1

Direct Identification of Continuous-time Models from Sampled Data: Issues, Basic Solutions and Relevance

Hugues Garnier[1], Liuping Wang[2] and Peter C. Young[3]

[1] Centre de Recherche en Automatique de Nancy, Nancy-Université, CNRS, France
[2] RMIT University, Melbourne, Australia
[3] Lancaster University, UK & Australian National University

1.1 Introduction

Mathematical models of dynamic systems are required in most areas of scientific enquiry and take various forms, such as differential equations, difference equations, state-space equations and transfer functions. The most widely used approach to mathematical modelling involves the construction of mathematical equations based on physical laws that are known to govern the behaviour of the system. Amongst the drawbacks to this approach are that the resulting models are often complex and not easily estimated directly from the available data because of identifiability problems caused by over-parameterisation. This complexity also makes them difficult to use in applications such as control system design.

If sufficient experimental or operational data are available, an alternative to physically-based mathematical modelling is data-based 'system identification', which can be applied to virtually any system and typically yields relatively simple models that can well describe the system's behaviour within a defined operational regime. Such models can be either in a 'black-box' form, which describes only the input–output behaviour, or in some other, internally descriptive form, such as state-space equations, that can be interpreted in physically meaningful terms. This book presents some recent developments in system identification applied to the modelling of continuous-time systems.

Dynamic systems in the physical world are naturally described in continuous-time (CT), differential equation terms because the physical laws, such as conservation equations, have been evolved mainly in this form. Paradoxically, however, the best known system identification schemes have been based on discrete-time (DT) models (sometimes referred to as sampled-data

models), without much concern for the merits of natural continuous-time model descriptions and their associated identification methods. In fact, the development of CT system identification techniques occurred in the the last century, before the development of the DT techniques, but was overshadowed by the more extensive DT developments. This was mainly due to the 'go completely discrete-time' trend that was spurred by parallel developments in digital computers.

Much less attention has been devoted to CT modelling from DT data and many practitioners appear unaware that such alternative methods not only exist but may be better suited to their modelling problems. The identification of continuous-time models is indeed a problem of considerable importance that has applications in virtually all disciplines of science and engineering. This book presents an up-to-date view of this active area of research and describes methods and software tools recently developed in this field.

This chapter is organised as follows. In the first section, the general procedure for system identification is reviewed. Thereafter, the basic features for fitting DT and CT models to sampled data are presented with the objective of highlighting issues in CT model identification. Basic solutions to the main issues are then presented. The main motivations for identifying CT models directly from sampled data are then discussed, before we present some specialised topics in system identification that deserve special attention. At the same time, this introductory chapter aims at tying together the different contributions of the book. In this regard, the outline of the book is presented in the last section.

1.2 System Identification Problem and Procedure

A linear time-invariant continuous-time system with input u and output y can always be described by

$$y(t) = G(p)u(t) + \xi(t) \tag{1.1}$$

where G is the transfer function, p the time-domain differential operator and the additive term $\xi(t)$ represents errors and disturbances of all natures. The source of $\xi(t)$ could be measurement errors, unmeasured process disturbances, model inadequacy, or combinations of these. It is assumed that the input signal $\{u(t),\ t_1 < t < t_N\}$ is applied to the system, with $u(t)$ and the output $y(t)$ both sampled at discrete times $t_1, \cdots, t_N$. The sampled signals are denoted by $\{u(t_k); y(t_k)\}$.

The identification problem can be stated as follows: determine a continuous-time model for the original CT system from N sampled measurements of the

input and output $Z^N = \{u(t_k); y(t_k)\}_{k=1}^N$.

There are three different kinds of parameterised models:

- ***grey-box models***, where the model is constructed in continuous-time from basic physical principles and the parameters represent unknown values of the system coefficients that, at least in principle, have a direct physical interpretation. Such models are also known as physically parameterised or tailor-made models;
- ***black-box models***, which are families of flexible models of general applicability. The parameters in such models, which can be continuous time or discrete time, have no direct physical interpretation (even though the CT version is closer to the physically parameterised model than the DT version), but are used as vehicles to describe the properties of the input–output relationships of the system. Such models are also known as ready-made models;
- ***data-based mechanistic (DBM) models***, which are effectively models identified initially in a black-box, generic model form but only considered credible if they can be interpreted in physically meaningful terms.

In this book, we restrict our attention to black-box model identification. The reader is referred, for instance, to [4] and the references therein, for grey-box model identification; and [53] and the references therein, for DBM model identification.

The basic ingredients for the system identification problem are as follows

- the data set;
- a model description class (the model structure);
- a criterion of fit between data and models;
- a way to evaluate the resulting models.

System identification deals with the problem of determining mathematical models of dynamical, continuous-time systems using measured input–output data. Basically this means that a set of candidate models is chosen and then a criterion of fit between model and data is developed. Finally, the model that best describes the data according to the criterion, within the model set, is computed using some suitable algorithm.

There are two fundamentally different time-domain approaches to the problem of obtaining a black-box CT model of a natural CT system from its sampled input–output data:

- the ***indirect approach***, which involves two steps. First, a DT model for the original CT system is obtained by applying DT model estimation methods to the available sampled data; and then the DT model is transformed into the required CT form. This indirect approach has the advantage that

it uses well-established DT model identification methods [23, 39, 52]. Examples of such methods, which are known to give consistent and statistically efficient estimates under very general conditions, are prediction error methods optimal instrumental variable techniques;

- the ***direct approach***, where a CT model is obtained immediately using CT model identification methods, such as those discussed in this book. Without relying any longer on analogue computers, the present techniques exploit the power of the digital tools. In this direct approach, the model remains in its original CT form.

Independent of how the identification problem is approached, a model parametrisation will lead to the definition of a predictor

$$\hat{y}(t_k, \boldsymbol{\theta}) = g(\boldsymbol{\theta}, Z^{k-1}) \tag{1.2}$$

that depends on the unknown parameter vector $\boldsymbol{\theta}$, and past data Z^{k-1}. The general procedure for estimating a parameterised model from sampled data, regardless of whether it is a CT or DT model, is as follows:

1. from observed data and the predictor $\hat{y}(t_k, \boldsymbol{\theta})$, form the sequence of prediction errors

$$\varepsilon(t_k, \boldsymbol{\theta}) = y(t_k) - \hat{y}(t_k, \boldsymbol{\theta}) \quad k = 1, \ldots, N \tag{1.3}$$

