Monograph:
Coca, D. and Billings, S.A. (2001) Identification of Finite Dimensional Models for Distributed-Parameter Systems. Research Report. ACSE Research Report 802 . Department of Automatic Control and Systems Engineering

Identification of Finite Dimensional Models for Distributed-Parameter Systems

D. Coca

S.A. Billings

Department of Automatic Control and Systems Engineering,
University of Sheffield
Sheffield, S1 3JD,
UK

Research Report No. 802
September 2001

IDENTIFICATION OF FINITE-DIMENSIONAL MODELS FOR DISTRIBUTED-PARAMETER SYSTEMS

Daniel Coca, Stephen A. Billings

Department of Automatic Control & Systems Engineering, University of Sheffield, Mappin Street, Sheffield, S1 3JD, UK

Abstract:
This paper addresses the problem of obtaining finite dimensional models of distributed-parameter systems from pointwise data using system identification. The data are first interpolated into a finite dimensional space and expanded in terms of a finite element basis. A discrete time model is estimated based on the resulting finite dimensional coordinate vector. The existence and convergence of such a representation is established for a class of abstract first order systems. The proposed approach is illustrated in practice using simulated noise contaminated data.

Keywords: Distributed-Parameter Systems, Finite Element Solutions, Identification Algorithms, Interpolation Approximation, System Identification.

1. INTRODUCTION

In most practical cases, the analysis, simulation and control of a distributed parameter system, which is described by partial differential equations (PDE's) and is characterised by an infinite dimensional state-space, cannot be solved using only analytical methods (Bensoussan *et al.*, 1992). The solution in these cases is to replace the original infinite dimensional PDE description with a finite dimensional approximation which captures, with sufficient accuracy, the properties of the original PDE.

Different techniques can be applied to transform the original PDE into an approximate system of ordinary differential or difference equations. The most commonly used approaches are the finite difference and the finite element methods. All methods require knowledge of the form and parameters of the PDE's describing the distributed parameter system.

This paper addresses this problem from a system identification perspective. The idea is to obtain the finite dimensional approximate model from pointwise observations without assuming any a priori knowledge of the structure or the parameters of the PDE's. The proposed approach involves two basic steps involving the finite element approximation of the variables in the spatial domain and the identification of the finite dimensional model from the time-dependent coordinate vector respectively. In the first stage the pointwise observations are interpolated onto a finite dimensional space and expanded in terms of a suitably conditioned finite element basis which accounts for the boundary conditions that are assumed known. If the number of degrees of freedom is too large the dimension of the coordinate vector can be reduced by projecting the initial interpolation onto a lower dimensional subspace subject to maintaining a certain degree of accuracy. System identification techniques are employed in the second stage to estimate a discrete-time model based on the resulting coordinate vector.

The theoretical aspects of the proposed identification approach are also investigated. In particular the existence, stability and convergence of the finite dimensional models are established for a class of first order systems. The proposed approach is illustrated using simulated noise contaminated data.

2. THE EVOLUTION EQUATION

Let H be a separable real Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and norm $|\cdot|$ and V another separable Hilbert space which is embedded continuously and densely in H. Here H is identified with its own dual space. Let V^* denote the dual space of V and $\|\cdot\|_*$ denote the norm on V^*. It follows that $V \subset H \subset V^*$ with continuous and dense embeddings. Specifically the following inequality is assumed to hold

$$|\varphi| \leq \lambda^{-1/2}\|\varphi\| \tag{1}$$

The notation $\langle\cdot,\cdot\rangle$ will also be used to denote the duality pairing between V and V^* where the pairing between $\varphi \in H$ and $\psi \in V$ agrees with the inner product $\langle\varphi,\psi\rangle$. It follows that $\|\varphi\|_* \leq \lambda|\varphi|$ and $\|\psi\|_* \leq \lambda^2\|\psi\|$. Often in practice it will be assumed that $H = L^2(\Omega)$, V is the Sobolev space $H^l(\Omega)$ with dual $V^* = H^{-l}(\Omega)$.

