



HAL
open science

Data-based continuous-time modelling of environmental systems

Hugues Garnier, Peter C. Young, Stéphane Thil, Marion Gilson

► **To cite this version:**

Hugues Garnier, Peter C. Young, Stéphane Thil, Marion Gilson. Data-based continuous-time modelling of environmental systems. 5ème conférence STIC & Environnement, Nov 2007, Lyon, France. pp.CDROM. hal-00167182

HAL Id: hal-00167182

<https://hal.science/hal-00167182>

Submitted on 26 Nov 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Data-based Continuous-time Modelling of Environmental Systems

H. GARNIER¹, P.C. YOUNG², S. THIL¹ et M. GILSON¹

- 1 Centre de Recherche en Automatique de Nancy, Nancy-Université, CNRS, BP 239, F-54506 Vandœuvre-lès-Nancy Cedex, France
{hugues.garnier, stephane.thil, marion.gilson}@cran.uhp-nancy.fr
- 2 Centre for Research on Environmental Systems and Statistics, Lancaster University, Lancaster LA1 4YQ, U.K. and Fenner School of Environment and Society, Australian National University, Canberra, Australia, p.young@lancaster.ac.uk

Abstract

In this paper, we argue that many environmental processes are more naturally defined in the context of continuous-time differential equation models, normally derived by the application of natural laws, such as mass and energy conservation. As a result, there are advantages if such models are estimated directly in this continuous-time differential equation form, rather than being formulated and estimated as discrete-time models. This paper discusses these advantages and briefly describes the key features of the two toolboxes for use with MATLAB[®] which support the advanced time-domain methods for directly identifying linear continuous-time models from discrete-time data. The arguments presented in the paper are finally illustrated by a practical example in which the proposed data-based approach to direct CT modelling is applied to a set of rainfall-flow data.

Keywords

continuous-time systems; environmental systems; linear differential equation; rainfall-flow modelling, sampled data, system identification.

INTRODUCTION

Most mathematical models of environmental systems are formulated on the basis of natural laws, such as dynamic conservation equations, often expressed in terms of continuous-time (CT), linear or nonlinear differential equations. Paradoxically, Transfer Function (TF) models, which have been growing in popularity over the last decade because of their ability to characterise environmental data in an efficient parametric (parsimonious) form, are almost always presented in the alternative, discrete-time (DT) terms. One reason for this paradox is that environmental data are normally sampled at regular intervals over time, so forming discrete-time series that are in a most appropriate form for DT modelling (Young 1984, Ljung 1999). Another is that most of the technical literature on the statistical identification and estimation (calibration) of TF models deals with these DT models. Closer review of the literature, however, reveals apparently less well-known publications dealing with estimation methods that allow for the direct identification of CT (differential equation) models from discrete-time, sampled data (Young 1981, Garnier *et al.* 2003, H. Garnier and L. Wang 2008).

This paper first discusses the formulation, identification and estimation¹ of linear CT models. It then briefly describes the main features of the user-friendly CAPTAIN and CONTSID toolboxes for MATLAB[®] which support the computationally efficient CT model identification algorithms. In addition, we present a practical example, where the proposed data-based approach to direct CT modelling is applied to a typical set of effective rainfall-flow data. Our main aim is to demonstrate the many advantages of direct continuous-time model estimation, in relation to its discrete-time alternative, and so encourage its practical application.

CONTINUOUS-TIME MODELS

For simplicity of presentation, the theoretical basis for the statistical identification and estimation of linear, continuous-time models from discrete-time, sampled data can be outlined by considering the case of a linear, single input, single output system. It should be noted, however, that the analysis extends straightforwardly to multiple input systems and, in a more complex manner, to full multivariable systems

$$x(t) = \frac{B(s)}{A(s)}u(t - \tau) \quad (1)$$

Here $A(s)$ and $B(s)$ are the following polynomials in the derivative operator $s = d/dt$