2. filter the prediction errors through a linear filter $F(\bullet)$ to enhance or attenuate interesting or unimportant frequency bands in the signals

$$\varepsilon_{\mathrm{f}}(t_k, \boldsymbol{\theta}) = F(\bullet)\varepsilon(t_k, \boldsymbol{\theta}) \tag{1.4}$$

   where $\bullet$ can be the shift operator if the filter is in discrete time or the differential operator when the filter is in continuous time;
3. choose a scalar-valued, positive function $l(\cdot)$ to measure the size or norm of the filtered prediction error

$$l(\varepsilon_{\mathrm{f}}(t_k, \boldsymbol{\theta})) \tag{1.5}$$

4. minimise the sum of these norms

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} V_N(\boldsymbol{\theta}) \tag{1.6}$$

   where

$$V_N(\boldsymbol{\theta}) = \frac{1}{N}\sum_{k=1}^{N} l(\varepsilon_{\mathrm{f}}(t_k, \boldsymbol{\theta})) \tag{1.7}$$

This procedure is general and pragmatic, in the sense that it is independent of the particular CT or DT model parametrisation used, although this parametrisation will affect the minimisation procedure. Indeed, as we will see, some peculiarities occur in CT model identification that do not occur in DT model identification. We deal with these aspects of the estimation problem in the following three sections. For simplicity of presentation, the formulation and basic solution of both CT and DT model identification problems will be restricted to the case of a linear, single-input, single-output system.

1.3 Basic Discrete-time Model Identification

1.3.1 Difference Equation Models

Perhaps the simplest model of a linear, discrete-time system is the linear difference equation

$$y(t_k) + a_1 y(t_{k-1}) + \ldots + a_{n_a} y(t_{k-n_a}) = b_1 u(t_{k-1}) + \ldots + b_{n_b} u(t_{k-n_b}) + v(t_k) \tag{1.8}$$

where the relationship between the input and output is expressed in terms of the sampled sequences $u(t_k)$ and $y(t_k)$ for $k = 1, 2, \ldots, N$.
Equation (1.8) can also be written as

$$A(q^{-1})y(t_k) = B(q^{-1})u(t_k) + v(t_k) \tag{1.9}$$

or

$$y(t_k) = \frac{B(q^{-1})}{A(q^{-1})}u(t_k) + \xi(t_k); \quad \xi(t_k) = \frac{1}{A(q^{-1})}v(t_k) \tag{1.10}$$

with

$$\begin{aligned} B(q^{-1}) =& b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b}, \\ A(q^{-1}) =& 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a} \end{aligned}$$

where q^{-1} is the backward shift operator, *i.e.*, $q^{-1}x(t_k) = x(t_{k-1})$. Equation (1.8) can be expressed in a vector form that is linear in the model parameters

$$y(t_k) = \boldsymbol{\varphi}^T(t_k)\boldsymbol{\theta} + v(t_k) \tag{1.11}$$

with

$$\boldsymbol{\varphi}^T(t_k) = \left[-y(t_{k-1}) \cdots - y(t_{k-n_a})\, u(t_{k-1}) \cdots u(t_{k-n_b})\right] \tag{1.12}$$

$$\boldsymbol{\theta} = \left[a_1 \ \ldots \ a_{n_a} \ b_1 \ \ldots \ b_{n_b}\right]^T \tag{1.13}$$

In this case, the predictor defined in (1.2) takes the form

$$\hat{y}(t_k, \boldsymbol{\theta}) = \boldsymbol{\varphi}^T(t_k)\boldsymbol{\theta} \tag{1.14}$$

1.3.2 The Traditional Least Squares Method

A traditional way of determining $\boldsymbol{\theta}$ is to minimise the sum of the squares of the prediction error $\varepsilon(t_k, \boldsymbol{\theta})$ by defining the criterion function

$$V_N(\boldsymbol{\theta}) = \frac{1}{N}\sum_{k=1}^{N}\left(y(t_k) - \hat{y}(t_k, \boldsymbol{\theta})\right)^2 \tag{1.15}$$

then minimising with respect to $\boldsymbol{\theta}$. In the present case, $\hat{y}(t_k, \boldsymbol{\theta})$ is linear in $\boldsymbol{\theta}$ and the criterion V_N is quadratic, so that $V_N(\boldsymbol{\theta})$ can be minimised analytically to give the least squares (LS) estimate

$$\hat{\boldsymbol{\theta}}_{\mathrm{LS}} = \left[\frac{1}{N}\sum_{k=1}^{N}\boldsymbol{\varphi}(t_k)\boldsymbol{\varphi}^T(t_k)\right]^{-1}\frac{1}{N}\sum_{k=1}^{N}\boldsymbol{\varphi}(t_k)y(t_k) \tag{1.16}$$

Once the regression vector $\boldsymbol{\varphi}(t_k)$ is constructed (from the measured sampled input–output data), the solution can be computed easily. In the statistical literature, this approach is usually referred to as 'linear regression analysis' and the linear, simultaneous equations that yield the solution (1.16) are termed the 'normal equations'. It is important to realise, however, that this is not a classical regression problem because the elements of the regression vector $\boldsymbol{\varphi}(t_k)$ are not exactly known, as required in regression analysis, but are measured variables that can be contaminated by noise. This has deleterious effects on the parameter estimates that are considered later in the book. It should also be noted that this basic LS method is a special case of the more general prediction error method discussed in Section 1.2, where the analytical solution does not exist and recourse has to be made to other optimisation approaches, such as gradient optimisation or iterative 'relaxation' estimation.

1.3.3 Example: First-order Difference Equation

The traditional LS method is outlined below for the case of a simple first-order DT model

$$y(t_k) + a_1 y(t_{k-1}) = b_1 u(t_{k-1}) + v(t_k) \tag{1.17}$$

which can be written in regression form as

$$y(t_k) = [-y(t_{k-1})\, u(t_{k-1})]\begin{bmatrix} a_1 \\ b_1 \end{bmatrix} + v(t_k) \tag{1.18}$$

Now, according to (1.16), from N available samples of the input and output signals observed at discrete times $t_1, \ldots, t_N$, uniformly spaced, the linear LS parameter estimates are given by

$$\begin{bmatrix} \hat{a}_1 \\ \hat{b}_1 \end{bmatrix} = \begin{bmatrix} \frac{1}{N}\sum\limits_{k=1}^{N} y^2(t_{k-1}) & -\frac{1}{N}\sum\limits_{k=1}^{N} y(t_{k-1})u(t_{k-1}) \\ -\frac{1}{N}\sum\limits_{k=1}^{N} y(t_{k-1})u(t_{k-1}) & \frac{1}{N}\sum\limits_{k=1}^{N} u^2(t_{k-1}) \end{bmatrix}^{-1} \begin{bmatrix} -\frac{1}{N}\sum\limits_{k=1}^{N} y(t_k)y(t_{k-1}) \\ \frac{1}{N}\sum\limits_{k=1}^{N} y(t_k)u(t_{k-1}) \end{bmatrix}$$

It is well known that, except in the special case when $v(t_k)$ in (1.8) is a white noise, simple LS estimation is unsatisfactory. Solutions to this problem led to the development of various approaches, as documented in many books (see *e.g.*, [23, 39, 52]).