Consider the following evolution equation

$$\frac{du}{dt} + Au = v(t) \tag{2}$$

$$u(0) = u_0 \tag{3}$$

and the equivalent variational formulation

$$\langle\frac{du}{dt},\varphi\rangle + \langle Au,\varphi\rangle = \langle v(t),\varphi\rangle, \tag{4}$$

$$u(0) = u_0, \forall\varphi \in V \tag{5}$$

where it is assumed that:

(A1) $A \in \mathcal{L}(V,V^*)$.
(A2) The operator A is coercive that is $\langle A\varphi,\varphi\rangle \geq \alpha\|\varphi\|^2$, $\forall\varphi \in V$ for some $\alpha > 0$.
(A3) The forcing function $v(t,x) \in C(\mathbb{R}_+;H) \cap L^2(0,T;H)$ is bounded in $L^\infty(\mathbb{R}_+;H)$.

A solution of the initial value problem (2), (3) is a function $u \in L^2(0,T;V)$ with $D_t u \in L^2(0,T;V^*)$ that satisfies (2) and (3) for all $T > 0$. Specifically, here it is assumed that

(A4) $u(t,x) \in C(\mathbb{R}_+,H) \cap L^2(\mathbb{R}_+;V)$ is a unique solution of (2), (3).

The equation (2) is usually complemented by boundary conditions which can be of the Dirichlet, Neumann or periodicity type for example. These can be accomodated by considering restrictions of A and v to corresponding closed subspaces V

3. THE IDENTIFICATION METHOD

In general, the numerical integration of evolution equations is based on a finite dimensional approximation of the original infinite dimensional system.

The idea is to reduce the infinite dimensional system to a system of ordinary differential or difference equations which can be used either to compute an approximate solution or to design the controller. One well known approach is the finite difference method which involves approximating the partial differential operator by finite differences. An alternative approach is the finite element Galerkin method (FEM) (Brenner and Ridgway Scott, 1994). Unlike the finite difference method the finite element method is essentially an approximation of the space V, where the solution of the partial differential equation is sought, using finite dimensional subspaces.

The problem addressed in this paper is that of estimating from pointwise data, a finite dimensional, discrete-time dynamical system which approximates with sufficient accuracy the unknown infinite dimensional system. This approach assumes no a priori knowledge of the PDE's which governs the distributed parameter system. The identification method can be viewed as an *inverse* finite element Galerkin approach where the solution is used to derive the finite dimensional model rather than the original PDE's.

The identification is performed in two stages. In the first stage the data is interpolated onto a finite element subspace V^n. This involves computing the input and output coordinate vectors relative to the finite element basis. The second stage involves estimating a finite dimensional, discrete-time model which approximates this input/output behaviour.

3.1 *Approximation Results*

Let $V^n \subset V$ with $n = 1, 2, \ldots$ be a sequence of finite dimensional subspaces of H which are dense in V and are spanned by a finite dimensional basis $\{\varphi_j^n\}_{j=0}^n$. Moreover, there exists a constant $C_1 > 0$ independent of n such that for any $f(x) = \sum_{j=0}^n c_j\varphi_j^n(x)$ in V^n

$$\sum_{j=0}^{n} |c_j|^2 \leq C_1|f|^2 \tag{6}$$

For example, this condition is satisfied in the case of the B-spline basis. Let

$$y_n(t,x) = \sum_{j=0}^{n} y_{n,j}(t)\varphi_j^n(x), \qquad t > 0 \tag{7}$$

$$v_n(t,x) = \sum_{j=0}^{n} v_{n,j}(t)\varphi_j^n(x) \qquad t>0 \quad (8)$$

denote the interpolation of u and v respectively in V^n.

The following theorem establishes the existence, stability and convergence of a finite dimensional dynamical system which provides an approximate realisation of the input/output behaviour $\mathbf{v}_n(t) = (v_{n,0}(t), ..., v_{n,n}(t))$, $\mathbf{y}_n(t) = (y_{n,0}(t), ..., y_{n,n}(t))$, $t>0$.