$$A(s) = s^n + a_1s^{n-1} + \dots + a_{n-1}s + a_n \quad (2)$$

$$B(s) = b_0s^m + b_1s^{m-1} + \dots + b_{m-1}s + b_m \quad (3)$$

in which n and m can take on any positive integer values with $m \leq n$; τ is any pure time delay in time units; $u(t)$ and $x(t)$ denote, respectively, the deterministic input and output signals of the system. Of course, the model (1) can also be written in the following differential equation form, which is often more familiar to physical scientists

$$\frac{d^n x(t)}{dt^n} + a_1 \frac{d^{n-1} x(t)}{dt^{n-1}} + \dots + a_n x(t) = b_0 \frac{d^m u(t - \tau)}{dt^m} + \dots + b_m u(t - \tau) \quad (4)$$

The structure of the CT model is denoted by the triad $[n \ m \ \tau]$. In most practical situations, the observed input and output signals $u(t)$ and $y(t)$ will be sampled in discrete-time. In the case of uniform sampling, at a constant sampling interval T_s , these sampled signals will be denoted by $u(t_k)$ and $y(t_k)$ and the output observation equation then takes the form,

$$y(t_k) = x(t_k) + \xi(t_k) \quad k = 1, \dots, N \quad (5)$$

where $x(t_k)$ is the sampled value of the unobserved, noise-free output $x(t)$. Given the discrete-time, sampled nature of the data, the measured output is assumed to be contaminated by a noise or stochastic, discrete-time disturbance signal $\xi(t_k)$. This noise is assumed to be independent of the input signal $u(t_k)$ and it represents the aggregate effect of all the stochastic inputs to the system, including distributed unmeasured inputs, measurement errors and modelling error. If $\xi(t_k)$ has rational spectral density, then it can be modelled as an Auto-Regressive (AR) or Auto-Regressive, Moving-Average (ARMA) process but this restriction is not essential.

¹The statistical meaning of these terms will be used here, where ‘identification’ is taken to mean the specification of an identifiable model structure and ‘estimation’ relates to the estimation of the parameters that characterize this identified model structure.

The objective is then to identify a suitable model structure defined by the triad $[n \ m \ \tau]$ for (1) and estimate the parameters that characterize this structure, based on the sampled input and output data $Z^N = \{u(t_k); y(t_k)\}_{k=1}^N$.

IDENTIFICATION METHODS FOR DATA-BASED CT MODELLING

The problem of identifying continuous-time models, based on discrete-time sampled data can be approached in two main ways:

- The *direct* approach: here, it is necessary to identify the most appropriate CT model structure ; and then directly estimate the TF parameters a_i, b_i , and τ that characterise this structure. Of course, some approximation will be incurred in this estimation procedure because the inter-sample behaviour of input signal $u(t)$ is not known and it must be interpolated over the sampling interval in some manner.
- The *indirect* approach: here, a DT model for the original CT system is first obtained by applying a conventional stochastic DT model identification and estimation procedure. This estimated DT model is then converted to a CT model, again using some assumption about the nature of the input signal $u(t)$ over the sampling interval T_s .

Various statistical methods of identification and estimation have been proposed to implement the two approaches outlined above and these have been formulated in both the time and frequency domains. However, only estimation in the time domain will be considered here, and only one direct, statistically efficient estimation method will be briefly discussed (and later used in the practical example section): the Refined Instrumental Variable method for Continuous-time Systems (RIVC) (Young *et al.* 2008). The RIVC method and its simplified version SRIVC (when the additive measurement noise is assumed to be white) are the only time-domain methods that can be interpreted in optimal statistical terms, so providing an estimate of the parametric error covariance matrix and, therefore, estimates of the confidence bounds on the parameter estimates. The SRIVC/RIVC algorithms are available in both the CAPTAIN and CONTSID toolboxes which are briefly described in the next section.