The simple difference equation model (1.8) and the well-known LS estimator (1.16) represent the simplest archetype of DT model identification.

1.3.4 Models for the Measurement Noise

In the previous section, we parameterised the description of dynamical systems in a particular form. There are many other possibilities that depend on the method used to model the measurement noise. A common approach used in DT model identification is to assume that the additive disturbance $\xi(t_k)$, contaminating the output measurement has a rational spectral density and can be represented as a DT white noise source $e(t_k)$ passed through a linear filter $H(q^{-1})$, *i.e.*,

$$\xi(t_k) = H(q^{-1})e(t_k) \tag{1.19}$$

When the system and noise models are combined, the standard discrete-time model of a linear dynamic system then takes the form

$$y(t_k) = G(q^{-1})u(t_k) + H(q^{-1})e(t_k) \tag{1.20}$$

In general, the estimation of the parameters in this model is a non-linear statistical estimation problem that can be solved in several ways: *e.g.*, gradient optimisation procedures, such as the maximum likelihood and prediction error methods; and iterative procedures, such as optimal instrumental variables.

1.4 Issues in Direct Continuous-time Model Identification

1.4.1 Differential Equation Models

A continuous-time model of the system takes the form of a constant coefficient differential equation

$$\frac{\mathrm{d}^n y(t)}{\mathrm{d}t^n} + a_1 \frac{\mathrm{d}^{n-1} y(t)}{\mathrm{d}t^{n-1}} + \cdots + a_n y(t) = b_0 \frac{\mathrm{d}^m u(t)}{\mathrm{d}t^m} + \cdots + b_m u(t) + v(t) \tag{1.21}$$

where $\frac{d^i x(t)}{dt^i}$ denotes the ith time derivative of the continuous-time signal $x(t)$. Equation (1.21) can be written alternatively as

$$y^{(n)}(t) + a_1 y^{(n-1)}(t) + \cdots + a_n y(t) = b_0 u^{(m)}(t) + \cdots + b_m u(t) + v(t) \tag{1.22}$$

where $x^{(i)}(t)$ denotes the ith time derivative of the continuous-time signal $x(t)$. Equation (1.21) or (1.22) can be written in the alternative time-domain differential operator form

$$A(p)y(t) = B(p)u(t) + v(t) \tag{1.23}$$

or

$$y(t) = \frac{B(p)}{A(p)}u(t) + \xi(t); \quad \xi(t) = \frac{1}{A(p)}v(t) \tag{1.24}$$

with

$$\begin{aligned} B(s) =& b_0 p^m + b_1 p^{m-1} + \cdots + b_m \\ A(s) =& p^n + a_1 p^{n-1} + \cdots + a_n, \quad n \geq m \end{aligned}$$

and p denoting the differential operator, *i.e.*, $px(t) = \frac{dx(t)}{dt}$.
At any time instant $t = t_k$, (1.22) can be rewritten in regression-like form as

$$y^{(n)}(t_k) = \boldsymbol{\varphi}^T(t_k)\boldsymbol{\theta} + v(t_k) \tag{1.25}$$

where the regressor and parameter vectors are now defined by

$$\boldsymbol{\varphi}^T(t_k) = \left[-y^{(n-1)}(t_k) \cdots - y(t_k)\, u^{(m)}(t_k) \cdots u(t_k)\right] \tag{1.26}$$

$$\boldsymbol{\theta}^T = [a_1 \ \ldots \ a_n \ b_0 \ \ldots \ b_m] \tag{1.27}$$

However, unlike the difference equation model, where only sampled input and output data appear, the differential equation model (1.25) contains input and output time derivatives that are not available as measurement data in most practical cases. When compared with DT model identification, direct CT model identification raises several technical issues that are discussed in the following sections.

1.4.2 Input–Output Time Derivatives

The first difficulty in handling CT models is due to the need for the (normally unmeasured) time derivatives of the input–output signals. Various methods have been devised to deal with the reconstruction of the time derivatives [8,37,40–42,45,56]. Each method is characterised by specific advantages, such as mathematical convenience, simplicity in numerical implementation and computation, handling of initial conditions, physical insight, accuracy and others.
One traditional approach that dates from the days of analogue computers [47] is known as the state-variable filter (SVF) method. This method will be reviewed in Section 1.5.1, with the objective to highlight the differences from DT model identification.

1.4.3 Models for the Measurement Noise

Another difficulty with CT model identification is due to continuous-time stochastic processes. Although the noise model can be given in a CT form, difficulties arise in the estimation because of the theoretical and practical problems associated with the use of CT white noise and its derivatives. A noise model in an equivalent discrete-time form is much more flexible and easier to implement in the estimation problem. Thus, a hybrid model parametrisation method has evolved that involves the identification of a CT model for the

process and a DT model for the noise [16,29,58]. The continuous-time hybrid model of a linear dynamic system then takes the following form,

$$x(t) = G(p)u(t) \tag{1.28a}$$
$$\xi(t_k) = H(q^{-1})e(t_k) \tag{1.28b}$$
$$y(t_k) = x(t_k) + \xi(t_k) \tag{1.28c}$$

or,

$$y(t_k) = G(p)u(t_k) + H(q^{-1})e(t_k) \tag{1.29}$$

where the operators have been mixed informally here in order to illustrate the nature of the estimation model. This approach alleviates the practical difficulties that may be encountered in the parameter estimation of the fully stochastic CT model.

1.5 Basic Direct Continuous-time Model Identification

1.5.1 The Traditional State-variable Filter Method

Let us first consider the transfer function (TF) model (1.23) in the simple noise-free case (the noise-free output is denoted as $x(t)$), *i.e.*,

$$A(p)x(t) = B(p)u(t) \tag{1.30}$$

Assume now that a SVF filter with operator model $F(p)$ is applied to both sides of (1.30). Then, ignoring transient initial conditions

$$A(p)F(p)x(t) = B(p)F(p)u(t) \tag{1.31}$$

The minimum-order SVF filter is typically chosen to have the following operator model form[4]

$$F(p) = \frac{1}{(p+\lambda)^n} \tag{1.32}$$

where λ is the parameter that can be used to define the bandwidth of the filter.
Equation (1.31) can then be rewritten, in expanded form, as

$$\left(\frac{p^n}{(p+\lambda)^n} + a_1\frac{p^{n-1}}{(p+\lambda)^n} + \ldots + a_n\frac{1}{(p+\lambda)^n}\right)x(t) = \left(b_0\frac{p^m}{(p+\lambda)^n} + \ldots + b_m\frac{1}{(p+\lambda)^n}\right)u(t) \tag{1.33}$$

