Proceeding of the 2004 American Control Conference
Boston, Massachusetts June 30 - July 2, 2004

Time-domain approaches to continuous-time model identification of dynamical systems from sampled data

Hugues Garnier and Peter Young

***Abstract*— This paper first gives a tutorial introduction to the main aspects of existing time-domain methods and software for identifying linear continuous-time models of dynamical systems from sampled input/output data. The second part of the paper demonstrates these approaches via simulated and real data examples.**

I. INTRODUCTION

The identification of continuous-time (CT) models is a problem of considerable importance that has applications in virtually all disciplines of science. Early research on this topic focussed on identification of CT models from CT data (e.g. [1], [2], [3]). Subsequently, however, rapid developments in digital data acquisition and computers have resulted in attention being shifted to the identification of discrete-time (DT) models from sampled data, as documented in many books (see e.g [4], [5] and [6]). 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.

In order to identify a continuous-time model from time-domain sampled data, two main time-domain approaches are possible. In the first, 'indirect' approach, a DT model is identified first using DT identification methods, and this is then converted into a CT model using a standard algorithm for discrete to continuous-time conversion (e.g. *d2cm* in Matlab$^{\text{TM}}$). In the second, 'direct' approach the CT model is identified directly from DT data. These latter approaches are often incorrectly presented as being too complicated but, as we will see, they are straightforward, reliable and have proven useful in many practical applications.

This tutorial concentrates on these 'direct' methods of CT identification. It is not intended to be comprehensive in bibliographic terms and does not attempt to review the literature on the identification of models based on stochastic differential equations (see e.g [7] and the references on this topic therein).

H. Garnier is with Centre de Recherche en Automatique de Nancy, CNRS UMR 7039, Université Henri Poincaré, Nancy 1, BP 239, F-54506 Vandœuvre-les-Nancy Cedex, France hugues.garnier@cran.uhp-nancy.fr

P.C. Young is with Centre for Research on Environmental Systems and Statistics - Systems and Control Group, CRES/IENS, Lancaster University, Lancaster LA1 4YQ, U.K. p.young@lancaster.ac.uk

II. PROBLEM FORMULATION

Consider a linear, single-input, single-output, CT system[1] whose input $u(t)$ and output $y(t)$ are related by a constant coefficient differential equation of order n

$$x^{(n)}(t) + a_1 x^{(n-1)}(t) + \ldots + a_n x^{(0)}(t) = b_0 u^{(m)}(t) + b_1 u^{(m-1)}(t) + \ldots + b_m u^{(0)}(t) \qquad (1)$$

where $x^{(i)}(t)$ denotes the ith time-derivative of the continuous-time signal $x(t)$. Equation (1) can also be written in the transfer function (TF) form:

$$x(t) = \frac{B(p)}{A(p)} u(t), \qquad (2)$$

with

$$B(p) = b_0 p^m + b_1 p^{m-1} + \cdots + b_m,$$
$$A(p) = p^n + a_1 p^{n-1} + \cdots + a_n,$$

where p is the differential operator, *i.e.* $px(t) = \frac{dx(t)}{dt}$. It is assumed that the input signal $\{u(t), t_1 < t < t_N\}$ is applied to the system and that the output $x(t)$ is sampled at discrete times $t_1, \cdots, t_N$, not necessarily uniformly spaced. The sampled signals are denoted by $\{u(t_k); x(t_k)\}$.

In order to obtain high quality estimation results, it is vital to also consider the inevitable errors that will affect the measured output signal. The measured output $y(t_k)$ is assumed to be corrupted by an additive measurement noise $v(t_k)$

$$y(t_k) = x(t_k) + v(t_k)$$

The identification problem can now be stated as follows: estimate the parameters of the differential equation model from N sampled measurements of the input and output $Z^N = \{u(t_k); y(t_k)\}_{k=1}^{N}$.

