N_{stable}	100	100	100	9	51	78	90
				MSE	2.7e-002	8.2e-002	4.1e-002	1.0e+002	2.9e+003	3.6e+006	1.3e+005
				σ_{SE}	1.6e-002	5.3e-002	2.5e-002	1.1e+001	1.9e+004	2.8e+007	6.4e+005
trial3	50ms	multi-sine	noise-free	SE	3.3e-002	3.9e-002	2.4e+000	1.3e-018	1.1e-002	3.6e-026	3.6e-026
trial4	50ms	multi-sine	white 10dB	N_{stable}	100	100	100	42	65	99	96
				MSE	2.0e-001	4.8e-001	2.3e+000	4.5e+001	9.6e+000	2.1e+000	1.6e+002
				σ_{SE}	1.2e-001	3.4e-001	1.1e+000	1.7e+001	1.1e+001	1.7e+000	5.6e+002
trial8	10ms	PRBS	noise-free	SE	1.1e-005	7.0e-002	4.1e-002	3.3e-005	8.1e-019	1.8e-022	1.8e-022
trial9	10ms	PRBS	white 10dB	N_{stable}	100	100	100	31	74	88	96
				MSE	2.9e-003	1.4e-001	4.7e-002	9.3e+000	1.1e+001	2.0e+000	7.0e+000
				σ_{SE}	1.9e-003	9.5e-002	1.7e-002	9.2e+000	1.2e+001	2.7e+000	3.1e+000
trial10	50ms	PRBS	noise-free	SE	1.2e-002	3.4e+000	9.9e-001	5.1e-018	9.7e-025	3.4e-027	3.4e-027
trial11	50ms	PRBS	white 10dB	N_{stable}	100	100	100	51	97	92	87
				MSE	4.0e-002	4.0e+000	1.1e+000	2.3e+001	4.8e+000	2.0e+000	1.0e+001
				σ_{SE}	2.1e-002	1.3e+000	2.1e-001	1.6e+001	1.0e+001	4.3e+000	7.3e+000

Table 3. Monte Carlo simulation results for the fourth-order system

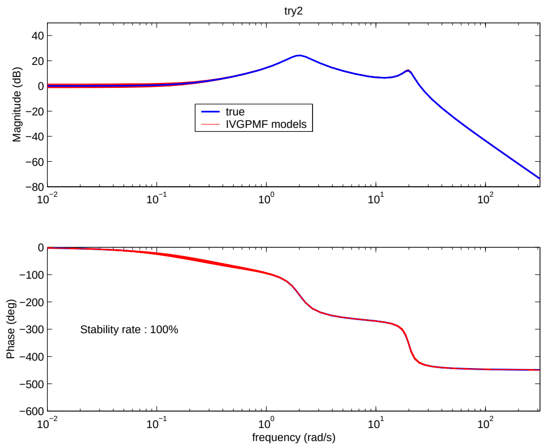

Fig. 1. Bode Diagrams for stable IVGPMF models - trial2 (Multisine - $T_s = 10ms$)

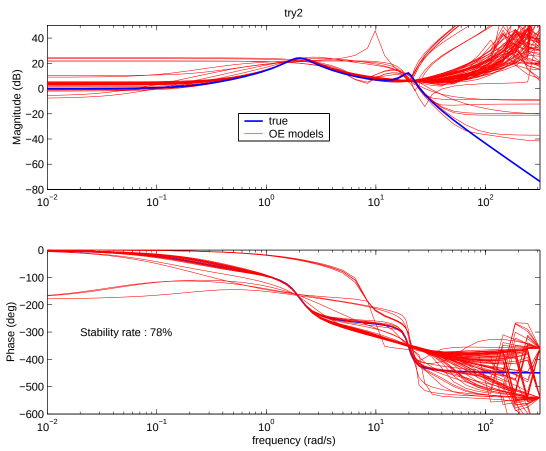

Fig. 2. Bode Diagrams for stable OE models - trial2 (Multisine - $T_s = 10ms$)

System IDentification. In: *12th IFAC Symposium on System Identification*. Santa Barbara (USA).

Ljung, L. (1999). *System identification. Theory for the user, second edition*. Prentice Hall.

Middleton, R.H. and G.C. Goodwin (1990). *Digital control and estimation - A unified approach*. Prentice Hall.

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

Rao, G.P. and N.K. Sinha (1991). *Identification of continuous-time systems. Methodology and computer implementation*. Chap. Continuous-time models and approaches, pp. 1–15. N.K. Sinha and G.P. Rao ed.. Kluwer Academic Publishers.

Schoukens, J., R. Pintelon and H. Van Hamme (1994). Identification of linear dynamic systems using piecewise constant excitations : use, misuse and alternatives. *Automatica* **30**(7), 1953–1169.

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

Söderström, T. and P. Stoica (1989). *System Identification*. Series in Systems and Control Engineering. Prentice Hall.

Sinha, N.K. and G.P. Rao (1991). *Identification of continuous-time systems. Methodology and computer implementation*. Kluwer Academic Publishers.

Unbehauen, H. and G.P. Rao (1987). *Identification of continuous systems*. Systems and control series. North-Holland. Amsterdam.

Unbehauen, H. and G.P. Rao (1990). Continuous-time approaches to system identification - a survey. *Automatica* **26**(1), 23–35.

Young, P. (1981). Parameter estimation for continuous-time models - a survey. *Automatica* **17**(1), 23–39.