Let $F_i(p)$ for $i = 0, 1, \ldots, n$ be a set of filters defined as

[4] The filter dc gain can be made unity if this is thought desirable.

$$F_i(p) = \frac{p^i}{(p+\lambda)^n} \tag{1.34}$$

By using the filters defined in (1.34), Equation (1.33) can be rewritten, as

$$\left(F_n(p) + a_1 F_{n-1}(p) + \ldots + a_n F_0(p)\right) x(t) = \left(b_0 F_m(p) + \ldots + b_m F_0(p)\right) u(t) \tag{1.35}$$

Equation (1.35) can also be written as

$$x_{\mathrm{f}}^{(n)}(t) + a_1 x_{\mathrm{f}}^{(n-1)}(t) + \ldots + a_n x_{\mathrm{f}}^{(0)}(t) = b_0 u_{\mathrm{f}}^{(m)}(t) + \ldots + b_m u_{\mathrm{f}}^{(0)}(t) \tag{1.36}$$

with

$$x_{\mathrm{f}}^{(i)}(t) = f_i(t) * x(t)$$
$$u_{\mathrm{f}}^{(i)}(t) = f_i(t) * u(t)$$

where $f_i(t)$, for $i = 0, \ldots, n$ represent the impulse responses of the filters defined in (1.34) and $*$ denotes the convolution operator. The filter outputs $x_{\mathrm{f}}^{(i)}(t)$ and $u_{\mathrm{f}}^{(i)}(t)$ provide *prefiltered* time derivatives of the inputs and outputs in the bandwidth of interest, which may then be exploited for model parameter estimation.

Consider now the situation where there is an additive noise on the output measurement. Then, at time instant $t = t_k$, substituting $x_{\mathrm{f}}(t)$ for $y_{\mathrm{f}}(t)$, (1.36) can be rewritten in standard linear regression-like form as

$$y_{\mathrm{f}}^{(n)}(t_k) = \boldsymbol{\varphi}_{\mathrm{f}}^T(t_k)\boldsymbol{\theta} + \eta(t_k) \tag{1.37}$$

where $\eta(t_k)$ is a filtered noise term arising from the output measurement noise $\xi(t_k)$ and the filtering operations, while

$$\boldsymbol{\varphi}_{\mathrm{f}}^T(t_k) = \left[-y_f^{(n-1)}(t_k) \cdots - y_f^{(0)}(t_k)\, u_{\mathrm{f}}^{(m)}(t_k) \cdots u_{\mathrm{f}}^{(0)}(t_k)\right] \tag{1.38}$$

$$\boldsymbol{\theta} = [a_1 \ \ldots \ a_n \ b_0 \ \ldots \ b_m]^T \tag{1.39}$$

Now, from N available samples of the input and output signals observed at discrete times $t_1, \ldots, t_N$, not necessarily uniformly spaced, the linear least-squares (LS)-based SVF parameter estimates are given by

$$\hat{\boldsymbol{\theta}}_{\mathrm{LSSVF}} = \left[\frac{1}{N}\sum_{k=1}^{N} \boldsymbol{\varphi}_{\mathrm{f}}(t_k)\boldsymbol{\varphi}_{\mathrm{f}}^T(t_k)\right]^{-1} \frac{1}{N}\sum_{k=1}^{N} \boldsymbol{\varphi}_{\mathrm{f}}(t_k) y_{\mathrm{f}}^{(n)}(t_k) \tag{1.40}$$

It is well known that, except in the special case where $\eta(t_k)$ in (1.37) is zero mean and uncorrelated (white noise), LS-based SVF estimation although simple, is unsatisfactory. For example, even if the noise term $\xi(t_k)$ in (1.24) is white, the resultant parameter estimates are asymptotically biased and inconsistent. Solutions to this problem are the subject of various chapters in this book.

1.5.2 Example: First-order Differential Equation

The LS-based SVF method is outlined below for the case of a simple first-order differential model given by

$$y^{(1)}(t) + a_1 y(t) = b_0 u(t) + v(t) \tag{1.41}$$

Applying a first-order SVF filter to both sides yields

$$\left(\frac{p}{p+\lambda} + a_1 \frac{1}{p+\lambda}\right) y(t) = b_0 \frac{1}{p+\lambda} u(t) + \frac{1}{p+\lambda} v(t) \tag{1.42}$$

which can be rewritten as

$$\left(F_1(p) + a_1 F_0(p)\right) y(t) = b_0 F_0(p) u(t) + F_0(p) v(t) \tag{1.43}$$

Equation (1.43) can be expressed for $t = t_k$ as

$$y_{\mathrm{f}}^{(1)}(t_k) + a_1 y_{\mathrm{f}}(t_k) = b_0 u_{\mathrm{f}}(t_k) + \eta(t_k) \tag{1.44}$$

and written in regression-like form as

$$y_{\mathrm{f}}^{(1)}(t_k) = \left[-y_{\mathrm{f}}(t_k)\, u_{\mathrm{f}}(t_k)\right] \begin{bmatrix} a_1 \\ b_0 \end{bmatrix} + \eta(t_k) \tag{1.45}$$

Now, according to (1.40), from N available samples of the input and output signals observed at discrete times $t_1, \ldots, t_N$, not necessarily uniformly spaced, the linear LS-based SVF parameter estimates are given by

$$\begin{bmatrix} \hat{a}_1 \\ \hat{b}_0 \end{bmatrix} = \begin{bmatrix} \frac{1}{N}\sum_{k=1}^{N} y_{\mathrm{f}}^2(t_k) & -\frac{1}{N}\sum_{k=1}^{N} y_{\mathrm{f}}(t_k) u_{\mathrm{f}}(t_k) \\ -\frac{1}{N}\sum_{k=1}^{N} y_{\mathrm{f}}(t_k) u_{\mathrm{f}}(t_k) & \frac{1}{N}\sum_{k=1}^{N} u_{\mathrm{f}}^2(t_k) \end{bmatrix}^{-1} \begin{bmatrix} -\frac{1}{N}\sum_{k=1}^{N} y_{\mathrm{f}}^{(1)}(t_k) y_{\mathrm{f}}(t_k) \\ \frac{1}{N}\sum_{k=1}^{N} y_{\mathrm{f}}^{(1)}(t_k) u_{\mathrm{f}}(t_k) \end{bmatrix}$$

The differential equation model (1.21) and the traditional LS-based SVF estimator (1.40) represent the simplest archetype of direct CT model identification.

1.6 Motivations for Identifying Continuous-time Models Directly from Sampled Data

There are many advantages in describing a physical system using CT models and also in identifying the CT models directly from sampled data. Here, we implicitly assume that the sampling rate is sufficiently fast to permit the identification of a continuous-time model from sampled data. It is true that DT models may be better suited to the design and simulation of control systems in a digital environment. However, because a DT model is estimated

from sampled data with a fixed sampling rate, it is only really valid for this chosen sampling rate in its later applications. On the other hand, if a CT model is obtained using data collected at a fast sampling rate, this CT model may be discretised into a DT model with any sampling rate (not necessarily related to the original sampling rate). This is particularly advantageous in the situation where the issue is one of choosing the appropriate sampling rate for discrete-time system modelling and control system design.