III. IDENTIFICATION METHODS

In comparison with the DT model identification, direct CT model identification raises several technical issues. Unlike the difference equation model, the differential equation model (1) contains time-derivative terms that may be required and are not normally available for measurement. Various methods have been devised to deal with the need to reconstruct these time-derivatives [8], [9], [10], [11], [12], [13], [14]. Each method is characterized by specific advantages, such as mathematical convenience, simplicity

[1] A time-delay on the system input is not considered for simplicity here but is easy to accommodate.

in numerical implementation and computation, handling of initial conditions, physical insight, and accuracy.

In this tutorial paper, we consider two specific methods that exemplify the historical development of direct CT identification. Initially, most 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. Instead, consistent estimates were obtained by using basic Instrumental Variable (IV) methods. One deterministic approach of this type, known as the state-variable filter (SVF) method [1], [2], [3], dates from the days of analog and hybrid computers. This is reviewed first, with the aim of highlighting some of the peculiarities that occur in comparison with DT model identification. Then, a more sophisticated IV method for direct CT stochastic model identification is outlined in order to demonstrate the advantages of the stochastic model formulation [15], [16], [17].

A. The traditional SVF method

Let us first consider the TF model (1) in the simple noise free-case. This latter can then be written in the linear equation form,

$$A(p)x(t) = B(p)u(t), \tag{3}$$

Assume now that a filter with operator model $F(p)$ is applied to both sides of (3). Then, ignoring transient initial condition effects

$$A(p)F(p)x(t) = B(p)F(p)u(t), \tag{4}$$

The minimum-order SVF filter is typically chosen to have the following form[2]

$$F(p) = \frac{1}{E(p)} = \frac{1}{(p+\lambda)^n} \tag{5}$$

where λ is the breakpoint frequency.

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

$$F_i(p) = \frac{p^i}{E(p)} = \frac{p^i}{(p+\lambda)^n} \tag{6}$$

and $f_i(t)$ be their corresponding functions in the time domain. By using the filters defined in (6), equation (4) can then be rewritten, in expanded form, as

$$\begin{aligned}&\{F_n(p) + a_1 F_{n-1}(p) + \ldots + a_n F_0(p)\}x(t)\\ &= \{b_0 F_m(p) + \ldots + b_m F_0(p)\}u(t)\end{aligned} \tag{7}$$

In terms of time-domain signals, (7) can be written as

$$\begin{aligned}&x_f^{(n)}(t) + a_1 x_f^{(n-1)}(t) + \ldots + a_n x_f^{(0)}(t)\\ &= b_0 u_f^{(m)}(t) + \ldots + b_m u_f^{(0)}(t)\end{aligned} \tag{8}$$

where

$$\begin{aligned}x_f^{(i)}(t) &= f_i(t) * x(t)\\ u_f^{(i)}(t) &= f_i(t) * u(t).\end{aligned}$$

[2]The filter d.c. gain can be made unity if this is thought desirable.

and $*$ denotes the convolution operator. The filter outputs $x_f^{(i)}(t)$ and $u_f^{(i)}(t)$ provide *prefiltered* time-derivatives of the inputs and outputs in the bandwidth of interest, which may then be exploited for model parameter estimation.

At time-instant $t = t_k$, considering now the situation where there is additive noise on the output measurement, equation (8) can be rewritten in standard linear regression form as

$$y_f^{(n)}(t_k) = \phi_f^T(t_k)\theta + \varepsilon(t_k), \tag{9}$$

with

$$\phi_f^T(t_k) = \left[-y_f^{(n-1)}(t_k) \cdots - y_f^{(0)}(t_k) u_f^{(m)}(t_k) \cdots u_f^{(0)}(t_k)\right] \tag{10}$$

$$\theta = [a_1 \; \ldots \; a_n \; b_0 \; \ldots \; b_m]^T \tag{11}$$

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 parameter estimates are given by

$$\hat{\theta}_N^{LS} = \left[\sum_{i=1}^{N} \phi_f(t_k)\phi_f^T(t_k)\right]^{-1} \sum_{i=1}^{N} \phi_f(t_k) y_f^{(n)}(t_k), \tag{12}$$

Unfortunately, it is well-known that, except in the special case where $\varepsilon(t_k)$ is zero mean and serially uncorrelated (white noise), LS estimation, such as (12), although simple, is unsatisfactory. For instance, even if the additive noise $v(t_k)$ is white, the resultant parameter estimates are asymptotically biased and inconsistent. One of the simplest solutions to this asymptotic bias problem is to use IV methods because they do not require *a priori* knowledge of the noise statistics.