Theorem 3.1. Assuming (A1)-(A4) to hold, let $u(t,x)$ be the unique solution of (2) with initial conditions (3) and forcing function $v(t,x)$. Let $\mathbf{v}_n(t)$, $\mathbf{y}_n(t)$ be the coordinate vectors of $v_n(t,x)$ and $y_n(t,x)$ defined in (8) and (7) respectively. An n+1-dimensional dynamical system exists such that if $\mathbf{u}_n(t) = (u_{n,0}(t), ..., u_{n,n}(t))$ is the trajectory of the system with input $\mathbf{v}_n(t)$ and initial conditions $\mathbf{u}_n(0) = \mathbf{y}_n(0)$ and $u_n(t,x) = \sum_{j=0}^{n} u_{n,j}(t)\varphi_j^n(x)$ then:

a) $u_n(t,x)$ is bounded in $L^\infty(\mathbb{R}_+;H)$ and $u_n(t,x) \to y_n(t,x)$ strongly in $L^2(0,T,H)$ and $L^2(0,T,V)$ as $n\to\infty$.
b) The trajectory $\mathbf{u}_n(t)$ is bounded in $L^\infty(\mathbb{R}_+;\mathbb{R}^{n+1})$ and $\mathbf{u}_n(t) \to \mathbf{y}_n(t)$ in $L^2(0,T;\mathbb{R}^{n+1})$ as $n\to\infty$.

Proof:

For each $n = 0,1,2,...$ define the operator $A_n : V^n \to V^n$ by

$$\langle Au_n, \varphi^n\rangle = \langle A_n u_n, \varphi^n\rangle, \qquad \varphi^n \in V^n \quad (9)$$

for any $u_n \in V^n$. From the Riesz Representation Theorem (Naylor and Sell, 1989) applied to the Hilbert space V^n it follows that A_n is a well defined operator given that $\langle Au_n, \cdot\rangle$ is a bounded linear functional.

Consider the initial value problem in V^n

$$\frac{du_n}{dt} + A_n u_n = v_n(t), \qquad (10)$$

$$u_n(0) = y_n(0) \qquad (11)$$

which is an ordinary differential equation. For each $n \geq 1$ the existence of solutions on some interval $(0,T_n)$ follows from standard theorems for ordinary differential equations. The *a priori* estimates below show that these solutions are defined for all $t>0$ (i.e $T_n = +\infty$). From (10) and (9) it follows that for any $\varphi^n \in V^n$, u_n is the solution in V^n of

$$\langle \frac{du_n}{dt}, \varphi^n\rangle + \langle Au_n, \varphi^n\rangle = \langle v_n, \varphi^n\rangle \qquad (12)$$

$$u_n(0) = y_n(0) \qquad (13)$$

For $\varphi^n = u_n$ it follows that

$$\frac{1}{2}\frac{d}{dt}|u_n|^2 + \langle Au_n, u_n\rangle = \langle v_n, u_n\rangle \qquad (14)$$

Since A is coercive

$$\frac{1}{2}\frac{d}{dt}|u_n|^2 + \alpha||u_n||^2 \leq |v_n||u_n| \qquad (15)$$

and subsequently, using the well known inequality,

$$ab \leq \frac{\epsilon}{2}a^2 + \frac{b^2}{2\epsilon} \qquad (16)$$

and (1) it follows that

$$\frac{d}{dt}|u_n|^2 + \alpha||u_n||^2 \leq \frac{1}{\lambda\alpha}|v_n|^2 \qquad (17)$$

and subsequently, after using (1) again that

$$\frac{d}{dt}|u_n|^2 + \alpha\lambda|u_n|^2 \leq \frac{1}{\lambda\alpha}|v_n|^2 \qquad (18)$$

Integrating (18) and using the classical Gronwall lemma yields

$$|u_n|^2 \leq |u_n(0)|^2 e^{-\alpha\lambda t} + \frac{|v_n|_s^2}{(\alpha\lambda)^2}(1-e^{-\alpha\lambda t}) \quad (19)$$

where $|v_n|_s^2 = \sup\{|v_n|^2\}_{t\in[0,\infty)}$ which is finite since $v(t)$ is bounded in $L^\infty(\mathbb{R}_+;H)$. Therefore $T_n \to +\infty$ as announced earlier i.e. the solution u_n is defined for all $t>0$. It remains to prove that u_n converges to the interpolate y_n as $n\to\infty$. However, it is well known that the interpolate $y_n \to u$ in $L^2(0,T,V)$ and $L^2(0,T,H)$ strongly as $n\to\infty$. From the triangle inequality it follows that it is sufficient to show that $u_n \to u$ strongly in in $L^2(0,T,V)$ as $n\to\infty$.

Equation (19) implies that u_n remains in a bounded set of $L^\infty(\mathbb{R}_+;H)$ as $n\to\infty$. Going back to (17) it also follows that $||u_n||$ is uniformly bounded for any $t>0$ so that for any $T>0$ u_n remains in a bounded set of $L^2(0,T;V)$ as $n\to\infty$.