The following subsections provide a discussion of the various issues of continuous-time versus discrete-time modelling.

1.6.1 Physical Insight into the System Properties

Most physical phenomena are more transparent in a CT setting, as the models of a physical system obtained from the application of physical laws are naturally in a CT form, such as differential equations. A continuous-time model is preferred to its discrete-time counterpart in the situation where one seeks a model that represents an underlying CT physical system, and wishes to estimate parameter values that have a physical meaning, such as time constants, natural frequencies, reaction times, elasticities, mass values, *etc.* While these parameters are directly linked to the CT model, the parameters of DT models are a function of the sampling interval and do not normally have any direct physical interpretation. For example, consider a mechanical system represented by the following second-order CT transfer function

$$\frac{1}{ms^2 + bs + k} \tag{1.46}$$

where the parameters represent the mass, elasticity and friction that have a direct physical meaning. Now, a DT model of the same process will take the following form

$$\frac{b_0 z + b_1}{a_0 z^2 + a_1 z + a_2} \tag{1.47}$$

where z denotes the Z-transform variable. The parameters of the corresponding DT model do not have a direct physical meaning.

In many areas such as, for example, astrophysics, economics, mechanics, environmental science and biophysics, one is interested in the analysis of the physical system [3, 18, 27, 56]. In these areas, the direct identification of CT models has definite advantages.

1.6.2 Preservation of *a priori* Knowledge

The *a priori* knowledge of relative degree (the difference between the orders of the denominator and numerator) is easy to accommodate in CT models and,

therefore, allows for the identification of more parsimonious models than in discrete time. This is obvious in the example of the second-order mechanical system, where additional parameters are introduced in the numerator of the DT transfer function by the sampling process.

1.6.3 Inherent Data Filtering

Explicit prefiltering strategies are recommended in the general approach to system identification [23, 52], where it is shown that these strategies improve the statistical efficiency of the parameter estimates. However, the prefiltering strategy is not inherent in DT model identification and the user is, therefore, confronted with a choice of whether to add prefiltering. This scenario is contrasted with the situation in CT identification, where the prefiltering is inherent and has two roles: in addition to its original use for reconstructing the filtered time derivatives within the bandwidth of the system to be identified, it became clear [58] that it can perform the same, statistically meaningful prefiltering role as in DT identification.

1.6.4 Non-uniformly Sampled Data

In some situations, it is difficult to obtain equidistant sampled data. This problem arises in medicine, environmental science, transport and traffic systems, astrophysics and other areas, where measurement is not under the control of the experimenter or where uniform sampling is practically impossible. For these non-uniformly sampled data systems, the standard DT linear, time-invariant models will not be applicable because the assumption of a uniformly sampled environment, as required for the existence of such discrete-time models, is violated. On the other hand, the coefficients of CT models are assumed to be independent of the sampling period and so they have a built-in capability to cope with the non-uniformly sampled data situation. With a small modification of the data handling procedure, the measurements are considered as points on a continuous line, which do not need to be equidistantly spaced.

1.6.5 Transformation between CT and DT Models

The parameter transformation between DT and CT representations is well studied [32]. The poles of a DT model are mapped according to the poles in the continuous-time model via the relation: $p_\mathrm{d} = \mathrm{e}^{p_\mathrm{c} T_\mathrm{s}}$, where p_d is the discrete-time pole, p_c is the continuous-time pole and T_s is the sampling interval. However, the zeros of the DT model are not as easily mapped as the poles. Even if the continuous-time system is minimum phase (*i.e.*, all zeros in the left half-plane), the corresponding discrete-time model can be non-minimum phase (*i.e.*, possesses zeros outside of the unit circle). In addition, due to the discrete nature of the measurements, the discrete-time

models do not capture all of the information about the CT signals.

Moreover, in order to describe the signals *between* the sampling instants, some additional assumptions have to be made: for example, assuming that the excitation signal is constant within the sampling intervals (the zero-order hold assumption). However, violation of these assumptions may lead to estimation errors [35].

1.6.6 Sensitivity Problems of DT Models at High Sampling Rates

It is well known that discrete-time models encounter difficulties when the sampling frequency is too high in relation to the dominant frequencies of the system under study [1]. In this situation, the DT poles lie too close to the unit circle in the complex domain and the parameter estimates can become statistically ill-defined.

1.6.7 Stiff Systems

Stiff systems are systems with eigenvalues that are of a different order of magnitude, *i.e.*, the system contains both slow and fast dynamics. Since a DT model is related to a single sampling rate, it is often difficult in such situations to select a sampling rate that captures the complete dynamics of the system without any compromise. In order to illustrate this scenario, suppose that there are two time constants in a stiff system, the fast time constant is 1 (s) and the slow time constant is 100 (s). Typically, the sampling interval T_s is selected approximately in the range of 0.1 to 0.25 of the time constant in order to capture the dynamics associated with this time constant. Assume that $T_\mathrm{s} = 0.25$ of the fast time constant, the poles in the discrete-time model are then $\mathrm{e}^{-0.01T_\mathrm{s}} = 0.9975$ and $\mathrm{e}^{-T_\mathrm{s}} = 0.7788$; and the slow pole is now very close to the unit circle in the complex plane (see previous subsection). As a result, a small estimation error could cause the estimated model to become unstable. However, if we now reduce the sampling rate, in order to avoid this difficulty, the dynamics associated with the fast time constant can become poorly identified. For example, suppose that $T_\mathrm{s} = 10$ (s), then poles in the discrete model are $\mathrm{e}^{-0.1} = 0.9048$ and $\mathrm{e}^{-10} = 4.54{\times}10^{-5}$; so that, although the slow pole moves away from the unit circle, the fast pole virtually disappears from the model structure. Thus, we see that DT models find it difficult, at a specified sampling interval, to deal with both the quick and slow dynamics. In contrast to this, a stiff system can be better captured by a continuous-time model estimated from rapidly sampled data and the coefficients of this model are independent of the sampling rate.

1.7 Specialised Topics in System Identification

The general framework of parameter estimation for linear, time-invariant CT models has to include the consideration of additional factors, such as the identification of the model structure (the order of the transfer function polynomials and the size of any pure time delay), the possible non-integral nature of any pure time delay, identification from data collected during closed-loop experiments, *etc.* The following subsections briefly introduce these other factors and how they are discussed in the present book.