Let us consider the most common IV method, where the instrumental variable is generated by an ‘auxiliary model’ which, as we see later, may be iteratively adapted [18]. In the simplest, non-iterative case, the IV vector is then given by,

$$\hat{\phi}_f^T(t_k) = \left[-\hat{y}_f^{(n-1)}(t_k) \cdots - \hat{y}_f^{(0)}(t_k) u_f^{(m)}(t_k) \cdots u_f^{(0)}(t_k)\right], \tag{13}$$

where

$$\hat{x}_f(t_k) = F(p)\hat{x}(t_k) \tag{14}$$

and $\hat{x}(t_k)$ is the estimated noise-free output calculated from,

$$\hat{x}(t_k) = \frac{B(p, \hat{\theta}_N^{LS})}{A(p, \hat{\theta}_N^{LS})} u(t_k). \tag{15}$$

The IV-based parameter estimates are then given by

$$\hat{\theta}_N^{IV} = \left[\sum_{i=1}^{N} \hat{\phi}_f(t_k)\phi_f^T(t_k)\right]^{-1} \sum_{i=1}^{N} \hat{\phi}_f(t_k) y_f^{(n)}(t_k), \tag{16}$$

provided that the inverse exists. Despite its simplicity, this IV technique is one of six methods that have proven successful in extensive Monte Carlo simulation studies [14].

Comments
1. In the above, it has been assumed that there are no transient initial condition effects on the data used in the estimation. If such effects are present, however, the IV implementation can still yield consistent estimates (the effects are treated as additive noise). However, the quality of the results depends on the severity and longevity of the effects and an alternative is to estimate the initial conditions concurrently with the model parameters (e.g. [1], [14], [19]). This can be advantageous in the case of transient signal data. Of course, treating the initial conditions as an additional set of unknowns in this manner complicates the parameter estimation.
2. Explicit prefiltering strategies are sometimes recommended in practice in DT model identification [6] and are an inherent part of Refined Instrumental Variable (RIV) estimation [4] (see later). Theses strategies improve the statistical efficiency of the parameter estimates. However, the prefiltering strategy is *essential* in direct CT model identification because it has two combined roles. In addition to performing the same prefiltering role as in DT identification, it is also required to reconstruct the time-derivatives within the bandwidth of the system to be identified.
3. The user must provide *a priori* the breakpoint frequency λ of the SVF (5). Intuitively, this can be chosen in order to emphasize the frequency band of interest and generally, it should be chosen equal to, or larger than, the bandwidth of the system to be identified (of course automatic selection of the prefilter characteristics on statistical grounds is much preferable and an algorithm that does this is discussed later in section III-B).
4. The digital implementation of the various CT operations can obviously influence the quality of the estimates (see e.g. [20]). For instance, the intersample nature of the filter input signals is clearly important and a high sampling frequency is often required, therefore, in order to produce accurate CT estimation. Note that if the data are non-uniformly sampled, more advanced variable time-step methods can be used in a straightforward way (see e.g. [21]).

B. Stochastic identification and the iterative IV method

Disregarding the noise properties, as in the deterministic approaches outlined in the previous section, leads to statistical inefficiency (increased variance of the estimates) and does not provide information on the estimated variance-covariance properties of the parameter estimates. The key idea of stochastic identification is to assume that the disturbing noise $v(t)$ can be written, at the sampling instances, as filtered, discrete-time, white noise $v(t_k)$. This avoids the mathematically difficult problem of treating CT random processes. There are only a few direct continuous-time model identification methods which explicitly model the unknown noise transfer function [15], [22], [23].