These estimates ensure the existence of an element u' and a subsequence $n'\to\infty$ such that for all $T>0$, $u_{n'} \to u'$ weakly in $L^2(0,T,V)$ $du_{n'}/dt \to du/dt$ weakly in $L^2(0,T,V^*)$ and $u_{n'} \to u'$ weak-star in $L^\infty(\mathbb{R}^+,H)$, as $n'\to\infty$. Owing to a classical compactness theorem (Temam, 1983) it follows that $u_{n'} \to u'$ strongly in $L^2(0,T;H)$ for all $T>0$ as $n'\to\infty$. By passing to the limit in (12) it follows that $u' = u$ and the whole sequence converges to u. The strong convergence result in $L^2(0,T;V)$ follows easily by showing that the expression

$$X_n = \frac{1}{2}|u_n(T) - u(T)|^2 + \int_0^T ||u_n - u||^2 dt \quad (20)$$

tends to zero as $n \to \infty$.

If we expand u_n in equation (10) in terms of the finite element basis in V^n and take the inner product with φ_j^n for $j = 0, ..., n$, this leads to the following system of differential equations

$$M^n \frac{d\mathbf{u}_n}{dt} + E^n \mathbf{u}_n = M^n \mathbf{v}_n(t) \qquad (21)$$

where $M_{i,j}^n = \langle \varphi_i^n, \varphi_j^n \rangle$ and $E_{i,j}^n = \langle A\varphi_i^n, \varphi_j^n \rangle$.

The second part of the theorem follows easily since according to (6)

$$\sum_{j=0}^{n} |u_{n,j}(t) - y_{n,j}(t)|^2 \leq C_1 |u_n - y_n|^2 \quad (22)$$

which after integrating with respect to t

$$\int_0^T \sum_{j=0}^{n} |u_{n,j}(t) - y_{n,j}(t)|^2 \leq C_1 \int_0^T |u_n - y_n|^2 (23)$$

leads to the convergence result in $L^2(0, T; R^{n+1})$.

3.2 *The Identification Problem*

Consider the evolution equation (2) with Dirichlet boundary conditions satisfing assumptions (A1)-(A3) and $u(t, x)$ a solution satisfying assumption (A4). Without loss of generality V is identified with $H^l(\Omega)$, the Sobolev space of order $l \geq 2$.

To account for the fact that in general it is not possible to measure the full state of the system, the following observation operator is introduced

$$\mathcal{Z} : C([0, T], C(\Omega)) \to \mathcal{Y} \qquad (24)$$

where $\mathcal{Y}$ is the *observation space* to which the measurements $y = \mathcal{Z}u$ belong.

In what follows it is assumed that point measurements are recorded from a finite number of locations distributed uniformly over the spatial domain i.e. the data is spatially sampled at the $n-1$ nodal points $\frac{1}{n}, \frac{2}{n}, ..., \frac{n-1}{n}$. Note that this is not a strong requirement. Non-uniform sampled data in the spatial domain can also be handled. For simplicity Ω is assumed one-dimensional, in particular $\Omega = (0, 1)$. The results however are also valid for $\Omega \subset \mathbb{R}^d$ with $d > 1$.

Specifically, in the case of discrete-discrete observations considered here, the observation operator is defined as

$$y_{N,n} = \mathcal{Z}u = \{u(t_i, x_j)\}_{i=1,...,N}^{j=1,...,n} \qquad (25)$$

and the observation space is $\mathcal{Y} = \mathbb{R}^{N \times n}$. It is assumed that in the time domain, the data is uniformly sampled over the interval $[0, T]$ of observation with a sampling time Δt. In practice it is assumed that both $\Delta x = \frac{1}{n}$ and Δt are sufficiently small so that the full behaviour of the solution u is captured.

Let V^n be a finite dimensional subspace of V. The identification problem is to determine, based only on the given set of discrete observations $y_{N,n} = \{u(t_i, x_j)\}_{i=1,...,N}^{j=1,...,n}$ and $v_{N,n} = \{v(t_i, x_j)\}_{i=1,...,N}^{j=1,...,n}$ a finite dimensional dynamical system whose solution $u_n \in C(\mathbb{R}_+; V^n) \cap L^2(0, T; H^j(\Omega))$ approximates the observed dynamical behaviour in V^n.