1.7.1 Identification of the Model Structure

Data-based modelling of a continuous-time model consists of model structure identification and the estimation of the parameters that characterise this structure. A continuous-time model structure is prescribed by its model order: *i.e.*, the order of its denominator polynomial and a relative degree. Due to the relative degree, there are many candidate model structures for a given model order. The objective of model structure identification is to select the 'best' model structure among all candidates, based on performance indices, which are often the sum of squares of prediction errors, the statistical properties of the errors and numerous statistical identification criteria. Model structure identification will be discussed in Chapter 6.

1.7.2 Identification of Pure Time (Transportation) Delay

An important additional part of the model structure is the existence of a pure time delay parameter. Unlike the situation in DT identification, where the time delay is assumed to be an integral number of sampling intervals and is often absorbed into the definition of the numerator polynomial (as leading zero-valued parameters), the time-delay parameter for CT system models is normally associated directly with the input signal and can have a non-integral value. As a result, the estimation of the time-delay parameter in CT identification deserves special attention. The interesting issues, in this regard, include simultaneously identifying the continuous-time model parameters and time-delays. Identification of systems with unknown time-delay is discussed in Chapters 11 and 12.

1.7.3 Identification of Continuous-time Noise Models

Identification of the system characteristics from output observations only is referred to as time-series analysis in econometrics, blind identification in signal processing, noise modelling in system identification, and operational modal analysis in mechanical engineering. A fundamental problem is how to model a continuous-time stochastic process based on sampled measurements. Several solutions are possible. One of the key issues is how to sample a continuous-time stochastic system. These issues are discussed in Chapter 2.

1.7.4 Identification of Multi-variable Systems

Systems with many input signals and/or many output signals are called multi-variable. Such systems are often more challenging to model. In particular, systems with several outputs can be difficult. A basic reason for the difficulties is that the couplings between several inputs and outputs lead to more complex models. The structures involved are richer and more parameters are required to obtain a good fit. A class of multi-variable system identification schemes, based on the subspace estimation and state-space realisations have emerged since the late 1980s. The use of these subspace methods to identify CT state-space models is discussed in Chapter 10.

1.7.5 Identification in Closed Loop

Many systems have feedback that cannot be interrupted for an identification experiment, as for example when an existing controller cannot safely be disconnected from an industrial process. In this situation, special procedures are necessary to avoid identifiability problems that can be induced by the feedback connection. Closed-loop identification schemes are described in Chapters 5 and 13.

1.7.6 Identification in the Frequency Domain

Linear dynamic systems have equivalent and complementary descriptions: in the time domain and in the frequency domain. Although the two descriptions are basically equivalent to each other, the formulation of the identification problem leads to different methods in the two domains. In many practical situations, parameter estimation in the frequency domain is of considerable interest [30]. Practical aspects of frequency-domain parametric identification methods are discussed in Chapter 8.

1.7.7 Software for Continuous-time Model Identification

System identification is typically an iterative procedure, where the insights and judgements of the user are mingled with formal considerations, extensive data handling and complex algorithms. To make the application of the identification procedure successful, it is almost always necessary to have some user-friendly software tools to facilitate the user's modelling. These software aspects are discussed in Chapters 8 and 9.

1.8 Historical Review

In contrast to the present day, the control world of the 1950s and 1960s was dominated by CT models as most control system design was concerned with

CT systems and most control system implementations employed analogue techniques. Moreover, almost all CT identification methods were largely deterministic, in the sense that they did not explicitly model the additive noise process nor attempt to quantify the statistical properties of the parameter estimates. Nevertheless, it is fascinating to see that some of these early papers introduced interesting concepts that foreshadowed later, important developments of a very similar nature. For instance, Valstar [43] and Young [47, 48] suggested the use of prefilters to solve the derivative measurement problem and this same 'state-variable filter' (SVF) approach[5] was re-discovered, some 20 years afterwards [34], under the title of 'Poisson-moment functionals' (PMF). Most early research also used completely analogue implementation, with both the prefilters and the estimation algorithm implemented in an analogue manner (*e.g.*, [14, 22, 48]); while some, adumbrating developments to come, utilised hybrid implementations where analogue prefiltering was combined with a digital identification algorithm [47, 49]. Indeed, two of these references [14, 22] also consider non-linear system identification, using a purely deterministic 'state-dependent parameter' approach that would emerge, many years later, in a stochastic, purely digital form (*e.g.*, [59]).

Also in the 1960s, it was realised that measurement noise could cause asymptotic bias on the parameter estimates when linear least squares (regression) methods were used to estimate the parameters in dynamic systems. Within a largely deterministic setting, papers appeared (*e.g.*, [20, 46, 49, 50]) that graphically demonstrated the value of the instrumental variable (IV) modification to both the recursive and *en-bloc* least squares algorithms that had been used for CT identification prior to this. Here, the instrumental variables were generated as the output of a recursively updated 'auxiliary model' of the system that, together with the prefilters, was implemented directly in continuous time.

Perhaps because of the dominant interest in DT identification and estimation since 1970, a stochastic formulation of the CT estimation problem did not appear until 1980. Then, Young and Jakeman [58], following the optimal prefiltering and recursive-iterative estimation procedures for DT systems (first presented in [51]), suggested an optimal 'hybrid' refined instrumental variable solution to the CT identification problem (RIVC). This involves a CT model of the system and a discrete-time ARMA model for the noise. However, at that time, it was only implemented in a simplified form (SRIVC) that yields consistent and statistically efficient parameter estimates when the additive noise $\xi(t)$ in (1.24) is white.

Responding to the research on RIVC estimation, Huang *et al.* [15] implemented an alternative hybrid solution that allowed for coloured noise and

[5] Also called the 'method of multiple filters' [49].

utilised a gradient optimisation algorithm, rather than the iterative solution used in the SRIVC algorithm and proposed in the RIVC algorithm. However, they chose to convert the problem into an entirely DT form and so did not implement the prefilters and auxiliary model explicitly in continuous time. Also, they did not present any stochastic simulation results and, as such, it is not possible to reach any clear conclusions about the statistically efficiency of the algorithm.

Despite these excursions into stochastic systems and full statistical estimation, most publications on CT identification during the 1970s and 1980s were deterministic in concept and suggested various methods of implementing prefilters (see [8] for a recent overview for example). Most of the deterministic approaches are available in the continuous-time system identification (CONTSID) toolbox[6] for MATLAB® (see Chapter 9 in this book). Since the deterministic methods have been documented so fully, it will suffice here merely to outline the main features of each approach.

Linear Filter Methods

These methods originated from the third author's early research in this area [47, 48, 50] where the method was referred to as the '*method of multiple filters*' (MMF). It involves passing the input and output signals through a chain of (usually identical) first-order prefilters with user-specified bandwidth, normally selected so that it spans the anticipated bandwidth of the system being identified (see Section 1.5.2 for the simplest example of this method). More recently this MMF approach has been re-named the *generalised Poisson moment functionals* (GPMF) approach [34, 40]. Recent MMF/GPMF developments have been proposed by the first author and his co-workers [2,7,9–11].