One particularly successful stochastic identification method is the iterative Simplified Refined Instrumental Variable method for Continuous-time model Identification (SRIVC: see [15], [16], [17]). This approach involves a method of adaptive prefiltering based on a quasi-optimal[3] statistical solution to the problem when the additive noise $v(t_k)$ is white. SRIVC is a logical extension of the more heuristically defined SVF and follows from the optimal RIV and SRIV algorithms for DT identification [4]. This technique presents the advantage of not requiring manual specification of prefilter parameters.

Following the usual Prediction Error Minimization (PEM) approach (Maximum Likelihood (ML) in the present situation because of the Gaussian assumptions), a suitable error function $\varepsilon(t)$ is given by the output error (OE),

$$\varepsilon(t) = y(t) - \frac{B(p)}{A(p)}u(t)$$

Minimization of a least squares criterion function in $\varepsilon(t)$, measured at the sampling instants provides the basis for the *output error* estimation methods. However $\varepsilon(t)$ can also be rewritten as

$$\varepsilon(t) = \frac{1}{A(p)}\big(A(p)y(t) - B(p)u(t)\big)$$

Since the operators commute in this linear case, the filter $F(p) = 1/A(p)$ can be taken inside the brackets to yield

$$\varepsilon(t) = A(p)y_f(t) - B(p)u_f(t) \tag{17}$$

or,

$$\begin{aligned}\varepsilon(t) = y_f^{(n)}(t) + a_{n-1}y_f^{(n-1)}(t) + \ldots + a_0 y_f^{(0)}(t) \\ - b_m u_f^{(m)}(t) - \ldots - b_0 u_f^{(0)}(t)\end{aligned} \tag{18}$$

where

$$\begin{cases} y_f^{(i)}(t) = f_i(t) * y(t), & i = 0, \ldots, n \\ u_f^{(i)}(t) = f_i(t) * u(t), & i = 0, \ldots, m. \end{cases} \tag{19}$$

and the set of filters now takes the form

$$F_i(p) = \frac{p^i}{A(p)} \tag{20}$$

The associated estimation model can be written at time-instant $t = t_k$ in the form:

$$y_f^{(n)}(t_k) = \phi_f^T(t_k)\theta + \varepsilon(t_k) \tag{21}$$

where $\phi_f^T(t_k)$ and θ are defined as in (10) and (11) respectively with $F_i(p)$ defined in (20). Thus, provided we assume that $A(p)$ is known, the estimation model (21) forms a basis for the definition of a likelihood function and ML estimation.

There are two problems with this formulation. The obvious one is, of course, that $A(p)$ is not known *a priori*. The less obvious one is that, in practical applications, we cannot assume that the noise $v(t_k)$ will have the nice white

[3]The method is quasi-optimal because true optimality would require optimal interpolation of the input signal $u(t)$ over the sampling interval, whereas only simple interpolation is used in the SRIVC implementation. However, this normally produces very good, near optimal estimation results.

noise properties assumed above: it is likely that the noise will be a coloured noise process, say $\xi(t_k)$. Both of these problems can be solved by employing a similar approach to that used in the *Refined Instrumental Variable* (RIV) algorithm for DT system identification and estimation (see [4] and the prior references therein). Here, a 'relaxation' optimization procedure is devised that adaptively adjusts an initial estimate $A(p, \hat{\theta}^0)$ of $A(p, \hat{\theta}^j)$ iteratively until it converges on an optimal estimate of $A(p)$. And the coloured noise problem is solved conveniently by exploiting IV estimation within this iterative optimization algorithm.

Of course, if the noise $v(t_k) = \xi(t_k)$ is coloured, then the method is not quasi-optimal in statistical terms. However, experience has shown that it is robust and normally yields estimates with reasonable statistical efficiency (i.e. low but not minimum variance). However, albeit at the cost of increased complexity, it is possible to use a hybrid approach in the coloured noise case, where the noise modelling, as well as the noise-derived parts of the prefiltering, are carried out in discrete-time terms [15], [22], [23].