3.3 *Finite Element Approximation*

A common choice of finite element subspaces V^n on Ω are the spaces of continuous piecewise polynomial functions defined with respect to a uniform mesh on Ω. Let $\{\varphi_j^n\}_{j=0}^n$ be the standard lth order B-spline base (de Boor, 1978). In this case $V^n = \text{span}\{\varphi_j^n\}_{j=0}^n$ and V is the Sobolev space $H^l(\Omega)$. Note that $\bigcup_{n \in \mathbb{Z}} V^n$ is dense in $H = L^2(\Omega)$ and $H^l(\Omega)$. Moreover, for the B-spline basis the inequality (6) holds.

When defining the approximation subspaces V^n and the associated basis elements respectively, it is important to take into account the boundary conditions. For instance for zero Dirichlet boundary conditions $V = H_0^l$ and V^n is the space of continuous, piecewise, lth order polynomial functions corresponding to the uniform partition $\{0, \frac{1}{n}, \frac{2}{n}, ..., 1\}$, which vanish at 0 and 1 and is denoted $S_0^{n,l}$. In parctice the standard B-spline basis functions $\{\varphi_j^n\}_{j=0}^n$ can be modified to account for specific boundary conditions.

Let $y_n(t, x) = I_{n,1}^x I_{N,1}^t u(t, x)$ be the linear B-spline interpolation of the pointwise data $y_{N,n} = \{u(t_i, x_j)\}_{i=1,...,N}^{j=1,...,n}$ where $I_{N,1}^t$ and $I_{n,1}^x$ are the linear interpolation operators in $S^{N,1}([0, T])$ and $S^{n,1}(\Omega)$ respectively. It follows that $y_n(t, x)$ can be expressed in terms of the two-dimensional linear tensor splines $\Phi^{N,n}(t, x) = \varphi^N(t) \otimes \varphi^n(x)$ such that

$$y_n(t, x) = \sum_{i=1}^{N} \sum_{j=0}^{n} y_{n,j}(t_i) \Phi_{i,j}^{N,n}(t, x) \quad (26)$$

The interpolation $v_n(t, x)$ of the perturbation function from the pointwise data $v_{N,n}$ can be defined in a similar manner.

Choosing the optimal approximation subspace V^n, that is the mesh size $h = \frac{1}{n}$ is very important. In practice, the initial mesh size could be selected based on the frequency content of the solution along the spatial and temporal coordinates.

As in the numerical integration of PDE's, if the mesh is too fine the dimension of the resulting finite dimensional model will be too large and computationally expensive. In practice however, if the sampling is too fine, the inital interpolation can be projected onto a coarser subspace.

In identification, the mesh size is related to the number of measurement locations in the spatial domain. Recent developments in sensor technology mean that the number of measurement locations can be quite large and still cost-effective. For example, the data could represent a video recording of patterns in a chemical reaction, a sequence of MRI scans of brain activity or the web tension profile measured using a full-web measurement system during the papermaking process.

The order of the identified dynamical system could be much smaller than the initial size of the coordinate vector. Finding a more economical representation for the initial interpolate y_n can be achieved using a simple iterative algorithm. Let ρ be a desired approximation error bound. Starting with $m = n-1$ the initial interpolate y_n is projected successively onto coarser spaces $V^{n-2}, V^{n-1}, \ldots$ in order to find the minimum number of basis functions $m = m_{min}$ for which the approximation error is less that the given threshold. This approach leads to a more economical approximation of the solution.

4. NUMERICAL EXAMPLE

This section illustrates the identification of a finite dimensional, discrete-time dynamical model for the following diffusion equation (heat equation)

$$\frac{\partial u(t,x)}{\partial t} - c\frac{\partial^2 u(t,x)}{\partial x^2} = 0, \tag{27}$$

with domain $\Omega = (0, 1)$, initial conditions

$$u(0,x) = \begin{cases} 2Bx & x \in (0,\, 0.5) \\ 2B - 2Bx & x \in (0.5,\, 1) \end{cases} \tag{28}$$

and Dirichlet boundary conditions. For $B = \pi^2$ the exact solution $u(t,x)$ of the initial value problem (27), (28) is given by the following series expansion

$$u(t,x) = \sum_{k=1}^{\infty} \frac{8(-1)^{k-1}}{(2k-1)^2} e^{-c(2k-1)^2\pi^2 t} \cdot \sin\left((2k-1)\pi x\right). \tag{29}$$

The solution, based on the first 50 terms of the expansion (29) with $c = 1.0$ was sampled uniformly in both the spatial and time domain with $\Delta x = 1/128$ and $\Delta t = 0.5 \times 10^{-3}$.