This paper also promotes the *Data-Based Mechanistic* (DBM) modelling philosophy (see (Young 1998) and the prior references therein) based on the direct identification and estimation of a CT model. This involves the application of an advanced method of direct continuous-time TF model estimation and identification and the interpretation of this estimated model in physically meaningful terms. This DBM modelling approach can be contrasted with ‘*black-box*’ modelling, since DBM models are only deemed credible if, in addition to explaining the time series data in a statistically efficient, parsimonious manner, they also provide an acceptable physical interpretation of the system under study. They can also be contrasted with ‘*grey-box*’ models, because the model structure is inferred inductively from the data, rather than being assumed *a priori* before model identification and estimation in a hypothetico-deductive manner (see the discussion in (Young 2002)).

SOFTWARE SUPPORT FOR DATA-BASED CT MODELLING

The CONTSID toolbox

The CONtinuous-Time System IDentification (CONTSID) toolbox² for MATLAB[®] supports continuous-time transfer function or state-space model identification directly from regularly or irregularly time-domain sampled data, without requiring the determination of a discrete-time model. The motivation for developing the CONTSID toolbox was first to fill in a gap, since no software support was available to both serve the cause of direct time-domain identification of continuous-time linear models and also provide the potential user with a platform for testing and evaluating all of the available data-based CT modelling techniques. In this latter regard, it includes most of the direct deterministic methods, the stochastic SRIVC/RIVC techniques, as well as 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 (Garnier *et al.* 2008).

The CAPTAIN toolbox

The Computer Aided Program for Time series Analysis and Identification of Noisy systems (CAPTAIN)³ for MATLAB[®] is a more general toolbox intended not only for the identification of CT *and* DT transfer function models but also for DBM modelling, including the extrapolation, interpolation and smoothing of nonstationary and nonlinear time series (see e.g. Taylor *et al.*, 2007). The CT (and DT) identification algorithms are all based on *Refined Instrumental Variable* (RIV) estimation (Young 1984). In particular, CT model identification is provided by the SRIVC/RIVC algorithms mentioned above.

ADVANTAGES OF DIRECT DATA-BASED CT MODELLING

As noted in the previous section, there are two fundamentally different time-domain approaches to the problem of estimating a CT model of a continuous-time system from its sampled input-output data. The indirect approach has the advantage that it uses well-established DT model identification methods (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.

On the surface, the choice between the two approaches may seem trivial. However, some recent studies have exposed some serious shortcomings of the indirect route through DT models. Indeed, an extensive analysis aimed at comparing direct and indirect approaches has been presented recently. The simulation model used in this analysis provides a very good test for CT and DT model estimation methods: it was first suggested by Rao and Garnier (Rao and Garnier 2002) (see also (Garnier *et al.* 2003, Rao and Garnier 2004)); and further, independent investigations confirmed the results (Ljung 2003). This simulation example illustrates some of the difficulties that may appear in DT modelling under less standard conditions, such as rapidly sampled data or relatively wide-band systems:

- relatively high sensitivity to the initialization. DT model identification often requires computationally costly minimization algorithms without even guaranteeing

²<http://www.cran.uhp-nancy.fr/contsid>.

³<http://www.es.lancs.ac.uk/cres/captain>.

convergence (to the global optimum). In fact, in many cases, the initialization procedure for the identification scheme can be a key factor in obtaining satisfactory estimation results when compared to the direct methods discussed in the present paper, which do not suffer from this problem;

- numerical issues in the case of fast sampling because the eigenvalues lie close to the unit circle in the complex domain, so that the model parameters are more poorly defined in statistical terms;
- *a priori* knowledge of the relative degree is not easy to accommodate;
- there is no inherent data prefiltering.

Further, the question of parameter transformation between a DT description and a CT representation is non-trivial. First, the zeros of the DT model are not as easily translatable to CT equivalents as the poles. Second, due to the discrete nature of the measurements, they do not contain all the information about the CT signals. 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 between the sampling intervals (the zero-order hold assumption). Violation of these assumptions may lead to severe estimation errors.