IV. Software and advantages

A. Software

CONTSID toolbox: The CONtinuous-Time System IDentification (CONTSID) toolbox contains most of the parametric modelling methods developed over the last thirty years which allow one to directly identify CT models of linear time-invariant SISO, MISO and MIMO systems from uniformly and non-uniformly sampled data. It comprises most of the direct deterministic methods, the SRIVC technique outlined above and also output error and subspace-based methods. The toolbox is designed as an add-on to the Mathwork's System IDentification (SID) toolbox and has been given a similar setup. It can be downloaded from: *http://www.cran.uhp-nancy.fr/*

CAPTAIN toolbox: The Computer Aided Program for Time series Analysis and Identification of Noisy systems (CAPTAIN) is a more general toolbox intended not only for the identification of DT and CT transfer function models but also for the extrapolation, interpolation and smoothing of nonstationary and nonlinear time series. The DT and CT identification algorithms are all based on *Refined Instrumental Variable* (RIV) estimation [4]. In particular, CT model identification is provided by the SRIVC algorithm outlined above. The toolbox can be downloaded from: *http://www.es.lancs.ac.uk/cres/captain/*

B. Advantages

The main advantage of the continuous-time methods over the alternative and better known discrete-time methods is that they provide differential equation models whose parameters can be interpreted immediately in physically meaningful terms. As a result, they are of direct use to scientists and engineers who most often derive models in differential equation terms based on natural laws and who are much less familiar with 'black-box' discrete-time models. The direct continuous-time methods can be adapted easily to handle the case of irregularly sampled data. As we shall see, they also offer advantages when applied to systems with widely separated modes and rapidly sampled data.

An extensive analysis aimed at comparing direct and indirect approaches has been discussed recently [14], [24], [25]. This example illustrates some of the well-known difficulties that may appear in DT modelling (sensitivity to the initialization, numerical issues in the case of fast sampling, *a priori* knowledge of the relative degree not easy to accommodate, non inherent data prefiltering), while the direct CT modelling approaches are free from these difficulties.

V. Illustrative simulation and real examples

A. Rainfall-flow modelling

This example concerns the modelling of the daily effective rainfall-flow data from the ephemeral River Canning in Western Australia, as shown in Figure 1. Effective rainfall is a nonlinear transformation of measured rainfall that is a function of the soil–water storage in the catchment and provides a measure of the rainfall that is effective in causing flow variations (rather than that retained by the soil). Further information on the modelling of rainfall-flow processes is given in [26] and the references therein. Another hydrological example is discussed in a recent, related tutorial paper [13] that reinforces the results reported here. The best

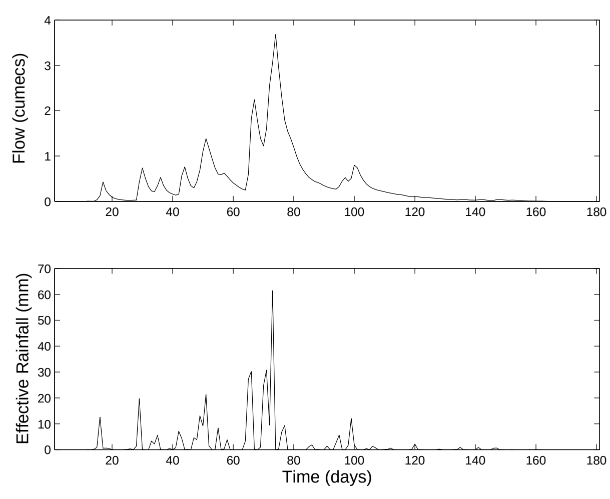

Fig. 1. A typical set of effective rainfall and flow data

identified model based on the data in Figure 1 takes the form:

$$y(t) = \frac{b_o p + b_1}{p^2 + a_1 p + a_2} u(t) + \xi(t) \quad (22)$$

with SRIVC parameter estimates,

$\hat{a}_1 = 0.457(0.032); \hat{a}_2 = 0.025(0.005)$

$\hat{b}_0 = 0.138(0.002); \hat{b}_1 = 0.051(0.002); \hat{b}_2 = 0.0046(0.0008)$

where the figures in parentheses are the estimated standard errors. The coefficient of determination based on the simulated output of this model is $R_T^2 = 0.980$ (i.e. 98% of the flow variance explained by the model output). The model can be interpreted in a physically meaningful manner as a parallel pathway process with an instantaneous pathway (rainfall affecting flow within one day); a 'quick' flow pathway reflecting surface water processes, modelled as a first order process with a time constant (residence time) of 2.5 hours; and a 'slow-flow' pathway, again modelled as a first order process reflecting ground water effects, this time with a much longer time constant of 15.9 hours.