Integration-based Methods

The main idea of these methods is to avoid the differentiation of the data by performing an order n integration. These integral methods can be roughly divided into two groups. The first group, using numerical integration and orthogonal function methods, performs a basic integration of the data and special attention has to be paid to the initial condition issue. The second group includes the *linear integral filter* (LIF: [33]) and the *re-initialised partial moments* (RPM: [6]) approaches. Here, advanced integration methods are used to avoid the initial condition problem either by exploiting a moving integration window (LIF) or a time-shifting window (RPM).

[6] http://www.cran.uhp-nancy.fr/contsid: the CONTSID toolbox also contains the SRIVC and RIVC algorithms.

Modulating Function Methods

This approach was first suggested almost half a century ago by Shinbrot in order to estimate the parameters of linear and non-linear systems [36]. Further developments have been based on different modulating functions. These include the Fourier-based functions [26], in either trigonometric or complex exponential form; spline-type functions; Hermite functions and, more recently, Hartley-based functions [42]. A very important advantage of using Fourier- and Hartley-based modulating functions is that the model estimation can be formulated entirely in the frequency domain, making it possible to use efficient DFT/FFT techniques.

Other Methods

Several other approaches have been suggested that cannot be classified directly into any of the categories discussed in the previous subsections. An interesting approach is reported in [16] where the idea is to replace the differentiation as represented by the Laplace operator s by the operator w. These operators are related via the bilinear relationship $w = \frac{s-a}{s+a}$. The new w-domain model can be estimated directly from sampled data, using filtered signals. Afterwards, the parameters of this model are translated back to the parameters of the ordinary continuous-time model, using simple algebraic relations. The w operator can be an all-pass filter. In this case, the filter does not alter the frequency content of the signals and only affects the phase. This setup is closely related to the SVF method (see also [5] for a related scheme where the filters take the form of CT Laguerre functions). Two other approaches that have attracted a lot of attention in the identification community in the 1990s are subspace-based methods (see [2, 13, 17, 21, 24, 25] but also Chapter 10 in this book) and finite difference methods [19, 28, 38]. This latter approach, which is based on replacing the differentiation operator with finite differences, will be considered in some depth in this book (see Chapters 2 and 3).

Most recently, attention has re-focused on stochastic model identification and statistically optimal CT estimation procedures. First, in discussing a paper on optimal CT identification by Wang and Gawthrop [44], Young [54] drew attention to the virtues of the existing SRIVC estimation algorithm and demonstrated its superiority in a simulation example. This encouraged, at last, the implementation of the full hybrid RIVC algorithm [57] that is presented and evaluated in Chapter 4 of this book, as well as the development of an associated closed-loop version of the algorithm [12], which is described in Chapter 5. A useful by-product of this renewed interest in these optimal algorithms is that optimal RIV algorithms are now available for Box–Jenkins-type stochastic transfer function models of CT and DT systems, providing a unified approach to the identification and estimation of transfer function models [55].

References

1. K. J. Aström. On the choice of sampling rates in parametric identification of time series. *Information Science*, 1(1):273–278, 1969.
2. T. Bastogne, H. Garnier, and P. Sibille. A PMF-based subspace method for continuous-time model identification. Application to a multivariable winding process. *International Journal of Control*, 74(2):118–132, 2001.
3. A.R. Bergström. *Continuous-time Econometric Modeling*. Oxford University Press, Oxford, UK, 1990.
4. T. Bohlin. *Practical Grey-box Process Identification. Theory and Applications*. Springer, Series: Advanced in Industrial Control, 2006.
5. C.T. Chou, M. Verhaegen, and R. Johansson. Continuous-time identification of SISO systems using Laguerre functions. *IEEE Transactions on Signal Processing*, 47(2):349–362, 1999.
6. M. Djamai, E. Tohme, R. Ouvrard, and S. Bachir. Continuous-time model identification using reinitialised partial moments. Application to power amplifier modeling. *14th IFAC Symposium on System Identification (SYSID'2006)*, Newcastle, Australia, March 2006.
7. H. Garnier, M. Gilson, and W.X. Zheng. A bias-eliminated least-squares method for continuous-time model identification of closed-loop systems. *International Journal of Control*, 73(1):38–48, 2000.
8. H. Garnier, M. Mensler, and A. Richard. Continuous-time model identification from sampled data. Implementation issues and performance evaluation. *International Journal of Control*, 76(13):1337–1357, 2003.
9. H. Garnier, P. Sibille, and T. Bastogne. A bias-free least-squares parameter estimator for continuous-time state-space models. *36th IEEE Conference on Decision and Control (CDC'97)*, volume 2, pages 1860–1865, San Diego, California, USA, December 1997.
10. H. Garnier, P. Sibille, and A. Richard. Continuous-time canonical state-space model identification via Poisson moment functionals. *34th IEEE Conference on Decision and Control (CDC'95)*, volume 2, pages 3004–3009, New Orleans, USA, December 1995.
11. H. Garnier, P. Sibille, and T. Spott. Influence of the initial covariance matrix on recursive LS estimation of continuous models via generalised Poisson moment functionals. *10th IFAC Symposium on System Identification (SYSID'94)*, pages 3669–3674, Copenhagen, Denmark, 1994.
12. M. Gilson, H. Garnier, P.C. Young, and P. Van den Hof. A refined IV method for closed-loop system identification. *14th IFAC Symposium on System Identification (SYSID'2006)*, pages 903–908, Newcastle, Australia, March 2006.
13. B. Haverkamp. *State Space Identification. Theory and Practice*. PhD thesis, TU Delft, The Netherlands, 2001.
14. L.L. Hoberock and R.H. Kohr. An experimental determination of differential equations to describe simple nonlinear systems. *Joint Automatic Control Conference*, Washington, Seattle, pages 616–623, 1966.
15. H. Huang, C. Chen, and Y. Chao. Augmented hybrid method for continuous process identification from sampled data with coloured noise. *International Journal of Control*, 46:1373–1390, 1987.
16. R. Johansson. Identification of continuous-time models. *IEEE Transactions on Signal Processing*, 42(4):887–896, 1994.