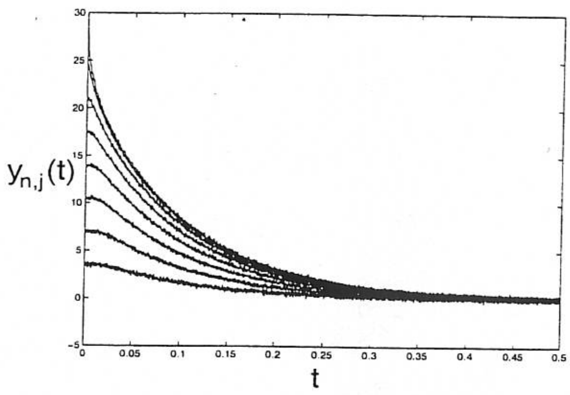

Fig. 1. Noise-corrupted coordinate vector $\tilde{\mathbf{y}}_n(t)$

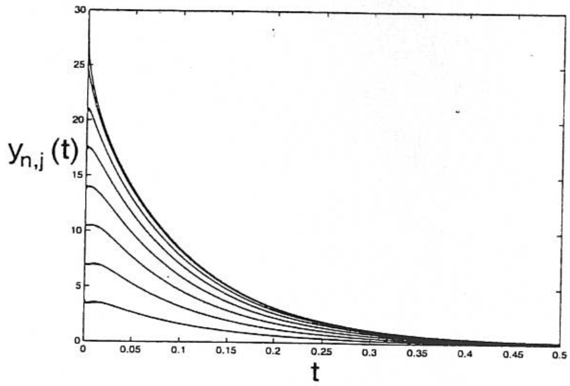

Fig. 2. Model predicted output (solid) $\widehat{\mathbf{y}}_n(t)$ of the MIMO-ARMA model superimposed on the original noise free trajectory (dash-dot) $\mathbf{y}_n(t)$.

From each location N=1000 data points were generated and superimposed with white noise with variance $\sigma^2 = 0.01$. The data were interpolated using linear B-spline functions. The initial interpolated solution involving 127 basis functions was subsequently projected on a lower approximation subspace and expressed in terms of only 15 basis functions.

One thousand samples shown in Figure 1, corresponding to the noisy coordinate vector $\tilde{\mathbf{y}}_n(t) = (\tilde{y}_{n,1}(t), \ldots, \tilde{y}_{n,15}(t))$, were used for identification. The data was used to estimate a MIMO-ARMA model (not given here for reasons of space) which included both deterministic and stochastic terms. The selection of the linear terms included in each of the 15 subsystems was performed using the Orthogonal Forward Regression algorithm (Billings *et al.*, 1988).

The deterministic part of the model was simulated and the resulting model predicted output

$$\widehat{\mathbf{y}}_n(t) = \mathbf{A}\widehat{\mathbf{y}}_n(t)) \tag{30}$$

was plotted in Figure 2 superimposed on the original noise-free trajectory $\mathbf{y}_n(t)$ The two trajectories are practically indistinguishable. In equation (30) $\mathbf{A}$ is the matrix of backward shift polynomials

$$\mathbf{A}(i,j) = a_{i,j}^{1}q^{-1} + ... + a_{i,j}^{n_y}q^{-n_y} \quad (31)$$

where $a_{i,j}^{k}$ represent the estimated coeeficients and q^{-k} is the backward shift operator.