Indirect estimation produced mixed results. Two cases were considered: without (OE) and with noise model parameter estimation. In the first case, the indirectly identified CT model, based on SRIV estimation of the DT model has virtually the same parameter estimates and R_T^2 as the SRIVC estimated CT model. The indirectly identified model based on PEM(OE) estimation is almost as good, with $R_T^2 = 0.979$. However, as expected, the indirectly identified model based on ARX estimation is quite poor, with $R_T^2 = 0.959$. In the second case, the RIV-based estimation with AIC identified AR(4) model is only a little different from that obtained without noise model parameter estimation, with $R_T^2 = 0.979$. However, all reasonable PEM-based estimation results (MA(2), AR(4) and ARMA (2,4) noise models) are worse, with $R_T^2 = 0.959$, $R_T^2 = 0.933$ and $R_T^2 = 0.932$, respectively. More importantly in practical terms, none of these PEM-based models identified the long time constant, so would be rejected on physical grounds. In the case of the MA(2) noise model (the ARMAX model form), the eigenvalues have different signs which cannot be interpreted at all in physically meaningful terms.

Based on this real example, a Monte Carlo Simulation (MCS) study was designed using the SRIVC estimated model and the input effective rainfall data sampled at sampling intervals from 5 minutes to 24 hours [27]. This study was based on 50 stochastic realizations with 20% white noise (by standard deviation) added to the simulated output for each realization. Only 50 realizations were used since the MCS in this case is computationally very intensive, with sample sizes ranging from $52,128$ to 181. For each realization, the identification was designated a failure if the error on the a_1 parameter estimate was greater than three standard deviations from the true value, with the standard deviation based on the SRIVC estimation results.

It is clear from the MCS results that CT identification using the SRIVC algorithm is much more reliable than either of the indirect methods considered. In particular, the direct CT identification has no failures for sampling intervals up to one hour and only 0.32% thereafter. By contrast, RIV-based indirect method has *mean* failure rates at short, medium and long sampling intervals of 7.1%, 2.5% and 1.5%, respectively; while the equivalent figures for the PEM-based indirect method are 8.2% 6.3% and 11.5%. The main reason for the rather poorer performance of the PEM-based indirect approach appears to be because the system is 'stiff', being characterized by widely spaced eigenvalues. This makes the PEM-based gradient optimization algorithm sensitive to the initial estimates and can result in convergence to non-global minima. In other MCS studies using heteroscedastic additive noise, similar to that on the real data, the results are worse than those reported here, with mean failure rates of the PEM-based indirect method of up to 45%[4].

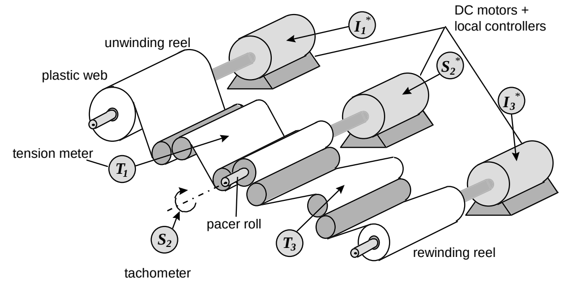

Fig. 2. Winding process

B. Winding process modelling

Winding systems are encountered in a wide variety of industrial plants such as rolling mills in the steel industry, plants involving web conveyance including coating, paper-making and polymer film extrusion processes. The main role of a winding process is to control the web conveyance in order to avoid the effects of friction and sliding, as well as the problems of material distortion and can also damage the quality of the final product. The plant is a prototype of an industrial winding process. A diagram of the process is presented in Figure 2. The main part is composed of a plastic web that is unwound from the first reel, goes over the traction reel and is finally rewound on the rewinding reel. Reel 1 and 3 are coupled with a DC-motor that is controlled with input setpoint currents I_1* and I_3*. The angular speed of reel 2 (S_2) and the tensions in the web between reel 1 and 2 (T_1) and between reel 2 and 3 (T_3) are measured by dynamo tachometers and tension meters. The process is described in more detail in [28]. The SRIVC method has been used to estimate the continuous-time parameters of a multi-input transfer function model with different denominators [29]. Cross-validation results are plotted on Figure 3 where it may be observed that there is a good agreement with high values of R_T^2. This and the previous example demonstrate that the SRIVC algorithm works well on practical and industrial examples.