The advantages of direct CT identification approaches over the alternative DT identification methods can be summarized as follows:

- they provide directly 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;
- they can estimate fractional time-delay systems;
- the estimated model is defined by a unique set of parameter values that are not dependent on the sampling interval T_s ;
- there is no need for conversion from discrete to continuous-time, as required in the indirect route based on initial DT model estimation;
- the direct continuous-time methods can handle easily the case of irregularly sampled data;
- Any *a priori* knowledge of the relative degree is easy to accommodate and, therefore, allows for the identification of more parsimonious models than in discrete-time;
- they offer distinct advantages when applied to systems with widely separated modes;
- the methods promoted in this paper for CT model estimation include inherent data pre-filtering;
- they are well suited to situations of over-sampling. This is particularly interesting since today’s data acquisition equipment can provide nearly continuous-time measurements and, therefore, make it possible to acquire data at very high sampling frequencies. Note that as mentioned in Ljung (2003), the use of prefiltering and

decimation strategies may lead to better results in the case of DT modelling but these are not so obvious or straightforward for a practitioner; while the direct CT methods are free of these difficulties (indeed, the estimation is superior in just these conditions)

Given these advantages, it is clear that, if a CT model is required, it is much better for the practitioner to utilise a direct approach to model identification and estimation.

PRACTICAL EXAMPLE: 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 (Young 2002) and the references therein. Another hydrological example is discussed in a recent, related tutorial paper (Young and Garnier 2006) that reinforces the results reported here. The best SRIVC identified TF model,

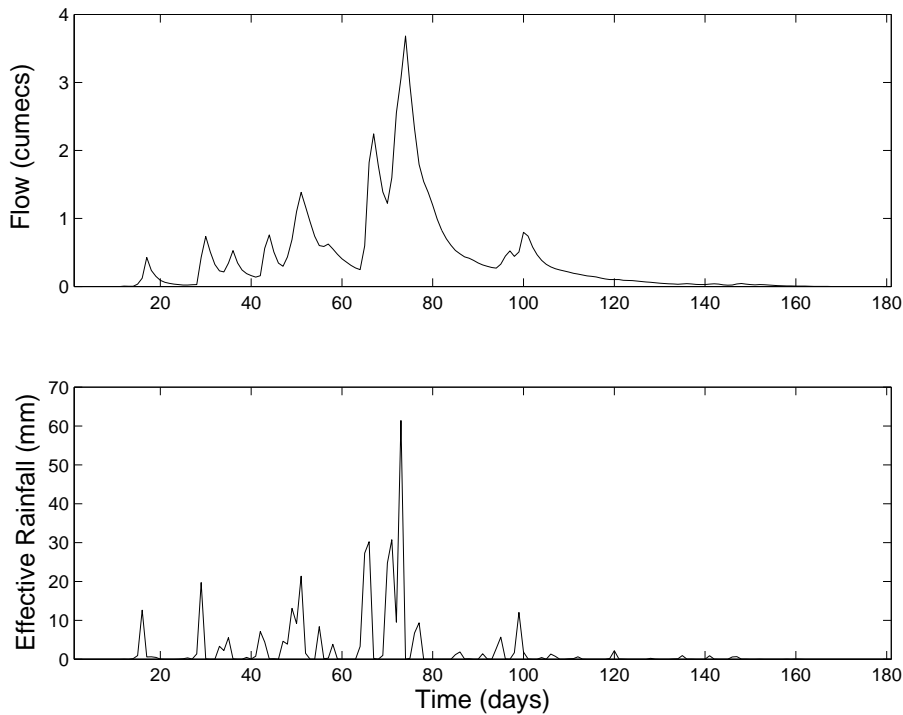


Figure 1: A typical set of effective rainfall and flow data

based on the data in Figure 1, takes the form

$$y(t) = \frac{b_0 s^2 + b_1 s + b_2}{s^2 + a_1 s + a_2} u(t) + \xi(t) \quad (6)$$

with SRIVC parameter estimates,

$$\begin{aligned} \hat{a}_1 &= 0.457(0.032); \hat{a}_2 = 0.025(0.005) \\ \hat{b}_0 &= 0.0138(0.002); \hat{b}_1 = 0.051(0.002); \hat{b}_2 = 0.0046(0.0008) \end{aligned}$$

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 is explained by the model output). Following the DBM modelling approach, the model (6) 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 days; 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 days (see the above references for a more detailed explanation).