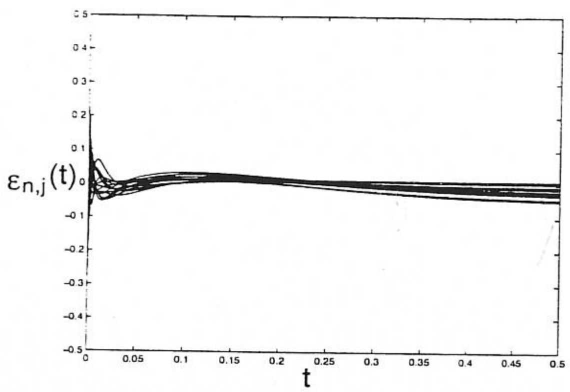

Fig. 3. Model prediction errors $\varepsilon_{n,k}(t)$

The model prediction errors $\varepsilon_{n,k}(t) = y_{n,k}(t) - \widehat{y}_{n,k}(t)$, relative to the original noise-free coordinate vector $\mathbf{y}_n(t)$, plotted in figure (3), are very small with a NRMSE of less than 1%. The model output was used to compute the approximate PDE solution $\widehat{y}(t,x)$ shown in Fig.(4).

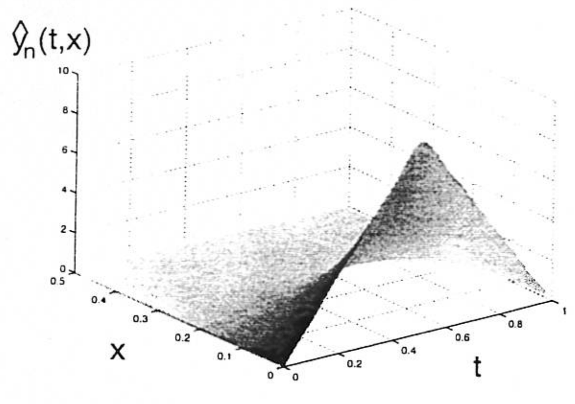

Fig. 4. Predicted PDE solution $\widehat{y}(t,x)$

The prediction errors $e(t,x) = y(t,x) - \widehat{y}(t,x)$ plotted in Fig.(5) have the same order of magnitude as the initial B-spline approximation errors.

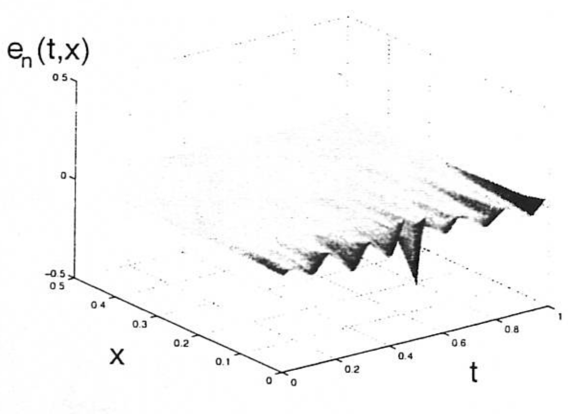

Fig. 5. Predicted PDE solution errors $e(t,x)$

5. CONCLUSIONS

Finite dimensional approximations of PDE's play an essential role in the control and simulation of distributed parameter systems. This paper has developed, analysed and tested a method for deriving the finite dimensional approximation directly from noisy data using system identification.

The proposed approach can be used to identify, directly from measurements, a finite dimensional approximate representation of a distributed parameter system for which the governing PDE's are not available. However, even when the equations are known, this approach can be used to provide a more economical and even more accurate representation than the one obtained by classical methods. Indeed, in a companion paper it will be shown both in theory and by means of an example that, for a given subspace V^n, the identified model is more accurate than the equivalent finite element Galerkin approximation derived from the original PDE's.

6. ACKNOWLEDGEMENTS

The authors fully acknowledge that this work was supported by the Engineering and Physical Sciences Research Council, UK.

7. REFERENCES

Bensoussan, A., G. DaPrato, M. C. Delfour and S. K. Mitter (1992). *Representation and Control of Infinite Dimensional Systems*. Vol. 1. Birkhäuser, Boston.

Billings, S. A., S. Chen and M. J. Kronenberg (1988). Identification of MIMO non-linear systems using a forward-regression orthogonal estimator. *Int. J. Control* **49**, 2157–2189.

Brenner, S. C. and L. Ridgway Scott (1994). *The Mathematical Theory of Finite Element Methods*. Vol. 1. Springer Verlag.

de Boor, C. (1978). *A practical guide to splines*. Springer Verlag.

Naylor, A. W. and G. R. Sell (1989). *Linear Operator Theory in Engineering and Science*. Vol. 40 of *Applied Mathematical Sciences*. Springer Verlag, New York.

Temam, R. (1983). *Navier-Stokes Equations and Nonlinear Functional Analysis*. Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, Philadelphia.