VI. Conclusion

This paper has provided a tutorial introduction to time-domain methods for directly identifying linear continuous-

[4]All of these identification results were computed in the latest version of Matlab using version 5.0.2 of the SID and CAPTAIN toolboxes. The PEM-based results were quite a lot worse when using version 4 of the SID toolbox.

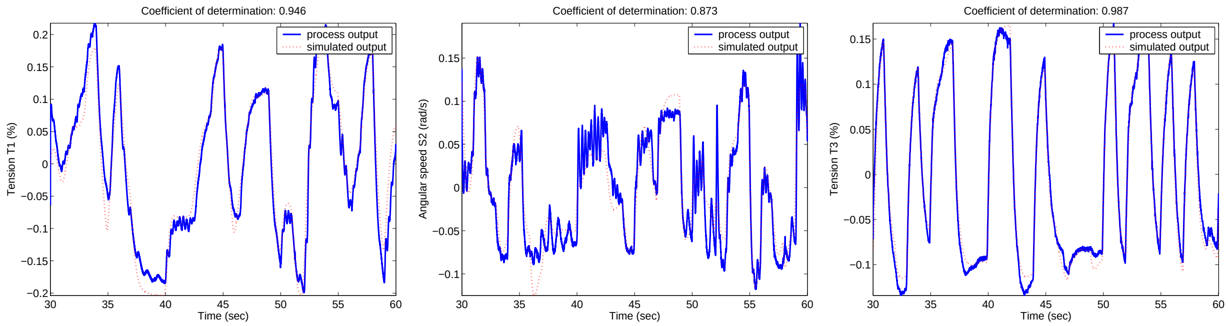

Fig. 3. Cross-validation results for the winding process

time models of stochastic systems from discrete-time sampled data and illustrated the practical utility of these methods. The main advantage of these methods is that they provide differential equation models whose parameters can be interpreted immediately in physically meaningful terms. These methods have proven successful in many practical applications and are available as user-friendly and computationally efficient algorithms in the CONTSID and CAPTAIN toolboxes for Matlab^{TM}.

ACKNOWLEDGMENT

The paper was completed whilst H. Garnier was visiting the Centre for Complex Dynamic Systems and Control (CDSC), University of Newcastle, Australia. He gratefully acknowledges the financial support of CDSC and the Henri Poincaré University of Nancy, France.