17. R. Johansson, M. Verhaegen, and C.T. Chou. Stochastic theory of continuous-time state-space identification. *IEEE Transactions on Signal Processing*, 47(1):41–50, 1999.
18. R.E. Kearney and I.W. Hunter. System identification of human joint dynamics. *Critical Reviews in Biomedical Engineering*, 18(1):55–87, 2000.
19. E. Larsson. *Identification of Stochastic Continuous-time Systems. Algorithms, Irregular Sampling and Cramér-Rao Bounds*. PhD thesis, Uppsala University, Sweden, 2004.
20. V.S. Levadi. Parameter estimation of linear systems in the presence of noise. *International Conference on Microwaves, Circuit Theory and Information Theory*, Tokyo, Japan, 1964.
21. W. Li, H. Raghavan, and S. Shah. Subspace identification of continuous time models for process fault detection and isolation. *Journal of Process Control*, 13(5):407–421, 2003.
22. P.M. Lion. Rapid identification of linear and nonlinear systems. *Joint Automatic Control Conference, Washington, Seattle*, pages 605–615, 1966.
23. L. Ljung. *System Identification. Theory for the User*. Prentice Hall, Upper Saddle River, USA, 2nd edition, 1999.
24. G. Mercère, R. Ouvrard, M. Gilson, and H. Garnier. Subspace based methods for continuous-time model identification of MIMO systems from filtered sampled data. *9th European Control Conference (ECC'07)*, Kos, Greece, July 2007.
25. A. Ohsumi, K. Kameyama, and K. Yamaguchi. Subspace identification for continuous-time stochastic systems via distribution-based approach. *Automatica*, 38(1):63–79, 2002.
26. A.E. Pearson, Y. Shen, and V. Klein. Application of Fourier modulating function to parameter estimation of a multivariable linear differential system. *10th IFAC Symposium on System Identification (SYSID'94)*, pages 49–54, Copenhagen, Denmark, 1994.
27. M.S. Phadke and S.M. Wu. Modelling of continuous stochastic processses from discrete observations with application to sunspots data. *Journal American Statistical Association*, 69(346):325–329, 1974.
28. D.T. Pham. Estimation of continuous-time autoregressive model from finely sampled data. *IEEE Transactions on Signal Processing*, 48:2576–2584, 2000.
29. R. Pintelon, J. Schoukens, and Y. Rolain. Box-Jenkins continuous-time modeling. *Automatica*, 36(7):983–991, 2000.
30. R. Pintelon and J. Schoukens. *System Identification: a Frequency Domain Approach*. IEEE Press, Piscataway, USA, 2001.
31. G.P. Rao and H. Unbehauen. Identification of continuous-time systems. *IEE Proceedings - Control Theory and Applications*, 153(2):185–220, 2006.
32. K.J. Åström, P. Hagander, and J. Sternby. Zeros of sampled systems. *Automatica*, 20(1):31–38, 1984.
33. S. Sagara and Z.Y. Zhao. Numerical integration approach to on-line identification of continuous-time systems. *Automatica*, 26(1):63–74, 1990.
34. D.C. Saha and G.P. Rao. *Identification of Continuous Dynamical Systems - The Poisson Moment Functionals (PMF) approach*. Springer-Verlag, Berlin, 1983.
35. J. Schoukens, R. Pintelon, and H. Van Hamme. Identification of linear dynamic systems using piecewise constant excitations: use, misuse and alternatives. *Automatica*, 30(7):1953–1169, 1994.
36. M. Shinbrot. On the analysis of linear and non linear systems. *Transactions of the ASME*, 79:547–552, 1957.

37. N.K. Sinha and G.P. Rao (eds). *Identification of Continuous-time Systems. Methodology and Computer Implementation*. Kluwer Academic Publishers, Dordrecht, 1991.
38. T. Söderström, H. Fan, B. Carlsson, and S. Bigi. Least squares parameter estimation of continuous-time ARX models from discrete-time data. *IEEE Transactions on Automatic Control*, 42(5):659–673, 1997.
39. T. Söderström and P. Stoica. *System Identification*. Series in Systems and Control Engineering. Prentice Hall, Englewood Cliffs, 1989.
40. H. Unbehauen and G.P. Rao. *Identification of Continuous Systems*. Systems and control series. North-Holland, Amsterdam, 1987.
41. H. Unbehauen and G.P. Rao. Continuous-time approaches to system identification - a survey. *Automatica*, 26(1):23–35, 1990.
42. H. Unbehauen and G.P. Rao. A review of identification in continuous-time systems. *Annual Reviews in Control*, 22:145–171, 1998.
43. J.E. Valstar. In flight dynamic checkout. *IEEE Transactions on Aerospace*, AS1(2), 1963.
44. L. Wang and P. Gawthrop. On the estimation of continuous-time transfer functions. *International Journal of Control*, 74(9):889–904, 2001.
45. P. Young. Parameter estimation for continuous-time models - a survey. *Automatica*, 17(1):23–39, 1981.
46. P. C. Young. Regression analysis and process parameter estimation: a cautionary message. *Simulation*, 10:125–128, 1968.
47. P.C. Young. In flight dynamic checkout - a discussion. *IEEE Transactions on Aerospace*, 2:1106–1111, 1964.
48. P.C. Young. The determination of the parameters of a dynamic process. *Radio and Electronic Engineering (Journal of IERE)*, 29:345–361, 1965.
49. P.C. Young. Process parameter estimation and self adaptive control. In P.H. Hammond (ed), *Theory of Self Adaptive Control Systems*, Plenum Press, pages 118–140, 1966.
50. P.C. Young. An instrumental variable method for real-time identification of a noisy process. *Automatica*, 6:271–287, 1970.
51. P.C. Young. Some observations on instrumental variable methods of time-series analysis. *International Journal of Control*, 23:593–612, 1976.
52. P.C. Young. *Recursive Estimation and Time-series Analysis*. Springer-Verlag, Berlin, 1984.
53. P.C. Young. Data-based mechanistic modeling of environmental, ecological, economic and engineering systems. *Journal of Modelling & Software*, 13:105–122, 1998.
54. P.C. Young. Optimal IV identification and estimation of continuous-time TF models. *15th IFAC World Congress on Automatic Control*, Barcelona, Spain, July 2002.
55. P.C. Young. The refined instrumental variable method: unified estimation of discrete and continuous-time transfer function models. *Journal Européen des Systèmes Automatisés*, 2008.
56. P.C. Young and H. Garnier. Identification and estimation of continuous-time data-based mechanistic (DBM) models for environmental systems. *Environmental Modelling and Software*, 21(8):1055–1072, August 2006.
57. P.C. Young, H. Garnier, and M. Gilson. An optimal instrumental variable approach for identifying hybrid continuous-time Box–Jenkins model. *14th IFAC*

Symposium on System Identification (SYSID'2006), pages 225–230, Newcastle, Australia, March 2006.

58. P.C. Young and A.J. Jakeman. Refined instrumental variable methods of time-series analysis: Part III, extensions. *International Journal of Control*, 31:741–764, 1980.
59. P.C. Young, P. McKenna, and J. Bruun. Identification of nonlinear stochastic systems by state dependent parameter estimation. *International Journal of Control*, 74:1837–1857, 2001.