Indirect estimation produced mixed results. Two cases were considered: without (OE) and with noise (PEM) 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 identify the long time constant and so would be rejected on physical grounds. Indeed, for the MA(2) noise model (the most common ARMAX model form), the eigenvalues have different signs which cannot be interpreted at all in physically meaningful terms.

CONCLUSIONS

This paper has outlined an approach to Data-Based Mechanistic (DBM) modelling of environmental systems based on methods for the direct identification and estimation of continuous-time (transfer function or differential equation) models. It has also argued that such models are more appropriate to the modelling of environmental systems than their more widely known discrete-time equivalents. The main advantage of these methods, over the alternative discrete-time equivalents, is that they provide differential equation models whose parameters can be interpreted immediately in physically meaningful terms, as our example has shown. Consequently, 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. Moreover, the continuous-time methods can be adapted easily to handle the case of irregularly sampled data or non-integral time delays that are often encountered in the modelling of environmental systems. They are also much superior when applied to rapidly sampled data, where discrete-time methods often perform poorly because the eigenvalues lie close to the unit circle in the complex domain, leading to estimation problems. These advantages, together with the parsimony that is a natural consequence of DBM CT transfer function modelling, should mean that any relationships between the CT model parameters and physical measures of the environmental process should be clearer and better defined statistically.

This data-based DBM CT modelling approach has proven successful in many practical applications and the most powerful statistical identification and estimation methods are available in the user-friendly CAPTAIN and CONTSID toolboxes.

References

- Garnier, H., M. Gilson, T. Bastogne and M. Mensler (2008). The CONTSID toolbox for Matlab: a software support for time-domain continuous-time model identification from sampled data. In *Identification of continuous-time models from sampled data*, H. Garnier and L. Wang (Eds.). Springer-Verlag (in press).
- Garnier, H., M. Mensler and A. Richard (2003). Continuous-time model identification from sampled data. Implementation issues and performance evaluation. *International Journal of Control* **76**(13), 1337–1357.
- H. Garnier and L. Wang (Eds.), (2008). *Identification of continuous-time models from sampled data*. Springer-Verlag (in press).
- Ljung, L. (1999). *System identification. Theory for the user*. 2nd ed. Prentice Hall. Upper Saddle River.
- Ljung, L. (2003). Initialisation aspects for subspace and output-error identification methods. *European Control Conference (ECC'2003)*. Cambridge (U.K.).
- Rao, G.P. and H. Garnier (2002). Numerical illustrations of the relevance of direct continuous-time model identification. *15th Triennial IFAC World Congress*. Barcelona (Spain).
- Rao, G.P. and H. Garnier (2004). Identification of continuous-time systems: direct or indirect? *Systems Science* **30**(3), 25–50.
- Taylor, C.J., D.J. Pedregal, P.C. Young and W. Tych (2007). Environmental time series analysis and forecasting with the CAPTAIN toolbox. *Environmental Modelling and Software* **22**(6), 797–814.
- Young, P. (1981). Parameter estimation for continuous-time models - a survey. *Automatica* **17**(1), 23–39.
- Young, P.C. (1984). *Recursive estimation and time-series analysis*. Springer-Verlag. Berlin.
- Young, P.C. (1998). Data-based mechanistic modeling of environmental, ecological, economic and engineering systems. *Journal of Modelling & Software* **13**, 105–122.
- Young, P.C. (2002). Advances in real-time flood forecasting. *Philosophical Trans. Royal Society, Physical and Engineering Sciences* **360**(9), 1433–1450.
- Young, P.C. and H. Garnier (2006). Identification and estimation of continuous-time data-based mechanistic (DBM) models for environmental systems. *Environmental Modelling and Software* **21**(8), 1055–1072.
- Young, P.C., H. Garnier and M. Gilson (2008). Refined instrumental variable identification of continuous-time hybrid Box-Jenkins models. In *Identification of continuous-time models from sampled data*, H. Garnier and L. Wang (Eds.), Springer-Verlag (in press).