REFERENCES

[1] P.C. Young, "In flight dynamic checkout - a discussion," *IEEE Transactions on Aerospace*, vol. 2, pp. 1106–1111, 1964.
[2] P.C. Young, "The determination of the parameters of a dynamic process," *Radio and Electronic Eng. (Jnl. Of IERE)*, vol. 29, pp. 345–361, 1965.
[3] P.C. Young, *Theory of self adaptive control systems*, chapter Process parameter estimation and self adaptive control, pp. 118–140, Plenum Press, P.H. Hammond edition, 1966.
[4] P.C. Young, *Recursive Estimation and Time-Series Analysis*, Springer-Verlag, Berlin, 1984.
[5] T. Söderström and P. Stoica, *System Identification*, Series in Systems and Control Engineering. Prentice Hall, 1989.
[6] L. Ljung, *System identification. Theory for the user*, Prentice Hall, Upper Saddle River, 2nd edition, 1999.
[7] N.R. Kristensen, H. Madsen, and S.B. Jørgensen, "Parameter estimation in stochastic grey-box models," *Automatica*, vol. 40, no. 2, pp. 225–237, 2004.
[8] P. Young, "Parameter estimation for continuous-time models - a survey," *Automatica*, vol. 17, no. 1, pp. 23–39, 1981.
[9] H. Unbehauen and G.P. Rao, *Identification of continuous systems*, Systems and control series. North-Holland, Amsterdam, 1987.
[10] H. Unbehauen and G.P. Rao, "Continuous-time approaches to system identification - a survey," *Automatica*, vol. 26, no. 1, pp. 23–35, 1990.
[11] N.K. Sinha and G.P. Rao (Editors), *Identification of continuous-time systems. Methodology and computer implementation*, Kluwer Academic Publishers, Dordrecht, 1991.
[12] H. Unbehauen and G.P. Rao, "A review of identification in continuous-time systems," *Annual reviews in Control*, vol. 22, pp. 145–171, 1998.
[13] P.C. Young, H. Garnier, and A. Jarvis, "The identification of continuous-time linear and nonlinear models: a tutorial with environmental applications," in *13th IFAC Symposium on System Identification (SYSID'2003)*, Rotterdam (The Netherlands), August 2003, pp. 618–628.
[14] H. Garnier, M. Mensler, and A. Richard, "Continuous-time model identification from sampled data. Implementation issues and performance evaluation," *Int. J. Control*, vol. 76, no. 13, pp. 1337–1357, 2003.
[15] P.C. Young and A.J. Jakeman, "Refined instrumental variable methods of time-series analysis: Part III, extensions," *Int. J. Control*, vol. 31, pp. 741–764, 1980.
[16] P.C. Young, "Optimal IV identification and estimation of continuous-time TF models," in *15th Triennial IFAC World Congress on Automatic Control*, Barcelona (Spain), July 2002.
[17] P.C. Young, "Comments on "On the estimation of continuous-time transfer functions"," *Int. J. Control*, vol. 75, no. 9, pp. 693–697, 2002.
[18] P.C. Young, "An instrumental variable method for real-time identification of a noisy process," *Automatica*, vol. 6, pp. 271–287, 1970.
[19] P.J. Gawthrop, "Parametric identification of transient signals," *IMA Journal of Mathematical Control and Information*, vol. 1, pp. 117–1289, 1984.
[20] C.T. Chou, M. Verhaegen, and R. Johansson, "Continuous-time identification of SISO systems using Laguerre functions," *IEEE Transactions on Signal Processing*, vol. 47, no. 2, pp. 349–362, 1999.
[21] E. Huselstein and H. Garnier, "An approach to continuous-time model identification from non-uniformly sampled data," in *41st IEEE Conference on Decision and Control (CDC'02)*, Las-Vegas, Nevada (USA), December 2002.
[22] R. Johansson, "Identification of continuous-time model," *IEEE Transactions on Signal Processing*, vol. 42, no. 4, pp. 887–896, 1994.
[23] R. Pintelon, J. Schoukens, and Y. Rolain, "Box-Jenkins continuous-time modeling," *Automatica*, vol. 36, pp. 983–991, 2000.
[24] G.P. Rao and H. Garnier, "Numerical illustrations of the relevance of direct continuous-time model identification," in *15th Triennial IFAC World Congress on Automatic Control*, Barcelona (Spain), July 2002.
[25] L. Ljung, "Initialisation aspects for subspace and output-error identification methods," in *European Control Conference (ECC'2003)*, Cambridge (U.K.), September 2003.
[26] P.C. Young, "Advances in real-time flood forecasting"," *Philosophical Trans. Royal Society, Physical and Engineering Sciences*, vol. 360, no. 9, pp. 1433–1450, 2002.
[27] P.C. Young, "Identification and estimation of continuous-time hydrological models from discrete-time data," in *International Conference on Hydrology: Science and Practice for the 21st century*, London (U.K.), July 2004.
[28] 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*, vol. 74, no. 2, pp. 118–132, 2001.
[29] E. Huselstein, M. Gilson, H. Garnier, and A. Richard, "Instrumental variable algorithms for multiple input systems described by multiple continuous-time TF models," in *Asian Control Conference*, Melbourne (Australia), July 2004.